Transient Kinetic Proofreading

G. Iyengar[1] and M. Perry[2]

[1]IEOR Department, Columbia University
[2]IEOR Department, Columbia University

This version: March 10, 2021

Abstract

We propose a 2-dimensional Markov chain model to understand and efficiently compute the transient behavior of the kinetic proofreading mechanism in a single T-cell. We show that a stochastic version of absolute ligand discrimination is a direct consequence of the finite number of receptors on the cell surface; thus, pointing to number control as being important for absolute ligand discrimination. We also develop 1-dimensional approximations for several limiting regimes that significantly decrease the computational time. We present results of numerical experiments that explore the behavior of the new model for a wide range of parameters, and its robustness to parameter errors.

1 Introduction

Kinetic proofreading is a mechanism proposed to explain the extremely low probability of error observed in a number of biological processes, e.g. microtubule growth, DNA replication, and the activation of T-cells by antigen presenting cells [7, 15]. Kinetic proofreading exploits the fact that cycling through intermediate complexes on the path to a final product can ensure that the path to undesired products have a significantly lower probability than the path to the desired final product.

T-cells are a component of the adaptive arm of the immune system [16]. In contrast to the innate immune system, which attacks a broad range of pathogens based on common cell-surface markers, each individual cell in the adaptive immune system is specially tailored to sense and react to antigens associated with a specific type of pathogen. Among other things, a T-cell is responsible for sensing and killing pathogens expressing the T-cell's target antigen on their cell surface. T-cells use small differences in the binding energy between antigens and T-cell receptors (TCRs) to activate in the presence of a specific "non-self" antigen, while at the same time refraining from activating in response to "self" antigens. Kinetic proofreading allows T-cells to discriminate with much better accuracy than that implied by the difference in binding energy between self and non-self antigens.

McKeithan [14] proposed the first kinetic proofreading-based model for TCR activation to explain the immune system's ability to accurately discriminate self and non-self antigens. This model is described by the following set of ODEs.

$$\begin{aligned}
\frac{dC_0}{dt} &= -\left(k_u + k_p\right) C_0(t) + k_b A R \\
\frac{dC_n}{dt} &= -\left(k_u + k_p\right) C_n(t) + k_p C_{n-1}(t), \;\; n = 1, 2, \ldots, N-1 \\
\frac{dC_N}{dt} &= -k_u C_N + k_p C_{N-1}
\end{aligned}$$

where A denotes the density of the antigen interacting with the TCR, R denotes the density of the TCR, $C_0(t)$ is the density of a TCR-antigen complex at time t, $C_n(t)$ is the density of a TCR-antigen complex with n phosphorylations at time t, k_b is the rate at which antigens bind to TCRs and is assumed to be the same for all antigens, k_u is the rate at which antigens unbind from TCRs and varies between different antigens,

and k_p is the rate at which TCR-antigen bindings are phosphorylated. This model assumes that a TCR-antigen complex with N phosphorylations is required for T-cell activation. The steady state concentration of the TCR-antigen complex is $C_N \propto L/k_u^N$. Suppose a self antigen with affinity $\tau_u^{(s)} = 1/k_u^{(s)}$ and a non-self antigen with affinity $\tau_u^{(o)} = 1/k_u^{(o)}$ are present in the same concentration R. Then the ratio of the steady-state concentration $C_N^{(s)}$ of the self-antigen-TCR complex and the steady-state concentration $C_N^{(n)}$ of the non-self-antigen-TCR complex is $C_N^{(s)}/C_N^{(n)} = (\tau_u^{(s)}/\tau_u^{(o)})^N$. By repeatedly dissociating the receptor-antigen bindings en route to a final product, the proofreading network ensures that the rate of error decreases exponentially with the size of the network.

The analysis above assumes that the concentration of self and non-self antigens are roughly equal. In reality, this is almost never the case. There is evidence of densities as low as ten non-self antigens per cell leading to T-cell activation [9], whereas self antigens are far more abundant. More specifically, the adaptive immune system is able to respond to non-self antigen at very low concentrations (referred to as *sensitivity*) while refraining from activating even at very high concentrations of self antigen (referred to as *specificity*), and it does so even for non-self antigens with unbinding rates that are very close to the unbinding rates of self antigens (referred to as *fidelity*). The ability of the adaptive immune system to distinguish between self and non-self antigens despite very large discrepancies in antigen density is known as *absolute ligand discrimination* [5].

Altan-Bonnet and Germain [2], Lalanne and Francoise [12] and Francois and Altan-Bonnet [5] proposed and analyzed deterministic mean-field models for absolute ligand discrimination that were based on a feedback loop that regulates the phosphorylation rate of antigen-bound TCRs via a kinase whose activity is itself regulated by the quantity of phosphorylated antigen-bound TCRs. This feedback loop, called *adaptive sorting*, decouples the behavior of the proofreading mechanism from antigen density, thereby achieving absolute ligand discrimination.

We propose a stochastic single-cell model for TCR-activation. Our main contributions are as follows:

(a) We propose a 2-dimensional Markov chain model for T-cell activation where the first dimension models the interaction between the TCR and the antigen, and the second dimension models the interaction of TCR-antigen pairs with a second downstream molecule that we refer to as ZAP[1]. The T-cell activates when the number of antigen-TCR-ZAP complexes is greater than a threshold. Thus, in our model activation is a stochastic phenomenon at the single-cell level. We show that the probability of activation with a specified delay can be computed as easily as computing the mean time to activation. We are therefore able to quantify the impact of the probability threshold α, or equivalently, the specificity of the proofreading network.

(b) In our model, absolute ligand discrimination is a consequence of the fact that the number of TCRs on the cell surface is finite. Thus, we show a connection between absolute ligand discrimination and controlling of the number of receptors on the cell surface [19]. This opens up the possibility that the properties of a single T-cell can be modulated by controlling the number of receptors.

(c) We show that the 2-dimensional Markov chain can be approximated by an appropriately defined 1-dimensional chain. The 1-dimensional approximation not only helps with speeding up the computations several fold but also helps identify the bottleneck dynamics that lead to absolute ligand discrimination. We employ the approximations to analyze the trade-offs between sensitivity, specificity, fidelity and delay in kinetic proofreading. Additionally, we also investigate the robustness of the model to biochemical errors.

The rest of this paper is organized as follows. In Section 2 we introduce the 2-dimensional Markov chain model for T-cell activation, discuss the behavior of this model in two asymptotic limits, and propose the 1-dimensional approximation that is used in subsequent numerical experiments. In Section 3, we discuss the results of the numerical experiments investigating the trade-off between sensitivity, specificity, fidelity and delay. Lastly, in Section 4 we discuss possible extensions of the model.

[1]The model can be extended to include more downstream molecules without changing the nature of the analysis or computational strategy.

2 Markov chain model for T-cell activation

Consider a single T-cell exposed to a single type of antigen with a uniform density L on the cell surface. The cell has N receptors spread uniformly across its surface that bind to the antigen. An antigen-bound receptor is able to recruit one or more other molecules needed to activate downstream reactions and activate the cell. We assume that a certain number of TCRs need to recruit *one* more molecule, and refer to the second molecule as ZAP. We note here that the details of the second molecule, i.e. whether it is phosphorus, Lck, ZAP, etc., are not relevant to our model. As will be clear from the discussion below, one could add more downstream molecules to the model without changing the analysis or the computation.

Following [8, 20], we model the response of a T-cell to an antigen by a two-dimensional continuous-time Markov chain, where the first dimension refers to the number of antigen-bound TCRs and the second dimensions to the number of antigen-TCR-ZAP complexes. The state and events in the model are as follows:

(1) A single T-cell has N TCRs available on its surface. The assumption that N is finite is a significant departure from [8, 20], and is essential for absolute ligand discrimination.

(2) An unbound TCR binds to an antigen with rate $k_b L$. We implicitly assume that the TCR only reacts to a single antigen, and are ignoring the effects of ligand antagonism [6].

(3) Antigen-bound TCRs recruit a single ZAP molecule at rate k_{on}.

(4) Antigens unbind from TCRs at rate $k_u = \tau_u^{-1}$. We assume that the antigen-TCR-ZAP complex loses the ZAP molecule if the antigen unbinds. In line with McKeithan [14], we assume that ZAPs do not unbind from antigen-TCR complexes unless the complex itself breaks apart. This assumption does not change the qualitative behavior of the model.

(5) The cell activates once $Z \geq 2$ TCRs have recruited ZAP molecules [2].

The state of a T-cell in our model is described by two variables: the number of antigen-TCR complexes $n \leq N$ and the number of antigen-TCR-ZAP complexes $z \leq Z$. We let (n_t, z_t) denote the state of the T-cell at any time $t \geq 0$. Since $n_t, z_t \geq 0$, $z_t \leq n_t$, and the cell activates when $z_t = Z$, it follows that the state

$$(n_t, z_t) \in \mathcal{S} = \{(n, z) : 0 \leq z \leq Z, 0 \leq n \leq N, z \leq n\},$$

and the dynamics of $\{(n_t, z_t) : t \geq 0\}$ is described by a two-dimensional continuous time Markov chain with an absorbing boundary at

$$\mathcal{B} = \{(n, z) \in \mathcal{S} : z = Z, Z \leq n \leq N\}, \tag{1}$$

and rate matrix Q defined as follows [18].

$$\begin{aligned}
Q_{(n,z),(n+1,z)} &= k_b L \mathbb{1}\left(0 \leq n < N, 0 \leq z < Z\right) \\
Q_{(n,z),(n-1,z)} &= (n-z) k_u \mathbb{1}\left(z < n \leq N, 0 \leq z < Z\right) \\
Q_{(n,z),(n,z+1)} &= (n-z) k_{on} \mathbb{1}\left(z < n \leq N, 0 \leq z < Z\right) \\
Q_{(n,z),(n-1,z-1)} &= z k_u \mathbb{1}\left(z \leq n \leq N, 0 < z < Z\right) \\
Q_{(n,z),(n,z)} &= -\left(Q_{(n,z),(n+1,z)} + Q_{(n,z),(n-1,z)} + Q_{(n,z),(n,z+1)} + Q_{(n,z),(n-1,z-1)}\right) \\
Q_{(n,z),(\eta,\zeta)} &= 0, \ (\eta, \zeta) \notin \{(n+1, z), (n-1, z), (n, z+1), (n-1, z-1), (n, z)\}
\end{aligned} \tag{2}$$

These rates follow from the assumptions above. For example, if the T-cell is in state (n, z) and it has not yet activated, i.e. $z < Z$, the antigen binds to an available TCR, i.e. if $n < N$, with rate $k_b L$; thus, $Q_{(n,z),(n+1,z)} = k_b L$ if $n < N$ and $z < Z$. All the other rates similarly follow from the assumptions above. Note that $Q_{(n,z),(n,z)}$ is set to ensure that $\sum_{(n',z') \in \mathcal{S}} Q_{(n,z),(n',z')} = 0$. See Figure 1 for a schematic of the Markov chain transitions. In our numerical simulations, we use $k_{on} = \frac{1}{51} s^{-1}$ and $k_b = \frac{1}{2250} s^{-1} \mu m^2$ [8, 20].

[2] For $Z = 1$, the dynamics for $((1+\epsilon)L, \tau_u/(1+\epsilon))$ are the same as that for (L, τ_u). Therefore, absolute ligand discrimination is not possible.

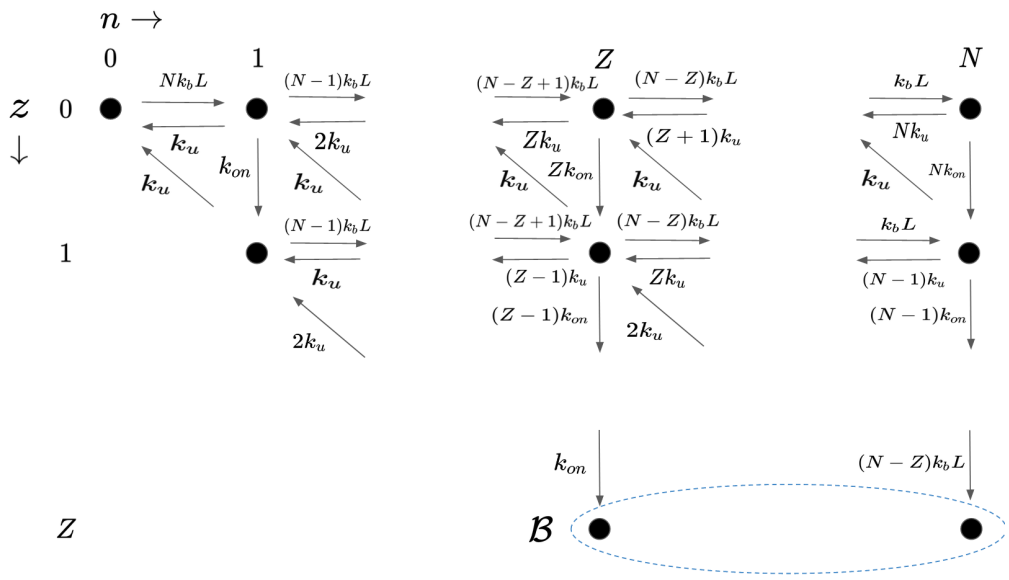

Figure 1: State space and transition rates for model of T-cell activation. Horizontal movements correspond to changes in the number of antigen-bound TCRs n, while vertical movements correspond to changes in the number of antigen-TCR-ZAP complexes z. $\mathcal{B}$ corresponds to states for which the cell activates.

We assume that the T-cell must make a decision within a specified delay D from the first presentation of the antigen i.e. if the cell does not have Z ZAPs recruited by time D, the cell has missed its opportunity to sense the antigen in time. We want to characterize the set $\mathcal{A}_\alpha$ of density-affinity pairs (L, τ_u) for which the T-cell activates before D with probability greater than a prescribed threshold $\alpha \in (0, 1)$. In order to characterize the operating characteristics of the T-cell, we define the first passage time $T_{(i,j)}$ from the state $(i, j) \in \mathcal{S}\backslash\mathcal{B}$ to a state $(n, Z) \in \mathcal{B}$ as follows:

$$T_{(i.j)}(L, \tau_u) = \min\{t \geq 0 : (n_t, z_t) \in \mathcal{B} \text{ s.t. } (n_0, z_0) = (i, j)\} \tag{3}$$

We emphasize that $T_{(i,j)}$ is a function of the density-affinity pair (L, τ_u) since the rate matrix Q depends on (L, τ_u). Since the random time to when a T-cell activates from the initial presentation of the antigen is given by $T_{(0,0)}(L, \tau_u)$, it follows that

$$\mathcal{A}_\alpha = \left\{(L, \tau_u) : \mathbb{P}(T_{(0,0)}(L, \tau_u) \leq D) \geq \alpha\right\}. \tag{4}$$

In order to characterize $\mathcal{A}_\alpha$, we need to characterize the entire distribution of $T_{(0,0)}(L, \tau_u)$ as opposed to just its mean, as was done in earlier works [8, 20]. We are interested in characterizing sets $\mathcal{A}_\alpha$ for a range of α in order to understand the impact of the probability threshold α. We show below that the computational complexity for computing $\mathcal{A}_\alpha$ for a given value of α is approximately the same as the complexity in computing the mean first passage time.

Since an analytical expression for the distribution of the first passage time $T_{(0,0)}$ is unknown, we compute the Laplace transform $\mathcal{L}_{(0,0)}(s) = \mathbb{E}[e^{-sT_{(0,0)}}]$ of $T_{(0,0)}$ using the following result.

Lemma 2.1 (Laplace Transform)**.** *Let $\mathcal{L}_{(i,j)} = \mathbb{E}\left[e^{-sT_{(i,j)}}\right]$ denote the Laplace transform of the density function for the first passage time $T_{(i,j)}$ from state (i, j) to the absorbing boundary $\mathcal{B}$. Then*

$$\mathcal{L}_{(i,j)}(s) = \begin{cases} 1 & (i,j) \in B, \\ \frac{1}{s - Q_{(i,j),(i,j)}} \sum_{(k,l)\neq(i,j)} Q_{(i,j),(k,l)} \mathcal{L}_{(k,l)}(s) & (i,j) \notin B. \end{cases} \tag{5}$$

Therefore, $\mathcal{L}_{(0,0)}(s)$ can be computed by solving the above set of equations. The complexity of computing $\mathcal{L}_{(0,0)}$ is $\mathcal{O}(N^2Z^2)$.

The CDF

$$\mathbb{P}\left(T_{(0,0)} \leq t\right) = \mathcal{L}^{-1}\left\{\frac{\mathcal{L}_{(0,0)}(s)}{s}\right\}(t) = \lim_{\sigma\to\infty}\frac{1}{2\pi i}\int_{b-i\sigma}^{b+i\sigma} e^{st}\frac{\mathcal{L}_{(0,0)}(s)}{s}ds, \tag{6}$$

where b is any real number to the right of all singularities of $\mathcal{L}_{(0,0)}(s)/s$. Let $s = c + id$, for $c, d \in \mathbb{R}$. Then

$$\left|\frac{\mathcal{L}_{(0,0)}(s)}{s}\right| = \frac{1}{\sqrt{c^2+d^2}}\left|\mathbb{E}\left[e^{-cT_{(0,0)}}e^{-idT_{(0,0)}}\right]\right| \leq \frac{1}{\sqrt{c^2+d^2}}\mathbb{E}\left[e^{-cT_{(0,0)}}\right].$$

Thus, it follows that any singularity of $\mathcal{L}_{(0,0)}(s)/s$ has $\Re(s) \leq 0$, therefore any $b > 0$ is sufficient. We use $b = 1$ in our numerical study. Next, using the trapezoidal rule with step size h, we approximate the Bromwich intergral in (6) with an infinite series [1]

$$\lim_{\sigma\to\infty}\frac{1}{2\pi i}\int_{b-i\sigma}^{b+i\sigma} e^{st}\frac{\mathcal{L}_{(0,0)}(s)}{s}ds \approx \frac{he^{bt}}{\pi} + \frac{2he^{bt}}{\pi}\sum_{k=1}^{+\infty}\mathrm{Re}(\mathcal{L}_{(0,0)}(b+ikh))\cos(kht),$$

and then appropriately truncate the series to approximate the integral. We use the specification in Abate et al [1] for h and the number of terms to keep. We found that the number of terms was independent of the parameters (N, Z, D), i.e. only a constant number of Laplace transform evaluations are needed for any parameter values. Thus, it follows that the computational complexity of computing the CDF in terms of (N, Z) is the same as that of computing the mean first passage time (MPFT), since computing the MFPT reduces to solving a system of equations very similar to those used to compute $\mathcal{L}_{(0,0)}$.

In Figure 2 we plot

$$\mathcal{L}_\alpha(\tau_u) = \inf\{L : (L, \tau_u) \in \mathcal{A}_\alpha\} = \inf\left\{L : \mathbb{P}(T_{(0,0)}(L, \tau_u) \leq D) \geq \alpha\right\}, \tag{7}$$

for $\alpha \in \{0.05, 0.5, 0.95\}$. The set $\mathcal{A}_\alpha$ is above, and to the right of the corresponding $\mathcal{L}_\alpha$-curve. We also plot the minimum density required for the MPFT to be below the delay D. The plots in this figure also clearly show the additional information in $\mathcal{L}_\alpha$ for different values of α. For example, the plots show that the threshold is more sensitive to the affinity τ_u as compared to the density L. This information is not available in the single plot corresponding to the MPFT.

We see that the model exhibits a probabilistic version of the absolute ligand discrimination described in [12]. For antigens with sufficiently large affinity, the cell activates with very high probability by the cutoff time D, and it does so for all antigen density L above some threshold $L_{\min}$. For a fixed density above this lower threshold, one can decrease the affinity until a critical affinity $\tau_{\min}$ is reached, at which point the probability of activation starts to dramatically decrease as the affinity is further decreased. Finally, this critical affinity is the same for almost all densities above the density threshold.

The $\mathcal{L}_\alpha$ curve can be very closely approximated by two asymptotes: a vertical asymptote at the affinity $\tau_{\min}$, and a horizontal asymptote at the density $L_{\min}$. See Figure 2 for the asymptotes corresponding to the $\mathcal{L}_{0.05}(\tau_u)$ curve. The thresholds $(L_{\min}, \tau_{\min})$ are functions of the parameters (N, Z), the allowed delay D, and the probability α. We argue that these asymptotes are a consequence of the fact that N is finite. First consider the vertical asymptote. For large enough L, all of the N TCRs are bound to the antigen, and therefore, the ZAP dynamics, and the activation probability, is independent of density L. We rigorously prove this in Section 2.2. The dynamics that lead to the horizontal asymptote are more subtle. We need to show that $L_{\min} = \lim_{\tau_u\to\infty}\mathcal{L}_\alpha(\tau_u) > 0$, i.e. there is a strictly positive density $L_{\min}$ below which the T-cell does not activate. It is clear that $\mathcal{L}_\alpha(\tau_u)$ is a decreasing function of τ_u; therefore, $L_{\min} \geq \mathcal{L}_\alpha(\infty)$. When $\tau_u = \infty$, both n_t and z_t are non-decreasing functions of time. Since $z_t \leq n_t$, it follows that $\mathbb{P}(z_t \geq Z) \leq \mathbb{P}(n_t \geq Z)$; furthermore, the transition rate from state $n \to n+1$ is bounded above by Nk_bL. Thus, it follows that the probability that the T-cell activates by delay D with probability α is bounded below by the probability that $\mathbb{P}(\sum_{i=1}^{Z-1} X_i \leq D) \geq \alpha$, where X_i are independent, exponential random variables with rate Nk_bL. Thus, it follows that $L_{\min} \geq \mathcal{L}_\alpha(\infty) \geq \inf\{L : \mathbb{P}(\sum_{i=1}^{Z-1} X_i \leq D) \geq \alpha\} > 0$. Note that the last term is strictly positive only if $N < \infty$. We rigorously establish this result in Section 2.3.

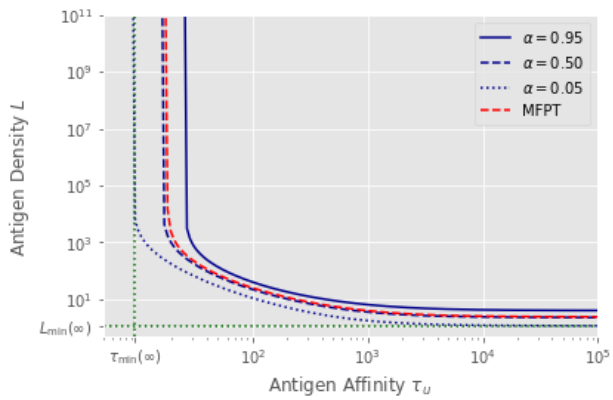

Figure 2: Let $\mathcal{L}_\alpha(\tau_u)$ denote the minimum density required to activate by delay D with probability α (see (7)). Plots of $\mathcal{L}_\alpha(\tau_u)$ for $Z = 8$ $N = 12$ and delay $D = 10^3$. For each affinity τ_u, the minimum density L was found using bisection over densities ranging from 10^{-4} to 10^{16}. The delay $D = 10^3$ seconds was chosen because it represents a realistic amount of time for human immune cells to operate over (10^3 seconds ≈ 15 minutes). Horizontal and vertical asymptotes $L_{\min}(\infty)$ and $\tau_{\min}(\infty)$ displayed for $\alpha = 0.05$. Also included is the minimum density L needed for a cell exposed to antigen with affinity τ_u to have a MFPT less than D. Time to compute $\alpha = 0.50$ curve was 30 seconds, time to compute MPFT curve was 2 seconds. This factor of 15 difference comes from using 15 summands to approximate the Bromwich integral. This same ratio holds when testing larger N and Z values.

In the rest of this section, we propose two one-dimensional Markov chains that approximate the behavior of the full two-dimensional chain for large (L, τ_u), are exact in the limit of density $L \to \infty$ or affinity $\tau_u \to \infty$. The goal of studying these chains is two-fold. The first is purely computational. We want to understand the dependence of $(L_{\min}, \tau_{\min})$ on (N, Z, D, α), and we get close to 100 times computational speed-up from using the 1-dimensional chains. The second reason for studying these chains is to rigorously establish the mechanisms leading to the asymptotes, or equivalently, absolute ligand discrimination.

2.1 Minimum Possible Activation Time

We begin with characterizing a lower bound for the delay D such that the first passage time $T_{(0,0)}(L, \tau_u) \leq D$ with probability at least α for some (L, τ_u). Since the forward transitions, i.e. those that increase (n, z), are increasing in L and the backward transitions, i.e. those that decrease (n, z), are decreasing in τ_u, it follows that the first passage time $T_{(0,0)}$ is monotonically decreasing in both L and τ_u. Thus, in order to ensure there exist (L, τ_u) for which the first passage time $T_{(0,0)}(L, \tau_u) \leq D$ with probability at least α, it is sufficient to show that $\lim_{\tau_u \to \infty} \lim_{L \to \infty} \mathbb{P}(T_{(0,0)}(L, \tau_u) \leq D) \geq \alpha$. In Section 2.2, we establish that the first passage time $T_{(0,0)}(\tau_u, L)$ in the limit $L \to \infty$ is approximated by the first passage time $T_0^{(B)}$ from state 0 to state Z in a birth-death process with birth rates $\lambda_z = (N - z)k_{on}$ and death rate $\mu_z = z/\tau_u$. Next, taking the limit $\tau_u \to \infty$, the limit

$$\begin{aligned} \lim_{\tau_u \to \infty} T_0^{(B)} &= \exp(Nk_{on}) + \exp((N-1)k_{on}) + \cdots + \exp((N - Z + 1)k_{on}), \\ &= \text{Erlang}((H_N - H_Z), k_{on}), \end{aligned}$$

where $\exp(\lambda)$ denotes an exponential random variable with rate λ, and $H_N = \sum_{k=1}^{N} \frac{1}{k}$ denotes the N-th Harmonic number. The same limit results from taking the limit $L \to \infty$ in the Markov chain $M^{(f)}$ that we discuss in Section 2.3. Thus, in order to ensure that the network reaches $\mathcal{B}$ with probability at least α, we need

$$D \geq D^*(N, Z, \alpha) \triangleq \inf \left\{ T : \mathbb{P}\left(\text{Erlang}((H_N - H_Z), k_{on} \leq T\right) \geq \alpha \right\}. \tag{8}$$

We ensure that $D \geq D^*(N, Z, \alpha)$ in all of our analyses so that $\tau_{\min}(\infty)$ and $L_{\min}(\infty)$ are well-defined.

2.2 Vertical Asymptote

Fix $D \geq D^*$ (see (8)), and define

$$;\tau_{\min}(L; N, Z, D, \alpha) \triangleq \inf\left\{\tau_u > 0 : \mathbb{P}\left(T_{(0,0)}(\tau_u, L; N, Z) \leq D\right) \geq \alpha\right\}. \tag{9}$$

where $T_{(0,0)}(\tau_u, L; N, Z)$ is the first passage time from state $(0,0)$ to the absorbing boundary $\mathcal{B}$ for an antigen with affinity τ_u and density L when the cell has N TCRs and activates when Z antigen-TCR-ZAP complexes form. Informally, $\tau_{\min}(L; N, Z, D, \alpha)$ denotes the minimum affinity τ_u needed to ensure that the cell activates with probability at least α by time D with antigen density L, when the cell has N TCRs on the cell surface, and Z phosphorylated TCRs needed for activation. For brevity, let $\tau_{\min}(L)$ denote $\tau_{\min}(L; N, Z, D, \alpha)$.

To study the behavior of $\tau_{\min}(L)$ as $L \to \infty$, we define

$$\tau_{\min}(\infty) \triangleq \tau_{\min}\left(\lim_{L\to\infty} L\right) = \inf\left\{\tau_u > 0 : \lim_{L\to\infty} \mathbb{P}\left(T_{(0,0)}(\tau_u, L) \leq D\right) \geq \alpha\right\} \tag{10}$$

where the last equality holds by Fatou's lemma and the fact that $\lim_{L\to\infty}\{T_{(0,0)}(\tau_u, L) \leq D\}$ exists and is equal to $\{T_{(0,0)}(\tau_u, \infty) \leq D\}$, so we can exchange the limit and probability.

Note that as $L \to \infty$, all the N TCRs instantly convert into TCR-antigen complexes, and the bottleneck event that determines activation is the recruitment of ZAP, i.e. activation is controlled by the ZAP dynamics. We formalize this in Theorem 2.2. Let $B = \{b_t : t \geq 0\}$ denote a birth-death process defined on $\{z : 0 \leq z \leq Z\}$ with birth rates $\lambda_z = (N - z)k_{on}$ and death rates $\mu_z = z/\tau_u$, and an absorbing boundary at Z. Let $T_0^{(B)}(\tau_u; N, Z)$ denote the first passage time from state 0 to state Z in a birth-death process B. Then, we have the following characterization for $\tau_{\min}(\infty)$.

Theorem 2.2 (Vertical Asymptote)**.** *The minimum affinity $\tau_{\min}(\infty)$ required to activate in the limit $L \to \infty$ is given by*

$$\tau_{\min}(\infty) = \tau_{\min}^{(B)}(N, Z, D, \alpha) \triangleq \inf\left\{\tau_u > 0 : \mathbb{P}\left(T_0^{(B)}(\tau_u; N, Z) \leq D\right) \geq \alpha\right\}.$$

Thus, Theorem 2.2 implies that, in order to characterize $\tau_{\min}(\infty)$ as a function of (N, Z, T, α), we only need to compute the distribution of the first passage time $T_0^{(B)}(\tau_u)$ for the birth-death process B. The Laplace transforms $\mathcal{L}_i(s)$ for the first passage time from state i to state Z for B can be computed in the same way as the Laplace transforms $\mathcal{L}_{(i,j)}(s)$ for the original model. Computing $\mathcal{L}_i(s)$ for $0 \leq i \leq Z$ reduces to solving a tri-diagonal system of equations,with a computational complexity of $\mathcal{O}(Z)$ as opposed to $\mathcal{O}(N^2Z^2)$. Figure 3 confirms that $\tau_{\min}(\infty) = \tau_{\min}^{(B)}$ and that $\tau_{\min}^{(B)}$ can be computed much faster than $\tau_{\min}(\infty)$ ($\sim$10-100 times faster for moderate-sized N and Z).

2.3 Horizontal Asymptote

Fix $D \geq D^*$ (see (8)), and define

$$L_{\min}(\tau_u; N, Z, D, \alpha) = \inf\left\{L > 0 : \mathbb{P}\left(T_{(0,0)}(\tau_u, L; N, Z) \leq D\right) \geq \alpha\right\}$$

denote the minimum density L needed to ensure that the T-cell activates with delay D with probability at least α. We want to show that $\lim_{\tau_u\to\infty} L_{\min}(\tau_u; N, Z, D, \alpha)$ exists, study the dependence of $\lim_{\tau_u\to\infty} L_{\min}(\tau_u; N, Z, D, \alpha)$ on the parameters (N, Z, D, α), and understand how quickly $L_{\min}(\tau_u; N, Z, D, \alpha)$ converges to $\lim_{\tau_u\to\infty} L_{\min}(\tau_u; N, Z, D, \alpha)$. For brevity, let $L_{\min}(\tau_u)$ denote $L_{\min}(\tau_u; N, Z, D, \alpha)$.

To study the behavior of $L_{\min}(\tau_u)$ as $\tau_u \to \infty$, we define

$$L_{\min}(\infty) \triangleq L_{\min}\left(\lim_{\tau_u\to\infty} \tau_u\right) = \inf\left\{L > 0 : \lim_{\tau_u\to\infty} \mathbb{P}\left(T_{(0,0)}(\tau_u, L) \leq D\right) \geq \alpha\right\}$$

where the last equality holds because the limit and probability can be exchanged for the same reason as in Section 2.2. In the limit of $\tau_u \to \infty$, the unbinding rate $k_u \to 0$, and therefore the rates $Q_{(n,z),(n-1,z)} = Q_{(n,z),(n-1,z-1)} \to 0$, i.e. the activation time should be approximated by a Markov chain where the state

(n, z) is always monotonically increasing. To this end, define the Markov chain $M^{(f)}$ with only forward transitions i.e. the rate matrix $Q^{(f)}$ is given by

$$\begin{aligned}
Q^{(f)}_{(n,z),(n+1,z)} &= (N-n)k_b L \mathbb{1}\left(n < N, z < Z\right) \\
Q^{(f)}_{(n,z),(n,z+1)} &= (n-z)k_z^{on} \mathbb{1}\left(z < \min\{n, Z\}\right) \\
Q^{(f)}_{(n,z),(n,z)} &= -\left(Q^{(f)}_{(n,z),(n+1,z)} + Q^{(f)}_{(n,z),(n,z+1)}\right) \\
Q^{(f)}_{(n,Z),(\cdot,\cdot)} &= 0
\end{aligned}$$

and all other rates not specified are 0. The only difference between the original rate matrix Q, and the rate matrix $Q^{(f)}$ is that $Q^{(f)}_{(n,z),(n-1,z)} = Q^{(f)}_{(n,z),(n-1,z-1)} = 0$. Let $T^{(f)}_{(0,0)}$ denote the first passage time from the state $(0, 0)$ to the boundary $\mathcal{B}$ for the chain $M^{(f)}$. Then we have the following result.

Theorem 2.3 (Horizontal Asymptote)**.** *The minimum density $L_{\min}(\infty)$ required to activate in the limit $\tau_u \to \infty$ is given by*

$$L_{\min}(\infty) = L^{(f)}_{\min}(N, Z, D, \alpha) \triangleq \inf\left\{L > 0 : \mathbb{P}\left(T^{(f)}_{(0,0)}(L; N, Z) \leq D\right) \geq \alpha\right\}.$$

Theorem 2.3 implies that we can compute $L_{\min}(\infty)$ and understand its dependence on (N, Z, D, α) by computing the first passage time distribution for the simpler Markov chain $M^{(f)}$. Computing the Laplace transforms $\mathcal{L}_{(i,j)}(s)$ for $M^{(f)}$ reduces to solving a sparse upper triangular set of equations. These equations can be solved via back-substitution resulting in $\mathcal{O}(NZ)$ complexity as compared to $\mathcal{O}(N^2Z^2)$ for computing the Laplace transforms for the first passage times for M. Figure 3 (e)-(h) confirms that $L_{\min}(\infty) = L^{(f)}_{\min}$ and that $L^{(f)}_{\min}$ can be computed much faster than $L_{\min}(\infty)$ ($\sim$10-100 times faster for moderate-sized N and Z). Computing the horizontal asymptote is faster for smaller N and Z because it only requires one pass through the problem data (of size $\mathcal{O}(NZ)$), whereas computing the vertical asymptote requires two passes through the problem data (of size $\mathcal{O}(Z)$). The difference in complexity starts to show up once N and/or Z become larger.

2.4 1-D Approximation for Finite L, τ_u

In Sections 2.2 and 2.3 we established that the vertical and horizontal asymptotes of the $L_\alpha(\tau_u)$ curve can be characterized in terms of simpler Markov chains whose first passage time distributions can be computed much more quickly. In this section, we study the shape of the $\mathcal{L}_\alpha$ curve for finite values of L and τ_u. We do this because the asymptotes do not fully determine if the proofreading network will discriminate between self and non-self antigens. Even if $\tau_{\min}(\alpha = 0.05)$ and $\tau_{\min}(\alpha = 0.95)$ are close together, and separate the self and non-self antigen affinity, the network will not discriminate the self and non-self antigen if the density L needed for $\tau_{\min}(L)$ to be close to $\tau_{\min}(\infty)$ is very large. Similarly, it could be that $L_{\min}(\infty)$ is quite small, but the affinity τ_u needed to ensure $L_{\min}(\tau_u) \approx L_{\min}(\infty)$ is very large. For a proofreading network to display absolute ligand discrimination, it must be that $\tau_{\min}(L) \approx \tau_{\min}(\infty)$ when $L \approx L_{\min}(\infty)$, i.e. $\mathcal{L}_\alpha$ should converge quickly to the horizontal asymptote, and $L_{\min}(\tau_u) \approx L_{\min}(\infty)$ whenever $\tau_u \approx \tau_{\min}(\infty)$, i.e. the $\mathcal{L}_\alpha$ should converge to the vertical asymptote quickly.

To this end, we define

$$\begin{aligned}
\tau_\epsilon(N, Z, D, \alpha) &\triangleq \tau_{\min}\left((1+\epsilon)L_{\min}(\infty)\right) \\
L_\epsilon(N, Z, D, \alpha) &\triangleq L_{\min}\left((1+\epsilon)\tau_{\min}(\infty)\right)
\end{aligned}$$

τ_ϵ (resp. L_ϵ) quantifies the steepness of the $L - \tau_u$ as it approaches the horizontal (resp. vertical) asymptote. See Figure 7 to see this. Informally, $L_{\min}(\infty)$ is the value of the density L corresponding to affinity $\tau_u = \infty$, and τ_ϵ corresponds to the affinity such that $((1+\epsilon)L_{\min}(\infty), \tau_\epsilon)$ is on the $\mathcal{L}_\alpha$ curve, i.e. at $\tau_u = \tau_\epsilon$ the density value is within an $(1+\epsilon)$-factor of the level at $\tau_u = \infty$. Therefore, if τ_ϵ is relatively small, the $\mathcal{L}_\alpha$ curve approaches the horizontal asymptote very quickly. A similar analysis motivates the definition of L_ϵ. These quantities help us understand the curvature of the $\mathcal{L}_\alpha$ curves.

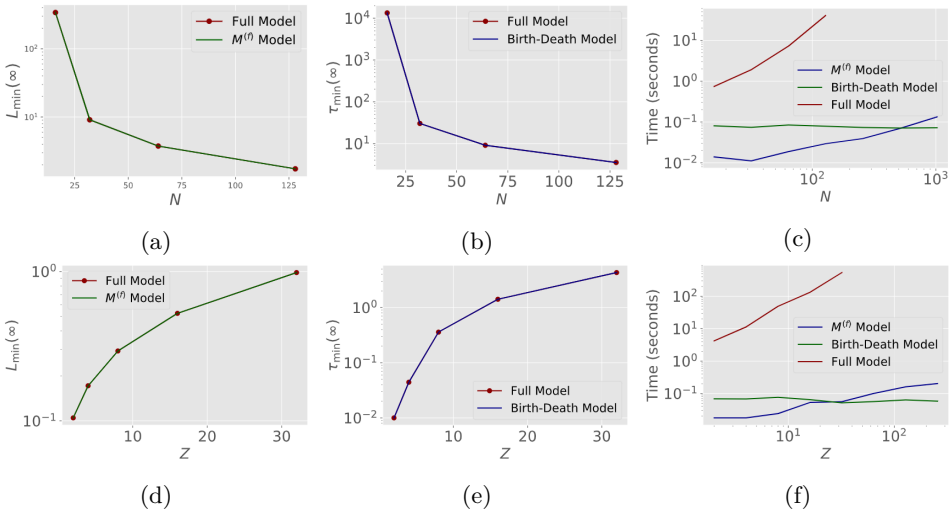

Figure 3: Comparison of location of horizontal and vertical asymptotes computed using the full model vs. using the simplified models, as well as the speed to compute them. (a) $L_{\min}(\infty)$ computed using full model (red dots) and simplified model $M^{(f)}$ (green line), as a function of $N \in \{16, 32, 64, 128\}$, $Z = 16$, $\alpha = 0.95$, $D = 300$s. (b) $\tau_{\min}(\infty)$ computed using full model and birth-death process (blue line). N, Z, α, D same as (a). (c) Speed to compute $L_{\min}(\infty)$ using $M^{(f)}$, speed to compute $\tau_{\min}(\infty)$ using birth-death process, and speed to compute $\tau_{\min}(\infty)$ using full model, as a function of N (speed to compute $\tau_{\min}(\infty)$ and speed to compute $L_{\min}(\infty)$ using full model are the same). For $L_{\min}(\infty)$ and $\tau_{\min}(\infty)$, N values extended to $N \in \{16, 32, 64, 128, 256, 512, 1024\}$. (d)-(f) Same as (a)-(c) but as a function of Z. $N = 256$, $\alpha = 0.95$, $D = 450$s. $Z \in \{2, 4, 8, 16, 32\}$ for (a), (b), and full model in (c); $Z \in \{2, 4, 8, 16, 32, 64, 128, 256\}$ in (c) for simplified models.

In order to efficiently simulate the entire $\mathcal{L}_\alpha$ curve and investigate its "distance" from ideal absolute ligand distribution behavior, we develop a 1-dimensional approximation of the full model M where the n- and $z-$dynamics are decoupled. We show that this decoupling significantly reduces the complexity of computing $(\tau_\epsilon, L_\epsilon)$.

Consider a birth-death process $B = \{b(t) : t \geq 0\}$ on the state space $\mathcal{Z} = \{0, 1, ..., Z\}$ with an absorbing boundary at Z. The transition rates are

$$\begin{aligned} \mu_{z,z-1} &= z/\tau_u, \\ \mu_{z,z+1} &= \frac{1}{\mathbb{E}\left[\gamma_z(L, \tau_u)\right]}, \end{aligned} \tag{11}$$

with $\gamma_z(L, \tau_u)$ being the absorption time in a z-dependent Markov chain defined below. See Figure 4.

Let $M^{(z)} = \{n^{(z)}(t) : t \geq 0\}$ denote a Markov chain on $\{z, z+1, ..., N\} \cup \{\mathcal{A}\}$, where the state $\mathcal{A}$ is an absorbing state. For state $i \in \{z, z+1, ..., N\}$ the transition rates are given by

$$\begin{aligned} \mu^{(z)}_{i,i-1} &= (i - z)/\tau_u, \quad i > z, \\ \mu^{(z)}_{i,i+1} &= (N - i)k_b L, \quad i < N, \\ \mu^{(z)}_{i,\mathcal{A}} &= (i - z)k_{on}. \end{aligned}$$

The Markov chain $M^{(z)}$ starts from the initial state $n^{(z)}(0)$ defined below. Before describing the procedure for selecting $n^{(z)}(0)$, we describe the rationale for this decoupling of the (n, z) dynamics.

The T-cell activates when $z = Z$. We would be able to efficiently estimate the quantities of interest if one could describe a Markov chain where the transition rates are only a function of the state z. The rate from

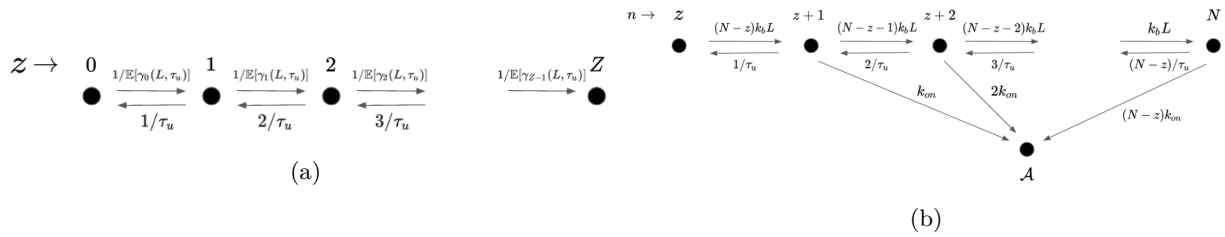

Figure 4: Proposed model used to approximate the convergence of $\tau_{\min}(L) \rightarrow \tau_{\min}(\infty)$ and $L_{\min}(\tau_u) \rightarrow L_{\min}(\infty)$. The first passage time to the boundary $\mathcal{B}$ in M is approximated with the first passage time to Z in the birth-death process shown in (a). $\gamma_z(L,\tau_u)$ denotes the first passage time to $\mathcal{A}$ starting value $n^{(z)}(0)$ in the chain displayed in (b). We replace $\gamma_z(L,\tau_u)$ by the maximum entropy distribution with mean $\mathbb{E}[\gamma_z(L,\tau_u)]$, i.e. an exponential with rate $1/\mathbb{E}[\gamma_z(L,\tau_u)]$. Once the model in subfigure (a) transitions to a new value of z (so either to $z-1$ or $z+1$), the model here changes to reflect the change in z and restarts.

state z to state $z-1$ in the 2-D Markov chain M is given by $\frac{z}{\tau_u}$, i.e. it only depends on z, and so we keep this rate as it is. However, the rate from state z to $z+1$ depends on the number of TCR-antigen complexes n, and therefore, cannot be written as a function of only the state z. We know that $n \geq z$, and so define a Markov chain $M^{(z)}$ with the state space consisting of $\{z, \ldots, N\}$ and an absorbing state $\mathcal{A}$ that represents the event $z \rightarrow z+1$. Note that in the 2-D Markov chain, it's possible that n is much larger than z when the chain makes a transition from $z \rightarrow z+1$, and so setting $n^{(z)}(0) = z$ could significantly overestimate the first passage time. To correct this, we set the initial state of $n^{(z)}(0)$ in the following recursive manner:

$$n^{(z)}(0) = \inf\left\{i \geq n^{(z-1)}(0) : \mu_{i,i+1}^{(z-1)} < \mu_{i,\mathcal{A}}^{(z-1)}\right\},$$

i.e. $n^{(z)}(0)$ is set to the value of $i \geq n^{(z-1)}(0)$ where the rate of transition to the stopping state $\mathcal{A}$ is larger than the rate to the state $i+1$. Setting $n^z(0)$ in this manner allows the approximation to mimic the behaviour of the 2-D chain M both for when $L\tau_u$ is very large and the transitions in M are likely to be close to $(0,0) \rightarrow (1,0) \rightarrow (2,0) \cdots \rightarrow (N,0) \rightarrow (N,1) \cdots \rightarrow (N,Z)$, and when $L\tau_u$ is very small and the the transitions in M are likely to be close to $(0,0) \rightarrow (1,0) \rightarrow (1,1) \rightarrow (2,1) \rightarrow \cdots \rightarrow (N,Z-1) \rightarrow (N,Z)$. This setting also allows for more complicated transitions associated with moderate values of $L\tau_u$. Finally, we replace the first passage time $\gamma_z(L,\tau_u)$ from $n_z(0)$ to $\mathcal{A}$ by the maximum entropy non-negative random variable with the same mean, i.e. an exponential random variable with rate $\frac{1}{\mathbb{E}[\gamma_z(L,\tau_u)]}$. This explains the forward rate in (11).

Inverting the Laplace transform for the birth-death process B allows us to approximate τ_ϵ and L_ϵ. Since calculating the mean first passage time to $\mathcal{A}$ for each of the $M^{(z)}$ processes involves solving a tri-diagonal system of equations, this step takes at most $\mathcal{O}(N)$ time; resulting in a total complexity of $\mathcal{O}(NZ)$. The Laplace transform for the birth-death process can be computed in $\mathcal{O}(Z)$ time. Thus, the approximate model results in an $\mathcal{O}(NZ)$ speedup compared to using the model M.

Figures 5 and 6 compare the values of $\tau_\epsilon(N,Z,D,\alpha)$ and $L_\epsilon(N,Z,D,\alpha)$ computed using the full model compared to the values computed using the approximation detailed in this section. We see that the true and approximated values of τ_ϵ and L_ϵ are typically within a factor of 2 from one another, and the time to approximate these two values is $10-100$ times faster for the values of N and Z tested.

3 Results

In this section, we explore several biologically relevant metrics using the models developed in Section 2. Recall that two metrics introduced earlier were *sensitivity* defined as $L_{\min}(\infty)$, and *specificity*, defined as the gap between the $\alpha = 0.95$ vertical asymptote and the $\alpha = 0.05$ vertical asymptote, i.e. reliability of the proof reading network.

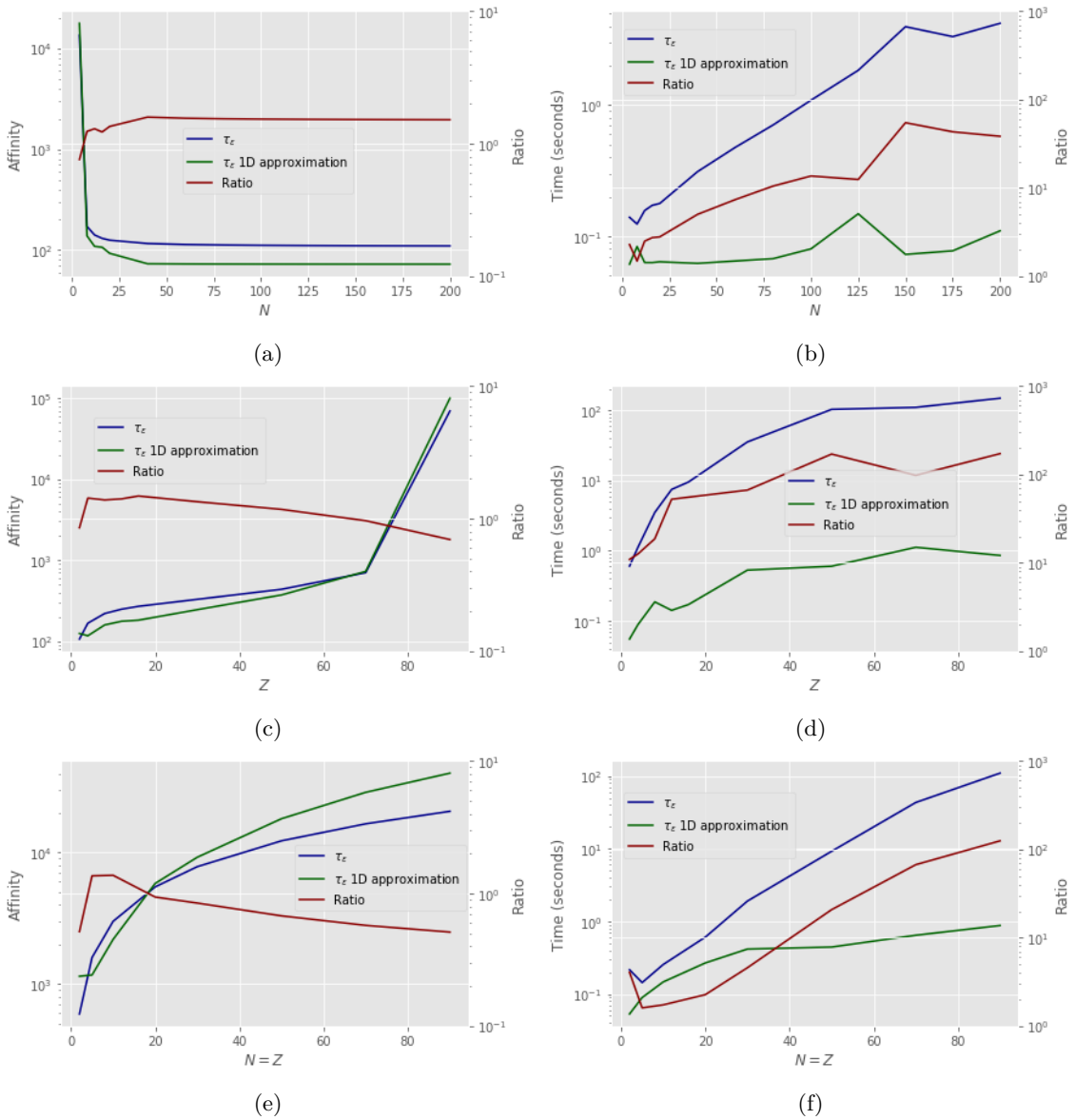

Figure 5: Comparison of τ_ϵ using the full model M and the approximation in Figure 4. $\epsilon = 1$ and $\alpha = 0.95$ for all figures. (a) τ_ϵ as computed by the full model M (blue line) and the approximation (green line), as well as the ratio of the two (red line), as a function of $N = 4, ..., 200$. $Z = 4$, $D = 225$. (b) Time to compute τ_ϵ using original model M (blue line) and approximation (blue line), as well as ratio between the two, as a function of N. Same parameters as (a). (c) Same as (a) but as a function of $Z = 2, ..., 90$. $N = 90$, $D = 400$. (d) Same as (b), but as a function of Z. Same parameters as (c). (e) Same as (a), but as a function of $N = Z = 2, ..., 90$, with D set to $1.25D^*(N, Z, \alpha)$. (f) Same as (b), but as a function of $N = Z$. Same parameters as (e).

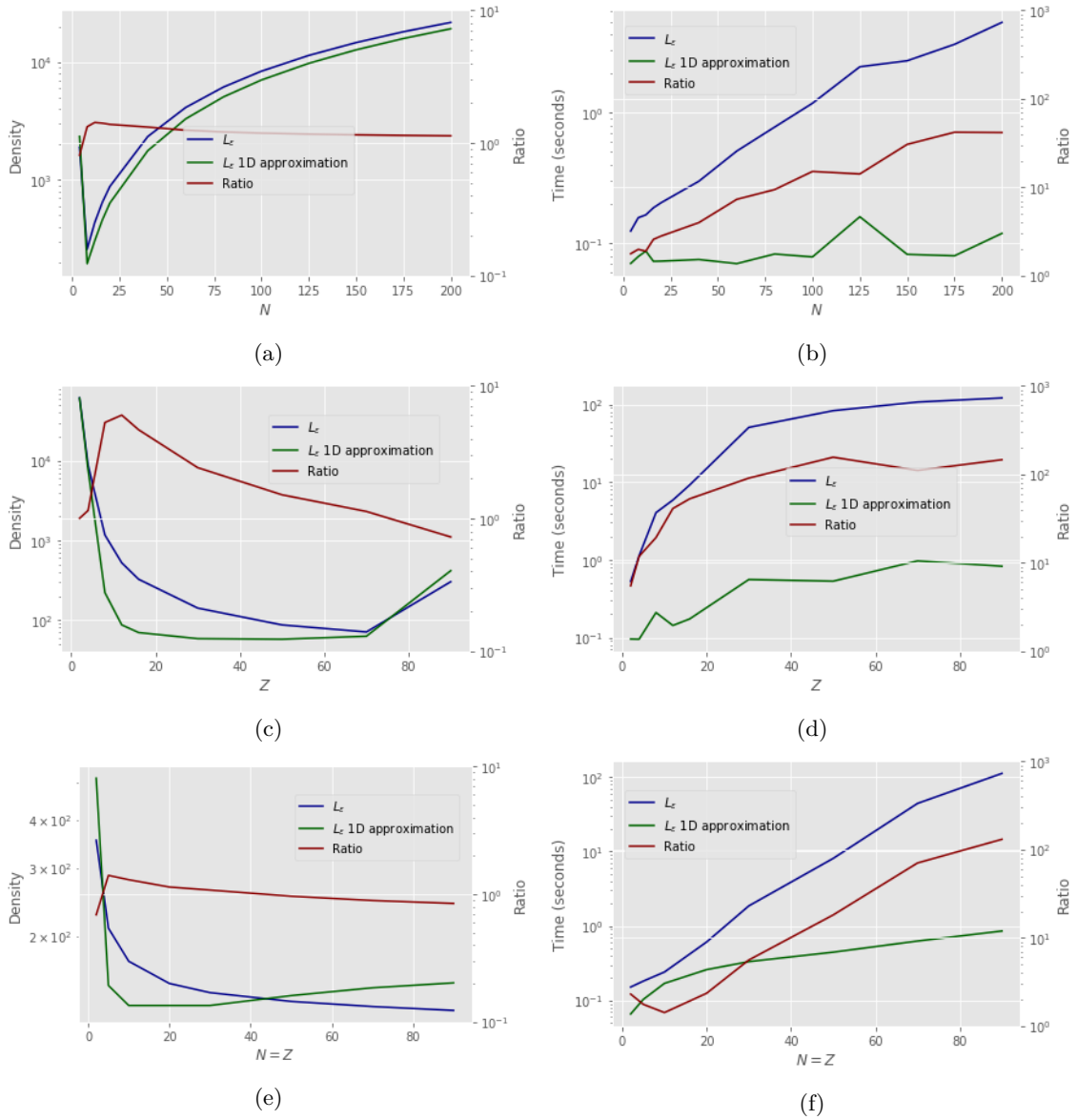

Figure 6: Comparison of L_ϵ using original model M and approximation from Figure 4. Same layout as Figure 5.

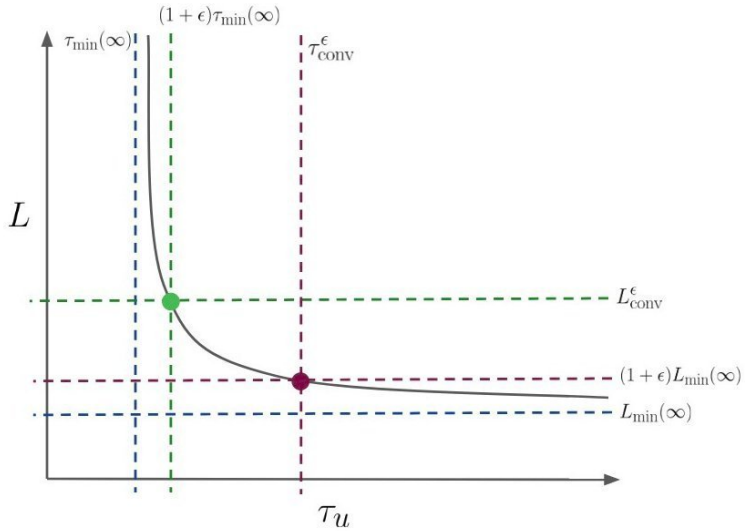

Figure 7: Examples of covering $(\tau_u, L_{\min}(\tau_u))$ curve with $C_\epsilon^{\tau_u}$ (in green) and C_ϵ^{L} (in red).

The monotonicity properties for the Markov chain M imply that the $\mathcal{L}_\alpha$ curve lies in both the sets

$$\mathcal{C}_\epsilon^{\tau_u} = \{(L, \tau_u) : L_{\min}(\infty) \leq L \leq L_\epsilon \text{ or } \tau_{\min}(\infty) \leq \tau_u \leq (1+\epsilon)\tau_{\min}(\infty)\},$$
$$\mathcal{C}_\epsilon^{L} = \{(L, \tau_u) : L_{\min}(\infty) \leq L \leq (1+\epsilon)L_{\min}(\infty) \text{ or } \tau_{\min}(\infty) \leq \tau_u \leq \tau_\epsilon\}.$$

See Figure 7. The set $C_\epsilon^{\tau_u}$ corresponds to the L-shaped set bounded by the blue and green lines, and the set C_ϵ^{L} corresponds the L-shaped set bounded by the blue and purple lines. These coverings provide an estimate for the "distance" of the $\mathcal{L}_\alpha$ curve from the ideal absolute ligand discrimination curve. We define the "distance" to be the maximum of the L and τ_u gap between the two boundaries of the L-shaped covering. This, the "distance" implied by the covering $C_\epsilon^{\tau_u}$ (resp. C_ϵ^{L}), is $d_\epsilon^{\tau_u} = \max\{\epsilon\tau_{\min}(\infty), L_\epsilon - L_{\min}(\infty)\}$ (resp. $d_\epsilon^{L} = \max\{\tau_\epsilon - \tau_{\min}(\infty), \epsilon L_{\min}(\infty)\}$). We define the "distance" $d_\epsilon = \min\{d_\epsilon^{\tau_u}, d_\epsilon^{L}\}$ to be the minimum of the "distance" implied by either of the two covering sets. We define the *fidelity* $F(N, Z, D, \alpha)$ as the inverse of the minimum "distance" over all possible choices for ϵ, i.e.

$$\begin{aligned} F(N, Z, D, \alpha) &= \frac{1}{\min_{\epsilon>0}\{d_\epsilon\}}, \\ &= \frac{1}{\min_{\epsilon>0}\{\min\{\max\{\epsilon\tau_{\min}(\infty), L_\epsilon - L_{\min}(\infty)\}, \max\{\tau_\epsilon - \tau_{\min}(\infty), \epsilon L_{\min}(\infty)\}\}\}} \end{aligned} \tag{12}$$

The next metric we are interested in is the robustness of the ligand discrimination. Due to the inherent noise in the copy number and transport of receptors in biological systems, the ideal kinetic proofreading network must be resilient to errors in the number of TCRs N on the cell surface, as well as errors in the number of ZAP molecules Z required for activation. We assume that the T-cell wants to discriminate between self antigens with affinity τ_{self} and a non-self antigen with affinity $\tau_{\text{non-self}}$ within a delay D. Let N and Z denote the nominal values describing a particular kinetic proofreading network. Since the actual number of TCRs $N+\delta N$ on the cell surface may differ from N because of transport noise, the corresponding $\mathcal{L}_\alpha$ curve may no longer be able to discriminate between the self and non-self antigens. We define the robustness $r_N(N, Z, D)$ of the kinetic proofreading network with parameters (N, Z, D) to errors in N as follows:

$$r_N(N, Z, D) \triangleq \frac{1}{N}\max\left\{\delta N : \begin{smallmatrix}\tau_{\text{self}} \leq \tau_{\min}(\infty; N\pm\delta N, Z, T, \alpha=0.05), \\ \tau_{\text{non-self}} \geq \tau_{\min}(\infty; N\pm\delta N, Z, T, \alpha=0.95)\end{smallmatrix}\right\} \tag{13}$$

$r_N(N, Z, D, \alpha)$ denotes the maximum relative change in N before the kinetic proofreading network is unable to differentiate between the self and non-self antigens.

The robustness with respect to $r_Z(N, Z, D)$ is similarly defined as

$$r_Z(N, Z, D) \triangleq \frac{1}{Z} \max \left\{ \delta Z : \begin{smallmatrix} \tau_{\text{self}} \leq \tau_{\min}(\infty; N, Z \pm \delta Z, T, \alpha = 0.05), \\ \tau_{\text{non-self}} \geq \tau_{\min}(\infty; N, Z \pm \delta Z, T, \alpha = 0.95) \end{smallmatrix} \right\} \tag{14}$$

It is possible that the robustness $r_N(N, Z, D, \alpha)$ is very sensitive to Z. In order to study this effect, we define robustness $r_{NZ}(N, Z, D, \alpha)$ to simultaneous perturbations in N and Z as follows:

$$r_{NZ}(N, Z, D) \triangleq \min_{|\delta Z| \leq m} \{ r_N(N, Z + \delta Z, D) \} \tag{15}$$

Using the computational strategies described in Section 2, we simulated the performance of the kinetic proofreading system for

$$\begin{aligned} (\tau_{\text{self}}, \tau_{\text{non-self}}) &\in \{(1, 10), (3, 10)\} \text{ seconds}, \\ N &\in \{2, \ldots, 100\}, \\ Z &\in \{2, \ldots, N\}, \\ D &\in [1.1 D^*(N, Z, \alpha = 0.95), 100 D^*(N, Z, \alpha = 0.95)] \,. \end{aligned} \tag{16}$$

For a given value of $(\tau_{\text{self}}, \tau_{\text{non-self}})$ we call (N, Z, D) *admissible* if the asymptote $\tau_{\min}(\infty; N, Z, D, \alpha = 0.95) \leq \tau_{\text{non-self}}$, i.e. the proofreading network will activate for the non-self antigen within delay D with probability $\alpha \leq 0.95$, and $\tau_{\min}(\infty; N, Z, D, \alpha = 0.05) \geq \tau_{\text{self}}$, i.e. the proofreading network does *not* trigger for the self-antigen with probability $\geq 1 - \alpha = 0.95$. This filtering step fixes the specificity of the proofreading network.

In Figure 8 we plot sensitivity $L_{\min}(\infty; N, Z, D, \alpha = 0.95)$ vs the delay D for all admissible values of (N, Z, D) – each point corresponds to an admissible triple (N, Z, D). The lower envelop of the set of points corresponds to values of (N, Z, D) that are Pareto optimal with respect to sensitivity and delay, i.e. these points represent the optimal trade-off between these two metrics. Note that on the Pareto front, the sensitivity $L_{\min}$ is a decreasing function of the delay D, i.e. the proofreading network can be designed to be more sensitive if a higher delay is tolerable.

Next, we want to visualize the (L, D, F) Pareto front. For any point (L, D), we define

$$F(L, D) = \max \{ F(N', Z', D') : ||(L_{\min}(\infty; N', Z'), D') - (L, D)||_2 / ||(L, D)||_2 \leq \delta = 0.05 \} \,.$$

The color of the point (L, D) denotes the value of $F(L, D)$. We take the maximum over a neighborhood of (L, D) in order to smooth the sampling error. On the Pareto front, the fidelity and sensitivity both increase as the allowed delay D increases, and begins to level off between $D = 100$ and $D = 1000$ seconds, a biologically realistic range of activation times. In the right-hand plot in Figure 8 we see that, as one would expect, increasing the required specificity by increasing τ_{self} limits the delay, sensitivity, and fidelity with which the cell can operate.

In Figure 9 the points are colored by robustness. Similar to Figure 8, the color recovers the projection of points from the space of sensitivity, delay, and robustness to the space of sensitivity and delay. In the first row the points are colored by robustness to perturbations in N, i.e. r_N defined in (13), in the second row the points are colored by robustness to perturbations in Z, i.e. r_Z defined in (14), and the third row the points are colored by the joint robustness r_{NZ} defined in (15). We see that the values of r_N and r_Z on the Pareto front increase as D increases. The trends in the N and Z-robustness r_{NZ} are slightly more complicated – on the Pareto front, r_{NZ} is maximized for intermediate values of D.

Increasing specificity by only a small amount (e.g. a 30 % increase in τ_{self}) can significantly reduce the achievable level of robustness in the kinetic proofreading model. Increasing specificity had the same effect on fidelity, although this change is not as drastic as that for robustness.

4 Conclusion

In this paper, we propose a kinetic proofreading model that achieves absolute ligand discrimination by leveraging the fact that the number of TCRs is finite. We model the T-cell dynamics by a 2-dimensional

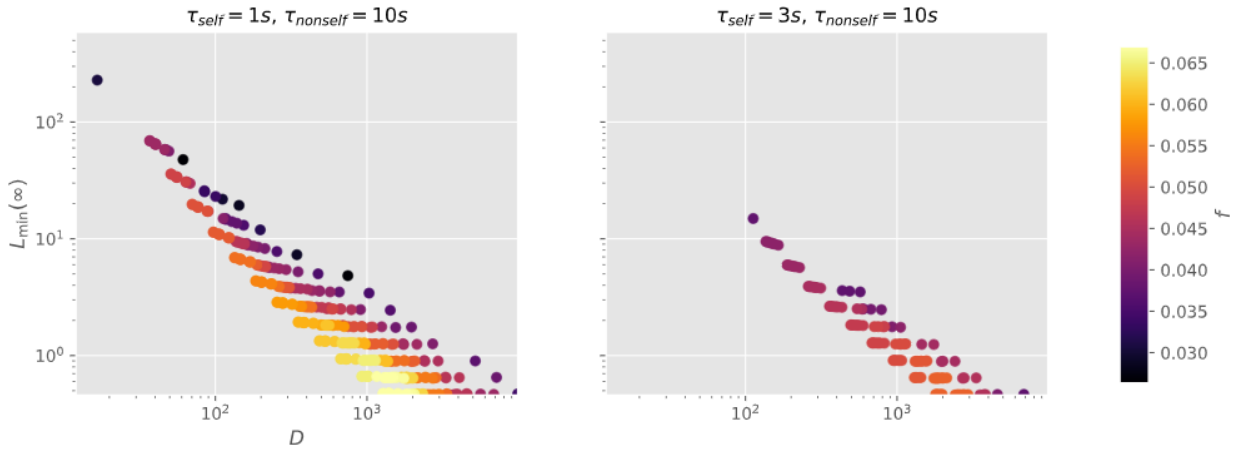

Figure 8: Achievable values of D and $L_{\min}(\infty)$ for fixed τ_{self} and $\tau_{\text{non-self}}$. Points are colored based on the value of $F(L, D)$.

Markov chain and describe how to compute the first passage time to activation in the model. We show that in the limit of $L \to \infty$ the 2-D model reduces to a 1-D birth-death process, and in the limit as $\tau_u \to \infty$ the model reduces to a simpler 2-D model with no backward transitions. These simplifications significantly reduce the time required to compute the first passage time distribution. We also provide a 1-D approximation of the full model for finite L and τ_u, and show via numerical simulations that it approximates the model fairly well. Next, we use this 1-D model to study the trade-offs between sensitivity, specificity, speed, fidelity, and robustness to errors in N and Z. We find that slower networks with higher sensitivity have greater fidelity and robustness to errors in N and Z, and that increasing the specificity of a model can greatly impact fidelity and robustness.

Possible extensions to our model include incorporating spatial dynamics into the model, and using the model to consider the effect of cell heterogeneity. The intent of the model developed in [8, 20], which was the motivation for this work, was to understand the role of receptor clustering in the activation of T-cells. The work in [20] found that clusters of antigen-bound TCRs are necessary for the recruitment of ZAP70, a molecule that activates downstream signaling in T-cells. Incorporating TCR clustering, and other types of spatial heterogeneity [3, 10], into this model could be an interesting direction for further study.

Biological systems are subject to large amounts of noise that can both help and hurt the system. Developing this model further to study how cell heterogeneity can help a population of T-cells be more robust to changing environmental conditions [11] and antigen variability [13], and how the noisy responses from heterogeneous cells can be aggregated in an effective manner [4], could be promising extensions.

References

[1] Abate, Joseph, Choudhury, Gagan L, and Whitt, Ward. "An introduction to numerical transform inversion and its application to probability models". In: *Computational probability*. Springer, 2000, pp. 257–323.

[2] Altan-Bonnet, Grégoire and Germain, Ronald N. "Modeling T cell antigen discrimination based on feedback control of digital ERK responses". In: *PLoS Biol* 3.11 (2005), e356.

[3] Burroughs, Nigel J and Van Der Merwe, P Anton. "Stochasticity and spatial heterogeneity in T-cell activation". In: *Immunological reviews* 216.1 (2007), pp. 69–80.

[4] Butler, Thomas Charles, Kardar, Mehran, and Chakraborty, Arup K. "Quorum sensing allows T cells to discriminate between self and nonself". In: *Proceedings of the National Academy of Sciences* 110.29 (2013), pp. 11833–11838.

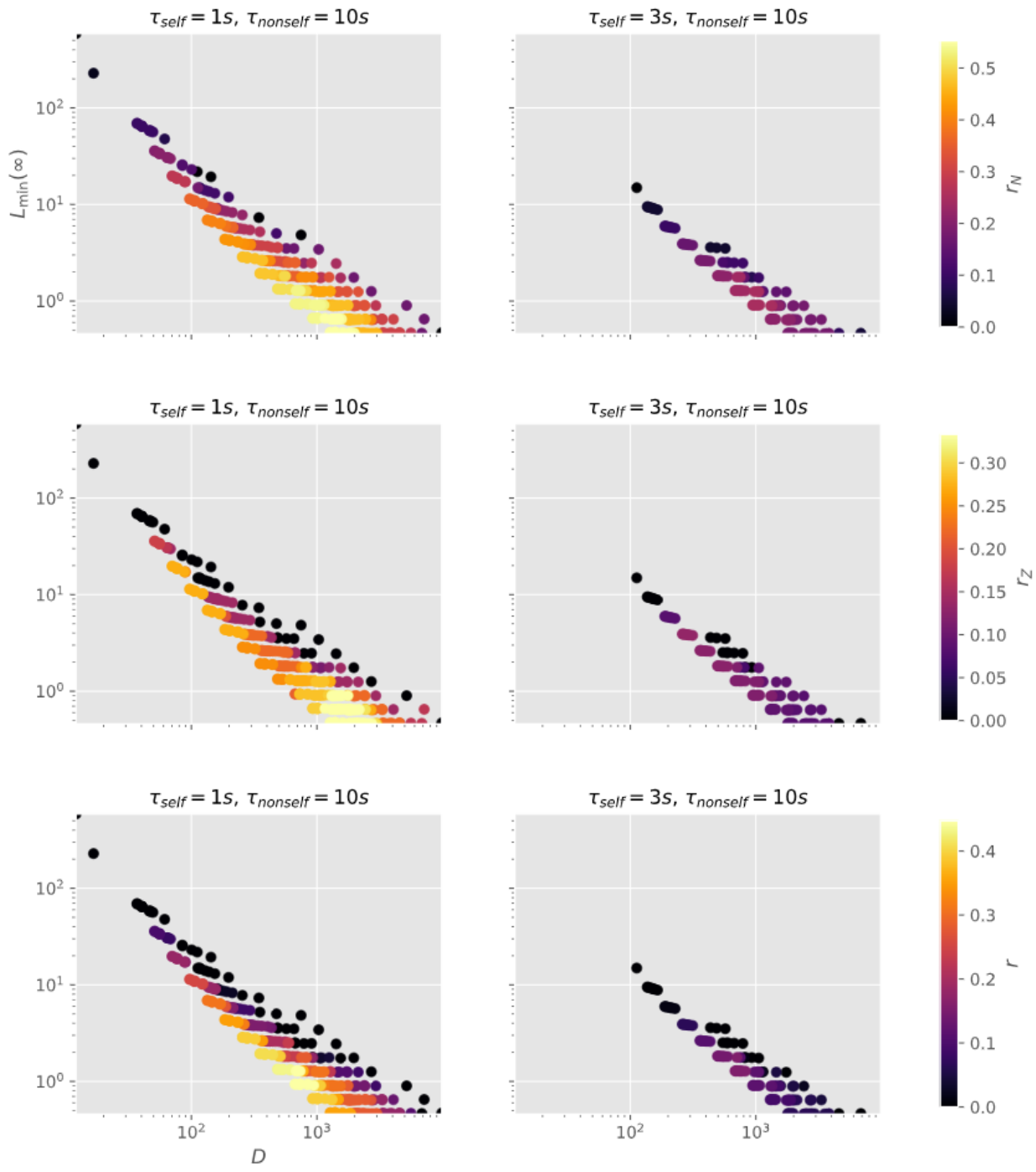

Figure 9: Achievable values of D and $L_{\min}(\infty)$ for fixed τ_{self} and $\tau_{\text{non-self}}$. Points in the top two plots are colored based on r_N. Points in the middle two plots are colored based on r_Z. Points in the bottom two plots are colored based on r_{NZ}. In all three rows, highest robustness among nearby points is used to color the points.

[5] François, Paul and Altan-Bonnet, Grégoire. "The case for absolute ligand discrimination: modeling information processing and decision by immune T cells". In: *Journal of Statistical Physics* 162.5 (2016), pp. 1130–1152.

[6] François, Paul et al. "Phenotypic spandrel: absolute discrimination and ligand antagonism". In: *Physical Biology* 13.6 (2016), p. 066011.

[7] Hopfield, John J. "Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity". In: *Proceedings of the National Academy of Sciences* 71.10 (1974), pp. 4135–4139.

[8] Husain, K. "The Mechanics of Cell Surface Clusters and Their Role in Signalling". PhD thesis. National Center for Biological Sciences, 2018.

[9] Irvine, Darrell J et al. "Direct observation of ligand recognition by T cells". In: *Nature* 419.6909 (2002), pp. 845–849.

[10] Iyengar, Garud and Rao, Madan. "A cellular solution to an information-processing problem". In: *Proceedings of the National Academy of Sciences* 111.34 (2014), pp. 12402–12407.

[11] Kussell, Edo and Leibler, Stanislas. "Phenotypic diversity, population growth, and information in fluctuating environments". In: *Science* 309.5743 (2005), pp. 2075–2078.

[12] Lalanne, Jean-Benoît and François, Paul. "Principles of adaptive sorting revealed by in silico evolution". In: *Physical review letters* 110.21 (2013), p. 218102.

[13] Mayer, Andreas et al. "How a well-adapted immune system is organized". In: *Proceedings of the National Academy of Sciences* 112.19 (2015), pp. 5950–5955.

[14] McKeithan, T. W. "Kinetic proofreading in T-cell receptor signal transduction". In: *Proceedings of the national academy of sciences* 92.11 (1995), pp. 5042–5046.

[15] Ninio, J. "Kinetic amplification of enzyme discrimination". In: *Biochimie* 57.5 (1975), pp. 587–595.

[16] Perelson, A. S. and Weisbuch, G. "Immunology for physicists". In: *Reviews of modern physics* 69.4 (1997), p. 1219.

[17] Privault, Nicolas. "Understanding Markov Chains". In: *Examples and Applications, Publisher Springer-Verlag Singapore* (2013).

[18] Sigman, K. "IEOR 6711: Continuous-Time Markov Chains". `http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-CTMC.pdf`. 2009.

[19] Sorkin, Alexander and Von Zastrow, Mark. "Endocytosis and signalling: intertwining molecular networks". In: *Nature reviews Molecular cell biology* 10.9 (2009), pp. 609–622.

[20] Taylor, M. J. et al. "A DNA-based T cell receptor reveals a role for receptor clustering in ligand discrimination". In: *Cell* 169.1 (2017), pp. 108–119.

A Proofs of results

A.1 Laplace Transform

Lemma 2.1 (Laplace Transform)**.** *Let $\mathcal{L}_{(i,j)} = \mathbb{E}\left[e^{-sT_{(i,j)}}\right]$ denote the Laplace transform of the density function for the first passage time $T_{(i,j)}$ from state (i,j) to the absorbing boundary $\mathcal{B}$. Then*

$$\mathcal{L}_{(i,j)}(s) = \begin{cases} 1 & (i,j) \in B, \\ \frac{1}{s - Q_{(i,j),(i,j)}} \sum_{(k,l)\neq(i,j)} Q_{(i,j),(k,l)} \mathcal{L}_{(k,l)}(s) & (i,j) \notin B. \end{cases} \tag{5}$$

Therefore, $\mathcal{L}_{(0,0)}(s)$ can be computed by solving the above set of equations. The complexity of computing $\mathcal{L}_{(0,0)}$ is $\mathcal{O}(N^2Z^2)$.

Proof. Since $\mathcal{B}$ is absorbing, it follows that $T_{(i,j)} = 0$ for all $(i.j) \in \mathcal{B}$. Therefore, $\mathcal{L}_{(i,j)}(s) = 1$ for all s.

Fix $(i,j) \notin \mathcal{B}$. Then the first-step analysis approach implies that [17]

$$\begin{aligned}
\mathcal{L}_{(i,j)}(s) &= \sum_{(k,l)\neq(i,j)} P_{(i,j),(k,l)} \mathbb{E}\left[e^{-sT_{(i,j)}} \mid \text{jump from state } (i,j) \text{ to state } (k,l)\right] \\
&= \sum_{(k,l)\neq(i,j)} P_{(i,j),(k,l)} \mathbb{E}\left[e^{-s\left(H_{(i,j)}+T_{(k,l)}\right)}\right] = \sum_{(k,l)\neq(i,j)} P_{(i,j),(k,l)} \mathbb{E}\left[e^{-sH_{(i,j)}}\right]\left[e^{-sT_{(k,l)}}\right] \\
&= \frac{1}{s - Q_{(i.j),(i,j)}} \sum_{(k,l)\neq(i,j)} Q_{(i,j),(k,l)} \mathcal{L}_{(k,l)}(s)
\end{aligned}$$

where $P_{(i,j),(k,l)} = \frac{Q_{(i,j),(k,l)}}{\sum_{(k,l)\neq(i,j)} Q_{(i,j),(k,l)}} = -\frac{Q_{(i,j),(k,l)}}{Q_{(i,j),(i,j)}}$ is the probability of transitioning to state $(k,l) \neq (i,j)$ upon leaving (i,j), and $H_{(i,j)}$ is the holding time for state (i,j) and is exponentially distributed with parameter $-Q_{(i,j),(i,j)}$. □

A.2 Vertical Asymptote

Theorem 2.2 (Vertical Asymptote)**.** *The minimum affinity $\tau_{\min}(\infty)$ required to activate in the limit $L \to \infty$ is given by*

$$\tau_{\min}(\infty) = \tau_{\min}^{(B)}(N, Z, D, \alpha) \triangleq \inf\left\{\tau_u > 0 : \mathbb{P}\left(T_0^{(B)}(\tau_u; N, Z) \leq D\right) \geq \alpha\right\}.$$

Proof. This result follows from showing that $\lim_{L\to\infty} \mathbb{P}\left(T_{(0,0)}(\tau_u, L) \leq D\right) = \mathbb{P}\left(T_0^{(B)}(\tau_u) \leq D\right)$. Let (n_t, z_t) be the state of the Markov chain at time t. Since the boundary $\mathcal{B}$ is absorbing,

$$\mathbb{P}\left(T_{(0,0)}(\tau_u, L) \leq D\right) = \mathbb{P}_{(0,0)}\left((n_D, z_D) \in \mathcal{B}\right) = \mathbb{P}_{(0,0)}\left(z_D = Z\right)$$

with $\mathbb{P}_{(n,z)}$ denoting the distribution of (n_t, z_t) given $(n_0, z_0) = (n, z)$. Therefore, it suffices to show $\lim_{L\to\infty} \mathbb{P}_{(0,0)}\left((n_T, z_T) \in \mathcal{B}\right) = \mathbb{P}\left(T_0^{(B)}(\tau_u) \leq D\right)$. We establish this by showing

$$\lim_{L\to\infty} \mathbb{P}_{(0,0)}\left(z_t = z\right) = \lim_{L\to\infty} \mathbb{P}_{(N,0)}\left(z_t = z\right) = \mathbb{P}_0\left(B_t = z\right)$$

The first equality is established in Lemma A.1, and the second equality is established in Lemma A.2. □

The intuition here is that in the limit of $L \to \infty$, all N TCRs are instantaneously bound to antigens, and the activation is determined the dynamics of the antigen bound TCRs acquiring ZAP molecules. In Lemma A.1 we rigorously establish the first result.

Lemma A.1. *For all $t > 0$, $\lim_{L\to\infty} \mathbb{P}_{(0,0)}(z_t = z) = \lim_{L\to\infty} \mathbb{P}_{(N,0)}(z_t = z)$.*

Proof. The rate out of each state (n, z) in the chain is upper bounded by $k_b L + \frac{N}{\tau_u} + \frac{Z}{\tau_u} + N k_{on}$. We define the following four independent Poisson processes.

$$P_t^{(E)} \sim Poisson(k_b L t), \qquad P_t^{(N)} \sim Poisson(N k_{on} t)$$
$$P_t^{(W)} \sim Poisson(\frac{N}{\tau_u} t), \qquad P_t^{(SW)} \sim Poisson(\frac{Z}{\tau_u} t)$$

We think of the arrivals for the process $P_t^{(E)}$ as "eastward" movements from (n, z) to $(n, z+1)$, arrivals for $P_t^{(W)}$ as "westward" movements from (n, z) to $(n-1, z)$, arrivals for $P_t^{(N)}$ as "northward" movements from (n, z) to $(n, z+1)$, and arrivals for $P_t^{(SW)}$ as "southwestward" movements from (n, z) to $(n-1, z-1)$. We can see that

$$\begin{aligned}
\mathbb{P}_{(0,0)}\begin{pmatrix} n_{K(\epsilon)/L}=N, z_{K(\epsilon)/L}=0, \\ n_t=n, z_t=z \end{pmatrix} &= \mathbb{P}_{(0,0)}\left(n_{K(\epsilon)/L} = N, z_{K(\epsilon)/L} = 0\right) \mathbb{P}_{(N,0)}\left(n_{t-K(\epsilon)/L} = n, z_{t-K(\epsilon)/L} = m\right) \\
&\leq \mathbb{P}_{(0,0)}\left(n_t = n, z_t = z\right) \\
&= \mathbb{P}_{(0,0)}\begin{pmatrix} n_{K(\epsilon)/L}=N, z_{K(\epsilon/L)}=0, \\ n_t=n, z_t=z \end{pmatrix} + \mathbb{P}_{(0,0)}\begin{pmatrix} (n_{K(\epsilon)/L}, z_{K(\epsilon/L)})\neq(N,0), \\ n_t=n, z_t=z \end{pmatrix} \\
&\leq \mathbb{P}_{(0,0)}\begin{pmatrix} n_{K/L}=N, z_{K/L}=0, \\ n_t=n, z_t=z \end{pmatrix} + \mathbb{P}_{(0,0)}\left(\left(n_{K/L}, z_{K/L}\right) \neq (N, 0)\right)
\end{aligned}$$

Next, we have

$$\begin{aligned}
\mathbb{P}_{(0,0)}\left(n_{K/L} = N, z_{K/L} = 0\right) &\geq \mathbb{P}\left(P_{K/L}^{(E)} \geq N, P_{K/L}^{(W)} = P_{K/L}^{(N)} = P_{K/L}^{(SW)} = 0\right) \\
&= \mathbb{P}(P_{K/L}^{(E)} \geq N) e^{-\frac{K}{L}\left(\frac{N}{\tau_u} + \frac{Z}{\tau_u} + N k_{on}\right)}
\end{aligned}$$

Since $e^{-x}/x^j \to 0$ as $x \to \infty$ for any finite j, it follows that $\mathbb{P}\left(P_{K/L}^{(E)} \geq N\right) = 1 - e^{-K k_b} \sum_{j=0}^{N-1} \frac{(K k_b)^j}{j!} \to 1$ as $K \to \infty$, and $e^{-\frac{K}{L}\left(\frac{N}{\tau_u} + \frac{Z}{\tau_u} + N k_z\right)} \to 1$ as $L \to \infty$ for any fixed K. Thus, it follows that for all $\epsilon > 0$, there exists some $K = K(\epsilon)$ such that for all $L \geq L(\epsilon)$, $\mathbb{P}(P_{K/L}^{(E)} \geq N) e^{-\frac{K}{L}\left(\frac{N}{\tau_u} + \frac{Z}{\tau_u} + N k_z\right)} \geq 1 - \epsilon$. Hence,

$$\begin{aligned}
(1-\epsilon)\mathbb{P}_{(N,0)}\left(n_{t-K(\epsilon)/L} = n, z_{t-K(\epsilon)/L} = z\right) &\leq \mathbb{P}_{(0,0)}\left(n_{K(\epsilon)/L} = N, z_{K(\epsilon)/L} = 0\right) \mathbb{P}_{(N,0)}\begin{pmatrix} n_t=z, \\ z_t=z \end{pmatrix} \mid \begin{matrix} n_{K(\epsilon)/L}=N, \\ z_{K(\epsilon)/L}=0 \end{matrix}\Big) \\
&= \mathbb{P}_{(0,0)}\left(n_t = n, z_t = z\right) \\
&= \mathbb{P}_{(0,0)}\begin{pmatrix} n_t=n, z_t=z, \\ n_{K(\epsilon)/L}=N, z_{K(\epsilon)/L}=0 \end{pmatrix} + \mathbb{P}_{(0,0)}\begin{pmatrix} n_t=n, z_t=z, \\ (n_{K(\epsilon)/L}, z_{K(\epsilon)/L})\neq(N,0) \end{pmatrix} \\
&\leq \mathbb{P}_{(0,0)}\begin{pmatrix} n_t=n, z_t=z, \\ n_{K(\epsilon)/L}=N, z_{K(\epsilon)/L}=0 \end{pmatrix}) + \mathbb{P}_{(0,0)}((n_{K(\epsilon)/L}, z_{K(\epsilon)/L}) \neq (n, 0)) \\
&\leq \mathbb{P}_{(N,0)}\left(n_{t-K(\epsilon)/L} = n, z_{t-K(\epsilon)/L} = z\right) + \epsilon
\end{aligned}$$

Now, choose any sequence of densities $\{L_k\}$ with $L_k \to \infty$, as well as a sequence $\{\epsilon_k\}$ such that $\epsilon_k \to 0$ and $K(\epsilon_k)/L_k \to 0$. Taking the limit along (L_k, ϵ_k) gives

$$\lim_{k\to\infty} \mathbb{P}_{(N,0)}\left(n_{t-K(\epsilon_k)/L_k} = n, z_{t-K(\epsilon)/L_k} = z\right) = \lim_{k\to\infty} \mathbb{P}_{(0,0)}\left(n_t = n, z_t = z\right)$$

In the limit $k \to \infty$, $K(\epsilon_k)/L_k \to 0$, and the state distribution of a continuous time Markov chain is continuous in time, so

$$\lim_{k\to\infty} \mathbb{P}_{(N,0)}\left(n_{t-K(\epsilon_k)/L_k} = n, z_{t-K(\epsilon_k)/L_k} = z\right) = \lim_{k\to\infty} \mathbb{P}_{(N,0)}\left(n_t = n, z_t = z\right)$$

and we can conclude that $\lim_{L\to\infty} \mathbb{P}_{(0,0)}\left(n_t = n, z_t = z\right) = \lim_{L\to\infty} \mathbb{P}_{(N,0)}\left(n_t = n, z_t = z\right)$. □

Next, we establish that dynamics of the 2D Markov chain in the limit of $L \to \infty$ is given by the 1D birth-death process in the ZAP dimension.

Lemma A.2. $\lim_{L\to\infty} \mathbb{P}_{(N,0)}(z_t = z) = \mathbb{P}_0(B_t = z)$.

Proof. Instead of tracking the aggregate behavior (n_t, z_t) of the chain, we keep track of individual arrivals and departures of TCRs and ZAPs, and show $\lim_{L\to\infty} \mathbb{P}_{(N,0)}(z_t \le z) = \mathbb{P}_0(B_t \le z)$ using the following coupling

(i) Let (a_t, d_t) denote the birth and death processes associated with the birth-death process B_t i.e. $B_t = a_t - d_t$.

(ii) Suppose there is an arrival at time s. Then, increase z_s by 1 with probability n_s/N and keep it the same with probability $1 - n_s/N$.

(iii) Suppose there is a departure at time s, decrease z_s by 1 only if the associated arrival had led to an increase in z_s.

By construction, z_t defined above has the same distribution as the number of ZAPs z_t in the Markov chain. Also by construction, $z_t \le B_t$ on a sample-path basis, and therefore $\mathbb{P}_{(N,0)}(z_t \le z) \ge \mathbb{P}_0(B_t \le z)$, i.e. $\lim_{L\to\infty} \mathbb{P}_{(N,0)}(z_t \le z) \ge \mathbb{P}_0(B_t \le z)$.

Additionally, $B_t - z_t \le A(T_N(t))$ on a sample-path basis, with $T_N(t)$ being the total amount of time the Markov chain spends in states with $n_s < N$ for $s \in (0, t]$, and $A(T_N(t))$ being the number of arrivals in B_t during the periods of time in $(0, t]$ when $n_s < N$. Thus, we have that

$$\begin{aligned}
\mathbb{P}_{(N,0)}\left(z_t \le z\right) &\le \mathbb{P}_0\left(B_t - A\left(T_N(t)\right) \le z\right) \\
&\le \mathbb{P}_0\left(B_t \le A\left(T_N(t)\right) + z \middle| T_N(t) \le u\right) \mathbb{P}_{(N,0)}\left(T_N(t) \le u\right) + \mathbb{P}_{(N,0)}\left(T_N(t) \ge u\right) \quad \forall u > 0 \\
&\le \mathbb{P}_0\left(B_t \le A(u) + z\right) \mathbb{P}_{(N,0)}\left(T_N(t) \le u\right) + \mathbb{P}_{(N,0)}\left(T_N(t) > u\right) \\
&\le \mathbb{P}_0\left(B_t \le A(u) + z\right) + \mathbb{P}_{(N,0)}\left(T_N(t) > u\right)
\end{aligned}$$

where the second-to-last inequality uses the fact that $T_N(t)$ is independent of B_t. Taking limits, we see

$$\begin{aligned}
\lim_{L\to\infty} \mathbb{P}_{(N,0)}\left(z_t \le z\right) &\le \mathbb{P}_0\left(B_t \le A(u) + z\right) + \lim_{L\to\infty} \mathbb{P}_{(N,0)}\left(T_N(t) > u\right), && \forall u > 0 \\
&\le \mathbb{P}_0\left(B_t \le A(u) + z\right), && \forall u > 0 \\
&\le \inf_{u>0} \mathbb{P}_0(B_t \le A(u) + z) \\
&= \mathbb{P}_0\left(B_t \le z\right)
\end{aligned}$$

The first line uses the fact that B_t and $A(u)$ are independent of L. The second inequality comes from the fact $\limsup_{L\to\infty} \mathbb{P}_{(N,\cdot)}\left(T_N(t) > u\right) = 0$ for all $u > 0$, with $\mathbb{P}_{(n,\cdot)}$ being the conditional probability when the chain starts with n TCRs and arbitrary $0 \le z \le n$ at time 0, and this result is established in Lemma A.3. The third inequality comes from the fact that, in general, the CDF of a random variable is monotonically increasing in its argument, and $A(u)$ is monotonically decreasing in a sample-path sense, so $\mathbb{P}_0\left(B_t \le A(u) + z\right)$ is minimized by taking $u \to 0$. The last equality uses the fact CDFs by definition are right-continuous, and $\lim_{u\to\infty} A(u) = 0$ in a sample-path sense. We conclude that $\lim_{L\to\infty} \mathbb{P}_{(N,0)}\left(z_t \le z\right) = \mathbb{P}_0\left(B_t \le z\right)$. □

Lemma A.3. $\limsup_{L\to\infty} \mathbb{P}_{(N,\cdot)}(T_N(t) \ge u) = 0$.

Proof. By Markov's inequality, we know

$$\limsup_{L\to\infty} \mathbb{P}_{(N,\cdot)}\left(T_N(t) \ge u\right) \le \limsup_{L\to\infty} \frac{\mathbb{E}_{(N,\cdot)}\left[T_N(t)\right]}{u}.$$

Next, let $D_N(t)$ be the number of transitions from a state with N TCRs to a state with $N-1$ TCRs in $(0, t]$ and S be the sojourn time from a state with $N-1$ TCRs to a state with N TCRs. By definition, $T_N(t) = \sum_{k=1}^{D_N(t)} \min\{S_k, t\}$ with the S_k's being iid copies of S, so by Wald's equation we have $\mathbb{E}_{(N,\cdot)}\left[T_N(t)\right] =$

$\mathbb{E}_{(N,\cdot)}\left[D_N(t)\right]\mathbb{E}_{(N-1,\cdot)}\left[\min\{S,t\}\right]$. We see that

$$\begin{aligned}
\mathbb{E}_{(N-1,\cdot)}\left[\min\{S,t\}\right] &= \mathbb{E}_{(N-1,\cdot)}\left[\min\{S,t\}\middle|\text{first transition to }(N,z)\right]\mathbb{P}_{(N-1,\cdot)}\left(\text{first transition to }(N,z)\right)\\
&\quad+\mathbb{E}_{(N-1,\cdot)}\left[\min\{S,t\}\middle|\text{first transition not to }(N,z)\right]\mathbb{P}_{(N-1,\cdot)}\left(\text{first transition not to }(N,z)\right)\\
&\le \mathbb{E}_{(N-1,\cdot)}\left[\min\{S,t\}\middle|\text{first transition to }(N,z)\right]+t\,\mathbb{P}_{(N-1,\cdot)}\left(\text{first transition not to }(N,z)\right)\\
&\le \mathbb{E}_{(N-1,\cdot)}\left[S\middle|S\le t,\text{first transition to }(N,z)\right]+t\frac{Nk_{on}+(N+Z)/\tau_u}{k_bL+Nk_{on}+(N+Z)/\tau_u}\\
&\le \frac{1}{k_bL}+t\frac{Nk_{on}+(N+Z)/\tau_u}{k_bL+Nk_{on}+(N+Z)/\tau_u}
\end{aligned}$$

Since the rates for the Poisson processes $P_t^{(W)}$, $P_t^{(SW)}$, and $P_t^{(N)}$ are all finite for all $\tau_u>0$ and independent of L, $\mathbb{E}_{(N,\cdot)}\left[D_N(t)\right]$ is finite for all L. Therefore, we conclude that $\limsup_{L\to\infty}\mathbb{E}_{(N,\cdot)}\left[T_N(t)\right]=0$, and $\limsup_{L\to\infty}\mathbb{P}_{(N,\cdot)}\left(T_N(t)\ge u\right)=0$. □

A.3 Horizontal Asymptote

Theorem 2.3 (Horizontal Asymptote)**.** *The minimum density $L_{\min}(\infty)$ required to activate in the limit $\tau_u\to\infty$ is given by*

$$L_{\min}(\infty)=L_{\min}^{(f)}(N,Z,D,\alpha)\triangleq\inf\left\{L>0:\mathbb{P}\left(T_{(0,0)}^{(f)}(L;N,Z)\le D\right)\ge\alpha\right\}.$$

Proof. We equivalently show that

$$\lim_{\tau_u\to\infty}\mathbb{P}\left(T_{(0,0)}(\tau_u,L)\le D\right)=\mathbb{P}\left(T_{(0,0)}^{(f)}\le D\right)$$

As in the analysis of the vertical asymptote, it suffices to show

$$\lim_{\tau_u\to\infty}\mathbb{P}_{(0,0)}\left(n_t=n,z_t=z\right)=\mathbb{P}_{(0,0)}^{(f)}\left(n_t=n,z_t=z\right)$$

for all $(n,z)\in\mathcal{S}$ and $t>0$, with $\mathbb{P}_{(0,0)}^{(f)}$ being the distribution of the modified Markov chain $M^{(f)}$ given $(n_0,z_0)=(0,0)$. We have

$$\mathbb{P}_{(0,0)}\left(n_t=n,z_t=z\right)=\mathbb{P}_{(0,0)}\left(\begin{smallmatrix}n_t=n,z_t=z,\text{ no backward}\\\text{transitions before time }t\end{smallmatrix}\right)+\mathbb{P}_{(0,0)}\left(\begin{smallmatrix}n_t=n,z_t=z,,\text{ at least one}\\\text{backward transition before }t\end{smallmatrix}\right)$$

We also see that

$$\mathbb{P}_{(0,0)}\left(\text{no backward transitions before time }t\right)=e^{-\frac{t}{\tau_u}(N+Z)}\to 1,$$

as $\tau_u\to\infty$, and

$$\mathbb{P}_{(0,0)}\left(n_t=n,z_t=z\middle|\text{no backward transitions by time }t\right)=\mathbb{P}_{(0,0)}^{(f)}\left(n_t=n,z_t=z\right)$$

Putting these two together tells us that

$$\lim_{\tau_u\to\infty}\mathbb{P}_{(0,0)}\left(\begin{smallmatrix}n_t=n,z_t=z,\\\text{no backward transitions before time }t\end{smallmatrix}\right)=\mathbb{P}_{(0,0)}^{(f)}\left(n_t=n,z_t=z\right).$$

A similar argument establishes that $\lim_{\tau_u\to\infty}\mathbb{P}_{(0,0)}\left(\begin{smallmatrix}n_t=n,z_t=z,\\\text{at least one backward transition before }t\end{smallmatrix}\right)=0$. Therefore, $\lim_{\tau_u\to\infty}\mathbb{P}_{(0,0)}\left(n_t=n,z_t=z\right)=\mathbb{P}_{(0,0)}^{(f)}\left(n_t=n,z_t=z\right)$. □