Biophysical Journal, Volume *113*

Supplemental Information

Untangling the Hairball: Fitness-Based Asymptotic Reduction of Biological Networks

Félix Proulx-Giraldeau, Thomas J. Rademaker, and Paul François

Supplement

1 Description of $\bar{\phi}$

In the following, we illustrate $\bar{\phi}$'s MATLAB implementation. The MATLAB code is available in the Supplementary Materials.

1.1 MATLAB Algorithm

Seven functions are used:
runLoop
paramMODELtype
calcFitness
odeMODELtype
evalLim
updateParam
catch_problems

runLoop is the main script. paramMODELtype and odeMODELtype define parameters and associated ODEs which are problem specific. Parameters associated to the model are initially stored in variable *default*, then later modified parameters are stored in variable *param* and the list of removed parameters is stored in variable *removed*

A flowchart of the algorithm is presented in Fig. S1. In the following five steps we probe (1 - 3), rank, select, evaluate, accept (4), reduce and repeat (5).

1. Assign the parameter vector (PV) that paramMODELtype returns to *default*. This point in parameter space is going to be probed.

2. The fitness landscape around the initial PV *default* is characterized by the symmetric matrix *fitnessmap*, containing the fitness for modified parameters or couples of parameters. The fitness function, on the contrary, is problem specific, and is computed by calcFitness. Row by row, *fitnessmap* is filled by multiplying/dividing the parameters per entry by a reacaling *factor* ($f = 10$ from the main text). The performance of each of these entries is measured by computing the fitness with the new parameter combination relative to the initial fitness. A network with N parameters has $2N^2$ independent entries.

3. Removing a parameter is done in evalLim. With an estimate of the fitness landscape at hand found via the previous steps, the algorithm takes the corresponding limit (to 0 or ∞) for the parameters that were rescaled by f. We consider only changes of parameters giving identical or improved fitness. There exist four groups of two parameter limits θ_i, θ_j. In Tab. 1, the groups are presented in order of importance. When several couples of parameters give favorable changes to the fitness, we evaluate the limit of all couples that fall in group 1 one by one.

4. When we encounter a parameter limit in which the fitness is improved, we eliminate corresponding parameters and return to step 1. If for none of the couples in the parameter limits the fitness is improved, we move to the members of group 2, the limits to infinity, and similarly when we find a parameter limit that improves the fitness, we reduce and move on. Otherwise we move to the parameter limits of groups 3 and, finally to group 4 with the same criteria. This natural order shows our preference for removing parameters one by one (set parameters values to zero), instead of simply rescaling them (as products). Notice that we take a very conservative approach where fitness can only be incrementally improved with this procedure.

The steps in evalLim are the following.

A Find the least nonnegative elements in *fitnessmap*

B Divide these in the groups defined above

C Pick a random element from the highest ranked nonempty group

D updateParam takes the PV *default* and a 2×2 block of *removed* as arguments and returns an updated PV to *param.*

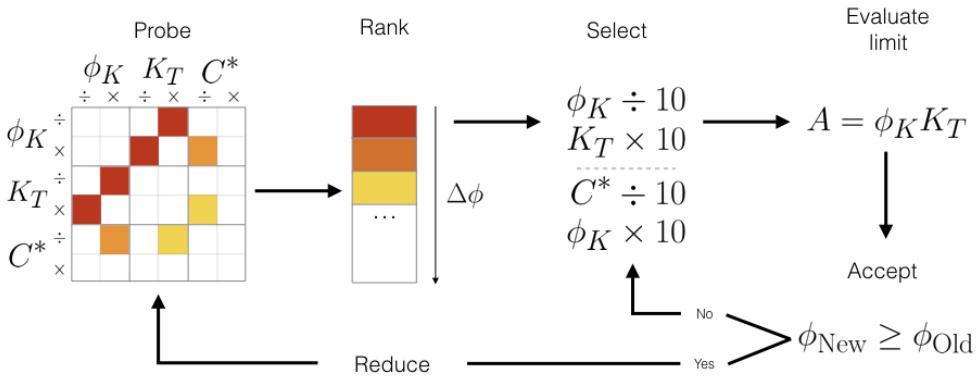

Figure S 1: A flowchart of the algorithm.

E Compute a temporary fitness ϕ_{new} with *param*.

F Decide as follows:
If $\phi_{\mathrm{new}} \geq \phi_{\mathrm{init}}$.
Accept removal
Return *param* and *removed*
If $\phi_{\mathrm{new}} < \phi_{\mathrm{init}}$.
Reject removal
Set *fitnessmap*(picked element) to inf .
Repeat cycle at step A

The method we use to take asymptotic limits is described in the next section.

5. The returned PV becomes the new initial point in an $(N-1)$-dimensional plane that is embedded in N-dimensional parameter space. Around this new initial point, we will probe the fitness landscape in the next round. In *removed*, the removed parameters and their limits are stored such that $\bar{\phi}$ ignores directions of reduced parameters in subsequent rounds.

This procedure is repeated until there are no free parameters left, or until all directions will decrease the fitness.

Table 1: Four groups of two-parameter limits

Group	Operation	Corresponding Limit taken
1	Division of two parameters by f	$(\theta_i, \theta_j) \to 0$
2	Multiplication of two parameters by f	$(\theta_i, \theta_j) \to \infty$
3	Division/multiplication by f	$\theta_i \to 0, \theta_j \to \infty$
4	Division/multiplication by f	Rescaling keeping product $\theta_i \cdot \theta_j =$ constant

1.2 Taking asymptotic limits

There are two kinds of asymptotic limits: parameters are either taken to 0 or to ∞. The 0 case is trivial to deal with: when a parameter is chosen to be 0, we simply put and maintain it to 0 in the subsequent steps of $\bar{\phi}$.

In evaluating a limit to infinity, one cannot simply numerically set this parameter to infinity, like in the case of a zero-limit. Instead, we consider a limit where this parameter is increased to such an extend that it dominates other terms in a sum that affect the same variable; these other terms are then removed from the equations. More precisely, consider the following equation:

$$\dot{y}_2 = ay_1 - (b + c + dy_1)y_2. \tag{1}$$

In the limit of $b \to \infty$ we replace this equation by the following differential equation:

$$\dot{y}_2 = ay_1 - by_2, \tag{2}$$

where $b \to b' = fb$, where f is our multiplicative factor defined in the previous section. This implements the idea that the c and dy_1 terms are negligible compared to b.

It is important to define a vector of parameter coefficients to keep track of these infinities. The vector of coefficients is attached to the parameter vector and updated in updateParam similarly. When the limit of a parameter is taken to infinity, its coefficient becomes zero, and the other terms in the sum will disappear. Practically, Eq. 1 is rewritten as

$$\dot{y}_2 = c_d a y_1 - (c_c c_d b + c_b c_d c + c_a c_b c_c d y_1)y_2. \tag{3}$$

The coefficients $c_{a,b,c,d}$ are initially set to 1. After evaluating the limit of $b \to \infty$, we set $c_b = 0$, and the simplification from Eq. 1 to Eq. 2 indeed takes place.

This can however create mass conservation problems in the rate equations. Consider the following equations for $\dot{y}_4$ and $\dot{y}_5$ where y_4 is turned into y_5 with rate r

$$\begin{aligned} \dot{y}_4 &= ay_3 - (r+q)y_4 \\ \dot{y}_5 &= ry_4 - dy_5 \end{aligned} \tag{4}$$

In the limit where parameter $q \to \infty$, parameter r will disappear from the equation of $\dot{y}_4$ potentially creating a divergence in the equations. A way to circumvent this is to impose global mass conservation: situations where y_4 is turned into y_5 correspond to signalling cascades where complexes are transformed into one another, so that we can impose that the total quantity of complex is conserved. This effectively adds a compensating term to the cascade. We also explicitly control for divergences and discard parameter sets for which variables diverge.

1.3 Choice of the path in parameter space

As shown in Section 1.1, the matrix *fitnessmap* is analyzed in the function evalLim. This matrix is symmetrical since the upper triangular part of the matrix corresponding to parameters (k_1, k_2) and the lower triangular part corresponding to parameters (k_2, k_1) give similar limits for groups 2 and 4 in Tab. 1. When given the choice between sending $(k_1, k_2) \to \infty$ or $(k_2, k_1) \to \infty$, FIBAR chooses randomly between the two, because the parameter combinations have the same change in fitness and in both cases a new parameter k_1/k_2 can be identified. However, because of FIBAR's design, choosing one will result in a different exploration of parameter space in the remaining steps. By choosing the first parameter combination, $\bar{\phi}$ will effectively freeze k_1 but allows $\bar{\phi}$ to keep exploring the logarithmic neighborhood of k_2. If the second combination is chosen, then the value of k_2 is frozen and it is the neighborhood of k_1 that will be probed. k_2 and k_1 may be present in different equations in the model, resulting in two not necessarily converging reductions.

A choice thus needs to be made in the final parameter reduced model. This allows for introduction of some kind of stochasticity in the produced networks in order to identify recurring patterns in the reduction. It can be a challenge in terms of reproducibility. One way to solve this problem is to set

a fixed rule in the function evalLim (using variable seed) which is called the deterministic method in the main article. The method of choice (random or deterministic) is left at the discretion of the user. We indeed see differences in the way networks are reduced, but the final structure of the reduced networks in all these cases can easily be mapped onto one another as described in the main text.

2 Mathematical definitions of fitness

In this section, we give mathematical definitions of the fitness functions used for both problems presented in the main text

2.1 Biochemical adaptation

For biochemical adaptation, we use a fitness very similar to the one proposed in [1].

We take as an initial input $I = 0.5$, then integrate for 1000 units of time and shift I to 2. After waiting another 1000 time steps, we measure ΔO_{ss} and ΔO_{max} as indicated in Figure 2 in the main text, and we take as a fitness $\Delta O_{max} + \frac{\epsilon}{\Delta O_{ss}}$ with $\epsilon = 0.01$ for the NFBL model and $\epsilon = 0.001$ for the first variant of the reduction of the IFFL model and $\epsilon = 0.01$ for the second variant.

2.2 Absolute discrimination

For absolute discrimination, we use a mutual information score, very similar to the one proposed in [2].

Imagine that some ligands with concentration L and binding time τ are presented to a cell. For ligand concentrations chosen from a specified distribution, we can now compute the typical output distribution $p_\tau(O)$ for a given τ. We consider log-uniformly distributed Input concentrations, integrate the system of equations, and generate a histogram of output values O corresponding to the input L. We take this as our approximation of $p_\tau(O)$.

Absolute discrimination works when the typical output values O across the range of L are unique for different τ. Intuitively, this means that the output distributions for two binding times should overlap as little as possible, as illustrated in Fig. 2. We use these distributions for two values of τ

to define marginal probability distributions $p_{\tau_i}(O) = p(O|\tau_i)$. Lastly, we consider equiprobable τ_is, and define $p(\tau_i, O) = p(O|\tau_i)p(\tau_i) = \frac{1}{2}p_{\tau_i}(O)$ and compute the mutual information between O and τ as

$$\mathcal{I}(O, \tau) = H(O) + H(\tau) - H(O, \tau) \tag{5}$$

where H is the classical Shannon entropy $H(O, \tau) = -\sum_{i,O} p(\tau_i, O) \log p(\tau_i, O)$.

Mutual information measures how much information we can recover from one variable knowing the other. For instance, when $\mathcal{I}(O, \tau) = 0$, it means we cannot recover information on the value of τ by observing O, which would be the case when both distributions are equal $p_{\tau_1}(O) = p_{\tau_2}(O)$. Conversely, when the two distributions are fully separated, this means we can fully recover τ by observing O. Thus, the the mutual information is at its maximum of 1 *bit*. For partially overlapping distributions, the mutual information varies gradually between 0 to 1. The choice of mutual information allows us to focus only on the respective positions of the distributions, and not on their shape, average values, etc... This allows us to focus on the discriminatory phenotype irrespective of other parameters. We very often obtain peaked distribution in O corresponding to horizontal lines (as in the lower panel of Fig. 2); the absolute level of these lines is arbitrary.

During the reduction, we typically sampled 50 log-uniformly distributed L on the interval $[1, 10^4]$ and binned the resulting outputs O in 40 log-uniformly distributed bins in the range $[10^{-2}, 10^2]$. The results are largely independent from the number of bins or the range of the bins, as long as O remains in the neighborhood of biologically feasible values, the working range of the initial networks. Partly due to this loose constraint, the output of the reduced networks was near the output of the initial networks.

3 Reductions of biochemical adaptation

In order to reproduce Michaelis-Menten kinetics for the Negative Feedback Loop and the Incoherent Feedforward Loop models from [3], we consider a model of linear ODEs where the dynamics of a node are mediated by an intermediary enzyme. In the models from [3], the production rates are of the form

$$rate_X = \frac{1 - X}{K + 1 - X} \tag{6}$$

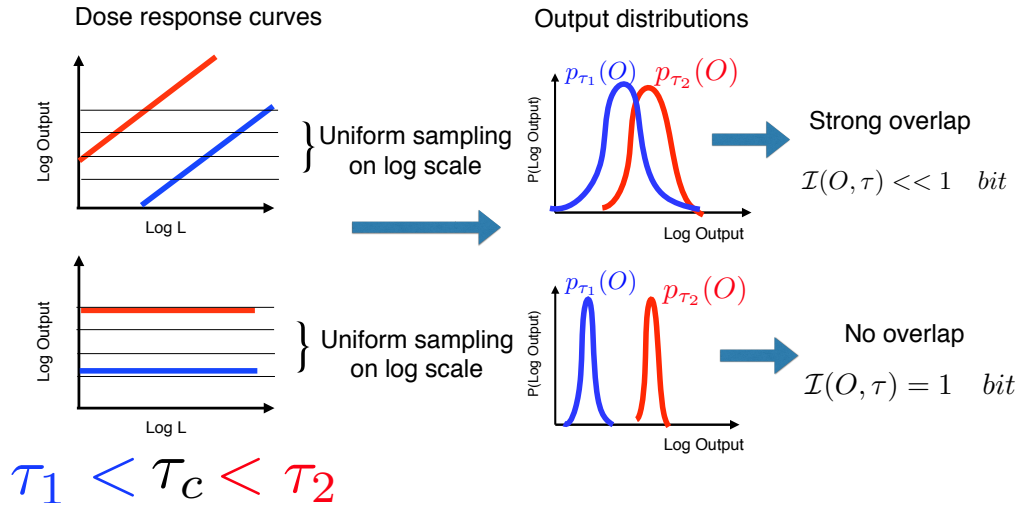

Figure S 2: Illustration of the mutual information score used for absolute discrimination. Upon sampling L from a log-uniform distribution, we integrate the equations, then compute histograms $p_{\tau_i}(O)$ for τ_1 and τ_2. From those joint distributions we compute the mutual information $\mathcal{I}(O,\tau)$. If the distribution overlap, it is not possible to distinguish well between the two τ_i, which means $\mathcal{I}(O,\tau)$ is low. If the distributions do not overlap, we can unambigously determine τ given O, thus $\mathcal{I}(O,\tau) = 1\, bit$.

The problem with this is that this imposes $X < 1$, which is not enforced by construction in [3]. To circumvent this difficulty, we slightly change their model by introducing a production node P_X and degradation node D_X to each node X. Both P_X and D_X are degraded by X. The quasi-static value for the production (P_X^{eq}) and degradation (D_X^{eq}) nodes are:

$$P_X^{eq} = \frac{K}{K+X}, \quad D_X^{eq} = \frac{K}{K+X} \tag{7}$$

so that the full equation for X is given by

$$\dot{X} = k_p P_X - k_d D_X X = k_p \frac{K}{K+X} - k_d X \frac{K}{K+X} \tag{8}$$

with k_d and k_p rates, which are potentially modulated by other proteins. Notice that we have included linear X dependency in the degradation. In particular, addition of P_X ensures that the production rate is a decreasing function of X, as was hypothesized in [3].

4 Negative feedback loop

Initial equations for the negative feedback loop model are given by

$$\begin{aligned}
\dot{A} &= k_1 I\, P_A - k_2 A B D_A \\
\dot{P}_A &= k_3(1 - P_A) - k_4 A P_A \\
\dot{D}_A &= k_5(1 - D_A) - k_6 A D_A \\
\dot{B} &= k_7 A\, P_B - k_8 B D_B \\
\dot{P}_B &= k_9(1 - P_B) - k_{10} B P_B \\
\dot{D}_B &= k_{11}(1 - D_B) - k_{12} B D_B.
\end{aligned}$$

Initial parameters are given in Tab. 2 and steps of the reduction of this model are given in Tab. 3.

Table 2: Negative feedback loop initial parameters

Parameter	Value
k_1	2
k_2	1
k_3	2
k_4	1
k_5	2
k_6	2
k_7	2
k_8	1
k_9	1
k_{10}	2
k_{11}	1
k_{12}	2

5 Incoherent feedforward foop

For this model, we got two kinds of reductions and present two variants. Initial equations for the incoherent feedforward loop model are

$$\begin{aligned}
\dot{A} &= k_1 I\, P_A - k_2 A D_A \\
\dot{P}_A &= k_3(1 - P_A) - k_4 A P_A \\
\dot{D}_A &= k_5(1 - D_A) - k_6 A D_A \\
\dot{B} &= k_7 A\, P_B - k_8 B D_B \\
\dot{P}_B &= k_9(1 - P_B) - k_{10} B P_B \\
\dot{D}_B &= k_{11}(1 - D_B) - k_{12} B D_B \\
\dot{C} &= k_{13} A\, P_C - k_{14} C B D_C \\
\dot{P}_C &= k_{15}(1 - P_C) - k_{16} C P_C \\
\dot{D}_C &= k_{17}(1 - D_C) - k_{18} C D_C
\end{aligned}$$

Initial parameters for the first variant of this model are given in Tab. 4. Steps and equations of the first variant of the reduction of this model are given in Tab. 5. Initial parameters for the second variant of this model are given in Tab. 6.

Table 3: Negative feedback loop reduction

Step	ϕ_{init}	Parameters		Limit	Description per group
1	1.6794	(k_{12}, k_8)	$\rightarrow$	∞	$D_B = k_{11}/k_{12}B$
2	1.6951	(k_{11}, k_8)	$\rightarrow$	$(0, \infty)$	$\Gamma_1 = k_7 k_{11}/k_{12}$
3	1.6957	(k_4, k_9)	$\rightarrow$	$0, \infty$	$P_A = 1$ and $P_B = 1$
4	54.4743	k_{10}	$\rightarrow$	0	
5	54.4743	(k_5, k_6)	$\rightarrow$	∞	
6	54.5648	k_3	$\rightarrow$	∞	
7	54.6099	(k_1, k_2)	$\rightarrow$	∞	
			FINAL OUTPUT		$A\|_{B_{eq}} = \frac{\Gamma_1}{k_7} = \frac{k_8 k_{11}}{k_7 k_{12}}$

Steps of the second variant of the reduction of this model are given in Tab. 7. For this variant, we took $\epsilon = 0.01$ in the fitness function and slightly different initial parameters. Final equations for the second variant of the IFFL reduction are given by:

$$\begin{aligned}
\dot{A} &= k_1 I\, P_A - k_2 A \\
\dot{P}_A &= k_3(1 - P_A) - k_4 A P_A \\
\dot{B} &= k_7 A - k_8 B \\
\dot{C} &= k_{13} A - k_{14} B C
\end{aligned}$$

The main difference between the two variants is on the equations for B and C. On the example of Fig. 3B in the paper, B and C regulate their own production rate, and B titrates C, while on the system above, and B regulates the degradation rate of C.

6 Adaptive Sorting

We perform parameter reduction on the Adaptative Sorting model without any symmetry breaking process. Initial equations for the adaptive sorting model are given by

Table 4: Incoherent feedforward loop first variant initial parameters

Parameter	Value
k_1	2
k_2	0.5
k_3	3
k_4	1
k_5	2
k_6	3
k_7	2
k_8	1
k_9	3
k_{10}	1
k_{11}	2
k_{12}	1
k_{13}	2
k_{14}	2
k_{15}	3
k_{16}	2
k_{17}	1
k_{18}	3

Table 5: Incoherent feedforward loop first variant

Step	ϕ_{init}	Parameters		Limit	Description per group
1	0.5958	(k_{14}, k_{13})	$\to$	∞	
2	0.5960	(k_6, k_5)	$\to$	∞	
3	0.5961	(k_{15}, k_{16})	$\to$	∞	
4	0.5962	(k_{10}, k_9)	$\to$	∞	
5	0.5971	(k_{13}, k_{18})	$\to$	$(0, \infty)$	$D_C = k_{17}/k_{18}C$
6	0.6163	(k_{11}, k_8)	$\to$	$(0, \infty)$	$D_B = k_{11}/k_{12}B$
7	0.6169	(k_{18}, k_{17})	$\to$	∞	
8	0.6177	(k_5, k_2)	$\to$	$(0, \infty)$	$D_A = k_5/k_6A$
9	0.6699	(k_{17}, k_{16})	$\to$	$(0, \infty)$	$P_C = k_{15}/k_{16}C$
10	0.7144	(k_9, k_7)	$\to$	$(0, \infty)$	$P_B = k_9/k_{10}B$
11	0.8544	(k_4, k_1)	$\to$	∞	$P_A = k_3/k_4A$
12	0.8670	(k_3, k_1)	$\to$	$(0, \infty)$	
13	0.8698	(k_{16}, k_7)	$\to$	$(0, \infty)$	
14	1.6083	(k_{12}, k_8)	$\to$	∞	
15	1.6094	(k_8, k_7)	$\to$	∞	
16	2.2361	(k_2, k_1)	$\to$	∞	
			FINAL OUTPUT		$C_{eq} = \frac{k_8 k_{10} k_{11} k_{13} k_{15} k_{18}}{k_7 k_9 k_{12} k_{14} k_{16} k_{17}}$

$$
\begin{aligned}
\dot{K} &= \beta(K_T - K) - \alpha K C_0 \\
\dot{C_0} &= \kappa(L - \sum_i C_i)(R - \sum_i C_i) + bC_1 - (\phi K + \tau^{-1})C_0 \\
\dot{C_1} &= \phi K C_0 - (\tau^{-1} + b)C_1.
\end{aligned}
$$

Initial parameters are given in Tab. 8. Steps of the reduction of this model are given in Tab. 9.

7 SHP-1 model

7.1 SHP-1 model first reduction

We first perform parameter reduction on the SHP-1 model with parameter symmetry breaking. Initial equations for the SHP-1 model are given by

Table 6: Incoherent feedforward loop second variant initial parameters

Parameter	Value
k_1	3
k_2	0.5
k_3	3
k_4	1
k_5	2
k_6	3
k_7	5
k_8	1
k_9	3.5
k_{10}	1
k_{11}	5
k_{12}	1
k_{13}	2
k_{14}	1
k_{15}	3
k_{16}	2
k_{17}	1
k_{18}	3

Table 7: Incoherent feedforward loop second variant

Step	ϕ_{init}	Parameters		Limit	Description per group
1	0.8788	(k_{14}, k_{13})	$\to$	∞	
2	0.8789	(k_5, k_6)	$\to$	∞	
3	0.8793	(k_9, k_{10})	$\to$	∞	
4	0.8802	(k_3, k_4)	$\to$	∞	
5	0.8811	(k_4, k_1)	$\to$	∞	
6	0.8814	(k_{16}, k_{15})	$\to$	∞	
7	0.8822	(k_{18}, k_{15})	$\to$	$0, \infty$	$P_C = D_C = 1$
8	1.0372	(k_{11}, k_{12})	$\to$	∞	
9	1.0375	k_{17}	$\to$	∞	
10	1.0378	(k_2, k_1)	$\to$	∞	
11	1.3659	(k_6, k_8)	$\to$	$0, \infty$	$D_A = 1$
12	1.3736	k_{12}	$\to$	0	$D_B = 1$
13	1.4156	(k_{10}, k_1)	$\to$	$0, \infty$	$P_B = 1$
			FINAL OUTPUT		$C_{eq} = \frac{k_8 k_{13}}{k_7 k_{14}}$

$$
\begin{aligned}
\dot{S} &= \alpha C_1(S_T - S) - \beta S \\
\dot{C}_0 &= \kappa(L - \sum_i C_i)(R - \sum_i C_i) + \gamma_1 S C_1 - (\phi_1 + \tau^{-1})C_0 \\
\dot{C}_1 &= \phi_1 C_0 + \gamma_2 S C_2 - (\gamma_1 S + \phi_2 + \tau^{-1})C_1 \\
\dot{C}_2 &= \phi_2 C_1 + \gamma_3 S C_3 - (\gamma_2 S + \phi_3 + \tau^{-1})C_2 \\
\dot{C}_3 &= \phi_3 C_2 + \gamma_4 S C_4 - (\gamma_3 S + \phi_4 + \tau^{-1})C_3 \\
\dot{C}_4 &= \phi_4 C_3 + \gamma_5 S C_5 - (\gamma_4 S + \phi_5 + \tau^{-1})C_4 \\
\dot{C}_5 &= \phi_5 C_4 - (\gamma_5 S + \tau^{-1})C_5
\end{aligned}
$$

Initial parameters for this model are given in Tab. 10. Steps of the first reduction of this model are given in Tab. 11. The final system is given by the following equations when the reduction steps of Tab. 11 are applied.

As an example of how the simpler model is contained into the full system of equations, but with some zero parameters, this reduction corresponds to the simplified system

Table 8: Adaptative sorting initial parameters

Parameter	Value
ϕ	3×10^{-4}
K_T	10^3
α	1
β	1
κ	10^{-4}
R	10^4
b	5×10^{-2}

Table 9: Adaptive sorting

Step	I_{init}	Parameters		Limit	Description per group
1	0.8131	(α, β)	$\to$	∞	$C^* = \beta/\alpha$
2	0.8131	(K_T, ϕ)	$\to$	$(0, \infty)$	$A = \phi K_T$
3	0.8131	(κ, R)	$\to$	$(0, \infty)$	$\kappa R \to \infty$
4	0.8131	R	$\to$	∞	
5	0.8645	(C^*, A)	$\to$	$(0, \infty)$	$\lambda = AC^*$
6	1	α	$\to$	∞	To undo the effect $C_1 \propto L$ for $L \leq 2$
7	1	b	$\to$	0	Uncluttering τ
		FINAL OUTPUT			$C_1 = \lambda\tau = \phi K_T \beta\tau/\alpha$

$$\dot{S} = \alpha C_1 S_T - \beta S \tag{9}$$

$$\dot{C}_0 = \kappa(L - \sum_i C_i)R + \gamma_1 S C_1 - \phi_1 C_0 \tag{10}$$

$$\dot{C}_1 = \phi_1 C_0 - (\phi_2 + \gamma_1 S)C_1 \tag{11}$$

$$\dot{C}_2 = \phi_2 C_1 - \phi_3 C_2 \tag{12}$$

$$\dot{C}_3 = \phi_3 C_2 - \phi_4 C_3 \tag{13}$$

$$\dot{C}_4 = \phi_4 C_3 + \gamma_5 S C_5 - (\phi_5 + \tau^{-1})C_4 \tag{14}$$

$$\dot{C}_5 = \phi_5 C_4 - \gamma_5 S C_5 \tag{15}$$

Eqs. 14 and 15 correspond to Eqs. 2 and 3 in the main text, with proportionality of S to C_1 given by Eq. 9 and proportionality of C_3 to C_1 by Eqs. 12 and 13.

Table 10: SHP-1 model initial parameters

Parameter	Value
ϕ	9×10^{-2}
γ	1
S_T	7.2×10^{-1}
β	3×10^2
α	1
$\beta/\alpha = C^*$	3×10^2
κ	10^{-4}
R	3×10^4

Table 11: SHP-1 First reduction

Step	I_{init}	Parameters		Limit	Description per group
1	0.7369	(κ, R)	$\to$	$(0, \infty)$	
2	0.7369	γ_1	$\to$	∞	
3	0.8468	(ϕ_2, ϕ_1)	$\to$	$(0, \infty)$	
4	0.8583	R	$\to$	∞	
5	0.8583	(γ_4, γ_5)	$\to$	$0, \infty$	Kinetic sensing module
6	1.0000	γ_2	$\to$	0	Uncluttering τ
7	1.0000	γ_3	$\to$	0	
8	1.0000	(ϕ_3, S_T)	$\to$	∞	Rescaling
9	1.0000	(ϕ_1, ϕ_4)	$\to$	∞	
10	1.0000	β	$\to$	∞	Adaptation module
11	1.0000	(ϕ_4, S_T)	$\to$	∞	
12	1.0000	(S_T, α)	$\to$	$(0, \infty)$	
		FINAL OUTPUT			$C_5 = \frac{\phi_2 \phi_5 \beta}{\gamma_5 S_T \alpha} \tau$

7.2 SHP-1 model second reduction

We then perform another reduction of the same model using a different binning for the computation of the mutual information. Initial parameters and equations are identical as in the previous reduction presented in section 7.1. Steps for this reduction are given in Tab. 12.

The final system is given by the following equations when the reduction steps given in Tab. 12 are applied.

Table 12: SHP-1 Second reduction

Step	I_{init}	Parameters		Limit	Description per group
1	0.6328	(κ, R)	$\rightarrow$	$(0, \infty)$	
2	0.6328	R	$\rightarrow$	∞	
3	0.6375	(γ_4, α)	$\rightarrow$	$(0, \infty)$	
4	0.6464	(γ_2, γ_1)	$\rightarrow$	$0, \infty$	
5	0.7264	γ_5	$\rightarrow$	∞	Adaptive module
6	1.0000	γ_3	$\rightarrow$	0	
7	1.0000	(ϕ_1, β)	$\rightarrow$	∞	
8	1.0000	ϕ_4	$\rightarrow$	∞	Kinetic sensing module
9	1.0000	(ϕ_3, S_T)	$\rightarrow$	∞	
10	1.0000	(S_T, β)	$\rightarrow$	∞	
11	1.0000	(ϕ_5, β)	$\rightarrow$	$(0, \infty)$	
12	1.0000	(β, ϕ_2)	$\rightarrow$	$(0, \infty)$	
			FINAL OUTPUT		$C_5 = \frac{\phi_2 \phi_5 \beta}{\gamma_5 S_T \alpha} \tau$

$$\dot{S} = \alpha C_1 S_T - \beta S \tag{16}$$

$$\dot{C}_0 = \kappa R(L - \sum_i C_i) + \gamma_1 S C_1 - \phi_1 C_0 \tag{17}$$

$$\dot{C}_1 = \phi_1 C_0 - (\phi_2 + \gamma_1 S) C_1 \tag{18}$$

$$\dot{C}_2 = \phi_2 C_1 - \phi_3 C_2 \tag{19}$$

$$\dot{C}_3 = \phi_3 C_2 + \gamma_4 S C_4 - \phi_4 C_3 \tag{20}$$

$$\dot{C}_4 = \phi_4 C_3 + \gamma_5 S C_5 - (\phi_5 + \gamma_4 S + \tau^{-1}) C_4 \tag{21}$$

$$\dot{C}_5 = \phi_5 C_4 - \gamma_5 S C_5. \tag{22}$$

7.3 SHP-1 model third reduction

We perform another reduction of the same model using slightly different initial parameter values. All parameters are given in Tab. 10 with $S_T \rightarrow 5S_T$. Initial set of equations is identical as in Sections 7.1 and 7.2. Steps for this reduction are given in Tab. 13.

The final system is given by the following equations when the reduction steps given in Tab. 13 are applied.

Table 13: SHP-1 Third reduction

Step	I_{init}	Parameters		Limit	Description per group
1	0.4946	(β, α)	$\rightarrow$	∞	
2	0.4946	(R, κ)	$\rightarrow$	$(0, \infty)$	
3	0.4946	(γ_1, γ_5)	$\rightarrow$	0	Kinetic sensing module
4	1.0000	(κ, ϕ_1)	$\rightarrow$	∞	
5	1.0000	ϕ_1	$\rightarrow$	∞	
6	1.0000	(ϕ_2, ϕ_4)	$\rightarrow$	$(0, \infty)$	
7	1.0000	(ϕ_5, γ_3)	$\rightarrow$	∞	Adaptation module
8	1.0000	γ_2	$\rightarrow$	0	
9	1.0000	S_T	$\rightarrow$	∞	
10	1.0000	γ_4	$\rightarrow$	0	
11	1.0000	(ϕ_4, ϕ_3)	$\rightarrow$	$(0, \infty)$	Rescaling
12	1.0000	(α, γ_3)	$\rightarrow$	$(0, \infty)$	
			FINAL OUTPUT		$C_5 = \frac{\phi_2\phi_3\phi_4\beta}{\gamma_3 S_T \alpha}\tau^2$

$$
\begin{aligned}
\dot{S} &= \alpha C_1 S_T - \beta S \\
\dot{C}_0 &= \kappa(L - \sum_i C_i)(R - \sum_i C_i) - \phi_1 C_0 \\
\dot{C}_1 &= \phi_1 C_0 - (\phi_2 + \tau^{-1}) C_1 \\
\dot{C}_2 &= \phi_2 C_1 + \gamma_3 S C_3 - (\phi_3 + \tau^{-1}) C_2 \\
\dot{C}_3 &= \phi_3 C_2 - (\phi_4 + \gamma_3 S) C_3 \\
\dot{C}_4 &= \phi_4 C_3 - \phi_5 C_4 \\
\dot{C}_5 &= \phi_5 C_4 - \tau^{-1} C_5.
\end{aligned}
$$

7.4 SHP-1 model reduction without feedback

We perform a reduction of the SHP-1 model with the SHP-1 mediated feedback turned off. Parameter values are given in Tab. 10 with $S_T = 0$. The network topology is as in Fig. S 3A and the corresponding initial set of equations is identical as in Sections 7.1 and 7.2. Fig. S 3B shows that the reduction does not converge when crucial network elements (SHP-1 feedback) are missing.

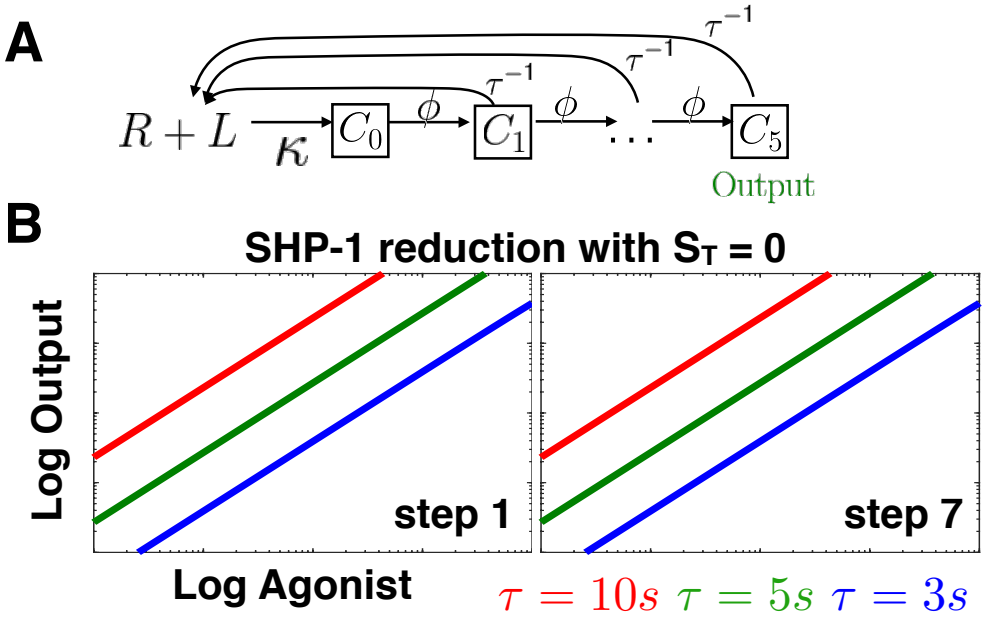

Figure S 3: "Negative control" for SHP-1 model: we attempt to reduce this model with $\bar{\phi}$ in absence of SHP-1 (corresponding to pure kinetic proofreading). The algorithm fails to optimize behavior and fitness, indicating that it is not possible to do so for arbitrary networks.

7.5 Analytical study

The full analytical study of this model is done in [4]. Assuming all $\phi_i = \phi$ and $\gamma_i = \gamma$ are equal, we get at lowest order

$$C_1 \simeq r_-(1 - r_-)\frac{\kappa RL}{\kappa R + \nu_1} \tag{23}$$

with

$$r_\pm = \frac{\phi + S + \nu_1 \pm \sqrt{(\phi + S + \nu_1)^2 - 4\phi S}}{2S} \tag{24}$$

.

We can use the previous expression to get a closed equation for S as a function of $r_-(S)$ and C_*.

$$S = S_T\frac{C_1}{C_1 + C_*} = S_T\frac{r_-(1 - r_-)}{r_-(1 - r_-) + \frac{C_*(\kappa R + \nu_1)}{\kappa RL}} \tag{25}$$

This is a 4th order polynomial equation in S in terms of the parameters that can be conveniently solved numerically. Once this is done, we get the following expression for C_N, the final complex in the cascade as a function of $r_\pm$ to the lowest order in r_-^N.

$$C_N \simeq \frac{\kappa RL}{\kappa R + \nu_1}\left(1 - \frac{r_-}{r_+}\right) r_-^N \tag{26}$$

To see why this feedback hinders perfect adaptation, it is useful to consider the limit of big L and big S_T. In this limit, it is shown in [4] that the parameter r_- becomes inversely proportional to the feedback variable $1/S$, thus giving at lowest order a S^{-N} contribution in Eq. 26, clearly coming from the coupling of N identical proofreading steps. Those equations can be approximately solved [4] so that

$$C_N \simeq \left(\frac{\phi\beta}{\alpha\gamma S_T}\right)^{N/2} (L)^{1-N/2}. \tag{27}$$

So we see that, unless $N = 2$, there is an unavoidable L dependency. The $L^{-N/2}$ dependency comes from the steady state value of the feedback variable $S \propto L^{1/2}$ appearing when we fully close this system.

8 Lipniacki model

In this section, we present two parameter reductions performed by $\bar{\phi}$. Initial equations for the Lipniacki model are:

$$
\begin{aligned}
\dot{X}_2 &= b_1\, pMHC_{free}\, TCR_{free} + (s_2 + s_3)X_{23} - (lb\, LCK_{free} + ly_1 X_{29} + \tau^{-1})X_2 \\
\dot{X}_3 &= lb\, LCK_{free}\, X_2 + ls_1 X_4 + (s_2 + s_3)X_{24} - (ly_2 + X_{37} + ly_1 X_{29} + \tau^{-1})X_3 \\
\dot{X}_4 &= X_{37}X_3 - (ly_2 + ls_1 + \tau^{-1})X_4 \\
\dot{X}_5 &= ly_2 X_3 + ls_1 X_6 - (tp + X_{37} + ly_1 X_{29} + \tau^{-1})X_5 \\
\dot{X}_6 &= ly_2 X_4 + X_{37}X_5 - (tp + ls_1 + \tau^{-1})X_6 \\
\dot{X}_7 &= tpX_5 + ls_1 X_8 - (tp + X_{37} + ly_1 X_{29} + \tau^{-1})X_7 \\
\dot{X}_8 &= tpX_6 + X_{37}X_7 - (tp + ls_1 + \tau^{-1})X_8 \\
\dot{X}_9 &= tpX_7 + ls_1 X_8 - (\tau^{-1} + X_{37} + ly_1 X_{29} + \tau^{-1})X_9 \\
\dot{X}_{10} &= tpX_8 + X_{37}X_9 - (ls_1 + \tau^{-1})X_{10} \\
\dot{X}_{22} &= ly_1 X_{29}\, TCR_{free} + \tau^{-1}(X_{23} + X_{24}) - (s_2 + s_3)X_{22} \\
\dot{X}_{23} &= ly_1 X_{29} X_2 - (s_2 + s_3 + \tau^{-1})X_{23} \\
\dot{X}_{24} &= ly_1 X_{29}(X_3 + X_5 + X_7 + X_9) - (s_2 + s_3 + \tau^{-1})X_{24} \\
\dot{X}_{29} &= s_1(X_5 + X_7 + X_9)\, SHP_{free} + s_3(X_{22} + X_{23} + X_{24}) + s_0\, SHP_{free} \\
&\quad - ly_1(X_2 + X_3 + X_5 + X_7 + X_9 + TCR_{free})X_{29} - s_2 X_{29} \\
\dot{X}_{31} &= z_1(X_9 + X_{10})(m_1 - X_{31}) + z_0 m_1 - (z_0 + z_2)X_{31} \\
\dot{X}_{33} &= 2X_{31}(e_1 - X_{34}) + 2m_2 X_{34} - (m_2 + 3X_{31})X_{33} \\
\dot{X}_{34} &= X_{31}X_{33} - 2m_2 X_{34} \\
\dot{X}_{36} &= 2X_{34}(ls_2 - X_{37}) + 2e_2 X_{37} - (e_2 + 2X_{34})X_{36} \\
\dot{X}_{37} &= X_{34}X_{36} - 2e_2 X_{37}
\end{aligned}
$$

To ensure physical behavior throughout the reduction process, we manually implement the following mass conservation laws.

$$
\begin{aligned}
pMHC_{free} &= pMHC - \left(\sum_{i=2}^{10} X_i + X_{23} + X_{24}\right) \\
TCR_{free} &= TCR - \left(\sum_{i=2}^{10} X_i + X_{22} + X_{23} + X_{24}\right) \\
LCK_{free} &= LCK - \left(\sum_{i=3}^{10} X_i + X_{24}\right) \\
SHP_{free} &= SHP - (X_{22} + X_{23} + X_{24} + X_{29}) \\
ZAP_{free} &= ZAP - X_{31} \\
MEK_{free} &= MEK - (X_{33} + X_{34}) \\
ERK_{free} &= ERK - (X_{36} + X_{37})
\end{aligned}
$$

We also perform initial rescaling of equations X_{31} to X_{37} to save $\bar{\phi}$ steps:

$$
\begin{aligned}
X_{31} &\to \frac{m_1 X_{31}}{ZAP} \\
X_{33} &\to \frac{e_1 X_{33}}{MEK} \\
X_{34} &\to \frac{e_1 X_{34}}{MEK} \\
X_{36} &\to \frac{ls_2 X_{36}}{ERK} \\
X_{37} &\to \frac{ls_2 X_{37}}{ERK}
\end{aligned}
$$

Initial parameters are given in Tab. 14.

8.1 Lipniacki model reduction: first variant

For this reduction, we used mutual information as a fitness function. We discarded all values of the output below the measurable threshold 10^{-2}, and used 40 log-uniformly distributed bins on the interval $[10^{-2}, 10^2]$ for the computation of the Output distribution. The Input concentrations were given by 50 log-uniformly distributed values on the interval $[1, 10^4]$.

Steps of the first biochemical reduction of this model (odeLIPbasic.m in the MATLAB code) are given in Tab. 15. The results of the biochemical reduction are given by

$$
\begin{aligned}
\dot{X}_2 &= b_1\, pMHC_{free}\, TCR + s_{22}X_{23} - lbX_2 \\
\dot{X}_3 &= lbX_2 + s_{23}X_{24} - ly_{21}X_3 \\
\dot{X}_4 &= X_{37}X_3 - ly_{22}X_4 \\
\dot{X}_5 &= ly_{21}X_3 - (tp_1 + X_{37} + ly_{14}X_{29} + \tau^{-1})X_5 \\
\dot{X}_6 &= ly_{22}X_4 + X_{37}X_5 - (tp_2 + \tau^{-1})X_6 \\
\dot{X}_7 &= tp_1X_5 - (tp_3 + X_{37} + ly_{15}X_{29} + \tau^{-1})X_7 \\
\dot{X}_8 &= tp_2X_6 + X_{37}X_7 - (tp_4 + \tau^{-1})X_8 \\
\dot{X}_9 &= tp_3X_7 - (\tau^{-1} + X_{37} + ly_{16}X_{29})X_9 \\
\dot{X}_{10} &= tp_4X_8 + X_{37}X_9 - \tau^{-1}X_{10} \\
\dot{X}_{23} &= ly_{12}X_{29}X_2 - (s_{22} + \tau^{-1})X_{23} \\
\dot{X}_{24} &= (ly_{13}X_3 + ly_{14}X_5 + ly_{15}X_7 + ly_{16}X_9)X_{29} - (s_{23} + \tau^{-1})X_{24} \\
\dot{X}_{29} &= s_{11}X_5 + s_{12}X_7 + s_{13}X_9 - ly_{11}TCR\, X_{29}.
\end{aligned}
$$

We then perform parameter symmetry breaking (odeLIPadvanced in the MATLAB code). Steps of reduction are given in Tab. 16.

Parameter symmetry breaking results in the following system that can be again extracted analytically and is a subset of the initial full system:

$$
\begin{aligned}
\dot{X}_2 &= b_1\, pMHC_{free}\, TCR - lbX_2 \\
\dot{X}_3 &= lbX_2 - ly_2X_3 \\
\dot{X}_5 &= ly_2X_3 - ly_{13}X_{29}X_5 \\
\dot{X}_7 &= tp_1X_5 - (tp_2 + ly_{14}X_{29} + \tau^{-1})X_7 \\
\dot{X}_9 &= tp_2X_7 - (ly_{15}X_{29} + \tau^{-1})X_9 \\
\dot{X}_{23} &= ly_{12}X_{29}X_2 - \tau^{-1}X_{23} \\
\dot{X}_{24} &= ly_{13}X_{29}X_5 - \tau^{-1}X_{24} \\
\dot{X}_{29} &= s_1X_5 - ly_{11}TCR\, X_{29}
\end{aligned}
$$

Only four more steps of reduction are needed to reach perfect adaptation, namely $(ly_{13}, ly_{15}) \to 0$, $(ly_{11}, ly_{14}) \to \infty$, $(tp_1, ly_{14}) \to \infty$ and finally $ly_{14} \to \infty$. We apply those steps of reduction by hand and reach the final following system.

$$\dot{X}_2 = b_1\, pMHC_{free}\, TCR - lbX_2 \tag{28}$$
$$\dot{X}_3 = lbX_2 - ly_2X_3 \tag{29}$$
$$\dot{X}_5 = ly_2X_3 - tp_1X_5 \tag{30}$$
$$\dot{X}_7 = tp_1X_5 - ly_{14}X_{29}X_7 - tp_2X_7 \tag{31}$$
$$\dot{X}_9 = tp_2X_7 - \tau^{-1}X_9 \tag{32}$$
$$\dot{X}_{24} = ly_{14}X_7X_{29} - \tau^{-1}X_{24} \tag{33}$$
$$\dot{X}_{29} = s_1X_5 - ly_{11}TCR\, X_{29} \tag{34}$$

8.2 Lipniacki model reduction: second variant

Initial equations, parameters, mass conservation laws and equation transformations for this reduction are the same as for the previous Lipniacki reduction. For this reduction, we chose mutual information as the fitness with 40 bins log-uniformly distributed on the interval $[10^{-2}, 10^2]$, plus a lower bin for concentrations below 10^{-2} and a higher bin for concentrations above 10^2. We chose 50 log-uniformly distributed Input concentrations on the interval $[1, 10^4]$. Because of the binning choice, the fitness, was optimized quicker, while most reduction took place in the neutral fitness landscape of maximum fitness of 1 *bit*. The details of this biochemical reduction are given in Tab. 17. After the first reduction, the system is reduced to

$$
\begin{aligned}
\dot{X}_2 &= b_1\, pMHC_{free}\, TCR_{free} + (s_{22} + s_{32})X_{23} - lb\, LCK\, X_2 \\
\dot{X}_3 &= lb\, LCK\, X_2 + ls_{11}X_4 + (s_{23} + s_{33})X_{24} - (ly_{21} + X_{37} + ly_{11}X_{29} + \tau^{-1})X_3 \\
\dot{X}_4 &= X_{37}X_3 - (ly_{22} + ls_{11} + \tau^{-1})X_4 \\
\dot{X}_5 &= ly_{21}X_3 + ls_{12}X_6 - (tp_1 + X_{37} + ly_{12}X_{29} + \tau^{-1})X_5 \\
\dot{X}_6 &= ly_{22}X_4 + X_{37}X_5 - (tp_2 + ls_{12} + \tau^{-1})X_6 \\
\dot{X}_7 &= tp_1X_5 + ls_{13}X_8 - (tp_3 + X_{37} + ly_{13}X_{29} + \tau^{-1})X_7 \\
\dot{X}_8 &= tp_2X_6 + X_{37}X_7 - (tp_4 + ls_{13} + \tau^{-1})X_8 \\
\dot{X}_9 &= tp_3X_7 + ls_{14}X_8 - (X_{37} + ly_{14}X_{29} + \tau^{-1})X_9 \\
\dot{X}_{10} &= tp_4X_8 + X_{37}X_9 - (ls_{14} + \tau^{-1})X_{10} \\
\dot{X}_{22} &= ly_{15}X_{29}\, TCR_{free} + \tau^{-1}(X_{23} + X_{24}) - (s_{21} + s_{31})X_{22} \\
\dot{X}_{23} &= ly_{16}X_{29}X_2 - (s_{22} + s_{32} + \tau^{-1})X_{23} \\
\dot{X}_{24} &= X_{29}(ly_{11}X_3 + ly_{12}X_5 + ly_{13}X_7 + ly_{14}X_9) - (s_{23} + s_{33} + \tau^{-1})X_{24} \\
\dot{X}_{29} &= (s_{11}X_5 + s_{12}X_7 + s_{13}X_9)\, SHP + (s_{31}X_{22} + s_{32}X_{23} + s_{33}X_{24}) \\
&\quad - (ly_{16}X_2 + ly_{11}X_3 + ly_{12}X_5 + ly_{13}X_7 + ly_{14}X_9 + ly_{15}TCR_{free})X_{29} - s_{24}X_{29} \\
X_{37} &= 0.05
\end{aligned}
$$

We then perform parameter symmetry breaking on this system. Steps of the biochemical reduction of this model are given in Tab. 18. We can remove equations for X_4, X_6, X_{23} and X_{24} as they are dead ends in the network. $X_{37} = 0.5$ is held constant. The final expression of the output given in Tab. 18 is extracted from remaining equations at steady-state; expanding the equations for the relevant cascade we get

$$\dot{X}_2 = b_1\, pMHC_{free}\, TCR_{free} - lb\, LCK\, X_2 \tag{35}$$
$$\dot{X}_3 = lb\, LCK\, X_2 - ly_{21} X_3 \tag{36}$$
$$\dot{X}_5 = ly_{21} X_3 - (tp_1 + ly_{12} X_{29}) X_5 \tag{37}$$
$$\dot{X}_7 = tp_1 X_5 - X_{37} X_7 \tag{38}$$
$$\dot{X}_8 = X_{37} X_7 - tp_4 X_8 \tag{39}$$
$$\dot{X}_{10} = tp_4 X_8 - \tau^{-1} X_{10} \tag{40}$$
$$\dot{X}_{22} = ly_{15}\, TCR_{free}\, X_{29} - s_{21} X_{22} \tag{41}$$
$$\dot{X}_{29} = s_{11}\, SHP\, X_5 - ly_{15} TCR_{free} X_{29} \tag{42}$$
$$X_{37} = 0.5 \tag{43}$$

The output here is X_{10}. Variables X_{22}, X_{29} and X_5 respectively correspond to variables R_p, S and C_5 in the main text. The structure of Eqs. 35 to 42 is clearly very similar to the equations of the previous reduction 28 to 34, with a linear cascade for the second reduction $X_2 \to X_3 \to X_5 \to X_7 \to X_8 \to X_{10}$ and $X_2 \to X_3 \to X_5 \to X_7 \to X_9$ for the first reduction, modulated by a parallel loop via X_{29} and X_5. As described in the main text, the structural difference comes from the mechanism of this loop, the first reduction giving an effective feedforward adaptive system, while the second reduction is an integral feedback mechanism.

9 Antagonism

The models we reduce have all captured the phenomenon of ligand antagonism, where the response of agonist ligands in the presence of high amounts of well chosen subthreshold ligands (i.e. with binding time lower than critical binding time τ_c triggering response) is antagonized. With our fitness, we have quantified and selected for absolute discrimination in the networks, and through the reduction, ligand antagonism has remained, but the hierarchy of antagonism has changed. In the simplest systems, antagonism is maximum for minimum τ, while for more complex models there maximum antagonism is reached closer to threshold τ_c (see discussion in [5]). It turns out we can recover this property by adding two terms to the final reduced equations.

An overview of antagonism is presented in Fig. S4. We draw the response line as a binary activation by choosing a threshold of the final output for ac-

tivation (we know from our previous works [2, 4] that adding stochasticity to have a more probabilistic view does not fundamentally change this picture). The earliest response (in terms of ligand concentration) always comes from the agonist alone. Note how T cells in Fig. S4 A presented to OVA agonists + strong G4 antagonists are activated at higher agonist concentration than the weak E1 antagonists, and G4 have a binding time close to threshold than E1. This hierarchy is typical for experimentally observed antagonism: antagonism strength is large just below τ_c, the critical binding time above which a response is elicited.

Similarly, in the full models for SHP-1 and Lipniacki (Fig. S4 B - C), we have the same hierarchy. However, for the same binding times in reduced SHP-1 (Fig. S4 E) and reduced Lipniacki (Fig. S4 F), we have an inverted hierarchy, where ligands further below are more antagonizing, so closer to the naive models discussed in [5].

It turns out that the position of the adaptive module m in the kinetic proofreading cascade of N complexes, defined as the complex on which the variable S implements the negative "tug-of-war" term described in the main text, determines the antagonism strength, like in Fig. 4 of ref. [5]. We can rescue the correct hierarchy of antagonism by adding kinetic terms τ^{-1} to the equations. We illustrate this on the second variant of SHP-1 reduction. The antagonism hierarchy is initially absent from the reduced model (Fig. S4 G). When we add τ^{-1} terms to Eqs. 17 and 18, it is retrieved, Fig. 4, because $m = 4$ is large enough. When m is too low ($m = 2$, Figs. S4 E - F), antagonism behavior peaks for $\tau \ll \tau_c$ and we can not recover the hierarchy observed experimentally.

References

[1] Paul François and Eric D Siggia. A case study of evolutionary computation of biochemical adaptation. *Physical Biology*, 5(2):26009, 2008.

[2] Jean-Benoît Lalanne and Paul François. Principles of adaptive sorting revealed by in silico evolution. *Physical Review Letters*, 110(21):218102, May 2013.

[3] Wenzhe Ma, Ala Trusina, Hana El-Samad, Wendell A Lim, and Chao Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4):760–773, August 2009.

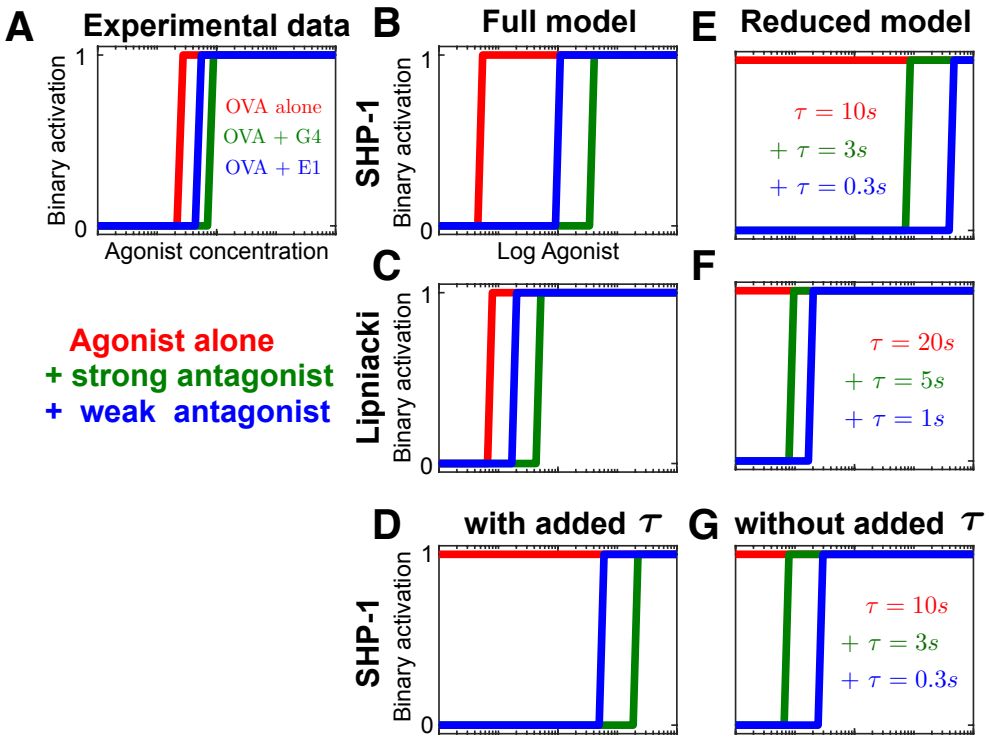

Figure S 4: Overview of antagonism. Red corresponds to agonists alone, green to agonists in the presence of a fixed number of strong antagonist ligands and blue to agonists with weak antagonists. The output is shown as an binary activation depending on threshold crossing. (A) Experimental data, reproduced from Fig. 5B of [4]. (B)-(C) Full SHP-1/Lipniacki model, showing typical antagonistic hierarchy with binding times as in (E),(F), which show reduced variants of the SHP-1/Lipniacki model via parameter symmetry breaking. (D),(G) Second variant of the reduced SHP-1 model. Upon adding back terms in τ to Eqs. 17-18, we retrieve the proper hierarchy of antagonism. We added 10^4 antagonist ligands to SHP-1 models, 10^2 antagonist ligands to Lipniacki models and 10 μmol antagonist ligand concentration in the experiments.

[4] Paul François, Guillaume Voisinne, Eric D Siggia, Grégoire Altan-Bonnet, and Massimo Vergassola. Phenotypic model for early T-cell activation displaying sensitivity, specificity, and antagonism. *Proc Natl Acad Sci U S A*, 110(10):E888–97, March 2013.

[5] Paul François, Mathieu Hemery, Kyle A Johnson, Laura N Saunders. Phenotypic spandrel: absolute discrimination and ligand antagonism. Physical Biology. 2016;13(6):066011.

Table 14: Lipniacki model initial parameters

Parameter	Value	Details
TCR	3×10^4	
LCK	10^5	
SHP	3×10^5	
ZAP	10^5	
MEK	10^5	Can't be modified by $\bar{\phi}$
ERK	3×10^5	
b_1	$3 \times 10^{-1}/TCR$	Agonist peptide binding
lb	$3 \times 10^{-1}/LCK$	LCK(s) binding
ly_1	$5/SHP$	pSHP complex binding
ly_2	3×10^{-1}	Theorine phosphorylation at complex
ls_1	10^{-1}	Spontaneous serine dephosphorylation
ls_2	5×10^{-1}	ppERK catalyzed serine phosphorylation
tp	5×10^{-2}	TCR phosphorylation
s_0	10^{-5}	Spontaneous SHP phosphorylation
s_1	$3 \times 10^2/SHP$	SHP phosphorylation
s_2	6×10^{-4}	SHP dephosphorylation
s_3	5×10^{-2}	SHP dissociation
z_0	2×10^{-6}	Spontaneous ZAP phosporylation
z_1	$5/ZAP$	ZAP phosphorylation
z_2	2×10^{-2}	ZAP dephosphorylation
m_1	$5 \times ZAP/MEK$	MEK phosphorylation
m_2	2×10^{-2}	MEK dephosphorylation
e_1	$5 \times MEK/ERK$	ERK phosphorylation
e_2	2×10^{-2}	ERK dephosphorylation

Table 15: Lipniacki basic first variant

Step	I_{init}	Parameters		Limit	Description per group
1	0.45	(m_1, m_2)	$\rightarrow$	∞	
2	0.47	b_1	$\rightarrow$	∞	
3	0.47	(lb, LCK)	$\rightarrow$	$(0, \infty)$	$lb' = lb\,LCK$
4	0.47	ls_1, z_1	$\rightarrow$	$0, \infty$	Turning off the positive feedback
5	0.50	e_1, e_2	$\rightarrow$	0	
6	0.50	z_0, z_2	$\rightarrow$	0	
7	0.50	m_1	$\rightarrow$	0	
8	0.50	s_3	$\rightarrow$	0	
9	0.50	(ls_2, LCK)	$\rightarrow$	$(0, \infty)$	
10	0.50	ly_2	$\rightarrow$	∞	
11	0.50	(TCR, SHP)	$\rightarrow$	∞	
12	0.5017	s_0	$\rightarrow$	0	
13	0.5017	(s_2, ly_1)	$\rightarrow$	$(0, \infty)$	Products
14	0.5017	(SHP, s_1)	$\rightarrow$	$(0, \infty)$	$s_1' = s_1 SHP$
15	0.5017	(LCK, ly_1)	$\rightarrow$	$(0, \infty)$	
16	0.5216	s_1	$\rightarrow$	∞	

Table 16: Lipniacki advanced first variant

Step	I_{init}	Parameters		Limit	Description per group
1	0.5837	(s_{22}, s_{23})	$\rightarrow$	0	
2	0.5837	(b_1, ly_{22})	$\rightarrow$	∞	$ly_{22} \rightarrow \infty$ makes no change
3	0.5837	(s_{21}, ly_{22})	$\rightarrow$	∞	
4	0.5837	ly_{22}	$\rightarrow$	∞	
5	0.5837	(TCR, ly_{11})	$\rightarrow$	$(0, \infty)$	$ly'_{11} = ly_{11}TCR$
6	0.5837	s_{12}	$\rightarrow$	0	
7	0.6097	s_{13}	$\rightarrow$	0	
8	0.6147	(tp_2, tp_3)	$\rightarrow$	$(0, \infty)$	
9	0.6231	(ly_{13}, lb)	$\rightarrow$	$0, \infty$	
10	0.6245	(tp_4, ls_2)	$\rightarrow$	$(0, \infty)$	Products
11	0.6246	(tp_3, ly_{16})	$\rightarrow$	$(0, \infty)$	
12	0.6354	(ly_{16}, ly_{15})	$\rightarrow$	$(0, \infty)$	
13	0.6563	(ly_{15}, ly_{13})	$\rightarrow$	$(0, \infty)$	
14	0.6699	ly_{21}	$\rightarrow$	∞	
15	0.6749	ls_2, ly_{12}	$\rightarrow$	$0, \infty$	
16	0.7405	(ly_{14}, s_{11})	$\rightarrow$	∞	

Table 17: Lipniacki basic second variant

Step	I_{init}	Parameters		Limit	Description per group
1	0.7583	(m_2, m_1)	$\rightarrow$	∞	Shutting down positive feedback
2	0.8337	(z_2, m_1)	$\rightarrow$	∞	
3	0.8337	LCK	$\rightarrow$	∞	
4	0.8777	lb	$\rightarrow$	∞	
5	0.8777	(ls_1, ly_2)	$\rightarrow$	$(0, \infty)$	
6	0.8777	(z_0, s_0)	$\rightarrow$	0	
7	0.8777	(m_1, s_1)	$\rightarrow$	∞	
8	0.8777	(ly_2, z_1)	$\rightarrow$	$(0, \infty)$	
9	0.8915	e_2	$\rightarrow$	0	
10	0.8915	z_1	$\rightarrow$	0	
11	0.8915	e_1	$\rightarrow$	∞	
12	0.8915	(s_1, SHP)	$\rightarrow$	$(0, \infty)$	Rescaling
13	0.8954	(b_1, SHP)	$\rightarrow$	∞	
14	0.9029	(s_3, s_2)	$\rightarrow$	$(0, \infty)$	
15	0.9278	(ls_2, tp)	$\rightarrow$	$(0, \infty)$	
16	0.9351	(tp, TCR)	$\rightarrow$	$(0, \infty)$	
17	0.9725	(s_2, ly_1)	$\rightarrow$	$(0, \infty)$	
18	1	(SHP, ly_1)	$\rightarrow$	$(0, \infty)$	

Table 18: Lipniacki advanced second variant

Step	I_{init}	Parameters		Limit	Description per group
1	1	(m_{21}, ly_{22})	$\rightarrow$	0	Cleaning unnecessary parameters
2	1	(m_1, s_{24})	$\rightarrow$	0	
3	1	(ly_{11}, ls_{13})	$\rightarrow$	0	
4	1	(s_{12}, s_{31})	$\rightarrow$	0	
5	1	(ly_{13}, ls_{11})	$\rightarrow$	0	
6	1	(e_1, ls_{14})	$\rightarrow$	0	
7	1	(tp_2, s_{33})	$\rightarrow$	0	
8	1	(m_{22}, ls_{12})	$\rightarrow$	0	
9	1	(z_2, s_{22})	$\rightarrow$	0	
10	1	(s_{32}, s_{13})	$\rightarrow$	0	
11	1	ly_{16}	$\rightarrow$	0	
12	1	(s_{23}, ly_{14})	$\rightarrow$	0	
13	1	ls_2	$\rightarrow$	∞	
14	1	(s_{11}, tp_4)	$\rightarrow$	∞	Strengthening remaining reactions
15	1	(ly_{21}, ly_{15})	$\rightarrow$	∞	
16	1	(b_1, s_{21})	$\rightarrow$	∞	
17	1	tp_3	$\rightarrow$	0	Turning off one output
18	1	(lb, ly_{15})	$\rightarrow$	∞	Strengthening remaining reactions
19	1	(SHP, s_{21})	$\rightarrow$	∞	
20	1	(LCK, ly_{12})	$\rightarrow$	∞	
21	1	(ly_{12}, tp_4)	$\rightarrow$	∞	
22	1	(ly_{15}, tp_4)	$\rightarrow$	∞	
23	1	tp_4	$\rightarrow$	∞	
24	1	(TCR, tp_1)	$\rightarrow$	$(0, \infty)$	
			FINAL OUTPUT		$X_{10} = \frac{tp_1 s_{21} TCR}{s_{11} ly_{17}} \tau$