Adaptive Dynamics of Intra- and Inter-Cellular Networks: Emergence of Memory and Learning in Cell Signaling and Immune Systems

By

Tanmay Mitra

PHYS10201005004

The Institute of Mathematical Sciences, Chennai

A thesis submitted to the

Board of Studies in Physical Sciences

In partial fulfillment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of
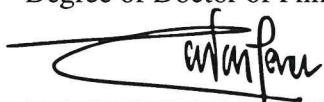
HOMI BHABHA NATIONAL INSTITUTE

December, 2018

Homi Bhabha National Institute

Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by **Tanmay Mitra** entitled "Adaptive Dynamics of Intra- and Inter-Cellular Networks: Emergence of Memory and Learning in Cell Signaling and Immune Systems" and recommend that it maybe accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.
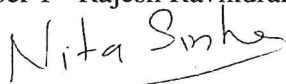
_____ Date: 22|8|19

Chair - Gautam I. Menon

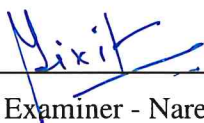_____ Date: 22.08.19

Supervisor/Convener - Sitabhra Sinha

_____ Date: 22|8|19

Member 1 - Rajesh Ravindran

_____ Date: 22/08/2019

Member 2 - Nita Sinha

_____ Date: 22|8|2019

External Examiner - Narendra M. Dixit

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it may be accepted as fulfilling the dissertation requirement.

Date: 22.08.19

Place: IMSc Chennai

Supervisor

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Tanmay Mitra.

Tanmay Mitra

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier in whole or in part for a degree / diploma at this or any other Institution / University.

Tanmay Mitra

ACKNOWLEDGEMENTS

PUBLICATIONS

a. <u>Published:</u>

1. **Mitra, T.**, Menon, S. N. and Sinha, S. (2018). *Emergent memory in cell signaling: Persistent adaptive dynamics in cascades can arise from the diversity of relaxation time-scales.*
Scientific Reports, **8**, 13230. Arxiv Preprint, arXiv:1801.04057.

b. <u>Accepted:</u>

c. <u>Communicated:</u>

1. **Mitra, T.**, Menon, S. N. and Sinha, S. (2018). *Non-associative learning in intra-cellular signaling networks.*
Arxiv Preprint, 1807.01243.

Tanmay Mitra.
[TANMAY MITRA]

Contents

List of Figures

List of Tables

Synopsis

Complex networks of interacting components are seen across a broad range of spatio-temporal scales in biology – from the levels of molecules and cells to that of ecosystems. Two examples of such systems are the intra-cellular signaling network (comprising molecules such as kinase proteins, second messengers, cell-surface receptors, etc.) and the immune system (comprising cells that coordinate the immunological response of the body, such as T-cells, B-cells, etc.). The aim of the thesis is to obtain a clearer understanding of aspects of the dynamical evolution of such systems using modelling and data-analysis techniques. As both of the biological networks mentioned above are involved in functions that are vital for the continued survival of an organism, explicating the dynamical processes underlying their behavior may play a crucial role in developing better and/or more effective treatment of diseases arising through disruptions of these networks.

In **Chapter 1** we begin with a short overview of the literature on intra-cellular signaling pathways, as well as immune networks. We discuss the biological networks in the post translational levels ranging from the single cell intra-cellular network to cell-cell communication networks. We also focus on various biological roles of MAPK cascade and how decision making at the level of a single cell is shaped by the temporal responses of this pathway. A different level of complexity is seen when one moves from the scale of molecular communication within a single cell to that of inter-cellular interactions which is seen, for example, in the immune system. We briefly review the literature that pertains to the development of the immune system as it evolves over the life-time of an individual.

In **Chapter 2** we consider the mitogen-activated protein kinase (MAPK) signaling cascade, an evolutionarily conserved motif present in all eukaryotic cells. Intra-cellular signaling networks coordinate the entire range of biological processes that are necessary for the cell to provide appropriate responses, even in the presence of a high degree of noise, to a wide variety of environmental signals. As a breakdown in communication between different members of this network leads to pathological outcomes, it is important to understand the underlying processes that allow robust information processing in the system. An ubiquitous component (motif) of the signaling network is the MAPK cascade, comprising the three kinase proteins, viz., MAPK, MAPK-Kinase (MAP2K), and MAP2K-Kinase (MAP3K). This signaling pathway is found to exist in all eukaryotic cells, and is involved in many critical cellular functions including cell cycle control, stress response, differentiation and growth. Its crucial importance for the proper functioning of an organism is underscored by the fact that it is affected in many diseases including cancer, auto-immune disorders and degenerative syndromes, thereby making it an important drug target. The basic linear cascade dynamics involves regulation of the activity of the MAP3K protein

by an upstream signal. On being activated, MAP3K acts as the enzyme for activating the MAP2K protein, which in turn controls the activation of the MAPK protein. The activated MAPK is known to be involved in many functions, such as the initiation of transcription of genes or stimulating the activity of other kinases.

While the steady-state behavior of the MAPK pathway stimulated by a time-invariant signal is relatively well-understood, we show using a computational model that it exhibits a rich repertoire of transient adaptive responses to changes in stimuli. When the signal is switched on, the response is characterized by long-lived modulations in frequency as well as amplitude. On withdrawing the stimulus, the activity decays over time-scales much longer than that of phosphorylation-dephosphorylation processes, exhibiting reverberations characterized by repeated spiking in the activated MAPK concentration. The long-term persistence of such post-stimulus activity suggests that the cascade retains memory of the signal for a significant duration following its removal, even in the absence of any explicit feedback or cross-talk with other pathways. We find that the molecular mechanism underlying this behavior is related to the existence of distinct relaxation rates for the different cascade components. This results in the imbalance of fluxes between different layers of the cascade, with the repeated reuse of activated kinases as enzymes when they are released from sequestration in complexes leading to one or more spike events following the removal of the stimulus. The persistent adaptive response reported here, indicative of a cellular "short-term" memory, suggests that this ubiquitous signaling pathway plays an even more central role in information processing by eukaryotic cells. In addition, we consider the dependence of the reverberations on the concentrations of the kinases and phosphatases, as well as on the variations of individual parameters. This allows us to establish the robustness of the results reported here, which is essential for future experimental validation.

Nonlinear systems driven by recurrent signals are known to exhibit complex dynamical responses which, in the physiological context, can have important functional consequences. In **Chapter 3** we consider response of the MAPK cascade to a periodic train of pulses. The resulting response of the cascade, which shows integrative capability over several successive pulses, is characterized by complex adaptive behavior. Depending on circumstances, periodic stimulation can result in an enhanced response (sensitization) or a diminished one (adaptation). We also observe responses characterized by alternate high and low MAPK activity, or an attenuated response even for strong stimuli. The ensemble of these responses provides a basis for signal integration and non-associative learning by the cellular signaling network. Learning is usually associated with multi-cellular organisms, e.g., those possessing a nervous system. However, our computational study shows that simpler forms of learning may appear in more rudimentary systems - as simple as the canonical MAPK motif at the level of a single eukaryotic cell. Taken in conjunction with previous analogies made between the intra-cellular signaling network and the nervous system, our study provides an intriguing perspective on how this signaling cascade can "remember" and "learn" even in the absence of any explicit feedback or cross-talk with other pathways. As the MAPK signaling motif is involved in crucial cellular functions in all eukaryotic cells, the potential consequences of such emergent memory and adaptive response are far-reaching. In addition, the existence of a response threshold of the cascade, an apparent refractory behavior following stimulation with short inter-pulse

interval, and an alternans-like response under certain conditions suggest an analogy with excitable media.

In **Chapter 4** we move from the intra-cellular to the inter-cellular scale and consider a different type of adaptive complex system, viz., the immune system. A salient characteristic of the human immune system is that it adapts over time in response to foreign antigens and distress signals. The immune system of a child not yet born would have little experience of the factors that can stimulate it, which manifests in the relative insignificance of memory cells. Adults, in contrast, are more likely to have had considerable exposure to a plethora of signals that activate the immune system. Consequently, the cellular composition of the adaptive component of the system will be expected to differ significantly, in particular in terms of possessing a significant complement of memory cells. In addition, correlations between different immune system cell-types are established over time as a result of their interactions. In this chapter, we investigate the genetic, environmental and developmental signatures of the functional network inferred from the correlations by analyzing the populations of various cell types in both umbilical cord and adult blood samples. We find that the latter exhibits a higher degree of correlation in the proportions of cells of the adaptive immune system, suggesting a strong role played by maturation in the evolution of the system. Our analysis validates several correlations between different cell types that have been alluded to in the literature and also suggests a few previously unreported relations. The results reported here have been explicitly verified to be not severely affected by the sizes of the samples investigated. The empirical data used for our analysis has been provided by Dr Vineeta Bal (National Institute of Immunology, New Delhi).

We conclude in **Chapter 5** with a summary and general discussion of the implications of our results. We also briefly outline possible future extensions of our studies presented in the thesis.

1

Introduction

Biological systems have an astounding ability to adapt to their environment, adjusting their response to stimuli over time [1]. The phenomenon wherein repeated exposure of such systems to stimulation eventually results in the alteration of their response is often referred to as learning and usually associated with multi-cellular organisms, e.g. those possessing a nervous system [2–7]. In general, the capacity to 'remember' previous experiences and 'learning' from them in order to alter subsequent responses is a key survival trait in the animal kingdom [8]. Such capabilities have been investigated in several model organisms, such as the sea slug *Aplysia*, by neuroscientists. It is observed that when *Aplysia* is exposed to a light tactile stimulation, for example, by a gentle touch, it show only a weak manifested in siphon and gill withdrawal. However, if the stimulus is paired with a strong electrical shock to its tail, it results in a marked defensive reflex. Following a few trials it is seen that the touch alone can evoke an enhanced siphon and gill withdrawal reflex [9]. This association is seen to emerge relatively fast (within 15 trials) and last for several days. It can therefore be seen as an instance of classical conditioning, a form of associative learning, with the touch stimulus as the conditioned stimulus and the electrical shock as the unconditioned one [9–11].

Rudimentary forms of learning have been seen even in organisms which do not possess

neurons [12, 13]. Even single-celled organisms can exhibit adaptation in response to changes in their environment - which is distinct from evolutionary changes occurring over a long time-scale involving several generations such as in bacteria developing antibiotic resistance [14]. For instance, a simple form of habituation has recently been identified in the single-celled protist *Physarum Polycephalum* (popularly known as slime mold) which traces back the origin of learning in biological systems to much earlier evolutionarily time-scales (almost 500 million years ago) than previously believed [13]. Habituation refers to phenomenon where the behavioral response to a stimulus diminishes when the organism is repeatedly exposed to the stimulus. It is a form of non-associative learning [15, 16] (which is marked by the absence of conditioning) with the response reverting to its initial nature when further stimulation is stopped. Sensitization provides another example of non-associative learning where the response to a stimulus is amplified upon reinforcement.

Moving away from behavioral responses, a different form of learning is seen to shape the vertebrate immune system [17–21]. This system comprises both innate and adaptive immune cells that respond to a broad range of invaders detected via means of pathogen-associated molecular patterns (PAMPs) and which subsequently generates an immune response to specific antigens associated with the pathogens [22, 23]. The immune system has to learn to *not attack* host cells (failure of which would lead to autoimmune disorders) through several mechanisms such as negative selection of T cells in the thymus [18] and peripheral anergy [24]. Further, adaptive mechanisms throughout development from infancy to adulthood shape the interactions and functions of various immune components [25]. Learning from previous exposures to a pathogen enable the immune system to mount a more effective response aimed at eliminating the pathogen rapidly during subsequent invasions over the lifetime of the host. Whereas learning in the adaptive components of the immune system provides protection against specific antigens through the existence of memory cells [20, 25], learning in the innate components (for example, Natural Killer cells) changes their response on the basis of previous exposures [17]. For instance, in

6

invertebrates which lack adaptive immune cell types (T cells and B cells), immunological 'memory' is deployed solely by the innate components. As the immune system adapts over time in response to exposure to antigens and distress signals, learning in the innate as well as adaptive immune system shapes the functional network connecting its different components. As a result, new correlations between different immune system cell types are established in the course of development.

Be it a single cell, a non-neural organism, a multi-cellular organism having a nervous system, or the immune system, each of these biological systems are capable of complex adaptive responses to a plethora of external and internal stimuli in their micro-environment [26–28]. In general, such complex adaptive systems are characterized by a non-trivial repertoire of behavioral responses emerging from typically non-linear spatio-temporal interactions between their components [28]. Such systems can often *remember* prior exposures to stimuli and *learn* to exhibit altered response following repeated exposures. *Memory* of such a system alludes to its capability to retain information about a signal which has subsequently been withdrawn from its environment. Emergence of memory and learning is observed in various contexts in eukaryotic cells, ranging from directional memory in a single chemotactic cell (that enables cells to orient themselves even under low signal gradients or when the gradient changes over time) to neurobiological memory displayed by an assembly of neurons connected to each other through chemical synapses and electrical gap junctions. The detailed mechanisms underlying the formation of memory in each of these systems are context specific and can involve nonlinear interactions among a large set of components, that could be embedded at multiple levels in biological networks. In this thesis, we will demonstrate the emergence of memory and non-associative learning in an evolutionarily conserved intra-cellular signal transduction motif, the Mitogen Activated Protein Kinase (MAPK) cascade, which is ubiquitous among eukaryotic cells. Moving over to the scale of inter-cellular networks, we will also show the changes that take place in the network of functional interactions between different component cell-types of human immune systems between birth and adulthood, as

a result of various antigenic and distress triggers as well as environmental, genetic and developmental factors.

1.1 Intra-cellular Networks

Determining the appropriate response to a signal from its environment requires non-trivial information processing by a cell [29]. For the proper functioning of the cell, it requires a wide range of biochemical molecules that mediate the myriad dynamical processes going on within it [30]. Many of these molecules interact among themselves to perform one or more tasks, thereby creating intra-cellular networks of nucleic acids, metabolites, proteins, etc. [31]. The diversity of functional processes inside the cell are now known to be controlled by means of a number of such complex networks [32] involved in the regulation of gene expression [33], metabolism [34] and signal transduction [31, 35], among others. Intrinsic fluctuations and perturbative influences from the environment can affect the interaction affinities thereby making these intra-cellular networks dynamic in nature [29, 36, 37]. Over the past couple of decades, there has been great interest among physicists to understand the dynamics of these networks by considering them as systems operating out of equilibrium. Often these complex dynamical networks involve regulation of the activity of their constituent nodes through activation or inhibition [37]. For instance, in gene regulatory networks, genes express proteins that either enhance or reduce the activity of other genes by acting as promoters or suppressors, respectively [33, 38]. In the case of signal transduction network which enables the cell to encode the information present in its micro-environment and transfer this information through a cascade of linked enzyme-substrate reactions, the activated kinase proteins activate other proteins by phosphorylating them while phosphatases act as inhibitors that dephosphorylate the activated proteins [39]. While it had taken several years of effort to uncover the glycolytic pathway (a metabolic network), modern experimental methodologies allow the study of physical interactions between pairs of proteins with ease [40]. The ability to perform

high-throughput experiments have contributed to the reconstruction of several complex networks of biochemical entities inside the cells, which has prompted further theoretical and experimental research aimed at understanding how signals are encoded and decoded in these biological complex dynamic networks [41]. We now give a brief summary of such efforts focused on the intra-cellular signal transduction networks.

Living cells are being constantly bombarded by large numbers of chemical signals that impinge on their cell membranes. The receptors present on the membrane can bind with ligands upon encountering them and thereby act as sensors to the extra-cellular environment. A wide variety of receptors are embedded on the cell membrane in order to identify chemical signals that they are specifically designed to detect [42]. Different cells have different types of receptors, with their frequency counts also varying from cell to cell. Signal transduction mediates the process by which conformational changes in receptors on binding with ligands are converted into information to be sent inside the cell nucleus for subsequent decoding in order to regulate intra-cellular processes by initiating a relay of linked biochemical reactions [29, 41, 42]. Thus, the activated receptor causes a cascade of reactions where each downstream signaling molecule is activated by the activated signaling protein on the level immediately above it. Eventually, each reaction pathway results in the activation of transcription factors that modifies expression of certain genes thereby determining the eventual response of the cell [43]. The signal transduction pathways can involve metabolic enzymes. Regulation of the intra-cellular pathways is mediated by post-translational modifications like phosphorylation and dephosphorylation, and are carried out by a wide range of protein molecules such as kinases and phosphatases [39]. Adaptor molecules can act as scaffolding, i.e., bringing various molecules in physical proximity so that they can interact [44]. There can be several signaling molecules which can take part in different pathways, thereby effectively creating a cross-talk between these signaling pathways. The connections within the signaling networks are generated through biological evolution [31]. The signal transduction networks have been claimed to have properties such as small-world-ness and scale-free topology [45]. While protein-protein interaction

9

networks, which are also intra-cellular, are isotropic and undirected [45], signal transduction networks are anisotropic and directed [31, 45]. As we move up the evolutionary tree, the complexity, degeneracy and occasional redundancy in the functions of these intra-cellular signal transduction networks make them extremely daunting to analyze and understand [35, 46, 47].

1.2   Inter-Cellular Networks

Efficient communication between cells is one of the fundamental necessities for robust functioning in an organ, as well as, a key step for development in multicellular organisms [48–50]. Information processing between cells is crucial in mediating homeostasis and adaptation in inter-cellular networks [49, 51]. There is a plethora of complex biophysical mechanisms by which a cell can send or receive signal to or from another cell. It can be implemented through diffusion of chemical messengers [52] (for instance, hormones or growth factors in paracrine signaling), transportation of extracellular vesicles [53], cellular bridge of membrane nanotubes [54] and gap junctions [55]. Two intriguing examples of adaptive complex biological networks that operate at the scale of multiple cells are the nervous system and the immune system. Various types of cytokines and chemokines mediate the communication among components of the immune system comprising primarily of innate and adaptive immune cells [23]. In the case of the nervous system, several types of neurotransmitters coordinate the activity of neurons which are the nodes of the inter-cellular neural network [56]. On the other hand, inter-cellular signaling in animal development is regulated by several positive and negative feedback interactions between participating cells, as well as, a set of inter-cellular signaling molecules which enable dynamical regulation. For example, the inter-cellular WNT proteins control tissue organization and formation of body axis in vertebrates during development through complex interactions mediated by membrane bound frizzled receptors [50]. An intriguing role of macrophage (a cell type present in the innate immune system) that has recently been un-

10

covered implicates it in the establishment of long-range interactions in the inter-cellular network of non-immune cells during postembryonic tissue remodelling [57]. There is also growing evidence suggesting complex interplay between brain, hypothalamic–pituitary–adrenocortical (HPA) axis and the immune system [58]. The presence of certain immune cells inside the brain and central nervous system, direct influence by the nerve pathways on immune system organs (lymph nodes, spleen), stress mediated immune response, emergence of autoimmunity in phases of emotional turmoil, etc. hint towards a far more complex set of interactions between cells within an organism than previously believed [26]. Thus, having a complete understanding of these inter-cellular networks and their interplay can be daunting task. The non-equilibrium collective dynamics of these networks resulting in adaptability and maintenance of their functional robustness under small perturbations is therefore a biophysical problem of great interest.

1.3   Overview of The Immune System

"Yet it was with those who had recovered from the disease that the sick and the dying found most compassion. These knew what it was from experience, and had now no fear for themselves; for the same man was never attacked twice – never, at least, fatally." – Historian Thucydides in his description of the plague of Athens (430 BC) in *History of the Peloponnesian War*

In nature, every form of life is armed with some mechanism to provide it defence against potentially harmful agents. For vertebrates, this defensive mechanism is implemented in an immune system which can be subdivided into an innate and an adaptive component [22, 23]. As the name suggests, the cell types which are part of the innate system are already encoded with some broad range of pathogen specific receptors in the germline and can be activated by encountering a diverse array of PAMPs to rapidly eliminate foreign substances [22, 59]. For example, the phagocytes that consist of basophils, macrophages,

monocytes, etc. are one of the fundamental cell types of the innate immune systems [60]. These cells can destroy microbes by simply engulfing them, a process known as phagocytosis which is generally initiated when these cells recognize generic structures shared by a large class of microbes [59]. For instance, lipopolysaccharides, which are observed in cell walls of many bacteria, can trigger the response of innate immune cells [22, 59]. Evolutionarily, innate immunity is the older form of defense mechanism which is predominant in invertebrates as well as in fungi and plants [61, 62]. Apart from offering first-line rapid response against foreign infectious microbes, the innate immune cells also play role in antigen presentation (for example, by dendritic cells), recruitment of immune cells to the sites of inflammation, production of cytokines that act to mediate biological inter-cellular communication in the immune system, and regulation of adaptive immune response by the complement proteins (for example, C3 and C4) [23, 60, 63–65]. It is worth noting that some components of the blood coagulation system in the vertebrates are also capable of generic defense and can participate as chemotactic signal mediators for phagocytes [66]. Recent discoveries have shown links between irregularity of various components in the coagulation system and several autoimmune, as well as, neuroinflammatory disorders [67, 68].

The adaptive component of the immune system which evolved in vertebrates, on the other hand, can 'acquire' traits or 'adapt' to changing circumstances according to the history of its exposure to foreign substances [23]. Although evolutionarily it can be seen first in jawed fishes [69, 70], the presence of Ig gene homologs in invertebrates point towards its link with ancient immune precursors [19, 71]. The immune cells that show adaptive immunity are orchestrated with receptors highly specific to fragments of different infectious agents, non-infectious foreign molecules (such as, pollen grains or toxins) and even altered self-molecules (for example, in cancer). In general, any substance that can trigger the immune response of adaptive immune cells is considered as an antigen [72]. In the case of autoimmune disorders, following failure in one or multiple regulatory mechanisms of the immune system, the adaptive components develop immune response to the

bio-molecules belonging to one's own body and the proteins that initiate the autoimmune responses are called self-antigens [73]. When innate immunity fails to clear an invasion by foreign molecules, the adaptive immune cells mount an effective immune response by recognizing antigens through their specific molecular signatures (such as, a particular protein present on the surface of bacteria) [22, 23]. The primary function of adaptive immune system is the recognition of specific antigens by lymphocytes which can be broadly classified into T cells and B cells [23]. Both of these adaptive cell types are successors of their common lymphoid progenitors which are derived from hematopoietic stem cells (HSC) in the bone marrow [74]. Before coming into circulation in the peripheral blood as mature T cells, the immature T cells undergo 'training' in the thymus (hence the name 'T') where potential autoreactive T cells are killed by apoptosis through negative selection [18]. Activation of mature T cells in periphery is mediated by the recognition process of specific antigens which are presented to them by the innate immune cells in association with Major Histocompatibility Complex (MHC) [22, 23]. However, B Cell receptors can recognize antigens directly or through antigen presentation by innate cells, and then get activated to differentiate into large number of specific Plasma B cells (a process called Clonal Expansion) which generate enormous amount of immunoglobulin or antibody, a soluble form of BCR specific to the recognized antigen [75]. The antibodies play a role in marking the pathogens that express the specific type of antigen to be subsequently killed by immune cells. A priori, there is no way that the adaptive immune system would know the specific types of antigens which it would encounter. Hence, it prepares millions of cells capable of recognizing specific set of molecular signatures, thereby practically making the host able to elicit immune response to a diverse range of antigens. [76]

A salient feature of the adaptive immunity is that it changes the adaptive immune response based on prior exposure(s) [23]. Some of the differentiated cells derived from the activated T cells and B cells form memory cells and can be extremely long-lived. Immunological memory of a previously encountered antigen makes the adaptive immune cells capable of mounting a stronger immune response on subsequent encounters with the

13

same antigen in order to eliminate it rapidly [77, 78]. During the initial encounter, only a few lymphocytes can elicit effective immune response which is known as the primary response. However, the formations of long-lived memory T cells and memory B cells during this primary response make sure that the secondary response to the same antigen(s) would be of higher affinity and more effective. Thus, these cells function 'adaptively' to 'acquire' effective defence mechanism in the lifetime of the host [77, 78]. Adaptive immunological memory can last for decades and in some cases, for the entire lifetime of the host. Recent evidence suggests that the innate immune cells can also learn to mount improved response through earlier exposures [17, 79]. This innate learning can have a broad generic effect on immunological processes to deal with microbes that they encounter up to few months afterwards. In recent times, growing evidence suggests that various forms of stress might be one of the underlying players in deciding how the immune system will react to different exposures and distress signals, and how the inter-cellular interactions among the immune cells will evolve with time [26, 80]. When a child is about to be born, the immune cells present in its umbilical cord is not yet exposed to environmental triggers, pathogens and other foreign substances. In addition, during development, the host body produces many substances to which the immune system of an unborn child is yet to react. However, in the case of adults, the immune system has already undergone various changes triggered by genetic, developmental, environmental and pathogenic factors [25]. As the immune system is a complex network of all these innate and adaptive components, the alteration that happens during the development will carry signatures of how different cell types of this network evolve. The triggers throughout one's life will change the inherent non-equilibrium dynamical interplay between its nodes, i.e., the immune cells. With every factor that elicits immune reaction, the functional relation between different cell types may be modified. Hence, during development, correlations can emerge (or be altered) within the immune network. For healthy individuals, the number of total immune cells remains in a non-equilibrium steady state (homeostasis). In the absence of a bio-physical non-equilibrium model of the immune system, analyzing the data of relative abundances

14

of various immune cells in the cord blood as well as the adult peripheral blood can provide us with a partial understanding of the long-time-scale adaptive dynamics of the immune network.

1.4 Existence of Multiple Time-scales in Adaptive Dynamics

Whereas we use the term "adaptation" to indicate alteration in the response of a biological system due to repetition of external stimulation, it is also used in literature to characterize the scenario where changes in the micro-environment drive the state variables of a biological system to respond to the altered external condition temporarily but eventually they return either exactly to their original states (perfect adaptation [81]) or close to the original states (near-perfect adaptation [82]). The existence of such a property in biological systems have been often attributed to the presence of several feedback mechanisms and disparity of time-scales in their state variables [81, 83, 84]. For instance, the internal adaptive dynamics encompassing multiple time-scales can be seen in chemotaxis where the chemotactic cells (such as, prokaryotic bacteria, eukaryotic phagocytes, etc) follow the gradient set up by a chemo-attractant by altering their tumbling (re-orientation) for random walk [81, 85]. In chemotaxis, the time scale of reorientation is greater than the time-scale of the state variables that are responsible for sensing the chemical gradient but lower than those responsible for adaptation to changes in external stimuli [85]. Often, the biological systems respond to fold changes in external stimulation, where the degree of biological response proportionately varies with the relative change in stimulation with respect to the intensity of a prior stimulation. This phenomenon is commonly known as Weber's law [86–88]. There can be systems as simple as autocatalytic reactions having positive feedback schemes with time-scale differences in their state variables that obey Weber's law while adapting to external stimulation [89].

Changes in external situations operate as a perturbation to the biological systems and lead to different strategies for taking important decision at the level of a single cell (such as, proliferation of immune cells following antigen recognition [23]) to coordinated function at the level of an organ (such as, release of insulin by the pancreatic beta cells [90]). At the level of single cell, the external information is transcribed to the nucleus for decision making by the intra-cellular signal transduction pathways. The state variables which are operational in information processing can have different time-scale which may lead to non-trivial consequences. While adjusting to the alterations in external environment, different state variables can take different times to adapt to changes. Because of the complex interactions and multiple time-scales of operation, the resultant dynamics of the system can potentially induce a rich repertoire of temporal patterns of the state variables. In the context of coordinated strategy or synchronized function in biological systems, the dynamics of the individual entities (nodes) need to be properly coordinated. Generally, the perturbative effects are manifest in complex adaptive networks with a longer time scale than that in the dynamics of individual entities (nodes) [32, 91]. It can appear as an adaptive slow variation in the network topology as compared to faster dynamics in the constituent nodes [92] to maintain homeostasis (such as, in immune system) or reach synchronization [93]. At a finer scale, adaptive strategies to reinforce coupling between individual nodes can produce complex emergent phenomenon [94–97]. In the case of neural networks, the plasticity of the neurons within a constrained neighborhood is the basis of memory and learning in the brain. The constraint of a limited resource for a node to establish connections with the other units of the network in the presence of adaptive coupling between the nodes can lead to a complex network with meso- and macro-scale organization as observed, for instance, in the nervous system [94]. One can speculate that such a scenario is even possible in the case of immune networks where adaptive strategies are present in individual cells. In addition to this, overall adaptive changes in the immune network are expected to occur over a much longer developmental time-scale.

1.5 Aim and Overview of the Thesis

Complex networks of interacting components are seen across a broad range of scales in biology – from molecules and cells to populations and ecosystems. Two examples of such systems are the intra-cellular signaling network (comprising molecules such as kinase proteins, second messengers, cell-surface receptors, etc.) and the immune system (comprising cells involved in coordinating the immunological response of the body such as T-cells, B-cells, etc.). The aim of this thesis is to obtain a clearer understanding of the non-linear dynamics or temporal evolution of such systems using modelling and data-analysis techniques. As both of the biological networks mentioned above are involved in functions that are vital for the continued survival of an organism, explicating the dynamical processes underlying their behavior may play a crucial role in developing better and/or more effective treatment of diseases involving disruptions of the normal function of these networks.

Intra-cellular signaling networks coordinate the entire range of biological processes that are necessary for the cell to provide appropriate response, even in the presence of a high degree of noise, to a wide variety of environmental signals [29, 31]. As breakdown in communication between different members of this network lead to disease, it is important to understand the underlying processes which allow robust information processing in the system. A ubiquitous component (motif) of the signaling network is the Mitogen-Activated Protein Kinase (MAPK) cascade, comprising the three kinase proteins, viz., MAPK, MAPK-Kinase or MAP2K, and MAP2K-Kinase or MAP3K [98,99]. This signaling pathway is seen to exist in all eukaryotic cells and is involved in many critical cellular functions including cell cycle control, stress response, differentiation and growth [98,99]. Its crucial importance for the proper functioning of an organism is underscored by the fact that it is affected in many diseases including cancer, auto-immune and neuro-degenerative disorders and it is, therefore, an important drug target [100]. The basic linear cascade dynamics involves regulation of the activity of the MAP3K protein by an upstream signal.

On being activated, MAP3K acts as the enzyme for activating the MAP2K protein, which in turn controls the activation of the MAPK protein. The activated MAPK is known to be involved in many functions, such as the initiation of transcription of genes or stimulating the activity of other kinases [98, 99].

While the behavior of the cascade has been investigated through modelling studies earlier, these have mostly focused on its asymptotic behavior when subjected to constant stimulation. From this 'asymptotic' perspective, the linear chain structure of the reaction cascade implies a rigid relation between stimulus and response, belying the diversity of behavior that such a system can be capable of in a framework that focuses on its time-varying or dynamic properties. Thus, in contrast to most earlier studies we focus on modelling the dynamics of the cascade responding to temporal variations in the stimulus, such as, an abrupt change in the amplitude of the external signal. As in reality, the cellular micro-environment is in a state of constant flux, investigating the response of the cascade to such time-varying stimulus provides one with a better appreciation of the functioning of the signaling network in vivo. The interplay of diverse time-scales that are involved in the enzyme-substrate complex formation and disassociation, as well as, product formation rates, at different stages of the cascade, suggests that the system may respond with an extremely complex set of behavioral patterns to such time-varying signals. During their adaptation to a stimulus, how the system responds and whether it can process significant biological information downstream is a key aspect that we look at. In particular, we focus on the emergence of long-lived memory, i.e., the influence of past events on the present activity of the cascade. We show how the cascade can retain the memory of an already withdrawn stimulus through the persistence of the activity of MAPK protein by generating post-stimulus reverberatory dynamics of the double phosphorylated MAPK. The persistent adaptive response that we observe is indicative of a form of cellular memory and suggests that this ubiquitous signaling pathway plays an even more central role in information processing by eukaryotic cells. In addition, we consider the dependence of the reverberations on the concentrations of the kinases and phosphatases, as well as on

18

the variations of individual parameters. This allows us to establish the robustness of the results reported here, which is essential for future experimental validation.

In a related problem, we study whether a signaling motif as simple as MAPK cascade can function as a memory module following its exposure to environmental cues and show rudimentary forms of learning such as non-associative learning due to repetitive stimulation. In intra-cellular biological scenario, the MAP3K protein often encounter repetitive stimulation from its upstream signaling components. Thus, it is important to find if the system can show altered response when exposed to subsequent stimulation, and thereby forming a basis of learned response. We investigate the possibility of different types of non-associative learning, such as, adaptation and sensitization. We also observe responses characterized by alternate high and low MAPK activity (alternans), or an attenuated response even for strong stimuli. The ensemble of these responses provides a mechanism by which the intra-cellular signaling network achieves signal integration in the case of application of repetitive pulse-trains. Learning is usually associated with multi-cellular organisms, e.g., those possessing a nervous system. However, our computational study shows that simpler forms of learning may appear in more rudimentary systems - as simple as the canonical MAPK motif at the level of a single eukaryotic cell. Taken in conjunction with previous analogies made between the intra-cellular signaling network and the nervous system, our study provides an intriguing perspective on how this signaling cascade can 'remember' and 'learn' even in the absence of any explicit feedback or cross-talk with other pathways. As the MAPK signaling motif is involved in crucial cellular functions in all eukaryotic cells, the potential consequences of such emergent memory and adaptive response are far-reaching. In addition, the existence of a response threshold of the cascade, an apparent refractory behavior following stimulation with short inter-pulse interval, and an alternans-like response under certain conditions suggest an analogy with excitable media. The non-associative learning that we observe in the MAPK cascade in response to periodic variations of the stimulation strength, can be potentially used to explain a variety of biological phenomena, including, immunological anergy in peripheral

19

blood and regulation of positive and negative selection of immature T cells in thymus.

Moving from the intra-cellular scale to the inter-cellular scale, we focus on learning that is occurring over an entire network of cells comprising the mammalian immune system as the evolving system develops to respond to the plethora of environmental cues it is exposed to over the course of time. Thus, the dynamical evolution of the system as it adapts to foreign antigens and distress signals will result in the relative populations of different cell types in the system to change drastically over time. For example, a newborn child, has a negligible population of the different immune memory cells compared to adult individuals whose immune systems have had considerable exposure to a range of stimuli. In addition, correlations in the populations of different immune cell types may be established over time as a result of the interaction of the immune system with the environment. We have investigated this evolution by analyzing empirical data on populations of different immunological cell types collected from newborns and adults. In principle, through such data-analysis it is possible to uncover how genetic, environmental and developmental factors affect the evolution of this network. It also allows identification of clusters of different cell types of the immune system that are related in terms of whether they evolve in a correlated manner over time. Thus, a systematic analysis of the empirical data using quantitative tools, can help in understanding how exposure to the environment governs the dynamics of the inter-cellular network as the immune system evolves over time.

In the research reported in this thesis we have investigated the dynamical properties of the biological networks discussed above using tools from nonlinear physics and statistical data-analysis. The primary focus is to understand the genesis of memory and learning in these systems that span a range of length-scales.

2

Emergent memory in cell signaling: Persistent adaptive dynamics in cascades can arise from the diversity of relaxation time-scales

2.1 Introduction

Intra-cellular signaling networks are paradigmatic of complex adaptive systems that exhibit a rich repertoire of responses to stimuli [101]. Such networks mediate the response of a cell to a wide variety of extra- and intra-cellular signals primarily through a sequence of enzyme-substrate biochemical reactions [102, 103]. While the complexity of the entire signaling system is daunting [35], it is possible to gain an insight into how it functions by focusing on a key set of frequently occurring motifs. These often take the form of linear signaling cascades, referred to as pathways. One of the best known of these pathways is the mitogen-activated protein kinase (MAPK) cascade that is present in all eukaryotic cells [98, 99]. It is involved in regulating a range of vital cellular functions, in-

21

cluding proliferation and apoptosis [99], stress response [104] and gene expression [105]. This signaling module comprises a sequential arrangement of three protein kinases, viz., MAPK, MAPK kinase (MAP2K) and MAPK kinase kinase (MAP3K). Modular function is initiated when extracellular signals stimulate membrane-bound receptors upstream of the cascade, with the information being relayed to MAP3K by a series of intermediaries. Activated kinases in each layer of the module function as enzymes for phosphorylating (and thereby activating) the kinase in the level immediately downstream, with the subsequent deactivation being mediated by corresponding dephosphorylating enzymes known as phosphatases (PPase). The terminal kinase in this cascade, i.e., MAPK, transmits the signal further downstream by phosphorylating various proteins including transcription regulators [42]. Extensive investigations into the asymptotic dynamical behavior of the cascade have contributed towards an in-depth understanding of several emergent features including ultrasensitivity [106], and oscillations [107, 108] that arise through retrograde propagation of activity [109–111] or explicit feedback [112]. One of the striking features of the cascade is the occurrence of bistability which allows the system to switch between two possible states corresponding to low and high activity [108,113–116]. This provides a post-transcriptional mechanism for obtaining a sustained response from transient signals, i.e., cellular memory [117, 118].

Memory can be understood as long-term alterations in the state of a system in response to environmental changes, which allow the system to retain information about transient signals long after being exposed to them [117]. This can arise in the cell through mechanisms such as auto-regulatory transcriptional positive feedback [119] and nucleosomal modifications [120]. In the context of cell-fate determination, it has been shown that an irreversible biochemical response can be generated from a short-lived stimulus through feedback-based bistability [118]. This corresponds to a permanent alteration of the state of the system, thereby actively maintaining 'memory' of the signal. As bistability has also been observed to arise through multi-site phosphorylation in signaling modules, protein phosphorylation has been suggested as a plausible post-transcriptional mechanism for cel-

lular memory [117, 121, 122]. In particular, there have been extensive investigations of the MAPK cascade as it integrates a large range of signals received by the cell in order to control numerous cellular decisions [123–129]. While majority of these investigations have considered the asymptotic dynamical behavior of the system, one may also observe transitory modulations in the response of the cascade in a changing environment [130, 131]. The latter could encode information about prior stimuli to which the system was exposed, and can be a potential mechanism for imparting a form of "short-term" memory to the signaling cascade.

In this chapter we show that a linear MAPK cascade can indeed exhibit short-term memory through transient modulations in its response to an environmental change. Crucially, this can arise even in the absence of explicit feedback between different layers or crosstalk with other pathways. These modulations can persist long after the initial trigger, lasting for durations that are several orders of magnitude longer than the time-scales associated with phosphorylation-dephosphorylation processes. We demonstrate that this occurs both when a signal begins activating the MAPK cascade, as well as when it is withdrawn. On application of the stimulus, the module exhibits long-lived frequency and amplitude modulations in the activation profile of the constituent kinases. Following the withdrawal of stimulus, activity in the cascade decays over an extremely long time-scale, during which reverberatory dynamics, characterized by large-amplitude spiking in MAP Kinase activity, can be observed. We explain the emergence of such long-lived memory of the withdrawn stimulus in terms of the imbalance of fluxes between different layers of the cascade, which results from the diversity of relaxation time-scales of the cascade components, and the reuse of activated kinases as enzymes when they are released from sequestration. This phenomenon is seen to be robust with respect to variations in the model parameters, including the kinetic rate constants and the molecular concentrations of the constituent kinases and phosphatases. Our results reveal that a biochemical signaling module as simple as the MAPK cascade is capable of exhibiting short-term memory that is manifested as persistent modulations in the adaptive response of the system to

changes in stimuli.

2.2 Model

The dynamics of the three layer MAPK signaling cascade has been simulated using the Huang-Ferrell model [106]. Each of the constituent kinase and phosphatase-mediated enzyme-substrate reactions comprise (i) a reversible step corresponding to the formation of the enzyme-substrate complex and (ii) an irreversible product formation step corresponding to the activation/deactivation of a kinase, as described below.

Table 2.1: Components of the MAPK Cascade

Component	Notation	Symbol
Mitogen-activated Protein Kinase Kinase Kinase	MAP3K	3K
Singly Phosphorylated Mitogen-activated Protein Kinase Kinase Kinase	MAP3K*	3K*
Mitogen-activated Protein Kinase Kinase	MAP2K	2K
Singly Phosphorylated Mitogen-activated Protein Kinase Kinase	MAP2K*	2K*
Doubly Phosphorylated Mitogen-activated Protein Kinase Kinase	MAP2K**	2K**
Mitogen-activated Protein Kinase	MAPK	K
Singly Phosphorylated Mitogen-activated Protein Kinase	MAPK*	K*
Doubly Phosphorylated Mitogen-activated Protein Kinase	MAPK**	K**
MAP3K-Phosphatase	3K PPase	P_{3K}
MAP2K-Phosphatase	2K PPase	P_{2K}
MAPK-Phosphatase	K PPase	P_K

The three layer MAPK cascade comprises the following enzyme-substrate reactions:

$$S + 3K \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} S.3K \xrightarrow{k_2} S + 3K^*$$

$$P_{3K} + 3K^* \underset{kp_{-1}}{\overset{kp_1}{\rightleftharpoons}} 3K^*.P_{3K} \xrightarrow{kp_2} P_{3K} + 3K$$

24

$$3K^* + 2K \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} 3K^*.2K \xrightarrow{k_4} 3K^* + 2K^*$$

$$P_{2K} + 2K^* \underset{kp_{-3}}{\overset{kp_3}{\rightleftarrows}} 2K^*.P_{2K} \xrightarrow{kp_4} P_{2K} + 2K$$

$$3K^* + 2K^* \underset{k_{-5}}{\overset{k_5}{\rightleftarrows}} 3K^*.2K^* \xrightarrow{k_6} 3K^* + 2K^{**}$$

$$P_{2K} + 2K^{**} \underset{kp_{-5}}{\overset{kp_5}{\rightleftarrows}} 2K^{**}.P_{2K} \xrightarrow{kp_6} P_{2K} + 2K^*$$

$$2K^{**} + K \underset{k_{-7}}{\overset{k_7}{\rightleftarrows}} 2K^{**}.K \xrightarrow{k_8} 2K^{**} + K^*$$

$$P_K + K^* \underset{kp_{-7}}{\overset{kp_7}{\rightleftarrows}} K^*.P_K \xrightarrow{kp_8} P_K + K$$

$$2K^{**} + K^* \underset{k_{-9}}{\overset{k_9}{\rightleftarrows}} 2K^{**}.K^* \xrightarrow{k_{10}} 2K^{**} + K^{**}$$

$$P_K + K^{**} \underset{kp_{-9}}{\overset{kp_9}{\rightleftarrows}} K^{**}.P_K \xrightarrow{kp_{10}} P_K + K^*$$

The above enzyme-substrate reactions can be expressed in terms of the following coupled ordinary differential equations (ODEs):

$$\frac{d[3K]}{dt} = k_{-1}.[S.3K] + kp_2.[3K^*.P_{3K}] - k_1.[S].[3K],$$

$$\frac{d[S.3K]}{dt} = k_1.[S].[3K] - (k_{-1} + k_2).[S.3K],$$

$$\frac{d[3K^*.P_{3K}]}{dt} = kp_1.[P^f_{3K}].[3K^*] - (kp_2 + kp_{-1}).[3K^*.P_{3K}],$$

$$\frac{d[3K^*]}{dt} = k_2.[S.3K] + kp_{-1}.[3K^*.P_{3K}] - kp_1.[P^f_{3K}].[3K^*]$$
$$+ (k_{-3} + k_4).[3K^*.2K] - k_3.[3K^*].[2K]$$
$$+ (k_{-5} + k_6).[3K^*.2K^*] - k_5.[3K^*].[2K^*],$$

$$\frac{d[2K]}{dt} = k_{-3}.[3K^*.2K] + kp_4.[2K^*.P_{2K}] - k_3.[3K^*].[2K],$$

$$\frac{d[3K^*.2K]}{dt} = k_3.[3K^*].[2K] - (k_{-3} + k_4).[3K^*.2K],$$

$$\frac{d[2K^*.P_{2K}]}{dt} = kp_3.[P^f_{2K}].[2K^*] - (kp_4 + kp_{-3}).[2K^*.P_{2K}],$$

$$\frac{d[2K^*]}{dt} = k_4.[3K^*.2K] + kp_{-3}.[2K^*.P_{2K}] - kp_3.[P^f_{2K}].[2K^*]$$
$$+ k_{-5}.[3K^*.2K^*] - k_5.[3K^*].[2K^*] + kp_6.[2K^{**}.P_{2K}],$$

$$\frac{d[3K^*.2K^*]}{dt} = k_5.[3K^*].[2K^*] - (k_6 + k_{-5}).[3K^*.2K^*],$$

$$\frac{d[2K^{**}.P_{2K}]}{dt} = kp_5.[P^f_{2K}].[2K^{**}] - (kp_6 + kp_{-5}).[2K^{**}.P_{2K}],$$

$$\frac{d[2K^{**}]}{dt} = k_6.[3K^*.2K^*] + kp_{-5}.[2K^{**}.P_{2K}] - kp_5.[P^f_{2K}].[2K^{**}]$$
$$+ (k_{-7} + k_8).[2K^{**}.K] - k_7.[2K^{**}].[K]$$
$$+ (k_{-9} + k_{10}).[2K^{**}.K^*] - k_9.[2K^{**}].[K^*],$$

$$\frac{d[K]}{dt} = k_{-7}.[2K^{**}.K] + kp_8.[K^*.P_K] - k_7.[2K^{**}].[K],$$

$$\frac{d[2K^{**}.K]}{dt} = k_7.[2K^{**}].[K] - (k_8 + k_{-7}).[2K^{**}.K],$$

$$\frac{d[K^*.P_K]}{dt} = kp_7.[P_K^f].[K^*] - (kp_{-7} + kp_8).[K^*.P_K],$$

$$\frac{d[K^*]}{dt} = k_8.[2K^{**}.K] + kp_{-7}.[K^*.P_K] - kp_7.[P_K^f].[K^*]$$

$$+ k_{-9}.[2K^{**}.K^*] - k_9.[2K^{**}].[K^*] + kp_{10}.[K^{**}.P_K],$$

$$\frac{d[2K^{**}.K^*]}{dt} = k_9.[2K^{**}].[K^*] - (k_{-9} + k_{10}).[2K^{**}.K^*],$$

$$\frac{d[K^{**}.P_K]}{dt} = kp_9.[P_K^f].[K^{**}] - (kp_{-9} + kp_{10}).[K^{**}.P_K],$$

$$\frac{d[K^{**}]}{dt} = k_{10}.[2K^{**}.K^*] + kp_{-9}.[K^{**}.P_K] - kp_9.[P_K^f].[K^{**}].$$

where

$$[S] = [S]_{\text{tot}} - [S.3K],$$

$$[P_{3K}^f] = [P_{3K}] - [3K^*.P_{3K}],$$

$$[P_{2K}^f] = [P_{2K}] - [2K^*.P_{2K}] - [2K^{**}.P_{2K}],$$

$$[P_K^f] = [P_K] - [K^*.P_K] - [K^{**}.P_K].$$

2.2.1 Methods

The time-evolution of the molecular concentrations of the different components of the cascade are modeled using a set of coupled ordinary differential equations (see above) that are integrated using the stiff solver `ode15s` implemented in *MATLAB Release 2010b*. As the concentrations of the different molecular species can vary over several orders of magnitudes, low values of relative and absolute tolerances have been used in order to ensure the accuracy of the resulting time-series. Note that the quasi-steady-state hypothesis has not been invoked [132]. To ensure that initially all kinases are non-phosphorylated we prepare the initial resting state of the system by simulating it for a long duration ($\sim 10^6$

27

mins) in the absence of any signal. Subsequently MAP3K is exposed to a stimulus of amplitude S and duration 5000 minutes. On using the base values for the parameter set as given in Ref. [106], the system exhibits ultrasensitivity (as reported earlier) which provides verification of the correct numerical implementation of the model (see Appendix). Following the removal of the stimulus, we continue to simulate the system until it returns to the resting state or the simulation duration exceeds 10^4 minutes.

We have analyzed the long-lived reverberatory activity of the cascade after the removal of the stimulus by using the following measures:

The primary recovery time (τ_{PR}). Following the activation of the cascade by introducing a stimulus, the maximum concentration R_{max} of MAPK** is recorded. On removing the stimulus, MAPK activity starts to decay. The time taken for MAPK** to monotonically decrease to half of R_{max} is defined as the primary recovery time (τ_{PR}).

Number of spikes during relaxation (N_r). Following primary recovery, MAPK activity may exhibit a series of spikes, which are defined to be occurring whenever MAPK** concentration exceeds 70% of R_{max}. The number of such spikes that are observed before the cascade reaches its resting state is designated as N_r.

The total duration of reverberatory activity (τ_r). When spiking is observed in MAPK activity following the removal of the applied stimulus, the reverberatory activity duration is defined as the interval between the termination of primary recovery and the final spike event, i.e., $\tau_r = t_{final} - \tau_{PR}$. The time of the ith spike t_i is defined as the instant when MAPK activity reaches maximum during that particular event. For $\tau_{PR} > 6000$ mins, the total duration of the reverberatory activity may not be measured accurately as the total simulation duration does not exceed 10^4 minutes.

The total memory time (τ_m). The total duration of memory activity following removal of the applied stimulus is defined as the sum of the primary recovery time and the total duration of reverberatory activity, i.e., $\tau_m = \tau_{PR} + \tau_r$. Note that when the asymptotic

dynamical behavior of the cascade in presence of the signal is oscillatory, on withdrawing the signal the activity may decay extremely rapidly resulting in $\tau_m \approx 0$.

Relaxation time (τ_x). For the situations where the steady state corresponds to a fixed-point attractor we define a relaxation time τ_x for each constituent of the cascade. This is the time required by its concentration to evolve to the half-way point between the resting state and steady state values.

Robustness Analysis. In order to investigate the robustness of the results reported here with respect to variations in the parameter values, we have performed simulations over an ensemble of cascade models whose parameter sets are obtained by uniform random sampling over a physiologically plausible range (see Appendix). The deviation of such a randomly sampled parameter set from the base values used by Huang & Ferrell [106] is measured by the Total Parameter Variation (TPV) [133]: $TPV = \sum_{i=1}^{n} |\log_{10}(p_i/p_{i,HF})|$, where p_i represents the value of the *i-th* parameter in the given sample and $p_{i,HF}$ is the corresponding base value (*i* denotes any one of the 32 system parameters whose values are varied in this analysis). We have measured different characteristics of reverberatory activity, viz., N_r, τ_m and τ_r, for each realization of the cascade (corresponding to a particular random set of parameter values) and have observed them as a function of the TPV for four different values of the stimulus strength. We observe that qualitatively similar results to those reported here are observed for many different realizations.

2.3   Results

For the results reported in this chapter we consider the Huang-Ferrell model of the MAPK signaling cascade [106], schematically shown in Fig. 2.1 (a). Typically, investigations into the dynamics of this model focus on the asymptotic response to sustained stimulation. For all the parameter sets used in this study, the system exhibits a characteristic sequence of transitions in its asymptotic dynamical state upon increasing the strength of the stimulus
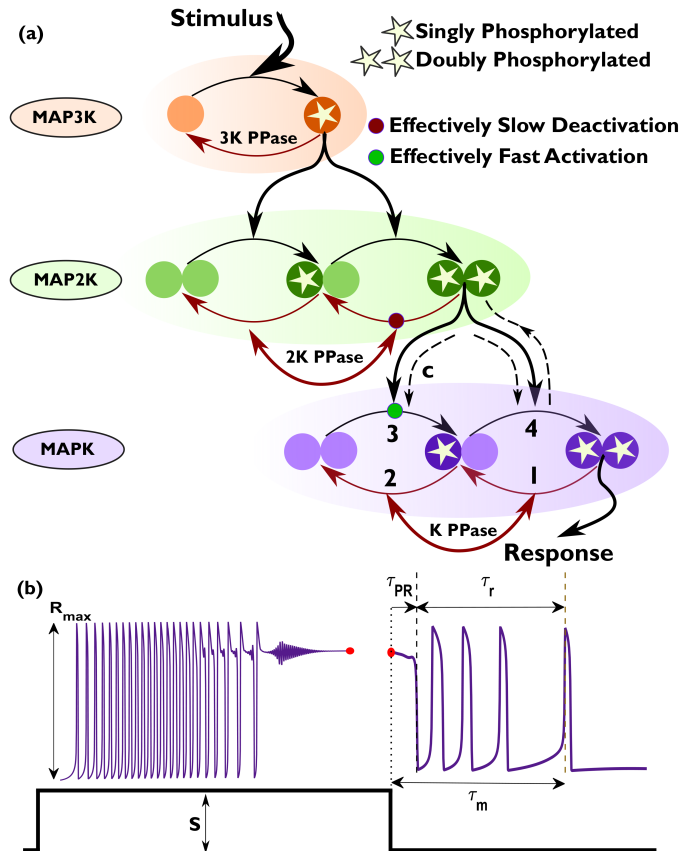
Figure 2.1: (a) Schematic representation of a linear MAPK cascade comprising three layers. Signaling is initiated by a stimulus S activating MAPK kinase kinase (MAP3K). Activation/deactivation of kinases is achieved by adding/removing phosphate groups, which is referred to as phosphorylation/dephosphorylation respectively. The activated MAP3K regulates the phosphorylation of MAPK kinase (MAP2K). Doubly phosphorylated MAP2K, in its turn, controls the activation of MAPK. The response of the cascade to the signal is measured in terms of MAPK activity, viz., the concentration of doubly phosphorylated MAPK. Deactivation of a phosphorylated kinase is regulated by the corresponding phosphatase (indicated by PPase) in the corresponding layer of the cascade. The numbers $1 - 4$ represent the sequence of events that lead to the emergence of a large amplitude spiking response following the withdrawal of the stimulus. The enzyme-substrate protein complex formed during activation of MAPK by doubly phosphorylated MAP2K is indicated by "c". Broken lines have been used to highlight the principal processes that drive the reverberatory dynamics, which functions as a memory of the signal (see chapter for details). (b) Schematic illustrating the emergence of long-lived transient modulations of MAPK activity in response to initiation of a signal of optimal strength S. Withdrawing the stimulus can result in persistent large-amplitude spiking in the response of MAPK, suggestive of a form of "short-term" memory. The maximum response of MAPK to the stimulus is denoted by $R_{\max}$. The primary recovery time (τ_{PR}) is characterized as the duration following withdrawal of stimulus after which MAPK activity decreases to its half-maximum value ($R_{\max}/2$) for the first time. The duration over which reverberatory dynamics occurs is indicated by τ_r, while the total duration for which memory of the withdrawn stimulus persists is $\tau_m = \tau_{PR} + \tau_r$.

30

(see Appendix). For the parameter sets used in our study, there is a lower critical value of the stimulus strength that demarcates a steady state regime characterized by low levels of MAPK activation from a large-amplitude oscillatory regime. Beyond an upper critical value of the stimulus strength, the oscillatory regime gives way to another steady state regime marked by high levels of MAPK activation. In contrast to such asymptotic dynamics, here we report on the transient activity of the system responding to a change in the stimulus. Specifically, we describe the response immediately following the introduction of a signal of amplitude S and that following its removal.

Emergence of persistent modulations in kinase activity. Our results reveal that the transient dynamics can be unexpectedly long-lived, lasting for durations that are much longer compared to the time-scales associated with the phosphorylation and dephosphorylation processes in the cascade (Fig. 2.1, b). While most of the detailed results reported here were obtained using a set of parameter values that differ only marginally from the base values used by Huang & Ferrell [106], qualitatively similar behavior can be observed for many other parameter sets drawn from a physiologically plausible range, as explicitly shown in Figs. 2.2 and 2.3 (see Methods and Appendix for a discussion on robustness).

We first report the behavior of a cascade that is initially in the resting state (characterized by the absence of any phosphorylated components) when it is exposed to a signal. The transient activity that immediately follows exhibits several non-trivial features such as regular spiking in the activity of MAP2K and MAPK depending on the total concentrations of the kinases (Fig. 2.2, b-e) and the signal strength. For a fixed initial state and signal strength, the spikes can further show modulation in their frequency (Fig. 2.2, c-e) as well as amplitude (Fig. 2.2, b and d). In certain cases, both types of modulation can be observed (Fig. 2.2, d). In the representative time series of MAPK activity shown in Fig. 2.2(a-e), the system dynamics eventually converges to a stable fixed point (Fig. 2.2, a-d) or a stable limit cycle (Fig. 2.2, e). Corresponding phase-space projections are shown in Fig. 2.2 (f-i). The complex modulations seen in many of these figures can arise as a

result of coexisting attractors. For example, in Fig. 2.2 (d) the system state appears to spend a considerable time in the basin of attraction of a limit cycle before approaching a stable fixed point (see Appendix for details). Note that if the cascade components are already phosphorylated to an extent when the stimulus is switched on, the system will reach the asymptotic state much more rapidly thereby reducing the duration of transient activity.

An even more intriguing set of complex modulations is observed in the response of the cascade when the signal is withdrawn any time after the stimulated system has converged to the corresponding asymptotic state (which can be as short as a few minutes). Specifically, on switching off the signal, the cascade exhibits large-amplitude spiking behavior in the MAPK activity before eventually relaxing to the resting state (Fig. 2.3). The duration of the spiking activity and the inter-spike intervals can have a wide variety of time-scales as shown in Fig. 2.3 (a-e). Corresponding phase-space projections of the dynamics are shown in Fig. 2.3 (f-i). This post-stimulus reverberatory activity is seen over a range of stimuli strengths and is indicative of a form of memory that can be achieved without explicit feedback or inter-pathway crosstalk. An essential condition for observing this phenomenon is that prior to withdrawing the applied stimulus, the system should have reached an asymptotic state corresponding to either large-amplitude oscillations or a steady state characterized by high MAPK activity. While the reverberatory activity shown in the different panels of Fig. 2.3 persist over durations ranging from less than an hour to a few hours, even longer periods of reverberation can be obtained depending on system parameters (see Appendix).

Processes underlying long-lived memory and reverberatory dynamics. When the stimulus is withdrawn from the MAPK cascade, the decline in MAP Kinase activity comes about through MAPK** binding to MAPK PPase which dephosphorylates it, resulting in an increased concentration of MAPK* [Step 1, Figs. 2.1(a) and 2.4(a)]. In turn, the phosphatase binds to MAPK* thereby deactivating it to MAPK which results in an

Table 2.2: Signal amplitudes and system parameters for the panels in Figs. 2.2 and 2.3

Parameter	(a), (f)	(b), (g)	(c), (h)	(d), (i)	(e), (j)	Units
$[S]$	1.00	1.00	5.00	1.00	1.20	$10^{-6}\mu M$
$[K]_{tot}$	2.43	4.51	2.66	3.10	3.20	μM
$[2K]_{tot}$	3.96	4.64	2.87	1.63	1.80	μM
$[3K]_{tot}$	1.26	1.48	0.32	1.25	0.24	$10^{-2}\mu M$
$[P_K]$	4.88	2.75	5.88	1.75	0.50	$10^{-1}\mu M$
$[P_{2K}]$	0.30	1.40	0.80	1.30	0.30	$10^{-3}\mu M$
$[P_{3K}]$	1.96	3.72	2.70	1.02	1.00	$10^{-4}\mu M$
k_1	1.40	1.24	1.47	3.67	1.00	$10^{3}(\mu M.\text{min})^{-1}$
k_{-1}	1.84	2.32	4.91	3.68	1.50	10^{2}min^{-1}
k_2	6.28	2.32	4.01	7.26	1.50	10^{2}min^{-1}
kp_1	1.35	2.76	2.98	1.97	1.00	$10^{3}(\mu M.\text{min})^{-1}$
kp_{-1}	4.71	5.60	1.67	6.76	1.50	10^{2}min^{-1}
kp_2	1.94	0.71	0.49	3.14	1.50	10^{2}min^{-1}
k_3	4.52	2.37	2.21	2.71	1.00	$10^{3}(\mu M.\text{min})^{-1}$
k_{-3}	3.07	3.92	6.80	2.82	0.30	10^{2}min^{-1}
k_4	6.46	4.14	2.66	1.53	0.30	10^{2}min^{-1}
kp_3	0.92	0.69	2.61	2.96	1.00	$10^{3}(\mu M.\text{min})^{-1}$
kp_{-3}	5.86	2.90	6.37	1.81	1.50	10^{2}min^{-1}
kp_4	1.97	6.72	5.20	3.59	1.50	10^{2}min^{-1}
k_5	4.44	1.84	4.55	2.83	1.00	$10^{3}(\mu M.\text{min})^{-1}$
k_{-5}	1.23	4.47	2.75	5.79	0.30	10^{2}min^{-1}
k_6	6.12	6.07	1.10	1.97	0.30	10^{2}min^{-1}
kp_5	3.55	2.08	2.26	4.82	1.00	$10^{3}(\mu M.\text{min})^{-1}$
kp_{-5}	1.38	6.76	2.23	7.35	1.50	10^{2}min^{-1}
kp_6	7.09	5.52	3.17	6.82	1.50	10^{2}min^{-1}
k_7	4.41	4.28	2.70	4.69	1.00	$10^{3}(\mu M.\text{min})^{-1}$
k_{-7}	1.70	2.59	5.24	2.35	0.30	10^{2}min^{-1}
k_8	2.06	3.42	1.38	1.27	0.30	10^{2}min^{-1}
kp_7	1.03	2.13	3.25	2.46	1.00	$10^{3}(\mu M.\text{min})^{-1}$
kp_{-7}	2.76	1.28	4.84	1.18	1.50	10^{2}min^{-1}
kp_8	4.39	6.17	7.14	5.00	1.50	10^{2}min^{-1}
k_9	4.56	1.15	2.87	1.66	1.00	$10^{3}(\mu M.\text{min})^{-1}$
k_{-9}	6.30	2.36	2.66	5.56	1.50	10^{2}min^{-1}
k_{10}	5.29	7.01	5.08	5.14	1.50	10^{2}min^{-1}
kp_9	2.72	8.14	1.86	3.39	1.00	$10^{3}(\mu M.\text{min})^{-1}$
kp_{-9}	3.04	7.26	0.49	6.39	1.50	10^{2}min^{-1}
kp_{10}	2.23	3.97	3.57	6.49	1.50	10^{2}min^{-1}
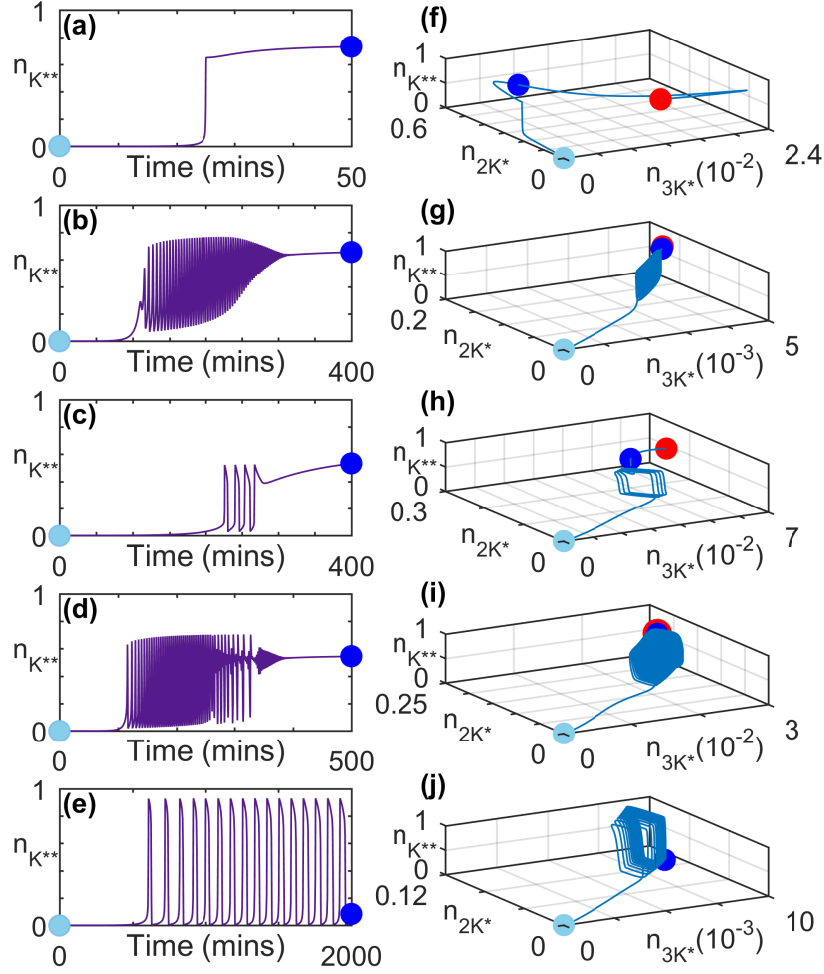
Figure 2.2: Transient activity in MAPK cascade immediately following the application of a stimulus at $t = 0$. (a-e) Characteristic time series for the normalized concentration of doubly phosphorylated MAPK ($n_{K^{**}}$) shown for different total concentrations of kinases. (f-j) Trajectories representing the evolution of the systems in panels (a-e) in the projection of the phase-space on the planes comprising normalized concentrations of active MAP3K (n_{3K^*}), singly phosphorylated MAP2K (n_{2K^*}) and active MAPK ($n_{K^{**}}$). The concentrations have been normalized by the total concentration of MAP3K ($[3K]_{tot}$), MAP2K ($[2K]_{tot}$) and MAPK ($[K]_{tot}$), respectively. The light blue and dark blue markers in each of the panels (f-j) demarcate the portion of the trajectories that correspond to the time series shown in panels (a-e). The steady state of the system is represented by a red marker in panels (f-i). In panels (e) and (j), the system converges to a stable limit cycle. For details of parameter values and signal amplitudes for the systems shown in each of the panels see Table 2.2.

extremely rapid decline in the concentration of MAPK* (Step 2). Concurrently, the deactivation of MAP2K** is delayed, as most of it is bound in the complex MAP2K**.MAPK that has a long time-scale of disassociation. To proceed further we can analyze the con-
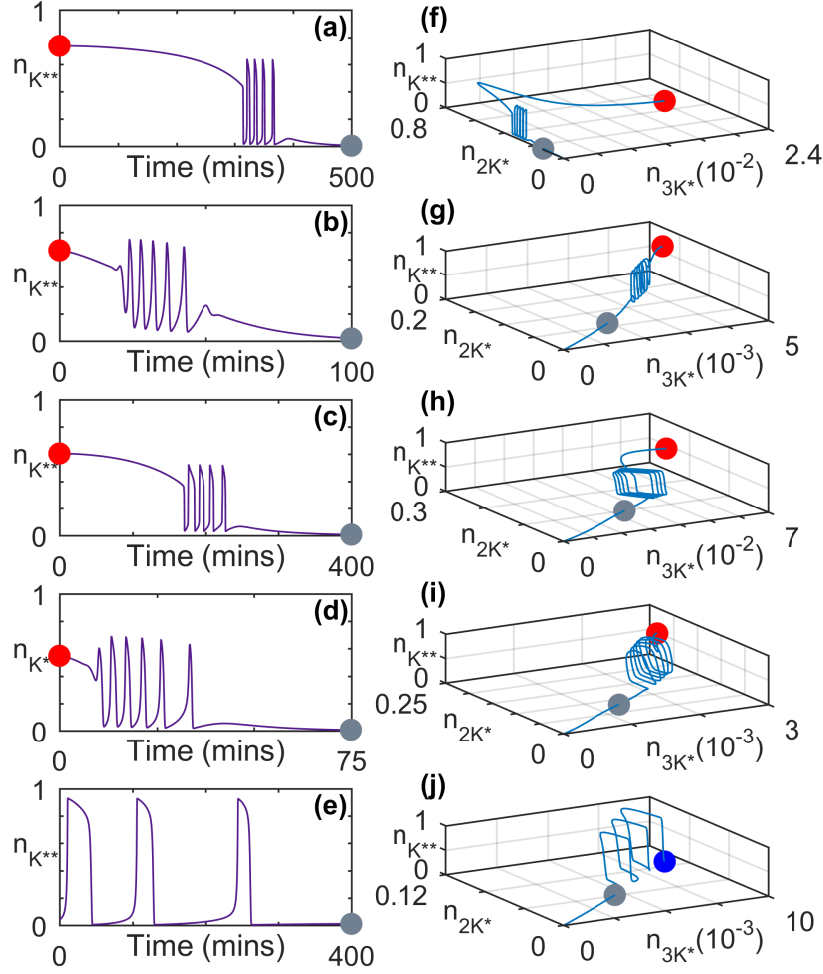
Figure 2.3: Transient activity in MAPK cascade immediately following the withdrawal (at $t = 0$) of an applied stimulus. (a-e) Characteristic time series for the normalized concentration of doubly phosphorylated MAPK ($n_{K^{**}}$) shown for different total concentrations of kinases. (f-j) Trajectories representing the evolution of the systems in panels (a-e) in the projection of the phase-space on the planes comprising normalized concentrations of active MAP3K (n_{3K^*}), singly phosphorylated MAPK (n_{2K^*}) and active MAPK ($n_{K^{**}}$). The concentrations have been normalized by the total concentration of MAP3K ($[3K]_{tot}$), MAP2K ($[2K]_{tot}$) and MAPK ($[K]_{tot}$), respectively. The steady state of the system prior to the withdrawal of the stimulus is represented by a red marker (panels f-i). The system in panels (e) and (j) is seen to relax from a state characterized by stable limit cycle oscillations (represented by the blue marker). In each trajectory shown in (f-j) the grey marker denotes the state of the system corresponding to the final time point in panels (a-e). The concentration of active MAPK is close to its resting state value following the time period shown in (a-e). The parameter values and signal amplitudes for each panel are same as those for the corresponding panels in Fig. 2.2.

stituent processes in terms of the normalized chemical flux N_{Flux} of a molecular species, i.e., its rate of growth expressed relative to the maximum rate of growth of MAPK**. We

35

observe that the suppression of MAP2K** deactivation mentioned above results in its normalized chemical flux exceeding that of MAPK [Fig. 2.4(b)]. Thus, there is a net growth in activity in the MAP Kinase layer as whenever MAP2K** is released from the complex, it is available to phosphorylate MAPK which results in an increase in the concentration of MAPK* (Step 3). The resulting rise in MAPK* manifests as a spike in its concentration [Fig. 2.4(a)], and it subsequently gets phosphorylated again to increase MAPK** concentration even in the absence of any stimulation (Step 4). When the net difference between the normalized flux of MAP2K** and MAPK reaches a maximum, the normalized chemical flux of MAPK** attains its highest value and consequently peak activity of MAP Kinase is observed [Fig. 2.4(c)]. Thus, steps 1-4 represent one complete cycle of MAP Kinase reverberatory activity characterized by an initial decline and a subsequent rise in MAPK** concentration. These steps are subsequently repeated a number of times resulting in a series of spikes in MAPK activity [Fig. 2.4(d)]. The abrupt nature of the rise and fall of MAP Kinase activity that manifests as spikes is a consequence of the bistable nature of the dynamics in the MAPK layer of the cascade [113, 115]. In other words, MAP2K** can be bound either to the corresponding phosphatase (resulting in its subsequent deactivation) or to MAPK/MAPK* (which protects it from deactivation by being inaccessible to its phosphatase). This competition between the phosphatase and the downstream kinase results in a part of the available MAP2K** being sequestered for long times and thus being available for activating MAPK (on being released from the complex) long after the withdrawal of the original stimulus. This results in the post-stimulus repeated spiking activity in the MAPK reported here. We note that similar spiking behavior is also observed in the activity of MAP2K, with the phase of the MAP2K** spikes shifted slightly forward with respect to the corresponding ones in MAPK**, which suggests that they result from retrograde propagation of activity from the MAPK to the MAP2K layer [110]. On the other hand, MAP3K shows a monotonic decline in its activity following the removal of the stimulus.

In order to characterize in detail the memory of prior activity retained by the cascade

which is manifested as long-lived transient reverberations following the withdrawal of stimulus, we use the following measures (see Methods): (i) the primary recovery time (τ_{PR}), (ii) the number of spikes (N_r) that occur during the relaxation process, (iii) the temporal intervals between successive spikes ($t_i - t_{i-1}$, where t_i is the time of occurrence of the ith spike event) and (iv) the total duration of reverberatory activity (τ_r) following primary recovery. The total memory time (τ_m) is the sum of τ_{PR} and τ_r as indicated in Fig. 2.1 (b). In the following we use these measures to present a detailed characterization of the behavior of the cascade components over a range of parameter values (Figs. 2.5-2.7).

MAP Kinase cascade components have different recovery time-scales. As mentioned earlier, the emergence of long-lived reverberatory activity of MAPK following the withdrawal of an applied stimulus can be linked to the flux imbalance of different cascade components which suggests significant differences in their rates of relaxation. As shown in Fig. 2.5 (a), this is indeed the case, even for parameter regimes where no spiking activity of MAPK is observed (i.e., $N_r = 0$). As can be seen, the nature of increase of the relaxation time with increasing total concentrations of kinase protein MAP2K is distinct for the different molecular species and also depends on the state of their phosphorylation. In the lower layers of the cascade, we also find a crossover between two regimes seen at lower and higher values of $[2K]_{tot}$ respectively. These regimes are characterized by relatively slow and rapid increases (respectively) in the recovery times with increasing $[2K]_{tot}$, and appear to be related to the steady-state value attained by MAPK activity upon sustained stimulation of the cascade for the corresponding value of $[2K]_{tot}$ [Fig. 2.5 (b)]. The crossover between the two regimes is seen to occur for a value of $[2K]_{tot}$ for which $\sim 17\%$ of MAPK is activated for the parameter values used in Fig. 2.5 (b).

The distinct regimes are also observed in the dependence of the primary recovery time τ_{PR} on $[2K]_{tot}$ [Fig. 2.5(c)]. As can be observed, the difference between the regimes becomes more pronounced with an increase in the total concentration of MAP3K. An important
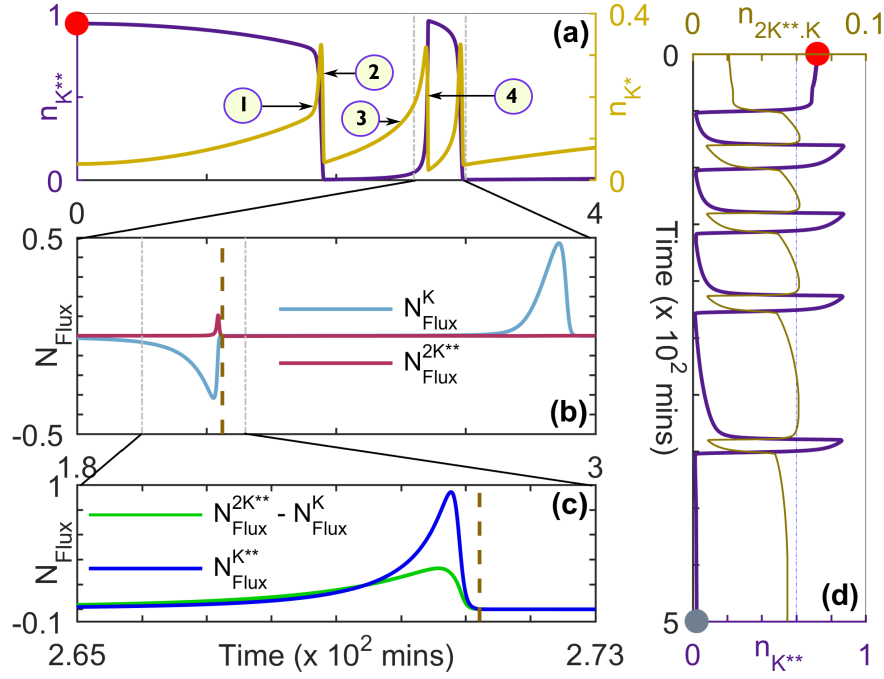
Figure 2.4: Processes underlying emergent memory and reverberatory dynamics in the MAPK cascade. (a) A characteristic time-series for the normalized concentrations of singly and doubly phosphorylated MAPK (n_{K^*} and $n_{K^{**}}$, respectively) following the removal of an applied stimulus of amplitude $S = 2.0 \times 10^{-6} \mu M$ at $t = 0$. The numbers $(1 - 4)$ represent the sequence of events that lead to the emergence of the post-stimulus large-amplitude spiking activity shown schematically in Fig. 2.1 (b). (b) Normalized chemical flux N_{Flux} of MAPK and MAP2K** shown for the segment of the time-series where the spiking behavior in $n_{K^{**}}$ is observed following the withdrawal of the stimulus to MAP3K [demarcated by broken vertical lines in (a)]. (c) Normalized chemical flux N_{Flux} of MAPK** shown along with the difference between the normalized fluxes of MAP2K** and MAPK for the duration indicated by broken vertical lines in (b) corresponding to the peak in the spiking activity of MAPK**. For both panels (b) and (c), normalization of flux is with respect to the maximum of the flux for MAPK**. (d) Characteristic time-series for the reverberatory activity of MAPK following the withdrawal of a stimulus of amplitude $S = 1.2 \times 10^{-6} \mu M$ at $t = 0$, showing the normalized concentration of MAPK** ($n_{K^{**}}$) along with that of the protein complex MAP2K**.MAPK ($n_{2K^{**}.K} = [\text{MAP2K}^{**}.\text{MAPK}]/[2K]_{tot}$). The reference line shows that the peak normalized concentration of the protein complex eventually decreases over time. For details of parameter values see Appendix . The steady state of the system prior to the withdrawal of the stimulus is represented by a red marker [panels (a) and (d)] while the grey marker in (d) corresponds the final time point in Fig. 2.3(d).
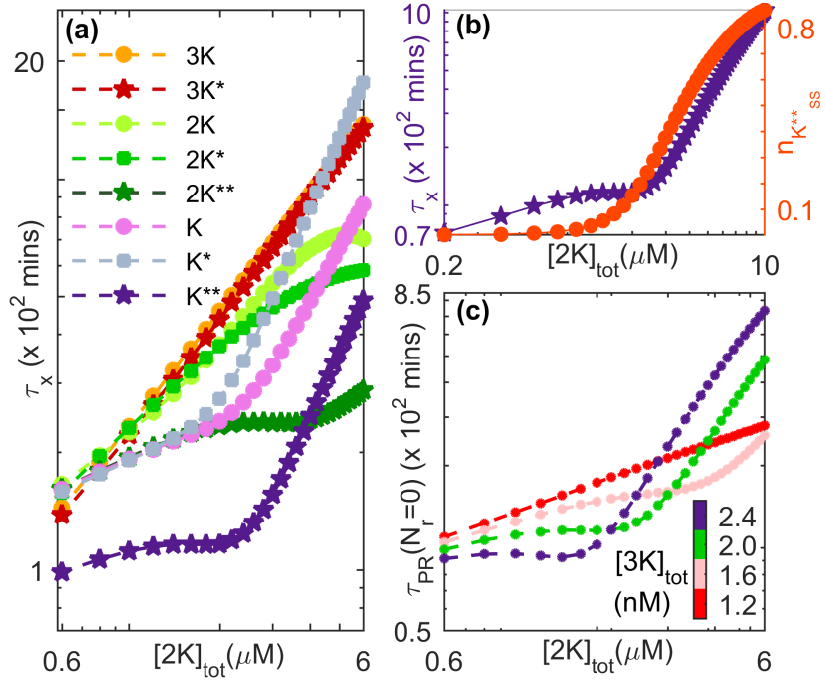
Figure 2.5: Components of the MAPK cascade exhibit relaxation behavior occurring over a broad range of time-scales. Decay of activity is shown after withdrawing an applied stimulus of amplitude $S = 1.2 \times 10^{-6} \mu M$. (a) The relaxation times τ_x of the different molecular species (non, singly and doubly phosphorylated kinase proteins) in each of the layers of the cascade vary with the total concentration of MAP2K. The nature of this dependence is distinct for lower and higher values of $[2K]_{tot}$, which is most prominently observed in the lower layers of the cascade. (b) The occurrence of distinct regimes in the relaxation behavior of MAPK** for different $[2K]_{tot}$ is related to the corresponding increase in the steady state value attained by MAPK** concentration under sustained stimulation of the cascade. At a specific value of the steady-state normalized MAPK activity $n_{K^{**}}$, we observe a crossover from the regime characterized by slowly increasing τ_x seen at lower total concentrations of MAP2K to a regime where τ_x increases relatively rapidly for higher $[2K]_{tot}$. (c) The crossover behavior is also observed in the dependence of the closely related measure τ_{PR}, the primary recovery time (see Methods), on $[2K]_{tot}$. The difference between the two regimes become more prominent upon increasing the total concentration of MAP3K ($[3K]_{tot}$). For both panels (a) and (b) $[K]_{tot} = 0.8 \mu M$ and $[3K]_{tot} = 2.0 nM$, while for panel (c), $[K]_{tot} = 0.8 \mu M$. For details of all other parameter values see Appendix .

39

point to note is that for lower values of $[2K]_{tot}$, the recovery time decreases with increasing $[3K]_{tot}$ while the reverse trend is seen for higher values of $[2K]_{tot}$. We have verified that increasing the stimulus amplitude S while keeping the total MAP3K concentration fixed has a similar effect on the relaxation behavior of activated MAPK (see Appendix). As increasing total concentration of MAP2K results in increased steady-state activity of MAPK, we conclude that, in general, higher activity states of MAPK are associated with increasing relaxation time when either the signal or the substrate (MAP3K) is increased. Conversely, for states characterized by much lower MAPK activity, larger values of S or $[3K]_{tot}$ results in reduced relaxation periods.

Dependence of reverberatory activity on total kinase concentrations. Diverse cellular environments are characterized by different total concentrations of the various molecular components of the MAPK cascade. Thus, in order to determine the robustness of spiking and reverberatory activity following the removal of an applied stimulus, it is important to see how they are affected by varying total kinase concentrations. Such a study will also indicate the ease with which these phenomena can be experimentally observed. Fig. 2.6 shows the variation of different measures of reverberatory activity on the total concentrations of MAPK, MAP2K and MAP3K. While there is a complex dependence on these parameters for the exact number of spikes N_r and the duration of the total memory time τ_m, the phenomenon of reverberatory activity following withdrawal of stimulation can be observed over a large range of the parameter space, underlining its robustness.

We also observe that on increasing $[3K]_{tot}$, the response of N_r to variation in $[K]_{tot}$ and $[2K]_{tot}$ becomes relatively homogeneous. Increasing the stimulus amplitude S [compare panels (a,c,e) with (b,d,f) of Fig. 2.6] does not seem to alter the qualitative nature of the variation in N_r and τ_m over the parameter space in general, although we do observe that the domains corresponding to different values of N_r occupy different regions [Fig. 2.6(e and f)]. Note that for low $[3K]_{tot}$, high values of N_r are observed to coexist with low values of τ_m [Fig. 2.6(a,c and b,d)]. While it may appear surprising that these two measures of mem-

40
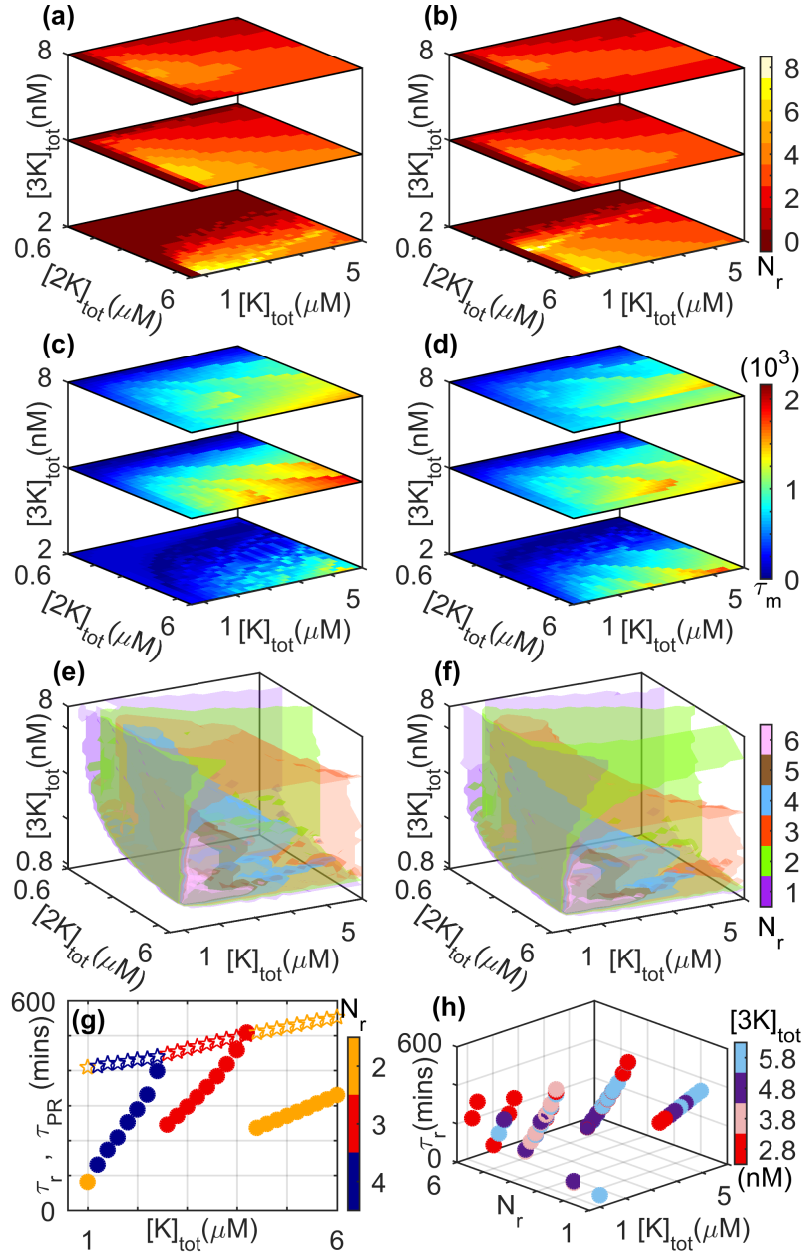
Figure 2.6: Dependence of reverberatory activity on the total kinase concentrations, viz., $[K]_{tot}$, $[2K]_{tot}$ and $[3K]_{tot}$. (a-b) The number of spikes N_r, (c-d) the total memory time τ_m (in minutes) and (e-f) isosurfaces for N_r observed on withdrawing an applied stimulus of amplitude S [$= 0.8 \times 10^{-6} \mu M$ for (a,c,e) and $1.2 \times 10^{-6} \mu M$ for (b,d,f)] are shown as functions of total concentrations of the three kinases. (g) The primary recovery time τ_{PR} (stars) and the total duration of reverberatory activity τ_r (filled circles) are shown for different values of N_r (indicated by the color bar). While τ_{PR} increases monotonically with increasing total MAPK concentration, τ_r shows a more complex dependence ($[2K]_{tot} = 3\mu M$ and $[3K]_{tot} = 4nM$). (h) The dependence of τ_r on $[K]_{tot}$ for different values of N_r has a similar nature for different choices of $[3K]_{tot}$ (indicated by the color bar, $[2K]_{tot} = 3\mu M$). Note that for panel (h), we consider only situations where the system attains a steady state on maintaining stimulation. For details of all other parameter values see Appendix .

41
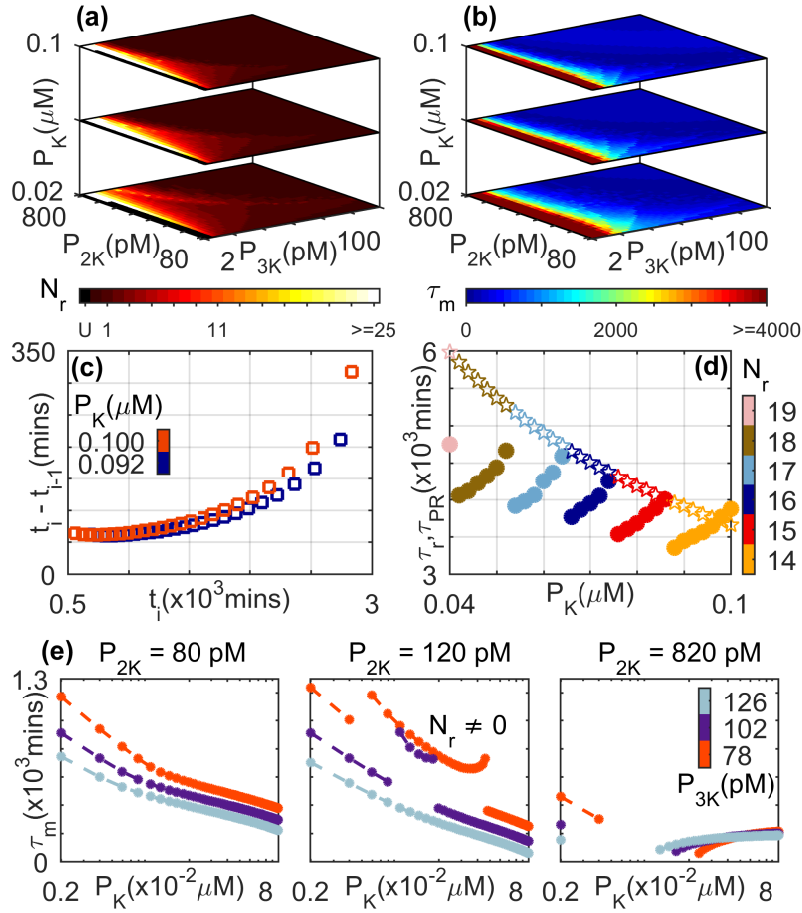
Figure 2.7: Dependence of reverberatory activity on the total concentrations of the phosphatases MAPK PPase ($[P_K]$), MAP2K PPase, ($[P_{2K}]$) and MAP3K PPase ($[P_{3K}]$). (a) The number of spikes N_r and (b) the total memory time τ_m (in minutes) observed on withdrawing an applied stimulus of amplitude $S = 0.8 \times 10^{-6} \mu M$. Situations where the primary recovery time is longer than a maximum or cut-off value (see Methods), such that the duration of the reverberatory dynamics cannot be properly measured, are indicated by the color corresponding to "U". (c) The interval between successive spikes $i - 1$ and i increases with time (t_i being the time of occurrence of the ith spike). As the MAPK PPase concentration is increased, the durations of these intervals are seen to increase. The total concentrations of the other two phosphatases are maintained at $[P_{2K}] = 680 pM$ and $[P_{3K}] = 10 pM$. (d) The variation of primary recovery time τ_{PR} (stars) and the total duration of reverberatory activity τ_r (filled circles) as a function of total MAPK PPase concentration are shown for different values of N_r (indicated by the color bar). While τ_{PR} decreases monotonically with increasing $[P_K]$, τ_r shows a more complex dependence ($[P_{2K}] = 200 pM$ and $[P_{3K}] = 6 pM$). (e) Dependence of the total memory time τ_m on total MAPK PPase concentration ($[P_K]$ shown in log scale) for different total concentrations of MAP2K PPase (values indicated above each of the three panels) and MAP3K PPase (indicated using different colors as shown in the color bar). Note that we consider only situations where the system attains a steady state on maintaining stimulation. For details of all other parameter values see Appendix .

42

ory are not in consonance in this region of parameter space, it can be explained by noting that the stimulated system is in an oscillatory state, and following the removal of the signal these relatively high-frequency oscillations cease after a short duration. Fig. 2.6 (g) suggests that the variation seen in τ_m as a function of the total MAPK concentration for a specific N_r is mostly governed by τ_r, the total duration of reverberatory activity, with the corresponding dependence of τ_{PR} on $[K]_{tot}$ being weak.

As the total MAPK concentration is increased, we observe that while the primary recovery time increases almost linearly, the nature of the reverberatory dynamics as reflected in τ_r shows a more complex dependence on $[K]_{tot}$ [Fig. 2.6 (g)]. If for a given value of $[K]_{tot}$ the MAPK activity following withdrawal of the stimulus shows N_r spikes over a duration of τ_r, then on increasing $[K]_{tot}$ the time-interval between the spikes increases (thereby resulting in an increase of τ_r) until a critical value beyond which the last of the N_r spike no longer appears. Thus, at this point N_r reduces by unity with a concomitant drop in τ_r. This series of events is repeated for steadily decreasing values of N_r as the total MAPK concentration is increased further. Each value of N_r is associated with a characteristic rate of increase in τ_r with $[K]_{tot}$. With a reduction in N_r (as a result of increasing $[K]_{tot}$), this rate is found to decrease as well, which suggests a saturation of the system response. These results are robust with respect to different choices of total MAP3K concentration as can be seen from Fig. 2.6(h), suggesting that similar behavior will be seen for a range of strengths for the applied signal (see Appendix).

Dependence of reverberatory activity on total phosphatase concentrations. We have also investigated the role that phosphatase availability plays on the reverberatory activity of the cascade following the withdrawal of the stimulus. As is the case for total kinase concentrations shown in Fig. 2.6, we see from Fig. 2.7 (a-b) that the number of spikes N_r and the duration of total memory time τ_m depend on the total concentrations of the phosphatases MAPK PPase, MAP2K PPase, and MAP3K PPase. For larger values of the concentrations, viz., $[P_K]$, $[P_{2K}]$ and $[P_{3K}]$, respectively, the system operates in the low-

amplitude response regime. As mentioned earlier, the reverberatory MAPK dynamics during recovery following withdrawal of the applied stimulus will not be seen in this regime. As the phosphatase concentrations are decreased, spiking behavior of MAPK activity is observed with both τ_m and N_r attaining high values in an optimal range. The large variation seen in τ_m [Fig. 2.7 (b)] arises as regions in $[P_{2K}]$-$[P_{3K}]$ parameter space characterized by the same value of N_r are seen to exhibit a range of different values of τ_r and τ_{PR} [Fig. 2.7 (d)]. For reverberatory activity associated with a specific N_r, we observe that the duration τ_r increases with increasing total MAPK PPase concentration. This is a consequence of the intervals between successive spikes $(t_i - t_{i-1})$ increasing with $[P_K]$ as is shown in Fig. 2.7 (c). Note that the results are qualitatively similar for different amplitudes of the applied stimulus (see Appendix). However, increasing $[P_K]$ results also in decreased time for primary recovery τ_{PR} [Fig. 2.7 (d)], which in conjunction with the previously mentioned result leads to non-monotonic dependence of the total memory time τ_m on phosphatase availability. While this non-monotonicity is suggested in Fig. 2.7 (b), it is shown clearly in Fig. 2.7 (e) where the central panel corresponds to situations where spiking behavior is observed in MAPK activity.

Investigation into the dependence of τ_m on P_K [Fig. 2.7(e)] reveals that the range of $[P_K]$ over which reverberatory activity (i.e., $N_r \neq 0$) occurs is demarcated by discontinuities in the functional dependence of τ_m on P_K. For intermediate P_{2K} [Fig. 2.7(e), central panel] where the system attains a steady state on maintaining stimulation, the spiking activity following withdrawal of the stimulus becomes more prominent for low total concentration of MAP3K PPase. For higher P_{2K} [Fig. 2.7(e), right panel] where the system becomes oscillatory over an intermediate range of $[P_K]$, reverberatory activity is observed over a broader range of $[P_{3K}]$. While we have assumed that the same phosphatase acts on both the singly and doubly phosphorylated forms of the kinase in a particular layer of the cascade (as in the canonical Huang-Ferrell model), we have explicitly verified that our results are not sensitively dependent on this.
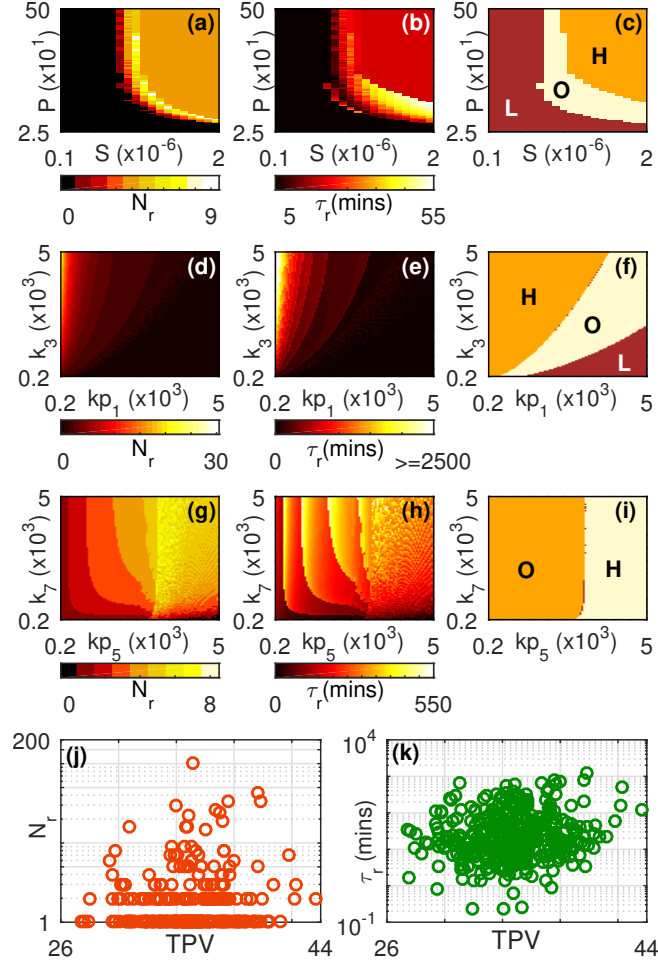
Figure 2.8: Dependence of the post-stimulus reverberatory activity of the MAPK cascade on extrinsic and intrinsic parameters. The corresponding dynamical attractors of the system under sustained stimulation are also shown. (a,d,g) The number of post-stimulus spikes N_r and (b,e,h) the total duration of reverberatory activity τ_r (in minutes) observed on withdrawing the stimulus, as well as, (c,f,i) the corresponding asymptotic dynamical states of the cascade under sustained stimulation, are shown. They are displayed as a function of (a-c) the stimulus strength S and its duration P (in mins), (d-f) the kinetic rates kp_1 and k_3 (both measured in $\mu M.\text{min}^{-1}$) which govern the enzyme-substrate complex formation steps in the dephosphorylation of MAP3K and single phosphorylation of MAP2K (respectively), and (g-i) the kinetic rates kp_5 and k_7 (both measured in $\mu M.\text{min}^{-1}$) which govern the enzyme-substrate complex formation steps in the dephosphorylation of activated MAP2K and single phosphorylation of MAPK (respectively). The strength of the signal used to stimulate the cascade in all cases is $S = 2 \times 10^{-6} \mu M$. For values of all other parameter values see Appendix . (j-k) Robustness of the observed reverberatory activity in MAPK cascade following withdrawal of applied stimulus having strength $S = 10^{-6} \mu M$ with respect to variation in the system parameters. The panels show (j) the number of spikes during relaxation N_r and (k) the total duration of reverberatory activity τ_r, on the Total Parameter Variation (TPV, as described in Methods). The circles in each panel represent an individual realization of the cascade dynamics where each parameter set is chosen by uniform random sampling from a physiologically plausible range (see Appendix for details).

Dependence of reverberatory activity on model parameters. In addition to the effects of kinase and phosphatase concentrations considered above, we have studied how the reverberatory dynamics depends on the different extrinsic and intrinsic model parameters. Fig. 2.8 (a-b) shows how the number of spikes N_r and the total duration of reverberatory activity τ_r depend on the signal strength S and the time interval P for which it is applied prior to withdrawal. From Fig. 2.8 (c) we observe that for large values of S and P the system operates in the high-amplitude response (H) regime. On lowering these stimulation parameters sufficiently a transition to an oscillatory dynamical regime (O) occurs. Further decrease in S and P results in yet another transition, this time to a low-amplitude response (L) regime. Note that, while reduction in the duration of stimulation can be compensated by increasing the signal strength in order to drive the system to the O or H regime starting from L, it appears that there is a critical value of S below which dynamical transitions cannot be achieved even if P is increased indefinitely.

As can be seen, the reverberatory MAPK dynamics following the withdrawal of the stimulus occurs only in the O and H regimes of the $S - P$ parameter space. Intriguingly, the peak of reverberatory activity appears in the region corresponding to the boundary between these two dynamical regimes. This is analogous to *critical slowing down* in the response of physical systems which is manifested in the divergence of relaxation times near a phase transition [134]. However, there does not appear to be a simple relation between the asymptotic dynamical regimes observed on sustained stimulation of the cascade and the nature of the transient reverberatory activity, when we consider the dependence on intrinsic parameters such as the kinetic rate constants that govern the different reaction steps in the cascade. For example, in Fig. 2.8 (d-e) we observe the variation of N_r and τ_r as a function of the rates kp_1 and k_3. As can be seen by comparing with the dynamical regimes in parameter space shown in Fig. 2.8 (f), the peak of reverberatory activity occurs quite far from any of the transition zones. On the other hand, when we consider the variation of N_r and τ_r as a function of the rates kp_5 and k_7 [Fig. 2.8 (g-h)], the peak of N_r does appear to coincide with the boundary between the O and H regimes [Fig. 2.8 (i)], while

46

the τ_r dependence is more complex, showing a number of local maxima. The dependence of the reverberatory activity on several other reaction rates of the system are shown in Appendix.

Apart from looking at the role that individual parameters play in the post-stimulus activity of the cascade, we have also considered how simultaneous variation of all of the parameters affect the reverberatory dynamics. This allows us to investigate whether the phenomena are robust, an essential property if they are to be observed experimentally, as environmental variations, polymorphisms or mutations can often cause multiple parameters of the signaling cascade to be altered. We have verified that the reverberations are not sensitively dependent on system parameters by simulating the dynamics of the cascade using a large number of different parameter sets, each being obtained by randomly sampling the parameter values from their respective physiologically plausible ranges. Fig. 2.8 (j-k) shows that both in terms of the number of spikes N_r, as well as, the total duration of reverberatory activity τ_r, the phenomena of post-stimulus repeated spiking in the activated MAPK concentration we report here is not confined to a very small region of the parameter space but can be seen for a wide variety of choices for the parameter values. Thus, these results establish the robust nature of the emergent "short-term" cellular memory.

2.4   Discussion

In this chapter we have shown that an isolated MAPK signaling module can serve as a fundamental motif in the intra-cellular signaling network for imparting a form of short-term memory to the cell. The emergence of long-lived reverberatory activity reported here arises from the diversity of relaxation time-scales for the different components of the MAP Kinase cascade, which results in flux imbalance between activation of the MAPK layer and deactivation in the MAP2K layer. One may therefore expect to observe results qualitatively similar to what has been reported here whenever the system has disparate

47

time-scales regardless of the actual molecular concentrations and kinetic rates which can vary substantially across different cells [135–137]. Thus, as the MAPK cascade is present in all eukaryotic cells [98, 99], the mechanism for short-term memory in such a signaling cascade that is presented here may hold for such cells in general. As the duration of MAPK** activity is critical for many cellular decisions [138], e.g., the prolonged activation of ERK resulting in its translocation to the nucleus [139], the persistent reverberatory activity seen here may play a non-trivial role in regulation of cellular functions. In addition, it has recently been shown that frequency modulation of ERK activity pulses can encode information controlling cellular proliferation [140], suggesting that the pulsatile nature of the reverberatory MAPK activity shown here may need to be taken into account when considering such processes.

The basal level activity of MAPK in a normal cell is maintained at a low proportion of the total MAPK concentration and serves several biological functions [141]. We observe a crossover between two qualitatively distinct regimes of relaxation behavior of MAPK** occurring at a steady state that is characterized by relatively low proportion of activation of the available MAPK [$\sim$ 17% in Fig. 2.5 (b)]. Thus, there appears to be an effective threshold for MAPK activity (which may be related to its basal state level) that demarcates the different relaxation regimes following the removal of the applied stimulus. A similar crossover is also observed for the primary recovery time τ_{PR}.

One of the most challenging aspects of computational modeling of the dynamics of biological networks is correctly assigning the values of the large number of parameters associated with these models [142]. For the MAPK cascade investigated here, several of the kinetic rates associated with different reactions have never been measured experimentally. There are also a variety of values for the system parameters that have been reported in the literature [107]. For our investigations we have primarily used a reference set of parameter values (see Appendix for details) that differ only marginally from the base values originally used by Huang & Ferrell [106], and yet which allow the cascade to exhibit

three distinct dynamical regimes under different conditions. These are characterized by a time-invariant low-amplitude response (L), oscillatory dynamics (O) and a time-invariant high-amplitude response (H), respectively. We note that the phenomena we observe by using this reference parameter set in the cascade model is robust with respect to variations in the values of the parameters. Indeed, as reported above, the post-stimulus reverberatory activity of the cascade is observed for a large number of different parameter sets that have been randomly sampled from a physiologically plausible range.

It is known that ERK MAPK isoforms (e.g., p42 and p44) are abundantly expressed in non-dividing terminally differentiated neurons [124]. Activation of MAPK by spaced stimulation is known to be responsible for morphological changes in dendrites [123]. Studies also suggest that the activation of the MAPK pathway is linked with associative learning in the mammalian nervous system, synaptic plasticity and neurological memory [123, 124, 129, 143, 144]. An intriguing possibility suggested by the results reported here is that the observed repeated spiking in MAPK activity may function as an effective temporally spaced signal to the nucleus of a neuron. This can then facilitate subsequent changes in the cell required for memory formation [123, 129].

Another well-known example of eukaryotic cellular memory is observed during chemotactic migration along the gradient of a chemical signal [145, 146]. The directionality of migration is known to persist for a certain duration, even if the chemical gradient is altered or becomes static. Studies show that the protein Moesin contributes to the long-lived rigidity of the cytoskeleton assembly that subsequently leads to the directional memory in polarized migrating cells [146]. However, the intra-cellular processes that underlie the persistent activity of Moesin in the absence of a gradient mediated signal are still largely unknown. Evidence suggests that the regulation of Moesin and other ERM proteins are linked with the activity of the MAPK pathway [147, 148]. The long-term reverberatory activity of MAPK following the withdrawal of a stimulus that is reported here may be a possible mechanism underlying such persistent cellular behavior.

To the best of our knowledge the post-stimulus reverberatory activity described here is yet to be reported in the experimental literature. This could possibly be because, typically, during experiments, recording of ERK activity is stopped soon after the withdrawal of the stimulus. Note that in all cases where we observe reverberations, it was preceded by an apparent monotonic decay of MAPK activity immediately following the withdrawal of the stimulus. Experiments would necessarily have to be carried out for much longer durations beyond this initial decay of MAPK activity (i.e., the primary recovery time τ_{PR}) in order to observe the phenomenon reported here. It is of course possible that in the cellular milieu, coupling of the MAPK cascade to other intra-cellular signaling pathways, as well as, the possible presence of explicit feedback connections, may mask the response that is seen here in the case of an isolated MAPK cascade. Also, *in vivo* the cascade will be subject to a variety of signals that will often arrive in close succession. This will make it unlikely to observe extremely long-lived post-stimulus responses lasting over tens or hundreds of minutes. However, we hope that the results reported here will stimulate experiments specifically designed to test the existence of an emergent "short-term" memory in intra-cellular signaling.

To conclude, we have shown the possibility of long-lived reverberatory activity in a signaling cascade following the withdrawal of external stimuli. Our results suggest a mechanism through which the intra-cellular signaling system can encode short-term memory of signals to which the cell was previously exposed. The large-amplitude spiking activity of MAPK following the removal of a prior stimulus may also provide a mechanism for signal integration and learning when the cascade is repeatedly stimulated. We note that there may be additional factors not considered here that may lengthen the persistence of reverberatory activity, including scaffold proteins that increase the lifetime of kinase complexes. Our results suggest that the MAPK cascade potentially has a key role in shaping the information processing capabilities of eukaryotic cells in diverse environments.

2.A   Appendix: System Parameters

The numerical values for the reaction rates used for most of the results reported here (viz., panels (e) & (j) of Fig. 2.2–2.3 and Figs. 2.4–2.7) have been obtained from Jesan *et al.* [111]. A comparison between these values used in our study (MMS) and the base values of the parameters proposed by Huang & Ferrell [106] is shown in Table 2.3. Note that the values for the kinetic rate constants used here differ only marginally from the base values (the deviating values are indicated in red).

The signal amplitudes and system parameters used to generate the representative time-series shown in all panels of Figs. 2.2 and 2.3 are listed in Table 1 of the main text. Note that the kinetic rates for panels (e, j) of Fig. 2.2 and Fig. 2.3 are same as the MMS values listed in Table 2.3.

In order to investigate the robustness of the results reported in the chapter with respect to variations in the parameter values, we have performed simulations over an ensemble of cascade models whose parameter sets are obtained by uniform random sampling over a physiologically plausible range given in Table 2.4. The random values thus obtained are further subject to the constraint that the resulting Michaelis-Menten constant K (defined as $K = (k_r + k_{cat})/k_f$, where k_f, k_r and k_{cat} are the forward, reverse and catalytic rate constants, respectively) should not exceed 1500 nM as per Huang & Ferrell (1996) [cited in main text].

Table 2.3: Reaction Rates

Rate constant	HF	MMS	Units
k_1	1000	1002	$(\mu M.\text{min})^{-1}$
k_{-1}	150	150	min^{-1}
k_2	150	150	min^{-1}
kp_1	1000	1002	$(\mu M.\text{min})^{-1}$
kp_{-1}	150	150	min^{-1}
kp_2	150	150	min^{-1}
k_3	1000	1002	$(\mu M.\text{min})^{-1}$
k_{-3}	150	30	min^{-1}
k_4	150	30	min^{-1}
kp_3	1000	1002	$(\mu M.\text{min})^{-1}$
kp_{-3}	150	150	min^{-1}
kp_4	150	150	min^{-1}
k_5	1000	1002	$(\mu M.\text{min})^{-1}$
k_{-5}	150	30	min^{-1}
k_6	150	30	min^{-1}
kp_5	1000	1002	$(\mu M.\text{min})^{-1}$
kp_{-5}	150	150	min^{-1}
kp_6	150	150	min^{-1}
k_7	1000	1002	$(\mu M.\text{min})^{-1}$
k_{-7}	150	30	min^{-1}
k_8	150	30	min^{-1}
kp_7	1000	1002	$(\mu M.\text{min})^{-1}$
kp_{-7}	150	150	min^{-1}
kp_8	150	150	min^{-1}
k_9	1000	1002	$(\mu M.\text{min})^{-1}$
k_{-9}	150	150	min^{-1}
k_{10}	150	150	min^{-1}
kp_9	1000	1002	$(\mu M.\text{min})^{-1}$
kp_{-9}	150	150	min^{-1}
kp_{10}	150	150	min^{-1}

Table 2.4: Biologically plausible range of the parameters used for random sampling

Parameter	Range of values	Units
$[K]_{tot}$	$0.075 - 6$	μM
$[2K]_{tot}$	$0.075 - 6$	μM
$[3K]_{tot}$	$1.875 \times 10^{-4} - 1.5 \times 10^{-2}$	μM
$[P_K]$	$7.5 \times 10^{-3} - 0.6$	μM
$[P_{2K}]$	$1.875 \times 10^{-5} - 1.5 \times 10^{-3}$	μM
$[P_{3K}]$	$1.875 \times 10^{-5} - 1.5 \times 10^{-3}$	μM
k_f	$62.5 - 5000$	$(\mu M.\text{min})^{-1}$
k_r	$9.36 - 750$	min^{-1}
k_{cat}	$9.36 - 750$	min^{-1}

Table 2.5: Total concentration (in μM) of the kinase proteins for Fig. 2.4 (panels a–c) and Fig. 2.7

$[K]_{tot}$	$[2K]_{tot}$	$[3K]_{tot}$
4.8	1.2	0.0030

Table 2.6: Total concentration (in μM) of the phosphatase proteins for Figs. 2.4–2.6 and Figs. 2.14-2.15

$[P_{3K}]$	$[P_{2K}]$	$[P_K]$
1×10^{-4}	3×10^{-4}	0.05

Table 2.7: Total concentration (in μM) of the phosphatase proteins for Figs. 2.14 and 2.15

Panels	$[K]_{tot}$	$[2K]_{tot}$	$[3K]_{tot}$
(a) and (f)	3.0	3.0	0.0080
(b) and (g)	1.0	2.4	0.0024
(c) and (h)	1.2	6.0	0.0028
(d) and (i)	2.0	2.2	0.0024
(e) and (j)	4.8	6.0	0.0014

2.B   Appendix: Supplementary Figures
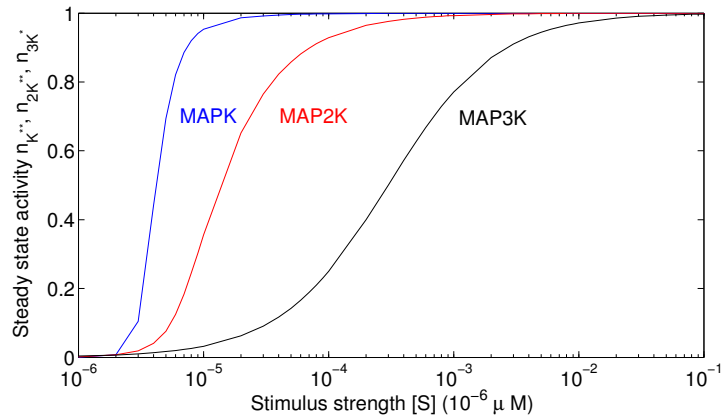
Figure 2.9: Steady-state kinase activity obtained in our numerical implementation of the MAPK cascade dynamics reproducing the results obtained by Huang & Ferrell [106], using their base values for the parameters.
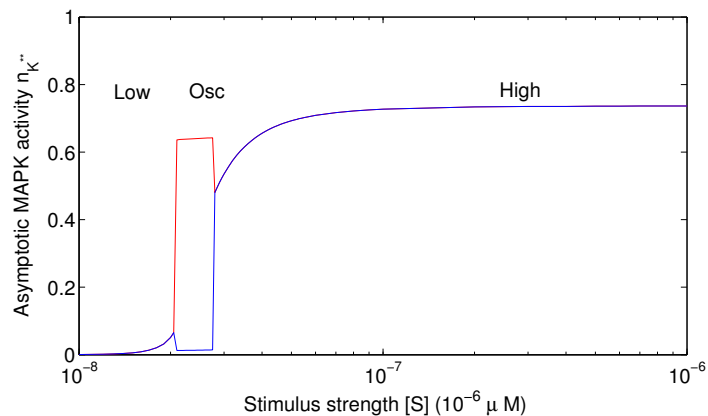
Figure 2.10: Representative bifurcation diagram of the MAPK cascade dynamics showing the asymptotic activity of MAPKinase as a function of the strength of the applied stimulus S. As the strength is increased, the behavior shows successive transitions from a low-response steady state regime ("Low") to large-amplitude oscillations ("Osc") and finally to a high-response steady state ("High") regime. The red and blue curves represent, respectively, the maxima and minima of MAPK activity in the oscillatory regime. The parameter set used is identical to that used for generating the time-series shown in Fig. 2.2 (a) [see Table 1 in the main text]. Qualitatively similar bifurcation diagrams are seen for all other parameter sets that give rise to reverberatory activity.
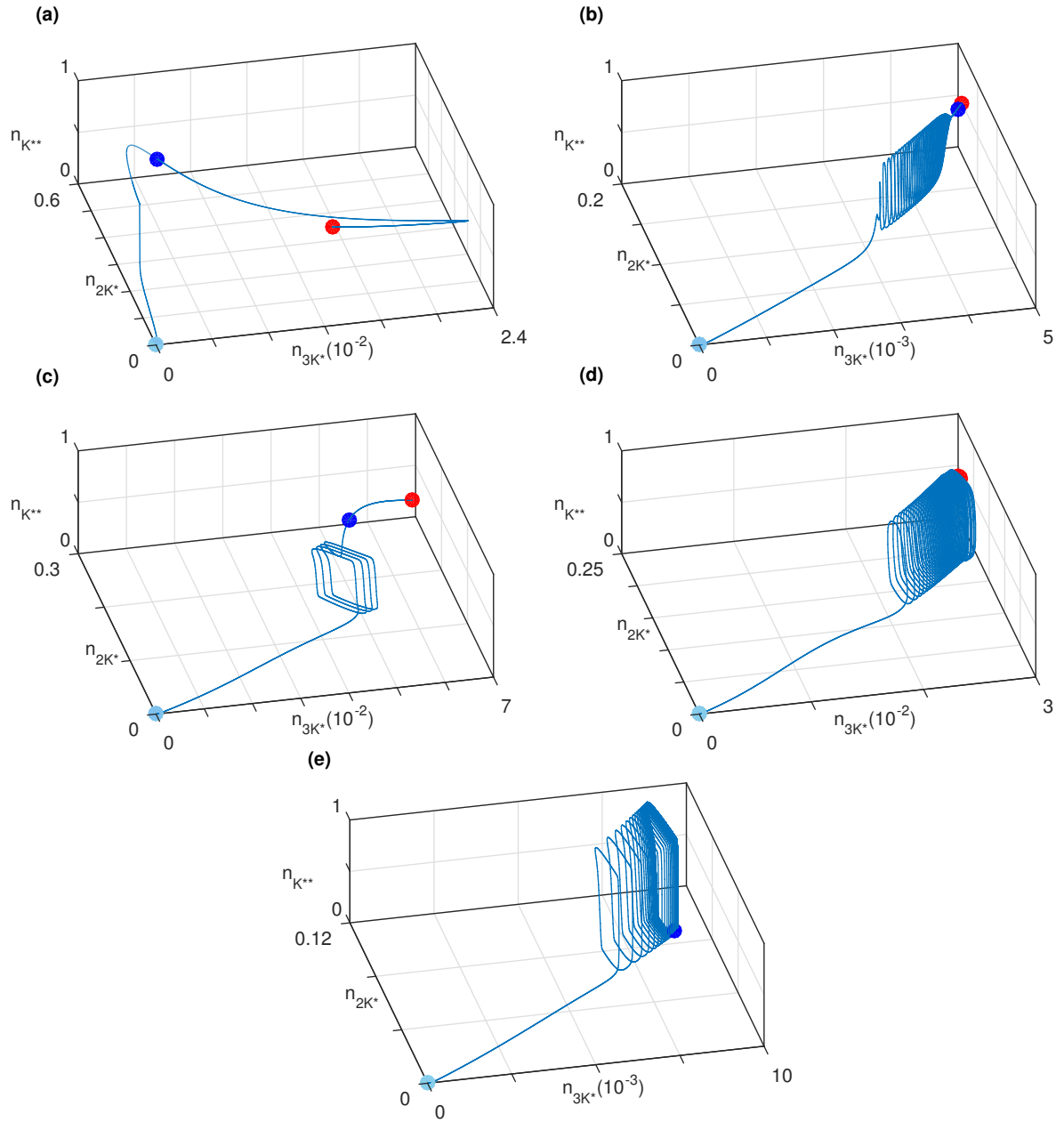
Figure 2.11: Alternative representation of each of the phase-space trajectories shown in Figure 2.2 (f-j) in the main text. The size and orientation of these panels have been adjusted for clarity.
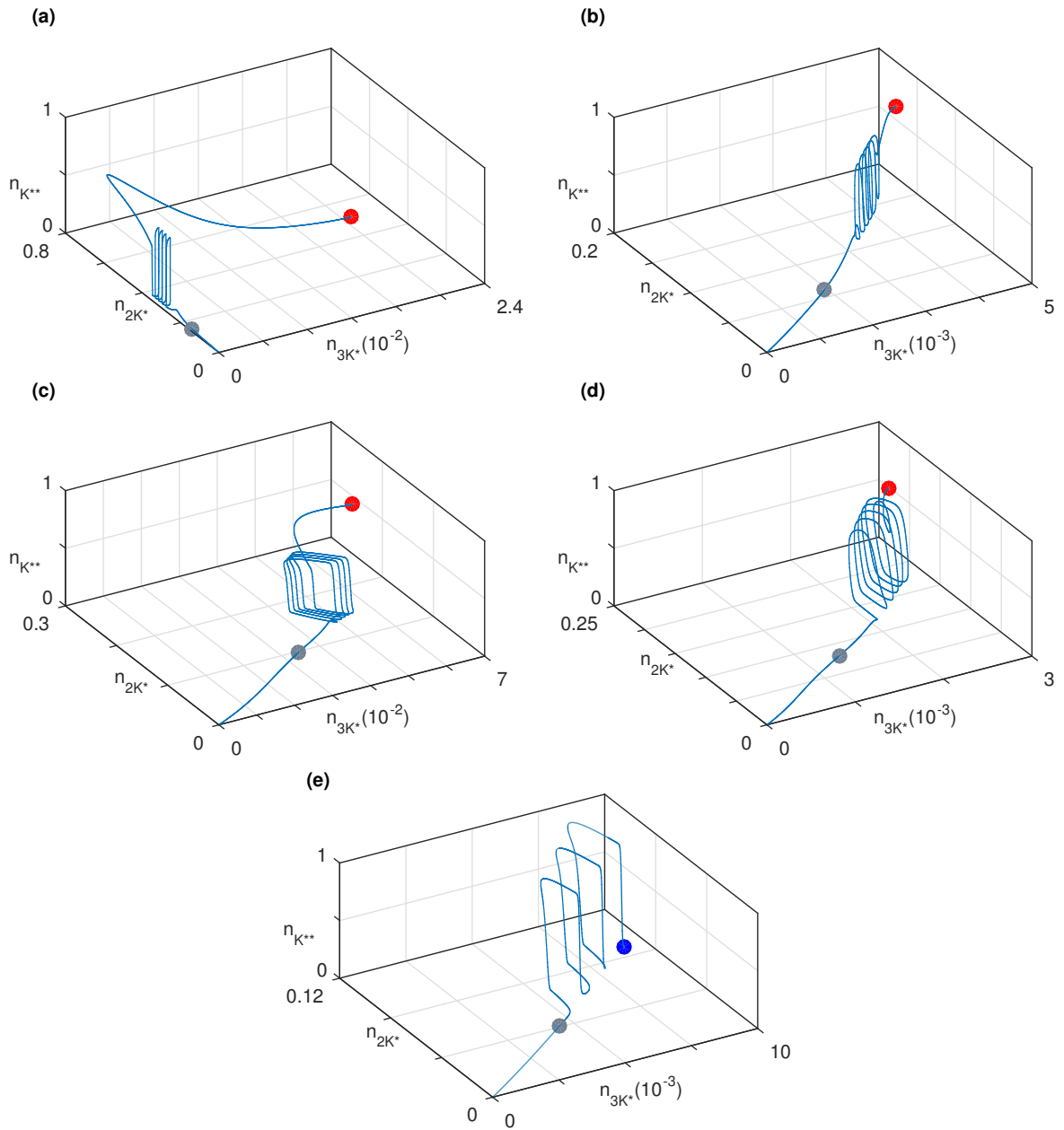
Figure 2.12: Alternative representation of each of the phase-space trajectories shown in Figure 2.3 (f-j) in the main text. The size and orientation of these panels have been adjusted for clarity.

57

Figure 2.13: Magnified views of the phase-space trajectory shown in Fig. 2.2 (i). The blue markers correspond to the final point in the time series displayed in Fig. 2.2 (d), while the red markers indicate the fixed point of the dynamical system in the presence of stimulus. (a) Magnified view of the trajectory beginning from the black marker shown in Fig. 2.2 (i). The pink marker denotes the starting point of the segment of the trajectory displayed in panel (b). (b) Further magnification of a section of the phase-plane trajectory shown in panel (a) corresponding to the duration when the system moves away from the unstable limit cycle and converges to the stable fixed point.
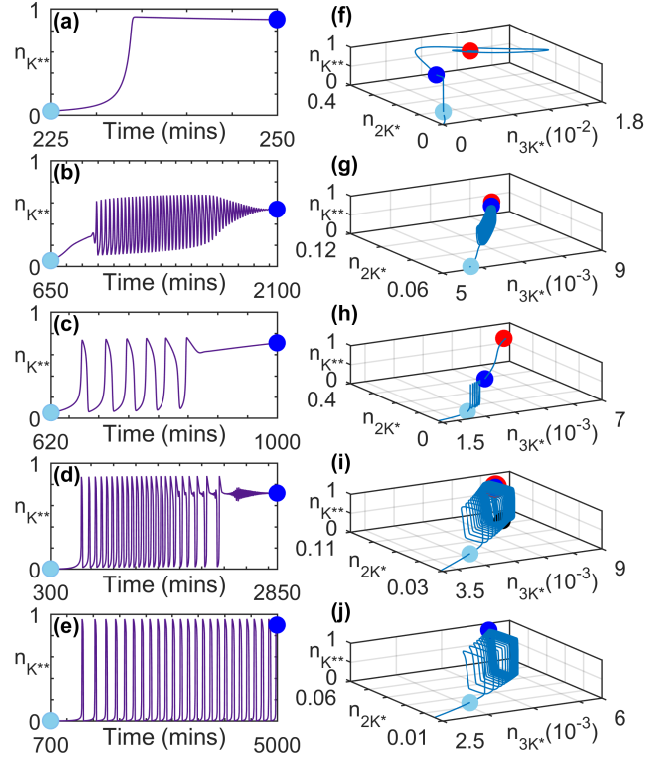
Figure 2.14: Transient activity in MAPK cascade immediately following the application of a stimulus having amplitude $S = 1.2 \times 10^{-6} \mu M$ at $t = 0$. (a-e) Characteristic time series for the normalized concentration of doubly phosphorylated MAPK ($n_{K^{**}}$) for different total concentrations of kinases (see Table 2.7. For all panels, the kinetic rates are identical to the MMS set (see Table 2.3) and phosphatase concentrations are as in Table 2.6. The concentration of active MAPK is insignificant prior to the time periods shown in panels (a-e). (f-j) Trajectories representing the evolution of the systems in panels (a-e) in the projection of the phase-space on the planes comprising normalized concentrations of active MAP3K (n_{3K^*}), singly phosphorylated MAP2K (n_{2K^*}) and active MAPK ($n_{K^{**}}$). The concentrations have been normalized by the total concentration of MAP3K ($[3K]_{tot}$), MAP2K ($[2K]_{tot}$) and MAPK ($[K]_{tot}$), respectively. The light blue and dark blue markers in each of the panels (f-j) demarcate the portion of the trajectories that correspond to the time series shown in panels (a-e). The steady state of the system is represented by a red marker in panels (f-i). In panels (e) and (j), the system converges to a stable limit cycle.

59

Figure 2.15: Transient activity in MAPK cascade immediately following the withdrawal (at $t = 0$) of an applied stimulus having amplitude $S = 1.2 \times 10^{-6} \mu M$. (a-e) Characteristic time series for the normalized concentration of doubly phosphorylated MAPK ($n_{K^{**}}$) shown for different total concentrations of kinases (see Table 2.7). (f-j) Trajectories representing the evolution of the systems in panels (a-e) in the projection of the phase-space on the planes comprising normalized concentrations of active MAP3K (n_{3K^*}), singly phosphorylated MAPK (n_{2K^*}) and active MAPK ($n_{K^{**}}$). The concentrations have been normalized by the total concentration of MAP3K ($[3K]_{tot}$), MAP2K ($[2K]_{tot}$) and MAPK ($[K]_{tot}$), respectively. The steady state of the system prior to the withdrawal of the stimulus is represented by a red marker (panels f-i). The system in panels (e) and (j) is seen to relax from a state characterized by stable limit cycle oscillations (represented by the blue marker). In each trajectory shown in (f-j) the grey marker denotes the state of the system corresponding to the final time point in panels (a-e). The concentration of active MAPK is close to its resting state value following the time period shown in (a-e). The parameter values for each panel are same as those for the corresponding panels in Fig. 2.14. Note that panel (d) is obtained using the same parameter values as Fig. 2.4 (d) of the main text.
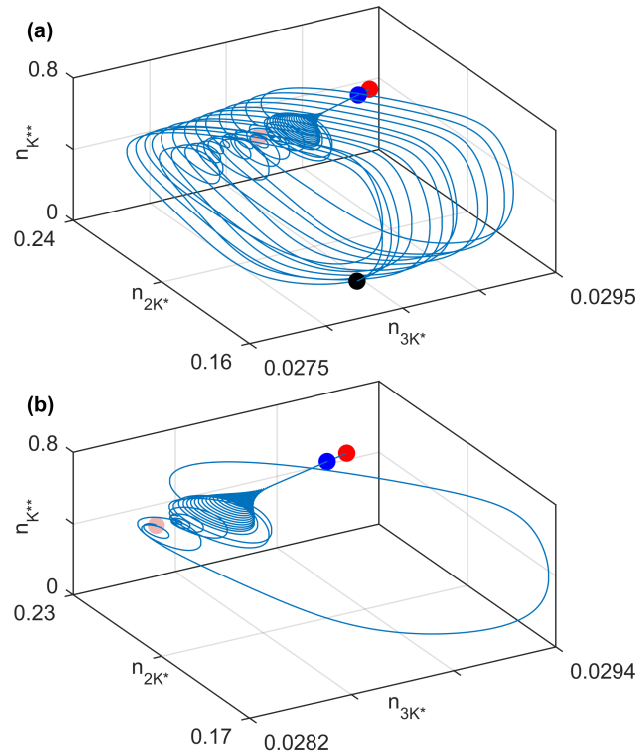
60
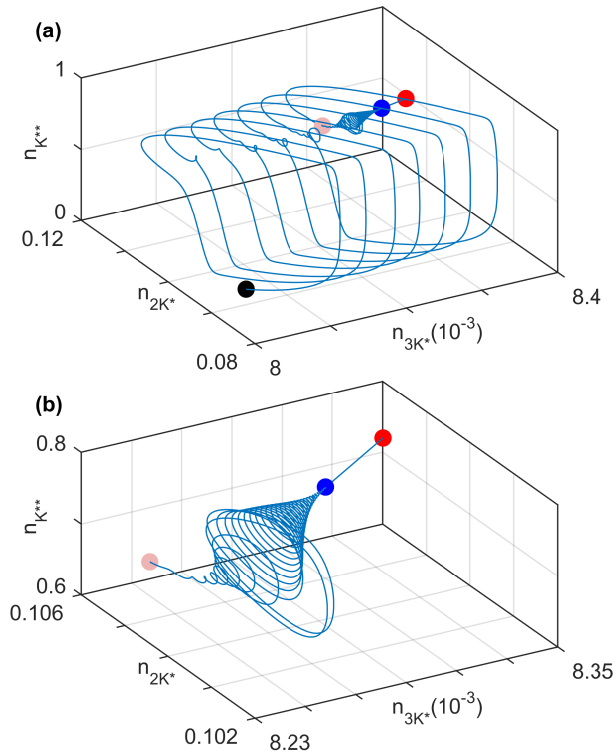
Figure 2.16: Magnified views of the phase-space trajectory shown in Fig. 2.14 (i). The blue markers correspond to the final point in the time series displayed in Fig. 2.14 (d), while the red markers indicate the fixed point of the dynamical system in the presence of stimulus. (a) Magnified view of the trajectory beginning from the black marker shown in Fig. 2.14 (i). The pink marker denotes the starting point of the segment of the trajectory displayed in panel (b). (b) Further magnification of a section of the phase-plane trajectory shown in panel (a) corresponding to the duration when the system moves away from the unstable limit cycle and converges to the stable fixed point.
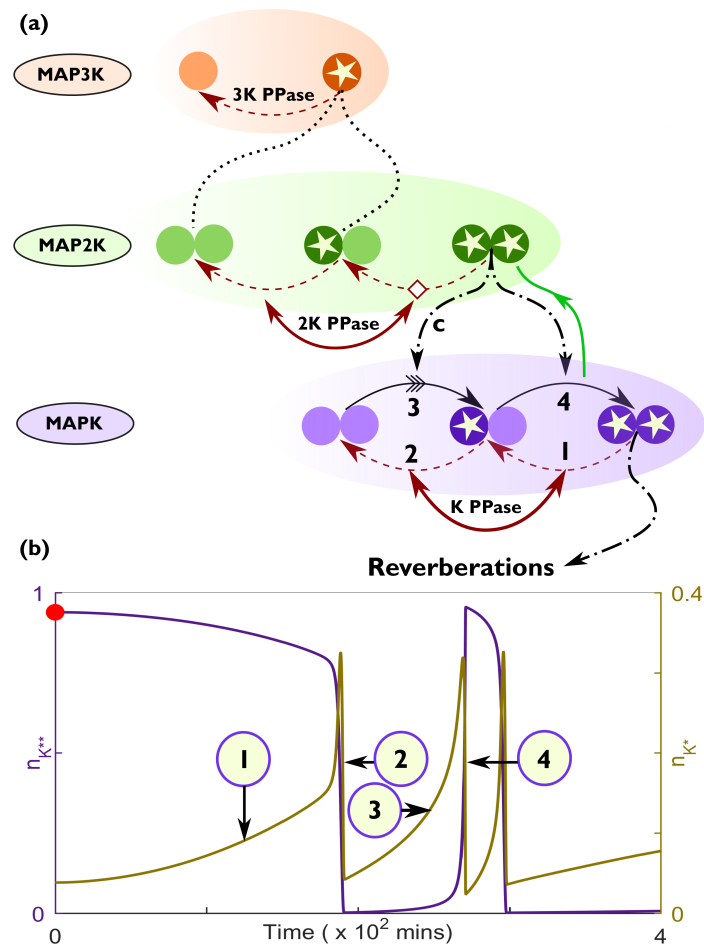
Figure 2.17: Detailed processes underlying long-lived memory and reverberatory dynamics. (a) Schematic representation of MAPK cascade showing the processes that occur subsequent to removing a stimulus. The numbers $(1-4)$ represent the sequence of events that lead to the emergence of the post-stimulus large-amplitude spiking activity shown in (b). The enzyme-substrate protein complex formed during activation of MAPK by doubly phosphorylated MAP2K is indicated by "c". The green arrow from the MAPK layer to the MAP2K layer represents the release of doubly phosphorylated MAP2K from downstream complexes. (b) A characteristic time-series for the normalized concentration of singly and doubly phosphorylated MAPK (n_{K^*} and $n_{K^{**}}$, respectively) following the removal of an applied stimulus of amplitude $S = 2.0 \times 10^{-6} \mu M$ at $t = 0$. The numbers $(1-4)$ represent the same events shown in (a). The total concentrations of the kinases and phosphatases used for generating the time-series are provided in Tables 2.5 and 2.6, respectively.
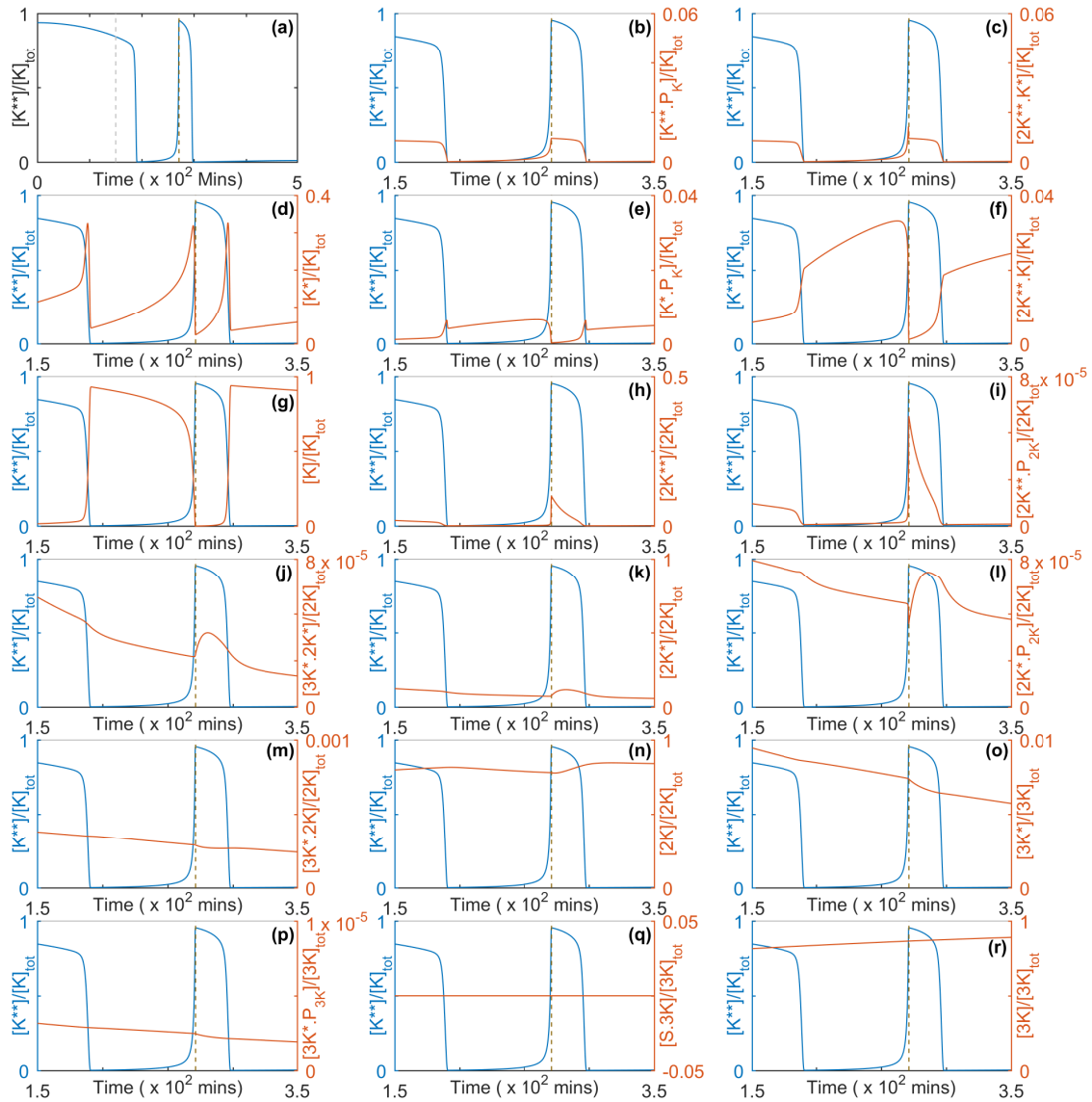
Figure 2.18: Characteristic dynamics of the molecular components of the MAP Kinase cascade following withdrawal of a stimulus. (a) The time-series of the normalized concentration of doubly phosphorylated MAPK ($[K^{**}]/[K]_{tot}$) following removal of an applied stimulus with amplitude $S = 2.0 \times 10^{-6} \mu M$ at $t = 0$. (b-r) Time-series of the normalized concentrations of the different components of the MAPK cascade, shown starting from $t = 150$ minutes after withdrawing the stimulus, displayed together with the time-series of normalized MAPK activity $[K^{**}]/[K]_{tot}$. The total concentrations of the kinases and phosphatases used for generating the figures are provided in Tables 2.5 and 2.6, respectively.

63
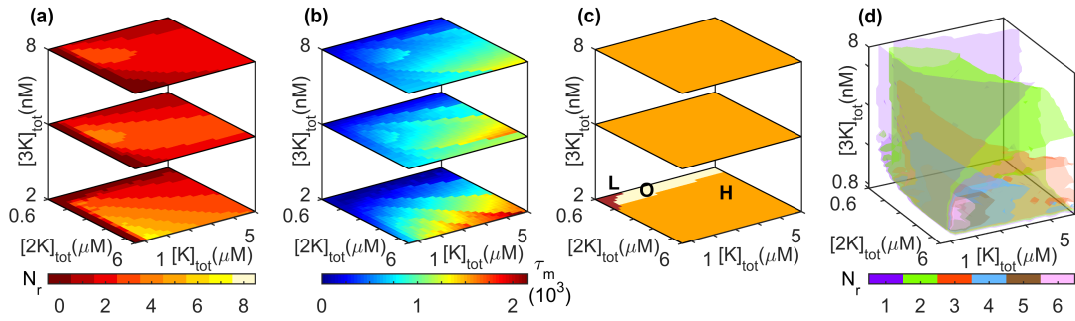
Figure 2.19: Dependence of the reverberatory activity on the total kinase concentrations, viz., MAPK ($[K]_{tot}$), MAP2K ($[2K]_{tot}$) and MAP3K ($[3K]_{tot}$). The corresponding dynamical attractors of the system under sustained stimulation are also shown. (a) The number of post-stimulus spikes N_r, (b) the total memory time τ_m (in minutes), (c) the corresponding asymptotic dynamical states of the cascade under sustained stimulation, and (d) isosurfaces for N_r observed on withdrawing an applied stimulus of amplitude $S = 2.0 \times 10^{-6} \mu M$, are shown as functions of the three total kinase concentrations. The total concentrations of the phosphatases are held fixed for (a-d) and are provided in Table 2.6. The kinetic rates used are given in Table 2.3.
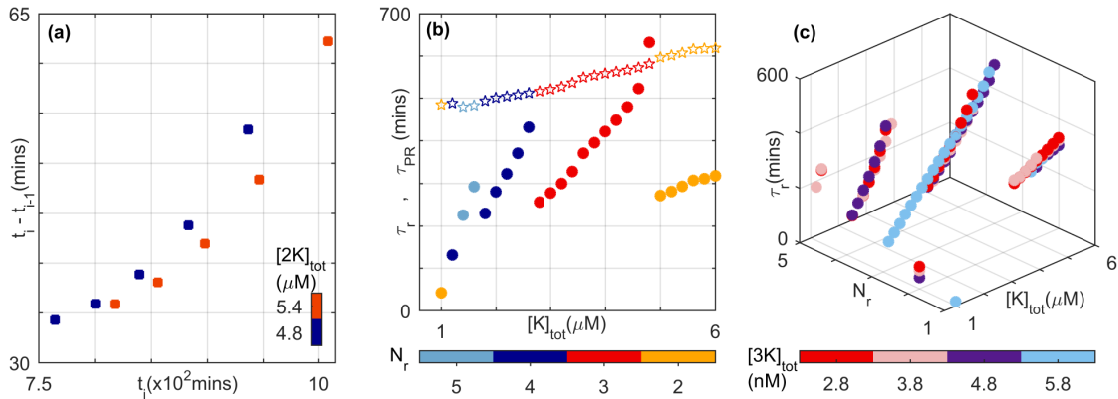
Figure 2.20: Characterization of the reverberatory dynamics observed after withdrawing a stimulus having amplitude $S = 2.0 \times 10^{-6} \mu M$. (a) The interval between successive spikes $i - 1$ and i increases with time (t_i being the time of occurrence of the ith spike) for two distinct total concentrations of MAP2K. The total concentrations of MAPK and MAP3K are $[K]_{tot} = 1.2 \mu M$ and $[3K]_{tot} = 2.8 nM$, respectively. (b) The primary recovery time τ_{PR} (stars) and the total duration of reverberatory activity τ_r (filled circles) are shown for different values of N_r (indicated by the color bar). While τ_{PR} increases monotonically with increasing total MAPK concentration, τ_r shows a more complex dependence ($[2K]_{tot} = 3\mu M$ and $[3K]_{tot} = 4nM$). (c) The dependence of τ_r on $[K]_{tot}$ for different values of N_r has a similar nature for different choices of $[3K]_{tot}$ (indicated by the color bar, $[2K]_{tot} = 3\mu M$). Note that for panel (c), we consider only situations where the system attains a steady state on maintaining the stimulation. For the total concentrations of the phosphatases see Table 2.6.
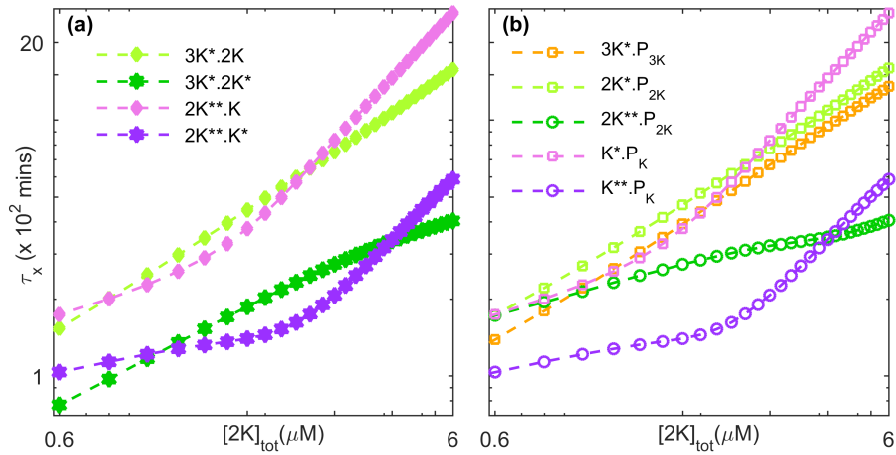
Figure 2.21: Protein complexes in the MAPK cascade exhibit relaxation behavior occurring over a broad range of time-scales. Decay of activity is shown after withdrawing an applied stimulus of amplitude $S = 1.2 \times 10^{-6} \mu M$. The relaxation times τ_x of the different molecular species, viz., (a) the protein complexes between non-phosphorylated and singly phosphorylated (non-active) kinase proteins and the doubly phosphorylated (active) kinase protein of the preceding layer, and (b) the protein complexes between the phosphorylated (singly- or doubly-) kinase proteins and the phosphatase that carries out dephosphorylation in the corresponding layer of the MAPK cascade, vary with the total concentration of MAP2K. The nature of this dependence is distinct for lower and higher values of $[2K]_{tot}$. For both panels, $[K]_{tot} = 0.8 \mu M$ and $[3K]_{tot} = 0.0020 \mu M$. The total concentrations of the phosphatases are provided in Table 2.6.

65
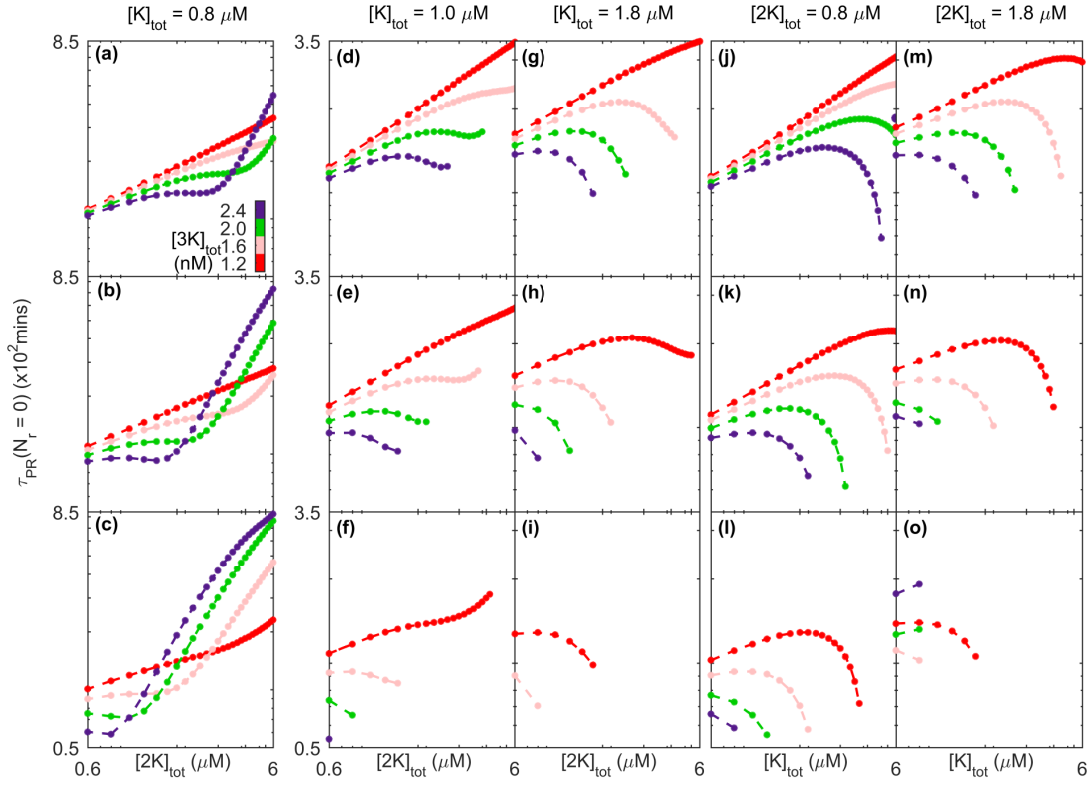
Figure 2.22: Dependence of the primary recovery time τ_{PR} on (a-i) the total concentration of MAP2K ($[2K]_{tot}$) and on (j-o) the total concentration of MAPK ($[K]_{tot}$) for different values of the total concentration of MAP3K ($[3K]_{tot}$), obtained upon removing stimuli having different amplitudes S. Panels (a,d,g,j,m) are for $S = 0.8 \times 10^{-6} \mu M$, panels (b,e,h,k,n) are for $S = 1.2 \times 10^{-6} \mu M$, and panels (c,f,i,l,o) are for $S = 2.0 \times 10^{-6} \mu M$. We have only considered situations where the system reaches a steady state upon application of a time-invariant stimulus, and that do not show any reverberatory activity ($N_r = 0$) during relaxation to the resting state. The curves in panels (a-i) are obtained for different values of $[K]_{tot}$, namely, (a-c) $[K]_{tot} = 0.8 \mu M$, (d-f) $[K]_{tot} = 1.0 \mu M$, and (g-i) $[K]_{tot} = 1.8 \mu M$. The curves in panels (j-o) are obtained for different values of $[2K]_{tot}$, namely, (j-l) $[2K]_{tot} = 0.8 \mu M$, and (m-o) $[2K]_{tot} = 1.8 \mu M$. The total concentrations of the phosphatases for all panels are given in Table 2.6.

66
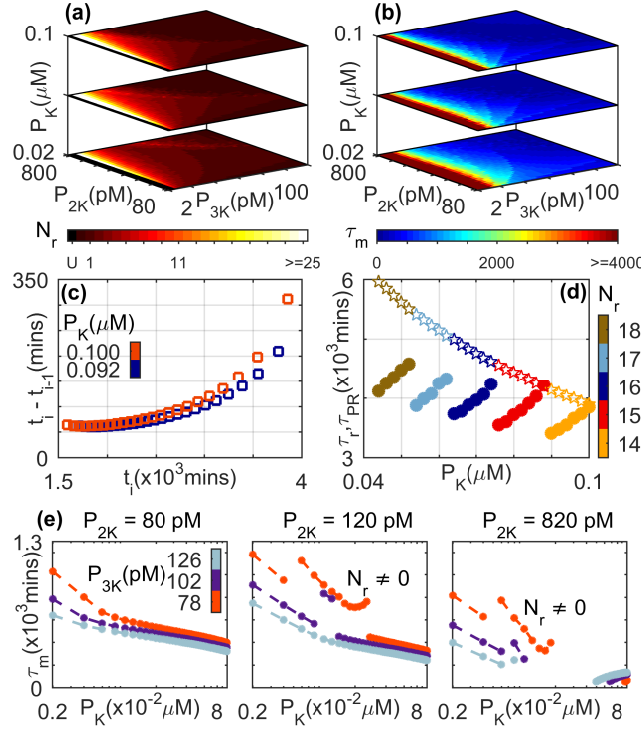
Figure 2.23: Dependence of reverberatory activity on the total concentrations of the phosphatases MAPK PPase ($[P_K]$), MAP2K PPase, ($[P_{2K}]$) and MAP3K PPase ($[P_{3K}]$). (a) The number of spikes N_r and (b) the total memory time τ_m (in minutes) observed on withdrawing an applied stimulus of amplitude $S = 2.0 \times 10^{-6} \mu M$. Situations where the primary recovery time is longer than a maximum or cut-off value (see Methods), such that the reverberatory nature of the dynamics cannot be properly measured, are indicated by the color corresponding to "U". (c) The interval between successive spikes $i - 1$ and i increases with time (t_i being the time of occurrence of the ith spike). As the MAPK PPase concentration is increased, the durations of these intervals are seen to increase. The total concentrations of the other two phosphatases are maintained at $[P_{2K}] = 680 pM$ and $[P_{3K}] = 10 pM$. (d) The variation of primary recovery time τ_{PR} (stars) and the total duration of reverberatory activity τ_r (filled circles) as a function of total MAPK PPase concentration are shown for different values of N_r (indicated by the color bar). While τ_{PR} decreases monotonically with increasing $[P_K]$, τ_r shows a more complex dependence ($[P_{2K}] = 200 pM$ and $[P_{3K}] = 6 pM$). (e) Dependence of the total memory time τ_m on total MAPK PPase concentration ($[P_K]$ shown in log scale) for different total concentrations of MAP2K PPase (values indicated above each of the three panels) and MAP3K PPase (indicated using different colors as shown in the color bar). Note that we consider only situations where the system attains a steady state on maintaining stimulation. For details of the total concentrations of the kinases, see Table 2.5.

67
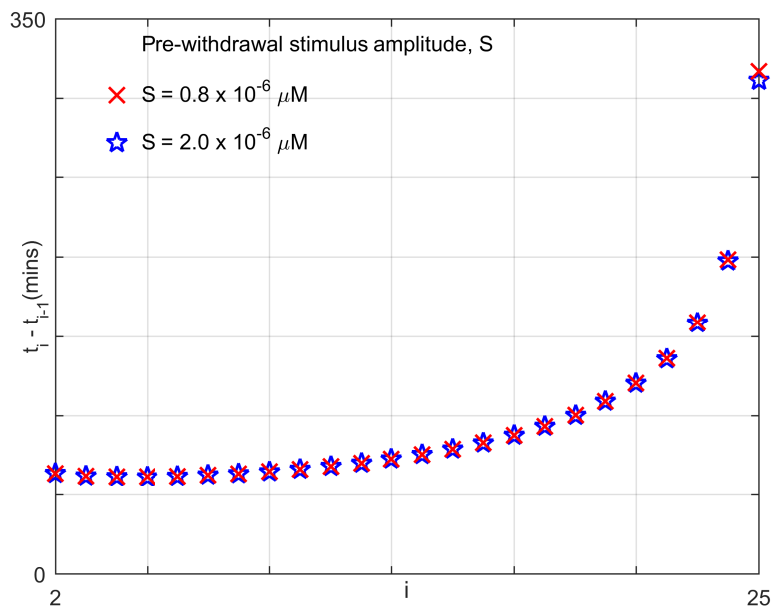
Figure 2.24: The time interval between successive spikes $i - 1$ and i obtained after removing a stimulus, increases with the number of spike events (i being the event number of the ith spike). The trend appears to be independent of the stimulus amplitude S. The total concentrations of the phosphatases are $P_K = 0.1\mu M$, $P_{2K} = 680pM$ and $P_{3K} = 10pM$, respectively. The total concentrations of the kinases are provided in Table 2.5.
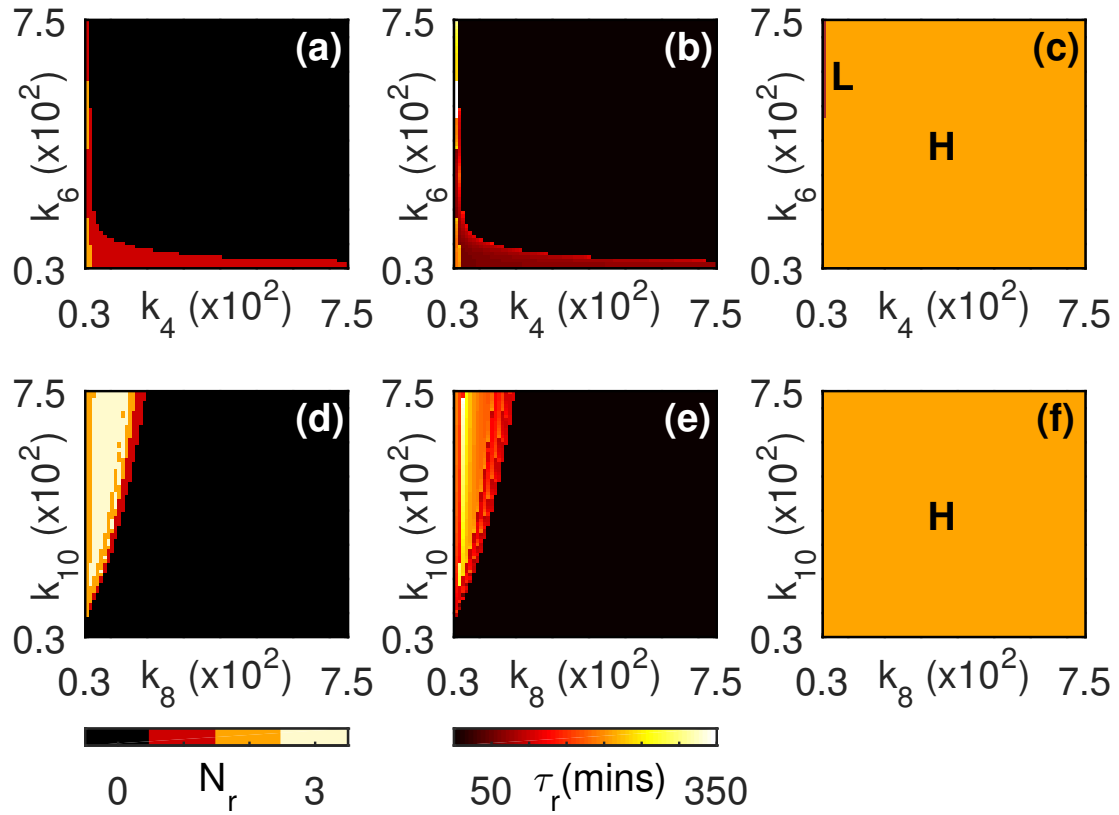
Figure 2.25: Dependence of reverberatory activity on specific kinetic rates governing the dynamics of the MAPK cascade. The corresponding dynamical attractors of the system under sustained stimulation are also shown. (a,d) The number of post-stimulus spikes N_r and (b,e) the total duration of reverberatory activity τ_r (in minutes) observed on withdrawing the stimulus, as well as, (c,f) the corresponding asymptotic dynamical states of the cascade under sustained stimulation, are shown. They are displayed as a function of the kinetic rates k_4 and k_6 which govern the product formation steps in the single and double phosphorylation reactions (respectively) during MAP2K activation (a-c) and of the kinetic rates k_8 and k_{10} which govern the product formation steps in the single and double phosphorylation reactions (respectively) during MAPK activation (d-f). The rates are expressed in units of min^{-1}. The values of the other reaction rates are chosen to be identical to the reference MMS set (Table 2.3). The total concentrations of the kinase and phosphatase molecules are same as in Fig. 2.2 (e) in the main text. The strength of the signal used to stimulate the cascade in all cases is $S = 2 \times 10^{-6} \mu M$. For comparison note that using the MMS reference set for all parameter values will yield $N_r = 2$, $tau_r = 320.7$ minutes and an asymptotic steady state corresponding to high MAPK activity (H) under sustained stimulation.
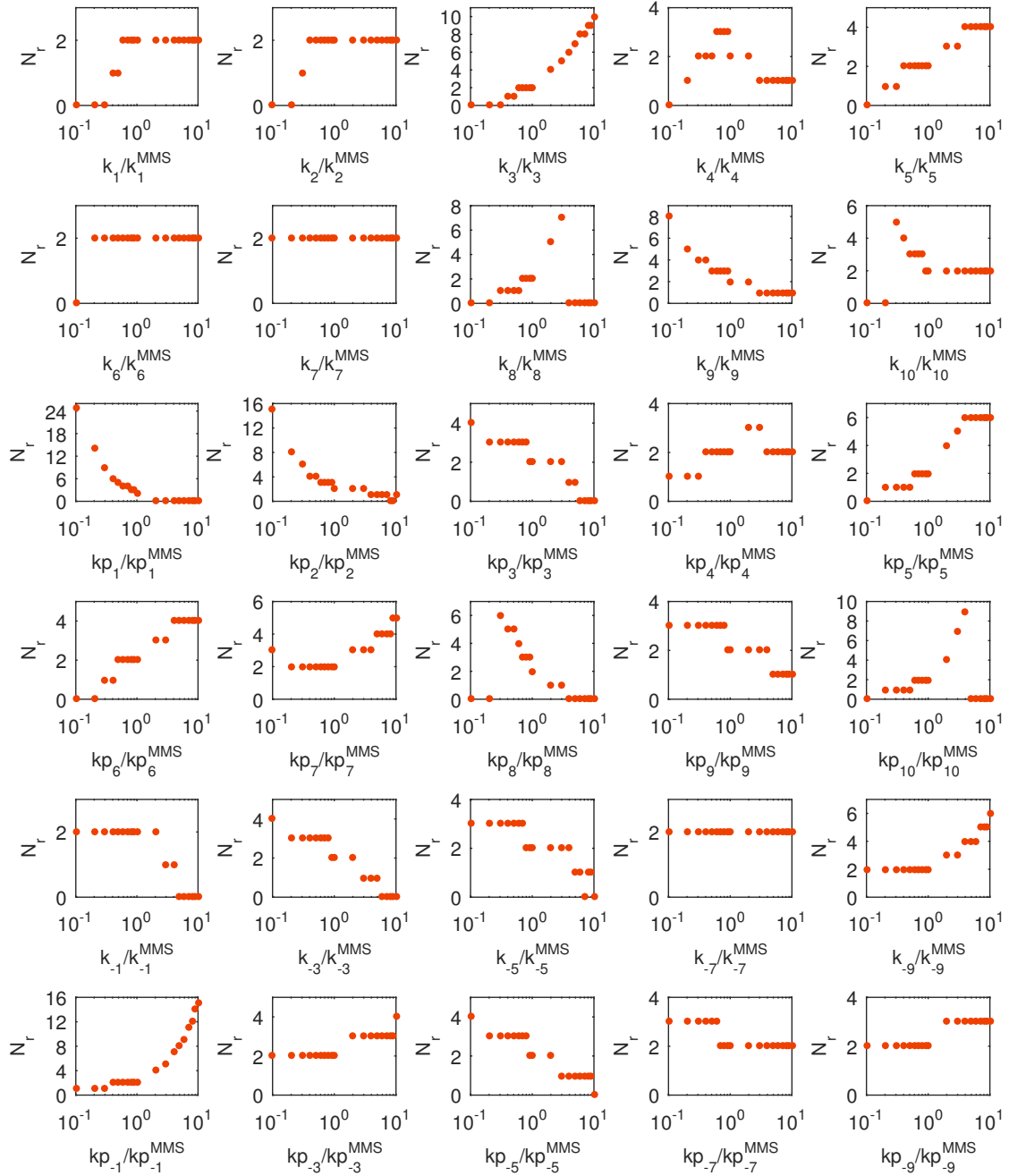
Figure 2.26: Dependence of the reverberatory activity measured in terms of the number of post-stimulus spikes N_r on each of the kinetic rates governing the MAPK cascade dynamics. In each panel, a specific kinetic rate is varied ten-fold either way from the corresponding MMS reference value while keeping all other parameters fixed at the respective value in the MMS reference set (Table 2.3). The total concentrations of the kinase and phosphatase molecules are same as in Fig. 2.2 (e) in the main text. The strength of the signal used to stimulate the cascade in all cases is $S = 2 \times 10^{-6} \, \mu M$. For comparison note that using the MMS reference set for all parameter values will yield $N_r = 2$.
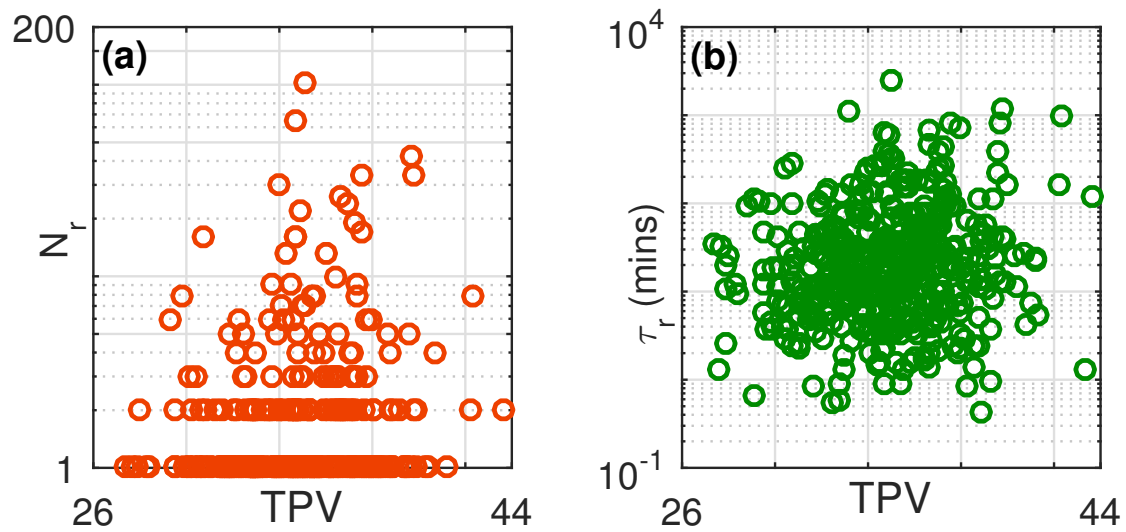
Figure 2.27: Robustness of the observed reverberatory activity in MAPK cascade following withdrawal of applied stimulus having strength $S = 5 \times 10^{-6} \mu M$ with respect to variation in the system parameters. The panels show (a) the number of spikes during relaxation N_r and (b) the total duration of reverberatory activity τ_r, on the Total Parameter Variation (TPV) as defined in the main text (see Methods). The circles in each panel represent an individual realization of the cascade dynamics where each parameter set is chosen by uniform random sampling from a physiologically plausible range (see Table 2.4).

3

Non-associative learning in intra-cellular signaling networks

3.1 Introduction

Nonlinear systems can respond to variations in their environment by exhibiting a wide range of complex dynamical patterns [149–154] that may often be functionally significant [102, 140, 155–159]. These variations are commonly associated with natural cycles such as the diurnal rhythm. In particular, biological systems are typically subjected to periodic stimuli with frequencies that can vary over a wide range of time-scales, viz., from ultradian to infradian rhythms [160–162]. Examples include the entrainment of the circadian clock to the day-night cycle [163], variations in hormonal levels over a period of a month that drive the menstrual cycle [164] and calcium oscillations at the time-scale of minutes which modulate the efficiency and specificity of gene expression [165]. Of all the biological systems capable of exhibiting complex functionally significant responses when driven by periodic stimuli, perhaps one of the simplest is the intra-cellular signaling network [166]. In its natural environment, the membrane-bound receptors of a cell may repeatedly be stimulated on encountering ligands, for instance as a consequence of

pulsatile variations in hormones [167]. Cellular functions may also be modulated by internal cues that vary periodically, e.g., oscillations in the concentrations of intra-cellular messengers such as Ca^{2+} [168, 169] and cyclic AMP [170, 171]. It is therefore important to investigate how key components of the signaling network in the cell respond when subjected to repeated stimulation.

An ubiquitous motif of this network is the mitogen-activated protein kinase (MAPK) cascade, which is found across all eukaryotic cells [98, 99]. It consists of a sequential arrangement of three types of protein kinase, viz., MAPK, MAPK kinase (MAP2K) and MAPK kinase kinase (MAP3K). The activated kinase in each layer of the cascade functions as an enzyme for phosphorylating (and thus activating) the kinase in the layer immediately downstream. The subsequent deactivation is mediated by the corresponding dephosphorylating enzyme known as phosphatases (P'ase). Despite its structural simplicity this motif is involved in regulating a wide array of vital cellular functions, including proliferation and apoptosis [99], stress response [104] and gene expression [105]. Activation of the cascade is initiated when extracellular ligands stimulate membrane-bound receptors, or when intra-cellular cues occur upstream of the cascade, with the information being relayed to MAP3K through a series of intermediaries. The terminal kinase of the motif (MAPK), transmits the signal further downstream by phosphorylating various proteins including transcription regulators [42]. The behavior of the cascade when subjected to sustained stimulation has been extensively investigated in earlier studies, and the existence of several emergent features has been observed. These include ultrasensitivity [106], bistability which allows the system to switch between two states corresponding to low and high activity [108, 113–116] and oscillations [107–110, 112, 131]. In earlier work we have shown that the cascade stimulated with a pulse of finite duration responds with a rich variety of transient behavior, including phenomena indicative of the presence of short-term memory [172]. The complex modulations seen in the response of the cascade are crucially dependent on the interactions between the time-scales of the intrinsic processes and that of the applied stimulus. It is thus intriguing to consider how the system

will respond to repeated stimulation.

In this chapter, we investigate the dynamics of the MAPK cascade upon being stimulated by periodic trains of pulses. Despite the absence of any explicit feedback, under suitable conditions we find that the system displays adaptive behavior including non-associative learning [15, 16], viz., habituation (desensitization) and sensitization. These allow plasticity in the behavioral repertoire of the intra-cellular signaling motif by enabling modification of the strength, duration and even the qualitative nature of its response to recurrent stimulation. In addition to these, we report the occurrence of a temporal sequence of strong and weak responses to successive pulses, reminiscent of the phenomenon of "alternans" [173, 174] in excitable cells (examples of such cells include ventricular myocytes and neurons [175]). This, coupled with the existence of a response threshold and an apparent refractory behavior when subjected to high-frequency stimulation strongly suggests an analogy with excitable media [175]. While learning is commonly associated with behavior at the level of organisms [3–7, 12], it is intriguing that rudimentary forms of such complex adaptive responses can be seen in a simple network of sub-cellular components. As the MAPK signaling cascade is involved in coordinating diverse processes in all eukaryotic cells, these results point to the potential functional utility of such emergent dynamical phenomena in these systems.

3.2 Methods

We have simulated the dynamics of the three layer kinase cascade using the Huang-Ferrell model of the MAPK signaling motif [106], schematically illustrated in Fig. 3.1 (a). This model consists of 10 enzyme-substrate reactions described by 18 coupled differential equations (see Appendix for details). Each of the several kinase and phosphatase-mediated enzyme-substrate reactions in the cascade consist of (i) a reversible enzyme-substrate complex formation step, and (ii) an irreversible step corresponding to the activa-

tion/deactivation of a kinase. The ratio of the activation and deactivation rates ranges over four orders of magnitude [108], underlining the vast diversity of dynamical time-scales present in the system. The equations are numerically solved without invoking the quasi-steady-state hypothesis [132]. We explicitly ensure that the total concentrations of each of the constituent kinases in the system are conserved. In our simulations, we assume that the cascade is initially in the resting state, where the kinases are completely non-phosphorylated. Following the exposure of the cascade to a train of pulses, we record the resulting response pattern, viz., the MAPK activity.

3.3 Results

Investigations into the dynamics of the Huang-Ferrell model [106] have typically focused on the asymptotic response of the cascade to sustained stimulation. In contrast, here we investigate the response of the system when it is subjected to recurrent activation by periodic stimuli. Specifically, we consider a signal comprising a train of pulses, each having amplitude S, duration P and separated from each other by an inter-pulse interval I [Fig. 3.1 (a)]. The cascade is released from stimulation between two successive pulses, and attempts to relax back to its resting state. On arrival of the next pulse, the cascade is activated once more, albeit before it has completely relaxed. This, coupled with the multiple time-scales of activation and relaxation present in the system, results in non-trivial adaptive temporal response. Selected examples of such behavior are shown in Fig. 3.1 (b-d). These different time series of the activated MAPK concentration (normalized with respect to the total MAPK concentration) correspond to the cascade being subjected to pulse trains characterized by different parameter values of P and I.

Fig. 3.1 (b) displays the response of the system subjected to high-frequency stimulation by short-duration pulses. Here, starting from its resting state value, each subsequent pulse elicits a slightly higher response of n_K^{**} until the peak activation suddenly spikes to a value
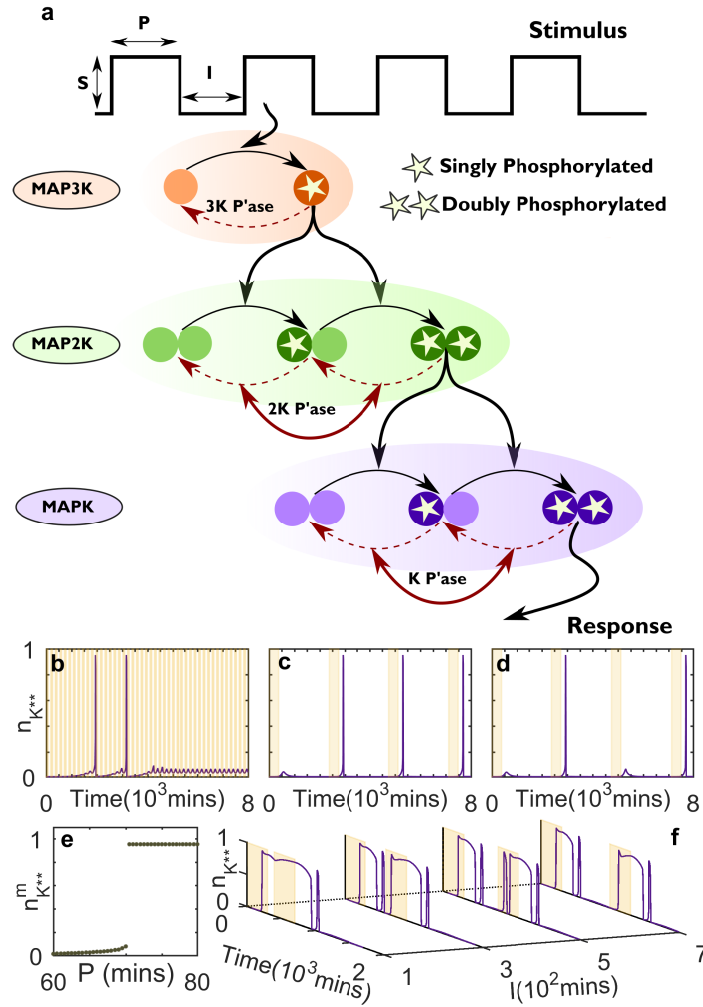
Figure 3.1: Non-associative learning in a MAPK cascade stimulated by a pulse train. (a) Schematic representation of a linear three-layer MAPK cascade whose component kinases are activated/deactivated by the addition/removal of phosphate groups through phosphorylation/dephosphorylation respectively. Signaling is initiated when MAPK kinase kinase (MAP3K) is activated by a periodic signal comprising a series of pulses having amplitude S and duration P, separated by inter-pulse interval I. For the cases investigated here, the cascade receives no stimulus between two successive pulses. The response of the cascade to the signal is measured in terms of MAPK activity, viz., the normalized concentration n_K^{**} of doubly phosphorylated MAPK. (b-d) Time series representing qualitatively different adaptive responses of the cascade to pulse trains characterized by a range of S, P and I. The shaded bars correspond to the intervals during which MAP3K is stimulated. (b) Desensitization behavior of the cascade corresponding to an attenuated response on persistent exposure to the periodic stimulus. (c) Sensitization of the cascade characterized by a low level of MAPK activity on initial exposure followed by stronger responses upon repeated stimulation. (d) Alternating high and low levels of MAPK activity ("alternans") in response to successive pulses. (e) Threshold-like response to the pulse duration P of the maximum MAPK activity for a fixed set of values of the signal strength S and inter-pulse duration I of the pulse train. (f) Nonlinear dependence of the MAPK cascade response on the inter-pulse interval for a pair of pulses (shaded bars). For details of system and signal parameter values used see Appendix.

77
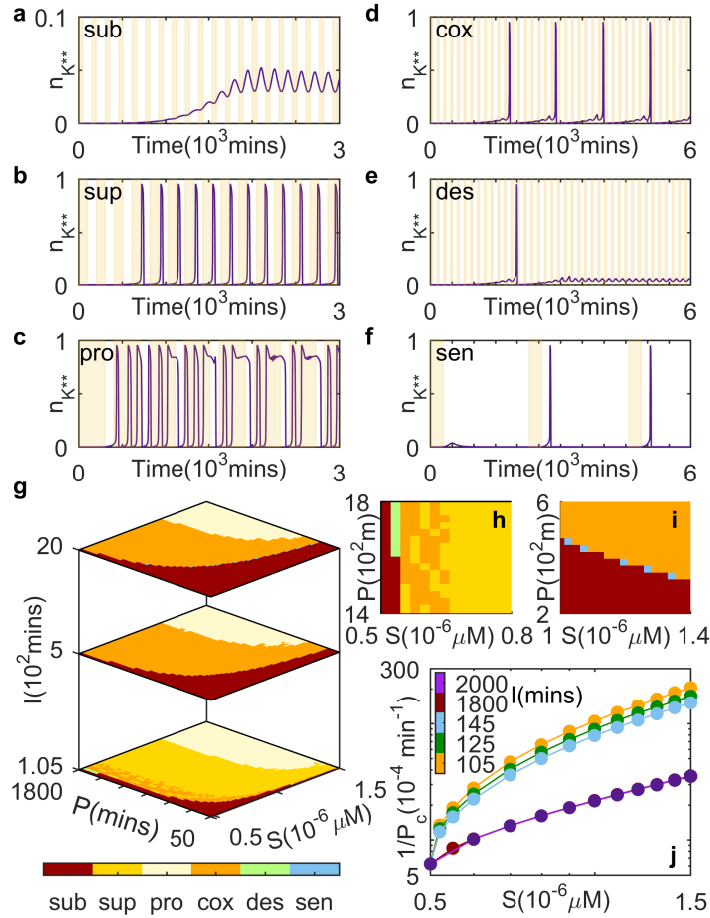
Figure 3.2: Characterization of different responses of MAPK cascade to stimulation of MAP3K by a train of pulses. (a-f) Characteristic responses of the MAPK cascade to stimulation of MAP3K by a train of pulses, each of amplitude S having duration P, with inter-pulse interval I: (a) attenuated response of the cascade characterized by sub-threshold activity (sub), (b) large-amplitude spiking responses characterizing supra-threshold activity (sup), (c) prolongation of supra-threshold activity duration (pro) on application of a signal having pulses with longer duration, (d) coexistence (cox) of sub- and supra-threshold activity which, for a range of I, results from the integration of responses over several preceding pulses, (e) desensitization (des), where integration over multiple successive pulses results in a supra-threshold spiking response but subsequently only exhibits sub-threshold activity, and (f) sensitization (sen), where sub-threshold activity in response to the initial pulse gives way to supra-threshold activity for all subsequent pulses. The shaded bars correspond to the intervals during which MAP3K is stimulated. (g) Dependence of the cascade response on the pulse strength S and duration P for three different values of the inter-pulse interval I. The colors represent the nature of the response [classified into the categories (a-f) mentioned above]. (h-i) Magnified views of the $P - S$ planes for (h) $I = 105$ and (i) 2000 minutes show the regions corresponding to desensitization and sensitization, respectively. Each grid point in panels (g-i) represents the response of the cascade starting from an initial condition corresponding to the resting state. (j) The variation of the critical value of pulse duration P_c, above which the cascade exhibits supra-threshold response, with pulse amplitude S. The curves correspond to pulse trains having different inter-pulse intervals I (as shown in the colorbar).

78

close to its saturation. This behavior can be interpreted as a form of signal integration, and may be repeated multiple times as the pulse train is continued. However, for an appropriate range of P and I (as in the figure), after a given number of pulses we observe behavior analogous to *desensitization* when the system no longer shows spiking activity, even for sustained periodic stimulation. Thus, following an initial large amplitude response, the subsequent activity of the system is attenuated even though the nature of the received signal remained unchanged. When the cascade is stimulated instead by low-frequency pulse trains having relatively longer pulse durations, we observe a phenomenon analogous to *sensitization*. Here the cascade exhibits low-level activity on receiving the initial pulse but switches to high-amplitude spiking in response to all subsequent pulses [Fig. 3.1 (c)]. Thus, the initial low-level activity effectively "primes" the cascade to reach response levels close to saturation. This occurs because of the existence of long relaxation time-scales in certain components of the cascade, allowing for response accumulation over successive stimulations. Decreasing P by a small amount gives rise to a qualitatively distinct phenomenon characterized by alternating low and high peaks of MAPK activity, reminiscent of *alternans* [173, 174].

As alternans is a phenomenon that is associated with excitable media [173], it is intriguing to consider whether the periodically stimulated cascade exhibits other characteristics of such systems, in particular, the existence of a response threshold [175]. As seen in Fig. 3.1 (e), there is indeed a large discontinuous change in the peak activation $n_{K^{**}}^m$ of MAPK when the pulse duration P crosses a specific value P_c that depends on the choice of S and I. Extending the analogy with excitable media, we find that the cascade also exhibits a nonlinear relation between its response to successive pulses and the inter-pulse interval. This can be seen from the behavior displayed in Fig. 3.1 (f), where the cascade is stimulated by a pair of pulses separated by an interval I. When I is reduced, the response duration resulting from the second pulse increases in comparison to the duration of the response caused by the first. As an aside, we note that for the parameter regime considered here, the system exhibits post-stimulus reverberatory activity [172].

Fig. 3.2 (a-f) depicts a set of representative time-series showing the activity of the cascade on either side of the response threshold, obtained for different choices of the periodic stimulation parameters. We note that in all of the cases shown here, the system shows a gradual build-up of activity over multiple pulses before reaching asymptotic peak activity levels. This corresponds to signal integration (mentioned earlier), where the response of the system to successive stimuli is modulated by the preceding stimuli. Fig. 3.2 (a) shows a typical subthreshold response (*sub*), where the peak MAPK activity is highly attenuated ($< 5\%$ of the saturation response value). Note that the nature of the response (i.e., whether it is sub- or supra-threshold) is a function of all three stimulation parameters S, P and I. For instance, for the same signal strength S considered in panel (a), the steady-state response of the cascade would have been close to saturation if the stimulation had been applied in a sustained fashion (i.e., $I \to 0$).

Panels (b-f) of Fig. 3.2 depict a variety of suprathreshold temporal behavior, the simplest of which is characterized by a 1:1 spiking response to the periodically applied pulses [*sup*, Fig. 3.2 (b)]. On varying the different stimulation parameters we observe other types of suprathreshold activity. For example, on increasing P alone (or alternatively, S alone), the system exhibits prolongation of the peak activity close to saturation [*pro*, Fig. 3.2 (c)]. For high-frequency stimulation (i.e., low I) after a transient period we observe suprathreshold peak responses only after every N pulses for values of P and S that lie between those giving rise to *sub* and *sup* responses [see the lowest plane of Fig. 3.2 (g)]. This response behavior, which corresponds to the coexistence (*cox*) of peak activity levels having different amplitudes (ranging from values just above zero to near-saturation) is shown in Fig. 3.2 (d). For lower frequency stimuli, the *cox* regime corresponds to $M : 1$ response where multiple peaks in MAPK activity, whose amplitudes can again vary widely, are observed in response to each pulse [not shown]. Apart from these, we also observe behavior corresponding to non-associative learning, viz., desensitization [*des*, Fig. 3.2 (e)] and sensitization [*sen*, Fig. 3.2 (f)], as described earlier. Specifically, at the interface of the *cox* and *sub* regions in the stimulation parameter space, the *des* response regime is ob-

served for high-frequency pulse trains [Fig. 3.2 (h)] while for low-frequency stimulation we obtain *sen* [Fig. 3.2 (i)]. We note that for higher values of S, the transition from *cox* to *sub* gets sharper thereby reducing the range of P over which the *des* and *sen* phenomena are observed. An overview of the responses observed in the stimulation parameter space is given in Fig. 3.2 (g) indicating the conditions for which each of the responses described above can be obtained.

The "learning" behavior associated with the periodically stimulated cascade is seen in the vicinity of the response threshold mentioned earlier corresponding to the boundary of the *sub* regime [Fig. 3.2 (g)]. Hence, we examine the dependence of the threshold on the stimulation parameters in Fig. 3.2 (j). The reciprocal relation between the signal strength S and the critical pulse duration P_c necessary for suprathreshold response, seen over a wide range of S, suggests that the threshold is determined by the total signal intensity of a pulse, which is measured as the product of S and P. Deviation from this simple relation is observed for sufficiently low signal strength. This implies that a minimal value of S is required to observe a suprathreshold response, regardless of the duration for which the pulse is maintained. We note that in the limit of $I \rightarrow 0$, this minimal signal strength corresponds to the lower critical value required to observe a transition from a low level of MAPK activity to high-amplitude oscillations in the case of a cascade subjected to sustained stimulation [108, 172]. As I is increased, we observe that the response threshold (measured in terms of the critical pulse duration P_c) increases, which suggests that the excitability of the system reduces as the frequency of the periodic stimulus decreases.

The phenomena reported here are robust with respect to variations in the model parameters around the values used in this chapter, including the kinetic rate constants and the molecular concentrations of the constituent kinases and phosphatases. We have also observed similar behavior with cascades having branched architecture, e.g., MAP3K activating two different types of MAP2K [111]. While we have assumed that the same phosphatase acts on both the singly and doubly phosphorylated forms of the kinase in a

particular layer of the cascade (as in the canonical Huang-Ferrell model), we have explicitly verified that our results are not sensitively dependent on this.

3.4 Discussion

A mechanistic understanding of the phenomena reported here is made difficult by the large number of coupled dynamical variables in the model that operate across different time-scales. This complexity may be untangled by using the framework of excitable systems. As alluded to earlier, many of the characteristic features associated with excitability are present in the system investigated here. These include the existence of two qualitatively distinct states of activation separated by a threshold [Fig. 3.1 (e)], a nonlinear response to repeated stimulation [Fig. 3.1 (f)], an apparently refractory behavior as seen most prominently during desensitization [Fig. 3.1 (b)] and phenomena analogous to alternans [Fig. 3.1 (d)]. This appealing analogy provides a means by which a phenomenological understanding of the emergent behavior of this complex system might be achieved. We note that the excitability paradigm has been invoked earlier to explain aspects of cellular activity in the context of antigen recognition by T cells [176, 177]. Our results show that the emergent dynamics of MAPK cascade, which is known to mediate immune response [178], provides an explicit mechanistic basis for such a theoretical framework to explain the adaptive response of the immune system to its microenvironment.

Among the functionally significant dynamical phenomena reported here, the phenomenon of learning is perhaps the most intriguing. It confers on the system the ability to modify its behavior in response to information, which is critical for adapting to a changing environment. The capability to learn often presupposes the existence of a feedback that allows bidirectional communication between the components associated with receiving a signal and those that initiate a corresponding response [179]. In the kinase cascade investigated here, an explicit feedback is absent as each layer activates the one immediately

downstream. However, an implicit feedback results from the inherent features of kinase activation, viz., sequestration and multi-site phosphorylation [108, 111, 113, 115]. This can have non-trivial consequences, such as the appearance of short-term memory, even when the MAPK cascade is subjected to a single pulse [172].

To conclude, in this chapter we have shown that a rich repertoire of responses can be obtained when the system is exposed to a train of pulses. This results from the implicit feedback, which orchestrates an interplay between the periodic stimulus and the diverse activation and relaxation time-scales of the signaling components. In particular, the system can exhibit sensitization and desensitization, which are examples of non-associative learning. These may play an important role in the cell's ability to function in its natural environment, where it is continually exposed to signals of varying intensity and duration. This necessitates an ability to respond selectively to the received stimuli. Such adaptive mechanisms allow the cell to ignore persistent background stimuli through habituation (desensitization) but respond strongly to infrequent signals to which it has been primed through earlier exposure (sensitization). Given that a single linear cascade exhibits such complex adaptive behavior, it is intriguing to speculate about the potential capabilities inherent in the coordinated action of multiple subcellular processes [101]. The mechanism through which learning at the sub-cellular scale can impact adaptive behavior in an organism at cellular and possibly higher scales remains an intriguing question.

3.A Appendix: The Model Equations

Table 3.1: Components of the MAPK Cascade

Component	Notation	Symbol
Mitogen-activated Protein Kinase Kinase Kinase	MAP3K	3K
Singly Phosphorylated Mitogen-activated Protein Kinase Kinase Kinase	MAP3K*	3K*
Mitogen-activated Protein Kinase Kinase	MAP2K	2K
Singly Phosphorylated Mitogen-activated Protein Kinase Kinase	MAP2K*	2K*
Doubly Phosphorylated Mitogen-activated Protein Kinase Kinase	MAP2K**	2K**
Mitogen-activated Protein Kinase	MAPK	K
Singly Phosphorylated Mitogen-activated Protein Kinase	MAPK*	K*
Doubly Phosphorylated Mitogen-activated Protein Kinase	MAPK**	K**
MAP3K-Phosphatase	3K P'ase	P_{3K}
MAP2K-Phosphatase	2K P'ase	P_{2K}
MAPK-Phosphatase	K P'ase	P_K

The three layer MAPK cascade comprises the following enzyme-substrate reactions:
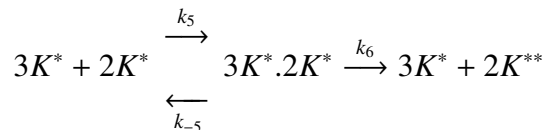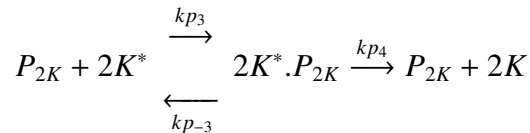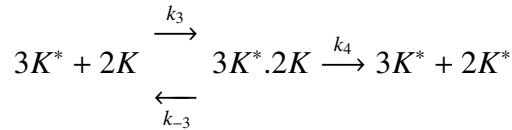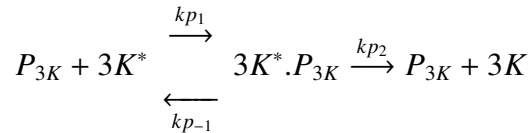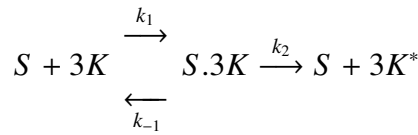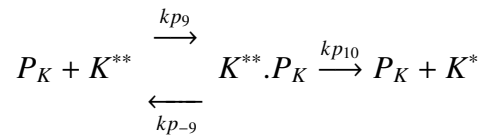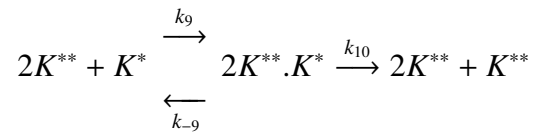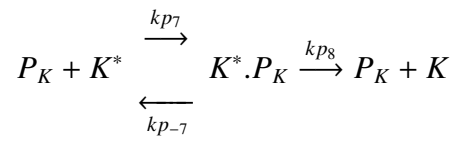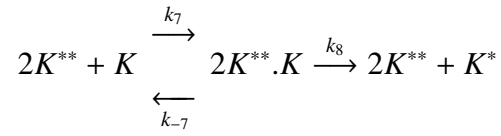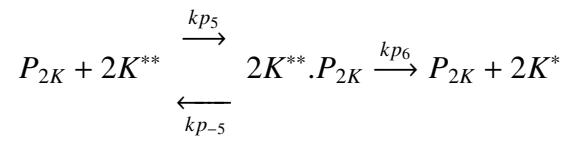
$$S + 3K \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} S.3K \xrightarrow{k_2} S + 3K^*$$

$$P_{3K} + 3K^* \underset{kp_{-1}}{\overset{kp_1}{\rightleftharpoons}} 3K^*.P_{3K} \xrightarrow{kp_2} P_{3K} + 3K$$

$$3K^* + 2K \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} 3K^*.2K \xrightarrow{k_4} 3K^* + 2K^*$$

$$P_{2K} + 2K^* \underset{kp_{-3}}{\overset{kp_3}{\rightleftharpoons}} 2K^*.P_{2K} \xrightarrow{kp_4} P_{2K} + 2K$$

$$3K^* + 2K^* \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} 3K^*.2K^* \xrightarrow{k_6} 3K^* + 2K^{**}$$

84

$$P_{2K} + 2K^{**} \underset{kp_{-5}}{\overset{kp_5}{\rightleftharpoons}} 2K^{**}.P_{2K} \xrightarrow{kp_6} P_{2K} + 2K^*$$

$$2K^{**} + K \underset{k_{-7}}{\overset{k_7}{\rightleftharpoons}} 2K^{**}.K \xrightarrow{k_8} 2K^{**} + K^*$$

$$P_K + K^* \underset{kp_{-7}}{\overset{kp_7}{\rightleftharpoons}} K^*.P_K \xrightarrow{kp_8} P_K + K$$

$$2K^{**} + K^* \underset{k_{-9}}{\overset{k_9}{\rightleftharpoons}} 2K^{**}.K^* \xrightarrow{k_{10}} 2K^{**} + K^{**}$$

$$P_K + K^{**} \underset{kp_{-9}}{\overset{kp_9}{\rightleftharpoons}} K^{**}.P_K \xrightarrow{kp_{10}} P_K + K^*$$

The above enzyme-substrate reactions can be expressed in terms of the following coupled ordinary differential equations (ODEs):

$$\frac{d[3K]}{dt} = k_{-1}.[S.3K] + kp_2.[3K^*.P_{3K}] - k_1.[S].[3K],$$

$$\frac{d[S.3K]}{dt} = k_1.[S].[3K] - (k_{-1} + k_2).[S.3K],$$

$$\frac{d[3K^*.P_{3K}]}{dt} = kp_1.[P_{3K}^f].[3K^*] - (kp_2 + kp_{-1}).[3K^*.P_{3K}],$$

$$\frac{d[3K^*]}{dt} = k_2.[S.3K] + kp_{-1}.[3K^*.P_{3K}] - kp_1.[P_{3K}^f].[3K^*]$$
$$+ (k_{-3} + k_4).[3K^*.2K] - k_3.[3K^*].[2K]$$
$$+ (k_{-5} + k_6).[3K^*.2K^*] - k_5.[3K^*].[2K^*],$$

$$\frac{d[2K]}{dt} = k_{-3}.[3K^*.2K] + kp_4.[2K^*.P_{2K}] - k_3.[3K^*].[2K],$$

$$\frac{d[3K^*.2K]}{dt} = k_3.[3K^*].[2K] - (k_{-3} + k_4).[3K^*.2K],$$

$$\frac{d[2K^*.P_{2K}]}{dt} = kp_3.[P_{2K}^f].[2K^*] - (kp_4 + kp_{-3}).[2K^*.P_{2K}],$$

$$\frac{d[2K^*]}{dt} = k_4.[3K^*.2K] + kp_{-3}.[2K^*.P_{2K}] - kp_3.[P_{2K}^f].[2K^*]$$
$$+ k_{-5}.[3K^*.2K^*] - k_5.[3K^*].[2K^*] + kp_6.[2K^{**}.P_{2K}],$$

$$\frac{d[3K^*.2K^*]}{dt} = k_5.[3K^*].[2K^*] - (k_6 + k_{-5}).[3K^*.2K^*],$$

$$\frac{d[2K^{**}.P_{2K}]}{dt} = kp_5.[P_{2K}^f].[2K^{**}] - (kp_6 + kp_{-5}).[2K^{**}.P_{2K}],$$

$$\frac{d[2K^{**}]}{dt} = k_6.[3K^*.2K^*] + kp_{-5}.[2K^{**}.P_{2K}] - kp_5.[P_{2K}^f].[2K^{**}]$$
$$+(k_{-7} + k_8).[2K^{**}.K] - k_7.[2K^{**}].[K]$$
$$+(k_{-9} + k_{10}).[2K^{**}.K^*] - k_9.[2K^{**}].[K^*],$$

$$\frac{d[K]}{dt} = k_{-7}.[2K^{**}.K] + kp_8.[K^*.P_K] - k_7.[2K^{**}].[K],$$

$$\frac{d[2K^{**}.K]}{dt} = k_7.[2K^{**}].[K] - (k_8 + k_{-7}).[2K^{**}.K],$$

$$\frac{d[K^*.P_K]}{dt} = kp_7.[P_K^f].[K^*] - (kp_{-7} + kp_8).[K^*.P_K],$$

$$\frac{d[K^*]}{dt} = k_8.[2K^{**}.K] + kp_{-7}.[K^*.P_K] - kp_7.[P_K^f].[K^*]$$
$$+k_{-9}.[2K^{**}.K^*] - k_9.[2K^{**}].[K^*] + kp_{10}.[K^{**}.P_K],$$

$$\frac{d[2K^{**}.K^*]}{dt} = k_9.[2K^{**}].[K^*] - (k_{-9} + k_{10}).[2K^{**}.K^*],$$

$$\frac{d[K^{**}.P_K]}{dt} = kp_9.[P_K^f].[K^{**}] - (kp_{-9} + kp_{10}).[K^{**}.P_K],$$

$$\frac{d[K^{**}]}{dt} = k_{10}.[2K^{**}.K^*] + kp_{-9}.[K^{**}.P_K] - kp_9.[P_K^f].[K^{**}].$$

where

$$[S] = [S]_{\text{tot}} - [S.3K],$$

$$[P_{3K}^f] = [P_{3K}] - [3K^*.P_{3K}],$$

$$[P_{2K}^f] = [P_{2K}] - [2K^*.P_{2K}] - [2K^{**}.P_{2K}],$$

$$[P_K^f] = [P_K] - [K^*.P_K] - [K^{**}.P_K].$$

It is explicitly ensured that the total concentrations of all individual kinases and phosphatases are conserved at all times. The concentrations of the different molecular species can vary over several orders of magnitudes. We have therefore numerically solved the

equations using low relative and absolute tolerances in order to ensure the accuracy of the resulting time-series.

3.B Appendix: System Parameters

The numerical values for the reaction rates used in all our simulations are obtained from Ref. [111], and are listed in Table 3.2. Please note that these values of kinetic rate constants are very close to that of Huang-Ferrell base values [106].

Table 3.2: Reaction Rates

Rate constant	Our base value	Huang-Ferrell value	Units
k_1	1002	1000	$(\mu M.min)^{-1}$
k_{-1}	150	150	min^{-1}
k_2	150	150	min^{-1}
kp_1	1002	1000	$(\mu M.min)^{-1}$
kp_{-1}	150	150	min^{-1}
kp_2	150	150	min^{-1}
k_3	1002	1000	$(\mu M.min)^{-1}$
k_{-3}	30	150	min^{-1}
k_4	30	150	min^{-1}
kp_3	1002	1000	$(\mu M.min)^{-1}$
kp_{-3}	150	150	min^{-1}
kp_4	150	150	min^{-1}
k_5	1002	1000	$(\mu M.min)^{-1}$
k_{-5}	30	150	min^{-1}
k_6	30	150	min^{-1}
kp_5	1002	1000	$(\mu M.min)^{-1}$
kp_{-5}	150	150	min^{-1}
kp_6	150	150	min^{-1}
k_7	1002	1000	$(\mu M.min)^{-1}$
k_{-7}	30	150	min^{-1}
k_8	30	150	min^{-1}
kp_7	1002	1000	$(\mu M.min)^{-1}$
kp_{-7}	150	150	min^{-1}
kp_8	150	150	min^{-1}
k_9	1002	1000	$(\mu M.min)^{-1}$
k_{-9}	150	150	min^{-1}
k_{10}	150	150	min^{-1}
kp_9	1002	1000	$(\mu M.min)^{-1}$
kp_{-9}	150	150	min^{-1}
kp_{10}	150	150	min^{-1}

The signal parameters used to generate representative time-series in Fig. 3.1–3.2 following the introduction of a signal are listed in Table 3.3 and Table 3.4 respectively.

Table 3.3: Signal parameters for the panels in Fig. 3.1

Parameter	(b)	(c)	(d)	(e)	(f)	Units
S	1.2	1.2	1.2	1.2	3	$\times 10^{-6} \mu M$
P	71	372	371	$60 - 80$	300	$mins$
I	106.75	2000	2000	105	$100 - 700$	$mins$

Table 3.4: Signal parameters for the panels in Fig. 3.2

Parameter	(a)	(b)	(c)	(d)	(e)	(d)	Units
S	1.5	1.5	1.5	1.5	1.5	1.5	$\times 10^{-6} \mu M$
P	49	100	300	50	49.5	287	$mins$
I	105	105	105	105	105	2000	$mins$

Table 3.5: Total concentration (in μM) of the kinases and phosphatase proteins for Figs. 3.1–3.2

Protein	Value
$[K]_{tot}$	4.8
$[2K]_{tot}$	1.2
$[3K]_{tot}$	0.0030
MAP3K-Phosphatase	1×10^{-4}
MAP2K-Phosphatase	3×10^{-4}
MAPK-Phosphatase	0.05

4

Inferring the network relating immune cell types in a human population: Correlation analysis of data from adult and cord blood samples

4.1 Introduction

An integral attribute of the human immune system is its ability to adapt upon exposure to foreign antigens [23]. The development of this "adaptive immunity" commences after birth, and manifests in terms of the memory components of the adult immune system, viz. the levels of certain cell types in the blood [20, 23]. In contrast, the immune system of newly born infants has considerably lower levels of cells that comprise the memory compartments of adaptive immune system (such as memory T-cells and B-cells). The immune system of them is broadly characterized by the general cell types of the adaptive immune system that hardly differentiate in making the memory pool upon first encounters of infectious agents [180], and the innate immune system whose role is to mount a defence

against general pathogens [22, 23, 59]. While there has been significant research into determining the roles of the different cell types of the innate [60] and adaptive immune systems [22, 23, 181], there remain intriguing questions related to the development of immunity over time. In particular, it is of interest to investigate how the relative levels of immune cell types change upon decades of exposure to pathogens in the environment.

In this chapter, we perform a statistical analysis of datasets related to the population of immune cell types in two different stages of maturation, namely post-parturition and in adults. This data was experimentally obtained by our collaborators at the National Institute of Immunology (NII) from a population of healthy adults and from the umbilical cord blood of a number of infants after birth. The data that we consider comprises information related to the total cell counts for all major types of immune cells. Our analysis probes the genetic, environmental and developmental signatures in the underlying networks of immune cells. We quantify the change in cross-correlation of cell abundance in cord and adult blood, and the corresponding probability density functions of the correlation matrix in each case. We observe that the cord and adult blood samples form two distinct clusters in an abstract space formed by the leading principal components obtained from spectral analysis, even without considering specific memory cell types. This suggests that the underlying network of cell types of the immune system evolves substantially over the course of maturation.

4.2 Methods

4.2.1 Data Description

The two datasets used in our study correspond to population counts for 26 distinct immune cell types for adult blood and umbilical cord blood, respectively. These counts are determined from stoichiometric analysis [182] of the blood drawn from 78 adult individ-

uals as well as from 76 umbilical cords during child birth. As there are very few entries for PB cells in cord blood samples, we exclude PB cells from our analysis of cord blood data. In addition, we exclude those samples for which cell counts are only partially available. Our reduced data set thus consists of 73 adult individuals and 75 cord samples. We then obtain a normalized data set by dividing the population size of each of the basic immune cell types by the total leucocyte count. Finally, we express this data in terms of the corresponding z-scores, and use the resulting dataset for all subsequent analysis.

4.2.2 Spearman Rank Correlation

Having datasets of several immune cell counts for different samples, one can first try to analyze the inter-dependence of these population sizes across the samples for both adults and cords. We have calculated this dependence in terms of the pairwise Spearman Rank Correlation coefficient $[C(i, j)]$ that measures the monotonic dependence between two variables i and j. Spearman's rank correlation is a type of non-parametric measure to probe how the rankings of two variables are related [183]. The usual Pearson Correlation Coefficient [184] if calculated between the ranked variables (the ranks that are obtained after converting the original data) gives the measure of $C(i, j)$. Let's say i_x and j_x are two variables of a dataset having n sample (i.e., x represents the index varying from 1 to n). If ri_x and rj_x are the new variables after converted to ranks, $C(i, j) = \text{cov}\,(ri_x, rj_x)/\sigma_{ri_x}.\sigma_{rj_x}$, where 'cov and '$\sigma$' denotes the covariance and standard deviation of the rank variables.

4.2.3 Principal Component Analysis

The immune cell components that we consider are biologically regulated by multiple common factors (such as chemokines and cytokines) and are often functionally interdependent. In order to account for this, we reduce the dimensionality of the data while attempting to preserve as much information from the original data as possible. To this

end, we perform a Principle Component Analysis (PCA) that orthogonally transforms the raw data containing the normalized cell counts of different immune cell types to obtain a new set of linearly uncorrelated variables [185]. We then obtain the principle components through an eigenvalue decomposition [186] of the cross-correlation matrices. The eigenmodes corresponding to the first four largest eigenvalues yield the first four principle components, and their corresponding eigenvector components represent the contributions from each of the original variables, viz. the immune cell types.

4.2.4 Relative Change in Correlation Measures

Relative change in cross correlation (RC_{ij}) has been calculated as follows: $RC_{ij} = (C_{ij}^{Adult} - C_{ij}^{Cord})/((C_{ij}^{Adult} + C_{ij}^{Cord})/2)$.

4.3 Results

We first consider the distribution of cell types (Table 4.1) across individuals in each of the data sets (viz., cord and adult blood samples). In order to compare the relative abundances of the different immune cell sub-types in the cord and adult blood, we limit our analysis to 25 cell types, discarding data relating to Plasma B cells (PB), which is nearly absent in the cord blood, We would like to emphasize that the results for the case of adult blood data are qualitatively and quantitatively similar when considering PB as well (the corresponding figure is included in the Appendix). The correlation between each pair of cell types in the two cases is obtained using the Spearman rank correlation (described in the Methods) and displayed in the form of Cross-correlation matrices **C**, for the cases of cord blood [Fig. 4.1(a)] and adult blood [Fig. 4.1(b)]. We find that relative change between the correlation matrices for cord and adult blood data displayed in Fig. 4.1, which is calculated in terms of the relative change in cross correlation (RC_{ij}), is most significant for the cell
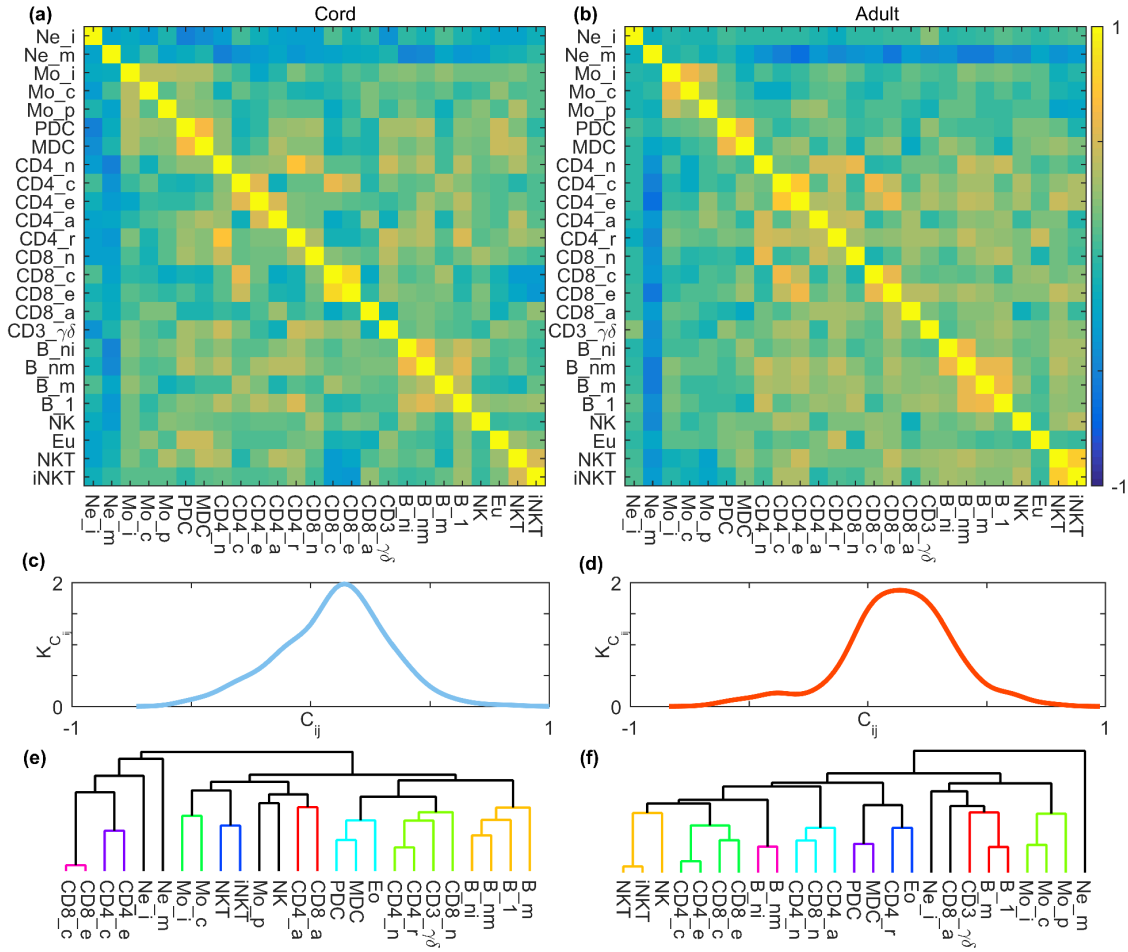
Figure 4.1: Investigating the inter-relation between different immune cell sub-types from cross-correlation of their relative abundance in cord (a,c,e) and adult (b,d,f) blood. (a-b) Cross-correlation matrix $\mathbf{C}$ obtained by measuring the Spearman rank correlation between the relative abundances of the i-th and j-th cell types, viz., $r_i(= n_i/N)$ and $r_j(= n_j/N)$, where $i, j = 1, \ldots, M$. Here M is the total number of distinct cell types being considered), while n_i and N are the cell counts of the i-th cell type and total cell count, respectively. The values of $C_{ij} \in [-1, 1]$ are represented by a color-scale indicated in the colorbar next to panel (b). The 25 individual cell types analyzed are indicated along the rows and columns. Data for an additional cell type, Plasma B cells (PB), have been removed from the data sets for ease of comparison between the two matrices as the cord data contain only negligible quantities of such cells. See the Appendix for a correlation matrix of the adult data that includes PB. (c-d) The probability density functions $P(C_{ij})$ of the correlation matrix elements showing that the distribution of C_{ij} for the adult blood data is more positively skewed ($\gamma_{adult} = 1.06$) than that of the cord blood ($\gamma_{cord} = 0.95$). (e-f) Dendrograms showing the hierarchical clustering between different cell types, constructed from the cross-correlation matrix using the complete linkage method. Distance d_{ij} between cell types i and j have been computed using the expression $d_{ij} = \sqrt{2(1 - C_{ij})}$. Colored branches indicate significant clustering (the threshold for significance corresponding to 70% of the maximum linkage).

95

Table 4.1: Cell types of the immune system

Cell type	Notation
Immature Neutrophil	Ne_i
Mature Neutrophil	Ne_m
Inflammatory Monocyte	Mo_i
Classical Monocyte	Mo_c
Patrolling Monocyte	Mo_p
Plasmacytoid Dendritic Cell	PDC
Myeloid or Conventional Dendritic Cell	MDC
CD4 Naive T Cell	CD4_n
CD4 Central Memory T Cell	CD4_c
CD4 Effector Memory T Cell	CD4_m
CD4 Effector Memory RA+ T Cell	CD4_a
CD4 Regulatory T cell	CD4_r
CD8 Naive T Cell	CD8_n
CD8 Central Memory T Cell	CD8_c
CD8 Effector Memory T Cell	CD8_e
CD8 Effector Memory RA+ T Cell	CD8_a
CD3 $\gamma\delta$ T Cell	CD3_$\gamma\delta$
Immature Naive B Cell	B_ni
Mature Naive B Cell	B_nm
Memory B Cell	B_m
Plasma B Cell	PB
B1 Cell	B_1
Natural Killer Cell	NK
Eosinophil	Eo
NKT Cell	NKT
iNKT Cell	iNKT

types iNKT and CD8_a (the corresponding figure is included in the Appendix). In order to quantify the change in cross-correlation of cell abundance, we consider the probability density functions $P(C_{ij})$ of the correlation matrix elements. As can be seen in Fig. 4.1(c-d) the distribution of C_{ij} is more positively skewed ($\gamma_{adult} = 1.06$) for the case of adult blood than that of cord blood ($\gamma_{cord} = 0.95$). In order to determine the nature of heirarchical clustering between cell types, we construct dendograms from cross-correlation matrix in each case. This is obtained by first converting the cross-correlation into a distance metric $d_{ij} = \sqrt{2(1 - C_{ij})}$ between each pair of cells i and j, and imposing a significance threshold corresponding to 70% of the maximum linkage.
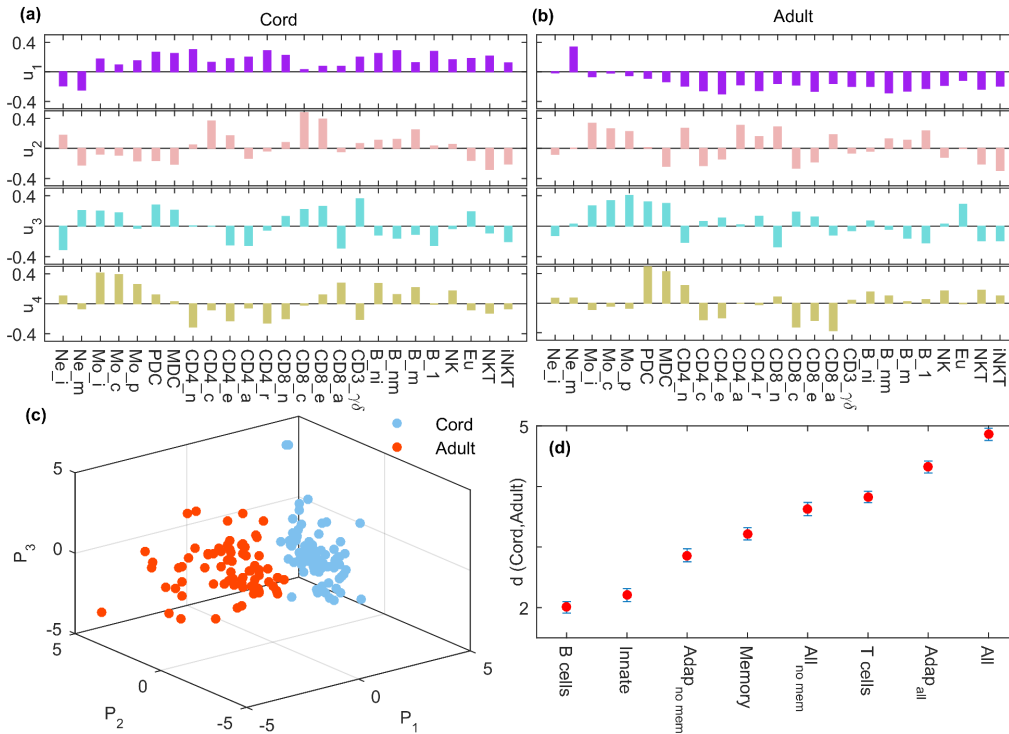
Figure 4.2: Distinguishing the inter-relations between different cell types in cord and adult blood using principal component analysis. (a-b) Eigenvector components corresponding to different cell types (indicated at the base of each panel) for the four largest eigenmodes $u_1 - u_4$ (arranged in decreasing order of the corresponding eigenvalues) of the cross-correlation matrices obtained from cord (a) and adult (b) data. (c) Scatter plot of the individual samples of cord (blue circles) and adult (red circles) blood shown in the subspace formed by the first three principal components. The coordinates of each sample k (= $1, \ldots, Q$, with $Q_{cord} = 75$ and $Q_{adult} = 73$) are given by $P_i(k) = \Sigma_j z_k(j) u_i(j)$ ($j = 1, \ldots, M$ being the index for the different cell types) where z_k is the z-score vector representing the relative abundance of the different cell types in the k-th sample, centered around the mean and scaled by the standard deviation for each cell type. As can be seen, samples of the two classes are clearly separated into two clusters. Note that memory cells (viz., CD4_c, CD4_e, CD8_c, CD8_e and B_m) are not included in the above calculations. (d) Distance between the centroids of the clusters corresponding to cord and adult individual samples calculated using all cell types or subsets thereof (indicated along the abcissae). Note that the panel (c) corresponds to the subset "All$_{no\ mem}$". As expected the distance between the two clusters is reduced when only innate components of the immune system are considered. The error bars are obtained by 10^4 bootstrap samples of each dataset.

In order to study the inter-relations between cell types in cord and adult blood in more detail, we perform a principal component analysis (described in the Methods) in each case. The resulting eigenvector components for the four largest eigenmodes $u_1 - u_4$ are displayed for the case of cord blood [Fig. 4.2 (a)] and adult blood [Fig. 4.2 (b)]. These are displayed in each case for all 25 cell types, with the rows of Fig. 4.2 (a-b) arranged in decreasing order of the corresponding eigenvalues. In order to visualize the quantitative changes between cord and adult blood data, we consider the sub-space formed by the first three principal components. The coordinates of each of the individual samples of the two data sets k ($= 1, \ldots, Q$, with $Q_{cord} = 75$ and $Q_{adult} = 73$) are specified for each cell type $j = 1, \ldots, M$ as $P_i(k) = \Sigma_j z_k(j) u_i(j)$. Here, the $z_k(j)$ is the z-score of the relative abundance of cell type j, and is centered around the mean and scaled by the standard deviation for the corresponding cell type. In order to screen out the role of memory cells in any difference that may arise between cord and adult blood, we do not consider these cells in the aforementioned calculations. As seen in Fig. 4.2 (c), the samples for the cord blood and adult blood form two clusters that are clearly seperated in the chosen sub-space. Thus, despite considering a subset of cells excluding memory cells, we find significant differences between cord and adult blood. In order to further quantify the differences that arise from subsets of cell types, we compute the distance between the centroids of the clusters corresponding to cord blood and adult blood. This quantity is calculated using all cell types and using specific subsets. As seen in Fig. 4.2(d), we find that the distance between the two clusters remains significant for most of the subsets considered.

A detailed view of the spectral decomposition of the cross-correlation between cell types is presented for the case of cord blood [Fig. 4.3(a-b)] and adult blood [Fig. 4.3(c-d)]. This is presented in the form of scatter plots shown along the planes formed by each consecutive pair of the second, third and fourth largest eigenmode, viz. the second and third [Fig. 4.3(a,c)] and the third and fourth [Fig. 4.3(b,d)].

Figure 4.3: Two-dimensional projections of the spectral decomposition of the cross-correlation between different immune cell types for cord (top row) and adult (bottom row) blood. Scatter plots are shown along the planes formed by the second and third largest eigenmodes (a,c), and third and fourth largest eigenmodes (b,d).

4.4   Discussion and Conclusion

Our investigations reported above illustrates the utility of statistical analysis in addressing fundamental questions related to the development of immune response in humans. Although there has been much experimental research done on uncovering the role of specific types of immune cells and how they respond to foreign antigens, our results highlight the information that can be acquired by analyzing cross-sectional data of different immune cells sampled across a population. We find that the correlations between the levels of immune cell types exhibit significant changes from cord to adult blood. This suggests that, barring pathological cases, there is a common signature for the development of the immune system across human populations.

Although there have been earlier studies showing some of the differences in the immune system of adults and neonates, the novelty of our results lie in its focus on the pairwise correlations between different immune cell types in neonates and adults. Earlier studies have mostly focused on the distributions of different cell types, but have not considered interactions between them. Furthermore, while the difference in neonate and adult immune systems was known in terms of absence and presence (respectively) of memory cells, we have shown that the two systems differ even in the correlations between their constituent cell types. Our results can also be used to infer the possible function of certain components of the immune system. For instance, a smaller centroid distance for B cells between the cord and adult data clusters in the space formed by the three leading principal components, suggest that B cells can have a role in innate immune response.

In addition to the insights gained from the cross-correlation matrices, principle component analysis reveals clear distinctions in the characteristics of the cord and adult blood data. Specifically, when considering the three most significant eigenvector components of each data set, we find that cord and adult blood form two distinct clusters. This separation between the clusters is apparent even when considering subsets of the data in each case, indicating that the relative levels of the underlying cell types are characterized by highly

distinct features for the two data sets.

Finally, the methodology employed in our analysis is sufficiently general to be applicable to a variety of other contexts and can be used to investigate additional questions related to changes in the relative levels of immune cell types over time. While there remain important questions related to the genetic and environmental factors affecting the development of the immune system, our study demonstrates that population-level data sets can yield intriguing observations related to the trajectories undertaken by the immune system as it develops.

4.A   Appendix: Supplementary Figures
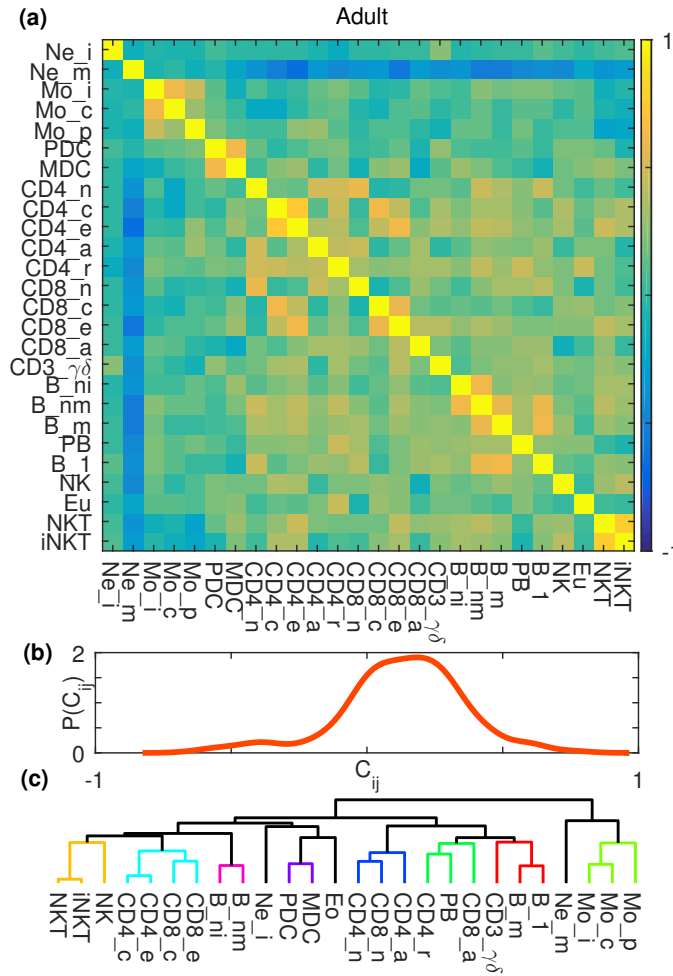
Figure 4.4: Investigating the inter-relation between different immune cell sub-types from cross-correlation of their relative abundance in adult blood with Plasma B cells (PB) included. (a) Cross-correlation matrix $\mathbf{C}$ obtained by measuring the Spearman rank correlation between the relative abundances of the i-th and j-th cell types, viz., $r_i(= n_i/N)$ and $r_j(= n_j/N)$, where $i, j = 1, \ldots, M$. Here M is the total number of distinct cell types being considered), while n_i and N are the cell counts of the i-th cell type and total cell count, respectively. The values of $C_{ij} \in [-1, 1]$ are represented by a color-scale indicated in the colorbar. The 26 individual cell types analyzed are indicated along the rows and columns. (b) The kernel density smoothed estimates $P(C_{ij})$ of the correlation matrix elements, represented using relative frequencies showing the distribution of C_{ij} for the adult blood data. (c) Dendrogram showing the hierarchical clustering between different cell types, constructed from the cross-correlation matrix using the complete linkage method. Distance d_{ij} between cell types i and j have been computed using the expression $d_{ij} = \sqrt{2(1 - C_{ij})}$. Colored branches indicate significant clustering (the threshold for significance corresponding to 70% of the maximum linkage).
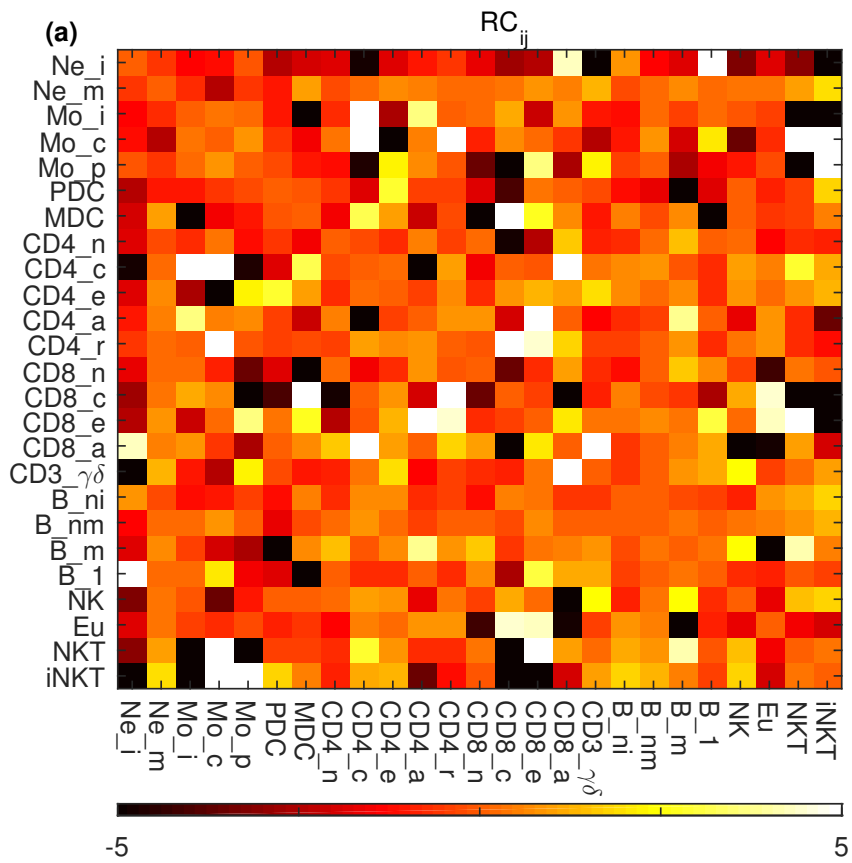
Figure 4.5: Matrix representing the relative change in the values of (a) cross-correlations between cell types between the cord blood and the adult blood.

5

Conclusions

The research described in this thesis is a contribution towards obtaining a better understanding of adaptive dynamics in two paradigmatic examples of complex adaptive systems, viz., the intra-cellular signaling network and the immune system. We observe the emergence of phenomena identified with memory and learning in these complex systems as an outcome of the interactions between different constituent elements and their interplay with environmental stimuli or extrinsic factors. In the following subsections, we summarize the important results and conclusions reported in this thesis. This is followed by a short discussion of possible extensions of the research presented here, as well as, broad contours of how the research field might develop in the future.

5.1   Summary of the Main Results

Emergence Memory In Intra-Cellular Signaling Networks

The mitogen-activated protein kinase (MAPK) signaling cascade, an evolutionarily conserved motif present in all eukaryotic cells, is involved in coordinating crucial cellular functions. While the asymptotic behavior of the pathway stimulated by a time-invariant

signal is relatively well-understood, we show using a computational model that it exhibits a rich repertoire of transient adaptive responses to changes in stimuli. When the signal is switched on, the response is characterized by long-lived modulations in frequency as well as amplitude. On withdrawing the stimulus, the activity decays over long time-scales, exhibiting reverberations characterized by repeated spiking in the activated MAPK concentration. The long-term persistence of such post-stimulus activity suggests that the cascade retains memory of the signal for a significant duration following its removal. The molecular mechanism underlying the reverberatory activity is related to the existence of distinct relaxation rates for the different cascade components. This results in the imbalance of fluxes between different layers of the cascade, with the reuse of activated kinases as enzymes when they are released from sequestration in complexes. The persistent adaptive response, indicative of a cellular short-term memory, suggests that this ubiquitous signaling pathway plays an even more central role in information processing by eukaryotic cells. Determining the appropriate response of a cell to a signal from its environment requires non-trivial information processing by the intra-cellular signaling network. This may be significantly aided by the memory of exposure to the stimulus. Our research shows that the MAPK motif can function as a memory device allowing the effect of a signal to persist even when the cell is no longer exposed to it. This is manifested as reverberatory spiking activity of the cascade, a consequence of the broad range of relaxation time-scales of its components. One may therefore expect to observe results qualitatively similar to what has been reported here whenever the system has disparate time-scales regardless of the actual molecular concentrations and kinetic rates which can vary substantially across different cells. We observe a crossover between two qualitatively distinct regimes of relaxation behavior of active form of MAPK occurring at a steady state that is characterized by relatively low proportion of activation of the available MAPK. Thus, there appears to be an effective threshold for MAPK activity (which may be related to its basal state level) that demarcates the different relaxation regimes following the removal of the applied stimulus. As this signaling motif is involved in crucial cellular functions,

the potential consequences of such emergent memory are far-reaching.

Non-associative Learning in Intra-Cellular Signaling Networks

Nonlinear systems driven by recurrent signals are known to exhibit complex dynamical responses which, in the physiological context, can have important functional consequences. One of the simplest biological systems that is exposed to such repeated stimuli is the intra-cellular signaling network. In this thesis we have investigated the periodic activation of an evolutionarily conserved motif of this network, viz., the mitogen-activated protein kinase (MAPK) signaling cascade, with a train of pulses. The resulting response of the cascade, which shows integrative capability over several successive pulses, is characterized by complex adaptive behavior. These include aspects of non-associative learning, in particular, habituation and sensitization, which are observed in response to high- and low-frequency stimulation, respectively. Learning, wherein repeated exposure to sustained stimulation eventually results in an altered response, is usually associated with multi-cellular organisms, e.g., those possessing a nervous system. However, simpler forms of learning, e.g., sensitization and adaptation, may appear in more rudimentary systems. In this work we report novel results on how the response of a crucial part of the eukaryotic intra-cellular signaling network, namely MAPK cascade can change on receiving a signal repeatedly over an extended period of time. Depending on circumstances, this can result in either an enhanced response (sensitization) or a diminished one (habituation). Taken in conjunction with previous analogies made between the intra-cellular signaling network and nervous systems, this provides an intriguing perspective on how the signaling network of a cell can function as an adaptive complex system. In addition, the existence of a response threshold of the cascade, an apparent refractory behavior following stimulation with short inter-pulse interval, and an alternans like response under certain conditions, suggest an analogy with excitable media.

107

Evolution of the Immune Network

Adaptive Immunity is an integral part of human immune system and is evolves over time as the system is exposed to foreign antigens and distress signals. A new-born child beign not exposed to such signals tend to have very few immune memory cells. Adults, on the other hand, show a variety of memory components in their adaptive immune system because of long exposure to the environment. With age the connections characterizing the functional network of associations between different types of immune cells keeps changeing. This is because, new correlations may be established between memory immune cells and other cell types. By looking at the population of various immune cells in umbilical cord blood samples as well as in adult blood samples, we have uncovered several new functional relationships between the immune cells. In particular, clusters of different cell types have been identified that are related in terms of whether they evolve in a correlated manner. We have characterized the similarity and dissimilarly of various immune components in cord and adult blood, which can enable us to discern possible genetic, environmental and developmental signatures in this dynamic network. A systematic analysis of the empirical data using quantitative tools can help in understanding how exposure to the environment governs the dynamics of the inter-cellular network as the immune system develops. We find that the adult immune system exhibits a higher degree of correlation in the proportions of cells of the adaptive immune system, suggesting a strong role played by maturation in the evolution of the system. Our analysis validates several correlations between different cell types that have been alluded to in the literature and also suggests a few previously unreported relations. We have also quantified the hierarchy of the differences of immune compartments (such as T cells, B cells, innate cells, etc) between umbilical cord and adult data.

5.2 Outlook and Future Direction

Mutations, dysregulations and altered expressions of the proteins embedded in MAPK cascade are often linked with various types of cancer [100, 187, 188], autoimmune disorders [189, 190], as well as, many inflammatory and neurodegenerative diseases [100]. Depending on the cellular context, the MAPK cascade can function both as a pro- and an anti-apoptotic pathway. We note that several candidate drugs for treating cancer have failed by not being able to attenuate the activity of MAPK** [191] when it acts as a modulator for cellular proliferation, tumorigenesis, tumor growth and maintenance. Thus, it is intriguing to consider the possibility that the memory displayed by MAPK** activity can subvert intra-cellular signaling networks to promote survival, proliferation and maintenance of cancer cells. As there is evidence of cross-talk between the MAPK and PI3K-AKT pathways [192], we can speculate that the memory displayed by MAPK pathway upon removal of a stimulus may play an important role in establishing a bi-directional signaling cross-talk between PI3K-AKT and MAPK pathways in several possible micro-environmental scenarios within the cell.

Another key aspect of the long-term memory in cell signaling is its direct impact on the fundamental limits on detectability for sensing extra-cellular chemical concentrations in diverse contexts [193–196]. It is worth noting that the input stimulus to a MAPK cascade is essentially a membrane-bound activated protein working in conjunction with the cell-surface receptors. The landmark studies that have set fundamental limits on sensing extra-cellular chemical concentrations through binding-unbinding events mediated by cell-surface receptors (such as the Burg-Purcell Limit [193] and Maximum Likelihood Estimator [194]) are based on the assumption that binding events are independent of each other. In addition, the effective downstream signaling is assumed to be associated with the corresponding binding event alone. As our work reported in this thesis demonstrates, cell signaling processes may exhibit memory wherein post-stimulus spiking in the effector protein concentrations can occur upon withdrawal of the applied stimulus which was

mediated by a lone binding event. The long memory of the signal retained by the cell may alter the subsequent signaling dynamics following repeated stimulation which is mediated by multiple binding events. Thus, one may expect that the fundamental detectability limits may change because of such cellular memory. As shown in the context of fluctuating environmental concentrations in Ref. [196], memory may enhance the detectability of the ligand concentration of the extra-cellular chemicals. Thus, a potential problem one can work on in the future involves exploring the role of transient reverberatory dynamics in modulating detectability limits for chemicals in the external environment of a cell.

Generally the framework of ordinary differential equations (ODEs) is used to model the dynamics of motifs of the intra-cellular signal transduction networks by deploying mass action principles and Michaelis-Menten kinetics to represent each of the constituent bio-chemical reactions [132]. However, when the number of molecules participating in these reactions are very low, the dynamics becomes intrinsically noisy. Such stochasticity can result in non-trivial qualitative changes in the behaviour of the models of intra-cellular signaling dynamics [197, 198]. Although we have carried out the computational studies of the signal transduction systems in the mean-field limit, describing each reaction by an ODE owing to the presence of quantitively larger copy numbers of the MAPK family proteins inside the cells, presence of such intrinsic noise can impact the response patterns for cells where the copy numbers of the proteins present in the MAPK family are of the order of 10 to 100. An obvious extension of the present work, therefore, will be to ask how stochasticity will affect the adaptive dynamics of MAPK motif when the signal is a short-duration pulse, as well as, when subject to repeated stimulation by a pulse train. This can be done by using, for instance, the Gillespie algorithm [199] or its variants for simulating the reaction dynamics.

In many signaling pathways, the signaling molecules are compartmentalized inside the cell, i.e., inside the cytosol the mobility of participating proteins may be restricted to particular spatial regions inside the cell. Often the geometry of the cell, the chemical

gradients of the chemicals, formation of microdomains, heterogeneity owing to presence of a myriad types of molecules, diverse time-scales for diffusion of the molecules, etc., can contribute to these spatial dependencies in signal transduction [103]. In such cases, we can use either partial differential equations (PDEs) or compartmental models to incorporate the spatial information [200] and therey study the adaptive dynamics underlying cellular memory and non-associative learning in signal transduction networks. The chemical reactions that we use to describe events can often consist of many intermediate steps. Hence, encompassing the intermediate steps becomes a tedious computational job for large scale networks. Even without the intermediate states of the reactions, the analytical understanding is not always easy for a large number of coupled ODEs as in the case of the MAPK motif. The coupling and non-linearity in the interdependence of the state variables makes it very difficult to gain a mechanistic understanding of the associated dynamics. For example, we are yet to formulate a successful reduced phenomenological model of the MAPK system (e.g., involving only two or three dynamical variables) which will enable us to gain a deeper physical understanding of the system. A future aim will be to devise such a reduced model which preserves the qualitative features of the adaptive response, including the emergence of cellular memory and non-associative learning.

Another challenge in modeling signal transduction networks is in the presence of many kinetic parameters which are often difficult to experimentally measure exactly. Indeed, some of the parameters can be 'sloppy' in that varying them over a large range does not signfiicantly affect the qualitative nature of the response [201]. Although using such a 'sloppy' framework may help in deriving a reduced model, another complimentary approach can be to model the signaling cascades by a Boolean logic circuit where level of activity of a node is represented by binary state variables. Historically, Boolean networks encompass many different kinds of regulatory relationships citeKauffman1969. In the case of signal transduction, it has been shown that signaling networks can display retrograde propagation of information resulting in non-trivial outcomes, such as long-range interaction between two different pathways [110, 111]. Thus, one can conceive of a study

involving Boolean networks incorporating bi-directional regulatory interactions as well as sequestration effects.

The research work that is presented in this thesis demonstrating the emergence of memory and learning in a canonical MAPK cascade, a signal transduction motif ubiquitous in all eukaryotic cells, does not incorporate cross-talks with other components of the cell signaling machinery. Also, we do not explicitly implement several explicit negative and positive feedbacks within the MAPK cascade. While it is intriguing to observe complex adaptive behaviour in the MAPK motif in the absence of any explicit feedback or crosstalk, one can ask whether the adaptive dynamics reported here is robust with respect to these additional features. For instance, they can enable the system to operate at different time-scales depending on the system parameters. Given the complexity of such a system, we speculate that more complex dynamics (and possibly novel types of learning) might be observed. This will introduce a new dimension to our understanding of the capabilities of intra-cellular signal processing machinery. Although the regimes in the parametric space over which we observe desensitization, sensitization and alternans are quite limited in the case of a lone canonical MAPK motif, the complexity of cross-talks and feedback might broaden these parametric regimes making these phenomena more robust.

Our results of non-associative learning in MAPK cascade may also provide an explicit mechanistic basis to the excitability paradigm that has been theoretically suggested earlier in irder to explain the thymic selection and peripheral anergy in the context of tunability of activation thresholds during repetitive interactions of T cells with self-antigens [176, 177]. Although the mechanism of negative and positive selection of maturing T cells in Thymus and clonal deletion of B cells in bone-marrow have been investigated earlier it is not yet understood how general these are [18, 23, 24]. Our results show that the emergent dynamics of MAPK cascade, which is known to mediate immune response [178] can provides an explicit mechanistic basis for a theoretical framework to explain the adaptive response of the immune system. It is worth investigating how the maturing T cells

achieve certain cell fates, what role does stochastic binding with APCs carrying different MHCs of type I & II play and how the intra-cellular transduction pathways initiate cell differentiation, proliferation or apoptosis based on periodic or stochastic signals that the maturing T cells are receiving due to TCR binding with MHC ligands. In this thesis, we have demonstrated (in the context of signal transduction) that one can obtain completely different downstream activity over long time-scales if stimulated repeatedly. Tolerance, sensitization and memory can emerge in such systems even without explicit feedback. In some biological systems, for instance, in maturing T cells, this can potentially lead to understanding various cell fates achieved by the maturing cells and can explain the different decisions taken by cells based on their micro-environment. To the best of our knowledge, our results are the first to show that simple motifs, such as the MAPK cascade, in the intra-cellular signaling network can exhibit short-term memory and instances of non-associative learning, that may have biological significance. We would like to note that a recent experimental study [202] provides support for our proposal of the existence of short-term memory in cell signaling.

Moving on to the immune system, there are a wide variety of research directions that can be taken up in the future. For instance, one of the aims could be to understand the process underlying self/nonself discrimination and its relationship with the nature of temporal dynamics of key signal transduction proteins, for instance MAPK proteins. In particular, how does repeated interaction with antigens/self-antigens impart individual immune cells with different outcomes in terms of their function and cell fates. The differentiation of the immune cells into memory cells is yet another problem one can look at. In the case of T cells, the thymic output does not involve any memory T cell [23]. Through antigenic interactions, the T cells form memory compartments and live for extremely long time. How do they achieve such a fate? How do they live for such a long time? How do the regulatory interactions within the immune network contribute to their development? B cells also differentiate into memory compartments following antigenic exposures so that they can mount a stronger response upon subsequent exposure. In this thesis, we have restricted

ourselves to draw an overall picture of the immune network in cord and adult blood. One can ask how exposure to stimuli actually alter the connection structure of the network by considering explicitly cytokine or chemokine mediated molecular interactions take place during the maturation process of the immune system in the course of development. Similar techniques to those we use to differentiate the functional correlations between different immune cell types in the cord blood and adult blood can be used to understand differences in the immune system of diseased individuals.

To summarize, this thesis has tried to characterize, using the framework of dynamical systems theory, the broad features of adaptive dynamics (including memory and learning) displayed by two biological networks that function at very different levels (intra-cellular and inter-cellular). Our results may initiate new research initiatives in the future to look at similar dynamics in other biological systems.

Bibliography

[1] H. Lodish *et al.*, *Molecular Cell Biology* (4th Ed., W. H. Freeman, New York, 1999).

[2] E. R. Kandel *et al.*, *Principles of Neural Science* (5th Ed., McGraw-Hill, New York, 2012).

[3] T. M. Hennessey, W. B. Rucker, and C. G. McDiarmid, Anim. Learn. Behav. **7**, 417 (1979).

[4] R. D. Hawkins and E. R. Kandel, Psychol. Rev. **91**, 375 (1984).

[5] R. F. Thompson, Science **233**, 941 (1986).

[6] B. Brembs *et al.*, Science **296**, 1706 (2002).

[7] A. C. Giles, J. K. Rose, and C. H. Rankin, Int. Rev. Neurobiol. **69**, 37 (2006).

[8] M. H. Olson, *Introduction to theories of learning* (Routledge, 2015).

[9] T. J. Carew, R. D. Hawkins, and E. R. Kandel, Science **219**, 397 (1983).

[10] E. T. Walters, T. J. Carew, and E. R. Kandel, Proc. Natl. Acad. Sci. USA **79**, 6675 (1979).

[11] T. J. Carew, E. T. Walters, and E. R. Kandel, Proc. Natl. Acad. Sci. USA **1**, 1426 (1981).

[12] P. Lyon, Front. Microbiol. **6**, 264 (2015).

[13] R. P. Boisseau, D. Vogel, and A. Dussutour, Proc. R. Soc. B **283**, 1829 (2016).

[14] J. Davies and D. Davies, Microbiol. Mol. Biol. Rev. **74**, 417 (2010).

[15] B. D. Burrell and C. L. Sahley, Curr. Opin. Neurobiol. **11**, 757 (2001).

[16] C. H. Rankin *et al.*, Neurobiol. Learn. Mem. **92**, 135 (2009).

[17] M. A. Cooper, J. M. Elliott, P. A. Keyel, L. Yang, J. A. Carrero, and W. M. Yokoyama, Proc. Natl. Acad. Sci. USA **106**, 1915 (2009).

[18] K. A. Hogquist, T. A. Baldwin, and S. C. Jameson, Nat. Rev. Immunol. **5**, 772 (2005).

[19] G. W. Litman, J. P. Rast, and S. D. Fugmann, Nat. Rev. Immunol. **10**, 543 (2010).

[20] M. McFall-Ngai, Nature (Lond.). **445**, 153 (2007).

[21] Z. Pancer, and M. D. Cooper, Annu. Rev. Immunol. **24**, 497 (2006).

[22] R. Medzhitov, Nature (Lond.), **449**, 819 (2007).

[23] D. D. Chaplin, J. Allergy Clin. Immunol. **125**, S3 (2010).

[24] D. L. Mueller, Nat. Immunol. **11**, 21 (2010).

[25] A. K. Simon, G. A. Hollander, and A. McMichael, Proc. R. Soc. B **282**, 20143085 (2015).

[26] G. P. Chrousos, Nat. Rev. Endocrinol. **5**, 374 (2009).

[27] C. Koch, and G. Laurent, Science **284**, 96 (1999).

[28] J. H. Holland, Daedalus **121**, 17 (1992).

[29] J. Selimkhanov, B. Taylor, J. Yao, A. Pilko, J. Albeck, A. Hoffmann, L. Tsimring, and R. Wollman, Science **346**, 1370 (2014).

[30] A. L. Barabási and Z. N. Oltvai, Nat. Rev. Genet. **5**, 101 (2004).

[31] D. Bray, Science **301**, 1864 (2003).

[32] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[33] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969).

[34] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, Nature (Lond.) **407**, 651 (2000).

[35] G. Weng, U. S. Bhalla, and R. Iyengar, Science **284**, 92 (1999).

[36] R. G. Smock and L. M. Gierasch, Science **324**, 198 (2009).

[37] U. Alon, *An introduction to systems biology: design principles of biological circuits*, (Chapman and Hall/CRC, 2006).

[38] G. Karlebach and R. Shamir, Nat. Rev. Mol. Cell Biol. **9**, 770 (2008).

[39] T. Hunter, Cell **80**, 225 (1995).

[40] G. Karimova, J. Pidoux, A. Ullmann, and D. Ladant, Proc. Natl. Acad. Sci. USA **95**, 5752 (1998).

[41] J. J. Tyson, K. Chen, and B. Novak, Nat. Rev. Mol. Cell Biol. **2**, 908 (2001).

[42] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell, 6th Ed.* (Garland Science, New York, 2014).

[43] S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber, Phys. Rev. Lett. **94**, 128701 (2005).

[44] T. Pawson and J. D. Scott, Science **278**, 2075 (1997).

[45] B. H. Junker and F. Schreiber, *Analysis of biological networks* (John Wiley and Sons, 2011).

[46] G. M. Edelman and J. A. Gally, Proc. Natl. Acad. Sci. USA **98**, 13763 (2001).

[47] Z. N. Oltvai and A. L. Barabási, Science **298**, 763 (2002).

[48] W. R. Loewenstein, Biochim. Biophys. Acta **560**, 1 (1979).

[49] R. Bruzzone, T. W. White, and D. L. Paul, Eur. J. Biochem. **238**, 1 (1996).

[50] A. Wodarz and R. Nusse, Annu. Rev. Cell Dev. Biol. **14**, 59 (1998).

[51] M. Freeman, Nature (Lond.) **408**, 313 (2000).

[52] M. Demarque, A. Represa, H. Becq, I. Khalilov, Y. Ben-Ari, and L. Aniksztejn, Neuron, **36**, 1051 (2002).

[53] M. Simons and G. Raposo, Curr. Opin. Cell Biol. **21**, 575 (2009).

[54] L. Marzo, K. Gousset, and C. Zurzolo, Front. Physiol. **3**, 72 (2012).

[55] N. M. Kumar and N. B. Gilula, Cell **84**, 381 (1996).

[56] J. Masson, C. Sagne, M. E. Hamon, and S. El Mestikawy, Pharmacol. Rev. **51**, 439 (1999).

[57] D. S. Eom and D. M. Parichy, Science **355**, 1317 (2017).

[58] R. Dantzer, Physiol. Rev. **98**, 477 (2017).

[59] S. W. Brubaker, K. S. Bonham, I. Zanoni, and J. C. Kagan, Annu. Rev. Immunol. **33**, 257 (2015).

[60] D. Artis and H. Spits, Nature (Lond.) **517**, 293 (2015).

[61] K. Buchmann, Front. Immunol. **5**, 459 (2014).

[62] A. C. Kushalappa, K. N. Yogendra, and S. Karre, Crit. Rev. Plant Sci. **35**, 38 (2016).

[63] A. Iwasaki and R. Medzhitov, Nat. Immunol. **16**, 343 (2015).

[64] N. S. Merle, S. E. Church, V. Fremeaux-Bacchi, and L. T. Roumenina, Front. Immunol. **6**, 262 (2015).

[65] N. S. Merle, R. Noe, L. Halbwachs-Mecarelli, V. Fremeaux-Bacchi, and L. T. Roumenina, Front. Immunol. **6**, 257 (2015).

[66] T. van der Poll and H. Herwald, Thromb. Haemostasis **112**, 640 (2014).

[67] J. K. Ryu, M. A. Petersen, S. G. Murray, K. M. Baeten, A. Meyer-Franke, J. P. Chan, E. Vagena, C. Bedard, M. R. Machado, P. E. R. Coronado, T. Prod'homme, I. F. Charo, H. Lassmann, J. L. Degen, S. S. Zamvil, and K. Akassoglou, Nat. Commun. **6**, 8164 (2015).

[68] H. J. Ahn, J. L. Chen, D. Zamolodchikov, E. H. Norris, and S. Strickland, Curr. Opin. Hematol. **24**, 427 (2017).

[69] L. Tort, J. C. Balasch, and S. Mackenzie, Inmunol. **22**, 277 (2003).

[70] T. Matsunaga, Cytogen. Genome Res. **80**, 138 (1988).

[71] M. Kasahara, T. Suzuki, and L. Du Pasquier, Trends Immunol. **25**, 105 (2004).

[72] P. Matzinger, Science **296**, 301 (2002).

[73] M. J. Shlomchik, J. E. Craft, and M. J. Mamula, Nat.Rev. Immunology **1**, 147 (2001).

[74] A. Galy, M. Travis, D. Cen, and B. Chen, Immunity **3**, 459 (1995).

[75] F. Sablitzky, G. Wildner, and K. Rajewsky, EMBO J. **4**, 345 (1985).

[76] S. Tonegawa, Nature (Lond.) **302**, 575 (1983).

[77] R. M. Zinkernagel, M. F. Bachmann, T. M. Kündig, S. Oehen, H. Pirchet, H. Hengartner, Annu. Rev. Immunol. **14**, 333 (1996).

[78] R. Ahmed and D. Gray, Science **272**, 54 (1996).

[79] M. G. Netea, Eur. J. Clin. Invest. **43**, 881 (2013).

[80] R. Dantzer and K. W. Kelley, Life Sci. **44**, 1995 (1989).

[81] T. M. Yi, Y. Huang, M. I. Simon, and J. Doyle, Proc. Nat. Acad. Sci. USA **97**, 4649 (2000).

[82] B. A. Mello and Y. Tu, Biophys. J. **84**, 2943 (2003).

[83] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang, Cell **138**, 760 (2009).

[84] M. Behar, N. Hao, H. G. Dohlman, and T. C. Elston, Biophys. J. **93**, 806 (2007).

[85] M. Inoue and K. Kaneko, Phys. Rev. E **74**, 011903 (2006).

[86] L. Goentoro, O. Shoval, M. W. Kirschner, and U. Alon, Mol. Cell **36**, 894 (2009).

[87] J. E. Ferrell, Mol. Cell **36**, 724 (2009).

[88] V. Sourjik and N. S. Wingreen, Curr. Opin. Cell Biol. **24**, 262 (2012).

[89] M. Inoue, K. Kaneko, Phys. Rev. Lett. **107**, 048301 (2011).

[90] Z. Fu, E. R. Gilbert, and D. Liu, Curr. Diabetes Rev. **9**, 25 (2013).

[91] T. Gross and H. Sayama, *Adaptive networks*, (Springer, Berlin, Heidelberg, 2009).

[92] C. Meisel and T. Gross, Phys. Rev. E **80**, 061917 (2009).

[93] F. Sorrentino and E. Ott, Phys. Rev. Lett. **100**, 114101 (2008).

[94] R. Gutiérrez, A. Amann, S. Assenza, J. Gómez-Gardenes, V. Latora, and S. Boccaletti, Phys. Rev. Lett. **107**, 234103 (2011).

[95] M. Zanin, D. Papo, I. Sendina-Nadal, and S. Boccaletti, Phys. Rev. E **84**, 060102 (2011).

[96] B. J. Berry, L. D. Kiel, and E. Elliott, Proc. Nat. Acad. Sci. USA **99**, 7187 (2002).

[97] D. R. Chialvo, Nat. Phys. **6**, 744 (2010).

[98] C. Widmann, S. Gibson, M. B. Jarpe, and G. L. Johnson, Physiol. Rev. **79**, 143 (1999).

[99] R. Seger and E. G. Krebs, FASEB J. **9**, 726 (1995).

[100] E. K. Kim and E. J. Choi, Biochim. Biophys. Acta **1802**, 396 (2010).

[101] U. S. Bhalla and R. Iyengar, Science **283**, 381 (1999).

[102] J. E. Purvis and G. Lahav, Cell **152**, 945 (2013).

[103] B. N. Kholodenko, Nat. Rev. Mol. Cell Bio. **7**, 165 (2006).

[104] A. J. Waskiewicz and J. A. Cooper, Curr. Opin. Cell Biol. **7**, 798 (1995).

[105] B. Su and M. Karin, Curr. Opin. Immunol. **8**, 402 (1996).

[106] C. Y. Huang and J. E. Ferrell, Proc. Natl. Acad. Sci. USA **93**, 10078 (1996).

[107] H. Shankaran, D. L. Ippolito, W. B. Chrisler, H. Resat, N. Bollinger, L. K. Opresko, and H. S. Wiley, Mol. Sys. Biol. **5**, 332 (2009).

[108] L. Qiao, R. B. Nachbar, I. G. Kevrekidis, and S. Y. Shvartsman, PLOS Comput. Biol. **3**, e184 (2007).

[109] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag, Mol. Syst. Biol. **4**, 161 (2008).

[110] A. C. Ventura, J. A. Sepulchre, and S. D. Merajver, PLOS Comput. Biol. **4**, e1000041 (2008).

[111] T. Jesan, U. Sarma, S. Halder, B. Saha, and S. Sinha, PLOS One **8**, e64409 (2013).

[112] B. N Kholodenko, Eur. J. Biochem. **267**, 1583 (2000).

[113] N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, J. Cell Biol. **164**, 353 (2004).

[114] S. Legewie, B. Schoeberl, N. Blüthgen, and H. Herzel, Biophys. J. **93**, 2279 (2007).

[115] F. Ortega, J. L. Garcés, F. Mas, B. N. Kholodenko, and M. Cascante, FEBS J. **273**, 3915 (2006).

[116] J. E. Ferrell, Curr. Opin. Cell Biol. **14**, 140 (2002).

[117] M. C. Inniss and P. A. Silver, Curr. Biol. **23**, R812 (2013).

[118] W. Xiong and J. E. Ferrell, Nature (Lond.) **426**, 460 (2003).

[119] C. M. Ajo-Franklin, D. A. Drubin, J. A. Eskin, E. P. Gee, D. Landgraf, I. Phillips, and P. A. Silver, Gene. Dev. **21**, 2271 (2007).

[120] D. R. Burrill and P. A. Silver, Cell **140**, 13 (2010).

[121] J. Gunawardena, Proc. Natl. Acad. Sci. USA **102**, 14617 (2005).

[122] O. Hadac, I. Schreiber, and M. Pribyl, J. Chem. Phys. **138**, 065102 (2013).

[123] G. Y. Wu, K. Deisseroth, and R. W. Tsien, Nat. Neurosci. **4**, 151 (2001).

[124] J. D. Sweatt, J. Neurochem. **76**, 1 (2001).

[125] C. Dong, R. J. Davis, and R. A. Flavell, Ann. Rev. Immunol. **20**, 55 (2002).

[126] C. Huang, K. Jacobson, and M. D. Schaller, J. Cell Sci. **117**, 4619 (2004).

[127] G. Altan-Bonnet and R. N. Germain, PLOS Biol. **3**, e356 (2005).

[128] P. Smolen, D. A. Baxter, and J. H. Byrne, Am. J. Physiol. Cell Physiol. **294**, C503-515 (2008).

[129] G. T. Philips, X. Ye, A. M. Kopec, and T. J. Carew, J. Neurosci. **33**, 7565 (2013).

[130] H. Ryu *et al.* Mol. Syst. Biol. **11**, 838 (2015).

[131] M. Kochanczyk *et al.*, Sci. Rep. **7**, 38244 (2017).

[132] T. Millat, E. Bullinger, J. Rohwer, and O. Wolkenhauer, Math. Biosci. **207**, 40 (2007).

[133] U. Alon, M. G. Surette, N. Barkai, and S. Leibler, Nature (Lond.) **397**, 168 (1999).

[134] J. P. Sethna, *Statistical Mechanics: Entropy, Order Parameters and Complexity* (Oxford University Press, Oxford, 2006).

[135] S. Filippi, C. P. Barnes, P. D. Kirk, T. Kudo, K. Kunida, S. S. McMahon, T. Tsuchiya, T. Wada, S. Kuroda, and M. P. Stumpf, Cell Rep. **15**, 2524 (2016).

[136] M. Jeschke, S. Baumgärtner, and S. Legewie, PLOS Comput. Biol. **9**, e1003357 (2013).

[137] S. Mukherjee, S. C. Seok, V. J. Vieland, and J. Das, Proc. Natl. Acad. Sci. USA **110**, 18531 (2013).

[138] C. J. Marshall, Cell **80**, 179 (1995).

[139] A. Plotnikov, E. Zehorai, S. Procaccia, and R. Seger, BBA-Mol. Cell Res. **1813**, 1619 (2011).

[140] J. G. Albeck, G. B. Mills, and J. S. Brugge, Mol. Cell **49**, 249 (2013).

[141] Q. Lu, M. Paredes, J. Zhang, and K. S. Kosik, Mol. Cell. Biol. **18**, 3257 (1998).

[142] K. Nienałtowski, M. Włodarczyk, T. Lipniacki, and M. Komorowski, BMC Syst. Biol. **9**, 65 (2015).

[143] X. Ye, J. L. Shobe, S. K. Sharma, A. Marina, and T. J. Carew, Proc. Natl. Acad. Sci. USA **105**, 20511 (2008).

[144] C. M. Atkins, J. C. Selcher, J. J. Petraitis, J. M. Trzaskos, and J. D. Sweatt, Nat. Neurosci. **1**, 602 (1998).

[145] M. Skoge, H. Yue, M. Erickstad, A. Bae, H. Levine, A. Groisman, W. F. Loomis, and W. J. Rappel, Proc. Natl. Acad. Sci. USA **111**, 14448 (2014)

[146] H. V. Prentice-Mott, Y. Meroz, A. Carlson, M. A. Levine, M. W. Davidson, D. Irimia, G. T. Charras, L. Mahadevan, and J. V. Shah, Proc. Natl. Acad. Sci. USA **113**, 1267 (2016).

[147] X. Guo, L. Wang, B. Chen, Q. Li, J. Wang, M. Zhao, W. Wu, P. Zhu, X. Huang, and Q. Huang, Am. J. Physiol.-Heart C. **297**, H238 (2009).

[148] M. Koss, G. R. Pfeiffer, Y. Wang, S. T. Thomas, M. Yerukhimovich, W. A. Gaarde, C. M. Doerschuk, and Q. Wang, J. Immunol. **176**, 1218 (2006).

[149] J. Testa, J. Pérez, and C. Jeffries, Phys. Rev. Lett. **48**, 714 (1982).

[150] L. Glass, A. L. Goldberger, M. Courtemanche, and A. Shrier, Proc. R. Soc. Lond. A **413**, 9 (1987).

[151] M. C. Cross and P. C. Hohenberg, Rev. Mod. Phys. **65**, 851 (1993).

[152] A. L. Lin *et al.*, Phys. Rev. Lett. **84**, 4240 (2000).

[153] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, 2003).

[154] M. I. Rabinovich, P. Varona, A. I. Selverston and H. D. Abarbanel, Rev. Mod. Phys. **78**, 1213 (2006).

[155] J. Aschoff, in *Biological Rhythms*, edited by J. Aschoff (Springer, Boston, MA, 1981).

[156] L. Glass, Nature (Lond.) **410**, 277 (2001).

[157] D. E. Nelson *et al.*, Science **306**, 704 (2004).

[158] D. A. Stavreva *et al.*, Nat. Cell Biol. **11**, 1093 (2009).

[159] R. Heinrich and S. Schuster, *The Regulation of Cellular Systems* (Springer, Boston, MA, 2012).

[160] E. Van Cauter, Horm. Res. Paediat. **34**, 45 (1990).

[161] D. Lloyd and M. Stupfel, Biol. Rev. **66**, 275 (1991).

[162] U. Schibler and F. Naef, Curr. Opin. Cell Biol. **17**, 223 (2005).

[163] K. P. Wright Jr. *et al.*, Curr. Biol. **23**, 1554 (2013).

[164] B. M. Sherman and S. G. Korenman, J. Clin. Investig. **55**, 699 (1975).

[165] R. E. Dolmetsch, K. Xu, and R. S. Lewis, Nature (Lond.) **392**, 933 (1998).

[166] Y. Li and A. Goldbeter, Biophys. J. **61**, 161 (1992).

[167] G. Leng (ed.), *Pulsatility in Neuroendocrine Systems*, (CRC Press, Boca Raton, FL, 1988)

[168] M. J. Berridge, P. H. Cobbold, and K. S. R. Cuthbertson, Phil. Trans. R. Soc. Lond. B **320**, 325 (1988).

[169] R. W. Tsien and R. Y. Tsien, Annu. Rev. Cell Biol. **6**, 715 (1990).

[170] O. Dyachok, Y. Isakov, J. Sågetorp, and A. Tengholm, Nature (Lond.) **439**, 349 (2006).

[171] L. N. Borodinsky and N. C. Spitzer, Sci. STKE **2006**, pe22 (2006).

[172] T. Mitra, S. N. Menon, and S. Sinha, Sci. Rep. **8**, 13230 (2018).

[173] B. Echebarria and A. Karma, Phys. Rev. Lett. **88**, 208101 (2002).

[174] J. I. Goldhaber *et al.*, Circ. Res. **96**, 459 (2005).

[175] S. Sinha and S. Sridhar, *Patterns in Excitable Media* (CRC Press, Boca Raton, FL, 2015).

[176] Z. Grossman and W. E. Paul, Proc. Natl. Acad. Sci. USA **89**, 10365 (1992).

[177] Z. Grossman and A. Singer, Proc. Natl. Acad. Sci. USA **93**, 14747 (1996).

[178] C. Dong, R. J. Davis, and R. A. Flavell, Annu. Rev. Immunol. **20**, 55 (2002).

[179] J. E. R. Staddon, *Adaptive Behavior and Learning*, 2nd edition (Cambridge Univ. Press, Cambridge, 2016).

[180] N. L. Smith, E. Wissink, J. Wang, J. F. Pinello, M. P. Davenport, A. Grimson, and B. D. Rudd, J. Immunol. **193**, 177 (2014).

[181] K. Hoebe, E. Janssen, and B. Beutler, Nat. Immunol. **5**, 971 (2004).

[182] S. B. Prabhu *et al.*, PLoS One **11**, e0162242 (2016)

[183] C. Spearman, Am. J. Physiol. **15**, 72 (1904).

[184] K. Pearson, Proc. R. Soc. London **58**, 240 (1985).

[185] K. Pearson, Philos. Mag. **2**, 559 (1901).

[186] J. N. Franklin, *Matrix Theory* (Dover Publications, 1968).

[187] E. F. Wagner and A. R. Nebreda, Nat. Rev. Cancer **9**, 37 (2009).

[188] A. S. Dhillon, S. Hagan, O. Rath, and W. Kolch, Oncogene **26**, 3279 (2007).

[189] D. N. Krementsov, T. M. Thornton, C. Teuscher, and M. Rincon, Mol. Cell. Biol. **33**, 3728 (2013).

[190] G. Gorelik and B. Richardson, Autoimmun. **43**, 17 (2010).

[191] K. H. T. Paraiso, I. V. Fedorenko, L. P. Cantini, A. C. Munko, M. Hall, V. K. Sondak, J. L. Messina, K. T. Flaherty, and K. S. M. Smalley, Brit. J. Cancer **102**, 1724 (2010).

[192] E. Aksamitiene, A. Kiyatkin, and B. N. Kholodenko, Biochem. Soc. T. **40**, 139 (2012).

[193] H. C. Berg and E. M. Purcell, Biophys. J. **20**, 193 (1977).

[194] R. G. Endres and N. S. Wingreen, Phys. Rev. Lett. **103**, 158101 (2009).

[195] C. C. Govern and P. R. ten Wolde, Phys. Rev. Lett. **109**, 218103 (2012).

[196] G. Aquino, L. Tweedy, D. Heinrich, and R. G. Endres, Sci. Rep. **4**, 5688 (2014).

[197] D. B. Forger and C. S. Peskin, Proc. Natl. Acad. Sci. USA **102**, 321 (2005).

[198] J. Das, M. Ho, J. Zikherman, C. Govern, M. Yang, A. Weiss, A. K. Chakraborty, and J. P. Roose, Cell **136**, 337 (2009).

[199] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[200] S. R. Neves and R. Iyengar, J. Biol. Chem. **284**, 5445 (2009).

[201] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, PLoS Comput. Biol. **3**, e189 (2007).

[202] P. C. Spinosa, B. A. Humphries, D. L. Mejia, J. M. Buschhaus, J. J. Linderman, G. D. Luker, and K. E. Luker, Sci. Signal. **12**, eaaw4204 (2019).