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

Many intracellular signaling pathways are composed of molecular switches, proteins that transition between two states—*on* and *off*. Typically, signaling is initiated when an external stimulus activates its cognate receptor that in turn causes downstream switches to transition from *off* to *on* using one of the following mechanisms: activation, in which the transition rate from the *off* state to the *on* state increases; derepression, in which the transition rate from the *on* state to the *off* state decreases; and concerted, in which activation and derepression operate simultaneously. We use mathematical modeling to compare these signaling mechanisms in terms of their dose-response curves, response times, and abilities to process upstream fluctuations. Our analysis elucidates several general principles. First, activation increases the sensitivity of the pathway, whereas derepression decreases sensitivity. Second, activation generates response times that decrease with signal strength, whereas derepression causes response times to increase with signal strength. These opposing features allow the concerted mechanism to not only show dose-response alignment, but also to decouple the response time from stimulus strength. However, these potentially beneficial properties come at the expense of increased susceptibility to upstream fluctuations. In addition to above response 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) exhibits a sustained response whereas derepression (activation) adapts. In total, we show how the architecture of molecular switches govern their response properties. We also discuss the biological implications of our findings.

Introduction

Several 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 *derepression*, 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]. In many of these cases, derepression occurs through decrease in the degradation rate of a component instead of its deactivation rate. Concerted mechanisms are found in bacterial two component systems, wherein the same component acts as kinase and phosphatase [28–35].

Mathematical modeling has proven to be a useful tool for understanding the design principles of signaling pathways, and,

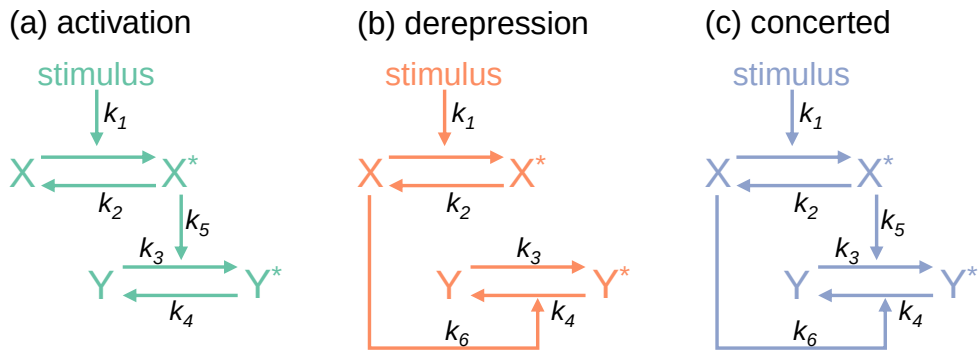

Figure 1: **Mechanisms for signaling through molecular switches**. Presentation of a stimulus activates a receptor ($X \rightarrow X^*$). The reverse reaction causes deactivation of the receptor ($X^* \rightarrow X$). These transitions govern the activity of a molecular switch downstream. (**a**) In 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.

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 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 these mechanisms should enable our understanding of why different organisms have chosen different mechanisms. To this end, we specifically choose four metrics for the comparison: a) dose-response, b) response time, c) ability to suppress or filter stochastic fluctuations in upstream components, and d) effect of receptor removal. The rationale behind comparing 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 signaling pathways evince receptor removal [11, 42, 54–56]. We study these properties by constructing 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 with the stimulus strength. In contrast, derepression makes the switch response less sensitive than the receptor occupancy and slows down the response speed as stimulus strength increases. 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 mechanism (equal activation and repression), the dose-response curve of the pathway aligns with the receptor occupancy and the response time 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 *off* state to the *on* state increases as the number of active receptor molecules increases (Fig. 1(a)). In the derepression mechanism, the transition rate from the *on* state to the *off* 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 active receptor, and k_6 is the strength of repression of an individual inactive receptor. Thus, the (total) activation strength is $k_5 X_T$ and (total) repression strength is $k_6 X_T$. Lastly, we assume that X_T and Y_T are 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. 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 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 (1) requires rate constants and initial conditions to be specified. 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 (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 (1) is

$$\overline{X^*} = \frac{S X_T}{S + \frac{k_2}{k_1}}, \tag{2a}$$

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

Here $\overline{X^*}$ and $\overline{Y^*}$ are the number of active (occupied) receptors and the number of active switches 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. (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. Furthermore, the fractional receptor occupancy ($\overline{X^*}/X_T$) for a given stimulus (S) is determined by $k_1 S / k_2$. As for the switch, the response ($\overline{Y^*}$) 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-response of the switch depends 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 (4) provides the following insights:

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

(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 $k_4 \ll k_3 + k_5 X_T$. The limiting case of $k_3 = 0$ leads to $Y_0^* = 0$; likewise, $k_4 = 0$ results in $Y_\infty^* = Y_T$. 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 ($k_5 X_T$) 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 the switch (Y^*) is determined by the ratio of active receptors to the total number of

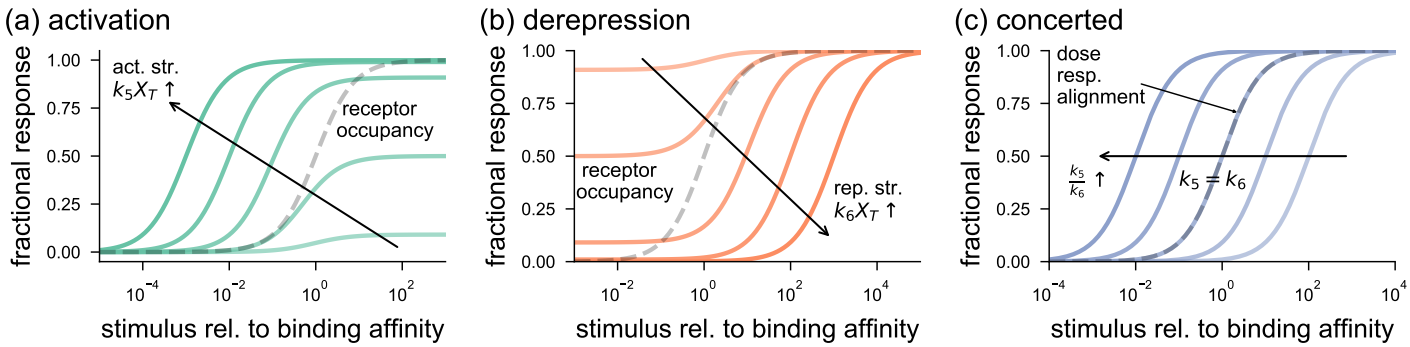

Figure 2: **Dose-response curves for signaling mechanisms through molecular switches**. The response is measured in terms of fraction of active switches Y^*/Y_T as the stimulus level varies. The receptor occupancy curve denotes the fraction of active receptors X^*/X_T. The stimulus is normalized by its 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 as $k_3 = 1$, $k_4 = 0$, and $k_5 = 0$. **(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 $(0.01, 0.1, 1, 10, 100)$.

receptors (X^*/X_T) [13, 14]. The absolute value of the total number of receptors (X_T) has no bearing on $\overline{Y^*}$. This may be seen by setting $k_3 = 0$ and $k_4 = 0$ in the expression of $\overline{Y^*}$ in (2):

$$\overline{Y^*} = \frac{SY_T}{S + \frac{k_2k_6}{k_1k_5}}. \tag{5}$$

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 the decreased amount of the receptor (X_T) lowers the repression strength k_6X_T and shifts the dose response curve to the left in comparison to that of the wild-type system. 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 $R(0) \neq \overline{R}$. We may also think of $1/\mathcal{T}_R$ as the speed of the response in the sense that if the response is determined by a single kinetic step, $\mathcal{T}_R$ is reciprocal of the rate constant for that step. For example, the response time for the receptor is given by (section S1, SI):

$$\mathcal{T}_{X^*} = \frac{1}{k_1S + k_2}. \tag{7}$$

Thus, the response time decreases (i.e., response speeds up) if k_1S+k_2 increases. Because the response time depends upon the sum $k_1S + k_2$ and the steady-state receptor occupancy depends upon the ratio k_1S/k_2, these quantities can be tuned independently.

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

$$\mathcal{T}_{Y^*}\big|_{S=0} = \frac{1}{k_3 + k_4 + k_6X_T}. \tag{8}$$

When the stimulus is present, analytic solutions to the integrals in (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 ODE system in (1)

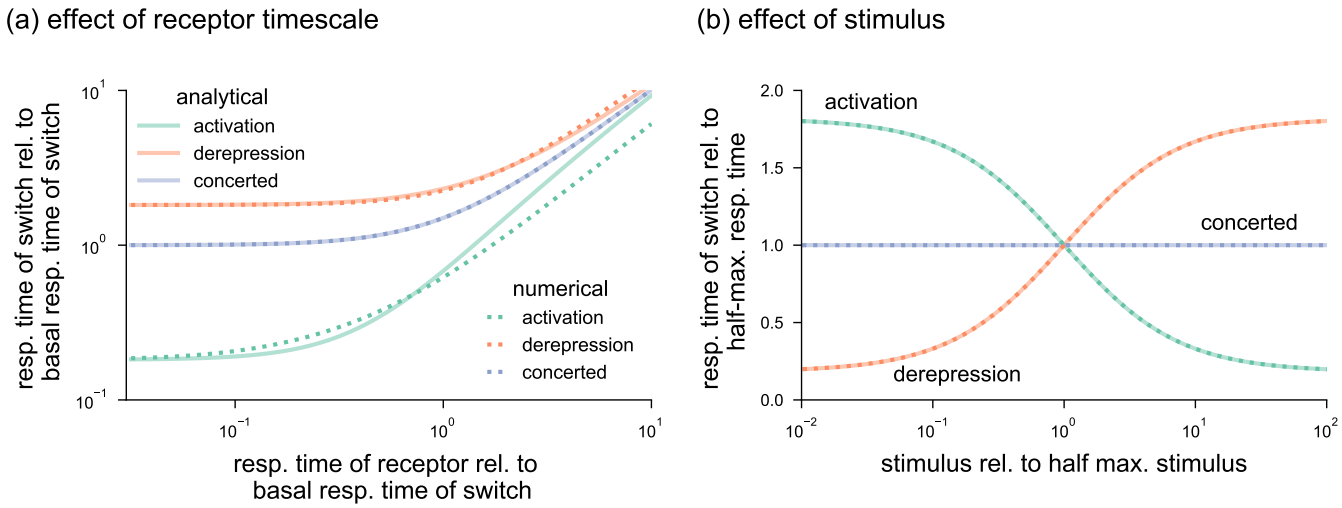

Figure 3: **Response times of molecular switches governed by activation, derepression, and concerted mechanisms**. (**a**) The response time of the switch increases as the response time of the receptor increases. Differences in the signaling mechanisms are more prominent when the receptor response is 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 (9) (solid lines) and numerically validated by using (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**) With increase in the stimulus level, response time decreases for activation, increases for derepression, and does not change 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.

around its steady-state:

$$\mathcal{T}_{Y^*} \approx \underbrace{\frac{1}{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k2}}}_{\text{resp. time of } Y^* \text{ when } X^* \text{ is in steady-state}} + \underbrace{\frac{1}{k_1 S + k_2}}_{\text{resp. time of } X^*} \times \underbrace{\frac{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k2}}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k2}}}_{\text{time-averaging}}. \quad (9)$$

This equation is exact for the special case when $k_5 = k_6$ (section S2-B, SI). The first term in (9) 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 $k_4 + k_6 X_T - k_6 \frac{k_1 S X_T}{k_1 S + k_2}$; so, inverse of their sum would give the response time. The second term represents the response time of the receptor ($\mathcal{T}_{X^*}$) multiplied by a time-averaging factor which computes the ratio of $\mathcal{T}_{X^*}$ to the sum of $\mathcal{T}_{X^*}$ and the response time of the switch when $X^* = \overline{X^*}$. The time-averaging term lies between 0 and 1; its value approaches 0 if the receptor response is much faster than the switch response when $X^* = \overline{X^*}$ and approaches 1 if the receptor response is much slower than that of the switch.

If the receptor response is much faster than that of the switch, we expect that the latter does not depend upon the former (time-averaging term $\to 0$). Indeed in this limit, (9) gives

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

Comparing (10) with the basal response time in (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 and a perfect 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 (time-averaging term $\to 1$). Indeed in this case, (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 (S). 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. We find that activation shortens the response time (speeds up the response) with increasing stimulus levels, whereas derepression increases the response time (slows down the response) (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. To better understand this behavior, consider the response time for the limiting case of fast receptor dynamics. (10) can be rewritten as

$$\mathcal{T}_{Y^*} \approx \frac{\frac{1}{k_3 + k_4 + k_6 X_T}\Theta_{Y^*} + \frac{1}{k_3 + k_4 + k_5 X_T} S}{S + \Theta_{Y^*}}, \quad (11)$$

which changes from $\frac{1}{k_3 + k_4 + k_6 X_T}$ at $S = 0$ to $\frac{1}{k_3 + k_4 + k_5 X_T}$ as $S \to \infty$. The half-maximal stimulus Θ_{Y^*} is same as defined in (4). For the

activation mechanism, $\frac{1}{k_3+k_4+k_6X_T} > \frac{1}{k_3+k_4+k_5X_T}$; so the response time decreases with stimulus. Moreover, $\frac{1}{k_3+k_4+k_6X_T} < \frac{1}{k_3+k_4+k_5X_T}$ for the derepression mechanism and $\frac{1}{k_3+k_4+k_6X_T} = \frac{1}{k_3+k_4+k_5X_T}$ for the perfect concerted mechanism. Therefore, the response time increases with stimulus for the derepression mechanism and is independent of the stimulus for the concerted case.It is also worth pointing out 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)).

Our analysis of dose-response properties for ratiometric signaling given in (5) reveals that this mechanism is independent of the total number of receptors X_T when the basal rates of the switch are zero ($k_3 = 0$ and $k_4 = 0$). Using these values in the expression for the response time in (9) demonstrates that this property does not hold for the response time. Specifically, the response time decreases with an increase in X_T (section S2–B–c, SI).

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 mechanism simultaneously decreases the lifetime of the *off* state and increases the lifetime of the *on* state. Therefore, 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. 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.

Reaction	Population 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^2 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 [62–64] (section S3-B, SI). In particular, moments for the number of active receptors (X^*) are given by

$$\langle X^* \rangle = \frac{k_1 S X_T}{k_1 S + k_2}, \tag{12a}$$

$$CV^2_{X^*} = \frac{\langle X^{*2} \rangle - \langle X^* \rangle^2}{\langle X^* \rangle^2} = \frac{1}{X_T}\frac{k_2}{k_1 S}. \tag{12b}$$

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 (1). The coefficient of variation squared increases as the number of receptors (X_T) decreases. Therefore the noise analysis is important 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

$$\langle Y^* \rangle \approx \frac{k_3 + k_5 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 \frac{k_2}{k_1 S + k_2} X_T} Y_T, \tag{13a}$$

$$\begin{aligned} CV^2_{Y^*} \approx & \underbrace{\frac{1}{Y_T}\frac{k_4 + k_6 X_T \frac{k_2}{k_1 S + k_2}}{k_3 + k_5 X_T \frac{k_1 S}{k_1 S + k_2}}}_{\text{contribution from act./deact. of } Y^*} + \\ & CV^2_{X^*} \times \underbrace{\frac{k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}{k_1 S + k_2 + k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}}_{\text{time-averaging}} \\ & \times \underbrace{\frac{\left(\frac{k_1 S}{k_1 S + k_2}\right)^2 (k_4 k_5 X_T + k_6 X_T (k_3 + k_5 X_T))^2}{\left(k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}\right)^2 \left(k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}\right)^2}}_{\text{coupling}}. \end{aligned} \tag{13b}$$

We validate these approximations using exact semi-analytical approach based on [65] (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 [66–68]. 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 (($CV^2_{X^*}$). The contribution from activation/deactivation of the switch in (13) has a similar form as $CV^2_{X^*}$ in (12). Accordingly, the contribution of this term decreases with increase in Y_T or increase in the ratio of the total activation rate $\left(k_3 + k_5 X_T \frac{k_1 S}{k_1 S + k_2}\right)$ with total deactivation rate $\left(k_4 + k_6 X_T \frac{k_2}{k_1 S + k_2}\right)$. This ratio increases if the activation strength increases or the repression strength decreases. The contribution of CV^2_X to $CV^2_{Y^*}$ is scaled by time-averaging and coupling terms. The time-averaging term is the same as that in (9); it varies between 0 and 1, depending upon the relative timescales of the receptor and the switch. Thus, 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^*. The coupling term in (13) determines how strongly X^* affects Y^*. For example, this term is zero when the stimulus is absent ($S = 0$) or when both k_5 and k_6 are zero. In both these cases, the switch is decoupled from the receptor.

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 the 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 differences in the noise properties, if any, are solely due to the architecture of the molecular switch and not dependent on the parameters. With this setup, we examine the effect of relative timescales (response times) of the receptor and the switch. We observe that in (13), varying $k_1 S + k_2$ while maintaining $k_1 S / k_2$ only affects the time-averaging term; all other terms are not affected. As shown in Fig. 4(a), 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 these properties, activation and derepression counteract to enable intermediate response for the concerted mechanism. Intuitively, the increase in fluctuations occurs because in the concerted mechanism, fluctuations in the upstream component affect both transitions $Y \rightarrow Y^*$ and $Y^* \rightarrow Y$. In the case of activation and derepression, however, only one of these transitions is coupled with the upstream component. As a result, the concerted mechanism performs worse in terms of noise. We further highlight this observation by varying the relative strengths of activation ($k_5 X_T$) and derepression ($k_6 X_T$) in Fig. 4(b). The noise is greatest for the concerted mechanism when $k_5 X_T = k_6 X_T$.

We also analyze the special case 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). However, similar to the response time, the $CV^2_{Y^*}$ also depends upon X_T through the time-averaging term and $CV^2_{X^*}$, both of which decrease with increases in X_T (section S3–B–c, SI). To summarize, ratiometric signaling only holds for the steady-state response. A cell that has higher X_T would respond faster as well as with less noise than a cell with a smaller X_T.

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 [69, 70], 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 [70] which is what we focus on.

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

$$\frac{dX}{dt} = k_p - k_d X - k_1 S X + k_2 X^*, \tag{14a}$$

$$\frac{dX^*}{dt} = k_1 S X - k_2 X^* - k_d^* X^*, \tag{14b}$$

$$\frac{dY^*}{dt} = \left(k_3 + k_5 X^*\right)\left(Y_T - Y^*\right) - \left(k_4 + k_6 X\right) Y^*. \tag{14c}$$

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 (1) and the ODEs in (14) 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 behavior is referred to as adaptation [70, 71].

What are appropriate parameter regimes where the switch response $Y^*(t)$ in (14) 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 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.

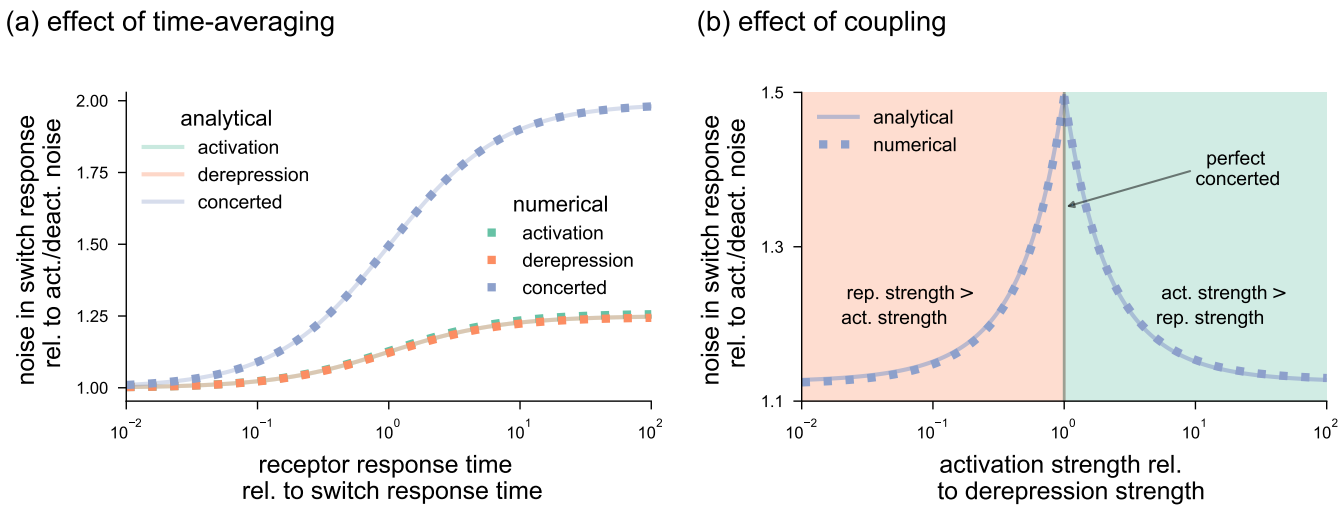

Figure 4: **Noise in the number of active switch molecules**. Noise is quantified using coefficient of variation squared (CV^2_{Y*}) as in (13). The overall CV^2_{Y*} is shown relative to the contribution from activation/deactivation of Y^*. The analytical result is computed using (13), which is validated numerically. (**a**) Noise with change in response time of the receptor. The noise increases as receptor response time increases, i.e., as the receptor slows down in comparison with the switch response time. 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. The receptor response time ($k_1S + k_2$) is varied by changing k_2 while keeping the same receptor occupancy through the ratio k_1S/k_2, so as to keep the same number of switches.The differences across signaling mechanisms are controlled by ensuring the same total activation rate of the switch $k_3 + k_5\frac{k_1SX_T}{k_1S+k_2}$ and same total deactivation rate $k_4 + k_6\frac{k_2SX_T}{k_1S+k_2}$. We used the following parameters: $k_3 = 0$, $k_4 = 1$, $k_5 = 0.02$, 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 were taken to be 100 each. The receptor occupancy was maintained by $k_1S/k_2 = 1$.(**b**) Noise with change in relative strengths of activation and derepression for a concerted mechanism. The noise is highest when the activation and derepression strengths match (perfect concerted mechanism). Deviating from the perfect concerted mechanism towards either stronger activation (shaded green region) or stronger derepression (shaded orange region) leads to smaller noise. Parameters were chosen such that total activation and the total deactivation rates were same across signaling mechanisms. For derepression, the activation strength was kept constant and the repression strength k_6X_T was varied with a commensurate change in the basal deactivation rate k_4. For activation, the repression strength was kept constant and the activation strength k_5X_T was increased with appropriate change in the basal activation rate k_3. We used the following parameters: $k_1 = 1$, $S = 1$, $k_2 = 1$, $k_3 + k_5\frac{k_1SX_T}{k_1S+k_2} = 1$, $k_4 + k_6\frac{k_2X_T}{k_1S+k_2} = 1$, $X_T = 100$, and $Y_T = 100$.

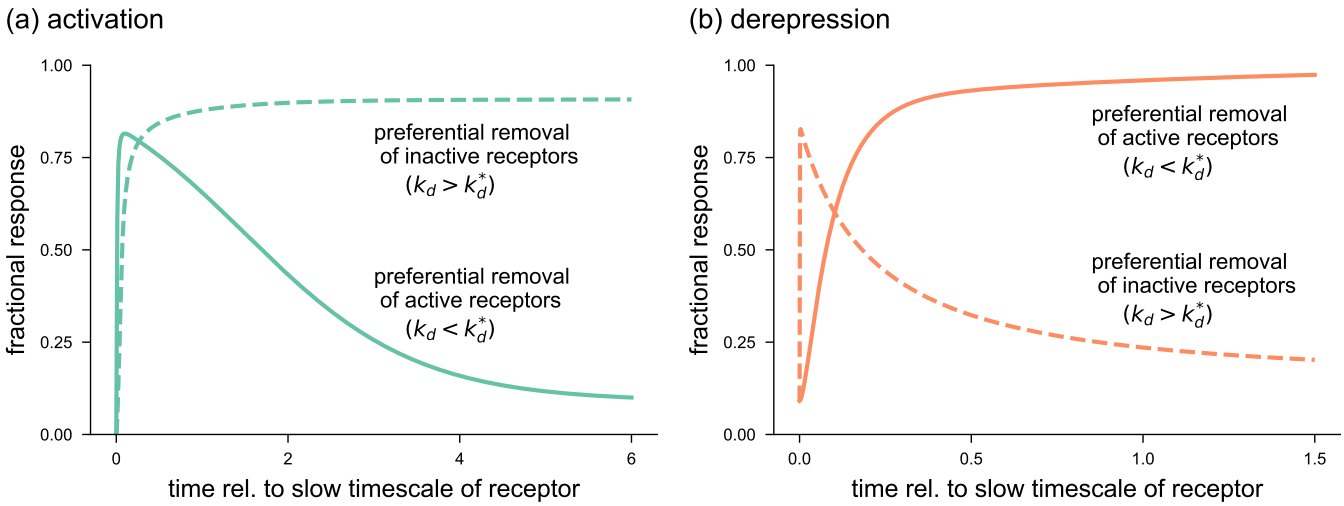

Figure 5: **Effect of receptor removal on responses of activation and derepression mechanisms**. The response is measured in terms of fraction of active receptors (Y^*/Y_T) over time which is normalized to the slow timescale of the receptor (see section S4, SI). (**a**) For an activation mechanism, the 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. For the adaptive response, we chose $k_p = 0.11$, $k_d = 0.0011$, $k_d^* = 0.11$. For the sustained response, we set $k_p = 101$, $k_d = 1.01$ and $k_d^* = 0.01$. Rest of parameters were selected as $k_1 = 1$, $k_2 = 1$, $k_3 = 0$, $k_4 = 10$, $k_5 = 1$, $k_6 = 0$, $S = 1$, and $Y_T = 100$ (**b**) For derepression mechanism, preferential removal of X results in adaptation whereas preferential removal of X^* causes sustained response.For the adaptive response, we used $k_p = 101$, $k_d = 1.01$, $k_d^* = 0.01$. For the sustained response, we chose $k_p = 1$, $k_d = 0.01$ and $k_d^* = 1$. Rest of parameters were taken as $k_1 = 1$, $k_2 = 1$, $k_3 = 10$, $k_4 = 0$, $k_5 = 0$, $k_6 = 1$, $S = 100$, and $Y_T = 100$

Discussion

Molecular switches are important components of most signaling pathways. Typically, these switches can exist in two states, *on* and *off*, and the presence of an external stimulus biases the switch toward the *on* state. This transition can occur either by increasing the *off*-to-*on* rate (activation), decreasing the *on*-to-*off* rate (derepression), or both (concerted). We characterized these three mechanisms in terms of their dose-response curves, response times, and ability to process upstream fluctuations. We further examined how these three mechanisms were affected by receptor removal. The following list summarizes key differences in the performance of switches based on activation, derepression and concerted mechanisms:

- Both activation and derepression cannot align signaling activity with receptor occupancy. In particular, activation reduces the stimulus level required for half-maximal signaling as compared to 50% receptor occupancy ($\Theta_{Y^*} < \Theta_{X^*}$), whereas derepression produces a rightward shift of the dose-response curve ($\Theta_{Y^*} > \Theta_{X^*}$). The dose-response curve aligns with the receptor occupancy curve ($\Theta_{Y^*} = \Theta_{X^*}$) for a perfect concerted mechanism (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.

- The response time for the activation mechanism decreases with signal strength, whereas it increases for the derepression mechanism. Importantly, the response time for a perfect concerted mechanism is independent of signal strength (Fig. 3).

- Activation and derepression mechanisms respond similarly to upstream fluctuations, whereas the concerted mechanism is more susceptible to fluctuations (Fig. 4). Unlike the mean steady state response, fluctuations in the output signal for the ratiometric signaling do depend on the total number of receptors.

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

These results suggest performance trade-offs in the operating characteristics for each mechanism. The activation mechanism can increase the sensitivity of the pathway and generate response times that decrease with signal strength, but at the cost of dose-response curves that do not align with receptor occupancy, potentially limiting the pathways ability to transfer information [72]. In this sense, the activation mechanism operates as an 'eager' system that is sensitive to small receptor occupancies and accelerates the response for stronger signals. Therefore, activation seems appropriate for situations in which the cost of a false negative is greater than a false positive. For example, the adrenaline response to imminent danger should be sensitive and fast because cost of a false positive is small but a false negative can be deadly.

Similar to the activation mechanism, derepression leads to misalignment of the dose-response curve and receptor occupancy. However, for derepression the dose-response curve is shifted to the right. Another difference between these mechanisms is that for derepression, the response time increases with signal strength. Therefore, derepression acts as a 'conservative' system that does not respond to low receptor occupancy, waiting for a strong signal before committing to a response. Derepression seems appropriate for scenarios where the cost of a false positive is greater than a false negative. Interestingly, derepression-based signaling is found in many plants pathways. We speculate that it happens because plants have to continually allocate their limited resources between growth in competition with its neighbors and immunity to survive pathogen attack [73,74]. For example, plants would perhaps ignore growth of a low level of pathogenic bacteria before allocating resources to fight them. Another possible scenarios where derepression may be used include irreversible cell-fate decisions such as the WNT pathway for embryo development [75], and fail-safe mechanisms such as the hypoxia-inducible factor in face of oxygen deprivation [76].

The concerted mechanism is better able to align with the receptor occupancy curve than either the activation or derepression mechanisms. Therefore, it has a better information fidelity [72]. The concerted mechanism also can generate response times that are independent of the strength of the input signal. However, these features come at the cost of higher susceptibility to upstream fluctuations. We note that in a recent study it was shown that ratiometric (concerted) signaling provided an advantage for gradient sensing, because it could compensate for spatial variations in the receptor concentration [14]. The system under consideration in that study was the mating response of yeast. For this case, the spatial fluctuations in the receptor concentration were larger than downstream fluctuations in signaling, allowing the concerted mechanism to outperform an activation-based mechanism.

While misalignment of the dose-response curve with receptor occupancy can cause loss of information, it may also offer some advantages. 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 feature 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 an adaptive response and thereby a fold-change detection [59]. A key feature of fold-change detection is that the sensitivity of the system decreases each time the system adapts [59,77]. Our results suggest that a relative sensing mechanism may be implemented with a derepression if the receptor removal operates on inactive receptors. We speculate that a negative feedback operating on inactive receptors would also yield the same effect.

Given that activation and derepression shift dose-response in opposite directions, a natural question to ask is whether dose-response alignment can occur in a signaling cascade where activation and derepression operate sequentially? To explore this possibility, we constructed 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). It is worth noting that nonlinear regulation, such as feedback and feedforward loops, can also be used to compensate for undesirable characteristics of a given signaling mechanism. For example, negative feedback can align the dose-response curve with receptor occupancy for signaling pathways that operate through activation [43, 72, 78].

The models considered here are based on mass action kinetics and therefore cannot capture saturation effects. Traditionally, signaling pathways are modeled using Michelis-Menten kinetics. While we believe the qualitative features of our results will hold in this case, investigating how the behavior of the three mechanisms changes when the effects of enzyme saturation are included will be the subject of future work. Another future direction is to extend the analysis to include feedback and feedforward regulation. Finally, while we have focused our investigations on signaling pathways, our results are likely to be relevant in other intracellular systems, such as gene regulatory networks and metabolic pathways.

Acknowledgments

The authors thank Daniel Lew (Duke University), Nicolas Buchler (North Carolina State University), Cesar A. Vargas-Garcia (Agrosavia), and members of the Jones and Elston labs for discussion and feedback. This work was supported by NSF grant MCB–1713880 to AMJ and TCE, NIH grant RO1 GM065989 to AMJ, and NIH grant R35 GM127145 to 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] D. Schnoerr, G. Sanguinetti, and R. Grima, "Approximation and inference methods for stochastic biochemical kinetics—a tutorial review," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 9, p. 093001, 2017.

[64] 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.

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

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

[67] M. Komorowski, J. Miękisz, and M. P. Stumpf, "Decomposing noise in biochemical signaling systems highlights the role of protein degradation," *Biophysical journal*, vol. 104, no. 8, pp. 1783–1793, 2013.

[68] 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.

[69] 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.

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

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] H. Clevers and R. Nusse, "Wnt/β-catenin signaling and disease," *Cell*, vol. 149, no. 6, pp. 1192–1205, 2012.

[76] S. Salceda and J. Caro, "Hypoxia-inducible factor 1α (HIF-1α) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions," *Journal of Biological Chemistry*, vol. 272, no. 36, pp. 22642–22647, 1997.

[77] 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.

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

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

[80] 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.

[81] S. Modi, M. Soltani, and A. Singh, "Linear noise approximation for a class of piecewise deterministic Markov processes," in *2018 Annual American Control Conference (ACC)*, pp. 1993–1998, IEEE, 2018.

[82] 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 (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 $A^*(t)$denote the number of molecules that are in states A and A^*, respectively, 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 (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, (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 $0 < f < 1$. More specifically, we define $\mathcal{T}_f$ as the solution to the following equation

$$\frac{A^*(\mathcal{T}_f) - A^*(0)}{\overline{A^*} - A^*(0)} = f \tag{S1.8a}$$

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

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

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

which has a straightforward solution

$$\mathcal{T}_f = -\frac{\log(1 - f)}{k_1 S + k_2}. \tag{S1.10}$$

Notably, the response time is set by $1/(k_1 S + k_2)$ up to a scale which depends on the specific value of f. We discuss three cases. First, setting $f = 1/2$ 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, $f = (e-1)/e \approx 0.632$ 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

$$\mathcal{T}_{90\%} - \mathcal{T}_{10\%} = \frac{\log 9}{k_1 S + k_2}. \tag{S1.11c}$$

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

S2-A Transient solution

Analytical solutions for nonlinear ODEs such as those in (S2.1) typically do not exist. However, a careful look at (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. The solutions for other cases can be computed numerically. We also provide an approximate solution using linearization around the steady-state solution $(\overline{X^*}, \overline{Y^*})$.

It turns out that the forms of the ODEs for the perfect concerted mechanism and the linearized system are similar. Therefore, we consider the following generic system first and compute its transient solution.

$$\frac{dR_1}{dt} = \eta_0 - \eta_1 R_1, \tag{S2.3a}$$

$$\frac{dR_2}{dt} = \xi_0 + \xi_1 R_1 - \xi_2 R_2, \tag{S2.3b}$$

with initial conditions $(R_1(0), R_2(0))$. Let $\mathcal{R}_1[\omega]$ and $\mathcal{R}_2[\omega]$ respectively denote the Laplace transforms of $R_1(t)$ and $R_2(t)$. Then we have that

$$\omega \mathcal{R}_1[\omega] - R_1(0) = \frac{\eta_0}{\omega} - \eta_1 \mathcal{R}_1[\omega], \tag{S2.4a}$$

$$\omega \mathcal{R}_2[\omega] - R_2(0) = \frac{\xi_0}{\omega} + \xi_1 \mathcal{R}_1[\omega] - \xi_2 \mathcal{R}_2[\omega]. \tag{S2.4b}$$

Solving for $\mathcal{R}_1[\omega]$ and $\mathcal{R}_2[\omega]$

$$\mathcal{R}_1[\omega] = \frac{\eta_0}{\omega(\omega + \eta_1)} + \frac{R_1(0)}{\omega + \eta_1}, \tag{S2.5a}$$

$$\mathcal{R}_2[\omega] = \frac{\xi_0}{\omega(\omega + \xi_2)} + \frac{\xi_1 \eta_0}{\omega(\omega + \eta_1)(\omega + \xi_2)} + \frac{\xi_1 R_1(0)}{(\omega + \xi_2)(\omega + \eta_1)} + \frac{R_2(0)}{\omega + \xi_2}. \tag{S2.5b}$$

Taking inverse Laplace transform gives

$$R_1(t) = \frac{\eta_0}{\eta_1}\left(1 - e^{-\eta_1 t}\right) + R_1(0)e^{-\eta_1 t}, \tag{S2.6a}$$

$$R_2(t) = \frac{\xi_0 + \xi_1 \frac{\eta_0}{\eta_1}}{\xi_2} + \left(R_2(0) - \frac{\xi_0}{\xi_2}\right) e^{-\xi_2 t} - \frac{\eta_0 \xi_1}{\eta_1 \xi_2}\left(\frac{\eta_1 e^{-\xi_2 t} - \xi_2 e^{-\eta_1 t}}{\eta_1 - \xi_2}\right) + \xi_1 R_1(0)\left(\frac{e^{-\xi_2 t} - e^{-\eta_1 t}}{\eta_1 - \xi_2}\right). \tag{S2.6b}$$

We can compute the steady-state solution by taking the limit $t \to \infty$:

$$\overline{R_1} = \frac{\eta_0}{\eta_1}, \tag{S2.6c}$$

$$\overline{R_2} = \frac{\xi_0 + \xi_1 \frac{\eta_0}{\eta_1}}{\xi_2}. \tag{S2.6d}$$

The solution for the limiting case when $\eta_1 = \xi_2$ may also be obtained by taking the limit $\eta_1 \to \xi_2$. Another special case, which is more relevant for our discussion in this manuscript, is when the initial conditions are specified as $R_1(0) = 0$ and $R_2(0) = \frac{\xi_0}{\xi_2}$. For this case, we have the following

$$R_1(t) = \frac{\eta_0}{\eta_1}\left(1 - e^{-\eta_1 t}\right), \tag{S2.7a}$$

$$R_2(t) = \frac{\xi_0 + \xi_1 \frac{\eta_0}{\eta_1}}{\xi_2} - \frac{\eta_0 \xi_1}{\eta_1 \xi_2}\left(\frac{\eta_1 e^{-\xi_2 t} - \xi_2 e^{-\eta_1 t}}{\eta_1 - \xi_2}\right). \tag{S2.7b}$$

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

A perfect concerted model is characterized by $k_5 = k_6$. Substituting $k_5 = k_6$ in (S2.1) results in

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^*, \tag{S2.8a}$$

$$\frac{dY^*}{dt} = k_3 Y_T + k_6 Y_T X^* - (k_3 + k_4 + k_6 X_T) Y^*, \tag{S2.8b}$$

with initial condition $(X^*(0), Y^*(0)) = \left(0, \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}\right)$. We note that the form of (S2.8) is same as that of (S2.3), with parameters $\eta_0 = k_1 S X_T$, $\eta_1 = k_1 S + k_2$, $\xi_0 = k_3 Y_T$, $\xi_1 = k_6 Y_T$, and $\xi_2 = k_3 + k_4 + k_6 X_T$. Thus, we can use (S2.7) to get the transient solution

$$X^*(t) = \frac{k_1 S X_T}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right), \tag{S2.9a}$$

$$Y^*(t) = \frac{k_3 + k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} Y_T - \frac{k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} Y_T \frac{(k_1 S + k_2) e^{-(k_3 + k_4 + k_6 X_T)t} - (k_3 + k_4 + k_6 X_T) e^{-(k_1 S + k_2)t}}{k_1 S + k_2 - (k_3 + k_4 + k_6 X_T)}. \tag{S2.9b}$$

For the special case when $k_1 S + k_2 = k_3 + k_4 + k_6 X_T$, we have

$$Y^*(t) = \frac{k_3 + k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} Y_T - \frac{k_6 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T} Y_T e^{-(k_3 + k_4 + k_6 X_T)t}\left(1 + (k_3 + k_4 + k_6 X_T)t\right). \tag{S2.9c}$$

S2–A–b Approximate transient solution using linearization

The ODE system in (S2.1) contains the nonlinear term $X^* Y^*$, which can be linearized around the steady-state solution $(\overline{X^*}, \overline{Y^*})$ as

$$X^* Y^* \approx \overline{Y^*} X^* + \overline{X^*} Y^* - \overline{X^*}\,\overline{Y^*}, \tag{S2.10a}$$

where

$$\overline{X^*} = \frac{k_1 S X_T}{k_1 S + k_2}, \tag{S2.10b}$$

$$\overline{Y^*} = \frac{k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}} Y_T. \tag{S2.10c}$$

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

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^*, \tag{S2.11a}$$

$$\frac{dY^*}{dt} = \left(k_3 Y_T + (k_5 - k_6)\overline{X^*}\,\overline{Y^*}\right) + \left(k_5 Y_T - (k_5 - k_6)\overline{Y^*}\right) X^* - (k_3 + k_4 + k_6 X_T + (k_5 - k_6)\overline{X^*}) Y^*. \tag{S2.11b}$$

These ODEs are similar to those in (S2.3). The parameters are: $\eta_0 = k_1 S X_T$, $\eta_1 = k_1 S + k_2$, $\xi_0 = k_3 Y_T + (k_5 - k_6)\overline{X^*}\ \overline{Y^*}$, $\xi_1 = k_5 Y_T - (k_5 - k_6)\overline{Y^*}$, and $\xi_2 = k_3 + k_4 + k_6 X_T + (k_5 - k_6)\overline{X^*}$. With the initial conditions $(X^*(0), Y^*(0)) = \left(0, \frac{k_3 Y_T}{k_3+k_4+k_6 X_T}\right)$, the solution same as that in (S2.7) and is given by.

$$X^*(t) = \frac{k_1 S X_T}{k_1 S + k_2}\left(1 - e^{-(k_1 S + k_2)t}\right), \tag{S2.12a}$$

$$Y^*(t) = \frac{k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}}{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}} Y_T - \frac{k_1 S X_T}{k_1 S + k_2} \frac{(k_4 k_5 + k_3 k_6 + k_5 k_6 X_T) Y_T}{\left(k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}\right)^2} \times$$
$$\frac{(k_1 S + k_2) e^{-\left(k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}\right)t} - \left(k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}\right) e^{-(k_1 S + k_2)t}}{k_1 S + k_2 - \left(k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}\right)}. \tag{S2.12b}$$

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

S2-B Response time

In this section, we compute the response times for the perfect concerted model and the linearized model. To this end, we recall that the response time for a response $R(t)$ is

$$\mathcal{T}_R = \frac{\int_0^\infty t \left|\overline{R} - R(t)\right| dt}{\int_0^\infty \left|\overline{R} - R(t)\right| dt}, \tag{S2.13}$$

where $\overline{R}$ is the steady-state response. We use this definition to compute the response times for the generic ODE system considered in (S2.3), whose solution is given by (S2.6). We then adapt the solution for our systems of interest, namely, the perfect concerted model and the linearized model.

We begin by computing the response time for $R_1(t)$. The term $\overline{R_1} - R_1(t) > 0$ is

$$\overline{R_1} - R_1(t) = \left(\frac{\eta_0}{\eta_1} - R_1(0)\right) e^{-\eta_1 t}. \tag{S2.14}$$

Note that

$$\int_0^\infty e^{-\eta_1 t} dt = \frac{1}{\eta_1}, \tag{S2.15a}$$

$$\int_0^\infty t\, e^{-\eta_1 t} dt = \frac{1}{\eta_1^2}. \tag{S2.15b}$$

Using these in (S2.13), we get

$$\mathcal{T}_{R_1} = \frac{1}{\eta_1}. \tag{S2.16}$$

It is worth noting that $\mathcal{T}_{R_1}$ does not depend upon the initial condition $R_1(0)$ and is only defined if $R_1(0) \neq \overline{R_1}$.

Next we compute the response time for $R_2(t)$. The term $\overline{R_2} - R_2(t) > 0$ is given by

$$\overline{R_2} - R_2(t) = -\left(R_2(0) - \frac{\xi_0}{\xi_2}\right) e^{-\xi_2 t} + \frac{\eta_0 \xi_1}{\eta_1 \xi_2}\left(\frac{\eta_1 e^{-\xi_2 t} - \xi_2 e^{-\eta_1 t}}{\eta_1 - \xi_2}\right) - \xi_1 R_1(0)\left(\frac{e^{-\xi_2 t} - e^{-\eta_1 t}}{\eta_1 - \xi_2}\right). \tag{S2.17}$$

The integrals of exponential terms in (S2.15) may be used to compute the integrals for the numerator and the denominator of the response time. In particular, we have that

$$\mathcal{T}_{R_2} = \frac{-\frac{R_2(0) - \frac{\xi_0}{\xi_2}}{\xi_2^2} + \frac{\eta_0 \xi_1}{\eta_1 \xi_2}\left(\frac{\frac{\eta_1}{\xi_2^2} - \frac{\xi_2}{\eta_1^2}}{\eta_1 - \xi_2}\right) - \xi_1 R_1(0)\frac{\frac{1}{\xi_2^2} - \frac{1}{\eta_1^2}}{\eta_1 - \xi_2}}{-\frac{R_2(0) - \frac{\xi_0}{\xi_2}}{\xi_2} + \frac{\eta_0 \xi_1}{\eta_1 \xi_2}\left(\frac{\frac{\eta_1}{\xi_2} - \frac{\xi_2}{\eta_1}}{\eta_1 - \xi_2}\right) - \xi_1 R_1(0)\frac{\frac{1}{\xi_2} - \frac{1}{\eta_1}}{\eta_1 - \xi_2}} \tag{S2.18a}$$

$$= \frac{-\frac{R_2(0) - \frac{\xi_0}{\xi_2}}{\xi_2^2} + \frac{\eta_0 \xi_1 (\eta_1^2 + \eta_1 \xi_2 + \xi_2^2)}{\eta_1^3 \xi_2^3} - \frac{\xi_1 R_1(0)(\eta_1 + \xi_2)}{\eta_1^2 \xi_2^2}}{-\frac{R_2(0) - \frac{\xi_0}{\xi_2}}{\xi_2} + \frac{\eta_0 \xi_1 (\eta_1 + \xi_2)}{\eta_1^2 \xi_2^2} - \frac{\xi_1 R_1(0)}{\eta_1 \xi_2}} \tag{S2.18b}$$

$$= \frac{-\eta_1^3 \xi_2 \left(R_2(0) - \frac{\xi_0}{\xi_2}\right) + \eta_0 \xi_1 (\eta_1^2 + \eta_1 \xi_2 + \xi_2^2) - \eta_1 \xi_1 \xi_2 (\eta_1 + \xi_2) R_1(0)}{-\eta_1^3 \xi_2^2 \left(R_2(0) - \frac{\xi_0}{\xi_2}\right) + \eta_0 \eta_1 \xi_1 \xi_2 (\eta_1 + \xi_2) - \eta_1^2 \xi_1 \xi_2^2 R_1(0)}. \tag{S2.18c}$$

We deduce several important insights from the above expression. First, we note that the response time $\mathcal{T}_{R_2}$ depends upon the initial conditions $R_1(0)$ and $R_2(0)$. Second, the dependence on $R_1(0)$ and $R_2(0)$ drops for the special case when $R_1(0) = 0$ and $R_2(0) = \frac{\xi_0}{\xi_2}$. In this case, $\mathcal{T}_{R_2}$ simplifies to

$$\mathcal{T}_{R_2} = \frac{\eta_1^2 + \eta_1\xi_2 + \xi_2^2}{\eta_1\xi_2(\eta_1 + \xi_2)}. \tag{S2.19}$$

Finally, if $R_1(0)$ is taken to be at the steady-state $\overline{R_1} = \frac{\eta_0}{\eta_1}$ and $R_2(0)$ is set as $R_2(0) = \frac{\xi_0}{\xi_2}$, then we get

$$\mathcal{T}_{R_2} = \frac{1}{\xi_2}. \tag{S2.20}$$

With this in mind, it is convenient to express (S2.19) as

$$\mathcal{T}_{R_2} = \frac{1}{\xi_2} + \frac{\xi_2}{\eta_1 + \xi_2}\frac{1}{\eta_1}, \tag{S2.21}$$

where the first-term is the response time if R_1 were at steady-state, and the second term is the time-averaged $\mathcal{T}_{R_1}$.

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

For this case, we can simply adapt the results of (S2.16) and (S2.21).

$$\mathcal{T}_{X^*} = \frac{1}{k_1 S + k_2}, \tag{S2.22a}$$

$$\mathcal{T}_{Y^*} = \frac{1}{k_3 + k_4 + k_6 X_T} + \frac{1}{k_1 S + k_2} \times \frac{k_3 + k_4 + k_6 X_T}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T}. \tag{S2.22b}$$

S2–B–b Response time for the linear approximation

As with the response time for the perfect concerted mechanism, here too we adapt the results of (S2.16) and (S2.21).

$$\mathcal{T}_{X^*} = \frac{1}{k_1 S + k_2}, \tag{S2.23a}$$

$$\mathcal{T}_{Y^*} = \frac{1}{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}} + \frac{1}{k_1 S + k_2} \times \frac{k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}}. \tag{S2.23b}$$

How good is the above approximation of response time? One check is to plug in $k_5 = k_6$ to obtain the approximation for the perfect concerted model for which we have the exact expression of the response time in (S2.22). Indeed, substituting $k_5 = k_6$ in (S2.23) yields

$$\mathcal{T}_{X^*} = \frac{1}{k_1 S + k_2}, \tag{S2.24a}$$

$$\mathcal{T}_{Y^*} = \frac{1}{k_3 + k_4 + k_6 X_T} + \frac{1}{k_1 S + k_2} \times \frac{k_3 + k_4 + k_6 X_T}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T}, \tag{S2.24b}$$

which is exactly same as (S2.22). Thus the linear approximation is exact for the perfect concerted model. This is not surprising because the perfect concerted model is linear by construction.A second check of how good the approximation in (S2.23) is through numerical computation, which is discussed in a later section.

S2–B–c Response time for ratiometric signaling

Ratiometric signaling is the special case where the signaling output does not depend upon the total number of receptors X_T. In the main text, we show that when $k_3 = 0$ and $k_4 = 0$, then the response is independent of X_T ((5)). Here we ask whether setting $k_3 = 0$ and $k_4 = 0$ also result in the response time indepedent from X_T. To this end, we plug these values in the expression of $\mathcal{T}_{Y^*}$ in (S2.23):

$$\mathcal{T}_{Y^*} = \frac{1}{k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}} + \frac{1}{k_1 S + k_2} \times \frac{k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}}{k_1 S + k_2 + k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}}. \tag{S2.25}$$

Clearly, the response time depends upon X_T, thereby establishing that the ratiometric signaling is only applicable for the dose-response. We further ask how X_T affects the response time. To this end, the most convenient limit to check is when the receptor dynamics is fast, i.e., $k_1 S + k_2 \gg k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}$, which gives us

$$\mathcal{T}_{Y^*} \approx \frac{1}{k_6 X_T + (k_5 - k_6)\frac{k_1 S X_T}{k_1 S + k_2}}. \tag{S2.26}$$

Thus, if everything else is constant then increasing X_T decreases the response time. Even when the receptor dynamics is not fast, we can verify this effect by looking at the sign of the derivative of $\mathcal{T}_{Y^*}$ with respect to X^*

$$\frac{d\mathcal{T}_{Y^*}}{dX_T} = -\frac{(k_1 S + k_2)^5 + 2(k_1 S + k_2)^3 (k_2 k_6 X_T + k_1 S k_5 X_T)}{X_T (k_2 k_6 X_T + k_1 S k_5 X_T)\left((k_1 S + k_2)^2 + k_2 k_6 X_T + k_1 S k_5 X_T\right)^2} < 0. \tag{S2.27}$$

Thus increasing X_T speeds up the response. Next, we discuss the numerical method to compute response time which we use to validate our approximations.

S2–B–d Numerical computation of the response time

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

$$\mathcal{T}_{Y^*} = \frac{\int_0^\infty t \left|\overline{Y^*} - Y^*(t)\right| dt}{\int_0^\infty \left|\overline{Y^*} - Y^*(t)\right| dt} \tag{S2.28}$$

is that it can be computer numerically via solution of an augmented ODE system

$$\frac{dX^*}{dt} = k_1 S(X_T - X^*) - k_2 X^*, \tag{S2.29a}$$

$$\frac{dY^*}{dt} = \left(k_3 + k_5 X^*\right)\left(Y_T - Y^*\right) - \left(k_4 + k_6\left(X_T - X^*\right)\right) Y^*, \tag{S2.29b}$$

$$\frac{dV_1}{dt} = \overline{Y^*} - Y^*, \tag{S2.29c}$$

$$\frac{dV_2}{dt} = 1, \tag{S2.29d}$$

$$\frac{dV_3}{dt} = V_1 V_2. \tag{S2.29e}$$

Here $V_1(t)$, $V_2(t)$ and $V_3(t)$ are the augmented states to the original ODE system. The initial conditions are given by

$$\left(X^*(0), Y^*(0), V_1(0), V_2(0), V_3(0)\right) = \left(0, \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}, \frac{\left(k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}\right) Y_T}{k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}} - \frac{k_3 Y_T}{k_3 + k_4 + k_6 X_T}, 0, 0\right). \tag{S2.29f}$$

Note that the state $V_1(t)$ computes the integral in the denominator upto a time horizon t, $V_2(t)$ tracks the time, and $V_3(t)$ computes the numerator up to 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. We can use the approximation of response time in (S2.23) to set a time for the integration.

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 Y^* at time t. Then, we can write the chemical master equation (CME) that describes the time evolution of $P_{m,n}$

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

where $m = 0, 1, \ldots, X_T$ and $n = 0, \ldots, Y_T$ [60,79]. It is often difficult to analytically solve the 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 $P_m(t)$ 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 (S3.2). Multiplying both sides by z^m and summing over m yields

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

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

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

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

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

Using Binomial theorem, the above expression can be written as

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

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

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

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

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

which is a Binomial distribution with parameters X_T and $\frac{k_1 S}{k_1 S + k_2}$ [80]. The stationary moments of this distribution are given by

$$\langle X^* \rangle = \frac{k_1 S X_T}{k_1 S + k_2}, \quad \text{(S3.10a)}$$

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

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

S3-B Moment dynamics

We are specifically concerned with moments of the two-tier model. To this end, we take the well-established approach of using the ODEs that govern the moment dynamics (e.g., see [62, 64]). A generic moment may be written as

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

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 nonlinearity $X^* Y^*$ in these equations, the moment dynamics is not closed in that a lower-order moment depends upon a higher-order moment [62–64]. 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, $k_5 = k_6$. Let us write moment dynamics for first two moments.

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

$$\frac{d\langle Y^* \rangle}{dt} = k_3 Y_T + k_6 Y_T \langle X^* \rangle - (k_3 + k_4 + k_6 X_T)\langle Y^* \rangle, \quad \text{(S3.12b)}$$

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

$$\frac{d\langle X^* Y^* \rangle}{dt} = k_3 Y_T \langle X^* \rangle + k_1 S X_T \langle Y^* \rangle + k_6 Y_T \langle X^{*2} \rangle - (k_1 S + k_2 + k_3 + k_4 + k_6 X_T)\langle X^* Y^* \rangle, \quad \text{(S3.12d)}$$

$$\frac{d\langle Y^{*2} \rangle}{dt} = k_3 Y_T + k_6 Y_T \langle X^* \rangle + (k_3(2Y_T - 1) + k_4 + k_6 X_T)\langle Y^* \rangle + 2k_6(Y_T - 1)\langle X^* Y^* \rangle - 2(k_3 + k_4 + k_6 X_T)\langle Y^{*2} \rangle. \quad \text{(S3.12e)}$$

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

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

$$\langle Y^* \rangle = \frac{k_3 + k_6 \langle X^* \rangle}{k_3 + k_4 + k_6 X_T} Y_T = \frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 X_T} Y_T. \tag{S3.13b}$$

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

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

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

$$\langle X^* Y^* \rangle = \frac{k_1 S k_2 k_6 X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k_6 X_T)} + \left(\frac{k_1 S}{k_1 S + k_2} X_T \right) \left(\frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 X_T} Y_T \right). \tag{S3.15}$$

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

$$\langle Y^{*2} \rangle = \frac{k_3 Y_T}{2(k_3 + k_4 + k_6 X_T)} + \frac{k_6 Y_T \langle X^* \rangle}{2(k_3 + k_4 + k_6 X_T)} + \frac{(k_3(2Y_T - 1) + k_4 + k_6 X_T) \langle Y^* \rangle}{2(k_3 + k_4 + k_6 X_T)} + \frac{2k_6(Y_T - 1)) \langle X^* Y^* \rangle}{2(k_3 + k_4 + k_6 X_T)} \tag{S3.16a}$$

$$= \frac{k_3 Y_T + k_4 + k_6 X_T}{k_3 + k_4 + k_6 X_T} \langle Y^* \rangle + \frac{k_6 (Y_T - 1) \langle X^* Y^* \rangle}{k_3 + k_4 + k_6 X_T}. \tag{S3.16b}$$

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

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

the centered cross moment is

$$\langle X^* Y^* \rangle - \langle X^* \rangle \langle Y^* \rangle = \frac{k_1 S k_2 k_6 X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k_6 X_T)}, \tag{S3.18}$$

and the variance of Y^* is

$$\langle Y^{*2} \rangle - \langle Y^* \rangle^2 = \left(\frac{k_3 Y_T + k_4 + k_6 X_T}{k_3 + k_4 + k_6 X_T} \right) \left(\frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 X_T} Y_T \right) + \frac{k_6 (Y_T - 1)}{k_3 + k_4 + k_6 X_T} \times$$
$$\left(\frac{k_1 S k_2 k_6 X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k_6 X_T)} + \frac{k_1 S X_T}{k_1 S + k_2} \frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k X_T} Y_T \right) - \left(\frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 X_T} Y_T \right)^2. \tag{S3.19}$$

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

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

and

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

$$= \left(\frac{k_3 Y_T + k_4 + k_6 X_T}{k_3 + k_4 + k_6 X_T} \right) \left(\frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 X_T} Y_T \right) \frac{1}{Y_T^2} \left(\frac{k_3 + k_4 + k_6 X_T}{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T} \right)^2$$
$$+ \frac{1}{Y_T^2} \left(\frac{k_3 + k_4 + k_6 X_T}{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T} \right)^2 \frac{k_6 (Y_T - 1)}{k_3 + k_4 + k_6 X_T} \left(\frac{k_1 S k_2 k_6 X_T Y_T}{(k_1 S + k_2)^2 (k_1 S + k_2 + k_3 + k_4 + k_6 X_T)} \right)$$
$$+ \frac{1}{Y_T^2} \left(\frac{k_3 + k_4 + k_6 X_T}{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T} \right)^2 \frac{k_6 (Y_T - 1)}{k_3 + k_4 + k_6 X_T} \left(\frac{k_1 S X_T}{k_1 S + k_2} \frac{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T}{k_3 + k_4 + k_6 X_T} Y_T \right) - 1. \tag{S3.22}$$

On simplifying, we get

$$CV^2_{Y^*} = \frac{1}{Y_T}\frac{k_3 Y_T + k_4 + k_6 X_T}{k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T} + \frac{Y_T - 1}{Y_T}\frac{k_6 \left(k_3 + k_4 + k_6 X_T\right)\left(k_1 S k_2 k_6 X_T\right)}{(k_1 S + k_2)^2 \left(k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T\right)^2 \left(k_1 S + k_2 + k_3 + k_4 + k_6 X_T\right)}$$
$$+ \frac{Y_T - 1}{Y_T}\frac{k_6 k_1 S X_T}{\left(k_3 + k_6 \frac{k_1 S}{k_1 S + k_2} X_T\right)(k_1 S + k_2)}. \tag{S3.23}$$

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

$$\frac{dX^*}{dt} = k_1 S X_T - (k_1 S + k_2) X^*, \tag{S3.24a}$$
$$\frac{d\left\langle Y^* \right\rangle}{dt} = k_3 Y_T + k_6 Y_T X^* - (k_3 + k_4 + k X_T)\left\langle Y^* \right\rangle, \tag{S3.24b}$$
$$\frac{d\left\langle Y^{*2} \right\rangle}{dt} = k_3 Y_T + k_6 Y_T X^* + (k_3(2Y_T - 1) + k_4 + k_6 X_T)\left\langle Y^* \right\rangle + 2k_6(Y_T - 1)X^*\left\langle Y^* \right\rangle - 2(k_3 + k_4 + k_6 X_T)\left\langle Y^{*2} \right\rangle. \tag{S3.24c}$$

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

$$CV^2_{Y^*}\Big|_{act./deact.} = \frac{1}{Y_T}\frac{k_4 + k_6 X_T \frac{k_2}{k_1 S + k_2}}{k_3 + k_6 X_T \frac{k_2}{k_1 S + k_2}}. \tag{S3.25}$$

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^*:

$$CV^2_{Y^*} - CV^2_{Y^*}\Big|_{act./deact.} = \frac{Y_T - 1}{Y_T}\frac{(k_3 + k_4 + k_6 X_T)(k_1 S k_2 k^2 X_T)}{(k_1 S + k_2 + k_3 + k_4 + k_6 X_T)(k_1 S + k_2)^2 \left(k_3 + k_6 X_T \frac{k_1 S}{k_1 S + k_2}\right)^2}. \tag{S3.26}$$

We expect that the term on the right hand side should have contribution from $CV^2_{X^*}$, which is time-averaged. Recall (S2.22) that $k_1 S + k_2$ is response time of the receptor and that $k_3 + k_4 + k_6 X_T$ is response time of the switch if the receptor dynamics is fast. Thus, $\frac{k_3 + k_4 + k_6 X_T}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T}$ can be interpreted as the timescale averaging. Therefore, we write

$$CV^2_{Y^*} = \underbrace{\frac{1}{Y_T}\frac{k_4 + k_6 X_T \frac{k_2}{k_1 S + k_2}}{k_3 + k_6 X_T \frac{k_1 S}{k_1 S + k_2}}}_{\text{contribution from act./deact. of } Y^*} + \underbrace{\frac{k_3 + k_4 + k_6 X_T}{k_1 S + k_2 + k_3 + k_4 + k_6 X_T}}_{\text{time-averaging}} \underbrace{\frac{Y_T - 1}{Y_T}\left(\frac{k_6 X_T \frac{k_1 S}{k_1 S + k_2}}{k_3 + k_6 X_T \frac{k_1 S}{k_1 S + k_2}}\right)^2}_{\text{coupling}} CV^2_{X^*}, \tag{S3.27}$$

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 [81]. 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.

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

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

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.

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 $CV^2_{Y^*}$ can be obtained in the same manner as done for the perfect concerted model and is given by

$$CV^2_{Y^*} \approx \underbrace{\frac{1}{Y_T} \frac{k_4 + k_6 X_T \frac{k_2}{k_1 S + k_2}}{k_3 + k_5 X_T \frac{k_1 S}{k_1 S + k_2}}}_{\text{contribution from act./deact. of } Y^*} + \underbrace{\frac{k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}{k_1 S + k_2 + k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}}}_{\text{time-averaging}} \times \underbrace{\frac{\left(\frac{k_1 S X_T}{k_1 S + k_2}\right)^2 (k_4 k_5 + k_6 (k_3 + k_5 X_T))^2}{\left(k_3 + k_5 \frac{k_1 S X_T}{k_1 S + k_2}\right)^2 \left(k_3 + k_4 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}\right)^2}}_{\text{coupling}} CV^2_{X^*}. \tag{S3.31}$$

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 $k_5 = k_6$ shows that the noise approximation above differs from (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.

S3–B–c Coefficient variation squared for ratiometric signaling

For ratiometric signaling, in which the steady-state response does not depend upon the total number of receptors X_T, we need $k_3 = 0$ and $k_4 = 0$. Substituting these in the expression of $CV^2_{Y^*}$ in (S3.31), we get

$$CV^2_{Y^*} \approx \frac{1}{Y_T} \frac{k_2 k_6}{k_1 S k_5} + \frac{k_6^2}{\left(k_1 S + k_2 + k_5 \frac{k_1 S X_T}{k_1 S + k_2} + k_6 \frac{k_2 X_T}{k_1 S + k_2}\right)\left(\frac{k_1 S k_5}{k_1 S + k_2} + \frac{k_2 k_6}{k_1 S + k_2}\right)} \frac{k_2}{k_1 S}. \tag{S3.32}$$

Thus, increasing X_T decreases overall noise because X_T increases the denominator terms in the above above formula. Next, we provide exact computation of moments using a semi-analytical approach.

S3–B–d 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 [65]. Another closely related method is the method of conditional moments described in [82].

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

$$b_i = \begin{cases} 1, & X^* = i, \\ 0, & \text{otherwise}. \end{cases} \tag{S3.33a}$$

It then follows that

$$\sum_{i=0}^{X_T} b_i = 1, \quad b_i b_j = 0,\ i \neq j, \quad b_i^2 = b_i. \tag{S3.33b}$$

We now recast our original model in the new state-space $\begin{bmatrix} b_0 & b_1 & \ldots & b_{X_T} & Y^* \end{bmatrix}^\top$. 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

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} 0 & b_0 & b_1 & \dots & b_{X_T-2} & b_{X_T-1} + b_{X_T} & Y^* \end{bmatrix}^\top . \tag{S3.34a}$$

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

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 + b_1 & b_2 & b_3 & \dots & b_{X_T} & 0 & Y^* \end{bmatrix}^\top . \tag{S3.34b}$$

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

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* + 1 \end{bmatrix}^\top . \tag{S3.34c}$$

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

$$\begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* \end{bmatrix}^\top \mapsto \begin{bmatrix} b_0 & b_1 & b_2 & \dots & b_{X_T-1} & b_{X_T} & Y^* - 1 \end{bmatrix}^\top . \tag{S3.34d}$$

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 $\langle b_i \rangle$.

$$\frac{d \langle b_0 \rangle}{dt} = -k_1 X_T \langle b_0 \rangle + k_2 \langle b_1 \rangle , \tag{S3.35a}$$

$$\frac{d \langle b_i \rangle}{dt} = k_1 (X_T - i + 1) \langle b_{i-1} \rangle - k_1 (X_T - i) \langle b_i \rangle + k_2 (i + 1) \langle b_{i+1} \rangle - k_2 i \langle b_i \rangle , 1 \leq i \leq X_T - 1, \tag{S3.35b}$$

$$\frac{d \langle b_{X_T} \rangle}{dt} = k_1 \langle b_{X_T-1} \rangle - k_2 X_T \langle b_{X_T} \rangle . \tag{S3.35c}$$

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 (S3.8). Therefore, the solution to these ODEs is

$$\langle b_i \rangle = \binom{X_T}{i} \left(\frac{k_1}{k_1 + k_2} \left(1 - e^{-(k_1+k_2)t} \right) \right)^i \left(1 - \frac{k_1}{k_1 + k_2} \left(1 - e^{-(k_1+k_2)t} \right) \right)^{X_T - i} . \tag{S3.36}$$

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

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

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

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

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

$$\frac{d\left\langle b_0 Y^{*2}\right\rangle}{dt} = -k_1 X_T \left\langle b_0 Y^{*2}\right\rangle + k_2 \left\langle b_1 Y^{*2}\right\rangle + k_3 Y_T \left\langle b_0\right\rangle + (-k_3 + k_4 + k_6 X_T + 2k_3 Y_T)\left\langle b_0 Y^*\right\rangle - (2k_3 + 2k_4 + 2k_6 X_T)\left\langle b_0 Y^{*2}\right\rangle, \tag{S3.38a}$$

$$\begin{aligned}\frac{d\left\langle b_i Y^{*2}\right\rangle}{dt} &= k_1 (X_T - i + 1)\left\langle b_{i-1} Y^{*2}\right\rangle - k_1 (X_T - i)\left\langle b_i Y^{*2}\right\rangle + k_2(i+1)\left\langle b_{i+1} Y^{*2}\right\rangle - k_2 i \left\langle b_i Y^{*2}\right\rangle \\ &+ (k_3 Y_T + k_5 Y_T i)\left\langle b_i\right\rangle + (-k_3 - k_5 i + k_4 + k_6 X_T - k_6 i + 2k_3 Y_T + 2k_5 Y_T i)\left\langle b_i Y^*\right\rangle \\ &- (2k_3 + 2k_5 i + 2k_4 + 2k_6 X_T - 2k_6 i)\left\langle b_i Y^{*2}\right\rangle, \quad 1 \le i \le X_T - 1,\end{aligned} \tag{S3.38b}$$

$$\begin{aligned}\frac{d\left\langle b_{X_T} Y^{*2}\right\rangle}{dt} &= k_1 \left\langle b_{X_T-1} Y^{*2}\right\rangle - k_2 X_T \left\langle b_{X_T} Y^{*2}\right\rangle + (k_3 Y_T + k_5 Y_T X_T)\left\langle b_i\right\rangle + (-k_3 - k_5 X_T + k_4 + 2k_3 Y_T + 2k_5 Y_T X_T)\left\langle b_i Y^*\right\rangle \\ &- (2k_3 + 2k_5 X_T + 2k_4)\left\langle b_i Y^{*2}\right\rangle.\end{aligned} \tag{S3.38c}$$

These ODEs require initial condition to compute transient moments which we discuss 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 $\left\langle b_i Y^{*2}\right\rangle = \langle b_i\rangle\langle Y^* 2\rangle$. 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

$$\langle b_0 Y^*\rangle = \frac{k_3}{k_3 + k_4 + k_6 X_T} Y_T, \quad \left\langle b_0 Y^{*2}\right\rangle = \frac{k_3^2 Y_T^2 + k_3(k_4 + k_6 X_T) Y_T}{(k_3 + k_4 + k_6 X_T)^2}. \tag{S3.39}$$

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

$$\frac{d\mu_0}{dt} = M_0 \mu_0, \tag{S3.40}$$

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

$$M_0 \mu_0 = 0, \tag{S3.41}$$

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

$$\tilde{M}_0 \tilde{\mu_0} + c = 0, \tag{S3.42}$$

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

$$\mu_1 = \begin{bmatrix}\langle b_0\rangle & \dots & \langle b_{X_T}\rangle & \langle b_0 Y^*\rangle & \dots & \langle b_{X_T} Y^*\rangle.\end{bmatrix} \tag{S3.43}$$

The corresponding ODE system is then

$$\frac{d\mu_1}{dt} = \begin{bmatrix} M_0 & 0 \\ M_{10} & M_{11}\end{bmatrix}\mu_1 \tag{S3.44}$$

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

$$\mu_2 = \begin{bmatrix}\langle b_0\rangle & \dots & \langle b_{X_T}\rangle & \langle b_0 Y^*\rangle & \dots & \langle b_{X_T} Y^*\rangle & \left\langle b_0 Y^{*2}\right\rangle & \dots & \left\langle b_{X_T} Y^{*2}\right\rangle.\end{bmatrix} \tag{S3.45}$$

Then we can write the ODE system:

$$\frac{d\mu_2}{dt} = \begin{bmatrix} M_0 & 0 & 0 \\ M_{10} & M_{11} & 0 \\ M_{20} & M_{21} & M_{22}\end{bmatrix}\mu_2. \tag{S3.46}$$

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

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

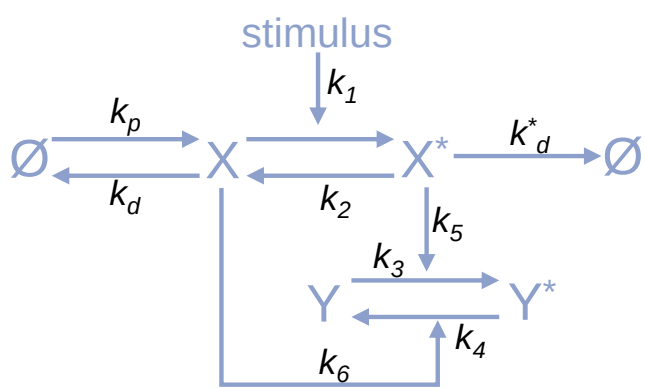

Figure S4.1: Concerted mechanism with receptor production and degradation.

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.

$$\frac{dX}{dt} = k_p - k_d X - k_1 S X + k_2 X^*, \tag{S4.4a}$$

$$\frac{dX^*}{dt} = k_1 S X - k_2 X^* - k_d^* X^*. \tag{S4.4b}$$

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

$$\omega \mathcal{X}[\omega] - \frac{k_p}{k_d} = \frac{k_p}{\omega} - (k_1 S + k_d)\mathcal{X}[\omega] + k_2 \mathcal{X}^*[\omega], \tag{S4.5a}$$

$$\omega \mathcal{X}^*[\omega] = k_1 S \mathcal{X}[\omega] - (k_2 + k_d^*)\mathcal{X}^*[\omega]. \tag{S4.5b}$$

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

$$\mathcal{X}[\omega] = \frac{\frac{k_p}{k_d}(\omega + k_2 + k_d^*) + k_p}{\omega^2 + 2\zeta\kappa\omega + \kappa^2} + \frac{k_p(k_2 + k_d^*)}{\omega(\omega^2 + 2\zeta\kappa\omega + \kappa^2)}, \tag{S4.6a}$$

$$\mathcal{X}^*[\omega] = \frac{\frac{k_1 k_p S}{k_d}}{\omega^2 + 2\zeta\kappa\omega + \kappa^2} + \frac{k_1 k_p S}{\omega(\omega^2 + 2\zeta\kappa\omega + \kappa^2)}, \tag{S4.6b}$$

where we have used the following notation

$$\zeta = \frac{k_1 S + k_2 + k_d + k_d^*}{2\sqrt{(k_1 S + k_d)(k_2 + k_d^*) - k_1 S k_2}}, \tag{S4.7a}$$

$$\kappa = \sqrt{(k_1 S + k_d)(k_2 + k_d^*) - k_1 S k_2}. \tag{S4.7b}$$

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

$$\omega_{1,2} = \kappa(-\zeta \pm \sqrt{\zeta^2 - 1}). \tag{S4.8a}$$

The following usual relations hold for ω_1 and ω_2:

$$\omega_1 + \omega_2 = -2\kappa\zeta, \tag{S4.8b}$$

$$\omega_1\omega_2 = \kappa^2, \tag{S4.8c}$$

$$\omega_1 - \omega_2 = 2\kappa\sqrt{\zeta^2 - 1}. \tag{S4.8d}$$

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

$$X^*(t) = \frac{k_1 k_p S}{\kappa^2} + c_1^* e^{\omega_1 t} + c_2^* e^{\omega_2 t}. \tag{S4.9a}$$

Here the terms c_1^* and c_2^* are

$$c_1^* = \frac{k_1 S}{2\kappa\sqrt{\zeta^2 - 1}}\left(\frac{k_p}{k_d} + \frac{k_p\omega_2}{\kappa^2}\right), c_2^* = -\frac{k_1 S}{2\kappa\sqrt{\zeta^2 - 1}}\left(\frac{k_p}{k_d} + \frac{k_p\omega_1}{\kappa^2}\right). \tag{S4.9b}$$

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

$$X(t) = \frac{1}{k_1 S}\frac{dX^*}{dt} + \frac{k_2 + k_d^*}{k_1 S}X^* \tag{S4.10a}$$

$$= \frac{1}{k_1 S}\left(c_1^*\omega_1 e^{\omega_1 t} + c_2^*\omega_2 e^{\omega_2 t}\right) + \frac{k_2 + k_d^*}{k_1 S}\left(\frac{k_1 k_p S}{\kappa^2} + c_1^* e^{\omega_1 t} + c_2^* e^{\omega_2 t}\right) \tag{S4.10b}$$

$$= \frac{k_p(k_2 + k_d^*)}{\kappa^2} + \left(\frac{c_1^*\omega_1 + c_1^*(k_2 + k_d^*)}{k_1 S}\right)e^{\omega_1 t} + \left(\frac{c_2^*\omega_2 + c_2^*(k_2 + k_d^*)}{k_1 S}\right)e^{\omega_2 t}. \tag{S4.10c}$$

Having determined these solutions, we next provide a lower bound on ζ. It is worth noting that $\zeta > 1$ implies that the roots $\omega_{1,2}$ are real.

A lower bound for ζ. The parameter ζ defined in (S4.7) is always greater than one, regardless of the choice of parameters. To see this, we look at ζ^2

$$\zeta^2 = \frac{\left(k_1 S + k_2 + k_d + k_d^*\right)^2}{4\left(k_2 k_d + k_d k_d^* + k_1 k_d^* S\right)} \tag{S4.11a}$$

$$= \frac{(k_1 S + k_d)^2 + (k_2 + k_d^*)^2 + 2(k_1 S + k_d)(k_2 + k_d^*)}{4\left((k_1 S + k_d)(k_2 + k_d^*) - k_1 S k_2\right)}. \tag{S4.11b}$$

This implies that

$$(k_1 S + k_d)^2 + (k_2 + k_d^*)^2 + (2 - 4\zeta^2)(k_1 S + k_d)(k_2 + k_d^*) + 4\zeta^2 k_1 S k_2 = 0 \tag{S4.11c}$$

$$\implies \left((k_1 S + k_d) - (k_2 + k_d^*)\right)^2 + (4 - 4\zeta^2)(k_1 S + k_d)(k_2 + k_d^*) + 4\zeta^2 k_1 S k_2 = 0. \tag{S4.11d}$$

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 (S4.8a) are negative and satisfy

$$\omega_2 < \omega_1 < 0, \quad |\omega_1| < |\omega_2|. \tag{S4.12}$$

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. Because $|\omega_2| > |\omega_1|$, we refer to the timescale set by ω_2 as fast timescale and the one set by ω_1 as the slow timescale. 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 (S4.9a). At $t = 0$, the trajectory begins from $X^*(0) = 0$. Consider a scenario where $\zeta \to 1$, implying that $\omega_1 \approx \omega_2$

$$X^*(t) \approx \frac{k_1 k_p S}{\kappa^2} - \frac{k_1 k_p S}{\kappa^2} e^{-\kappa t}, \tag{S4.13}$$

which increases over time to reach the steady-state $\frac{k_1 k_p S}{\kappa^2}$. Suppose that ζ is now increased. The terms c_1^* and c_2^* relax with different timescales. Specifically, c_1^* relaxes at a slower timescale than c_2^*. Because $c_1^* + c_2^* = -\frac{k_1 k_p S}{\kappa^2} < 0$, at least one of has to be positive. 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

$$X^*(t) \approx \frac{k_1 k_p S}{\kappa^2} + c_1^* + c_2^* e^{-(\kappa\zeta - \kappa\sqrt{\zeta^2-1})t} = -c_2^* + c_2^* e^{-(\kappa\zeta - \kappa\sqrt{\zeta^2-1})t}. \tag{S4.14}$$

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

$$X^*(t') \approx -c_2^* > 0. \tag{S4.15}$$

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:

$$c_1^* + c_2^* = -\frac{k_1 k_p S}{\kappa^2}, \quad c_2^* < 0, \quad |c_2^*| > \frac{k_1 k_p S}{\kappa^2}. \tag{S4.16}$$

We substituted the values of ζ and κ from (S4.7) and used to symbolic solver to solve the above inequalities. We obtain that the following should be satisfied:

$$0 < k_d < k_d^*. \tag{S4.17}$$

We get 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

$$\left(\frac{c_1^*\omega_1 + c_1^*(k_2 + k_d^*)}{k_1 S}\right) + \left(\frac{c_2^*\omega_2 + c_2^*(k_2 + k_d^*)}{k_1 S}\right) = \frac{k_p}{k_d} - \frac{k_p(k_2 + k_d^*)}{\kappa^2}, \quad \left(\frac{c_2^*\omega_2 + c_2^*(k_2 + k_d^*)}{k_1 S}\right) > \frac{k_p}{k_d} - \frac{k_p(k_2 + k_d^*)}{\kappa^2} > 0, \tag{S4.18}$$

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

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.

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 (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

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

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

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

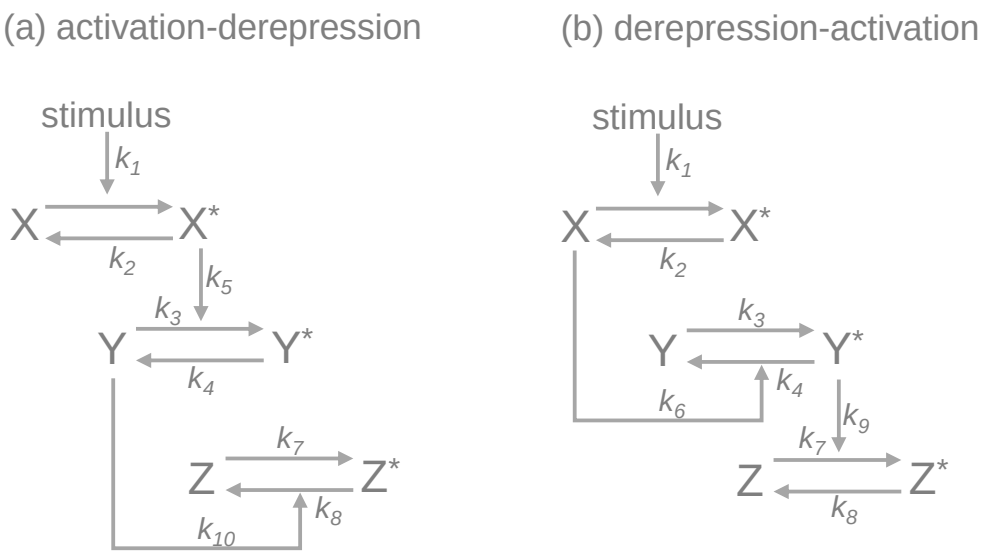

Figure S5.1: Three tier cascades with alternating activation and derepression mechanisms

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 (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.