PNAS
www.pnas.org

Supplementary Information for

How the T cell signaling network processes information to discriminate between self and agonist ligands.

Authors: Raman S. Ganti[1], Wan-Lin Lo[2], Darren B. McAffee[3], Jay T. Groves[3,4,5], Arthur Weiss[2,6,*], & Arup K. Chakraborty[1,7,8,9,10,*]

*Correspondence to: arthur.weiss@ucsf.edu or arupc@mit.edu

This PDF file includes:

Supplementary text

Figures S1 to S16

Tables S1 to S7

SI References

Supplementary Information Text

Derivation of Channel Capacity and Assumptions:

We can determine the conditional probability of the output given that inputs are normal or infected cells. For normal cells, we have $P(O \mid L_f = 0, L_s, \tau_s)$, where O is the downstream biochemical output, L_f is a two state variable that indicates whether the target cell is normal or infected, L_s is the concentration of self ligands, and τ_s is the

www.pnas.org/cgi/doi/10.1073/pnas.2008303117

lifetime of bonds between receptor and self ligands. Let us assume that the probability distribution of τ_s is given by a delta function peaked around a single value ($\tau_s = 0.5s$):

$$P(\tau_s) = \delta(\tau_s - 0.5). \tag{1}$$

Since τ_s, L_f, and L_s are independent, we can marginalize out τ_s:

$$P(O \mid L_f = 0, L_s) = \int P(O \mid L_f = 0, L_s, \tau_s) \times \delta(\tau_s - 0.5)\, \mathrm{d}\tau_s. \tag{2}$$

Of course, sampling τ_s from a distribution would be more realistic, but fixing τ_s simplifies the numerical procedure discussed in the main text. Furthermore, the value of τ_s only increases or decreases the number of steps needed for the output distributions to separate.

Assuming that $P(L_s)$ is given by a log-normal distribution (1), we can marginalize out L_s, which is also assumed to be independent of L_f.

$$P(O \mid L_f = 0) = \int P(O \mid L_f = 0, L_s) P(L_s)\, \mathrm{d}L_s \tag{3}$$

For infected cells, the conditional probability distribution is $P(O \mid L_f > 0, L_s, \tau_s, L_f, \tau_f)$. Since $\tau_f = 10\tau_s$, we ignore the dependency on τ_f. Doing so also assumes that τ_f is single valued. As before, we marginalize out τ_s and L_s giving $P(O \mid L_f > 0, L_f)$. Because it is difficult to determine $P(L_f \mid L_f > 0)$, we make the simplifying assumption that L_f also takes on a single value $L_{f,0} << L_s$ in infected cells. Using $P(O \mid L_f = 0)$,

$P(O \mid L_f > 0)$ and $p(L_f = 0) = p(L_f > 0) = 1/2$, we arrive at the expression for capacity shown in the main text.

The procedure to calculate the channel capacity is straightforward. We fixed ligand binding lifetimes: for self ligands, the half-life $\tau_s = 0.5$ s, and for agonist ligands, the half-life $\tau_f = 5.0$ s. Next, we sampled the values of the self ligand densities, L_s, from a log-normal distribution, as suggested by experiments (1). Spatial heterogeneities in the distribution of signaling molecules are not considered in our analyses. For each value of L_s, the concentration of downstream product was determined by numerically solving the system of ordinary differential equations describing the network (**Fig. 1B**). The concentration of downstream biochemical output was then binned into a histogram to give $P(O \mid L_f)$. The numerical procedure was carried out for two cases: $L_f = 0$ where equations corresponding to the biochemical reactions were solved in the absence of agonist ligands and $L_f = L_{f,0}$ where a small amount of agonists is present in a sea of much higher numbers of self ligands. The ODEs describing the reaction networks were solved numerically via the PySB software package (2). The "Real-valued Variable-coefficient Ordinary Differential Equation solver" (VODE) with fixed-leading-coefficient implementation was used to integrate the equations (3). We used a stiff solver based on backward differentiation formulas (BDF) with an internally generated full Jacobian. The number of internally defined steps allowed during one call to the solver is set to the maximum possible value.

Numerical Calculation of Channel Capacity:

The channel capacity as defined in the main text is given by:

$$\begin{aligned} C &= \max_{p(L_f)} I(L_f;O) \\ &= \frac{1}{2}\int P(O\,|\,L_f=0)\log\left(\frac{P(O\,|\,L_f=0)}{P(O)}\right) + P(O\,|\,L_f=L_{f,0})\log\left(\frac{P(O\,|\,L_f=L_{f,0})}{P(O)}\right)\mathrm{d}O \end{aligned} \quad (4)$$

where $P(O)=\frac{1}{2}\left(P(O\,|\,L_f=0)+P(O\,|\,L_f=L_{f,0})\right)$. Substituting $P(O)$ into Eq (4) gives

$$\begin{aligned} C &= \max_{p(L_f)} I(L_f;O) \\ &= \frac{1}{2}\int P(O\,|\,L_f=0)\log\left(\frac{P(O\,|\,L_f=0)}{\frac{1}{2}\left(P(O\,|\,L_f=0)+P(O\,|\,L_f=L_{f,0})\right)}\right) \\ &+ P(O\,|\,L_f=L_{f,0})\log\left(\frac{P(O\,|\,L_f=L_{f,0})}{\frac{1}{2}\left(P(O\,|\,L_f=0)+P(O\,|\,L_f=L_{f,0})\right)}\right)\mathrm{d}O \end{aligned} \quad (5)$$

If the conditional distributions completely overlap, $P(O\,|\,L_f=0)=P(O\,|\,L_f=L_{f,0})$ and we immediately see that $C=0$. If the distributions are completely separate, $P(O\,|\,L_f=0)=0$ when $P(O\,|\,L_f=L_{f,0})>0$ and $P(O\,|\,L_f=0)>0$ when $P(O\,|\,L_f=L_{f,0})=0$ so that Eq (5) reduces to

$$\begin{aligned} C &= \frac{1}{2}\int P(O\,|\,L_f=0)+P(O\,|\,L_f=L_{f,0})\,\mathrm{d}O \\ &= \int P(O)\,\mathrm{d}O = 1 \end{aligned} \quad (6)$$

The last condition is useful for numerical evaluation of Eq (4). Of course, capacity will be sensitive to the bin size used in the numerical integral. We use Eq (6) as a measure for the appropriate bin size. If $\int P(O)\,\mathrm{d}O > 0.99$, then $\mathrm{d}O$ is sufficiently small and Eq (4) is computed via the same bin size. We also use Eq (6) to determine when the distributions become completely separate. If $C = \int P(O)\,dO$, then $C = 1$.

Rate Constants for Simulations:

Tables 1 and 2 show the initial concentrations and rate constants used for the network drawn in **Fig. 1B** and the calculations presented in **Figs. 1-2**. **Fig. S1** provides nomenclature for all species used in **Fig. 1**. For simplicity, $k_{on} = k_p = k_{lck,on}[Lck]_0 = k_{zap,on}[Zap]_0 = k_{lat,on}[LAT]_0$, which varies in many calculations; k_p is the phosphorylation rate of signaling kinases, $k_{lck,on}$, $k_{zap,on}$, and $k_{lat,on}$ are rates at which Lck, ZAP-70, and LAT bind to the TCR-pMHC complex, and $[Lck]_0$, $[Zap]_0$, and $[LAT]_0$ are initial concentrations of the signaling proteins; $k_s = 2.0s^{-1}$ is the off-rate of self ligands, $k_f = 0.2s^{-1}$ is the off-rate of agonist ligands, $k_{off} = 0.005s^{-1}$ is the off-rate of reactions on the main path, $k_{L,on} = 0.0022(mol \cdot s)^{-1}$ is the on-rate for both ligands. Reaction rates of the side paths were fixed so that disassembly of the complex happens quickly after ligand unbinding. Total concentrations of different species are provided by (1).

Table 1: Initial Concentrations

Species	Total Concentration (number of molecules)
$[R]_0$	30000
$[Lck]_0$	30000
$[Zap]_0$	1200000
$[LAT]_0$	150000

$[L_f]_0 = 30\ mol$ and $[L_s]_0$ is sampled from a log-normal distribution:

$$P([L_s]_0) = \frac{1}{[L_s]_0 \sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln([L_s]_0) - \mu)^2}{2\sigma^2} \right] \tag{7}$$

where $\mu = 6.0$ and $\sigma = 1.0$. $\langle L_s \rangle = \exp[\mu + \sigma^2/2] \sim 665 mol$.

Table 2: Rate constants for T cell networks presented in **Figs. 1B** and **S3A**

Main Path Reactions	$k_{forward}$	$k_{reverse}$
$R + L \rightleftharpoons RL$	$k_{L,on}$	k_s or k_f
$RL + Lck \rightleftharpoons RL\_Lck$	$k_{on}/[Lck]_0$	k_{off}
$RL\_Lck \rightleftharpoons RPL\_Lck$	k_{on}	k_{off}
$RPL\_Lck + Zap \rightleftharpoons RPL\_Lck\_Zap$	$k_{on}/[Zap]_0$	k_{off}
$RPL\_Lck\_Zap \rightleftharpoons RPL\_Lck\_Zap\_P$	k_{on}	k_{off}
$RPL\_Lck\_Zap\_P + LAT \rightleftharpoons RPL\_Lck\_Zap\_P\_LAT$	$k_{on}/[LAT]_0$	k_{off}
$RPL\_Lck\_Zap\_P\_LAT \rightleftharpoons RPL\_Lck\_Zap\_P\_LATP$	k_{on}	k_{off}
$RPL\_Lck\_Zap\_P\_LATP \rightarrow RPL\_Lck\_Zap\_P + LATPP$	k_{on} or $k_{on}/10$	
$LATPP \rightarrow LAT$	k_{off}	

Side Path Reactions	$k_{forward}$	$k_{reverse}$

$RL\_Lck \rightleftharpoons R\_Lck + L$	k_s or k_f	$k_{L,on}$
$R\_Lck \rightleftharpoons R + Lck$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RPL\_Lck \rightleftharpoons RP\_Lck + L$	k_s or k_f	$k_{L,on}$
$RP\_Lck \rightleftharpoons RP + Lck$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RP \rightleftharpoons R$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$
$RPL\_Lck\_Zap \rightleftharpoons RP\_Lck\_Zap + L$	k_s or k_f	$k_{L,on}$
$RP\_Lck\_Zap \rightleftharpoons RP\_Zap + Lck$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RP\_Zap \rightarrow RP + Zap$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RPL\_Lck\_Zap\_P \rightleftharpoons RP\_Lck\_Zap\_P + L$	k_s or k_f	$k_{L,on}$
$RP\_Lck\_Zap\_P \rightleftharpoons RP\_Zap\_P + Lck$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RP\_Zap\_P \rightleftharpoons RP\_Zap$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$
$RPL\_Lck\_Zap\_P\_LAT \rightleftharpoons RP\_Lck\_Zap\_P\_LAT + L$	k_s or k_f	$k_{L,on}$
$RP\_Lck\_Zap\_P\_LAT \rightleftharpoons RP\_Zap\_P\_LAT + Lck$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RP\_Zap\_P\_LAT \rightleftharpoons RP\_Zap\_LAT$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$
$RP\_Zap\_LAT \rightleftharpoons RP\_Zap + LAT$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RPL\_Lck\_Zap\_P\_LATP \rightleftharpoons RP\_Lck\_Zap\_P\_LATP + L$	k_s or k_f	$k_{L,on}$
$RP\_Lck\_Zap\_P\_LATP \rightleftharpoons RP\_Zap\_P\_LATP + Lck$	$20\ s^{-1}$	10^{-5} $(mol \cdot s)^{-1}$
$RP\_Zap\_P\_LATP \rightleftharpoons RP\_Zap\_LATP$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$

$RP\_Zap\_LATP \rightleftharpoons RP\_Zap\_LAT$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$

Legend:

Cycle 1

RL: Receptor-Ligand
RL_Lck: Receptor/Co-receptor-Ligand
R_Lck: Receptor/Co-receptor

Cycle 2

RPL_Lck: Receptor/Co-receptor-Ligand with phosphorylated ITAMs
RP_Lck: Receptor/Co-receptor with phosphorylated ITAMs
RP: Receptor with phosphorylated ITAMs

Cycle 3

RPL_Lck_Zap: Receptor/Co-receptor-Ligand with Zap70 bound to ITAMs
RP_Lck_Zap: Receptor/Co-receptor with Zap70 bound to ITAMs
RP_Zap: Receptor with Zap70 bound to ITAMs

Cycle 4

RPL_Lck_Zap_P: Receptor/Co-receptor-Ligand with phosphorylated Zap70 bound to ITAMs
RP_Lck_Zap_P: Receptor/Co-receptor with phosphorylated Zap70 bound to ITAMs
RP_Zap_P: Receptor with phosphorylated Zap70 bound to ITAMs

Cycle 5

RPL_Lck_Zap_P_LAT: Receptor/Co-receptor-Ligand-Zap70P bound to LAT
RP_Lck_Zap_P_LAT: Receptor/Co-receptor-Zap70P bound to LAT
RP_Zap_P_LAT: Receptor-Zap70P bound to LAT
RP_Zap_LAT: Receptor-Zap70 bound to LAT

Cycle 6

RPL_Lck_Zap_P_LATP: Receptor/Co-receptor-Ligand-Zap70P bound to phosphorylated LAT
RP_Lck_Zap_P_LATP: Receptor/Co-receptor-Zap70P bound to phosphorylated LAT
RP_Zap_P_LATP: Receptor-Zap70P bound to phosphorylated LAT
RP_Zap_LATP: Receptor-Zap70 bound to phosphorylated LAT

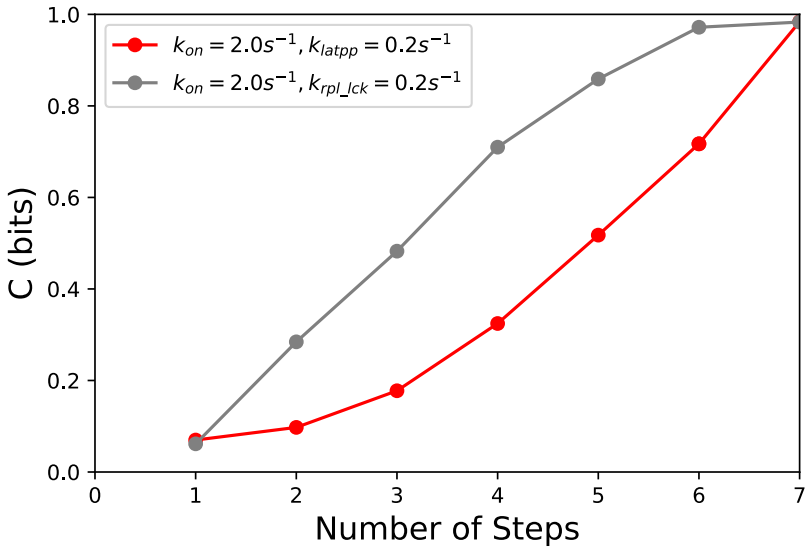
Step 7 (a)

LATPP: Fully phosphorylated LAT complex

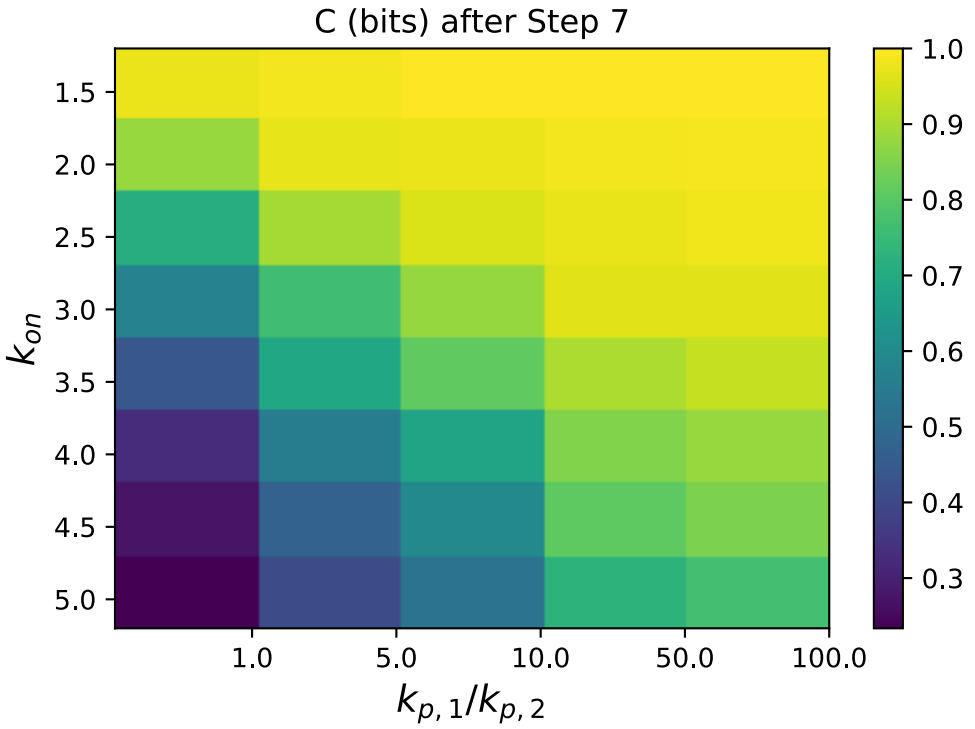
Cycle 7 (b)

RPL_Lck_Zap_P_LATPP: Receptor/Co-receptor-Ligand-Zap70P bound to fully phosphorylated LAT complex
RP_Lck_Zap_P_LATPP: Receptor/Co-receptor-Zap70P bound to fully phosphorylated LAT complex
RP_Zap_P_LATPP: Receptor-Zap70P bound to fully phosphorylated LAT complex
RP_Zap_LATPP: Receptor-Zap70 bound to fully phosphorylated LAT complex

Figure S1: The legend defines the nomenclature of all species used in **Fig. 1B** and **Fig. S3A**. Step 7(a) corresponds to the network topology shown in **Fig. S3A** and cycle 7(b) corresponds to the topology in **Fig. 1B**.

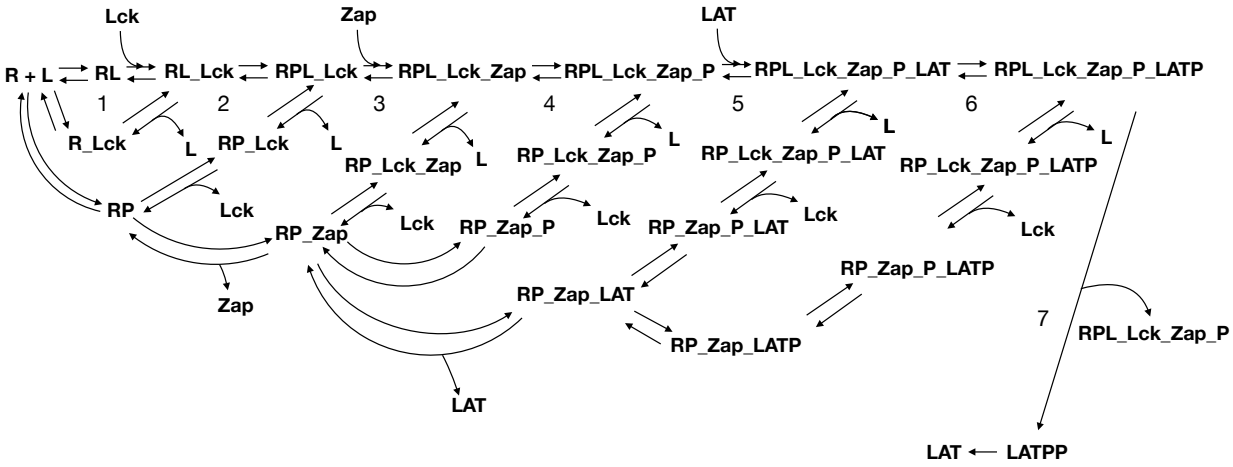

$k_{on} = 2.0s^{-1}, k_{latpp} = 0.2s^{-1}$
$k_{on} = 2.0s^{-1}, k_{rpl\_lck} = 0.2s^{-1}$
C (bits)
Number of Steps
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
7

A

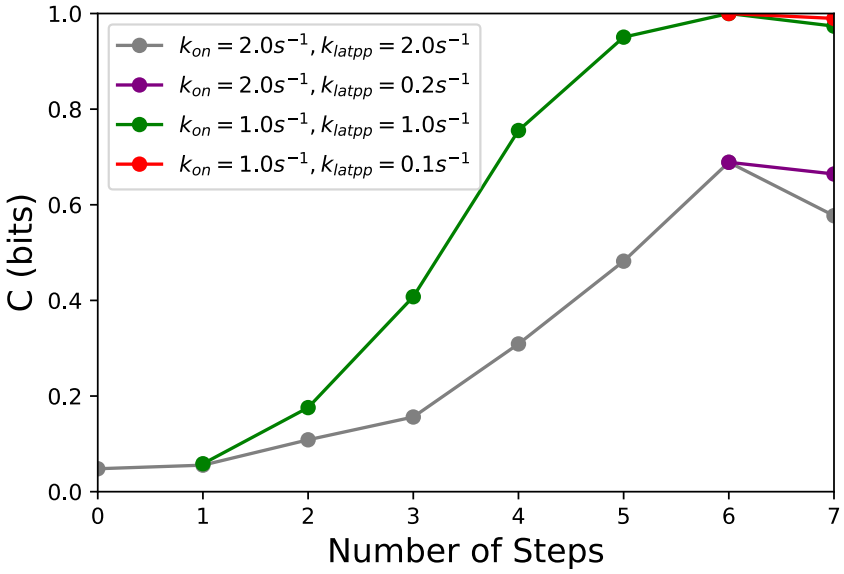


B

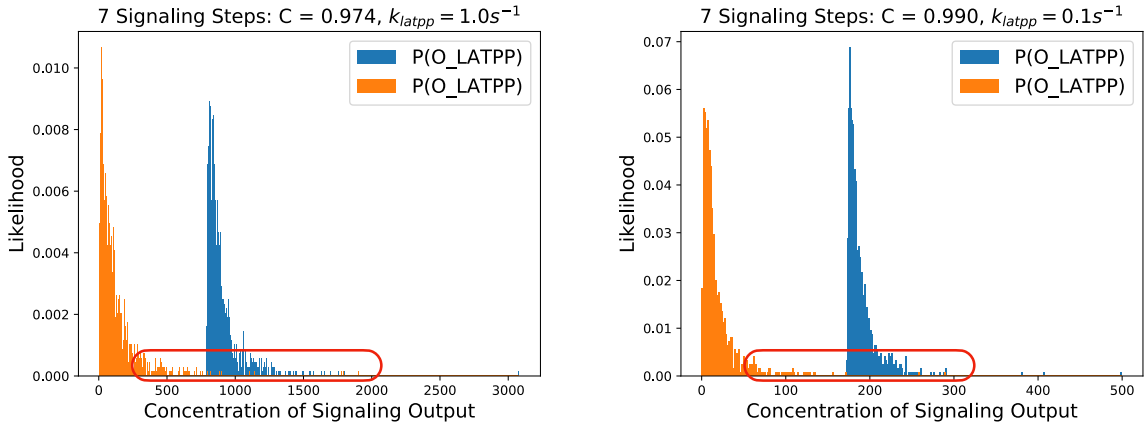
Figure S2: (**A**) Symmetry in discriminatory improvement: a slow phosphorylation step early in signaling (grey line) eventually leads to the same capacity as a slow step that happens later (red line). (**B**) Parameter search for the network described in **Fig. 1B** of the main text. For all cases, slowing the final phosphorylation step improves discrimination, suggesting that the slow step is the final opportunity for capacity to increase so that discrimination becomes perfect before downstream signaling ensues.


Lck
Zap
LAT
R + L
RL
RL_Lck
RPL_Lck
RPL_Lck_Zap
RPL_Lck_Zap_P
RPL_Lck_Zap_P_LAT
RPL_Lck_Zap_P_LATP
1
2
3
4
5
6
7
R_Lck
RP_Lck
RP_Lck_Zap
RP_Lck_Zap_P
RP_Lck_Zap_P_LAT
RP_Lck_Zap_P_LATP
L
RP
RP_Zap
RP_Zap_P
RP_Zap_P_LAT
RP_Zap_P_LATP
RP_Zap_LAT
RP_Zap_LATP
Lck
Zap
LAT
RPL_Lck_Zap_P
LAT ⟵ LATPP

A


$k_{on} = 2.0s^{-1}, k_{latpp} = 2.0s^{-1}$
$k_{on} = 2.0s^{-1}, k_{latpp} = 0.2s^{-1}$
$k_{on} = 1.0s^{-1}, k_{latpp} = 1.0s^{-1}$
$k_{on} = 1.0s^{-1}, k_{latpp} = 0.1s^{-1}$
C (bits)
Number of Steps
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
7

B



C

Figure S3: (**A**) Realistic biochemical network describing the early events of TCR signaling (see **Fig. S1** for species nomenclature) that considers LATPP dissociating from the receptor-ligand complex after Y132 phosphorylation. (**B**) Discriminatory capacity vs. number of signaling steps. (**C**) Histograms of signaling output generated after Y132 phosphorylation for the network: a drop in capacity (left panel) occurs when the Y132 step proceeds at the same rate ($k_{latpp} = 1.0s^{-1}$) as the other LAT Y phosphorylation steps; the effect is due to increased output from APCs that present high densities of only self ligands, causing the two distributions to overlap (see red insets). The drop is averted (right panel) when the final step is slowed down ($k_{latpp} = 0.1s^{-1}$) with respect to the previous step.

Table 3 shows the rate constants used to explore the consequences of slowing the final LAT phosphorylation step in the network shown in **Fig. 1B** of the main text, where LATPP remains bound to the TCR complex. If C < 1 before Tyr 132 phosphorylation happens, **Fig. S2A** shows that slowing the step down ($k_{latpp} = k_{latp}/10$, red line) increases channel capacity and enhances APC discrimination. Interestingly, slowing

down an earlier step (**Fig. S2A**, grey line) results in the same increase of capacity. **Fig. S2B** shows for a range of possible k_{on}, as $k_{p,1} / k_{p,2}$ increases, i.e. the second phosphorylation step is made slow with respect to the first, discrimination improves. Therefore, if LAT remains bound to the TCR-pMHC complex, the slow step proofreads and is likely necessary to increase capacity.

Fig. S3A shows the topology of the signaling network when fully phosphorylated LATPP dissociates from the TCR complex in a Markovian step. As a consequence of the data processing inequality, the capacity can either remain constant or decrease during this step. If the rate of phosphorylation is too fast, output from self only APCs is amplified, leading to a drop in capacity (**Fig. S3B**). The effect is more pronounced as k_{on} is increased in the model (**Fig. S3B**, gray and purple). Upon observing the histograms in **Fig. S3C**, we clearly observe the source of the capacity drop. APCs that present high concentrations of self ligands produce enough signaling output by step 6 to catalyze the formation of LATPP. The resulting product increases the overlap in the conditional output distributions (**Fig. S3C**, left panel, red box). Slowing the Y132 phosphorylation step mitigates the preceding effect and prevents a drop in capacity (**Fig. S3C**, right panel, red box).

Table 3: Rate constants for T cell network presented in **Fig. 1B**

Main Path Reactions	$k_{forward}$	$k_{reverse}$
All main reactions in Table 2		
$RPL\_Lck\_Zap\_P\_LATP \rightleftharpoons RPL\_Lck\_Zap\_P\_LATPP$	Varies	k_{off}

Side Path Reactions	$k_{forward}$	$k_{reverse}$
All side reactions in Table 2		
$RPL\_Lck\_Zap\_P\_LATPP \rightleftharpoons RP\_Lck\_Zap\_P\_LATPP + L$	k_s or k_f	$k_{L,on}$
$RP\_Lck\_Zap\_P\_LATPP \rightleftharpoons RP\_Zap\_P\_LATPP + Lck$	$20\ s^{-1}$	$10^{-5}\ (mol \cdot s)^{-1}$
$RP\_Zap\_P\_LATPP \rightleftharpoons RP\_Zap\_LATPP$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$
$RP\_Zap\_LATPP \rightleftharpoons RP\_Zap\_LATP$	$20\ s^{-1}$	$10^{-5}\ s^{-1}$

Toy Model:

We attempt to derive $N_{L,min}$ for the toy model shown in **Fig. S4**. For a fixed M total steps, there are N_L steps that ligand remains bound eventually leading to formation of non ligand-specific product RP_{N_L}. RP_{N_L} then acts as an enzyme and phosphorylates S leading to *SP*. The error rate can be defined as follows:

$$\eta = \frac{[SP \mid L_f = 0]}{[SP \mid L_f > 0]} \tag{8}$$

In the following derivations, we assume $[R] \approx R^T$ i.e. the receptor concentration is in excess.

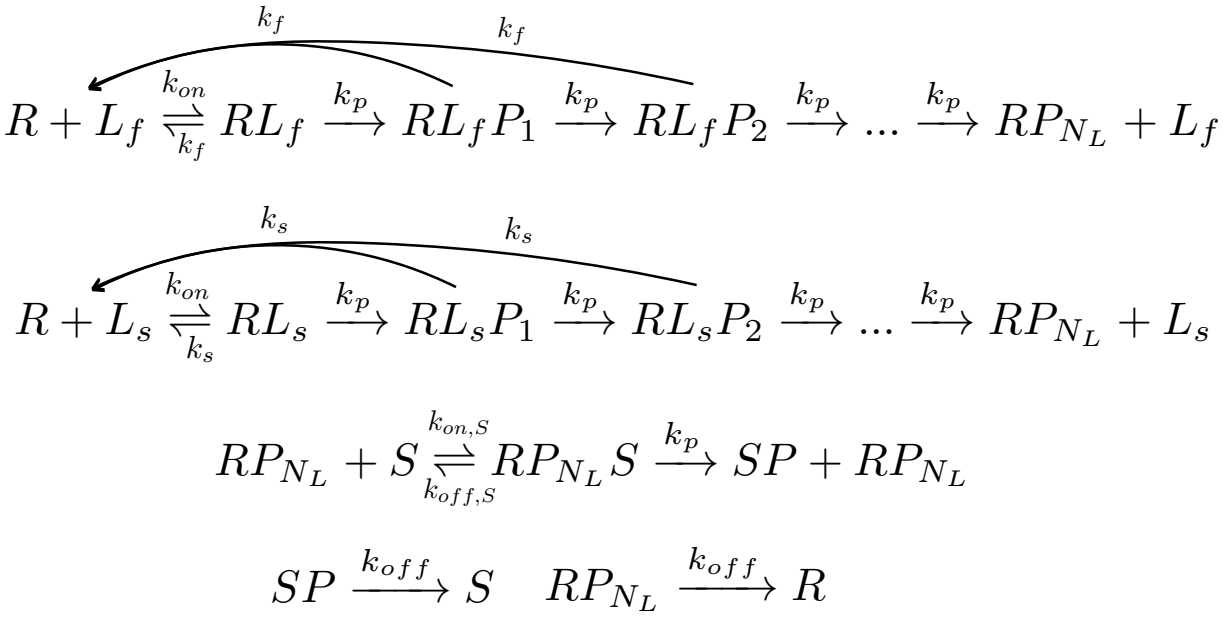


Figure S4: Toy model used to determine $N_{L,min}$. For a fixed total M steps leading to formation of SP, there are N_L steps where the ligand remains bound. N_L ranges from 1 to $M-3$. The network shows the case where $N_L = M-3$.

Solving the equations describing the toy model in **Fig. S4** at steady-state gives:

$$[SP] = \frac{k_p}{k_{off}} \left(\frac{k_{on,s}}{k_p + k_{off,s}} \right) [RP_{N_L}][S] \tag{9}$$

where

$$[S] = \frac{S_T}{\left(1 + \frac{k_{on,s}}{k_p + k_{off,s}} [RP_{N_L}] \left(1 + \frac{k_p}{k_{off}} \right) \right)} \tag{10}$$

We define:

$$\alpha = \frac{k_p}{k_p + k_s}$$
$$\beta = \frac{k_p}{k_p + k_f} \tag{11}$$
$$\gamma = \frac{k_p}{k_p + k_{off}}$$

We can further express:

$$[RP_{N_L}] = \frac{k_p}{k_{off}}[RP_{N_L-1}] = \frac{k_p}{k_{off}}\frac{k_{on}}{k_p}\gamma^{(M-3-N_L)}\alpha^{N_L}[R][L_s] \tag{12}$$

Substituting, we are left with:

$$[SP \mid L_f = 0] = \frac{\frac{k_p}{k_{off}}\left(\frac{k_{on,S}}{k_p + k_{off,S}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{(M-3-N_L)}\alpha^{N_L}[R][L_s]S_T}{\left(1+\left(\frac{k_{on,S}}{k_p + k_{off,S}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{(M-3-N_L)}\alpha^{N_L}[R][L_s]\left(1+\frac{k_p}{k_{off}}\right)\right)} \tag{13}$$

We can also solve for $[L_s]$ from $[L_s^T]$

$$[L_s] = \frac{L_s^T}{\left(1+[R]\left(\frac{k_{on}}{k_p}\right)\alpha\left(\frac{1-\alpha^{N_L}}{1-\alpha}\right)\right)} \tag{14}$$

We follow the same approach when a small concentration of agonist ligands is present:

$$[SP \mid L_f > 0] = \frac{\frac{k_p}{k_{off}}\left(\frac{k_{on,S}}{k_p + k_{off,S}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{(M-3-N_L)}(\alpha^{N_L}[R][L_s] + \beta^{N_L}[R][L_f])S_T}{\left(1 + \left(\frac{k_{on,S}}{k_p + k_{off,S}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{(M-3-N_L)}(\alpha^{N_L}[R][L_s] + \beta^{N_L}[R][L_f])\left(1 + \frac{k_p}{k_{off}}\right)\right)}$$

(15)

and

$$[L_f] = \frac{L_f^T}{\left(1 + [R]\left(\frac{k_{on}}{k_p}\right)\beta\left(\frac{1-\beta^{N_L}}{1-\beta}\right)\right)} \tag{16}$$

Then, we can solve for the error rate:

$$\eta = f \times \left(\frac{1}{1 + \left(\frac{\beta}{\alpha}\right)^{N_L}\left(\frac{[L_f]}{[L_s]}\right)}\right) \tag{17}$$

where

$$f = \frac{\left(1 + \left(\frac{k_{on,S}}{k_p + k_{off,S}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{(M-3-N_L)}(\alpha^{N_L}[R][L_s] + \beta^{N_L}[R][L_f])\left(1 + \frac{k_p}{k_{off}}\right)\right)}{\left(1 + \left(\frac{k_{on,S}}{k_p + k_{off,S}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{(M-3-N_L)}\alpha^{N_L}[R][L_s]\left(1 + \frac{k_p}{k_{off}}\right)\right)} \tag{18}$$

Since $[L_f] << [L_s]$, $f \sim 1$ and η reduces to

$$\eta \approx \left(\frac{1}{1+\left(\frac{\beta}{\alpha}\right)^{N_L}\left(\frac{[L_f]}{[L_s]}\right)} \right) \tag{19}$$

As $L_f << L_s$, the error rate only decreases if $\beta/\alpha \sim \tau_f/\tau_s > 1$. From Eq (19), it is clear that $N_{L,min}$ derived from η is not sensitive to M or S_T. **Fig. S5** shows that η will start at 1 and decrease as a function of N_L. Table 4 shows parameters used in the numerical calculations of $N_{L,min}$ for the toy model presented in **Fig. S4**.

Instead of using η as our discriminatory measure, we can derive an analytical expression for capacity via the toy model presented in **Fig. S4**. If $P(O \mid L_f)$ can be derived, evaluation of Eq (4) is possible.

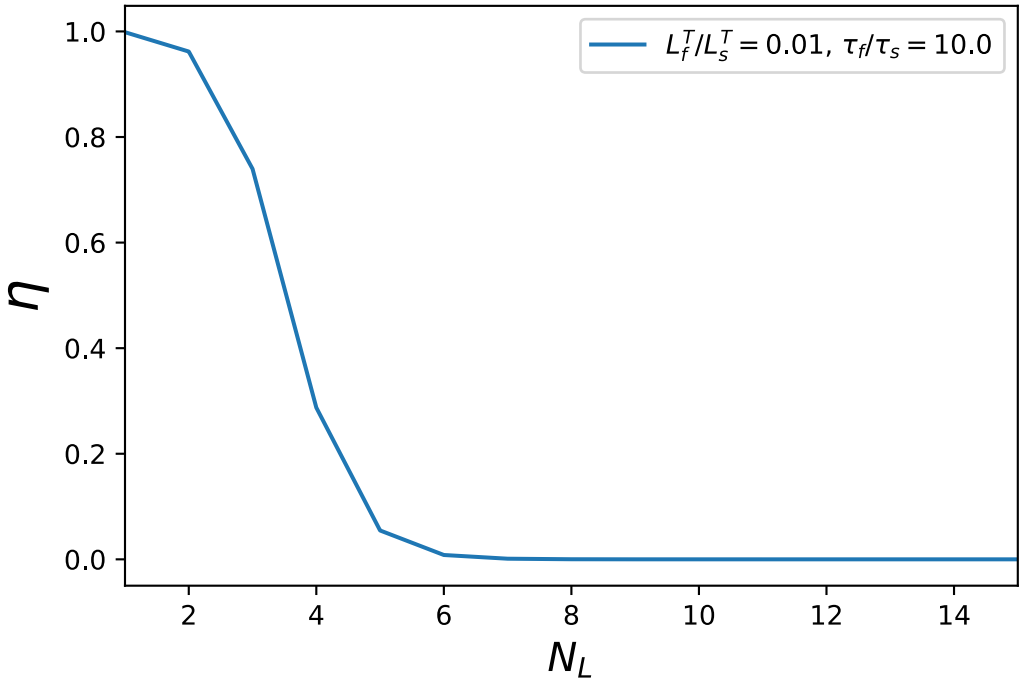


Figure S5: Numerical calculation of η vs. N_L.

Table 4: Rate constants and initial concentrations for toy model presented in **Fig. S4**

Species	Total Concentration (number of molecules)
$[R]_T$	30000
$[L_f]_T$	10
$[L_s]_T$	Varies
$[S]_T$	10000

Rate Constant	Value
k_{on}	$k_{L,on}$
k_f	$0.2\ s^{-1}$
k_s	Varies
k_p	$0.1\ s^{-1}$
k_{off}	$0.05\ s^{-1}$

$k_{on,S}$	$0.1\ (mol \cdot s)^{-1}$
$k_{off,S}$	$0.05\ s^{-1}$
M	15

$P(SP(L_s^T))$ can be related to $P(L_s^T(SP))$ using the following transformation:

$$P(SP(L_s^T)) = P(L_s^T(SP))\left|\frac{dL_s^T}{dSP}\right| \tag{20}$$

Substituting the variables

$$a = \left(\frac{k_{on,s}}{k_p + k_{off,s}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{M-3-N_L}\alpha^{N_L}[R] \tag{21}$$

and $h = \frac{k_p}{k_{off}}$ into Eq (13) gives:

$$SP = \frac{ha[L_s]S_T}{(1 + a[L_s](1+h))} \tag{22}$$

Rearranging gives

$$L_s = \frac{SP}{(haS_T - aSP(1+h))} \tag{23}$$

If we define

$$\omega = 1 + [R]\left(\frac{k_{on}}{k_p}\right)\alpha\left(\frac{1-\alpha^{N_L}}{1-\alpha}\right) \tag{24}$$

we arrive at the following expression for L_s^T:

$$L_s^T(SP) = \omega\left(\frac{SP}{(haS_T - aSP(1+h))}\right) \tag{25}$$

Differentiating with respect to *SP* gives:

$$\frac{dL_s^T}{dSP} = \omega\left(\frac{(haS_T - aSP(1+h)) + aSP(1+h)}{(haS_T - aSP(1+h))^2}\right) \tag{26}$$

Using Eq (25) and (26) with Eq (20), we can evaluate

$$P(SP \mid L_f = 0) = \frac{1}{L_s^T(SP)\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(L_s^T(SP)) - \mu)^2}{2\sigma^2}\right] \left|\frac{dL_s^T}{dSP}\right| \tag{27}$$

To determine $P(SP \mid L_f = L_{f,0})$, we follow the same procedure except instead use Eq (15). Defining

$$b = \left(\frac{k_{on,s}}{k_p + k_{off,s}}\right)\left(\frac{k_{on}}{k_{off}}\right)\gamma^{M-3-N_L}[R] \tag{28}$$

we derive

$$L_s^T(SP; L_f) = \omega\left(\frac{SP + bSP\beta^{N_L}[L_f](1+h) - hb\beta^{N_L}[L_f]S_T}{(hb\alpha^{N_L}S_T - bSP\alpha^{N_L}(1+h))}\right) \tag{29}$$

Differentiating gives:

$$\frac{dL_s^T(SP; L_f)}{dSP} = \omega \times$$

$$\left(\frac{(1+b\beta^{N_L}[L_f](1+h))(hb\alpha^{N_L}S_T - bSP\alpha^{N_L}(1+h)) + (SP + bSP\beta^{N_L}[L_f](1+h) - hb\beta^{N_L}[L_f]S_T)(b\alpha^{N_L}(1+h))}{(hb\alpha^{N_L}S_T - bSP\alpha^{N_L}(1+h))^2}\right)$$

(30)

Once again, using Eqs (29), (30) and (20) gives

$$P(SP \mid L_f = L_{f,0}) = \frac{1}{L_s^T(SP;L_f)\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(L_s^T(SP;L_f)) - \mu)^2}{2\sigma^2} \right] \left| \frac{dL_s^T(SP;L_f)}{dSP} \right| \quad (31)$$

Substituting Eqs (27) and (31) into Eq (5) gives an analytical expression for capacity. Unlike Eq (19), it is difficult to see the dependence of capacity on N_L. With the parameters listed in Table 4, we numerically evaluate the capacity using the preceding expressions. **Fig. S6** shows that the capacity's dependence on N_L is opposite to that of the error rate (see **Fig. S5**). $C \sim 0$ for $N_L = 1$ and eventually increases to 1 after six steps. Since C measures discrimination over a distribution of L_s^T, **Fig. S6** shows the results for $\ln(L_s^T) \sim \mathcal{N}(\mu = \ln(100[L_f^T]), \sigma = 1.0)$. It is worth noting that the magnitude of S_T and M affects the mean values of the output distributions, but does not affect the relative distance between them. As a consequence, capacity is insensitive to S_T and M.

Repeating the numerical procedure for a range of τ_f / τ_s and different values of μ, we compute $N_{L,min}$, the minimum number of steps that the ligand must remain bound to the TCR complex to reach perfect discrimination. After $N_{L,min}$ steps, ligand can dissociate from the complex and downstream signaling can proceed. As expected, **Fig. S7** shows that $N_{L,min}(C = 1)$ is a decreasing function of τ_f / τ_s and an increasing function of μ. The advantage of deriving $N_{L,min}$ from capacity is that $C = 1$ clearly defines the point at which perfect discrimination is reached.

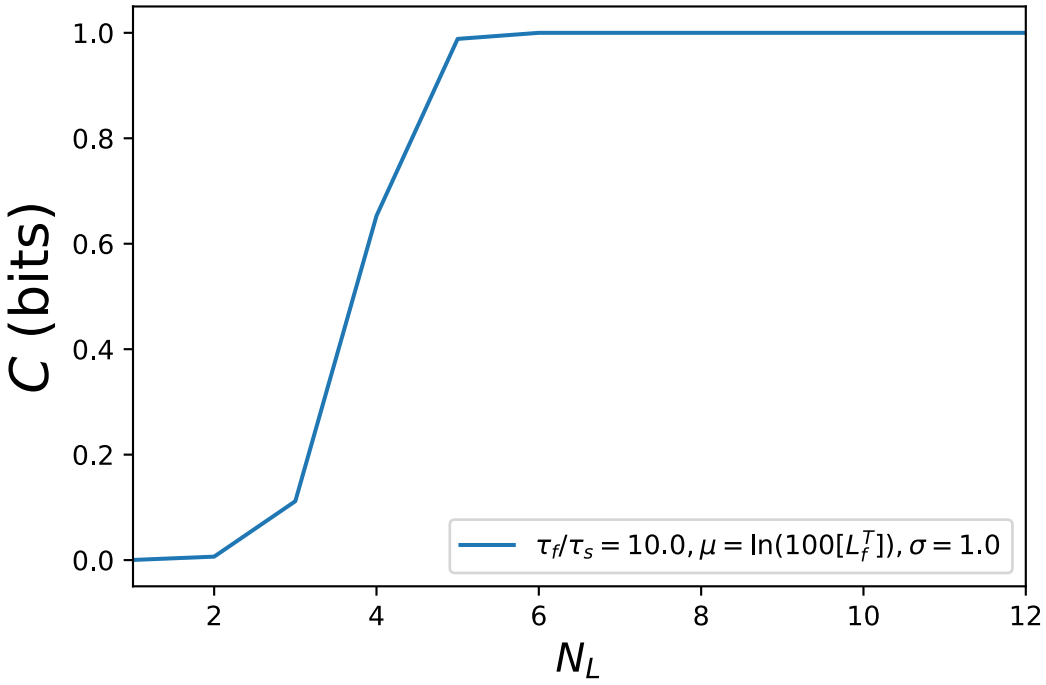


Figure S6: Numerical calculation of C vs. N_L.

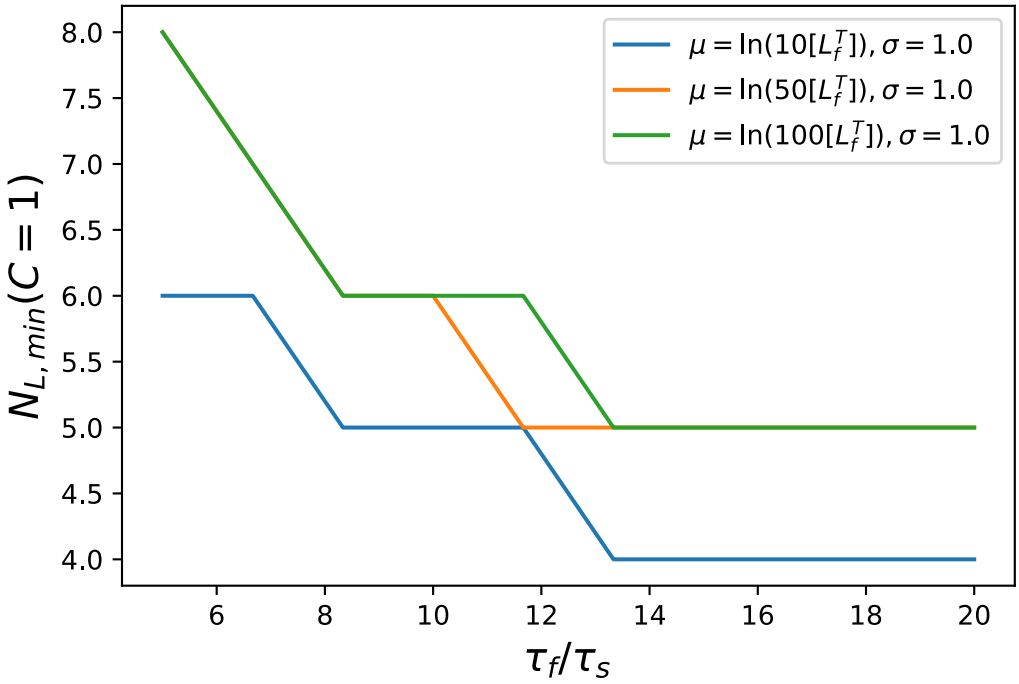


Figure S7: Numerical calculations of $N_{L,min}(C = 1)$. For a fixed set of rate constants (see Table 4), $N_{L,min}$ is calculated by finding the smallest value of N_L that gives $C = 1$. The calculation is repeated for a range of τ_f / τ_s and three values of μ.

Negative Feedback Loops:

We investigated the effects of negatively regulating free active Lck in the system by adding the reactions shown in **Fig. S8A** to those depicted in **Fig. S3A**. **Fig. S9A** shows that for different concentrations of L_f, discrimination improves over a range of feedback strengths.

$$L_f > 0:$$

$$RPL_s\_Lck\_Zap\_P + Lck \xrightarrow{k_{neg,fb}} RPL_s\_Lck\_Zap\_P + Lcki$$

$$RPL_f\_Lck\_Zap\_P + Lck \xrightarrow{k_{neg,fb}} RPL_f\_Lck\_Zap\_P + Lcki$$

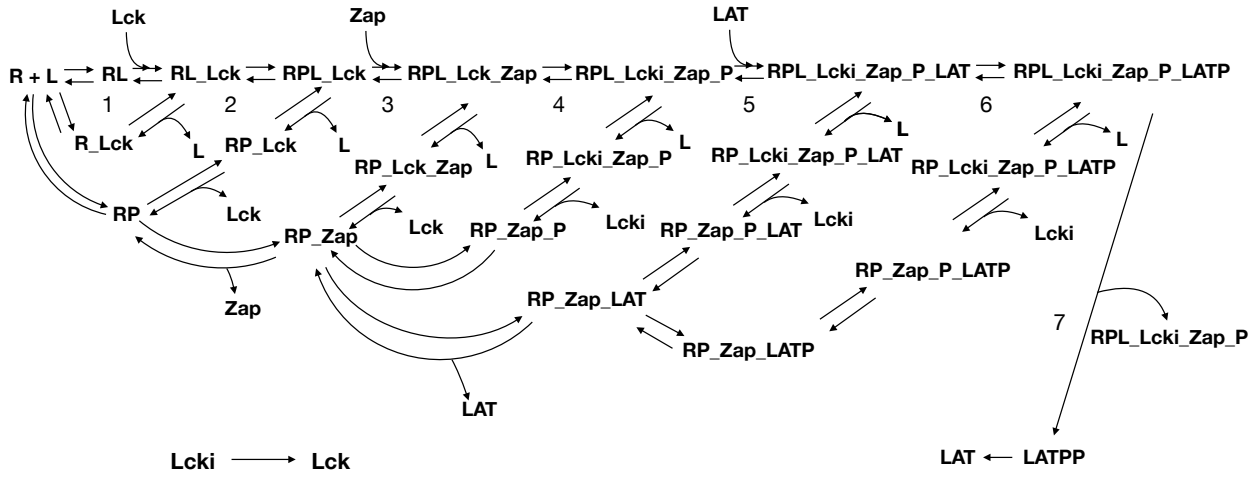
$$Lcki \xrightarrow{k_{lcki}} Lck$$

$$L_f = 0:$$

$$RPL_s\_Lck\_Zap\_P + Lck \xrightarrow{k_{neg,fb}} RPL_s\_Lck\_Zap\_P + Lcki$$

$$Lcki \xrightarrow{k_{lcki}} Lck$$

A

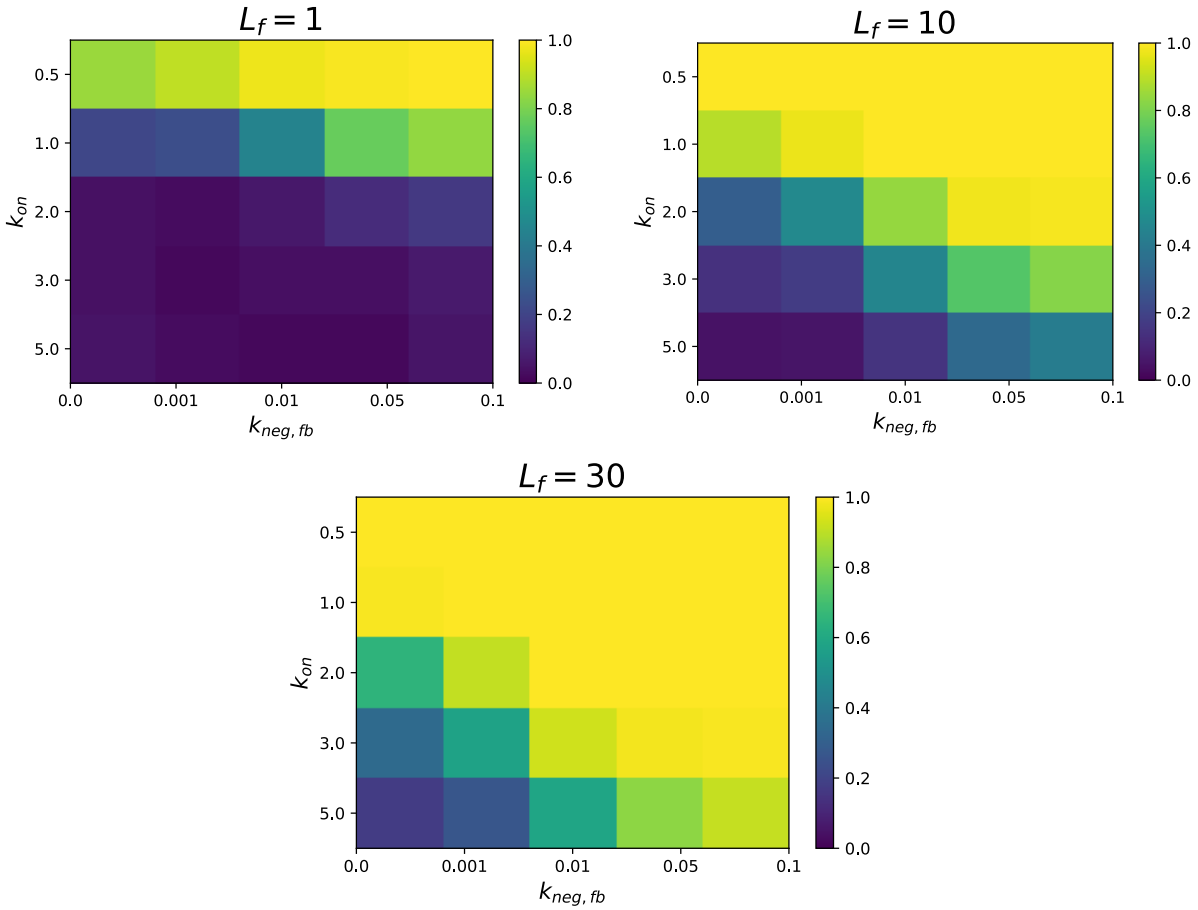


B

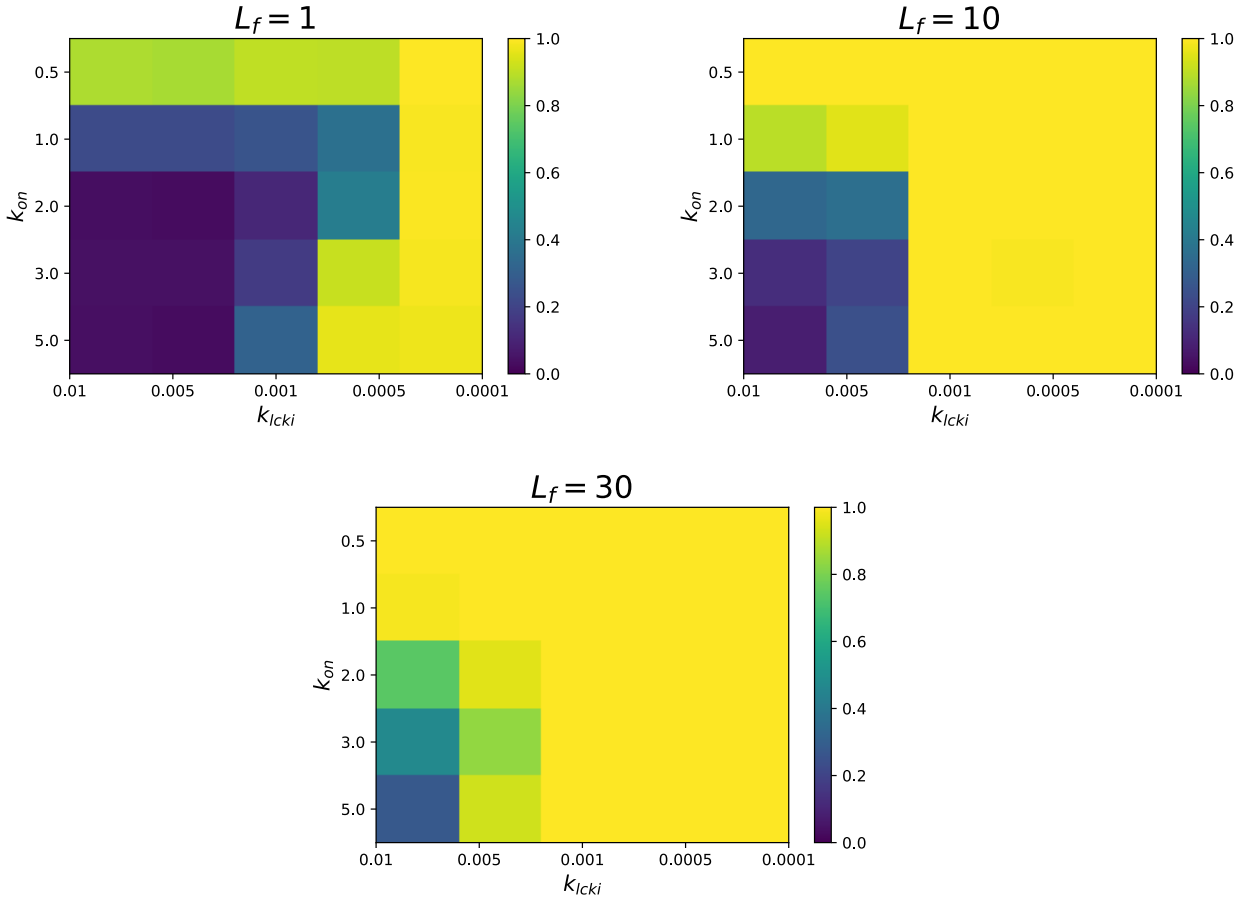
Figure S8: Two examples of early negative feedback loops in the T cell network. (**A**) shows the active form of ZAP-70 inactivating the free pool of Lck in the system. These reactions are combined with those shown in **Fig. S3A**. (**B**) offers a more realistic version of the negative feedback loop. Active ZAP-70 inactivates proximal Lck, which remains bound to

the TCR complex. Lck is inactive for the remainder of signaling. Eventually, Lcki dissociates from the complex in the side paths. Feedback strength depends on the reversion rate of Lcki to Lck.

We also investigated negative feedback of proximal, bound Lck due to activated ZAP-70. These reactions are shown in **Fig. S8B**. Active ZAP-70 converts the bound Lck to an inactive form, Lcki, by phosphorylating its negative regulatory residues. As a consequence, proximal Lck is inactivated and released in all sideways paths following cycle 3. The rate at which unbound Lcki is converted back to Lck determines the feedback loop strength. **Fig. S9B** measures the capacity for different main path on-rates (k_{on}) and feedback strengths (k_{lcki}). Surprisingly, we observe some differences in the realistic representation. There appears to be a sharp thresholding behavior: if the conversion of inactive to active Lck happens at an optimal rate, $k_{lcki} \sim 0.001$ for $L_f = 10$ or 30, discrimination suddenly becomes perfect. The latter behavior contrasts with **Fig. S9A**, which shows a gradual improvement in capacity as feedback strength increases.


$L_f = 1$
$L_f = 10$
$L_f = 30$
k_{on}
$k_{neg,\,fb}$
0.5
1.0
2.0
3.0
5.0
0.0
0.001
0.01
0.05
0.1
0.0
0.2
0.4
0.6
0.8
1.0

A



B

Figure S9: A full parameter search of the networks shown in **Fig. S8** for three different L_f values. (**A**) corresponds to the network in **Fig. S8A** while (**B**) corresponds to **Fig. S8B**. Feedback strength is inversely related to k_{lcki}. In (**A**), $k_{lcki} = 0.01s^{-1}$. (**A**) shows a gradual improvement of discrimination due to increasing feedback strength whereas (**B**) shows a sharp transition to perfect discrimination beyond a specific feedback strength. Surprisingly, the transition does not depend on k_{on}.

In the final variation, we explore the effect of a mock late negative feedback loop.

Suppose that RPL_Lck_Zap_P_LATP inactivates ZAP-70 thereby preventing its attachment to ITAMs and subsequent activation by Lck. Reactions associated with this scenario are shown in **Fig. S10A**. **Fig. S10B** compares the effect on capacity of early (**Fig. S8A**) and late (**Fig. S10A**) negative feedback. As expected, capacity is equivalent in both cases at $k_{neg,fb} = 0$. When $k_{neg,fb}$ is increased, for the same k_{on} and $k_{neg,fb}$, early negative feedback is more effective than late feedback at improving discrimination. The difference is even more pronounced for faster k_{on}.

$$L_f > 0:$$

$$RPL_s\_Lck\_Zap\_P\_LATP + Zap \xrightarrow{k_{neg,fb}} RPL_s\_Lck\_Zap\_P\_LATP + Zapi$$

$$RPL_f\_Lck\_Zap\_P\_LATP + Zap \xrightarrow{k_{neg,fb}} RPL_f\_Lck\_Zap\_P\_LATP + Zapi$$

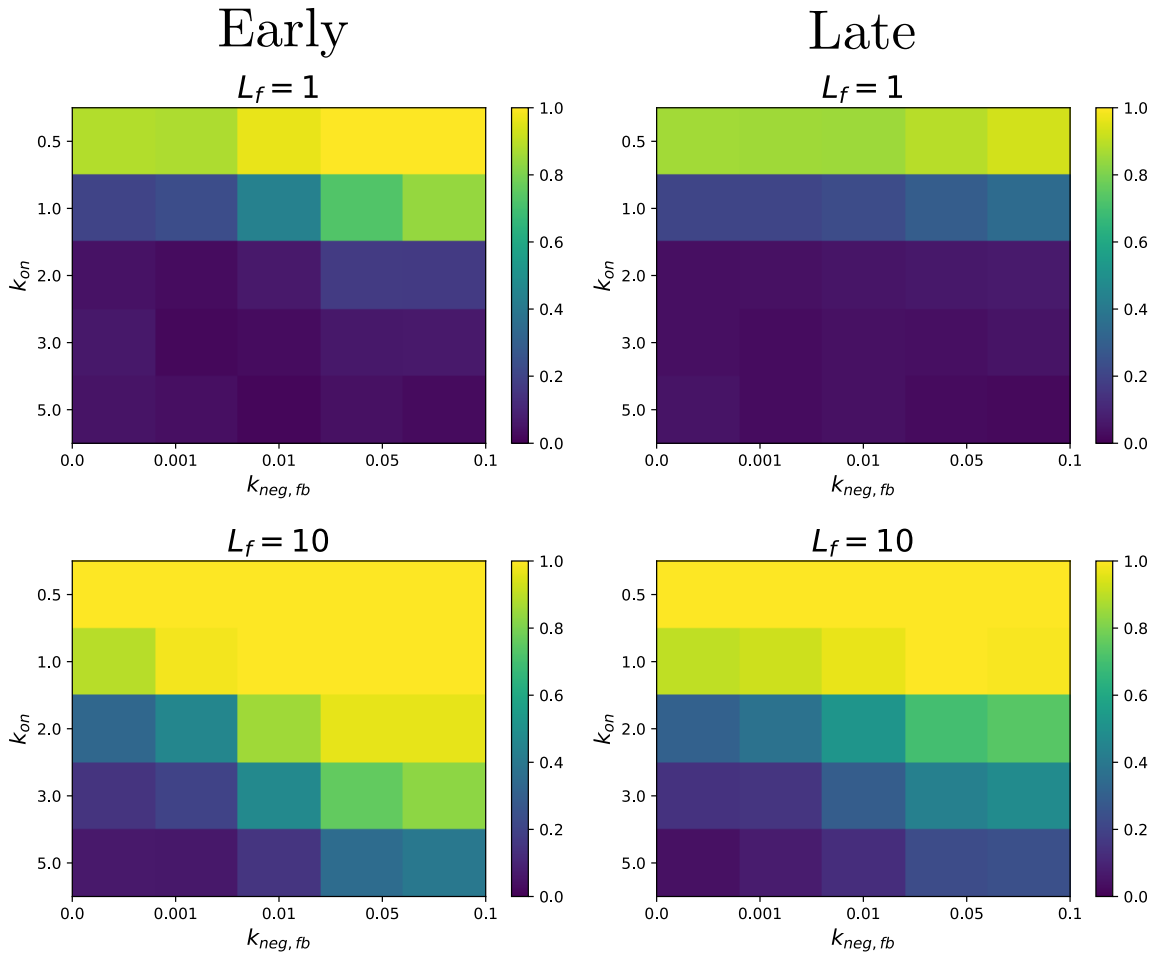
$$Zapi \xrightarrow{k_{zapi}} Zap$$

$$L_f = 0:$$

$$RPL_s\_Lck\_Zap\_P\_LATP + Zap \xrightarrow{k_{neg,fb}} RPL_s\_Lck\_Zap\_P\_LATP + Zapi$$

$$Zapi \xrightarrow{k_{zapi}} Zap$$

A



B

Figure S10: The effects of late negative feedback. (**A**) shows an essentially identical network to the one presented in **Fig. S8A** except that feedback occurs further downstream. We consider the consequences of RPL_Lck_Zap_P_LATP inactivating the free pool of ZAP-70. (**B**) shows a direct comparison of early and late negative feedback. Clearly, late feedback is less effective at improving discrimination.

For the calculations involving the negative feedback loops presented in **Figs. S8A** and **S9A**, rate constants are set to those detailed in Tables 2 and 5. For **Fig. S9B**, the rate of conversion from inactive to active Lck determines the strength of the negative feedback

loop. Therefore, in the latter case, k_{lcki} is varied. In the early and late negative feedback comparison presented in **Fig. S10**, $[Lck]_0 = [Zap]_0 = 30000\ mol$. Then, $k_{lck,on} = k_{zap,on} = k_{on} / [Lck]_0 = k_{on} / [Zap]_0$.

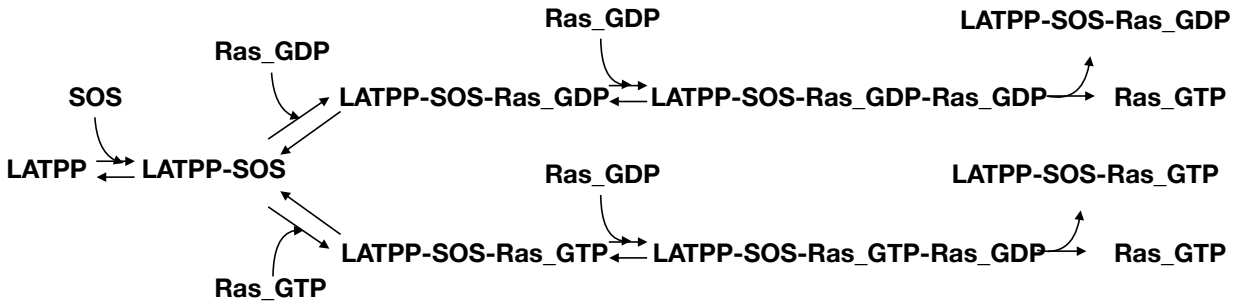
Table 5: Rate constants used for negative feedback loops.

Rate Constants **Figs. S8A** and **S10A**	Value
$k_{neg,fb}$	Varies
k_{lcki}	$0.01\ s^{-1}$
k_{zapi}	$0.01\ s^{-1}$

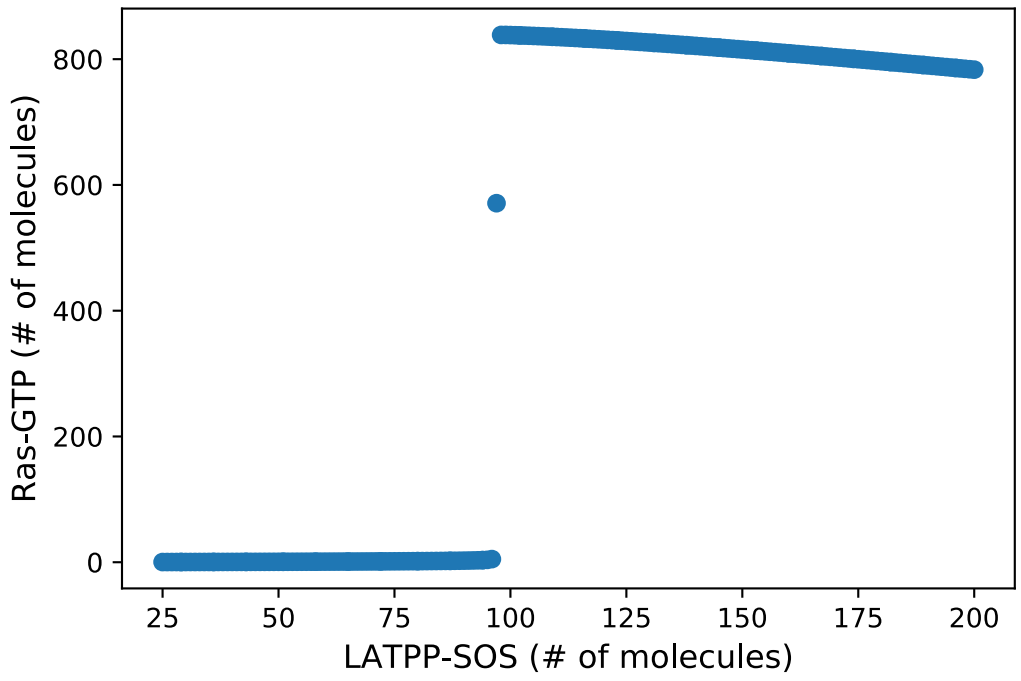
Positive Feedback Loops:

We explore the effects of a late positive feedback loop by connecting the SOS feedback loop (see **Fig. S11**) to the kinetic proofreading network shown in **Fig. S3A**, while including the negative feedback loop in **Fig. S8A**. T cells make the decision to activate extremely quickly: after the initial binding event, it takes $1-5$ minutes. **Fig. S12A** shows the dynamics of forming the LATPP-SOS complex for $k_{on} = 1.5 s^{-1}$ and $k_{neg,fb} = 0.05(mol \cdot s)^{-1}$. Within the typical time-scale of activation and across a range of self ligand concentrations, we observe that LATPP-SOS output is significantly greater when agonist ligand is present. **Fig. S12B** shows the output distributions at $t = 300s$. After 8 steps, discrimination becomes essentially perfect and little overlap remains between the output distributions. Assuming the bi-stability point is located at $[\text{LATPP-SOS}]_{crit} \approx 97$ as shown by the red dotted line in **Fig. S12B** we expect that output

from self in the presence of agonist ligands will be amplified, whereas self only output will be quenched.

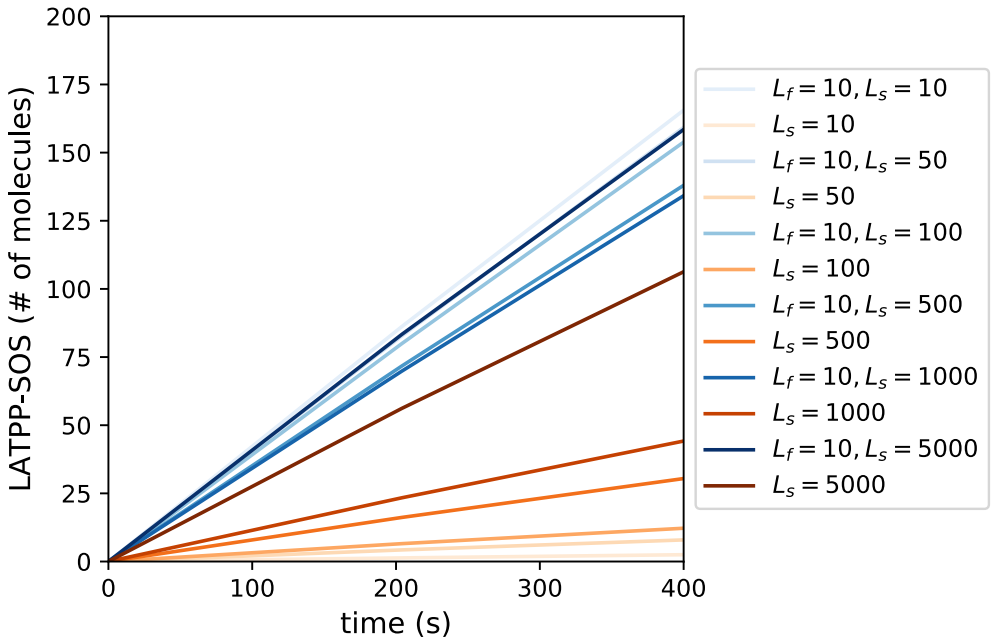


A

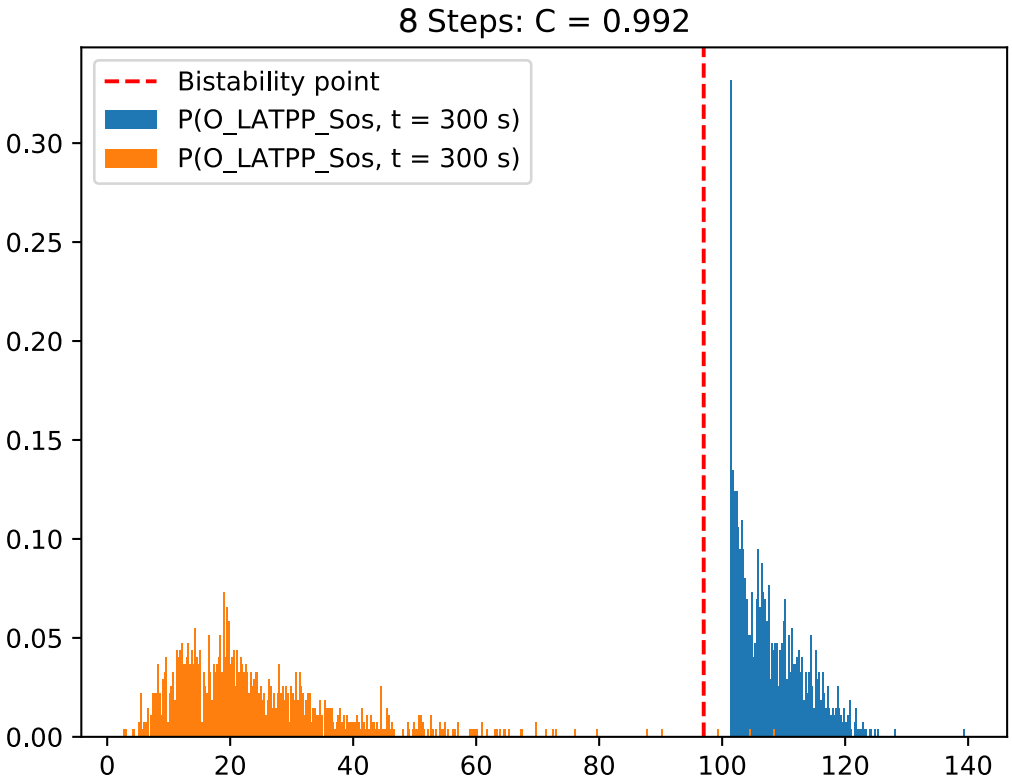


B

Figure S11: Ras-SOS positive feedback loop. (**A**) shows the reactions that lead to bi-stable dependence of Ras-GTP on SOS (4). If parameters are chosen so that Ras_GTP-bound SOS is more effective at phosphorylating Ras_GDP than Ras_GDP-bound SOS, there will be a bi-stability point such as the one shown in (**B**).



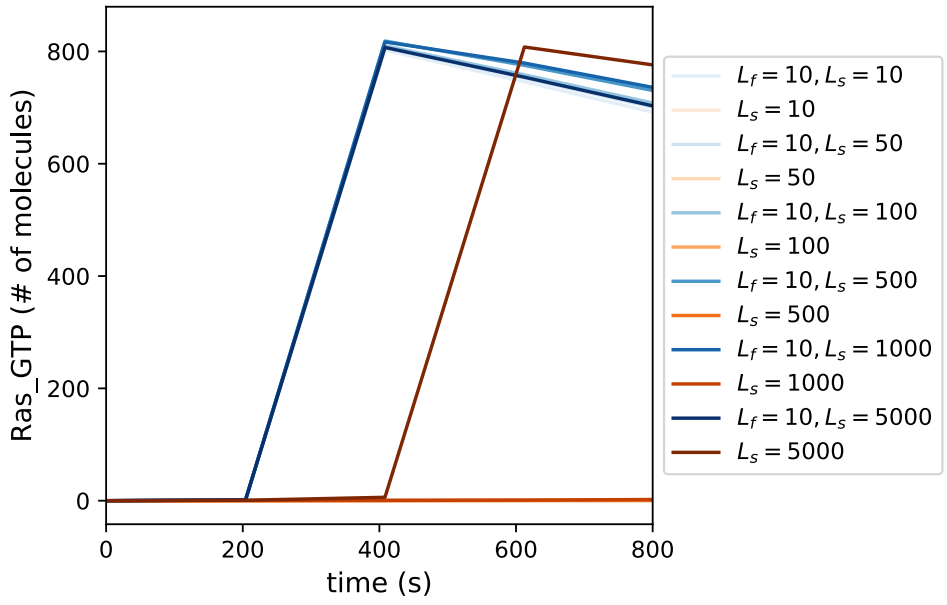
A



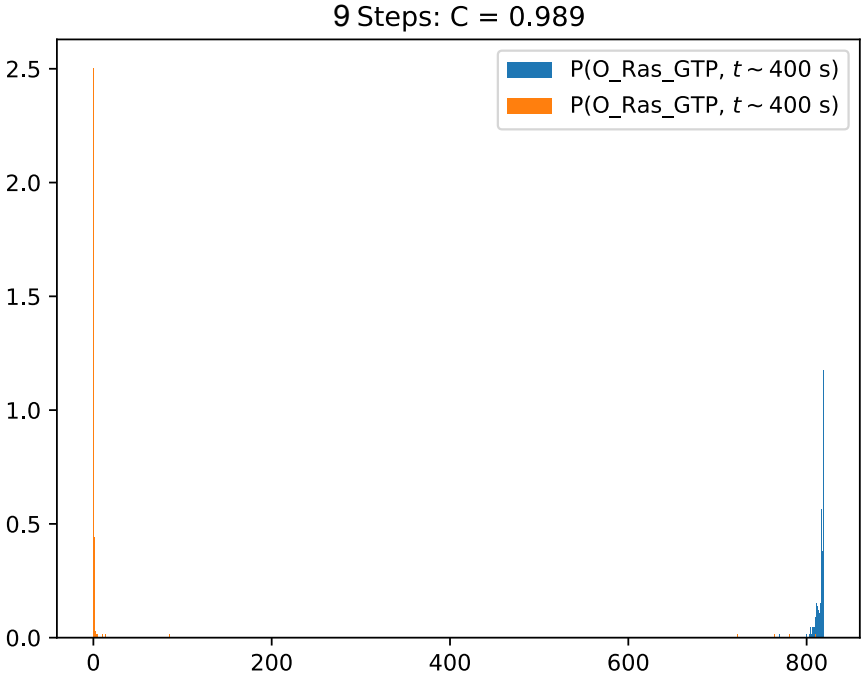
B

Figure S12: Dynamics of LATPP-SOS formation for $k_{on} = 1.5\ \ s^{-1}$ and $k_{neg,fb} = 0.05$ $(mol \cdot s)^{-1}$. (**A**) shows the production of LATPP-SOS after combining the first reaction in **Fig. S11A** with those illustrated in **Fig. S3A**. Production of LATPP-SOS is consistently greater when agonist ligands are present over a range of self concentrations. Line darkness is proportional to the number of self ligands in the simulation. (**B**) shows normal (orange) and infected (blue) cell output distributions at $t = 300s$ when sampling over the full self ligand distribution. The bi-stability point in **Fig. S11B** is shown here as a red dotted line.

Indeed, **Fig. S13A** shows that after $t \approx 200s$, the SOS feedback loop causes rapid catalytic conversion of Ras_GDP to Ras_GTP. Signal amplification (blue) happens within the time-scale of decision-making whereas positive feedback of noise (orange) would occur too late. The histogram in **Fig. S13B** shows two sharply peaked distributions corresponding to inactivation (orange) and activation (blue). We conclude that the SOS positive feedback loop converts the distribution of product from the upstream kinetic proofreading network into digital output.



A



B

Figure S13: Dynamics of Ras_GTP formation. (**A**) shows the production of Ras_GTP after combining all of the reactions in **Fig. S11A** with **Fig. S3A**. At $t \sim 200s$, there is a sudden increase in production of Ras_GTP, showing that infected cell output has exceeded the bi-stability point. Any amplification due to interaction with a normal APC happens outside the time-scale of T cell activation. (**B**) shows the resulting output distributions of normal (orange) and infected (blue) cells at $t \sim 400s$. The positive feedback loop converts output into two sharply peaked distributions corresponding to "Off" and "On".

In the positive feedback loop calculations presented in **Figs. S12B-S13B**, $k_{on} = 1.5 \ \ s^{-1}$, the second LAT phosphorylation step proceeds at $k_{on} = 0.15 \ \ s^{-1}$, $k_{neg,fb} = 0.05$

$(mol \cdot s)^{-1}$, and $k_{lcki} = 0.01\ s^{-1}$. Tables 6-7 show parameter values for the Ras-SOS reactions.

Table 6: Initial Concentrations for Ras-SOS feedback loop

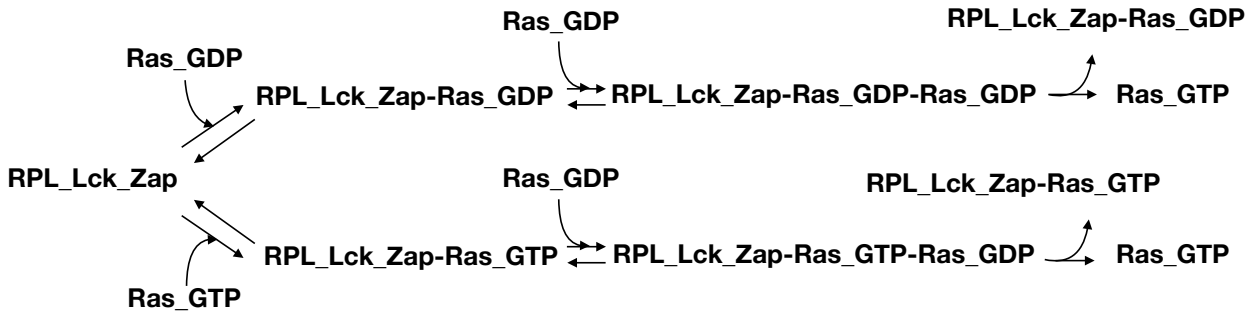
Species	Total Concentration (number of molecules)
All initial concentrations in Table 1	
$[SOS]_0$	1000
$[Ras\_GDP]_0$	1000
$[Ras\_GAP]_0$	10

Table 7: Rate constants for Ras-SOS positive feedback loop.

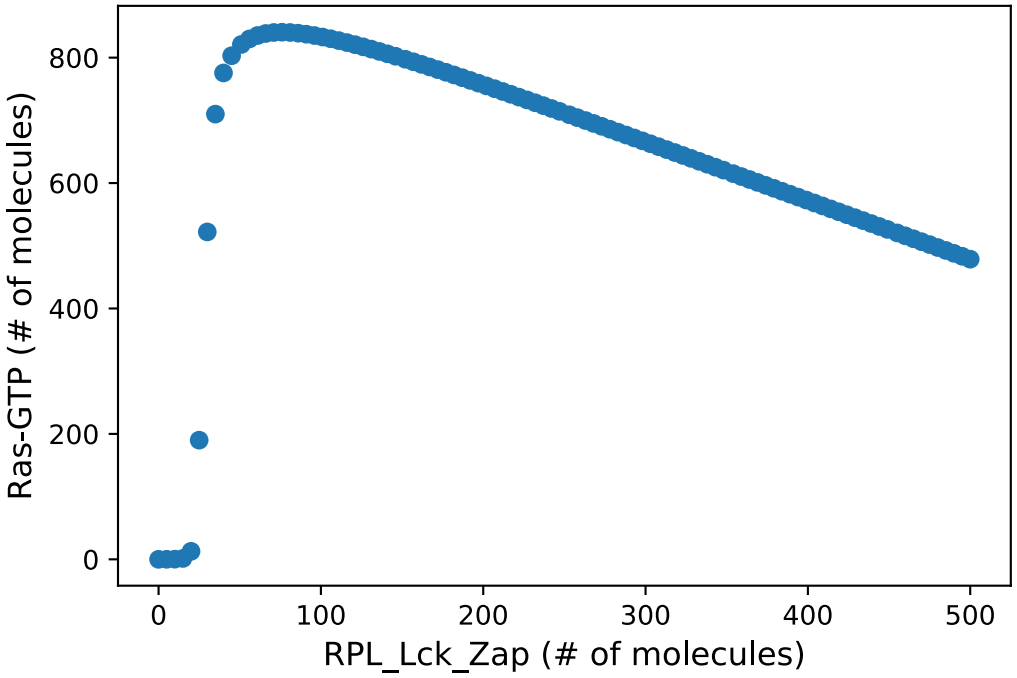
Reactions	$k_{forward}$	$k_{reverse}$
All reactions in Table 2 including negative feedback		
$LATPP + SOS \rightleftharpoons LATPP - SOS$	0.001 $(mol \cdot s)^{-1}$	0.005 s^{-1}
$LATPP - SOS + Ras\_GDP \rightleftharpoons LATPP - SOS - Ras\_GDP$	0.024 $(mol \cdot s)^{-1}$	30.0 s^{-1}
$LATPP - SOS - Ras\_GDP + Ras\_GDP \rightleftharpoons$ $LATPP - SOS - Ras\_GDP - Ras\_GDP$	0.014 $(mol \cdot s)^{-1}$	10.0 s^{-1}
$LATPP - SOS - Ras\_GDP - Ras\_GDP \rightarrow$ $LATPP - SOS - Ras\_GDP + Ras\_GTP$	0.03 s^{-1}	
$LATPP - SOS + Ras\_GTP \rightleftharpoons LATPP - SOS - Ras\_GTP$	0.022 $(mol \cdot s)^{-1}$	4.0 s^{-1}
$LATPP - SOS - Ras\_GTP + Ras\_GDP \rightleftharpoons$ $LATPP - SOS - Ras\_GTP - Ras\_GDP$	0.01 $(mol \cdot s)^{-1}$	1.0 s^{-1}

$LATPP - SOS - Ras\_GTP - Ras\_GDP \rightarrow LATPP - SOS - Ras\_GTP + Ras\_GTP$	$0.494\ s^{-1}$	
$Ras\_GAP + Ras\_GTP \rightleftharpoons Ras\_GAP - Ras\_GTP$	$0.348\ (mol \cdot s)^{-1}$	$2.0\ s^{-1}$
$Ras\_GAP - Ras\_GTP \rightarrow Ras\_GAP + Ras\_GDP$	$1.0\ s^{-1}$	

Lastly, we explore the consequences of an early positive feedback loop in **Figs. S14-S15**. Early in the kinetic proofreading process, capacity is low (see **Fig. S15A**) as one would expect. If a positive feedback loop similar to the actual one initiated by LAT (see **Fig. S11A**) is engaged instead by RPL_Lck_Zap (see **Fig. S14A**), self output (orange in **Fig. S15**) that exceeds a threshold value (red dotted line in **Fig. S15A**) will be rapidly amplified. Clearly, from **Fig. S15B** a drop in discrimination occurs if a bi-stable regime happens before the distributions completely separate.

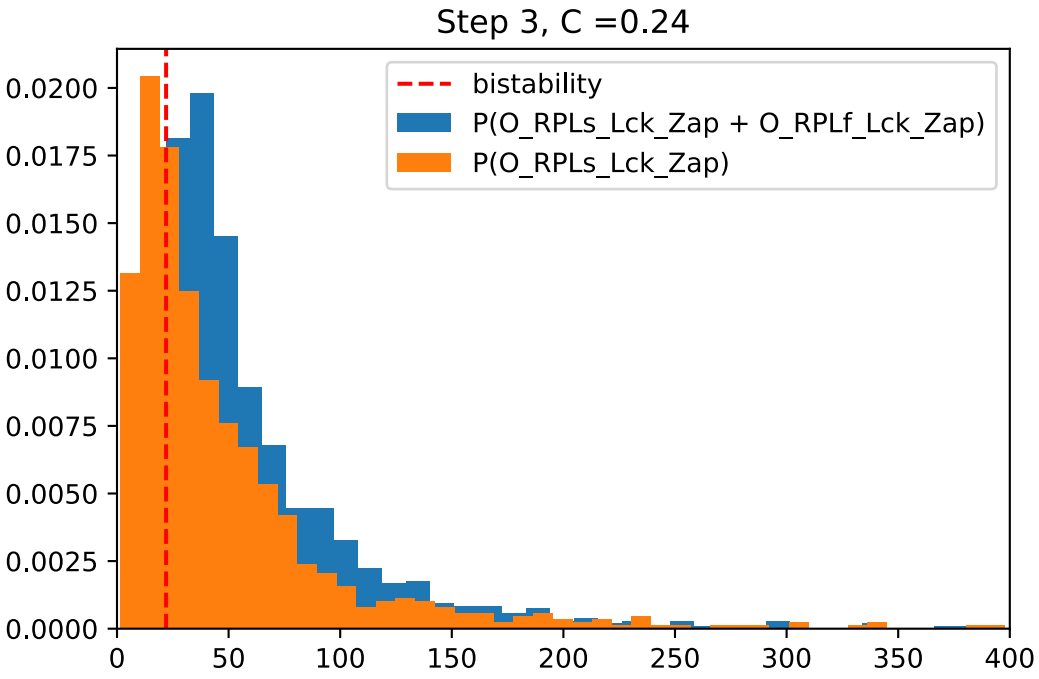


A

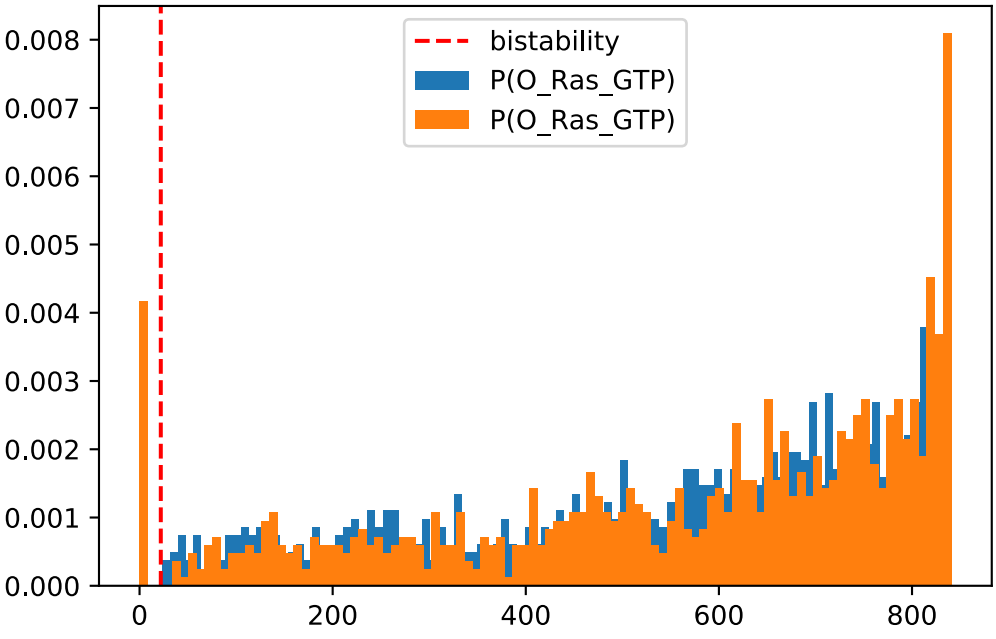


B

Figure S14: Early positive feedback loop. (**A**) shows an alternative set of reactions that lead to (**B**) bi-stable dependence of Ras-GTP on RPL_Lck_Zap.


Step 3, C =0.24
bistability
P(O_RPLs_Lck_Zap + O_RPLf_Lck_Zap)
P(O_RPLs_Lck_Zap)
0.0200
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
0
50
100
150
200
250
300
350
400

A

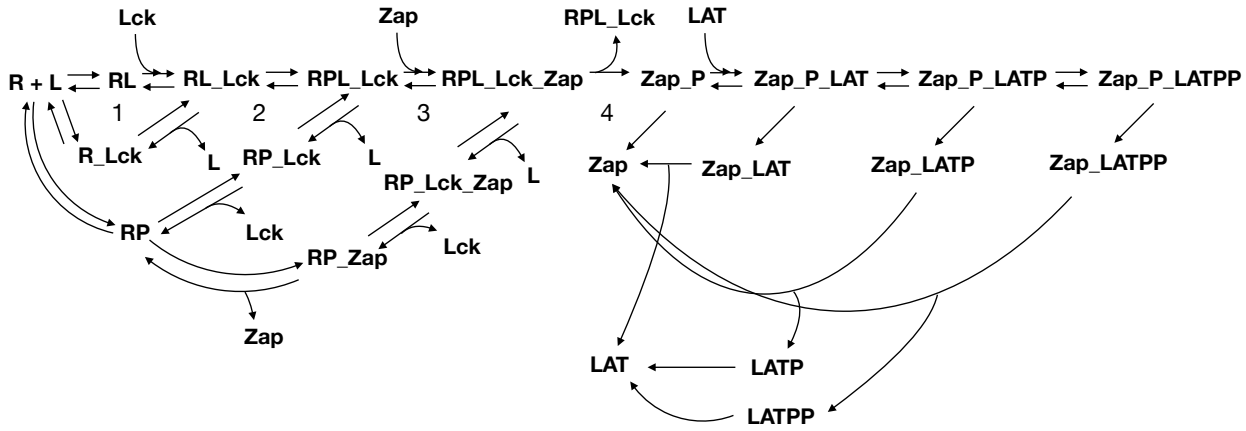

bistability
P(O_Ras_GTP)
P(O_Ras_GTP)
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
0
200
400
600
800

B

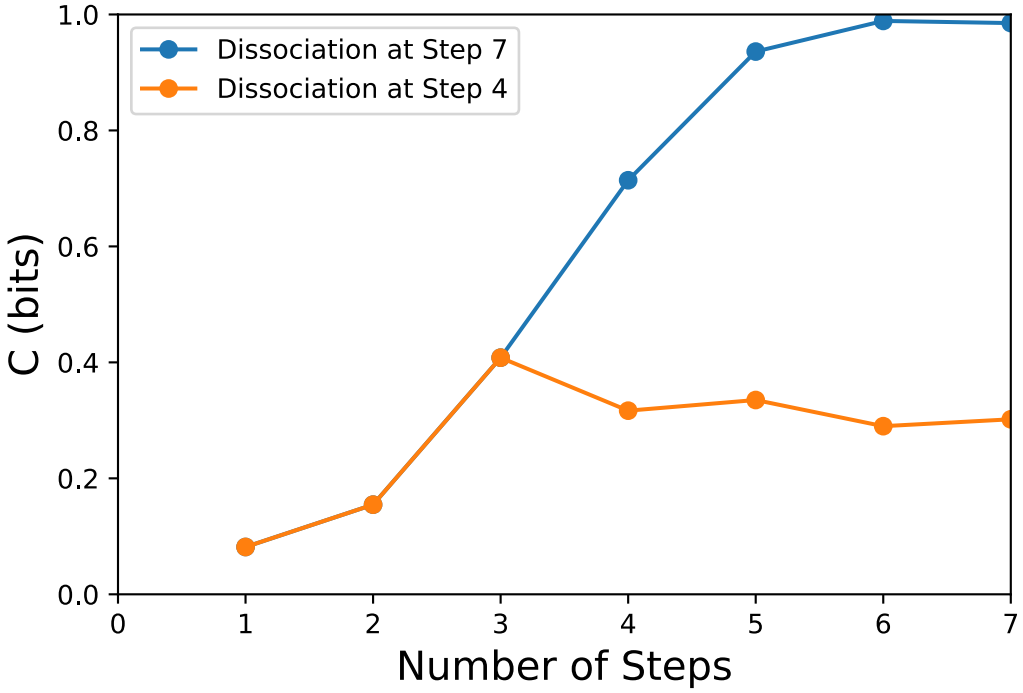
Figure S15: (**A**) shows histograms of RPL_Lck_Zap when the distributions have not completely separated ($C \sim 0.24$). Attempting positive feedback via a cubic dependence of Ras-GTP on RPL_Lck_Zap leads to amplification of self output and a discriminatory drop (**B**).

<u>"Catch-and-release" vs. "Bridge" model:</u>

We numerically investigate the "Catch-and-release" model suggested by Katz et al (5) using the network shown in **Fig. S16A**. After ZAP-70 binds to the ITAMs, Lck phosphorylates ZAP-70 and activated Zap_P dissociates from the receptor-ligand complex. Downstream biochemical steps proceed as before except discriminating between correct and incorrect product is not possible. Consequentially, after Zap_P dissociation, **Fig. S16B** shows the capacity drop and remain relatively constant as predicted by the data processing inequality. Early dissociation of active ZAP-70 from the receptor-ligand complex makes discriminatory improvement no longer possible.



A



B

Figure S16: (**A**) Reaction network showing early Zap dissociation from the receptor-ligand complex. (**B**) Numerical test of Zap dissociation after Lck activation. If unbound Zap_P proceeds with downstream signaling, the capacity can no longer increase (orange) resulting in a discriminatory drop with respect to the network (blue) in which Zap remains bound.

SI References

1. G. Altan-Bonnet, R. N. Germain, Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biol*. **3**, 1925–1938 (2005).

2. C. F. Lopez, J. L. Muhlich, J. A. Bachman, P. K. Sorger, Programming biological models in Python using PySB. *Mol. Syst. Biol*. (2013) https:/doi.org/10.1038/msb.2013.1.

3. P. N. Brown, G. D. Byrne, A. C. Hindmarsh, VODE: A Variable-Coefficient ODE Solver. *SIAM J. Sci. Stat. Comput*. (1989) https:/doi.org/10.1137/0910062.

4. J. Das, *et al*., Digital Signaling and Hysteresis Characterize Ras Activation in Lymphoid Cells. *Cell* (2009) https:/doi.org/10.1016/j.cell.2008.11.051.

5. Z. B. Katz, L. Novotná, A. Blount, B. F. Lillemeier, A cycle of Zap70 kinase activation and release from the TCR amplifies and disperses antigenic stimuli. *Nat. Immunol*. (2017) https:/doi.org/10.1038/ni.3631.