UNIVERSITY OF CALIFORNIA SAN DIEGO

Multiscale Spatiotemporal Probabilistic Graph Models for Neuropsychiatry Applications: Scaling Theoretical Frameworks to Data-Driven Diagnostics from Molecules to Minds

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Bioengineering with a Specialization in Computational Neuroscience

by

Margot Wagner

Committee in charge:

Professor Gert Cauwenberghs, Co-Chair
Professor Terrence J. Sejnowski, Co-Chair
Professor Terry Jernigan
Professor Andrew McCulloch
Professor Benjamin Smarr

2023

The Dissertation of Margot Wagner is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Gert Cauwenberghs as a chair of my committee. Gert has supported me through the more difficult emotional times of this experience while always maintaining an infectiously positive outlook. He is brilliant, an advocate, and a deeply caring person, who always aided me with honesty and kindness. When requested, he sat up late working with me, even on weekends. Gert is a man with a heart of gold.

I would also like to thank my committee chair Terry Sejnowski for his support. Terry has a true gift for inspiring people, regardless of age or background and often with a smile and joke. He is a giant in his field for a good reason. Working with and learning from him has been an amazing opportunity that I will not soon forget. Terry demonstrates daily the importance of communication and charisma in addition to genius to truly make an impact. Together, these two make an invaluable team. I am truly lucky to stand on the shoulders of giants.

In addition to my chairs, I thank my committee and Tom Bartol for their insight, time, and energy. There are many that I worked with that helped and made the journey a much more enjoyable one. All members of both the Sejnowski and Cauwenberghs labs have been exceptionally wonderful during my time. Tom Bartol made me feel at home in the CNL with the days we sat side-by-side working through MCell, him regaling me with stories and a gleam in his eye. Furthermore, Alessandra Camassa has been amazing and continuous support on the ABCD work, as well the rest of the ABCD team.

Lastly, I would like to acknowledge my family and friends, without whom I truly could not have made it here. My mother sacrificed so much to immigrate to the U.S. and provide me with a competitive educational background. My father has been an amazing sounding board and escape from the difficulties of research during a pandemic. He has also been an inspiration as a living example of the importance of this work in changing lives. My friends and partner have provided me with unwavering support and care during this time as well. I cannot thank them enough.

It truly takes a village.

Chapter 2, in full, is a reprint of the material as it appears in Frontiers in Neuroscience 2021. Wagner, Margot; Bartol, Thomas M.; Sejnowski, Terrence J.; Cauwenberghs, Gert., *Markov Chain Abstractions of Electrochemical Reaction-Diffusion in Synaptic Transmission for Neuromorphic Computing*, Frontiers, 2021. The dissertation author was the primary investigator and author of this paper.

Chapter 3 is coauthored with Wagner, Margot; Uppal, Abhinav; Rognes, Marie E.; Bartol, Thomas M.; Sejnowski, Terrence J.; and Cauwenberghs, Gert. The dissertation author was the primary author of this chapter.

Chapter 4 is coauthored with Wagner, Margot; Camassa, Alessandra; Liu, Brandon; Gano, Kameron; Cauwenberghs, Gert; and Sejnowski, Terrence J. The dissertation author was the primary author of this chapter.

Chapter 5, in full, is currently being prepared for submission for publication of material. Wagner, Margot; Jagayat, Jasleen; Kumar, Anchan; Shirazi, Amir; Alavi, Nazanin; Omrani, Mohsen. The dissertation author was the primary investigator and author of this paper.

VITA

2018	Bachelor of Chemical and Biomolecular Engineering, University of Delaware
2020	Master of Science, University of California San Diego
2023	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

"Markov Chain Abstractions of Electrochemical Reaction-Diffusion in Synaptic Transmission for Neuromorphic Computing" Front. Neurosci., vol. 15, November 2021

"EEG source derived salience network coupling supports real-world attention switching" Neuropsychologia., vol. 178, January 2023

FIELDS OF STUDY

Major Field: Bioengineering (Computational Neuroscience)

Studies in Computer Science
Professors Niema Moshiri and Sanjoy Dasgupta

Studies in Artificial Intelligence
Professors Lawrence Saul and Gary Cottrell

Studies in Data Science
Professors Arun Kumar and Julian McAuley

ABSTRACT OF THE DISSERTATION

Multiscale Spatiotemporal Probabilistic Graph Models for Neuropsychiatry Applications: Scaling Theoretical Frameworks to Data-Driven Diagnostics from Molecules to Minds

by

Margot Wagner

Doctor of Philosophy in Bioengineering with a Specialization in Computational Neuroscience

University of California San Diego, 2023

Professor Gert Cauwenberghs, Co-Chair
Professor Terrence J. Sejnowski, Co-Chair

Neuropsychiatric disorders, specifically mental disorders, generally lack a quantitative and biophysiological basis for diagnostics and treatment due to fundamental limitations in theoretical knowledge of the disorder and brain-mind duality. Due to the complex multiscale nature of the brain, large and multimodal datasets as well as biophysically-based simulations are required to elucidate its functioning. As such, computational methods exist to bridge these scales. Specifically, probabilistic graph models are utilized here to capture inherent uncertainty in the system while being computationally efficient and tractable to allow for scaling and combining biophysically and theoretical-based models with Big Data in order to model and diagnose known

mental disorders. First, a theoretically-driven computationally efficient reaction-diffusion model of synaptic transmission using Markov models and eigenmode decomposition is presented to scale molecular simulations to neural networks with applications in pharmacological simulations, artificial neural networks, and neuromorphic engineering. The second part connects the network level to behavior using deep learning, graph models, and manifold learning applied to neuroimaging data in adolescent depression using a combination of theory- and data-driven techniques. In addition to creating scalable models, this work interrogates structural and functional biomarkers and creates a neuroimaging pipeline resulting in automatic disorder detection. Finally, the focus shifts to diagnostics of anxiety and depression using behavioral data in the form of natural language processing, making use of transformer deep learning architectures.

Chapter 1

Introduction

1.1 The Problem: Mental Disorders

Mental disorders are pervasive with serious negative consequences. Nearly one in four adults in the United States have a mental illness at any given time. In 2021, there were 2.6 million suicide attempts in the US, and costs associated with mental disorders are in the trillions worldwide. Additionally, those with any mental disorder are significantly more likely to have substance abuse issues (Fig. 1.1) [21].

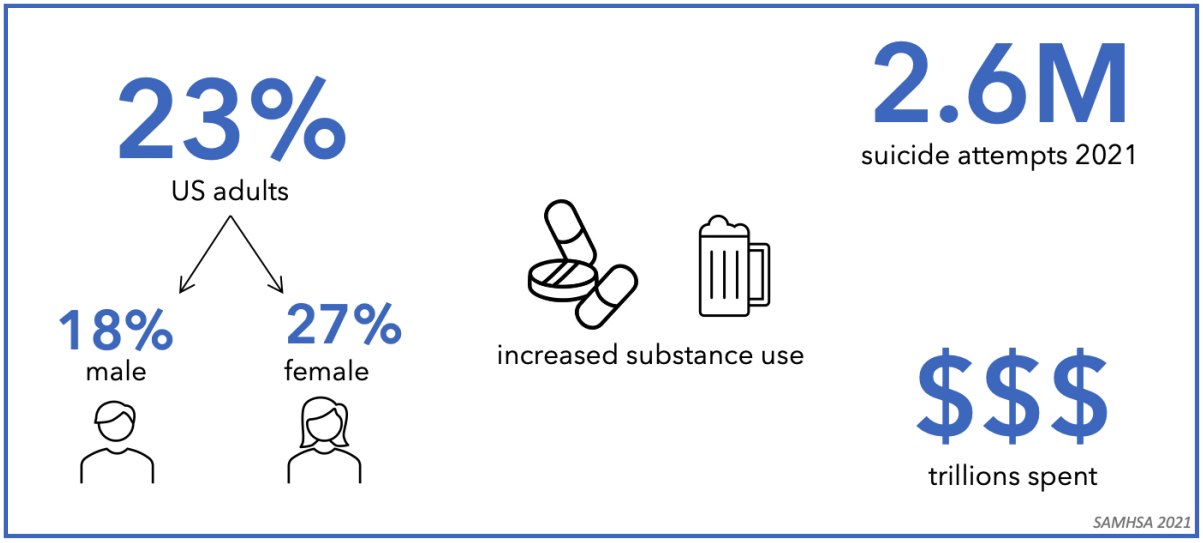

Figure 1.1. Pervasiveness of mental disorders in the United States, as obtained by the 2021 National Survey of Drug Use and Health (NSDUH) by the Substance Abuse and Mental Health Services Administration (SAMHSA) [21].

1.1.1 The Pervasive Impacts of Mental Illness

Mental illness stands as an undeniable issue in contemporary society, compounded by its inextricable ties to substance abuse. A comprehensive study conducted in 2010 revealed disconcerting estimates of over four million excess deaths attributed to mental disorders during that year [52]. Tragically, suicide accounts for over one in a hundred deaths [10], with a staggering projection of 703,000 suicides documented in 2019 globally [13]. Furthermore, for every suicide death, there are approximately 20 attempted suicides [5]. Suicide is the third and fourth leading cause of death among young individuals aged 15 to 29, females and males respectively. These disorders are also a primary contributor to global disability, responsible for one-sixth of all years lived with disability. Combined with substance use disorders and neurological conditions, it becomes a quarter. Depressive disorders alone are the second leading cause of disability after back and neck pain [10]. The impact of severe mental disorders, such as schizophrenia and bipolar disorder, is particularly disheartening, curtailing the average lifespan by 10 to 20 years [54]. Moreover, mental disorders significantly contribute to the occurrence of preventable noncommunicable diseases, most notably cardiovascular and respiratory diseases, as well as infections [7]. The global scale of suffering due to mental health issues is incontestable.

Beyond the profound social implications, mental disorders carry significant economic ramifications. As elucidated by The World Economic Forum, these conditions accounted for a staggering $2.5 trillion loss to the global economy in 2010. This encompassed $1.7 trillion due to diminished productivity and an additional $0.8 trillion incurred as direct costs of care. Projections estimate that by 2030, this economic burden will escalate to $6 trillion [38].

The ubiquity of mental disorders is deeply concerning. In 2019, the World Health Organization (WHO) estimated that approximately 1 in every 8 people worldwide (970 million people) were actively grappling with a mental disorder. Among these conditions, anxiety and depression were identified as the most prevalent afflictions, impacting an estimated 301 million and 280 million individuals, respectively [8]. These numbers have only worsened in the aftermath

of the COVID-19 pandemic, with initial estimates indicating a more than 25% increase in anxiety and depression cases [61]. Notably, women are more vulnerable to mental disorders, with a prevalence rate 50% higher than that of men, although men demonstrate a greater propensity for substance use disorders. [8]. While the peak burden is typically observed during early adulthood, mental disorders affect individuals across all age groups [10]. For instance, 1 in 7 individuals aged 10 to 19 has endured a mental disorder at some point in their lifetime [8].

These trends exist in all countries, regardless of their income level. In fact, mental disorders exhibit an increased prevalence in high-income countries (15.1% compared to 11.6%) [9] . However, this discrepancy is predominantly ascribed to differing attitudes surrounding mental health which lead to underreporting [120]. In the US, a recent survey conducted by SAMHSA revealed that more than one in five adults (22.7% or 57.8 million individuals) experienced mental illness in 2021. Again, females were affected at a higher prevalence compared to males (27.2% vs 18.1%),with the highest rates observed among young adults aged 18 to 25 (33.7%), followed by adults aged 26 to 49 (28.1%), and finally adults over 50 (15.0%). Perturbingly, out of the 57.8 million affected individuals, only 47.2% received any form of mental health services. Additionally, the survey also found that nearly half of US adolescents (49.5%) had encountered a mental disorder in their lifetimes, specifically within the age range of 13 to 18 [21].

It is worth acknowledging that the aforementioned study did not encompass individuals without a fixed address (e.g. homeless individuals), those serving in active military duty, or those residing in institutional group quarters (e.g. correctional facilities, nursing homes, mental institutions, long-term hospitals). These particular groups often exhibit elevated rates of mental illness, suggesting an overall underestimation of the prevalence [21].

Notwithstanding these challenges, there exists a burgeoning recognition of the pivotal role of mental health, providing a glimmer of hope for the future. Initiatives such as the WHO Special Initiative for Mental Health: Universal Health Coverage for Mental Health (2019) and the World Mental Health Report: Transforming Mental Health for All (2022) exemplify an increasing emphasis on addressing mental health concerns on a global scale [15].

The pervasive occurrence of mental disorders worldwide can be attributed to a myriad of interconnected factors. The WHO highlights a multitude of contributors, including economic downturns, social polarization, public health emergencies, widespread humanitarian emergencies, forced displacement and housing crises, protracted conflicts, violence and war, and the escalating impact of climate change [15, 136, 51]. The discussion of these macroscopic phenomena, while important, is out of the scope of the current work.

The imperative for mental health services is undeniably evident. However, the current state of care falls significantly short of effectiveness, plagued by persistent gaps and imbalances. These services remain chronically underfunded, with a disconcerting prioritization of physical disorders despite their inextricable connection to mental health [195, 14, 176, 11]. According to The Mental Health Atlas 2020, expenditures dedicated to mental health care constitute a mere 2% of health care budgets in high-income countries and a meager 1% in low- and middle-income countries (LMIC) [12]. Moreover, the allocation and utilization of these funds raise concerns, as a substantial 70% of the funding in middle-income countries is directed solely towards psychiatric hospitals [12].

Furthermore, a profound scarcity of services exacerbates the crisis. Many nations lack formalized mental health care systems, and even in countries where such services exist, they often remain inaccessible or unaffordable, resulting in a significant portion of those in need going undiagnosed and untreated. Shockingly, half of the global population resides in a country where there is only one psychiatrist for every 200,000 or more citizens. Even when access to mental health services is attainable, the quality of care exhibits considerable variability and is generally subpar. Patients may encounter challenges in affording prescribed medications which often come with undesirable side effects. Compounding these issues, there is a pervasive dearth of health literacy in mental health fostering a climate of widespread stigma and discrimination [15].

This confluence of factors highlights the pressing need for comprehensive and equitable reform in mental health care, spanning from adequate funding and resources allocation to improved accessibility, affordability, and quality of services. Additionally, addressing the

pervasive stigma and enhancing health literacy are critical components in fostering a supportive and inclusive environment for individuals grappling with mental health challenges.

The urgency to increase investments in mental health research is paramount, considering that a mere 7% of global health research funds are currently allocated to this domain. Moreover, the distribution of research efforts tilts heavily towards basic research, neglecting the requisite emphasis on clinical or applied research that directly addresses the impact of specific disorders. A stark example of this imbalance is evident in the allocation of mental health research funding, where suicide and self-harm receive less than 1% of the global funds dedicated to mental health research [223].

The benefits of improving mental health are indisputable. Enhanced mental well-being equips individuals with improved coping mechanisms to navigate life's stressors, strengthens their capacity to forge meaningful connections with others, and enhances their ability to learn and perform effectively. From a public health perspective, fostering mental health yields a multitude of advantages, including reduced suffering, enhanced physical well-being, elevated quality of life, improved functional capacity, increased life expectancy, closure of the care gap, and diminished inequalities. Many argue that mental health care is an inherent human right. Furthermore, from an economic standpoint, investing in mental health not only boosts productivity but also curtails indirect costs, thus yielding a substantial return on investment [15]. Indeed, a global modeling study estimated that expanding treatment for depression and anxiety alone would yield a remarkable benefit-cost ratio of 5 to 1 [55]. Overall, it is clear there is a pressing need for advancement of the mental health field from basic research all the way to policy-level implementations.

Collectively, these considerations emphasize the dire necessity for progress in the field of mental health, spanning from advancing basic research to implementing policies and interventions. There is an imperative to allocate adequate resources, enhance research efforts, and translate findings into tangible actions at the policy level. Addressing this pressing need is vital to ameliorating mental health outcomes, fostering societal well-being, and realizing the full

potential of individuals and communities.

1.1.2 A History of Mental Disorders and Neuropsychiatry

The historical presence of mental disorders can be traced back to the earliest written records, with references dating back to ancient Mesopotamia [182]. Paradoxically, despite the long-standing existence of these conditions, significant advancements in their diagnostics, treatment, and fundamental understanding have remained elusive. The intricate connection between the brain and the mind continues to bewilder researchers and remains an enigma to this day. The annals of neuropsychiatry bear witness to recurring cycles of hopeful "golden ages" that ultimately falter as well as complex politics stemming from contending and unanswered theories on the mind.

The brain has uniquely prompted the establishment of two distinct medical specialties: neurology and psychiatry, each addressing supposedly disparate disorders of the brain and mind, respectively. As mounting evidence suggests that the mind is an emergent property of the brain, the partition between these disciplines becomes increasingly questionable. The historical division between neurology and psychiatry arose due to diverging opinions among experts in the field over the past century as various approaches and theories were introduced in attempts to tackle the elusive field of mental illness with varying degrees of success. Indeed, the classification of mental illness as either rooted in biological factors or influenced by social elements is a viewpoint that has oscillated over time, reflecting the ebb and flow of perspectives within the field.

Across different regions of the world, varying degrees of distinction were observed between neurology and psychiatry, yet in many instances, the recognition of the interplay between the brain and the mind precluded a clear separation. However, in the United States, a notable division between these fields materialized in 1948 when Archives of Neurology and Psychiatry, a singular journal until then, was bifurcated into two distinct journals, reflecting the emerging schism between the two disciplines [143]. Psychiatry, as a dynamic and ever-evolving discipline, has undergone substantial transformations and redefinitions throughout the

past century and a half. The history of psychiatry in North America, in particular, has been marked by a mixture of distinction, controversy, and significant shifts in its conceptualization.

The division between neurologists and psychiatrists originated in the 1850s when neurologists emerged. Neurology gained significant momentum as a separate entity in the 1870s with the "brain localization project" – later called "brain mythology"– which sought to identify precise brain anatomy associated with mental disorders [90, 41]. Following the unsuccessful endeavors of anatomists, there was a brief era of "Big Data Psychiatry" catalyzed by Eric Kraepelin's pioneering longitudinal study in 1891, which popularized the field of neuropsychiatry in the US [1, 68, 143]. This empirical perspective on psychiatry persisted into the early 20th century with the influence of Adolf Meyer. As well as serving as the president of the American Psychiatric Association (APA), Meyer ardently advocated for empirically grounded psychobiology, emphasizing the necessity for rigorous study of both mind and body [18, 35, 66, 93]. He advocated for comprehensive collection of all available data, encompassing brain tissue and heredity studies, as well as developmental, social, and environmental factors shaping of a patients' life. Furthermore, famous neurologist Stanley Cobb asserted that "the mind is the living brain in action, and the brain is subject to physical and chemical changes just as any other cell or tissue in the body" [33]. Both Cobb and Kraepelin are widely considered to be founders of biological psychiatry in the US [199]. As the 20th century commenced, a conceptualization of mental disorders similar to the present understanding began to emerge, albeit with fewer technological advancements at their disposal.

Significant progress in understanding the brain and mind was hindered during this time by technological limitations paired with the intricate nature of these subjects. As a result, focus shifted towards the emergence and increasing popularity of psychoanalysis during the 1920s and '30s. Particularly in America, Neo-Freudianism gained prominence, presenting a perspective that moved away from predominantly sexual explanations and instead highlighted anxiety as the primary factor driving mental disorders [154]. The appeal of Freudianism can be attributed to its ability to provide a coherent causal explanation for disorders, consequently enabling actionable

treatment strategies that did seem to aid patients [233]. Moreover, into the 1940s, the field of psychiatry faced a general skepticism towards biological approaches due to the negative outcomes associated with practices such as electroconvulsive therapy (ECT) [2], lobotomies [3], and eugenics [93]; the last of which being met with particular revulsion after the second world war.

A significant turning point occurred in the field of psychiatry with the approval of chlorpromazine by the US Food and Drug Administration (FDA) in 1954, marking the first sanctioned drug for the treatment of mental disorders. This milestone ushered in an era of pharmaceutical therapy, propelled by the concurrent advertising revolution and the prevailing "age of anxiety" caused by the Cold War [93]. The impact was immediate and profound, as evidenced by the remarkable boost in sales experienced by Smith, Kline, and French (SKF, today GlaxoSmithKline or GSK), the manufacturers of chlorpromazine. Within a year of its approval, the drug exponentially increased their sales figures, and within a decade, an astounding 50 million prescriptions had been filled. This extraordinary success continued, with SKF's revenues doubling three times within a span of 15 years [4]. The advent of chlorpromazine paved the way to a flurry of other psychiatric drugs, including the widespread use of meprobamate. By the end of the 1960s, one in three prescriptions written in the US were meprobamate. Another notable drug was Valium, which became the most prescribed drug in the country within a year of its introduction, a position it held until 1982 [151].

The perceived efficacy of these pharmaceutical interventions, coupled with the prevailing neo-Freudian attitudes of the time, contributed to an unexpected consequence of mass deinstitutionalization. The process of deinstitutionalization gained momentum in 1963 with Congress passed the Mental Retardation Facilities and Community Mental Health Centers Construction Act, signaling a shift towards community-based care. This trend was further solidified during the economic recession of the 1970s [116, 117]. The consequences of this policy change reverberate to this day, with correctional facilities now serving as the three largest providers of mental health services in the US [171]. The lasting impact of these transformations highlights the intricate and

ongoing challenges faced by the modern mental healthcare system.

The shortcomings of neo-Freudian leaders of the time paired with the popularity of pharmaceutical interventions and emerging technological innovation ushered in the biological psychiatry revolution of the late 1970s and 1980s. In an effort to establish psychiatry as a legitimate medical field rather than a purely sociological field, psychiatrists increasingly emphasized emerging developments in biological research and pharmacology to persuade the public of the medical nature of mental disorders. This period was characterized by overly confident statements, exemplified by a 1982 article in The Washington Post proclaiming that "scientists have unlocked the doors to understanding the body's most complicated and baffling organ: the brain" [60]. Furthermore, in 1990, the US Congress in collaboration with President George H. W. Bush declared the subsequent decade the "Decade of the Brain", signaling an optimistic belief that the resolution of humanity's perplexity regarding the brain and mind was within reach [45].

Despite optimistic claims, the words of Thomas Insel – the director of the National Institutes of Mental Health (NIMH) at the time – prompted sobering contemplation. In 2010, Insel remarked that despite the proclaimed "Decade of the Brain," there was no discernible increase in the rate of recovery from mental illness, nor a detectable reduction in the burden of mental illness [107]. This statement serves as a reminder of the complex and multifaceted nature of mental health challenges, urging us to acknowledge the limitations of the current understanding and approaches.

While there was limited advancement in the diagnosis and treatment of mental disorders, significant discoveries emerged from diverse areas of basic research, ranging from the genomic revolution to a deeper comprehension of the cellular complexity to remarkable progress in the study of brain circuits. Technological innovation in the form of computerized tomography (CT), positron emission tomography (PET), magnetic resonance imaging (MRI), and functional MRI (fMRI) have revolutionized the field of neuroimaging. Additionally, in 1987 Deep Brain Stimulation (DBS) emerged as a breakthrough technique for modulating brain activity [201]. These

tools have expanded the understanding of the brain, unveiling novel concepts and complexity; however their translation to clinical applications has been limited.

Again, Thomas Insel emphasized the scarcity of novel therapeutic targets, primarily due to the continued limited understanding of the underlying pathophysiology of neurodevelopmental disorders [109]. While psychiatric medications have seen little improvement in efficacy, their sales have increased exponentially, largely driven by the relaxation of regulations surrounding direct-to-consumer (DTC) pharmaceutical advertising, perpetuating the trend of overmedication. Psychiatry, in particular, has been heavily impacted, with psychiatric medications being the most heavily advertised [26]. Despite the absence of functional innovation, revenue generated by psychiatric drugs increased sixfold from 1987 to 2001, surpassing the growth rate of general prescription drug sales [126]. Instead of truly embodying the era of biological psychiatry as heralded by scientist, journalists, and politicians alike, the 1980s witnessed an era dominated by pharmaceutical sales. The field still lacks the requisite knowledge to effectively translate fundamental neurobiology into innovative diagnostics and therapeutics [109]. This realization underscores the necessity for embracing alternative approaches and harnessing additional tools to advance the understanding and treatment of complex mental disorders.

1.1.3 The Issue of Psychiatric Diagnostics

A recurring theme throughout psychiatry's history is a dearth of quantitative biophysiological underpinnings for characterizing diseases, setting it apart from other domains within medicine. As such, diagnostics is a particular area of concern in the field. In parallel with prevailing views of empirical psychiatry as advocated by Kraepelin and Meyer during the time, the American Medico-Psychological Association published the Statistical Manual for the Use of Institutions for the Insane in 1918 [160]. Subsequently, through numerous revisions, it evolved into the inaugural edition of the Diagnostic and Statistical Manual (DSM), which was first introduced by the APA in 1952 [165]. Presently in its fifth iteration, the DSM remains the key tool for classification of mental disorders. Primarily utilized in the US, it is a qualitative checklist

that associates each mental disorder with predetermined criteria, with the APA assuming the role of defining, determining, and demarcating the disorders. Consequently, this classification system offers a binary perspective whereby a patient either possesses the disorder or does not. The advent of the biological psychiatry revolution in the 1980s fostered optimism that these checklists would eventually be supplanted by quantitative biomarkers, thereby facilitating the establishment of substantial and reproducible diagnostic criteria. However, this expectation remains unfulfilled.

Since its inception in 1952, the DSM has become a subject of increasing scrutiny. Following a significant gap since DSM-IV in 1994, the release of DSM-5 in 2013 sparked a crescendo of discontent. The APA had envisioned the fifth edition as the culmination of the long-awaited integration of biological principles into diagnostics, thereby realizing the aspirations of the biological revolution. Despite numerous experts cautioning against premature publication, the APA forged ahead. As anticipated, DSM-5 faced severe criticisms, as the promised scientific foundation and identifiable biomarkers were conspicuously absent. In response, Thomas Insel proclaimed that the NIMH would redirect its research endeavors away from DSM categories. Instead, the focus would shift towards the development of an improved system, one that rested on the bedrock of neuroscientific and cognitive science discovery, rather than relying solely on manifest symptoms [109, 107, 108]. Another esteemed former NIMH director, Steven Hyman, also voiced sharp criticism of the APA's misstep. He underscored that the field of psychiatry had failed to generate novel insights into molecular targets for diagnoses and treatments since the 1950s [93]. Regarding the DSM-5, he stated "the underlying science remains immature and... therapeutic development in psychiatry is simply too difficult and too risky" [103, 104, 105]. Since then, both Insel and Hyman have left the NIMH and have begun work in digital big data analytics with hopes of unveiling complex patterns. Psychiatric diagnostics still remain a persistent issue to this day.

1.1.4 Convergence of Neuropsychiatry

Today, the growing body of compelling evidence points towards the resurgence of a unified field known as neuropsychiatry, which embraces the intrinsic interconnectedness of the brain and mind. This inseparable dyad challenges the increasingly arbitrary and unsubstantiated division between neurology and psychiatry, which predominantly stems from theoretical schisms within professional guilds. The perpetuation of this division sustains the fallacy of brain-mind dualism, yielding negative consequences such as the stigmatization of individuals with mental illness, a general confusion around the biological basis of mental disorders, and inadequate insurance coverage for psychiatric and mental health services.

Historically, a crude demarcation between neurological and psychiatric conditions was established based on whether the disorder was considered organic (brain) or functional (mind). Neurology focused primarily on disorders originating from observable lesions, while psychiatric disorders were attributed to disturbances in affective function without clear-cut biological correlates. Specifically, neurology has predominantly relied on observational correlations, linking symptoms to structural changes or associating disorders with specific pathologies. This neurology-psychiatry classification scheme becomes increasingly tenuous in light of burgeoning research. For instance, epilepsy was considered a mental disorder until electroencephalography (EEG) was invented and used to localize epileptic attacks in the temporal and frontal lobe during the 1950s [183]. Afterwards, it was reclassified as a neurological disorder. Mounting research reveals that major psychiatric disorders manifest as brain diseases, further undermining the notion of a rigid separation between these domains.

The convergence of neurology and psychiatry has emerged through the biologization of psychiatry and softening of neurology. Disorders that were once considered purely psychological, lacking identifiable biological foundations, have undergone a transformative shift in their perception, owing to the undeniable link between biology and behavior. This paradigmatic shift has been largely facilitated by remarkable advancements in neuropharmacology, which have

elucidated the intricate interplay between brain chemistry and behavioral manifestations. Groundbreaking research on neurotransmitters and the demonstrated impact of existing pharmaceutical medications have unequivocally established this connection.

Noteworthy examples include the profound impact of interventions like Deep Brain Stimulation (DBS) and Transcranial Magnetic Stimulation (TMS) on mood and cognitive function, respectively, underscoring the physiological effects of these treatments in addressing depression. In parallel, neurology has witnessed a softening of its stance, shifting from its previous skepticism of concepts such as plasticity, regeneration, and recovery, largely due to the blanket dismissal of psychiatric principles including talk therapy [143]. The rise of cognitive neuroscience has provided a contextual framework for integrating psychotherapy within the domain of neural plasticity [109]. It is now widely recognized that the most optimal therapeutic outcomes for mental disorders often emerge from a comprehensive approach that combines medication and talk therapy. A crucial lesson gleaned from the retrospective examination of these disciplines over the past century is that many theories that have staunchly divided the field have frequently proved erroneous. Nonetheless, challenges persist regarding professional boundaries and territorial disputes, impeding the seamless integration of these fields.

1.2 The Strategy: Computational Neuropsychiatry

1.2.1 The Brain as a Complex Multiscale System

It is important to emphasize the biological complexity of the brain as an organ, particularly its multiscale nature. Multiscale here refers to different levels of resolution and complexity in which one can study the brain. Despite efforts, for most mental disorders, there is no single definitive underlying cause that can be determined. As was discovered by the pseudoscientific phrenologists of the 19th century, measurements of the skull and brain are not informative enough to predict behavior and mental traits [90, 41]. Unfortunately, there is no single gene mutation that causes most mental disorders. Instead, it becomes necessary to probe at varying levels

of resolution from genes to DNA to proteins, molecules, cells, circuits, networks, cognition, behavior, and the environment (Fig. 1.2). The latter is a particularly important scale to include in neuroscience due to the notion of embodiment – neural system functioning depends on feedback from the body and the world around it. Therefore, the system cannot be accurately described as closed loop unless the external environment is included.

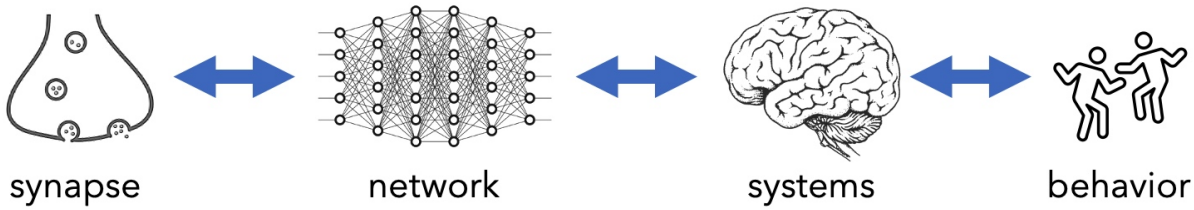

Figure 1.2. Multiscale nature of mental disorders spanning from the molecular and cellular level through networks and systems to behavior.

In theory, it should be possible to model continuously through the scales from the lowest level of genes to the highest level of behavioral presentations recognized clinically. Indeed, this is what would causally and mechanistically describe mental disorders, from which clear, concrete, and quantitative biomarkers and brain dysfunction pathways would emerge. Such a framework would truly exemplify an understanding of the brain and its function. Unfortunately, there lacks a fundamental lack of understanding, particularly of human cognition, to build such a bridge from the biophysical to the phenomenological. A major limitation in this is that tools have not yet been invented in order to experimentally probe and visualize brain dysfunction at every scale. Furthermore, precisely how t mind is an emergent property from the brain is not yet mechanistically understood making the link from networks to cognition and behavior particularly difficult. Translation from advances in neuroscience to clinical applications has been slow due to these factors. Eventually, the hope would be for a multiscale model that allows for the prediction of brain dysfunction effects, and this is something the field is working towards.

Being able to bridge these levels of investigation is important for a causal explanation around mental disorders. For example, in the case of depression, pharmaceutical intervention

strategies operate on the cellular and subcellular level, affecting neurotransmitters, such as serotonin and dopamine, but diagnostics as well as psychotherapy operate in the behavioral realm. Being able to fully bridge the observations between scales would allow for better mechanistic understanding of disorder development and improved treatment methods from beginning to end. Especially given that the best outcomes to date have been demonstrated with a combination of treatment methodologies. The question then becomes, especially given the limitations associated with experimentation, how do we do this?

1.2.2 The Era of Computational Neuropsychiatry

In the continuum of psychiatry's transformative revolutions, today is a new era–the era of computational neuroscience and psychiatry, marked by the inaugural international computational psychiatry meeting held in 2013 [82]. This computational revolution extends beyond psychiatry alone. Virtually every field from biology to finance to communications is undergoing a paradigm shift driven by computational advancements.

This age emphasizes the development and utilization of novel methodologies, rather than solely focusing on uncovering underlying causal mechanisms. Computational methods have gained popularity for two primary reasons: the "Big Data Revolution" and better, faster computers. "Big Data" here refers to large and complex datasets which are now being produced faster and at a larger scale than ever before, including in neuropsychiatry. Technological innovations have vastly expanded the availability of neuroscientific data, such as the "omics" revolution enabling large-scale -omics studies encompassing genomics, proteomics, and metabolomics. Additionally, the advent of structural and functional MRI in the 1990s has further contributed to the wealth of neurological data available. Recently, numerous initiatives have emerged, dedicated to collecting large-scale neurophysiological datasets, including the BRAIN Initiative, Human Connectome Project (HCP), and Adolescent Brain Cognitive Development (ABCD) study.

Furthermore, the exponential growth of computational power is a driving force behind the burgeoning fields of artificial intelligence (AI), computer science, and data science. Indeed,

neural networks used in AI today are rooted in computational neuroscience models conceived approximately half a century ago. However, it is only now that these fields are exploding due to the unprecedented magnitude of available data and processing capabilities.

In conjunction with the technological landscape shaping the backdrop for this revolution, it assumes a pivotal role as an indispensable instrument in its own right. The brain, being an intricately complex and multiscale organ, necessitates computational endeavors to bridge the explanatory gaps in fundamental neuroscience and elucidate profound insights that elude experimental means. Furthermore, computational modeling operates in tandem with experimentation, serving as an auxiliary tool to guide and optimize experimental pursuits, thereby enhancing their efficacy and precision. Sole reliance on experimentation proves exceptionally challenging when attempting to establish connections between lower-lever information-theoretic measurements and dysfunctions in cognition and semantic processing. Computational modeling permits quantitative exploration of intricate systems across various levels of interest, acknowledging that explanations are not confined to a solitary level but necessitate investigations that traverse multiple levels. A pertinent illustration of this is witnessed in the genomics revolution, which witnessed numerous unsuccessful endeavors to localize singular genes as biomarkers for psychiatric disorders [220].

By employing computational models, researchers gain the capacity to traverse different testing environments, transitioning from animal models to human models, thereby widening the scope and applicability of their findings. Moreover, contemporary methodologies suffer from imprecision, as they lack both biomarkers and a methodological framework to quantify treatment outcomes, rendering the efficacy of interventions uncertain. Finally, software tools emerge as a compelling avenue for translating neuroscientific discoveries into practical instruments that ameliorate the lives of patients, effectively bridging the gap between theoretical advancements and tangible clinical applications.

There are numerous applications that show promise in this field including computational phenotyping, diagnostic classification, treatment selection, and biophysical modeling, although there are and will be many more. Generally, there are two primary approaches to computational

methods, both within neuropsychiatry as well as other fields: theory-driven and data-driven. Theory-driven approaches are limited by existing prior knowledge and mechanistic understanding while data-driven approaches aim to be theory-agnostic and look to ascertain insights from data. The hallmark of theory-driven approaches in neuropsychiatry are biophysically-detailed models. These models utilize simulations and mathematical analysis informed by scientifically-backed data, model, and parameter choices as well as physically accurate models of dynamics. These models are appropriate for studying systems in which there exists detailed knowledge of the structure and function of the area of interest. Such models are particularly useful for understanding causal relationships at specific levels of analysis or to bridge scales. Of note, biophysical models do typically involved substantial simplifications or abstractions. Nonetheless, they have been shown to be useful in linking biological abnormalities involved in psychiatric disorders to certain levels of dysfunction.

Navigating between the molecular and network level tends to be the primary focus of such biophysical models due to their relatively high level of mechanistic understanding in the field relative to other scales. Since the efficacy of pharmaceuticals, neurotransmitter studies have been particularly popular. One analysis built a biophysically realistic neural network model to investigate the impacts of serotonin and glutamate disruptions implicated in obsessive-compulsive disorder (OCD), in which patients typically display decreased serotonin and increased glutamate relative to controls [138]. The model found that OCD network abnormalities could be assuaged by increasing serotonin amounts, regardless of whether the disruptions were originally caused by serotonin or glutamate, suggesting a possible mechanism for the relative efficacy of selective serotonin reuptake inhibitor (SSRI) usage for OCD treatment. Furthermore, they were able to implicate a specific serotonin receptor (5HT2A), demonstrating promise for precise treatments moving forward. Another study built similar networks to relate excitatory neurotransmitter receptor density to neuron activity signals obtained from fMRI. The focus was on two populations of neurons shown to be affected in psychosis – the resting state and task-based networks. It was found that when excitatory receptor function on inhibitor interneurons was reduced, neuron

activity signals matched those found in psychosis [27]. It was also shown that excitatory neurons in the cortex called pyramidal neurons were not implicated. Thus, biophysically realistic models provide a tool to bridge scales and probe mechanistic causes of psychiatric disorders that experimentation cannot. Working together, such computational models are able to motivate particular experimental studies.

In applications where prior knowledge is lacking, data-driven approaches provide an alternative method to uncover patterns in complex data and tend to be more oriented towards developing clinically useful applications. Typically, they utilize machine and deep learning methods to elucidate insights. Computational phenotyping seeks to find measurable neural or behavioral traits defined in terms of a computational model [156]. Similar to standard phenotyping, it is expected to show variation across individuals. Large-scale and multimodal computational phenotyping has not yet occurred but holds promise as a method to quantitatively capture psychiatric differences. Related is the task of automatic diagnosis classification, which is compounding into a substantial body of work [166, 224]. Given the limitations of current psychiatric diagnostics, this work seeks to automatically classify patients with disorders compared to controls, typically using machine learning. For example, a recent competition was launched to classify schizophrenia from healthy controls using MRI data, and the top performing entry achieved an area under the curve (AUC) of 0.89 [204]. The top three methods used different approaches and reached similar performance metrics [192], and an ensemble approach achieved 0.93 AUC [204]. This is significant given psychiatric overlap in diagnostics. While deep learning models now are not interpretable, there is promise for learning biomarkers based on automatic classification as well. Additionally, attention-based mechanisms provide some amount of insight to decision factor weighting for deep learning models.

Once diagnosed, psychiatric disorders generally have multiple treatment strategies per disorder, so treatment selection provides another area of application for computational methods. Similar to recommender systems used in movie recommendations by Netflix, patients' similarity can be used to predict effective treatment strategies. One study used EEG to predict a

pharmacologic treatment strategy for patients in a randomized clinical trial [67]. The dataset contained 17,000 depression treatment attempts for over 1,800 subjects, averaging a disturbing nearly 10 treatment attempts per subject. The treatment selection used EEG features to predict the medication that would be most likely to be successful significantly outperformed clinical selection [102]. Another study looked to predict treatment regimens based on patient data, finding that comorbid personality disorders responded better to antidepressants than cognitive behavioral therapy (CBT) while patients that were married, employed, had more life events, and failed more antidepressants attempts responded better to CBT [69]. Using data-driven treatment selection as opposed to standard treatment selection led to a larger reduction on depressive symptoms as quantified by the Hamilton Rating Scale for Depression [79]. It is clear utilizing data for treatment selection shows significant promise for psychiatry moving forward. This is a particularly important area for improvement given claims surrounding psychiatrist medication selection being heavily influenced by pharmaceutical companies [93].

While these methods are typically utilized in isolation, their true power emerges when they are artfully unified. Despite demarcation between theory and data-driven approaches, their collaboration seemingly yields the most optimal outcomes. Nonetheless, studies gravitate towards one of these two camps. Clinical applications, by and large, embrace data-driven methodologies, while the pursuit of theoretical understanding utilizes theory-driven approaches. An illustrative example of the benefit of merging the two lies in grappling with the "curse of dimensionality." When confronted with many dimensions of data–consider a scenario where an exhaustive collection of data on every topic for a given subject is available–there exists the peril of overfitting. Essentially, the model might excessively fit to the provided data, severely hampering its ability to generalize to novel examples. Put differently, the model will learn the subjects used for training too specifically such that it is unable to generalize to new subjects. In domains where mechanistic understanding is sufficient, theory-driven models can be used to obtain features with fewer dimensions than the original dataset. Subsequently, these new features can be employed in a data-driven manner for clinical applications.

One example of this approach combining theory and data methods was exemplified in a study that used EEG data to classify patients as on or off DBS. Raw data was compared with theory-driven measures obtained from the raw data as features. Specifically, a drift-diffusion model (DDM) was used to obtain a model of decision making. Indeed, classification improved when leveraging the fitted DDM model parameters rather than the raw data [221]. This demonstrates potential for combined theory- and data-driven methodologies to elucidate new insights to mental illnesses.

Overall, the current era of computational neuropsychiatry presents an opportunity to make significant gains in the understanding, diagnostics, and treatment of mental disorders by making use of a combination of theory- and data-driven approaches. It seeks to supplement experimental methods while providing objective and quantitative measures to move the field forward.

1.3 The Tools: Artificial Intelligence and Probabilistic Graph Models

1.3.1 Artificial Intelligence Today

Artificial intelligence (AI) is more popular now than ever, with the number of AI publications having more than doubled between 2010 (200K) to 2021 (500K). The number of AI code repositories associated with publications has grown almost 27 times in the past 12 years, and the total number of GitHub AI repositories has grown from just over 1,500 in 2011 to nearly 350,000 in 2022. Additionally, the AI conference attendance has been generally increasing, although the COVID-19 pandemic did affect this trend. Neural Information Processing Systems (NeurIPS), the largest international conference has over a 10-fold increase from 1300 participants to over 15,000 in 2022. Numbers even reached over 22,000 in 2020 when the conference was virtual-only. At the same time, AI performance is also increasing exponentially. The ImageNet annual challenge saw an accuracy of over 40% between 2011 to 2022, going from 63% to 89%

without additional training data and 91% with additional images. Demand for AI-related skills in the labor markets has also increased internationally. In the US alone, the percentage of all jobs requiring AI skills is 2.05% in 2022 compared to 0.58% in 2014. Overall corporate investment in AI saw a peak in 2021, reaching $276.14B. Likely due to post-pandemic economic downturn, this has decreased by approximately a third to $189.6B in 2022. Both of which are significantly higher than even 2013, in which the global investment was only $14.57B. The number of newly funded AI companies in the US in 2022 was 542, compared to 337 in 2013. The focus areas that received the most investments globally in 2022 were medical and healthcare ($6.1B), data tools ($5.9B), financial technology ($5.5B), security ($5.4B), and retail ($4.2B). Interestingly, medical and healthcare investments have risen 165% from $2.3B in 2017 to $6.1B in 2022, suggesting an increase in investments due to the COVID-19 pandemic [144].

Education further represents the increased prevalence of AI as the proportion of US computer science PhD graduates specializing in AI was 19.1% in 2021, compared to 10.2% in 2010 and even 14.9% in 2020. Of note though, these graduates are increasingly joining industry over academia. In 2011, the percentage of AI PhD graduates that stayed in academia was equal to industry at approximately 40% (41.6% and 40.9% respectively). A decade later, 65.4% took a job in industry compared to 28.2% in academia. This is likely related to the availability of faculty positions, which has actually slightly decreased since 2012 (710 hires in 2021, 733 hires in 2012). Concurrently, there has been an increase in number of students graduating with a computer science bachelor's degree, from 9K in 2010 to just over 33K in 2021, nearly four times more [144].

Today, there are even the beginnings of policy and governance surrounding AI, which was never before seen. Globally, the number of bills related to AI that were passed into law grew from only 1 in 2016 to 37 in 2022, with the majority of 9 from the US. US government spending on AI-related contracts has also increased approximately 2.5 times since 2017, from $0.66B to $1.7B in 2022 [144].

It seems virtually impossible to escape the conversations surrounding AI, especially with

the release of ChatGPT. Amongst all the media and public's outcry, it is clear there is a general sense of confusion as to what artificial intelligence is in practice.

1.3.2 The Interdisciplinary Nature of Artificial Intelligence

In theory, AI seeks to create machines that display intelligence where intelligence is defined primarily in the context of natural intelligence, using cognitive terminology such as perceiving, understanding, predicting, and manipulated the world around it. Indeed, the ultimate test of a successful AI is The Turing Test, in which a human tests a computer using written questions, and the machine passes if the tester cannot discern whether it is human or machine. In practice, artificial intelligence is a long way off from artificial general intelligence (AGI), although performance in human-like tasks gets better each year.

The field of artificial intelligence is commonly perceived as exclusively computer science, however it is inherently interdisciplinary. In building the AI models, numerous fields of math are represented, including logic, probability, statistics, and calculus to name a few. Economics helps to answer questions surrounding decision and game theory, tradeoffs, and utility optimization. Obviously, computer engineering is critically important to build efficient hardware systems to run software models on and control theory provides quantitative methods of feedback. To many, the engineering side of AI is clear. What is often forgotten about is the scientific side. In order to design an intelligent system, natural intelligence must, to some degree, be understood and implemented. Understandably, this introduces philosophy around the definitions of intelligence, reasoning, and knowledge. Then looking towards the fields of psychology and neuroscience to answer the questions of how humans and animals think and act becomes a natural next step. Indeed, this is where neural networks emerged from. Lastly, linguistics is important to relate language to thought. The field of cognitive science even emerged from all these as an attempt to house all the fields under one title with the goal of bringing together AI models and psychological experimentation to quantitatively describe the human mind. Data science has emerged similarly to attempt to handle the complexity of data surrounding AI questions [191]. It can be said with

certainty that these will not be the last fields that emerge in humanity's quest to understand and recreate its own intelligence. It is an elusive mystery that still pervades to this day despite each human having a brain.

1.3.3 A History of AI

The history of AI and natural intelligence has been understandably intertwined. AI as we know it today began in the 1940s with the creation of the first mathematical model of an artificial neuron by Warren McCulloch and Walter Pitts (1943) [147]. A neuron, the nerve cells that primarily compose the brain, takes in signals from surrounding neurons. In response, it will either fire itself, perpetuating the signal, or not. The McCulloch-Pitts neuron similarly took in multiple inputs from neighboring neurons and would respond by either turning "on" or "off" as a sort of biological logic gate. They were followed by Donald Hebb who, in 1949, developed a rule for changing connection strengths between the neurons based on firing, called Hebbian learning, which still remains popular today [96]. The first ever neural network was built soon after in 1950 by Marvin Minsky and Dean Edmonds out of 3000 vacuum tubes and a B-24 bomber automatic pilot mechanism. It could simulate 40 neurons in total [191, 111]. It was also the year that Alan Turing introduced the Turing test as well as the ideas of machine learning and reinforcement learning in his article "Computing Machinery and Intelligence" [213]. In 1956, ten top researchers in the field got together at Dartmouth with the goal of solving artificial intelligence, coining the first usage of the term "Artificial Intelligence" to describe this field of study [152]. While general intelligence was not solved, the 1950s and early 1960s did see AI successfully take on numerous tasks never before seen, including games, puzzles, and IQ tests. Another major event occurred in 1962 with the advent of the perceptron, a simple linear classifier artificial neuron still used in neural networks to this day, by Frank Rosenblatt [187, 188].

Research in the second half of the 1960s into the 1970s slowed down significantly due to fundamental limitations with the approaches being used. Many early AI systems were not general models but rather built from "informed introspection" on how humans perform a task,

generally utilizing logic tools. Furthermore, computational complexity theory was not yet developed leading to a naïve view of the intractable nature of the AI problems. It was believed that larger problems could easily be solved simply with faster and better hardware despite the combinatorially explosive nature of these problems [130]. Lastly, the tools available were limited. The final nail in the coffin seemed to be when Minsky and Seymour Papert published the book *Perceptrons* in 1969, which proved that perceptrons could represent very little [153]. Multilayer neural networks were not addressed in this book, but the damage had been done and funding dried up.

Neural networks returned with a vengeance in the mid-1980s when multiple groups, including Geoffrey Hinton, in parallel reinvented the back-propagation learning algorithm for use in neural networks and disseminated their results in *Parallel Distributed Processing* [189, 190]. Interestingly enough, the backpropagation algorithm was already developed in the early 1960s by Henry Kelley (1960) [115] and Arthur Bryson (1962) [43], but remained unused in neural networks for twenty years, again highlighting the importance of interdisciplinary collaboration. This sparked the connectionist movement centered around using neural networks as the basis of AI, which was in contrast to the previously used symbolic approaches by Newell and Simon in the 1970s [161] and logic by the Dartmouth men in the 1950s [152]. Thus, machine learning and probabilistic reasoning took over hard-coding and Boolean logic. Popular networks still used today were invented or popularized during this time, including hidden Markov Models (1980s), Bayesian networks (1988) [172], Markov decision processes [209], and convolutional neural networks (CNNs). It also set the stage for communally-sourced shared benchmarks as performance measures, including MNIST (handwritten number images), ImageNet, and COCO (image object recognition), datasets which are still used to this day.

After the connectionist heyday of the 1980s, there was a brief slow down in progress due primarily to lack of compute power and unavailability of large datasets. This changed at the turn of the century with the Big Data Revolution caused by advances in computing and the emergence of the World Wide Web. Since then, there has been exponential growth in the creation of large

and complex datasets in a wide variety of forms, including text, audio, video, and images. The connectionist movement paired with the Big Data Revolution paved the way for deep learning, the field of AI most common today. In 2011, the first CNN trained on a graphical processing unit (GPU) with a very large dataset by Geoffrey Hinton's lab swept the competition, launching the field into the Deep Learning Revolution, which we are still in today.

1.3.4 AI in Medicine

While more straight-forward problems have been the main focus for AI for quite some time, increasing attention is being paid to AI in medicine with the push for data standardization and the development of large medical datasets. Much like a digital thermometer or stethoscope, AI acts as another tool that can be used by doctors to better diagnose and treat their patients, rather than replacing doctors. Recently, AI has proven particularly useful in diagnostics, matching or surpassing doctors diagnostic accuracy. These diagnostics employ a wide array of AI techniques generally dependent on the disease and data available. The major areas utilized for diagnostics are computer vision applied to static images and time series or sequential data analysis.

Use of computer vision for diagnostics has the largest body of work. In fact, the FDA approved uses of AI for healthcare using computer vision applied to medical scans (MRI/PET) and pathology images. It functions theoretically similar to other applications where a model, typically a CNN, is trained using many images of a particular disease or disorder, allowing it to classify future images as representative of said disease or not. For example, applications include the detection and quantification of breast densities using mammography [127], detecting stroke and brain bleeds from computerized axial tomography [139, 6], and classification of dermatological conditions [78]. A systematic review and meta-analysis comparing deep learning performance to clinicians in detecting disease from medical images found them to be currently equivalent [131], although some studies have even surpassed clinicians. For example, the LYNA system which diagnoses metastatic breast cancer is more accurate than a clinician, achieving accuracy of 99.6%. It turns out that the most effective is the combination of doctor and AI

tool, which is able to surpass performance measures of either method independently, suggesting the features used to diagnose by clinicians is different to some extent than AI [133, 206]. This highlights to importance of machine and human expert partnerships, rather than the typical notion of AI as a replacement for humans.

Times series applications focus on temporal data, or more generally on ordered sequential data. It uses existing sequential data to classify a state for the data or detect anomalies within it. Popular models for sequential data types are recurrent neural networks (RNNs) or, more recently, transformers, although CNNs can also be used. Applications include detecting and classifying arrhythmias in electrocardiograms [92]. These models can also be applied to speech for detection of diseases that impact natural speaking. This method has been used for chronic pharyngitis [129], Parkinson's disease [227], Alzheimer's disease [80], post-traumatic stress disorder [142], and major depressive disorder [184]. A specific type of sequential data analysis is natural language processing (NLP), which seeks to make predictions using written language. A large area of research lies in applying NLP techniques to electronic health records (EHR). For example, one study used EHRs to predict patient mortality after hospitalization [179]. Another study sought to predict disease severity and therapy efficacy from EHRs combined with medical data [73].

This is just the beginning of AI's usage in medical diagnostics, and these examples provide exciting evidence for its usage in neuropsychiatry moving forward. In using models under the general class of AI, it becomes critically important to break the system into it's constituent parts and approach the problem using design principles.

1.3.5 Design of an Intelligent Agent and Task Environment

The pursuit of artificial intelligence revolves around the design and development of intelligent agents, defined as entities capable of perceiving their environment and taking action based on those perceptions. Like many other complex challenges, it is advantageous to approach the problem using design principles. Typically, the design task entails creating an agent with

appropriate software and hardware components, although this is not always the case. In general, an agent requires a computing architecture and a mechanism for executing actions. An agent's behavior is encapsulated by its agent function, which serves as an abstract mapping from percepts or sequences of percepts to actions. A percept here refers to any content an agent's is aware of. The primary focus of AI lies in the development of a concrete and internalized agent program that defines this agent function. This program establishes the link between percepts received as inputs and the corresponding actions generated as outputs. The resulting action undertaken by an agent is a product of both its inherent knowledge and the complete sequence of percepts it has observed. Historically, an input-output table approach was utilized, however this quickly becomes intractable as the complexity of the environment and the agent's interactions grow. In contrast, the field has increasingly shifted towards internalized computation, seeking more efficient and flexible methods of agent programming. Moreover, drawing insights from human intelligence, there is no analog of a fixed table mapping percepts to actions. The challenge therefore lies in devising programs that engender meaningful, rational, and intelligent behavior while maintaining computational tractability. By employing sophisticated design strategies, AI endeavors to create intelligent agents capable of exhibiting behavior that aligns with human-like intelligence while remaining computationally viable.

The simplest type of agents is a simple reflex agent, which selects actions based on its current percepts using condition-action or if-then rules. This means that if a certain condition is met, then an action occurs. Although not generally discussed in the context of AI, humans possess reflexes that parallel this agent archetypes, such as the knee-jerk or patellar reflex. When the patellar tendon is struck in the correct spot, an automatic kick response is elicited. Such an agent is said to lack autonomy as it relies entirely on its preprogrammed prior knowledge rather than its percepts or any form of learning. The efficacy of this agent program hinges on the environment being entirely observable and predictable, allowing the agent to have complete visibility of the environment's state at each moment and determine the subsequent state based on the current state and the agent's action. An illustrative example of such a program can be

found in the equations of motion in physics, where the agent in question is an object in motion. By providing the initial position and velocity of the object, it becomes feasible to predict its subsequent location at the next time point using the equation s_2–$s_1 = v \times t$. While this equation suffices for describing simple cases, its efficacy diminishes rapidly when additional factors such as friction and air resistance are introduced. Indeed, the real world encompasses a level of complexity that cannot be adequately captured by a simplistic and deterministic equation. In essence, reflex agents serve as the rudimentary building blocks of agent design, demonstrating the limitations of deterministic and rule-based approaches when confronted with the intricacies of real-world environments.

The aforementioned examples highlight the significance of accurately capturing the task environment, which is the surroundings in which the agents are problem-solving. Observability is a key factor, where environments are considered fully observable if the agent is able to detect all aspects relevant to its choice of action. If not, typically due to the noisy input or because parts of the environment's state are missing, it is considered partially observable. The extreme case is if the agent is unable to sense the environment at all, resulting in an unobservable system. If the environment is not fully observable, the next simplest agent program is that of a model-based reflex agent, which is able to keep track of parts of the world that are not currently observable as some sort of internal state. In order to do this, there needs to be a model of how the environment changes over time, typically including the effects of the agent's actions as well as how the environment independently evolves. This is called a transition model. If a particular goal is assigned to the agent, the model-based agent becomes a goal-based agent. The last category of agents is the utility-based agent, which has a more nuanced modeling of success than the goal-based agent in the form of performance measures. Typically, the model seeks to maximize overall expected utility, which often addresses tradeoffs [191].

The categories of agents previously discussed can be further extended into learning agents, which provide an alternative to programming everything *a priori*. Learning agents possess the ability to take in information from their surroundings and update their internal

model, enabling them to generalize more effectively to complicated environments and tasks. A substantial portion of contemporary AI research focuses on learning agents, particularly within the subfield of machine learning, where the machine assumes the role of the agent and learns from environmental data provided by the designer.

When designing AI problems, careful attention must be given not only to agent but also to the task environment design. In addition to previously mentioned observability, there are other relevant dimensions to consider. Related to but distinct from observability is determinism. If the next state of the environment can be completely determined or predicted based on the current state and the agent's actions, it is considered deterministic. The majority of real world situations are characterized by such complexity that they are inherently nondeterministic.

Another significant dimension is the episodic or sequential nature of the environment. In episodic environments, decisions made by the agent are independent of past decisions while in sequential environments, prior decisions impact subsequent ones. Furthermore, the concept of dynamics refers to whether the environment undergoes changes while the agent is in the process of decision-making, as opposed to remaining static. An environment's state space is yet another dimension to consider. A discrete environment comprises of a finite number of distinct states, whereas continuous environments involve an infinite range of possible states. A known environment refers to the agent (or the agent's designer) knowledge of the environment. If it is known, then the outcomes for all actions are given while unknown required learning. Lastly, it is crucial to consider whether the AI problem involves a single agent or multiple agents interacting within the environment.

By carefully considering these various dimensions of the task environment design, alongside agent design, researchers and practitioners can effectively tackle AI challenges and develop intelligent systems that operate optimally within their respective domains, including in neuropsychiatry.

1.3.6 Representing Uncertainty Using Probabilistic Graph Networks

In scenarios where an agent encounters challenges in navigating the world with absolute certainty, such as in environments that are nondeterministic or not fully observable, the ability to handle uncertainty becomes crucial. Uncertainty arises when it is impossible to exhaustively list all possible outcomes, often due to an abundance of possibilities resulting in a large state space or theoretical knowledge gaps. In this context, probability theory serves as the primary mathematical language for handling uncertainty. Rather than possessing unwavering certainty in actions, the agent's knowledge is characterized by the degree of belief in a particular outcome. For instance, when predicting weather, it is impossible to achieve 100% certainty. Instead, forecasts are expressed with a degree of belief in the form of probability (e.g. a 20% chance of rain tomorrow). However, maintaining probabilities, also referred to as knowledge representation, for all components in a system quickly becomes a daunting task without effective tools.

Knowledge representation in uncertain domains is commonly handled by employing probabilistic graph models (PGMs), also known simply as graph models. Graph models are a particularly useful mathematical and computational construct for handling the inherent uncertainty that exists due to the complex, nondeterministic, and not fully observable nature of the real world. In this context, a graph encompasses both the concept from graph theory in discrete mathematics and the abstract data type in computer science. A graph is characterized by a set of nodes (also referred to as vertices or points) connected by edges (also referred to as links or lines). In PGMs, the nodes represent random variables or states, while the edges signify dependencies or transition likelihoods, which are quantified using probabilities. Typically, the states are predetermined and the weights are either known *a priori* due to some theoretical understanding of the system or learned from data, as is the case in neural networks, which are a special case of graph models. In essence, the PGM can be used to model the agent program, making use of the task environment.

Leveraging PGMs offers a relatively efficient and tractable method of representing probabilistic distributions in a visually intuitive manner. That being said, the assembly of these

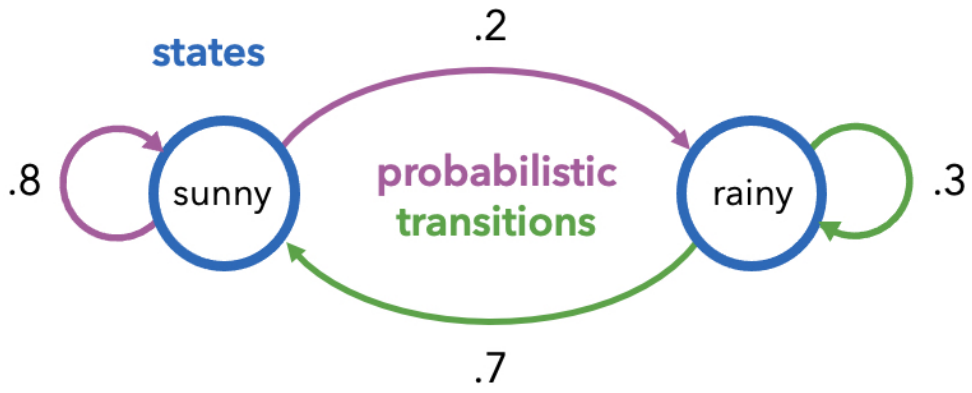

Figure 1.3. Simple 2-state probabilistic graph model for predicting weather where the blue circles represent the state of the weather (sunny or rainy), the purple arrows depict the transition probabilities given the current state is sunny, and the green arrows depict the probabilities associated with the rainy state.

models is a non-trivial task that can get complicated quickly. Despite the breadth of the field, it can broadly be categorized into two distinct types. If the edges are undirected, the model is referred to as a Markov random field or Markov network. On the other hand, if the edges are directed, the model is known as a Bayesian network. Bayesian networks possess the additional property of being acyclic. Machine learning models, such as neural networks, are special cases of Bayesian networks. These models can be intuitively extended to encompass temporal uncertainty by introducing an additional dimension, resulting in graph models associated with each time step. The effective representation of uncertainty through PGMs facilitates tasks such as inference–approximating inference utilizing known distributions and learning of unknown distributions–and simulation.

1.4 Unified Ideology

In the forthcoming chapters, there will be discussions of various agents spanning multiple scales with the goal being to scale specific levels of investigation for use in neuropsychiatric applications. To provide a comprehensive foundation, it is worth discussing each within the rigorous framework of AI design combined with computational neuropsychiatry ideologies

before proceeding further. In line with the previous discussion, all task environments are not fully observable and nondeterministic.

Beginning at the microscopic scale, **Chapters 2 and 3** focus on building theory-driven, efficient graph models of biophysical synaptic reaction-diffusion systems to connect the molecular and cellular levels to networks. This intricate system embodies a multi-agent problem, with the intelligent agents being the particles. As is customary in such scenarios, the system operates within a partially observable environment. While substantial knowledge has been amassed regarding the system's dynamics through *in vivo* and *in vitro* experimentation, it remains fundamentally nondeterministic. Therefore, the states and their probabilistic transitions are provided *a priori*, without learning. The system manifests sequential, dynamic, and discrete characteristics. Markov models act as the PGM framework operating primarily in the spectral domain to assess spatiotemporal dynamics in a computationally-efficient manner.

Both **Chapters 4 and 5** operate on the macroscopic scale. Specifically, **Chapter 4** looks to connect the networks level through systems to behavior using a combined theory and data-driven approach. Here, graphs are employed for segmentation, dimensionality reduction, and as usable biomarkers. Segmentation is done with a pretrained deep learning model, meaning it is a single agent system with previously learned weights. Dimensionality reduction is done using using low-dimensional graph-based manifold learning. Both of these data-driven task environments are episodic, static, and discrete in nature with machines assuming the role of the agent. Lastly, functional connectivity graphs are built for the subjects. These graphs are combined theory- and data-driven, sequential, dynamic, and discrete.

Finally, **Chapter 5** looks solely at scaling data-driven methods at the behavioral level using natural language processing (NLP). The learning agent is a transformer-based automatic diagnostics model, which is static, discrete and single-agent. Typically, NLP is implemented with sequential approaches, but transformers provide a spectral alternative that is episodic and readily scalable.

Overall, the following work seeks to bridge levels of investigation for neuropsychiatry

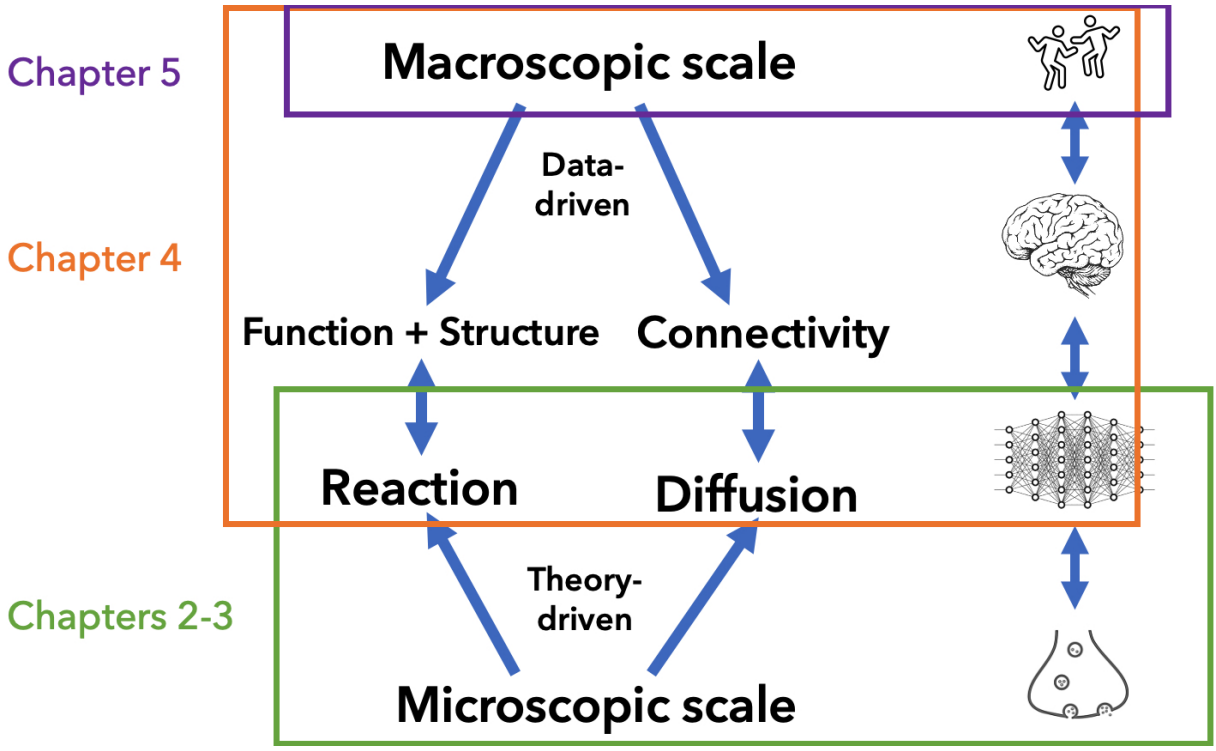

Figure 1.4. Structure of thesis and breakdown into specific chapters.

applications with computationally-efficient scaling acting as a primary point of optimization.

Chapter 2

Efficient Markov Chain Reactions for Scalable Models of Synaptic Transmission

2.1 Introduction

It has been known since the pioneering of computer architecture by John von Neumann that brains are far more effective and efficient in processing sensory information than digital computers, owing to the massively parallel distributed organization of neural circuits in the brain that tightly couple synaptic memory and computing at a fine grain scale [218]. Modern day computers still follow the "von Neumann" architecture where computing and memory are kept separate, incurring severe penalties in computing bandwidth due to the bottleneck in data flow between centralized processing and vast memory. Moore's law's relentless scaling of semiconductor technology, with a doubling of integration density every two years, has allowed the von Neumann architecture to remain fundamentally unchanged since its advent. As the shrinking dimensions of transistors supporting the progression of Moore's law are approaching fundamental limits, it has become essential to consider alternative novel computing architectures to meet increasing computational needs in this age of the deep learning revolution, which itself is driven by advances rooted in a deeper understanding of brain function [196]. At the forefront of this movement are neuromorphic systems, introduced by Carver Mead ([150]) as a solution to these limitations. Neuromorphic engineering looks towards human brains as inspiration for hardware systems due to their highly efficient computational nature. The human brain is regarded as the

pinnacle of efficient computing, operating at an estimated rate of 10^{16} complex operations per second while consuming less than 20 W of power [58, 48]. Therefore, neuromorphic engineering looks to mimic the function and organization of neural structures using hybrid analog and digital systems. This is possible because there is significant overlap in the physics of computation between the brain and neuromorphic engineering (Figure 2.1). In both systems, information is carried in the form of charge, and, in hardware, neuronal membrane dynamics are represented using metal-oxide-semiconductor field-effect transistors (MOSFETs) [149]. In the MOSFET sub-threshold region of operation, electrons and holes are the carriers of current between *n*- or *p*-type channels and behave akin to ions flowing through ion channels that mediate current across the neuronal cell membrane. Fundamentally, these hardware systems share analogous properties to their biological counterparts, including charge stochasticity, diffusion as the primary mechanism of carrier transport, and energy barriers modulated by gating voltage. Paired with Boltzmann distributions of charge, these systems are able to emulate current as an exponential function of the applied voltage, capturing the same biophysics underlying the neuronal dynamics [149, 42].

Since the introduction of neuromorphic engineering, computational models of different complexity have been introduced to describe neuronal dynamics, typically ranging from more detailed and realistic conductance-based Hodgkin-Huxley models to simpler integrate-and-fire models allowing for better scalability. Synaptic connectivity between neurons is of primary concern in the field currently because synaptic strength and plasticity are fundamental to learning and memory in both biological and artificial representations of neural networks [42, 106]. In neuromorphic architectures, synapses instantiate both computation and memory, and a new focus on compact electronic implementations of this computational memory has been emerging recently including the use of memristors [39]. Efficient representation of synapses is a crucial topic of concern as there are roughly 10^4 synapses for each neuron, totalling approximately 10^{16} in the human brain. They are diverse in nature and have highly complex temporal and spatial dynamics, which further complicates their representations [42]. Currently, there is a

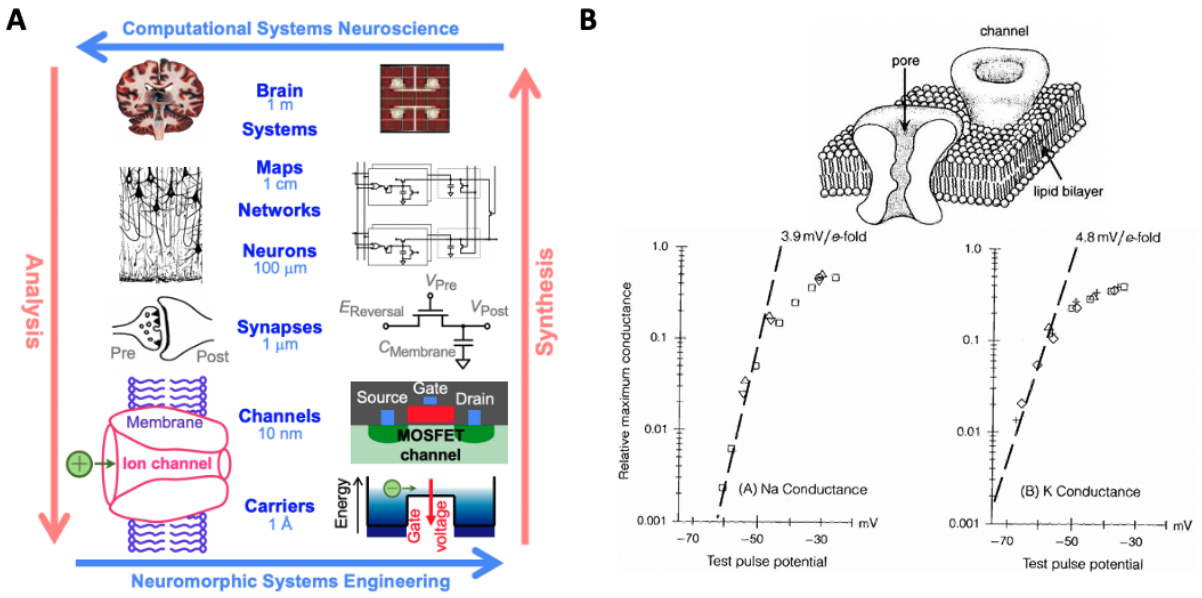

Figure 2.1. **(A)** Computational neuroscience and neuromorphic engineering relation. Reproduced with permission from [58, 48]. **(B)** Thermodynamic equivalence between MOSFETs and ion channels giving current as an exponential function of applied voltage for sodium *(left)* and potassium *(right)* [99, 149].

push for efficient synaptic models while maintaining the intricate dynamical behavior exhibited biophysiologically. Current models include time-multiplexing synapses, analog bistable synapses, and binary synapses to name a few, but the need for scalable and dynamically complex models of synaptic function and transmission is still existent and critical [42, 31].

Modelling synapses is a challenging task due to their intricacy and sheer quantity. As noted above, there are an estimated 10^{16} synapses in the human brain. They vary in function and type, including both chemical and electrical synapses and exhibit behavior spanning multiple different temporal and spatial scales, as well as being highly stochastic in nature [214, 219]. Additionally, synaptic plasticity causes changes in synaptic strength over time associated with learning and memory. Synaptic transmission involves a multitude of mechanisms and molecular components, making simulations including all components not readily scalable. In order to capture the sophisticated dynamics of synapses in a scalable manner, abstractions have to be made according to the research problem in question. The stochastic nature of synapses also makes

large scale simulations more complicated as modelling stochastic processes is typically more computationally demanding. It has been shown in multiple instances that the noise present due to the stochastic variability in synapses is highly integral to synaptic transmission, so this becomes an important feature to maintain [141]. For example, [155] proved that models including ion channel noise in Ca^{2+} channels paired with the existence of a presynaptic mechanism causing random delays in synaptic vesicle availability best capture the interspike interval behavior of auditory nerve fiber models. Additionally, multiple experimental works have found the existence of presynaptic vesicles that are released into the synaptic cleft with some probability [47, 121]. There are multiple similar conclusions found in modelling and experimental results as recently discussed by [148].

Synapses form the connections between neurons and the strength of these connections changes over time, forming the basis of learning and memory in both biological and artificial neural networks. The computations involved in accurately modelling the biophysics of synapses are complex due to the highly nonlinear nature of their dynamics, yet most of the neural network models in use today abstract synaptic strength to a single or small number of scalar values, tuned to a specific task. The learning rule for updating synaptic strength is then typically applied using abstractions of synaptic plasticity such as spike-time dependent plasticity and its causal extensions for scalable real-time hardware implementation [174]. Physical constraints and limitations in VLSI implementations restrict the functional form of synaptic representation. In turn, these abstractions restrict the potential computing power of neuromorphic systems and restrain achievable benchmarks in approaching the functional flexibility, resilience, and efficiency of neural computation in the biological brain. Our work addresses the need for a more biophysically realistic model of the synapse with biologically tunable parameters to represent synaptic dynamics while offering a path towards efficient real-time implementation in neuromorphic hardware.

Synaptic transmission is dictated by a series of events initiated by presynaptic stimulation in the form of action potentials. An action potential causes membrane depolarization which

leads to stochastic opening and closing of voltage-dependent calcium channels (VDCCs) lying on the presynaptic membrane and a resulting influx of calcium (Ca^{2+}) to the presynaptic terminal. Neurotransmitter release is modulated by Ca^{2+} binding to calcium sensors near the neurotransmitter filled vesicles at the active zone, but Ca^{2+} has other fates as it diffuses from the VDCCs. In addition to binding to the Ca^{2+} sensors, it can bind to calbindin (CALB), which acts as a buffer, or it can be removed by plasma membrane Ca^{2+} ATPase (PMCA) pumps. If enough Ca^{2+} is able to bind to the Ca^{2+} sensors, though, then neurotransmitters are released across the synaptic cleft and initiate downstream effects at the postsynaptic membrane [30]. This process of synaptic transmission is the basis of communication in the brain.

Abstracting this for computational efficiency, we created a series of Markov state transitions to realize the system with multiple internal states allowing for a biophysically tunable model of synaptic connectivity implementable in neuromorphic architectures. Markov models have a history of use as a stochastic discrete state alternative to Hodgkin-Huxley type formulations since their introduction [99, 28, 62]. Additional stochastic models have been introduced, including the Gillespie method (1977), which has been used to model neural channel noise [84, 202, 56]. Markov models have also found use in whole-cell models [222]. Further extensions utilize a particle model [119]. The importance of the inclusion of stochasticity in ion channel behavior and synaptic transmission generally cannot be understated. Its inclusion has been demonstrated time and time again in experimental work and is thought to be integral in the form and function of synaptic transmission [148]. This provides an additional complication in modelling synapses and has been handled at various different stages of transmission, including the stochastic models of vesicle release using probabilistically generated quantal components, stochastic models of transmitter diffusion, and stochastic models of receptors [47, 30, 214]. These simulations are computationally expensive due to the high transition rates paired with the small number of transitions necessitating a small timepoint. Specifically, Markov models have shown to be an effective method of modelling ion channels but require high computational cost to effectively do so.

This paper looks to abstract the computationally complex and nonlinear nature of synaptic transmission dynamics in a manner that is efficient and readily scalable for implementation in neuromorphic silicon very large-scale integrated (VLSI) circuits. This is done by introducing an efficient stochastic sampling scheme within a Markov chain representation of the components integral to stochastic presynaptic quantal transmission.

2.2 Materials and Methods

2.2.1 Markov Chain Models

The cascade of events from the action potential stimulus input to the presynaptic neurotransmitter release output can be equivalently modelled as a Markov chain to realize the system with multiple internal states instead of directly tracking all molecules and their kinetics in a computationally complex spatiotemporal 3D reaction-diffusion model. Each internal Markov state is assumed to be dependent solely on the state at the previous timepoint and is conditionally independent of all previous timepoints, simplifying simulations. Therefore, the fully biophysically complex system of synaptic transmission can be abstracted and sampled to create a Markov Chain Monte Carlo (MCMC) simulation which answers the same question of neurotransmitter release utilizing tunable biophysical parameters while providing scalability for implementation in neuromorphic architectures.

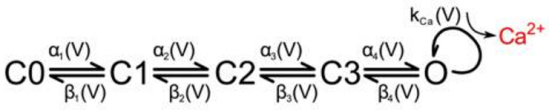

Figure 2.2. State diagram for voltage-dependent calcium channels and resulting calcium influx in the presynaptic membrane. Reproduced with permission from [30].

For any given stimulus input, the VDCCs are assigned transition probabilities between states based on a five-state kinetic model (Figure 2.2) found experimentally and validated computationally with four closed states and a single open state [37, 57, 30]. Prior to the stimulus,

all VDCCs begin in the initial closed state, C0, and the concentration of Ca^{2+} in the presynaptic terminal is at steady-state. The transition probabilities are voltage dependent akin to a Hodgkin-Huxley model where $\alpha_i(V) = \alpha_{io} \exp\left(\frac{V_m}{V_i}\right)$ and similarly $\beta_i(V) = \beta_{io} \exp\left(\frac{V_m}{V_i}\right)$ with parameter values from [37]. The number of open VDCCs at any given moment is used to determine the number of Ca^{2+} entering the presynaptic terminal based on experimental I-V curves and the resulting I-V equation found in [37] and used in [30], which gives the value for k_{Ca}. Ca^{2+} influx is captured by including transitions from the final closed VDCC state, C3, to the open VDCC state and an internal Ca^{2+} generation. Using this, influx of Ca^{2+} is modelled over the entire stimulus input due to the VDCCs opening.

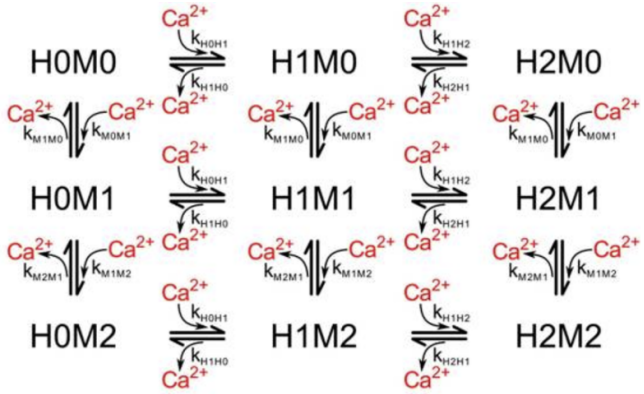

Figure 2.3. State diagram for calbindin binding where HaMb describes the a^{th} high-affinity binding state and the b^{th} medium-affinity binding state. Reproduced with permission from [30].

Once Ca^{2+} has entered the presynaptic terminal, much of it binds to CALB, which acts as a buffer and primarily modulates the amount of Ca^{2+} that is able to reach the Ca^{2+} sensors at the active zone. The state transitions are reversible first-order reactions, thus transition probabilities are dependent on the free Ca^{2+} in the system and updated as that amount changes over time. CALB has four binding sites, two of high affinity and two of medium affinity, leading to a nine-state Ca^{2+} concentration-dependent kinetic model (Figure 2.3) [159]. By modelling the binding and unbinding of Ca^{2+} to CALB as a loss or gain of free Ca^{2+} respectively, Ca^{2+} transients can also be elucidated.

Our Markov chain is a discrete-state chain in discrete time. Markov chains are modelled by a probability that the chain will move to another state given its current state and is conditionally independent of all previous timesteps. The probabilities are by nature only dependent on the current state of the Markov chain. The probability of the state of a molecule X can typically be predicted for a certain timepoint $t + \Delta t$ as some particular state x_j using the states at all previous timepoints from the start of the simulation, $t = 0$, to the timepoint just before that in question, t. For a Markov chain simulation solely dependent on the previous timepoint, it is possible to predict the probability that a molecule is in a given state, x_j at the timepoint $t + \Delta t$ using solely the state of the single timepoint just before, X_t, which is known to be a particular state x_i. Thus the probability of the molecule being in state x_j given that at the previous timepoint it was in state x_i is given as P_{ij}. Succinctly, this is written as

$$P(X_{t+\Delta t} = x_j | X_t = x_i, X_{t-\Delta t}, ..., X_0) = P(X_{t+\Delta t} = x_j | X_t = x_i) = P_{ij} \tag{2.1}$$

For state transitions, the probability of transitioning to an adjacent state is the transition rate inherent in the system (k_{ij} for the transition from state i to state j, and k_{ij} is not necessarily equal to k_{ji}) times the timepoint, Δt. In the case of CALB transitions, this is further multiplied by the amount of free unbound Ca^{2+} for forward reactions as it is a first-order reaction. For the VDCCs, the transition rates are the α, β, and k_{Ca}. The probability that a molecule stays in its current state is the sum of the probabilities it transitions to an adjacent state subtracted from unity. For a multi-state system, this gives a transition probability matrix for the likelihood of transition from a given state at the current timepoint to any other state at the next timepoint. This matrix is sparse, with nonzero probabilities only for adjacent states to which a transition is possible. In the case of the five-state VDCC system, this gives the probability of a transition from state i to to state j as

$$P_{ij} = \begin{cases} k_{ij}\,\Delta t & j = i \pm 1 \\ 1 - \sum_{k=0}^{N_{adj}} P_{ik} & j = i, k \neq i \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

where transitions to adjacent states are given by the transition rate k_{ij} times the timepoint, Δt; the probability of staying in the current state is the sum of probabilities of adjacent state transitions subtracted from unity, where N_{adj} is the number of possible adjacent states. The probability of transitioning to a non-adjacent state is set to zero.

Typically Markov state transitions are modelled via a discrete inverse transform method, where given a random variable X, the transition probabilities P_{ij} describe a partition of unity (Figure 2.4). Therefore, we can generate a random number uniformly, $R \sim U(0,1)$ and map it onto discrete values of X. For example, in a two state system, $X_j = 0$ if $R \leq P_{i0}$ or $X_j = 1$ if $P_{i0} < R \leq P_{i0} + P_{i1} = 1$. This involves searching the state space for the next state given the current state for each molecule in the system at each timepoint, which can be a slow process for systems with a large number of states and molecules.

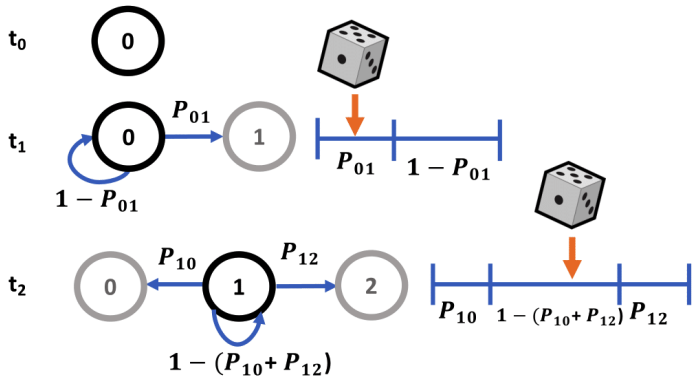

Figure 2.4. Markov sampling scheme for state transitions using partitions of unity.

Here we have implemented a more efficient MCMC sampling strategy involving sampling from a multinomial distribution. Therefore, instead of sampling from a uniform distribution for

each of n molecules, we sample from a multinomial distribution once for each state, using n molecules as the number of experiments, where $X \sim \text{Multi}(n, p_1, \ldots p_k)$. For simulations where the number of possible states is less than the number of molecules, this is a more efficient sampling strategy. Since we are particularly interested in the number of molecules in each state at each timepoint, this is an effective approach. Multinomial sampling thus describes the distribution of the n experiments across k possible outcomes each with a probability of p_k, where n_k is the number of experiments falling into the k^{th} outcome following a probability mass function of

$$f(n_1, \ldots, n_k; n, p_1, \ldots p_k) = \frac{n!}{n_1! \ldots n_k!} \, p_1^{n_1} \ldots p_k^{n_k} \tag{2.3}$$

In our model, for each state i, we have an initial number of molecules in that state at a given timepoint t, or $n_{i,t}$. As previously described, there exists a probability that the molecules will transition to any state at the next timepoint, including staying in the original state given by P_{ij}. Thus, to determine the distribution of molecules $n_{i,t}$ across all states at the next timepoint, we sample from a multinomial distribution according to

$$X_{i,t+\Delta t} \sim \text{Multi}(n_{i,t}, P_{i1}, \ldots P_{ik}) \tag{2.4}$$

for k possible states. We do this sampling for each state at each timepoint and sum accordingly. This expedites computation by only requiring a single computation at each timepoint, sampling the distribution of all n molecules at once. Algorithm 1 highlights the pseudocode for this process.

Algorithm 1: Markov Multinomial Reaction Sampling

Result: The number of molecules in each state at each timepoint for a simulation

initialize number of states;

initialize matrix of number of molecules per state per timepoint;

for *each timepoint* **do**

  for *each state* **do**

    calculate transition probability according to eqtn 2.1;

    sample multinomial distribution with current number of molecules in the state and transition probability;

  end

  update number of molecules in each state using samples;

end

Markov simulations for the VDCCs were run for 65 VDCCs all starting in the closed state, C0. CALB molecules were initiated in the different binding states according to the steady-state concentration of Ca^{2+} and at a baseline concentration of $4.5\,10^{-5}$ M. All simulations were run for 10 ms with a timestep of 1 μs. The simulations were repeated 1,000 times to obtain an average and standard deviation. Markov simulations were implemented using Python.

2.2.2 MCell Models

MCell is a modelling software that uses spatially realistic 3D geometries and Monte Carlo reaction-diffusion modelling algorithms, which allows for biophysically realistic simulations of high complexity as it specifically tracks the state of every molecules in space and time [30]. Due to the accuracy and specificity, it provides a ground truth for biological simulations but does so at the cost of computational complexity.

To validate and compare our Markov models of synaptic transmission, we built a biophysically realistic stochastic 3D reaction-diffusion system with all major components for presynaptic vesicle release variability in response to a stimulus input (Figure 2.5) based on the models of

[158] and [30]. The model includes realistic geometry for a CA3-CA1 *en passant* synapse focusing primarily on the presynaptic Schaffer collateral axon of a CA3 pyramidal cell found in the hippocampus with parameters set from experimental data [158, 30]. The CA3-CA1 synapse was chosen for the simulations as it is highly studied experimentally and is important for learning and memory. Furthermore, CA3-CA1 synapses are relatively small, containing one to two neurotransmitter release zones. Release from this region is also known to be highly stochastic in nature, necessitating the inclusion of stochasticity in biologically realistic models [158]. All kinetics and parameters match those used for the equivalent Markov models.

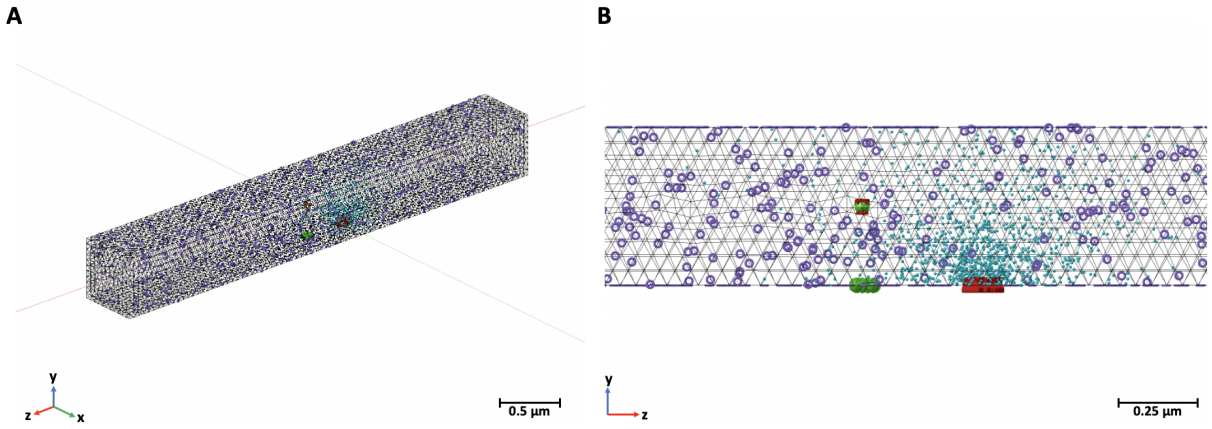

Figure 2.5. MCell model for synaptic transmission containing voltage-dependent calcium channels (*red*), calcium (*blue*), calcium sensors (*green*), and plasma membrane calcium ATPase pumps (*purple*). **(A)** Entire box representing one vesicular release site in CA3 and **(B)** a close-up of the release site.

The MCell model includes the canonical presynaptic geometry for an average CA3-CA1 synaptic terminal as a rectangular box measuring 0.5μm by 0.5μm by 4μm. This box captures the dynamics of a single synaptic active zone, referring to the region on the presynaptic membrane specialized for neurotransmitter release. Initially, the terminal contains the CALB buffer, steady-state Ca^{2+} concentration, PMCA pumps, VDCCs and Ca^{2+} sensors modulating neurotransmitter release [158]. The detailed diffusion dynamics and kinetics of these systems are based on experimental data and have been discussed in further detail in [30]. The active zone is based on that of an average presynaptic active zone containing seven docked neurotransmitter vesicle

release sites. The VDCCs, of type P/Q, are stationed at a biophysically realistic distance from the active zone. They transition states in response to the membrane depolarization. The location, number, and Ca^{2+} conductance of the VDCCs is replicated from experimental data [158]. PMCAs are homogenously placed across the presynaptic membrane while CALB molecules are in a uniform concentration within the volume. This is a flexible architecture that can respond to any stimulus input and allows for monitoring of the states of each molecule in the system. The MCell CA3-CA1 synaptic transmission models were originally created and validated in [158] and [30]. To compare with the Markov models, we used the same single action potential stimulus.

MCell models were also run 1,000 times for 10 ms with a timestep of 1 μs.

2.3 Results

2.3.1 Voltage-Dependent Calcium Channels

The efficient Markov chain implementation has strong agreement with the full MCell model in terms of the internal state transients in response to an external stimulus. The number of closed VDCCs (state C0) decreases over the duration of the stimulus (Figure 2.6A). The internal states (C1-C3) subsequently increase and decrease as the membrane voltage increases and the forward rates for the VDCCs increase (Figure 2.6B-D), leading to an exponential increase in the open VDCCs while the membrane depolarizes. Figure 2.6E shows the fraction of open VDCCs over time in response to the action potential, which controls the amount of Ca^{2+} influx to the system. At the maximum membrane potential, almost all VDCCs are in the open state. As the membrane repolarizes, the reverse reaction rate constants increase, and the VDCCs close. This leads to another increase and decrease in the internal VDCC states as the receptors go from their open to resting closed state (C0).

In its open state, VDCCs allow for the probabilistic influx of Ca^{2+} through the channels into the presynaptic bouton. This is exemplified in Figure 2.6F, where there is an increase in the

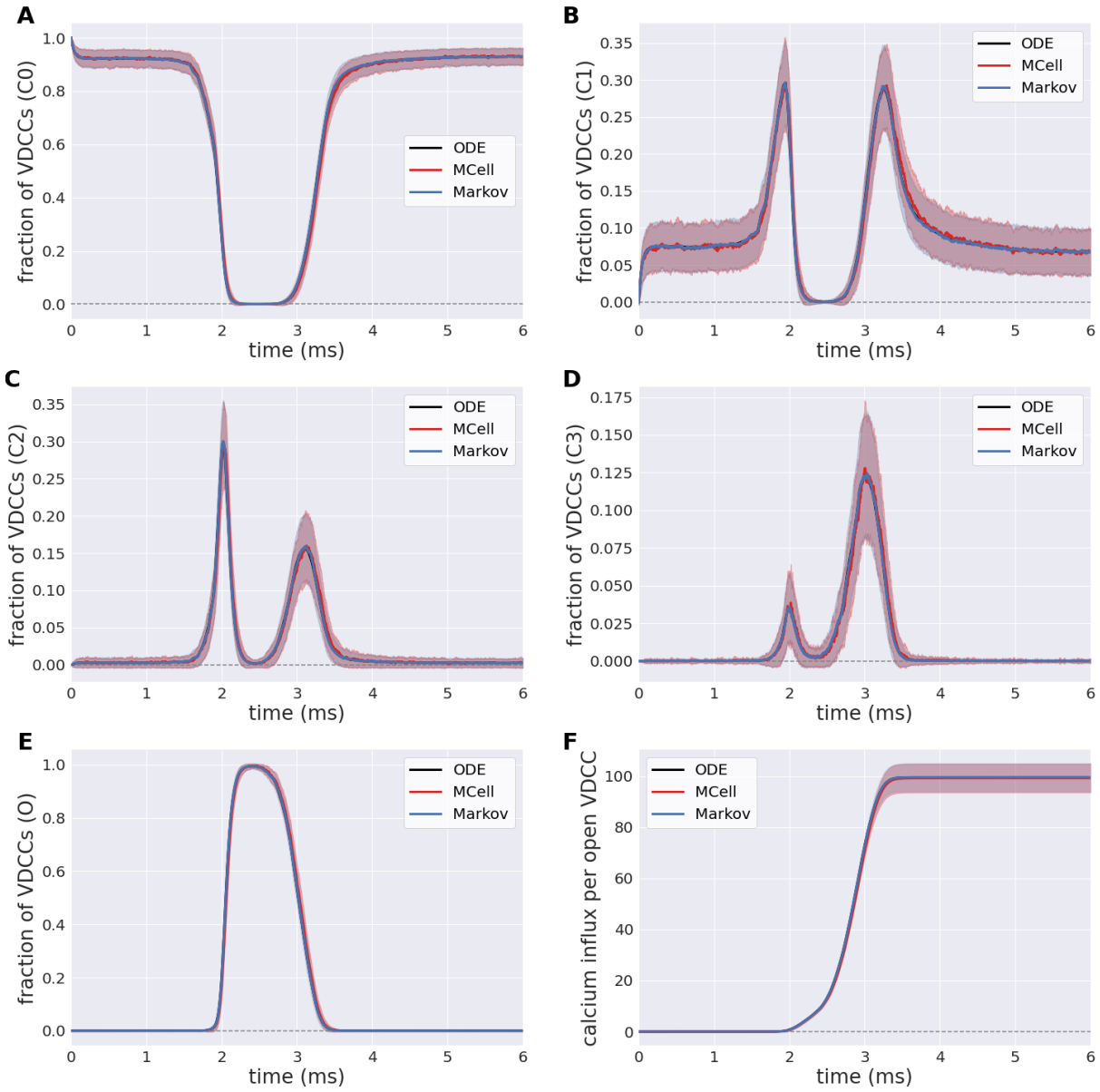

Figure 2.6. Fraction of voltage-dependent calcium channels (VDCCs) in each state: **(A)-(D)** internal closed channel states, C0-C3, and **(E)** open channel state, O. **(F)** Calcium influx through open VDCCs due to action potential stimulus for stochastic MCell, Markov, and deterministic ODE simulation.

Ca^{2+} influx through the open VDCCs over the course of the stimulus. Again, there is strong agreement between the more computationally complex MCell model and the computationally efficient Markov equivalent model.

2.3.2 Calbindin Buffer

Simulations of homogeneous Ca^{2+} and CALB were run using the MCell, Markov and deterministic simulation schemes. In the presence of Ca^{2+}, the forward binding reaction is heavily favored, and this is highlighted in Figure 2.7A where free Ca^{2+} exponentially decreases. A similar transient is apparent for the unbound state of CALB, as it quickly transitions to different stages of high and medium binding (Figure 2.7B). Over the course of the simulation, all the free Ca^{2+} is removed from the system, and CALB states reach a new steady-state where there is still CALB. Similarly, the fully bound state, H2M2, rapidly increases and reaches a new steady state that is still only 1% of all CALB (Figure 2.7C). This is due to the high concentration of CALB in the presynaptic bouton. Even once all the Ca^{2+} is in a bound state, there is still plenty of unbound or partially bound CALB remaining in the system. CALB acts as a strong buffer allowing for Ca^{2+} storage and asynchronous neurotransmitter release, so this and slow unbinding rates become an important feature of CALB. The rapid extent to which Ca^{2+} binds to CALB shows the impact of buffering on Ca^{2+}'s ability to diffuse and bind to the Ca^{2+} sensors in the active zone. The inclusion of CALB at such high concentrations becomes a key feature of maintaining the relatively low release rates of neurotransmitters even in the presence of a stimulus.

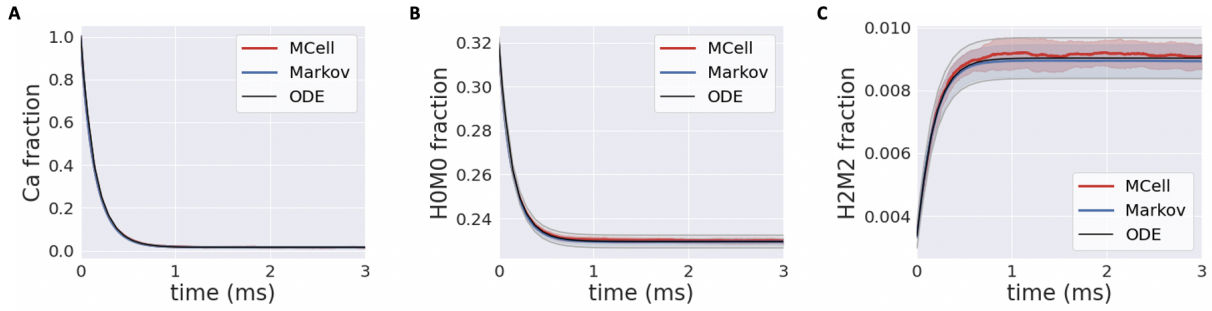

Figure 2.7. **(A)** Transients for homogeneous calcium-calbindin buffer binding in the presynaptic bouton for free calcium; **(B)** the calbindin state, H0M0; and **(C)** the fully bound calbindin state, H2M2 in all simulation types.

2.3.3 Complexity Analysis

MCell uses a scheduler which allows for only making changes to the scheduled particles, though in the worst-case, this still scales with the total number of particles in the simulation, n, where n_{VDCC} is 65 and n_{calb} is 2.7×10^4. It also scales with the length of the simulation, t, described by the number of time points for a discrete simulation. The simulations for the VDCC and CALB both use 10k timepoints. At each timepoint, a particle can transition to any of its adjacent or branched states, b, which is similarly described by a fan-out factor in electronic implementation. From the VDCC kinetic model described in Figure 2.2, b_{VDCC} is 1-2 depending on the state while the CALB kinetic model in Figure 2.3 gives b_{calb} of 2-4. The overall time complexity for MCell is $\mathcal{O}(bnt)$. The classical Markov representation tracks every particle. It also searches through the space of each adjacent state for potential state transitions at each time point. Therefore, classical Markov implementation similarly results in an $\mathcal{O}(bnt)$ time complexity, or $\mathcal{O}(nt \log_2 b)$ at best for implementation with an efficient search algorithm. Both the multinomial Markov model and the Euler ODE implementation describe the system in terms of the number of molecules in each state leading to a dependence on the total number of states, $s \geq b$, rather than the total number of particles. The total number of states for VDCC is 5 (Figure 2.2) while the number of states for CALB is 9 (Figure 2.3). Due to efficient sampling methods, the multinomial Markov method is independent of the number of adjacent states, leading to a time complexity of $\mathcal{O}(bst)$ for both the multinomial Markov and Euler ODE methods. Thus, our stochastic multinomial Markov model is equally amenable to large scale simulations as the deterministic ODE method that is typically used in simulations involving more synapses.

The traditional Markov sampling model and the MCell representation store the molecular states in bits for each particle as well as the states adjacent to the current state, leading to a space complexity of $\mathcal{O}(bn \log_2 s)$ The efficient Markov model and ODE solution both simply store the number of molecules represented by bits in each state at each timepoint as well as the branched states resulting in a space complexity of $\mathcal{O}(bs \log_2 n)$. There exists a trade off here

Table 2.1. Space and time complexity for the various simulation strategies.

Complexity	Time Complexity	Space Complexity
MCell	$O(b\ n\ t)$	$O(b\ n \log_2 s)$
Standard Markov	$O(b\ n\ t)$	$O(b\ n \log_2 s)$
Multinomial Markov	$O(b\ s\ t)$	$O(b\ s \log_2 n)$
ODE (Euler)	$O(b\ s\ t)$	$O(b\ s \log_2 n)$

n (total number of molecules), s (total number of states), b (number of branching states), t (number of timepoints)

between the number of particles in each state compared to the number of states where one is stored directly and one is stored as an index. Thus, for simulations where the number of states is less than the number of particles, the multinomial Markov model is an efficient representation of the system, which is typically the case for biochemical simulation. MCell is more efficient with large state-space systems, but the number of states could be sparsified in a multinomial Markov representation by implementing dynamic instantiation and annihilation of states. Additionally, unseen or rarely seen states could be ignored by truncating based on probability of a particle being in that state. This would functionally decrease the number of states in the system allowing for use of the multinomial Markov simulation method.

2.3.4 Benchmarks

Runtime and total floating point operations were used as metrics for comparison between the simulation methods (Table 2.2). We also looked at the number of pseudorandom number generator calls (n_{PRNG}) between the simulations as this provides a metric to elucidate the differences observed in execution time between the simulations. Here we compare MCell, the standard Markov model, and the multinomial Markov stochastic models. The deterministic Euler solution is included as well for a non-stochastic comparison. Again, it is valuable to note the importance of stochasticity in these models. Significant work has shown the necessity of stochasticity in models of synaptic transmission in order to match experimental work. It has been demonstrated that deterministic models at this scale generally underestimate quantal release

as concentration fluctuations are not captured [214, 148, 30]. Thus while deterministic ODE models provide efficient simulation techniques, they are not able to capture the full complexity of the dynamics of synaptic transmission, hence motivating the need for an efficient stochastic model.

In the VDCC simulations, the multinomial sampling MCMC model has a runtime on the order of the forward Euler deterministic solution. The MCell and the standard Markov stochastic models exemplify a runtime an order of magnitude higher. The number of operations is also higher for the MCell and the standard Markov models compared to the multinomial Markov model. The standard Markov case generates a pseudorandom number for each molecule and each timestep, so n_{PRNG} is equivalent to the number of molecules multiplied by the number of timesteps, $n_{VDCC}t$. In the multinomial Markov simulation, a pseudorandom number is generated for each occupied state and possible branching points at each timepoint, which gives $(bs)_{VDCC}t$ in the worst-case scenario. Therefore, n_{PRNG} is smaller for the multinomial case as long as $(bs)_{VDCC} < n_{VDCC}$, which is always the case here.

Table 2.2. Benchmarks for different simulation types for both voltage-dependent calcium channel and calbindin binding simulations.

VDCC	Runtime (sec)	No. Pseudorandom Number Generator Calls (n_{PRNG})	Total Floating-Point Operations (FLOPs)
MCell	139.20	1.5×10^9	1.7×10^{11}
Standard Markov	109.16	6.5×10^5	7.0×10^7
Multinomial Markov	9.40	1.0×10^5	1.3×10^7
ODE (Euler)	10.05	0	7.7×10^6
Calbindin	**Runtime**	**No. Pseudorandom Number Generator Calls (n_{PRNG})**	**Total Floating-Point Operations (FLOPs)**
MCell	29.34	6.9×10^7	8.7×10^9
Standard Markov	76.78	2.7×10^8	1.4×10^{10}
Multinomial Markov	2.37	3.6×10^5	2.7×10^7
ODE (Euler)	0.72	0	1.6×10^6

For the CALB model, the multinomial Markov method is again an order of magnitude faster than the MCell model although it is also an order of magnitude slower than the deterministic

model. The standard Markov model is an order of magnitude slower than the multinomial model. The operations are also fewer for the multinomial case than the standard case. Again, the standard Markov case gives n_{PRNG} equal to $n_{calb}t$ while the multinomial Markov simulation is $(bs)_{calb}t$. Again we see a smaller n_{PRNG} in the multinomial case because $(bs)_{calb} < n_{calb}$ even in the worst-case scenario where b is at its maximum value. Simulations are not currently optimized on hardware suggesting opportunities for further decreases in runtime. Overall, the multinomial Markov simulation provides a computationally efficient alternative to stochastic MCell simulations while maintaining the biological accuracy.

2.3.5 Neuromorphic Implementation

Thermodynamic foundations of neuromorphic engineering suggest direct biophysical implementation of populations of ion channels with individual stochastic opening and closing of gating variables driven by thermal noise fluctuations [149]. So it seems only natural to consider implementations using stochastic ODEs describing the rates of reaction kinetics under additive white Gaussian noise (AWGN):

$$\frac{dX_{i,t}}{dt} = \sum_j k_{ij} X_{j,t} + \xi_{i,t} \tag{2.5}$$

where $\xi_{i,t}$ is normally distributed with zero mean and variance dependent on the magnitude of $X_{i,t}$. Fully parallel, continuous-time analog implementation of reaction kinetic rate equations of the type (2.5) have been demonstrated in micropower integrated circuits, e.g., cytomorphic chips in BiCMOS integrated silicon technology [225]. Abundant intrinsic noise present in these micropower cytomorphic circuits can serve as AWGN, although its magnitude is determined by thermal processes that are hard to control and other non-white Gaussian sources of intrinsic noise contribute strongly colored low-frequency spectra. Thus, discrete-time implementation of the ODEs (2.5) through Euler integration on a digital computer offers greater control over the shape and amplitude of the AWGN distribution, limited by the quality of pseudo-random number

generation by deterministic algorithms.

Although purely digital algorithmic implementations go against foundational principles of neuromorphic engineering rooted in the physics of computation [149], the convenience of their programmability and reproducibility have made ODE-based digital emulation platforms such as Loihi a popular choice among more software-focused neuromorphic computer scientists [65]. The computation involved in such discrete-time ODEs (2.5) can be performed at varying degrees of parallelism in custom or reconfigurable digital hardware, with the variables X_i being updated in sequence through time-multiplexing a single processing core in one extreme case, or all X_i updated in parallel with dedicated processing elements for each in the other extreme case. Ultimately in practice, the energy efficiency is relatively independent of the compute implementation, and depends more critically on the available memory bandwidth in accessing the rate parameters defining network connectivity [173]. In essence, discrete-time Euler-integration ODE implementation of (2.5) amounts to sampling from a normal distribution

$$X_{i,t+\Delta t} \sim (n_{i,t+\Delta t}, \sigma_{i,t}) \tag{2.6}$$

with mean and standard deviation

$$n_{i,t+\Delta t} = \sum_j P_{ij} n_{j,t} \tag{2.7}$$

$$\sigma_{i,t} = \sqrt{n_{i,t}(n - n_{i,t})} \tag{2.8}$$

incurring computational complexity $\mathcal{O}(bst)$ (Sec. 2.3.3).

More fundamentally, the main disadvantage of implementing stochastic ODEs (2.5) or their discrete-time digital versions (2.6) is that they are primarily based on the Central Limit Theorem for very large number of variables, $n \to \infty$. As such, they have limited accuracy in approximating the reaction kinetics in systems with smaller numbers of molecular variables. While one may be tempted to assume that molecules are always excessively abundant, this

is not typically the case since reactions are rate limited by the least abundant of reagents. Low numbers in molecular dynamics are prevalent in biologically relevant settings, giving rise to significant amounts of biological noise that are critical in neural dynamics, the highly stochastic quantal release of neurotransmitter in synaptic transmission. Thus, there is need for a mathematical description of stochastic synaptic transmission dynamics able to capture the accuracy in simulations with relatively small numbers of variables.

Here we have shown that our multinomial Markov alternative, which directly samples the variables from the multinomial distribution (2.4) rather than the limiting normal distribution (2.6), produces accurate results for any value of n while offering nearly identical implementation complexity $\mathcal{O}(bst)$ (Sec. 2.3.3). Hence we see the Markov chain abstractions of reaction kinetics not only as a means to approach biophysical realism in modeling molecular cellular dynamics without molecular-scale representation, but also as a means towards efficient neuromorphic hardware without biophysical compromise. The key point is that the computational complexity of implementing our multinomial Markov model is essentially identical to that of stochastic ODEs (see Table 2.1), whether in software executing serially on a von Neumann programmable digital computer or in massively parallel digital or analog hardware. Hence, the neuromorphic circuit designer tasked to implement brain-inspired models of information processing faces an easy choice: more bio-realistic models that account for detailed stochasticity in reaction kinetics incur the same resource utilization and energy costs, and use similar design principles, as their stochastic ODE approximations.

In addition to field-programmable gate array (FPGA) reconfigurable [173] or custom-integrated neuromorphic programmable [65] instantiations in digital hardware, we envision physically neuromorphic instantiations in micropower analog continuous-time compute-in-memory hardware that obviate sampling from posterior distributions and directly implement Markov state transitions through parallel implementation of sum-product rules with self-normalizing probabilities [49, 50], at throughput density and energy efficiency that are orders of magnitude higher than today's most advanced general-programmable computational platforms.

2.4 Discussion

The goal of this work was to create a more computationally efficient model of biologically realistic synaptic transmission for use in large-scale neuromorphic systems. We created a multinomial MCMC sampling strategy for capturing the internal states of vital molecules in the system in response to stimulus where transition probabilities could be voltage- or concentration-dependent, and the next timestep could be predicted solely using the current timestep. This scheme was implemented to capture the dynamics of the stochastic opening and closing of VDCCs through multiple internal states as well as the resulting Ca^{2+} influx into the presynaptic bouton through the open VDCCs. Once Ca^{2+} has entered the presynaptic terminal, we also simulated Ca^{2+} binding to the CALB buffer which modulates Ca^{2+} levels in the bouton, directly impacting the amount of Ca^{2+} that reaches the Ca^{2+} sensors in the active zone. This amount impacts the neurotransmitter release from the presynaptic side and the resulting effects on the postsynaptic side.

All simulations were modelled using the multinomial Markov sampling method as well as a typical Markov sampling method and compared to highly detailed 3D geometric stochastic reaction-diffusion simulations done using MCell. The Markov simulations show agreement with the MCell simulations for the system dynamics including the number of open VDCCs and Ca^{2+} influx in response to an action potential stimulus as well as the binding of Ca^{2+} to the CALB buffer. Differences are observed from the deterministic solution to the stochastic simulations implying the importance of stochasticity in these simulations to capture more biologically-realistic systems.

Exemplified by runtime and total number of operations, the multinomial MCMC method of simulations was shown to be more efficient than the standard Markov model while also being faster than the MCell equivalents. This is hopeful for scaling these biologically-realistic models to large-scale systems while maintaining biological tunability.

Next steps involve modelling the remaining kinetics in a similar fashion including the

binding and removal of Ca^{2+} by the plasma membrane Ca^{2+} (PMCA) pumps as well as binding to the Ca^{2+} sensors. In addition, to capture the diffusion of Ca^{2+} through the presynaptic terminal but specifically to the Ca^{2+} sensors at the active zone, a diffusive kernel must be included to the system. Upon inclusion of these elements, the entire process from stimulus to neurotransmitter release can be captured as a series of Markov chains leading to powerful implications for synaptic transmission modelling. The whole synapse can be included as well with the inclusion of a diffusive kernel across the synaptic cleft as well as downstream effects on the postsynaptic side, of which many mirror similar kinetics and dynamics as the presynaptic side leading to a natural extension of this modelling framework. The resulting system would be a biologically tunable model of synaptic transmission for any stimulus input in a highly efficient manner. This opens the door for large-scale implementations of synaptic transmission and learning readily implementable into neuromorphic architectures with strong biological realism.

Through the utilization of Markov-based abstractions applied to biophysically realistic 3D reaction-diffusion models of a chemical synapse, we have created a compact and efficient internal state space representation of synaptic transmission. This is in response to the challenge presented by the high dimensionality and complex nature of molecular-scale interactions in synapses and across scales making implementation in very large-scale systems previously unattainable. The model is directly amenable to efficient emulation in parallel neuromorphic hardware systems while maintaining biophysically relevant and interpretable parameters that are readily tunable. This opens the door towards neuromorphic circuits and systems on very large scale that strike a greater balance between integration density and biophysical accuracy in modelling neural function at the whole-brain level.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This work was supported by a NSF Graduate Research Fellowship to Margot Wagner, the National Science Foundation, and the Office of Naval Research.

Acknowledgment

Chapter 2, in full, is a reprint of the material as it appears in Frontiers in Neuroscience 2021. Wagner, Margot; Bartol, Thomas M.; Sejnowski, Terrence J.; Cauwenberghs, Gert., *Markov Chain Abstractions of Electrochemical Reaction-Diffusion in Synaptic Transmission for Neuromorphic Computing*, Frontiers, 2021. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Grid-Free Scalable Synaptic Reaction-Diffusion Models Using Spectral Methods

3.1 Introduction

Synapses form the connections between neurons and are electrochemical in nature. They form the basis of learning and memory and are the building blocks for neural plasticity. Messages pass between neurons using neurotransmitters, which also affect the neuron connection strength. Furthermore, as the location of neurotransmitter activity, it is of primary importance for neuropharmacological intervention strategies. As such, the ability to scale biophysically realistic models of synaptic transmission is necessary to model the impact of cellular dysfunction and pharmaceutical intervention at scale. Unfortunately, synaptic transmission involves a multitude of mechanisms and molecular components, making biophysically-accurate simulations not readily scalable.

Creating accurate models of biophysical properties at the single-neuron level provides the foundation for understanding network dynamics. However, these detailed neuron descriptions require significant computational resources. Consequently, when studying large neural circuits, researchers often resort to using simplified models that overlook much of the biological complexity, representing neurons and synapses in an artificially simplistic manner. Therefore, there is a need to develop neuron models that can maintain essential biological characteristics while minimizing computational demands allowing for scaling to the network level. A significant

impediment to the implementation of large-scale, biologically realistic neural network lies in the intricate nature of these reaction-diffusion synaptic transmission systems.

Conventionally, when dealing with neural network models, stochasticity in neuronal reaction-diffusion systems is often neglected. However, extensive evidence highlights the crucial role played by stochasticity in neuronal information processing across various scales [70]. For instance, in experimental scenarios, cortical neurons exhibit slight variations in responses to time-dependent stimuli between repeated trials of the same input [203]. Furthermore, even in tightly controlled experiments where the input current is directly injected into the soma, the stochastic opening and closing of ion channels introduce channel noise, resulting in subtle fluctuations in the membrane potential and the precise timing of output spikes. Neurons also display subthreshold membrane potential fluctuations, which are not captured in deterministic simulations despite their demonstrated contribution to sensory signal processing and synchronization among neighboring neurons. The development of computational models to include stochasticity at the network and systems level include the utilization of noise-driven harmonic oscillators to capture oscillatory dynamics, but the source of the stochasticity is the noise rather than of biophysical origin. Regardless, the inclusion of stochasticity is an important consideration in these models.

Reaction-diffusion models describe the dynamics of chemical reactions and diffusion processes in space and time using partial differential equations (PDEs). Simulating these systems involves solving PDEs numerically with the primary methods used being grid-based or meshing schemes, including finite difference, finite element, and finite volume methods. All of these techniques involve discretizing space and solving the PDE in small domains then modeling simulations using time stepping. in order to accurately capture the intricate nature of molecular diffusion, it is imperative to have sufficiently small discrete volumes. Unfortunately, this approach faces challenges when dealing with large spatial volumes as the number of nodes rapidly increases, leading to computationally expensive models. Consequently, as the length and size of the simulation expand, both of these methods become increasingly burdensome in terms of computational resources. The issue becomes particularly formidable when scaling synaptic

reaction-diffusion processes as there are trillions of synapses in the human brain. The methods mentioned are deterministic but can easily be made stochastic by adding noise or sampling probabilistic distributions.

Simulations incorporating stochasticity in neural networks involve an approximation of the continuous diffusion process through discretization of both the state and time domains. Diffusion, being inherently spatiotemporal, necessitates modeling approaches that effectively capture its dynamics. Two primary methods have emerged for modeling diffusion. The first method involves approximating the diffusion process through a random walk with infinitesimally small time steps, which the most commonly employed approach. Mathematically, this approximation entails constructing a random walk with very small displacements and time intervals approaching zero. The accuracy of this algorithm relies on the chosen value for the time interval. Diffusion occurs with probabilistic movements to adjacent spatial locations. The second method involves employing a discrete approximation to the stochastic differential equation (SDE) that describes the diffusion process where the SDE is governed by Fick's law of diffusion with an added noise term. The discretized version of the SDE can be conceptualized as a random walk characterized by normally distributed displacements.

Spectral methods offer an alternative approach to solving dynamical systems, characterized by representing the solution as a summation of specific "basis functions", as introduced by Steven Orszag [167]. For example, Fourier series utilize sinusoids as basis functions. The coefficients of these basis functions are then determined to satisfy the given differential equation. Various techniques fall under the purview of spectral methods, including Fourier series methods for problems involving periodic geometries, polynomial spectral methods for both finite and unbounded geometries, and spectral iteration methods for efficiently solving steady-state problems (add citations). While finite element methods rely on basis functions that are only nonzero within localized subdomains, spectral methods employ basis functions that have nonzero values across the entire domain. This global scope allows spectral methods to capture variable dynamics comprehensively, in contrast to the local nature of finite elements. While this characteristic

may pose limitations depending on the specific application, it proves advantageous for scaling biological simulations.

An alternative approach to modeling synaptic transmission would be to create an efficient graph representation of the components in synaptic transmission that would allow for a biophysically meaningful yet scalable models (Fig. 3.1). The system would then act as an input-output system with tunable experimental parameters allowing for elucidation around synaptic connectivity changes, including in the presence of pharmaceuticals. Specifically, it could be used to interrogate the effect of changes at the molecular scale on neural network connectivity. The idea is that of a lumped compartmental model where the focus is on amounts of particles in particular states rather than individually tracking each component in the system. This can be done by making use of spectral methods.

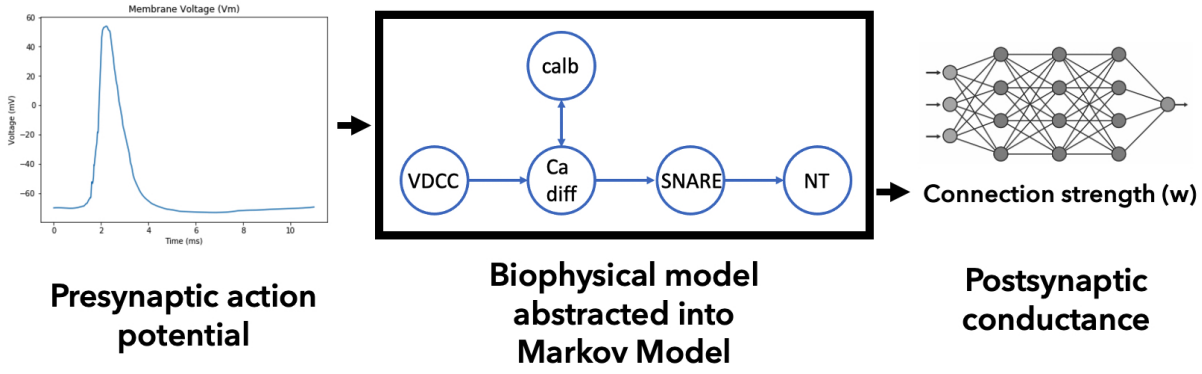

Figure 3.1. Synaptic transmission input-output system using efficient graph representations to model molecular components.

The biological system of interest is the synapse, specifically the presynaptic terminal. When an action potential reaches the presynaptic terminal, it is depolarized, leading to influx of calcium (Ca^{2+}) into the terminal. Neurotransmitter release is modulated by Ca^{2+} sensors in the active zone region. Therefore, release is impacted by calcium dynamics. In addition to the diffusive profile of Ca^{2+}, its dynamics are primarily affected by calbindin (CALB), which is a volumetric buffer that diffuses through the terminal. CALB is at high concentrations with fast forward binding meaning it quickly binds to the available Ca^{2+} (CaCALB) and acts as a

store for neurotransmitter release on longer time scales. Calcium dynamics therefore directly mediate neurotransmitter release with modulation according to reaction-diffusion with the CALB volumetric buffer.

3.1.1 Eigenmode Decomposition as a Dimensionality Reduction Technique

Eigenvectors or principal components, are a fundamental concept in linear algebra and are widely used in various fields. They represent the characteristic patterns or modes of variation in a system or dataset and are orthogonal in nature, meaning each eigenvector corresponds to a distinct and non-overlapping pattern. The first vector captures the largest amount of variance in the data, and subsequent vectors capture decreasing amounts of variance in orthogonal directions. They can be used for dimensionality reduction by projecting data onto a subspace spanned by top few vectors to obtain a lower-dimensional representation of the data that preserves most of its variance. A system, therefore, can be decomposed into its fundamental components in the form of eigenvectors

In the case of complex dynamical systems, these eigenvectors become eigenmodes where eigenmodes are used to describe the stable modes of oscillation of a system. When dealing with linear partial differential equations, such as in the case of reaction-diffusion systems, the eigenmodes take the form of cosine and sine functions, similar to those use in Fourier series and are referred to as Fourier modes (Fig. 3.2 *left*). Because they use sine-type functions, they can be described in terms of frequency of oscillation. The idea here is that all dynamical systems can be approximated as oscillating systems, typically with some decay, and any dynamical oscillating system can be described by a collection of normal modes. The intuition here is that these normal modes of oscillation are similar to harmonics in a string instrument.

Furthermore, operator splitting is a numerical technique employed to solve PDEs, including the reaction-diffusion equation. It is used for equations that are difficult or computationally expensive to solve directly, including complex physical systems with multiple interacting pro-

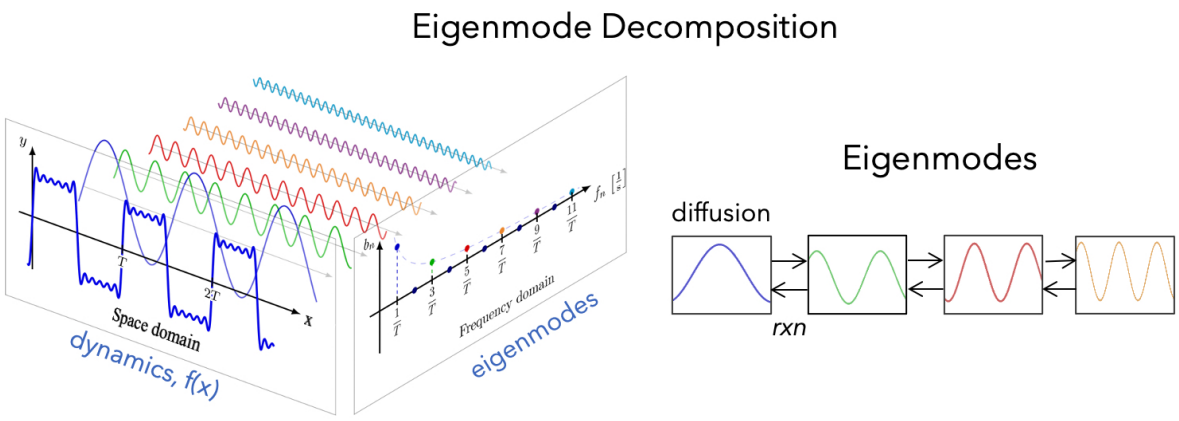

Figure 3.2. (*Left*) Eigenmode decomposition of a dynamical system. (*Right*) Markov model of synaptic transmission reaction-diffusion system where eigenmodes model diffusion and reactions mediate transitions between eigenmodes.

cesses. This method breaks down the original PDE into simpler subproblems which can then be solved individually. For example, in reaction-diffusion, it becomes possible to separate the PDE into a linear superposition space and time.

In the synaptic transmission system, diffusion is a linear process while the reactions are non-linear, causing the majority of the complexity. Diffusion can then be handled using the aforementioned eigenmode decomposition techniques while reactions are handled otherwise. The eigenmodes then capture the spatial dynamics of the system in the spectral domain due to the use of Fourier modes. Afterwards, it is possible to build a graphical model of the system with eigenmode states and transitions between eigenmodes mediated by reactions (Fig. 3.2 *right*). As long as the number of eigenmodes remains less than the number of spatial locations in a grid-based method, this provides a computationally-efficient alternative method for modeling reaction-diffusion.

In the following formulation, first we build a deterministic reaction-diffusion model describing Ca^{2+} influx into and diffusion through the presynaptic terminal in addition to its reversible binding with CALB. This is done using an eigenmode formulation. Afterwards, we build a stochastic model of Ca^{2+} diffusion using Markov models in the spectral domain. Both models are validated against grid-based methods where the number of eigenmodes used is

equivalent to the number of spatial locations in the grid-based formulation.

3.2 Deterministic Spectral Reaction-Diffusion

Ca^{2+} dynamics directly determines neurotransmitter activity making it the focus of these simulations. Here, we look to simulate Ca^{2+} influx to the presynaptic terminal with reversible second-order binding to the CALB volume buffer. Ca^{2+} is initially introduced to the system as an impulse injection at time zero, which is described mathematically as a dirac delta. The length of the terminal is significantly longer than the cross-section, so a 1D simplifying assumption can be used. Furthermore, the presynaptic terminal is adjacent to neighboring similar presynaptic terminals. As such, the boundary conditions can best be described as reflective or zero-flux Neumann conditions. Therefore, the problem formulation seeks to find the amount of Ca^{2+} in the system, described as

$$u = u(x,t) \tag{3.1}$$

Initial Conditions

where the initial conditions is given by a general function over space:

$$u_\circ = u(x_\circ, t = 0) = \delta(x_\circ) \tag{3.2}$$

Boundary Conditions

$$\left.\frac{du}{dx}\right|_{x=0} = \left.\frac{du}{dx}\right|_{x=L} = 0 \tag{3.3}$$

3.2.1 Spectral 1D Diffusion with No Reactions

The following solution technique uses Fourier series to decompose the diffusion equation. Focusing initially on the 1D homogeneous general diffusion PDE formulation using Fick's law of diffusion with no sink or source term

$$\frac{\partial u}{\partial t} = D\frac{d^2u}{dx^2} \tag{3.4}$$

For a linear time-variant system, we can utilize operator splitting or separation of variables to formulate a solution to the homogeneous PDE factorized as a superposition of the product of independent variables of the form. Using an eigenmode decomposition paired with the principle of superposition further states that a linear system can be described by an infinite sum of eigenmodes

$$u(x,t) = \sum_n X_n(x)T_n(t) \tag{3.5}$$

Substituting (3.5) into (3.4) and solving gives the independent equations in both space and time as

$$X(x) = \mathrm{A}\cos\left(\frac{n\pi}{L}x\right) + \mathrm{B}\sin\left(\frac{n\pi}{L}x\right) \tag{3.6}$$

$$T(t) = \mathrm{C}e^{-D\lambda t} \tag{3.7}$$

where A, B, and C are constants determined by the initial conditions and $\lambda = \sqrt{\frac{n\pi}{L}}$

For Neumann boundary conditions, the *sin* terms become zero. Combining the independent equations for $X(x)$ (3.6) and $T(t)$ (3.7) into (3.5) gives the overall equation:

$$u(x,t) = \sum_{n=0}^{\infty} \mathrm{A}_n \cos\left(\frac{n\pi x}{L}\right) e^{-D\left(\frac{n\pi}{L}\right)^2 t} \tag{3.8}$$

where n indexes the eigenmodes and where A_n is a combined constant given by the initial conditions. For a dirac delta input function at some position, x_o, this is

$$\mathrm{A}_n = \mathrm{Z}\int_0^L \delta(x_\circ)\cos\left(\frac{n\pi x}{L}\right)dx = \mathrm{Z}\cos\left(\frac{n\pi x_\circ}{L}\right) \tag{3.9}$$

where the normalization constant Z is either $\sqrt{\frac{2}{L}}$ or $\sqrt{\frac{1}{L}}$ for the nonzero and zero eigenmodes respectively.

Therefore, the overall equation for diffusion with Neumann boundary conditions for concentration is

$$\boxed{u(x,t) = \frac{1}{L} + \frac{2}{L}\sum_{n=1}^{\infty} \cos\left(\frac{n\pi}{L}x_{\circ}\right)\cos\left(\frac{n\pi}{L}x\right)e^{-D\left(\frac{n\pi}{L}\right)^2(t-t_o)}} \tag{3.10}$$

3.2.2 Spectral 1D Diffusion with Reactions

In general, reaction-diffusion systems in 1D are described by

$$\frac{du}{dt} = D\frac{d^2u}{dx^2} + q(u) \tag{3.11}$$

Again, we begin with an equation for concentration separated into a sum of spatial and temporal components, using the premise of separation of variables.

$$u^{\mathrm{A}}(x,t) = \sum_n X_n(x)T_n^{\mathrm{A}}(t) \tag{3.12}$$

where $u^{\mathrm{A}}(x,t)$ is the concentration of species A, which can be described by the product of spatially-dependent and species-independent eigenmodes $X(x)$ for eigenmodes 0 to n and the temporal component, T_n^{A}, which is dependent on species A. Spatial eigenmodes $X(x)$ only depend on the boundary conditions, so they are the same between species if and only if the boundary conditions are the same between the chemical components, which is the case here. As such, both $X(x)$ and the resulting λ are not dependent on the chemical species and are of the same *cos* form as seen in (3.10) The temporal component is dependent on the reaction system in question.

Second Order Reversible Reaction

Upon entering the presynaptic terminal, Ca^{2+} binds reversibly to the volumetric buffer CALB. The binding can be described simply as a second-order reversible reaction with Ca^{2+} and CALB as reactants and the bound CaCALB complex as the product according to

$$\mathrm{A} + \mathrm{B} \underset{\mathrm{k}_r}{\overset{\mathrm{k}_f}{\rightleftharpoons}} \mathrm{C} \tag{3.13}$$

where A is Ca^{2+}, B is CALB, and C is CaCALB and the reaction rates are obtained experimentally.

This system is described by a flow or reaction term that is positive entering and negative leaving Ca^{2+} and CaCALB respectively. The differential equation describing this reaction using reaction rates is

$$\frac{d[\mathrm{C}]}{dt} = -\frac{d[\mathrm{A}]}{dt} = -\frac{d[\mathrm{B}]}{dt} = k_f[\mathrm{A}][\mathrm{B}] - k_r[\mathrm{C}] \tag{3.14}$$

In addition to the stoichiometric conversions between the species due to the reaction, the three species each undergo diffusion at species-dependent diffusivities resulting in the reaction-diffusion equation describing Ca^{2+} as:

$$\frac{\partial[\mathrm{A}]}{\partial t} = -k_f[\mathrm{A}][\mathrm{B}] + k_r[\mathrm{C}] - D^{\mathrm{A}}|\overrightarrow{\nabla}|^2_{\overrightarrow{r}}[\mathrm{A}] \tag{3.15}$$

Similar such equations can be written for CALB and CaCALB as well.

The concentrations can be written using separation of variables, as seen in (3.12). For Ca^{2+}, this is:

$$\sum_i X_i(x)\frac{d}{dt}T_i^{\mathrm{A}}(t) = -k_f\sum_i X_i(x)T_i^{\mathrm{A}}(t)\sum_j X_j(x)T_j^{\mathrm{B}}(t) + k_r\sum_l X_l(x)T_l^{\mathrm{C}}(t) - D^{\mathrm{A}}\sum_i \lambda_i X_i(x)T_i^{\mathrm{A}}(t) \tag{3.16}$$

To obtain an equation in terms of the spatial dependent component, it is possible to integrate the temporal components out of the equation. As a result of the temporal components forming an orthonormal basis, this becomes possible by multiplying the equation by any of the eigenmodes and integrating the result over the space.

Further simplification occurs as a results of the orthonormality of the eigenmodes given by

$$\int_L X_i(x)X_n(x)dx = \delta_{in} \tag{3.17}$$

Therefore, the temporal component of a given eigenmode is described simply by the summation of the temporal components across all eigenmodes for the forward reaction. The reverse and diffusion components are in terms of the temporal component of the given eigenmode.

$$\frac{d}{dt}T_n^{\mathrm{A}}(t) = -k_f\sum_i\sum_j \alpha_{ijn}T_i^{\mathrm{A}}(t)T_j^{\mathrm{B}}(t) + k_rT_n^{\mathrm{C}}(t) - D^{\mathrm{A}}\lambda_n T_n^{\mathrm{A}}(t) \tag{3.18}$$

$$\frac{d}{dt}T_n^{\mathrm{B}}(t) = -k_f\sum_i\sum_j \alpha_{ijn}T_i^{\mathrm{A}}(t)T_j^{\mathrm{B}}(t) + k_rT_n^{\mathrm{C}}(t) - D^{\mathrm{B}}\lambda_n T_n^{\mathrm{B}}(t) \tag{3.19}$$

$$\frac{d}{dt}T_n^{\mathrm{C}}(t) = k_f\sum_i\sum_j \alpha_{ijn}T_i^{\mathrm{A}}(t)T_j^{\mathrm{B}}(t) - k_rT_n^{\mathrm{C}}(t) - D^{\mathrm{C}}\lambda_n T_n^{\mathrm{C}}(t) \tag{3.20}$$

where

$$\alpha_{ijn} = \int_L X_i(x)X_j(x)X_n(x)dx \tag{3.21}$$

In the double summation, we are left with α_{ijn} nonlinear interaction terms as coupling coefficients from the product of the three orthonormal spatial components giving cross terms.

In particular, for 1D diffusion with Neumann boundary conditions, $X_n(x)$ is given by (3.6). Therefore, α_{ijn} is of the form

$$\alpha_{ijn} = Z_i Z_j Z_n \int_0^L \cos(ax)\cos(bx)\cos(cx)dx \tag{3.22}$$

From the cosine product rule, this simplifies to

$$\int_0^L \cos(ax)\cos(bx)\cos(cx)dx = \begin{cases} 1 & a \pm b \pm c = 0 \\ 0 & otherwise \end{cases} \tag{3.23}$$

Therefore

$$\alpha_{ijl} = \frac{Z_i Z_j Z_l}{4} L \quad \text{for } n = |i \pm j| \tag{3.24}$$

The overall equations for Ca^{2+} is then

$$\boxed{u^{\mathrm{A}}(x,t) = \frac{T_\circ^{\mathrm{A}}(t)}{\sqrt{L}} + \sqrt{\frac{2}{L}} \sum_{n=1}^{\infty} \cos\left(\frac{n\pi}{L}x\right) T_n^{\mathrm{A}}(t)} \tag{3.25}$$

with equations of the same form for CALB and CaCALB.

Simulation Setup

The prior derivation can then be applied to model Ca^{2+} and CALB in the synapse. Initially, Ca^{2+} is injected into the system at time $t = 0$. CALB is initially in its unbound state, uniformly distributed throughout the system, and there is no initial bound CaCALB. Therefore, the initial conditions of each species can be described as

Initial Conditions

$$
\begin{aligned}
u^{\mathrm{A}}(x_{\circ},0) &= [Ca^{2+}]_{\circ} \\
u^{\mathrm{B}}(x,0) &= [CALB] \\
u^{\mathrm{C}}(x,0) &= 0
\end{aligned}
\tag{3.26}
$$

The initial conditions for the temporal component of the overall equation are equivalent to the spatial components integrated over the volume in which the particle initial conditions exist. For Ca^{2+}, the simplifies to the value sampled at the location of the impulse injection. Since CALB is uniformly distributed, the zeroth eigenmode is the only nonzero component, which is given by the initial number across the spatial dimension. Bound CaCALB is initially at zero.

$$
T_n^{\mathrm{A}}(x_{\circ},t=0) = [\mathrm{Ca}^{2+}][\mathrm{Ca}^{2+}]_{[\mathrm{Ca}^{2+}]_{\circ} Z_n \cos\left(\frac{n\pi}{L} x_{\circ}\right)} \tag{3.27}
$$

$$
T_n^{\mathrm{B}}(x,t=0) = \begin{cases} [\mathrm{CALB}]Z_0 = \frac{[\mathrm{CALB}]}{L} & n=0 \\ 0 & n \neq 0 \end{cases} \tag{3.28}
$$

$$
T^{\mathrm{C}}(x,t=0) = 0 \tag{3.29}
$$

In order to model this simulation, first the system of ODEs describing the temporal component (Eqs. 3.18 - 3.20) is solved numerically using Runge-Kutta of order 5(4) with initial conditions given by (Eqs. 3.27 - 3.29). Afterwards, the temporal component time courses can be used to obtain the molecular concentration dynamics of each species using (3.12). See 3.2.2 for an algorithmic approach to the previous description.

Validation Against Finite Difference with Reactions

In order to validate the functionality of this method, it was compared to a finite difference (FDM) reaction-diffusion model. Both simulations used the biophysical constants as defined in

[30]. For the simplified CALB-binding reaction, the reaction rates for the first medium-affinity binding site in CALB were used as it is the largest forward reaction rate. Each method was run with 100 time points and 150 spatial points, which meets von Neumann stability criteria for FDM.

Algorithm 2: Deterministic Spectral Diffusion Algorithm

Result: The number of molecules in each state at each timepoint for a simulation

initialize temporal initial conditions T(t = 0) (Eqs. 3.27 - 3.29);

initialize temporal system of equations (Eqs. 3.18 - 3.20);

solve system of equations using RK45 **for** *each timepoint* **do**

  for *each spatial node* **do**

    for *each species* **do**

      sample (3.12)

    end

  end

end

3.3 Stochastic Spectral Markov Diffusion using Random Walk Eigenmode Decomposition

Markov models can be incorporated to add stochasticity to the diffusion solution. This can be done with an eigenmode decomposition of the diffusion solution, specifically of random walk diffusion, which is the typical method used to model diffusion with stochasticity. Random walk is a grid-based approach, meaning the spatial domain is converted to a discrete spatial mesh. After defining a spatial grid for diffusion, a general random walk diffusion model for a given particle can be described according to

$$\mathrm{n}_i \underset{\mathrm{k}_{ij}}{\overset{\mathrm{k}_{ji}}{\rightleftharpoons}} \mathrm{n}_j \tag{3.30}$$

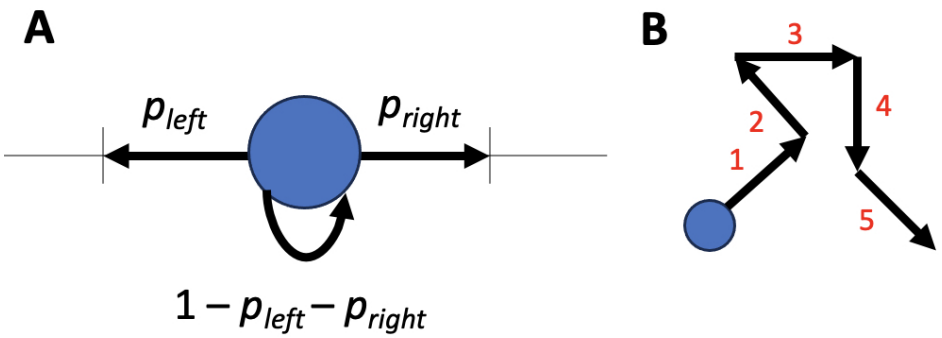

Figure 3.3. Random walk grid-based formulation

where n_i and n_j are the number of molecules in spatial location i and j, respectively with their associated rate constants k_{ij} and k_{ji}. For a small enough time step, the probability can be analyzed in terms of the change in number of particles in the i state in the context of the transfer of a single particle (Fig. 3.3). There are three possible cases: a particle being added to state i, no change in the number or one being removed from state i.

$$\lim_{\Delta t \to 0} P(n_i(t+\Delta t)|n_j(t)) = \begin{cases} \sum_j k_{ij} n_j(t)\Delta t, & \text{if } n_i(t+\Delta t) = n_i(t)+1 \\ 1 - \sum_j k_{ij} n_j(t)\Delta t & \text{if } n_i(t+\Delta t) = n_i(t) \\ \sum_j k_{ij} n_i(t)\Delta t, & \text{if } n_i(t+\Delta t) = n_i(t)-1 \end{cases} \tag{3.31}$$

In the limit of the time step going to zero, the derivative of the mean is then given by

$$\frac{d\mu_{n_i}(t)}{dt} = \lim_{\Delta t \to 0} \frac{\mu_{n_i}(t+\Delta t) - \mu_{n_i}(t)}{\Delta t} = \sum_j k_{ij} n_j(t) - \sum_j k_{ji} n_i(t) \tag{3.32}$$

which is just the difference of influx and efflux at the given node.

3.3.1 Eigenmode Decomposition

The derivative of the mean can equivalently be written in matrix form as

$$\frac{d}{dt}\overrightarrow{\mu}_{\overrightarrow{n}(t)} = \overline{\overline{\mathrm{A}}}\,\overrightarrow{\mu}_{\overrightarrow{n}(t)} \tag{3.33}$$

The matrix $\overline{\overline{\mathrm{A}}}$ represents the diffusive transitions (transition matrix) according to

$$\mathrm{A}_{ij} = k_{ij} - \sum_{h} k_{hi}\delta_{ij} \tag{3.34}$$

where the diagonals represent rates associated with staying in the current spatial location and off diagonals represent the transition to other spatial locations.

The transition matrix can be decomposed into its eigenmodes according to

$$\begin{aligned} \overline{\overline{\mathrm{A}}}\,\overline{\overline{\mathrm{U}}} &= \overline{\overline{\mathrm{U}}}\,\overline{\overline{\Lambda}} \\ \overline{\overline{\mathrm{A}}} &= \overline{\overline{\mathrm{U}}}\,\overline{\overline{\Lambda}}\,\overline{\overline{\mathrm{U}}}^{-1} \end{aligned} \tag{3.35}$$

where $\overline{\overline{\mathrm{U}}}$ represents the eigenmode matrix and $\overline{\overline{\Lambda}}$ are the associated eigenvalues.

Eigenmode decomposition allows navigation between the spatial and spectral domains. To eliminate space from the solution, it is critical to explicitly define the number of particles in the spatial domain, the number of particles in the spectral domain, and how to go between the domains. The number of particles in the spatial domain ($\overrightarrow{n}$) is given by the product of the eigenmodes and the number of particles in the spectral domain ($\overrightarrow{m}$) according to

$$\overrightarrow{\mu}_{\overrightarrow{n}} = \overline{\overline{\mathrm{U}}}\,\overrightarrow{\mu}_{\overrightarrow{m}} \tag{3.36}$$

3.3.2 Spectral Markov Diffusion Statistics

To capture the varying eigenmode dynamics, the modes can be described as a 2-state system where the states represent possible fates for particles – either removal or addition to a

given mode. The total particle count in a particular mode can then be given as the difference between these to states. Without reactions, each eigenmode is an independent system.

$$+\mathrm{m}_k \underset{\frac{\lambda}{2}}{\overset{\frac{\lambda}{2}}{\rightleftharpoons}} -\mathrm{m}_k \tag{3.37}$$

where m is the number of particles in mode k, either associated with particle gain (+) or loss (-) and λ is the eigenvalue for the mode given by the eigenmode decomposition.

Relationship Between Spatial and Spectral Domains

The probability p_i of finding a particle at spatial node i is given by the sum of the product of the eigenvector v_{ik}, representing spatial node i and eigenmode k, and the probability q_k of finding a particle at eigenmode k. In general, the eigenvectors can be used to transform variables between the Cartesian spatial representation and the spectral domain.

$$\begin{aligned} p_i &= \frac{n_i}{N} = \sum_k \mathrm{v}_{ik} q_k = \frac{1}{N} \sum_k \mathrm{v}_{ik} m_k \\ q_k &= \frac{m_k}{N} = \sum_i \mathrm{v}_{ik} p_i = \frac{1}{N} \sum_i \mathrm{v}_{ik} n_i \end{aligned} \tag{3.38}$$

where N is the total number of particles, n_i is the number of particles at spatial node i, and m_k is the number of particles in eigenmode k.

In this formulation, a particle has exactly two possible outcomes; it can be in the gain or loss state. As such, this is a Bernoulli process. Therefore, the probabilities of the two outcomes sum to unity according to

$$q_{+k} + q_{-k} = 1 \tag{3.39}$$

where q_{+k} and q_{-k} are the probabilities of a particle ending up in either the gain or loss state, respectively.

3.3.3 Model Statistics

There are multiple particles, so the overall diffusion process is captured by the Binomial distribution. Therefore, the mean and variance are given by the following:

Mean

$$\begin{aligned} \mathrm{E}[m_k] &= +\eta_k \cdot P(m_k = +\eta_k) - \eta_k \cdot P(m_k = -\eta_k) \\ &= \eta_k \left(q_{+k} - q_{-k}\right) = q_k \end{aligned} \tag{3.40}$$

Variance

Full variance derivations can be found in Appendix 3.5

$$\mathrm{E}[m_k^2] = \eta_k^2 = \sum_i p_i (\mathrm{v}_{ik})^2 \tag{3.41}$$

Multiple Particle Extension

These are true for a system with 1 particle. To extend this formulation to N particles, the sum becomes equal to N rather than 1, and η_k needs to be adjusted to scale properly according to:

$$\begin{aligned} q_{+k} + q_{-k} &= N \\ \eta_k \left(q_{+k} - q_{-k}\right) &= N q_k \end{aligned} \tag{3.42}$$

3.3.4 Transforming Between Spatial and Spectral Domains

From Spatial Particle Count to Spectral Domain

In order to obtain the initial conditions for use in the spectral simulation, it is necessary to transform the distribution of particles in the spatial domain to the spectral domain. This can be

done making use of the model statistics and eigenmode decomposition by solving the system of equations in 3.38 - 3.40 and 3.59. In terms of particle count, we obtain the resulting relationship to transform particles

$$\boxed{m_{\pm k} = \frac{1}{2}\left(1 \pm \frac{\sum_i n_i \mathrm{v}_{ik}}{\sqrt{N \sum_i n_i (\mathrm{v}_{ik})^2}}\right)} \tag{3.43}$$

where again m_k is the number of particles in eigenmode k, n_i is the number of particles at spatial node i, and N is the total number of particles.

If the simulation is an impulse of N particles at node x, then the number of particles at each spatial node is zero, except for the location of the impulse injection, simply giving

$$m_{\pm k} = \frac{1}{2}\left(\sqrt{N^2 (\mathrm{v}_{xk})^2} \pm N \mathrm{v}_{xk}\right) \tag{3.44}$$

From Spectral Particle Count to Spatial

The number of particles at each spatial node can be found by taking the difference between the number of particles gained and lost in each mode and transforming it back to the spatial domain according to

$$\boxed{n_i = \sum_k \mathrm{v}_{ik} m_k = \sum_k \mathrm{v}_{ik} (m_{+k} - m_{-k})} \tag{3.45}$$

3.3.5 Markov Validation Against Random Walk Formulation

To validate the Markov spectral method, we compare it to a random walk simulation using a toy example for 1D diffusion. This is a relatively simple diffusion problem but acts as a proof of concept. The random walk simulation has a given number of spatial nodes from which a transition matrix can be built using the rate of diffusion as defined by the rate constant for a given

particle. Then the eigendecomposition of the dynamics matrix is taken to obtain a set of linearly independent eigenvectors and eigenvalues with which the initial conditions can be transformed to the spectral domain. Afterwards, a Markov simulation is run using a 2-state system for each mode and probabilities associated with the eigenvalues. The simulation is then transformed to the spatial domain to obtain particle counts at each spatial location.

The System

Here, we are modeling a 1D Ca^{2+} diffusion system divided into 10 spatial locations. There are no reactions, and the boundaries are reflective Neumann. The time steps (1 μs) are small enough such that Ca^{2+} can diffuse or stay stationary in each step. Ca^{2+} is injected into the system as an impulse at the center at the beginning of the simulation. The length used was the 4 μm terminal distance with a spatial step of 0.36 μm and was run for a total of 1 ms. Again, the number of eigenmodes used was equivalent to the number of spatial locations and both simulations were run 1000 times to get the mean and standard deviations.

States

Each eigenmode is described by an independent two-state system, except the zeroth mode which just has one state. This is because the zeroth eigenmode represents the total number of particles in the system, and there is mass conservation. The total number of eigenmodes is initially equal to the number of spatial locations, although not all may be necessary to fully capture the dynamics of the system. Therefore, the overall number of states is $2M-1=2S-1$ where M is the number of eigenmodes and S is the number of spatial nodes.

Transitions

Transitions are only possible within an eigenmode. Particles can transition between states or stay in their current state through a self-loop, resulting in 4 possible transitions per eigenmode, except for the zeroth eigenmode which cannot transition. In the rest, the total number

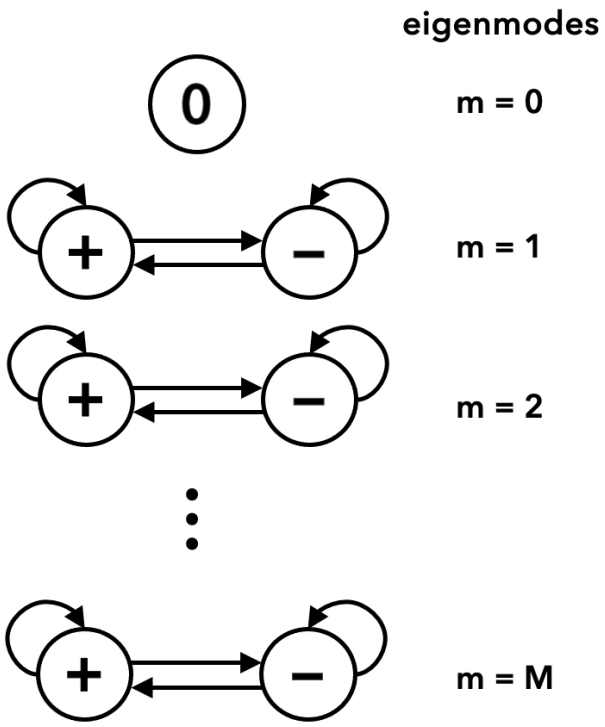

Figure 3.4. Graphs representing Markov states, where each eigenmode is described by a two state system capturing particle loss and particle gain. The zeroth eigenmode is one state as there is conservation of mass in the current formulation.

of transitions is $4(M-1) = 4(S-1)$. Transitions are given by the eigenmode decomposition of the transition matrix, and therefore depend on the spatial diffusion resolution.

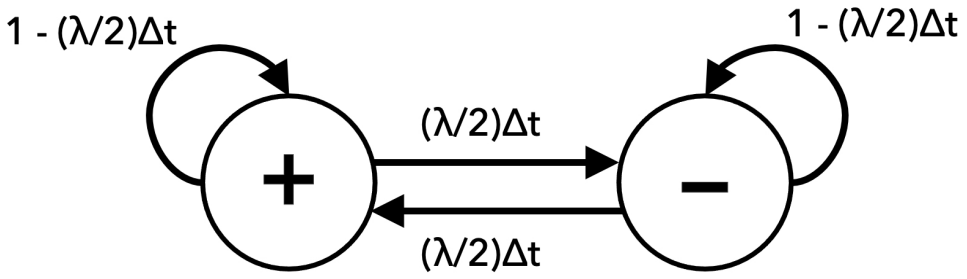

Figure 3.5. 2-state eigenmode Markov model showing transitions given as the half the eigenvalue multiplied by the time step where the eigenvalue is obtained from eigendecomposition of the dynamics matrix.

Transition Matrix

In 1D, there are three possible outcomes for where a particle could lie which dictate the particle's movement.

For particles in the middle of the line:

$$\frac{dn_i}{dt} = kn_{i-1} - 2kn_i + kn_{i+1} \tag{3.46}$$

At the left edge:

$$\frac{dn_i}{dt} = -kn_i + kn_{i+1} \tag{3.47}$$

At the right edge:

$$\frac{dn_i}{dt} = kn_{i-1} - kn_i \tag{3.48}$$

Therefore, for an example 5-node 1D diffusion system, the overall dynamics matrix is written as

$$A = \begin{bmatrix} -k & k & 0 & 0 & 0 \\ k & -2k & k & 0 & 0 \\ 0 & k & -2k & k & 0 \\ 0 & 0 & k & -2k & k \\ 0 & 0 & 0 & k & -k \end{bmatrix} \tag{3.49}$$

The eigenvalues and eigenvectors are then obtained according to (3.35). They are real and positive.

Initial Conditions

Initial conditions are given by the transformation of number of particles in the spatial domain to the spectral domain, according to (3.37).

Algorithm 3: Eigenmode Markov Diffusion Algorithm

```
Result: The number of molecules in each state at each timepoint for a simulation
initialize number of spatial nodes;
initialize initial conditions across spatial nodes;
initialize dynamics matrix (1.49);
solve eigendecomposition of dynamics matrix;
initialize transition probabilities (Fig. 1.3);
transform spatial initial conditions to spectral using (1.44);
for each timepoint do
    for each eigenmode state do
        for each particle do
            draw a random number, U;
            move particles based on transition probabilities;
        end
    end
end
transform spectral simulation to spatial domain using (1.45);
```

3.4 Deterministic Spectral Diffusion Results

Initially, Ca^{2+} is injected into the system at the location of the voltage-dependent Ca^{2+} channel (VDCC) with a magnitude equal to the number typically produced by an action potential, and CALB is uniformly distributed throughout the volume. The time step used is 1 microsecond and the diffusive length was the standard terminal distance of 4 μm, with a step of approximately 30 nm. Ca^{2+} dynamics are shown for various time steps in Fig. 3.4, looking at the distances near the VDCC, located at 2.35 μm. Additional figures for Ca^{2+} dynamics as well CALB and CaCALB can be found in Appendix 3.5.

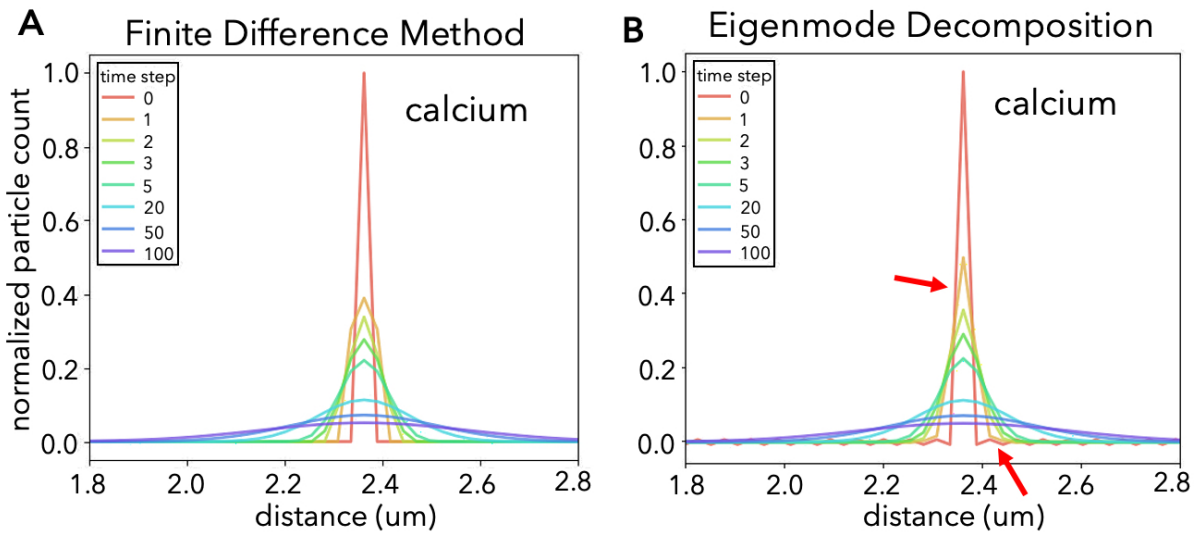

Figure 3.6. Calcium dynamics for deterministic calcium and calbindin simulation in 1D using (**A**) finite difference and (**B**) eigenmode decomposition methods. Red arrows highlight oscillations at time zero and differences in early time step dynamics.

Calcium influx is visible at the VDCC seen in Fig. 1.4 in red, at time zero. As time passes, Ca^{2+} diffuses through the length, creating smoother curves. Additionally, rapid binding to calbindin occurs. In this comparison, the number of eigenmodes is equivalent to the number of spatial locations, allowing assessment of the accuracy-efficiency trade off. At time zero, slight oscillations are visible in the eigenmode decomposition formulation (Fig. 1.4B red), and slight differences in the dynamics are visible at early time steps (Fig. 1.4B orange, yellow). This is a product of Gibbs phenomenon, a known feature of these spectral methods, where oscillations appear in the presence of jump discontinuities such as the impulse function if not enough eigenmodes are used to fully capture the dynamics. As expected, at time zero the series oscillates around the expected value. With fewer modes, these oscillations are larger. Because this is a product of the impulse function, these oscillations quickly disappear after time zero and the simulation is the same by the 5th time step at 5 μs. If calcium is introduced as a smooth function through the VDCC, this oscillation is expected to disappear.

The majority of the difference lies in the Ca^{2+} dynamics. Looking specifically at the difference between the two methods, it is clear that it is localized due to the impulse injections

(Fig. 3.4). The difference is specific to the location of the VDCC (Fig. 3.4A) and early time steps (Fig. 3.4B). In fact, the difference only reaches 4 spatial locations away from the injection site. Over time, it is apparent that the difference levels out pretty quickly to what is visible at time point 10.

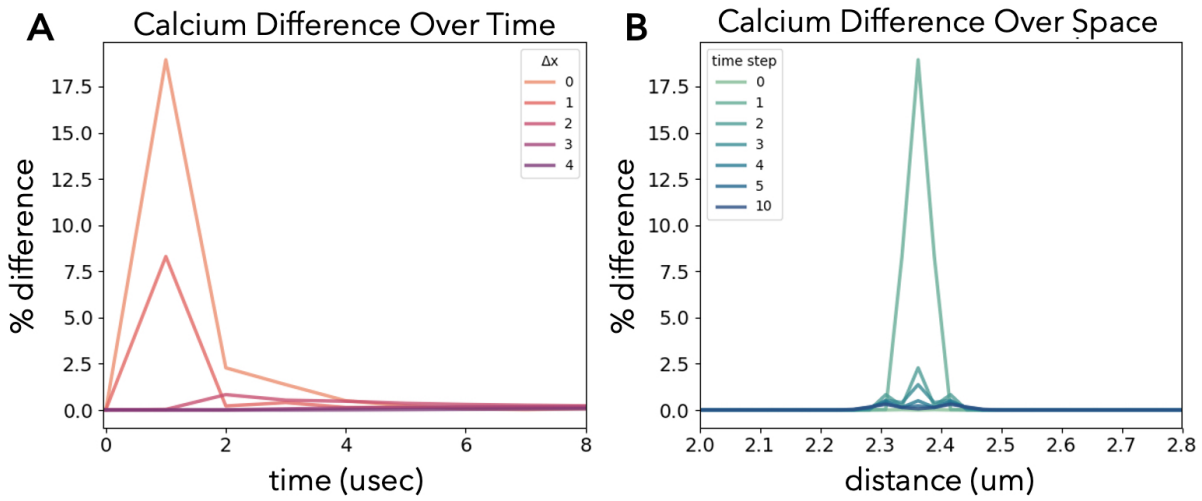

Figure 3.7. Percent difference between finite difference and eigenmode decomposition methods of reaction-diffusion for calcium dynamics plotted (**A**) over time for injection site and adjacent locations and (**B**) over space, looking at initial time steps.

These simulations set a strong foundation for spectral-based reaction-diffusion systems, but it is important to add stochasticity to these simulations. Furthermore, it is apparent that there is a trade off between accuracy and efficiency with more efficient models losing out on some biophysical accuracy, as is generally the case with scalable biophysical models.

3.4.1 Stochastic Markov Spectral Diffusion Results

There is clear agreement in the dynamics of the system between the Markov solution and the random walk solution, shown in Fig. 3.8, where the dark lines are the mean value and clouds are the standard deviation. Ca^{2+} is injected initally at the center point, node 5 (purple), and diffuses to the neighboring nodes. The means between the two methods are overlapped for approximately the entire simulation, but there is some disagreement in the variance (Figs. 3.8 and 3.9).

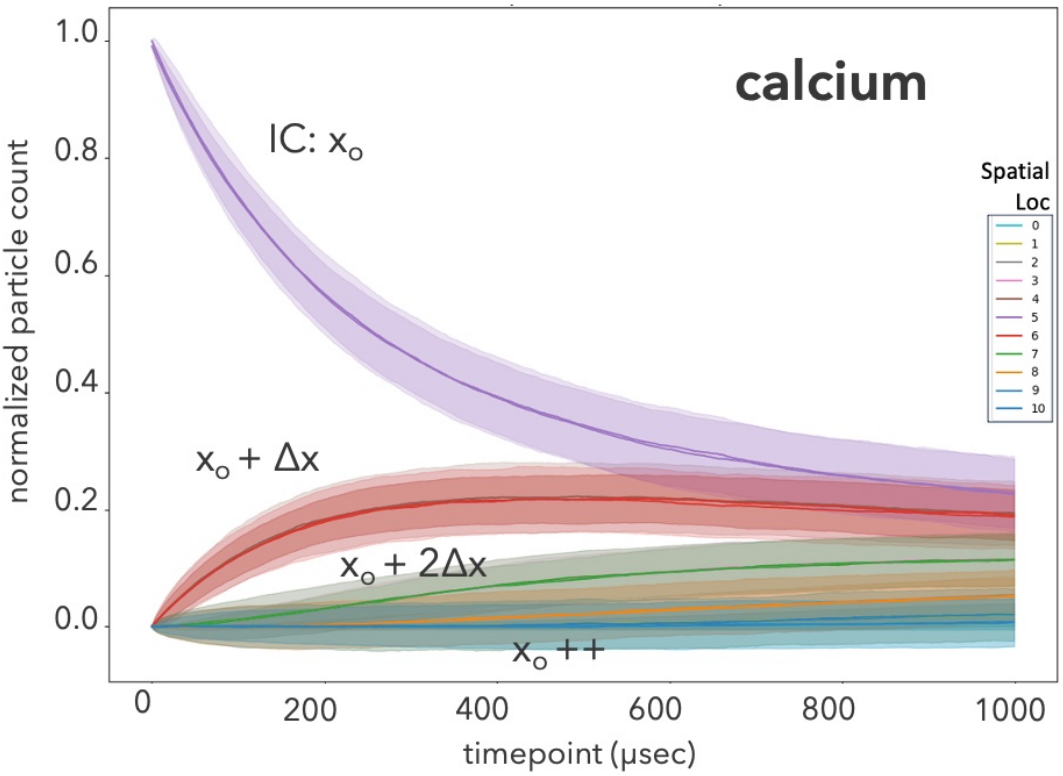

Figure 3.8. The normalized number of particles in each spatial location over time from the transformed spectral Markov simulation and random walk simulations.

Specifically, the variance of the spectral methods is larger than random walk when the number of particles in a given spatial node is close to zero. At small values of particles close to zero, there is a larger variance cloud in the eigenmode implementation. For a small number of modes, this method will again have small oscillations around zero. This can similarly be attributed to the Gibbs phenomenon causing numerical instability. Furthermore, there is currently no explicit handling of potential negative particles in this formulation.

3.5 Discussion

The work described here represents an alternative method for modeling reaction-diffusion systems by utilizing spectral methods as opposed to grid-based methods. The benefit of spectral methods is it allows for reducing the required number of dimensions to model the system by

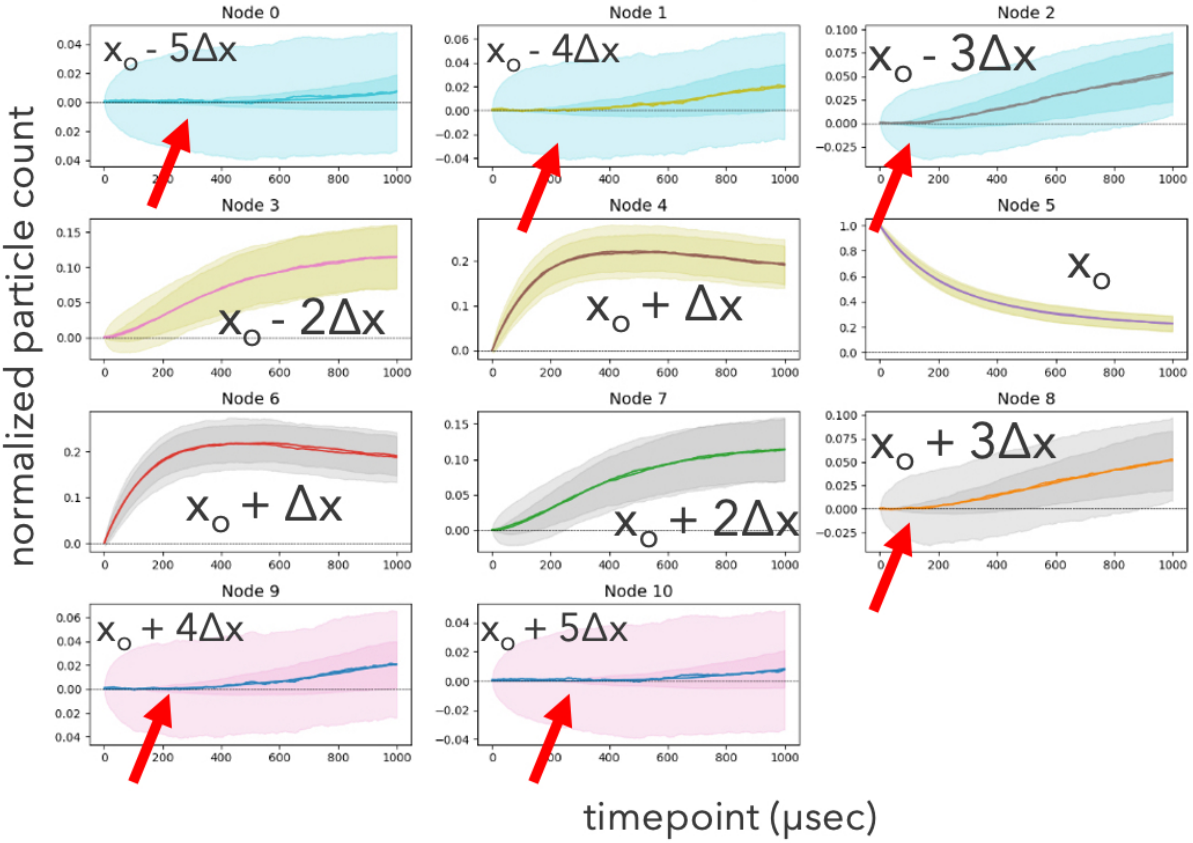

Figure 3.9. The normalized number of particles in each spatial location over time from the transformed spectral Markov simulation and random walk simulations. Red arrows indicate oscillations around zero associated with Gibbs phenomena.

only requiring a subset of eigenmodes to accurately capture the system's dynamics.

The primary limitation of this method, as is true with all spectral methods, is the presence of Gibbs phenomenon leading to oscillations around jump discontinuities. A formal and quantitative analysis of the impact of this limitation would be useful. Additionally, depending on the system, if the same number of eigenmodes as spatial locations are required to model the dynamics then the benefit of using spectral methods for computational efficiency is limited.

There is plenty of future work to be done. Particularly, the spectral reaction-diffusion method can be extended to three dimensions and applied to a more biophysically realistic model of synaptic transmission. Furthermore, it is worth testing the impact of adding stochasticity to this formulation using SDEs to describe the temporal components. Additionally, further

complexity can be added with VDCCs and the SNARE complex. The Markov model should additionally be extended to include reactions as well as diffusion. Furthermore, extending the results for Chapter 2, a binomial sampling scheme could be applied to the Markov simulations.

Overall, this study demonstrates the importance of spectral and graph methods in capturing probabilistic dynamics while allowing for scalability. There is promise in using such models going forward for bridging scales from molecular to network level models of neurodynamics.

Acknowledgment

Chapter 3 is coauthored with Wagner, Margot; Uppal, Abhinav; Rognes, Marie E.; Bartol, Thomas M.; Sejnowski, Terrence J.; and Cauwenberghs, Gert. The dissertation author was the primary author of this chapter.

Appendix

A.1 Derivations

A.1.1 Derivation of 2-State Spectral Diffusion System

Generally to model capture diffusion in the spectral domain, there is a Markov process associated with each eigenmode. An eigenmode Markov chain can be modeled in terms of the three possible methods of changes to the number of particles: +1, 0, -1. The eigenmodes have to reflect these three outcomes. Each eigenmode has three possible states with transitions given by the eigenmode. Thus, for a given eigenmode ℓ

$$\mathrm{m}_{+1\ell} \underset{\lambda_\ell}{\overset{\lambda_\ell}{\rightleftarrows}} \mathrm{m}_{0\ell} \underset{\lambda_\ell}{\overset{\lambda_\ell}{\rightleftarrows}} \mathrm{m}_{-1\ell} \tag{3.50}$$

Mean Rate

The mean rate is given by

$$\mu_{m\ell}(t) = \mu_{m_{+1\ell}}(t) - \mu_{m_{-1\ell}}(t) \tag{3.51}$$

The derivative with respect to time is then

$$\begin{aligned} \frac{d\mu_{m\ell}}{dt} &= \frac{d\mu_{m_{+1\ell}}}{dt} - \frac{d\mu_{m_{-1\ell}}}{dt} \\ = \lambda_\ell \left(\mu_{m_{0\ell}} - \mu_{m_{+1\ell}}\right) - \lambda_\ell \left(\mu_{m_{0\ell}} - \mu_{m_{-1\ell}}\right) \\ = -\lambda_\ell \left(\mu_{m_{+1\ell}} - \mu_{m_{-1\ell}}\right) \\ = -\lambda_\ell \mu_{m\ell} \end{aligned} \tag{3.52}$$

This derivation proves the model can be equivalently captured by a reduced, two-state form containing the $+1$ and -1 states according to

$$\mathrm{m}_{+1\ell} \underset{\frac{\lambda}{2}}{\overset{\frac{\lambda}{2}}{\rightleftarrows}} \mathrm{m}_{-1\ell} \tag{3.53}$$

The two state model's derivative is similarly given by

$$\begin{aligned} \frac{d\mu_{m\ell}}{dt} &= \frac{d\mu_{m_{+1\ell}}}{dt} - \frac{d\mu_{m_{-1\ell}}}{dt} \\ = \frac{\lambda}{2}\left(\mu_{m_{-1\ell}} - \mu_{m_{+1\ell}}\right) - \frac{\lambda}{2}\left(\mu_{m_{+1\ell}} - \mu_{m_{-1\ell}}\right) \\ = \frac{\lambda}{2}(\mu_{m_{-1\ell}} - \mu_{m_{+1\ell}} - \mu_{m_{+1\ell}} + \mu_{m_{-1\ell}}) \\ = -\lambda_\ell \left(\mu_{m_{+1\ell}} - \mu_{m_{-1\ell}}\right) = -\lambda_\ell \mu_{m\ell} \end{aligned} \tag{3.54}$$

This proves that the 3-state system can be equivalently described using a 2-state representation where the mean is the difference between the +1 and -1 values.

A.1.2 Eigenmode Variance

Within an eigenmode

$$\begin{aligned} \mathrm{E}[m_k^2] &= \eta_k^2 \cdot P(m_k = +\eta_k) + (-\eta_k)^2 \cdot P(m_k = -\eta_k) \\ &= \eta_k^2 \left(q_{+k} + q_{-k}\right) = \eta_k^2 \end{aligned} \tag{3.55}$$

Between eigenmodes

$$\begin{aligned} \mathrm{E}(m_k m_\ell) &= \sum_i \sum_j \mathrm{v}_{ik} \mathrm{v}_{j\ell} \mathrm{E}(n_i n_j) \\ &= \sum_i \sum_j \mathrm{v}_{ik} \mathrm{v}_{j\ell} p_i \delta_{ij} \\ &= \sum_i \mathrm{v}_{ik} \mathrm{v}_{i\ell} p_i \end{aligned} \tag{3.56}$$

At equilibrium, the particles are diffused homogeneously to equal concentrations, so $p_i = \frac{1}{N}$

$$\mathrm{E}(m_k m_\ell) = \frac{1}{N} \sum_i \mathrm{v}_{ik} \mathrm{v}_{i\ell} = \frac{1}{N} \delta_{k\ell} \tag{3.57}$$

Therefore, the modal variables will all have the same variance as the nodal variables, simply $\frac{1}{N}$.

Out of equilibrium, $p_i = \sum_h \mu_{ih} q_h$

$$\mathrm{E}(m_k m_\ell) = \sum_i \sum_h \mathrm{v}_{ik} \mathrm{v}_{i\ell} \mathrm{v}_{ih} q_h \tag{3.58}$$

Given $\sum_i p_i \mathrm{v}_{ik} \mathrm{v}_{i\ell} = \delta_{k\ell} \mathrm{v}_k$, we can obtain a closed-form solution for the variance

$$\eta_k^2 = \sum_i p_i (\mathrm{v}_{ik})^2 \tag{3.59}$$

In terms of particle counts, the variance can be written as

$$\eta_k^2 = \frac{1}{N} \sum_i n_i (\mathrm{v}_{ik})^2 \tag{3.60}$$

A.1.3 Particle Transformation to Spectral Domain

Using Probability

Using the following equations 3.39, 3.40, 3.38, 3.59

$$\begin{aligned} q_{+k} + q_{-k} &= 1 \\ \eta_k(q_{+k} - q_{-k}) &= q_k \\ q_k &= \sum_i p_i \mathrm{v}_{ik} \\ \eta_k^2 &= \sum_i p_i (\mathrm{v}_{ik})^2 \end{aligned} \tag{3.61}$$

We can substitute the definitions of q_k (Eq. 3.38) and η_k (Eq. 3.59) into Eq. 3.40.

$$\begin{aligned} \sqrt{\sum_i p_i (\mathrm{v}_{ik})^2}(q_{+k} - q_{-k}) &= \sum_i \mathrm{v}_{ik} p_i \\ q_{+k} - q_{-k} &= \frac{\sum_i p_i \mathrm{v}_{ik}}{\sqrt{\sum_i p_i (\mathrm{v}_{ik})^2}} \end{aligned} \tag{3.62}$$

We are left with an easily solvable system of equations for q_{+k} and q_{-k}

$$\begin{aligned} q_{+k} + q_{-k} &= 1 \\ q_{+k} - q_{-k} &= \frac{\sum_i p_i \mathrm{v}_{ik}}{\sqrt{\sum_i p_i (\mathrm{v}_{ik})^2}} \end{aligned} \tag{3.63}$$

The resulting equations are

$$q_{\pm k} = \frac{1}{2}\left(1 \pm \frac{\sum_i p_i \mathrm{v}_{ik}}{\sqrt{\sum_i p_i (\mathrm{v}_{ik})^2}}\right) \tag{3.64}$$

This equation describes the probability of particles in eigenmodes as a function of particles in spatial nodes and can be used to solve for the eigenmode simulation initial conditions. It also scales linearly with the number of total particles, N.

Using Particle Count

We would like to get an equation that relates the total number of particles between the spatial node and eigenmode simulations. This can be obtained from equation 3.64, scaling for the total number of particles, N:

$$m_{\pm k} = \frac{1}{2}\left(1 \pm \frac{\sum_i n_i \mathrm{v}_{ik}}{\sqrt{N \sum_i n_i (\mathrm{v}_{ik})^2}}\right) \tag{3.65}$$

This relationship is used to translate the initial conditions from the spatial simulation to the eigenmode space.

A.2 Deterministic Spectral Diffusion Results

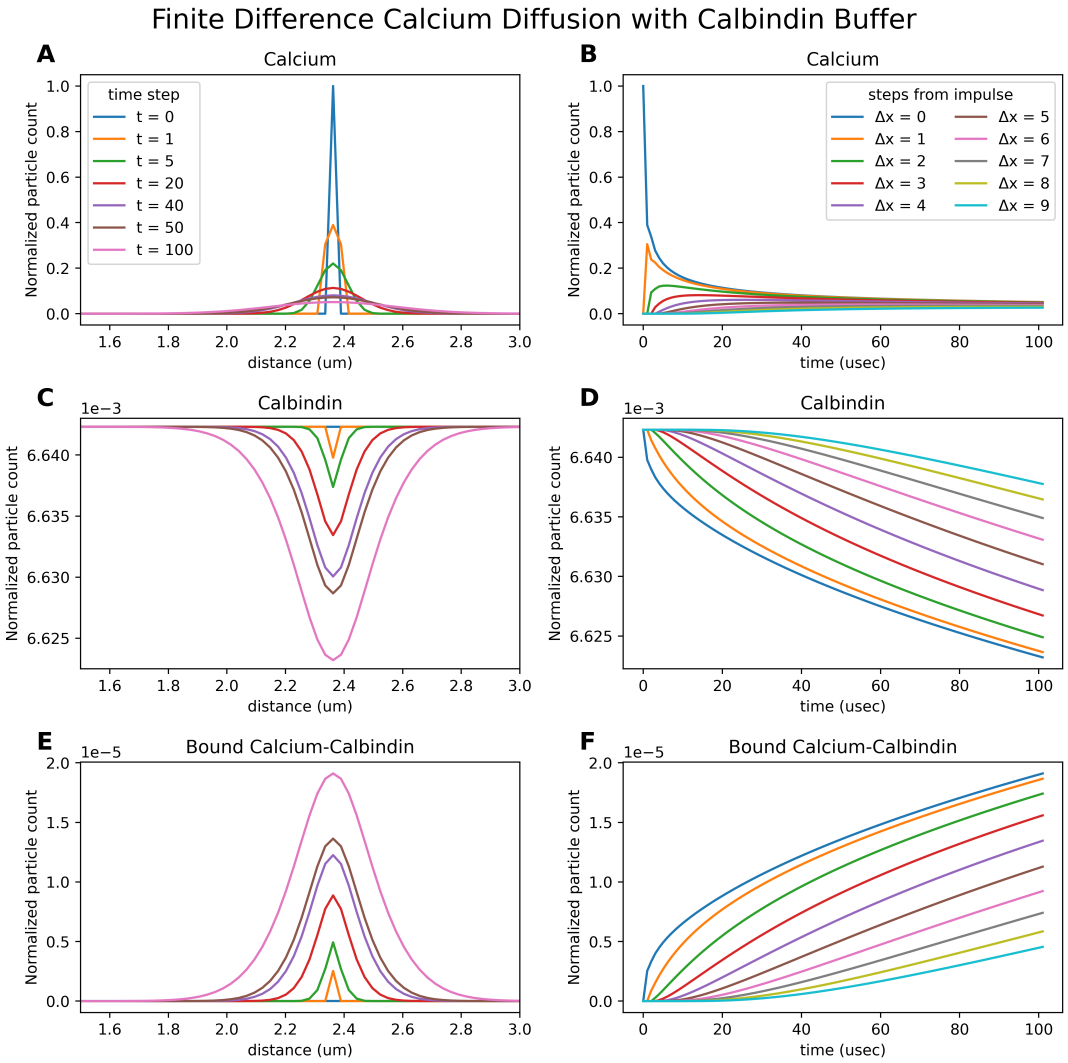

Figure 3.10. Normalized particle count across space (**A**-**C**) and time (**D**-**F**) for 1D Ca^{2+} (A/D), CALB (B/E), and CaCALB (C/F) reaction-diffusion simulation modeled using Finite Difference Method.

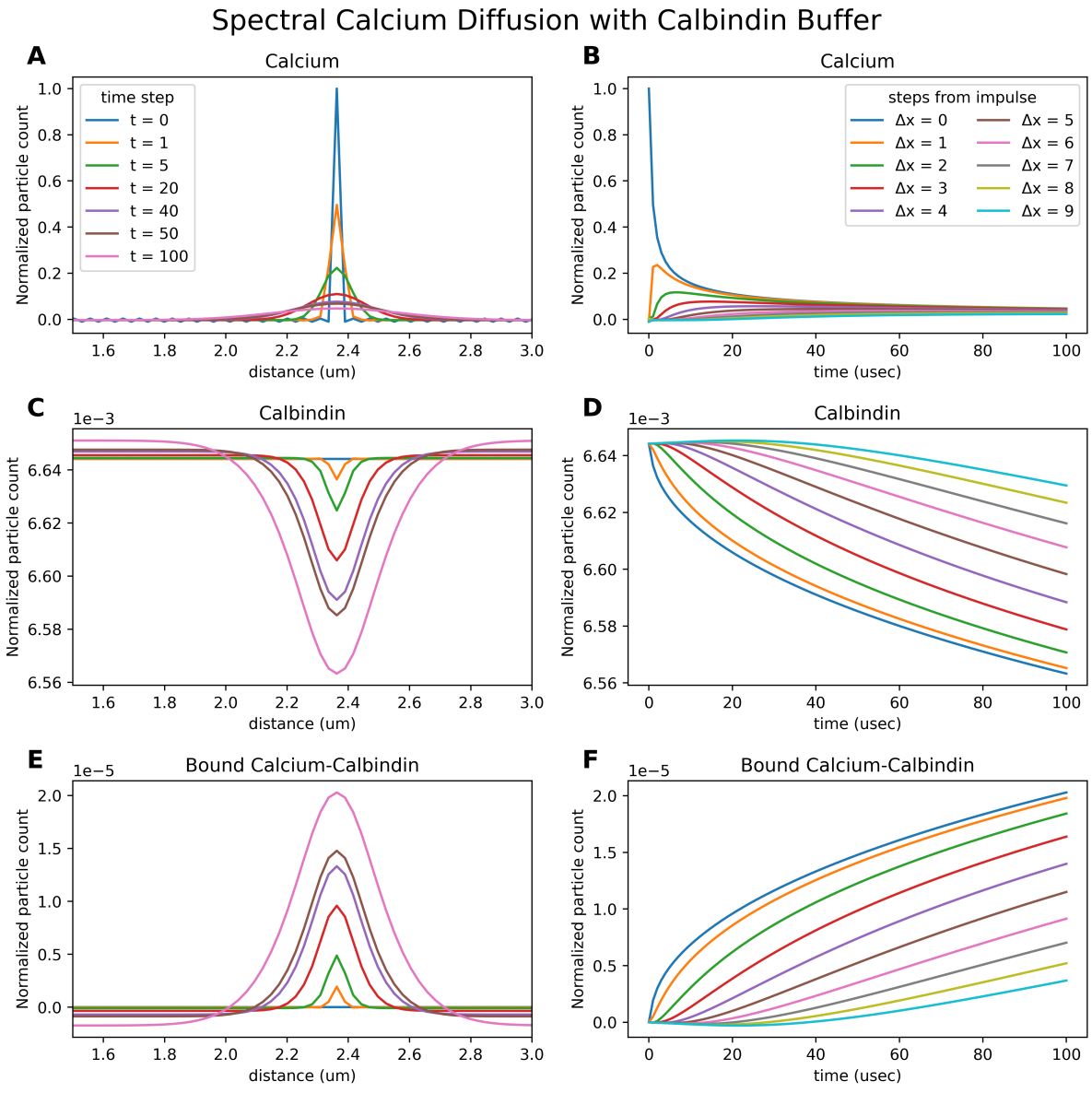

Figure 3.11. Normalized particle count across space (**A**-**C**) and time (**D**-**F**) for 1D Ca^{2+} (A/D), CALB (B/E), and CaCALB (C/F) reaction-diffusion simulation modeled using Spectral Reaction-Diffusion Method.

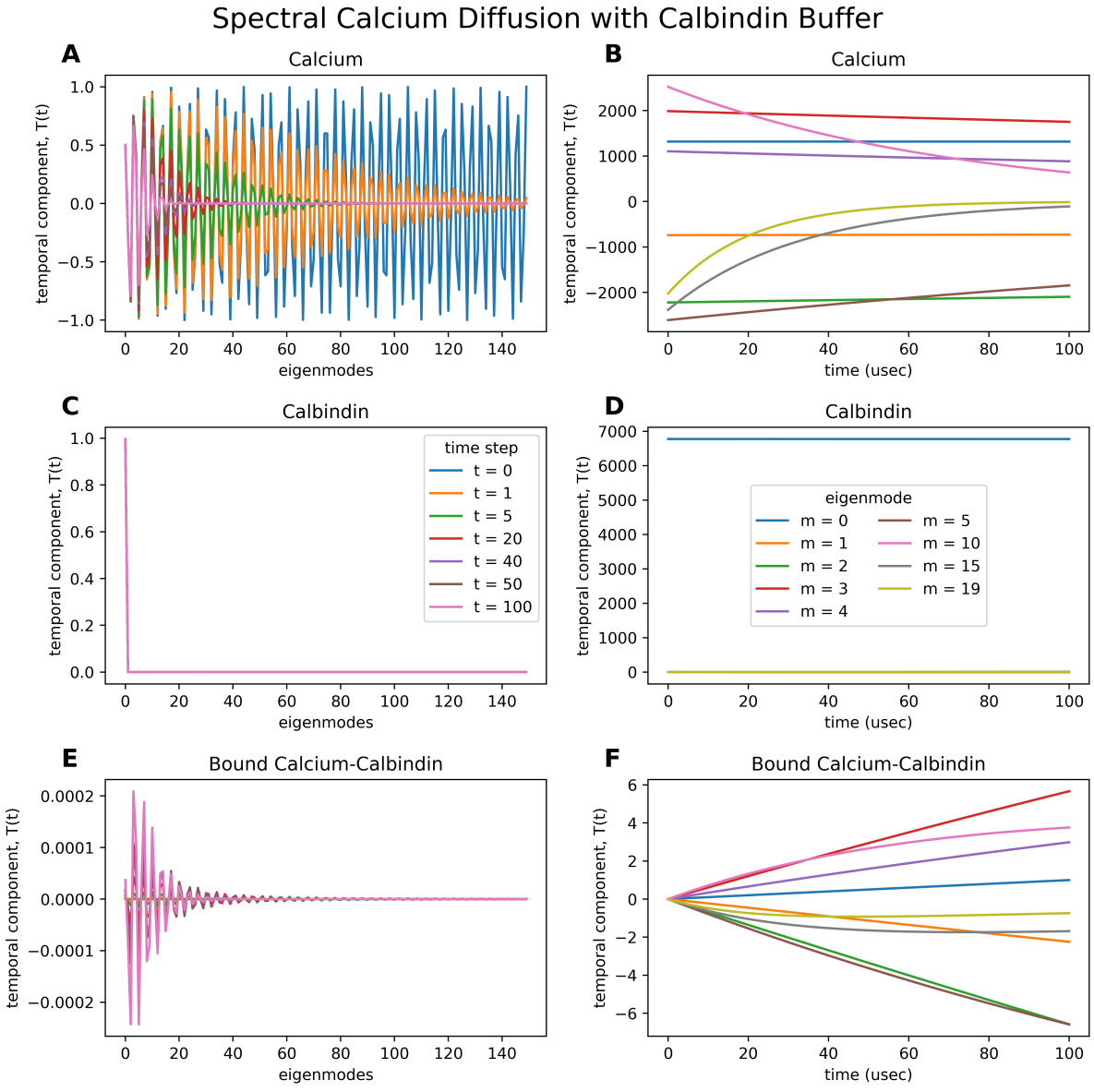

Figure 3.12. Temporal component of spectral methods solution across eigenmodes (**A**-**C**) and time (**D**-**F**) for 1D Ca^{2+} (A/D), CALB (B/E), and CaCALB (C/F) reaction-diffusion simulation modeled using Spectral Reaction-Diffusion Method.

Chapter 4

Structural and Functional Connectivity Biomarkers for Adolescent Major Depressive Disorder using Automated Neuroimaging Pipeline

4.1 Introduction

Depressive disorders are severe psychiatric disorders characterized by low moods, lethargy, and a high risk of suicide. They are pervasive and exhibit severe social and economic impacts globally. In 2019, an estimated 280 million people (3.8%) were living with depressive disorder globally, making up 29% of all mental disorders [10], the second leading disorder after anxiety disorders. The number further increased by an estimated 26% to 28% in response to the COVID-19 pandemic in 2020 [61]. Suicide, an unfortunately common outcome of depressive disorders, accounts for more than one in every 100 deaths globally [10], and for every one suicide, there are an estimated 20 attempts [5]. As of 2019, suicide was the third and fourth leading cause of death in 15-29 females and males respectively [10]. World Economic Forum researchers further estimated that mental health costs the world economy approximately $2.5 trillion in 2010 due to combined loss of economic productivity and direct cost of care, a number expected to rise to $6 trillion by 2030 [38].

Adolescents also suffer from depressive disorders, with 1 in 100 adolescents (1.1%)

ages 10-14 diagnosed. The number increases to 2.8% for ages 15-19 [10]. Moreover, 20% of adolescents age 12-17 have a major depressive episode each year (Fig. 4.1) [21]. A seminal United States study found that half of the mental disorders found in adulthood had appeared by the age of 14 and three quarters by the age of 24 years old [178]. This demonstrates the importance of early risk detection and intervention in adolescents with depressive disorders to protect against suicide or other negative outcomes.

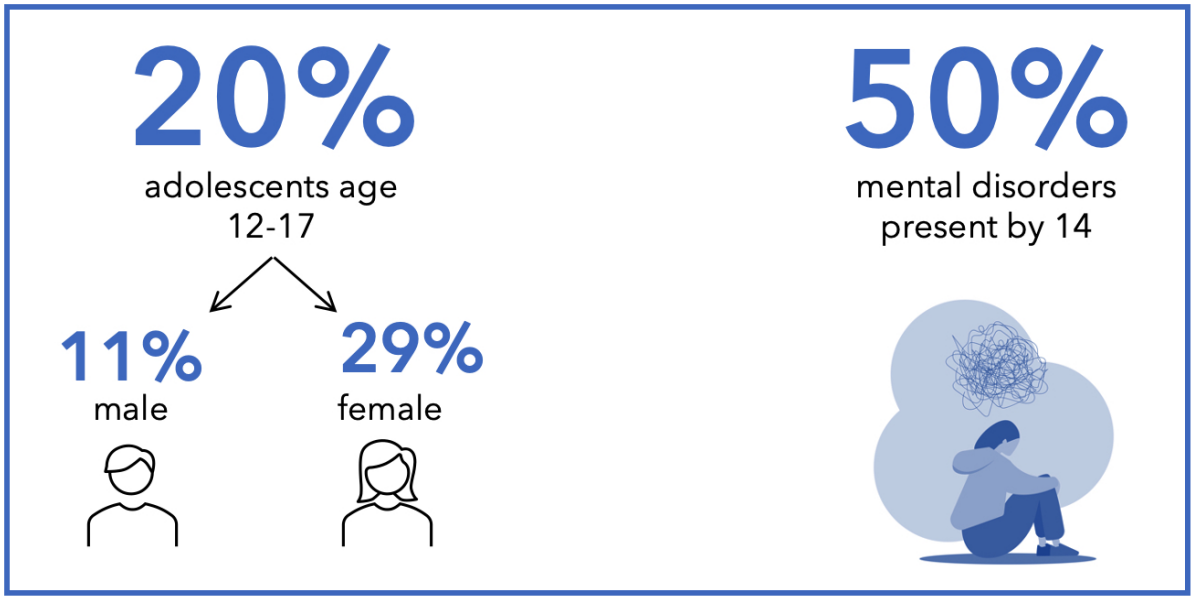

Figure 4.1. (*Left*) Prevalence of major depressive episode per year amongst US adolescents age 12-17 [21]. (*Right*) Results of study finding 50% of lifelong mental disorders are present by the age of 14 [178].

Despite the clear negative repercussions of depressive disorders, there does not yet exist a reliable and biologically-backed method of diagnoses, leading to further downstream issues of treatment validity. In other medical specialties, diagnostic tests are typically used, but of the over 3000 standardized laboratory diagnostic tests that exist, few, if any, are used routinely in psychiatric diagnostics [114]. The lack of biological underpinnings in mental disorders is due to a number of different factors including the use of underpowered studies without large enough sample sizes; approximate replications of studies to to a lack of standardized analysis methods; the tendency to average across patients to strengthen the statistical power of a study;

heterogeneity that is not well-handled, including in demographics and in the severity and quality of mental disorders; confounding factors that are not explicitly addressed, such as disease comorbidity and medication effects; and a missing diagnostic gold standard for mental disorders leading to unclear divides in populations between disordered and control groups. Historically, there has been a fundamental lack of large, standardized studies and datasets making downstream analyses difficult.

The standard used for psychiatric diagnostics in the United States is the Diagnostics and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), which is a symptomatically taxonomy of mental disorders, where symptoms are qualitative in nature and based on patient or care-taker expressed feelings and behaviors. For example, some of the symptoms listed for Major Depressive Disorder include "Depressed mood, Insomnia or hypersomnia, Fatigue or loss of energy." They are fundamentally non-quantitative and archaic methods of diagnosing. Its use has undergone criticisms due to its lack of biological basis for disorders. The most notable being Thomas Insel in 2013, the then director of the National Institute of Mental Health (NIMH), stating the DSM-5 was "at best, a dictionary" and that the NIMH will be moving its research away from DSM categories in order to begin to develop a better system [108]. Furthermore, as a result of limited diagnosis accuracies, a significant portion of patients fail to respond to treatment with antidepressants, and the response can only be determined after about 1 month [32].

With each new release of the DSM, the American Psychiatric Association (APA) tests its reliability across multiple sites in the US and Canada. Reliability is tested in the form of a inter-rater reliability, which measures the degree to which two clinicians could independently agree on the presence of a DSM-5 diagnoses. The DSM-5 was found to do a particularly poor job of detecting, classifying, and diagnosing major depressive disorder (MDD) in both adults and adolescents. Of the adult patients that were considered to have MDD via a previous diagnosis or qualifying symptom profile, only 49% were correctly identified as having MDD by one or both the clinicians using the DSM-5. Additionally, the degree to which the clinicians agreed about the diagnoses of the population given as kappa was 0.25. For adolescents, it was a similar

kappa of 0.28. The interpretation of these diagnoses given their interrater overlap was listed as "Questionable" by the APA [181].

Naturally, improved diagnostics and identifying biomarkers is an important area of research. In 2010, the NIMH began the Research Domain Criteria (RDoC) initiative to create a new data-driven taxonomy for mental disorders that seeks to connect behavioral and cognitive symptoms to a neurobiological and genetic basis [114]. It therefore attempts to identify quantitative biomarkers (i.e. genetics, imaging) or cognitive tests for diagnostics and treatment. Previous attempts to find useful biological psychiatry tests have failed due to limited accuracy or generalizability in clinical settings [134, 163]. This is primarily due to previously underpowered and heterogeneous studies. For example, a review that evaluated 30 clinical studies using working memory and fMRI in schizophrenia found a median sample size of 12 subjects [24]. Such studies are too small to have much statistic power and are unable to be combined due to their study differences. Numerous meta-analyses have concluded it is impossible to obtain an accurate quantitative conclusion due to the differences in protocols, tools, patients selected, and outcomes measured [24, 97]. As such, there is a push for large and standardized datasets for analysis, such as the Allen Brain Atlas [208] and the Alzheimer's Disease Neuroimaging Initiative (ADNI) [175].

4.1.1 Automatic Diagnosis

A promising area of research is in using a data-driven methods for diagnostic classification to elucidate complex statistical patterns in large and multimodal datasets. Computer vision applied to medical scans shows particular promise with nearly 400 Food and Drug Administration (FDA) approved AI and Machine Learning (ML)-enabled medical devices as of October 2022 [95]. While these data-driven methods are powerful on their own, they can also be combined with theory-driven mechanistic approaches in the form of prior knowledge. Applying theory to raw data can be used to obtain lower-dimensional feature representations of that data while also allowing room for testing scientific queries while maintaining higher interpret ability and

reliability. One example of this method was a study that used EEG data to classify patients as on or off deep brain stimulation (DBS). Raw data was compared with theory-driven measures obtained from the raw data as features. Specifically, a drift-diffusion model (DDM) was used to obtain a model of decision making. Indeed, classification was better when using fitted DDM model parameters rather than the raw data, suggesting the power of combining theory and data (i.e. machine learning) approaches [221]. This demonstrates potential for combined theory- and data-driven methodologies to elucidate new insights to mental illnesses.

4.1.2 Neuroimaging Data Types

Magnetic resonance imaging (MRI) is a method used to image the brain. Structural MRI (sMRI) is takes a static image of the brain at a higher resolution to provide anatomical data. It is a series of 2D images stacked to form a 3D image of the brain. Functional MRI (fMRI) indirectly measures brain activity through its hemodynamic response. That is, active neurons receive oxygen from blood at a higher rate than inactive neurons. The protein in blood associated with oxygen transport, hemoglobin, displays different magnetic properties, allowing for blood-oxygen-level-dependent (BOLD) imaging used in fMRI [164]. As such, fMRI takes multiple 3D images over time at a lower resolution to capture brain activity, resulting in 4D data, the fourth dimension being time. It can be measured either in resting state or during a task. In both data types, each (x, y, z) coordinate in 3D space is called a voxel and contains a numerical measurement.

MRI images contain a wealth of information, yet they are highly dimensional data types with upwards of millions of voxels per image. As such, dimensionality reduction is an important aspect. Segmentation of images into "Regions of Interest" (ROI) is a common method, where the regions will be given a single representative value. ROIs vary and depend on the scientific question at hand, but a common way to segment is by brain structure. Therefore, the image can be segmented into structures, each of which can be assigned a representative value (i.e. average intensity). Such segmentation can also be useful to assign labels to raw data voxels, so further

analysis is grounded in anatomically meaningful areas.

Neuroimaging data allows for estimation of brain connectivity. Structural connectivity defines the existence of white matter tracts physically interconnecting brain regions whereas functional connectivity (FC) describes the statistical dependencies between neural signals acquired from different brain areas using measures such as correlation and coherence. It is a derived measure from fMRI that measures the connectivity between brain regions that share function properties and are typically spatially distanced or anatomically distinct. The most popular method to obtain FC is pairwise covariance, which is a symmetric measure meaning it cannot detect directional coupling or disambiguate a common input. Furthermore, it assumes the data is stationary, which is often not the case.

Structure as a feature

The structure and function of numerous brain regions has been implicated in depression, but few findings have been reliably and consistently reported. Structural abnormalities, including volumes of brain regions, have been associated with multiple brain disorders, including depression, suggesting their promise as potential biomarkers for diagnosis. A 2012 review analysing neuroimaging and neuropsychiatric studies found that the more commonly observed differences include decreased volumes of the frontal lobe [170, 59, 123, 194], orbitofrontal cortex [40], hippocampus [194, 135], amygdala [198], and caudate [135] have been implicated in depression. Additionally, the insular volume has been found to correlate with depression scores [205]. Overall, structural differences in the form of region volumes have been implicated frequently enough to suggest them as potential biomarkers for depression worth interrogating further in large datasets.

Function and functional connectivity as features

Functional differences have been implicated depression in existing research. Decreased metabolism in the prefrontal cortex, specifically the dorsolateral and dorsoventral regions has

been frequently replicated [118, 170]. Abnormalities in the anterior cingulate cortex (ACC) and insula have also been demonstrated [75, 146, 25]. The limbic regions also exemplify dysregulated activation, including the amygdala and thalamus [145, 74, 211, 124, 83]. There is significant disparity in existing work on the exact functional differences associated with depression, which has led many researchers to believe the abnormalities are associated with connectivity differences, rather than function exclusively [170, 76, 81, 145].

Dysfunctional brain networks have been highly associated with MDD [100, 112] leading to multiple studies seeking to elucidate functional connectivity as a potential biomarker for an MDD diagnosis [186, 231]. Between 1990 and 2017, over 200 papers were published using FC, and its been used for a multitude of brain and mental disorders including major depressive disorder, schizophrenia, bipolar disorder, autism spectrum disorder, ADHD, Alzheimer's disease, and mild cognitive impairment [77]. Corticolimbic and intracortical connectivity abnormalities have frequently been found present in depression, although different studies report an increase or decrease [205, 146, 177, 89, 88]. These studies primarily focus on adults, so little is known about adolescents' brains neurological function between healthy and disordered populations.

To capture function, the brain can be depicted as a graph or network, where the nodes represent the brain regions and the edges are the FC between them. FC, unlike structural connectivity, does not represent physical connections but rather similar activity in different regions. To obtain FC between two regions, the similarity of the time series of the two regions is analyzed, which can be done through multiple different methods that assess the similarity matrix. A major limitation of these studies is their focus on multivariate statistical analyses of static derived data, neglecting available rich data on dynamics and topology, which have also been highly associated with the pathophysiology of MDD [101, 180, 197].

4.2 Methods

4.2.1 Adolescent Brain Cognitive Development Dataset

This study utilized the Adolescent Brain Cognitive Development (ABCD) baseline dataset, which obtained data from 21 different collection sites across the United States [217]. THE ABCD study is the largest longitudinal study of brain development and child health in the United States. Currently, $440 million has been put into this study from various arms of the NIH, including the National Institute on Alcohol Abuse and Alcoholism (NIAA), the National Institute of Drug Abuse (NIDA), National Institute of Mental Health (NIMH) and National Institute of Neurological Disorders and Stroke (NINDS) to name a few [17, 162]. The overall structure of the study collects data from multiple different domains annually from both the participating adolescents as well as their primary caretakers. There are 11,880 adolescents participating in the study. Data collection began in 2017 and 2018, when the subjects were 9-10 years old. Currently, there is funding through 2027, when the subjects will be 19-20, but the goal is to continue data collection until they are 25. There are a multitude of subgoals to be achieved by this study, but the primary goal is to discover which factors, biological and/or environmental, cause adolescents to diverge from normal brain development resulting in issues such as mental disorders and substance use. Early risk detection is a major focus of this work. Previous studies have been unable to definitively and replicably answer these questions due to the lack of size, scope, and methodological standardization seen in the ABCD study [217, 46].

The primary methodological reasons for discrepancies in previous work is due to data heterogeneity, confounding factors, underpowered studies, and averaging across patients. Specifically, heterogeneity is often seen in demographics, severity and quality of mental disorders, data acquisition techniques, and analysis methods. Confounding factors to account for are medication effects, effects of substance abuse, and comorbid disorders. In general, there has previously been a fundamental lack of large, standardized datasets.

At a baseline, the domains that the ABCD study addresses for its adolescent subjects

includes mental and physical health, brain imaging, biospecimens, neurocognition, substance use, and culture and environment. In year 2, mobile data in the form of FitBit was also added. Physical and mental health, substance use, and culture and environment data is also collected from the primary care takers. All data modalities are collected annually except for neuroimaging data which is collected semi-annually [17]. The true power of this study lies in the magnitude of data. As such, to truly take advantage of it, this study must be treated as a Big Data problem.

The analysis uses the labels provided by the Achenbach System of Empirically Based Assessment (ASEBA). ASEBA is a dimensional and empirically-based response to the DSM, although it also lacks a biological basis. It was formulated using data on numerous items that describe particular behavioral and emotional problems. The data was then analyzed via multivariate statistical analysis, including principal component analysis (PCA) to obtain syndromes. It is bottom-up method of classifying disorders rather than top down [22].

The following analysis has been done with the baseline dataset as a proof of concept, but it is important to note that the rates of depressive disorders are expected to be at their lowest in the baseline year. At the baseline, subjects in the ABCD study were ages 9 to 10. Patients with MDD were included if they met the ASEBA threshold for clinical diagnosis of MDD. Control subjects were selected if they did not meet the ASEBA threshold for clinical diagnosis of any mental disorder screened. Subjects with missing data were excluded. The resulting number of subjects were 353 and 1,429 for depressive and control populations respectively.

In the ABCD study, all measures are taken in one session. First, sMRI with T1-weighting is done at a 1 mm isotropic resolution for cortical and subcortical segmentation. The T1w acquisition is a 3D T1w inversion prepared RF-spoiled gradient echo using prospective motion correction, when available (Siemens and GE scanners, not Philips scanners). After, two 5 minute sessions of resting state fMRI data are collected at a resolution of 2.4mm (TR = 800 ms). Then, T2-weighted sMRI is done for white matter, lesions, and cerebral spinal fluid analysis followed by two more 5 minute sessions of resting state fMRI. Then diffusion-weighted MRI is done at 1.7mm isotropic resolution to allow for DTI analysis and fractional anisotropy. Lastly, the

task fMRI is collected. Prospective motion correction was not implemented for dMRI or fMRI acquisition. The fMRI acquisitions used multiband EPI with a slice acceleration factor of 6 as well as field map scans for B_0 distortion correction. Data is then collected by ABCD's Data Analysis, Informatics, and Resource Center (DAIRC) [91].

Preprocessing was primarily handled by DAIRC and involved several steps outlined in Hagler et al., 2019, *Image processing and analysis methods for the Adolescent Brain Cognitive Development Study*. Neuroimage, 202:116091 with updates to the preprocessing pipeline specified in the ABCD Data Release 4.0 Release Notes. Quality control (QC) was handled through a combination of manual and automated methods. Preprocessing followed fMRIPrep methodologies. The data used here was from ABCD Release 4.0 minimally preprocessed dataset. In the segmentation section, a comparison was made with the data segmented by the DAIRC, which used FreeSurfer v7.1.1. In addition to the minimal preprocessing done by the DAIRC, the fMRI data was band-pass filtered and Z-scored. For this study, only the T1w sMRI and resting state fMRI (rs-fMRI) data types are used.

4.2.2 Automated Neuroimaging Pipeline

The neuroimaging pipeline is composed of multiple stages (Fig. 4.2). We start with the sMRI images as input to the pipeline. First, we automatically segment the high resolution sMRI using SynthSeg+, a robust deep learning MRI segmentation tool. This segmented sMRI can be directly interrogated for biomarkers or used for automatic diagnosis. For diagnosis, the dimensionality is reduced followed by training a classifier.

Alternatively, the pipeline can be extended to fMRI by downsampling the segmentation and apply it to the fMRI data to extract the average brain activity in each brain region (Fig. 4.3). Afterwards, DDC is applied to obtain functional connectivity per subject as biomarkers. From this, a graph is obtained for each subject with anatomical regions as the nodes and FC as the weights between them.

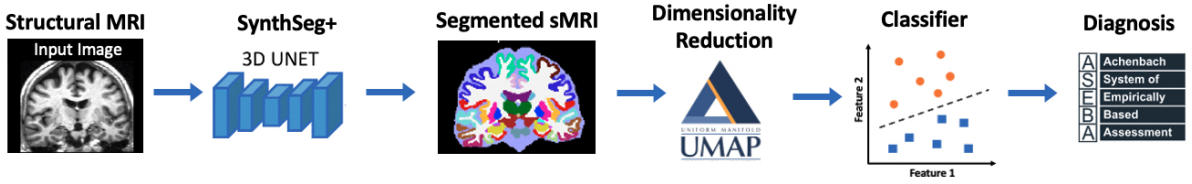

Figure 4.2. Automated neuroimaging pipeline for large-scale structural analysis. Structural MRI images are automatically segmented using deep learning. The brain region volumes act as biomarkers or their dimension can be reduced using manifold projections followed by classification to diagnoses.

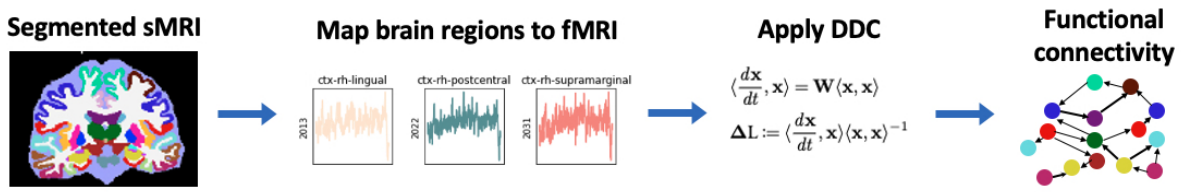

Figure 4.3. Functional connectivity neuroimaging pipeline. After segmentation shown in Fig. 4.2, the regions are mapped onto functional MRI data. Dynamical differential covariance is then used to obtain directed functional connectivity graphs for each subject as biomarkers.

4.2.3 SynthSeg+: A robust, automated segmentation tool

The first step of the pipeline involves the automatic segmentation of structural MRI data. Segmentation is an important step in MRI analysis as it breaks the image into its anatomical regions or other meaningful ROIs. It also reduces the dimension of the data, allowing for averaging or other representative measures per region. Traditionally, manual delineation has served as the gold standard for MRI segmentation. However, this approach is laborious, expensive, and requires specialized expertise, making it untenable for large scale usage. Existing software packages designed to automatically segment MRI data are unable to handle heterogeneous data types encountered in clinical settings. Instead, they work best with high-quality MRI scans at high resolution. To address these challenges, SynthSeg+ has emerged as the first automated algorithm robust enough to handle high variability found in MRI acquisitions, including variations in contrasts, resolutions, orientations, artifacts, and subject populations. By encompassing a wide range of MR sequences and contrasts, as well as accounting for real-life artifacts such as low

signal-to-noise ratio or incomplete field of view, SynthSeg+ serves as an AI segmentation suite that facilitates robust analysis of heterogeneous clinical datasets. It is able to provide whole-brain and cortical segmentation, intracranial volume estimation, and automatic detection of faulty segmentation.

The underlying framework of SynthSeg+ is composed of hierarchical networks and denoisers. For the segmentation task, three-dimensional U-Net convolutional neural networks (3D U-Net CNNs) are employed, trained on synthetic MRI scans possessing fully randomized contrast and resolution. Specifically, the image is fed into a hierarchical system of 3D U-NET CNN segmenters and denoisers. This training strategy enables the networks to acquire domain-agnostic representations, thereby exhibiting excellent generalization capabilities when confronted with heterogeneous MRI data. T The automated quality control (QC) is implemented with a regressor that provides QC scores for 10 regions. Importantly, SynthSeg+ has demonstrated its efficacy in effectively handling large datasets, further solidifying its practicality and reliability for widespread use. It was applied to 14,752 scans from 1,367 patients and was able to reproduce well-known aging atrophy patterns [36].

SynthSeg+ was used to segment the overall baseline ABCD sMRI dataset, including the subset used for depression analysis. It was segmented using the Desikan-Killany atlas consisting of 98 ROIs, including 68 cortical and 30 subcortical regions. For analysis including automatic QC, the same threshold was used as in the original SynthSeg+ analysis where a segmentation is rejected if at least one region obtains an automated QC score below 0.65.

4.2.4 Dynamical Differential Covariance

Functional connectivity is a measure of connectivity between brain regions that share functional properties. There are multiple statistical methods that are used to estimate functional connectivity. Pairwise covariance is most commonly used. It is a symmetric measure which cannot detect directional coupling or disambiguate unconnected nodes with a common input. It also assumes stationary data, which is not good for changing brain states. Another approach,

time series interactions such as Granger causality and cross convergent mapping (CCM), are the degree to which one time series can predict another time series, but it requires model selection and is computationally intensive. Lastly, generative models such as dynamic causal modeling and Bayes network models search all possible causal graphs and fit the dataset to every hypothesis. They make statistical assumptions and are also computationally intensive. Recently, dynamical differential covariance has been introduced as an improved measure of functional connectivity which links differential covariance and dynamical models of network activity. It has been experimentally validated to accurately emulate connectivity information and is computationally efficient thus readily scalable. Unlike standard covariance, it can uncover directional connectivity as well as disambiguating unconnected nodes with a common input. Additionally, it is statistically robust with low bias and high noise tolerance. It also is able to handle nonstationary data [53].

At the macroscopic level measured with fMRI, the collective activity of a population of neurons and interactions between brain regions can be approximated by linear dynamics because of ensemble averaging. For example, a simple linear auto-regressive model performed best on modeling fMRI. Therefore, we can use a Linear dynamical model (4.1) to describe global recordings. Where the neural activity over time can be described by the product of the square connectivity matrix, W, and the neural activity (such as a fMRI signal) given as a column vector, x.

$$\frac{d\mathrm{x}}{dt} = \mathbf{W}x \tag{4.1}$$

where x is the neural activity given by the BOLD signal and $\mathbf{W}$ is the functional connectivity matrix.

From (4.1), we can obtain an estimate for the linear connectivity matrix, given here as ΔL. The DDC estimator for the connectivity matrix is the product of two matrices. The first is differential covariance, which is the time averaged outer product of the derivative of the signal

and the signal. The second is an entry in the partial covariance matrix, which is the inverse of the time averaged outer product of the signal and itself.

$$\begin{aligned} \langle \frac{d\mathrm{x}}{dt}, \mathrm{x} \rangle &= \mathbf{W} \langle \mathrm{x}, \mathrm{x} \rangle \\ \Delta \mathrm{L} &:= \langle \frac{d\mathrm{x}}{dt}, \mathrm{x} \rangle \langle \mathrm{x}, \mathrm{x} \rangle^{-1} \end{aligned} \tag{4.2}$$

where ΔL is the DDC estimator for a linear connectivity $\mathbf{W}$. Inverting the partial covariance allows DDC to account for confounding effects such as correlation due to common input. Differential covariance carries information about sources and sinks, therefore giving the directionality.

DDC has be proven to be able to recover common input connectivity, a case where pairwise covariance fails. It has also been shown to perform better than covariance in recovering the underlying connectivity matrix of a network, especially in nonstationary data which covariance is not able to do. Lastly, when applied to the resting state fMRI (rs-fMRI) recordings obtained from the large-scale Human Connectome Project (HCP) dataset, DDC is shown to obtain accurate connectivity compared to ground truth [53].

4.2.5 Biomarker Detection

Structural Volumes

In addition to brain volume segmentation, SynthSeg+ automatically detect the ROI volumes. In order to test the volume accuracy, the obtained ROI volumes were compared with volumes obtained using FreeSurfer segmentation, as FreeSurfer is known to underestimate obtained volumes [140]. Furthermore, robustness across different collection conditions by comparing differences between MRI manufacturers. This is an appropriate proxy as ground truth segmentations should not show significant differences between manufacturer types. The comparisons were run on the segmentations for the entire ABCD dataset, which is composed

of 11,402 subjects, as well as the depression subset used. For subsets of data less than 5,000, the Shapiro-Wilk test was used to determine normalcy. Between methods of segmenting, the data is considered dependent. As such, a related T-test or Wilcoxon test was utilized to compare the null hypothesis that the two repeated samples have identical expected values for normally and non-normally distributed datasets, respectively. The number of statistically different ROI volumes between manufacturer type was also analysed for each segmentation method using the aforementioned statistical tests.

After validating SynthSeg+, the volumes were compared between control and depressed populations. Again, the Shapiro-Wilk test was used to determine if the volumes followed a normal distribution. In this analysis, the populations are considered independent. As such, the comparison was done using independent T-test and Mann-Whitney U tests for normal and non-normal distribution, respectively.

Functional Connectivity

Using regularized DDC, we are able to obtain functional connectivity graphs for each subject which we are then able to compare between populations. The DDC dataset is non-normally distributed, so the Mann-Whitney U test is used to obtain statistically significant connection differences. In addition to comparing whole brain connectivity differences, it is useful to analyze meaningful subnetworks in the cortex.

Specifically, three subnetworks were compared across subject populations. The first is the Default Mode Network (DMN), which is composed of the medial prefrontal cortex, posterior cingulate, and precuneus cortex. It is mostly employed during rest, spontaneous or self-generated cognition, self-referential processing and emotion regulation. As such, it is expected to be most active during rs-fMRI. The second is the Central Executive Network (CEN), composed of the lateral and dorsomedial prefrontal cortex and posterior parietal cortex. The CEN is primarily activated during cognitive tasks including attention and working memory and is thus implicated in cognitive functioning. The last subnetwork is the Salience Network (SN), which consists of

the insula, dorsal anterior cingulate, amygdala and temporal poles. This subnetwork is activated in response to salient stimuli and plays a central role in emotional control. Furthermore, it is implicated in switching between DMN and CEN dominated states [157].

4.2.6 Dimensionality Reduction and Automatic Detection

As mentioned, the ability to use biological features for automatic disease detection is an important area of research in neuropsychiatry. Here, a dimensionality reduction technique is applied to the volumetric data to allow for classification. Specifically, Uniform Manifold Approximation and Projection (UMAP) was used as a manifold learning technique for dimensionality reduction. It is similar in nature to t-distributed stochastic neighbor embedding (t-SNE), but it has demonstrated a superior ability to preserve global structure with better run time performance. UMAP functions based on Riemann geometry and algebraic topology by first building a high dimensional graph representation of the data then optimizing it to a low-dimensional space while maintaining as much information as possible. The main differences from t-SNE are that UMAP uses exponential probability distributions in high dimensions but not necessarily Euclidean distances, the probabilities are not normalized, it uses the number of nearest neighbors rather than perplexity, and it uses a slightly different symmetrization of the high-dimensional probability.

UMAP was applied to the SynthSeg+ derived volumes to obtain a lower dimensional representation of the dataset. UMAP uses weighted graphs as discrete representations of manifolds where the nodes are the brain region volumes and the connectivity is obtained by a nearest neighbor probabilistic algorithm. In this work, supervised manifold learning was applied which did make use of the labels. The number of nearest neighbors used was 50 and various target weights were tested where target weight scales the weight labels are assigned in the learning process. The goal was to maintain separability while minimizing label weight.

4.3 Results

4.3.1 SynthSeg+ compared to FreeSurfer

Compared to SynthSeg+, FreeSurfer consistently and significantly underestimates the intracranial volume (Fig. 4.4A and B). The datasets with and without subjects that failed QC have the same mean intracranial volumes. This suggests that, even though the subjects did not pass QC, their segmentations overall are relatively similar to the whole dataset. Furthermore, larger differences exist in intracranial volume estimation between MRI manufacture types for FreeSurfer-obtained segmentations compared to SynthSeg+ (Fig. 4.4B). Further analysis shows that the number of statistically different ROI volumes between manufacture types is largest for the FreeSurfer dataset 4.5 yellow). In two of the three manufacture comparisons, the FreeSurfer depression subset data varies the second most between scanner types as well. The SynthSeg+-derived segmentations for the depressive subset vary the least between scanner types, exemplifying a large robustness against scanners. In the Siemens manufacture comparisons, the SynthSeg+ dataset with automated QC removal was the most robust. Between GE and Philips, though, the dataset without QC removal is slightly more robust.

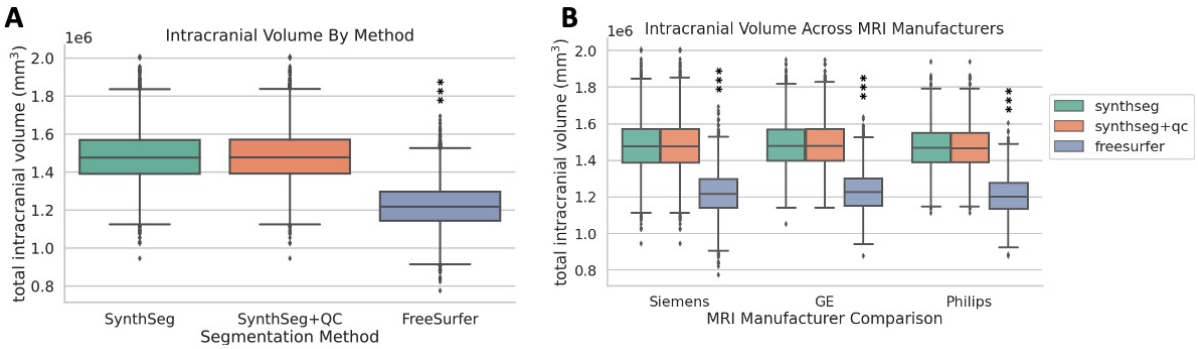

Figure 4.4. Comparison of ROI segmentation methods across the entire baseline ABCD dataset. (**A**) Total intracranial volume estimated SynthSeg, SynthSeg with automated quality control, and FreeSurfer. (**B**) Intracranial volume separated by manufacturer type.

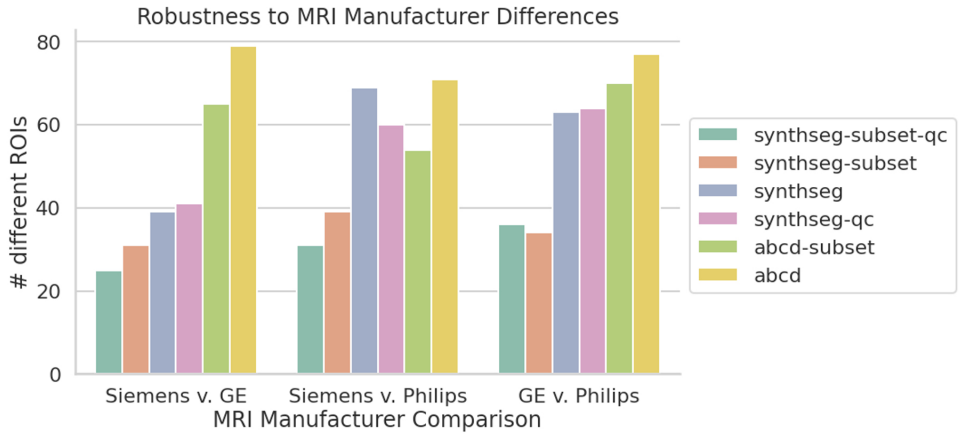

Figure 4.5. Number of statistically significantly different ROI volumes across MRI manufacturers using various segmentation methods. Subset is the depression subset; ABCD are FreeSurfer-derived segmentations. QC denotes removal of subjects that failed automated SynthSeg+ QC.

4.3.2 Structural Volume Biomarkers for Depressed Adolescents

Across all the subjects in the control and depressed populations, there were seven primary regions that displayed differences between groups (Fig. 4.6). These were the left and right postcentral gyrus, right supramarginal gyrus, right cerebrellar white matter, right precentral gyrus, left caudate, and left hippocampus. All regions showed decreased volumes in depressed patients.

Using SynthSeg+'s automated QC, 352 subjects were removed from the dataset, leaving 1,430 subjects in total (1,156 control and 274 depressed). All the aforementioned regions were still different between the populations, but there were an additional four regions as well. These regions included the brain stem, right ventral diencephalon (DC), left cerebrellar white matter, and left cerebral cortex. Again, all volumes demonstrated a decrease in depressed patients.

Across both analyses, the top two differing regions were the left and right postcentral gyrus. It is on the lateral surface of the parietal lobes and contains the primary somatosensory cortex, which is responsible for proprioception. Specifically, this region perceives sensations from the body including pain, touch, pressure, and temperature [72].

Previous studies that focused on depressed adults have also shown decreased hippocampal

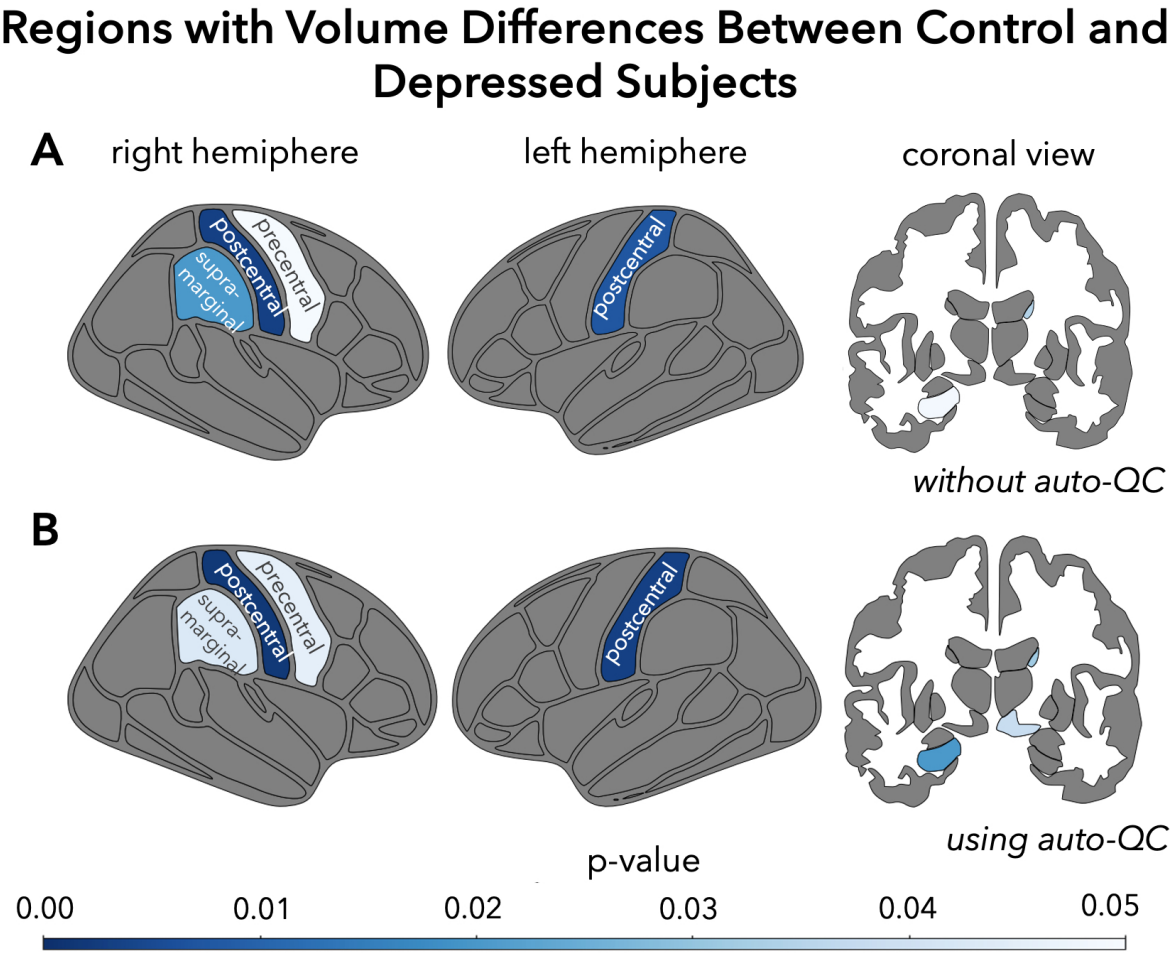

Figure 4.6. Regions with brain region volume differences between the control and depressed subject populations for all subjects (*left*) and only subjects that passed SynthSeg+ automatic segmentation quality control

and cerebrellar white matter volumes. Hippocampal volume reduction is the most commonly cited brain anatomical change in patients with depression. This region plays a vital role in memory processing and emotional management as an essential part of the limbic system. Decreased frontal lobe is also one of the most frequently cited differences, but adult studies implicate the prefrontal and frontal cortex primarily rather than the precentral gyrus [64].

Focusing on the common different regional structures, the precentral gyrus is associated voluntary motion. The supramarginal gyrus's main function is in phonological processing, or the processing of spoken and written language and emotional responses. Lastly, the caudate is

involved with movement planning as well as learning, memory, reward, motivation, and emotion [72]. Other regions that have been implicated elsewhere and not here include the cingulate cortex and putamen [64].

Abnormal white matter is also heavily implicated in depression although there are inconsistent findings in these studies as well. Right cerebrellar white matter in addition to changes in the corpus callosum are most typically cited. Here, changes in the right cerebrellar white matter are observed, although the left cerebrellar white matter also noticed differences once the subjects were removed with failed segmentations, suggesting possible skew in previously published works [64]. The left cerebrellar cortex also showed an associated decreased volume.

In addition to the left cerebrellar matter, the brain stem had a decreased volume in depressed patients after QC removal. This has been found previously and is considered to be related to the nuclei of neurotransmitters implicated in depression and used for antidepressants being located in the brain stem. This includes the raphe nucleus for serotonergic neurons, the substantia nigra and ventral tegmentum for dopaminergic neurons, and the locus coeruleus for norepinephrine neurons [169].

The right ventral DC was also smaller in depressed adolescents. This region is located between the telencephalon and the midbrain as part of the forebrain. It consists of the thalamus, hypothalamus, epithalamus, and subthalamus. Thalamus reduction has been cited in depressed patients previously and is related to memory, information transmission, and emotional management [64].

4.3.3 Dimensionality Reduction and Automatic Detection

UMAP applied to the brain volumes obtained from the segmented sMRI displays clear separation between class types when using label weighting (Fig. 4.7), even with low label weighting. The completely unsupervised representation (Fig. 4.7 *left*) does not show visible separability between classes. With even a moderate label weighting of 0.5, the classes are completely distinct with no outliers (Fig. 4.7 *middle*). In the low label weighting, there are four

outliers from the depressed grouping in the control cluster. These outliers are worth interrogating further to assess the true nature of their disorder. Such outliers shown here could, in theory, represent inaccurate or variant diagnoses. Additionally, there is a slight break in the depressed group resulting in two almost distinct clusters further worth interrogating.

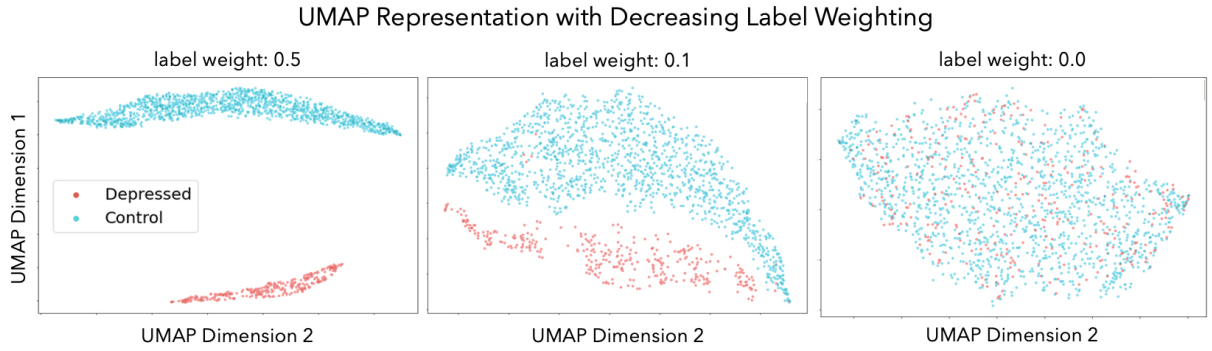

Figure 4.7. Low dimensional representation of depressive structural volume dataset using Uniform Manifold Approximation and Projection (UMAP) with label weighting of 0.5 (*left*), 0.1 (*middle*), and 0.0 (*right*)

This methodology is worth experimenting with and optimizing further as a data-driven mental disorder clustering technique. In addition to this dataset, it would be interesting to apply a similar analysis to the entire ABCD dataset as well as at later timepoints.

4.3.4 Functional Connectivity Biomarkers for Depressed Adolescents

There are observed differences in cortical connectivity differs between healthy and depressed adolescents (Fig. 4.8). The top five regions that exhibit disruptions are connections from the lingual (ROI 11) and rostral middle frontal (ROI 25) gyri in the left hemisphere; and the caudal middle frontal gyrus (ROI 36), temporal pole (ROI 65), and entorhinal cortex (ROI 38) in the right hemisphere. Fig 4.8A shows the top regions colored by the number of different connections observed between population. Fig 4.8B shows a matrix representation of the statistically different functional connections between populations and, notably, the top regional differences are all in the outgoing direction.

The lingual gyrus, also known as the medial occipitotemporal gyrus, lies in the occiptal

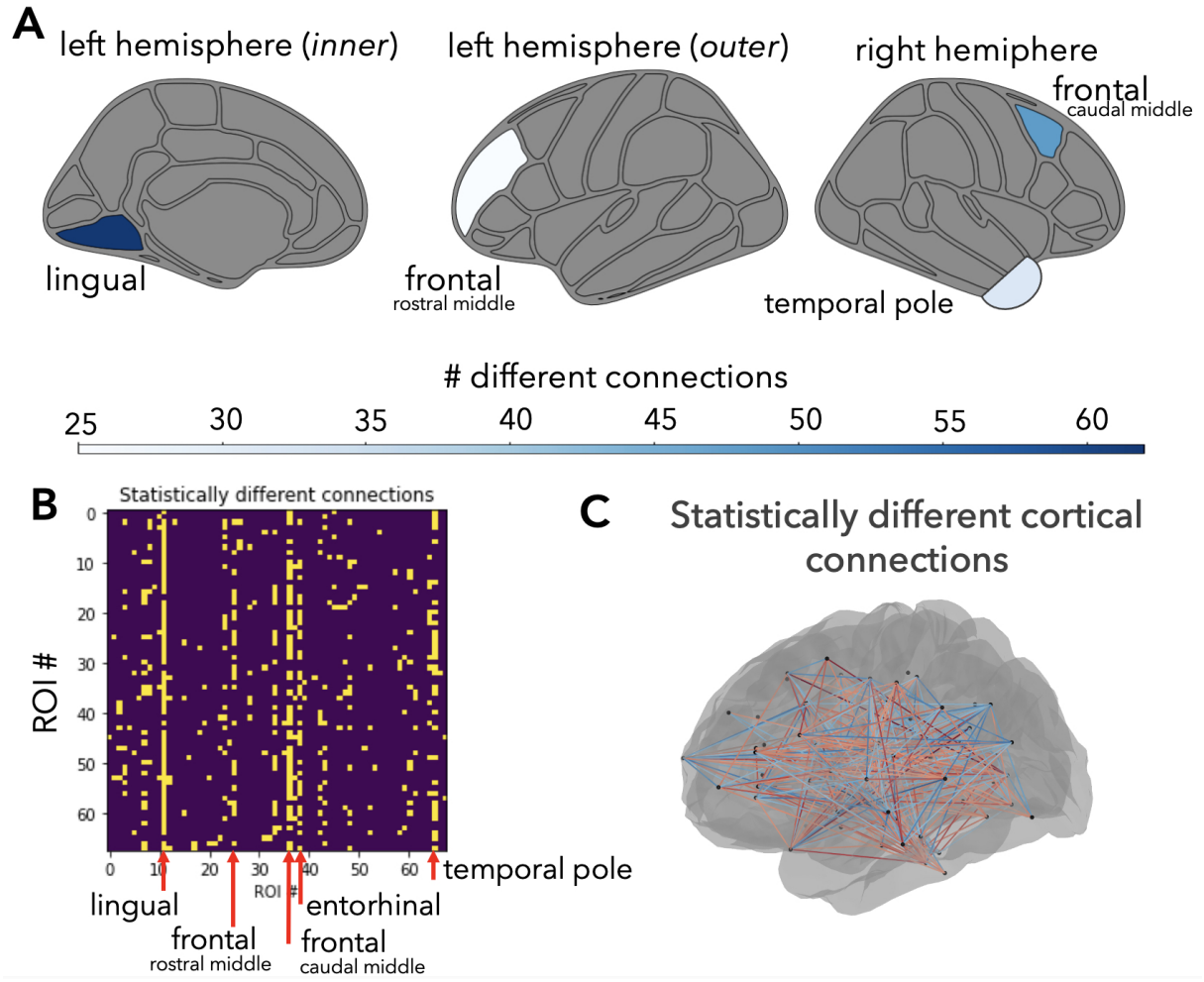

Figure 4.8. (A) Top cortical regions with disrupted functional connectivity (FC) colored by the number of different connections. Statistically different FC between populations shown in matrix (B) and graph format (C). Outgoing connections are on the horizontal axis and incoming are on the vertical.

lobe. It is considered to play an important role in vision, dreaming, and word processing [168]. The middle frontal gyrus is located in the frontal lobe, and its function is still an ongoing area of research. In general, the frontal lobe is considered to be the seat of higher-order executive functioning. The left (dominant) middle frontal gyrus has been shown to play a role in literacy development and the right (nondominant) in numeracy [122]. The frontal eye fields, which control eye movement, are also located in the caudal middle frontal gyrus [210]. Disruption in middle frontral gyrus connectivity has been observed in other depression studies and is

considered to be closely related to patients' low mood symptoms as well as maladaptive mood regulation [64].

The entorhinal cortex lies in the medial temporal lobe and is the main interface between the hippocampus and neocortex. As such, it is important in declarative memories including spatial navigation in memory formation, memory consolidation, and memory optimization [110]. Lastly, the temporal pole is in the temporal cortex and is a paralimbic region involved in high level semantic representation and socio-emotional processing. Moreover, it is heavily implicated in discussions of the mind [64].

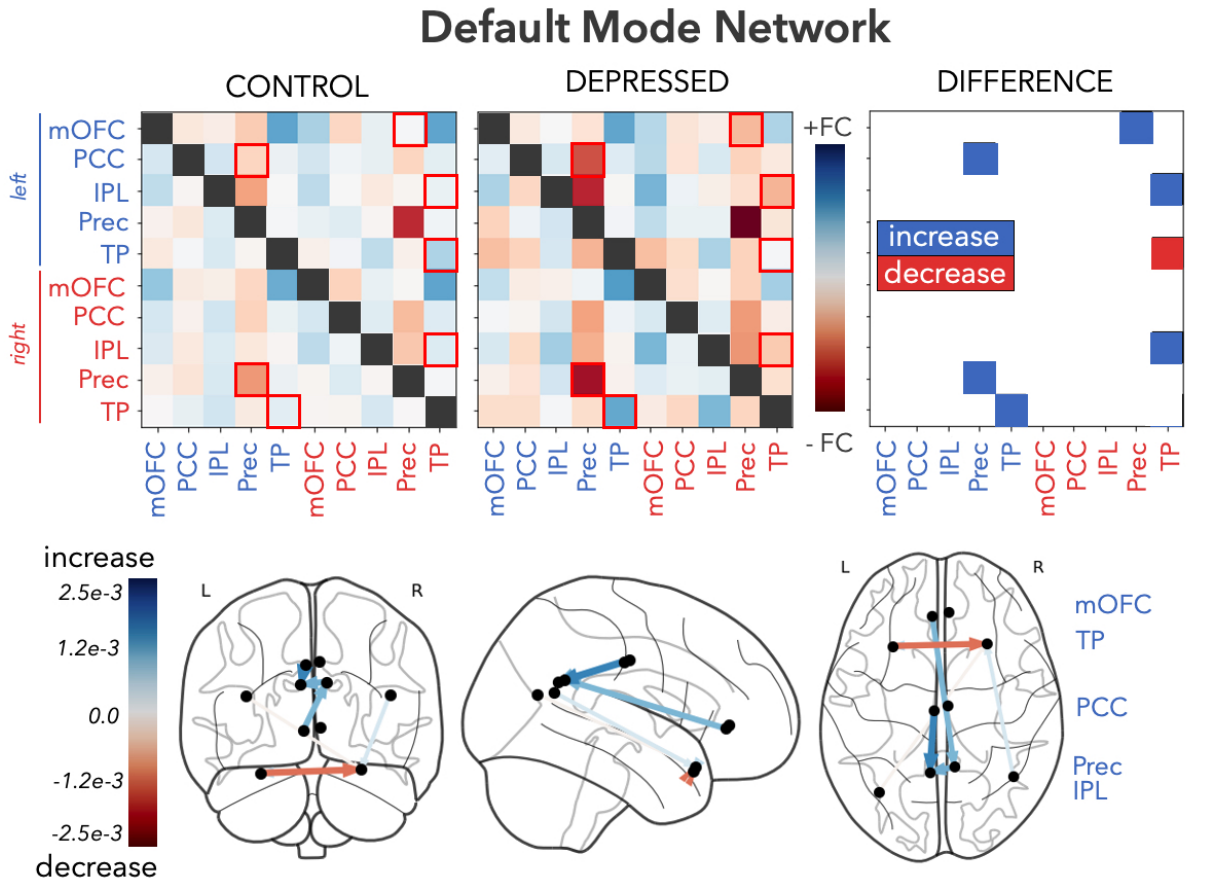

Figure 4.9. Functional connectivity in the Default Mode Network. (*Top*) Connectivity matrix for DMN in the control (*left*) and depressed subjects (*middle*). (*Right*) Significantly different FC as an increase (blue) or decrease (red) in depressed subjects. (*Bottom*) Changed connectivity in depressed patients.

In the DMN, the connectivity of multiple regions is increased in depressed patients (Fig.

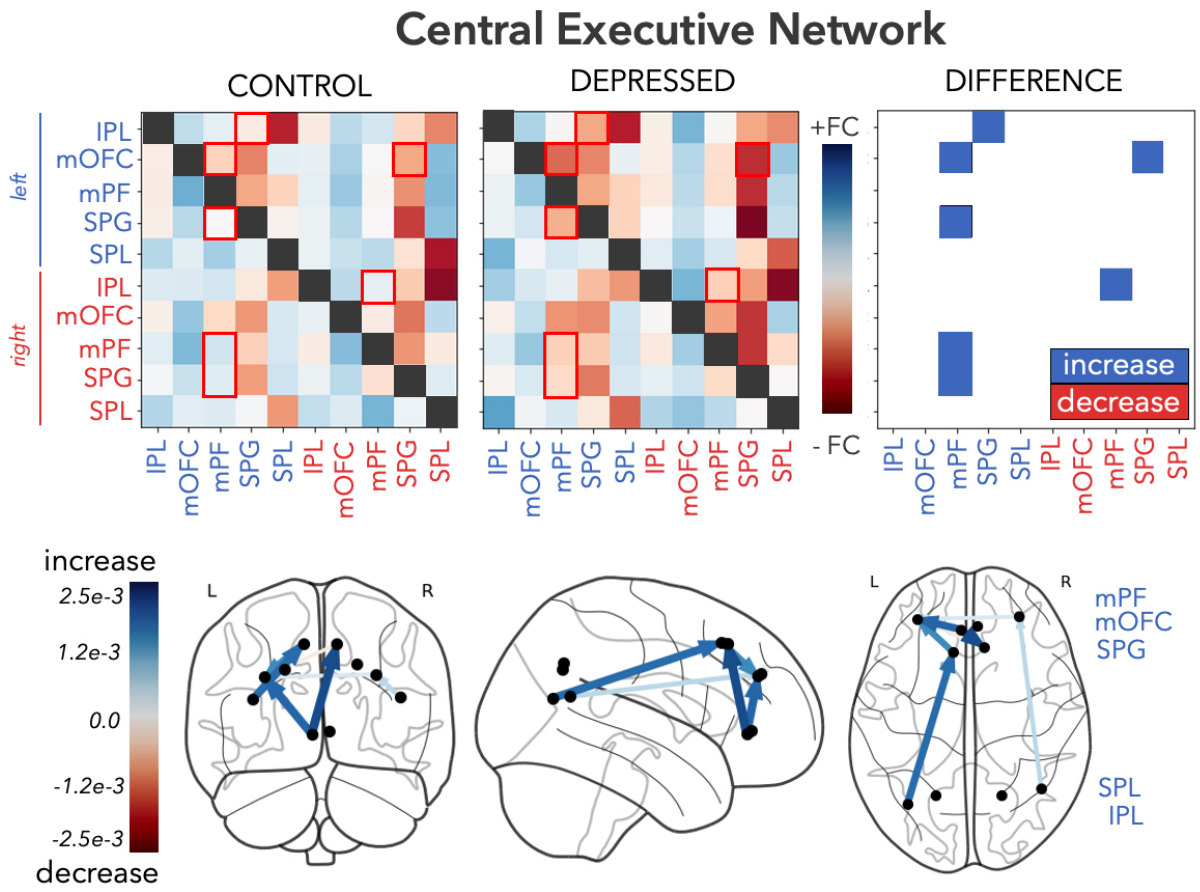

Figure 4.10. Functional connectivity in the Central Executive Network. (*Top*) Connectivity matrix for DMN in the control (*left*) and depressed subjects (*middle*). (*Right*) Significantly different FC as an increase (blue) or decrease (red) in depressed subjects. (*Bottom*) Changed connectivity in depressed patients.

4.9). This phenomena has been seen in other studies and is considered to be closely related to rumination symptoms [64]. Specifically here, the connectivity from the left temporal pole (TP) to the right increased while the reverse decreased. There was an increase from the right TP to both inferior parietal lobes (IPL). Lastly, there was an increase from the left and right precuneus cortex (Prec) to the posterior cingulate (PCC) and medial orbitofrontal cortex (mOFC), respectively.

In the central executive network, there is also increased connectivity between a number of regions (Fig. 4.10). This is contrast to some prior studies which found decreases in connectivity within this network in depressed patients, although again primarily from adult studies and using

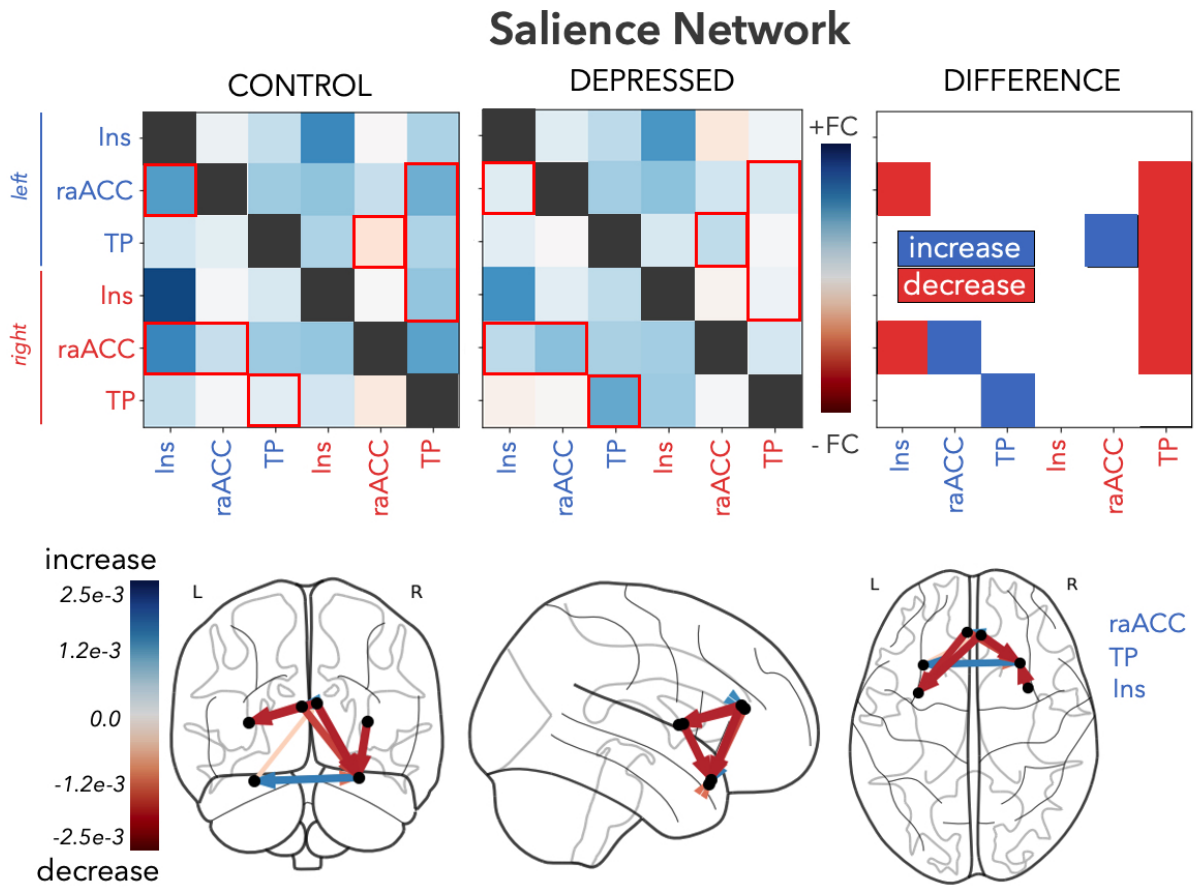

Figure 4.11. Functional connectivity in the Salience Network. (*Top*) Connectivity matrix for DMN in the control (*left*) and depressed subjects (*middle*). (*Right*) Significantly different FC as an increase (blue) or decrease (red) in depressed subjects. (*Bottom*) Changed connectivity in depressed patients.

pairwise covariance [64]. The region exhibiting the largest changes are the outgoing connections from the middle frontral gyrus (mPF), which is the same as found in prior works as mentioned previously. Connectivity is increased from the left mPF to the left mOFC, right mPF and both superior frontal gyri (SPG). Additionally, there is increased connectivity from the left SPG to the left IPL, from the right mPF to the right IPL, and from the right SPG to the left mOFC.

Lastly, the salience network showed the largest change between populations (Fig. 4.11), which is in agreement with what has been found elsewhere leading to abnormal salience network connectivity being considered one of the most crucial links in the pathogenesis of depression

[64, 169]. Prior works have also found increased connectivity between the anterior cingulate cortex (raACC) and insula (Ins), but here we find a decrease that is specific to the outgoing connections from the left Ins. This change is suspected to impact the connectivity between the default mode network and salience network, thus hindering the transition from rest state to task state.

Again, there significant changes in the TP connectivity with increased connectivity from the left TP to the right, but the outgoing connections from the right TP to all regions except the left Ins is markedly decreased. As mentioned previously, this region is involved with socio-emotional processing and heavily implicated in the concept of the mind.

Overall, this analysis demonstrates significant network disruption for depressed adolescents. The majority of the changed regions corroborate previous findings, but directionality is different in some cases. This is likely due to the usage of DDC, a directional measure, while most studies use pairwise covariance. Of the three network, the salience network showed the most differences, suggesting future investigations should focus on this subnetwork.

4.4 Discussion

Overall, we have built an automated pipeline for biomarker analysis and automatic disease detection for neuroimaging data in adolescent depression. There are multiple benefits of developing such a tool. It automates tasks that were previously manually intensive, which is expensive, tedious, and typically requires domain expertise. As a result, it is scalable to large datasets such as this. It further provides a standardized and thus reliable and replicable analysis methodology. Furthermore, it is generalizable to other scientific questions of interest. For example, here we investigate depression, but it could be easily used to investigate any mental disorder screened for in the ABCD study. Lastly, it is a precision medicine approach meaning that, instead of averaging across patients, it utilizes per subject information, creating a truly personalized framework.

In order to obtain a realistic understanding of the biophysiology associated with mental disorders, including depression, large and multimodal datasets are required. Naturally, these datasets become difficult to utilize with existing manual tools. As such, there is a need for automated methodologies to preprocess and analyze the Big Neuroscience datasets. Presented here is an automated neuroimaging pipeline for obtaining structural and functional biomarkers associated with adolescent MDD with associated depressive biomarkers and automatic diagnosis.

Given the automated and modular nature of the pipeline, it can easily be generalized to any mental disorder screened for in the ABCD baseline dataset, including anxiety, ADHD, oppositional defiant problems, conduct problems, avoidant personality problems, and antisocial personality problems. Furthermore, follow-up neuroimaging studies after the baseline can also be analyzed. The current pipeline solely utilizes resting fMRI data, so a natural extension would be to include the task fMRI data types, including the Monetary Incentive Delay Task, Stop Signal Task, and Emotional N-Back Task. fMRI can be similarly assessed using dimensionality reduction techniques and combined with the structural data to provide additional predictive power for automatic diagnosis. Furthermore, the ABCD dataset provides many other data types that can be integrated into this pipeline, including genetics and neurocognition.

The ABCD is particularly powerful due to its longitudinal nature. Therefore, utilizing multiple time points in this dataset would allow temporal analysis and early risk detection. It would also provide insight to the time course of disease progression in mental disorders. Furthermore, as the subjects undergo potential treatment as the study progresses, analysis of treatment efficacy can be integrated as well. Additionally, while better than DSM labels, using supervised ASEBA labels is a limitation of this work. The unsupervised aspect of UMAP dimensionality reduction can be used to assess the entire dataset to analyze how these biological biomarkers cluster in relation to existing diagnosis labels, both DSM and ASEBA.

Lastly, such a large dataset is rich for usage with deep learning models. Specifically, attention-based deep learning methods, such as 3D CNNs with attention [200, 229] for sMRI and Graph Attention Networks (GAT) [216] or attention-based spatio-temporal graph convolutional

networks (STGCN) [226]. These models would operate on the data without dimensionality reduction while the attention aspect would provide explainable insight to the biological features being used for automatic detection. In addition to more complicated deep learning models, different machine learning models could be tested in place of the current SVM as well as hyperparameter optimization.

This is an exciting area of research that holds significant promise for translational applications to clinicians. The field of psychiatry is on the precipice of a quantitative and biological revolution bringing into the modern age of medicine, where it will be possible to identify, diagnose, and treat mental disorders using concrete and actionable items.

Acknowledgment

Chapter 4 is coauthored with Wagner, Margot; Camassa, Alessandra; Liu, Brandon; Gano, Kameron; Cauwenberghs, Gert; and Sejnowski, Terrence J. The dissertation author was the primary author of this chapter.

Chapter 5

Using Natural Language Processing as a Scalable Mental Disorder Evaluation Technique

5.1 Introduction

There is currently a mental health crisis as the demand for mental health services vastly outstrips the availability of quality care. In the United States and Canada, approximately 60 million people grapple with mental health concerns, yet regrettably, over two-thirds of those in need are unable to receive care despite a staggering $250 billion in healthcare expenditures [20, 19]. The prevailing model of care delivery relies on one-on-one interactions between clinicians and patients, which is labor-intensive and thus lacks scalability. The paucity of mental healthcare clinicians hampers the efficacy of this traditional format. The World Health Organization estimates around half the world's population lives in countries where there is one psychiatrist for 200,000 or more citizens [16]. Furthermore, the considerable degree of clinician involvement further inflates the cost of conventional treatment. The COVID-19 pandemic has exacerbated this mismatch in care by increasing demand for services, fostering mental health awareness and diminishing the stigma associated with seeking care. This has created a persistent disparity between the supply and demand for mental healthcare, enduring even beyond the pandemic's cessation. Consequently, it is imperative to prioritize the development of innovative

solutions that streamline and automate the delivery of mental healthcare in order to address this public health concern.

The field of mental health care has lagged behind broader innovation in the healthcare industry. This discrepancy primarily arises from the dearth of robust and precise quantitative measurements, biomarkers, and evidence-based practices in this domain compared to other medical specialties. Thus, it becomes essential for the field to center its efforts on the cultivation of novel data-driven measurement tools which can effectively evaluate the mental status of patients. Thomas Insel, former head of the National Institutes of Mental Health, argues that a new diagnostic system based on emerging research that incorporates multiple layers of information is a pressing need in psychiatry [114, 108]. By doing so, the aim is to construct comprehensive and objective protocols that are capable of adeptly handling mental health issues, thereby bridging the existing care gap and ensuring optimal treatment for individuals in need.

Establishing objective measurement tools for mental health evaluation is challenging due to their inherent complexity. Unlike other chronic diseases which typically utilize one or few biophysiological target variables, such as blood glucose in diabetes or blood pressure in hypertension, mental health disorders lack definitive biomarkers. As a result, there is a reliance on indirect measurements, including changes in sleep pattern or activity level [230] or changes in body posture [87] and speech tone [113]. While these biophysical symptom proxies hold significance in a clinician's decision-making process, the patient's thought content remains the most influential factor in diagnosis and treatment determination. The assessment of thought form and content, mood status, stressors and anxiety level predominantly relies on the patient's verbal expression. Consequently, it becomes natural that the pursuit of a readily scalable technology for appraising a patient's mental status would incorporate the patient's linguistic utilization as a reliable information source. However, patients' speech is unstructured data, which makes the process of extracting clinically relevant data a challenging endeavor. Even in clinical settings, therapists and clinicians must attentively and carefully listen to their patients to extract well-defined structured data suitable for diagnosis, intervention and monitoring purposes.

Deep learning methods have gained prominence in the field of mental illness detection, specifically in the task of natural language processing (NLP) [228, 85, 44, 23, 98]. Unlike traditional statistical and machine learning methods, deep learning methods do not heavily rely on feature engineering and can process longer, more complex sentences in a context-dependent manner. They also exhibit enhanced capabilities in learning languages structures, allowing for effective transfer learning with limited data. Transformers, newer than convolutional and recurrent neural networks, show promise in handling sequential and textual data, making them suitable for mental health applications[71]. In this work, the focus is on fine-tuning a pretrained transformer model to detect symptomatic sentences related to depression and anxiety in a client's narrative. The paper acknowledges the state-of-the-art position of machine learning in NLP and includes a comparison with commonly used models in the results section.

5.2 Methods

5.2.1 Datasets

Training Dataset

The non-clinical training data for this study was collected from online mental health forums where individuals share their personal experiences and challenges related to mental health. To prepare the data, the stories were segmented into sentences, and each sentence was carefully examined and labeled by an expert clinician (Expert A) as either neutral or exhibiting signs of anxiety and/or depression. In addition, sentences that relied on the previous or following sentence for context were flagged as dependent examples. An illustration of this dependency can be seen in the following pair of sentences: "Would I say my life is perfect and I am happy every day? No.". In isolation, each sentence does not provide a clear indication of symptoms, but when considered together, they reveal relevant information. Therefore, both sentences were marked as dependent examples. Furthermore, sentences that were unrelated, such as asides or emphatic statements, were also identified and flagged. These two categories of sentences–dependent and

unrelated–were subsequently removed from the dataset as they were not suitable for labeling in the context of this research question. Out of the initial 3,780 sentences, only 97 sentences (2.6%) were removed based on criteria, resulting in a dataset of 3,683 sentences. Sentences labeled as displaying signs of anxiety, depression, or both were categorized as symptomatic (positive) examples while neutral sentences served as non-symptomatic (negative) examples. The resulting training dataset was 36% symptomatic and 64% non-symptomatic sentences. Once the training data was prepared, it was shuffled and split into a typical 80% training (2,946 sentences) and 20% validation (737 sentences) dataset. The test set for evaluation purposes consisted of a clinical dataset.

The training sentences used in this study had an average length of 19 words and 102 characters for symptomatic sentences, whereas non-symptomatic sentences had shorter averages of 16 words and 83 characters. In terms of data sources, each user contributed an average of 26 sentences. The highest number of sentences from a single patient in the dataset was 180 while the lowest contribution was 2 sentences.

Testing Dataset

The clinical data utilized in this study was collected from a board-reviewed and ethically-compliant online psychotherapy clinical trial conducted at Queen's University between 2020 and 2021. The study underwent a thorough review process by the Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board to ensure adherence to ethical standards (File #: 6020045). As part of their participation, patients provided written informed consent for the utilization of their anonymized data in academic research and publications. There were 55 subjects that participated in the trial. During the trial, participants diagnosed with major depressive disorder (MDD) received 12 sessions of therapist-supported electronic cognitive behavioral therapy (e-CBT) in asynchronous format. This asynchronous therapy involved engaging with weekly interactive online modules, which were delivered through a secure cloud-based online platform. During the initial week of the trial, participants were invited

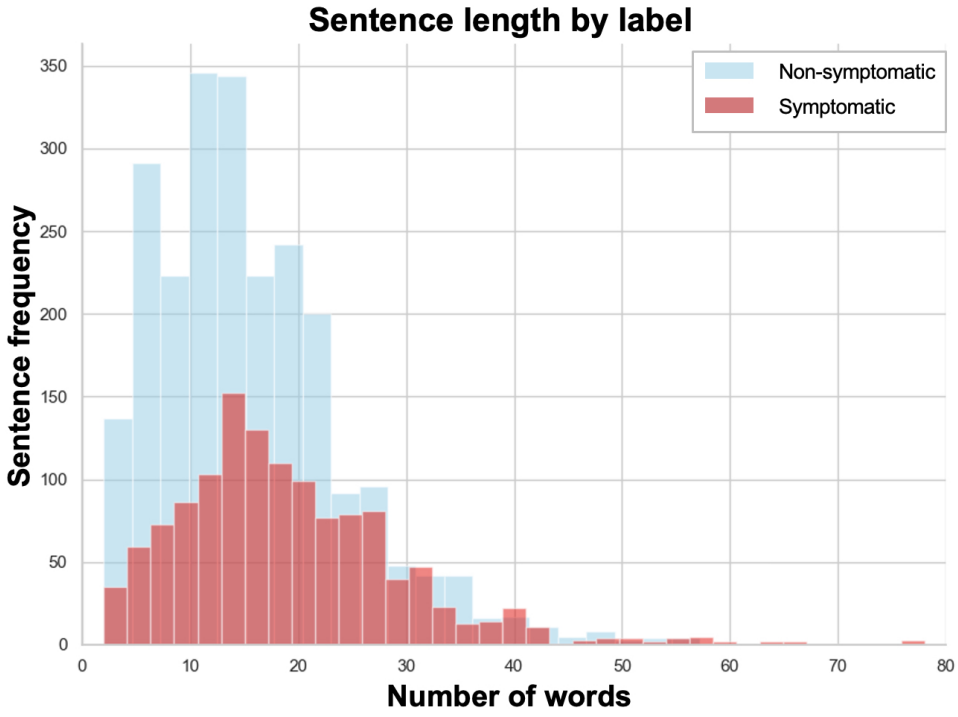

Figure 5.1. The sentence length by label type, symptomatic and non-symptomatic where symptomatic sentences were less frequent and shorter.

to share their personal narratives detailing their experiences with mental health challenges.

Participant narratives were segmented into 930 total sentences. Following the same inclusion criteria as the training dataset, 31 sentences (3.3%) were excluded, leaving 899 sentences for testing the algorithm performance. These sentences were similarly labeled as neutral or containing signs of anxiety and/or depression by two expert clinicians (Expert J and Expert M), who were different from the clinician involved in labeling the training dataset. Among the 899 sentences, Expert J considered 28% as symptomatic and 72% as non-symptomatic, while Expert M categorized 41.5% as symptomatic and 58.5% as non-symptomatic. Notably, this resulted in an inter-rater overlap (i.e. proportion of sentences having similar labels from both Experts J and M) of 76%, indicating a significant level of agreement between the two experts regarding sentence labeling. To assess label consistency across datasets, Expert J was also tasked with labelling the training dataset. The inter-rater overlap between Expert A and Expert J for the training dataset was 80%. These findings highlight the subjective nature of diagnostics in this

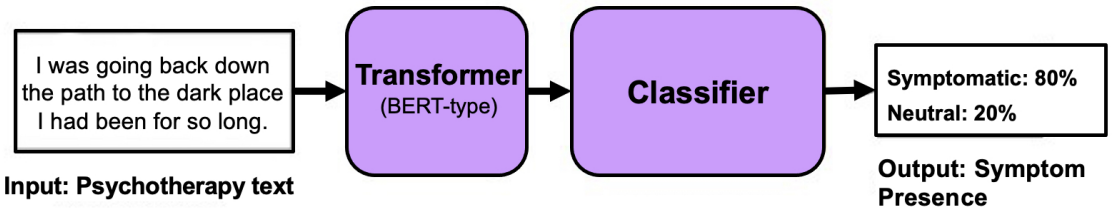

Figure 5.2. Natural language processing mental health task process figure where a narrative sentence is input to a transformer model with a classifier to predict status as either symptomatic of non-symptomatic.

field within this field.

5.2.2 Model Design

Due to the vast array of tasks in NLP, it is crucial to clearly define the specific task of interest prior to any modeling work. In this study, the task is a classification task, where the objective is to classify the text input based on a predefined label. Specifically, it is a binary sentence classification task, where the aim is to categorize a sentence input as one of two labels, symptomatic or non-symptomatic. Two particularly relevant subtasks in text classification are emotion recognition and sentiment analysis. Emotion recognition aims to assign a specific emotion (e.g. happy, sad, angry) to the input sentences. On the other hand, sentiment analysis focuses on capturing the overall attitude expressed in an input sentence (i.e. positive, negative, neutral). Given the nature of mental health, particularly anxiety and depression, these two subtask categories exhibit strong interrelationships and relevance to the current study.

5.2.3 Model Training

The Transformer model class was selected due to its superior performance in NLP tasks related to emotion detection, surpassing previous models that lacked contextual understanding [71, 215]. A number of models were selected from the HuggingFace transformer model library, including a standard Bidirectional Encoder Representations from Transformers (BERT) model as well other BERT-based models including RoBERTa [132], DistilBERT [193], ALBERT [125],

DeBERTa [94], and XLM-RoBERTa [63]. BERT was chosen as it has established itself as the de facto and widely adopted baseline for NLP experiment [185]. In addition to training the aforementioned baseline models, training was conducted on commonly employed baseline model variants, such as cased and uncased versions, as well as large-sized models, ensuring a comprehensive exploration of the model landscape.

A selection of transformer models that had undergone further fine-tuning for text classification and specific subtasks were also included in the initial model selection phase. While models from subtasks relevant to the current task (emotion recognition, sentiment analysis) were of particular interest, a wide range of subtasks were considered. These models were chosen based on their popularity within the HuggingFace community (as determined by the number of downloads) and their demonstrated performance in various tasks. A total of 44 unique models were tested using a standardized set of hyperparameters. Each model was trained for 5 epochs. The remaining training hyperparameters were set to their default values provided by HuggingFace.

The majority of the models (75%) were baseline models that had been fine-tuned to another task before our training. Among these fine-tuned models, 57% were specifically fine-tuned for a text classification task, 16% for token classification, and 2% for fill mask tasks. The token classification models were exclusively tuned for name entity recognition (NER) subtasks while the fill mask task was for a biological subtask. The text classification models were primarily trained on subtasks of sentiment analysis (36%) and emotion recognition (32%) subtasks.

Prior to model training, the training examples underwent tokenization using HuggingFace's tokenizer class, which employs base model-specific (i.e. BERT, DistilBERT, RoBERTa) tokenization techniques. It is worth noting that the training data exhibits an imbalance, with only 36% of sentences meeting the criteria for being symptomatic. To address this imbalance, a weighted cross-entropy loss function was employed during model training, where the weights were determined based on the distribution of the two classes. Model accuracy was used as the sole criteria for model selection due to the sensitive nature of the given task, which necessitates

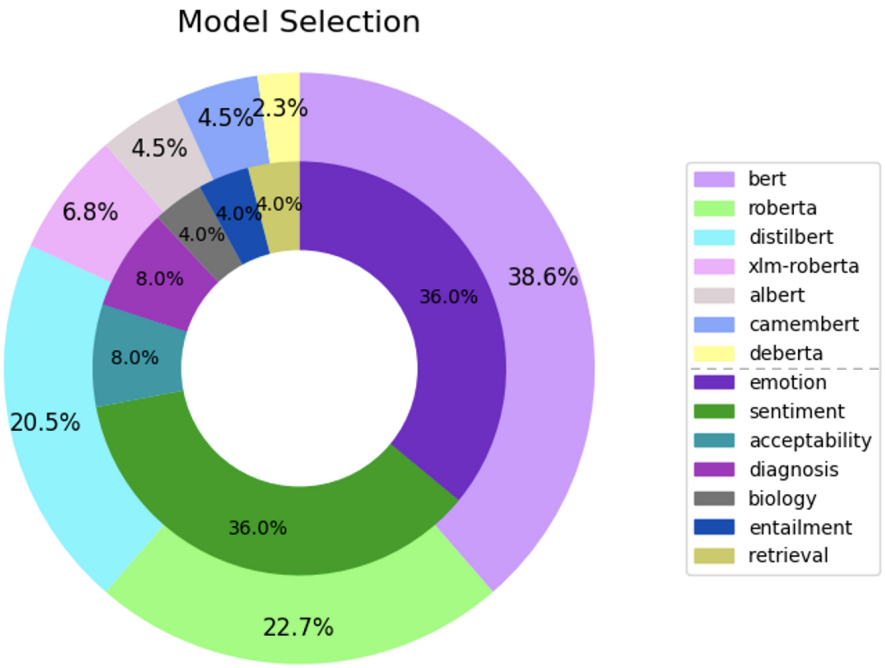

Figure 5.3. Task and subtask distributions of models used in model selection. BERT models pretrained on an emotion recognition task prior to current study training.

maximizing correct predictions even if it comes at the expense of computational efficiency, such as model size or latency. Specifically, the F1 score was used as the model accuracy metric, as it maintains a balance between precision and recall. For the given task, it is crucial to try to accurately predict as many symptomatic cases as possible (recall) while also maintaining a high level of confidence in the positive predictions (precision). Therefore, both metrics were considered essential in evaluating the model performance.

Data augmentation serves as a valuable approach to enhance the training dataset by introducing additional examples through slight modifications of existing ones. While in image datasets this can be achieved through simple techniques such as scaling, rotation, or color manipulation, text datasets require methods that preserve sentence meaning. Simple text augmentations like random word swap or insertion are not sufficient for complex tasks such as emotion detection, where sentence meaning plays a pivotal role. Therefore, alternative methods that better preserve sentence meaning were considered and compared, including the use of back translation.

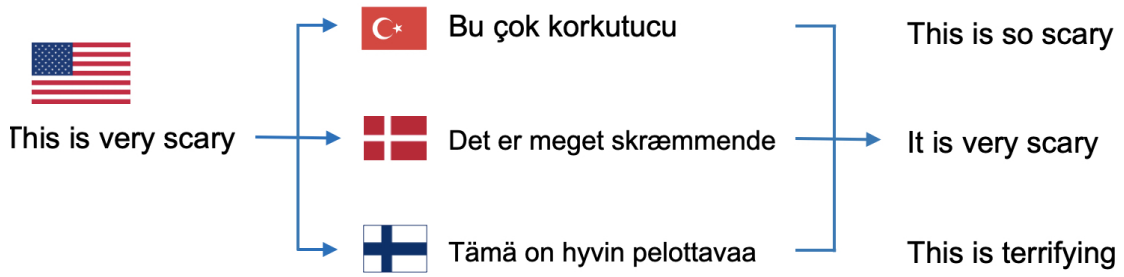

Figure 5.4. Back-translation method of data augmentation. An English sentence is translated to another language – here Turkish, Danish, and Finnish – then translated back to English with slight variations.

These methods offer more effective ways to generate synthetic training examples that contribute meaningfully to the training process.

Back-translation, the technique employed in this study, involves translating the original sentence to another language then back to English. It was implemented here with the NLPAug library for textual augmentation, utilizing HuggingFace transformer-based translation models. This approach is expected to maintain the sentiment of the sentence more accurately. The number and type of language intermediates used in the translation process were treated as additional tunable hyperparameters. Each available language intermediate on HuggingFace (as of 2022) was individually tested, effectively doubling the size of the training dataset with the inclusion of the back-translated synthetic counterparts for each sentence. Moreover, various ratios of back-translated synthetic sentences to the original sentence were explored, along with different combinations of language intermediates. A comprehensive investigation was conducted across 20 individual languages, followed by experimentation with the top 11 performing models in different combinations on the complete dataset. The number of languages used was incrementally increased until a drop in performance was observed. For details on this analysis, consult Appendix 1.B.

Hyperparameter tuning was conducted using the Tree Parzen Estimation (TPE) method with HyperOpt, employing an Asynchronous Successive Halving (ASHA) scheduling algorithm implemented with Ray Tune [34, 128]. The selected hyperparameters and their respective ranges

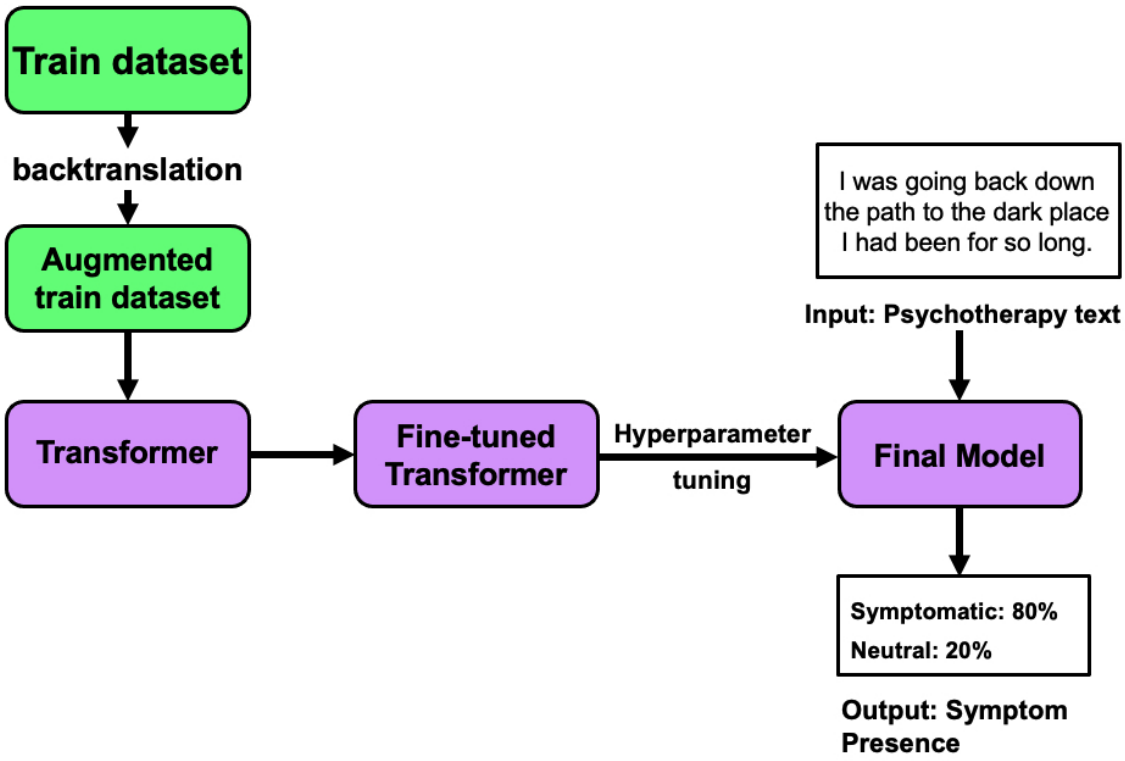

Figure 5.5. Schematic of overall training process. Train dataset is augmented then used to fine-tune a transformer model. The model is optimized through hyperparameter tuning before the final model is attained.

for tuning were chosen based on their ability to significantly enhance model performance in prior works [207]. The hyperparameters chosen included the number of training epochs, the random seed, the number of training examples per batch, and the learning rate. The final hyperparameters used can be found in Table 5.5.

The fine-tuned model trained on non-clinical data was then evaluated on the clinical dataset to assess its diagnostic capabilities. The model's performance was calculated and compared against the labels of Experts J and M, and the results are presented in Table 5.1.

5.3 Results

Performance evaluation was conducted across 44 different models by comparing their F1score (harmonic mean of precision and recall) and balanced accuracy (bA) on the training

Table 5.1. Performance of trained models across datasets.

	Hyperparameter-tuned	Augmented data	Fine-tuned	fastText*
Public dataset	F1: 79.3 bA: 84.0	F1: 79.1 bA: 83.6	F1: 77.2 bA: 82.3	F1: 68.1 bA: 75.2
Clinical Dataset (Expert J Labels)	F1: 73.0 bA: 72.0			
Clinical Dataset (Expert M Labels)	F1: 75.0 bA: 75.0			

** All other non-DL models performed worse than fastText. For more details, please refer to the appendix.*

dataset. The top 11 models, based on their performance, were fine-tuned using the augmented training dataset. A detailed comparison of these models can be found in Table 5.A.1 in the appendix.

$$\frac{\partial u}{\partial t} = D\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + q(x,y,z,t) \tag{5.1}$$

$$u(x,y,z,t) = \sum_n X_n(x,y,z)T_n(t) \tag{5.2}$$

The model that achieved the highest performance in our evaluation was a fine-tuned RoBERTa model that was previously trained on the TweetEval [29] dataset prior to this dataset. TweetEval is a dataset specifically designed for multi-class emotion recognition in Tweets, consisting of over 5000 Tweets categorized into four emotions: anger, joy, sadness, optimism. By leveraging this dataset, the RoBERTa model demonstrated superior performance, outperforming word-based, context-free algorithms like fastText by approximately 10%. For a detailed analysis of each model's performance, please consult Appendix 1.A.

Subsequently, we assessed the performance of the best-performing model, trained on the public dataset, using the clinical dataset that was labeled by two additional experts. The model

achieves an accuracy of approximately 74%. It is worth noting that even human experts do not exhibit complete consistency. The inter-rater overlap between Experts J and M was found to be 76%, which is comparable to the accuracy of the model.

The results table displays results for the baseline model, which refers to the Emotion RoBERTa model. The Augmented column refers to the Emotion RoBERTa model applied to the augmented training dataset. Finally, the Tuned performance is after hyperparameter tuning.

Both the F1 score (F1) and balanced accuracy (bA) are crucial metrics included in our evaluation. These metrics are specifically employed for classification tasks involving imbalance datasets, such as this. It is important to highlight that the F1 score maintains a balance between precision and recall, whereas the balanced accuracy assess specificity and recall. The balanced accuracy metric directly considers true negatives, making it particularly useful when both true positives and true negatives are equally significant. On the other hand, the F1 score prioritizes the positive cases, emphasizing the accuracy of identifying positive instances. By utilizing both metrics, we ensure a comprehensive assessment of the model's performance in handling imbalanced classification scenarios.

5.4 Discussion

In this project, we conducted fine-tuning on a variety of transformer models to identify symptomatic sentences in a client's mental health narrative, specifically those related to depression and anxiety. We aimed to compare the performance of these models and address the limited availability of labeled data by employing augmentation techniques to expand our dataset.

Our findings demonstrate that our most effective model achieved an impressive accuracy of approximately 80% (F1: 79.3%, bA 84.0%) when distinguishing between symptomatic and non-symptomatic sentences, which is comparable to the performance of human experts. Despite being trained on a public dataset, this model showcased a similar level of accuracy when classifying sentences from a clinical dataset collected from 55 patients participating in a separate

clinical trial (average F1=74%, bA=73.5%). Notably, our model's performance is in line with the level of agreement between different expert raters, indicating its reliability.

Through the fine-tuning of various transformer models, augmentation techniques, and comparative analyses, our project has successfully developed a model capable of accurately classifying symptomatic sentences. Its accuracy on both public and clinical datasets, combined with its performance comparable to interrater agreement, highlights its effectiveness and potential for practical applications.

Our study aimed to identify the most effective model for the given task, and our analysis revealed that the RoBERTa model, fine-tuned on the TweetEval benchmark, outperformed the other models examined. RoBERTa is an enhanced version of the original BERT transformer model, benefiting from robust optimization during pretraining, which ultimately resulted in improved model performance [71, 132]. Unlike BERT, RoBERTa underwent pretraining using an expanded dataset, comprising of five English-language corpora that totaled over 160 GB of uncompressed text. These corpora include BOOKCORPUS [232], WIKIPEDIA, CC-NEWS [137], OPENWEBTEXT [86], STORIES [212].

The model was further fine-tuned on the TweetEval benchmark prior to our task. Specifically, it utilized the Emotion Recognition dataset, which contains of over 5000 text statements sourced from Twitter. Each statement is associated with one of four emotions: anger, joy, sadness, and optimism. The exceptional performance of the RoBERTa model highlights the significance of emotion classification as a valuable precursor for mental health diagnostics. It is worth noting that the BERT base model, XLM-RoBERTa pretrained on a sentiment analysis task, and DistilBERT base models closely trailed in terms of performance.

By highlighting the superior performance of the RoBERTa model fine-tuned on Emotion TweetEval, our study underscores its efficacy as the optimal choice for the given task. This model's success indicated the potential utility of emotion classification in facilitating health diagnostics.

One effective strategy for enhancing model training involves employing diverse data

augmentation techniques to expand the training dataset. In our study, we explored the back-translation technique and observed a notable improvement in model performance as a result. Interestingly, we discovered that using intermediate languages from the Indo-European, Turkic, or Uralic language families yielded superior results compared to other language families, such as Sino-Tibetan, Japonic, Austronesian, or Afro-Asiatic. This can be attributed to the fact that languages belonging to the same language family as English tend to capture sentence structure and meaning more effectively due to their greater similarity.

To maximize the benefits of back-translation, we identified the combination of Turkish and Danish as particularly effective, allowing us to triple the size of the training dataset by generating two additional augmented sentences for each origin training sentence. This language combination produced the highest model performance, achieving an approximate 2% increase in the F1 score. For a detailed analysis of the language combination, corresponding performance, and the impact of increasing the ratio of synthetic sentences to original sentences, please refer to the appendix A2, where a table is provided.

By strategically implementing back-translation and specifically leveraging the Turkish and Danish languages, we successfully amplified the training dataset and enhanced the model's performance. The observed improvements validate the effectiveness of this approach for training models for the given task.

During the hyperparameter tuning process, we focused on optimizing several key hyper-parameters, including the number of training epochs, random seed, number of training examples per batch, and the learning rate. While these hyperparameters were carefully tuned, it is worth noting that there may still be room for further optimization. After thorough experimentation, we observed only a modest improvement in the overall model performance, with the F1 scores increasing by a mere 0.2%. For a comprehensive list of the final set of hyperparameters employed in this study, please refer to the appendix A3. Moving forwarded, further exploration and optimization of hyperparameters may garner additional model success. It is an avenue that warrants future investigation and can potentially yield more substantial improvements.

The findings from this study highlight the promising potential of training a transformer model for a nuanced and intricate clinical task, specifically the detection of symptomatic language use, even when faced with limited labeled data. Furthermore, the transferability of the model's knowledge to diverse datasets collected in distinct clinical settings is a crucial outcomes. Ultimately, these transformer models have the capacity to revolutionize the field by enabling scalable and objective mental status evaluations based on patients' language usage.

As we move forward, it becomes imperative to envision future clinical trials that leverage these objective measurements to predict essential clinical outcomes. These outcomes could expand to include factors such as patient engagement, symptom reduction, or even relapse prediction. By incorporating these objective measurements into the design of future trials, we can potentially enhance our understanding of the complex interplay between language use and clinical outcomes.

By utilizing the power of transformer models and their ability to accurately analyze language patterns, we pave the way for more precise and comprehensive evaluations of mental health. This has the potential to significantly impact clinical practice and improve patient care. The next critical step is to strategically integrate these models into clinical trials, enabling the generation of invaluable insights that can inform treatment decisions and interventions, ultimately enhancing patient outcomes and well-being.

Acknowledgment

Chapter 5, in full, is currently being prepared for submission for publication of material. Wagner, Margot; Jagayat, Jasleen; Kumar, Anchan; Shirazi, Amir; Alavi, Nazanin; Omrani, Mohsen. The dissertation author was the primary investigator and author of this paper.

Appendix

A.1 Model Selection

Table 5.2. Performance of trained models across datasets.

Model Name (*source*)	**Model Type**	**F1 Score**	**Balanced Acc**
TweetEmotionEval (*elonzo*)	RoBERTa	77.26%	82.37%
BERT base uncased	BERT	77.04%	82.00%
TwitterSentiment (*cardiffnlp*)	XLM-RoBERTa	76.92%	81.95%
DistilBERT base uncased	DistilBERT	76.81%	81.81%
TwitterSentiment (*cardiffnlp*)	RoBERTa	76.52%	81.60%

Table 5.2 displays the top 5 performing transformer models using F1 score as the sorting metric. These scores are without data augmentation, hyperparameter tuning, or any other optimizations. Balanced accuracy is seen to monotonically increase as well as F1 score. Interestingly, if the models are sorted by balanced accuracy performance, the top four models remain the same, and the fifth best model swaps with the sixth best model (F1: 76.50% BA: 81.72%, model type: RoBERTa). Three of the top five models are a RoBERTa model or model variant. Two are baseline models while three are models fine-tuned on tweets for another related emotion or sentiment task.

Nine of the forty-eight models performed no better than chance, with an F1 and BA of 0.5 and were thus removed from the following analyses as outliers.

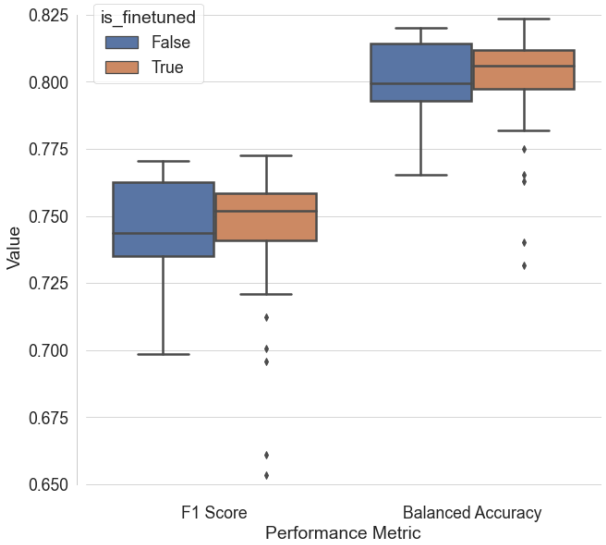

Figure 5.6. Performance of all transformer models trained during model selection. Models previously fine-tuned on related subtasks (*orange*) tended to perform slightly better on average than base models (*blue*) .

Table 5.3. Performance of Classical Machine Learning Models

Model Type	**F1 Score**	**Balanced Accuracy**
Logistic Regression	68.12%	75.22%
Multilayer Perceptron	62.64%	71.96%
Linear SVM	62.36%	70.64%
XGBoost	61.47%	70.96%
RBF SVM	57.76%	69.35%
AdaBoost	55.28%	67.18%
Decision Tree	54.82%	64.73%

A.2 Back-translation Data Augmentation Model Performance

Table 5.4. Performance of top 5 language combinations.

Languages	F1 Score	Balanced Acc
Turkish + Danish	79.10%	83.64%
Bengali + Turkish + Finnish + Spanish	78.55%	83.38%
Bengali + Uralic + Finnish + Portuguese	78.24%	83.11%
Bengali + Turkish + Dutch + Estonian	78.18%	83.09%
Spanish	78.15%	82.97%

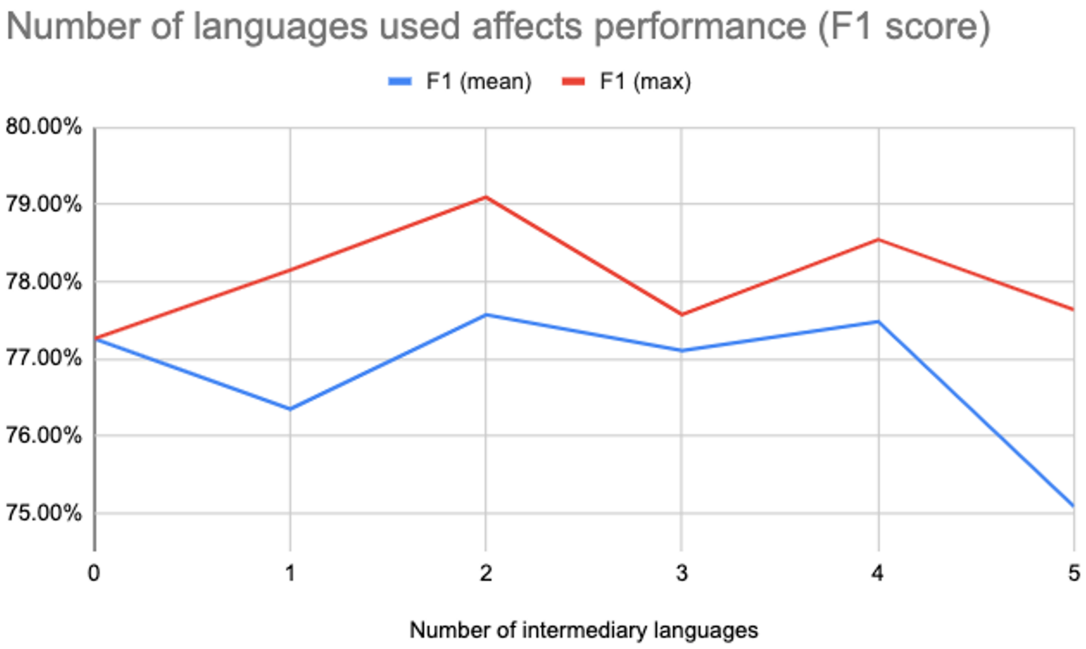

Figure 5.7. Effect of number of intermediary languages on model performance in back-translation data augmentation. Models performed best with two intermediary languages and worse with five or more languages.

A.3 Final Hyperparameters

Table 5.5. Final hyperparameter values after tuning.

Parameter	Value
Learning rate	2.528e-5
Random seed	33
Batch size	8
Epochs	1

Chapter 6

Conclusion and Outlook

Throughout this work, multiple scales have been investigated with scaling and bridging these scales being of primary concern and probabilistic graph models presenting as particularly useful constructs in this regard. **Chapter 1** provided an extended introduction to the fields of neuropsychiatry, artificial intelligence, and probabilistic graph models. **Chapters 2** and **3** focused on computationally efficient reaction-diffusion models of synaptic transmission using Markov models and spectral methods. **Chapter 4** used deep learning, graph models, dimensionality reduction, and machine learning techniques to create an automated neuroimaging pipeline for biomarker detection and automatic diagnosis of major depressive disorder. **Chapter 5** used transformers for natural language processing as a scalable diagnosis system for anxiety and depression.

There is a clear need for research into mental disorders due to their pervasive nature and dire consequences. As such, the benefits of this work are multifaceted. It builds multiscale tools allowing for movement between multiple scales of the brain that emphasizes computational efficiency in doing so with the goal of bidirectional movement from biology to behavior. Furthermore, emphasis is placed on the development of quantitative and objective measures to aid in the diagnosis of such disorders. Tools from the burgeoning field of artificial intelligence and engineering seek to elucidate insights from molecules to minds.

The future of this work and this field is exciting. Envision a world where a patient

feels off and goes to the neuropsychiatry office. The patient may first speak with a nurse to tell them the issues they have be facing, and this story is automatically converted to text and fed into a system train, on millions of prior psychotherapy stories. From that, the patient is flagged as requiring further intervention. It is then possible to take a brain scan and build a 3D reconstruction of the patient's brain with full granularity, from molecules to systems level. Using mass collections of similar scans, a new set of disorders has emerged with clear, quantitative, and objective biomarkers. The doctor can then use the model to pinpoint a disorder with a clear diagnosis and problem area. After which, it is possible to run simulations in the model to assess the effects of various pharmacological interventions to determine the best fit for the patient, specifically. This is the future of mental disorder diagnostics and treatment.

Bibliography

[1] The discovery of alzheimer's disease. 5(1).

[2] Shock therapy. Technical Report 1, Group for the Advancement of Psychiatry, sept 1947.

[3] Research on prefrontal lobotomy. Technical Report 6, Group for the Advancement of Psychiatry, june 1948.

[4] The year in review: Annual report smith, kline, french laboratories. Technical report, Smith, Kline, French Laboratories, Philadelphia, Pennsylvania, 1962.

[5] Preventing suicide: a global imperative. Technical report, World Health Organization, Geneva, 2014. WHO-124.

[6] Fda approves stroke-detecting ai software. *Nature Biotechnology*, 36(290), 2018. dias-11.

[7] Management of physical health condtions in adults with severe mental disorders: Who guidelines. Technical report, World Health Organization, Geneva, 2018. WHO-114.

[8] Gbd results tool. http://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/cb9c37d9454c80df77adaed394d7fc0f, 2019. WHO-96.

[9] Gbd results tool. http://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/5066348dc958b095cb6ceb4bfd9c3e07, 2019. WHO-99.

[10] Global health estimates 2019: Deaths by cause, age, sex, by country and by region, 2000-2019. Technical report, World Health Organization, Geneva, 2020. WHO-123, 126, 128, 129.

[11] Mental health of people with neglected tropical diseases: Towards a person-centered approach. Technical report, World Health Organization, Geneva, 2020. WHO-168.

[12] Mental health atlas 2020. Technical report, World Health Organization, Geneva, 2021. WHO-5.

[13] Suicide worldwide in 2019. Technical report, World Health Organization, Geneva, 2021. WHO-125.

[14] Integration of mental health and hiv interventions: key considerations. Technical report, UNAIDS and World Health Organization, Geneva, 2022. WHO-166.

[15] World mental health report: transforming mental health for all. Technical report, World Health Organization, Geneva, 2022. WHO.

[16] World mental health report: transforming mental health for all. Technical report, World Health Organization, Geneva, 2022.

[17] Adolescent Brain Cognitive Development℠ Study (ABCD Study®), June 2023.

[18] Former apa presidents. https://www.apa.org/about/governance/president/former-presidents, 2023. Accessed: 2023-06-08.

[19] Mental health and substance use state fact sheets. Technical report, Kaiser Family Foundation, San Fransisco, CA, March 2023.

[20] The state of mental health in america. Technical report, Mental Health America, Alexandria, VA, 2023.

[21] Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the united states: Results from the 2021 national survey on drug use and health. Technical report, Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, 2022. SAMHSA.

[22] T. M. Achenbach. *The Achenbach System of Empirically Based Assessemnt (ASEBA): Development, Findings, Theory, and Applications*. University of Vermont Research Center for Children, Youth, Families, Burlington, VT, 2009.

[23] Muhammad Shamim Ahmed, Daisy Kornblum, Dominic Oliver, Paolo Fusar-Poli, and Rashmi Patel. Associations of remote mental healthcare with clinical outcomes: a natural language processing enriched electronic health record data study protocol. *BMJ open*, 13(2), 2023.

[24] Allyssa J. Allen, Mélina E. Griss, Bradley S. Folley, Keith A. Hawkins, , and Godfrey D. Pearlson. Endophenotypes in schizophrenia: A selective review. *Schizophrenia Research*, 109(1-3):24–37, 2010.

[25] Amit Anand, Yu Li, Yang Wang, Jingwei Wu, Sujuan Gao, Lubna Bukhari, Vincent P. Mathews, Andrew Kalnin, and Mark J. Lowe. Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study. *Biological Psychiatry*, 57(10):1079–1088, May 2005.

[26] Pauline Anderson. Direct-to-consumer ads boost psychiatric drug use. *Medscape Medical News*, September 2016.

[27] Alan Anticevic, Mark Gancsos, John D Murray, Grega Repovs, Naomi R Driesen, Debra J Ennis, Mark J Niciu, Peter T Morgan, Toral S Surti, Michael H Bloch, Ramachandran Ramani, Mark A Smith, Xiao-Jing Wang, John H Krystal, and Philip R Corlett. Nmda receptor function in large-scale anticorrelated neural systems with implications for cognition and schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41):16720–5, 2012.

[28] Clay M. Armstrong. Interaction of tetraethylammonium ion derivatives with the potassium channels of giant axons. *Journal of General Physiology*, 58(4):413–37, 1971.

[29] Francesco Barbieri, Jose Camacho-Collados, Leonardo Neves, and Luis Espinosa-Anke. Tweeteval: Unified benchmark and comparative evaluation for tweet classification, 2020.

[30] Thomas M. Bartol, Daniel X. Keller, Justin P. Kinney, Chandrajit L. Bajaj, Kristen M. Harris, Terrence J. Sejnowski, and Mary B. Kennedy. Computational reconstitution of spine calcium transients from individual proteins. *Frontiers in Synaptic Neuroscience*, 7(17), 2015.

[31] Chiara Bartolozzi and Giacomo Indiveri. Synaptic dynamics in analog VLSI. *Neural Computation*, 19(10):2581–2603, 2007.

[32] Michael Bauer, Emanuel Severus, Stephan Köhler, Peter C Whybrow, Jules Angst, Hans-Jürgen Möller, and Wfsbp Task Force on Treatment Guidelines for Unipolar Depressive Disorders. World federation of societies of biological psychiatry (wfsbp) guidelines for biological treatment of unipolar depressive disorders. part 2: maintenance treatment of major depressive disorder. *The World Journal of Biological Psychiatry*, 16, 2015.

[33] William Bentinck-Smith and Elizabeth Bentinck-Smith. *Harvard University History of Named Chairs: Sketches of Donors and Donations: Professorships of the Faculties of Medicine and Public Health*. Harvard Univerity, Cambridge, Mass, 1993.

[34] James Bergstra, Dan Yamins, and David D Cox. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Python in Science*, SciPy, Austin, TX, 2013. Springer, Cham.

[35] Kalyan B Bhattacharyya. Johann bernhard aloys von gudden and the mad king of bavaria. *Annals of Indian Academy of Neurology*, 20(4):348–351, 2017.

[36] Benjamin Billot, Colin Magdamo, You Cheng, Steven E. Arnold, Sudeshna Das, and Juan Eugenio Iglesias. Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain MRI datasets. *Proceedings of the National Academy of Sciences*, 120(9):e2216399120, February 2023. Publisher: Proceedings of the National Academy of Sciences.

[37] Josef Bischofberger, Jörg R. P. Geiger, and Peter Jonas. Timing and efficacy of Ca^{2+} channel activation in hippocampal mossy fiber boutons. *The Journal of Neuroscience*, 22(24), 2002.

[38] D.E. Bloom, E.T. Cafiero, E Jané-Llopis, S Abrahams-Gessel, L.R Bloom, S Fathima, A.B Feigl, T Gaziano, M Mowafi, A Pandya, K Prettner, L Rosenberg, B Seligman, A.Z Stein, and C Weinstein. The global economic burden of noncommunicable diseases. Technical report, World Economic Forum, Geneva, 2011. WHO-131.

[39] Irem Boybat, Manuel Le Gallo, S. R. Nandakumar, Timoleon Moraitis, Thomas Parnell, Tomas Tuma, Bipin Rajendran, Yusuf Leblebici, Abu Sebastian, and Evangelos Eleftheriou. Neuromorphic computing with multi-memristive synapses. *Nature Communications*, 9(2514), 2018.

[40] J. Douglas Bremner, Meena Vythilingam, Eric Vermetten, Ahsan Nazeer, Jahangir Adil, Sarfraz Khan, Lawrence H. Staib, and Dennis S. Charney. Reduced volume of orbitofrontal cortex in major depression. *Biological Psychiatry*, 51(4):273–279, February 2002.

[41] Paul Broca. Du siège de la faculté du langage articulé. *"Bulletins de la Société d'Anthropologie"*, 6:377– 393, 1865.

[42] Frédéric D Broccard, Siddharth Joshi, Jun Wang, and Gert Cauwenberghs. Neuromorphic neural interfaces: from neurophysiological inspiration to biohybrid coupling with nervous systems. *Journal of Neural Engineering*, 14(4), 2017.

[43] Arthur E. Bryson. A gradient method for optimizing multi-stage allocation processes. In *Proc. of a Harvard Symposium on Digital Computers and Their Applications*, 1962.

[44] Franziska Burger, Mark A Neerincx, and Willem-Paul Brinkman. Natural language processing for cognitive therapy: Extracting schemas from thought records. *PloS one*, 16(10), 2021.

[45] George H W Bush. Presidential proclamation 6158: Project on the decade of the brain. pages 729–37, July 1990.

[46] B. J. Casey, Tariq Cannonier, May I. Conley, Alexandra O. Cohen, Deanna M. Barch, Mary M. Heitzeg, Mary E. Soules, Theresa Teslovich, Danielle V. Dellarco, Hugh Garavan, Catherine A. Orr, Tor D. Wager, Marie T. Banich, Nicole K. Speer, Matthew T. Sutherland, Michael C. Riedel, Anthony S. Dick, James M. Bjork, Kathleen M. Thomas, Bader Chaarani, Margie H. Mejia, Donald J. Hagler, M. Daniela Cornejo, Chelsea S. Sicat, Michael P. Harms, Nico U. F. Dosenbach, Monica Rosenberg, Eric Earl, Hauke Bartsch, Richard Watts, Jonathan R. Polimeni, Joshua M. Kuperman, Damien A. Fair, and Anders M. Dale. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, 32:43–54, August 2018.

[47] J Castillo and B Katz. Quantal components of the end-plate potential. *The Journal of Physiology*, 124(3):560–573, 1954.

[48] Gert Cauwenberghs. Reverse engineering the cognitive brain. *PNAS*, 110(39):15512–15513, 2013.

[49] S. Chakrabartty and G. Cauwenberghs. Margin normalization and propagation in analog VLSI. In *2004 IEEE International Symposium on Circuits and Systems (ISCAS)*, volume 1, pages I–901, 2004.

[50] Shantanu Chakrabartty and Gert Cauwenberghs. Sub-microwatt analog VLSI support vector machine for pattern classification and sequence estimation. In L. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 17, Cambridge, MA, 2005. MIT Press.

[51] Shu-Sen Chang, David Stuckler, Paul Yip, and David Gunnell. Impact of 2008 global economic crisis on suicide: time trend study in 54 countries. *British Medical Journal*, 347, 2013. WHO-57.

[52] Fiona J Charlson, Amanda J Baxter, Tarun Dua, Louisa Degenhardt, Harvey A Whiteford, Theo Vos Vikram Patel, Dan Chisholm, Tarun Dua, Ramanan Laxminarayan, and María Elena Medina-Mora. Excess mortality from mental, neurological, and substance use disorders in the global burden of disease study 2010. *Epdemiol Psychiatry Sci*, 24(2):121–140, 2014. WHO-121.

[53] Yusi Chen, Qasim Bukhari, Tiger W. Lin, and Terrence J. Sejnowski. Functional connectivity of fMRI using differential covariance predicts structural connectivity and behavioral reaction times. 6(2):614–633.

[54] Edward Chesney, Guy M Goodwin, and Seena Fazel. Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry*, 13(2):153–160, 2014. WHO-115.

[55] Dan Chisholm, Kim Sweeny, Peter Sheehan, Bruce Rasmussen, Filip Smit, Pim Cuijpers, and Shekhar Saxena. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry*, 3(5):415–424, 2016. WHO-217.

[56] Carson C. Chow and John A. White. Spontaneous action potentials due to channel fluctuation. *Biophysical Journal*, 71:3013–3021, 1996.

[57] PJ Church and EF Stanley. Single L-type calcium channel conductance with physiological levels of calcium in chick ciliary ganglion neurons. *The Journal of Physiology*, 496(1), 1996.

[58] Patricia Smith Churchland and Terrence J. Sejnowski. *The Computational Brain*. The MIT Press, 1992.

[59] C. E. Coffey, W. E. Wilkinson, R. D. Weiner, I. A. Parashos, W. T. Djang, M. C. Webb, G. S. Figiel, and C. E. Spritzer. Quantitative cerebral anatomy in depression. A controlled magnetic resonance imaging study. *Archives of General Psychiatry*, 50(1):7–16, January 1993.

[60] Victor Cohn. Charting 'the soul's frail dwelling-house': Brain waves: Scientists stalk the lodging of the mind. *Washington Post*, Sept 1982.

[61] COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 202 due to the covid-19 pandemic. *Lancet*, S0140-6736(21):02143–7, 2021. WHO-68.

[62] D. Colquhoun. *The relation between classical and cooperative models for drug action*. Drug receptors. Macmillan Press, London, 1973.

[63] Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. Unsupervised cross-lingual representation learning at scale, 2020.

[64] Lisong Dai, Hongmei Zhou, Xiangyang Xu, and Zhentao Zuo. Brain structural and functional changes in patients with major depressive disorder: a literature review. *PeerJ*, 7, 2019.

[65] Mike Davies, Narayan Srinivasa, Tsung-Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, Shweta Jain, Yuyun Liao, Chit-Kwan Lin, Andrew Lines, Ruokun Liu, Deepak Mathaikutty, Steven McCoy, Arnab Paul, Jonathan Tse, Guruguhanathan Venkataramanan, Yi-Hsin Weng, Andreas Wild, Yoonseok Yang, and Hong Wang. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 38(1):82–99, 2018.

[66] Matthew Day. Bavaria's 'mad king' ludwig may not have been so mad after all, February.

[67] Charles DeBattista, Gustavo Kinrys, Daniel Hoffman, Corey Goldstein, John Zajecka, James Kocsis, Martin Teicher, Steven Potkin, Adrian Preda, Gurmeet Multani, Len Brandt, Mark Schiller, Dan Iosifescu, and Maurizio Fava. The use of referenced-eeg (reeg) in assisting medication selection for the treatment of depression. *Journal of psychiatric research*, 45(1):64–75, 2011.

[68] Hannah S Decker. How kraepelinian was kraepelin? how kraepelinian are the neo-kraepelinians?–from emil kraepelin to dsm-iii. *History of psychiatry*, 18(17), Sept 2007.

[69] Robert J DeRubeis, Zachary D Cohen, Nicholas R Forand, Jay C Fournier, Lois A Gelfand, and Lorenzo Lorenzo-Luaces. The personalized advantage index: translating research on prediction into individualized treatment recommendations. a demonstration. *PloS one*, 9(1), 2014.

[70] Alain Destexhe and Michelle Rudolph-Lilith. *Neuronal Noise*. Springer, New York, NY, 2012.

[71] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, pages 4171–4186, Minneapolis, MN, 2019. ACL.

[72] Joseph DiGuiseppi and Prasanna Tadi. Neuroanatomy, Postcentral Gyrus. In *StatPearls*. StatPearls Publishing, Treasure Island (FL), 2023.

[73] Gerhard-Paul Diller, Aleksander Kempny, Sonya V Babu-Narayan, Marthe Henrichs, Margarita Brida, Anselm Uebing, Astrid E Lammers, Helmut Baumgartner, Wei Li, Stephen J Wort, Konstantinos Dimopoulos, and Michael A Gatzoulis. Machine learning

algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. *European heart journal*, 40(13):1069–1077, 2019. dias-47.

[74] W. C. Drevets. Neuroimaging and neuropathological studies of depression: implications for the cognitive-emotional features of mood disorders. *Current Opinion in Neurobiology*, 11(2):240–249, April 2001.

[75] W. C. Drevets, J. L. Price, J. R. Simpson, R. D. Todd, T. Reich, M. Vannier, and M. E. Raichle. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature*, 386(6627):824–827, April 1997.

[76] W. C. Drevets, T. O. Videen, J. L. Price, S. H. Preskorn, S. T. Carmichael, and M. E. Raichle. A functional anatomical study of unipolar depression. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 12(9):3628–3641, September 1992.

[77] Yuhui Du, Zening Fu, and Vince D. Calhoun. Classification and Prediction of Brain Disorders Using Functional Connectivity: Promising but Challenging. *Frontiers in Neuroscience*, 12, 2018.

[78] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, 2017. dias-15.

[79] National Collaborating Centre for Mental Health (UK). Depression: The treatment and management of depression in adults (updated edition). *British Psychological Society*, 2010.

[80] Kathleen C Fraser, Jed A Meltzer, and Frank Rudzicz. Linguistic features identify alzheimer's disease in narrative speech. *Journal of Alzheimer's disease*, 49(2):407–422, 2018. dias-3.

[81] Karl Friston. Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annual Review of Neuroscience*, 25:221–250, 2002.

[82] Karl J Friston, Klaas Enno Stephan, Read Montague, and Raymond J Dolan. Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry*, 1(2):148–158, July 2014.

[83] Cynthia H. Y. Fu, Steven C. R. Williams, Anthony J. Cleare, Michael J. Brammer, Nicholas D. Walsh, Jieun Kim, Chris M. Andrew, Emilio Merlo Pich, Pauline M. Williams, Laurence J. Reed, Martina T. Mitterschiffthaler, John Suckling, and Edward T. Bullmore. Attenuation of the neural response to sad faces in major depression by antidepressant treatment: a prospective, event-related functional magnetic resonance imaging study. *Archives of General Psychiatry*, 61(9):877–889, September 2004.

[84] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[85] Aziliz Le Glaz, Yannis Haralambous, Deok-Hee Kim-Dufor, Philippe Lenca, Romain Billot, Taylor C Ryan, Jonathan Marsh, Jordan DeVylder, Michel Walter, Sofian Berrouiguet, and Christophe Lemey. Machine learning and natural language processing in mental health: Systematic review. *Journal of medical Internet research*, 23(5), 2021.

[86] Aaron Gokaslan and Vanya Cohen. Openwebtext corpus.

[87] Franz Gravenhorst, Amir Muaremi, Jakob Bardram, Agnes Grünerbl, Oscar Mayora, Gabriel Wurzer, Mads Frost, Venet Osmani, Bert Arnrich, and Paul Lukowicz Gerhard Tröster. Mobile phones as medical devices in mental disorder treatment: an overview. *Personal and Ubiquitous Computing*, 19:335–353, 2015.

[88] Michael Greicius. Resting-state functional connectivity in neuropsychiatric disorders. *Current Opinion in Neurology*, 21(4):424–430, August 2008.

[89] Michael D. Greicius, Benjamin H. Flores, Vinod Menon, Gary H. Glover, Hugh B. Solvason, Heather Kenna, Allan L. Reiss, and Alan F. Schatzberg. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, 62(5):429–437, September 2007.

[90] Katja Guenther. Recasting neuropsychiatry: Freud's 'critical introduction' and the convergence of french and german brain science. *Psychoanalysis and History*, 14(2):203–226, 2012.

[91] Donald J. Hagler, SeanN Hatton, M. Daniela Cornejo, Carolina Makowski, Damien A. Fair, Anthony Steven Dick, Matthew T. Sutherland, B. J. Casey, Deanna M. Barch, Michael P. Harms, Richard Watts, James M. Bjork, Hugh P. Garavan, Laura Hilmer, Christopher J. Pung, Chelsea S. Sicat, Joshua Kuperman, Hauke Bartsch, Feng Xue, Mary M. Heitzeg, Angela R. Laird, Thanh T. Trinh, Raul Gonzalez, Susan F. Tapert, Michael C. Riedel, Lindsay M. Squeglia, Luke W. Hyde, Monica D. Rosenberg, Eric A. Earl, Katia D. Howlett, Fiona C. Baker, Mary Soules, Jazmin Diaz, Octavio Ruiz de Leon, Wesley K. Thompson, Michael C. Neale, Megan Herting, Elizabeth R. Sowell, Ruben P. Alvarez, Samuel W. Hawes, Mariana Sanchez, Jerzy Bodurka, Florence J. Breslin, Amanda Sheffield Morris, Martin P. Paulus, W. Kyle Simmons, Jonathan R. Polimeni, Andre van der Kouwe, Andrew S. Nencka, Kevin M. Gray, Carlo Pierpaoli, John A. Matochik, Antonio Noronha, Will M. Aklin, Kevin Conway, Meyer Glantz, Elizabeth Hoffman, Roger Little, Marsha Lopez, Vani Pariyadath, Susan Rb Weiss, Dana L. Wolff-Hughes, Rebecca DelCarmen-Wiggins, Sarah W. Feldstein Ewing, Oscar Miranda-Dominguez, Bonnie J. Nagel, Anders J. Perrone, Darrick T. Sturgeon, Aimee Goldstone, Adolf Pfefferbaum, Kilian M. Pohl, Devin Prouty, Kristina Uban, Susan Y. Bookheimer, Mirella Dapretto, Adriana Galvan, Kara Bagot, Jay Giedd, M. Alejandra Infante, Joanna Jacobus, Kevin Patrick, Paul D. Shilling, Rahul Desikan, Yi Li, Leo Sugrue, Marie T. Banich, Naomi Friedman, John K. Hewitt, Christian Hopfer, Joseph Sakai, Jody Tanabe, Linda B. Cottler, Sara Jo Nixon,

Linda Chang, Christine Cloak, Thomas Ernst, Gloria Reeves, David N. Kennedy, Steve Heeringa, Scott Peltier, John Schulenberg, Chandra Sripada, Robert A. Zucker, William G. Iacono, Monica Luciana, Finnegan J. Calabro, Duncan B. Clark, David A. Lewis, Beatriz Luna, Claudiu Schirda, Tufikameni Brima, John J. Foxe, Edward G. Freedman, Daniel W. Mruzek, Michael J. Mason, Rebekah Huber, Erin McGlade, Andrew Prescot, Perry F. Renshaw, Deborah A. Yurgelun-Todd, Nicholas A. Allgaier, Julie A. Dumas, Masha Ivanova, Alexandra Potter, Paul Florsheim, Christine Larson, Krista Lisdahl, Michael E. Charness, Bernard Fuemmeler, John M. Hettema, Hermine H. Maes, Joel Steinberg, Andrey P. Anokhin, Paul Glaser, Andrew C. Heath, Pamela A. Madden, Arielle Baskin-Sommers, R. Todd Constable, Steven J. Grant, Gayathri J. Dowling, Sandra A. Brown, Terry L. Jernigan, and Anders M. Dale. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage*, 202:116091, November 2019.

[92] Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H. Tison, Codie Bourn, Mintu P. Turakhia, and Andrew Y. Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25:65–69, 2019. dias-22.

[93] Anne Harrington. *Mind Fixers: psychiatry's troubled search for the biology of mental illness*. W. W. Norton Company, New York, NY, 2019.

[94] Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. Deberta: Decoding-enhanced bert with disentangled attention, 2021.

[95] Center for Devices and Radiological Health. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. *FDA*, October 2022. Publisher: FDA.

[96] Donald O. Hebb. *The Organization of Behavior*. Wiley, 1949.

[97] RW Heinrichs. Meta-analysis and the science of schizophrenia: variant evidence or evidence of variants? *Neurosci Biobehav Rev*, 28:379–394, 2004.

[98] Oliver Higgins, Brooke L Short, Stephan K Chalup, and Rhonda L Wilson. Machine learning and natural language processing in mental health: Systematic review. *Journal of medical Internet research*, 23(5), 2021.

[99] Alan Hodgkin and Andrew Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, 1952.

[100] Zhenghua Hou, Youyong Kong, Yingying Yin, Yuqun Zhang, and Yonggui Yuan. Identification of first-episode unmedicated major depressive disorder using pretreatment features of dominant coactivation patterns. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 104:110038, January 2021.

[101] Rainbo Hultman, Kyle Ulrich, Benjamin D. Sachs, Cameron Blount, David E. Carlson, Nkemdilim Ndubuizu, Rosemary C. Bagot, Eric M. Parise, Mai-Anh T. Vu, Neil M. Gallagher, Joyce Wang, Alcino J. Silva, Karl Deisseroth, Stephen D. Mague, Marc G. Caron,

Eric J. Nestler, Lawrence Carin, and Kafui Dzirasa. Brain-wide Electrical Spatiotemporal Dynamics Encode Depression Vulnerability. *Cell*, 173(1):166–180.e14, March 2018.

[102] Quentin J. M. Huys, Tiago V. Maia, and Michael J. Frank. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3):404–413, March 2016. Number: 3 Publisher: Nature Publishing Group.

[103] Steven E. Hyman. The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology*, 6:155–79, 2010.

[104] Steven E. Hyman. Revolution stalled. *Science Translational Medicine*, 4(155), October 2012.

[105] Steven E. Hyman. Psychiatric drug development: Diagnosing a crisis. *Cerebrum: Dana Forum on Brain Science*, April 2013.

[106] Giacomo Indiveri, Bernabé Linares-Barranco, Tara Julia Hamilton, André van Schaik, Ralph Etienne-Cummings, Tobi Delbruck, Shih-Chii Liu, Piotr Dudek, Philipp Häfliger, Sylvie Renaud, Johannes Schemmel, Gert Cauwenberghs, John Arthur, Kai Hynna, Fopefolu Folowosele, Sylvain Saighi, Teresa Serrano-Gotarredona, Jayawan Wijekoon, Yingxue Wang, and Kwabena Boahen. Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience*, 5(73), 2011.

[107] Thomas Insel. Understanding mental disorders as circuit disorders. *National Institute of Mental Health*, 2010.

[108] Thomas Insel. Transforming diagnosis. *NIMH Director's BlogPosts from 2013*, April 2013.

[109] Thomas R. Insel and Story C. Landis. Twenty-five Years of Progress: The View from NIMH and NINDS. *Neuron*, 80(3):10.1016/j.neuron.2013.09.041, October 2013.

[110] Joshua Jacobs, Michael J. Kahana, Arne D. Ekstrom, Matthew V. Mollison, and Itzhak Fried. A sense of direction in human entorhinal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 107(14):6487–6492, April 2010.

[111] Dean Edmonds Jr. Recollections of my wartime and university experiences. https://ethw.org/First-Hand:Recollections_of_my_Wartime_and_University_Experiences_in_Nuclear_Physics. Accessed: 2021-05-12.

[112] Roselinde H. Kaiser, Jessica R. Andrews-Hanna, Tor D. Wager, and Diego A. Pizzagalli. Large-Scale Network Dysfunction in Major Depressive Disorder: A Meta-analysis of Resting-State Functional Connectivity. *JAMA psychiatry*, 72(6):603–611, June 2015.

[113] Mitchel Kappen, Marie-Anne Vanderhasselt, and George M Slavich. Speech as a promising biosignal in precision psychiatry. *Neuroscience and biobehavioral reviews*, 148, 2023.

[114] S Kapur, AG Phillips, and TR Insel. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17:1174–1179, 2012.

[115] Henry J. Kelley. Gradient theory of optimal flight paths. *ARS Journal*, 30:947–954, 1960.

[116] John F. Kennedy. Mental illness and mental retardation: Message from the president of the united states relative to mental illness and mental retardation. *American Psychologist*, 18(6):280, 1963.

[117] John F. Kennedy. Message from the president of the united states relative to mental illness and mental retardation. *American Journal of Psychiatry*, 120(8):729–37, 1964.

[118] Tim A. Kimbrell, Terence A. Ketter, Mark S. George, John T. Little, Brenda E. Benson, Mark W. Willis, Peter Herscovitch, and Robert M. Post. Regional cerebral glucose utilization in patients with a range of severities of unipolar depression. *Biological Psychiatry*, 51(3):237–252, February 2002.

[119] Christof Koch. *Biophysics of Computation*. Oxford University Press, New York, New York, 1999.

[120] Brandon A Kohrt, Andrew Rasmussen, Bonnie N Kaiser, Emily E Haroz, Sujen M Maharjan, Byamah B Mutamba, Joop T V M de Jong, and Devon E Hinton. Cultural concepts of distress and psychiatric disorders: literature revew and research recommendations for global mental health epidemiology. *International Journal of Epidemiology*, 43(2):365–406, 2014. WHO-111.

[121] H Korn and DS Faber. Quantal analysis and synaptic efficacy in the CNS. *Trends in Neuroscience*, 14(10):439–45, 1991.

[122] Maki S. Koyama, David O'Connor, Zarrar Shehzad, and Michael P. Milham. Differential contributions of the middle frontal gyrus functional connectivity to literacy and numeracy. *Scientific Reports*, 7:17548, December 2017.

[123] A. Kumar, W. Bilker, Z. Jin, and J. Udupa. Atrophy and high intensity lesions: complementary neurobiological mechanisms in late-life major depression. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 22(3):264–274, March 2000.

[124] Veena Kumari, Martina T. Mitterschiffthaler, John D. Teasdale, Gin S. Malhi, Richard G. Brown, Vincent Giampietro, Michael J. Brammer, Lucia Poon, Andrew Simmons, Steven C. R. Williams, Stuart A. Checkley, and Tonmoy Sharma. Neural abnormalities during cognitive generation of affect in treatment-resistant depression. *Biological Psychiatry*, 54(8):777–791, October 2003.

[125] Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, and Radu Soricut. Albert: A lite bert for self-supervised learning of language representations, 2020.

[126] Caroline Lappetito. Ims report 11.5% dollar growth in '03 us prescription sales. Technical report, IMS Health, February 2004.

[127] E P V Le, Y Wang, Y Huang, S Hickman, and F J Gilbert. Artificial intelligence in breast imaging. *Clinical radiology*, 74(5):357–366, 2019. dias-9.

[128] Lisha Li, Kevin Jamieson, Afshin Rostamizadeh, Katya Gonina, Moritz Hardt, Benjamin Recht, and Ameet Talwalkar. Massively parallel hyperparameter tuning, 2018.

[129] Zhichao Li, Jilin Huang, and Zhiping Hu. Screening and diagnosis of chronic pharyngitis based on deep learning. *International journal of environmental research and public health*, 16(10):1688, 2019. dias-35.

[130] J. Lighthill. Artificial intelligence: A general survey. In J. Lighthill, N. S. Sutherland, R. M. Needham, H. C. Longuet-Higgins, and D. Michie, editors, *Artificial Intelligence: A Paper Symposium*. Science Research Council of Great Britain, 1973.

[131] Xiaoxuan Liu, Livia Faes, Aditya U Kale, Siegfried K Wagner, Dun Jack Fu, Alice Bruynseels, Thushika Mahendiran, Gabriella Moraes, Mohith Shamdas, Christoph Kern, Joseph R Ledsam, Martin Schmid, Konstantinos Balaskas, Eric Topol, Lucas M Bachmann, Pearse A Keane, and Alastair K Denniston. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health*, 2019.

[132] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Ro{bert}a: A robustly optimized {bert} pretraining approach. In *International Conference on Learning Representations*, ICLR. ICLR, 2020.

[133] Yun Liu, Timo Kohlberger, Mohammad Norouzi, George Dahl, Jenny L Smith, Arash Mohtashamian, Niels Olson, Lily Peng, Jason D Hipp, and Martin C Stumpe. Artificial intelligence-based breast cancer nodal metastasis detection: Insights into the black box for pathologists. *Archives of Pathology & Laboratory Medicine*, 143:859–868, 2018.

[134] P T Loosen, J C Garbutt, and A J Prange. Evaluation of the diagnostic utility of the trhinduced tsh response in psychiatric disorders. *Pharmacopsychiatry*, 20(3):90–5, 1987.

[135] Valentina Lorenzetti, Nicholas B. Allen, Alex Fornito, and Murat Yücel. Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. *Journal of Affective Disorders*, 117(1-2):1–17, September 2009.

[136] Crick Lund, Carrie Brooke-Sumner, Florence Baingana, Emily Claire Baron, Erica Breuer, Prabha Chandra, Johannes Haushofer, Helen Herrman, Mark Jordans, Christian Kieling, Maria Elena Medina-Mora, Ellen Morgan, Olayinka Omigbodun, Wietse Tol, Vikram Patel, and Shekhar Saxena. Social determinants of mental disorders and the sustainable development goals: a systematic review of reviews. *Lancet Psychiatry*, 50(4):357–369, 2018. WHO-52.

[137] Joel Mackenzie, Rodger Benham, Matthias Petri, Johanne R. Trippas, J Shane Culpepper, and Alistair Moffat. Cc-news-en: A large english news corpus, 2020.

[138] Tiago V Maia and Maria Cano-Colino. The role of serotonin in orbitofrontal function and obsessive-compulsive disorder. *Clinical Psychological Science*, 3(3):460–482, 2015.

[139] Arjun Majumdar, Laura Brattain, Brian Telfer, Chad Farris, and Jonathan Scalera. Detecting intracranial hemorrhage with deep learning. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 583–587. IEEE Engineering in Medicine and Biology Society, 2018. dias-10.

[140] Carolina Makowski, Sophie Béland, Penelope Kostopoulos, Nikhil Bhagwat, Gabriel A. Devenyi, Ashok K. Malla, Ridha Joober, Martin Lepage, and M. Mallar Chakravarty. Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: Comparing automated approaches to manual delineation. *NeuroImage*, 170:182–198, April 2018.

[141] Gerardo Malagon, Takafumi Miki, Isabel Llano, Erwin Neher, and Alain Marty. Counting vesicular release events reveals binomial release statistics at single glutamatergic synapses. *The Journal of Neuroscience*, 36(14):4010–4025, 2016.

[142] Charles R Marmar, Adam D Brown, Meng Qian, Eugene Laska, Carole Siegel, Meng Li, Duna Abu-Amara, Andreas Tsiartas, Colleen Richey, Jennifer Smith, Bruce Knoth, and Dimitra Vergyri 4. Speech-based markers for posttraumatic stress disorder in us veterans. *Depression and anxiety*, 36(7):607–616, 2019. dias-38.

[143] Joseph B. Martin. The Integration of Neurology, Psychiatry, and Neuroscience in the 21st Century. *American Journal of Psychiatry*, 159(5):695–704, May 2002. Publisher: American Psychiatric Publishing.

[144] Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault. The ai index 2023 annual report. Technical report, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, 2023.

[145] H. S. Mayberg, S. K. Brannan, J. L. Tekell, J. A. Silva, R. K. Mahurin, S. McGinnis, and P. A. Jerabek. Regional metabolic effects of fluoxetine in major depression: serial changes and relationship to clinical response. *Biological Psychiatry*, 48(8):830–843, October 2000.

[146] H. S. Mayberg, M. Liotti, S. K. Brannan, S. McGinnis, R. K. Mahurin, P. A. Jerabek, J. A. Silva, J. L. Tekell, C. C. Martin, J. L. Lancaster, and P. T. Fox. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *The American Journal of Psychiatry*, 156(5):675–682, May 1999.

[147] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–137, 1943.

[148] Mark D. McDonnell, Joshua H. Goldwyn, and Benjamin Lindner. Neuronal stochastic variability: Influences on spiking dynamics and network activity. *Frontiers in Computational Neuroscience*, 10, 2016.

[149] Carver Mead. *Analog VLSI and Neural Systems*. Addison-Wesley, Boston, MA, 1989.

[150] Carver Mead. Neuromorphic electronic systems. In *Proceedings of the IEEE*, volume 78, pages 1629–1636, 1990.

[151] Jonathan Michel Metzl. page 73. Duke University Press, Durham, NC, 2003.

[152] Marvin Minsky, John McCarthy, Nathaniel Rochester, and Claude Shannon. A proposal for the dartmouth summer research project on artificial intelligence, 1955.

[153] Marvin L. Minsky and Seymour Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, 1969.

[154] Stephen A Mitchell and Adrienne Harris. What's american about american psychoanalysis. *Psychoanalytic Dialogues*, 14(2):165–191, 2004.

[155] Bahar Moezzi, Nicolangelo Iannella, and Mark D. McDonnell. Modeling the influence of short term depression in vesicle release and stochastic calcium channel gating on auditory nerve spontaneous firing statistics. *Frontiers in Computational Neuroscience*, 8, 2014.

[156] P. Read Montague, Raymond J. Dolan, Karl J. Friston, and Peter Dayan. Computational psychiatry. *Trends in Cognitive Sciences*, 16(1):72–80, January 2012.

[157] Peter C. Mulders, Philip F. van Eijndhoven, Aart H. Schene, Christian F. Beckmann, and Indira Tendolkar. Resting-state functional connectivity in major depressive disorder: A review. *Neuroscience and Biobehavioral Reviews*, 56:330–344, September 2015.

[158] Suhita Nadkarni, Thomas M. Bartol, Charles F. Stevens, Terrence J. Sejnowski, and Herbert Levine. Modelling vesicular release at hippocampal synapses. *PLoS Computational Biology*, 6(11), 2010.

[159] U. Balentin Nagerl, David Novo, Istvan Mody, and Julio L. Vergara. Binding kinetics of calbindin-D28k determined by flash photolysis of caged Ca^{2+}. *Biophysical Journal*, 79(6):3009–3018, 2000.

[160] The National Committee for Mental Hygiene, Bureau of Statistics, New York, NY. *Statistical Manual for the Use of Institutions for the Insane*, 1918.

[161] A. Newell and Herbert A. Simon. Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 19:113–126, 1976.

[162] NIDA. Landmark study of adolescent brain development renews for additional seven years. 2020.

[163] MR Nuwer. On the controversies about clinical use of eeg brain mapping. *Brain Topography*, 3:103–111, 1990.

[164] S Ogawa, T M Lee, A R Kay, , and D W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, 1990.

[165] The Committee on Nomenclature and Statistics of the American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, Washington, D.C., 1952.

[166] Graziella Orrù, William Pettersson-Yeo, Andre F Marquand, Giuseppe Sartori, and Andrea Mechelli. Nmda receptor function in large-scale anticorrelated neural systems with implications for cognition and schizophrenia. *Neuroscience and biobehavioral reviews*, 36(4):1140–52, 2012.

[167] Steven A. Orszag. Numerical methods for the simulation of turbulence. *Physics of Fluids*, 12(12):250–257, 1969.

[168] Ali H Palejwala, Nicholas B Dadario, Isabella M Young, Kyle O'Connor, Robert G Briggs, Andrew K Conner, Daniel L O'Donoghue, and Michael E Sughrue. Anatomy and white matter connections of the lingual gyrus and cuneus. *World neurosurgery*, 151:e426–e437, 2021.

[169] Mayur Pandya, Murat Altinay, Jr. Donald A. Malone, , and Amit Anand. Where in the brain is depression? *Current psychiatry reports*, 14(6):632–642, 2012.

[170] Mayur Pandya, Murat Altinay, Donald A. Malone, and Amit Anand. Where in the Brain Is Depression? *Current psychiatry reports*, 14(6):634–642, December 2012.

[171] Anne E. Parsons. *From Asylum to Prison: Deinstitutionalization and the Rise of Mass Incarceration After 1945*. University of North Carolina Press, Chapel Hill, NC, 2018.

[172] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.

[173] Bruno U. Pedroni, Stephen R. Deiss, Nishant Mysore, and Gert Cauwenberghs. Design principles of large-scale neuromorphic systems centered on high bandwidth memory. In *2020 International Conference on Rebooting Computing (ICRC)*, pages 90–94, Atlanta, GA, 2020.

[174] Bruno U. Pedroni, Siddharth Joshi, Stephen R. Deiss, Sadique Sheik, Georgios Detorakis, Somnath Paul, Charles Augustine, Emre O. Neftci, and Gert Cauwenberghs. Memory-efficient synaptic connectivity for spike-timing-dependent plasticity. *Frontiers in Neuroscience*, 13(357), 2019.

[175] R C. Petersen, P S. Aisen, L A. Beckett, M C. Donohue, A C. Gamst, D J. Harvey, C R. Jack, W J. Jagust, L M. Shaw, A W. Toga, J Q. Trojanowski, and M W. Weiner. Alzheimer's Disease Neuroimaging Initiative (ADNI). *Neurology*, 74(3):201–209, January 2010.

[176] Fujiwara PI. The links between tuberculosis and mental health: evidence and best practice incorporating guidance to usaid. Technical report, USAID, Washington DC, 2022. WHO-167.

[177] Joseph L. Price and Wayne C. Drevets. Neurocircuitry of Mood Disorders. *Neuropsychopharmacology*, 35(1):192–216, January 2010. Number: 1 Publisher: Nature Publishing Group.

[178] Kessler R, Berglund P, Demler O, Jin R, and Merikangas K. Lifetime prevalence and age-of-onset distribution of dsm-iv disorders in the national comorbidity survey replication. *Arch Gen Psychiatry*, 62(6):593–602, 2005. WHO-108.

[179] Alvin Rajkomar, Eyal Oren, Kai Chen, Andrew M. Dai, Nissan Hajaj, Michaela Hardt, Peter J. Liu, Xiaobing Liu, Jake Marcus, Mimi Sun, Patrik Sundberg, Hector Yee, Kun Zhang, Yi Zhang, Gerardo Flores, Gavin E. Duggan, Jamie Irvine, Quoc Le, Kurt Litsch, Alexander Mossin, Justin Tansuwan, De Wang, James Wexler, Jimbo Wilson, Dana Ludwig, Samuel L. Volchenboum, Katherine Chou, Michael Pearson, Srinivasan Madabushi, Nigam H. Shah, Atul J. Butte, Michael D. Howell, Claire Cui, Greg S. Corrado, and Jeffrey Dean. Scalable and accurate deep learning with electronic health records. *npg digital medicine*, 1(18), 2018. dias-4.

[180] Juan P. Ramirez-Mahaluf, Alexander Roxin, Helen S. Mayberg, and Albert Compte. A Computational Model of Major Depression: the Role of Glutamate Dysfunction on Cingulo-Frontal Network Dynamics. *Cerebral Cortex*, 27(1):660–679, January 2017.

[181] Darrel A. Regier, William E. Narrow, Diana E. Clarke, Helena C. Kraemer, S. Janet Kuramoto, Emily A. Kuhl, and David J. Kupfer. DSM-5 Field Trials in the United States and Canada, Part II: Test-Retest Reliability of Selected Categorical Diagnoses. *American Journal of Psychiatry*, 170(1):59–70, January 2013. Publisher: American Psychiatric Publishing.

[182] Edward H. Reynolds and James V. Kinnier Wilson. Neurology and psychiatry in Babylon. *Brain*, 137(9):2611–2619, September 2014.

[183] Edward H. Reynolds and Michael R. Trimble. Epilepsy, psychiatry, and neurology. *Epilepsia*, 50:50–55, 2009.

[184] Fabien Ringeval, Björn Schuller, Michel Valstar, NIcholas Cummins, Roddy Cowie, Leili Tavabi, Maximilian Schmitt, Sina Alisamir, Shahin Amiriparian, Eva-Maria Messner, Siyang Song, Shuo Liu, Ziping Zhao, Adria Mallol-Ragolta, Zhao Ren, Mohammad Soleymani, and Maja Pantic. Avec 2019 workshop and challenge: State-of-mind, detecting depression with ai, and cross-cultural affect recognition. In *Proceedings of the 9th*

International on Audio/Visual Emotion Challenge and Workshop, pages 3–12, 2019. dias-37.

[185] Anna Rogers, Olga Kovaleva, and Anna Rumshisky. A primer in bertology: What we know about how bert works. *Transactions of the Association for Computational Linguistics*, 8:842–866, 2020.

[186] Maria J. Rosa, Liana Portugal, Tim Hahn, Andreas J. Fallgatter, Marta I. Garrido, John Shawe-Taylor, and Janaina Mourao-Miranda. Sparse network-based models for patient classification using fMRI. *NeuroImage*, 105:493–506, January 2015.

[187] Frank Rosenblatt. The perceptron: A perceiving and recognizing automaton. Report, Project PARA, Cornell Aeronautical Laboratory, January 1957.

[188] Frank Rosenblatt. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan, 1962.

[189] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.

[190] David E. Rumelhart and James L. McClelland, editors. *Parallel Distributed Processing, publisher* =. 1986.

[191] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education, Inc., Hoboken, NJ, fourth edition, 2021.

[192] Mert R Sabuncu and Ender Konukoglu. Clinical prediction from structural brain mri scans: a large-scale empirical study. *Neuroinformatics*, 13(1):31–46, 2015.

[193] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, 2020.

[194] I. Schweitzer, V. Tuckwell, D. Ames, and J. O'Brien. Structural neuroimaging studies in late-life depression: a review. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry*, 2(2):83–88, April 2001.

[195] Kate M Scott, Michael Von Korff, Matthias C Angermeyer, Corina Benjet, Ronny Bruffaerts, Giovanni de Girolamo, Josep Maria Haro, Jean-Pierre Lépine, Johan Ormel, José Posada-Villa, Hisateru Tachimori, and Ronald C Kessler. Association of childhood adversities and early-onset mental disorders with adult-onset chronic physical conditions. *Arch Gen Psychiatry*, 68(8):838–844, 2011. WHO-163.

[196] Terrence J. Sejnowski. *The Deep Learning Revolution*. The MIT Press, Cambridge, MA, 2020.

[197] Mohammad S. E. Sendi, Elaheh Zendehrouh, Jing Sui, Zening Fu, Dongmei Zhi, Luxian Lv, Xiaohong Ma, Qing Ke, Xianbin Li, Chuanyue Wang, Christopher C. Abbott, Jessica A. Turner, Robyn L. Miller, and Vince D. Calhoun. Abnormal Dynamic Functional Network Connectivity Estimated from Default Mode Network Predicts Symptom Severity in Major Depressive Disorder. *Brain Connectivity*, 11(10):838–849, December 2021.

[198] Y. I. Sheline, M. H. Gado, and J. L. Price. Amygdala core nuclei volumes are decreased in recurrent major depression. *Neuroreport*, 9(9):2023–2028, June 1998.

[199] Edward Shorter. *A History of Psychiatry: From the Era of the Asylum to the Age of Prozac*. Wiley, New York, NY, 1998.

[200] Labina Shrestha, Shikha Dubey, Farrukh Olimov, Muhammad Aasim Rafique, and Moongu Jeon. 3D Convolutional with Attention for Action Recognition, June 2022. arXiv:2206.02203 [cs].

[201] Jean Siegfried and J Shulman. Deep brain stimulation. *Pacing and Clinical Electrophysioly*, 10(1):271–2, 1987. DBS.

[202] E Skaugen and L Walloe. Firing behavior in a stochastic nerve membrane model based upon the Hodgkin-Huxley equations. *Acta Physiologica Scandinavica*, 107:343–363, 1979.

[203] William R. Softky and Christof Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *The Journal of Neuroscience*, 13(1):334–350, 1993.

[204] Arno Solin and Simo Särkkä. The 10th annual mlsp competition: First place. In *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–3, 2014.

[205] Reiner Sprengelmeyer, J. Douglas Steele, Benson Mwangi, Poornima Kumar, David Christmas, Maarten Milders, and Keith Matthews. The insular cortex and the neuroanatomy of major depression. *Journal of Affective Disorders*, 133(1-2):120–127, September 2011.

[206] David F. Steiner, Robert MacDonald, Yun Liu, Peter Truszkowski, Jason D. Hipp, Christopher Gammage, Florence Thng, Lily Peng, and Martin C Stumpe. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *Am. J. Surgical Pathology*, 42:1636–1646, 2018.

[207] Chi Sun, Xipeng Qiu, Yige Xu, and Xuanjing Huang. How to fine-tune bert for text classification? In *Chinese Computational Linguistics*, CCL. Springer, Cham, 2019.

[208] Susan M. Sunkin, Lydia Ng, Chris Lau, Tim Dolbeare, Terri L. Gilbert, Carol L. Thompson, Michael Hawrylycz, and Chinh Dang. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Research*, 41(Database issue):D996–D1008, January 2013.

[209] R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44, August 1988.

[210] Pichet Termsarasab, Thananan Thammongkolchai, Janet C. Rucker, and Steven J. Frucht. The diagnostic value of saccades in movement disorder patients: a practical guide and review. *Journal of Clinical Movement Disorders*, 2:14, October 2015.

[211] K. M. Thomas, W. C. Drevets, R. E. Dahl, N. D. Ryan, B. Birmaher, C. H. Eccard, D. Axelson, P. J. Whalen, and B. J. Casey. Amygdala response to fearful faces in anxious and depressed children. *Archives of General Psychiatry*, 58(11):1057–1063, November 2001.

[212] Trieu H. Trinh and Quoc V. Le. A simple method for commonsense reasoning, 2019.

[213] A. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.

[214] Mark C.W. van Rossum, BJ O'Brien, and RG Smith. Effects of noise on the spike timing precision of retinal ganglion cells. *The American Physiological Society*, 89(5):2406–2419, 2003.

[215] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, Conference on Neural Information Processing Systems, Long Beach, CA, 2017. NeurIPS.

[216] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph Attention Networks, February 2018. arXiv:1710.10903 [cs, stat].

[217] Nora D. Volkow, George F. Koob, Robert T. Croyle, Diana W. Bianchi, Joshua A. Gordon, Walter J. Koroshetz, Eliseo J. Pérez-Stable, William T. Riley, Michele H. Bloch, Kevin Conway, Bethany G. Deeds, Gayathri J. Dowling, Steven Grant, Katia D. Howlett, John A. Matochik, Glen D. Morgan, Margaret M. Murray, Antonio Noronha, Catherine Y. Spong, Eric M. Wargo, Kenneth R. Warren, and Susan R. B. Weiss. The conception of the ABCD study: From substance use to a broad NIH collaboration. *Developmental Cognitive Neuroscience*, 32:4–7, August 2018.

[218] John von Neumann. *The Computer and the Brain*. Yale University Press, 1958.

[219] Jun Wang, Gert Cauwenberghs, and Frederic D Broccard. Neuromorphic dynamical synapses with reconfigurable voltage-gated kinetics. *IEEE Transactions on Biomedical Engineering*, 67(7):1831–1840, 2020.

[220] Xiao-Jing Wang and John H. Krystal. Computational Psychiatry. *Neuron*, 84(3):638–654, November 2014.

[221] Thomas V Wiecki, Jeffrey Poland, and Michael J Frank. Model-based cognitive neuroscience approaches to computational psychiatry: Clustering and classification. *Clinical Psychological Science*, 3(3):378–399, 2015.

[222] Raimond L. Winslow, Jeremy Rice, Saleet Jafri, Eduardo Marbán, and Brian O'Rourke. Mechanisms of altered excitation-contraction coupling in canine tachycardia-induced heart failure, ii. *Circulation Research*, 84:571–586, 1999.

[223] Eva Woelbert, Rory White, Kierstin Lundell-Smith, Jonathan Grant, and Danielle Kemmer. Inequities of mental health research funding. Technical report, International Alliance of Mental Health Research Funders, Montreal, 2020. WHO-137.

[224] Thomas Wolfers, Jan K Buitelaar, Christian F Beckmann, Barbara Franke, and Andre F Marquand. From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neuroscience and biobehavioral reviews*, 57:328–49, Oct 2015.

[225] Sung Sik Woo, Jaewook Kim, and Rahul Sarpeshkar. A digitally programmable cytomorphic chip for simulation of arbitrary biochemical reaction networks. *IEEE Transactions on Biomedical Circuits and Systems*, 12(2):360–378, 2018.

[226] Jihua Ye, Shengjun Xue, and Aiwen Jiang. Attention-based spatio-temporal graph convolutional network considering external factors for multi-step traffic flow prediction. *Digital Communications and Networks*, 8(3):343–350, June 2022.

[227] Andong Zhan, Srihari Mohan, Christopher Tarolli, Ruth B. Schneider, Jamie L. Adams, Saloni Sharma, Molly J. Elson, Kelsey L. Spear, Alistair M. Glidden, Max A. Little, Andreas Terzis, E. Ray Dorsey, and Suchi Saria. Using smartphones and machine learning to quantify parkinson disease severity the mobile parkinson disease score. *JAMA neurology*, 75(7):876–880, 2018. dias-36.

[228] Tianlin Zhang, Annika M. Schoene, Shaoxiong Ji, and Sophia Ananiadou. Natural language processing applied to mental illness detection: a narrative review. *npj Digital Medicine*, 5(46), 2022.

[229] Xin Zhang, Liangxiu Han, Wenyong Zhu, Liang Sun, and Daoqiang Zhang. An Explainable 3D Residual Self-Attention Deep Neural Network FOR Joint Atrophy Localization and Alzheimer's Disease Diagnosis using Structural MRI. *IEEE Journal of Biomedical and Health Informatics*, 26(11):5289–5297, November 2022. arXiv:2008.04024 [cs, eess].

[230] Yuezhou Zhang, Amos A Folarin, Shaoxiong Sun, Nicholas Cummins, Rebecca Bendayan, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Petroula Laiou, Faith Matcham, Katie M White, Femke Lamers, Sara Siddi, Sara Simblett, Inez Myin-Germeys, Aki Rintala, Til Wykes, Josep Maria Haro, Brenda Wjh Penninx, Vaibhav A Narayan, Matthew Hotopf, and Richard Jb Dobson. Relationship between major depression symptom severity and sleep collected using a wristband wearable device: Multicenter longitudinal observational study. *JMIR Mhealth Uhealth*, 9(4), 2021.

[231] Jianlong Zhao, Jinjie Huang, Dongmei Zhi, Weizheng Yan, Xiaohong Ma, Xiao Yang, Xianbin Li, Qing Ke, Tianzi Jiang, Vince D. Calhoun, and Jing Sui. Functional network connectivity (FNC)-based generative adversarial network (GAN) and its applications in

classification of mental disorders. *Journal of Neuroscience Methods*, 341:108756, July 2020.

[232] Yukun Zhu, Ryan Kiros, Richard Zemel, Ruslan Salakhutdinov, Raquel Urtasun, Antonio Torralba, and Sanja Fidler. Aligning books and movies: Towards story-like visual explanations by watching movies and reading books, 2015.

[233] Gregory Zilboorg. *A History of Medical Psychology*, pages 486–487. Norton, New York, 1941.