UNIVERSITY OF CALIFORNIA
RIVERSIDE

An Adaptive Model of Self/Nonself Recognition by Innate Immune Cells

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Yawei Qin

September 2023

Dissertation Committee:

Dr. John Barton, Chairperson
Dr. Roya Zandi
Dr. Thomas Edward Kuhlman

The Dissertation of Yawei Qin is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

Throughout my PhD journey, I have been fortunate to receive invaluable support and guidance from numerous individuals. Their collective influence has played a significant role in shaping me into the person I am today. I have gained profound insights into research and life from their teachings, and I would like to take this opportunity to express my heartfelt gratitude to them all.

First, I would like to thank my advisor Prof. John Barton. John is a great advisor, who has a broad knowledge and always come up with great ideas. When I have some ideas, he encourages me to try to explore them. When I got some results, he always provided great insights into the results. Sometimes when I got some bad results, he always provided useful ideas and guided me in a good direction. In the past few years, John not only taught me how to do research and how to do presentation, but also showed me how to be a good researcher and the attitude towards research and life. Thanks to the help of John, I finished my PhD program and learned a lot. I am really grateful to have John being my PhD advisor.

I would like to thank Dr. Emily Mace, who provided good advice to the research, worked together to discuss the model and simulation result, and wrote paper together. Emily worked on experiments and provided great insight into our quantitative models from the perspective of a researcher conducting an experiment.

I am grateful to Dr. Roya Zandi, who taught me biophysics course and helped me a lot during my PhD program. In ther biophysics class, Roya encouraged us to do presentation on our research and taught us how to do research. Her passion for science is a great example for me.

I would like to thank Dr. Shen-Wen Tsai, Dr. Yanou Cui, Dr. Flip Tanedo, Dr. Hai-bo Yu and Dr Kirill Shtengel, who taught me physics classes. I would like to thank Dr. Thomas Edward Kuhlman, Dr. Heyrim Cho and Dr. Allen Mills for serving on my qualify exam committee and for providing useful feedback on my research.

I would like to thank Dr. Derek Beving. He helped me a lot in my PhD life. We can ask him for help for any problems we meet in our PhD life, including but not limited to registering classes, finding a roommate, and TA/RA assignment. Sometime even he cannot help directly, he will help me find the right person to contact with.

I would like to thank Dr. Yunxiao Li, Zhenchen Hong, Marco Garcia, Kevin Yang, Yirui Gao, Dr. Kai Shimagaki, Dr. Edwin Rodriguez, Brian Lee, and Liz Finney my colleagues in John Barton's lab. We formed a good group and shared research ideas together. We helped each other both in research and in campus life.

I would like to thank Yinan Dong, Dr. Mengyuan Xi, Dr. Nan Hua and Ding Chen. We study and live together for years. We had a great time in the past few years. Also thanks to Dr. Joshua Lui and Dr. Siyu Li, who helped me to learn the balance of research and life. My PhD life will be very different without any one of them.

To my parents for all the support.

ABSTRACT OF THE DISSERTATION

An Adaptive Model of Self/Nonself Recognition by Innate Immune Cells

by

Yawei Qin

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, September 2023
Dr. John Barton, Chairperson

The immune system plays a vital role in protecting our bodies from a wide range of pathogens. As we continue to identify new diseases that challenge our immune system, immunotherapy has gained prominence as a potential medical solution. It utilizes or enhances the body's own immune system to detect and destroy cancer cells. Immunotherapy presents advantages over traditional treatments, although its successful application demands a more extensive understanding of the immune system. While progress has been made, our understanding of the immune system, specifically the innate immune component, is still not comprehensive.

Some key questions that pique my curiosity include: How do innate immune cells distinguish between self and non-self targets? When transplanted to a different host, how do these cells adapt to their new surroundings? Can this process be elucidated through equations or quantitative models? I have always been deeply interested in exploring these questions. In this thesis, we apply statistical methods to tackle these issues. We develop a quantitative model to elucidate the regulatory mechanisms governing the activity of one type

of innate immune cell, namely natural killer cells. Furthermore, inspired by the behavior of natural killer cells, this model holds potential for application to other types of innate immune cells.

We develop a model to simulate the interactions between natural killer (NK) cells and their target cells, illuminating how NK cells learn to identify unhealthy signals from their environment. We apply our model across a range of experimental scenarios, showcasing the algorithm that we have designed to mirror specific experimental settings and delving into the relationship between model and experimental parameters.

Further, we provide evidence for the efficacy of our model by demonstrating that it is possible to identify rational values for model parameters that yield accurate estimations of the experimental data. This effectively validates our approach.

Progressing from this foundation, we present a high-dimensional extension of our model, contributing insights into immune protection at a population level. A key part of our discussion involves the distribution of receptor numbers on natural killer cells and the benefits that this distribution bestows.

Finally, we put our model to the test in the arena of anomaly detection. Through this exploration, we display the versatility and applicability of our mathematical framework, proving its potential to address other complex, real-world problems.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The interdisciplinary collaboration between physicists and biologists has made significant progress in the new century, as evidenced by an increasing number of publications in the field. This partnership leads to a more comprehensive understanding of complex biological systems, as physicists contribute quantitative and computational approaches, while biologists provide domain-specific knowledge and experimental data. And many biological phenomena remain poorly understood, offering physicists opportunities to apply their analytical skills and mathematical models to help decipher these mysteries. Concepts from physics, such as statistical mechanics, non-linear dynamics, and network theory, have already found applications in biology, leading to new insights and research directions. Studying biological systems can inspire the development of novel theoretical frameworks, expanding the horizons of physics itself. Moreover, the development of new experimental techniques and computational tools in both physics and biology has facilitated the study of biological systems at unprecedented levels of detail and complexity. This progress rep-

resents a significant breakthrough in the field. It has created new research opportunities, particularly for physicists with an interest in exploring biological phenomena.

As a physicist, I find the study of immune cells to be an intriguing and rewarding area of research. I am particularly fascinated by the complex interactions and behaviors of these cells and their role in maintaining the delicate balance of our immune system. My interest lies in applying quantitative and analytical skills to better understand the underlying principles governing their functions and responses to various challenges in biological systems.

Exploring the diverse array of receptors expressed by immune cells, such as natural killer cells, and their ability to detect and eliminate foreign or abnormal cells presents a unique opportunity to investigate the adaptability and precision of the immune system. By employing concepts from physics, such as statistical mechanics, non-linear dynamics, and network theory, I aim to develop novel models and frameworks that can shed light on the intricacies of immune cell regulation and communication.

1.1 Motivation

To gain a deeper understanding of the complex interactions and behaviors governing immune system dynamics, we examine natural killer (NK) cells as a representative example. Natural killer cells are a type of lymphocyte, which is a type of white blood cell that can be found in lymphatic fluid. Other examples of lymphocytes include B cells and T cells. However, unlike B cells and T cells, which belong to the adaptive immune system, NK cells are a part of the innate immune system. The innate and adaptive immune systems

are two primary components of the body's overall immune response. They work together to defend the body against harmful pathogens, but they function in fundamentally different ways. The innate immune system, as the name implies, is present at birth and provides the first line of defense against pathogens. It responds in a generic way to all pathogens, without specificity. Key elements of the innate immune system include physical barriers such as the skin, chemical barriers such as stomach acid, and immune cells like neutrophils, macrophages, and natural killer (NK) cells. This system responds rapidly to invaders, often within minutes or hours, but it does not have the capacity to remember past infections.

The adaptive immune system, also known as the acquired immune system, is highly specific and has the ability to remember past infections. This allows for a more rapid and effective response upon re-exposure to the same pathogen, a feature known as immunological memory. The adaptive immune system primarily involves T cells and B cells, which can recognize and respond to specific antigens. This system takes longer to respond than the innate immune system, often multiple days to weeks, but its response is more targeted and effective. The innate immune system provides immediate but general protection against pathogens. The adaptive immune system provides a delayed but highly specific and long-lasting immune response. Both systems are crucial for maintaining the body's health and defense against disease. The primary focus of this dissertation centers around the innate immune system's comprehensive defense mechanism, utilizing NK cells as a representative example to elucidate this intricate system.

NK cells, a key component of the innate immune system, are well-known for their ability to eliminate virus-infected and tumor cells without prior exposure to them. NK

cells accomplish this by expressing a diverse array of activating and inhibitory receptors. These receptors recognize and bind to ligands on the surface of target cells, regulating NK cell activities through various mechanisms. The ability of a cell to identify abnormal cells or microbes that it has never encountered before is remarkable. It is still not completely understood how NK cells utilize the combined signals from various receptors to identify unhealthy/non-self signals. NK cells are also able to adapt to new environments. Experiments [45] show that NK cells will lose the killing ability when moved from a normal environment to a new environment that lacks inhibitory signals and regain the killing ability against targets with less inhibitory ligands when moved back to a normal environment. Generally speaking, NK cells learn and adapt to the environment. How NK cells learn the environment and respond to environmental changes is a major topic of research.

1.2 Thesis overview

Our background in physics equips us with unique perspectives and analytical tools that can be used to study the immune system from a different angle compared to traditional biological approaches. Rather than concentrating solely on the intracellular level, we broaden our focus to include the intercellular and population levels.

To explain the adaption of NK cells, we developed a quantitative model to simulate NK cell processing of cellular signals and learning the environment. Since NK cells kill tumor cells without previous activation, we have an idea that NK cells could "learn" a healthy signal distribution from interacting with normal cells, and therefore identify abnormal signals as non-self/unhealthy ones. Here we build a Bayesian inference model to

simulate the process of NK cells learning the normal signal distribution. For simplicity, we assume the healthy signal received by receptors satisfies a Gaussian distribution. We use a Bayesian inference approach to model how NK cells learn the normal signal distribution. In particular, we model estimating the mean and variance of an unknown signal distribution. Due to the stochastic nature of cell-cell interactions and receptor-ligand combinations, the shift in estimated mean and variance can be described by stochastic differential equations. From statistical mechanics, we applied the Fokker-Plank equation to get the time evolution of the probability density functions of the parameters. We also built stochastic models to simulate outcomes of experiments and predict collective behaviors, which can be compared with experimental results. There are tens of immune-related receptors on the surface of immune cells, making the space of receptor-ligand interactions high-dimensional. We use statistical methods to simulate the receptor expression distributions and signal processing to extend our model to high dimensions.

In our model, NK cells learn the typical distribution of ligands on *healthy* cells through repeated encounters, allowing them to respond to significant deviations from typical ligand expression in rare, unhealthy targets. Our study complements mechanistic studies of immunity and signaling [17, 71, 104, 37, 73] by focusing on the principles underlying immune regulation. We show that our model is consistent with known NK cell behaviors and provides a good quantitative fit to data from three separate experiments in NK cells. Our results point to a tradeoff between the precision of immune responses and the ability to adapt to different environments. We also extend our model to consider multiple pairs of receptors and ligands. Both NK cells and macrophages use a wide array of re-

ceptors [55, 95]. Intriguingly, recent studies have revealed dramatic heterogeneity in the complement of receptors that individual NK cells express [43, 90]. The biological function of this heterogeneity, however, has remained unclear. Here we show that populations of NK cells with 'sparse' patterns of receptor expression, like those observed in experiments, are better able to respond to a variety of different targets than ones with 'dense' receptor expression profiles. Sparse receptor expression is especially important for separating a signal of aberrant ligand expression due to infection or transformation from noise due to background fluctuations in ligand concentrations. This suggests that heterogeneity in the NK cell repertoire may improve the immune system's ability to respond to multiple threats.

This model can also be applied on other field like outlier detection. We will show that this simple model have a comparable performance with some commercially used/open source anomaly detection algorithms.

1.3 Connections to statistical physics

Statistical physics is a branch of physics that employs statistical methods and mathematical tools to study the behavior of large ensembles of particles. It aims to describe the collective behavior of complex systems by connecting macroscopic properties, such as temperature and pressure, to the underlying microscopic interactions between individual particles. Central to statistical physics is the concept of equilibrium and the use of statistical mechanics, which allows the calculation of macroscopic quantities by averaging over the microscopic states of a system. This field, with its far-reaching implications, has successfully discovered applications in an array of disciplines, such as condensed mat-

ter physics, materials science, and surprisingly, even biological systems. By facilitating a more profound understanding of the emergent properties characteristic of complex systems, it plays a pivotal role in the demystification and comprehensive portrayal of intricate concepts. In doing so, it significantly contributes to the discourse within these scientific realms, enhancing our overall understanding and mastery over these complex phenomena.

Here, statistical physics has been applied to describe collective behaviors of biological systems at multiple levels.

At the cellular level, we study the interactions between an NK cell and a target cell, which involve random processes such as receptor-ligand binding. Instead of focusing on individual receptor-ligand pairs or single NK cell-target cell interactions, we consider the overall signal from all receptor-ligand connections and the behavior of NK cells in a specific environment. To achieve this, we develop a quantitative model that emphasizes statistical results, using probability distributions to describe the signal distribution for an NK cell interacting with a particular type of target cells. Bayesian inference and stochastic differential equations are employed to simulate the learning process and estimate parameter changes.

At the population level, we treat each NK cell individually and examine their collective behaviors. In our stochastic model, even NK cells with identical receptor distributions may respond differently to the same target in the future due to varying interaction paths. The focus is on the statistical results for the behavior of a group of NK cells, utilizing the Fokker-Planck equation to describe the evolution of probability distributions of parameters.

In high-dimensional systems, receptor expression exhibits randomness. We analyze the receptor distribution on each NK cell rather than the expression of individual receptors. Probability distributions are used to describe receptor distributions and compare the behaviors of groups of NK cells with varying receptor distributions.

By comparing our results with experimental data, we can establish connections between our model parameters and real biological information. Our statistical model allows us to link the macroscopic properties of biological systems to their underlying microscopic interactions, such as individual cell behavior or receptor-ligand binding, enhancing our understanding of complex biological activities.

Chapter 2

Mathematical & Statistical Methods

In this chapter, we introduce several statistical models to analyze complex biological systems. First, we discuss Bayesian inference, demonstrating how to estimate distributions of parameters based on data. Next, we explore stochastic differential equations, which provide a continuous approximation for parameter updates. We will discuss the roles of prior knowledge and new measurements in the drift and diffusion of parameter updates. Finally, we introduce the Fokker-Planck Equation, which describes the evolution of probability distributions for parameters over time.

2.1 Bayesian Inference

The main idea of our model is that NK cells learn from normal environments, knowing what are healthy cells and thus being able to identify the abnormal targets whose

information is different from what NK cells learned. The first step is learning the normal environment. Since it's impossible for each NK cell to interact with every healthy cell to learn the normal environment, there must be a process that NK cells are able to estimate the normal environment from a finite number of interactions. This is the statistical inference where we assume the finite interactions are a sample from a population, and we use statistical analysis results from the sample data to estimate the properties of the population. This is how NK cells learn the normal environment. Here we use Bayesian inference to simulate the learning process.

Bayesian inference is a statistical method that is used in estimating model parameters. It updates the probability of model parameters as more information becomes available. For example, a (potentially) biased coin has an unknown probability p to be heads and $1-p$ to be tails. If we throw the coin 10 times and we get 4 heads and 6 tails, we may estimate $p = 0.4$. While if we throw the coin more times and we get 500 heads and 500 tails in total, we may update the estimation of p. We can use Bayesian inference to estimate the probability distribution of parameter p as more and more data becomes available. One advantage of Bayesian inference is that it gives a probability distribution of parameters that need to be determined. The variance of the probability distribution, to some extent, can provide some information about how uncertain the estimation of the parameter is. Another advantage of Bayesian inference is that we do not need any evidence for the initial guess of the probability distributions of parameters. As more and more data are available, the probabilities of parameter settings that are more likely to generate the data will increase automatically. However, if we have prior knowledge of the parameter distribution, we can

apply previous research results on the prior distribution which might help to estimate the parameters since more information is used, compared to using the new measured data only. If the model has many parameters, the computation of the updated probability distribution is usually complex. However, such problems can often be solved numerically, which has lead to the broad application of Bayesian inference methods in many areas. Next, we will show how to do calculations of Bayesian inference, from basic Bayes' theorem to how NK cells estimate the environment.

2.1.1 Introduction to Bayesian inference

Bayesian inference is based on Bayes' theorem which was developed in the 18th century. Before diving into Bayesian inference, here I will first give a short introduction to Bayes' theorem. Bayes' theorem describes the relation between two conditional probabilities. It is used to calculate the probability of one event occurring given the condition that another event occurred.

The equation for Bayes' theorem is simply

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where A and B are events. $P(A)$ and $P(B)$ are probabilities of events A and B occurring. $P(A|B)$ is the conditional probability that event A occurs given that B has occurred. $P(B|A)$ is the conditional probability of event B occurring given that A has occurred.

Bayesian inference uses Bayes' theorem to update the probability distribution of model parameters. For example, B is the observation/new data and A is the model parameters, which are related to the data. $P(A)$ is named as the prior probability, which

is determined before new data is measured. $P(B|A)$ is named as likelihood, which is the probability that new data B is observed, assuming that it was generated from a model with parameters A. $P(B)$ is the marginal likelihood, which is usually unknown. Since it does not affect the determination of the relative probabilities of parameters, we can rewrite the above equation as

$$P(A|B) \propto P(B|A)P(A)$$

Thus even $P(B)$ is usually unknown, we can still get $P(A|B)$ by normalizing $P(B|A)P(A)$.

2.1.2 Bayesian inference for a Gaussian distribution with unknown mean and precision

To simulate NK cells learning the environment, we need to firstly quantify what is the environment and what is learned by NK cells. The learning process is done by interacting with normal cells. During the interaction, activating receptors on NK cell bind to their corresponding activating ligands on normal cells and inhibitory receptors on NK cells bind to their corresponding inhibitory ligands on normal cells.

We use a parameter $\mathbf{x}$ to denote the signal from those receptors during binding. To simplify, here we build a model which uses a real number to describe the signal from a receptor-ligand binding. The value of the number describes how activating the signal is: the larger the value is, the more activating the signal is. A positive number is activating signal and negative value is inhibitory signal. The total signal is just the sum of signals from all receptors that bound to ligands, which is a real number x. In this way, the 'environment' is the distribution of x when an NK cell interacts with all cells in a population. Since it's

impossible for NK cells to interact with all cells in the population, we assume NK cells learn the probability distribution of total signal x with finite interactions with targets in a sample. Since the ligands expressed on a cell should be similar to other cells in the same population, signals from a population of cells should be concentrated. Here we use Gaussian distribution to model the signal distribution of a population which is determined by two parameters, mean and variance. We use Bayesian inference to simulate a NK cell learning the two parameters.

In the Bayesian inference, we use the normal distribution as the likelihood since we assume the signals of a population floow a Gaussian distribution. It can be written as $P(x|\mu,\lambda) \sim N(\mu,\lambda)$, where x is the total signal observed during an interaction, μ is the mean and λ is the precision which is the inverse of variance, $\lambda = 1/\sigma^2$. For the prior distribution, which is an initial guess of the probability distribution of parameters (μ,λ), we use a prior function in the form of normal-gamma distribution since it is the conjugate prior of a normal distribution with unknown mean and precision. Here conjugate prior means that the prior distribution has the same functional form as the posterior. In other words, the posterior will be a normal-gamma distribution as well, which will be shown in the following calculation. Using conjugate prior greatly simplifies the calculation. Not only can we write the posterior distribution analytically, but also in the future when new data is available and we use it as the prior distribution for the new update, the calculation will follow the same form. A normal-gamma function has the form of

$$P(\mu,\lambda|m,\beta,\alpha,\kappa) = \frac{\beta^\alpha\sqrt{\kappa}}{\Gamma(\alpha)\sqrt{2\pi}}\lambda^{\alpha-1/2}e^{\lambda\left[-\beta-\kappa\frac{(\mu-m)^2}{2}\right]},$$

where (m,β,α,κ) are four parameter that determine the distribution and $\beta > 0$, $\alpha > 0$

and $\kappa > 0$. For a given λ, the conditional distribution of μ is a normal distribution, $N(\mu|\lambda) = N(m, \frac{1}{\beta\kappa})$. For a given μ, the conditional distribution of λ is a Gamma function, $Gamma(\alpha, \beta)$. We will see m is the mean of μ, and the ratio of α and β, $\frac{\alpha}{\beta}$, determines the mean of precision, λ. In the subsequent calculations, we will demonstrate that the estimated mean and variance (their mean values) are determined by m and β, respectively. The other two parameters, κ and α, primarily influence the weight of the data in determining the values of m and β. In our modified Bayesian inference model, we refer to κ and α as memory parameters, as they determine the number of recent data points used in calculating the estimated mean and variance. In other words, these parameters dictate how many past encounters are retained in the memory of an NK cell.

Given a normal-gamma prior distribution for the Gaussian mean μ and precision λ, the posterior distribution for μ and λ after taking a new measurement x is

$$\begin{aligned} P_{\text{post}}(\mu, \lambda|x) &\propto P(x|\mu, \lambda) \times P_{\text{prior}}(\mu, \lambda|m, \beta, \alpha, \kappa) \\ &= \lambda^{\alpha} e^{\lambda\left[-\beta - \frac{1}{2}\frac{\kappa}{\kappa+1}(x-m)^2 - (\kappa+1)\frac{\left(\mu - \frac{\kappa m + x}{\kappa+1}\right)^2}{2}\right]}, \end{aligned}$$

where

$$P(x|\mu, \lambda) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\lambda(x-\mu)^2}{2}}$$

is the likelihood function, where $\lambda = 1/\sigma^2$ is the precision, and

$$P_{\text{prior}}(\mu, \lambda|m, \beta, \alpha, \kappa) = \frac{\beta^{\alpha}\sqrt{\kappa}}{\Gamma(\alpha)\sqrt{2\pi}} \lambda^{\alpha - 1/2} e^{\lambda\left[-\beta - \kappa\frac{(\mu-m)^2}{2}\right]}$$

is the normal-gamma prior. The posterior thus also follows a normal-gamma distribution with modified parameters m', β', α', and κ', which are related to those of the prior distri-

bution by

$$\begin{aligned} m' &= \frac{\kappa m + x}{\kappa + 1} & \kappa' &= \kappa + 1 \\ \beta' &= \beta + \frac{\kappa}{\kappa + 1}\frac{(x-m)^2}{2} & \alpha' &= \alpha + \frac{1}{2}\,. \end{aligned} \tag{2.1}$$

2.1.3 Estimating the signal mean and variance

Estimated mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ are the expected values of μ and $\sigma^2 = \frac{1}{\lambda}$ for the distribution $P(\mu, \lambda)$. Prior to the measurement, $\hat{\mu}$ and $\hat{\sigma}^2$ are

$$\begin{aligned} \hat{\mu} &= \iint d\mu\, d\lambda\; \mu\, P(\mu, \lambda | m, \beta, \alpha, \kappa) = m\,, \\ \hat{\sigma}^2 &= \left\langle \frac{1}{\lambda} \right\rangle = \iint d\mu\, d\lambda\; \frac{1}{\lambda} P(\mu, \lambda | m, \beta, \alpha, \kappa) = \frac{\beta}{\alpha - 1}\,. \end{aligned} \tag{2.2}$$

After receiving a signal x, the new estimated mean and variance are

$$\begin{aligned} \hat{\mu}' = m' &= \frac{\kappa m + x}{\kappa + 1} = \frac{\kappa \hat{\mu} + x}{\kappa + 1}\,, \\ \left(\hat{\sigma}^2\right)' &= \frac{\beta'}{\alpha' - 1} = \frac{\beta + \frac{1}{2}\frac{\kappa (x-m)^2}{(\kappa+1)}}{\alpha - \frac{1}{2}} = \frac{\hat{\sigma}^2 (\alpha - 1) + \frac{1}{2}\frac{\kappa (x-m)^2}{(\kappa+1)}}{\alpha - \frac{1}{2}}\,. \end{aligned} \tag{2.3}$$

2.2 Stochastic differential equations

The update of the parameters in our model can be described as a Markov process, which means that the future state of the process depends only on the present state and not on its past history. The value of a parameter at a given time depends solely on its value at the immediately preceding time step. Additionally, we assume that the measured signal during an interaction between an NK cell and a target cell follows a fixed distribution, which implies that the signals are independent of one another in different steps. This assumption simplifies the analysis and makes it easier to develop and apply mathematical models, such as stochastic differential equations and Fokker-Planck equations, to describe the evolution

of the parameters and their joint probability distributions. In the following sections, we will provide an introduction to these equations.

2.2.1 Brownian motion

A stochastic differential equation (SDE) is a differential equation that includes terms associated with random processes. It was first developed in the context of Brownian motion, a physical phenomenon characterized by the random movement of particles suspended in a fluid. Albert Einstein and Marian Smoluchowski were pioneers in this field, utilizing linear stochastic differential equations to describe the motion of an object subjected to random forces.

The Wiener process is a mathematical model of Brownian motion, representing the random movement of particles suspended in a fluid. Suppose we have a particle suspended in a fluid. If the mass of the small particle is comparable to the mass of the molecules, the particle will be bombarded from all sides by the fluid molecules, and these bombardments will cause the particle to move in a random manner. We say there is a random force on the small particle that determines the movement. In an ideal scenario, we would solve the interconnected motion equations for all the fluid molecules and the small particle, in which case, no stochastic force would manifest. However, considering the astronomical number of molecules in the fluid — on the order of 10^{23} — we are usually unable to solve these intertwined equations. Additionally, given our lack of knowledge about the initial states of all the fluid molecules, we cannot precisely compute the small particle's motion in the fluid. If we were to use another identical system (particle and fluid), but with different initial conditions for the fluid, the motion of the small particle would vary.

Following standard thermodynamic practice, we consider a collection of such systems, known as a Gibbs ensemble [36]. In this context, the force would vary across systems, and our best recourse would be to consider ensemble averages of this force [87]. And the force is characterized as a stochastic or random force, the properties of which are discernible only through averaging.

The equation of particle motion is:

$$\dot{v} = \Gamma(t)\,,$$

where v is the velocity of particle, and $\Gamma(t)$ is the stochastic force. The stochastic force exhibits certain properties. Firstly, we posit that its ensemble average equals zero, denoted as $\langle \Gamma(t) \rangle = 0$. This implies that there's no systematic bias in the force; it's equally likely to push in any direction. Moreover, when we consider the product of two stochastic forces at different times, we assume that their average value equals zero for time differences $t\prime - t$ exceeding the collision duration time τ_0,

$$\left\langle \Gamma(t')\Gamma(t) \right\rangle = 0 \quad for \quad |t\prime - t| > \tau_0\,.$$

In other words, the effects of individual collisions are assumed to be uncorrelated if they are separated by more than the characteristic collision time. This reflects the memoryless nature of the stochastic force.

2.2.2 Langevin equation

The Wiener process serves as a cornerstone for exploring more intricate stochastic processes. Among these is the Langevin equation, developed to characterize the movement

of a particle subjected to both deterministic and random forces in a viscous medium. The Langevin equation expands upon the Wiener process by incorporating deterministic aspects such as friction and external impetus, alongside the random forces. This endows the Langevin equation with a broader scope and flexibility, making it an instrumental tool for simulating various physical and biological systems that exhibit stochastic dynamics.

A typical instance of a one-dimensional stochastic differential equation is the simple Langevin equation. This equation effectively encapsulates the behavior of a system under the dual influence of deterministic and random forces, as outlined in Gardiner's "Stochastic Methods" [32]. The general form of the Langevin equation is:

$$\frac{dx}{dt} = a(x,t) + b(x,t)\xi(t)\,,$$

where x is the parameter that changing with time, in the example of particle's Brownian motion, x can be the velocity of particle. $a(x,t)$ and $b(x,t)$ are two functions. $\xi(t)$ is the random term. The random term plays an important role in the calculation of SDE. Usually we assume it is a simple stochastic process, where for $t \neq t'$, $\xi(t)$ and $\xi(t')$ are independent with each other. Also we can set $\langle \xi(t) \rangle = 0$, where if it is not, we can move the non zero term to $a(x,t)$. Thus we have $\langle \xi(t)\xi(t') \rangle = \delta(t-t')$.

The integrated variable $W(t)$:

$$W(t) = \int_0^t dt'\, \xi(t')\,,$$

is a smooth function. The derivation of this equation can be found in Gardiner's "Stochastic Methods" [32]. Thus, we have $\xi(t)dt = dW(t)$. Since $\xi(t)$ and $\xi(t')$ are independent, thus $W(t)$ is a Markov process and $W(t)$ and $W(t') - W(t)$ are independent for $t' > t$. It has

been shown that $dW(t)$ possesses some unique properties. Two key properties are

$$dW(t)^2 = dt\,, \quad dW(t)^{2+N} = 0\,.$$

For a detailed explanation, please see Gardiner's "Stochastic Methods" [32].

Thus the corresponding stochastic integral equation is

$$x(t) - x(t_0) = \int_{t_0}^{t} dt'\, a[x(t'), t'] + \int_{t_0}^{t} dW(t')\, b[x(t'), t']\,.$$

And we can get a discretised version of the SDE by taking discrete time point, the equation can be written as

$$x_{i+1} = x_i + a[x_i, t_i]\Delta t_i + b[x_i, t_i]\Delta W_i\,.$$

2.2.3 Change of function of $x(t)$

For an arbitrary function of $x(t)$, the stochastic differential equation can be calculate using Taylor expansion, we expand $df[x(t)]$ to second order in $dW(t)$:

$$\begin{aligned} df[x(t)] &= f[x(t) + dx(t)] - f[x(t)] \\ &= f'[x(t)]dx(t) + \frac{1}{2}f''[x(t)]dx(t)^2 + ... \\ &= f'[x(t)]\left\{a[x(t), t]dt + b[x(t), t]dW(t)\right\} + \frac{1}{2}f''[x(t)]b[x(t), t]^2[dW(t)]^2 + ...\end{aligned}$$

Ignoring higher order terms, the equation can be simplified as

$$df[x(t)] = \{a[x(t), t]f'[x(t)] + \frac{1}{2}b[x(t), t]^2 f''[x(t)]\}dt + b[x(t), t]f'[x(t)]dW(t)\,, \qquad (2.4)$$

where $dW(t)^2 = dt$ from above is used.

2.2.4 The evolution of probability distribution of $x(t)$

From Eq. 2.4, we have the time evolution of an arbitrary $f(x(t))$:

$$\begin{aligned}\frac{\langle df[x(t)]\rangle}{dt} &= \left\langle \frac{df[x(t)]}{dt} \right\rangle = \frac{d}{dt}\langle f[x(t)]\rangle \\ &= \left\langle a[x(t),t]\partial_x f + \frac{1}{2}b[x(t),t]^2\partial_x^2 f \right\rangle .\end{aligned}$$

Since $x(t)$ has a probability distribution $p(x,t)$, thus we have

$$\begin{aligned}\frac{d}{dt}\langle f[x(t)]\rangle &= \int dx\, f(x)\partial_t p(x,t) \\ &= \int dx\, a[x(t),t]\partial_x f + \frac{1}{2}b[x(t),t]^2\partial_x^2 f p(x,t) .\end{aligned}$$

After applying the integration by parts method and disregarding the boundary terms, we obtain the following:

$$\int dx\, f(x)\partial_t p(x,t) = \int dx\, f(x)\{-\partial_x[a(x,t)p] + \frac{1}{2}\partial_x^2[b(x,t)^2 p]\} .$$

Since $f(x)$ is arbitrary, we have

$$\partial_t p(x,t) = -\partial_x[a(x,t)p(x,t)] + \frac{1}{2}\partial_x^2[b(x,t)^2 p(x,t)] .$$

This equation is recognizably in the form of the Fokker-Planck equation, which is used to describe the time evolution of a probability distribution under the influence of deterministic and random forces.

2.2.5 Multivariable system

In general, stochastic differential equations for a n variables system can be defined by

$$d\mathbf{x} = \mathbf{A}(\mathbf{x},t)dt + \mathbf{B}(\mathbf{x},t)d\mathbf{W}(t) ,$$

where x is the n variables, $d\mathbf{W}(t)$ is an n variable Wiener process. The probability density $p(\mathbf{x}, t)$ satisfies:

$$\partial_t p(\mathbf{x}, t) = -\sum_i \partial_i [A_i(\mathbf{x}, t)p] + \frac{1}{2}\sum_{i,j} \partial_i \partial_j \{[\mathbf{B}(\mathbf{x}, t)\mathbf{B}^T(\mathbf{x}, t)]_{ij} p\} .$$

2.3 Fokker-Planck Equation

The Fokker-Planck equation plays a pivotal role in the investigation of stochastic processes, as it delineates the temporal progression of a probability distribution. This equation serves as a robust mechanism for acquiring distribution functions. After the resolution of these functions, the averages of macroscopic variables can be deduced via integration. Notably, the Fokker-Planck equation's utility transcends systems in close proximity to thermal equilibrium; it is equally valuable in the analysis of systems that are significantly removed from thermal equilibrium.

2.3.1 Fokker-Planck equation for one variable

For a one-dimensional variable x with a distribution function $W(x, t)$, the general form of the Fokker-Planck equation is given as[86]:

$$\frac{\partial W}{\partial t} = [-\frac{\partial}{\partial x} D^{(1)}(x) + \frac{\partial^2}{\partial x^2} D^2(x)]W . \tag{2.5}$$

Here, $D^2(x) > 0$ is referred to as the diffusion coefficient and $D^{(1)}(x)$ is known as the drift coefficient. Both coefficients may depend on time. The Fokker-Planck equation is mathematically a linear second-order partial differential equation of parabolic type. Essentially, it is a diffusion equation with an additional first-order derivative term with respect to x. In

the mathematical literature, equation 2.5 is also often referred to as a forward Kolmogorov equation.

2.3.2 Fokker-Planck equation for N variables

The Fokker-Planck equation can be generalized to accommodate multiple variables. Specifically, for N variables $x_1, ..., x_N$, the equation can be extended as follows:

$$\frac{\partial W}{\partial t} = [-\sum_{i=1}^{N} \frac{\partial}{\partial x_i} D_i^{(1)}(\mathbf{x}) + \sum_{i,j=1}^{N} \frac{\partial^2}{\partial\, x_i x_j} D_{ij}^2(\mathbf{x})]W\,. \tag{2.6}$$

The drift vector $D_i^{(1)}$ and the diffusion tensor D_{ij}^2 generally depend on the N variables $\mathbf{x} = (x_1, \ldots, x_N)$. The Fokker-Planck equation, equation 2.6, describes the time evolution of the distribution function $W(\mathbf{x}, t)$ for N macroscopic variables $\mathbf{x}$.

2.3.3 Solution of Fokker-Planck equation

The Fokker-Planck equation, a cornerstone in the field of stochastic processes, can be analytically solved in certain specific scenarios. These special cases often involve conditions such as a linear drift vector and a constant diffusion tensor. Another scenario where solutions are attainable is when the drift vector and the diffusion matrix conform to distinct potential conditions. These conditions allow for the simplification of the equation, making an analytical solution feasible. However, these situations are limited and don't encompass all the complexities that can be encountered in the study of stochastic processes.

In the general case, obtaining solutions for the Fokker-Planck equation can be quite challenging. One of the factors that can contribute to this complexity is the inability to separate variables, a common technique used to simplify multi-variable differential equa-

tions. In addition, problems can arise when the number of variables in the equation is large. These conditions often necessitate the use of more advanced mathematical techniques or computational methods.

To overcome these complexities, researchers use a variety of solution strategies. Simulation methods, for instance, provide an effective means of estimating solutions. These methods typically involve creating a computational model of the system described by the equation, and then running simulations to observe the evolution of the system over time. This can provide valuable insights into the behavior of the system and can help generate approximate solutions to the equation.

A more theoretical approach involves transforming the Fokker-Planck equation into a Schrödinger equation, a key equation in quantum mechanics. This transformation, often referred to as the "Feynman-Kac formula", allows for the use of techniques developed in quantum mechanics to solve the Fokker-Planck equation. This method can be particularly useful when dealing with potential energy landscapes, which are common in many physical and chemical systems.

Lastly, numerical integration methods serve as a powerful tool in deriving solutions to the Fokker-Planck equation. These methods, which include techniques like finite-difference methods and finite-element methods, involve discretizing the equation and then solving the resulting system of equations. Even though these methodologies may require substantial computational power and can be resource-intensive, their importance cannot be underestimated. In many situations, especially those where traditional analytic solutions are out of reach or simply do not exist, they often emerge as the only feasible, and in-

deed indispensable, approach to successfully obtaining the needed solutions. Thus, despite their computational demands, these methods hold the key to unlocking answers in complex scenarios.

To sum up, deriving analytic solutions for the Fokker-Planck equation is feasible under specific conditions. However, the equation's complexity frequently calls for the application of sophisticated mathematical strategies or computational methodologies. In this dissertation, we've chosen to employ numerical integration methods to solve the Fokker-Planck equation. For a detailed explanation of our process and findings, kindly refer to the subsequent chapter.

Chapter 3

Modified Bayesian inference model for simulating cell-cell interactions

3.1 Biology background

The human immune system consists of both the adaptive and innate immune systems. The adaptive immune system, comprised of T and B cells, binds specifically to antigens. These antigens either originate from pathogens or, in the context of cancer, are altered versions of self proteins [44]. There's a selective process in place for T cells and B cells, which eliminates those that respond to self cells, preventing them from maturing. This process ensures that these immune cells specifically recognize and respond to signals that aren't derived from normal, healthy host cells.

The immune system also employs defense mechanisms that don't rely on the specific identification of foreign substances. Immune cells like Natural Killer (NK) cells and

macrophages discern their targets through a vast array of activating and inhibitory receptors, many of which bind to self-ligands. NK cells' activating receptors mainly bind to surface ligands, which are triggered during infection, transformation, or stress [34]. On the other hand, their inhibitory receptors largely bind to ligands expressed by healthy cells, thereby preventing the inadvertent destruction of healthy self cells. There are numerous types of inhibitory ligands, with major histocompatibility complex (MHC) class I molecules being a prominent example [55]. Macrophages function similarly. Many of their activating receptors respond to self-ligands indicative of cellular damage. However, the activity of macrophages is inhibited by ubiquitously expressed surface proteins such as CD47 [95]. This strategic array of receptors allows these immune cells to effectively respond to abnormal or potentially harmful conditions, while also avoiding unnecessary damage to healthy cells.

For both NK cells and macrophages, the integration of activating and inhibitory signals is a complex, context-dependent process. NK cells were originally noted for their ability to kill target cells that express low or undetectable levels of MHC class I [47]. However, subsequent studies showed that NK cells from mice and humans that express low levels of MHC class I are self-tolerant, though they also respond less vigorously to typical target cells [58, 42, 7, 19, 113]. Patients with defects in the transporter associated to antigen processing (TAP) protein provide one such example. In these individuals, the ability of TAP to load MHC class I proteins with peptides is reduced, causing little MHC class I to reach the cell surface [19]. Macrophages from CD47-deficient hosts are also tolerant of cells with low levels of CD47 expression, which would usually be phagocytosed [103]. These observations point towards an adaptive process through which the threshold for activation

is controlled by the local environment, which in NK cells is referred to as 'education' [50, 4]. This process must be extremely robust, as autoimmunity due to direct attack by NK cells and macrophages on healthy tissues is essentially unknown [75, 64].

How do these immune cells achieve self-tolerance while retaining the ability to respond to unhealthy targets? Qualitative models of education have been developed to explain NK cell behaviors [112, 84, 10], but no overarching quantitative framework currently exists. Most prior quantitative work on self/nonself recognition has focused primarily on adaptive immunity and the specific recognition of nonself [77, 78, 21, 88, 51], rather than the nonspecific recognition of 'missing' or 'altered' self detected by the innate immune system.

3.2 Modified Bayesian inference model

Innate immune cells combine the signals from activating and inhibitory receptors to make a decision whether the target signal is healthy or not. The combination process of signals from many receptors might be complex. Here we represent the net signal that an individual immune cell receives from a target cell with a single variable x. Larger values of x represent more activating signals, while smaller values represent more inhibitory signals.

To prevent autoimmunity, immune cell responses must be tuned to avoid activation against normal, healthy cells. We propose that this can be achieved by the immune cell through the construction of an internal representation $P_r(x)$ of the signal distribution (**Fig. 3.1**). Under normal conditions, the great majority of targets that an immune cell encounters should be healthy. Signals that are substantially more activating than those from typical cells, such that $P_r(x)$ is very small, are likely to originate from outlier cells that may

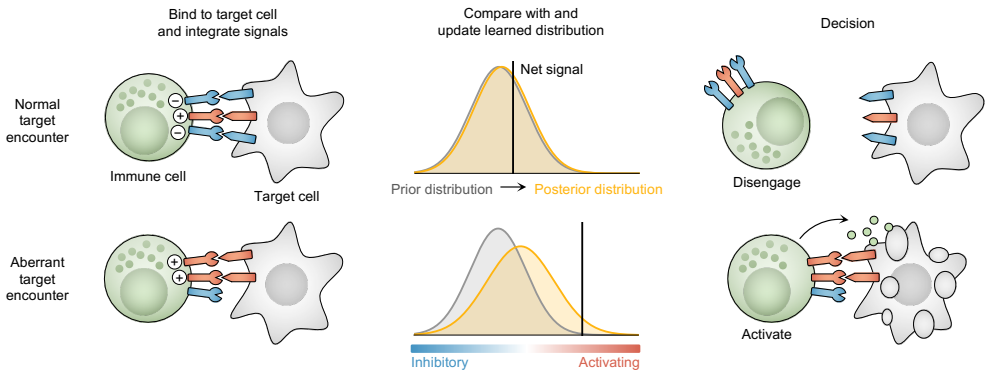

Figure 3.1: **Model overview.** Immune cells receive both activating and inhibitory signals from target cells that they encounter. Net signals are used to update an internal estimate of the signal distribution, reflecting the balance of activating and inhibitory ligands on target cell surfaces in the current environment. Signals from target cells that are far more activating than typical ones stimulate an immune response.

be infected, stressed, or transformed. Biologically, information needed to represent the signal distribution could be encoded by the level of intracellular proteins involved in signaling cascades, or by adjusting the distribution or spatial organization of surface receptors on the immune cell membrane. Here, we characterize the signal distribution that an immune cell receives from targets in its local environment with a Gaussian function, $P(x|\mu, \lambda)$. The *true* mean μ_t and precision $\lambda_t = 1/\sigma_t^2$ are unknown, and must be estimated through multiple encounters with target cells. Estimating the variance σ_t^2 in signal values, as well as the mean, is crucial to differentiate true outlier target cells from ones with normal variation in surface ligand expression.

In Bayesian inference, one begins a prior distribution $P_{\text{prior}}(\mu, \lambda)$ which constitutes an initial guess for the unknown parameters. This distribution is then updated each time data (i.e., a signal) is received following Bayes' rule,

$$P(\mu, \lambda|x) = \frac{P(x|\mu, \lambda)P_{\text{prior}}(\mu, \lambda)}{\int d\mu' \int d\lambda' \, P(x|\mu', \lambda')P_{\text{prior}}(\mu', \lambda')} \,. \tag{3.1}$$

In general, the true signal mean and variance may be nonstationary. For example, the distribution of ligands expressed on target cells may vary as an immune cell migrates from one tissue to another. To accommodate time-varying signals, we employ a modified Bayesian update rule where the strength of the prior distribution is fixed. We begin with a normal-gamma prior for the signal mean μ and precision λ,

$$P(\mu, \lambda | m, \kappa, \alpha, \beta) = \frac{\beta^\alpha \sqrt{\kappa}}{\Gamma(\alpha)\sqrt{2\pi}} \lambda^{\alpha - 1/2} e^{-\beta\lambda - \kappa\lambda \frac{(\mu - m)^2}{2}}, \tag{3.2}$$

which is the conjugate prior for a Gaussian distribution with unknown mean and variance. This is a choice of mathematical convenience that allows us to easily write down analytical expressions for how parameters are updated as new signals are received. Here $\Gamma(\alpha)$ represents the gamma distribution. The parameters κ and α in our modified version of the Bayesian inference are proportional to the number of measurements used for estimating the mean and variance. Unlike in standard Bayesian inference, we consider these parameters to be fixed. This constitutes our key modification to the standard method. Following equation 3.2, the mean value of μ is m, and the mean value of λ is α/β. When the immune cell binds to a new target cell and receives a signal x, the internal representation of the signal distribution is updated with new parameters

$$m \to \frac{\kappa m + x}{\kappa + 1}, \quad \beta \to \frac{\alpha - 1}{\alpha - \frac{1}{2}} \left[\beta + \frac{\kappa}{\kappa + 1} \frac{(x - m)^2}{2} \right]. \tag{3.3}$$

This expression differs from the standard Bayesian update in that the parameters κ and α are held fixed, and the updated value of β is shrunk by a factor of $(\alpha - 1)/\left(\alpha - \frac{1}{2}\right)$ to compensate.

3.2.1 Derivation of equation 3.3

Given a normal-gamma prior distribution for the Gaussian mean μ and precision λ, see equation 3.2, the posterior distribution for μ and λ after taking a new measurement x also follows a normal-gamma distribution with modified parameters m', β', α', and κ':

$$P_{\text{post}}(\mu, \lambda | m', \beta', \alpha', \kappa') = P(x|\mu, \lambda) P_{\text{prior}}(\mu, \lambda | m, \beta, \alpha, \kappa)\,.$$

The parameters of the posterior distribution are related to those of the prior distribution by

$$\begin{aligned} m' &= \frac{\kappa m + x}{\kappa + 1} & \kappa' &= \kappa + 1 \\ \beta' &= \beta + \frac{\kappa}{\kappa + 1} \frac{(x - m)^2}{2} & \alpha' &= \alpha + \frac{1}{2}\,. \end{aligned}$$

While this standard Bayesian approach described above works well for static signal distributions, it is not suitable for a signal distribution that changes over time. This is because the parameters κ and α are incremented after every measurement in equation 2.1, thereby minimizing the contribution of recent data after many measurements have been collected.

To enable adaptation to dynamic environments, we modify the update rules by fixing κ and α. These parameters count the effective number of samples used to estimate the mean and precision, respectively (see equation 2.1 and equation 2.3). Fixing α creates a potential problem because β and $(x-m)^2$ are nonnegative, which means that β will increase with each update (see equation 2.1) and thus $\hat{\sigma}^2$ will grow without bound. β' must therefore be shrunk by an α-dependent factor δ_α to keep estimates of the standard deviation finite. We can determine this factor by requiring that the update of the standard deviation follow

the same form as in the original case (equation 2.3), where now $\alpha' = \alpha$,

$$\left(\hat{\sigma}^2\right)' = \frac{\beta'}{\alpha' - 1} = \frac{\beta + \frac{1}{2}\frac{\kappa(x-m)^2}{(\kappa+1)}}{\alpha - \frac{1}{2}} = \delta_\alpha \frac{\beta + \frac{1}{2}\frac{\kappa(x-m)^2}{(\kappa+1)}}{\alpha - 1}\,, \tag{3.4}$$

which yields $\delta_\alpha = (\alpha - 1)\,/\,(\alpha - 1/2)$.

Rescaling β' in this way while keeping κ and α fixed,

$$\beta' = \frac{\alpha - 1}{\alpha - \frac{1}{2}}\left[\beta + \frac{\kappa}{\kappa + 1}\frac{(x-m)^2}{2}\right],$$

not only ensures that β is finite, it also gives the same expression for updating the estimated variance as previously,

$$\left(\hat{\sigma}^2\right)' = \frac{\beta'}{\alpha' - 1} = \frac{\alpha - 1}{\alpha - \frac{1}{2}}\hat{\sigma}^2 + \frac{\frac{1}{2}}{\alpha - \frac{1}{2}}\frac{\kappa}{\kappa + 1}(x - m)^2\,.$$

In summary, the adaptive update rules are given by

$$\begin{aligned} m' &= \frac{\kappa m + x}{\kappa + 1} & \kappa' &= \kappa \\ \beta' &= \frac{\alpha - 1}{\alpha - \frac{1}{2}}\left[\beta + \frac{\kappa}{\kappa + 1}\frac{(x-m)^2}{2}\right] & \alpha' &= \alpha\,. \end{aligned} \tag{3.5}$$

3.2.2 Modified inference as exponential weight decay

In standard Bayesian inference, all measurements contribute equally to parameter estimation. Let $\{x_1, x_2, \ldots, x_n\}$ represent a set of n signals, where the subscript indicates the order in which each signal was measured. Assuming starting prior parameter values of m_0, β_0, α_0 and κ_0, and following the standard update rules given in equation 2.1, after n signal measurements we have

$$\begin{aligned} m_n &= \frac{\kappa_0}{\kappa_0 + n}m_0 + \frac{1}{\kappa_0 + n}\sum_{i=1}^{n} x_i\,, \\ \beta_n &= \beta_0 + \frac{1}{2}\frac{\kappa_0 n}{\kappa_0 + n}\left(\frac{1}{n}\sum_{i=1}^{n} x_i - m_0\right)^2 + \frac{1}{2}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j\right)^2\,. \end{aligned}$$

If we instead follow the adaptive update rules given in equation 3.5, the parameter estimates become

$$\begin{aligned} m_n &= \frac{\kappa}{\kappa+1} m_{n-1} + \frac{1}{\kappa+1} x_n \\ &= \left(\frac{\kappa}{\kappa+1}\right)^2 m_{n-2} + \frac{\kappa}{\kappa+1}\frac{1}{\kappa+1} x_{n-1} + \frac{1}{\kappa+1} x_n \\ &= \delta_\kappa^n m_0 + \frac{1}{\kappa+1} \sum_{i=1}^{n} \delta_\kappa^{n-i} x_i \,, \end{aligned}$$

where $\delta_\kappa = \kappa / (\kappa+1)$, and

$$\begin{aligned} \beta_n &= \frac{\alpha-1}{\alpha-\frac{1}{2}} \left[\beta_{n-1} + \frac{1}{2}\frac{\kappa}{\kappa+1} \left(x_n - m_{n-1}\right)^2 \right] \\ &= \delta_\alpha^n \beta_0 + \frac{1}{2}\delta_\kappa \sum_{i=1}^{n} \delta_\alpha^{n-i+1} \left(x_i - m_{i-1}\right)^2 \,, \end{aligned}$$

where $\delta_\alpha = (\alpha-1)/(\alpha-1/2)$. Here, each of the n measurements no longer contributes equally to the estimated parameters. New measurements are emphasized more strongly than old ones. With each successive measurement, the effective weight of older signal measurements decreases by a factor of δ_κ for m, and by a factor of δ_α for β.

The argument above facilitates the interpretation of κ and α as parameters controlling the memory length for estimating the signal distribution. Though it would be natural to take $\kappa \sim 2\alpha$, in principle the memory length for m and β can be separately controlled by adjusting κ and α, respectively. Larger values of κ and α result in longer memory lengths and slower adaptation.

Thus our modified Bayesian inference framework introduces a *memory length* for adaptation, encoded by the parameters κ and α. The larger κ and α are, the less the estimated mean and variance will shift when a new signal is received. The update rule described in equation 3.3 is exactly equivalent to placing a weight on measurements of the

signal distribution that decays exponentially as new measurements are taken, emphasizing more recent signals over older ones. Larger values of κ and α result in a slower decay, and thus the memory of the signal distribution becomes longer. We emphasize that this notion of memory is distinct from the concept of memory in adaptive immunity. In this context, memory introduces a tradeoff between precision and adaptability. When the memory is long, it is possible to adapt to a specific distribution precisely, but adaptation to new environments is slow.

3.2.3 Internal representaton $P_r(x)$

Integrating over the unknown mean and precision in equation 3.2, the internal representation of the signal distribution $P_r(x)$ takes the form of a shifted, scaled t-distribution. In our model, the internal representation of the signal distribution $P_r(x)$ is calculated by integrating over Gaussian distributions with different means μ and precisions λ, weighted by the normal-gamma prior distribution $P(\mu, \lambda|m, \kappa, \alpha, \beta)$:

$$\begin{aligned} P_r(x) &= \int_{-\infty}^{\infty} d\mu \int_0^{\infty} d\lambda \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\lambda}{2}(x-\mu)^2} P(\mu, \lambda|m, \kappa, \alpha, \beta) \\ &= \beta^{\alpha} \sqrt{\frac{\kappa}{2\pi(\kappa+1)}} \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)} [\beta + \frac{1}{2}\frac{\kappa}{\kappa+1}(x-m)^2]^{-\alpha-\frac{1}{2}} \end{aligned}$$

This is a scaled, shifted t-distribution

$$P(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} ,$$

where $t = \sqrt{\frac{\alpha\kappa}{\beta(\kappa+1)}}(x-m)$ and $\nu = 2\alpha$.

3.2.4 Activation threshold

During each target cell encounter, we the immune cell activates if it receives a signal x that is substantially more activating than typical signals from the estimated distribution. Specifically, we assume that activation occurs when

$$\begin{aligned} P_r(y > x) &= \int_x^\infty dy \int_{-\infty}^\infty d\mu \int_0^\infty d\lambda \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\lambda}{2}(y-\mu)^2} P(\mu, \lambda) \\ &< \theta \,, \end{aligned} \tag{3.6}$$

for some threshold value θ. The threshold θ must be small to avoid autoimmunity. If an immune cell were to learn the exact signal distribution in an environment where all target cells are healthy, then θ would be the probability that the immune cell activates against a healthy cell. Here we will treat θ as constant, but in principle θ could be modulated adaptively by the immune system, for example to provide heightened surveillance during infection.

We propose that some behaviors of immune cells like as NK cells and macrophages can be understood by considering how they adapt to typical levels of stimulation from target cells in the local environment. As we show in Results, this perspective is consistent with experiments.

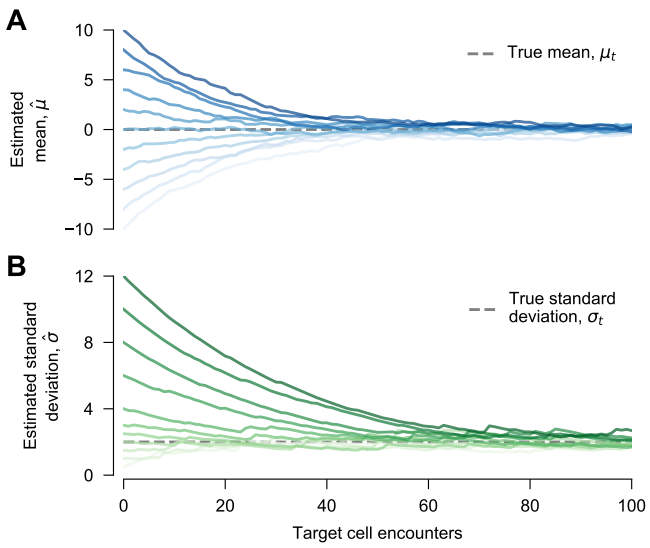

Figure 3.2: **Immune cells adapt to a static environment.** The mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ of the internal representation of the signal distribution converge toward the true mean ($\mu_t = 0$) and standard deviation ($\sigma_t = 2$) of the signal distribution in the environment. **A**, Convergence to μ_t from various initial values of m. The initial value of $\beta = (\alpha - 1)$ is the same in all cases. **B**, Convergence to σ_t from various initial values of β. The initial value of $m = 0$ is the same in all cases. For both sections, the memory length is set by $\kappa = 20$ and $\alpha = 10$.

3.3 Model results for simulated data

3.3.1 Immune cells adapt to the signal distribution in a static environment

We first tested the ability of our model to recover the true parameters of a test signal distribution. We compared the estimated values of μ and σ ($\hat{\mu} = \langle \mu \rangle = m$, $\hat{\sigma} = \sqrt{\langle 1/\lambda \rangle} = \sqrt{\beta/(\alpha - 1)}$, where $\langle \cdot \rangle$ denotes an average over equation 3.2) to those of the true signal distribution. **Figure 3.2** shows that immune cells adapt to the distribution of signals in the environment, even when the initial parameters of the prior are far from the true ones. The number of target cell encounters required to approach the true parameters depends on

the initial distance from them, and on the memory length. The shorter the memory length, the faster the convergence (Fig. 3.3). Adaptation to the true signal mean and variance is

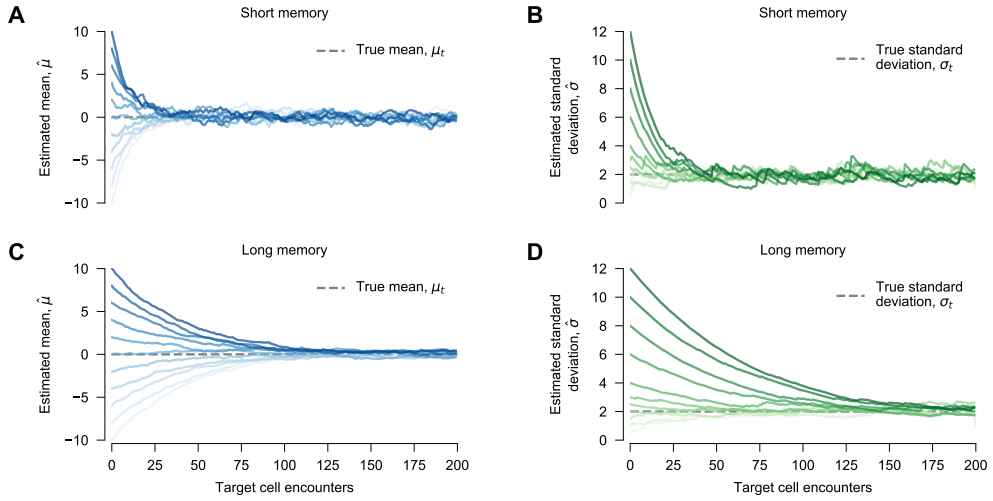

Figure 3.3: **Immune cells with shorter memory lengths exhibit faster but noisier adaptation.** The estimated mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ approach the true mean ($\mu_t = 0$) and true standard deviation ($\sigma_t = 2$) in a finite number of encounters, regardless of the initial value and memory length. When the memory length is shorter ($\kappa = 10$ and $\alpha = 5$, panels **A** and **B**), estimated values approach the true ones faster than for immune cells with longer memory lengths ($\kappa = 40$ and $\alpha = 20$, panels **C** and **D**). However, adaptation is less precise when the memory length is shorter. **A**, Convergence to true mean ($\mu_t = 0$) from various initial values of m. The initial value of $\beta = (\alpha - 1)$ is the same in all cases. **B**, Convergence to the true standard deviation ($\sigma_t = 2$) from various initial values of β. The initial value of $m = 0$ is the same in all cases. The memory length for both **A** and **B** is set by $\kappa = 10$ and $\alpha = 5$. (**C**, **D**) Display convergence toward the true mean and standard deviation as in (**A**, **B**). Initial parameters are the same as those in (**A**, **B**), except for longer memory lengths $\kappa = 40$ and $\alpha = 20$.

not perfect, however. This is because the memory length is finite. Shorter memory lengths result in greater 'noise' in the inferred parameters of the signal distribution.

3.3.2 Immune cells adapt to the signal distribution in dynamic environments

NK cells show a remarkable ability to adapt to changing environments. As noted above, MHC class I is a powerful inhibitory ligand for NK cells. NK cells in hosts that naturally express low levels of MHC class I are self-tolerant. Yet, they are also hyporesponsive to MHC class I-deficient target cells that would usually be killed by NK cells from hosts with normal levels of MHC class I expression [7, 58]. A pair of experiments showed that these behaviors are dynamic, rather than being fixed during development. When mature NK cells were transferred from MHC class I-deficient mice into mice with normal MHC class I expression, they regained their ability to kill MHC-deficient targets [45, 24]. Conversely, NK cells from normal mice that were transferred into the MHC class I-deficient environment became hyposensitive [45]. Importantly, this shift in behavior occurs without the need for cell division or changes in the composition of receptors on the NK cell surface [45]. Additional experiments have confirmed that the manipulation of MHC class I expression in mice leads to analogous results [23, 6].

Adaptation also occurs when NK cells or macrophages undergo prolonged exposure to activating stimuli. For macrophages, extended exposure to lipopolysaccharide (LPS) results in blunted responses to additional stimulation by LPS, a phenomenon known as endotoxin tolerance [106, 30]. When LPS is withdrawn, macrophages gradually recover normal function [106]. In NK cells, a similar phenomenon has been observed with prolonged exposure to ligands for the activating receptor NKG2D, resulting in hyposensitivity [74, 16, 11]. NK cells desensitized through long-duration exposure to NKG2D ligands were also

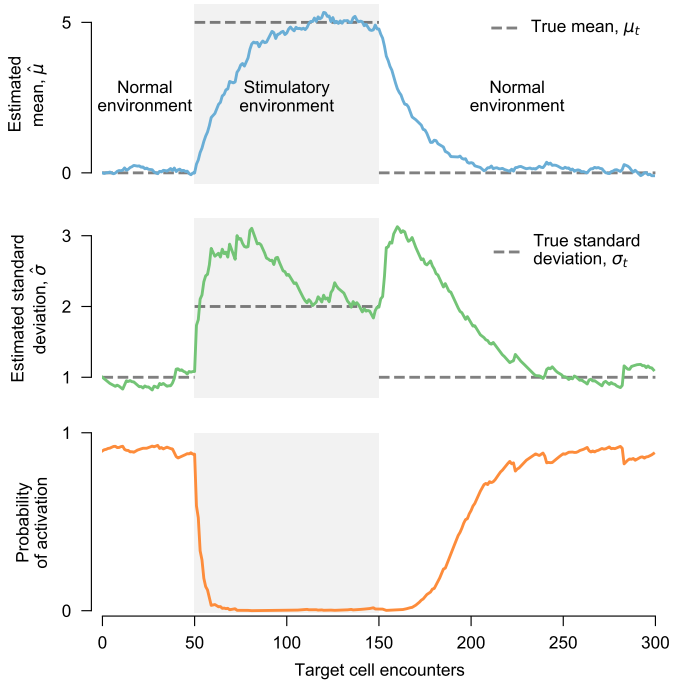

Figure 3.4: **Immune cells adapt to changing environments, mimicking experimentally observed development of hyposensitivity and recovery.** Adaptation of the estimated signal mean $\hat{\mu}$, standard deviation $\hat{\sigma}$, and probability of activation against an aberrant target as an immune cell is transferred between different environments. In the initial, 'normal' environment, the level of activating stimulus is low and the immune cell is primed to respond to aberrant targets. After transfer to a new, more stimulating environment (shaded region), the immune cell adapts to the new signal distribution and progressively loses the ability to respond to aberrant targets. When the immune cell is returned to the 'normal' environment, responsiveness is gradually restored. To compute the probability of activation, we used a Gaussian signal distribution from aberrant targets with mean 5 and standard deviation 2, which matches the environment in the shaded region. Initial parameters of the estimated signal distribution are $m = 0$ and $\beta = 9$ (leading to $\hat{\mu} = 0$, $\hat{\sigma} = 1$), with memory parameters $\kappa = 20$ and $\alpha = 10$.

hyposensitive to other activating stimuli [16]. Similar results have also been observed for prolonged exposure to ligands for other activating receptors [35, 92, 97].

We ran a series to simulations to mimic the transient exposure of immune cells to different environments or levels of stimulus. In our simulations, immune cells began in a 'normal' environment with $\mu_t = 0$ and $\sigma_t = 1$. To quantify their ability to activate against

targets that express high levels of activating ligands and/or low levels of inhibitory ligands, we computed the probability that an individual immune cell activates in response to an aberrant target cell with a signal drawn from a Gaussian distribution with mean $\mu_a = 5$ and standard deviation $\sigma_a = 2$, using a threshold $\theta = 0.01$. Because μ_a is substantially higher than μ_t, due to the presence (lack) of activating (inhibitory) ligands on the target cell surface, the probability of activation is initially high (**Fig. 3.4**). To simulate a change in environment, after 50 encounters with normal targets we switched the distribution of signals to match the signal distribution from 'aberrant' targets described above. As the immune cell adapts to this new environment, its probability of activation by aberrant target cells decreases. The probability of activation settles near zero once the immune cell fully adapts to the environment. However, this hyposensitivity is not permanent. After a total of 150 target cell encounters, we return the immune cell to the normal environment ($\mu_t = 0$, $\sigma_t = 1$). As the immune cell encounters more healthy cells, it regains its capacity to activate against aberrant targets. This trajectory mimics the development of hyposensitivity and restoration of normal function described in experiments [45, 24, 106, 30, 74, 16, 11, 35, 92, 97], and is observed consistently in our simulations (**Fig. 3.5**).

The number of encounters required to develop hyposensitivity in **Fig. 3.4**, and to recover it after the immune cell returns to the normal environment, depends on the memory parameters κ and α. This behavior also depends on the difference between the normal and aberrant environments. Nevertheless, one of the robust predictions of our model, that stands independent of the exact parameter selection, is the development and subsequent loss of hyposensitivity. From a biological standpoint, this memory length could

be adaptively modulated by the immune system, potentially through mechanisms such as cytokine signaling. This process serves as a critical protective measure, enhancing the system's ability to defend against infectious agents more effectively, thereby optimizing the body's immune response.

3.3.3 Finite memory results in heterogeneous immune cell responses

As shown in **Fig. 3.2**, immune cells in our model do not adapt perfectly to the true signal distribution in the environment due to finite memory. Together with the stochastic nature of target cell encounters, this implies that there will be a range of immune cell responses, even among immune cells with identical receptors. **Figure 3.5** shows an example of this behavior in a population of immune cells transferred between different environments, as shown for a single cell in **Fig. 3.4**.

To systematically explore this heterogeneity, we sought to characterize the distribution of estimated signal parameters m and β for a population of identical immune cells with finite memory, which inhabit the same environment. As an analytical approach, we developed a continuous approximation of the discrete modified Bayesian update dynamics described in equation 3.3. Assuming that parameter updates with each target cell encounter are small, equation 3.3 is described by a set of stochastic differential equations(See following sections). We then derived a Fokker-Planck equation that describes the evolution of the distribution of learned m and β values for a population of immune cells with identical receptors in the same environment(See following sections).

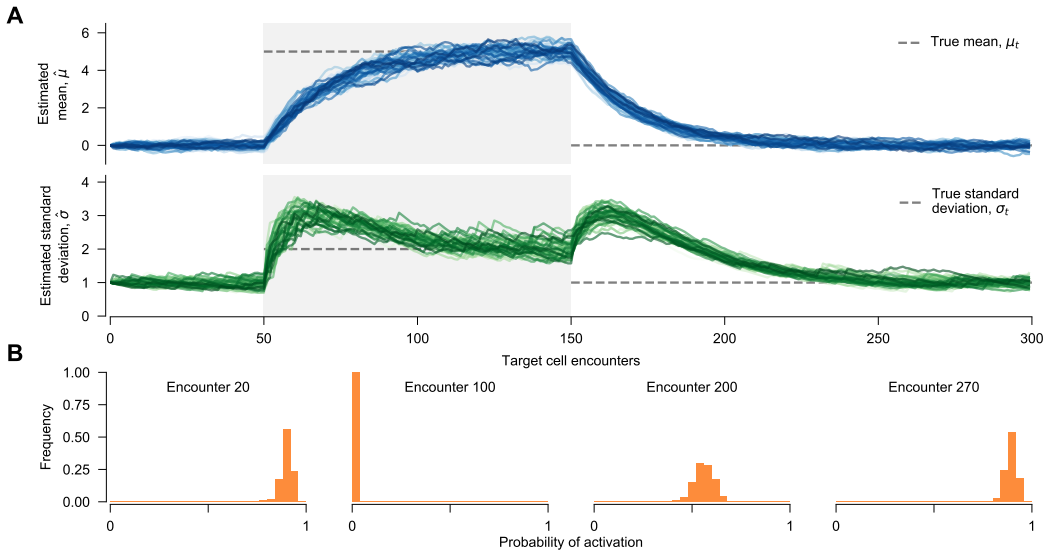

Figure 3.5: **Finite memory of past interactions with target cells results in heterogeneous immune cell behaviors. A**, Adaptation of 500 immune cells to changing environments, following the same conventions as in **Fig. 3**. Even though all immune cells start with the same initial values ($m = 0$ and $\beta = 9$, with memory parameters $\kappa = 20$ and $\alpha = 10$), the learned distribution for each immune cell evolves differently over time. This is due to the stochastic nature of signals from target cells, and the finite memory length of past target encounters. **B**, The heterogeneity of learned signal distributions also leads to heterogeneous immune cell responses, characterized by the probability of activation against an aberrant target cell. The signal distribution for aberrant targets is the same as in **Fig. 3**.

Derivation of the stochastic differential equations

Here, our aim is to characterize the evolution of the estimated signal distribution parameters across a vast population of cells, all equipped with identical receptors. To accomplish this, we initiate the process by applying the parameter update rules, as provided in equation 3.5. This enables us to calculate the difference between the updated parameters and their previous values.

$$\Delta m = m' - m = \frac{1}{\kappa + 1}\left(x - m\right) ,$$
$$\Delta \beta = \beta' - \beta = \frac{1}{2}\left[\delta_\alpha \delta_\kappa \left(x - m\right)^2 - \frac{\beta}{\alpha - \frac{1}{2}}\right] .$$

For an immune cell that is well-adapted to the environment, Δm and $\Delta\beta$ should be small, especially when the memory parameters κ and α are large. Writing the time interval Δt as a single encounter, we can develop stochastic differential equations that describe the evolution of m and β,

$$\begin{aligned}
\frac{\Delta m}{\Delta t} &= \frac{1}{\kappa+1}\left(x-m\right) = \frac{1}{\kappa+1}\left(\mu_t - m + \sigma_t\xi\right) , \\
\frac{\Delta\beta}{\Delta t} &= \frac{1}{2}\left[\delta_\alpha\delta_\kappa\left(x-m\right)^2 - \frac{\beta}{\alpha-\frac{1}{2}}\right] \\
&= \frac{1}{2}\left\{\delta_\alpha\delta_\kappa\left[\left(\mu_t-m\right)^2 + 2\left(\mu_t-m\right)\sigma_t\xi + \sigma_t^2\xi^2\right] - \frac{\beta}{\alpha-\frac{1}{2}}\right\} .
\end{aligned}$$

Here we assume a constant environment, where the signal distribution is Gaussian with mean μ_t and standard deviation σ_t. We thus write $x = \mu_t + \sigma_t\xi$, where ξ is a Gaussian white noise.

Collectively, the evolution of $\boldsymbol{\theta} = (m, \beta)$ follows

$$d\boldsymbol{\theta} = \mathbf{A}\left(\boldsymbol{\theta}, t\right)\, dt + \sqrt{\mathbf{D}\left(\boldsymbol{\theta}, t\right)}dW\,, \tag{3.7}$$

where $\mathbf{A}$ and $\mathbf{D}$ are referred to as the drift vector and the diffusion matrix, respectively. The drift vector describes the expected change in $\boldsymbol{\theta}$ parameters, and the diffusion matrix describes the covariance of changes in $\boldsymbol{\theta}$,

$$\mathbf{A} = \frac{\langle\Delta\boldsymbol{\theta}\rangle}{\Delta t}\,, \qquad \mathbf{D} = \frac{\left\langle\left(\Delta\boldsymbol{\theta}\right)^2\right\rangle - \langle\Delta\boldsymbol{\theta}\rangle^2}{\Delta t}\,.$$

The components of $\mathbf{A}$ and $\mathbf{D}$ can thus be computed by finding the first and second moments of $\Delta\boldsymbol{\theta}$, which are

$$
\begin{aligned}
\frac{\langle \Delta m \rangle}{\Delta t} &= \frac{1}{\kappa+1}\left(\mu_t - m\right), \\
\frac{\langle \Delta \beta \rangle}{\Delta t} &= \frac{1}{2}\left\{ \delta_\alpha \delta_\kappa \left[\left(\mu_t - m\right)^2 + \sigma_t^2 \right] - \frac{\beta}{\alpha - \frac{1}{2}} \right\}, \\
\frac{\left\langle \left(\Delta m\right)^2 \right\rangle}{\Delta t} &= \frac{1}{\left(\kappa+1\right)^2}\left[\left(\mu_t - m\right)^2 + \sigma_t^2 \right], \\
\frac{\left\langle \left(\Delta \beta\right)^2 \right\rangle}{\Delta t} &= \frac{1}{4}\left\{ \left[\delta_\alpha \delta_\kappa \left[\left(\mu_t - m\right)^2 + \sigma_t^2 \right] - \frac{\beta}{\alpha - \frac{1}{2}} \right]^2 + 2\delta_\alpha^2 \delta_\kappa^2 \sigma_t^2 \left[2\left(\mu_t - m\right)^2 + \sigma_t^2 \right] \right\}, \\
\frac{\langle \Delta m \Delta \beta \rangle}{\Delta t} &= \frac{1}{\kappa+1}\left(\mu_t - m\right) \times \frac{1}{2}\left\{ \delta_\alpha \delta_\kappa \left[\left(\mu_t - m\right)^2 + 3\sigma_t^2 \right] - \frac{\beta}{\alpha - \frac{1}{2}} \right\}.
\end{aligned}
$$

This allows us to write down the cumulants

$$
\begin{aligned}
\frac{\left\langle \left(\Delta m\right)^2 \right\rangle - \langle \Delta m \rangle^2}{\Delta t} &= \frac{1}{\left(\kappa+1\right)^2}\sigma_t^2, \\
\frac{\left\langle \left(\Delta \beta\right)^2 \right\rangle - \langle \Delta \beta \rangle^2}{\Delta t} &= \frac{1}{2}\delta_\alpha^2 \delta_\kappa^2 \sigma_t^2 \left[2\left(\mu_t - m\right)^2 + \sigma_t^2 \right], \\
\frac{\langle \Delta m \Delta \beta \rangle - \langle \Delta m \rangle \langle \Delta \beta \rangle}{\Delta t} &= \frac{1}{\kappa+1}\left(\mu_t - m\right) \times \delta_\alpha \delta_\kappa \sigma_t^2 .
\end{aligned}
$$

Finally we obtain expressions for the vector $\mathbf{A}$ and matrix $\mathbf{D}$,

$$
\mathbf{A} = \begin{bmatrix} \frac{1}{\kappa+1}\left(\mu_t - m\right) \\ \frac{1}{2}\left\{ \delta_\alpha \delta_\kappa \left[\left(\mu_t - m\right)^2 + \sigma_t^2 \right] - \frac{\beta}{\alpha - \frac{1}{2}} \right\} \end{bmatrix},
$$

$$
\mathbf{D} = \begin{bmatrix} \frac{1}{\left(\kappa+1\right)^2}\sigma_t^2 & \frac{\mu_t - m}{\kappa+1} \times \delta_\alpha \delta_\kappa \sigma_t^2 \\ \frac{\mu_t - m}{\kappa+1} \times \delta_\alpha \delta_\kappa \sigma_t^2 & \frac{1}{2}\delta_\alpha^2 \delta_\kappa^2 \sigma_t^2 \left[2\left(\mu_t - m\right)^2 + \sigma_t^2 \right] \end{bmatrix}.
$$

Fokker-Planck equation

Equation 3.7 is a stochastic differential equation that describes the evolution of $\boldsymbol{\theta} = (m, \beta)$ for a single immune cell. To understand the distribution of m and β at the population

level, we derive the Fokker-Planck equation corresponding to equation 3.7. Fokker-Planck equation gives the evolution of the probability distribution of the parameter $\boldsymbol{\theta}$. Following standard methods[86], the form of the Fokker-Planck equation is

$$\frac{\partial P\left(\boldsymbol{\theta},t\right)}{\partial t}=-\left(\frac{\partial}{\partial m}+\frac{\partial}{\partial\beta}\right)\mathbf{A}\left(\boldsymbol{\theta},t\right)P\left(\boldsymbol{\theta},t\right)+\left(\frac{1}{2}\frac{\partial^2}{\partial m^2}+\frac{1}{2}\frac{\partial^2}{\partial\beta^2}+\frac{\partial^2}{\partial m\partial\beta}\right)\mathbf{D}\left(\boldsymbol{\theta},t\right)P\left(\boldsymbol{\theta},t\right)\,.$$

Substituting in the expressions for $\mathbf{A}$ and $\mathbf{D}$ derived above, we obtain

$$\begin{aligned}\frac{\partial P\left(m,\beta,t\right)}{\partial t}=&\left(\frac{1}{\kappa+1}+\frac{1}{2\alpha-1}\right)P-\frac{\mu_t-m}{\kappa+1}\frac{\partial P}{\partial m}\\&-\left\{\frac{1}{2}\left[\delta_\alpha\delta_\kappa\left[\left(\mu_t-m\right)^2+\sigma_t^2\right]-\frac{\beta}{\alpha-\frac{1}{2}}\right]+\frac{\delta_\alpha\delta_\kappa}{\kappa+1}\sigma_t^2\right\}\frac{\partial P}{\partial\beta}\\&+\frac{1}{2}\frac{1}{\left(\kappa+1\right)^2}\sigma_t^2\frac{\partial^2P}{\partial^2m}+\frac{\mu_t-m}{\kappa+1}\times\delta_\alpha\delta_\kappa\sigma_t^2\frac{\partial^2P}{\partial m\partial\beta}\\&+\frac{1}{4}\delta_\alpha^2\delta_\kappa^2\sigma_t^2\left[2\left(\mu_t-m\right)^2+\sigma_t^2\right]\frac{\partial^2P}{\partial^2\beta}\,.\end{aligned}\tag{3.8}$$

Numerical integration of the Fokker-Planck equation

We used the central difference method to solve equation 3.8 numerically. Here we simulated adaptation to a true signal distribution with mean $\mu_t = 2$ and standard deviation $\sigma_t = 1$, using the memory parameters $\kappa = 20$ and $\alpha = 10$. We anticipated that the learned signal parameters should be centered around $m = 2$ and $\beta = 9$, which would give $\hat{\mu} = m = 2 = \mu_t$ and $\hat{\sigma}^2 = \frac{\beta}{\alpha-1} = 1 = \sigma_t^2$. Thus we selected a rectangular domain $[1.5, 2.5] \times [5, 15]$ in the (m, β) space to numerically evaluate the equations (**Fig. 4**). We set $h = w = 0.01$ as the spatial discretization size of m and β, respectively, dividing the domain into a 100×1000 grid. The discrete first and second derivatives at each grid point

can be expressed as

$$\begin{aligned}
\frac{\partial P}{\partial m}(m,\beta) &= \frac{P(m+h,\beta)-P(m-h,\beta)}{2h}, \\
\frac{\partial P}{\partial \beta}(m,\beta) &= \frac{P(m,\beta+w)-P(m,\beta-w)}{2w}, \\
\frac{\partial^2 P}{\partial^2 m}(m,\beta) &= \frac{P(m+h,\beta)+P(m-h,\beta)-2P(m,\beta)}{h^2}, \\
\frac{\partial^2 P}{\partial^2 \beta}(m,\beta) &= \frac{P(m,\beta+w)+P(m,\beta-w)-2P(m,\beta)}{w^2}, \\
\frac{\partial^2 P}{\partial m \partial \beta}(m,\beta) &= [P(m+h,\beta+w)+P(m-h,\beta-w)- \\
&\qquad P(m+h,\beta-w)-P(m-h,\beta+w)]/(4hw)\ .
\end{aligned}$$

After applying the spatial discretization to the right hand side, equation 3.8 can be simplified to

$$\frac{\partial \mathbf{P}}{\partial t} = \mathbf{M}\mathbf{P} \tag{3.9}$$

where the discrete probability distribution $\mathbf{P}$ is a 10^5 dimensional vector representing values of the function at all grid points. $\mathbf{M}$ is a $10^5 \times 10^5$ matrix with entries determined by equation 3.8, using the discrete derivative formulas above. $\mathbf{M}$ does not change over time.

We used the Crank–Nicolson method[25] to solve equation 3.9. The Crank–Nicolson discretization is

$$\frac{\mathbf{P}^{n+1}-\mathbf{P}^n}{dt} = \mathbf{M}\frac{\mathbf{P}^{n+1}+\mathbf{P}^n}{2}$$

where n is the discrete time index, proportional to the number of interactions between immune cells and target cells. This is a linear equation with variables $\mathbf{P}^{n+1}$ and $\mathbf{P}^n$. It can be written in the form of $\left(\mathbf{I}-\frac{\mathbf{M}}{2}dt\right)\mathbf{P}^{n+1} = \left(\mathbf{I}+\frac{\mathbf{M}}{2}dt\right)\mathbf{P}^n$, where $\mathbf{P}^{n+1}$ is solved for and normalized to 1 at each step. The initial values of the probability distribution $P(m,\beta,0)$

were set to be uniform among the internal grid points and zero at the boundaries. The time step was set as $dt = 0.001$.

Numerical result of Fokker-Planck equation

Figure 3.6A shows one example of the learned distribution of m and β parameters, derived from numerical integration of the Fokker-Planck equation. As expected, the distribution of learned parameters is centered around the true ones. Due to finite memory, some immune cells that have recently interacted with targets that provide unusually low levels of stimulus underestimate the true signal mean in this environment, for example.

Heterogeneous adaptation to the environment results in heterogeneous responses to target cells. **Figure 3.6B** shows the probability that an immune cell with particular values of m and β responds to a set of aberrant target cells. Lower values of m and β result in estimated signal distributions $P_r(x)$ that are more concentrated around smaller values of the signal x, making these immune cells more sensitive to aberrant targets (see equation 3.6). Thus, even for immune cells that express identical sets of receptors, stochastic encounters lead to a range in responsiveness to targets (**Fig. 3.6C**). This is a broad distribution of probability of activation, instead of a peak distribution of probability of activating. The main reason is that the antigen signal distribution $N(4.5, 1)$ here is close to the signal distribution NK cells adapt to, $N(2, 1)$, responses of NK cells are more 'stochastic' than that in **Fig. 3.5**, where the antigen signal distribution is $N(5, 2)$ and the healthy signal distribution is $N(0, 1)$. The distribution of probability of activating is determined by the difference between healthy signal distribution and unhealthy signal distribution, and the

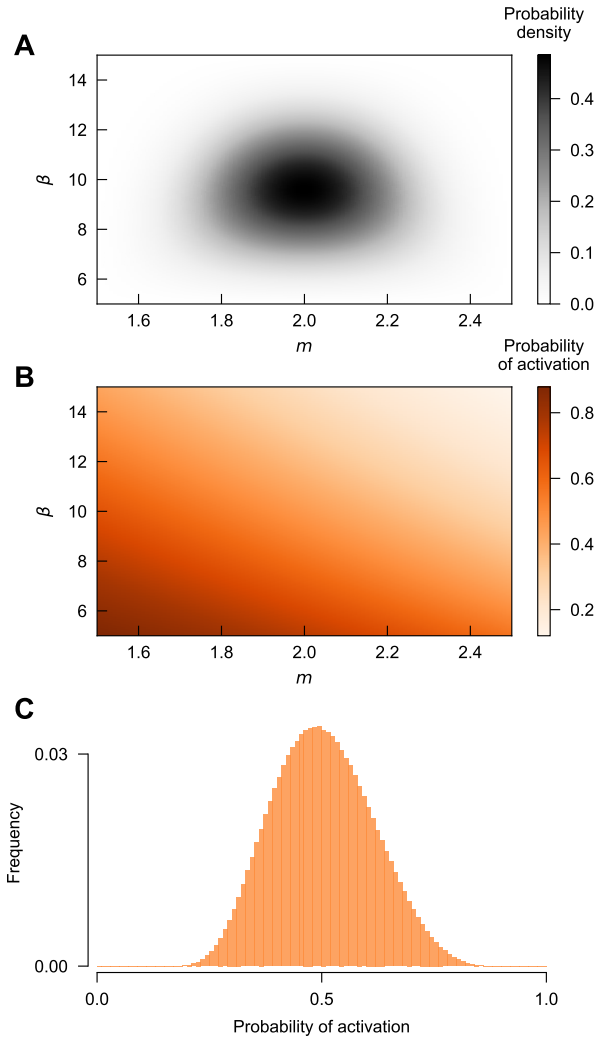

Figure 3.6: **Steady state distribution of immune cell adaptation and responsiveness due to finite memory.** **A**, Steady state joint probability distribution of learned (m, β) parameters, estimated by numerical solution of the Fokker-Planck equation. Here the true signal mean is $\mu_t = 2$, and the true standard deviation is $\sigma_t = 1$. Here we used memory values $\kappa = 20$ and $\alpha = 10$. We observe that the learned parameters are indeed concentrated around the true ones ($\hat{\mu} = m \approx 2$, $\hat{\sigma}^2 = \beta/(\alpha - 1) \approx 1$). **B**, Probability of activation against aberrant targets with signal distribution $\mu_a = 4.5$ and $\sigma_a = 1$ as a function of (m, β). Immune cells that happen to have lower values of both m and β have higher confidence that target cell signals should be more inhibitory, and thus they are more sensitive to stimulus from aberrant targets. **C**, Net distribution of probabilities of activation against aberrant targets, obtained by combining the distributions in **A** and **B**. A wide range of responses exist: some immune cells have a high probability of recognizing aberrant targets, while others are relatively unlikely to respond.

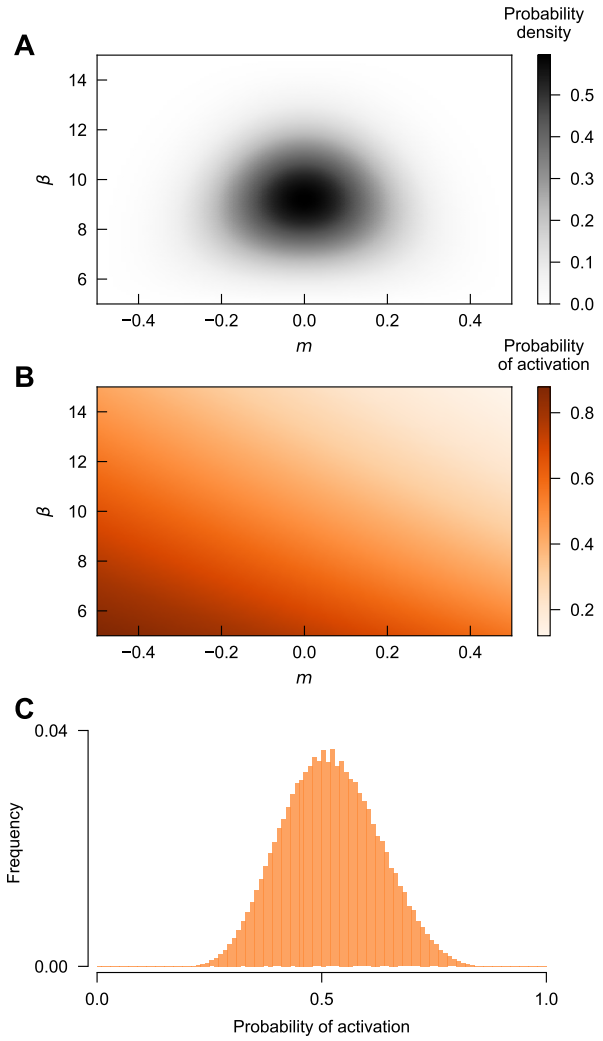

Figure 3.7: **The heterogeneous responses of NK cell population hold in all individuals.** Even for the same phenotype of NK cells and same target signal distribution, NK cells can react heterogeneously due to the finite memory and randomness of signal. This does not depend on the mean of healthy signal distribution. When we choose a different healthy signal mean, $\mu_t = 0$ instead of $\mu_t = 2$ (Fig. 3.6), there still would be some target signal distributions that NK cells response heterogeneously after training. Here we used $\mu_a = 2.5$ as the mean of target signal distribution.

absolute value of the healthy signal mean does not matter. **Figure 3.7** shows the similar result if we use the true mean of healthy signal to $\mu_t = 0$ and calculate the responses of NK cell to targets with mean $\mu_a = 2.5$.

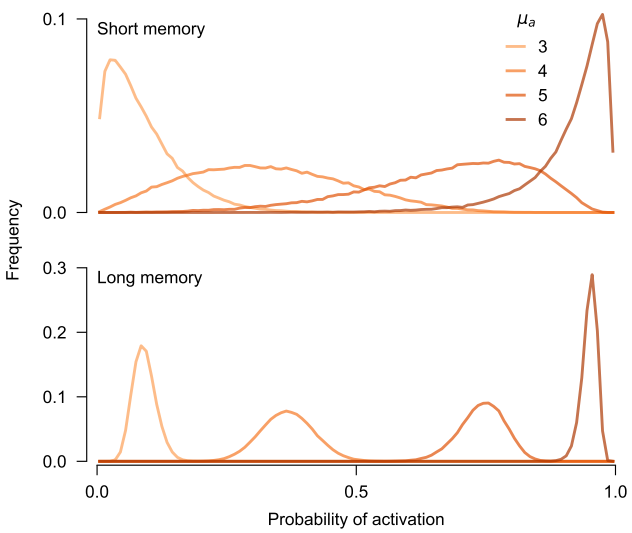

Figure 3.8: **Immune cell responses are diverse, but follow predictable trends depending on the level of stimulation received from target cells.** Distribution of probabilities of activation for populations of immune cells with short ($\kappa = 10$, $\alpha = 5$, *top*) and long ($\kappa = 100$, $\alpha = 50$, *bottom*) memory lengths. Each immune cell population consisted of 10^5 cells with random initial values of m and β uniformly distributed between $[0, 5]$ and $[0, 40]$, respectively. Immune populations were evolved through 10^3 target cell encounters with signal distributions $\mu_t = 2$ and $\sigma_t = 1$. Immune cells were then tested for their probability to activate against aberrant targets that provide different levels of stimulus ($\mu_a = (3, 4, 5, 6)$). Here $\sigma_a = 1$ in all cases. For both shorter and longer memory lengths, the probability of activation is low when μ_a is close to μ_t and high when $\mu_a \gg \mu_t$. However, the spread in activation probabilities is significantly larger for immune cell populations with short memories.

Memory length controls the degree of heterogeneity in immune cell adaptation to the environment, and consequently, the degree of heterogeneity in immune cell responses to targets. **Figure 3.8** shows responses of a panel of immune cell populations with different values of the memory parameters κ and α. Shorter memory values result in more heterogeneous responses. Responses against aberrant targets are more reliable for cells with longer memories. But importantly, even fairly short memory lengths result in behavior that is qualitatively similar to immune cells with longer memories and greater precision. Short memory lengths do not necessarily lead to pathological responses such as increased activa-

tion against healthy targets or hyposensitivity to strongly activating target cells. And as noted previously, long memories result in slow adaptation to different environments.

Our results on immune cell heterogeneity are consistent with experimental studies. Investigations of the patterns of target cell killing have observed widely varying responses for individual NK cells: some kill many targets efficiently, while others are inactive [98, 15, 40, 13, 111, 41]. This finding holds not only for diverse primary cells [98, 15, 40], but also for NK cell lines [13, 111, 41], where cells in the population would be predicted to have homogeneous expression and density of receptors.

3.4 Quantitative comparisons with experiments

To test the ability of our model to recapitulate experimentally-observed immune cell behaviors, we analyzed three data sets involving both in vitro and in vivo experiments.

3.4.1 Comparisons with experimental data of exposing to two different types of target cells

First, we considered a recent experiment in which primary human NK cells were exposed to two different types of target cells [89]. The target cells were Daudi (a malignant B cell line) cells that were either opsonized with rituximab, an antibody that can activate NK cells through CD16, or transfected to express MICA, a ligand for the activating NKG2D receptor. In this experiment, NK cells were presented with two targets in different orders (i.e., Daudi-rituximab/Daudi-rituximab, Daudi-rituximab/Daudi-MICA, Daudi-MICA/Daudi-rituximab, and Daudi-MICA/Daudi-MICA) and it was recorded whether the

NK cells killed neither target, the first target only, the second target only, or both target cells. The authors observed that the MICA-expressing target cells appeared to be more strongly activating and that the frequency of NK cell activation depended on past encounters.

To compare our model with the data of [89], we inferred parameters for the signal distribution for each type of target cell (healthy cells, which NK cells encountered prior to exposure to the unhealthy targets, Daudi-rituximab, and Daudi-MICA). For simplicity, we assumed that the means of these signal distributions differed, but that they all had unit standard deviations. We also set the mean of the healthy signal distribution to zero without loss of generality. We then fit the parameters α (the memory length, with κ set to 2α), μ_R (signal mean for Daudi-rituximab), and μ_M (signal mean for Daudi-MICA) *only* to data where NK cells encountered two Daudi-rituximab or two Daudi-MICA targets in a row. This would then allow us to *predict* how NK cells would respond to encounters with mixed targets, Daudi-rituximab/Daudi-MICA or Daudi-MICA/Daudi-rituximab.

Because the expression for the probability of activation in our model is complicated, we used approximate Bayesian computation (ABC) to fit the model parameters to data. This method draws parameter values from a prior distribution and uses them to generate simulated data that is compared with the real data, iteratively moving toward sets of parameters that produce results that are most consistent with the real data. We used a Python implementation of this method called Engine for Likelihood-Free Inference (ELFI) [60]. We chose a uniform prior distribution between 1 and 50 for α, and Gamma prior distributions with shape and scale parameters $k = 2$ and $\theta = 2$ for μ_R and μ_M. We used the

sum of the squared differences between the simulated and experimental response frequencies as our metric for assessing model fit to data. The best fit parameters were $\alpha = 21.5$, $\mu_R = 2.2$, $\mu_M = 3.2$. Although we only used data from NK cell encounters with two Daudi-rituximab or two Daudi-MICA targets, the model accurately predicts the outcome of Daudi-rituximab/Daudi-MICA and Daudi-MICA/Daudi-rituximab interactions. The total squared error of the best-fit model is 0.08 (**Fig. 3.9**).

As a reference, we also fit a simple 'null' model that includes constant probabilities that an NK cell responds to Daudi-rituximab or Daudi-MICA, p_R or p_M, respectively. When fit against *all* the data, we find $p_R = 0.33$ and $p_M = 0.77$. The null model fits the data with a total mean squared error of 0.15. Thus, this null model has a higher error than for our model, even though our model was only fit to a subset of the data.

3.4.2 Comparisons with series killing experiments

Second, we analyzed data from an experiment in which IL-2-activated primary NK cells were presented with HeLa-CD48 targets [80]. HeLa-CD48 cells are HeLa cells transfected to express CD48, a ligand for the activating receptor 2B4. In this experiment, NK cells encountered up to six HeLa-CD48 targets sequentially, and the fraction of encounters in which NK cells killed their targets was recorded. The authors observed that the fraction of targets killed declined with subsequent encounters.

Our model is able to simulate the regulation of immune cells activation as NK cell interact with same type of targets for multiple times. Next we are going to build a connection between parameters and the experimental variables. The simulation is determined

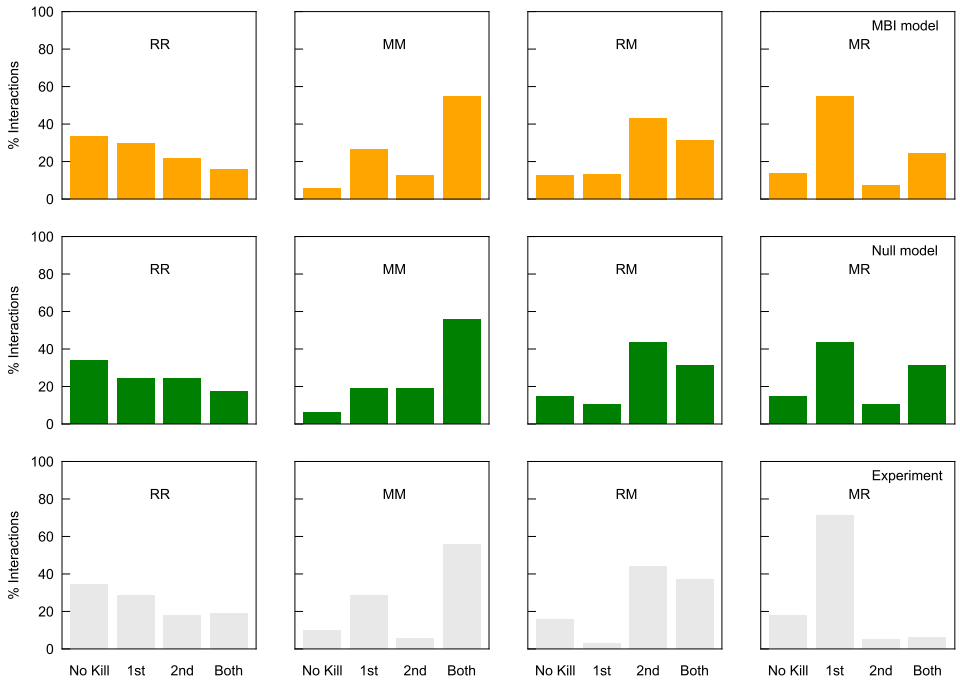

Figure 3.9: **Our model could reproduce the changes in killing ability of NK cells when interacting with targets in different orders.** The first row is the simulation result of our modified Bayesian model. The second row is result from "null" model where the probabilities of killing R and M are constant. The third row is the experimental result presented in the paper [89]. We simulated 8000 NK cells interacting with two types of targets, M and R, twice. It's 2000 NK cells for each of [RR, MM, RM, MR]. RR means NK cell encounters another R type target after encountering R; RM means NK cell encounters an M type target after encountering R. Then we group NK cells by ['No kill', 'kill the 1st target', 'kill the 2nd target' and 'kill both targets']. We assume the healthy signals satisfy $N(0, 1^2)$. We used data from RR and MM interactions for the parameter fitting. Since signals of R and M are more activating than healthy signals, the means must be positive and comparable to variance. We set the prior for mean as gamma function $P(\mu_{a,d}) \sim \frac{x}{4}e^{-\frac{x}{2}}$. We set prior for α as uniform distribution with domain $(1, 50]$ since α must be larger than 1 and should be comparable with 2, the number of interaction in the experiment. We got $\alpha = 21.5, \mu_R = 2.2, \mu_M = 3.2$. For the "null" model, we fitted p_R and p_M using all the data and got $p_R = 0.33, p_M = 0.77$. It shows that our model could provide a more similar result to experiments than the "null" model. The squared error of our model is 0.08 and the squared error of "null" model is 0.15.

by parameters $(\alpha, \kappa, \mu_t, \sigma_t, \mu_a, \sigma_a)$. α and κ are the memory parameters which determine the number of interactions required to adapt to a new environment, or in other word, time spent to adapt to a new environment. We assume that $\kappa = 2\alpha$ before, which ensures m and β have nearly the same exponential decay factor. μ and σ_t determine the healthy (true) signal distribution which is a normal distribution. μ_t is the mean and σ_t is the standard deviation of the healthy signal distribution. μ_a and σ_a determine the antigen signal distribution which is also a normal distribution. μ_a is the mean and σ_a is the standard deviation of the antigen signal distribution. Because the scale of the normal distributions is arbitrary, we can choose the scale that the healthy signal distribution is standard normal distribution $N(0, 1)$. Thus we only need to determine three parameters in total, α, μ_a, σ_a, to determine the activation of immune cells against targets which has signal distribution $N(\mu_a, \sigma_a)$.

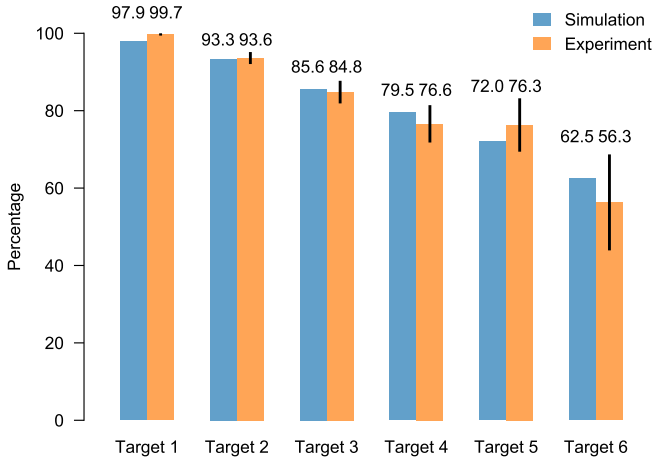

Figure 3.10: Our model could provide a similar serial killing result as experiment. The killing ratio decreases as expected in a serial killing event. The blue bar is the result from this model and the red bar is the result from the experiment. In this simulation, the number of immune cells is 1000. The error bar is calculated from experimental data $w_{bar} = \sqrt{p(1-p)/n}$, p is the probability of killing and n is the number of cells. The fitted parameters are $\alpha = 24.82, \mu = 6.85, \sigma = 2.18$.

To test our ability to recover these results, we fit the memory length α and the mean of the signal distribution for the HeLa-CD48 cells, μ_H, to the experimental data. As above, we assumed that the healthy signal distribution was normal with zero mean and unit variance, and we also fixed the variance of the distribution for HeLa-CD48 cells to one. We used the same prior distributions for α and μ as above. Our metric for assessing model fit was the squared difference between the simulated and experimental response frequencies, normalized by the uncertainty in the experimental response frequencies due to finite sampling. We used $\sqrt{p(1-p)/n}$ as the experimental uncertainty in response frequencies, where p is the empirical response frequency and n is the number of individual cells from which this frequency was determined. We found a set of best-fit parameters $\alpha =$

25.3 and $\mu_H = 4.7$, which lead to an excellent fit with the experimental data. Remarkably, we infer almost exactly the same value of the memory length α as we inferred for the first experiment. Here, a 'null' model with a constant probability of response will of course be unable to reproduce the observed decline in response frequency as an individual NK cell encounters more target cells.

Thus in our model, immune cells will gradually reduce the reaction against one type of targets if they keep interacting with those targets since they are changing their internal representation of the environmental signal distribution and are adapting to the new environment. This agrees with the experimental result that the killing rate decreases in a serial killing event [80]. According to the experimental data, we can get a set of parameters, which could provide a similar serial killing result (**Figure 3.10**).

3.4.3 Comparisons with experiments that NK cells transferred from MHC I deficient mice to normal mice

Finally, we analyzed data from an experiment in which NK cells were transferred between between MHC-deficient mice and normal mice [24]. After transferred to the new host, transferred NK cells were then measured for the production of IFN_γ in response to anti-NK1.1 stimulation. The authors observed that NK cell responses on the host mouse: MHC-deficient mice were hyporesponsive compared to normal mice. Mature NK cells transferred from one type of mouse to another adapted over the course of several days to respond like those that had originally developed in the mouse to which they were transferred, demonstrating the plasticity of NK cell responses.

To compare our model with this data, we fit the mean of the signal distribution in the MHC-deficient mice μ_D and the anti-NK1.1 stimulus μ_A, assuming that the signal distribution in normal mice follows a normal distribution with mean zero. As in the previous cases, we fixed the standard deviations of the signals in each environment to one. Here, we fit these parameters simply based on the typical probability of response to anti-NK1.1 stimulation for NK cells from normal or MHC-deficient environments, yielding $\mu_A = 1.5$ and $\mu_D = 0.6$. For consistency with with the inferred memory parameters in the past two experiments, we set $\alpha = 20$. Assuming roughly one NK-target cell interaction per hour, these parameters fit very well with the experimental data (**Figure 3.11**).

3.5 Comparison with explicitly time-dependent inference methods

The model that we have described above is able to capture signals in changing environments through the use of finite, fixed memory parameters α and κ. However, this approach does not explicitly model a time-varying signal distribution (for example, one characterized by a time-dependent mean μ_t and standard deviation σ_t). One may consider how the approach that we have used would compare with inference methods that attempt to estimate parameters for a time-varying distribution.

3.5.1 Dynamic Bayesian network estimation

To develop such a model, we used a dynamic Bayesian network (DBN) to estimate time-dependent signal means and precisions from data. In this framework, we assume that

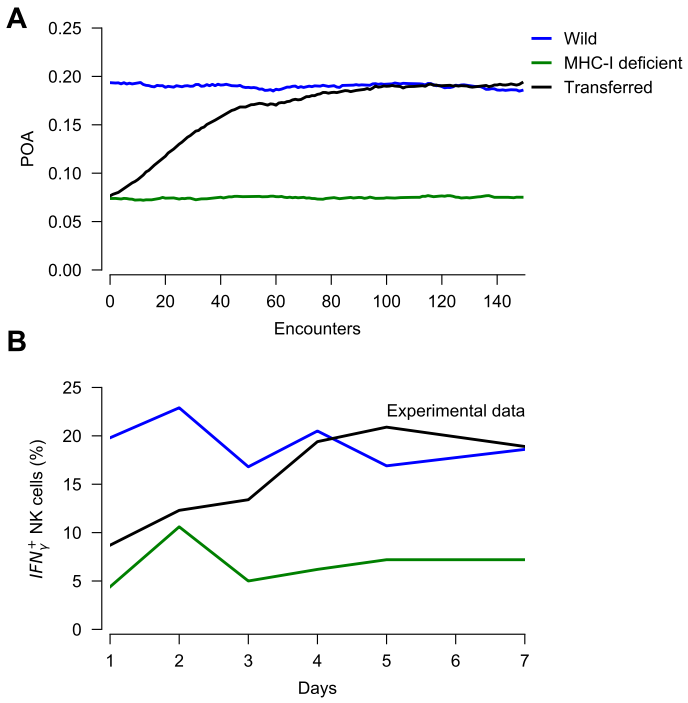

Figure 3.11: **Our model could predict the changes of killing ability when NK cells from MHC-I deficient mice are transferred to wild type mice.** **A** is the simulation result, which shows the changes of the probability of activating(POA) when NK cells encounter more and more targets in a new environment. **B** is the experimental result, which shows the changes in the proportion of NK cells that responded to stimuli. The blue line corresponds to transferring NK cells from wild-type mice to wild-type mice. The green line corresponds to transferring NK cells from MHC-I deficient mice to MHC-I deficient mice. The black line corresponds to transferring NK cells from MHC-I deficient mice to wild type mice. Here we assume signals from wild type mice satisfy the standard normal distribution,$N(0, 1^2)$, and memory $\alpha = 20$. We assume signals from MHC-I deficient mice and testing stimulus signals satisfy normal distributions as well but with different means. We use the ELFI for the parameters fitting. Since signals from MHC-I deficient mice and stimuli signals are more activating than healthy signals, the means must be positive and comparable to variance. We set the prior function for the mean as $P(\mu_{a,d}) \sim \frac{x}{4}e^{-\frac{x}{2}}$, whose domain is $(0, \infty)$. And we got $N(0.59, 1^2)$ for MHC-I deficient environment and $N(1.49, 1^2)$ for stimulus signals. Then we simulated the response of NK cells in these two environments and the response of NK cells after being transferred from MHC-I deficient environment to wild type environment. It shows that our model could capture the recovery of killing ability when NK cells are transferred from MHC-I deficient environment to a wild environment.

data is drawn from a normal distribution with mean μ_t and precision λ_t. For mathematical convenience, we will use a change of variables to work with the natural logarithm of the

precision $\theta_t = \ln(\lambda_t)$ instead of λ_t directly. The posterior probability for the parameters μ_t, θ_t conditioned on all past data $x_{1:t} = \{x_1, x_2, \ldots, x_t\}$ is then

$$P(\mu_t, \theta_t | x_{1:t}) = P(x_t | \mu_t, \theta_t) \iint d\mu_{t-1}\, d\theta_{t-1}\, \frac{P(\mu_t, \theta_t | \mu_{t-1}, \theta_{t-1}) P(\mu_{t-1}, \theta_{t-1} | x_{1:t-1})}{Z},$$

where Z is a normalization factor. ...

We also need to define how the (unobserved) signal mean and log-precision vary over time. Here for simplicity we will assume that both variables are related to each other by independent normal distributions, that is,

$$P(\mu_t, \theta_t | \mu_{t-1}, \theta_{t-1}) \propto e^{-\frac{\kappa_D}{2}(\mu_t - \mu_{t-1})^2} e^{-\frac{\kappa_D}{2}(\theta_t - \theta_{t-1})^2}.$$

Here κ_{DBN} is a parameter that describes how closely-related signal means and precisions from adjacent time points are to one another. In other words, this parameter describes a natural 'memory length' for signals in the environment.

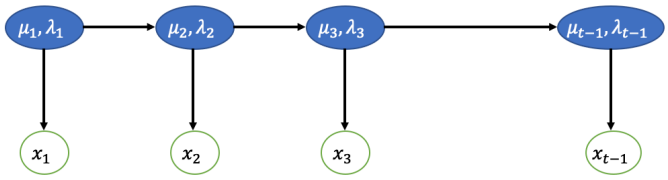

Figure 3.12: **Inference in Dynamic Bayesian Networks.** This shows the structure of the network. The value of observation in time point i, x_i, is determined by the hidden parameters μ_i, λ_i, which change over time. We have to estimate the changing hidden parameters according to what we have observed.

3.5.2 Derivative of posterior derivation

To model the time-varying signal distribution, we used dynamic Bayesian network (DBN) to estimate the mean $\mu_{1:t}$ and precision $\lambda_{1:t}$ sequences from observed signals sequence

$x_{1:t}$. Given the network described in **Fig. 3.12** that signal at time t, x_t, depends on (μ_t, λ_t) and (μ_t, λ_t) depends on previous values $(\mu_{t-1}, \lambda_{t-1})$, the joint probability would be

$$P(\mu_{0:t}, \lambda_{0:t}, x_{1:t}) = P(\mu_0, \lambda_0) \prod_{k=1}^{t} P(\mu_k, \lambda_k | \mu_{k-1}, \lambda_{k-1}) \prod_{k=1}^{t} P(x_k | \mu_k, \lambda_k) \,.$$

$P(\mu_0, \lambda_0)$ would be the prior distribution which is the initial guess of parameters distribution and does not depend on observations. $P(x_k|\mu_k, \lambda_k)$ describes the how likely x_k is observed given μ_k, λ_k which is a normal distribution $x_k \sim N(\mu_k, \lambda_k)$. $P(\mu_k, \lambda_k|\mu_{k-1}, \lambda_{k-1})$ describes the relation between parameters in the previous step $(k-1)$ and step k.

Without further information, a reasonable restriction is that values of one parameter in adjacent steps should be close to each other. For simplicity, we used normal distributions that the present value satisfy a normal distribution which is centered around the previous value. For the mean and precision, we assume $\mu_k \sim N(\mu_{k-1}, 1/\kappa_D)$ and $\lambda_k \sim N(\lambda_{k-1}, 1/\kappa_D)$, where κ_D is the parameter determines the "closeness" between values in adjacent steps. The full posterior is

$$P(\mu_{0:t}, \lambda_{0:t}|x_{1:t}) = P(\mu_0, \lambda_0) \prod_{k=1}^{t} P(\mu_k, \lambda_k | \mu_{k-1}, \lambda_{k-1}, x_{1:k}) \,.$$

The posterior at time $t (t \geq 1)$ is

$$\begin{aligned} P(\mu_t, \lambda_t | x_{1:t}) &= \iint d\mu_{t-1}\, d\lambda_{t-1} P(\mu_{t-1}, \lambda_{t-1}|x_{1:t-1}) P(\mu_t, \lambda_t | \mu_{t-1}, \lambda_{t-1}, x_{1:t}) \\ &= \iint d\mu_{t-1}\, d\lambda_{t-1} P(\mu_{t-1}, \lambda_{t-1}|x_{1:t-1}) P(\mu_t, \lambda_t | \mu_{t-1}, \lambda_{t-1}, x_{1:t-1}) P(x_t|\mu_t, \lambda_t) \\ &= P(x_t|\mu_t, \lambda_t) \iint d\mu_{t-1}\, d\lambda_{t-1} P(\mu_{t-1}, \lambda_{t-1}|x_{1:t-1}) P(\mu_t, \lambda_t | \mu_{t-1}, \lambda_{t-1}) \,. \end{aligned} \tag{3.10}$$

For a given prior $P(\mu_0, \lambda 0)$ and observations $x_{1:t}$, we can numerically solving equation 3.10 to calculate the posterior distribution for each step. We use the location of the peak point as

the estimated mean and logarithm of precision. For instance, we generated observations $x_{1:t}$ from normal distributions $N(\mu_t, 1/\lambda_t)$ where $\mu_t = 4\sin\frac{2\pi t}{T}$, $\lambda_t = 1 + \frac{1}{2}\sin\frac{2\pi t}{T}$, and $T = 100$ is the period. We used normal distributions for prior distribution, $P(\mu_0) = N(0,1), P(\lambda_0) = N(2.5,1)$. We solved the equation 3.10 in the domain of $\mu \times \lambda = [-10,10] \times (0,5]$. The results for different κ_D are shown in **Fig. 3.13** where for smaller κ_D (κ_D is much smaller than the period of true mean and precision), the estimation captures the change of true values. For larger κ_D, the estimation is more stable which is reasonable since this agrees with the definition of κ_D. The step length $dx = 0.1$ whcih divided the domain to a 200×50 grid. This step length is small enough to give good result. Using a smaller step length $dx = 0.05$ increases the computational cost a lot but does not increase accuracy much.

Our model has similar performance as the DBN model (shown in next section.) Now we are curious what is the update mechanism of estimated mean and precision in DBN model. How is it comparing with our model? Could we build a connection between the memory parameter κ in our model and κ_D in the DBN model? Next we are going to make a comparison of these two models. We will see that the update of parameters of two models are kind of similar to each other and we can establish a rough connection between κ and κ_D by using variational Bayesian approximation.

3.5.3 Comparison of the update mechanism of our model and the DBN model

For DBN model, the posterior at time t is given by equation 3.10. The integral is difficult to calculate after the first step ($t \geq 2$), which results in the posterior cannot be

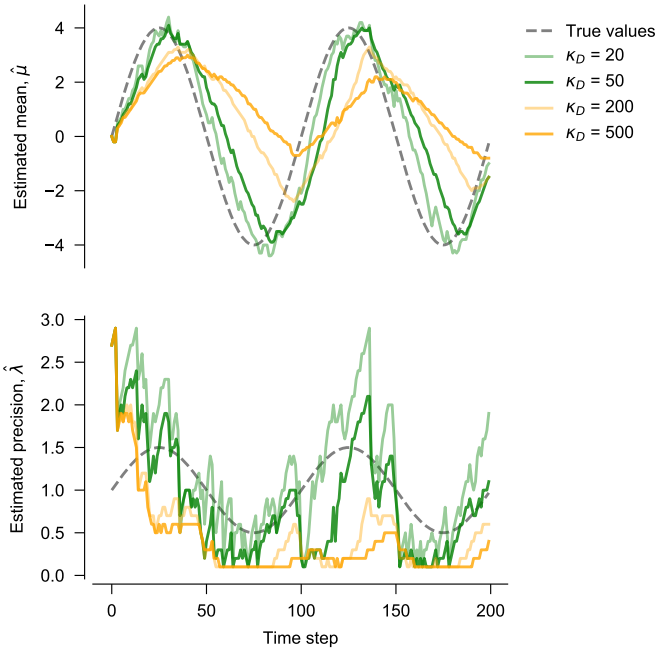

Figure 3.13: **DBN gives a good estimation of the mean and precision.** For sinusoidal mean and precision, small κ_D gives good adaption. Large κ_D gives more stable estimation. To some extent, κ_D is similar to the memory parameters, α and κ. Here we run the simulation for 200 steps. At each step, one data point is drawn from normal distribution with true mean μ and true precision λ. The dash lines are the curves for true values. We use $4\sin\frac{2\pi t}{100}$ as the dynamic true μ and $1+\frac{1}{2}\sin\frac{2\pi t}{100}$ as the true λ.

solved analytically. However, since the distribution is convex, proved in numerical result, we can calculate the posterior peak point, see how it is connected to the previous distribution, and use this as an indicator of the DBN model update mechanism. The posterior at time t is

$$P(\mu_t, \lambda_t | x_t) = P_L(x_t|\mu_t, \lambda_t) \iint d\mu_{t-1}\, d\lambda_{t-1} P_0(\mu_{t-1}, \lambda_{t-1}) P_T(\mu_t, \lambda_t | \mu_{t-1}, \lambda_{t-1}) \,,$$

where

$$P_L(x_t|\mu_t, \lambda_t) = \sqrt{\frac{\lambda_t}{2\pi}} e^{-\frac{1}{2}\lambda_t(x_t-\mu_t)^2}$$

and

$$P_T(\mu_t, \lambda_t|\mu_{t-1}, \lambda_{t-1}) = \sqrt{\frac{\kappa_D}{2\pi}} e^{-\frac{\kappa_D}{2}(\mu_t-\mu_{t-1})^2} \sqrt{\frac{\kappa_D}{2\pi}} e^{-\frac{\kappa_D}{2}(\lambda_t-\lambda_{t-1})^2}.$$

At the peak point, $\frac{\partial P(\mu_t,\lambda_t|x_t)}{\partial \mu_t} = 0, \frac{\partial P(\mu_t,\lambda_t|x_t)}{\partial \lambda_t} = 0$. Substituting the posterior into the equation, we will get the result:

$$\mu_t = \frac{\lambda_t}{\lambda_t + \kappa_D} x_t + \frac{\kappa_D}{\lambda_t + \kappa_D} \frac{\iint \mu_{t-1} P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}{\iint P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}.$$

$$\lambda_t = \frac{1}{2}\left[\langle\lambda_{t-1}\rangle - \frac{(x_t-\mu_t)^2}{2\kappa_D} + \sqrt{\frac{2}{\kappa_D} + \left[\langle\lambda_{t-1}\rangle - \frac{(x_t-\mu_t)^2}{2\kappa_D}\right]^2}\right],$$

where $\langle\lambda_{t-1}\rangle = \frac{\iint \lambda_{t-1} P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}{\iint P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}$. Calculation details are attached at the end of this section. For the update of μ, the peak of posterior is determined by the linear combination of the new observation and some transition of historical learned distribution. The update of precision is a little bit complicated, while for large κ_D, the last term under the square root can be simplified as:

$$\begin{aligned}
&\sqrt{\frac{2}{\kappa_D} + \left[\langle\lambda_{t-1}\rangle - \frac{(x_t-\mu_t)^2}{2\kappa_D}\right]^2} \\
&= \sqrt{\langle\lambda_{t-1}\rangle^2 - 2\langle\lambda_{t-1}\rangle\frac{(x_t-\mu_t)^2}{2\kappa_D} + (\frac{(x_t-\mu_t)^2}{2\kappa_D})^2 + \frac{2}{\kappa_D}} \\
&\approx \langle\lambda_{t-1}\rangle + \frac{1}{2\langle\lambda_{t-1}\rangle}\left[-2\langle\lambda_{t-1}\rangle\frac{(x_t-\mu_t)^2}{2\kappa_D} + (\frac{(x_t-\mu_t)^2}{2\kappa_D})^2 + \frac{2}{\kappa_D}\right] \\
&\approx \langle\lambda_{t-1}\rangle - \frac{(x_t-\mu_t)^2}{2\kappa_D} + \frac{1}{\kappa_D\langle\lambda_{t-1}\rangle}
\end{aligned}$$

The second step used Taylor expansion, $\sqrt{x+d} \approx \sqrt{x} + \frac{1}{2\sqrt{x}}d, d \ll 1$. Thus the update of precision can be approximated as

$$\lambda_t = \langle\lambda_{t-1}\rangle + \frac{1}{2\kappa_D}\left[\frac{1}{\langle\lambda_{t-1}\rangle} - (x_t-\mu_t)^2\right].$$

In summary, the update of the DBN model is

$$\begin{aligned}\mu_t &= \frac{\lambda_t}{\lambda_t+\kappa_D}x_t + \frac{\kappa_D}{\lambda_t+\kappa_D}\frac{\iint \mu_{t-1}P_0P_T d\mu_{t-1}\,d\lambda_{t-1}}{\iint P_0P_T d\mu_{t-1}\,d\lambda_{t-1}}\,. \\ \lambda_t &= \langle\lambda_{t-1}\rangle + \frac{1}{2\kappa_D}\left[\frac{1}{\langle\lambda_{t-1}\rangle} - (x_t-\mu_t)^2\right]\end{aligned} \tag{3.11}$$

In our model, we used the mean value of μ and λ as the estimated mean and precision equation 2.2, which is $\hat{\mu} = m, \hat{\sigma^2} = \frac{\beta}{\alpha-1}$. Since the normal-gamma function (distribution of posterior in our model) is convex, we can show that the peak point is close to the mean point. The peak point of our model is calculated by $\frac{\partial P(\mu,\lambda|m,\beta,\alpha,\kappa)}{\partial\mu} = 0, \frac{\partial P(\mu,\lambda|m,\beta,\alpha,\kappa)}{\partial\lambda} = 0$, which gives $\mu = m, \lambda = \frac{\alpha-\frac{1}{2}}{\beta}$. The update of the peak is $\mu' = m' = \frac{\kappa\mu+x}{\kappa+1}$ and $\lambda' = \frac{\alpha-\frac{1}{2}}{\beta'}$, substituting β' from equation 3.5 into it we will get $\frac{1}{\lambda'} = \frac{\alpha-1}{\alpha-\frac{1}{2}}\frac{1}{\lambda} + \frac{\frac{1}{2}}{\alpha-\frac{1}{2}}\frac{\kappa}{\kappa+1}(x-m)^2$ which is the same as the update of estimated mean and precision since $\sigma^2 = \frac{1}{\lambda}$. The update of precision can be approximated by

$$\begin{aligned}\frac{1}{\lambda'} &= \frac{\alpha-1}{\alpha-0.5}\frac{1}{\lambda} + \frac{0.5}{\alpha-0.5}\frac{\kappa}{\kappa+1}(x-m)^2 \\ \lambda' &= \frac{1}{\frac{\alpha-1}{\alpha-0.5}\frac{1}{\lambda} + \frac{0.5}{\alpha-0.5}\frac{\kappa}{\kappa+1}(x-m)^2} \\ &= \frac{\alpha-0.5}{\alpha-1}\lambda\frac{1}{1+\frac{0.5}{\alpha-1}\frac{\kappa}{\kappa+1}(x-m)^2\lambda} \\ &\approx \frac{\alpha-0.5}{\alpha-1}\lambda\left[1-\frac{0.5}{\alpha-1}\frac{\kappa}{\kappa+1}(x-m)^2\lambda\right] \\ &= \lambda + \frac{0.5}{\alpha-1}\lambda^2\left[\frac{1}{\lambda} - \frac{\alpha-0.5}{\alpha-1}\frac{\kappa}{\kappa+1}(x-m)^2\right]\,.\end{aligned}$$

In summary, the update of our model is

$$\begin{aligned}\mu' &= \frac{\kappa\mu+x}{\kappa+1} \\ \lambda' &= \lambda + \frac{0.5}{\alpha-1}\lambda^2\left[\frac{1}{\lambda} - \frac{\alpha-0.5}{\alpha-1}\frac{\kappa}{\kappa+1}(x-m)^2\right]\,.\end{aligned} \tag{3.12}$$

The formats of update of our model and DBN model, equation 3.11 and equation 3.12, are similar to each other. The main difference is that the change in DBN model is mainly

determined by κ_D while in our model, it mainly determined by α, κ. In the next section, we are trying to build a connection between them.

Following is the calculation details for the peak point of posterior distribution. For the partial difference over μ_t, we have

$$
\begin{aligned}
0 &= \frac{\partial P(\mu_t, \lambda_t | x_t)}{\partial \mu_t} \\
0 &= \iint d\mu_{t-1}\, d\lambda_{t-1} P_0(\mu_{t-1}, \lambda_{t-1}) \frac{\partial}{\partial \mu_t} [P_L(x_t|\mu_t, \lambda_t) P_T(\mu_t, \lambda_t | \mu_{t-1}, \lambda_{t-1})] \\
0 &= \iint d\mu_{t-1}\, d\lambda_{t-1} P_0 [(\frac{\partial}{\partial \mu_t} P_L) P_T + P_L (\frac{\partial}{\partial \mu_t} P_T)] \\
0 &= \iint d\mu_{t-1}\, d\lambda_{t-1} P_0 [-\lambda_t (\mu_t - x_t) P_L P_T - \kappa_D (\mu_t - \mu_{t-1}) P_L P_T]
\end{aligned}
$$

This gives :

$$
\begin{aligned}
&(\lambda_t + \kappa_D)\mu_t \iint d\mu_{t-1}\, d\lambda_{t-1} P_0 P_L P_T \\
&\qquad = \lambda_t x_t \iint d\mu_{t-1}\, d\lambda_{t-1} P_0 P_L P_T + \kappa_D \iint d\mu_{t-1}\, d\lambda_{t-1} \mu_{t-1} P_0 P_L P_T \\
&\mu_t = \frac{\lambda_t}{\lambda_t + \kappa_D} x_t + \frac{\kappa_D}{\lambda_t + \kappa_D} \frac{\iint \mu_{t-1} P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}{\iint P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}} \,.
\end{aligned}
$$

For the partial difference over λ_t, we have

$$
\begin{aligned}
0 &= \frac{\partial P(\mu_t, \lambda_t | x_t)}{\partial \lambda_t} \\
0 &= \iint d\mu_{t-1}\, d\lambda_{t-1} P_0(\mu_{t-1}, \lambda_{t-1}) \frac{\partial}{\partial \lambda_t} [P_L(x_t|\mu_t, \lambda_t) P_T(\mu_t, \lambda_t | \mu_{t-1}, \lambda_{t-1})] \\
0 &= \iint d\mu_{t-1}\, d\lambda_{t-1} P_0 [(\frac{\partial}{\partial \lambda_t} P_L) P_T + P_L (\frac{\partial}{\partial \lambda_t} P_T)] \\
0 &= \iint d\mu_{t-1}\, d\lambda_{t-1} P_0 [\frac{1}{2} \lambda_t^{-1} - \frac{1}{2} (\mu_t - x_t)^2 P_L P_T - \kappa_D (\lambda_t - \lambda_{t-1}) P_L P_T]
\end{aligned}
$$

This gives

$$
\frac{1}{2} \lambda_t^{-1} - \frac{1}{2} (\mu_t - x_t)^2 - \kappa_D (\lambda_t - \frac{\iint \lambda_{t-1} P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}{\iint P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}) = 0 \,.
$$

Considering $\lambda_t > 0$, the solution is

$$\lambda_t = \frac{1}{2}\left[\langle\lambda_{t-1}\rangle - \frac{(x_t-\mu_t)^2}{2\kappa_D} + \sqrt{\frac{2}{\kappa_D} + \left[\langle\lambda_{t-1}\rangle - \frac{(x_t-\mu_t)^2}{2\kappa_D}\right]^2}\right], \tag{3.13}$$

where $\langle\lambda_{t-1}\rangle = \frac{\iint \lambda_{t-1} P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}{\iint P_0 P_T d\mu_{t-1}\, d\lambda_{t-1}}$.

3.5.4 Connection between memory parameters in Modified Bayesian inference and transition factor in Dynamic Bayesian Network

Since the update mechanisms are similar to each other, we wondered if we could build a connection between the memory parameter κ and the transition factor κ_D, i.e. for given κ_D to find the equivalent κ, then run the simulation and compare the result which gives the comparison of performance between our model and the DBN model.

Since the posterior of DBN model is difficult to calculate, here we used variational Bayesian approximation to approximate the posterior. Variational Bayesian methods is often used in dealing with complicated posterior in Bayesian inference. Instead of calculating the true posterior, it finds a approximate distribution Q to approximate the posterior $P(\mu_t, \theta_t|x_{1:t})$. The Kullback–Leibler(KL) divergence, $D_{KL} = \int Q \ln \frac{Q}{P}$, describes the 'closeness' of P and Q. We used the mean-field form of variational Bayesian in which Q is factorized into single-variable factors, $Q(Z) = \prod_i q(z_i)$ where $Z = \{z_1, ..., z_i\}$ are the unknown variables and for every $q(z_i)$, it is normalized function $\int dz_i q(z_i) = 1$. It has been proved that the KL divergence gets the minimum value when $q(z_i) = \frac{1}{Z} exp\,\langle E(z_i, \bar{z}_i, D)\rangle_{Q(\bar{z}_i)}$ for all the variables[31], where D is the observed data, $\bar{z}_i$ means all other variables except z_i, $E(z_i, \bar{z}_i, D) = \ln P(Z, D)$ and $\langle E(z_i, \bar{z}_i, D)\rangle_{Q(\bar{z}_i)}$ means integrating $E(z_i, \bar{z}_i, D)$ and $Q(\bar{z}_i)$ over $\bar{z}_i$.

In our case, the factorization can be written as $Q(\mu_t, \theta_t) = q(\mu_t)q(\theta_t)$. The optimal Q is achieved when $\ln q(\mu_t) = \int q(\theta_t) \ln P(\mu_t, \theta_t, x_t)\, d\theta_t + \ln Z_{\mu_t}$ and $\ln q(\theta_t) = \int q(\mu_t) \ln P(\mu_t, \theta_t, x_t)\, d\mu_t + \ln Z_{\theta_t}$. The Z_{μ_t} and Z_{θ_t} are for normalisation. We are going to prove that if we use a normal distribution for $q(\mu_{t-1})$, the optimal $q(\mu_t)$ would be a normal distribution as well. Thus if we use a normal distribution as the prior for $P_0(\mu_0)$ in $P_0(\mu_0, \lambda_0) = P_0(\mu_0)P_0(\lambda_0)$, we will get normal distributed $q_{t=1,..,T}(\mu_t)$.

We assume that $q(\mu_{t-1}) \sim N(m_{t-1}, 1/r_{t-1})$, thus $P(\mu_{t-1}, \lambda_{t-1})$ is replaced with $q(\mu_{t-1})q(\lambda_{t-1})$. The posterior $P(\mu_t, \lambda_t|x_t)$ can be written as :

$$\begin{aligned}
P(\mu_t, \lambda_t|x_t) &= P(x_t|\mu_t, \lambda_t) \iint d\mu_{t-1}\, d\lambda_{t-1} P(\mu_{t-1}, \lambda_{t-1}) P(\mu_t, \lambda_t|\mu_{t-1}, \lambda_{t-1}) \\
&\propto \sqrt{\lambda_t} e^{-\frac{1}{2}\lambda_t(x_t-\mu_t)^2} \iint d\mu_{t-1}\, d\lambda_{t-1} \{e^{-\frac{1}{2}[r_{t-1}(\mu_{t-1}-m_{t-1})^2 + \kappa_D(\mu_t-\mu_{t-1})^2]} \\
&\quad \times e^{-\frac{\kappa_D}{2}(\lambda_t-\lambda_{t-1})^2} q(\lambda_{t-1})\} \\
&\propto \sqrt{\lambda_t} e^{-\frac{1}{2}\lambda_t(x_t-\mu_t)^2} \times e^{-\frac{1}{2}\frac{\kappa_D r_{t-1}}{\kappa_D + r_{t-1}}(\mu_t - m_{t-1})^2} \times \int d\lambda_{t-1}\, e^{-\frac{\kappa_D}{2}(\lambda_t-\lambda_{t-1})^2} q(\lambda_{t-1})\,.
\end{aligned}$$

Since $\ln P(\mu_t, \lambda_t, x_t) = \ln P(\mu_t, \lambda_t|x_t) - \ln P(x_t)$ we first calculate the log-posterior

$$\begin{aligned}
\ln\left(P(\mu_t, \lambda_t|x_t)\right) \propto \frac{1}{2}\ln(\lambda_t) - \frac{1}{2}\lambda_t(x_t-\mu_t)^2 - \frac{1}{2}\frac{\kappa_D r_{t-1}}{\kappa_D + r_{t-1}}(\mu_t - m_{t-1})^2 \\
+ \ln(\int d\lambda_{t-1}\, e^{-\frac{\kappa_D}{2}(\lambda_t-\lambda_{t-1})^2} q(\lambda_{t-1}))\,.
\end{aligned}$$

For $q(\mu_t)$,

$$\begin{aligned}
\ln(q(\mu_t)) &= \int q(\lambda_t) \ln P(\mu_(t), \lambda_(t), x_(t))\, d\lambda_t + \ln Z_{\mu_t} \\
&\propto -\frac{1}{2}\langle\lambda_t\rangle(x_t-\mu_t)^2 - \frac{1}{2}\frac{\kappa_D r_{t-1}}{\kappa_D + r_{t-1}}(\mu_t - m_{t-1})^2
\end{aligned} \tag{3.14}$$

where $\langle\lambda_t\rangle = \int \lambda_t q(\lambda_t)\, d\lambda_t$. It is obvious that $q(\mu_t)$ is a normal distribution, we can calculate the location of peak value(mean) m_t and precision r_t by calculating the first and second order derivative of equation 3.14, $\frac{d\ln q(\mu_t)}{d\mu_t} = 0(\mu_t = m_t)$ and $\frac{d^2\ln(q(\mu_t))}{d^2\mu_t} = -r_t$. Substituting

$\ln q(\mu_t)$ into these two equations, we will get

$$m_t = \frac{\langle\lambda_t\rangle x_t + \frac{\kappa_D r_{t-1}}{\kappa_D + r_{t-1}} m_{t-1}}{\langle\lambda_t\rangle + \frac{\kappa_D r_{t-1}}{\kappa_D + r_{t-1}}}$$

$$r_t = \langle\lambda_t\rangle + \frac{\kappa_D r_{t-1}}{\kappa_D + r_{t-1}}.$$

The update of mean is similar to our model, where the update is

$$\langle\mu\rangle = m, \quad m^{'} = \frac{\kappa m + x}{\kappa + 1}$$

which means if κ is equivalent to

$$k = \frac{\kappa_D r_t}{\kappa_D + r_t} \frac{1}{\langle\lambda_t\rangle}, \tag{3.15}$$

they should give similar results. We calculated r_t as the inverse of variance of the mean $r_t = \frac{1}{var(\int P(\mu_t, \lambda_t | x_t)\, d\lambda)}$, and $\langle\lambda_t\rangle$ as the mean of precision, $\langle\lambda_t\rangle = \iint d\mu_t\, d\lambda_t\, \lambda_t P(\mu_t, \lambda_t | x_t)$. Here if we using the mean of sequence k defined by equation 3.15 as κ,

$$\kappa = \frac{1}{T}\sum_{t=1}^{T} k = \frac{1}{T}\sum_{t=1}^{T} \frac{\kappa_D r_t}{\kappa_D + r_t} \frac{1}{\langle\lambda_t\rangle},$$

our modified Bayesian inference and Dynamic Bayesian Network would have similar performance in estimating signal distribution of a changing environment(**Fig. 3.14**). In other word, we can always find a κ in our model which is equivalent to the κ_D in DBN and they have similar performance in estimating dynamic signal environments. Both models work well for a slow-varying distribution.

3.6 Inference with signal saturation

In the analysis above, the range of the signal x is in principle unbounded. However, immune cells have a finite number of receptors, limiting the magnitude of activating

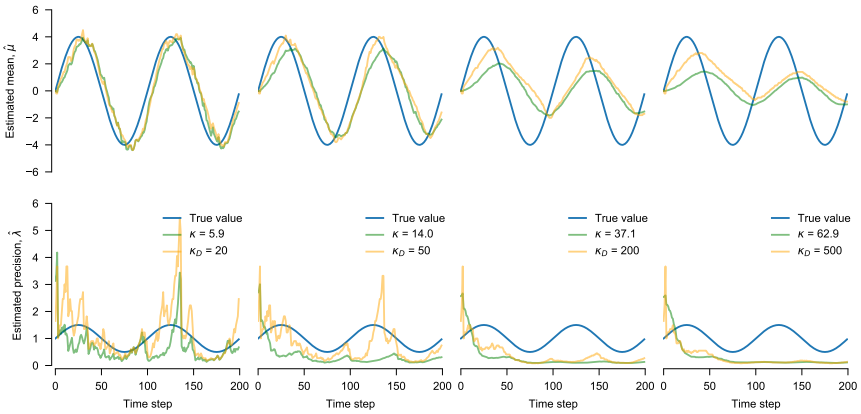

Figure 3.14: **Our model gives similar results if proper κ is chosen.** Using the mean value of sequence k defined by equation 3.15 as κ, our model has similar performance to DBN whether it is short memory or long memory. Here we use same parameters setting as Fig. 3.13. We run the simulation for 200 steps. Each step, one signal is generated from the normal distribution, $N(4\sin(\frac{2\pi t}{100}), 1/[1+\frac{1}{2}\sin(\frac{2\pi t}{100})])$. We first run DBN and calculate the equivalent κ. Using the equivalent κ and interact with the same signal sequence, our model performs as well as DBN in general. Thus we can always find a κ which is equivalent to κ_D in DBN. Our model is as good as DBN when estimating dynamic signal environments.

or inhibitory stimulus that they can receive. To mimic this saturation effect, we considered sigmoidally transforming signals x relative to the current estimated signal mean m. Specifically, we considered a transformed signal x' given by

$$x' = m + r\left[\frac{2}{1+e^{-(x-m)}} - 1\right],$$

where r sets the range of signals that can be received, which lies in $[m-r, m+r]$. Because strong activating and inhibitory signals are clipped to a finite range, adaptation to environments that differ substantially from the current one is slower (**Figs. 3.15, 3.16**). In this model, the estimated mean ultimately converges to the true signal mean but the standard deviation converges to a different, smaller value (which we call the transformed standard deviation, σ_{tf}) because the sigmoidal transformation reduces the signal variance.

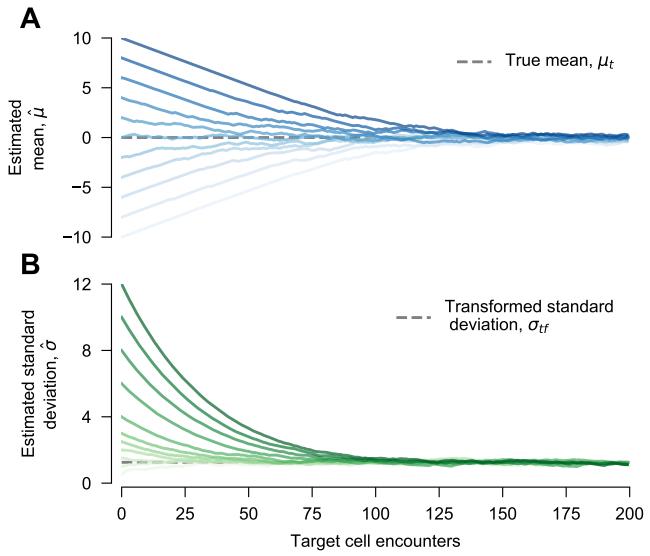

Figure 3.15: **Immune cell adaptation is more gradual with signal saturation.** The estimated mean (**A**) and standard deviation (**B**) converge toward the true mean and transformed standard deviation in a model with signal saturation. Parameters and initial conditions are the same as those in **Fig. 2**. Signal saturation changes the expected standard deviation of the signal when the model is perfectly adapted because large deviations are suppressed. We have thus replaced the true standard deviation σ_t with the transformed standard deviation σ_{tf} in **B**.

3.7 Conclusion

Theoretical analyses of immunity have often focused on the adaptive immune system and antigen-specific recognition of foreign material to distinguish self from nonself [26, 29, 78, 69]. Here, we described an 'algorithm' for self/nonself discrimination that operates in a very different way. Rather than learning to detect specific pathogens, our model immune cells learn the properties of *healthy* cells in their current environment, which allows them to respond to aberrant cells that may be infected, stressed, or transformed. Learning in our model is 'unsupervised' in the sense that immune cells do not have external information about whether the targets that they encounter are healthy or not. Nonetheless, it operates reliably following the simple assumption that the great majority of targets that

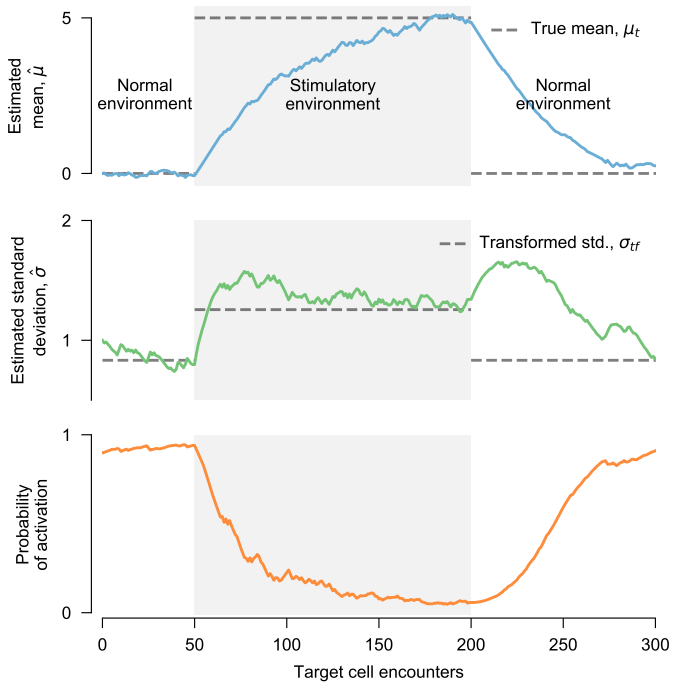

Figure 3.16: **Immune cell adaptation is more gradual with signal saturation.** As in **Fig. 3.4**, an immune cells adapts to different signal distributions in 'normal' and 'stimulatory' environments in a model with signal saturation. Adaptation is more gradual due to signal saturation, so we have extended the number of encounters in the stimulatory environment to allow for complete adaptation. As in **Fig. 3.4**, the immune cell is able to reliably activate against aberrant targets when it is adapted to a normal environment, but it loses this ability after long times in a stimulatory environment. Compared to a model without signal saturation, adaptation is more gradual and less noisy. In addition, the immune cell retains the ability to activate against aberrant targets for a longer time after being placed in the stimulatory environment.

immune cells encounter are likely to be normal, healthy cells. Our model captures multiple experimentally observed behaviors of innate immune cells. These include adapting responses to different environments, the development of hyposensitivity after prolonged exposure to stimulus, and the eventual recovery of normal function after the stimulus is withdrawn. At present, however, little data exists to quantitatively test predictions for how past encounters with target cells affect future responses. Future measurements of the kinetics of adaptation would be of great interest.

Recent work has also applied ideas Bayesian inference to immunity, focusing in particular on rules for optimizing the adaptive immune repertoire [68, 69]. There, models were constructed to optimally allocate immune cells with different antigen specificities to combat a shifting environment of pathogens. Our work is similar in that we also consider adaptation to varying environments. However, our model concerns the adaptation of single immune cells, and do not make detailed about how the local environment will vary over time. The parameters κ and α, however, do set a natural time scale over which memory of the environment is retained. Other intriguing studies have developed analogies between machine learning and immunity [83, 109], including a model of negative selection in T cells, where encounters with self peptides are central [109]. More generally, the problem of estimating time-varying signal distributions has some similarities with estimation using Kalman filters [105, 46]. An important difference in the present case is that the signal variance must also be estimated, and the way that the signal mean and variance change is unknown.

Similarly, our model can be compared with the discontinuity theory of immunity [79], which posits that immune cells in general respond to sharp changes in the environment. One of the main differences between our model and the discontinuity theory is that we explicitly consider the *variance* of signals in the environment. Though additional experiments will be needed to explore these models in quantitative detail, there is some evidence that variance in ligand expression is important. A recent experiment showed that MHC class I-deficient hematopoietic cells are spared in mice that also have hematopoietic cells with normal levels of MHC class I expression, but only if the MHC class I-deficient cells

comprise a substantial fraction of all cells [6]. Our model can be extended to incorporate additional features of immune-target interactions. One important extension of the model would be to explicitly consider signaling through multiple pairs of receptors and ligands. NK cells and macrophages use a wide array of receptors [55, 95]. Intriguingly, recent studies have revealed dramatic heterogeneity in the complement of receptors that individual NK cells express [43, 90]. Quantifying the ability of populations of immune cells with different patterns of receptor expression to collectively recognize target cells could shed light on principles governing the innate immune repertoire. Though we have focused on cells of the innate immune system, inspired especially by NK cells and macrophages, the model we developed may also apply more broadly to other cell types. There are some conceptual similarities between our model and the 'tunable threshold' model, originally applied to T cell signaling [38, 39]. Recent work demonstrated that T cells adapt to basal levels of T cell receptor (TCR) signaling, and that cells with stronger basal signaling were less responsive [114]. Importantly, this work also demonstrated that even T cells with identical TCRs exhibit heterogeneous responses to stimulus [114], which is one of the main predictions of our model.

3.8 Data and code

Data and code used in our analysis is available at the GitHub repository `https://github.com/bartonlab/paper-innate-immune-adaptation`. This repository also contains Jupyter notebooks that can be run to reproduce the results presented here.

Chapter 4

Extend the model to high dimensional signal processing

We have talked about our one-dimensional model, which we assumed the signal is a real number and did not take the signal generation process into consideration. However, we know that there are dozens of known receptors that play a role in activating or inhibiting NK cells. Next we would like to talk about the signal generation process. We studied the distribution of receptors in single NK cell and the receptor distributions in NK populations, which gives interesting results. From experimental single cell sequencing, we know that the number of expressed receptor(types) on each NK cell is small. We compared this sparse distribution with other types distribution like express all and uniform distribution, we found that sparse expression has some advantages, such as increasing the signal to noise ratio when reacting to unhealthy signals, increasing diversity and costing less energy.

4.1 Introduction to receptors distribution of NK cells

Researchers in the field have identified numerous types of activating and inhibitory receptors. However, the expression of these receptors on natural killer (NK) cells is not yet fully understood. Typically, in a single experiment, only a subset of NK cells expressing specific receptors is investigated. As more and more studies reveal the large number of NK cell phenotypes, this diversity has sparked considerable interest among researchers.

It's important to clarify that NK cells do not express all these receptors simultaneously. Rather, each individual NK cell expresses a limited fraction of the available receptors. When considered at the population level, this results in a complex assemblage of NK cells, each expressing a different set of receptors and forming a diverse, multifaceted population. Collectively, these varied NK cells provide robust protection for the host, with each subset of NK cells targeting specific types of threats.

The distribution of receptors on NK cells is likely a complex and dynamic process. For the purpose of simplicity in our discussion, we'll set aside the specifics of which receptors are expressed by individual NK cells. Instead, we'll focus on the fact that NK cells do not express all activating and inhibitory receptors. We'll consider the diversity in receptor expression across NK cells, and in conjunction with the adaptive ability of these cells, we aim to explore why this is the case and what advantages this setup may confer.

4.2 Sparse receptors distribution

In the paper, "Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry" by Horowitz et al. [43], they used mass cytometry to research

on the diversity of receptor expression of NK cells. Given the 28 receptors expression on 12 healthy and independent donors and 5 pairs of monozygotic twins, the results showed that the expression of NK cell receptors was super diverse. No phenotype accounts for more than 7% of total NK cells. In the top 50 most frequent phenotypes, only 14 out of the 28 receptors were expressed. The top 50 most frequent phenotypes account for only 15% of total cells. Not only the activating receptors, the expression of inhibitory receptors is also low[43].

As mentioned before, instead of focusing on the specific receptor expressed by a NK cells, we focused more on the distribution of receptors. First, we looked at the distribution of number of receptors expressed by each NK cell. Thanks to Amir Horowitz, we got the experimental data used in the paper[43]. The data comes from mass cytometry technique. Next I will first given an introduction to the data and mass cytometry, then I will show the result of distribution of receptors.

4.2.1 Introduction to mass cytometry

Mass cytometry represents a novel technology employed in the examination of receptors and small molecules in individual cells. It operates on the principle of using antibodies that bind to distinct elements. As these elements possess different mass, the measurement of this mass via mass spectrometry can reveal the quantity of antibodies binding to a particular element.

This innovative technology stands in contrast to traditional flow cytometry, which uses fluorescent dyes to detect the presence of specific substances. Instead, mass cytometry utilizes elements or metal tags for detection, providing several benefits. For one, the over-

lap between different colors, a common issue in traditional flow cytometry, is no longer a concern with mass cytometry. Additionally, metal tags are typically more stable compared to fluorescent dyes, enhancing the reliability of detection.

Another significant advantage of mass cytometry is the sheer range of different masses it can detect. The number of elements with distinct masses greatly outnumbers the variety of fluorescent dyes available. Consequently, mass cytometry is capable of providing a high-dimensional perspective of receptor expression. It enables the simultaneous measurement of many receptor expressions on a single cell, offering a more comprehensive view of cellular function and composition.

Here we are going to use the data from mass cytometry to find the receptor distribution of NK cells. The high-dimensional data generated through this technology will allow us to gain a deeper understanding of the diversity and specificity of receptor expression within the NK cell population. By harnessing the power of mass cytometry, we aim to unlock new insights into the complex world of cellular interactions and functions.

4.2.2 Introduction to the experimental data

The provided data set comprises 22 "*.fcs" files, each of which corresponds to data from a unique individual. These files encompass data from 12 independent subjects and five sets of twins. With the use of RStudio, we're able to import this data and transform it into a more manageable dataframe format. In the dataframe, rows represent individual cells and columns correspond to measured parameters. These parameters include the 28 receptors under investigation. Each row provides the receptor readout for a single cell, offering a detailed view into the individual cellular responses within the larger dataset.

4.2.3 Data analysis

The first thing to do with the data is identifying NK cells. Because the cells are from donors, not cell lines, thus we need to first select NK cells out of all the cells. In the paper, they mainly used a sequential gating to identify NK cells. Here we use the same method to define NK cells. Please look at the work flow demonstrated in Figure S2 in the paper[43] for details. First by selecting rows that satisfy $2000 > DNA1.Ir191 > 100, 2000 > DNA2.Ir193 > 100$, we select all intact cells. Here $Ir191, Ir193$ are two particles attached to DNA in this experiment, and thus wrote as $DNA1.Ir191$. In the following, $A.B$ format means particle B is attached to A and in experiment, we detect the particle to show the abundant of parameter A. Second, we select cells with short cell length $celllength < 65$ and smaller $In115 < 200$ to select those live singlet cell since live cells will exclude DOTA-maleimide and in the experiment DOTA-maleimide is attached with particle $In115$. Third, we select those $CD3.Cd112 < 100$ to exclude T cells. Fourth, we select cells with smaller $CD33$ to avoid macrophages, $CD33.Tb159 < 100$. Fifth, we select cells with low expression of $CD19$ to exclude B cells, $CD19.Nd142 < 5$. Finally, NK cells are defined as expressing either $CD56$ or $CD16$, excluding those $CD56$ negative cells expressing $(HLA - DR,\ HLA.DR.Yb172 < 150 | CD56.Yb174 > 20)$ and $(2500 > CD56.Yb174 > 5 | 1000 > CD16.Sm149 > 15)$.

For each donor, we used the same conditions to select NK cells. Finally, we got the number of NK cells identified from donors ranged from 4000 to 28000. The number of NK cells from different donors are shown in tables below. Table 4.1 shows the number of NK cells from the 12 independent donors.Table 4.2 shows the number of NK cells from

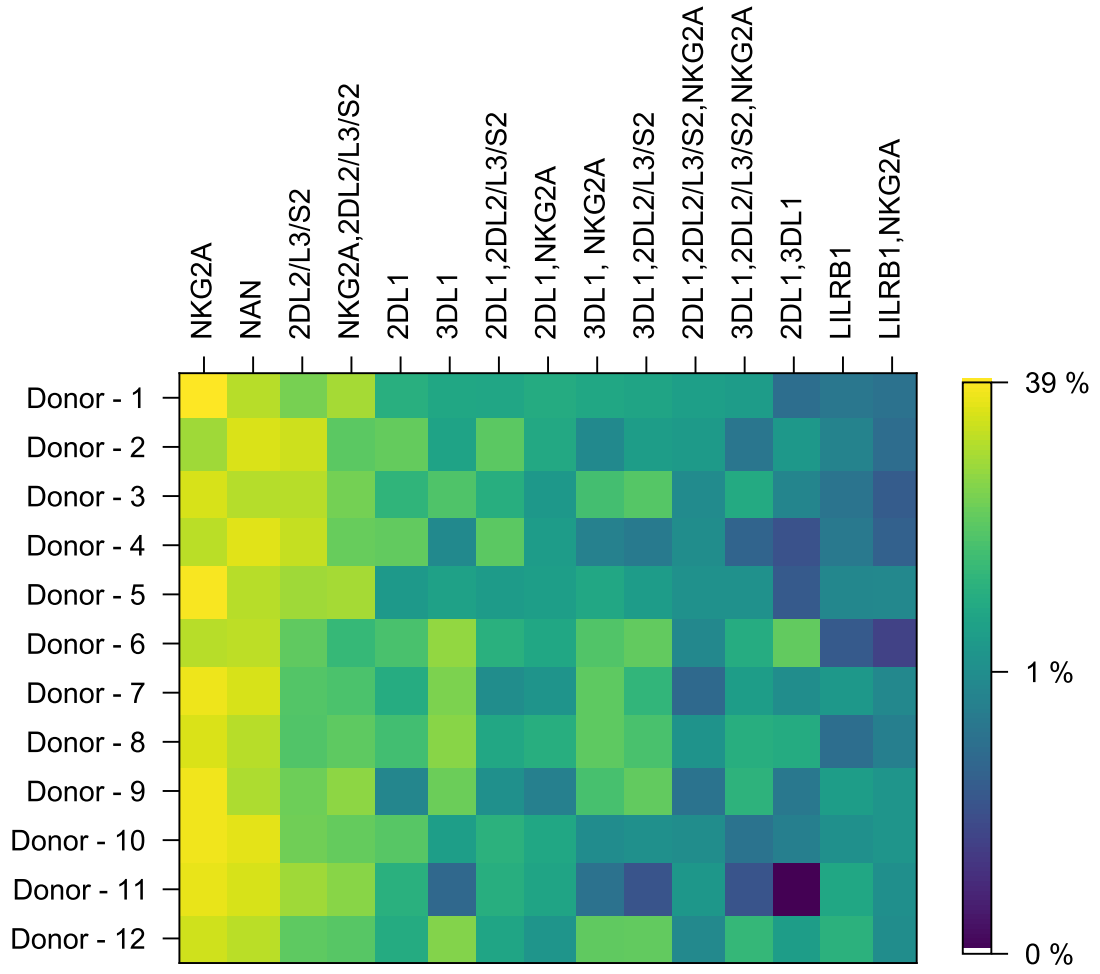

Figure 4.1: Our research findings displayed a phenotype distribution closely resembling that outlined in the referenced paper by Horowitz et al.[43]. This similarity suggests that the NK cells chosen for our study align closely with those used in the previous investigation. An intriguing observation from our study was that amongst the six inhibitory receptors considered, most NK cells expressed only one or two, rather than all of them.

the 5 sets of twins. In total we got the receptor distributions of 295270 NK cells. Next, to determine if one receptor is expressed or not, we need to set a threshold value of each receptor expressed to do the Boolean analysis. Here we used the same threshold used in the

paper [43]. Comparing with the result in the paper, our selection is good, we got similar phenotypes distributions to the result in the paper, **Fig. 4.1**. It proves that we successfully selected NK cells and the phenotype distribution of NK cells agrees with the result from the author.

Table 4.1: Number of NK cells from 12 independent donors

Donors	donor1	donor2	donor3	donor4	donor5	donor6
Number of NK cells	6829	6736	10082	14031	10722	9807
Donors	donor7	donor8	donor9	donor10	donor11	donor12
Number of NK cells	21045	19926	18658	22100	9779	28388

Table 4.2: Number of NK cells from 5 sets of twins donors

Donors	donor1	donor2	donor3	donor4	donor5
Number of NK cells	20124	15953	7084	12229	16729
Donors	donor6	donor7	donor8	donor9	donor10
Number of NK cells	8250	10194	14519	7784	4301

4.2.4 Distribution of number of receptor expressed by each NK cell

After successfully got the receptors expressed by NK cells. We know that NK cells do not express all receptors, then one interesting topoic is how many receptors expressed by one NK cell. Here we plot the histogram of distributions of the total number of receptors expressed on each NK cell for 12 independent donors (**Fig. 4.2**). We will see the distributions for different donors are similar to each other. The distribution is close to a Gaussian distribution or binomial distribution. If we use the data to fit binomial probability distribution, it fits very well. The optimal value for p in the binomial distribution is around 0.3 if we use $n = 28$ which is the total number of receptors, **Figure 4.3**. This is an interesting

result. Considering the meaning of binomial distribution, it seems that a simple model to explain this is each receptor is independently expressed and the probability of one receptor being expressed is around 0.3.

To verify if this is a good idea, we first need to check if all receptors are independent correlated. We calculated the connected correlation between the receptors. The connected correlation is defined as $Corr_{ij} = p_{ij} - p_i p_j$, where p_{ij} is the probability of receptor i and receptor j are expressed at the same time, p_i, p_j denote the probability of receptor i and receptor j are expressed, respectively. In the calculation, we used the fraction of total number of NK cells that express a receptor as the probability of that receptor being expressed. We went through all the pairs of two receptors, and found the connected correlation between receptors are relatively small. In Table 4.3, connected correlations of different pairs of receptors are listed in descending order. The highest one is CD94 and NKG2A, having a connected correlation equals 0.1. It agrees with the experiment result that CD94 is a type of transmembrane protein expressed on the surface of NK cells, and NKG2A is a most common subunit it paired to form a complex. NKG2A is an inhibitory receptor which binds to non-classical major histocompatibility complex (MHC) class I molecule HLA-E. Also CD94 binds to many other activating/inhibitory receptors.

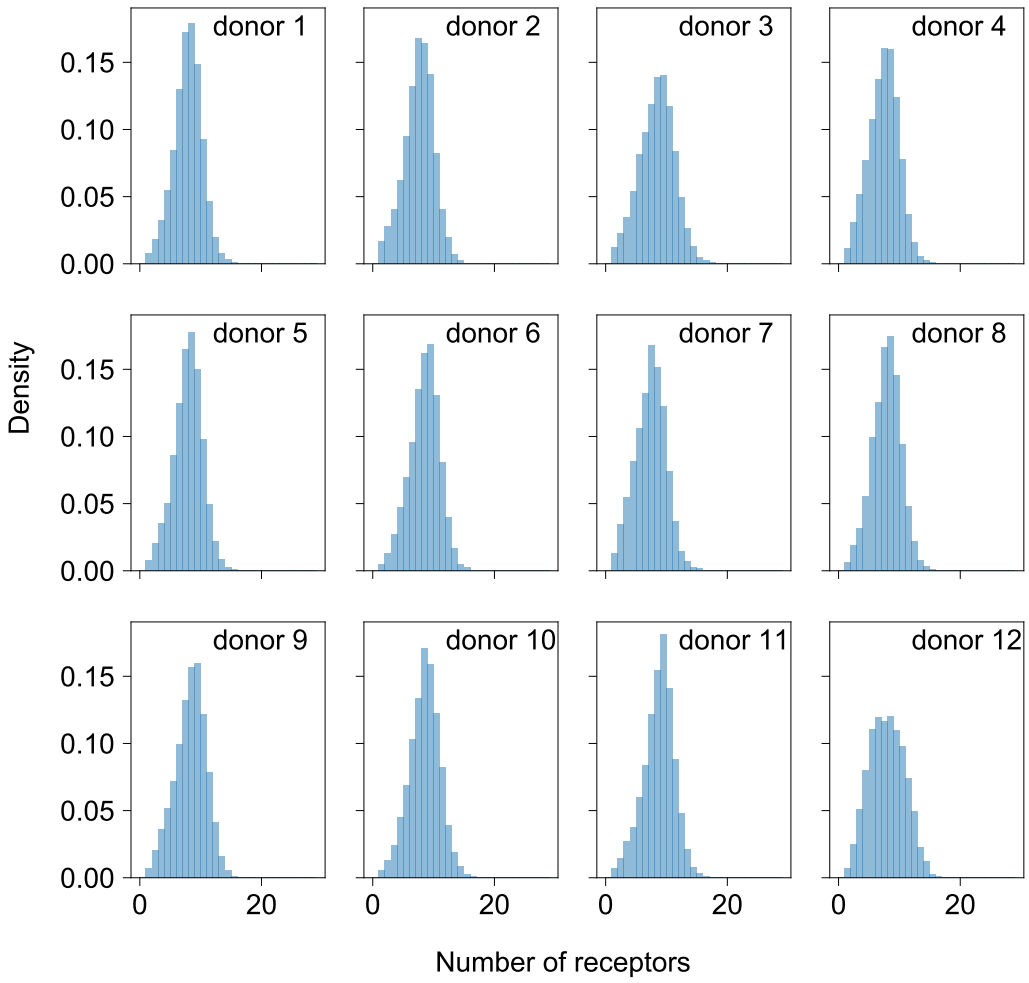

Figure 4.2: The distribution of the number of receptors expressed by a single NK cell was examined across all 12 donors. We found that this distribution was similar and highly concentrated across the donors. Most NK cells were observed to express a specific number of receptors, with only a small fraction of NK cells expressing an extreme number of receptors, either very few or many. Given the total count of 28 receptors, the mean and median values of expressed receptors were approximately around 8, demonstrating a typical expression level well below the maximum potential.

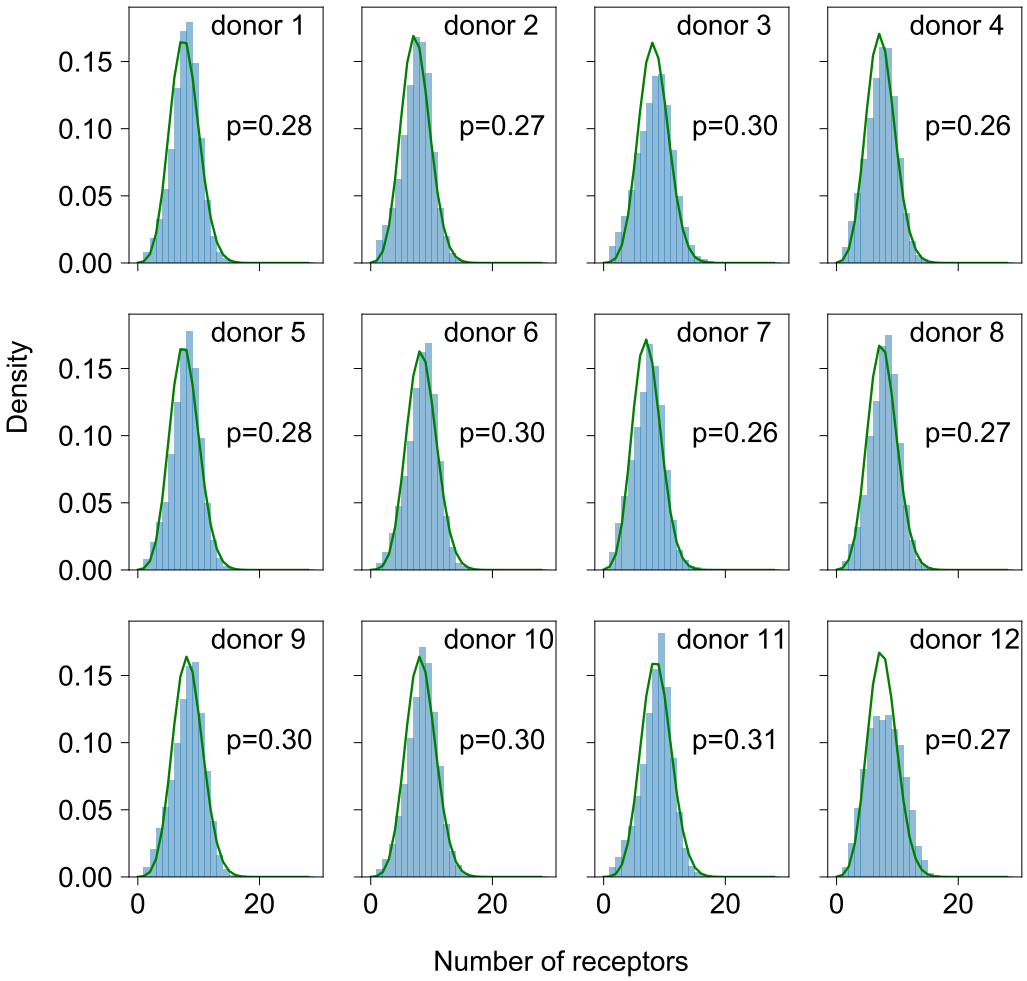

Figure 4.3: We fitted the distribution of the number of receptors expressed by a single NK cell to a binomial distribution, depicted by the green line in our results. Upon analysis, it was observed that the p values for different donors exhibited a high degree of similarity, with an approximate value of 0.3.

Table 4.3: Connected correlation between receptors

$receptor_i$	$receptor_j$	$p_{i,j}$	p_i	p_j	connected correlation
CD94	NKG2A	0.433	0.700	0.463	0.10871
NKp30	NKp46	0.338	0.581	0.483	0.05736
NKG2A	NKp46	0.261	0.463	0.483	0.03759
NKG2D	NKp46	0.279	0.503	0.483	0.03600
CD94	NKp46	0.372	0.700	0.483	0.03358
CD94	NKp30	0.439	0.700	0.581	0.03220
CD94	NKG2D	0.384	0.700	0.503	0.03194
NKG2A	NKp30	0.300	0.463	0.581	0.03143
NKG2D	NKp30	0.322	0.503	0.581	0.02970
NKG2A	NKG2D	0.255	0.463	0.503	0.02236
2B4	NKp46	0.306	0.589	0.483	0.02133
2B4	NKp30	0.363	0.589	0.581	0.02048
2B4	NKG2D	0.316	0.589	0.503	0.01999
KIR2DL2L3S2	NKp30	0.203	0.321	0.581	0.01641
CD94	NKG2C	0.058	0.700	0.060	0.01637
KIR3DL1	NKp44	0.021	0.152	0.032	0.01627
2B4	KIR2DL2L3S2	0.205	0.589	0.321	0.01537
KIR2DL2L3S2	NKp46	0.168	0.321	0.483	0.01325
KIR2DL2L3S2	NKG2C	0.032	0.321	0.060	0.01293

$receptor_i$	$receptor_j$	$p_{i,j}$	p_i	p_j	connected correlation
KIR2DL1	KIR2DL2L3S2	0.053	0.126	0.321	0.01237
KIR2DS4	KIR3DL1	0.032	0.139	0.152	0.01124
2B4	CD94	0.424	0.589	0.700	0.01116
NKG2C	NKG2D	0.040	0.060	0.503	0.00993
2B4	NKG2C	0.045	0.589	0.060	0.00976
KIR2DL1	NKp30	0.082	0.126	0.581	0.00828
KIR2DL2L3S2	NKG2D	0.169	0.321	0.503	0.00808
KIR2DS4	NKG2D	0.078	0.139	0.503	0.00798
2B4	KIR2DL1	0.082	0.589	0.126	0.00768
KIR3DL1	NKG2D	0.084	0.152	0.503	0.00756
KIR3DL1	NKp30	0.095	0.152	0.581	0.00640
KIR2DL1	KIR3DL1	0.026	0.126	0.152	0.00633
KIR2DL2L3S2	KIR3DL1	0.055	0.321	0.152	0.00626
KIR2DL1	NKp46	0.067	0.126	0.483	0.00610
CD94	KIR2DS4	0.103	0.700	0.139	0.00597
KIR2DS4	NKp30	0.087	0.139	0.581	0.00587
KIR2DL1	KIR2DS4	0.023	0.126	0.139	0.00505
2B4	KIR3DL1	0.094	0.589	0.152	0.00440
KIR2DL2L3S2	KIR2DS4	0.049	0.321	0.139	0.00414
KIR2DL1	NKG2D	0.067	0.126	0.503	0.00390

$receptor_i$	$receptor_j$	$p_{i,j}$	p_i	p_j	connected correlation
KIR2DL2L3S2	NKp44	0.013	0.321	0.032	0.00296
NKp30	NKp44	0.021	0.581	0.032	0.00278
KIR2DS4	NKp44	0.007	0.139	0.032	0.00262
KIR2DS4	NKG2A	0.067	0.139	0.463	0.00254
KIR2DL5	NKp46	0.010	0.017	0.483	0.00246
KIR2DL2L3S2	KIR2DL5	0.008	0.321	0.017	0.00238
KIR2DL2L3S2	LILRB1	0.014	0.321	0.036	0.00237
NKp44	NKp46	0.018	0.032	0.483	0.00226
NKG2D	NKp44	0.018	0.503	0.032	0.00218
KIR2DL1	KIR2DL5	0.004	0.126	0.017	0.00215
2B4	LILRB1	0.023	0.589	0.036	0.00188
KIR2DL1	NKG2C	0.009	0.126	0.060	0.00133
2B4	NKp44	0.020	0.589	0.032	0.00121
KIR2DL1	NKp44	0.005	0.126	0.032	0.00117
KIR2DL4	NKG2A	0.005	0.008	0.463	0.00117
KIR2DL4	LILRB1	0.001	0.008	0.036	0.00112
KIR2DL5	NKp30	0.011	0.017	0.581	0.00094
LILRB1	NKG2C	0.003	0.036	0.060	0.00087
NKG2C	NKp30	0.035	0.060	0.581	0.00085
KIR2DL4	NKp46	0.005	0.008	0.483	0.00068

$receptor_i$	$receptor_j$	$p_{i,j}$	p_i	p_j	connected correlation
KIR2DL4	NKG2C	0.001	0.008	0.060	0.00063
KIR2DL4	NKG2D	0.005	0.008	0.503	0.00060
2B4	KIR2DL4	0.005	0.589	0.008	0.00052
CD94	KIR2DL4	0.006	0.700	0.008	0.00042
2B4	KIR2DL5	0.010	0.589	0.017	0.00027
KIR2DL4	NKp30	0.005	0.008	0.581	0.00012
KIR2DL2L3S2	KIR2DL4	0.003	0.321	0.008	0.00009
KIR2DL5	LILRB1	0.001	0.017	0.036	0.00009
KIR2DL4	NKp44	0.000	0.008	0.032	0.00006
KIR2DL4	KIR2DS4	0.001	0.008	0.139	0.00003
CD94	NKp44	0.022	0.700	0.032	0.00001
KIR2DL4	KIR2DL5	0.000	0.008	0.017	0.00000
KIR2DL5	NKG2D	0.008	0.017	0.503	-0.00001
KIR2DL1	KIR2DL4	0.001	0.126	0.008	-0.00005
KIR2DL5	NKp44	0.000	0.017	0.032	-0.00011
KIR2DL4	KIR3DL1	0.001	0.008	0.152	-0.00017
LILRB1	NKp44	0.001	0.036	0.032	-0.00017
NKG2C	NKp44	0.002	0.060	0.032	-0.00018
KIR2DL5	NKG2C	0.001	0.017	0.060	-0.00044
KIR2DL1	LILRB1	0.004	0.126	0.036	-0.00049

$receptor_i$	$receptor_j$	$p_{i,j}$	p_i	p_j	connected correlation
2B4	NKG2A	0.272	0.589	0.463	-0.00058
KIR2DS4	NKG2C	0.007	0.139	0.060	-0.00092
KIR2DS4	LILRB1	0.004	0.139	0.036	-0.00097
CD94	KIR2DL2L3S2	0.224	0.700	0.321	-0.00125
KIR3DL1	LILRB1	0.004	0.152	0.036	-0.00128
KIR2DL5	KIR3DL1	0.001	0.017	0.152	-0.00129
2B4	KIR2DS4	0.081	0.589	0.139	-0.00132
NKG2A	NKp44	0.013	0.463	0.032	-0.00146
KIR2DL5	KIR2DS4	0.001	0.017	0.139	-0.00157
CD94	KIR2DL5	0.010	0.700	0.017	-0.00158
LILRB1	NKG2A	0.014	0.036	0.463	-0.00268
LILRB1	NKp46	0.015	0.036	0.483	-0.00306
KIR3DL1	NKp46	0.070	0.152	0.483	-0.00309
KIR2DL5	NKG2A	0.005	0.017	0.463	-0.00315
LILRB1	NKG2D	0.015	0.036	0.503	-0.00320
KIR2DS4	NKp46	0.064	0.139	0.483	-0.00337
NKG2C	NKp46	0.025	0.060	0.483	-0.00338
KIR3DL1	NKG2C	0.006	0.152	0.060	-0.00347
CD94	LILRB1	0.022	0.700	0.036	-0.00375
LILRB1	NKp30	0.017	0.036	0.581	-0.00411

$receptor_i$	$receptor_j$	$p_{i,j}$	p_i	p_j	connected correlation
CD94	KIR2DL1	0.084	0.700	0.126	-0.00437
CD94	KIR3DL1	0.097	0.700	0.152	-0.00947
NKG2A	NKG2C	0.014	0.463	0.060	-0.01355
KIR3DL1	NKG2A	0.055	0.152	0.463	-0.01578
KIR2DL1	NKG2A	0.036	0.126	0.463	-0.02185
KIR2DL2L3S2	NKG2A	0.121	0.321	0.463	-0.02767

Since our study mainly focused on the regulation of activating of NK cells. Here we may only consider those activating and inhibitory receptors. The connected correlation among those receptors are small, no greater than 0.05, which receptors are independent expressed might be an acceptable assumption. The probability of each receptor being expressed is smaller than 0.5. It means NK cells prefer not expressing all receptors. And not all phenotypes have the same probability being expressed since all phenotypes have the same probability being expressed is corresponding to $p = 0.5$. Instead, NK cells prefer a small number of receptors, we named it as sparse receptor expression.

4.3 Project high-dimensional signals to low-dimensional

Base on the experimental data, to be simple, we assume each type of receptor has the same probability that expressed on an NK cell surface. We assume the total signal from all receptors depends on the number of types of receptors binding to ligands. Here we ignore the effect of number of receptors for the same type since the effect of number of the same type of receptors can be included in the strength of signal received by the type of receptors.

In 1-D case, we used a Gaussian distribution as the the signal distribution, the distance from the mean to zero determined how activating/inhibitory those signals are for this type of target cells. We focused on the total signal instead of signals from each receptor. Now, in high dimensional case, we need to determine how is the signal from one receptor like. First, for activating and inhibitory receptors, the signal should have different signs. Same as before, we say positive signals are activating and negative signals

are inhibitory. Next, we need think about the ligand distribution. To test the killing ability of NK cells, we may assume target cells express all ligands, and the performance of NK cells are only determined by the receptor expression on them. Also, we can generate the ligand distribution for different types of targets. For healthy cells, we may generate more inhibitory ligands and for unhealthy cells we may generate more activating signals.

The total signal is the combine of signals from all receptors. Using the formula,

$$x_{total} = \sum_{i=1}^{n_r} a_i e_i x_i$$

where x_{total} is the total signal, i is i_{th} type of receptors, n_r is the total number of types of receptors, a_i denote either it's an activating receptor or inhibitory receptor, a_i can be $[1, -1]$, e_i denotes the expression of this type of receptor, e_i can be $[0, 1]$, and x_i is the signal from this type of receptor, it is proportional to the density of ligands.

4.4 Comparing the protection of different distributions of receptors

Since we are able to project the signals from all receptors to one dimensional combined signal, then we can use the one dimensional modified Bayesian inference model to process the combined signal and learn the environment. Here we compare the performance of different receptor distributions to explore the protection they could provided. First, we start comparing the sparse receptor distribution with other two distributions, one is expression all receptors which we name it as identical distribution, the other is uniform distribution where each phenotype has equal probability being expressed.

4.4.1 Receptor expression models

In the context of a sparse receptor distribution, we operate under the assumption that each receptor has an equal probability, denoted as p, of being expressed. This means that, on the surface of an NK cell, receptor 1 has a p probability of expression, as does receptor 2, with their expressions being independent of each other. Given the sparse nature of this distribution, p is a relatively small number, falling within the range of $[0, 0.5]$. Typical values for p might be $0.1, 0.22, 0.3$. The number of receptors expressed on an NK cell follows a binomial distribution, closely aligning with experimental results.

For identical distribution, all receptors are being expressed, which means every NK cell express all receptors.

For uniform distribution, all phenotypes have the same probability being expressed, which equivalent to each receptor has a probability, $p = 0.5$, being expressed independently.

In our simulations, we construct a 2D array to act as the expression matrix for generating the receptor expression across a number of NK cells. The number of rows in this array corresponds to the number of NK cells, while the number of columns represents the total possible receptors (or dimensions). Each NK cell (represented by each row) is required to express at least one receptor. The elements within this matrix hold a value of either 1 or 0, where 1 signifies that a specific receptor is expressed by the cell, and 0 indicates the opposite. In this context, we don't distinguish between activating and inhibitory receptors. However, if needed, we could represent an inhibitory receptor with a -1 value to reflect its opposite effect on the signal value.

A distinguishing feature among these distributions is the number of receptors expressed. The uniform distribution tends to express more receptors on each NK cell, with the identical model expressing the highest number of receptors. Figure 4.4 illustrates the differing distribution of the number of receptors expressed by each NK cell across various dimensions. If we denote the number of dimensions we are considering as d, this would imply that we assume there are d receptors in total. To investigate the performance changes of these receptor distributions, we start from a low dimensionality of $d = 2$ and gradually increase it up to $d = 32$.

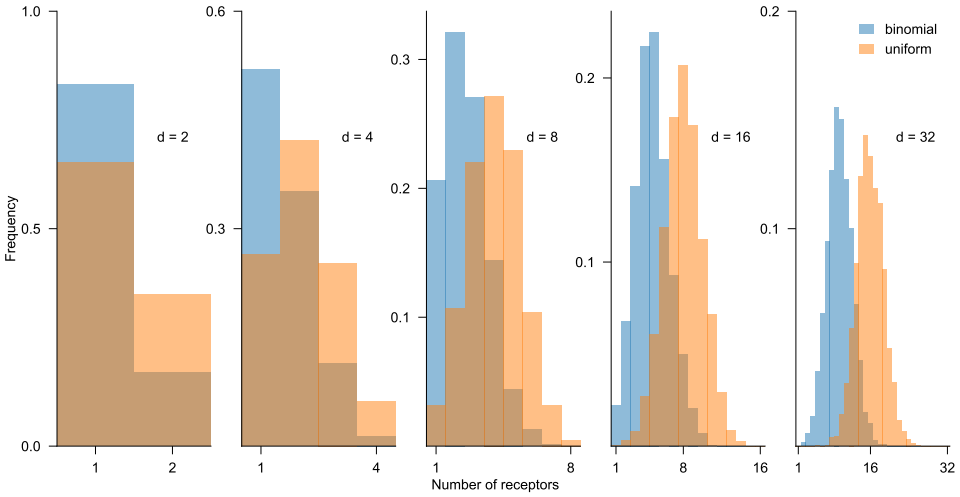

Figure 4.4: Receptor expression of binomial distribution and uniform distribution. The number d is the total number of receptors. For binomial distribution, most NK cells express less receptors than uniform distribution.

In the following sections, we will explore how these differing receptor distributions can impact the learning and killing abilities of NK cells.

4.4.2 Test on healthy signals

To evaluate the performance of NK cells possessing multiple receptors, we first need to establish the training and testing environments. As in the one-dimensional case, we set both the healthy (training) environmental signal distribution and the target (testing) environmental signal distribution as Gaussian distributions, but with differing means and variances. For instance, we could set the healthy signal for each receptor-ligand binding to satisfy $N(0,1)$, while the unhealthy signal for a receptor-ligand binding satisfies $N(4,1)$. This implies that unhealthy signals have a higher mean (i.e., are more activating) but maintain the same variance. After training the NK cells on the healthy environment, we can test the probability distribution of their activation.

First, we test the activation probability against healthy signals. It's crucial to ensure that NK cells do not eliminate a substantial number of healthy cells. For this test, we maintain the same settings as before. As for the memory parameters, we set $\kappa = 2\alpha$ and vary the memory length. We set the activation threshold at $\theta = 0.01$. For a trained NK cell, an activation threshold signal is established. Using this threshold and the target signal distribution, we can calculate the probability of activation (POA) for this NK cell relative to a particular target signal distribution. For instance, consider an NK cell with two receptors binding to a healthy cell. Each receptor binds to its corresponding ligand, generating a signal that follows the distribution $N(0,1)$. As a result, the combined signal follows the distribution $N(0,2)$. The probability of activation for this NK cell, given the target, is computed by integrating this combined signal distribution from the threshold to

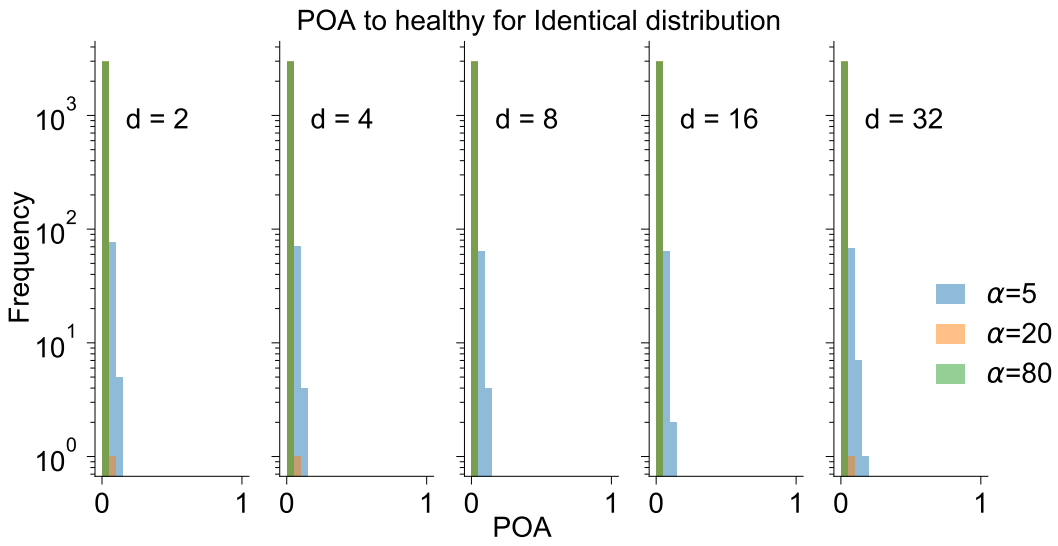

Figure 4.5: Distribution of the probability of activating when each NK cell express all receptors. The probability of activating is small and nearly zero. While for smaller α, the memory length is short, NK cells might kill some healthy cells since the fluctuation of the signal.

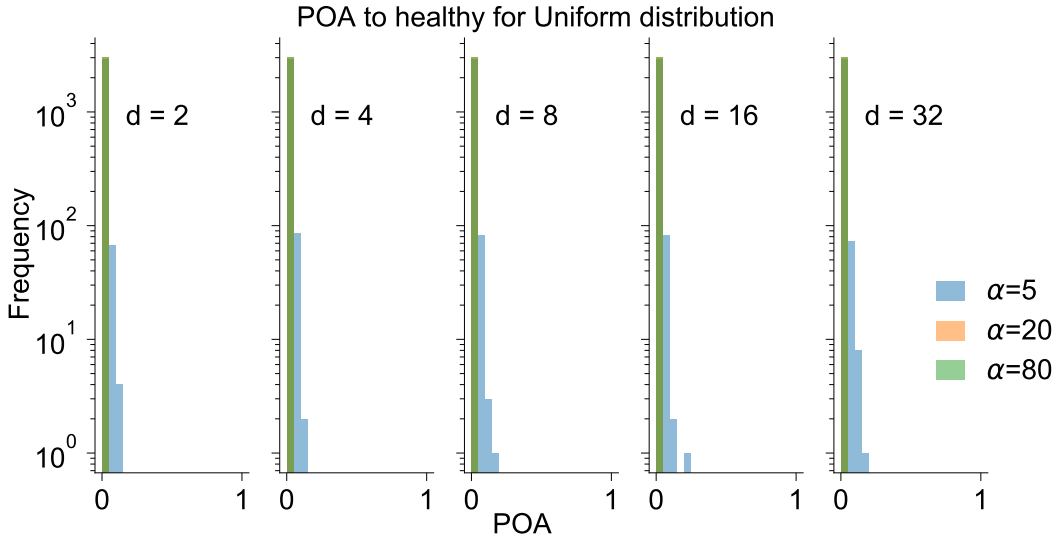

Figure 4.6: Distribution of the probability of activating when all phenotypes of receptor expression have equal probability being expressed. The probability of activating is small and nearly zero. While for smaller α, the memory length is short, NK cells might kill some healthy cells since the fluctuation of the signal.

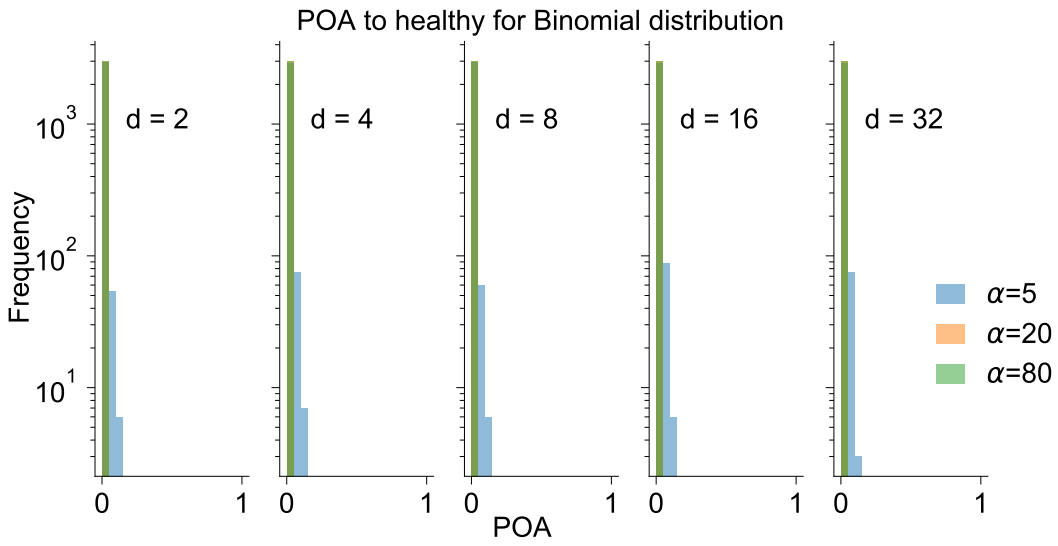

Figure 4.7: Distribution of the probability of activating when receptors on NK cells are sparsely expressed. The probability of activating is small and nearly zero. While for smaller α, the memory length is short, NK cells might kill some healthy cells since the fluctuation of the signal.

infinity. This integral represents the probability of this type of target being eliminated by this NK cell.

Ideally, after training within a healthy signal environment, the POA against healthy cells should approach zero as the NK cell adapts to the healthy signal environment and thus avoids eliminating healthy cells. The POA may not be precisely zero because we've employed a threshold, implying that a small proportion of encountered targets are always deemed untrustworthy. This is determined by the value of θ. Our results corroborate our expectation that for all three receptor distribution models, the POA against healthy cells is relatively low(**Fig. 4.5, 4.6, 4.7**).

4.4.3 Test on unhealthy signals

When dealing with an unhealthy signal distribution, we presume that in one "direction," the signal is more activating, signifying that it adheres to a Gaussian distribution with a higher mean. In our simulation, we stipulate that if an NK cell expresses a receptor in this direction, it will receive a signal from a $N(8, 1)$ distribution. This represents the scenario where an unhealthy cell tends to express a particular ligand that can bind with some activating receptor on an NK cell, leading to the NK cell receiving an activating signal. We conducted this performance testing by simulating the performance of a group of 3000 NK cells.

In the model where each NK cell expresses all receptors, known as the identical receptor expression model, we notice that as the number of dimensions decreases, the probability of activation also decreases, as illustrated in Fig. 4.8. This pattern is logical because as the number of dimensions decreases, so does the signal-to-noise ratio. For instance, an NK cell with two receptors might receive an activating signal from $N(8, 1)$ for one receptor, while the other receives a typical signal from $N(0, 1)$, resulting in a combined signal satisfying $N(8, 2)$. Conversely, if an NK cell has eight receptors, one might receive an activating signal from $N(8, 1)$, while the other seven get a typical signal from $N(0, 1)$, leading to a combined signal satisfying $N(8, 8)$. When the noise level (represented by the standard deviation) is comparable to the detected signal (the mean), distinguishing the signal from the noise becomes challenging.

Moreover, the performance varies depending on the memory length. NK cells with a shorter memory length, represented by a smaller α, demonstrate a broader distribution.

As the memory is brief, these NK cells are more likely to behave differently based on their unique encounter paths. Conversely, NK cells with a longer memory tend to behave similarly as they hold more information about the standard environment and thus have a similar comprehension of it.

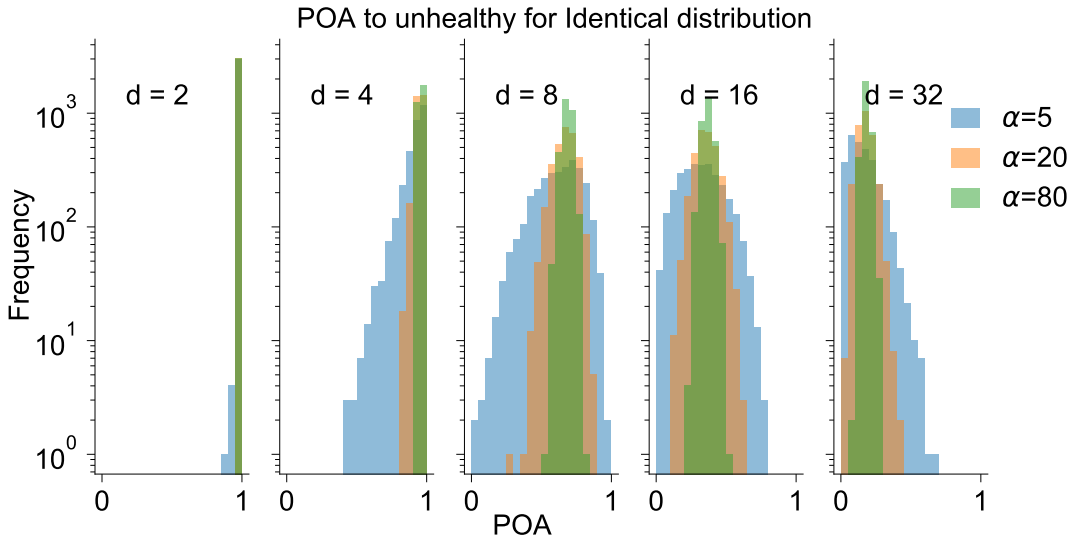

Figure 4.8: Distribution of the probability of activating against unhealthy signals when each NK cells express all receptors. As dimention d increases, the probability of activating decreases.

In the cases of uniform and binomial receptor distributions, not all receptors are expressed by the NK cells, thus maintaining a relatively high signal-to-noise ratio. This leads to a higher probability of activation against unhealthy signals, even when the dimensionality is high.

In our simulation, the binomial distribution model displays a smaller mean, implying that each NK cell expresses fewer receptors than in the uniform distribution model. Consequently, if they express the corresponding receptor, they will exhibit a higher signal-

to-noise ratio. Although many NK cells may not express the corresponding receptor, collectively, all the NK cells provide effective protection against unhealthy targets. This is due to the fact that for every type of unhealthy target, there will always be a subset of NK cells equipped to respond.

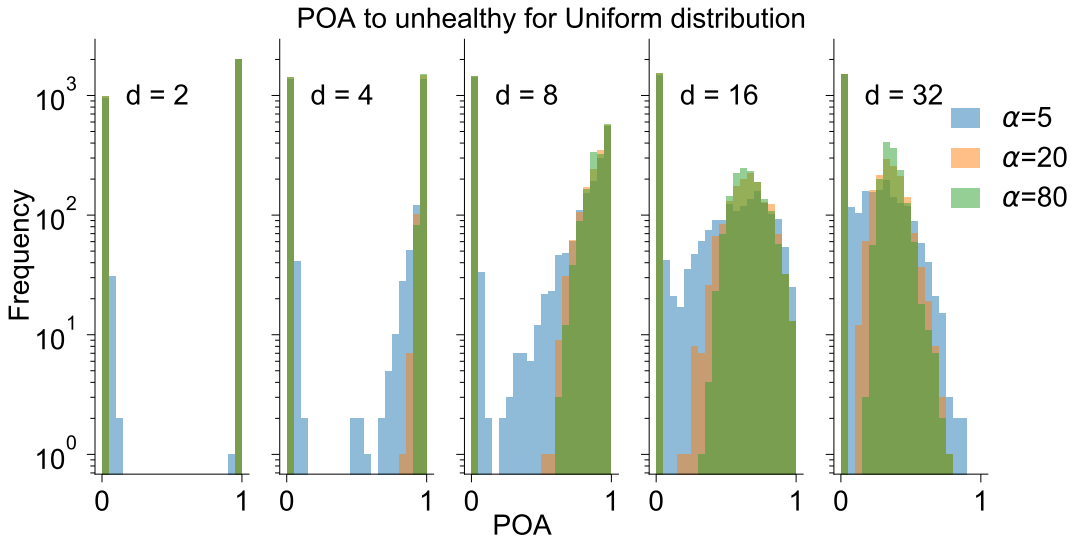

Figure 4.9: Distribution of the probability of activating against unhealthy signals when receptors on NK cells are uniformly distributed. Generally speaking, most NK cells still have a relatively high POA when the number of dimension is high comparing with the identical distribution model.

4.4.4 Sparse receptor expression's performance on different types of perturbation

Next, we will discuss the performance of our model under various perturbations. We initiated our simulations by allowing every immune cell to interact with 200 healthy targets. To generate the expression levels for each ligand of a healthy target, we drew them independently from a Gaussian distribution with a mean of 3 and a standard deviation of

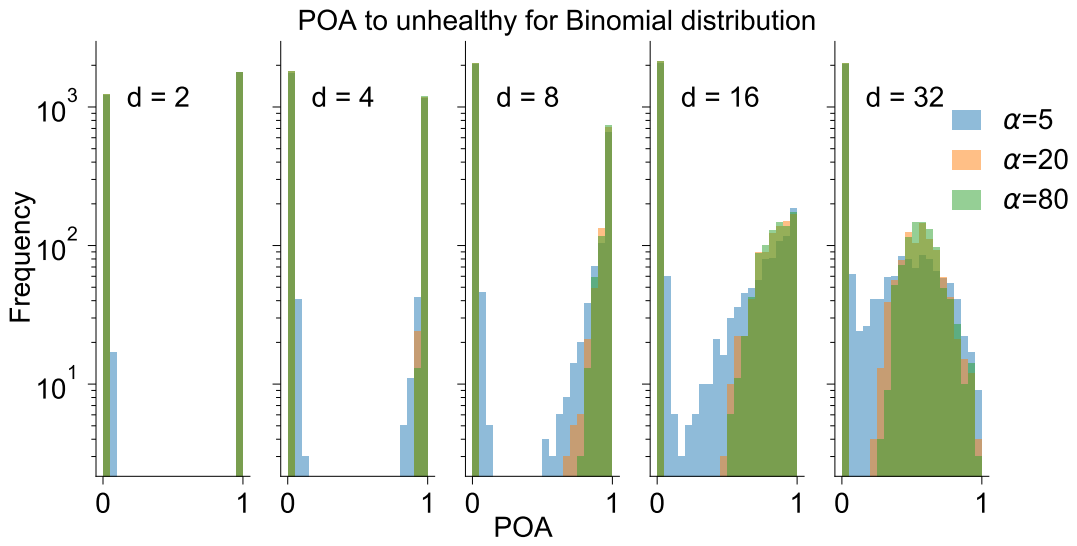

Figure 4.10: Distribution of the probability of activating against unhealthy signals when receptors on NK cells are sparsely expressed. At high dimension, there is still some fraction of NK cells have a high killing ability to the targets.

1. However, this approach sometimes resulted in a few rare cells having negative values for ligand expression. To address this issue, we repeated our simulations by setting the expression values to zero for any negative values, and we observed no variations in the outcomes presented below.

After training the immune cell population, we explored various models for aberrant target cells by introducing different perturbations to their surface ligand expression. Specifically, we considered two types of changes in ligand expression: sparse and dense. In sparse perturbations, changes were focused on a single ligand, leaving the expression of all other ligands normal. In contrast, dense perturbations involved modifying the expression of all ligands, with the magnitude of the change determined by generating a random n_r-dimensional vector with Gaussian entries and calculating the relative magnitude of each entry.

Moreover, we examined two types of perturbations: "pure" and "hidden". Pure perturbations resulted in changes that either increased the expression of activating ligands or decreased the expression of inhibitory ligands. On the other hand, hidden perturbations aimed to achieve a net change in activating - inhibitory ligand expression of zero. For sparse hidden perturbations, we randomly selected another ligand and adjusted its average expression to counterbalance the perturbation, while for dense hidden perturbations, we partitioned the group of all ligands randomly into two equally sized sets and favored activation for one set and inhibition for the other.

The magnitude for perturbation is set to be 5 after testing variety of values for magnitude. This is the best to show the difference between NK cells with sparse and dense receptor expression. If the perturbations are too small, aberrant cells may be very similar to healthy ones, making them difficult to distinguish. Conversely, if the perturbations are too large, aberrant cells may be so different from healthy cells that immune cells with the appropriate receptor(s) can easily recognize them. We generated 10^4 aberrant target cells. The performance of each group of NK cells is measured by the alignment distribution of NK cells. We considered the perturbations in ligand expression of aberrant cells and receptor expression patterns of immune cells as vectors, and calculated their normalized inner product. The threshold of recognizing a abnormal cells is set to $\theta = 0.01$ the same as we did in 1-D case. We plot the performance of population of NK cells with sparse and dense receptor expression on different types of perturbation in **Fig. 4.11**.

We will see that sparse receptor expression has a better performance to recognize diversity aberrant targets because they usually have a higher ratio of signal to noise. To

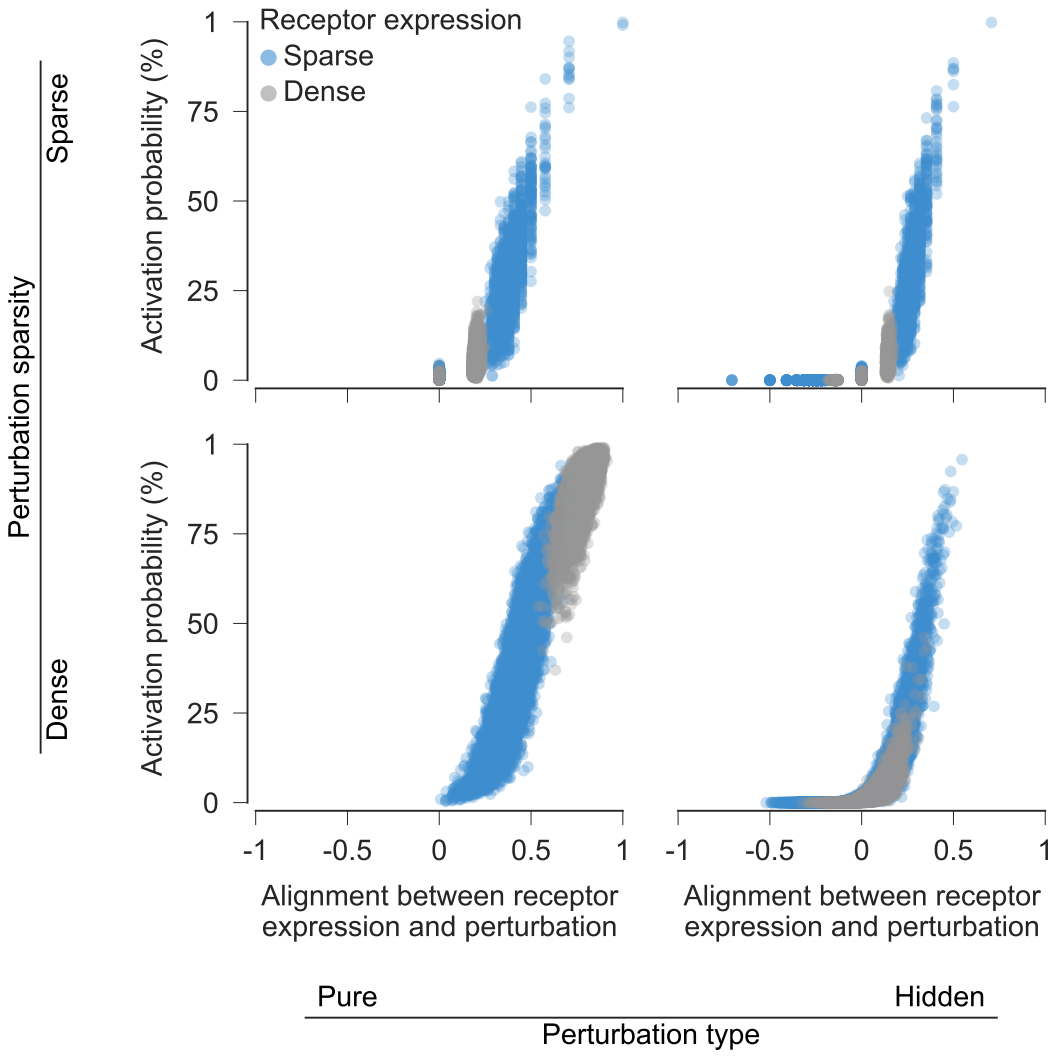

Figure 4.11: **Immune cells best recognize aberrant targets when they express the specific combination of receptors that recognize ligands with perturbed expression.** Immune cells with sparse receptor expression are more likely to align with changes in ligand expression for a variety of aberrant target cells, allowing them to kill these targets more efficiently. To quantify the alignment between immune cell receptors and perturbations, we treated the perturbations in ligand expression in aberrant cells and receptor expression patterns in immune cells as vectors and computed their normalized inner product.

tolerate to healthy cells, NK cells will express some inhibitory receptors. To be able to recognize aberrant targets, NK cells express some activating receptors. When receptor is sparsely expressed, one NK cell is able to respond to some aberrant targets expressing corresponding ligands precisely. If the receptors are dense expressed, it might received too many signals that are irrelevant for target detection.

Chapter 5

Performance of the model in anomaly detection

5.1 Introduction

As technology continues to advance, an increasing amount of time series data is being generated. This data spans various areas, from the number of website visits and real-time machine temperatures to bank card transaction records and personal heart rate performance curves. Such data plays a pivotal role in shaping the development of individuals, businesses, and even nations. Monitoring these time series data and identifying potential outliers is crucial for maintaining stability and minimizing potential losses. For instance, detecting outliers in bank consumption records can assist in identifying credit card fraud, allowing for immediate account freezing to curb losses for both consumers and banks. Similarly, monitoring machine temperatures in real-time and pinpointing abnormal

readings can ensure a safe production environment, thus preventing any production mishaps linked to extreme temperatures. In the medical field, timely detection of anomalies in heart rate performance curves can be lifesaving for patients with heart-related conditions, enabling them to receive prompt treatment. Therefore, the importance of outlier detection is growing and becoming more apparent in our increasingly data-driven world.

To detect outliers, firstly we have to define what is an outlier, which is also a key problem in building an outlier detection model. Generally speaking, outliers can be divided into two types, point outliers and subsequence/collective outliers [9]. While more specifically, the standard for outliers varies by data. For example, a high temperature detected in a running machine indicates dangers and must be reported as outliers immediately. While in the heart rate performance curve, a decrease in frequency must be an outlier and get noticed. On the other hand, the data changes over time, and the definition of outliers may also change over time. Outliers in one period may be normal in another period, in other words, outliers may become the new normal. Thus the model should be adaptable.

Now many techniques have been developed to detect outliers. Based on the input data type, they can be classified as univariate and multivariate types. Based on the outliers they focus on, they can be classified as point outlier detection methods and subsequence outlier detection methods. Based on the method they used, they can be classified as supervised and unsupervised methods. Here we introduce a new method, modified Bayesian inference. It's an unsupervised method and focuses on point outliers. It has been applied to biological models, and it helps to explain some experimental results [82]. Here we are going to prove that this simple model can be applied to general outlier detection tasks and

it gains good performance on the two popular benchmarks, Yahoo S5 [56] and Numenta Benchmarks [57].

The structure of this chapter is as follows. In section II, we introduce the algorithm of our model. In section III, two widely used Benchmark data sets, Yahoo website traffic data, and Numenta outlier Benchmark, are introduced. In section IV, we will describe the evolution method used in this paper. In section V, we evaluate the performance of our model on the benchmark data sets and compare our results with other algorithms. In section VI, a conclusion of advantages and disadvantages and how to further develop this model is discussed.

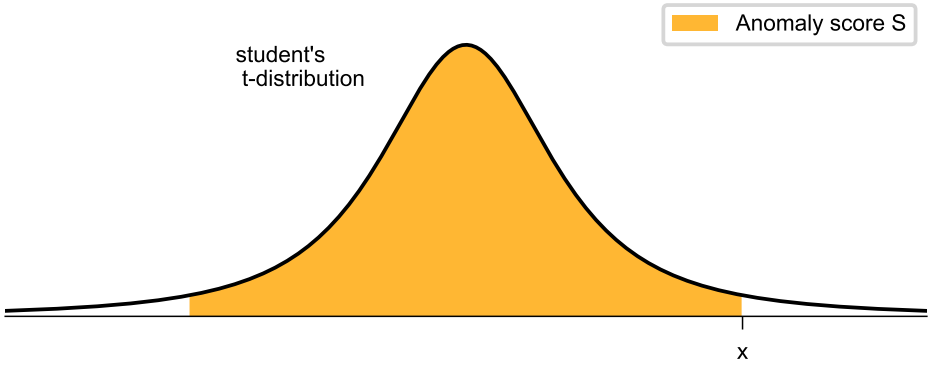

Figure 5.1: **Definition of anomaly score.** The black curve is the estimated student's t-distribution. The area of orange region is defined as anomaly score, S. Larger S is, x is more unlikely satisfying this estimated distribution, which means more likely to be an anomaly.

5.2 Model

In the paper [82], to understand the regulation of innate immune activation, we introduced a method that uses modified Bayesian inference to simulate the process that innate immune cells use to identify unhealthy cells. It estimated the environment signal distribution and made some changes to the parameters to make it adaptable. We refer to this model as Modified Bayesian Inference (MBI). The main idea is how a single immune cell knows which target cell is unhealthy and adapts to the new environment when interacting with the same type of unhealthy cells many times. For a single immune cell, it keeps interacting with surrounding cells, receiving signals from them, which can be treated as time series data. This time series data helps it learn the environmental signal distribution and identify unhealthy signals according to this estimated distribution. This is similar to what the outlier detection does, identifying outliers in time series data. Thus, we further developed this method and applied it to outlier detection. Following, we will introduce the modified Bayesian inference model.

5.2.1 Modified Bayesian inference

For the simplest, a one-dimensional time series that satisfies a static distribution, such as data that satisfies a certain Gaussian distribution, any data far from the mean is likely to be an outlier. Here we can use Bayesian inference to estimate the mean and variance of this static distribution, the estimated mean and variance can be used to assess how close the new signal is to the mean, or in other words, the probability of being an outlier. For example, for a sequence whose data points satisfy a fixed Gaussian distribution

$N(\mu_{True}, 1/\lambda_{True})$, where μ_{True} is the mean and λ_{True} is the precision, which is the inverse variance, $\lambda_{True} = 1/\sigma^2_{True}$, where σ^2_{True} is the variance. We would like to estimate the mean and precision. It is known from Bayesian theory that if we have a hypothesis of the probability distribution of mean and precision, $P(\mu, \lambda)$, and evidence x which is an observation related to the parameters mean and precision, then the updated probability distribution of mean and precision is

$$P(\mu, \lambda|x) = \frac{P(x|\mu, \lambda)P(\mu, \lambda)}{P(x)}\,, \tag{5.1}$$

where $P(\mu, \lambda)$ is named as the prior distribution which describes the probability distribution of mean and precision before signal x is detected, $P(x|\mu, \lambda)$ is the likelihood function which is a normal function, $P(x)$ is the total probability of x being observed and can be treated as a normalized factor which is $P(x) = \int P(x|\mu, \lambda)P(\mu, \lambda)\,\mathrm{d}\mu\,\mathrm{d}\lambda$, and $P(\mu, \lambda|x)$ is the posterior probability distribution which describes the probability distribution of mean and precision after detection of signal x. If we use the conjugate prior for the prior distribution, the posterior distribution will be in the same probability distribution family which makes life easier. The conjugate prior used for a normal distribution with unknown mean and precision is a normal-gamma function,

$$\begin{aligned} &P(\mu, \lambda|m, \kappa, \alpha, \beta) \\ &= \frac{\beta^\alpha\sqrt{\kappa}}{\Gamma(\alpha)\sqrt{2\pi}}\lambda^{\alpha-\frac{1}{2}}e^{-\beta\lambda-\kappa\lambda\frac{(\mu-m)^2}{2}}\,, \end{aligned}$$

where $\Gamma(*)$ is a gamma function. The normal-gamma function is uniquely determined by its four parameters, namely $(m, \kappa, \alpha, \beta)$. When the normal-gamma distribution is chosen as the prior distribution in Bayesian analysis, interestingly, the posterior distribution also assumes a normal-gamma distribution. This posterior distribution can be expressed as

$P(\mu, \lambda | m', \kappa', \alpha', \beta')$. Consequently, this implies that the process of updating the probability distribution, in essence, becomes a straightforward task of updating these four distinct parameters $(m, \kappa, \alpha, \beta)$.

The Bayesian inference method mentioned above has a good performance in estimating the mean and precision of a static distribution while it is not for dynamic signal distribution(see SI). However, we found that a small modification on the update rule will enable it to estimate dynamic mean and precision. The modified update rules are

$$\begin{aligned} m' &= \frac{\kappa m + x}{\kappa + 1} & \kappa' &= \kappa \\ \beta' &= \frac{\alpha - 1}{\alpha - \frac{1}{2}} \left[\beta + \frac{\kappa}{\kappa + 1} \frac{(x - m)^2}{2} \right] & \alpha' &= \alpha \,, \end{aligned} \tag{5.2}$$

where κ and α are fixed and they can be named as "memory length" since they are related with the number of data points took into account in determining the probability distribution of the mean and precision. We can show that using the modified update rule, the estimated mean and precision will always approach the true mean and precision within a period which is determined by the memory parameters κ and α(see SI). Now we have a good estimation of the signal distribution, the next step would be determining how likely a data point being an outlier which is named as anomaly score.

5.2.2 Anomaly score

With a good estimation of the distribution of the parameters, we will be able to calculate the estimation of the data distribution. Here we calculate the estimated data distribution, $P_{est}(x)$, by integrating over all possible μ and λ,

$$P_{est}(x) = \iint P(x|\mu, \lambda) \, P(\mu, \lambda) \, \mathrm{d}\mu \, \mathrm{d}\lambda \,,$$

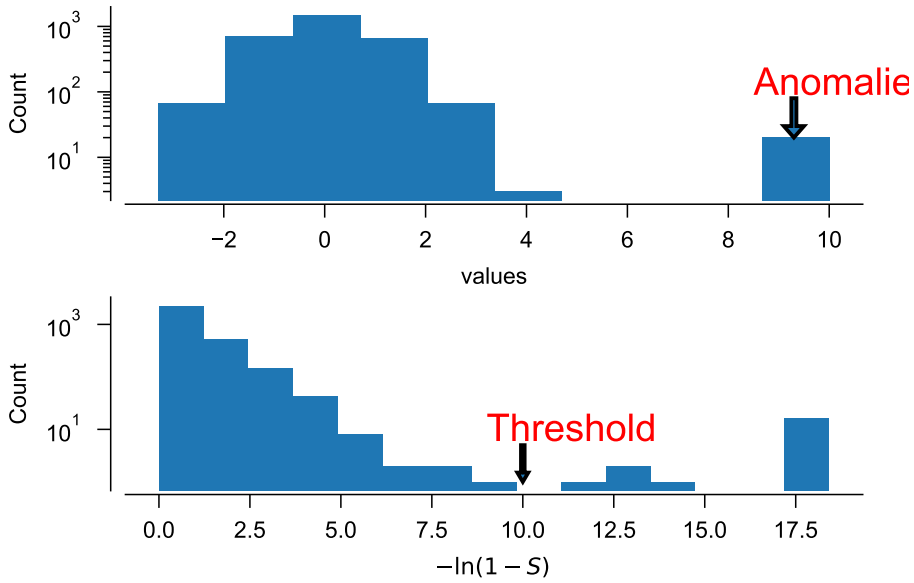

Figure 5.2: **Distribution of $S2$.** The top figure shows the data distribution. Here we randomly generated 3000 data points satisfying normal distribution $N(0,1)$ and 20 of them are modified as a fixed value 10. Those 20 anomalies are far from normal data distribution. The bottom figure shows the distribution of $-ln(1-S)$. That is how the threshold being determined.

where the $P(\mu,\lambda)$ is normal-gamma distribution determined by (m,κ,α,β) and $P(x|\mu,\lambda)$ is normal distribution. The integral gives a student's t-distribution,

$$P_{est}(t|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2},$$

where $t = \sqrt{\frac{\alpha\kappa}{\beta(\kappa+1)}}(x-m)$ and $\nu = 2\alpha$.

With the estimated data distribution, we define the anomaly score as how far away a data point is from the majority of the distribution. For example, the cumulative distribution function describes the percentage of data points from a distribution smaller than a given value. If the percentage is super small or large, it means the given value is far from the majority, in other words, it is less likely a data point selected from this distribution

which means it might be an outlier. Thus here we can define the anomaly score, S, as:

$$S = 2 * |t.cdf\,(x) - 0.5|\,,$$

where $t.cdf(x)$ means cumulative distribution function of data x, which is calculated by

$$\int_{-\infty}^{x} P_{est}\,(t(z)|\nu)\,\mathrm{d}z\,.$$

The range of the anomaly score is $[0, 1]$. The higher the anomaly score is, the more likely it is for a data point to be an outlier(see **Fig. 5.1**). Then it would be possible to choose a threshold of the anomaly score, larger than which the data points would be classified as outliers.

5.2.3 Threshold determination

The simplest way to set a threshold is choosing a value which is large enough. For example, we can set threshold as 0.995, which means if the anomaly score of a data point is larger than 0.995, we can say it is far from the majority and it can be classified as an outlier. This method has the lowest complexity which makes it suit for real time detection. However, it introduces some false positive rate and there will always be some outliers no matter it is true or not. Also it does not take the data into account and the same fixed threshold for all types of data which lower the overall performance.

Another way is to consider the anomaly score distribution. If all the data points comes from the same normal distribution, then the probability distribution of S should be uniform. If there are some outliers whose values at the ends of normal distribution, there would be a peak in the distribution of S around $S \sim 1$, because those outliers tend

to have larger S values. Our goal is to find the position of the peak if it exists. For mathematical convenient, we define "rarity" as $R = 1 - S$. Then we plot the distribution of parameter $S2 = -\log R$. We can show that for a normal distribution, this is a monotonically decreasing function(see SI). Whenever there are some outliers with extreme values, there will be a peak in the tail of the probability distribution of $S2$ because they tend to have smaller R, thus larger $S2$ values.(see **Fig. 5.2**). Thus the problem becomes to find where is the local minimum at the tail of $S2$ distribution. In this way, each data type has its own threshold which improves the performance.

Next, we can test the performance of this model. First we will test on some simulated data and show the performance is as expected. Then we will apply this model on two popular benchmark, Yahoo S5 website traffic data and Numenta Outlier Benchmark(NAB). We will show that this simple, low complex model has acceptable performance when comparing other models on the same benchmark. The code is in python and available in github[1].

5.3 Results

5.3.1 Adaptation of Modified Bayesian Inference

One key feature of the Modified Bayesian Inference model is adaptation. The estimated data distribution varies according to the observations of new data. Thus it has the ability to adapt to a new normal and detect local outliers at the same time. When adapting to a new normal, the changing points should be detected as outliers however

[1]https://github.com/ywqin/MBI-Anomaly-Detection

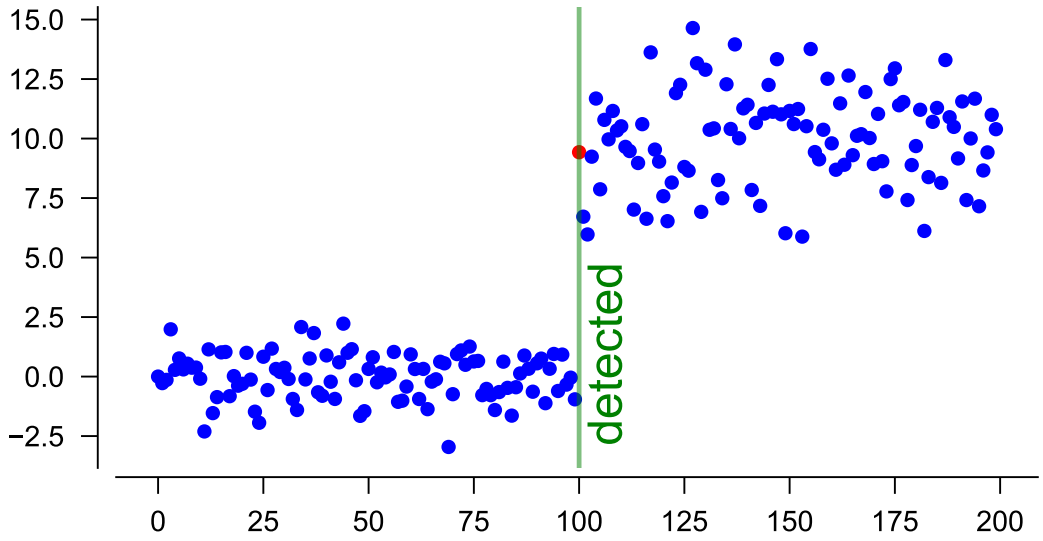

Figure 5.3: **Detection of changing point.** Here we generate 200 data points, the first 100 from normal distribution $N(0, 1)$ and the second 100 from $N(10, 2)$. We will see MBI model can detect the changing point. The red dot in the figure is the changing point and it is labeled as outlier by MBI.

the following data points should not be classified as outliers. For example, if the data distribution changes from a normal distribution $N(0, 1)$ to a new normal distribution $N(5, 4)$ at time point $t = 100$, the estimated mean and variance will adapt to the new values and it will not 'overreact'. (see **Fig. 5.3**). This has some important application in some fields which requires no overreaction to reduce the loss, just like the immune system detecting pathogens.

When the data changes over time, the Modified Bayesian Inference is able to detect outliers in local environments. For example, for periodic time series data, like the weather temperature in some places, there might be some missing data or some extreme data points due to some errors. While it's not a global extreme value, like a low temperature in summer while not as low as that in winter, the Modified Bayesian Inference still can find it out. (see

Fig. 5.4) Estimated the local data distribution make it "smarter" that those methods that only detect global outliers. Next, to standardize the performance, precision, recall and F_1 score are introduced.

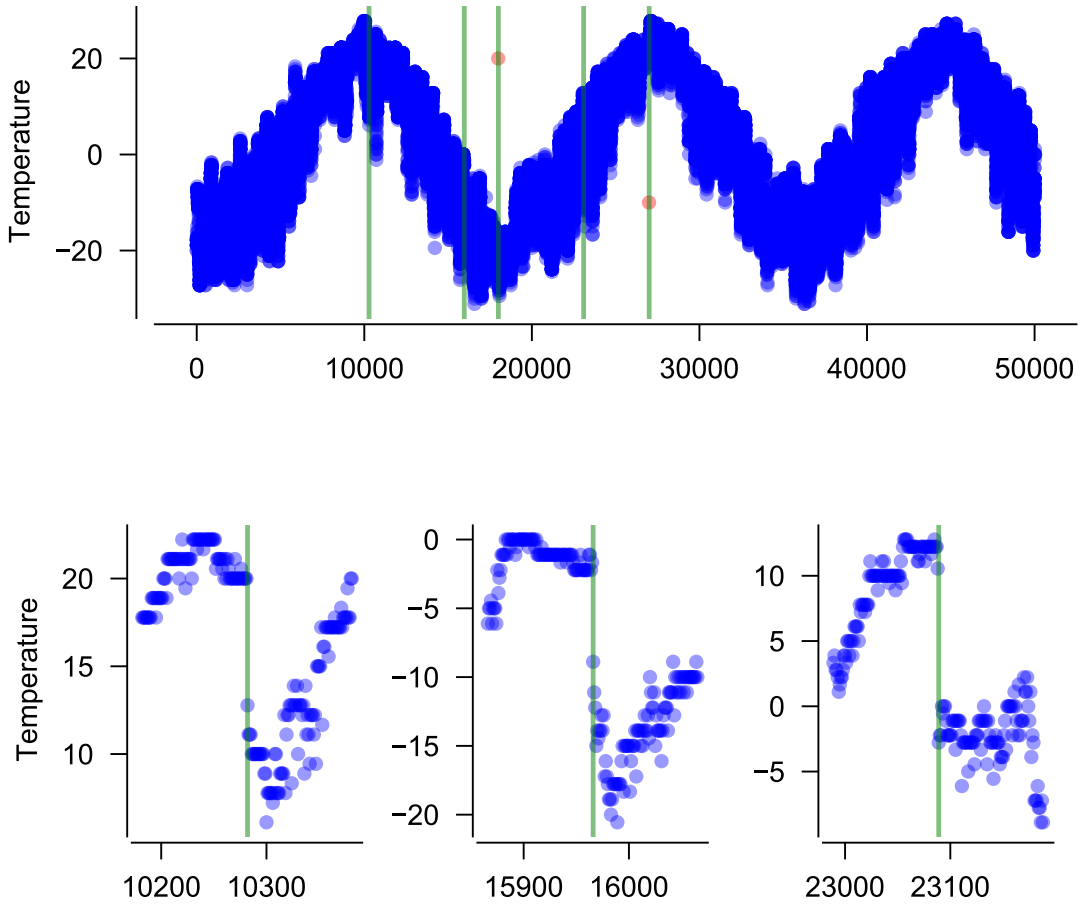

Figure 5.4: **Detect local anomalies.** Here we use the temperature data collected in Beijing. We inserted two anomalies which are red dots in the top figure. Using MBI, 5 anomalies are found. Two inserted anomalies are correctly labeled, and the other three labeled points correspond to temperature drop, which are shown in the bottom three sub-figures.

5.3.2 Definition of Precision, Recall and F_1 score

In order to assess and compare the performance of different algorithms, various metrics have been established. Among these, precision, recall, and the F1 score are frequently used. Precision is defined as Precision $= \frac{TP}{TP+FP}$, where 'TP' stands for true positives and 'FP' stands for false positives. In this context, 'positive' and 'negative' refer to whether a data point is predicted as an outlier or not, while 'true' and 'false' indicate if the predictions align with actual observations. Precision, ranging from 0 to 1, represents the proportion of accurate predictions among all positive predictions. On the other hand, recall is given by Recall $= \frac{TP}{TP+FN}$, where 'FN' denotes false negatives. Like precision, recall also ranges from 0 to 1. However, it measures the fraction of correct predictions among all actual positive points. These two metrics provide a comprehensive view of an algorithm's performance, considering both its precision and its ability to correctly identify positive instances. Then we can calculate the F_1 score,

$$F_1 = 2 \times \frac{Precision * Recall}{Precision + Recall},$$

which is a combination of precision and recall. The range of F_1 is $[0, 1]$. The value of F_1 is high only when both precision and recall are high. F_1 reaches maximum value 1 only if both precision and recall equal 1. F_1 reaches minimum value 0 when either precision or recall equals 0. F_1 score is widely used in measuring classification performance.

Since there are many different types of time series data and outliers, researchers have established some benchmarks to make it possible to compare the performance of different algorithms in detecting outliers.

5.3.3 Performance on Yahoo S5 Benchmark

Yahoo S5 Benchmark is provided by Yahoo [56]. It has 367 labeled time series containing many different types of outliers. The time series are classified into 4 groups, A1-A4. The A1 dataset includes some real time series that are based on the traffic to some Yahoo properties. The anomalies are labeled by humans. A2, A3, and A4 have some synthetic time series with randomly inserted anomalies. The number of outliers in each group is listed in Table 5.1. The types of outliers are single anomaly points, changing points, trend changes, etc. Here we mainly focus on single anomaly point detection. Thus we treat subsequence anomaly as one single anomaly, and if any of the data points within the subsequence are identified as an anomaly point, this anomaly is correctly identified.

Table 5.1: Yahoo S5 dataset info

Group	# time series	Type	# outliers
A1	67	real data	1669
A2	100	Synthetic data	466
A3	100	Synthetic data	943
A4	100	Synthetic data	1044

Applying MBI model on the Yahoo S5 dataset, the performance which described by precision, recall and F_1 scores, on the four groups is shown in Table 5.2. The parameters used here are $\alpha = 4, \kappa = 4$. Generally speaking, the precision is high and the recall is sometimes low, which is reasonable since MBI model is designed for identifying point anomalies only and is not good at identifying other types of outliers, like temporal outliers,

trend change, etc. The precision on real data, A1 group, is relatively small comparing with precisions on other groups.

However, this result is comparable to the result of other algorithms. For example, Twitter's Anomaly Detection(ADVec) and Windowed Gaussian [57]. ADVec is an algorithms provided by Twitter and applied on many benchmarks. It is a R package which is used for detecting outliers from a statistical standpoint or with seasonality. The performance of ADVec on Yahoo S5 dataset is shown in Table 5.3. The two parameters, alpha and period are set to 0.05 and 150 respectively. Windowed Gaussian is an algorithm provided by Numenta, which calculates the fraction from the data to the tail of the Gaussian probability distribution as anomaly score, where the mean and standard deviation of Gaussian distribution is determined by the data in a moving window. The performance of Windowed Gaussian on Yahoo S5 data is shown in Table 5.4. Compared with Twitter's ADVec algorithms and Windowed Gaussian, MBI model has higher recall and F1 score. While the precision on group A1 is smaller than that of ADVec, the recall is almost ten times larger. On A2 and A3, the precisions are similar, while MBI have higher recalls. On A4, both precision and recall of MBI are higher than those of ADVec and Windowed Gaussian. Generally speaking, MBI model has a relatively good performance on Yahoo S5 dataset. Next, we'll dig into the data and figure out what types of anomalies MBI can identify.

Table 5.2: Performance of MBI on Yahooo S5 dataset

	Precision	Recall	F1
A1	0.47	0.42	0.44
A2	0.94	0.94	0.94
A3	0.94	0.32	0.48
A4	0.81	0.19	0.31

Table 5.3: Performance of ADVec on Yahooo S5 dataset

	Precision	Recall	F1
A1	0.72	0.29	0.42
A2	0.97	0.17	0.28
A3	1.00	0.015	0.029
A4	0.56	0.038	0.070

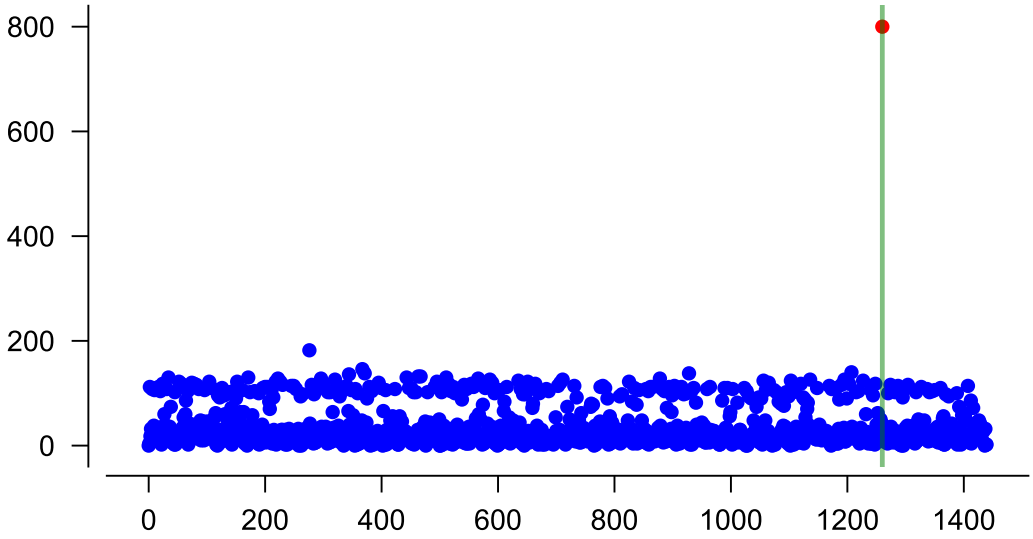

Figure 5.5: **Best performance of MBI.** There is only one data point labeled as outlier which is in red color. MBI finds this outlier correctly and no false positive prediction.

Table 5.4: Performance of Windowed Gaussian on Yahoo S5 dataset

	Precision	Recall	F1
A1	0.22	0.59	0.32
A2	1.00	0.62	0.76
A3	1.00	0.02	0.04
A4	0.16	0.05	0.08

In group A1, MBI model has good performance on file $'real\_45.csv'$. There is only one anomaly in this time series data and its value is much larger than others(**Fig. 5.5**). MBI model is able to find this type of point anomalies. Within the total 367 files, MBI

detect 112 files perfectly, of which all anomalies are identified and no false positive. MBI model does not have a good performance on file $'real\_17.csv'$, where there are 227 points are labeled as outliers which are in three subsequences and MBI model identified one out of them(**Fig. 5.6**). The one is identified because its value is much larger than precious ones. Since MBI model is able to adapt to new environment, points following that point are not identified. The other two subsequence are not identified because the values of their start points are comparable to previous normal points. MBI model is good at identifying sudden changes in value, not gradual changes. Also, in some files, all models have low performances because not all outliers are labeled [110](**Fig.5.7**). Some points around step 200 should be considered outliers because they are far away from normal points while only some points with extremely large values are labeled as outliers. This happens because there is no common definition of an outlier. Also the data provider mentioned that the labeling might not be consistent since they are labeled by humans. That is why we should focus on the performance of a model on each file, not just the overall performance on a benchmark. Next we will show that MBI has the advantage of low complexity and runs fast, which is an advantage for processing real-time data. For the time complexity, the average computation time for some models on each group [96] is shown in Table 5.5. The computation time is much shorter than The Numenta Platform for Intelligent Computing (NuPIC) which is a well known software for anomaly detection developed by numenta[2]. Also most of other algorithms take more time on A4 than A3 while ours is shorter. The difference between A3 and A4 is that A4 contains some fraction of change-points where the distribution changes

[2]https://www.numenta.com/

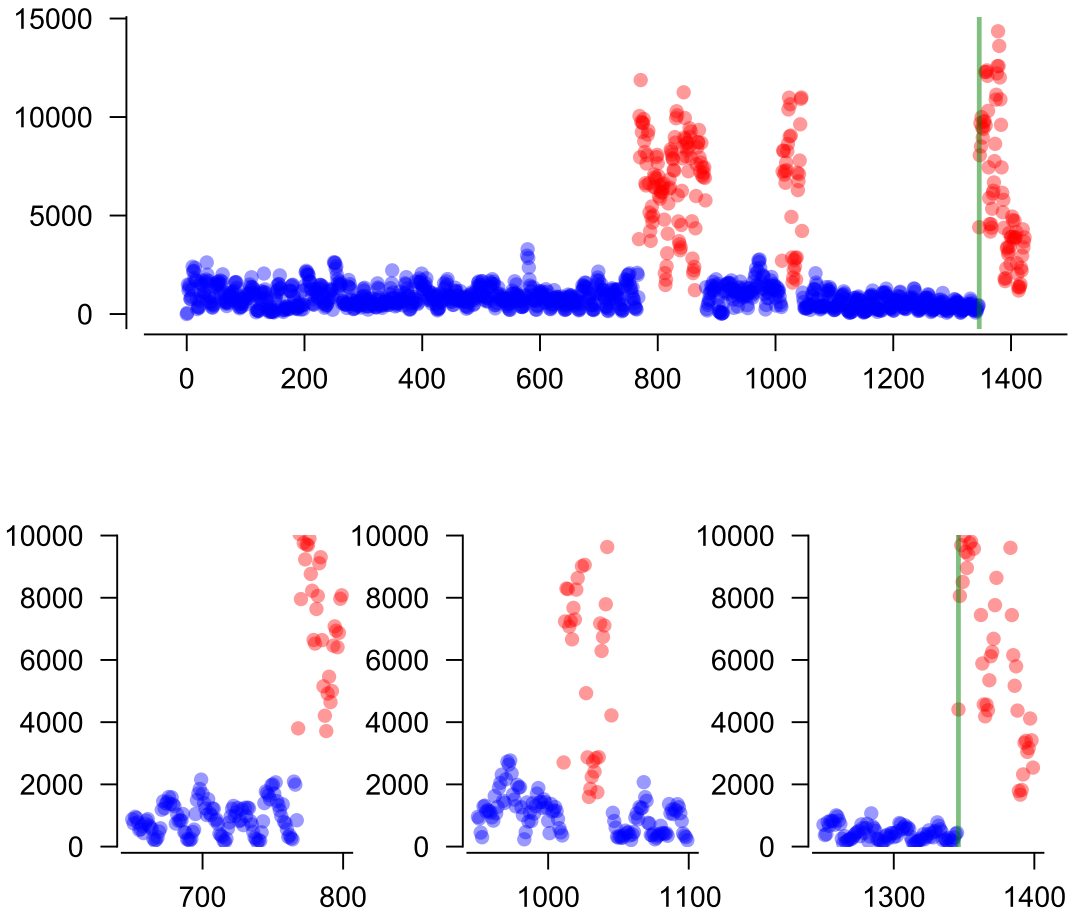

Figure 5.6: **Pattern anomalies.** There are three groups of data points are labeled as outliers which are in red color. MBI identified one of them (vertical green line) and no false positive prediction.

to a new normal. What's more, MBI also has a relatively higher F_1 score in A4 than other algorithms. It seems our model is good at dealing with change-points than others.

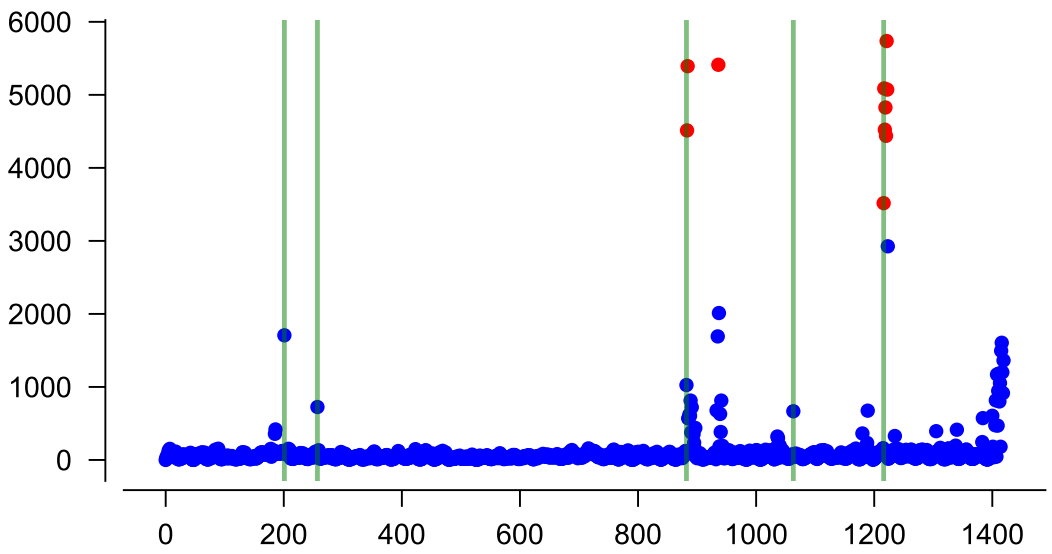

Figure 5.7: **Labeling might not be consistent.** MBI model identified all the labeled outliers. MBI model is also be able to identify the beginning of the noisy part, which should also be an outlier.

Table 5.5: Computation Time(s)

	MBI	NuPic(HTM)	ADVec
A1	11.6	368	3.3
A2	16.6	693	4.8
A3	20.0	813	5.6
A4	19.1	828	6.0

5.3.4 Performance on NAB benchmark

Another popular used benchmark is Numenta Anomaly Benchmark (NAB) which contains 57 labeled time series. It includes some real data like cpu utilization, the clicking rates of advertisement, room temperature, etc. It also has some artificial data with or without anomalies. Compared with Yahoo S5, NAB has more outlier types and longer time series, but fewer total time series. Besides, NAB provides a scoreboard to compare different algorithms using NAB score. By changing the weights of false positive and false negative,

the overall NAB score will be able to reward low false positive or low false negative which might be important for application in some fields[2].

Table 5.6: NAB score

Detector	Standard	Reward low FP	Reward low FN
HTM	64.7	56.5	69.3
ADVec	47.1	33.6	53.5
MBI	37.0	22.0	45.6
Skyline	35.7	27.1	44.5
Random	16.8	5.8	25.9

The performances of MBI and some other algorithms are shown in Table 5.6. The parameter settings of MBI model are $\alpha = 60, \kappa = 60$. Comparing with HTM(NumPic) and ADVec, performance of MBI on NAB is not that good. The main reason is many of the outliers in NAB are temporal outliers, not point outliers. MBI model is not good at dealing with subsequence outlier detection. For example, the performance of MBI on time series of 'realAWSCloudwatch/grok_asg_anomaly.csv' is good(**Fig. 5.8**), all three anomalies are identified (predictions are within the detection windows). The anomalies include one changing point and two extreme values. MBI is good at this. However, performance of MBI on 'realKnownCause/nyc_taxi.csv' data set is bad since all the outliers in the time series are either temporal outliers or values of anoomalies are comparable to normal points(**Fig. 5.9**). Even though NAB has only 57 datasets, it has many different kinds of outliers. Although MBI can only detect point outliers, the overall NAB score of MBI is similar to the commercially used algorithm Skyline. Some other reasons could be that the threshold is fixed for all time series data because of which it might miss some outliers with relatively small anomaly score, and some data are not close to normal distribution which the estimated

data distribution would be far away from the true data distribution and it might result in a large number of false positive.

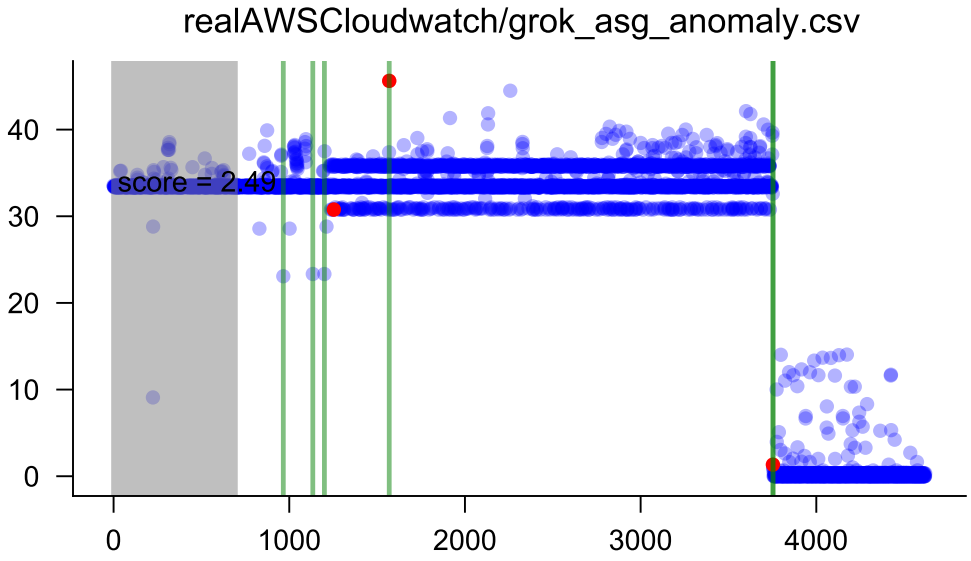

Figure 5.8: **NAB performance on AWS server metrics.** All three anomaly points (red dots) in the AWS server metrics collected by AmazonCloudwatch service, are identified.

Althought the score is similar to Skyline and smaller than HTM, the time cost of MBI is much smaller. It take less than one minutes for MBI to complete all the procedures including detection and scoring. While HTM takes around 20-30 minutes on a 2-4 cores laptop and Skyline takes much longer, several hours, Table 5.7. Generally speaking, MBI model runs fast and have a good performance for single point anomaly detection.

Table 5.7: Computation Time(s) on NAB

	Time(s)
MBI	within one minutes
HTM	20-30 minutes
Skyline	hours

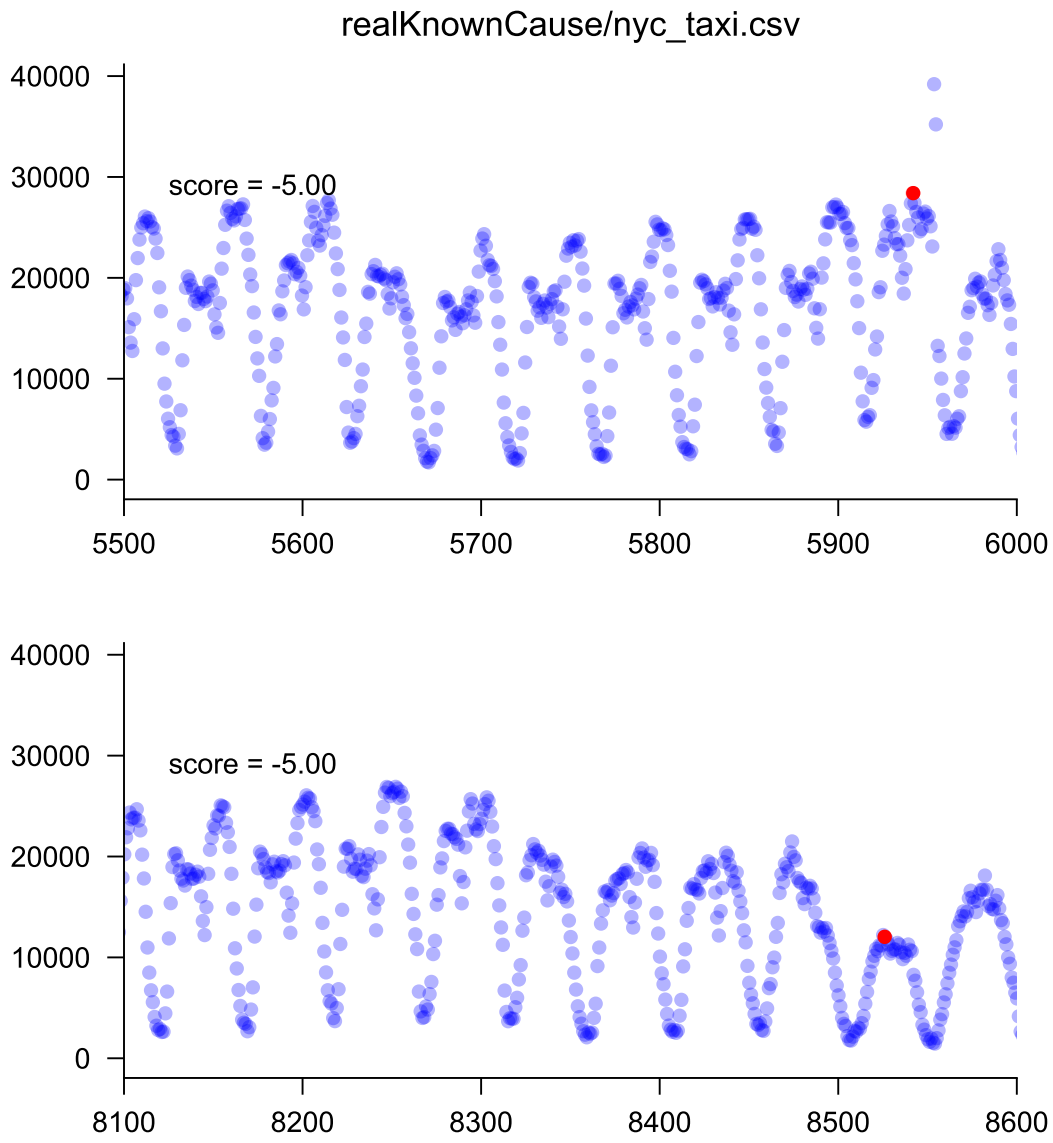

Figure 5.9: **NAB performance on New York taxi data set.** These are two anomaly points (red dots) in the number of NYC taxi passengers data set which are caused by NYC marathon and Christmas holidays. The score -5 means there are 5 anomalies in total and all of them are not identified. Here we just plot two of them.

5.4 Conclusion

We introduce a novel method for anomaly detection, the MBI model, which is based on Bayesian inference. By utilizing past data to construct an internal representation

of the normal environment, the MBI model can effectively distinguish between normal data and anomalies. With the application of weight decay for past data, the model can adapt to new environments within a few steps. This straightforward model demonstrates impressive performance for single anomaly detection. Although it is limited to detecting anomaly points, it exhibits comparable performance on benchmark tests such as Yahoo S5 and NAB benchmark. Most importantly, its fast execution allows for real-time anomaly detection tasks, making it suitable for our pathogen recognition immunity model. The parameters (α, κ) are representative of memory for past data; however, their optimal selection has not been extensively explored. Future research could focus on determining memory parameters based on the distribution of past data.

Chapter 6

Conclusions

In this dissertation, we introduced a new model describing the NK cells identifying unhealthy signals and adapting to a 'new normal'. Instead of specifically recognizing targets, we developed a new idea that innate immune cells can learn from normal signals and build an internal representation of the signal distribution of healthy cells. Thus, it is able to gain the ability to identify unhealthy cells and is able to adapt to new signal environments at the same time. This simple quantitative model can explain some behaviors of NK cells well.

After comparing with some experimental data, we proved that our model can generate similar results as the experimental data. Our model helps to explain the adaptation of NK cells in new environments and the behavior of "serial killing."

Also, we extended our model to high dimensional by taking into account the receptor expression of NK cells. We noticed that individual NK cells do not express all receptors. Instead, each NK cell expresses a small fraction of receptors. In this way, the receptor expression distribution of NK cells is diverse. Collectively, they provide better protection

than expressing a lot of receptors for each NK cell. This is because the signal to noise ratio of sparsely expressed receptors is higher than densely expressed receptors. This finding agrees with the experimental data, that each NK cell expresses a small number of activating/inhibitory receptors, rather than all of them, and instead, the number of phenotypes of NK cells is extremely large.

We then showed that this modified Bayesian inference model can also be used in anomaly detection in other fields. We applied this model to other types of time series data, like weather data, and we got a good performance.

In the future, this model can be further developed by considering the relationship between receptors. There are some experimental data that reveals the relation between receptors which considering the chemical mechanics behind them will give more precise predictions. Also, this model has the potential to simulate single NK cell behaviors, which could be compared with future experiments.

Bibliography

[1] *Computing Fourier series and power spectrum with Matlab*, 2000.

[2] Subutai Ahmad, Alexander Lavin, Scott Purdy, and Zuha Agha. Unsupervised real-time anomaly detection for streaming data. *Neurocomputing*, 262:134–147, 2017.

[3] Galit Alter, Maureen P Martin, Nickolas Teigen, William H Carr, Todd J Suscovich, Arne Schneidewind, Hendrik Streeck, Michael Waring, Angela Meier, Christian Brander, et al. Differential natural killer cell–mediated inhibition of hiv-1 replication based on distinct kir/hla subtypes. *Journal of Experimental Medicine*, 204(12):3027–3036, 2007.

[4] Nicolas Anfossi, Pascale André, Sophie Guia, Christine S Falk, Sophie Roetynck, C Andrew Stewart, Violette Breso, Coralie Frassati, Denis Reviron, Derek Middleton, et al. Human nk cell education by inhibitory receptors for mhc class i. *Immunity*, 25(2):331–342, 2006.

[5] Mikhail Belkin and Partha Niyogi. Using manifold stucture for partially labeled classification. In *Advances in neural information processing systems*, pages 929–936, 2002.

[6] Michael D Bern, Bijal A Parikh, Liping Yang, Diana L Beckman, Jennifer Poursine-Laurent, and Wayne M Yokoyama. Inducible down-regulation of mhc class i results in natural killer cell tolerance. *Journal of Experimental Medicine*, 216(1):99–116, 2019.

[7] Mark Bix, Nan-Shih Liao, Maarten Zijlstra, Janet Loring, Rudolf Jaenisch, and David Raulet. Rejection of class i mhc-deficient haemopoietic cells by irradiated mhc-matched mice. *Nature*, 349(6307):329, 1991.

[8] Jeanette E Boudreau and Katharine C Hsu. Natural killer cell education and the response to infection and cancer therapy: stay tuned. *Trends in immunology*, 39(3):222–239, 2018.

[9] Azzedine Boukerche, Lining Zheng, and Omar Alfandi. Outlier detection: Methods, models, and classification. *ACM Computing Surveys (CSUR)*, 53(3):1–37, 2020.

[10] Petter Brodin, Klas Kärre, and Petter Höglund. Nk cell education: not an on-off switch but a tunable rheostat. *Trends in immunology*, 30(4):143–149, 2009.

[11] Marine Champsaur and Lewis L Lanier. Effect of nkg2d ligand expression on host immune responses. *Immunological reviews*, 235(1):267–285, 2010.

[12] Laura Chiossone, Pierre-Yves Dumas, Margaux Vienne, and Eric Vivier. Natural killer cells and other innate lymphoid cells in cancer. *Nature Reviews Immunology*, page 1, 2018.

[13] Paul J Choi and Timothy J Mitchison. Imaging burst kinetics and spatial coordination during serial killing by single natural killer cells. *Proceedings of the National Academy of Sciences*, 110(16):6488–6493, 2013.

[14] S. Choi. Investigation on the pressure drop characteristics of cryocooler regenerators under oscillating flow and pulsating pressure conditions. *Cryogenics*, 44(3):203–210, March 2004.

[15] Athanasia E Christakou, Mathias Ohlin, Bruno Vanherberghen, Mohammad Ali Khorshidi, Nadir Kadri, Thomas Frisk, Martin Wiklund, and Björn Önfelt. Live cell imaging in a micro-array of acoustic traps facilitates quantification of natural killer cell heterogeneity. *Integrative Biology*, 5(4):712–719, 2013.

[16] Jerome D Coudert, Léonardo Scarpellino, Frederic Gros, Eric Vivier, and Werner Held. Sustained nkg2d engagement induces cross-tolerance of multiple distinct nk cell activation pathways. *Blood*, 111(7):3571–3578, 2008.

[17] Jayajit Das. Activation or tolerance of natural killer cells is modulated by ligand quality in a nonmonotonic manner. *Biophysical journal*, 99(7):2028–2037, 2010.

[18] Lucas Ferrari de Andrade, Rong En Tay, Deng Pan, Adrienne M Luoma, Yoshinaga Ito, Soumya Badrinath, Daphne Tsoucas, Bettina Franz, Kenneth F May, Christopher J Harvey, et al. Antibody-mediated inhibition of mica and micb shedding promotes nk cell–driven tumor immunity. *Science*, 359(6383):1537–1542, 2018.

[19] Henri de La Salle, Daniel Hanau, Dominique Fricker, Arlette Urlacher, Adrian Kelly, Jean Salamero, Stephen H Powis, Lionel Donato, Huguette Bausinger, Michel Laforet, et al. Homozygous human tap peptide transporter mutation in hla class i deficiency. *Science*, 265(5169):237–241, 1994.

[20] Weiwen Deng, Benjamin G Gowen, Li Zhang, Lin Wang, Stephanie Lau, Alexandre Iannello, Jianfeng Xu, Tihana L Rovis, Na Xiong, and David H Raulet. A shed nkg2d ligand that promotes natural killer cell activation and tumor rejection. *Science*, 348(6230):136–139, 2015.

[21] Vincent Detours, Ramit Mehr, and Alan S Perelson. A quantitative theory of affinity-driven t cell repertoire selection. *Journal of theoretical biology*, 200(4):389–403, 1999.

[22] T. H. Dracos. *Three-Dimensional Velocity and Vorticity Measuring and Image Analysis Technique: Lecture Notes from the short course held in Zurich, Switzerland*. Kluwer Academic Publisher, 1996.

[23] Takashi Ebihara, A Helena Jonsson, and Wayne M Yokoyama. Natural killer cell licensing in mice with inducible expression of mhc class i. *Proceedings of the National Academy of Sciences*, 110(45):E4232–E4237, 2013.

[24] Julie M Elliott, Joseph A Wahle, and Wayne M Yokoyama. Mhc class i–deficient natural killer cells acquire a licensed phenotype after transfer into an mhc class i–sufficient environment. *Journal of Experimental Medicine*, 207(10):2073–2079, 2010.

[25] SE Fadugba, OH Edogbanya, and SC Zelibe. Crank nicolson method for solving parabolic partial differential equations. *IJA2M*, 1(3):8–23, 2013.

[26] J Doyne Farmer, Norman H Packard, and Alan S Perelson. The immune system, adaptation, and machine learning. *Physica D: Nonlinear Phenomena*, 22(1-3):187–204, 1986.

[27] Nadine C Fernandez, Emmanuel Treiner, Russell E Vance, Amanda M Jamieson, Suzanne Lemieux, and David H Raulet. A subset of natural killer cells achieves self-tolerance without expressing inhibitory receptors specific for self-mhc molecules. *Blood*, 105(11):4416–4423, 2005.

[28] E. A. Finol and C. H. Amon. Flow-induced wall shear stress in abdominal aortic aneurysms: Part II–pulsatile flow hemodynamics. *Computer methods in biomechanics and biomedical engineering*, 5(4):319–28, August 2002.

[29] Stephanie Forrest, Alan S Perelson, Lawrence Allen, and Rajesh Cherukuri. Self-nonself discrimination in a computer. In *Proceedings of 1994 IEEE computer society symposium on research in security and privacy*, pages 202–212. Ieee, 1994.

[30] Simmie L Foster, Diana C Hargreaves, and Ruslan Medzhitov. Gene-specific control of inflammation by tlr-induced chromatin modifications. *Nature*, 447(7147):972, 2007.

[31] Charles W Fox and Stephen J Roberts. A tutorial on variational bayesian inference. *Artificial intelligence review*, 38:85–95, 2012.

[32] Crispin Gardiner. *Stochastic methods*, volume 4. Springer Berlin, 2009.

[33] Carolina Garrido, Maria Abad-Fernandez, Marina Tuyishime, Justin J Pollara, Guido Ferrari, Natalia Soriano-Sarabia, and David M Margolis. Interleukin-15-stimulated natural killer cells clear hiv-1-infected cells following latency reversal ex vivo. *Journal of virology*, 92(12):e00235–18, 2018.

[34] Stephan Gasser and David H Raulet. Activation and self-tolerance of natural killer cells. *Immunological reviews*, 214(1):130–142, 2006.

[35] Thaddeus C George, John R Ortaldo, Suzanne Lemieux, Vinay Kumar, and Michael Bennett. Tolerance and alloreactivity of the ly49d subset of murine nk cells. *The Journal of Immunology*, 163(4):1859–1867, 1999.

[36] Josiah Willard Gibbs. *Elementary principles in statistical mechanics: developed with especial reference to the rational foundations of thermodynamics*. C. Scribner's sons, 1902.

[37] Rachel A Gottschalk, Andrew J Martins, Bastian R Angermann, Bhaskar Dutta, Caleb E Ng, Stefan Uderhardt, John S Tsang, Iain DC Fraser, Martin Meier-Schellersheim, and Ronald N Germain. Distinct nf-κb and mapk activation thresholds uncouple steady-state microbe sensing from anti-pathogen inflammatory responses. *Cell systems*, 2(6):378–390, 2016.

[38] Zvi Grossman and William E Paul. Adaptive cellular interactions in the immune system: the tunable activation threshold and the significance of subthreshold responses. *Proceedings of the National Academy of Sciences*, 89(21):10365–10369, 1992.

[39] Zvi Grossman and William E Paul. Dynamic tuning of lymphocytes: physiological basis, mechanisms, and function. *Annual Review of Immunology*, 33:677–713, 2015.

[40] Karolin Guldevall, Ludwig Brandt, Elin Forslund, Karl Olofsson, Thomas W Frisk, Per E Olofsson, Karin Gustafsson, Otto Manneberg, Bruno Vanherberghen, Hjalmar Brismar, et al. Microchip screening platform for single cell assessment of nk cell cytotoxicity. *Frontiers in immunology*, 7:119, 2016.

[41] Lavesh A Gwalani and Jordan S Orange. Single degranulations in nk cells can mediate target cell killing. *The Journal of Immunology*, 200(9):3231–3243, 2018.

[42] Petter Hoglund, Claes Ohlen, Ennio Carbone, Lars Franksson, Hans-Gustaf Ljunggren, Ann Latour, Beverly Koller, and Klas Karre. Recognition of β2-microglobulin-negative (β2m-) t-cell blasts by natural killer cells from normal but not from β2m-mice: nonresponsiveness controlled by β2m-bone marrow in chimeric mice. *Proceedings of the National Academy of Sciences of the United States of America*, 88:10332–10336, 1991.

[43] Amir Horowitz, Dara M Strauss-Albee, Michael Leipold, Jessica Kubo, Neda Nemat-Gorgani, Ozge C Dogan, Cornelia L Dekker, Sally Mackey, Holden Maecker, Gary E Swan, et al. Genetic and environmental determinants of human nk cell diversity revealed by mass cytometry. *Science translational medicine*, 5(208):208ra145–208ra145, 2013.

[44] Charles A Janeway, Paul Travers, Mark Walport, Mark Shlomchik, et al. *Immunobiology: the immune system in health and disease*, volume 7. Current Biology London, 1996.

[45] Nathalie T Joncker, Nataliya Shifrin, Frédéric Delebecque, and David H Raulet. Mature natural killer cells reset their responsiveness when exposed to an altered mhc environment. *Journal of Experimental Medicine*, 207(10):2065–2072, 2010.

[46] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

[47] Klas Kärre, Hans Gustaf Ljunggren, Gerald Piontek, and Rolf Kiessling. Selective rejection of h–2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature*, 319(6055):675, 1986.

[48] U. Kertzscher, P. Debaene, L. Goubergrits, and K. Affeld. Experimental Assessment of Wall Shear Flow. In *Blood Flow Modeling and Diagnostics*, pages 109–134, 2005.

[49] Salim I Khakoo, Chloe L Thio, Maureen P Martin, Collin R Brooks, Xiaojiang Gao, Jacquie Astemborski, Jie Cheng, James J Goedert, David Vlahov, Margaret Hilgartner, et al. Hla and nk cell inhibitory receptor genes in resolving hepatitis c virus infection. *Science*, 305(5685):872–874, 2004.

[50] Sungjin Kim, Jennifer Poursine-Laurent, Steven M Truscott, Lonnie Lybarger, Yun-Jeong Song, Liping Yang, Anthony R French, John B Sunwoo, Suzanne Lemieux, Ted H Hansen, et al. Licensing of natural killer cells by host major histocompatibility complex class i molecules. *Nature*, 436(7051):709, 2005.

[51] Andrej Košmrlj, Abhishek K Jha, Eric S Huseby, Mehran Kardar, and Arup K Chakraborty. How the thymus designs antigen-specific and self-tolerant t cell receptor sequences. *Proceedings of the National Academy of Sciences*, 105(43):16671–16676, 2008.

[52] M. Kraizer. *Real-time image processing for particle tracking velocimetry.* PhD thesis, Tel aviv university, 2010.

[53] Matthew F Krummel, Frederic Bartumeus, and Audrey Gérard. T cell migration, search strategies and mechanisms. *Nature Reviews Immunology*, 16(3):193–201, 2016.

[54] Himanshu Kumar, Taro Kawai, and Shizuo Akira. Pathogen recognition by the innate immune system. *International reviews of immunology*, 30(1):16–34, 2011.

[55] Lewis L Lanier. Nk cell receptors. *Annual review of immunology*, 16(1):359–393, 1998.

[56] N. Laptev and S. Amizadeh. ahoo anomaly detection dataset s5, 2015.

[57] Alexander Lavin and Subutai Ahmad. Evaluating real-time anomaly detection algorithms–the numenta anomaly benchmark. In *2015 IEEE 14th international conference on machine learning and applications (ICMLA)*, pages 38–44. IEEE, 2015.

[58] Nan-Shih Liao, Mark Bix, Maarten Zijlstra, Rudolf Jaenisch, and David Raulet. Mhc class i deficiency: susceptibility to natural killer (nk) cells and impaired nk activity. *Science*, 253(5016):199–202, 1991.

[59] Eric Libby, Theodore J Perkins, and Peter S Swain. Noisy information processing through transcriptional regulation. *Proceedings of the National Academy of Sciences*, 104(17):7151–7156, 2007.

[60] Jarno Lintusaari, Henri Vuollekoski, Antti Kangasrääsiö, Kusti SkytÃ©n, Marko Järvenpää, Pekka Marttinen, Michael U. Gutmann, Aki Vehtari, Jukka Corander, and Samuel Kaski. Elfi: Engine for likelihood-free inference. *Journal of Machine Learning Research*, 19(16):1–7, 2018.

[61] Hans-Gustaf Ljunggren and Klas Kärre. In search of the 'missing self': Mhc molecules and nk cell recognition. *Immunology today*, 11:237–244, 1990.

[62] Eric O Long, Hun Sik Kim, Dongfang Liu, Mary E Peterson, and Sumati Rajagopalan. Controlling natural killer cell responses: integration of signals for activation and inhibition. *Annual review of immunology*, 31:227–258, 2013.

[63] Z. Lou, W. J. Yang, and P . D. Stein. Errors in the estimation of arterial wall shear rates that result from curve fitting of velocity profiles. *Journal of biomechanics*, 26(4-5):383–90, 1993.

[64] Wen-Tao Ma, Fei Gao, Kui Gu, and De-Kun Chen. The role of monocytes and macrophages in autoimmune diseases: A comprehensive review. *Frontiers in immunology*, 10, 2019.

[65] H. G. Maas. Contribution of digital photogrammetry to 3D PTV.

[66] Maureen P Martin, Ying Qi, Xiaojiang Gao, Eriko Yamada, Jeffrey N Martin, Florencia Pereyra, Sara Colombo, Elizabeth E Brown, W Lesley Shupert, John Phair, et al. Innate partnership of hla-b and kir3dl1 subtypes against hiv-1. *Nature genetics*, 39(6):733, 2007.

[67] Ludovic Martinet and Mark J Smyth. Balancing natural killer cell activation through paired receptors. *Nature Reviews Immunology*, 15(4):243, 2015.

[68] Andreas Mayer, Vijay Balasubramanian, Thierry Mora, and Aleksandra M Walczak. How a well-adapted immune system is organized. *Proceedings of the National Academy of Sciences*, 112(19):5950–5955, 2015.

[69] Andreas Mayer, Vijay Balasubramanian, Aleksandra M Walczak, and Thierry Mora. How a well-adapting immune system remembers. *Proceedings of the National Academy of Sciences*, 116(18):8815–8823, 2019.

[70] Ruslan Medzhitov and Charles A Janeway. Decoding the patterns of self and nonself by the innate immune system. *Science*, 296(5566):298–300, 2002.

[71] Sven Mesecke, Doris Urlaub, Hauke Busch, Roland Eils, and Carsten Watzl. Integration of activating and inhibitory receptor signaling by regulated phosphorylation of vav1 in immune cells. *Sci. Signal.*, 4(175):ra36–ra36, 2011.

[72] J. E. Moore, J. C. Xu, S. Glagov, C. K. Zarins, and D. N. Ku. Fluid wall shear stress measurements in a model of the human abdominal aorta: oscillatory behavior and relationship to atherosclerosis. *Atherosclerosis*, 110(2):225–240, October 1994.

[73] Sayak Mukherjee, Helle Jensen, William Stewart, David Stewart, William C Ray, Shih-Yu Chen, Garry P Nolan, Lewis L Lanier, and Jayajit Das. In silico modeling identifies cd45 as a regulator of il-2 synergy in the nkg2d-mediated activation of immature human nk cells. *Sci. Signal.*, 10(485):eaai9062, 2017.

[74] David E Oppenheim, Scott J Roberts, Sarah L Clarke, Renata Filler, Julie M Lewis, Robert E Tigelaar, Michael Girardi, and Adrian C Hayday. Sustained localized expression of ligand for the activating nkg2d receptor impairs natural cytotoxicity in vivo and reduces tumor immunosurveillance. *Nature immunology*, 6(9):928, 2005.

[75] Mark T Orr and Lewis L Lanier. Natural killer cell education and tolerance. *Cell*, 142(6):847–856, 2010.

[76] Jens Pahl and Adelheid Cerwenka. Tricking the balance: Nk cells in anti-cancer immunity. *Immunobiology*, 222(1):11–20, 2017.

[77] Alan S Perelson and George F Oster. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of theoretical biology*, 81(4):645–670, 1979.

[78] Alan S Perelson and Gérard Weisbuch. Immunology for physicists. *Reviews of modern physics*, 69(4):1219, 1997.

[79] Thomas Pradeu, Sébastien Jaeger, and Eric Vivier. The speed of change: towards a discontinuity theory of immunity? *Nature Reviews Immunology*, 13(10):764–769, 2013.

[80] Isabel Prager, Clarissa Liesche, Hanna Van Ooijen, Doris Urlaub, Quentin Verron, Niklas Sandström, Frank Fasbender, Maren Claus, Roland Eils, Joël Beaudouin, et al. Nk cells switch from granzyme b to death receptor–mediated cytotoxicity during serial killing. *Journal of Experimental Medicine*, 216(9):2113–2127, 2019.

[81] SP Preston, SL Waters, OE Jensen, PR Heaton, and DI Pritchard. T-cell motility in the early stages of the immune response modeled as a random walk amongst targets. *Physical Review E*, 74(1):011910, 2006.

[82] Yawei Qin, Emily M Mace, and John P Barton. A quantitative model for the regulation of innate immune activation. *bioRxiv*, pages 2020–04, 2020.

[83] Thomas J Rademaker, Emmanuel Bengio, and Paul François. Attack and defense in cellular decision-making: Lessons from machine learning. *Physical Review X*, 9(3):031012, 2019.

[84] David H Raulet and Russell E Vance. Self-tolerance of natural killer cells. *Nature Reviews Immunology*, 6(7):520, 2006.

[85] David H Raulet, Russell E Vance, and Christopher W McMahon. Regulation of the natural killer cell receptor repertoire. *Annual review of immunology*, 19(1):291–330, 2001.

[86] H. Risken. *The Fokker-Planck Equation: Methods of Solution and Applications.* Springer-Verlag, 1989.

[87] Hannes Risken and Hannes Risken. *Fokker-planck equation.* Springer, 1996.

[88] Almut Scherer, André Noest, and Rob J de Boer. Activation–threshold tuning in an affinity model for the t–cell repertoire. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1539):609–616, 2004.

[89] Katja Srpan, Ashley Ambrose, Alexandros Karampatzakis, Mezida Saeed, Adam NR Cartwright, Karolin Guldevall, Gabriela Dos Santos Cruz De Matos, Björn Önfelt, and Daniel M Davis. Shedding of cd16 disassembles the nk cell immune synapse and boosts serial engagement of target cells. *Journal of Cell Biology*, 217(9):3267–3283, 2018.

[90] Dara M Strauss-Albee, Julia Fukuyama, Emily C Liang, Yi Yao, Justin A Jarrell, Alison L Drake, John Kinuthia, Ruth R Montgomery, Grace John-Stewart, Susan Holmes, et al. Human nk cell repertoire diversity reflects immune experience and correlates with viral susceptibility. *Science translational medicine*, 7(297):297ra115–297ra115, 2015.

[91] Dara M Strauss-Albee, Amir Horowitz, Peter Parham, and Catherine A Blish. Coordinated regulation of nk receptor expression in the maturing human immune system. *The Journal of Immunology*, 193(10):4871–4879, 2014.

[92] Joseph C Sun and Lewis L Lanier. Tolerance of nk cells encountering their viral ligand during development. *Journal of Experimental Medicine*, 205(8):1819–1828, 2008.

[93] Osamu Takeuchi and Shizuo Akira. Pattern recognition receptors and inflammation. *Cell*, 140(6):805–820, 2010.

[94] Nidale Tarek, Jean-Benoit Le Luduec, Meighan M Gallagher, Junting Zheng, Jeffrey M Venstrom, Elizabeth Chamberlain, Shakeel Modak, Glenn Heller, Bo Dupont, Nai-Kong V Cheung, et al. Unlicensed nk cells target neuroblastoma following anti-gd2 antibody treatment. *The Journal of clinical investigation*, 122(9):3260–3270, 2012.

[95] Philip R Taylor, Luisa Martinez-Pomares, Martin Stacey, Hsi-Hsen Lin, Gordon D Brown, and Siamon Gordon. Macrophage receptors and immune recognition. *Annu. Rev. Immunol.*, 23:901–944, 2005.

[96] Markus Thill, Wolfgang Konen, and Thomas Bäck. Online anomaly detection on the webscope s5 dataset: A comparative study. In *2017 Evolving and Adaptive Intelligent Systems (EAIS)*, pages 1–8. IEEE, 2017.

[97] Sandeep K Tripathy, Peter A Keyel, Liping Yang, Jeanette T Pingel, Tammy P Cheng, Achim Schneeberger, and Wayne M Yokoyama. Continuous engagement of a self-specific activation receptor induces nk cell tolerance. *Journal of Experimental Medicine*, 205(8):1829–1841, 2008.

[98] Bruno Vanherberghen, Per E Olofsson, Elin Forslund, Michal Sternberg-Simon, Mohammad Ali Khorshidi, Simon Pacouret, Karolin Guldevall, Monika Enqvist, Karl-Johan Malmberg, Ramit Mehr, et al. Classification of human natural killer cells based on migration behavior and cytotoxic response. *Blood*, 121(8):1326–1334, 2013.

[99] Eric Vivier, Jacques A Nunès, and Frédéric Vély. Natural killer cell signaling pathways. *Science*, 306(5701):1517–1519, 2004.

[100] Eric Vivier, David H Raulet, Alessandro Moretta, Michael A Caligiuri, Laurence Zitvogel, Lewis L Lanier, Wayne M Yokoyama, and Sophie Ugolini. Innate or adaptive immunity? the example of natural killer cells. *science*, 331(6013):44–49, 2011.

[101] Eric Vivier, Elena Tomasello, Myriam Baratin, Thierry Walzer, and Sophie Ugolini. Functions of natural killer cells. *Nature immunology*, 9(5):503, 2008.

[102] Inja Waldhauer and Alexander Steinle. Nk cells and cancer immunosurveillance. *Oncogene*, 27(45):5932, 2008.

[103] Hui Wang, Maria Lucia Madariaga, Shumei Wang, Nico Van Rooijen, Per-Arne Oldenborg, and Yong-Guang Yang. Lack of cd47 on nonhematopoietic cells induces split macrophage tolerance to cd47null cells. *Proceedings of the National Academy of Sciences*, 104(34):13744–13749, 2007.

[104] Carsten Watzl, Michal Sternberg-Simon, Doris Urlaub, and Ramit Mehr. Understanding natural killer cell regulation by mathematical approaches. *Frontiers in immunology*, 3:359, 2012.

[105] Greg Welch and Gary Bishop. An introduction to the kalman filter. *University of North Carolina at Chapel Hill, Chapel Hill, NC*, 127–132, 1995.

[106] Michael A West and Wyrta Heagy. Endotoxin tolerance: a review. *Critical care medicine*, 30(1):S64–S73, 2002.

[107] C. E. Willert and M. Gharib. Digital particle image velocimetry. *Experiments in fluids*, 10(4):181–193, 1991.

[108] J. R. Womersley. Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. *The journal of physiology*, 127(3):553, 1955.

[109] Inge Wortel, Can Keşmir, Rob J de Boer, Judith N Mandl, and Johannes Textor. Is t cell negative selection a learning algorithm? *Cells*, 9(3):690, 2020.

[110] Renjie Wu and Eamonn Keogh. Current time series anomaly detection benchmarks are flawed and are creating the illusion of progress. *IEEE Transactions on Knowledge and Data Engineering*, 2021.

[111] Yuanhao Xu, Shufan Zhou, Yun Wah Lam, and Stella W Pang. Dynamics of natural killer cells cytotoxicity in microwell arrays with connecting channels. *Frontiers in immunology*, 8:998, 2017.

[112] Wayne M Yokoyama and Sungjin Kim. How do natural killer cells find self to achieve tolerance? *Immunity*, 24(3):249–257, 2006.

[113] Jacques Zimmer, Lionel Donato, Daniel Hanau, Jean-Pierre Cazenave, Marie-Marthe Tongio, Alessandro Moretta, and Henri de la Salle. Activity and phenotype of natural killer cells in peptide transporter (tap)-deficient patients (type i bare lymphocyte syndrome). *Journal of Experimental Medicine*, 187(1):117–122, 1998.

[114] Wendy M Zinzow-Kramer, Arthur Weiss, and Byron B Au-Yeung. Adaptation by naïve cd4+ t cells to self-antigen–dependent tcr signaling induces functional heterogeneity and tolerance. *Proceedings of the National Academy of Sciences*, 116(30):15160–15169, 2019.