Biophysical Journal, Volume *114*

Supplemental Information

Minimal Network Topologies for Signal Processing during Collective Cell Chemotaxis

Haicen Yue, Brian A. Camley, and Wouter-Jan Rappel

Minimal network topologies for signal processing during collective cell chemotaxis

Haicen Yue[1], Brian A. Camley[2,3], Wouter-Jan Rappel[1*]

1. Department of Physics, University of California, San Diego, La Jolla, CA, US
2. Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD, US
3. Department of Biophysics, Johns Hopkins University, Baltimore, MD, US

*rappel@physics.ucsd.edu

Supporting Material

Derivation of Equation (1) in the main text

According to the chain rule of differentiating a function with multiple variables,

$$\frac{\mathrm{dProb}}{\mathrm{d[S]}} = \frac{\partial \mathrm{Prob}}{\partial \mathrm{[S]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\mathrm{d[G]}}{\mathrm{d[S]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[R]}}\frac{\mathrm{d[R]}}{\mathrm{d[S]}} \tag{2}$$

$$\frac{\mathrm{d[G]}}{\mathrm{d[S]}} = \frac{\partial \mathrm{[G]}}{\partial \mathrm{[S]}} + \frac{\partial \mathrm{[G]}}{\partial \mathrm{[R]}}\frac{\mathrm{d[R]}}{\mathrm{d[S]}} + \frac{\partial \mathrm{[G]}}{\partial \mathrm{Prob}}\frac{\mathrm{dProb}}{\mathrm{d[S]}} \tag{3}$$

$$\frac{\mathrm{dProb}}{\mathrm{d[R]}} = \frac{\partial \mathrm{Prob}}{\partial \mathrm{[R]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\mathrm{d[G]}}{\mathrm{d[R]}} \tag{4}$$

$$\frac{\mathrm{d[G]}}{\mathrm{d[R]}} = \frac{\partial \mathrm{[G]}}{\partial \mathrm{[R]}} + \frac{\partial \mathrm{[G]}}{\partial \mathrm{Prob}}\frac{\mathrm{dProb}}{\mathrm{d[R]}} \tag{5}$$

Substitute $\frac{\mathrm{d[G]}}{\mathrm{d[S]}}$ in equation (2) with equation (3) and we get:

$$\frac{\mathrm{dProb}}{\mathrm{d[S]}} = \frac{\partial \mathrm{Prob}}{\partial \mathrm{[S]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\left(\frac{\partial \mathrm{[G]}}{\partial \mathrm{[S]}} + \frac{\partial \mathrm{[G]}}{\partial \mathrm{[R]}}\frac{\mathrm{d[R]}}{\mathrm{d[S]}} + \frac{\partial \mathrm{[G]}}{\partial \mathrm{Prob}}\frac{\mathrm{dProb}}{\mathrm{d[S]}}\right) + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[R]}}\frac{\mathrm{d[R]}}{\mathrm{d[S]}}$$

After rearrangement, we can get:

$$\left(1 - \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\partial \mathrm{[G]}}{\partial \mathrm{Prob}}\right)\frac{\mathrm{dProb}}{\mathrm{d[S]}} = \frac{\partial \mathrm{Prob}}{\partial \mathrm{[S]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\partial \mathrm{[G]}}{\partial \mathrm{[S]}} + \left(\frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\partial \mathrm{[G]}}{\partial \mathrm{[R]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[R]}}\right)\frac{\mathrm{d[R]}}{\mathrm{d[S]}} \tag{6}$$

Substituting $\frac{\mathrm{d[G]}}{\mathrm{d[R]}}$ in equation (4) with equation (5) results in:

$$\left(1 - \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\partial \mathrm{[G]}}{\partial \mathrm{Prob}}\right)\frac{\mathrm{dProb}}{\mathrm{d[R]}} = \frac{\partial \mathrm{Prob}}{\partial \mathrm{[R]}} + \frac{\partial \mathrm{Prob}}{\partial \mathrm{[G]}}\frac{\partial \mathrm{[G]}}{\partial \mathrm{[R]}} \tag{7}$$

Then we substitute $\frac{\partial \mathrm{Prob}}{\partial[\mathrm{G}]}\frac{\partial[\mathrm{G}]}{\partial[\mathrm{R}]}+\frac{\partial \mathrm{Prob}}{\partial[\mathrm{R}]}$ in the third term of the right-hand-side of equation (6) with equation (7) and get equation (1) in the main text as below:

$$\left(1-\frac{\partial \mathrm{Prob}}{\partial[\mathrm{G}]}\frac{\partial[\mathrm{G}]}{\partial \mathrm{Prob}}\right)\frac{\mathrm{dProb}}{\mathrm{d}[\mathrm{S}]}=\frac{\partial \mathrm{Prob}}{\partial[\mathrm{S}]}+\frac{\partial \mathrm{Prob}}{\partial[\mathrm{G}]}\frac{\partial[\mathrm{G}]}{\partial[\mathrm{S}]}+\left(1-\frac{\partial \mathrm{Prob}}{\partial[\mathrm{G}]}\frac{\partial[\mathrm{G}]}{\partial \mathrm{Prob}}\right)\frac{\mathrm{dProb}}{\mathrm{d}[\mathrm{R}]}\frac{\mathrm{d}[\mathrm{R}]}{\mathrm{d}[\mathrm{S}]}$$

Mean First Passage Time (MFPT)

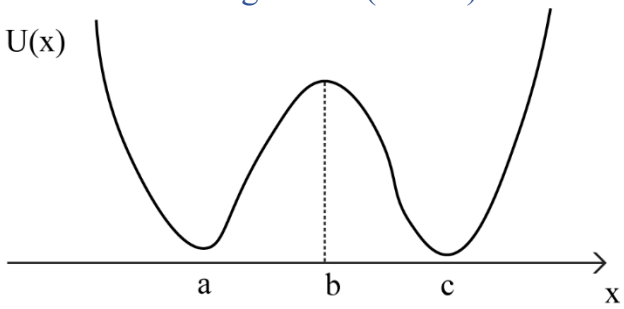

Fig. S1. Sketch of the double-well potential U(x).

The diffusion process in an external potential U(x) can be described by the quasilinear Fokker-Planck equation:

$$\frac{\partial \mathrm{P}(\mathrm{x},\mathrm{t})}{\partial \mathrm{t}}=\frac{\partial}{\partial \mathrm{x}}\mathrm{U}'(\mathrm{x})\mathrm{P}+\sigma^2\frac{\partial^2 \mathrm{P}}{\partial \mathrm{x}^2}.$$

If the potential U(x) is bistable with shape like that in Fig. S1, the MFPT from point "a" to "c" is (1):

$$\tau_{\mathrm{ca}}=\frac{1}{\sigma^2}\int_{\mathrm{a}}^{\mathrm{c}}\mathrm{e}^{\frac{\mathrm{U}(\mathrm{x}')}{\sigma^2}}\mathrm{dx}'\int_{-\infty}^{\mathrm{x}'}\mathrm{e}^{-\frac{\mathrm{U}(\mathrm{x}'')}{\sigma^2}}\mathrm{dx}''.$$

If the shape of U(x) satisfies that $\mathrm{e}^{\frac{\mathrm{U}(\mathrm{x}')}{\sigma^2}}$ is large when x' is near b and otherwise exponentially smaller, it can be replaced with $\exp\left[\frac{\mathrm{U}(\mathrm{b})}{\sigma^2}-\frac{|\mathrm{U}''(\mathrm{b})|}{2\sigma^2}(\mathrm{x}'-\mathrm{b})^2\right]$ in the first integral. Similarly, in the second integral, as $\mathrm{x}'\approx\mathrm{b}$, the main contribution comes from the neighborhood $\mathrm{x}''\approx\mathrm{a}$ and may be approximated with:

$$\int_{-\infty}^{\infty}\exp\left[-\frac{\mathrm{U}(\mathrm{a})}{\sigma^2}-\frac{\mathrm{U}''(\mathrm{a})}{2\sigma^2}(\mathrm{x}''-\mathrm{a})^2\right]\mathrm{dx}''=\sqrt{\frac{2\pi\sigma^2}{\mathrm{U}''(\mathrm{a})}}\exp\left[-\frac{\mathrm{U}(\mathrm{a})}{\sigma^2}\right].$$

So, the MFPT $\tau_{\mathrm{ca}}\approx\frac{2\pi}{\sqrt{\mathrm{U}''(\mathrm{a})|\mathrm{U}''(\mathrm{b})|}}\exp\left[\frac{\mathrm{U}(\mathrm{b})-\mathrm{U}(\mathrm{a})}{\sigma^2}\right]$. This is the Kramers approximation.

If the shape of U(x) does not satisfy the above prerequisite, we numerically calculate the double integrals for MFPT. The time cost for numerical integrations is large if we repeatedly implement this for every set of parameters and every state transition (total number is 64×64). We can make significant savings by applying the Kramers approximation where appropriate, and where the prerequisite for the approximation is not satisfied, by pre-computing a table of these rates numerically. We find that the MFPT only directly depends on $[\mathrm{R}]^{\mathrm{s}}$, $\mathrm{P}^{\mathrm{tot}}$ and y (a function of $[\mathrm{R}]^{\mathrm{s}}$, [S] and $[\mathrm{G}]^{\mathrm{s}}$). The superscript s here means the steady state. So, we make a table of the numerical computation results of the double integrals on a three-

dimensional $25 \times 200 \times 200$ grid of $[R]^s$, P^{tot} and y and use linear interpolation to get the MFPTs when needed.

Fitting procedure

Our model contains 10 parameters to be fitted with two of them being PA and PI which only are relevant in light-treatment condition. We use simulated annealing as our method of fitting. During the fitting, this method accepts a trial step with some probability dependent on an artificial temperature T, even when this step does not improve the fit. This can help avoid being trapped in a local minimum (2). The initial temperature is usually high to have a larger searching area and then, the temperature is gradually decreased, leading to more selective sampling towards the error decreasing direction. We use the simulannealbnd function in MATLAB (R2015b, The MathWorks, Natick, MA) with default settings.

Even with the simulated annealing method, it is still not guaranteed that the samplings are not trapped in the local minima. So, we run three rounds of fitting for each network. In the first round, the error function we use is:

$$\mathrm{error}_1 = \frac{1}{1000}\left(\frac{\left(\mathrm{NP}_{\mathrm{wild}}^{\mathrm{sim}} - \mathrm{NP}_{\mathrm{wild}}^{\mathrm{exp}}\right)^2}{0.2^2} + \frac{\left(\mathrm{NP}_{\mathrm{Rec^{DN}}}^{\mathrm{sim}} - \mathrm{NP}_{\mathrm{wild}}^{\mathrm{sim}} - \left(\mathrm{NP}_{\mathrm{Rec^{DN}}}^{\mathrm{exp}} - \mathrm{NP}_{\mathrm{wild}}^{\mathrm{exp}}\right)\right)^2}{0.1^2} + \mathrm{H}\left(0.1 - \mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}}\right)\frac{\left(\mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}} - \mathrm{Dir}_{\mathrm{wild}}^{\mathrm{exp}}\right)^2}{0.01^2}\right).$$

Here, H is the Heaviside step function ensuring that only a directionality that is not significantly positive (smaller than 0.1) increases the error function. Specifically, when $\mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}} > 0.1$, $\mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}}$'s value is not important at all as the whole term is zero and when $\mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}} < 0.1$, the term $\left(\mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}} - \mathrm{Dir}_{\mathrm{wild}}^{\mathrm{exp}}\right)^2$ gives a bias towards larger $\mathrm{Dir}_{\mathrm{wild}}^{\mathrm{sim}}$ during the fitting. As there are no data for light-treatment experiment in the first round, we only fit for the eight parameters except for PA and PI. We randomly choose 100 starting points in the parameter space and run 100 fittings with $\mathrm{T}_{\mathrm{initial}} = 100$. Then we pick the first five fitting results with the smallest error functions and then use them as the starting points for the next round of fitting. The error function used in the second round includes the data for light-treatment conditions:

$$\mathrm{error}_2 = \mathrm{error}_1 + \frac{1}{1000}\left(\frac{\left(\mathrm{NP}_{\mathrm{PA_{back}}}^{\mathrm{sim}} - \mathrm{NP}_{\mathrm{PA_{back}}}^{\mathrm{exp}}\right)^2}{0.3^2} + \frac{\left(\mathrm{NP}_{\mathrm{PA_{RecDN}}}^{\mathrm{sim}} - \mathrm{NP}_{\mathrm{PA_{RecDN}}}^{\mathrm{exp}}\right)^2}{0.4^2} + \frac{\left(\mathrm{NP}_{\mathrm{PI}}^{\mathrm{sim}} - \mathrm{NP}_{\mathrm{wild}}^{\mathrm{sim}} - \left(\mathrm{NP}_{\mathrm{PI}}^{\mathrm{exp}} - \mathrm{NP}_{\mathrm{wild}}^{\mathrm{exp}}\right)\right)^2}{0.1^2} + \mathrm{H}\left(\mathrm{Dir}_{\mathrm{PA_{back}}}^{\mathrm{sim}} + 0.1\right)\frac{\left(\mathrm{Dir}_{\mathrm{PA_{back}}}^{\mathrm{sim}} - \mathrm{Dir}_{\mathrm{PA_{back}}}^{\mathrm{exp}}\right)^2}{0.01^2}\right).$$

In the second round, we keep the eight parameters in the first round constant and only fit for PA and PI and $\mathrm{T}_{\mathrm{initial}} = 1$. Then we put all the ten parameters together for the third round of fitting with the five results from the second round as the new starting points. Then, $\mathrm{error}_3 = 100 \times \mathrm{error}_2$ and $\mathrm{T}_{\mathrm{initial}} = \mathrm{error}_3(\mathrm{starting\ point})$ for the third round and among the five results, we choose the one with the smallest error function as the final fitting result. The fitting results are listed in the Table S1 and Table S2.

Table S1. Parameter fitting results for six-link networks												
	Network A		Network B		Network C		Network D		Network E		Network F	
S_{grad}	0.4109		0.4814		0.8746		0.5147		0.1590		0.2290	
S_{mean}	3.982		0.01205		0.1686		0.06754		0.02587		0.01093	
$basal_{AR}$	2.064		0.04961		0.5324		0.01124		0.7171		0.02782	
k_{RP}	52.40		9.356		3.825		6.417		0.2470		9.535	
P^{tot}	0.8892		2.411		7.475		5.097		96.32		5.223	
k_1	k_{SG}	0.1228	k_{-SG}	0.05322	k_{RG}	47.81	k_{-RG}	2.897	k_{RG}	62.25	k_{-RG}	77.79
k_2	k_{RG}	1.490	k_{-RG}	92.65	k_{-GR}	0.8540	k_{GR}	0.09783	k_{-GP}	0.1259	k_{GP}	0.7983
k_3	k_{-GP}	2.125	k_{GP}	5.101	k_{-SP}	1.159	k_{-SP}	8.702	k_{-SP}	1.229	k_{-SP}	10.22
PI	1034		4.939		7561		4580		0.05685		27.14	
PA	0.2777		0.02406		0.05794		0.1339		0.004729		0.01876	
σ	0.1 (not fitted but fixed)											

Other sampling results

In Figs. S2-4 we present additional sampling results. Specifically, Fig. S2 and Fig. S3 show the results for networks with five interactions while Fig. S4 shows the sampling outcomes for networks with six interactions. Note that all networks in Figs. S2 and S3 can be excluded even though the network in the black box in Fig. S3 appears to work well for both experiments. But in fact, this network should be excluded. In this network, R is only positively dependent on S and R is the only entrance of the external signal. This means that $\frac{dProb}{d[S]} = \frac{dProb}{d[R]}\frac{d[R]}{d[S]}$ and $\frac{d[R]}{d[S]} > 0$. So the requirement of the $Receptor^{DN}$ experiments that $\frac{dProb}{d[S]} < 0$ is equivalent to $\frac{dProb}{d[R]} < 0$, which is contradictory to the prerequisite $\frac{dProb}{d[R]} > 0$ based on other experiments. For the six possible networks that we have selected, there is not this problem. Because in these six networks, R is not the only entrance of S for the network so that P's dependence on R is not necessarily equivalent to P's dependence on S. So, for these networks it is possible that $\frac{dProb}{d[S]} < 0$ while $\frac{dProb}{d[R]} > 0$.

Table S2. Parameter fitting results for two examples of seven-link networks				
	Network 1		Network 2	
S_{grad}	0.04922		0.1402	
S_{mean}	23.26		0.5970	
$basal_{AR}$	19.63		2.696	
k_{RP}	0.7761		8.232	
P^{tot}	137.9		2.723	
k_1	k_{SG}	0.2138	k_{SG}	0.3587
k_2	k_{RG}	154.7	k_{PG}	0.3165
k_3	k_{-GP}	0.008102	k_{RG}	2.324
k_4	k_{-GR}	205.00	k_{-GP}	1.397
PI	98.64		4.583	
PA	0.01032		0.1249	
σ	0.1 (not fitted but fixed)			

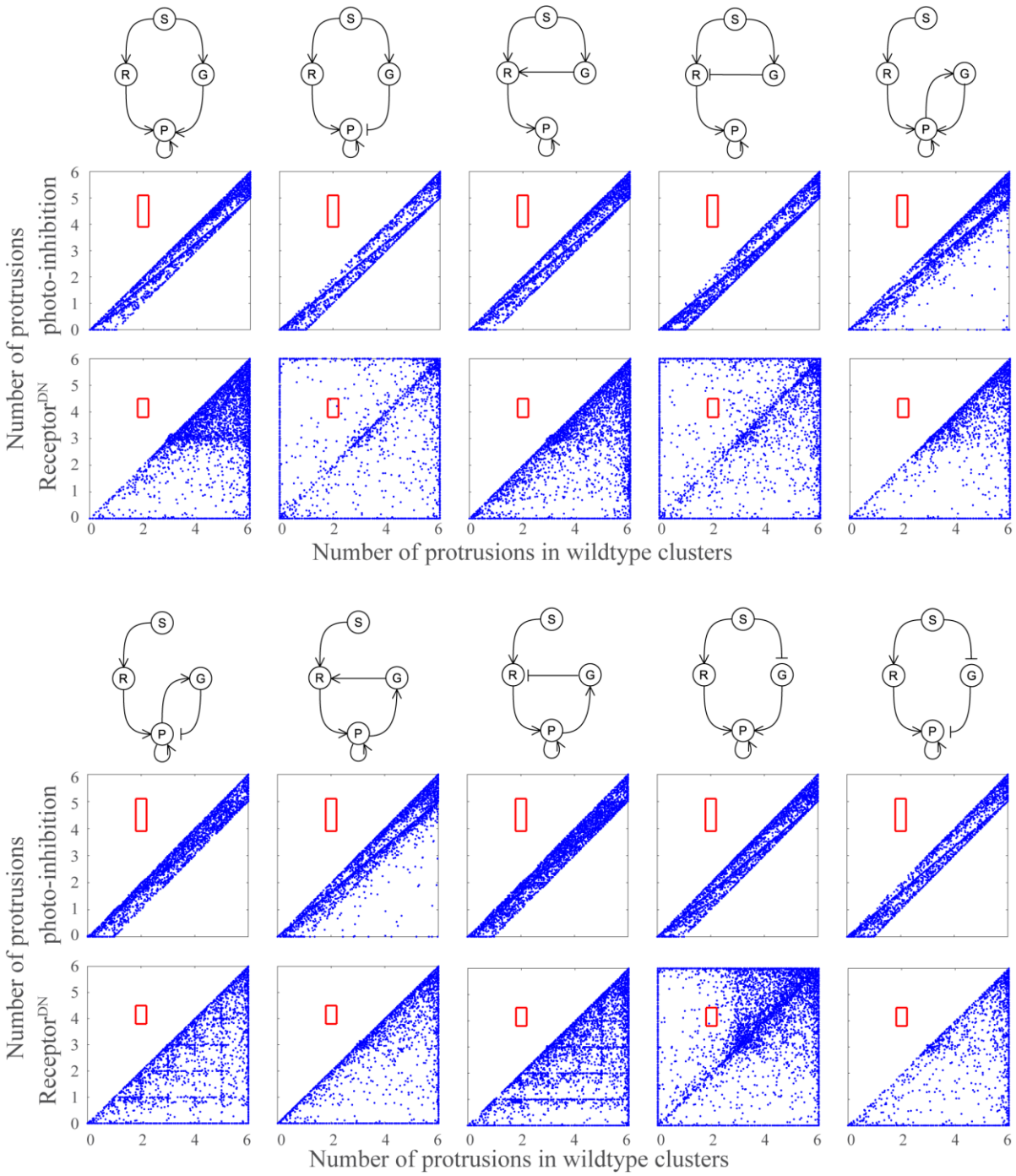

Fig. S2. The first ten sampling results for networks with five interactions. The other twelve are in Fig. S3.

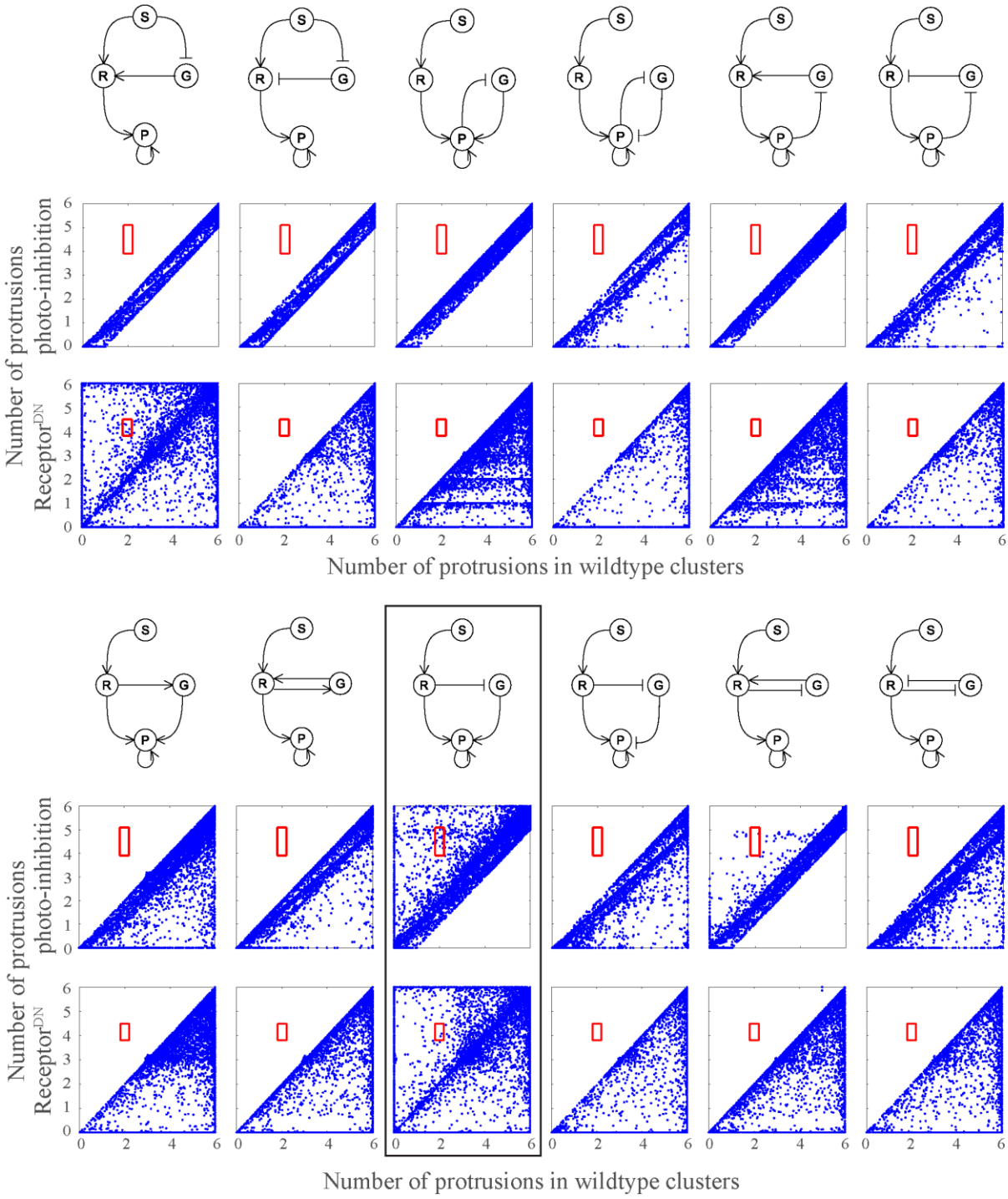

Fig. S3. The other twelve sampling results for networks with five interactions. The first ten are in Fig. S2. It's worth noting that the network in the black box appears to fit both experiments. However, it can be ruled out when considering the $\text{Receptor}^{\text{DN}}$ experiments and the $\frac{\text{dProb}}{\text{d[R]}} > 0$ prerequisite together. The details are in the supplementary text.

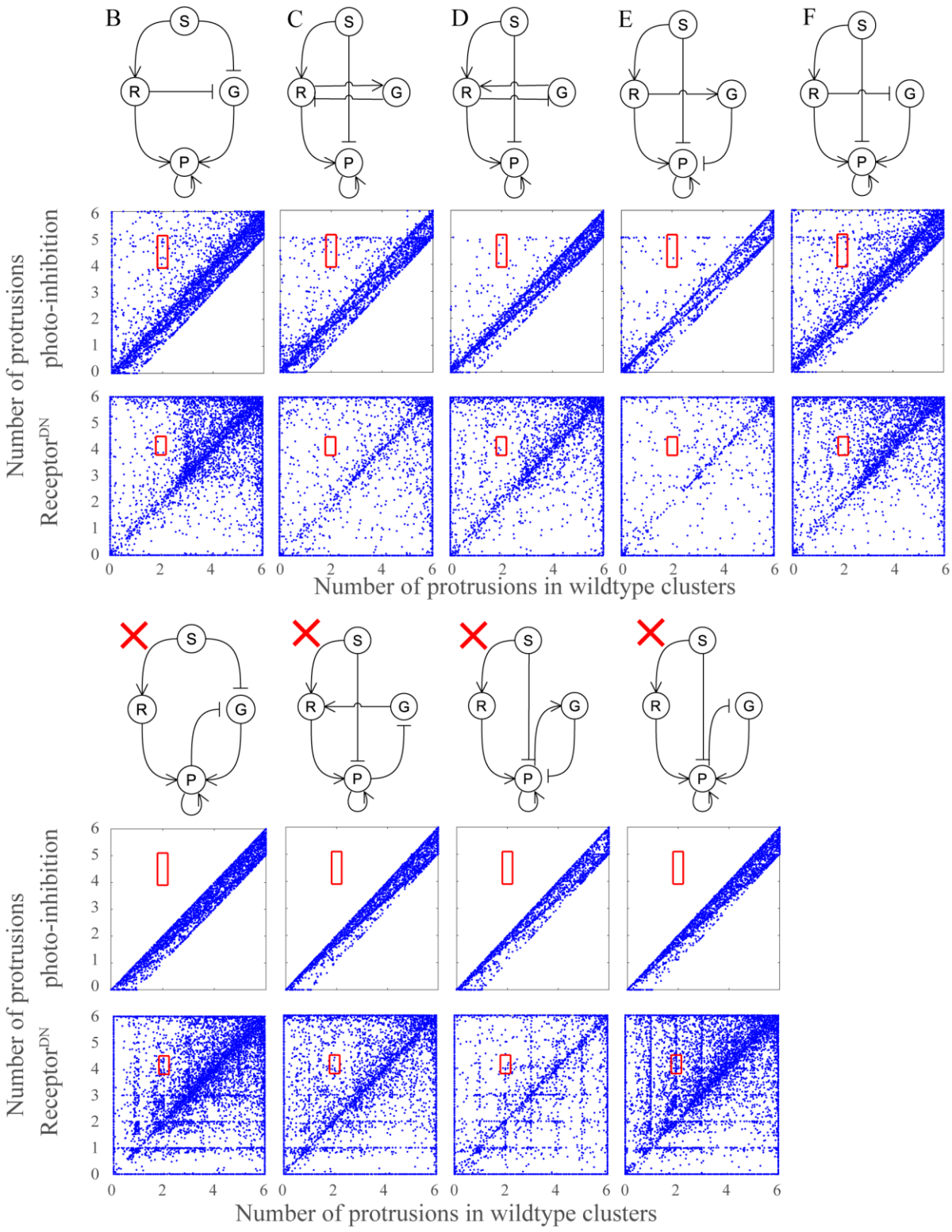

Fig. S4. The sampling results for networks with six interactions. The upper row corresponds to the networks B-F of Fig. 2c that are not excluded by the qualitative selection. The lower row shows examples of networks that are not consistent with the experimental data.

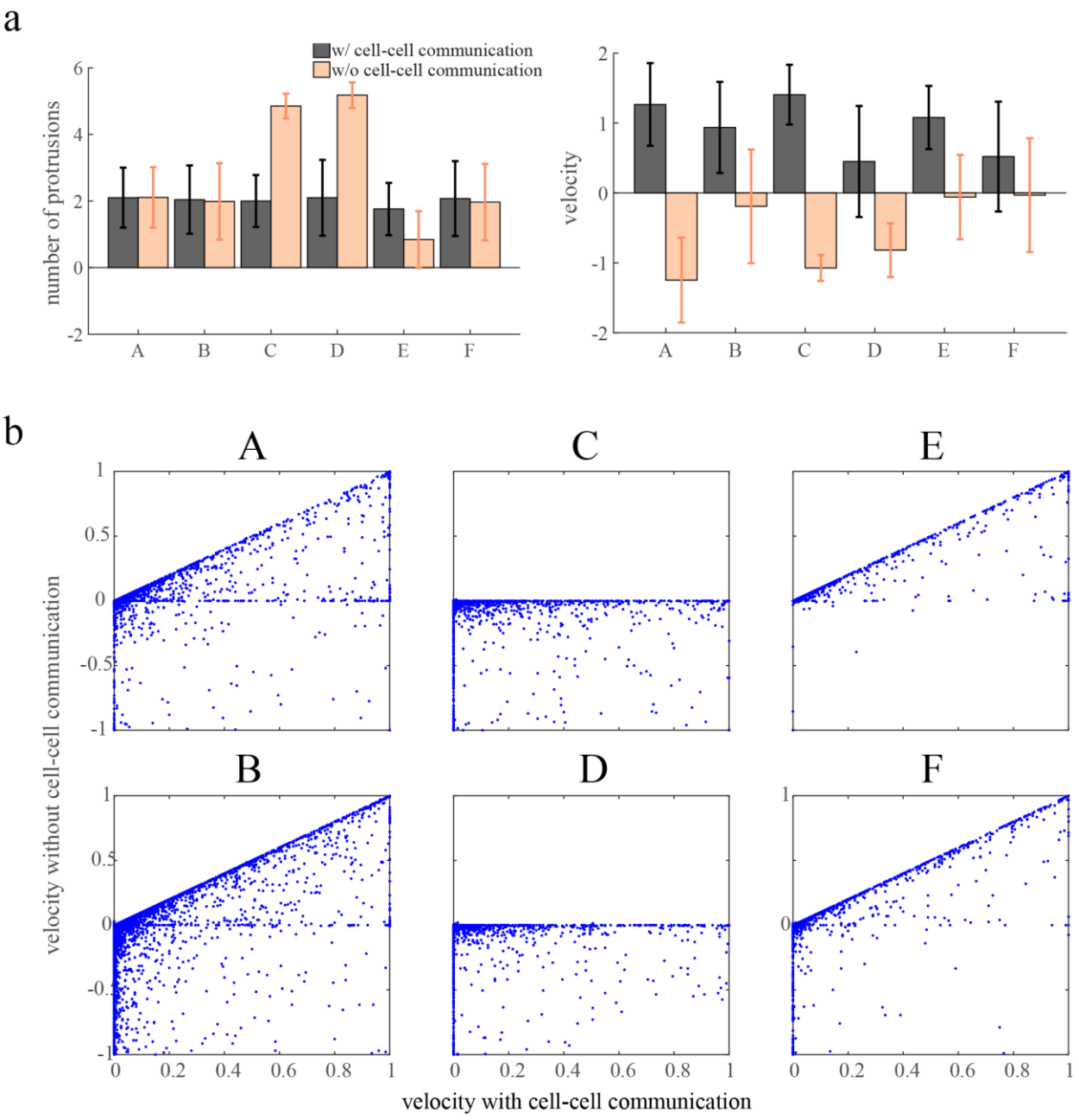

Fig. S5. The role of cell-cell communication. (a) The number of protrusions and cluster velocity for wildtype clusters with and without cell-cell communication for network A-F of Fig. 2c. Simulations were carried out using the fitted parameters. (b) Comparison of velocity between wildtype clusters with (x-axis) and without cell-cell communication (y-axis) using a large sample of randomly chosen parameters that result in positive cluster velocities in the presence of cell-cell communication. Each blue dot represents the result for one set of parameters. All blue dots are below the y=x line, indicating that the cluster velocity is always reduced when cell-cell communication is blocked.

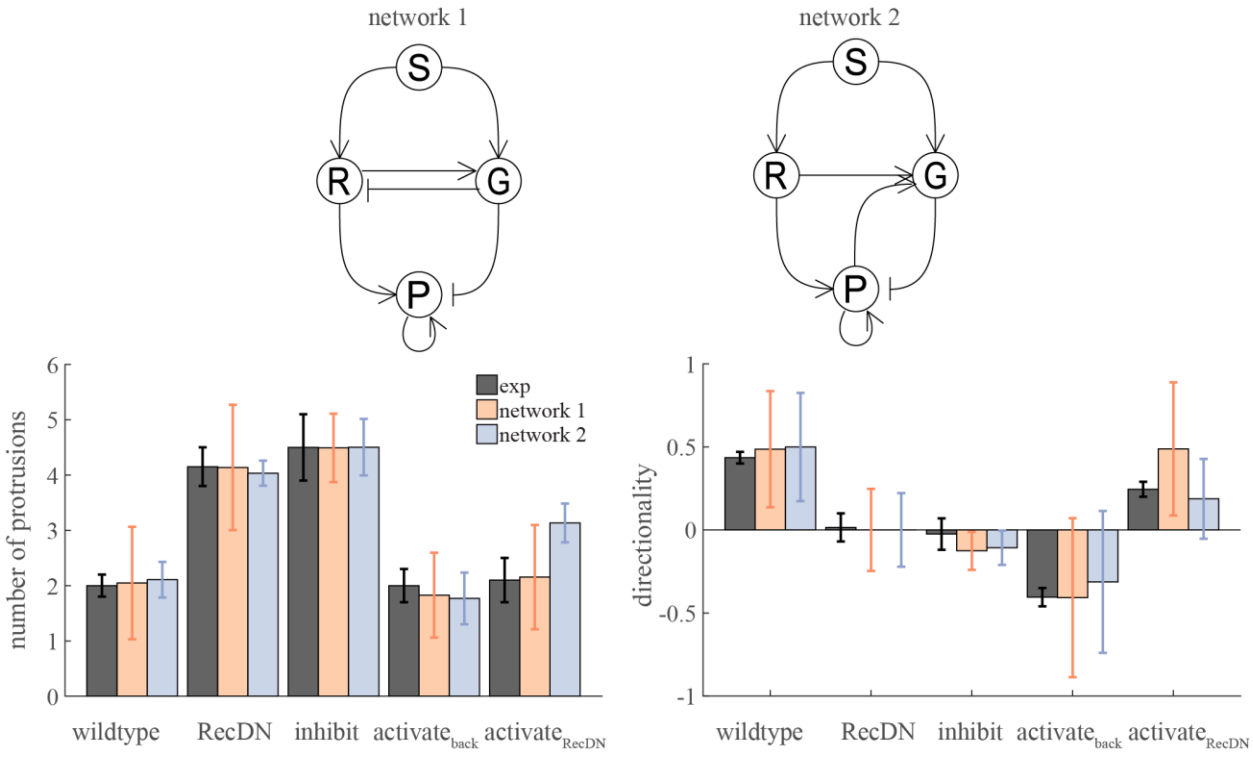

Fig. S6. Simulation results for two networks with 7 interactions.

SUPPORTING REFERENCES

1. Kampen, N.G. Van. 2011. Stochastic Processes in Physics and Chemistry. Elsevier.

2. Press, W.H. 2007. Numerical recipes 3rd edition: The art of scientific computing. Cambridge University Press.