Supplementary material for

Response dynamics of phosphorelays suggest their potential utility in cell signaling

Attila Csikász-Nagy, Luca Cardelli and Orkun S. Soyer

Contents

1. Extended model, assuming the protein in the first layer to be bifunctional.
2. Analytical steady state solution for *L2p*.
3. Analysis of "absolute concentration robustness" in the last layer, using the extended model with a bifunctional protein.
4. Script for the basic and extended models, executable in XPPAUT and Oscill8 packages.
5. Stochastic version of the basic model
6. Script for the stochastic version of the model, executable in the SPiM package
7. Interaction partners of the *Bacillus subtilis* sporulation phosphorelay proteins
8. Supplementary Figures 1-11.

1. Extended model, assuming the protein in the first layer to be bifunctional.

As discussed in the main text, the basic phosphorelay model (Eq. 1) can be extended by assuming the protein in the first layer to have dephosphorylation activity towards the protein in the last layer. In this case, and assuming enzymatic reactions for such dephosphorylation activity we can write the ordinary differential equations describing a 4-layer system as (using the *L* notation);

$$\frac{dL1p}{dt} = k_1 \cdot (L1_{tot} - L1L4p - L1p) - k_2 \cdot L1p \cdot (L2_{tot} - L2p) - k_{22} \cdot L1p$$

$$\frac{dL2p}{dt} = (k_2 \cdot L1p + k_{23}) \cdot (L2_{tot} - L2p) - k_3 \cdot L2p \cdot (L3_{tot} - L3p) - k_{32} \cdot L2p$$

$$\frac{dL3p}{dt} = (k_3 \cdot L2p + k_{33}) \cdot (L3_{tot} - L3p) - k_4 \cdot L3p \cdot (L4_{tot} - L1L4p - L4p) - k_{42} \cdot L3p \qquad \text{Eq. S1}$$

$$\frac{dL4p}{dt} = (k_4 \cdot L3p + k_{43}) \cdot (L4_{tot} - L1L4p - L4p) + k_{k51r} \cdot L1L4p - k_5 \cdot L4p - k_{51} \cdot L4p \cdot (L1_{tot} - L1L4p - L1p)$$

$$\frac{dL1L4p}{dt} = k_{51} \cdot L4p \cdot (L1_{tot} - L1L4p - L1p) - (k_{k51r} + k_{52}) \cdot L1L4p$$

where *Li* and *Lip* denote the unphosphorylated and phosphorylated form of the protein at the *i*'th layer and Li_{tot} denotes total protein concentration at that layer. *L1L4p* stands for the complex between the bifunctional enzyme of the first layer and the phosphorylated protein from the last layer. Note that the Michealis Menten coefficient K_m is given by $(k_{52} + k_{51r})/k_{51}$. For the analysis shown on Figure 2, K_m is varied by changing k_{51}.

2. Analytical steady state solution for *L2p*.

The following analytical solution for *L2p* is derived for steady states, using the base model (as shown in Eq. 1 of the main text) and by setting self-phosphorylation (k_{23}, k_{33}, k_{43}) and self-dephosphorylation rates (k_{22}, k_{32}, k_{42}) to zero;

$$L2p^3 \cdot A + L2p^2 \cdot B + L2p \cdot C + D = 0 \qquad \text{Eq. S2}$$

where;

$$A = k_5 k_4 k_3 k_2^2 L4_{tot} L3_{tot} - k_4 k_3 k_2^2 k_1 L1_{tot} L3_{tot} - k_5 k_3 k_2^2 k_1 L1_{tot}$$

$$B = -k_5 k_4 k_2^2 k_1 L1_{tot} L4_{tot} + k_4 k_2^2 k_1^2 L1_{tot}^2 - 2k_5 k_4 k_3 k_2^2 L2_{tot} L4_{tot} L3_{tot} + 2k_4 k_3 k_2^2 k_1 L1_{tot} L2_{tot} L3_{tot} + 2k_5 k_3 k_2^2 k_1 L1_{tot} L2_{tot}$$
$$-2k_5 k_4 k_3 k_2 k_1 L4_{tot} L3_{tot} + k_4 k_3 k_2 k_1^2 L1_{tot} L3_{tot} + k_5 k_3 k_2 k_1^2 L1_{tot}$$

$$C = 2k_5 k_4 k_2^2 k_1 L1_{tot} L4_{tot} L2_{tot} - 2k_4 k_2^2 k_1^2 L1_{tot}^2 L2_{tot} + k_5 k_4 k_2 k_1^2 L1_{tot} L4_{tot} + k_5 k_4 k_3 k_2^2 L2_{tot}^2 L4_{tot} L3_{tot} + 2k_5 k_4 k_3 k_2 k_1 L2_{tot} L4_{tot} L3_{tot}$$
$$-k_4 k_3 k_2^2 k_1 L1_{tot} L2_{tot}^2 L3_{tot} - k_5 k_3 k_2^2 k_1 L1_{tot} L2_{tot}^2 + k_5 k_4 k_3 k_1^2 L4_{tot} L3_{tot} - k_4 k_3 k_2 k_1^2 L1_{tot} L2_{tot} L3_{tot} - k_5 k_3 k_2 k_1^2 L1_{tot} L2_{tot}$$

$$D = k_4 k_2^1 k_1^2 L1_{tot}^2 L2_{tot}^2 - k_5 k_4 k_2 k_1^2 L1_{tot} L4_{tot} L2_{tot} - k_5 k_4 k_2^2 k_1 L1_{tot} L4_{tot} L2_{tot}^2$$

3. Analysis of "absolute concentration robustness" in the last layer, using the extended model with a bifunctional protein.

To test whether the phosphorylated form of the protein in the final layer of a phosphorelay can display robustness against variations in the concentrations of other proteins in the relay (i.e. absolute concentration robustness), we solved Eq. S1 for steady state expression for *L4p* (i.e. system output). At steady state the phosphate flow through the system is constant, thus, we can take the rate of incoming phosphate groups equal to the rate of phosphate loss from the final layer. Using Eq. S1, we can formulate this mathematically as;

$$k_1 \cdot (L1_{tot} - L1L4p - L1p) = k_5 \cdot L4p + k_{51} \cdot L4p \cdot (L1_{tot} - L1L4p - L1p) \qquad \text{Eq. S3}$$

solving for *L4p*, we get;

$$L4p = \frac{k_1 \cdot (L1_{tot} - L1L4p - L1p)}{k_5 + k_{51} \cdot (L1_{tot} - L1L4p - L1p)} \qquad \text{Eq. S4}$$

When k_5 is negligible compared to k_{51}, that is when self-dephosphorylation rate of the last layer is small compared to dephosphorylation by the bifunctional enzyme, Equation S4 would simplify to $L4p = k_1 / k_{51}$. Hence, under this condition, the concentration of phosphorylated protein at the last layer of the relay (i.e. system output) would depend only on the signal and efficiency of the bifunctional enzyme and would not depend on any of the protein concentrations in the system. It would show absolute concentration robustness.

To provide further support to this mathematical finding, we calculated input – response curves for models with different enzymatic efficiency of *L1* on *L4p*. As shown on Figure S6, absolute concentration robustness holds when the enzyme is very effective (K_m = 0.01) – but with such effective enzyme the steady state activity of *L4p* is rather low. This might be used in phosphorelays that should have a robust low output and that should transiently respond for input changes. When *L1* works as a slow enzyme (K_m = 1) then the absolute robustness does not hold (second column on Fig. S6), unless the background phosphate loss rate is also greatly reduced (third column on Fig. S6). In this latter case, *L4p* is still dependent on changes in the total concentration of *L4* due to saturation effects.

4. Scripts for the base and extended models, executable in XPPAUT and Oscill8 packages.

XPPAUT (http://www.math.pitt.edu/~bard/xpp/xpp.html) and OSCILL8 (http://sourceforge.net/projects/oscill8) executable script for the phosphorelay model as shown in Eq. S1. The script allows reducing the model to that given by Eq. 1 of the main text, simply by setting appropriate parameters to zero (see comments indicated by the dash symbol).

```
# 4 layer relay with possible crosstalk at all levels
# bifunctional enzyme can be turned on by changing k51 from 0
# crosstalk can be induced by changing  rates:
```

```
# k22, k32, k42 for inhibition or k23, k33, k43 for activation

# input parameter (this is varied for most analysis)
p k1=1

# Relay processing rates (numbers correspond to the order of reactions counted from the top)
p k2=1, k3=1, k4=1, k5=1

# Inhibiting crosstalk by dephosphorylation – "talk out"
p k22=0, k32=0, k42=0

# activating crosstalk by phosphorylation – "talk in"
p k23=0, k33=0, k43=0

# total protein levels
p L1tot=10, L2tot=10, L3tot=10, L4tot=10

# bifunctional enzyme rates (in order: binding of L1 and L4p, unbinding, product release)
p k51=0, k51r=1, k52=1


# equations
L1p' = k1*(L1tot-L1p-L4pL1) - k2*L1p*(L2tot-L2p) - k22*L1p
L2p' = (k2*L1p + k23)*(L2tot-L2p) - k3*L2p*(L3tot-L3p) - k32*L2p
L3p' = (k3*L2p + k33)*(L3tot-L3p) - k4*L3p*(L4tot-L4p-L4pL1) - k42*L3p
L4p' =(k4*L3p + k43)*(L4tot-L4p-L4pL1) - k5*L4p - k51*L4p*(L1tot-L1p-L4pL1) + k51r*L4pL1
L4pL1'= k51*L4p*(L1tot-L1p-L4pL1) - (k51r+k52)*L4pL1

done
```

5. Stochastic version of the basic model.

The reaction scheme used for stochastic simulation with clear input (no input noise) is as follows;

$$
\begin{aligned}
&L1 + B \xrightarrow{0.01} L1p + B \\
&L1p + L2 \xrightarrow{0.1} L1 + L2p \\
&L2p + L3 \xrightarrow{0.1} L2 + L3p \\
&L3p + L4 \xrightarrow{0.1} L3 + L4p \\
&L4p \xrightarrow{1.0} L4 + Phos
\end{aligned}
\qquad \text{Eq. S5}
$$

where B stands for a signaling molecule and the numbers on the arrows indicate the reaction propensities of dimension s^{-1}. The latter and the total number of molecules for stochastic simulation, are chosen so to match the mass action rates used in the ODE model. The first reaction propensity is set low (0.01) to diminish the impact of each discrete B increment (see below). The quantity of B and the rate of the first reaction together represent a certain level of influence of the input on the system.

The simulations start with 100 protein copies for each layer. A simulation runs sequentially with B incremented from 0 by 1, allowing the system to run for a certain time (until steady state) at each B level, over which time various statistics are taken of the fluctuations of the various species in the system. The simulation stops when B reaches 200 (twice the Li levels). This allows us explore 200 B levels in a single simulation and is equal to running separate simulations at each fixed B level, until steady state.

To model noisy input, we introduce an additional birth-death process for B;

$$A \xrightarrow{1.0} A + B$$
$$B \xrightarrow{1.0}$$

Eq. S6

where A stands for another pool of signaling molecules (or process) that give rise to B, which is then degraded. In this case, the simulation involves sequentially incrementing A from 0 by 1; that is, the simulation is now driven by A instead of B. Because of the birth-death type process and the stochastic nature of the simulation, the B level fluctuates with a Poisson distribution with instantaneous mean given by the A level. The B level then forms the (noisy) input into the system given in Eq. S5.

6. Script for the stochastic model, executable in SPiM package.

SPiM (http://research.microsoft.com/en-us/projects/spim/) executable script for the stochastic model of the 4-layer phosphorelay, embedding noisy input.

```
(* 4-layer + noise *)
directive sample 25000.0
directive plot L4p();L3p();L2p();L1p();B();A()
val poissrate = 1.0
val uRate = 1.0 val bRate = 0.1 val stimRate = 0.01
new L1tot:chan new k0@stimRate:chan new k1@bRate:chan
 new k2@bRate:chan  new k3@bRate:chan

let B() = do !k0; B() or delay@poissrate; ()

let L1() =  ?k0; L1p()
and L1p() = !k1; L1()
and L2() = ?k1; L2p()
and L2p() = !k2; L2()
and L3() = ?k2; L3p()
and L3p() = !k3; L3()
and L4() = ?k3; L4p()
and L4p() = delay@uRate; L4()

let raising(p:proc(), t:float) =
```

```
(* Produce one p() every t sec with precision dt *)
  (val dt= 100.0 run step(p, t, dt, dt))
   and step(p:proc(), t:float, n:float, dt:float) =
     if n<=0.0 then (p()|step(p,t,dt,dt))
     else delay@dt/t; step(p,t,n-1.0,dt)

let A() = delay@poissrate; (B()|A())

run 100 of (L1() | L2() | L3() | L4())
run raising(A,100.0)
```

7. Interaction partners of the *Bacillus subtilis* sporulation phosphorelay proteins

We collected the interaction partners of the members of the *Bacillus subtilis* sporulation phosphorelay from the String database (http://string.embl.de), with probability cutoff at 0.75. In the top layer of the relay stand five HK's: (KinA-E), the other layers are as follows: Spo0F-Spo0B-Spo0A. Here we report the names, brief descriptions and probability of existence of interaction between the investigated proteins (links are actively direct readers to the String database). Grey background label interaction partners that are members of the phosphorelay (so these are expected interactions). Red background label interaction partners that are kinases or phosphatases that are known to be involved in other pathways, but seem to have some cross-talking ability to this sporulation relay.

Layer 1:

KinA

spo0F	Sporulation initiation phosphotransferase F; Key element in the phosphorelay regulating sporula [...] (124 aa)	0.911
spo0B	Sporulation initiation phosphotransferase B; Key element in the phosphorelay regulating sporula [...] (192 aa)	0.786

KinB

spo0F	Sporulation initiation phosphotransferase F; Key element in the phosphorelay regulating sporula [...] (124 aa)	0.923
kapB	Kinase-associated lipoprotein B; May play a role in the activation or the expression of kinB (128 aa)	0.892

KinC

spo0A	Stage 0 sporulation protein A; May play the central regulatory role in sporulation. It may be a [...] (267 aa)	0.947
spo0F	Sporulation initiation phosphotransferase F; Key element in the phosphorelay regulating sporula [...] (124 aa)	0.929
ykqA	UPF0131 protein ykqA (277 aa)	0.809

KinD

spo0F	Sporulation initiation phosphotransferase F; Key element in the phosphorelay regulating sporula [...] (124 aa)	0.766

KinE

ogt	Methylated-DNA--protein-cysteine methyltransferase; Involved in the cellular defense against th [...] (165 aa)	0.869

spo0F	Sporulation initiation phosphotransferase F; Key element in the phosphorelay regulating sporula [...] (124 aa)	0.788

Layer 2 (Spo0F):

spo0B	Sporulation initiation phosphotransferase B; Key element in the phosphorelay regulating sporula [...] (192 aa)	0.994
kinC	Sporulation kinase C; Phosphorylates the sporulation-regulatory protein spo0A (428 aa)	0.929
kinB	Sporulation kinase B; Phosphorylates the sporulation-regulatory proteins spo0A and spo0F. Spo0F [...] (429 aa)	0.923
cheA	Chemotaxis protein cheA; Involved in the transmission of sensory signals from the chemoreceptor [...] (671 aa)	0.919
kinA	Sporulation kinase A; Phosphorylates the sporulation-regulatory proteins spo0A and spo0F. It al [...] (606 aa)	0.911
rapE	Response regulator aspartate phosphatase E (375 aa)	0.867
kinE	two-component sensor histidine kinase (738 aa)	0.788
kinD	two-component sensor histidine kinase (506 aa)	0.766
rocR	Arginine utilization regulatory protein rocR; Positive regulator of arginine catabolism. Contro [...] (461 aa)	0.766
phoR	Alkaline phosphatase synthesis sensor protein phoR; Member of the two-component regulatory syst [...] (579 aa)	0.751

Layer 3 (Spo0B):

spo0F	Sporulation initiation phosphotransferase F; Key element in the phosphorelay regulating sporula [...] (124 aa)	0.994
spo0A	Stage 0 sporulation protein A; May play the central regulatory role in sporulation. It may be a [...] (267 aa)	0.946
obg	Spo0B-associated GTP-binding protein; Essential GTP-binding protein (428 aa)	0.836
kinA	Sporulation kinase A; Phosphorylates the sporulation-regulatory proteins spo0A and spo0F. It al [...] (606 aa)	0.786
spoIIP	required for dissolution of the septal cell wall (stage II sporulation) (401 aa)	0.776

Layer 4 (Spo0A):

abrB	Transition state regulatory protein abrB; Ambiactive repressor and activator of the transcripti [...] (96 aa)	0.982
kinC	Sporulation kinase C; Phosphorylates the sporulation-regulatory protein spo0A (428 aa)	0.947
spo0B	Sporulation initiation phosphotransferase B; Key element in the phosphorelay regulating sporula [...] (192 aa)	0.946
sinR	HTH-type transcriptional regulator sinR; Negative as well as positive regulator of alternate de [...] (111 aa)	0.924
soj	Sporulation initiation inhibitor protein soj; Inhibits the initiation of sporulation, spo0J ant [...] (253 aa)	0.894
spoIVB	SpoIVB peptidase; Plays a central role in the sigma-K checkpoint which coordinates gene express [...] (425 aa)	0.892
sigA	RNA polymerase sigma factor rpoD; Sigma factors are initiation factors that promote the attachm [...] (371 aa)	0.882
spoIIE	Stage II sporulation protein E; Normally needed for pro-sigma E processing during sporulation b [...] (827 aa)	0.876

sigE	Sporulation-essential sigma factor; Sigma factors are initiation factors that promote the attac [...] (239 aa)	0.811
qk	Thermostable fibrinolytic enzyme Nk1 (Nattokinase); Subtilisin is an extracellular alkaline ser [...] (381 aa)	0.800

It can be concluded, that the first layer proteins interact only with the corresponding downstream elements, the second layer integrates the input from the five Kin HK's and also cross talks with other pathways through three other regulators. The third layer has no such cross talking capability, while the last layer spreads the signal to various transcriptional regulators.

8. Supplementary figures

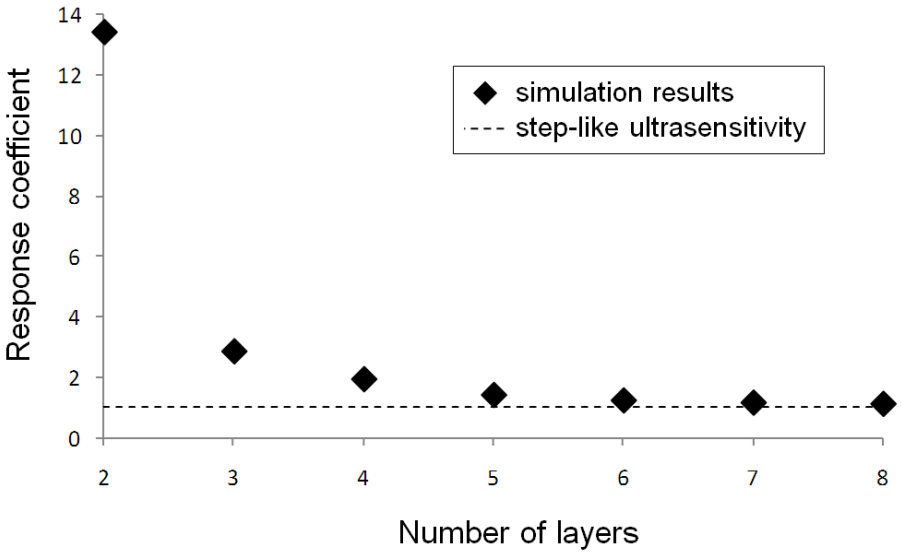

Figure S1. Response coefficient of the most ultrasensitive layer of the relay against increasing relay length. The response coefficient gives the ratio of input needed to achieve 90% vs. 10% response. Values closer to 1 means the system is more ultrasensitive; a minimal difference in input can cause a switch from 10% to 90% activity. Dashed line at response coefficient = 1 shows maximal, step-like ultrasensitivity.

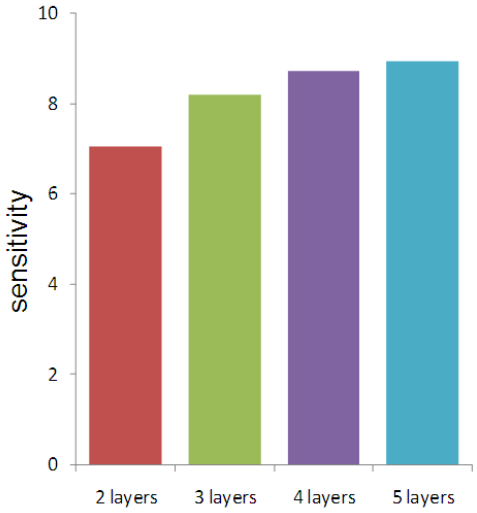

Figure S2. Sensitivity for relays with different length between input values 0, 2. For the details of calculation of sensitivity, see *Methods* section and legend of Figure 2.

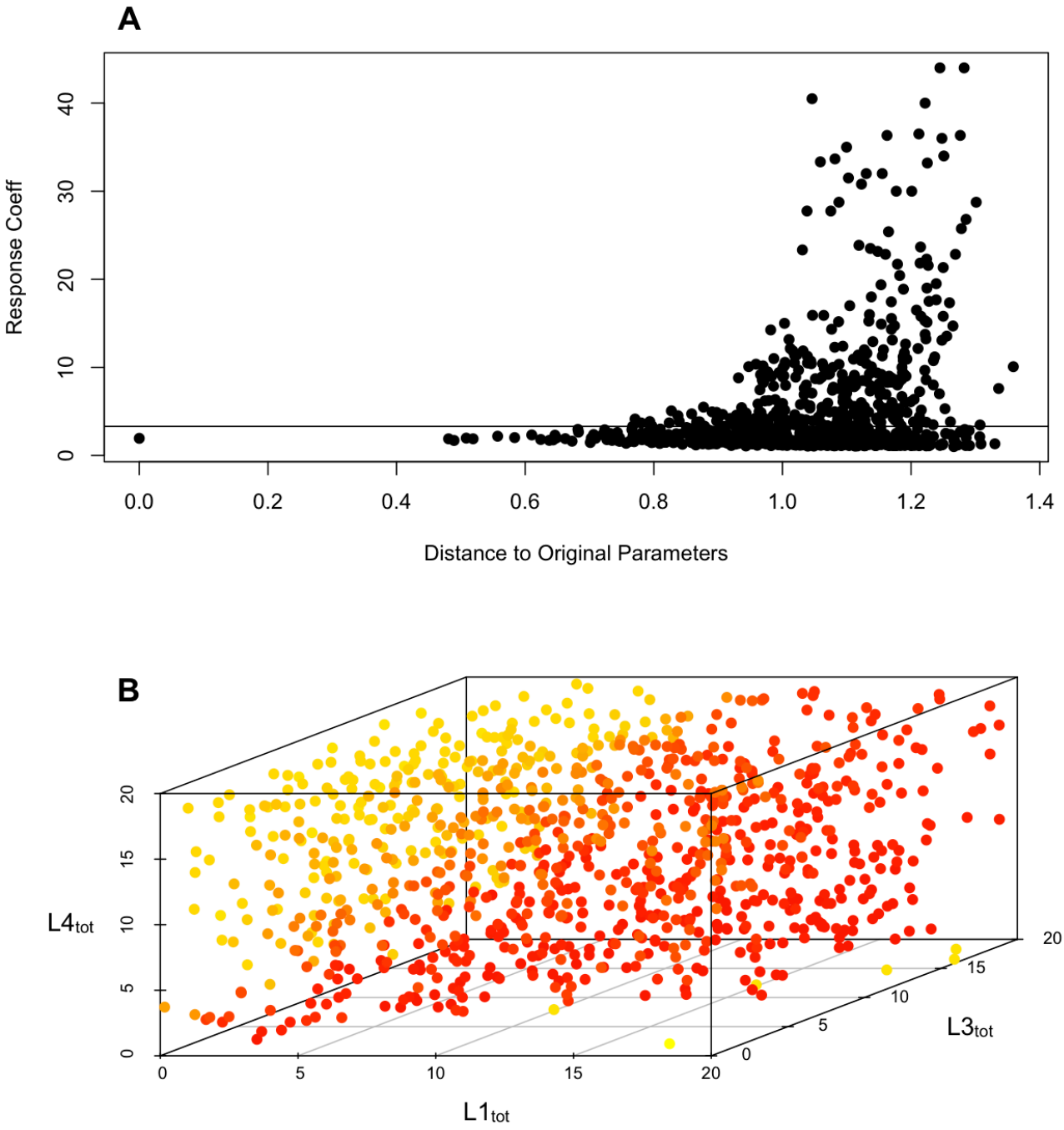

Figure S3. Response coefficient, calculated from input-response curve of the second layer for randomly selected parameter sets. A: Response coefficient versus changes in kinetic parameters (k_2, k_3, k_4 and k_5). Each of the 1000 dots corresponds to a different parameter set, with the far-left dot representing the original parameter set (all rates equal to one). The x-axis gives the Euclidian distance between a given parameter set and the original. The y-axis gives the response coefficient calculated as the input level for reaching 90% over 10% of maximum response. The horizontal line indicates a response coefficient of 3.03, equal to that obtained from a Hill function with Hill coefficient 2. **B:** Response coefficient versus changes in protein total concentrations ($L1_{tot}$, $L3_{tot}$, $L4_{tot}$). Each of the 1000 dots represents a combination of these parameters, while the color coding shows the value of the response coefficient. The maximum (red) and minimum (yellow) response coefficients observed were 11.53 and 1.07 respectively. Note that for both analyses input-response curves are derived from the analytical steady state solution of Eq. 1 from the main text and by setting self-phosphorylation and dephosphorylation rates to zero (see *Methods*).

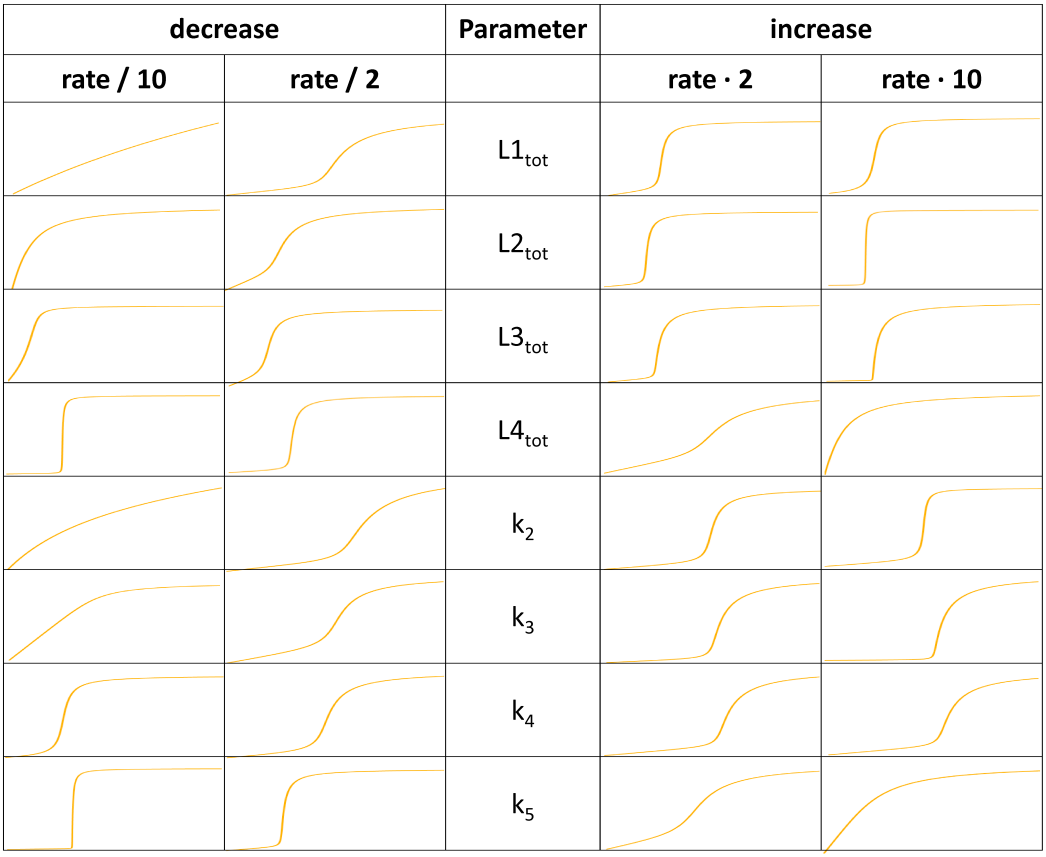

Figure S4. Input-response curves for the second layer (*L2p*) for systematic parameter variations. For the curves on the left (right), the parameter indicated in the middle is decreased (increased) by the amount shown in the header. The base value of the parameters is 1 (for kinetic rates) or 10 (for protein concentrations). X- and Y-axes on different panels are scaled so to visualize existence (or absence) of ultrasensitivity in *L2p* input-response curves.

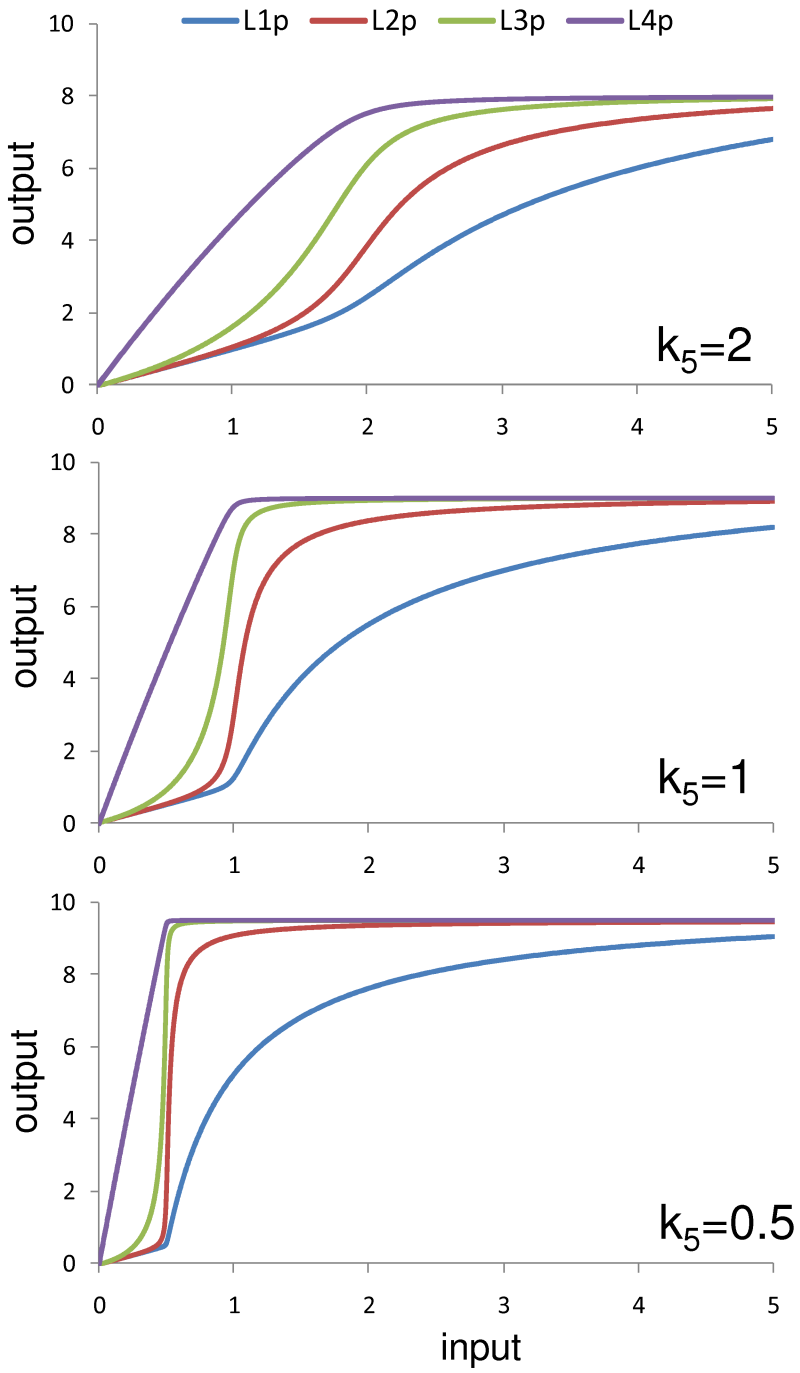

Figure S5. Input-response curves at various phosphate loss rates. Input-response curves were calculated at various phosphate loss rates (k_5) indicated on the different panels. Middle panel corresponds to the basal model. Note that the inflection of the *L2p* response curve and the saturation point of the *L4p* curve happens at an input value equal to k_5.

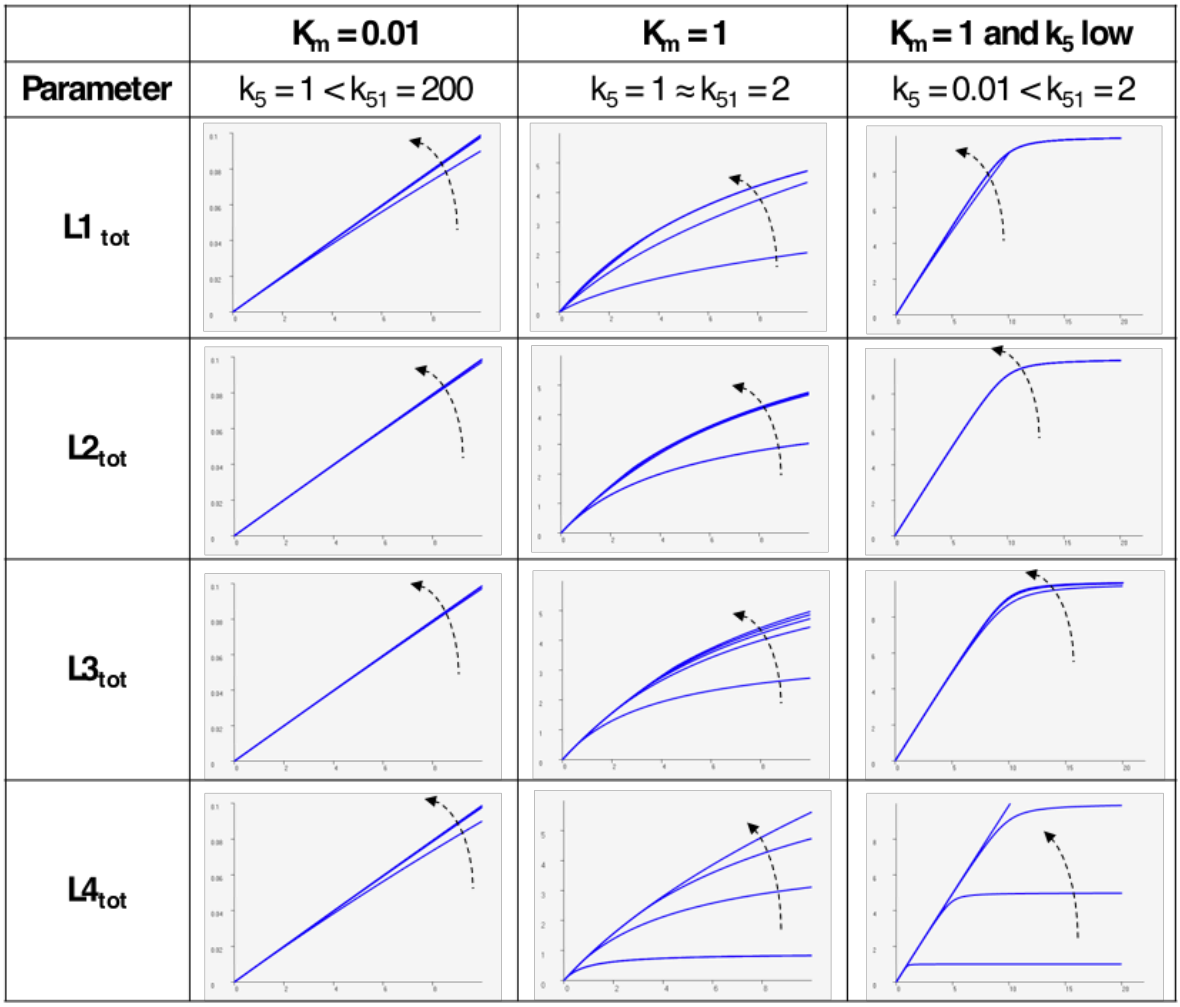

Figure S6. Absolute concentration robustness in L4p. The input – response curves for the last layer of the phosphorelay (*L4p*) with bifunctional enzyme present as *L1*. Enzyme efficiency (i.e. the value of K_m) is shown on the top row, while second row gives the values of other relevant parameters (see section 2). For each model, the total concentrations of proteins from each layer (rows) are set to 1, 5, 10, 20, and 100, and the resulting input-response curve for *L4p* are plotted. Dashed arrows on each panel indicate increasing total protein concentration. Absolute protein concentration for *L4p* is visible when *L1* is an efficient enzyme or - in a limited input regime - when the non-enzymatic, auto-dephosphorylation rate is low .

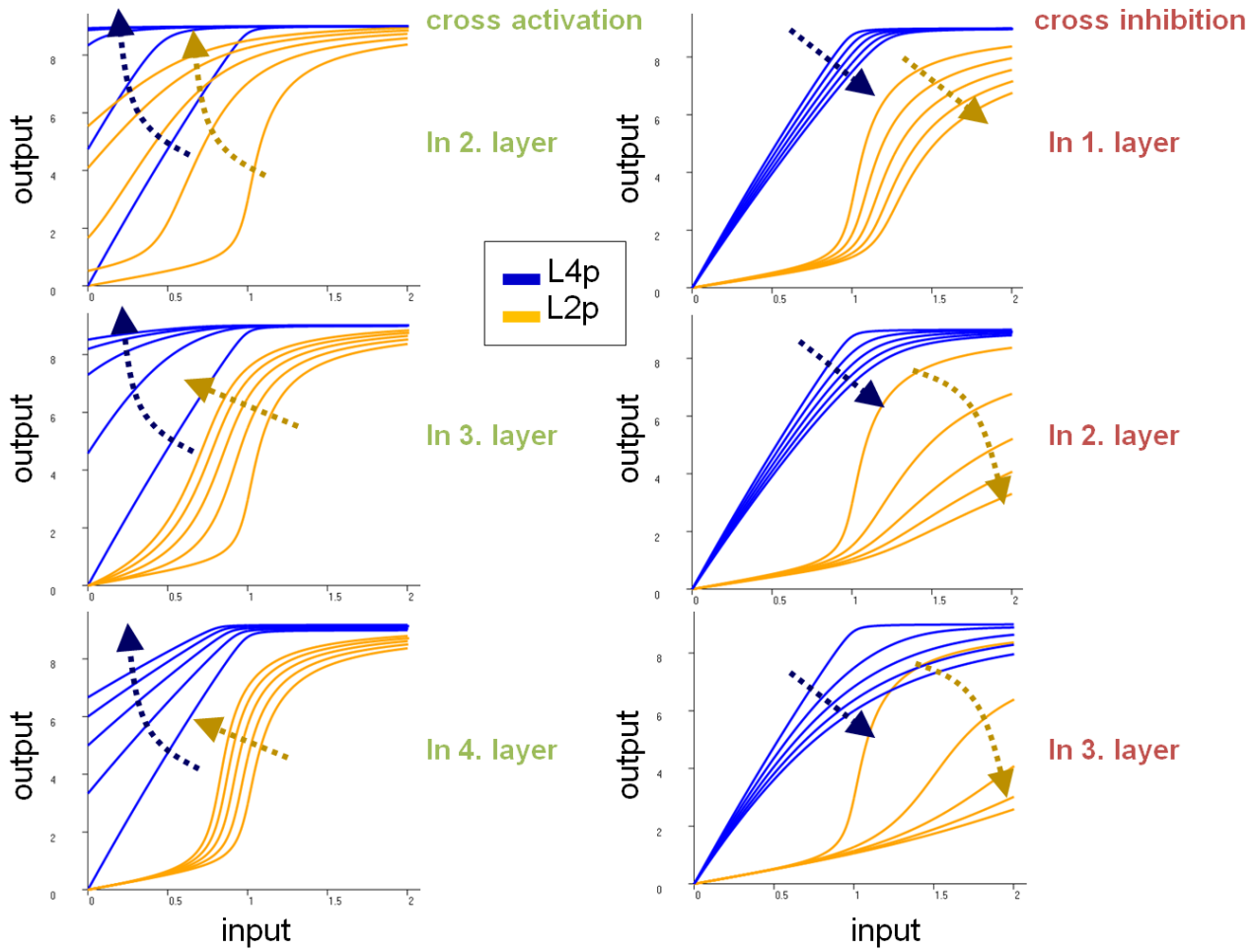

Figure S7. Effect of crosstalk on input-response curves. For relays with different length, we first derive the signal-response curves for the output layer (*L4p* - blue) and the most sensitive layer (*L2p* - yellow). We notice a clear decrease in ultrasensitivity of this middle layer as a consequence of crosstalk. We then repeat this analysis for increasing values of crosstalk parameters (see *Methods*). We calculate the sensitivities of the output (*L4p*) at each crosstalk level from the slopes of the blue curves. Weak cross-activation in the second layer leads to sharp activation (at crosstalk = *0.5*), but in a very narrow input regime, while the same crosstalk in lower layers of the relay lead to a bit shallower slope, but in a wider response range. Still these small differences show up as this case would go against the general trend, where activating crosstalk arriving in lower layers cause less sensitivity decrease (see Fig. S8).

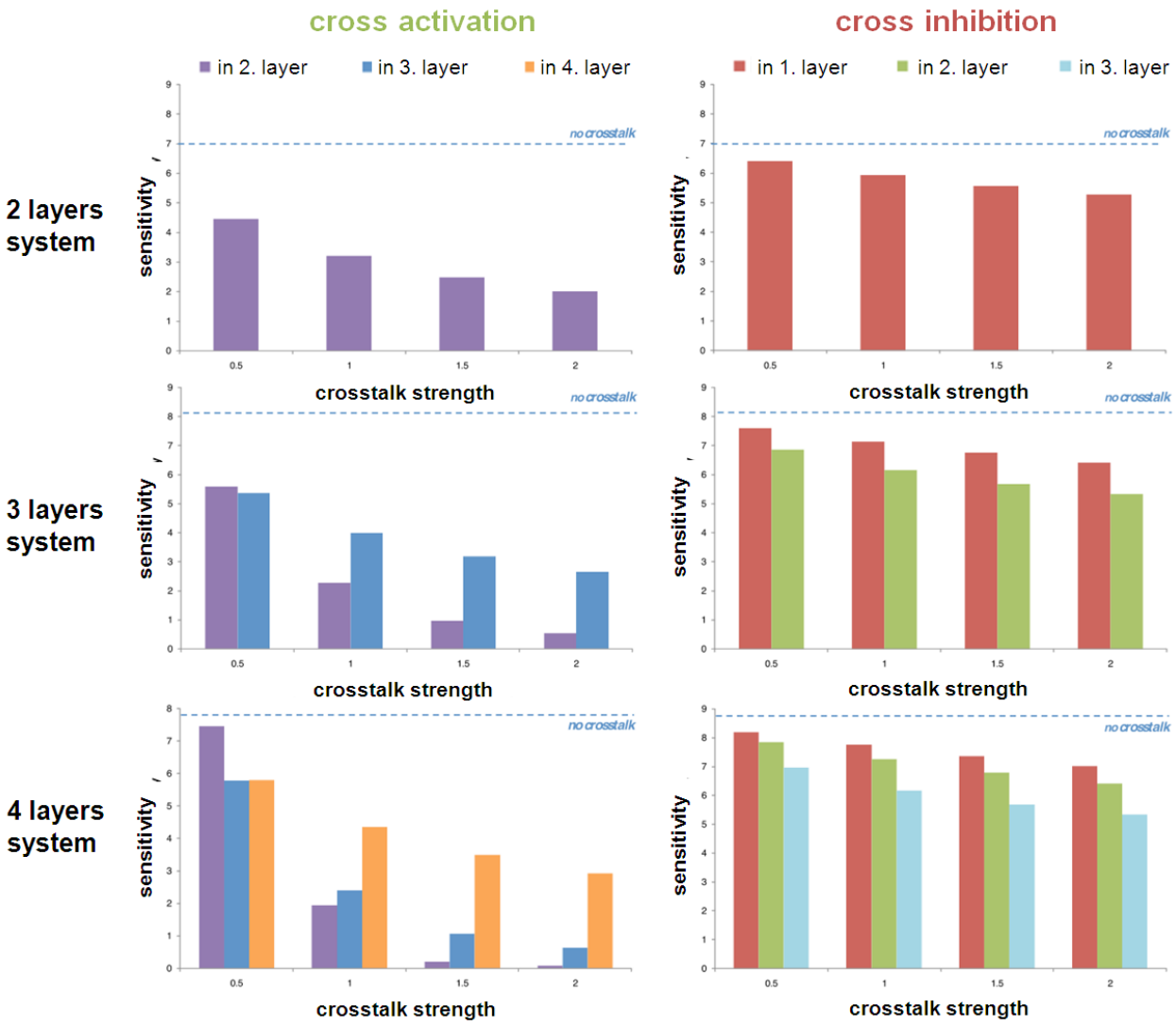

Figure S8. Effects of cross talk on sensitivity. Sensitivity of the response to the input (between input values 0 - 2) under different crosstalk strength in different system. The location of crosstalk entering the relay and its nature are indicated on both panels. Colored bars correspond to sensitivity at different layers as indicated on each panel. See legend of Figure 2 for details of the calculations.

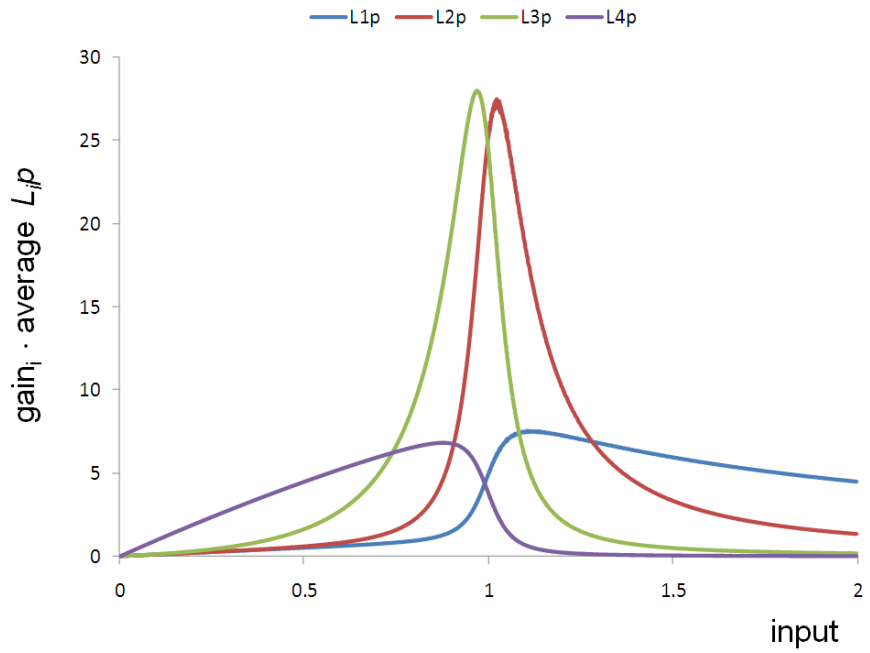

Figure S9. Gain · mean for different layers of a 4-layer system. Gain for layer *i* is calculated as (Δ*Lip*/*Lip*)/(Δ*input*/i*nput*) - similarly to sensitivity from the input-response curves obtained from the deterministic model.

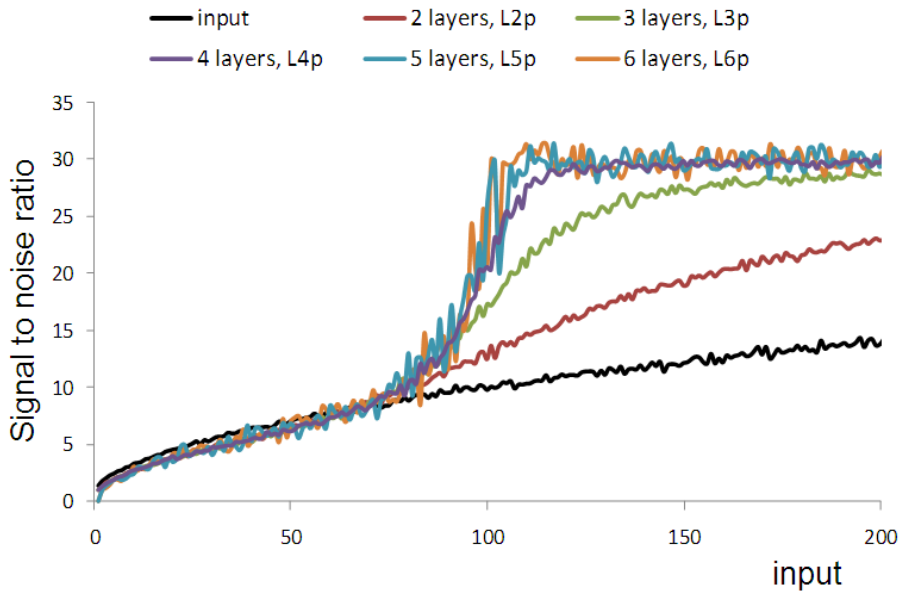

Figure S10. Effect of relay length on noise. Signal to noise ratio (SNR) of the response from the last layer for relays of different length. Note that increasing relay length behind four does not change SNR dynamics significantly.

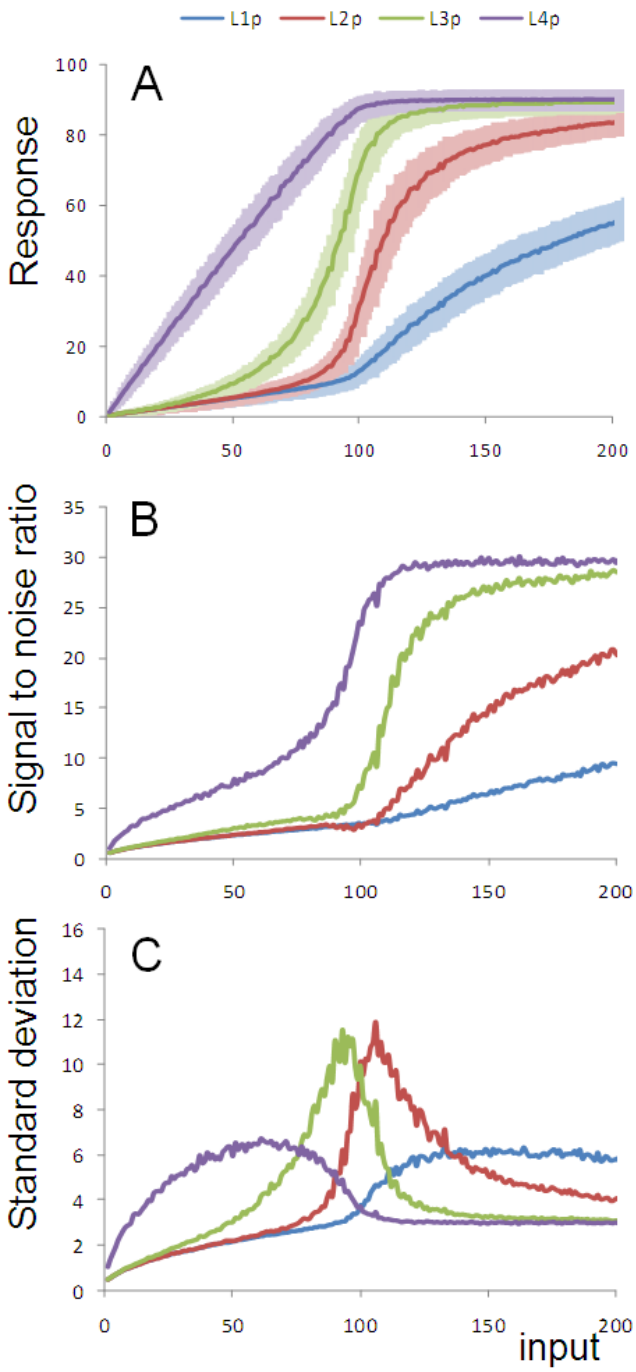

Figure S11. Effect of intrinsic noise on a 4 layers phosphorelay. Analysis of stochasticity in response dynamics of a 4-layer relay under "clear" input (see *Methods*). In this analysis, the input is increased in deterministically allowing us to analyze system dynamics just under intrinsic noise. As in Figure 4, we plot stochastic input-response curves (panel A), SNR for each layer (panel B) and variance in the phosphorylated protein at each layer (panel C). A comparison of these plots with those of Figure 4 reveal that they almost perfectly match, indicating that under the model parameters used the system is dominated by intrinsic noise only without any additional effects from external noise.