Rare Event Simulation for Tcells recognizing foreign antigens

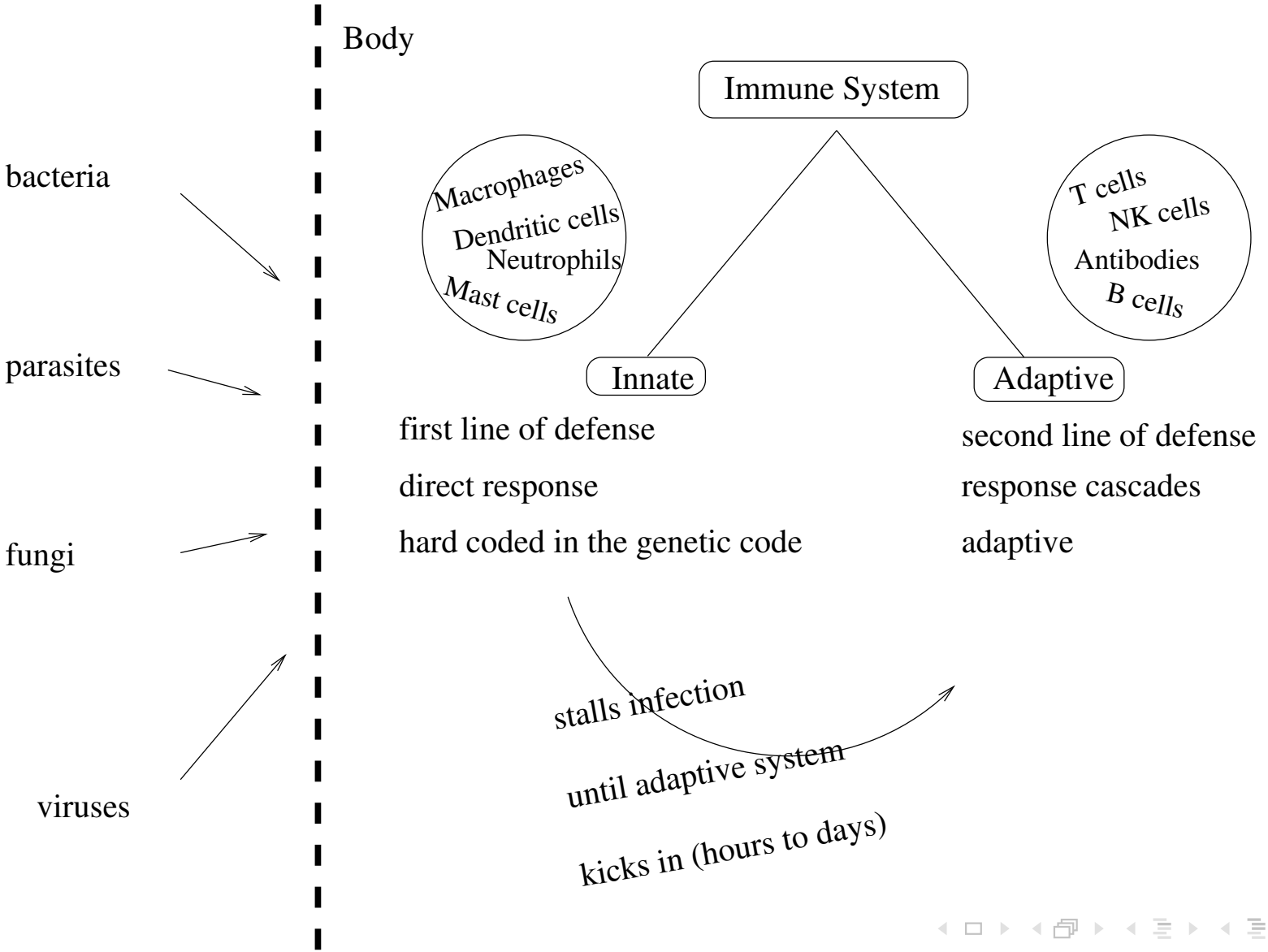
Florian Lipsmeier

joint work with Ellen Baake

Biomathematics and theoretical Bioinformatics Group
Bielefeld University

RESIM2008

immune system



self and nonself

antigen

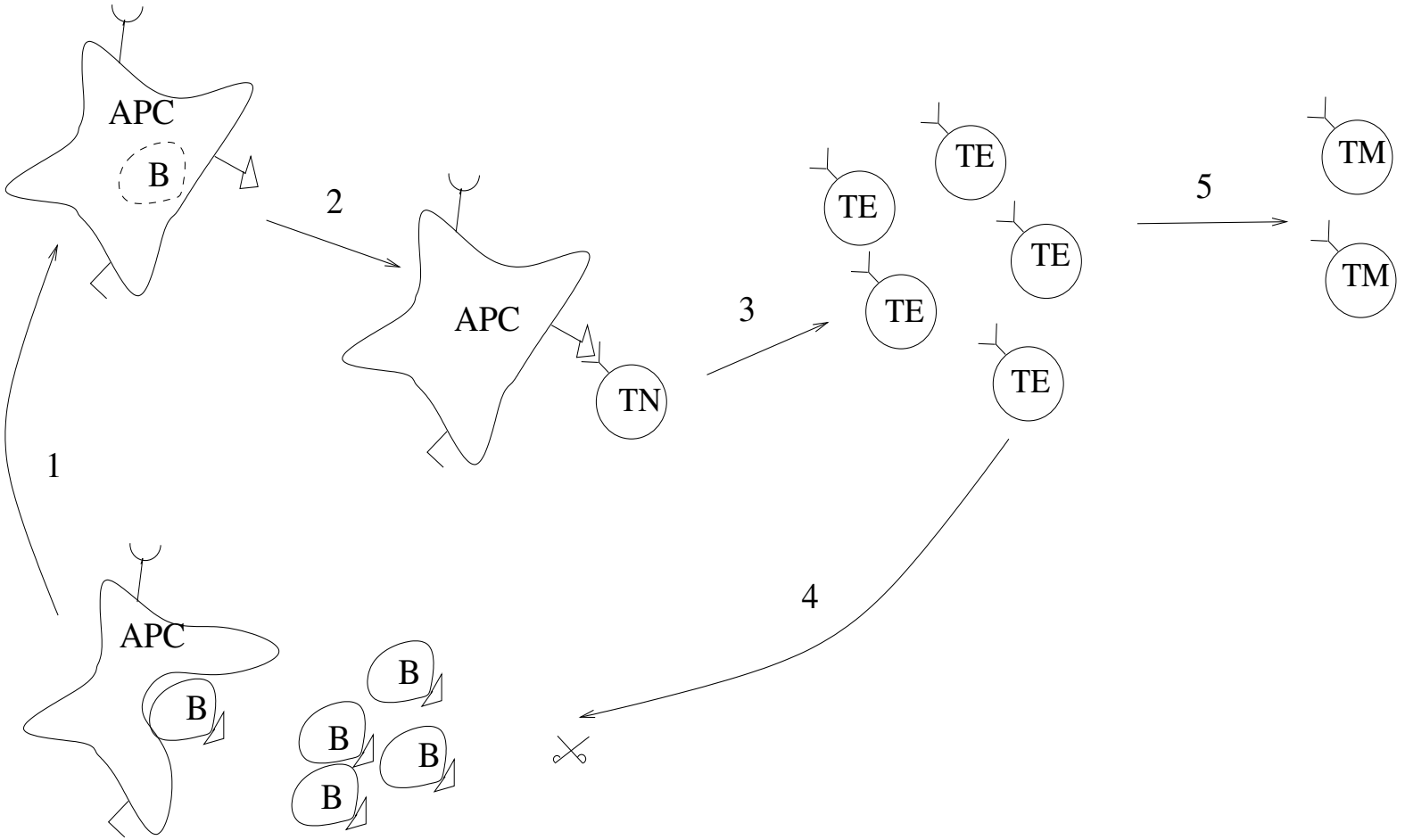
An antigen is anything that can trigger an immune response, most of the time protein fragments on cell surfaces

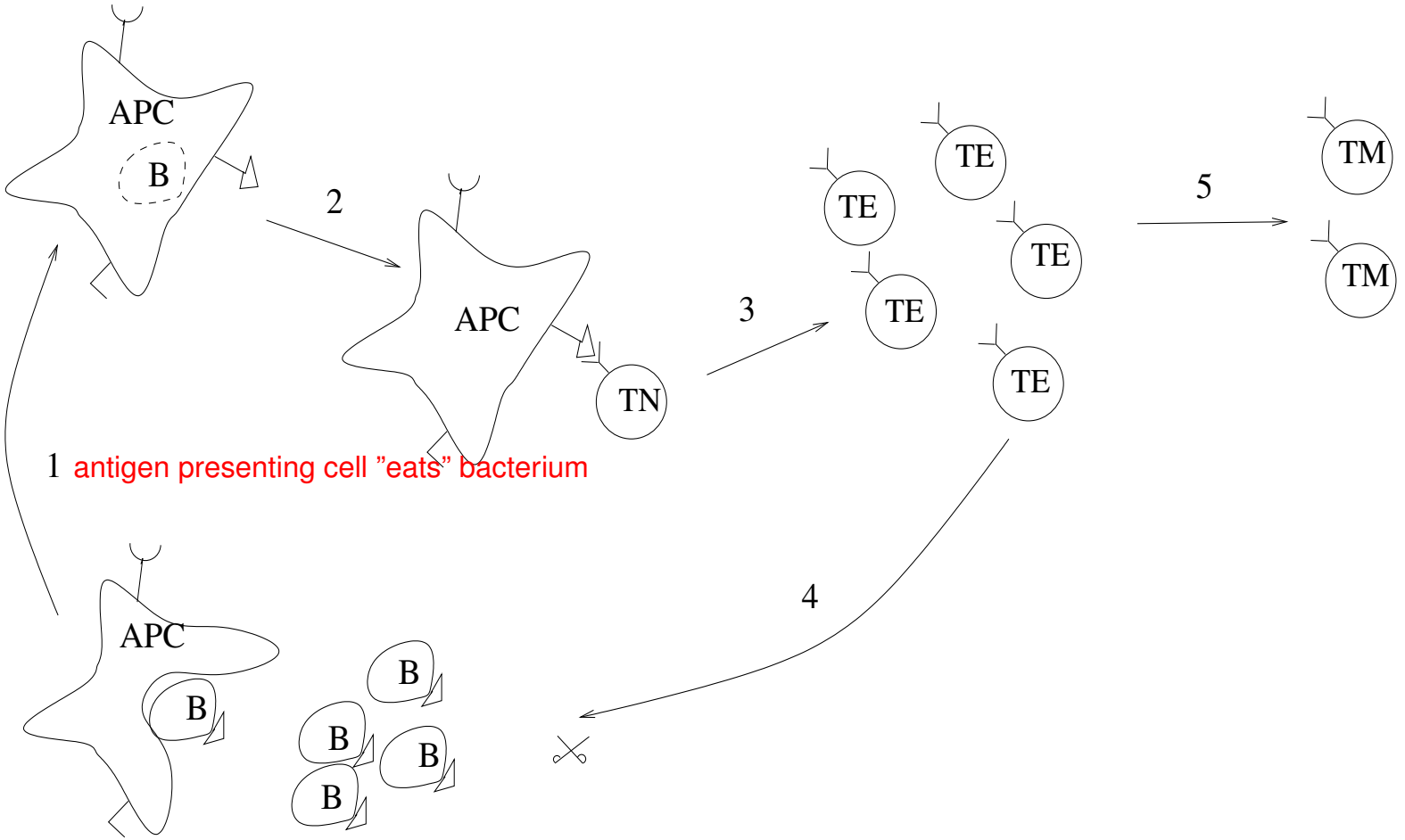
- self antigen (fragments of the body's own cells)
- foreign antigen (fragments of microbes or viruses, foreign tissue,...)

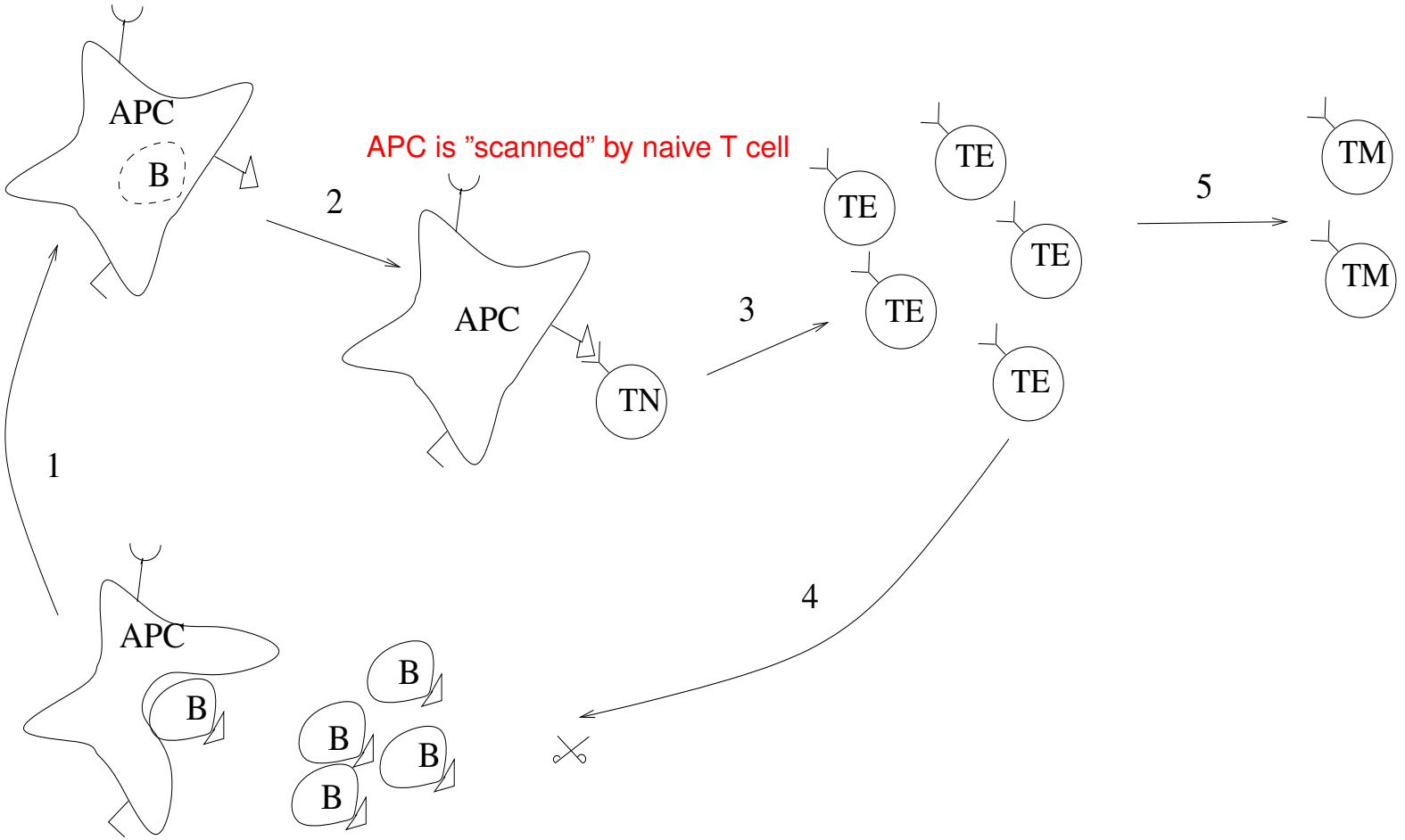
receptor

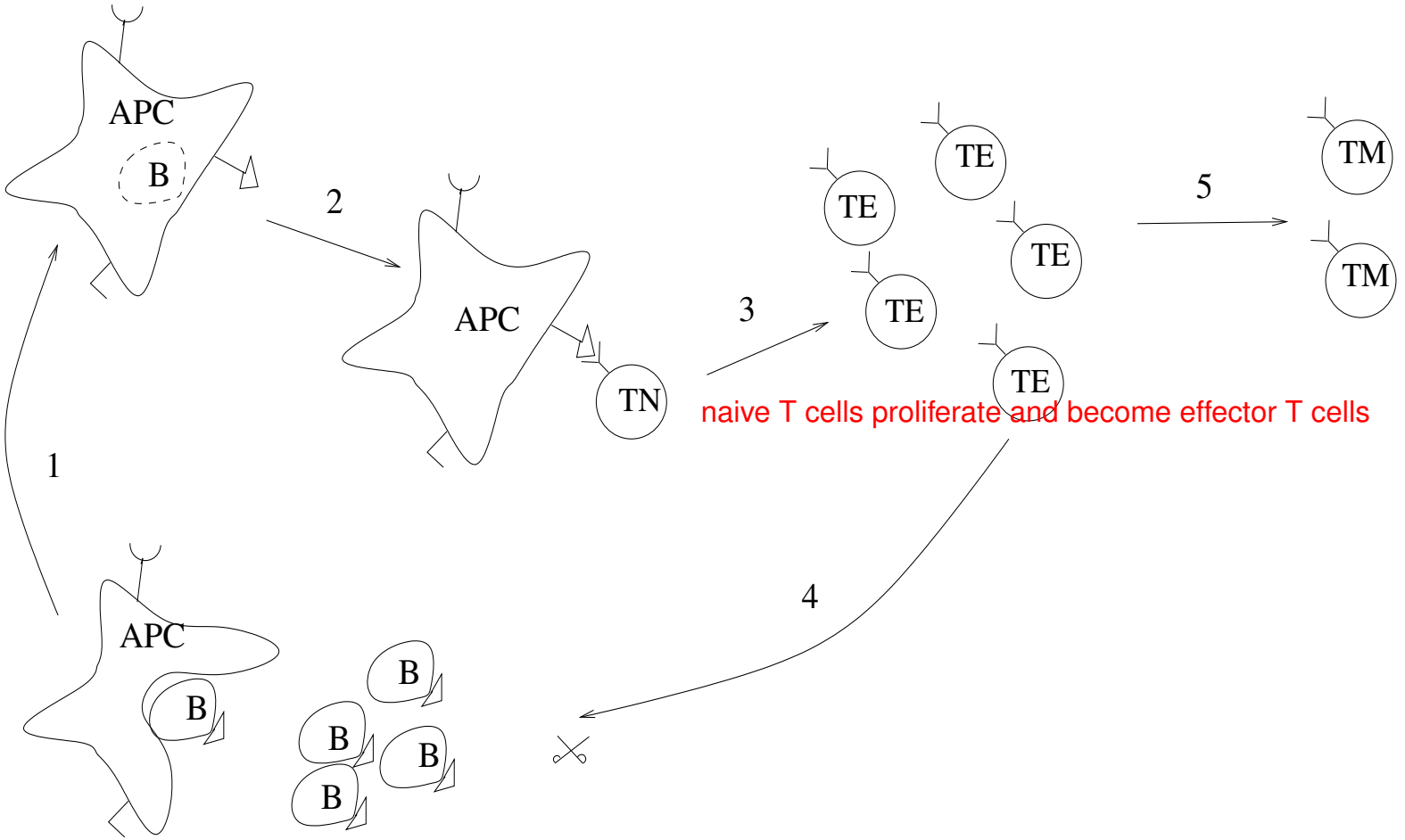
- binds to special classes of cell surface molecules
- innate immune system: several pre-defined receptors
- adaptive immune system: new generation of receptor types via shuffling of genes

"Key-Lock-System"

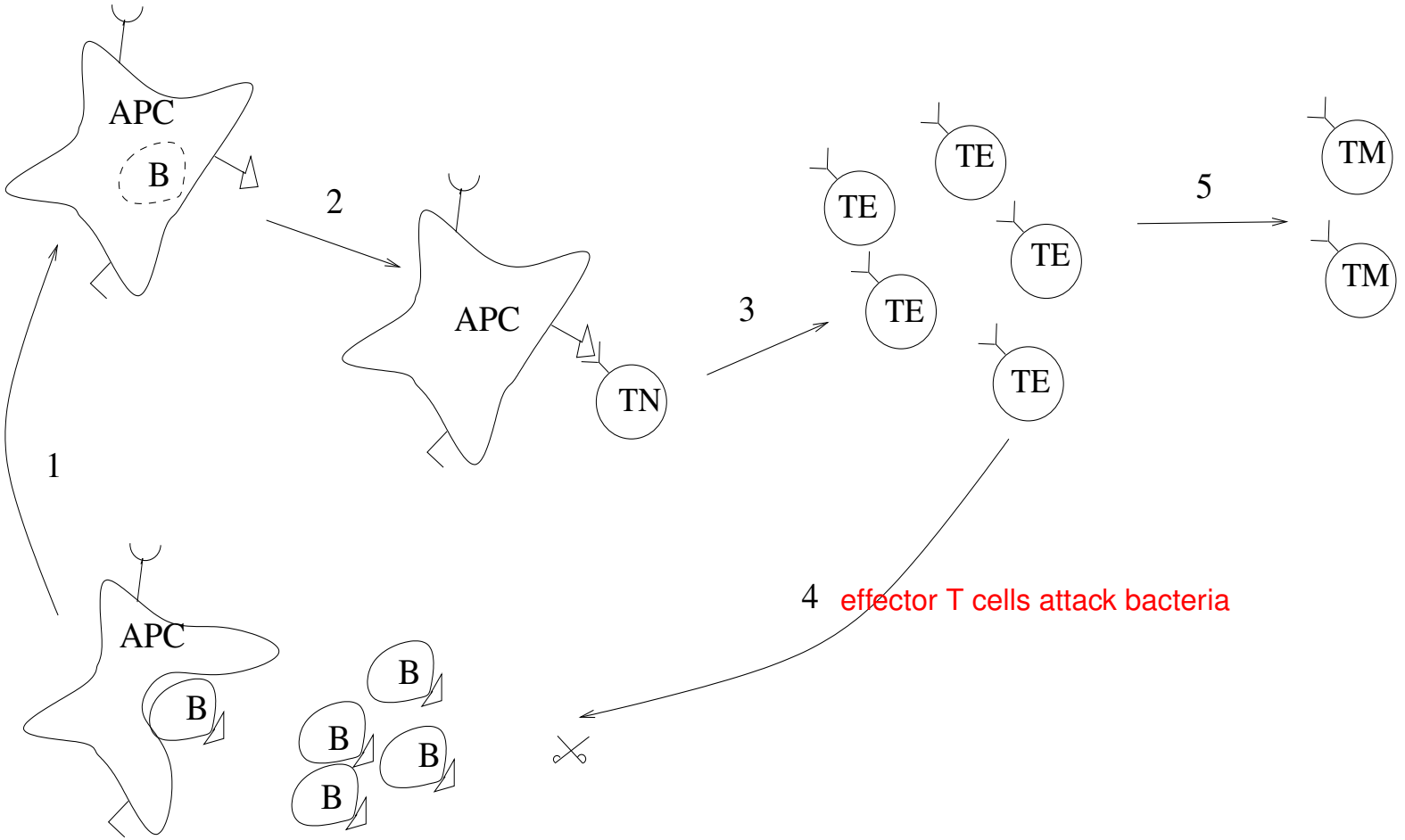

APC
B
2
APC
TN
3
TE
TE
TE
TE
5
TM
TM
1
APC
B
B
B
B
B
4

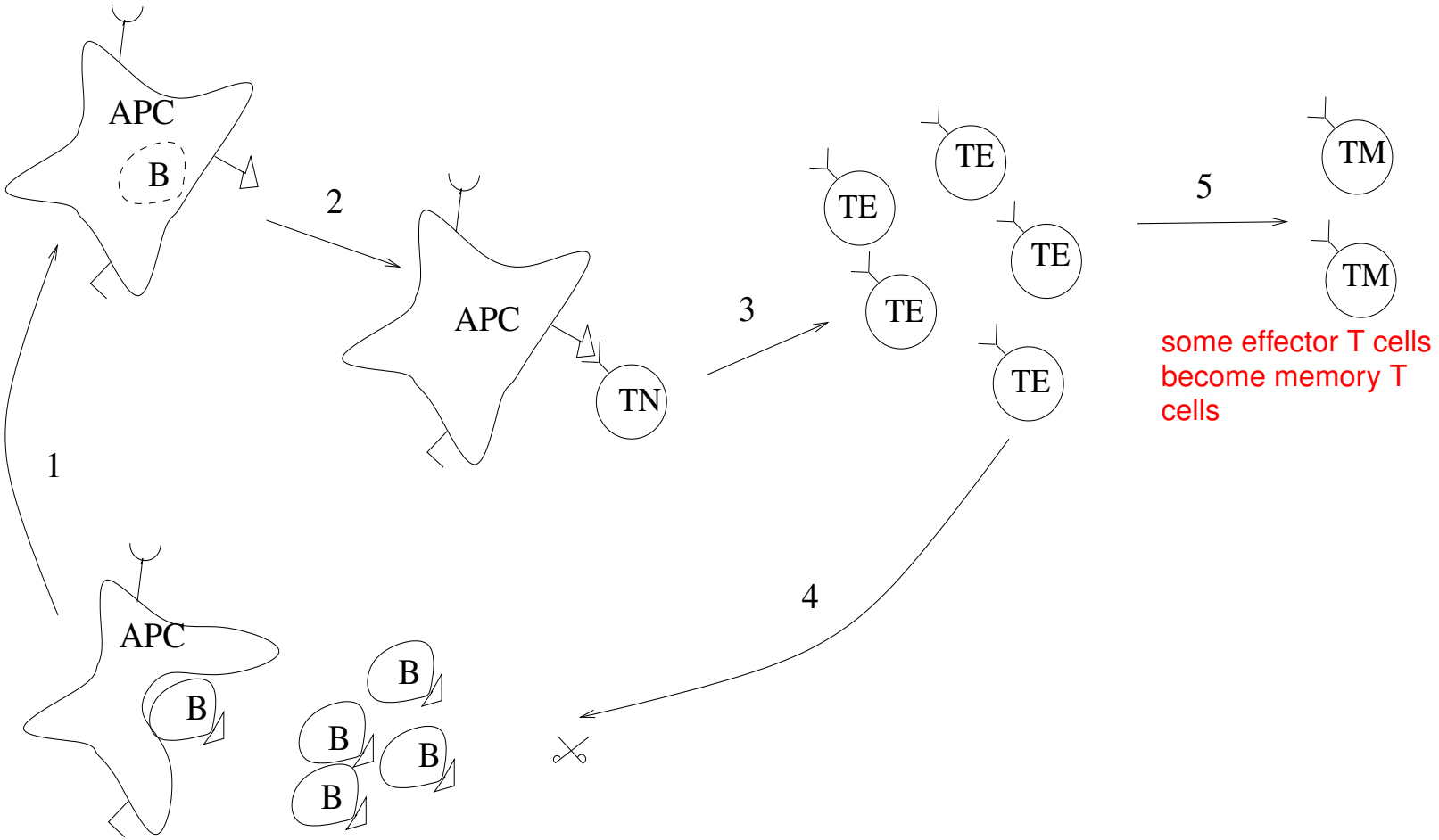

APC
B
2
APC
TN
3
TE
TE
TE
TE
TE
5
TM
TM
1 antigen presenting cell ”eats” bacterium
4
APC
B
B
B
B
B


APC
B
APC is ”scanned” by naive T cell
2
APC
TN
3
TE
TE
TE
TE
5
TM
TM
1
APC
B
B
B
B
B
4

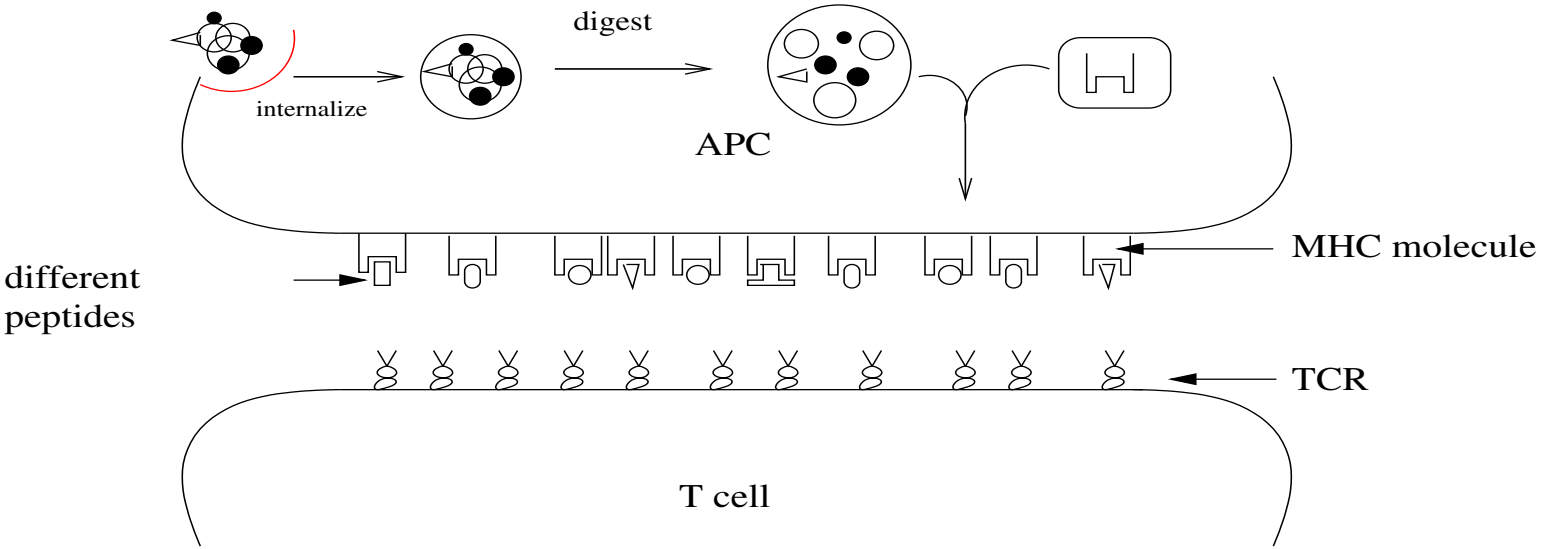


naive T cells proliferate and become effector T cells


APC
B
2
APC
TN
3
TE
TE
TE
TE
5
TM
TM
1
APC
B
B
B
B
B
4 effector T cells attack bacteria

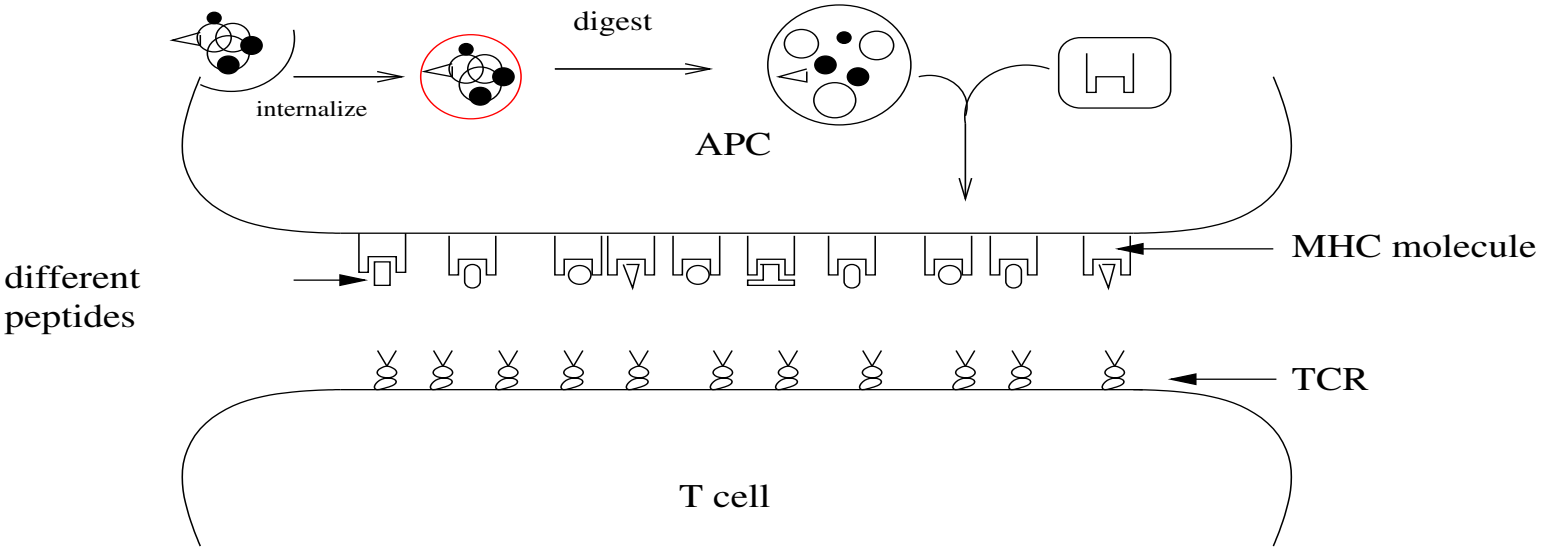

APC
B
1
2
APC
TN
3
TE
TE
TE
TE
TE
4
APC
B
B
B
B
B
5
TM
TM
some effector T cells become memory T cells

caricature model of T cell-Dynamics



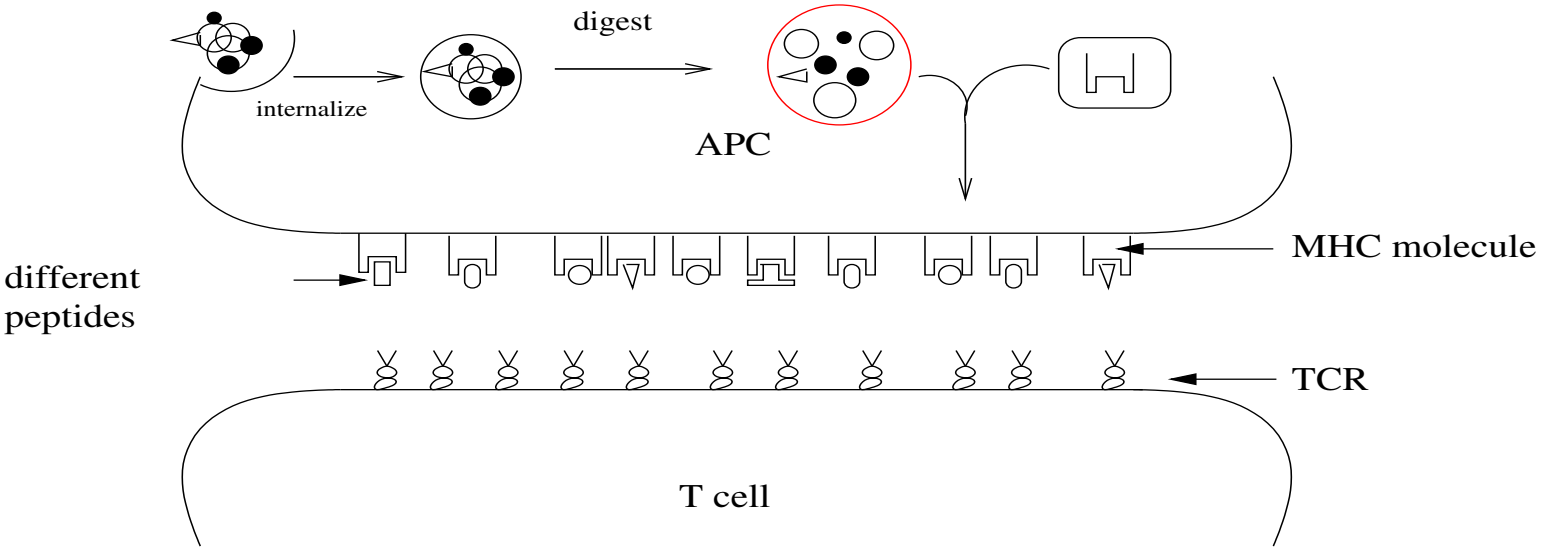
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



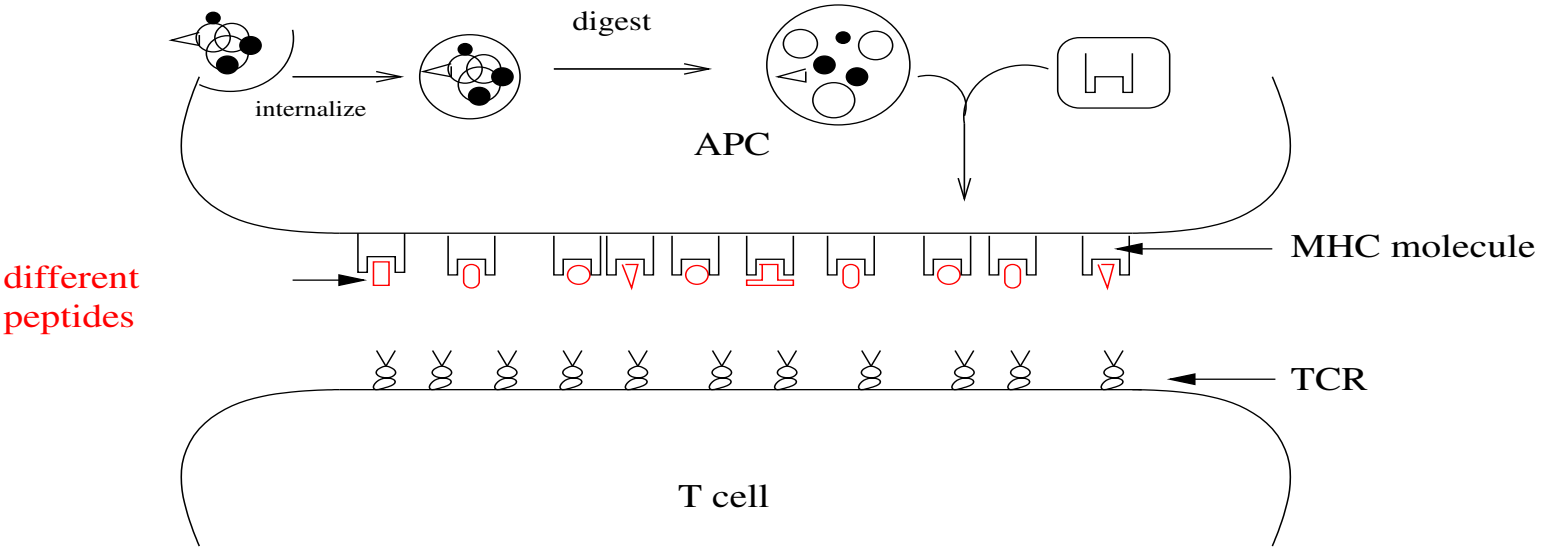
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



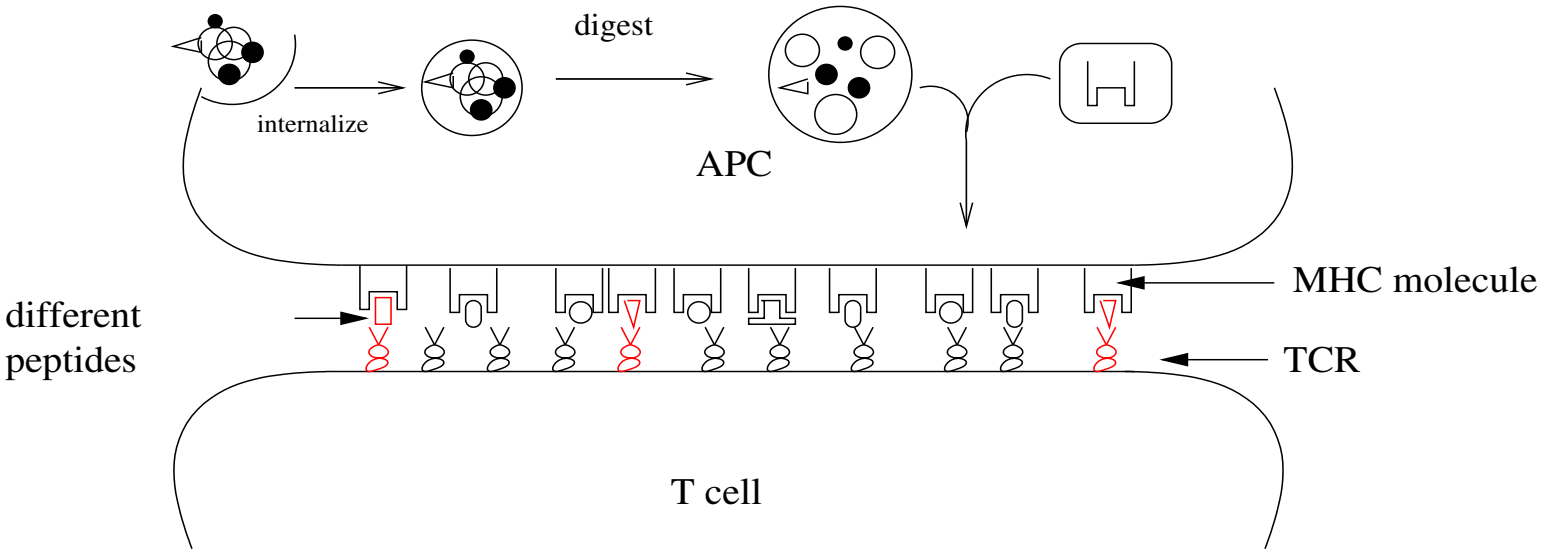
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



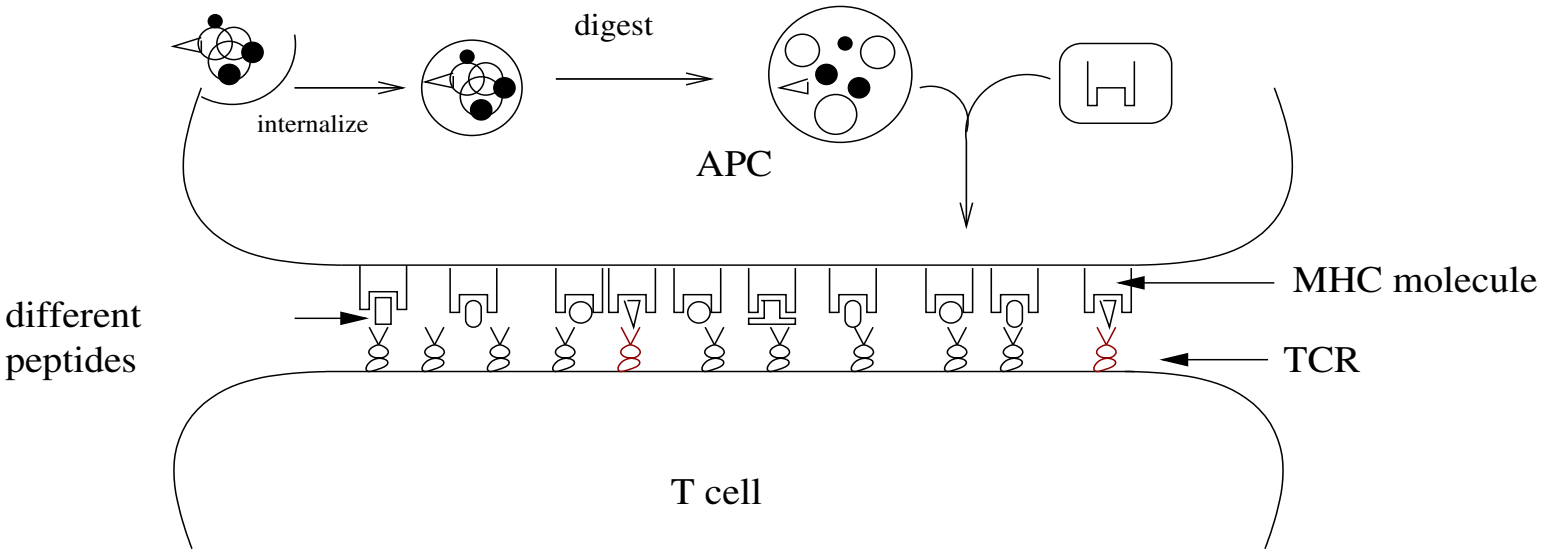
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



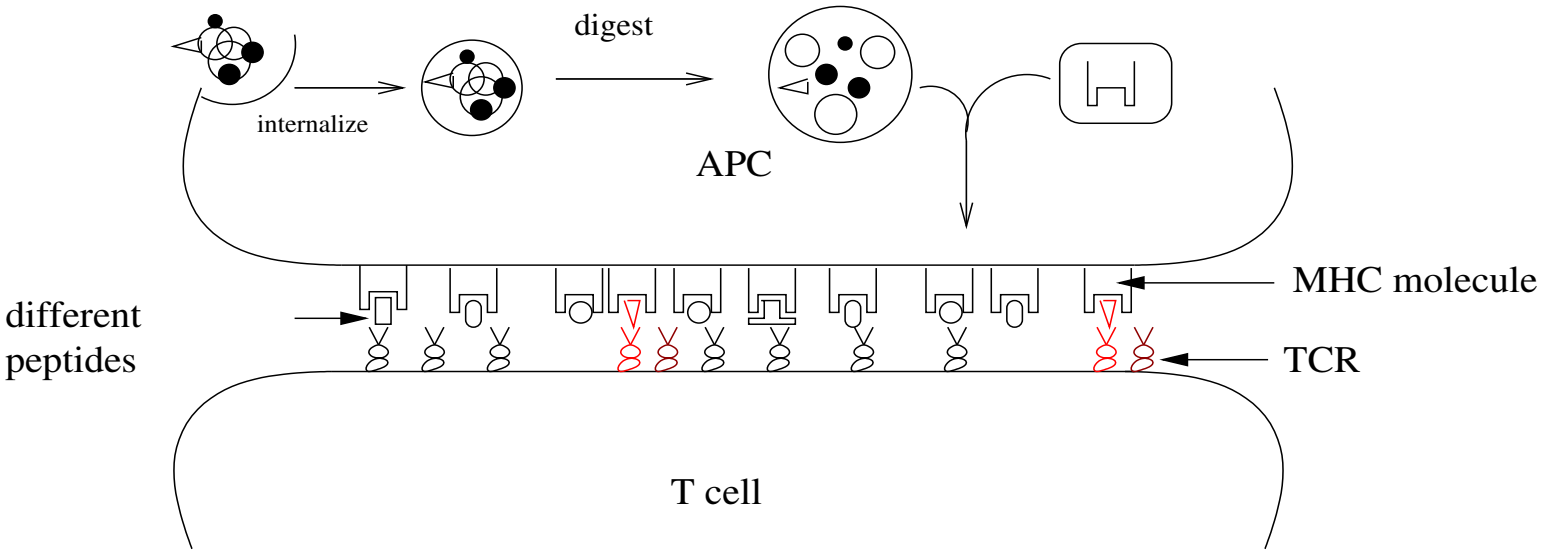
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



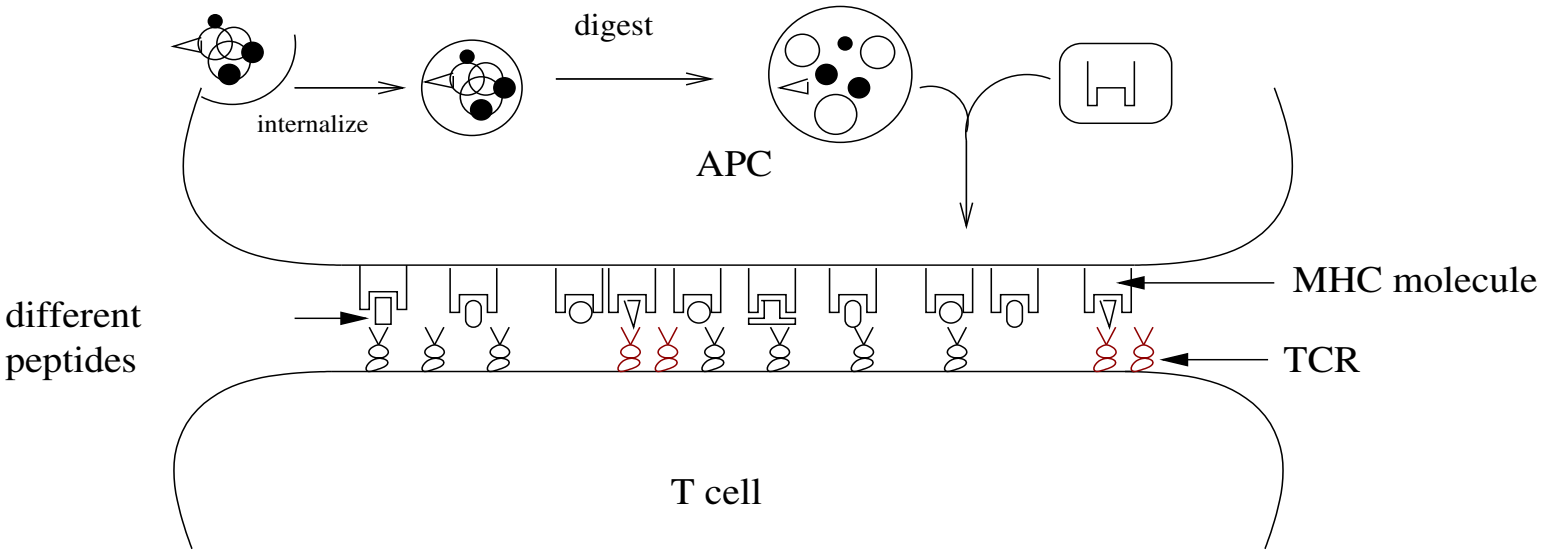
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



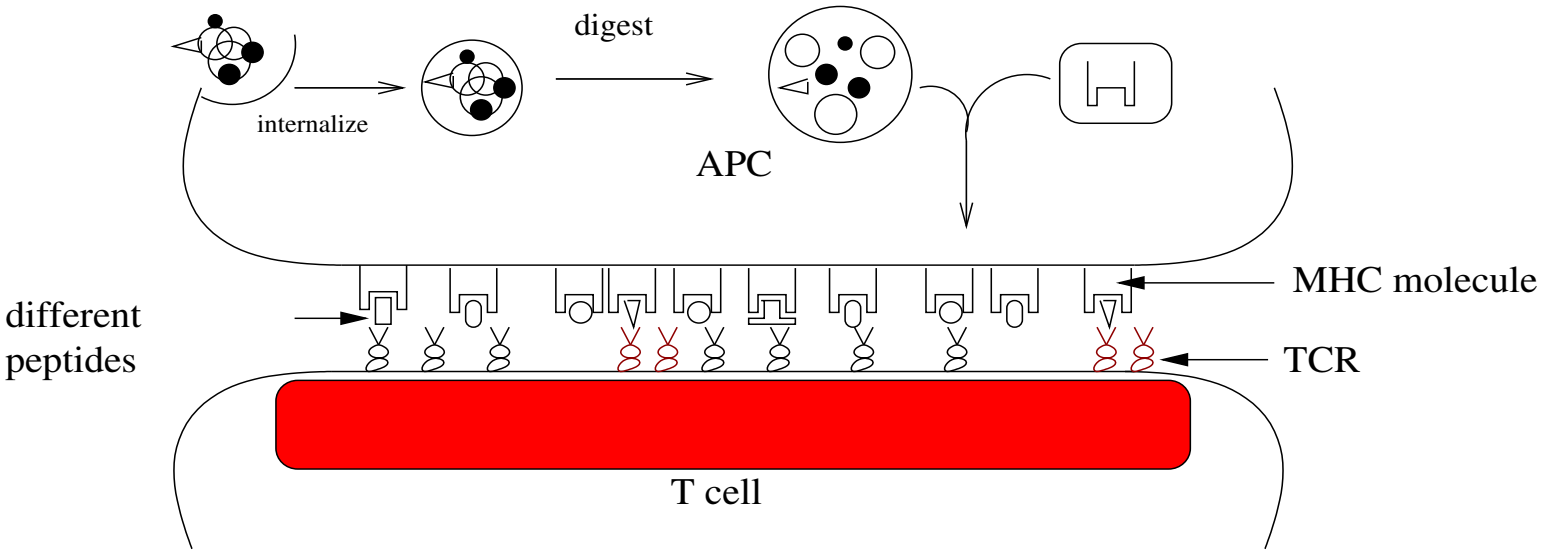
- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics



- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

caricature model of T cell-Dynamics

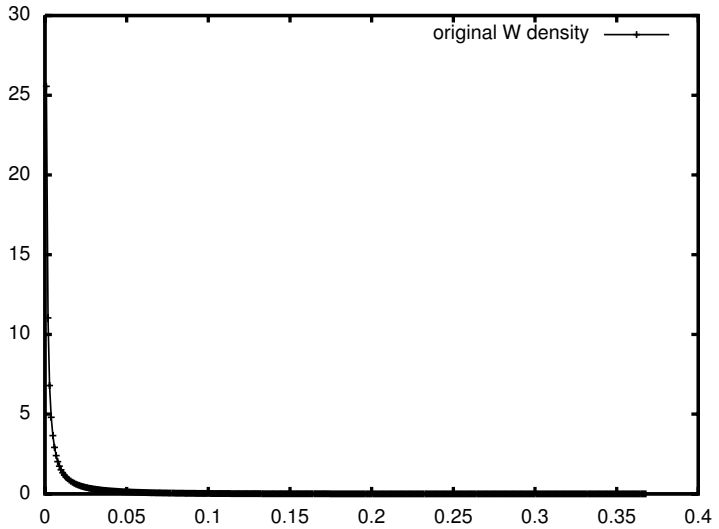


- APC digests and presents antigens → random sample from huge pool of different antigens
- T cells from the enviroment bind → random T cell with random TCR
- TCR is triggered after certain binding time → mean binding time describes antigen-TCR binding
- T cell is activated if certain number of TCRs is triggered

- *Stochastic approach* developed by van den Berg, Rand and Burroughs in 2001 [Van Den Berg et al., 2001]
- meeting of TCR with random antigen j: described by mean binding time $\rightarrow \mathcal{T}_j \sim Expo(\lambda)$ (iid)
- activation rate of antigen with binding time τ

$$h(\tau) = \frac{1}{\tau} \exp(-\frac{1}{\tau})$$

- activation rate of an antigen as a random variable $W = h(\mathcal{T})$



- activation rates of antigens are summed up

$$G(z^{(f)}) = \sum_{j=1}^{m^{(v)}} z^{(v)} h(\mathcal{T}_j) + \sum_{j=m^{(v)}+1}^{m^{(v)}+m^{(c)}} z^{(c)} h(\mathcal{T}_j) + z^{(f)} h(\mathcal{T}_{m^{(v)}+m^{(c)}+1})$$

3 classes of antigen types: constitutive, variable, foreign with different copy numbers

$$z^\star = \frac{M - z^{(f)}}{M} z_\star, \qquad \star \in \{v, c\}, \quad M, z_\star, \lambda \text{ parameters}$$

and type numbers $m^{(c)}, m^{(v)}$ ($m^{(v)} \gg m^{(c)}$, $z^{(f)} > z^{(c)} \gg z^{(v)}$)

- activation rates of antigens are summed up

$$G(z^{(f)}) = \sum_{j=1}^{m^{(v)}} z^{(v)} h(\mathcal{T}_j) + \sum_{j=m^{(v)}+1}^{m^{(v)}+m^{(c)}} z^{(c)} h(\mathcal{T}_j) + z^{(f)} h(\mathcal{T}_{m^{(v)}+m^{(c)}+1})$$

3 classes of antigen types: constitutive, variable, foreign with different copy numbers

$$z^\star = \frac{M - z^{(f)}}{M} z_\star, \qquad \star \in \{v, c\}, \quad M, z_\star, \lambda \text{ parameters}$$

and type numbers $m^{(c)}, m^{(v)}$ ($m^{(v)} \gg m^{(c)}$, $z^{(f)} > z^{(c)} \gg z^{(v)}$)

- activation rates of antigens are summed up

$$G(z^{(f)}) = \sum_{j=1}^{m^{(v)}} z^{(v)} h(\mathcal{T}_j) + \sum_{j=m^{(v)}+1}^{m^{(v)}+m^{(c)}} z^{(c)} h(\mathcal{T}_j) + z^{(f)} h(\mathcal{T}_{m^{(v)}+m^{(c)}+1})$$

3 classes of antigen types: constitutive, variable, foreign with different copy numbers

$$z^\star = \frac{M - z^{(f)}}{M} z_\star, \qquad \star \in \{v, c\}, \quad M, z_\star, \lambda \text{ parameters}$$

and type numbers $m^{(c)}, m^{(v)}$ ($m^{(v)} \gg m^{(c)}$, $z^{(f)} > z^{(c)} \gg z^{(v)}$)

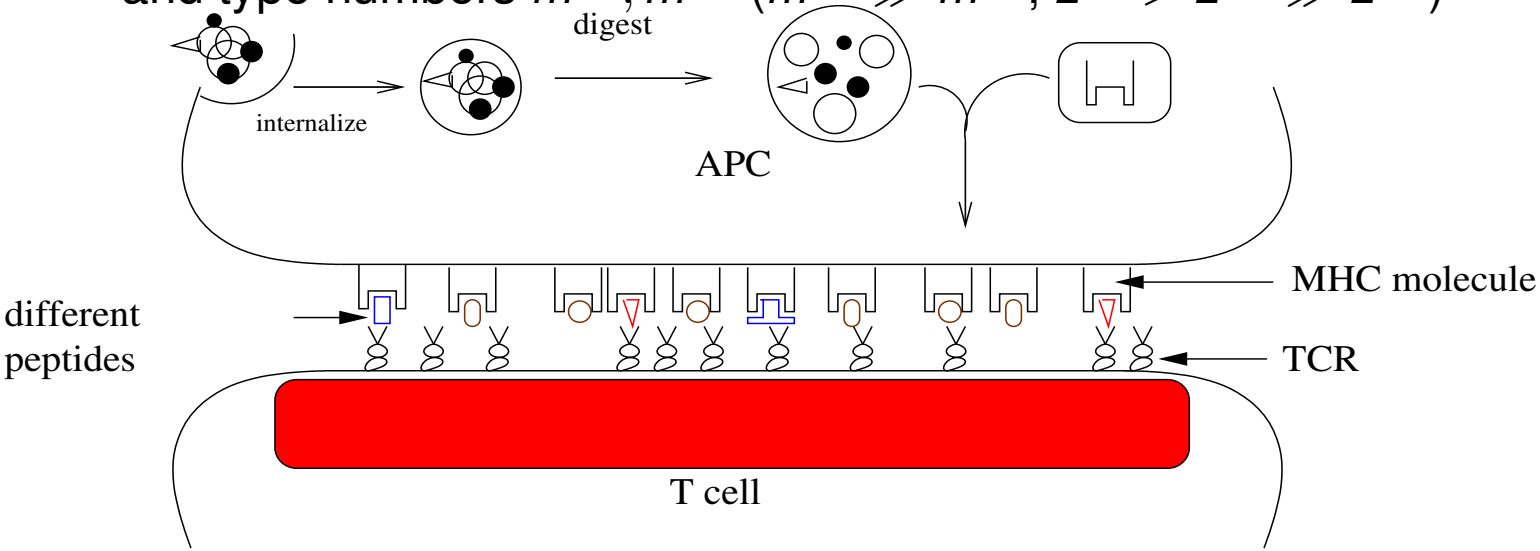
- activation rates of antigens are summed up

$$G(z^{(f)}) = \sum_{j=1}^{m^{(v)}} z^{(v)} h(\mathcal{T}_j) + \sum_{j=m^{(v)}+1}^{m^{(v)}+m^{(c)}} z^{(c)} h(\mathcal{T}_j) + z^{(f)} h(\mathcal{T}_{m^{(v)}+m^{(c)}+1})$$

3 classes of antigen types: constitutive, variable, foreign with different copy numbers

$$z^\star = \frac{M - z^{(f)}}{M} z_\star, \qquad \star \in \{v, c\}, \quad M, z_\star, \lambda \text{ parameters}$$

and type numbers $m^{(c)}, m^{(v)}$ ($m^{(v)} \gg m^{(c)}$, $z^{(f)} > z^{(c)} \gg z^{(v)}$)



Wanted: $\mathbb{P}(G(z^{(f)}) \geq g_{act})$, the probability to reach the activation threshold g_{act}.

- for a good foreign-self discrimination $\mathbb{P}(G(z^{(f)}) \geq g_{act}) \gg \mathbb{P}(G(0) > g_{act})$ should hold
- no apriori-distinction between foreign and self : W_j iid $\rightarrow \mathbb{E}[G(0)] = \mathbb{E}[G(z^{(f)})]$
- activation threshold is high $\rightarrow g_{act} \gg \mathbb{E}[G(z^{(f)})]$
- $\Rightarrow$ T cell activation is a rare event

Methods of analysis

- no direct analytic approach available
- asymptotic estimation with large deviation theory [Zint et al., 2008]
- validation via simple sampling hardly possible
- need of a better sampling method

asymptotically efficient rare event simulation

Dieker and Mandjes [2005], Bucklew [2004]

- $\{P_n\}$ sequence of probability measures with LDP, and rare function I
- take $\{P_n^\vartheta\}$ as IS distribution
- P_n^ϑ tilted measure:

$$\frac{dP_n^\vartheta}{dP_n}(x) = \frac{e^{n\vartheta x}}{\mathbb{E}(e^{n\vartheta X})} \tag{1}$$

- $\{P_n^\vartheta\}$ asymptotically efficient for rare event A if

$$\inf_{x\in\mathbb{R}}[I(x) - \vartheta x] + \inf_{x\in\overline{A}}[I(x) + \vartheta x] = 2 \inf_{x\in A^\circ} I(x) \tag{2}$$

+ technical conditions

asymptotically efficient rare event simulation

Dieker and Mandjes [2005], Bucklew [2004]

- $\{P_n\}$ sequence of probability measures with LDP, and rare function I
- take $\{P_n^\vartheta\}$ as IS distribution
- P_n^ϑ tilted measure:

$\lim_{n\to\infty} \frac{1}{n} \log N^\star_{P_n^\vartheta} = 0$

$$\frac{dP_n^\vartheta}{dP_n}(x) = \frac{e^{n\vartheta x}}{\mathbb{E}(e^{n\vartheta X})} \tag{1}$$

- $\{P_n^\vartheta\}$ asymptotically efficient * for rare event A if

$$\inf_{x\in\mathbb{R}} [I(x) - \vartheta x] + \inf_{x\in\overline{A}} [I(x) + \vartheta x] = 2 \inf_{x\in A^\circ} I(x) \tag{2}$$

+ technical conditions

asymptotically efficient rare event simulation

Dieker and Mandjes [2005], Bucklew [2004]

- $\{P_n\}$ sequence of probability measures with LDP, and rare function I
- take $\{P_n^\vartheta\}$ as IS distribution
- P_n^ϑ tilted measure:

$$\lim_{n\to\infty} \frac{1}{n} \log N^\star_{P_n^\vartheta} = 0$$

$$\eta = \sqrt{\frac{\text{variance of estimate}}{\text{estimate}}}$$

$$\frac{dP_n^\vartheta}{dP_n}(x) = \frac{e^{n\vartheta x}}{\mathbb{E}(e^{n\vartheta X})} \tag{1}$$

- $\{P_n^\vartheta\}$ asymptotically efficient * for rare event A if

$$\inf_{x\in\mathbb{R}} [I(x) - \vartheta x] + \inf_{x\in\overline{A}} [I(x) + \vartheta x] = 2 \inf_{x\in A^\circ} I(x) \tag{2}$$

+ technical conditions

our case:

- conditions fulfilled, with suitable embedding of $\mathbb{P}(G^{(f)} \geq g_{\text{act}})$
- and ϑ chosen so that $\mathbb{E}^{\vartheta}(G^{\vartheta}(z^{(f)})) = g_{\text{act}}$
- tilting of G with ϑ achieved by tilting each W_j with ϑ individually (independence!)

problem:

- no explicit density/distribution of W
- tilting W numerically unstable (poles)

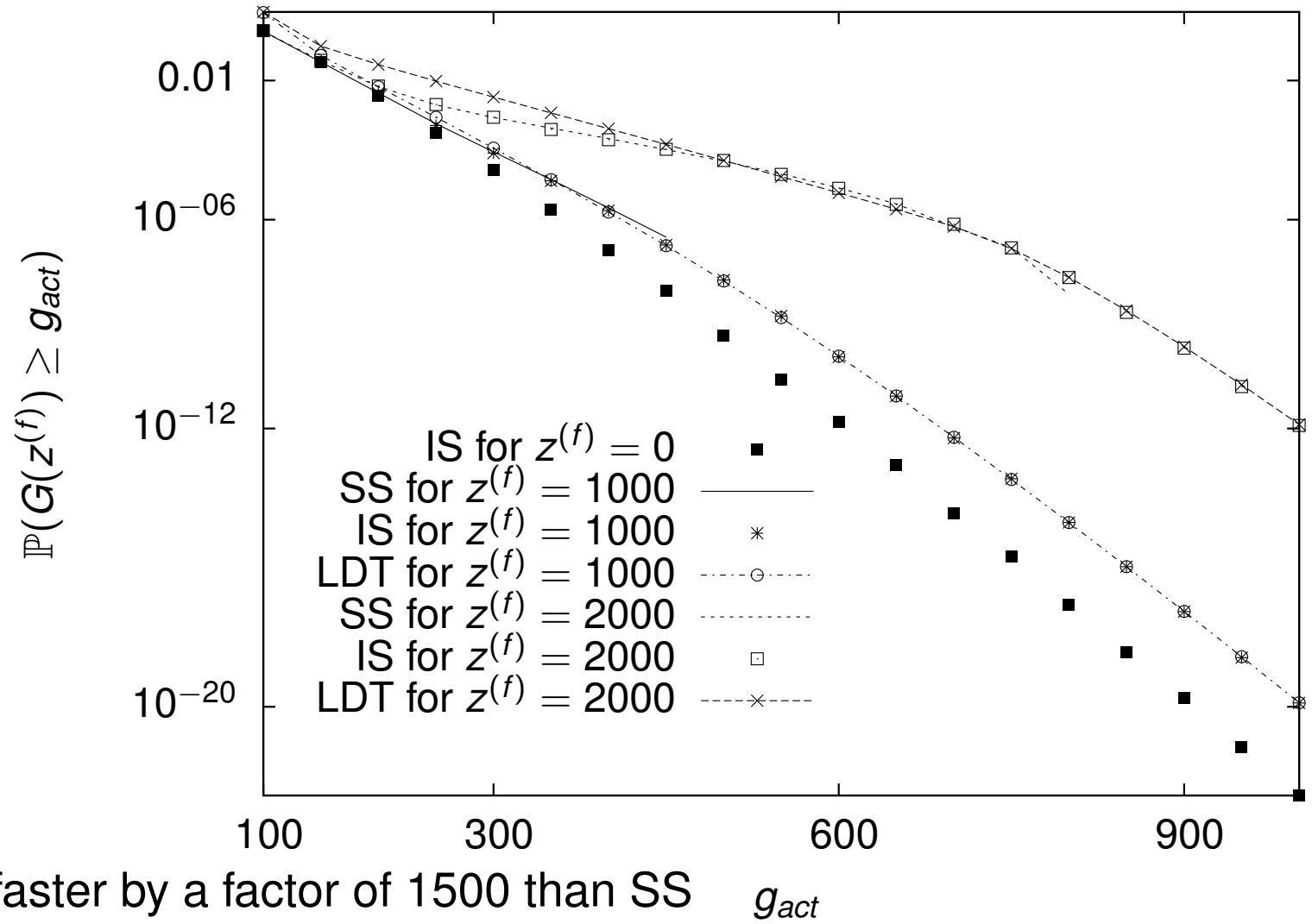
solution:

- tilting at level of $\mathcal{T}$ (with exponential density, f) via

$$\tilde{f}^{\vartheta}(\tau) = \frac{\exp(\vartheta h(\tau)) f(\tau)}{\mathbb{E}(\exp(\vartheta h(\mathcal{T}))} \tag{3}$$

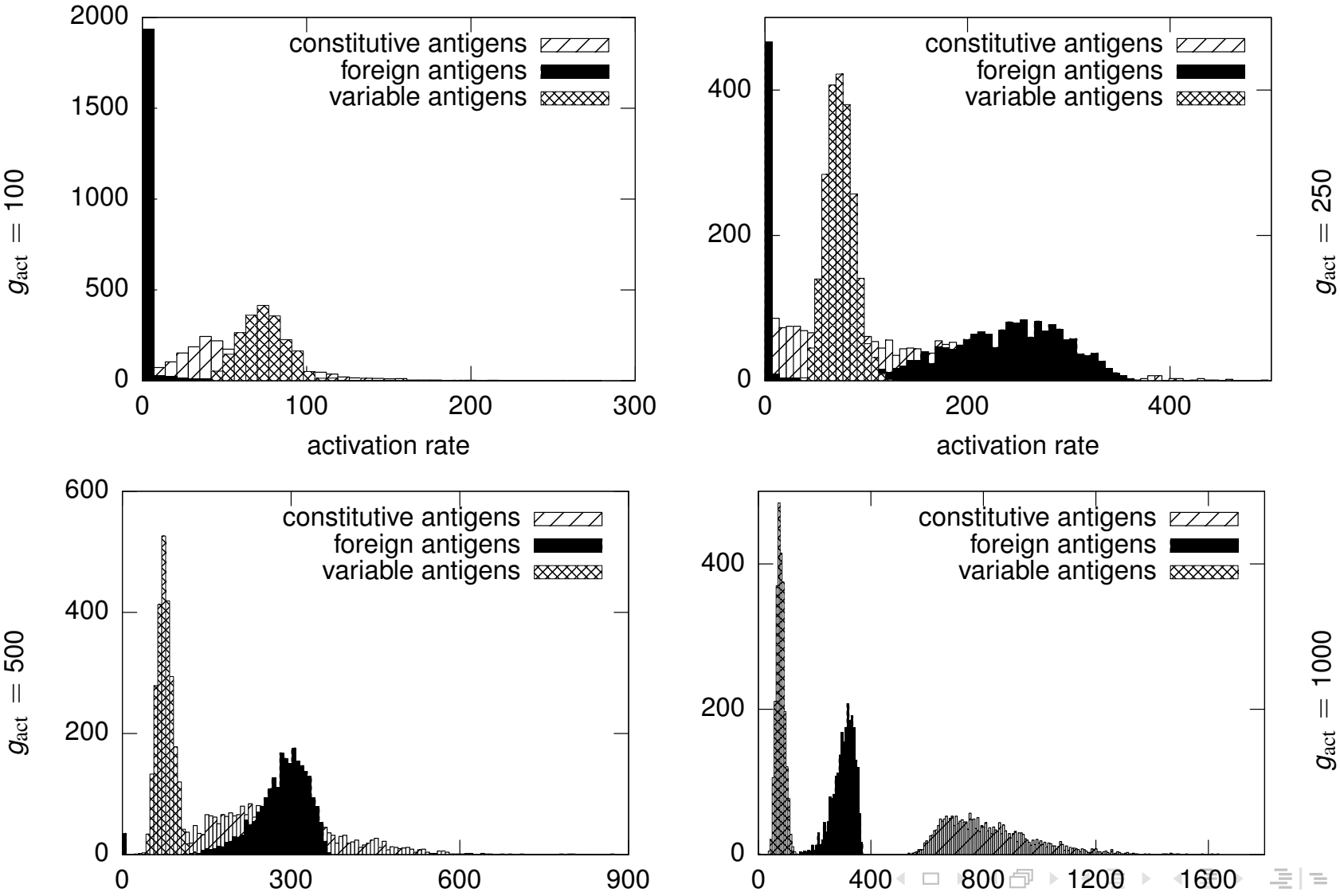
- tilted W: $W^{\vartheta} = h(\tilde{\mathcal{T}}^{\vartheta})$
- search table via alias method

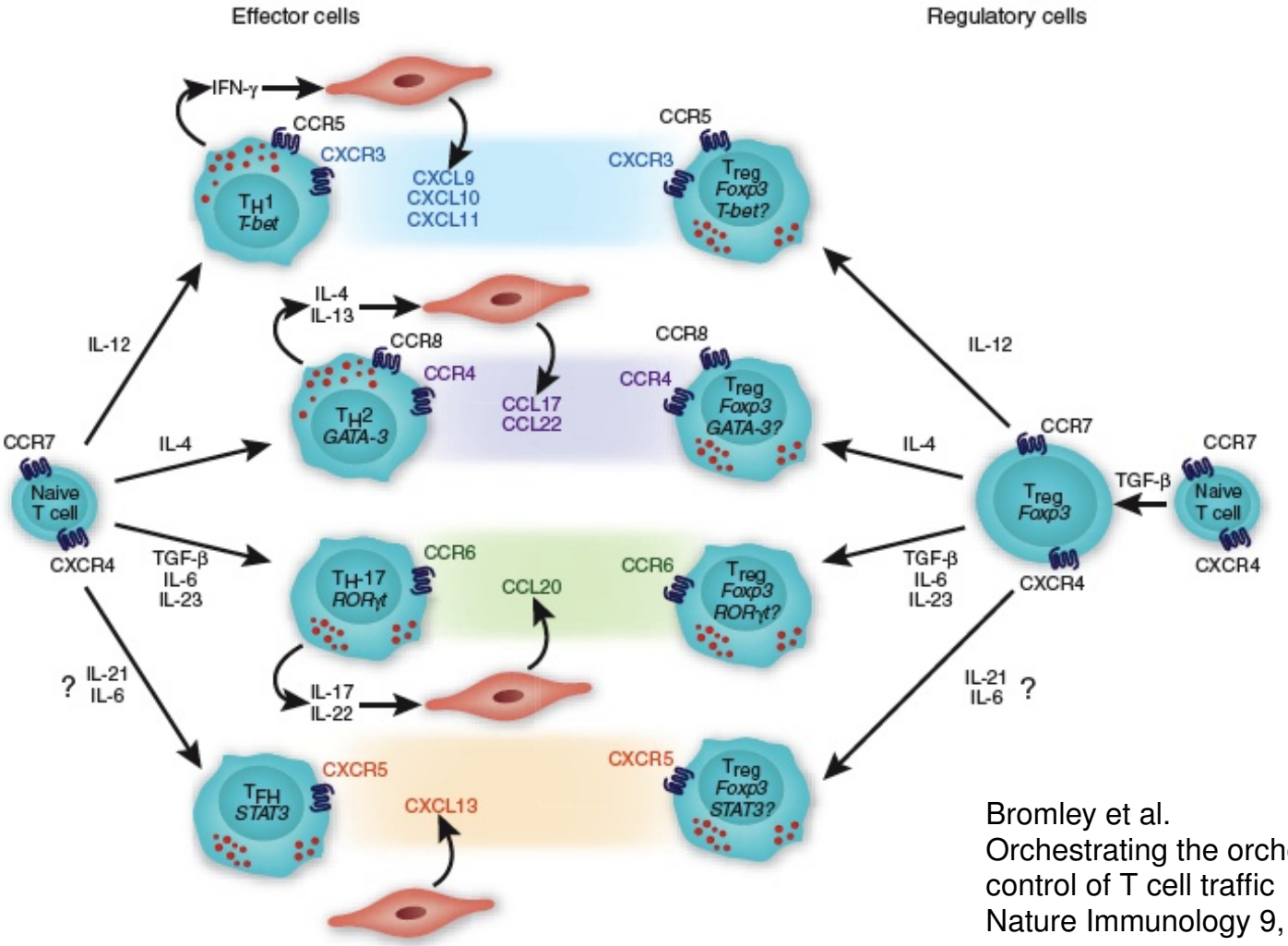
Activation curves



IS faster by a factor of 1500 than SS

activation rate histograms





Bromley et al.
Orchestrating the orchestrators: chemokines in control of T cell traffic
Nature Immunology 9, 970 - 980 (2008)

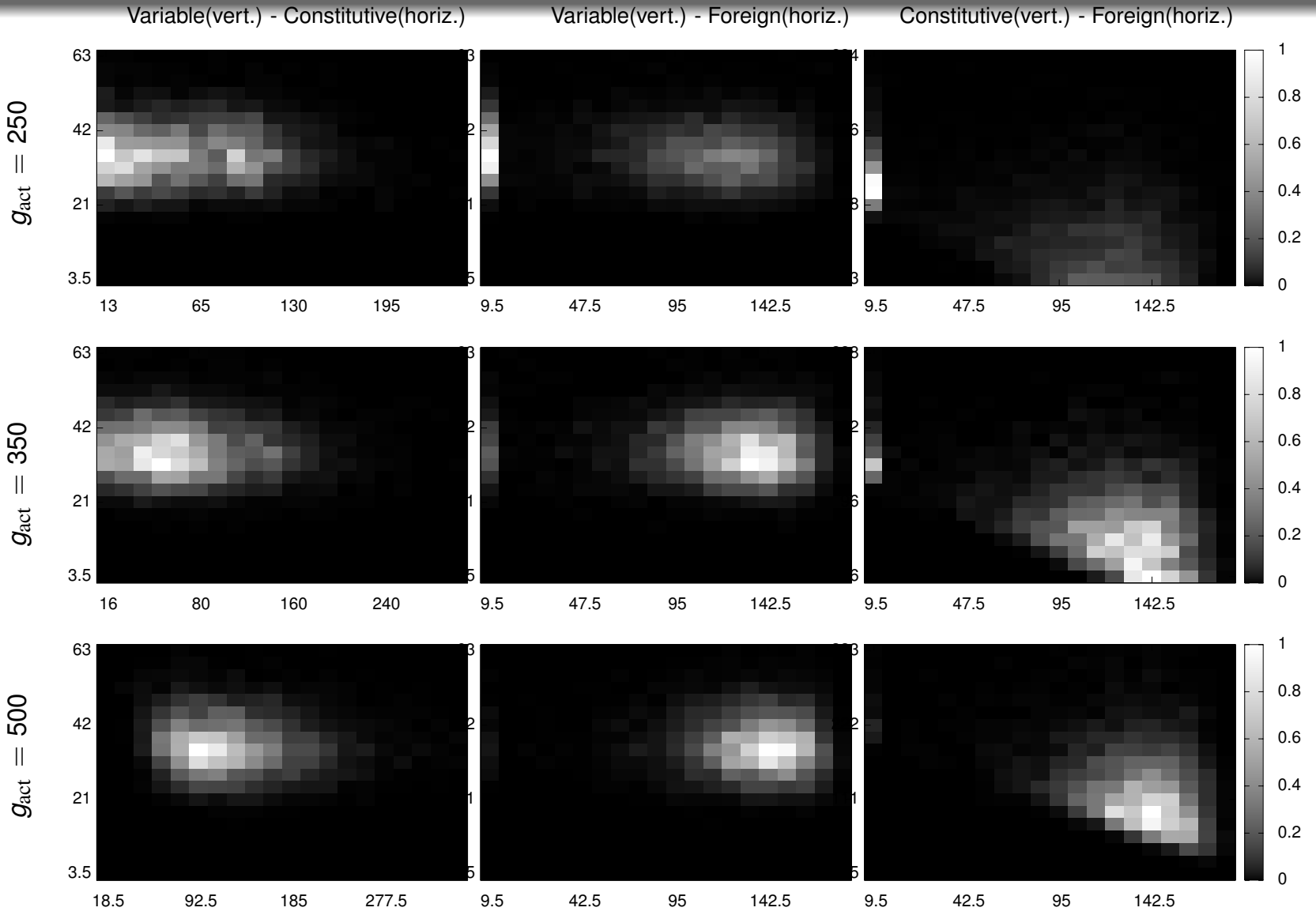
James A. Bucklew. *Introduction to Rare Event Simulation*. Springer, 2004.

A.B. Dieker and M. Mandjes. On asymptotically efficient simulation of large deviation probabilities. *Advanced applied probability*, 37:539–552, 2005.

H A Van Den Berg, D A Rand, and N J Burroughs. A reliable and safe t cell repertoire based on low-affinity t cell receptors. *J Theor Biol*, 209(4):465–486, 2001. ISSN 0022-5193 (Print). doi: 10.1006/jtbi.2001.2281.

Natali Zint, Ellen Baake, and Frank den Hollander. How t-cells use large deviations to recognize foreign antigens. *J. Math. Biol., in press*, 2008.

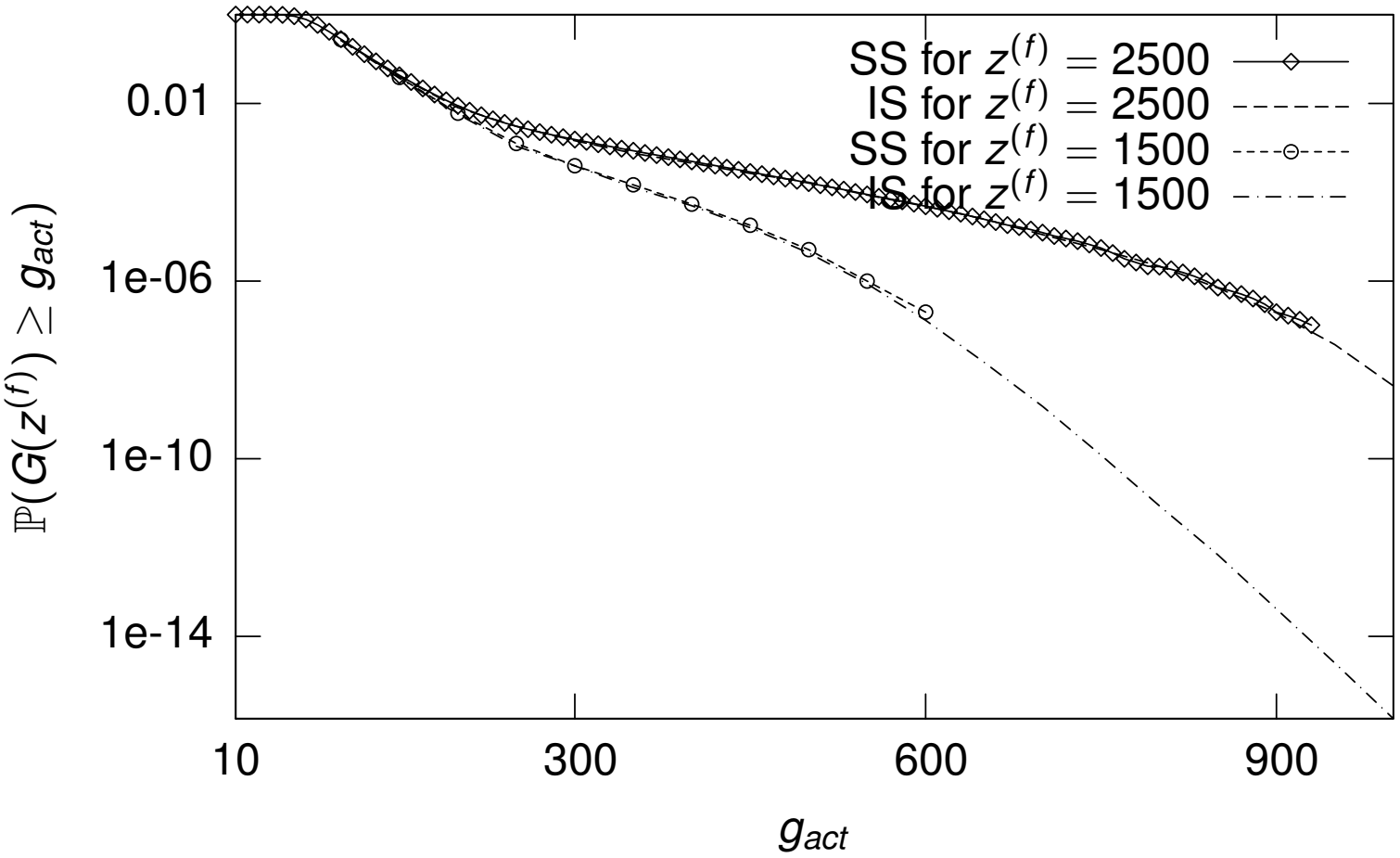
correlation analysis



Extension 1

$$G(z^{(f)}) := \left(\sum_{j=1}^{m^{(c)}} q Z_j^{(c)} W_j \right) + \left(\sum_{j=m^{(c)}+1}^{m^{(c)}+m^{(v)}} q Z_j^{(v)} W_j \right) + z^{(f)} W_{m^{(c)}+m^{(v)}+1}. \tag{4}$$

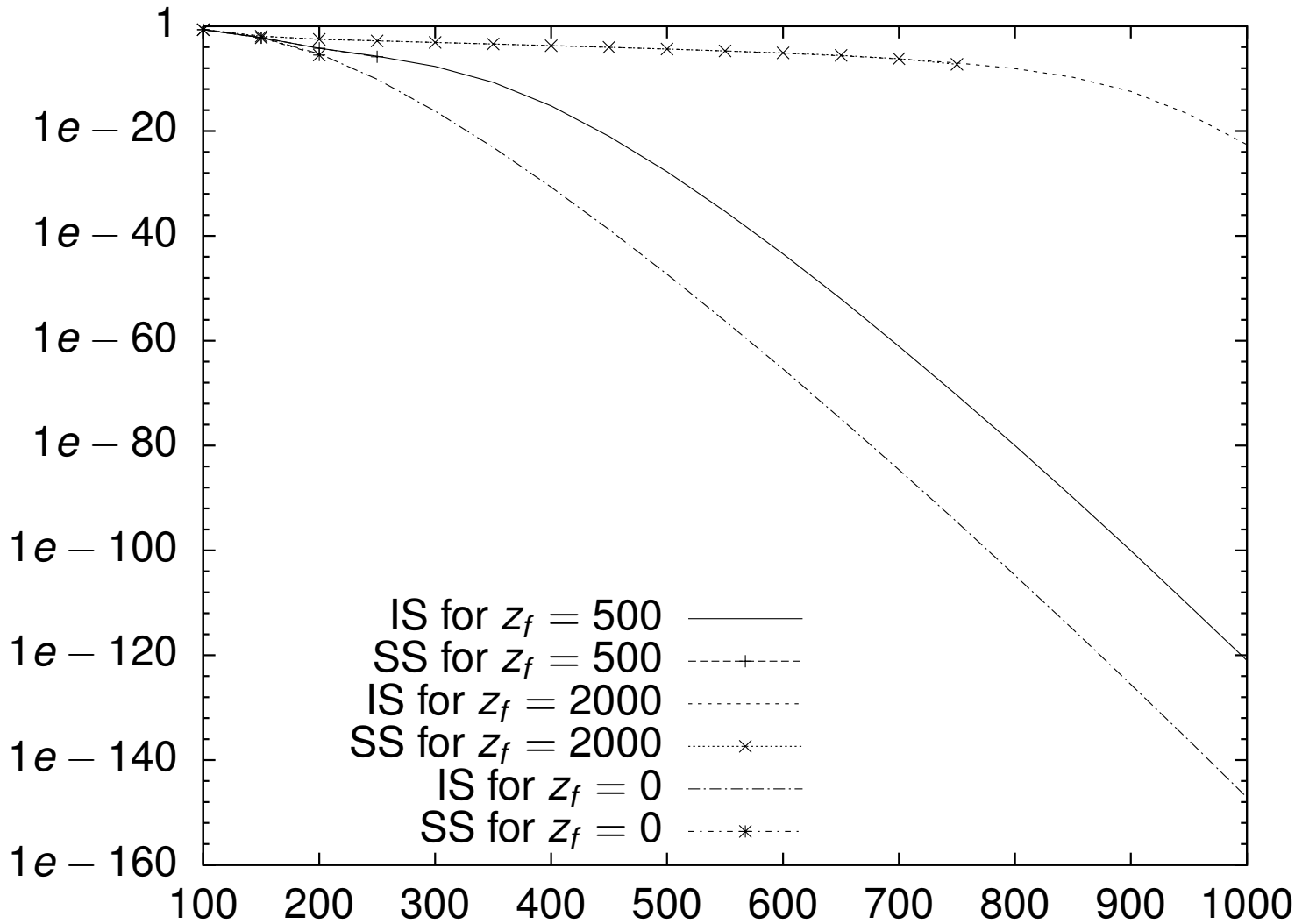
$Z_j^{(c)}, Z_j^{(v)}$ distributed according to binomial distributions with $\mathbb{E}(Z_j^{(c)}) = z^{(c)}$, $\mathbb{E}(Z_j^{(v)}) = z^{(v)}$ (so the expected number of antigens per APC is still M)

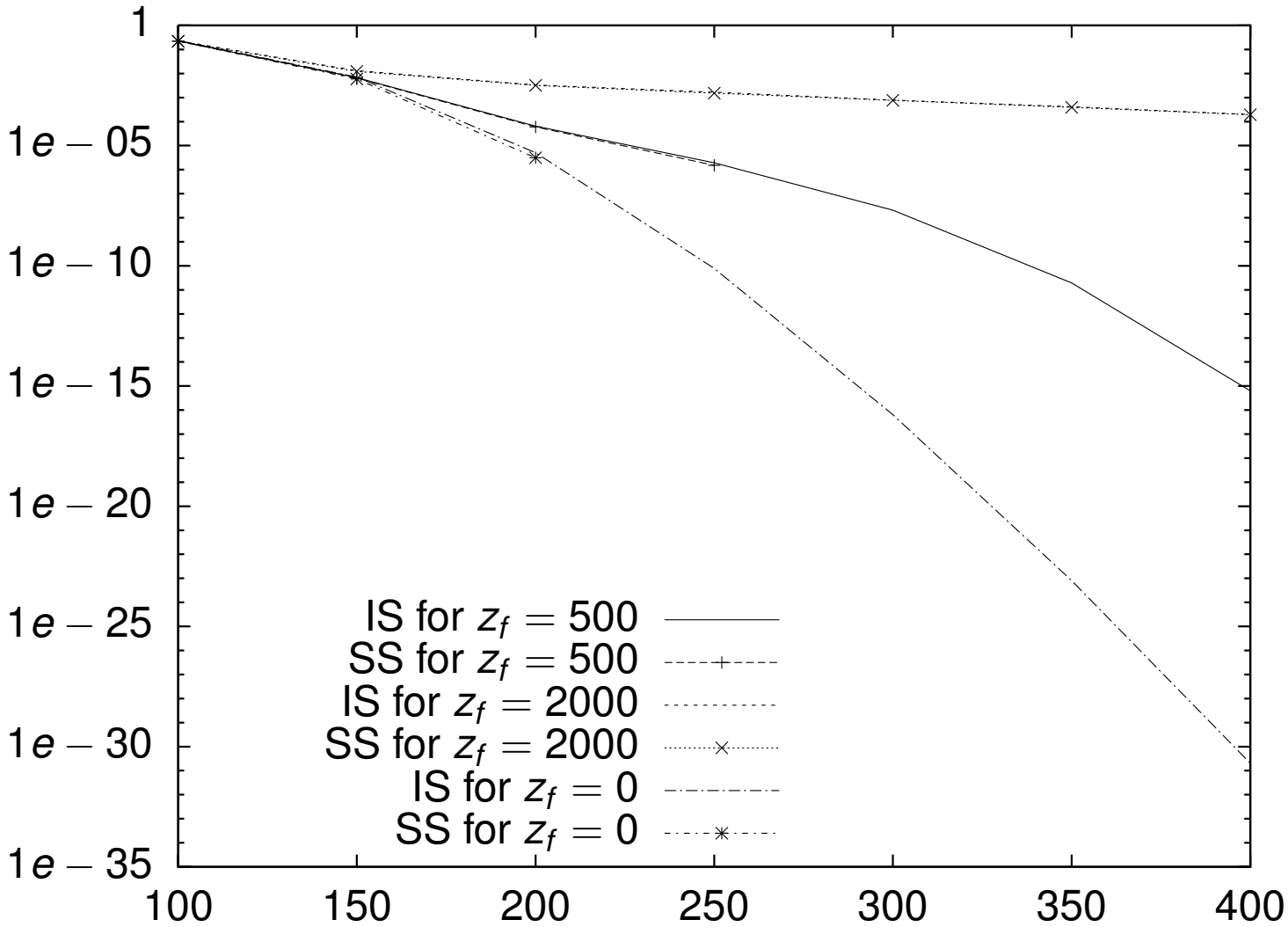

SS for $z^{(f)} = 2500$
IS for $z^{(f)} = 2500$
SS for $z^{(f)} = 1500$
IS for $z^{(f)} = 1500$
0.01
1e-06
1e-10
1e-14
$\mathbb{P}(G(z^{(f)}) \geq g_{act})$
10
300
600
900
g_{act}

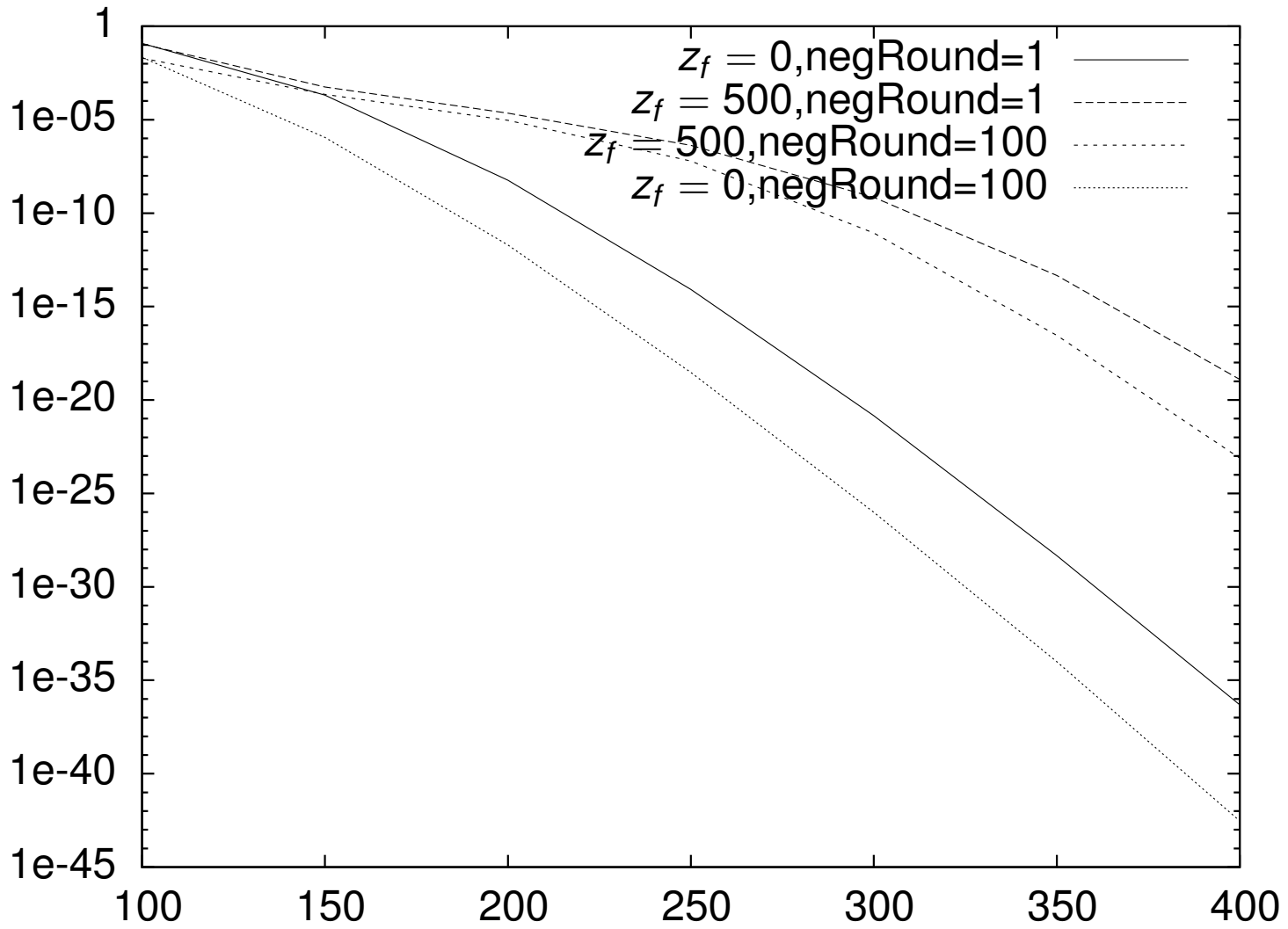
Negative selection

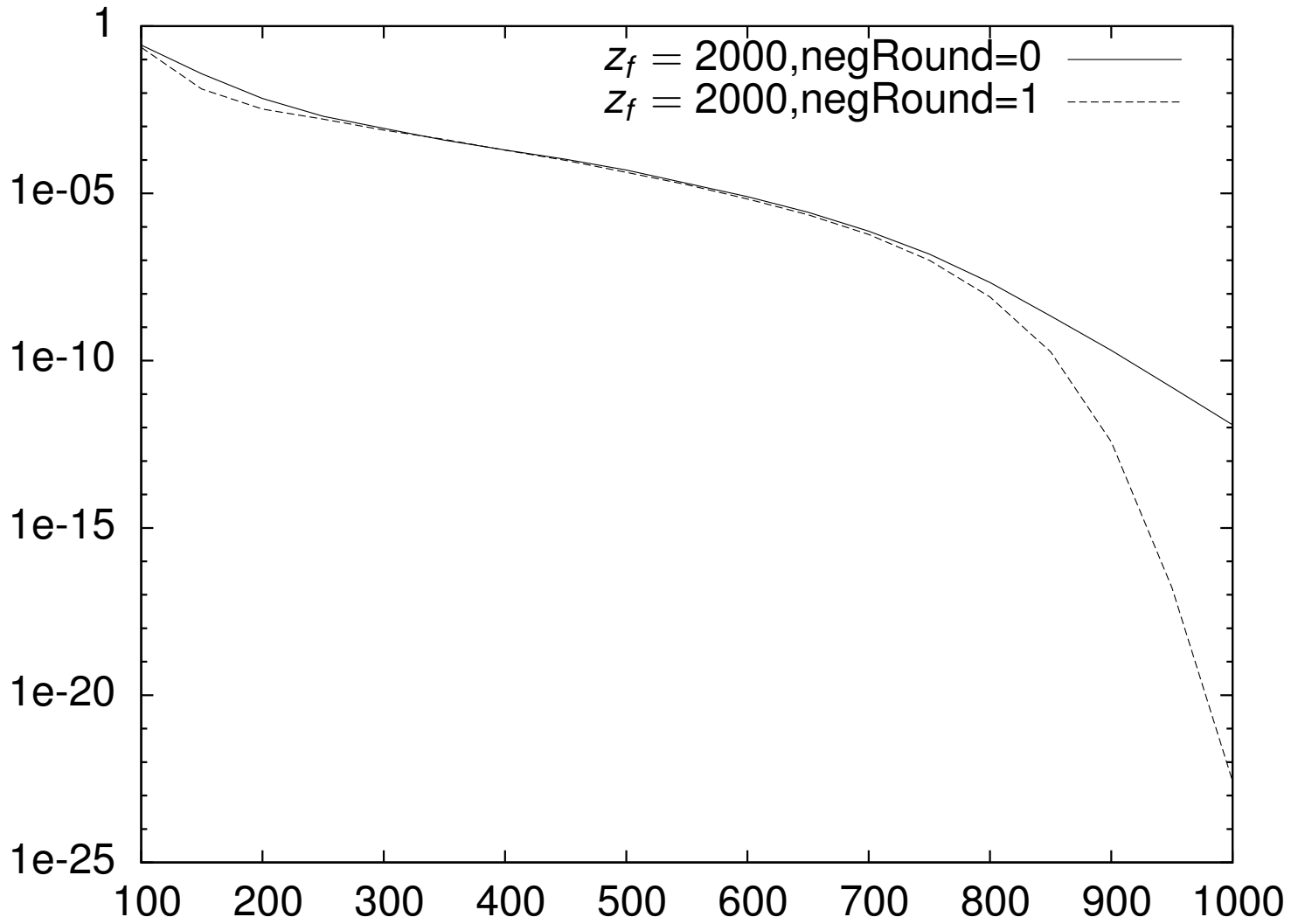
Idea:

$$\mathbb{P}(\text{survival of a T cell of type i}) = \mathbb{P}\left(\sum_{j=1}^{m^{(v)}} z^{(v)} h(\mathcal{I}_j) + \sum_{j=m^{(v)}+1}^{m^{(v)}+m^{(c)}} z^{(c)} h(\mathcal{I}_j) < g_{\text{thy}}\right) \tag{5}$$


1
1e − 20
1e − 40
1e − 60
1e − 80
1e − 100
1e − 120
1e − 140
1e − 160
100
200
300
400
500
600
700
800
900
1000
IS for $z_f = 500$
SS for $z_f = 500$
IS for $z_f = 2000$
SS for $z_f = 2000$
IS for $z_f = 0$
SS for $z_f = 0$


1
1e − 05
1e − 10
1e − 15
1e − 20
1e − 25
1e − 30
1e − 35
100
150
200
250
300
350
400
IS for $z_f = 500$
SS for $z_f = 500$
IS for $z_f = 2000$
SS for $z_f = 2000$
IS for $z_f = 0$
SS for $z_f = 0$


1
1e-05
1e-10
1e-15
1e-20
1e-25
1e-30
1e-35
1e-40
1e-45
100
150
200
250
300
350
400
$z_f = 0$,negRound=1
$z_f = 500$,negRound=1
$z_f = 500$,negRound=100
$z_f = 0$,negRound=100


1
1e-05
1e-10
1e-15
1e-20
1e-25
100
200
300
400
500
600
700
800
900
1000
$z_f = 2000$,negRound=0
$z_f = 2000$,negRound=1