Uncovering the Connectome

by

Anish K. Simhal

Department of Electrical and Computer Engineering
Duke University

Date: ______________________
Approved:

__________________________
Guillermo Sapiro, Supervisor

__________________________
Lawrence Carin

__________________________
Henry Pfister

__________________________
Martin A. Brooke

__________________________
Richard J. Weinberg

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Electrical and Computer Engineering
in the Graduate School of Duke University
2019
Abstract

Uncovering the Connectome

by

Anish K. Simhal

Department of Electrical and Computer Engineering
Duke University

Date: __________________________

Approved:

______________________________
Guillermo Sapiro, Supervisor

______________________________
Lawrence Carin

______________________________
Henry Pfister

______________________________
Martin A. Brooke

______________________________
Richard J. Weinberg

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical and Computer Engineering in the Graduate School of Duke University 2019
Abstract

Over the past two decades, there has been an explosion in the number of tools and technology available to neuroscientists. With the advent of array tomography (AT) in the last decade, our ability to study synapses and their proteometric composition in the mammalian cortex has skyrocketed. However, unlike electron microscopy (EM) data which is the gold standard for synapse detection, AT data presents a variety of challenges in visualizing and characterizing synapses. There are many sources of noise, no singular definition of a synapse, and no standardized approach for data processing. In this work, our goal is to study synapse anatomy by combining array tomography with novel image processing methods. First, we started by creating a probabilistic synapse detector, which detects synapses based on their proteometric subtype with no training data. Then, we created a tool to characterize the efficacy of antibodies for array tomography applications. We end by expanding the probabilistic synapse detection method for tripartite synapses and explore the differences in synapses between wild-type and FMR1 knockout mice. This analysis lead to the discovery of several new effects of the FMR1 gene on astrocytic synapse density including the observation that there is a significant decrease in the density of excitatory glutamatergic synapses and in their association with astrocytes, while the changes in inhibitory GABAergic synapses are less pronounced. Our results suggest that that in Fragile X Syndrome astrocytes may mediate at least some of the pathological effects on glutamatergic synapses, while GABAergic synapses are likely
influenced by a different mechanism.
Contents

Abstract iv
List of Tables x
List of Figures xi
Acknowledgements xiv

1 Introduction 1
  1.1 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
  1.2 Probabilistic synapse detection . . . . . . . . . . . . . . . . . . . . 2
  1.3 Characterizing antibodies for array tomography . . . . . . . . . . . 3
  1.4 Exploring astrocytic synapses in FMR1 knockout mice . . . . . . . 3
  1.5 Document structure and key contributions . . . . . . . . . . . . . . 4

2 Probabilistic Synapse Detection 5
  2.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
  2.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
  2.3 Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
    2.3.1 Connectomics . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
    2.3.2 What are synapses? . . . . . . . . . . . . . . . . . . . . . . . 9
    2.3.3 Electron microscopy . . . . . . . . . . . . . . . . . . . . . . . 10
    2.3.4 Array tomography . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.4 Methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
4 Exploring Astrocytic Contributions to Synaptic Changes in a Mouse Model of Fragile X Syndrome

4.1 Abstract

4.2 Introduction

4.3 Methods

4.3.1 Overview

4.3.2 Data generation

4.3.3 Animals

4.3.4 Array tomography

4.3.5 Immunolabeling

4.3.6 Imaging method

4.3.7 Computational analysis

4.3.8 Synapse detection

4.3.9 Synapse type definitions

4.3.10 Volume calculation

4.3.11 Statistical analysis

4.3.12 Data and code availability

4.4 Results

4.4.1 Overview of the datasets and detected synapses

4.4.2 Single channel analysis

4.4.3 Overall synapse densities

4.4.4 Changes in synaptic densities in FMR1 KO mice

4.4.5 Involvement of glia (astrocytes)

4.5 Discussion
5 Conclusion  101

Bibliography  104

Biography  117
List of Tables

2.1 Synaptic markers used in this work across the various datasets. . . . . . 25
2.2 The cAT datasets used for analysis (Collman et al., 2015). . . . . . . 25
2.3 Excitatory synapse detection queries for the cAT data. . . . . . . . . 26
2.4 Inhibitory synapse detection queries for the cAT data. . . . . . . . . 26
2.5 Excitatory and inhibitory synapse detection results. . . . . . . . . . 27
2.6 State of the art detection results for excitatory synapses from (Coll-
man et al., 2015). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
2.7 Excitatory synapse detection queries for the AT data. . . . . . . . . 33
2.8 Inhibitory synapse detection queries for the AT data. . . . . . . . . 33
3.1 Antibodies used in this study . . . . . . . . . . . . . . . . . . . . . 51
3.2 Results from pairwise antibody comparisons. . . . . . . . . . . . . . 53
3.3 Five antibodies evaluated at different concentrations. . . . . . . . . 55
3.4 Summary of candidate antibody comparisons. . . . . . . . . . . . . . 58
4.1 Mice used for the experiments and their condition . . . . . . . . . . . 77
4.2 Antibodies used for the experiments . . . . . . . . . . . . . . . . . . 79
4.3 Queries used for this analysis . . . . . . . . . . . . . . . . . . . . . 84
4.4 Density distribution of excitatory synapses across layers as calculated
by two methods. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 90
# List of Figures

2.1 The currently available methods to enable mapping of brain circuits at the scale levels .......................... 9  
2.2 Excitatory and inhibitory synapses imaged via electron microscopy 10  
2.3 Synaptic vesicle diagram ........................................ 11  
2.4 Overview pipeline of the array tomography process ............ 13  
2.5 Outline of the proposed method in (Collman et al., 2015) ........ 15  
2.6 Molecular architecture of a excitatory PSD-95 expressing synapse depicted in a simplified cartoon form (Weiler et al., 2014) 16  
2.7 Series of data cutouts showing a synapse in both electron microscopy data and array tomography data .................. 16  
2.8 Fundamental steps of the proposed probabilistic synapse detection algorithm. ............................................. 17  
2.9 Histograms showing the distribution of signal in the datasets 19  
2.10 Step 1: Calculate foreground probability maps ................. 20  
2.11 Step 2: Calculate probability of a pixel belonging to a 2D blob 21  
2.12 Step 3: Calculate probability of belonging to a 3D blob ........ 22  
2.13 Step 4: Determine if two blobs are adjacent to each other .... 24  
2.14 The relationship between the precision and recall values across a series of thresholds for each cAT dataset. ...................... 28  
2.15 Synaptogram showing the distribution of IF data for an EM identified synapse. ............................................. 30  
2.16 Synaptogram showing a ‘false positive.’ .......................... 31
2.17 Synaptogram showing a ‘false negative.’ ........................................ 32
2.18 These plots show the variation of punitive synapse density across dif-
ferent thresholds. ................................................................. 34
2.19 The difference in synaptic density between Layer IV and Layer V for
specific queries. ................................................................. 35
3.1 Challenges in evaluating synaptic antibodies. .............................. 41
3.2 Pipeline of the Synaptic Antibody Screening Tool (SACT). ............ 42
3.12 Scatterplot of the homer1 antibody subclones ......................... 61
3.3 Schematic diagram of input datasets. ...................................... 65
3.4 Automated punctum detection pipeline. .................................... 66
3.5 Example of erratic labeling. .................................................. 67
3.6 Pairwise comparison of immunofluorescence on single sections from
mouse brain. ................................................................. 68
3.7 Impact of punctum size requirements on antibody comparisons. ... 69
3.8 Changes in target synapse density and target specificity ratio as a
function of antibody concentration. ........................................ 70
3.9 Comparison of multiple candidate antibodies. ......................... 71
3.10 Comparison of multiple candidate antibodies using two reference synap-
tic antibodies. ................................................................. 71
3.11 Comparison of multiple candidate antibodies. ......................... 72
4.1 Overview of a tripartite synapse and a query .............................. 82
4.2 Probabilistic synapse detection pipeline ................................. 83
4.3 Overview of the datasets .................................................... 88
4.4 Overview of wild-type synapse density distributions ................... 91
4.5 Changes in synapse densities across layers ............................... 93
4.6 Postsynaptic marker size distributions ................................... 94
4.7 Astrocytic synapse densities in wild-type mice ........................... 99
4.8 Summary of the astrocytic synapse density differences in the knockout mice. 100
Acknowledgements

Words cannot adequately convey how grateful I am to be part of Guillermo Sapiro’s lab and even more grateful for his mentorship. Guillermo’s wisdom, patience, and thoughtfulness is beyond reproach. Watching him interact with other students, collaborators, and myself has made me a better scientist. His enthusiasm for travel and his encouragement to travel myself has helped me foster better collaborations that I could have imagined. Furthermore, I’m incredibly grateful for the amount of time Guillermo has spent coaching me — both in lab and virtually. Guillermo’s ability to effectively lead a research group is something I hope to emulate one day.

Without Kristina Micheva, this dissertation would not have been possible. As the inventor of array tomography, she has been instrumental for not just conducting and preparing the many datasets we have analyzed together over the years, but also for teaching me the imaging science I needed to learn. She has been a wonderful mentor and I am forever grateful for her guidance over the past five years.

Richard Weinberg’s mentorship over the past five years has been an incredibly blessing. I am especially grateful for his time and patience during our morning meetings at Foster’s Market. He has taught me an incredible amount about synapse biology and his constructive criticism of my early manuscripts has hopefully made me a better writer.

I would like to thank the lab members, current and past, whose support and friendship has meant a tremendous amount and makes Gross Hall feel like a sec-
ond home. My thanks to go Jordan Hashemi, Zhuoqing Chang, Matias Di Martino, Steven Espinosa, Qiang Qiu, Dmitry Isaev, Martin Bertran, Natalia Gil, Ze Wang, Chris Tralie, Jinyoung Kim, Marcelo Fiori, Alasdair Newson, Mariano Tepper, Cecilia Aguerrebere, Mauricio Delbracio, Raja Giryes, and Sira Ferradans.

I would especially like to thank Cecilia Aguerrebere for her guidance when I first started at Duke. She provided invaluable advice for building the probabilistic synapse detection method. I would also like to thank Raja Giryes for inviting me to visit him and give a talk at Tel Aviv University and thank Federico Lecumberry for inviting me to visit and give a talk at the Universidad de la República in Uruguay.

None of the work presented here would be possible without the wonderful collaborators that I have had the privilege of working with. At the Allen Institute for Brain Sciences, I would like to thank Stephen Smith for teaching me about synapse biology and Forrest Colman for sharing his initial dataset with me. At Johns Hopkins University, I’d like to thank Will Gray Roncal, Joshua Vogelstein, and Randal Burns for their friendship — it was my pleasure to be an honorary member of their group at the Society for Neuroscience conferences. At UC San Francisco, I’d like to thank Bill Seeley for his conversations and insights about human synapse biology. At UC Davis, I’d like to thank James Trimmer, Belvin Gong, and Karl Murray for teaching me the process of creating and evaluating synaptic antibodies. At UC Santa Cruz, I’d like to thank Yi Zuo for her insights about the Fragile X Syndrome.

I would like to thank the members of the Gross Hall community for creating such a welcoming and fun workspace. I would especially like to thank Kathy Peterson for her support and encouragement. I would also like to thank Robert Calderbank for his time, especially for serving on my qualification exam committee and preliminary exam committee and Lawrence Carin for serving on my defense committee. I would also like to thank Martin Brooke and Henry Pfister for their invaluable feedback and for serving on all three examination committees.
I would also like to thank my wonderful friends from both Duke and home — their laughter helped make each day brighter. Moreover, I would like to thank Swathi Bolneni for her tireless support. Finally, I would like to thank my parents, Sunil and Sunita, and my siblings, Rishi and Sonali for their endless love. I would not be here without them.
1

Introduction

1.1 Overview

The brain is the most complex organ in the human body and, arguably, the least understood. Since Santiago Ramon y Cajal, the father of modern neuroscience and co-founder of the neuron doctrine, began his exploration of the brain in the late 19th century, the field of neuroscience has grown exponentially. The technology and tools used to study the brain today were simply unimaginable twenty years ago. Harnessing these new tools, such as cloud computing, petabyte scale data collection, and machine learning has allowed for new neuroscience discoveries to be made at a breathtaking pace. The beauty of modern neuroscience is that it cannot live in a single traditional discipline, such as biology. It is, by design, an incredibly interdisciplinary field encompassing, but not limited to, biology, chemistry, physics, engineering, mathematics, and computer science.

Neuroscience is an extremely broad field; our focus is to develop tools to increase our understanding of synapse anatomy. Understanding synapses, the junction between two neurons where information is transmitted, is vital to both figuring out how the brain works and how to address and treat a variety of diseases related to
synapses (commonly referred to as synaptopathies) such as autism spectrum disorder (ASD), Alzheimer’s, and fragile X syndrome (FXS). To study these topics, immense quantities of data must be collected and analyzed. As datasets regularly approach tens of gigabytes per set (and even larger for newer studies), efficient computational analysis tools are the bottleneck for neuroscientists.

Our goal is to study synapse anatomy by combining array tomography, an imaging method for brain tissue, with novel image processing methods. Previous studies examining synapses look at limited population subsets (~1000 synapses) due to the limitations of using electron microscopy to image synapses in the cortex. In this dissertation, we work towards creating the computational tools needed to conduct relatively large scale (>1,000,000 synapses) studies. We begin by presenting a tool to detect synapses in array tomography data, a tool to develop better antibody markers to detect synapses, and a framework for wholistic synapse analysis and apply it to a large scale study of synapses in the upper layers of the somatosensory cortices in FMR1 knockout mice.

1.2 Probabilistic synapse detection

Array tomography (AT) is an exciting new method for imaging mammalian brain tissue developed ten years ago by our collaborators at the Allen Institute for Brain Sciences and Stanford University and a discussion about it is in Chapter 2. AT involving embedding small portions of tissue in a plastic resin, creating ultrathin slices, and the staining each slice with a set of antibodies. These slices are imaged, registered, and aligned, thus creating hyperspectral three dimensional datasets. In these datasets, each channel is a protein, usually a synaptic protein, that is measured using a target antibody. The juxtaposition of multiple synaptic protein markers can define not just the location of a potential synapse but also the synapse subtype itself. Unlike electron microscopy (EM) data where synapse identification is relatively
straight forward, AT data is a challenging medium for visualizing and characterizing synapses. The probabilistic synapse detector is our answer to those challenges. It is a flexible framework which incorporates the uncertainty expert observers feel when looking for synapses in AT data. It is described in detail in Chapter 2.

1.3 Characterizing antibodies for array tomography

The success in collecting excellent array tomography data is wholly depending on the quality of the antibodies used to label the issue. Our collaborators at UC Davis and Stanford University approached us with an interesting proposal — to use the basic ideas of the probabilistic synapse detector and create a new tool to evaluate antibody performance at scale. Antibodies are notoriously difficult to create and even more difficult to adequately validate for a given application. Many antibodies for the same target protein may produce visually similar results, thus making visual examination difficult. Quantitative measurements of antibody quality is critical to compare scientific results across studies. While there is an effort to address this issue, there still lacks a reliable method for evaluating antibodies for array tomography. Our solution, the Synaptic Antibody Characterization Tool (SACT), automatically detects and characterizes antibody staining for array tomography and presents the user with descriptive, quantitative measures. We demonstrate this tool’s performance by evaluating 146 different antibody clones for the same target antigen and recreating the same observations indicated by expert observers. After validating the efficacy of this tool, we deployed it at UC Davis for their synaptic antibody evaluation pipeline. The SACT is described in detail in Chapter 3.

1.4 Exploring astrocytic synapses in FMR1 knockout mice

After creating novel tools for detecting and analyzing synapses in array tomography data, we focused on applying these tools to discover new synapse biology. To do so, we
collaborated with UC Santa Cruz and Stanford University to characterize the effect of the FMR1 gene on synapses. The FMR1 gene is believed to be responsible for Fragile X Syndrome, a genetic disorder that may cause intellectual disabilities and is often associated with autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD). For this work, we examined layers 1-4 of the mouse somatosensory cortex for four FMR1 knockout mice and four wild-type mice and explored the relationships of the synapses between the classes, cortex layers, proteometric composition, and astrocytic processes. To analyze the astrocytic markers, we expanded the probabilistic synapse detection method to encompass tripartite synapses. We analyzed eight immunofluorescence channels and detected approximately 1.6 million synapses comprising of fifteen different subtypes. We found previously unreported changes in the density of VGluT2 synapses in the upper layers of the cortex and changes in the density of astrocytic synapses between knockout and wild-type mice. Our results are described in detail in Chapter 4.

1.5 Document structure and key contributions

The dissertation continues as follows. Chapter 2 describes the biological background for detecting synapses with array tomography and describes AT data collection process in detail. It also outlines the probabilistic synapse detection method and validates it with electron microscopy based manual annotations as ground truth. Chapter 3 describes the current issues with antibody validation and proposes a novel framework for quantitative antibody characterization for array tomography. Chapter 4 expands the probabilistic synapse detection method to include tripartite synapses and describes the results of a study we performed that examines the effect of the Fragile X Syndrome on synapse development. We hope the insights from this study will go towards better characterizing the changes this genetic condition creates.
2.1 Abstract

The systematic examination of synaptic organization within large (> 1mm³) regions of the brain containing millions of synapses requires robust techniques for synapse detection. Electron microscopy (EM) is the gold standard modality for synapse detection, but data acquisition is much slower and more expensive than with immunofluorescence (IF) light microscopy. The introduction of immunofluorescence array tomography (AT) enables acquisition of hyperspectral volumetric proteometric data with IF, allowing identification of multiple synapse subtypes. This creates a need for effective synapse detection algorithms suitable for AT data. The molecular diversity of synapses makes creating a canonical synapse model very challenging. Traditional machine learning techniques need large amounts of labeled training data for each potential subclass, but this is often not available, especially for rare or even unknown subclasses. Therefore, synapse detection algorithms must incorporate both synaptic proteometric diversity and neuroscience knowledge.
In this paper, we present a query-defined probabilistic algorithm for synapse detection. The output is a probability map; each voxel represents the probability that it belongs to a synapse. For each synapse subtype, the user selects synaptic markers and synapse punctum sizes to be considered by the method. The proposed model-based algorithm incorporates fundamental biological knowledge of synapses and how they are manifested in the immunofluorescence data, providing detections with confidence values. We first evaluated the proposed approach on two conjugate array tomography datasets, for which synaptic identification has been confirmed with EM. Using a series of different query structures characterizing different known synapse subtypes, we detected 93% of PSD-95 expressing excitatory synapses, a level of accuracy comparable to that of human experts. We also evaluated this method on one of the largest open-source array tomography datasets available, confirming previous EM-based reports that showed differences in synaptic density in various layers of the mouse cortex. This fundamental approach opens the door to data-driven discovery of new synapse types and their density in a probabilistic fashion. The methods developed for this application are readily applicable to other datasets. To this end, all data, code, and data derivatives are available through the NeuroData website, neurodata.io.

2.2 Introduction

Neurons communicate with each other at synapses, highly-specialized cell-cell contacts. Mammalian synapses are very small (the area of contacts is typically < 0.1µm²) but with an extremely complex molecular composition: an individual synapse may contain hundreds of different proteins (O’Rourke et al., 2012). A better understanding of synapses will provide valuable insights into neurological diseases and their potential cures. An important step towards this goal is the systematic examination of excitatory and inhibitory synapses within large (> 1mm³) brain regions, which con-
tain many millions of synapses; to accomplish this will require both high-throughput imaging and data analysis techniques designed for efficient synapse detection. Electron microscopy (EM) provides the gold standard for synapse detection, since its nanometer resolution allows easy identification of key morphological features, including presynaptic vesicles, clefts, and postsynaptic densities. Unfortunately, EM data acquisition is slow, expensive, and time consuming, in contrast to immunofluorescence (IF) light microscopy. Physical constraints limit the resolution of classical light microscopy to > 200nm in the X – Y plane, and considerably more in the Z-axis, making it difficult to identify synapses. These limits have been overcome with a variety of super-resolution techniques, but each approach to super-resolution faces its own difficulties, especially for acquiring the kinds of massive datasets required to study cubic millimeters of tissue at the resolution necessary to identify synapses.

The molecular diversity of synapses makes labeling of specific synapse subtypes at scale extremely challenging and time consuming, and virtually impossible for rare synapse subtypes. To address these concerns, we present a query-defined probabilistic synapse detection algorithm inspired by neuroscientists’ manual synaptic identification methods. For each synapse subtype, the user selects which synaptic markers and minimum punctum sizes to consider (the query). The output of the proposed algorithm is a probability map, where each voxel value is the probability that it belongs to a synapse. Considering both the imperfections inherent to AT-IF and the potential presence of unknown synapse subtypes, considering a probabilistic approach is critical both for robustness and for new discoveries. This proposed model-based algorithm incorporates synapse biology and experience from manual synapse annotation techniques for IF data to provide detections with confidence values. The algorithm is validated with the most comprehensive AT datasets currently available, and compared with ground truth when available (via EM) or with published empirical data.
on the density of synaptic subtypes.

2.3 Background

2.3.1 Connectomics

In modern neuroscience, one of the greatest challenges is to understand the connections between the billion of neurons in the mammalian brain and thus create a complete wiring diagram of the brain, referred to as the ‘connectome.’ Motivation for the study of the connectome, referred to as ‘connectomics’, follows from the idea that understanding the complete wiring diagram of the brain will lead to unprecedented revolutions in medicine (Collman et al., 2015). For example, neurological disorders result in prolonged suffering and dramatically destroy an individual’s quality of life. For many neurological diseases (such as Alzheimer’s), it is well known fact that there is a direct correlation with the number of connections between neurons. Healthier neurons tend to form more dense connections, whereas Alzheimer’s patients exhibit a sparse connectivity distribution (Roncal et al., 2014). Therefore, understanding the difference in connectome between healthy and unhealthy patients will provide invaluable insights into the effect of Alzheimer’s and provide clues towards cures (Burette et al., 2015).

Creating the connectome is a daunting, almost impossible challenge. The human brain has approximately 100 billion neurons, each connecting to anywhere from one to ten thousand other neurons. Therefore, scientists estimate the human brain has approximately one quadrillion synapses. To put that in perspective, if each synapse was a ‘byte’, one would need a thousand terabytes of hardware storage to store the data. The first step towards creating a connectome is to image the brain itself. Imaging the brain and collecting information about it is a challenging task. Different imaging techniques provide different insights about the mammalian brain - Figure
2.1 shows the various image acquisition techniques commonly used at different scale levels.

2.3.2 What are synapses?

Classically, a synapse is defined as the junction between two neurons where information is transferred. There are two main types of synapses, electrical and chemical; however, chemical synapses constitute the majority. Hence, for this report, the term ‘synapse’ will be used to describe chemical synapses. A chemical synapse takes the incoming electrical signal and releases calcium ions (neurotransmitters) after which a response is triggered on the receiving end. There are two types of responses, inhibitory or excitatory. Excitatory synapses, when activated, increase the likelihood of the associated neuron triggering. Inhibitory synapses do the opposite. Synapses have two sides - presynaptic (the side with vesicles, the round circular objects, in an EM image) and post-synaptic (the fuzzy side of a Glutamatergic (glut) synapse in an EM image). These two types of synapses are shown in Figure 2.2.

The proteometric diversity present in synapses is astounding. Each synaptic vesicle
(the round circles around the red arrows in Figure 2.2) consists of hundreds of different types of proteins and there are over a thousand different proteins that are thought to span the synaptic cleft (the area the arrow is pointing to in Figure 2.2) (O’Rourke et al., 2012). Furthermore, inhibitory and excitatory synapses contain different sets of proteins. Figure 2.3 illustrates that idea by depicting a single synaptic vesicle (the sphere) and its associated proteins (attached to the sphere). Furthermore, different proteins are associated with different sides of a synapse, as highlighted in Figure 2.6.

2.3.3 Electron microscopy

As mentioned before, electron microscopy provides the gold standard for synapse detection because its nanometer resolution allows for easy identification of key biological features, such as presynaptic vesicles, clefts, and postsynaptic densities. Therefore, there have been numerous papers regarding detection synapses using only EM data (Becker et al., 2013) (Kreshuk et al., 2014) (Roncal et al., 2014). Generally speaking, these methods have two steps - extract features from labeled synapses and then train a classifier. These features take advantage of contextual clues associated with
known biological properties; however, they are dependent on various properties of the data itself, including contrast and resolution (Becker et al., 2013). Once features have been extracted, (Roncal et al., 2014) trains both a random forest classifier and a convolutional neural network, averaging a 82% accuracy rate. While these methods perform well, they are computationally very expensive. For example, one of the preprocessing steps in (Roncal et al., 2014) takes approximately ‘three weeks when using 27 Titan GPUs’ (Roncal et al., 2014).

Furthermore, creating electron microscopy data is both time-expensive and monetarily expensive. Hence, for the problem of large scale synapse detection, a different data modality, array tomography, is more effective for large scale discovery.
Array tomography (AT), a recently-introduced modification of IF performed on sequential arrays of ultrathin (< 70nm) sections of tissue embedded in a plastic block, enables the acquisition of volumetric stacks of hyperspectral proteometric IF data. The thinness of the sections ensures optimal X/Y resolution, while the sectioning process itself provides superior Z-axis resolution. AT-IF provides an excellent platform for characterization of synapses, allowing the identification of multiple synapse subtypes (Micheva and Smith, 2007). This new data acquisition technique allows efficient tissue imaging at scale, whose analysis requires specialized synapse detection algorithms.

The first steps of AT are to fixate the tissue in question, slice into ultrathin sections, and then stain with fluorescent antibodies. Fluorescent antibodies are molecules which attach to a specific protein type and fluoresces when activated by a specific wavelength of light. Because synapses contain a deep diversity of proteins, scientists use specific antibodies to target proteins specifically associated with synapses. After each round of imaging, the stains are removed (eluted) and the tissue is stained again with a different set of fluorescent antibodies. Following the imaging session, the image volumes are deconvolved by the appropriate point spread function (PSF) and then aligned, stitched, and registered to each other. The full pipeline is shown in detail in Figure 2.4.

The current state-of-the-art method for IF-based synapse detection uses machine learning algorithms to classify synapses based on large training datasets (Collman et al., 2015). Other synapse detection approaches have relied on extensive human annotation (Knott et al., 2002), machine learning of human-annotated IF puncta
Figure 2.4: Top Row: Array production. Fixed brain tissue is dissected into a small sample, in this case a tissue punch, and embedded in resin (generally LR White). This embedded sample is sectioned into ribbons of ultrathin serial sections, which are each affixed to a microscope coverglass to form a stable array. Middle Row: imaging and image processing. A ribbon array is stained with antibodies against selected antigens, and indirect immunofluorescence (IF) is imaged using a high-resolution objective. The antibodies can be removed from the ribbon using a high-pH elution solution, and the array can then be used again for multiple cycles of immunostaining and imaging. Image processing software improves the resolution of the resulting images. Bottom Row: volume reconstruction. Custom software is used for stitching, registration, and alignment of acquired images into volumetric reconstructions of the original tissue sample. Source: Nick Weiler (Weiler et al., 2014)
(Busse and Smith, 2013), or with rigid ad hoc rules about the relative positions of local maxima in the data (Wang and Smith, 2012). For (Busse and Smith, 2013), the algorithm requires the user to label hundreds of synapses of the same class, a tedious task. The current state-of-the-art synapse detection using array tomography data is presented in (Collman et al., 2015). This method uses associated electron microscopy data to label synapses, extracts features based on the brightness and size of associated puncta, and then trains a support vector machine (SVM). The method is outlined in further detail in Figure 2.5. While this method performs extremely well, averaging 90% accuracy rates (Collman et al., 2015), it still requires massive amounts of training data per synapse subtype. Thus, our proposed method seeks to address these issues by not requiring any training data or any extremely computationally expensive steps.

2.4 Methods

2.4.1 Overview

The proposed algorithm is inspired from biological knowledge of synapse characteristics. Detecting synapses using data from immunofluorescence imaging involves identifying adjacent presynaptic and postsynaptic antibody markers, as shown in Fig 2.6 which shows the approximate locations of excitatory synaptic proteins. Fig 2.7 is an example of an excitatory synapse with images of presynaptic and postsynaptic antibody markers overlaid upon an EM image.

Manual synapse identification involves relating in one channel the punctum size, brightness, and adjacency to puncta in other channels. Without corresponding EM data, detections using only IF data have an associated degree of uncertainty. Thus, we propose a query-based probabilistic synapse detection method which reflects the thought process underlying manual synapse detection.
Figure 2.5: Outline of the proposed method in (Collman et al., 2015)
Figure 2.6: Molecular architecture of a excitatory PSD-95 expressing synapse depicted in a simplified cartoon form (Weiler et al., 2014)

Figure 2.7: A series of $1.277\mu m \times 1.186\mu m$ cutouts of EM and IF data. The leftmost image has both PSD-95 (red) and Synapsin (green) data overlaid, marked by the colored boundary lines. The presence of both the presynaptic and postsynaptic channels indicates the presence of a synapse with high probability. The center cutout shows the PSD-95 IF image and the rightmost cutout shows the Synapsin IF image. On both images, the EM identified synaptic cleft is marked by a blue bounding box.
The first step is to compute the probability of each pixel in a slice being ‘bright.’ This indicates the likelihood that the antibody is detecting its corresponding protein antigen. Then, we compute the probability that it belongs to a 2D punctum. Depending on the synapse subtype and on the IF dataset resolution, these bright puncta may span multiple slices. Often, puncta which span multiple slices have a higher probability of belonging to a synapse than those which do not span multiple slices. To capture this phenomenon, a factor is computed that attenuates the probability of a pixel being a synapse, depending on whether the prospective punctum spans multiple slices or not. The last step in the computation of the synapse probability map consists in verifying the presence of adjacent presynaptic and postsynaptic puncta by correlating the corresponding IF volumes. This algorithm provides a general framework for the evaluation of a wide variety of synapse subtypes, defined by setting the presynaptic and postsynaptic antibodies and puncta size.

The following sections provide a detailed description of each step in the process, visualized in Fig 2.8.

2.4.2 Step 1: Foreground probability

Raw immunofluorescence image data is noisy; for example, speckles of the antibody markers often bind with cellular structures not associated with synapses, such as mi-
tochondria. In addition, fluorescence imagery contains signal from sources other than fluorescently-labelled antibodies, e.g. from background autofluorescence. Finally, all digital imagery contains fluctuations in intensity from sources such as camera read noise and photon shot noise. The noise produced by these sources is usually smaller in magnitude than that originating from authentic synaptic labeling. On the other hand, they cannot simply be filtered out and dismissed from consideration, since the signal may originate from a true synaptic site, and we want to allow for the possibility that a concordance of weak evidence will accrue and lead to the detection of a synapse. Thus, the first step of the algorithm consists of differentiating the bright voxels, the foreground, from the noisy background in a probabilistic fashion.

IF data volumes, when stained for synaptic markers, are also extremely sparse - approximately 2% of the voxels in the dataset belong to the foreground, as indicated in Fig 2.9. Therefore, the IF image volume can be used to approximate the distribution of the background noise.

Let $v(x, y, z)$ be the intensity value of a voxel at position $(x, y)$ in slice $z$, for a given channel of the IF data. For this purpose, a probabilistic model, $p_B$, is computed which characterizes all the pixels that belong to the background, which includes approximately 98% of the voxels. The background noise model is computed independently per slice to account for variations in autofluorescence. The background model $p_B$ is assumed to be a Gaussian distribution, whose mean and variance $(\mu_B, \sigma^2_B)$ are empirically computed from each slice $z$ (the $z$ index is omitted in Eq 1 for simplicity of notation). Then, the probability of a voxel belonging to the background, i.e., not being ‘bright,’ is given by

$$p_B(x, y, z) = \frac{1}{\sigma_B \sqrt{2\pi}} \int_{v(x,y,z)}^{\infty} e^{-(t-\mu_B)^2/2\sigma^2_B} dt.$$  

(2.1)
Figure 2.9: These histograms show that, on average, 98% of voxels in the dataset lie below the threshold line indicated in red. The threshold lines are estimates based on visual inspection of the data.

Therefore, the probability of a voxel associated with the foreground, \( p_F \), is computed as

\[
p_F(x, y, z) = 1 - p_B(x, y, z).
\]  
(2.2)

Fig 2.10 shows an example of the transformation from the raw data to the foreground probability map.

2.4.3 Step 2: 2D Puncta probability

Once foreground pixels have been identified in a probabilistic fashion, the next step is to determine if they form a 2D punctum. Since synapses appear as bright puncta in the IF image data, voxels which form puncta should have a higher probability of being associated with a synapse than those which do not. The probability of a voxel belonging to a 2D punctum, \( p_P \), is computed by multiplying the voxel’s foreground probability by that of its neighbors in a predefined neighborhood region,

\[
p_P(x, y, z) = \prod_{i=x-W}^{x+W} \prod_{j=y-W}^{y+W} p_F(i, j, z),
\]  
(2.3)
where $W$ is the neighborhood size, defined by the smallest expected punctum size. These operations are analogous to applying a box filter on the logarithm of the probability map. In our experiments, $W$ was set to be slightly larger than the size of the point spread function (PSF) of the microscopes used.

Fig 2.11 shows an example of the foreground probability map and the corresponding 2D puncta probability map. This step transforms the original point-wise decision into a region-based decision, and maintains the probabilistic aspect of the method.

2.4.4 Step 3: 3D Puncta probability

Potential synaptic puncta can span multiple slices of a given channel; puncta that span multiple slices have a higher probability of being associated with a synapse than those which do not. Therefore, we propose a factor $f(x, y, z)$ which diminishes the probability values associated with voxels which do not maintain a similar probability value in adjacent slices:
The pixel's 2D puncta probability is compared to that of its neighbor in slice(s) before, $j_{\text{start}}$, and slice(s) after, $j_{\text{end}}$. The number of slices compared is dependent on the input size parameter for each antibody. The factor attenuates values for 2D puncta that do not span the required number of slices, as shown in Figure 2.12.

The 3D puncta probability map is then computed by multiplying the 2D puncta probability map by this factor,

$$p_{3DP}(x, y, z) = p_P(x, y, z)f(x, y, z), \quad (2.5)$$

which further improves the probability of a detection by considering the slice-to-slice spatial distribution, going from spatial 2D to spatial 3D.
Figure 2.12: Top: Three consecutive slices of the 2D puncta probability. Bottom: Factor image given by Eq. (2.4) (left) and the corresponding 3D puncta probability (right) of the center slice in the top row. Notice how only those 2D puncta that actually span multiple slices are kept with high intensity (probability) in the combined result (bottom right). The green arrow points to an example of a probable punctum which spans multiple slices. The red arrow points to an example of a relatively less probable punctum which does not span multiple slices and therefore is diminished in the output image. Each image is a cutout of size 2261 × 2501 pixels or 5.268 × 5.827µm.

2.4.5 Step 4: Presynaptic and postsynaptic puncta adjacency

In electron microscopic images, synapses are identified by the presence of synaptic vesicles on the presynaptic side, the close adjacency of the membranes of the presynaptic axon terminal to a postsynaptic dendrite or dendritic spine, and the presence of a distinct postsynaptic specialization, as shown in Fig 2.6. Synapses are identified in IF data by the close spatial arrangement of pre- and postsynaptic antibody markers, which correspond to proteins associated with synapses. Therefore, the next step in our approach is to look for the presence of presynaptic puncta in the neighborhood
of postsynaptic puncta. More precisely, for each postsynaptic antibody voxel (i.e., PSD-95), we search in the adjacent 3D neighborhood of the corresponding presynaptic (i.e., synapsin) volume for a high intensity probability signal. To accomplish this, a rectangular grid is defined in the presynaptic channels around each postsynaptic voxel, as shown in Fig 2.13. The size of the grid is defined by the initial query parameters, which depend on the inherent biology and microscope resolution. The logarithm of the 3D puncta probability map (Eq. 2.5) is integrated in each grid location and the maximum is taken as presynaptic signal level around the given postsynaptic location,

\[
\log p_{\text{pres}} = \max \left( \sum_{G_k} \log p_{3DP}(\text{presynaptic}) \right), \tag{2.6}
\]

where the grid \( G \) is centered at the current voxel \((x, y, z)\) and divided into \( K \times K \times K \) subregions \( G_k \). To search in a grid around a defined voxel location for the presynaptic signal, \( K \) is set to 3. When searching for the postsynaptic signal, \( K \) is set to 1 since postsynaptic signals are expected to loosely co-localize. These values can be adopted to the data resolution, if appropriate. The postsynaptic antibody pixel probability \( p_{\text{post}} \) is multiplied by the presynaptic probability to obtain the desired probability map:

\[
p_{\text{post}} = p_{3DP}(\text{postsynaptic}), \tag{2.7}
\]

is multiplied by the presynaptic probability to obtain the desired probability map:

\[
p_{\text{synap}}(x, y, z) = p_{\text{pres}}(x, y, z)p_{\text{post}}(x, y, z). \tag{2.8}
\]

Again, the probability information is here maintained, this time including the morphological relationship between the channels. This ‘grid’ like approach allows the method to be robust to slight image alignment and registration issues.
Figure 2.13: The first row contains a cutout showing a PSD-95 punctum with a pixel highlighted in the center of the image. The second row contains synapsin cutouts with the search grid overlaid. For this example, $K = 3$, so shown is a $3 \times 3$ grid spanning 3 slices. The brightest box is highlighted in green. Thus, the output value of the synaptic probability map at the pixel specific in the PSD-95 image is the average pixel value of the green box multiplied by the intensity value of the PSD-95 pixel.

2.5 Experimental evaluation

The proposed method was evaluated on a series of array tomography (AT) datasets published in (Collman et al., 2015) and (Weiler et al., 2014). These datasets were acquired using the AT methodology described in (Micheva and Smith, 2007). Each dataset was stained and imaged with antibodies for presynaptic and postsynaptic proteins and then aligned and registered. In the conjugate AT (cAT) dataset (Collman et al., 2015), the tissue samples were also imaged with a scanning electron microscope (SEM) and then the IF data was up-sampled, aligned, and registered to the EM data. Synapses identifiable in the EM image data were labeled and used to
Table 2.1: Synaptic markers used in this work across the various datasets. Not all markers were present in each dataset (Collman et al., 2015; Weiler et al., 2014).

<table>
<thead>
<tr>
<th></th>
<th>Presynaptic</th>
<th>Postsynaptic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Excitatory</strong></td>
<td>Synapsin</td>
<td>PSD-95</td>
</tr>
<tr>
<td></td>
<td>VGluT1</td>
<td>NR1</td>
</tr>
<tr>
<td></td>
<td>VGluT2</td>
<td>NR2B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GluR1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GluR2</td>
</tr>
<tr>
<td><strong>Inhibitory</strong></td>
<td>Synapsin</td>
<td>Gephyrin</td>
</tr>
<tr>
<td></td>
<td>GABA</td>
<td>GABAAR</td>
</tr>
<tr>
<td></td>
<td>vGAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GAD</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: The cAT datasets used for analysis (Collman et al., 2015).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Dimension (Pixels)</th>
<th>Resolution</th>
<th>Labeled Synapses</th>
</tr>
</thead>
<tbody>
<tr>
<td>KDM-SYN-120905</td>
<td>4508 × 6306 × 27</td>
<td>2.33 × 2.33 × 70 nm/pixels</td>
<td>236</td>
</tr>
<tr>
<td>KDM-SYN-140115</td>
<td>7936 × 9888 × 39</td>
<td>3.72 × 3.72 × 70 nm/pixels</td>
<td>1457</td>
</tr>
</tbody>
</table>

provide ground truth. Table 2.1 lists the synaptic markers used.

2.5.1 Evaluation on Conjugate array Tomography

Experimental setup: The proposed method was first evaluated on the cAT dataset published in (Collman et al., 2015) using the associated EM image data to create the ‘ground truth’ needed for evaluation. The datasets themselves are described in detail in Table 2.2. To evaluate the method’s performance on excitatory synapses, the set of query parameters in Table 2.3 were used. For inhibitory synapse detection, the queries listed in Table 2.4 were used. These parameters were based on prior literature concerning synaptic proteins and their respective antibodies (Busse and Smith, 2013) (Weiler et al., 2014). Only 20 inhibitory synapses were manually identifiable in the KDM-SYN-120905 dataset; therefore, the inhibitory synapse detection performance is only reported for the larger KDM-SYN-140115 dataset. For evaluation and visualization purposes, the output probability map, \( p_{\text{synap}}(x, y, z) \), from each query was thresholded and adjacent voxels which lie over the threshold were grouped into
Table 2.3: Excitatory synapse detection queries for the cAT data.

<table>
<thead>
<tr>
<th>Query</th>
<th>Antibody</th>
<th>Presynaptic Puncta Size ((x,y,z) \mu m)</th>
<th>Postsynaptic Antibody</th>
<th>Postsynaptic Puncta Size ((x,y,z) \mu m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.21</td>
</tr>
<tr>
<td>2</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGluT1</td>
<td>0.2 x 0.2 x 0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>NR1</td>
<td>0.2 x 0.2 x 0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGluT1</td>
<td>0.2 x 0.2 x 0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGluT1</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.07</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGluT1</td>
<td>0.2 x 0.2 x 0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>NR1</td>
<td>0.2 x 0.2 x 0.21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4: Inhibitory synapse detection queries for the cAT data.

<table>
<thead>
<tr>
<th>Query</th>
<th>Antibody</th>
<th>Presynaptic Puncta Size ((x,y,z) \mu m)</th>
<th>Postsynaptic Antibody</th>
<th>Postsynaptic Puncta Size ((x,y,z) \mu m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>Gephyrin</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>GABA</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>Gephyrin</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGAT</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.21</td>
<td>Gephyrin</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>GAD</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Performance Metrics: The ground truth is created with EM data; therefore, some of the synapses can only be identified in the EM data because they do not appear in the IF data. In the same way, there are examples of false detections that experts cannot differentiate from true detections without using the EM data. It is expected that the proposed approach will not be able to handle those particular cases correctly.
Table 2.5: Excitatory and inhibitory synapse detection results. Precision is defined as the number of true positives detections / (true positive detections + false positive detections). Recall is defined as the number of true synapses detected / (true synapses detected + missed synapses).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Excitatory</th>
<th></th>
<th>Inhibitory</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EM Precision</td>
<td>Recall</td>
<td>IF Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>KDM-SYN-120905</td>
<td>0.88</td>
<td>0.91</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>KDM-SYN-140115</td>
<td>0.92</td>
<td>0.94</td>
<td>0.93</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table 2.6: State of the art detection results for excitatory synapses from (Collman et al., 2015).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Excitatory, EM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dataset Precision Recall</td>
</tr>
<tr>
<td>KDM-SYN-120905</td>
<td>0.96</td>
</tr>
<tr>
<td>KDM-SYN-140115</td>
<td>0.94</td>
</tr>
</tbody>
</table>

We report in Table 2.5 the precision and recall values obtained for these two tested datasets. We differentiate two cases: first, considering all synapses manually identified in the EM data and counting all false positives returned by the program (referred to in Table 2.5 as “EM”); and second, considering the subsection of detections which can be manually verified by an expert using only IF data (referred to in Table 2.5 as “IF”). For example, detections which the EM data lists as a false positives but are impossible to verify using only IF data are removed from evaluation. Similarly, manually-identified synapses in the EM data which do not appear in IF data are also removed from secondary evaluation.

**Results:** Once the final probability map for each query was computed, maps for excitatory synapses were thresholded at 0.6 for the KDM-SYN-140115 dataset and 0.55 for the KDM-SYN-120905 dataset. The maps for inhibitory synapses were thresholded at 0.9. The discrepancy between threshold values is attributed to the different noise distributions of each antibody. Note that this threshold, the only
Figure 2.14: The relationship between the precision and recall values across a series of thresholds for each cAT dataset. As detailed in the text, the threshold is mostly for validation purposes, the proposed framework outputs a confidence/probability.

non-biological parameter of the system, can be ignored when working directly on the output (Eq. 2.8) or easily set for the entire dataset by visually inspecting a few detections. Fig 2.14 shows the relationship between the final threshold and accuracy in greater detail. As shown in Table 2.5, the proposed algorithm successfully detects most synapses in both datasets, with only a small fraction of false positive detections. Based on the *IF only* indicator, we observe that the algorithm performs at human level (approximately 90% accuracy), with false positives and false negatives limited to cases which human experts (including co-authors of this manuscript) are also not confident of their own result (Collman et al., 2015).

As shown in Table 2.6, the proposed algorithm performs on par with the state-of-the-art method (Collman et al., 2015) for excitatory synapse detection, while eliminating the need to undergo the labor-intensive process of cultivating a training dataset. Furthermore, due to the approximate ten-to-one ratio of excitatory to inhibitory synapses, creating training sets for inhibitory synapses is notoriously difficult. Our method is insensitive to the number of synapses per class as it only returns possible synapses which match the query parameters. Finally, the fact that we can skip training also makes the proposed system more applicable to diverse datasets.
Fig 2.15 shows an example of a true positive detection of excitatory synapses in the KDM-SYN-120905 dataset. The figure shows an example of a synaptogram, where each row (third to sixth rows) shows a different channel and each column is a 2D slice. The first row marked as Label shows the manual annotation of the synaptic cleft, i.e., the ground-truth, and the second row, marked as Result, corresponds to the output of the proposed synapse detection algorithm. Rows 3-6 are corresponding sections of each channel’s foreground probability map (the output of Step 2). The seventh row, marked as EM, shows the corresponding EM data. Figure 2.16 shows an example of a false positive which cannot be differentiated from a real detection by an expert without the assistance of EM data. Figure 2.17 shows a similar situation for a false negative detection.

2.5.2 Evaluation on array tomography

The proposed method was evaluated on the array tomography dataset published in (Weiler et al., 2014), which contains a portion of the mouse barrel cortex from Layer 3 to Layer 5. Unlike the conjugate array tomography dataset, there is no associated EM imagery and therefore there is no associated EM-derived ground truth. This is also a larger series of datasets, consisting of 11 volumes which total 2,306,233µm³ of cortex data. Since no ground truth is available, the proposed method was evaluated by verifying known properties of the dataset: there is an approximately ten-to-one ratio of excitatory to inhibitory synapses (Knott et al., 2002) and there are more inhibitory synapses in Layer 4 than Layer 5 in the mouse barrel cortex (De Felipe et al., 1997).

For this dataset, the query parameters were adjusted to reflect the different synaptic
Figure 2.15: Synaptogram showing the distribution of IF data for an EM identified synapse. The first row (from the top) shows a manually labeled synaptic cleft, as identified in the EM volume. The EM data was used for only for validation, since the method operates solely on the IF data. The second row shows the thresholded output of the proposed method. Rows 3-6 show the corresponding foreground probability maps for each channel. PSD-95 is Postsynaptic Density 95, VGluT1 is Vesicular Glutamate Transporter 1, and NR1 is Anti-Glutamate Receptor N-methyl-D-aspartate Receptor 1. PSD-95 and NR1 are both postsynaptic markers and tend to co-localize, while synapsin and VGluT1 are both presynaptic markers and tend to co-localize. Each ‘block’ is $1.221 \mu m \times 1.233 \mu m$.

markers used. Tables 2.7 and 2.8 list the query parameters used for detecting both inhibitory and excitatory synapses.

Once the probability maps were computed, they were thresholded for evaluation purposes only. Thresholds for each dataset were determined by examining the synaptic
Figure 2.16: Synaptogram showing a ‘false positive.’ Presynaptic and postsynaptic proteins are visible and experts would often label this a synapse presented with IF data alone. Therefore, the algorithm makes the same mistake as a human expert. However, no synapse was identified in the corresponding EM section. Each ‘block’ is 1.069 µm × 1.011 µm.

density values across various thresholds, as shown in Fig 2.18. As the figure shows, the appropriate thresholds for each dataset exist in a narrow band, consistent with the results in the cAT dataset. Thresholding at the optimal value for each dataset, as set by plots in Fig 2.18, amounted to 2,326,692 excitatory synapses and 252,833 inhibitory synapses. This amounts to approximately 1.12 synapses per cubic micron and an overall ratio of 9.2 excitatory to inhibitory synapses, which is consistent with results in literature (Beaulieu et al., 1994) (Schüz and Palm, 1989). Findings

\footnote{As before, the threshold can be ignored when directly working on the final output probability map (Eq. 2.8) and is set here only for validation purposes.}
Figure 2.17: Synaptogram showing a ‘false negative.’ While the corresponding EM sections shows a synapse, there is a lack of synaptic IF signal available to justify the presence of a synapse using solely IF data. Once again, the algorithm makes the same mistake a human expert would make when working only with the IF data. Each ‘block’ is $1.086 \mu m \times 1.130 \mu m$.

in (Micheva and Beaulieu, 1995) indicate that there should be a difference in the synaptic density between Layer IV and Layer V. Specifically, the synapse density should be higher in Layer IV than in Layer V and the graphs in Fig 2.19 confirm it. For all three inhibitory synaptic queries, there is a synapse density difference of over 50% between Layer IV and Layer V. There is also a greater than 50% synaptic density difference between Layer IV and Layer V for excitatory synapses containing vGluT2, as supported by (Busse and Smith, 2013) (Graziano et al., 2008) (Nakamura et al., 2007). These results further demonstrate the validity of the proposed method.
Table 2.7: Excitatory synapse detection queries for the AT data.

<table>
<thead>
<tr>
<th>Query</th>
<th>Antibody</th>
<th>Presynaptic Puncta Size (x,y,z) μm</th>
<th>Postsynaptic Antibody</th>
<th>Postsynaptic Puncta Size (x,y,z) μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query 1</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.14</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.14</td>
</tr>
<tr>
<td>Query 2</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.14</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGluT1</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query 3</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.14</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>VGluT2</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query 4</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.14</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>GluR1</td>
<td>0.2 x 0.2 x 0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query 5</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.14</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>GluR2</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query 6</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.07</td>
<td>PSD-95</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NR2B</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
</tbody>
</table>

Table 2.8: Inhibitory synapse detection queries for the AT data.

<table>
<thead>
<tr>
<th>Query</th>
<th>Antibody</th>
<th>Presynaptic Puncta Size (x,y,z) μm</th>
<th>Postsynaptic Antibody</th>
<th>Postsynaptic Puncta Size (x,y,z) μm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query 1</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.07</td>
<td>Gephyrin</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GABAAR</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td>Query 2</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.07</td>
<td>Gephyrin</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>vGAT</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query 3</td>
<td>Synapsin</td>
<td>0.2 x 0.2 x 0.07</td>
<td>Gephyrin</td>
<td>0.2 x 0.2 x 0.07</td>
</tr>
<tr>
<td></td>
<td>GAD</td>
<td>0.2 x 0.2 x 0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

by confirming known biological properties of the data.

2.6 Discussion and conclusion

This work reported a model-based synapse detection algorithm that incorporates both fundamental biological knowledge of synapses and how they are identified by experts in the immunofluorescence data. For this purpose, we created a series of probabilistic excitatory detectors for various subtypes of synapses, and included the 3D spatial relationships typical of synaptic structures. This is a novel approach to provide a probabilistic-based detection algorithm yielding not only detection but
Figure 2.18: These plots show the variation of punitive synapse density across different thresholds. Each curve represents a dataset in (Weiler et al., 2014) and the red lines show the expected synaptic density. For excitatory synapses, the expected density is 0.9 synapses $\mu m^3 \pm 0.15 \mu m^3$. For inhibitory synapses, the expected density is 0.1 synapses $\mu m^3 \pm 0.05 \mu m^3$ (Calverley and Jones, 1987) (Schüz and Palm, 1989).

detection with confidence values. The implementation of synapse detection as a probability map (i.e., probability of each pixel belonging to a synapse), as opposed to a binary detection / no-detection result may provide a powerful tool to assist experts throughout the exploratory process to gain new insights from the immunofluorescence data, including potentially discovering new synapse subtypes. The influence of different biological and AT components on the actual probability values, from the noise of the system to the expression level of the proteins and the subclass of the synapses, is a very interesting subject of investigation, becoming possible when the proposed algorithm is applied to a larger amount of data currently being created.

The algorithm is basically parameter-free and computationally very simple, rendering it ready to be applied to massive datasets. The results were validated with the best available cAT and AT data, producing state-of-the-art results without the need for training. As demonstrated with the experiments in this paper, the proposed
Figure 2.19: The difference in synaptic density between Layer IV and Layer V for specific queries.

framework can be exploited for the explicit detection of synapses or their properties, the latter being critical for the discovery of new subtypes as well as the patterns of distributions. These are the subjects of our current efforts.
3.1 Abstract

Application-specific validation of antibodies is a critical prerequisite for their successful use. Here we introduce an automated framework for characterization and screening of antibodies against synaptic molecules for high-resolution immunofluorescence array tomography (AT). The proposed Synaptic Antibody Characterization Tool (SACT) is designed to provide an automatic, robust, flexible, and efficient tool for antibody characterization at scale. SACT automatically detects puncta of immunofluorescence labeling from candidate antibodies and determines whether a punctum belongs to a synapse. The molecular composition and size of the target synapses expected to contain the antigen is determined by the user, based on biological knowledge. Operationally, the presence of a synapse is defined by the colocalization or adjacency of the candidate antibody punctum to one or more reference antibody puncta. The outputs of SACT are automatically computed measurements such as target synapse density and target specificity ratio that reflect the sensitivity
and specificity of immunolabeling with a given candidate antibody. These measurements provide an objective way to characterize and compare the performance of different antibodies against the same target, and can be used to objectively select the antibodies best suited for AT and potentially for other immunolabeling applications.

3.2 Introduction

Antibodies are an indispensable tool for the modern biologist. Their high-affinity binding to specific target molecules makes it possible to detect, isolate, and manipulate the function of these molecules. A staggering number of antibodies are available to the research community, as are many options to make new antibodies. However, since antibodies are biological tools employed in complex systems, they can be very difficult to evaluate and to use in a predictable and reproducible way. A large volume of misleading data has been published based on results from antibodies that did not perform as assumed (Anderson and Grant, 2006; Baker, 2015; Rhodes and Trimmer, 2006). Recognizing this problem, there has been substantial progress in optimizing antibody production and validation (Nilsson et al., 2005; Gong et al., 2016). The importance of establishing reliable practices for antibody use is now widely accepted, and many companies are adopting transparent practices for rigorous antibody validation (Fritschy et al., 1998; Uhlen et al., 2016). The performance of antibodies, however, is application-specific (Lorincz and Nusser, 2008), and the reliable performance of an antibody in one application does not guarantee its suitability for another application. For example, an antibody that yields a single band on an immunoblot analysis of a tissue homogenate may prove wholly unsatisfactory for immunohistochemistry on sections of fixed tissue. Moreover, the same antibody that yields a robust and specific signal in immunohistochemical labeling of tissue sections prepared under one set of conditions may yield a weak or noisy signal on comparable samples prepared under different conditions (Fritschy et al., 1998; Fukaya and
Therefore, it is up to the individual user to validate antibodies for other applications and conditions. This task is especially crucial for applications whose chemistry differs substantially from standard immunoblots.

Array tomography (AT) is a technique that involves immunolabeling and imaging of serial arrays of ultrathin (~70 nm) plastic-embedded tissue slices of aldehyde-fixed tissue (Micheva and Smith, 2007; Micheva et al., 2010). While embedding tissue in resin has multiple advantages, the embedding process requires tissue dehydration, infiltration in plastic resin, and subsequent resin polymerization, all of which can modify the protein structure and chemical state and thus have a major impact on its immunoreactivity. Identifying antibodies that yield robust and specific immunolabeling of target proteins in plastic sections is a daunting task. AT is a high-resolution/high-throughput tool well-suited for the study of synapses in the mammalian brain; unfortunately, finding antibodies that selectively label synapses presents additional challenges, due to their small size, high density, and overall neurochemical complexity and diversity.

The primary criterion for evaluating antibody performance for immunohistochemistry is determining whether the labeling pattern is consistent with the known tissue characteristics of its target protein. For an antibody against a synaptic protein, the immunolabeling must be localized at synapses. Though conceptually straightforward, the practical evaluation of this criterion is difficult and often involves a number of subjective and time-consuming decisions. A synapse can be unambiguously identified via electron microscopy, but this approach is too time-consuming and expensive to be practical for large scale (~ 100 antibodies against a target protein) antibody screening tests. A more efficient strategy is to double label the same sample with another antibody already known to localize at synapses, and measure colocalization (Micheva et al., 2010; Weiler et al., 2014). While effective, this method presents a number of challenges. Synaptic proteins are typically expressed in high
concentrations at synapses; however, these proteins are also present at other subcellular locations. Furthermore, synapses display a high level of protein diversity, so many synapses may completely lack a particular synaptic protein. Adding to the uncertainty, other sources of fluorescence can confuse the interpretation of the images. These sources of ‘noise’ include signal arising from autofluorescent tissue constituents such as lipofuscin granules, blood cells, contamination, and defects such as tissue folds created during sample processing (Figure 3.1). The trained eye of an expert can usually discern the different fluorescence sources and pick out the specific immunolabel, but this is a subjective and non-quantitative assessment. Furthermore, it can be extremely time consuming, especially when examining a large number of antibodies against the same antigen, as may be required during antibody production (Gong et al., 2016).

Accordingly, there is an urgent need for an efficient and robust framework for evaluation of synaptic antibodies. Here, we introduce the Synaptic Antibody Characterization Tool (SACT), which provides automatic and quantitative measurements of the intensity and specificity of immunolabel and enables the objective characterization and comparison of multiple synaptic antibodies for AT at scale. Because the terms used in this paper are specific to the domain, below is a selected set of definitions.
Definitions

Candidate Antibody — Antibody being tested for the antigen of interest.
Reference Antibody — A previously validated antibody for an antigen known to colocalize with or lie adjacent to the candidate antibody’s antigen of interest.
Colocalization — When two or more antibody puncta occupy the same physical space, as shown in Box B. This is often the case when both antibodies have presynaptic targets, or both have postsynaptic targets.
Adjacency — When two or more antibody puncta are physically next to each other in 3D space, as shown in Box A. This may be the case when one antibody has a presynaptic target and the other a postsynaptic target.
Punctum — A small blob of signal defined in 2D or 3D space resulting from imaging an antibody applied on a tissue. In Box A, one punctum is the green circle, another is the blue triangle.
Target synapse — A synapse that contains the candidate antibody target, based on biological knowledge. Operationally, the presence of a synapse is defined by the adjacency (Box A) or colocalization (Box B) of the candidate antibody punctum to one or more reference antibody puncta. Box C is an example where the two puncta are neither adjacent nor colocalize, and thus do not form a synapse.
Target specificity ratio — The proportion of candidate antibody puncta that lie at target synapses.

3.3 Methods

3.3.1 Overview

The proposed Synaptic Antibody Characterization Tool (SACT) was developed for the quantitative assessment of antibodies against synaptic targets used for AT. It automatically detects puncta of immunofluorescence labeling from candidate antibodies and computes their density, size and size variability. SACT then determines whether
**Figure 3.1:** Challenges in evaluating synaptic antibodies. *Left.* Synaptic proteins are found not only at synapses, but also outside of synapses at sites of synthesis and transport; many synaptic proteins also lie in other subcellular compartments reflecting other functions (such as transcriptional regulators). *Right.* Immunofluorescence detection of synaptic proteins is confounded by nonspecific binding of antibodies, low efficiency of target protein detection, fluorescent contaminants such as dust particles, and tissue autofluorescence (e.g., lipofuscin granules and blood cells).

A punctum belongs to a target synapse or not, by using the previously described probabilistic synapse detector (Simhal et al., 2017). The presence of a target synapse is defined by the colocalization or adjacency of the candidate antibody punctum to one or more reference antibody puncta. This allows the output of two additional measures: density of target synapses, that is the synapses containing candidate antibody puncta, and target specificity ratio, which is the fraction of candidate antibody puncta that are at the target synapses. These measurements provide an objective way to characterize the sensitivity and specificity of a candidate antibody and to compare its performance to other antibodies. The approach is outlined in Figure 3.2. We should note that SACT is a framework; additional measurements can be added and adapted as needed depending on the desired antibody characterization.
features.

Figure 3.2: Pipeline of the Synaptic Antibody Screening Tool (SACT). SACT combines a probabilistic punctum detector (top row) with a probabilistic synapse detector (Simhal et al., 2017) to determine the properties of the candidate antibody.

The data used for validating SACT were derived from serial sections of plastic-embedded tissue that was immunofluorescently labeled with a candidate antibody alongside one or more reference antibodies, chosen depending on the antigen. ‘Candidate antibody’ refers to the antibody whose performance is being evaluated, and ‘Reference antibody’ refers to an antibody previously validated for AT that marks a synaptic protein expected to colocalize with or be adjacent to the target of the candidate antibody. The colocalization or adjacency of these two (or more) markers indicates the presence of a target synapse with high probability. The selected area was imaged on at least 3 consecutive sections, the images were aligned into stacks (Figure 3.3), and the performance of the tested antibody was assessed using SACT.

Importantly, SACT is applicable to a variety of synaptic antigens with very different distributions, because the user defines the expected molecular composition and size of synapses where the antigen is present. Furthermore, the algorithm can be applied to new datasets without creating extensive manual annotations for each synapse subtype, unlike traditional classifiers such as support vector machines and deep learning used by other synapse detection algorithms (Bass et al., 2017; Busse
and Smith, 2013; Collman et al., 2015; Fantuzzo et al., 2017; Kreshuk et al., 2014).

3.3.2 Puncta detection

Immunolabeling for synaptic proteins appears as small blobs or ‘puncta,’ typically less than 1 $\mu$m diameter. Because synaptic structures are generally larger than the typical thickness of the individual sections used in our datasets (70 nm), the puncta corresponding to proteins that are abundant throughout the presynaptic or postsynaptic side span several sections and thus form three-dimensional puncta.

The punctum detection method (Figure 3.4) is a special case of the synapse detection method and is adapted from it. It provides a way to take the input raw IF images from the microscope and output segmented 3D puncta, without having to set a threshold unique to every imaging session. The input is the volumetric image data and a user-defined query which includes the minimum expected 3D punctum size. Requiring a minimum 3D size minimizes the impact of random specks of noise generated during the image acquisition process and ensures that the immunolabeling is appropriately expressed across slices. For instance, a target protein that is abundantly expressed at a synapse (e.g., synapsin) should be detected across multiple slices at the current working resolution. Therefore, the presence of a punctum in only one slice likely indicates random noise, nonspecific labeling or fluorescent contaminant. On the other hand, there is little reason to assume that less abundant target proteins or those present at isolated nanodomains within synapses (e.g., many receptors or ion channels) need to span multiple slices through a synapse.

The probabilistic punctum detection algorithm involves three main steps. The first step transforms the data from the input raw IF images to a probability space. To do so, we create a Gaussian model for the background noise (other distribution models can be used as well) by assuming the entire input channel to be noise and the signal itself to be an outlier. For the Gaussian model, we use the input data’s
mean intensity value and standard deviation of the intensity values. The probability of a pixel belonging to the foreground is one minus the probability of belonging to the background. Next, we compute the probability that each pixel belongs to a 2D punctum. To do so, we multiply the probability values in a region defined by the user - the minimum expected 2D blob size. Usually, this is $0.2\mu m \times 0.2\mu m$, which corresponds to two pixels by two pixels. Then, we see if these 2D puncta exist in consecutive z slices. The number of expected slices is part of the user-defined minimum punctum size and is usually $0.14\mu m$, which corresponds to two slices. This minimum punctum size criteria reduces the effect of random noise. The output of this third step is a probability map, where the values at each pixel are the probability it belongs to a 3D punctum.

This probability map is thresholded to segment out 3D puncta detections. The threshold is established by manual observation (calibrated, if needed, with a small region of the data). The threshold was set to 0.9 for all the datasets in this project. Figure 3.4 shows the output of each step in this punctum detection pipeline. The initial panel shows a random slice / region of IF data. Each progressive panel shows a new requirement added, thus the number of “puncta” decrease accordingly. The last panel in Figure 3.4 shows the blobs/regions that have met the criteria necessary for being a punctum; not every blob shown in the first panel meets those requirements.

3.3.3 Synapse detection

Characterizing synaptic antibodies for AT immunolabeling of brain sections requires detecting synapses. Over the past few years, several synapse detection methods have been presented that use traditional machine learning paradigms for detection (Bass et al., 2017; Busse and Smith, 2013; Collman et al., 2015; Fantuzzo et al., 2017; Kreshuk et al., 2014). While they perform well, each requires the user to supply manually-labeled synapse annotations for training – an often impractical and tedious
requirement for antibody validation, for which manual annotations would have to be created for each antibody. The probabilistic synapse detection method introduced in Simhal et al. (2017) does not require training data for synapse detection, making it an ideal synapse detector to use for antibody characterization. This approach extends the probabilistic puncta detection method to look for colocalization or adjacency between puncta from different synaptic proteins. For colocalization, the method looks for signal in a two by two pixel window. For adjacency, the method looks for blobs in a six by six pixel window. The algorithm takes as input the immunofluorescence data from the candidate and reference antibodies and the expected target synapse size (together referred to as the 'query') and outputs a probability map, where the value of each voxel represents the probability that it belongs to a synapse. This output can be thresholded to obtain the putative 3D synapse detections (see (Simhal et al., 2017) for a detailed discussion). This algorithm is very flexible; as detailed below, the queries can be adapted to different data characteristics and analysis goals, further rendering it appropriate for antibody validation.

Generally, at least two known synaptic markers are required to unequivocally detect a synapse; therefore, some of the datasets contain two reference antibodies for synapse detection (Figure 3.3B, D). However, when screening multiple candidate antibodies, it is important to minimize the time and cost of the screen. When performed with caution using appropriate antibody combinations, even a single pre-validated antibody can generate interpretable synapse-specific data. For example, if the target protein for which an antibody is being evaluated is localized at the postsynaptic sites of inhibitory synapses, as is the case with gephyrin, (Sheng and Kim, 2011), a reasonable strategy would be to label the tissue with an antibody against glutamic acid decarboxylase (GAD), a protein known to be specific to presynaptic terminals of GABAergic inhibitory synapses for which well-validated antibodies have already been identified. The corresponding query would then be to look for synapses con-
taining immunolabeling for gephyrin and GAD (Figure 3.3A,C). If the target protein of interest is instead localized to excitatory synapses, the query may include proteins specific to excitatory synapses such as the postsynaptic protein PSD-95 (postsynaptic density 95) or the presynaptic protein VGluT1 (vesicular glutamate transporter 1). This flexibility makes this framework ideal for antibody characterization.

3.3.4 SACT measurements

In order to evaluate an antibody, a series of measurements are computed; each measurement captures an aspect of the antibody’s performance that would be sought by an expert observer when manually interacting with the data. Each of these measurements provides unbiased quantitative information to help evaluate the intensity, specificity and sensitivity of immunolabeling obtained with a given antibody. These include the density of puncta, and the average punctum volume and standard deviation for each antibody, as well as the target synapse density (number of detected target synapses belonging to the specified subclass per volume), and target specificity ratio (the ratio of detected synapses to detected candidate antibody puncta). These measures provide the user a useful quantitative assessment of the data.

**Detected Density of Puncta.** The density of the 3D puncta detected reflects the biological properties of the tissue (i.e., the abundance and distribution of the target protein in the tissue studied), as well as the intensity and specificity of the immunolabeling with a given antibody at the concentration used. If the labeling is unexpectedly sparse, it suggests that the antibody is insensitive or too highly diluted. If it is unexpectedly dense, it may indicate that the antibody is nonspecific or binds to many non-synaptic sites such mitochondria or other ‘sticky’ subcellular sites. To determine the number of puncta detected in each channel, probability maps, where the value at every voxel is the probability it belongs to a 3D punctum, are computed.
as described in the previous section, and then thresholded.\textsuperscript{1} The process is outlined in Figure 3.4. A density value of 0 indicates no puncta were found. The units (in this work) are in puncta per cubic micrometer. The 3D punctum density is then calculated:

\[
3D \text{ Punctum Density} = \frac{\text{Number of Detected 3D Puncta}}{\text{Total Volume}}
\]

**Average Punctum Volume.** After segmenting the immunolabeling data, the average volume of puncta and the standard deviation of the punctum volume are calculated. If the average punctum size is smaller or larger than expected, it may indicate a lack of efficacy or specificity for immunolabeling with that antibody. If the standard deviation of punctum volume is large, it may indicate erratic labeling. Either way, it is important to quantify the size distribution when comparing multiple candidate antibody clones for the same target. Figure 3.5 shows an example of an antibody clone displaying a very large standard deviation, making it unlikely that the candidate antibody on the left will serve as a satisfactory marker for inhibitory (collybistin-positive) synapses. We could compute fuzzy volumes if we prefer to work with the probability maps instead of the thresholded data. In this work, the values are in pixels.

**Target Synapse Density.** When evaluating a synaptic antibody, it is essential to confirm that the immunolabeling localizes at the expected population of target synapses. Target synapses are operationally defined by the relationship between two or more synaptic protein markers. This definition includes the size (in three dimensions) of each individual synaptic marker and a defined relationship (adjacent or colocalized) between the markers. Thus, the proposed SACT incorporates a proba-

\textsuperscript{1} The threshold could be avoided by changing the formula to be the sum of the probabilities instead of the thresholded number of 3D puncta. This applies to the other measurements described next as well.
The thresholded output of which is the number of synapses detected with the candidate antibody. The target synapse density of a given volume is computed as

\[
\text{Target Synapse Density} = \frac{\text{Number of Detected Synapses}}{\text{Total Volume}}
\]

This measure is useful for evaluating antibodies against targets with a known distribution at synapses, where the density of synapses containing the target protein can be estimated. For example, ubiquitous markers of inhibitory synapses like gephyrin should be present at the large majority of inhibitory synapses, and should therefore have a synaptic density in rodent neocortex on the order of 0.15 synapses per \(\mu m^3\) (Knott et al., 2002). A computed target synapse density substantially lower than expected may indicate low sensitivity of the antibody and/or insufficient concentration, while a target synapse density considerably above that expected likely reflects nonspecific (off-target) binding of the antibody. In this work, the units used are synapses per cubic micrometer. Note, the data is not thresholded until after searching for adjacency between puncta from different channels. The threshold for segmenting the resulting probability maps was set to 0.9 and was held constant for every experiment in the paper.

**Target Specificity Ratio.** The target specificity ratio (TSR) represents the fraction of immunofluorescent puncta of the candidate antibody that are associated with a target synapse, detected as explained above. TSR is computed as

\[
\text{TSR} = \frac{\text{Number of Detected Synapses}}{\text{Number of Detected 3D Puncta}}
\]

Thus, the target specificity ratio is a measure of how many times the candidate antibody being evaluated colocalizes with or is adjacent to the reference antibody compared to how many times it does not colocalize or lie adjacent with the refer-
ence antibody. TSR values range from 1 (every punctum detected has an associated detected synapse) to 0 (no detected punctum has an associated detected synapse); the higher the TSR, the lower is the magnitude of the nonsynaptic labeling obtained with that candidate antibody. Interpretation of this measurement will depend on the specific target; some synaptic proteins are almost exclusively present at synapses (e.g., synapsin), while others are also found at extrasynaptic locations (e.g., glutamate receptors). Therefore, the TSR reflects both the biological distribution of the target protein, and non-specific binding of the candidate antibody, but when comparing two antibodies against the same antigen, differences in TSR reflect differences in their specificity.

3.4 Materials

3.4.1 Datasets

The datasets presented here were created from adult mouse neocortical tissue that was prepared, immunolabeled, and imaged using standard methods of AT (Micheva et al., 2010; Collman et al., 2015). We used adult (3 to 4 months old) C57Bl/6J mice of both sexes for these experiments. Briefly, the tissue was chemically fixed using 4% paraformaldehyde in PBS, embedded in LR White resin, and cut into serial ultrathin sections (70 nm) that were mounted onto coverslips. One of the datasets from the automated ranking of the candidate monoclonal antibodies experiments (IRSp53) used tissue prepared in a different way: chemical fixation with 2% paraformaldehyde and 2% glutaraldehyde in PBS, followed by freeze-substitution and embedding in Lowicryl HM20 (Collman et al., 2015). The sections were labeled with indirect immunofluorescence using Alexa conjugated secondary antibodies (highly cross-adsorbed goat secondary antibodies against the relevant species, conjugated to Alexa 488, 594 or 647). Only Alexa 488 and 594 were used for the candidate primary antibodies. The difference in the theoretical lateral resolution of these two secondary antibodies cal-
culated using Abbe’s equation is 33 nm, which has little influence on our analysis for which the search area for colocalization is a 200 nm square. The samples were imaged on an automated wide-field fluorescence microscope (Zeiss AxioImager Z1, Zeiss, Oberkochen, Germany) with a 63x Plan-Apochromat 1.4 NA oil objective. The resulting images are not affected by many of the commonly occurring optical aberrations inherent to other immunofluorescence methods, because array tomography imaged sections are only 70 nm thick and the high-NA objective is always used at its design condition at the immediate contact of specimen and coverslip. While the exact size of the datasets varies, their general structure is consistent. Each dataset is composed of multichannel stacks of images from serial sections with (for the current data acquisition protocol) 100 \times 100 nm pixel size and 70 nm slice thickness.

3.4.2 Primary antibodies

The antibodies used for the experiments presented here are listed in Table 3.1. Some of the antibodies used were tested in conjunction with screening of monoclonal antibody projects at the UC Davis/NIH NeuroMab Facility, consistent with our goal to facilitate testing of large panels of candidate antibodies in an efficient and objective fashion.

3.5 Results

3.5.1 Framework evaluation

The proposed synaptic antibody characterization and screening framework was evaluated via three different tasks. Each task demonstrates an aspect of the framework necessary for validating synaptic antibodies.

1. **Pairwise comparisons.** Comparing the performance of two previously validated antibodies against the same synaptic reference protein.
Table 3.1: **Antibodies used in this study.** *RRID: Research Resource Identifier. For the NeuroMab projects, the RRID of the antibody finally selected is listed; this selection was based on other factors in addition to the antibody performance evaluated using the current method.*

<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Host</th>
<th>Antibody Source</th>
<th>RRID*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bassoon</td>
<td>Mouse</td>
<td>NeuroMab L124 project</td>
<td>RRID:AB_2716712</td>
</tr>
<tr>
<td>Cav3.1</td>
<td>Mouse</td>
<td>NeuroMab 75-206</td>
<td>RRID:AB_2069421</td>
</tr>
<tr>
<td>Cav3.1</td>
<td>Rabbit</td>
<td>Synaptic Systems 152 503</td>
<td>RRID:AB_2619850</td>
</tr>
<tr>
<td>Collybistin</td>
<td>Mouse</td>
<td>NeuroMab L120 project</td>
<td>RRID:AB_2650452</td>
</tr>
<tr>
<td>GAD2</td>
<td>Rabbit</td>
<td>Cell Signaling 5843</td>
<td>RRID:AB_10835855</td>
</tr>
<tr>
<td>Gephyrin BD</td>
<td>Mouse</td>
<td>BD Biosciences 612632</td>
<td>RRID:AB_399669</td>
</tr>
<tr>
<td>Gephyrin</td>
<td>Mouse</td>
<td>NeuroMab L106 project</td>
<td>RRID:AB_2617120</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RRID:AB_2617121</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RRID:AB_2632414</td>
</tr>
<tr>
<td>GluA1</td>
<td>Rabbit</td>
<td>Millipore AB1504</td>
<td>RRID:AB_2113602</td>
</tr>
<tr>
<td>GluA2</td>
<td>Mouse</td>
<td>Millipore MAB397</td>
<td>RRID:AB_2113875</td>
</tr>
<tr>
<td>GluA3</td>
<td>Rabbit</td>
<td>Abcam ab40845</td>
<td>RRID:AB_776310</td>
</tr>
<tr>
<td>Homer1</td>
<td>Mouse</td>
<td>NeuroMab L113 project</td>
<td>RRID:AB_2629418</td>
</tr>
<tr>
<td>IRSp53</td>
<td>Mouse</td>
<td>NeuroMab L117 project</td>
<td>RRID:AB_2619741</td>
</tr>
<tr>
<td>GluN1</td>
<td>Mouse</td>
<td>Millipore MAB363</td>
<td>RRID:AB_94946</td>
</tr>
<tr>
<td>PSD95</td>
<td>Mouse</td>
<td>NeuroMab 75-028</td>
<td>RRID:AB_2292909</td>
</tr>
<tr>
<td>PSD95</td>
<td>Rabbit</td>
<td>Cell Signaling 3450</td>
<td>RRID:AB_2292883</td>
</tr>
<tr>
<td>Synapsin</td>
<td>Guinea pig</td>
<td>Synaptic Systems 106 004</td>
<td>RRID:AB_1106784</td>
</tr>
<tr>
<td>Synapsin</td>
<td>Rabbit</td>
<td>Cell Signaling 5297</td>
<td>RRID:AB_2616578</td>
</tr>
<tr>
<td>VGAT</td>
<td>Mouse</td>
<td>NeuroMab L118 project</td>
<td>RRID:AB_2650550</td>
</tr>
<tr>
<td>VGlut1</td>
<td>Guinea pig</td>
<td>Millipore AB5905</td>
<td>RRID:AB_2301751</td>
</tr>
<tr>
<td>VGlut1</td>
<td>Mouse</td>
<td>NeuroMab 75-066</td>
<td>RRID:AB_2187693</td>
</tr>
</tbody>
</table>

2. **Concentration comparisons.** Comparing the performance of a single antibody at different concentrations.

3. **Evaluating candidate monoclonal antibodies.** Comparing the performance of multiple candidate antibodies against the same synaptic target protein.

The first and second task validate the measurements proposed in the framework and involve only antibodies previously validated for array tomography. The third task evaluates the efficacy in a ‘real world’ application — characterizing multiple
antibody candidates whose suitability for AT has not yet been determined, and whose concentration is not known.

3.5.2 Pairwise comparisons

When two effective antibodies are available for use in a specific application, a common question is ‘which one is better?’ To answer this question, we created five AT datasets, each with two previously validated antibodies used at concentrations previously determined to yield optimal immunolabeling. These antibody pairs were evaluated alongside an antibody for a different synaptic target protein, thoroughly validated for AT in prior studies (Micheva et al., 2010; Weiler et al., 2014). An example slice of each dataset is shown in Figure 3.6. The higher-scoring antibody was judged to be the one that had more puncta associated with the reference antibody (i.e., labels more synapses, true positives) and/or fewer puncta not associated with the reference antibody (false positives). In a dataset comparing different PSD-95 clones, the PSD95R antibody clone had more puncta adjacent to synapsin, without noticeably more synapsin-unrelated puncta, and was therefore evaluated as better performing than PSD95M. Cav3.1R had more puncta that are not adjacent to VGluT1, and also displayed nonspecific labeling of the cell nucleus, and was therefore judged to perform worse than Cav3.1M. In the VGluT1 dataset, differences between the two antibodies were more subtle, as shown by the measurements of target synaptic density and target specificity ratio (see Table 3.2).

For each dataset, the minimum expected marker size was set at $0.2 \times 0.2 \times 0.14 \mu m$, corresponding to 2 pixels by 2 pixels by 2 slices. Each dataset was also independently evaluated and ranked by an expert observer (KDM) blind to the automatically computed results, based on visual examination of the immunolabeling. Two measures, target synapse density and target specificity ratio, were used to rank the two candidate antibodies (Table 3.2). When the antibodies are used at their optimal con-
Table 3.2: **Results from pairwise antibody comparisons.** Punctum density, average punctum size, and standard deviation of punctum size were all computed for each antibody tested. The names in bold in the second column represent the antibodies preferred by an expert based on visual examination. In each case, the measures automatically computed by the framework agree with the expert observer’s judgement.

<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Candidate antibodies</th>
<th>Reference Antibody</th>
<th>Punctum density / ( \mu m^3 )</th>
<th>Average punctum size / standard deviation (pixels)</th>
<th>Target Synapse Density / ( \mu m^3 )</th>
<th>Target Specificity Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synapsin</td>
<td>Synapsin GP</td>
<td>PSD-95</td>
<td>2.01</td>
<td>15.9 / 59.4</td>
<td>0.55</td>
<td>0.275</td>
</tr>
<tr>
<td>Synapsin R</td>
<td></td>
<td></td>
<td>1.09</td>
<td>27.0 / 48.6</td>
<td>0.5</td>
<td>0.457</td>
</tr>
<tr>
<td>VGluT1</td>
<td>VGluT1GP</td>
<td>Synapsin</td>
<td>1.54</td>
<td>12.1 / 30.2</td>
<td>0.44</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>VGluT1M</td>
<td></td>
<td>1.27</td>
<td>11.0 / 22.4</td>
<td>0.33</td>
<td>0.257</td>
</tr>
<tr>
<td>PSD-95</td>
<td>PSD-95M</td>
<td>Synapsin</td>
<td>0.95</td>
<td>10.7 / 101.7</td>
<td>0.54</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>PSD-95R</td>
<td></td>
<td>1.14</td>
<td>25.4 / 39.9</td>
<td>0.79</td>
<td>0.691</td>
</tr>
<tr>
<td>Gephyrin</td>
<td>GephyrinL106</td>
<td>GAD</td>
<td>0.72</td>
<td>12.5 / 94.1</td>
<td>0.13</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>GephyrinBD</td>
<td></td>
<td>0.46</td>
<td>17.9 / 263.1</td>
<td>0.08</td>
<td>0.175</td>
</tr>
<tr>
<td>Cav 3.1</td>
<td>Cav3.1M</td>
<td>VGluT1</td>
<td>0.58</td>
<td>7.4 / 14.0</td>
<td>0.33</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>Cav3.1R</td>
<td></td>
<td>1.22</td>
<td>8.6 / 52.3</td>
<td>0.33</td>
<td>0.271</td>
</tr>
</tbody>
</table>

Concentration, a higher measured target synapse density implies higher sensitivity of the antibody (since it detects the target protein at more synapses). A higher target specificity ratio (i.e., a higher proportion of detected immunolabeled puncta that are associated with synapses) indicates higher selectivity of the antibody for the protein of interest. The expert-preferred VGluT1 and PSD-95 antibodies scored higher on both sensitivity (target synapse density) and specificity (TSR), while others scored higher on only one of these measures. For both Cav3.1 and synapsin, the higher-scoring antibodies had higher TSR but gave target synapse densities comparable to the other antibody. In the case of gephyrin, both antibodies had similar TSR, but the expert picked antibody had a higher target synapse density. These results illustrate the importance of using complementary measurements for antibody evaluation.

The proposed framework provides multiple objective computations, and the user can
pick the most suitable one(s) for a given task.

To evaluate the robustness of the framework, the same comparisons were performed using queries with smaller and larger minimum synapse size requirements (requiring puncta to span only one slice vs three slices), as shown in Figure 3.7. All queries gave consistent results for all five antibody pairs, except for Query 1, which defined a synapse as spanning only one slice; in two out of the five cases, Query 1 failed to unequivocally identify the otherwise highest scoring antibody. Thus, the use of even limited three-dimensional information from immunolabeling on serial sections enabled robust quantification of antibody performance.

This experiment illustrates the power and breadth of the proposed method. The queries can be designed by the user to take into account resolution, synapse type, and antibody binding target. Multiple queries can be run, and the antibody performance can be objectively evaluated with multiple measurements.

3.5.3 Concentration comparisons

The optimal concentration of an antibody, which is dependent on both its binding affinity for the target protein and the abundance of the target protein in the particular sample, must be determined experimentally. Too high a concentration of the antibody will lead to high background labeling (false positives), while too low a concentration will lead to sparse labeling (false negatives). The proposed framework quantifies the effects of antibody concentration on immunolabeling of AT sections, as evaluated by the target synapse density and target specificity ratio measures. As the antibody concentration decreases, the target synapse density is also expected to decrease.

For this experiment, datasets were generated from a series of dilutions, as shown in Table 3.3 and Figure 3.8. For each dataset except GluN1 the minimum expected punctum size was $0.2 \times 0.2 \times 0.14\mu m$, corresponding to 2 pixels by 2 pixels by 2 slices.
Table 3.3: **Five antibodies evaluated at different concentrations.** In all experiments, a reference presynaptic antibody was included at its optimal concentration. The detected target synapse density decreases as the antibody concentration decreases. For GluN1, the background noise model changed to a Rayleigh distribution to better suit that specific dataset. All other datasets used the a Gaussian model for the background. See (Simhal et al., 2017) for a detailed discussion.

<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Candidate Antibodies</th>
<th>Reference Antibody</th>
<th>Punctum Density/ µm³</th>
<th>Average Punctum Size / standard deviation (pixels)</th>
<th>Target Synapse Density/ µm³</th>
<th>Target Specificity Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synapsin</td>
<td>SynCS 1:100</td>
<td>VGlut1</td>
<td>1.23</td>
<td>41.6 / 133.6</td>
<td>0.84</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>SynCS 1:1000</td>
<td></td>
<td>1.31</td>
<td>35.9 / 95.9</td>
<td>0.83</td>
<td>0.63</td>
</tr>
<tr>
<td>GluA1</td>
<td>GluR1 1:25</td>
<td>VGlut1</td>
<td>0.62</td>
<td>7.9 / 51.8</td>
<td>0.12</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>GluR1 1:125</td>
<td></td>
<td>0.43</td>
<td>9.5 / 41.4</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>GluR1 1:625</td>
<td></td>
<td>0.18</td>
<td>19.3 / 180.9</td>
<td>0.05</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>GluR1 1:3125</td>
<td></td>
<td>0.13</td>
<td>38.7 / 1117.6</td>
<td>0.03</td>
<td>0.21</td>
</tr>
<tr>
<td>GluA2</td>
<td>GluR2 1:25</td>
<td>Synapsin</td>
<td>0.89</td>
<td>10.5 / 62.3</td>
<td>0.45</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>GluR2 1:125</td>
<td></td>
<td>0.46</td>
<td>13.9 / 29.5</td>
<td>0.30</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>GluR2 1:625</td>
<td></td>
<td>0.27</td>
<td>15.5 / 85.8</td>
<td>0.17</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>GluR2 1:3125</td>
<td></td>
<td>0.23</td>
<td>15.9 / 103.8</td>
<td>0.13</td>
<td>0.57</td>
</tr>
<tr>
<td>GluA3</td>
<td>GluR3 1:25</td>
<td>VGlut1</td>
<td>1.99</td>
<td>6.5 / 11.8</td>
<td>0.42</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>GluR3 1:125</td>
<td></td>
<td>0.81</td>
<td>8.4 / 16.4</td>
<td>0.23</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>GluR3 1:625</td>
<td></td>
<td>0.26</td>
<td>10.8 / 23.5</td>
<td>0.07</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>GluR3 3:1325</td>
<td></td>
<td>0.17</td>
<td>20.6 / 154.1</td>
<td>0.05</td>
<td>0.26</td>
</tr>
<tr>
<td>GluN1</td>
<td>GluN1 1:25</td>
<td>VGlut1</td>
<td>1</td>
<td>37.2 / 1541.0</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>GluN1 1:125</td>
<td></td>
<td>0.88</td>
<td>17.8 / 208.2</td>
<td>0.63</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>GluN1 1:625</td>
<td></td>
<td>0.58</td>
<td>22.0 / 748.9</td>
<td>0.37</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>GluN1 1:3125</td>
<td></td>
<td>0.64</td>
<td>9.8 / 54.5</td>
<td>0.39</td>
<td>0.62</td>
</tr>
</tbody>
</table>

The minimum punctum size for GluN1 was \((0.2 \times 0.2 \times 0.07 \mu m)\) due to inaccuracies in the alignment of this dataset that caused inconsistencies in the positions of synapses on adjacent slices (the proposed algorithm can be easily adapted to challenges in the data by changing the query).

The first dataset tested the effect of a 10-fold change in concentration of an antibody against the general presynaptic marker synapsin, imaged in conjunction...
with immunolabeling for VGluT1, a presynaptic marker of excitatory synapses. The remaining datasets tested four sequential 5-fold concentration changes on immunolabeling with different glutamate receptor antibodies, and were evaluated against immunolabeling with antibodies for general presynaptic markers (synapsin) or markers of glutamatergic synapses (VGluT1). Synapsin, previously identified as a robust synaptic antibody for AT, performed equally well over the 10-fold concentration range as evaluated by both the target synapse density and target specificity ratio measurements. For each of the glutamate receptor antibodies, the measured target synapse density value decreased with increasing dilutions, as expected, while the target specificity value showed no consistent changes. Using the framework, we estimated that the optimal working range of the glutamate receptor antibodies tested lies within a dilution range of 1:25 to 1:125; further dilutions led to missing too many synapses without a substantial improvement in target specificity.

3.5.4 Automated ranking of candidate monoclonal antibodies

The generation of monoclonal antibodies begins with a high-throughput screening procedure that identifies numerous candidate antibodies, all of which must then be further investigated. Since only a small fraction of these candidate antibodies will exhibit robust and specific immunoreactivity in any given condition, it is important to screen as many candidate antibodies as possible for a given application. While some common antibody screens have been effectively automated (e.g., ELISA screens), screening on plastic sections from mammalian brain for antibodies that immunolabel specific populations of synapses must still be performed and analyzed manually by an expert observer, a difficult and labor-intensive process. Reasoning that the framework proposed here might facilitate the analysis of large-scale screens on tissue sections, we tested its performance by screening candidate monoclonal antibodies against synaptic target proteins generated at the UC Davis/NIH NeuroMab facility.
This procedure is especially challenging because the concentration of antibody in hybridoma tissue culture supernatants is unknown, so immunolabeling must be performed at antibody concentrations that may differ for different candidate antibodies, and these concentrations may not be optimal.

Arrays from mouse neocortex were prepared using standard AT methods. For each dataset, we imaged sections immunolabeled with a set of candidate antibodies against the same target protein. For each dataset, at least two antibodies were applied: the candidate antibody raised against the target protein of interest and a validated antibody at its optimal concentration. The ranking of candidate antibodies was determined based on two measurements provided by the framework: target synapse density and the target specificity ratio. The target specificity ratio was the deciding factor in most cases. Target synapse density was used to exclude candidate antibodies with unreasonably high values based on previous biological knowledge: excitatory synapses are expected to have a density of $\sim 1$ per $\mu m^3$, and the inhibitory synapses a density of $\sim 0.15$ per $\mu m^3$ (Calverley and Jones, 1987; Schüz and Palm, 1989; Knott et al., 2002). Each dataset was blindly evaluated and ranked by an expert observer, based on visual examination of the images. Screening of the Bassoon candidate antibodies was performed in two rounds; the second round included only those candidates identified as best or unclear in the first round. These experiments addressed several questions: 1) Can the framework be used to correctly rank the performance of multiple candidate antibodies? 2) What is the minimum number of reference antibodies required to accomplish this? and 3) What is the optimal minimum punctum size needed? Six datasets, ranging from 4 to 19 different candidate antibodies each, were analyzed.

Can the framework correctly rank the performance of multiple candidate antibodies?

Analysis of the six datasets demonstrated an excellent correspondence between
the framework’s ranking and expert evaluation of candidate antibody performance. The results are summarized in Table 3.4 and Figure 3.9. The computed measures not only allow the relative ranking of antibody performance, but also give an indication of the absolute utility of an antibody. For example, in the case of IRSp53, the two expert-preferred antibodies were ranked higher by the SACT than the other antibodies. However, the SACT measurements indicated that overall none of the IRSp53 antibodies tested were performing sufficiently well, because the TSR was extremely low (< 0.03).

Table 3.4: **Summary of candidate antibody comparisons.** Multiple candidate antibodies against the target protein were ranked using measurements from the proposed SACT. The candidate antibodies were independently evaluated by visual inspection of the immunofluorescence images. For each target protein, the reference antibodies were chosen according to the known characteristics of synapses expressing the target protein. Without an objective and automated tool to quantify antibody performance, all antibody evaluations are subjective (as is the standard today). The framework proposed in this work is a first step towards having synaptic antibody quantification be a routine part of antibody evaluation.

<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Reference antibodies</th>
<th>Total # candidate antibodies tested</th>
<th># good candidates chosen by expert</th>
<th># good candidates chosen by SACT</th>
<th>SACT false negatives</th>
<th>SACT false positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gephyrin</td>
<td>GAD</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Homer</td>
<td>PSD-95</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IRSP53</td>
<td>PSD-95, VGluT1</td>
<td>12</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0 (+1 unclear)</td>
</tr>
<tr>
<td>VGAT</td>
<td>GAD</td>
<td>17</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0 (+2 unclear)</td>
</tr>
<tr>
<td>Collybistin</td>
<td>GAD</td>
<td>19</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Bassoon 1st exp.</td>
<td>Synapsin</td>
<td>19</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0 (+1 unclear)</td>
</tr>
<tr>
<td>Bassoon 2nd exp.</td>
<td>VGluT1, PSD-95</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
What is the minimum number of reference antibodies?

The previous experiments with pairwise or concentration comparisons were performed using only one reference antibody. In those cases, the tested antibodies were already known to recognize their target in plastic sections; it is therefore reasonable to assume that the combination of one reference synaptic antibody with one tested synaptic antibody will generate synapse-specific data. To verify whether an additional reference antibody may offer an advantage when screening new antibodies, two of the datasets included both presynaptic and postsynaptic reference antibodies. In these two datasets, the performance of the query containing an additional reference antibody was compared to the standard single reference antibody query used in the previous experiments. In both cases, the results of the two queries were very similar (compare Figure 3.10 with Figure 3.9), suggesting that inclusion of a second reference synaptic antibody is unnecessary for the purpose of screening large sets of candidate antibodies.

What is the optimal minimum punctum size?

The pairwise antibody comparison experiments showed that the results were not affected by the stringency of the query, except in cases when the minimum puncta requirements were too permissive (smallest synapse size: labels present on 1 slice). Therefore, to screen multiple candidate antibodies, we generally chose queries of medium stringency, requiring the labels to be present in two consecutive slices. This strategy worked very well for candidate antibodies directed against abundant synaptic proteins (gephyrin, Homer1, IRSp53, VGAT, collybistin). In contrast, the permissive query, which required the labels to be present on only one section, gave inconclusive results in most of these cases (Figure 3.11). The first round of screening for Bassoon antibodies was an exception, because it yielded clearer results with a one-section query. This is likely due to the wide variations in concentration of the candidate antibodies present in the tissue culture supernatants used for screening,
many of which required subsequent dilution, as performed in the second round of testing. In this second round with adjusted concentrations, the two-section query performed well, as seen for the other abundant synaptic target proteins. These experiments suggest that it is best to start an antibody evaluation using a query that requires the labels to be present in two sections. The top-ranking antibodies based on such a query can then be selected and visually examined by experts to confirm their performance.

3.6 Homer1 antibody characterization

The Synaptic Antibody Screening Tool (SACT) was developed to characterize synaptic antibodies for array tomography at scale. To test its efficacy for datasets with more than a hundred antibody clones, we used a dataset that was created independently of the development of the SACT and features the Homer1 protein.

The Homer1 protein is among the most abundant scaffolding proteins in the post synaptic density (PSD) region of a synapse and is thought to play an important role in synaptic protein assembly (Hayashi et al., 2009). Therefore, creating effective antibodies will help unlock the mysteries of the PSD. The Homer1 dataset consists of 145 antibody clones against the Homer1 protein plus two known Homer1 antibodies. Each tissue sample was prepared and imaged using standard array tomography techniques (Micheva and Smith, 2007). Each antibody clone file is approximately 13.88 µm by 10.40 µm by 2.1 µm in size (or 1388 x 1040 x 3 in pixels) and consists of three channels: the target channel (a Homer1 subclone), the reference channel (PSD-95), and DAPI (used for image alignment). The objective was to run all 145 antibodies through the SACT and see if the subclones rankings would match the framework’s rankings. To determine optimal antibody performance, experts looked at both the detected synapse density, and more importantly, the target specificity ratio (TSR). The TSR is the ratio of detected synapses to detected puncta and lets the user quan-
tify the amount of ‘extraneous’ staining. Figure 3.12 presents our current results in detail. The current results match what we were expecting. The antibody clones that were manually ranked high (green) but received a poor TSR score tend to indicate alignment errors, which cause blobs to not span multiple slices.

**Figure 3.12**: Scatterplot of the homer1 antibody subclones. The subclones were ranked from 0 (worst) to 3 (best) by expert observers. The highest ranked subclones percolate to the top while the poorly ranked clones clump together in the bottom left corner of the plot.

### 3.7 Discussion

The present report introduces an effective framework for automated characterization and screening of antibodies for AT. The framework provides a number of automatically computed characteristics, such as target synapse density and target specificity.
ratio, that reflect the sensitivity and specificity of immunolabeling with a given antibody. Taken together, these computed characteristics provide an objective way to characterize and compare the performance of different antibodies against the same target, simplifying the process for selecting antibodies best suited for AT. When evaluating multiple candidate antibodies, this represents an efficient method to identify a small number of promising antibodies for further evaluation, which includes assays in knockout mouse and other relevant controls (Gong et al., 2016).

The Synaptic Antibody Characterization Tool (SACT, the implementation of the framework) is designed to be a flexible tool for antibody screening. Because it is query-based, it allows the user to define the molecular composition and size of synapses expected to contain the antigen. This flexibility is advantageous for synaptic antibody screening because the query can be designed to focus on different synapse subtypes, depending on prior biological knowledge (e.g., what combinations of proteins are likely to be present, and where the antibody target is expected to be located). Its inherent flexibility should allow this approach to be used also to validate antibodies that target other subcellular structures, ranging from the nodes of Ranvier on myelinated axons, to mitochondria, to histone marks in the nucleus. The method works with a wide selection of reference antibodies, which need not colocalize with the tested antibody. For example, antibodies to gephyrin and collybistin, both postsynaptic proteins, were evaluated using the presynaptic marker GAD as reference. The flexibility in reference antibody selection enables users to optimize the use of their available antibody stocks. With further practical experience we anticipate that a restricted group of well-characterized antibodies will be adopted as controls for each target category.

Our experiments demonstrate that SACT provides a robust method for antibody screening, ranking antibodies based on quantitative measures of their performance. In the pairwise comparisons of antibodies, there was 100% agreement between the
expert ranking and the automated antibody ranking based on target synapse density and target selectivity ratio. Variations in the size requirement did not affect the ranking, as long as synapse detection was based on more than one slice. Even when a synapse was required to be present on only one slice, performance was only modestly degraded, such that the outcome measures for some antibody pairs were ambiguous. The present approach accommodates variations in antibody concentration, as demonstrated by the experiments with multiple candidate antibodies from monoclonal antibody projects, which showed a high correlation between the ranking by algorithm and by expert evaluation of candidate antibody performance in all six datasets, even though the concentration of antibodies in the hybridoma tissue culture supernatants used for screening was unknown and intrinsically variable. This insensitivity to antibody concentration is very important in practice when evaluating multiple antibodies; by eliminating the need for immunolabeling with series of antibody dilutions, it substantially reduces the amount of work involved.

There are some limitations to the use of the proposed framework for antibody validation. This is not a stand-alone tool for generic antibody validation; it is designed to specifically address the performance of the antibody for immunofluorescence AT, and must be used along with other tests and controls. For example, SACT does not test for cross-reactivity with other proteins. A second limitation is that this approach requires prior knowledge of the expected distribution of the antigen (or some other characteristic to use for reference), especially if it is found only in a small population of synapses. In such cases it will be important to ensure that the tissue sample used for immunolabeling contains such synapses at a reasonable density and/or includes a reference marker to independently identify this population.

A number of technical issues can interfere with performance. Proper alignment of the sections in the imaged series is required to ensure that position of synapses is consistent on adjacent sections. In one of the concentration comparison experiments
with GluN1, inaccuracies in the alignment led to poor performance of the algorithm when using the standard size requirement of a synapse to be present on at least two consecutive sections. In this case, a one-slice size requirement was successfully used, but we show that this approach will not always work. To fully benefit from the advantages of using three-dimensional information from multiple serial slices, one must ensure that the datasets are well aligned. Another technical issue to consider is possible bleed-through during the fluorescent imaging, which can cause the false impression of colocalization between the tested and reference antibodies.

When carefully planned and executed to avoid pitfalls, the automated framework described here can be used to identify and characterize antibodies against a wide assortment of synaptic target proteins that yield robust and specific immunolabeling in plastic sections of brain tissue. This is particularly important because the nature of synaptic processing is still poorly understood, and many basic questions remain. For example, how many different types of synapse exist? How do these different types vary over different brain areas? How does their distribution change over time? With experience? Under pathological conditions? For questions of this nature, it is important to objectively assess a large number of individual synapses, and a large number of different molecules at each synapse, as can be done using AT. Identifying reliable synaptic antibodies for AT will remove a major limitation for such studies and allow a better understanding of synapses.
Figure 3.3: Schematic diagram of input datasets. A. The target protein, gephyrin, is a postsynaptic protein at inhibitory synapses, expected to be adjacent to GAD, a presynaptic protein abundant at inhibitory synapses. The set of squares represents a stack of images from serial ultrathin sections from mouse neocortex, double labeled with a gephyrin candidate antibody (brown dots) and a previously validated antibody against GAD (red dots). The large blue blobs represent DAPI, a marker for cell nuclei. B. Identical setup, but with Bassoon (a presynaptic protein present in excitatory synapses) as the target protein. In this example, the tissue is labeled with two reference antibodies to excitatory synapses, the presynaptic protein VGlut1 (purple dots) and the postsynaptic protein PSD-95 (red dots). C. Different combinations of puncta are detected on sections ($z=0$ to $z=3$) through a synapse. The candidate and the reference antibody can be present alongside each other on the same section (a), they can lie adjacent in the $z$-direction (b), or they can be adjacent both in the same section and across multiple sections (c). D. Identical setup as C with two reference antibodies depicted.
Figure 3.4: **Automated punctum detection pipeline.** This example illustrates the pipeline for antibody characterization. Each processed image shows the blobs/regions which have met criteria for being a punctum, and each successive panel adds a new requirement; the number of blobs considered as puncta decreases accordingly. In the final thresholded image, the blobs shown have met the requirement of spanning 3 slices and are centered on the slice shown. Other blobs which may appear ‘missing’ are centered either on the slice before or after. The first box shows raw single-label immunofluorescence from a single slice. The second box is the output of a ‘foreground probability’ step; the intensity value of each pixel encodes the probability it belongs to the foreground. The third box is the output of a ‘2D Punctum Probability’ step (each pixel coding the probability that it belongs to a 2D blob). Pixels in the 4th box display the probability that a voxel belongs to a blob which spans a minimum number of slices. The final thresholded image is shown below. The threshold is established by visual observation; for this work, the threshold was set to 0.9 for the entire project. Red scale bars represent 5 µm.
Figure 3.5: Example of erratic labeling. (Left) Immunolabeling for collybistin (associated with GABAergic synapses) on a raw IF slice; note the very broad distribution of sizes for puncta. Clusters of immunofluorescent label are detected as one large punctum; for antibodies that give such labeling pattern, the average punctum size will be large and with a large size variance. (Right) Relatively ‘normal’ pattern of immunolabeling on a raw IF slice, using a different collybistin candidate antibody. This difference is automatically quantified by computing the average three dimensional punctum size and size variance. The image on the left has an average punctum size of 124 pixels and standard deviation of 1350. The image of the right has an average punctum size of 10 pixels and standard deviation of 89 pixels. Each red scale bar is 5 μm.
Figure 3.6: Pairwise comparison of immunofluorescence on single sections from mouse brain. Each column represents an experiment where two antibodies against the same target protein (magenta) were evaluated by double labeling with a reference antibody (green). The expert’s visually-based preference is marked for each column. The sections are also labeled with the nuclear stain DAPI (blue). For each experiment, the two images are from the same section, except for the gephyrin results, where immunolabeling with the two gephyrin antibodies was performed on different sections. The SACT measurements for these images are shown in Table 3.2. Each image is 16x18 µm.
Figure 3.7: Impact of punctum size requirements on antibody comparisons. Black dots represent the higher scoring antibody. Each scatter plot shows the results of the comparison of two candidate antibodies against the same reference protein while varying the minimum synapse size requirements (queries numbered 1 through 4). The following are the minimum punctum size for the four queries. Query 1 (blue): candidate antibody - $0.2 \times 0.2 \times 0.07\mu m$, reference antibody - $0.2 \times 0.2 \times 0.07\mu m$. Query 2 (orange): candidate antibody - $0.2 \times 0.2 \times 0.14\mu m$, reference antibody - $0.2 \times 0.2 \times 0.07\mu m$. Query 3 (green): candidate antibody - $0.2 \times 0.2 \times 0.14\mu m$, reference antibody - $0.2 \times 0.2 \times 0.14\mu m$. Query 4 (red): candidate antibody - $0.2 \times 0.2 \times 0.21\mu m$, reference antibody - $0.2 \times 0.2 \times 0.07\mu m$. All queries gave consistent results for all five antibody pairs, except for Query 1 (see text for details). A TSR value of greater than one is an artifact of thresholding incorrectly splitting a punctum into two, it is remedied with simple morphological operations.
Figure 3.8: Changes in target synapse density (in synapses per cubic micrometers) and target specificity ratio as a function of antibody concentration. Each plot shows the target synapse density and target specificity ratio at different concentrations of the same antibody.
**Figure 3.9**: Comparison of multiple candidate antibodies. Each scatter plot shows the computed target synapse density and target specificity ratio of multiple candidate antibodies, with the best-ranking candidates circled. Expert ranking is color-coded: green - best, orange - unclear, red - fail. The outlier in the VGAT scatter plot was not included in the best candidate antibodies selection, because of the abnormally high synaptic density (0.7 per $\mu m^3$ compared to target max density of 0.15 per $\mu m^3$). Screening of the Bassoon project was performed in 2 rounds: the candidate antibodies identified as best or unclear in the first round were screened again with adjusted concentrations.

**Figure 3.10**: Comparison of multiple candidate antibodies using two reference synaptic antibodies. Left - Bassoon with PSD-95 and VGlut1. Right - IRSp53 with PSD-95 and VGlut1. Expert ranking is color coded: green - best, orange - unclear, red - fail. Compare with Figure 3.9.
Figure 3.11: Comparison of multiple candidate antibodies. These plots compare multiple candidate antibodies using an alternative query requiring the puncta to be present on only 1 slice, instead of 2 as in Figure 3.9. Each scatter plot shows the target synapse density and target specificity ratio of multiple candidate antibodies against the same target protein. Expert ranking is color coded: green - best, orange - unclear, red - fail. In many of these cases (gephyrin, VGAT, collybistin), the more permissive 1-slice query does not allow correct selection of the best performing candidate antibodies.
Exploring Astrocytic Contributions to Synaptic Changes in a Mouse Model of Fragile X Syndrome

4.1 Abstract

Fragile X Syndrome (FXS), a common inheritable form of intellectual disability, is known to alter neocortical synaptic connectivity. However, its impact on diverse synapse types and astrocytic processes is not well understood. Quantitatively characterizing the changes FXS causes on the synaptome will augment our ability to create targeted medications. To characterize the effect of FXS on the synaptome, we used immunofluorescent array tomography (IF-AT) to investigate different synaptic populations and their association with astrocytes in the adult somatosensory cortex of a FXS mouse model, the FMR1 knockout mouse. The multi-channel data collected from IF-AT was analyzed using an expanded version of a probabilistic synapse detector, which allowed for synapse type level analysis. We analyzed approximately 1.6 million synapses split into fifteen different synapse types. Our study reveals complex, synapse-type and layer specific changes in the neocortical circuitry of FMR1 knock-out mice. There is a significant decrease in the density of excita-
tory glutamatergic synapses and their association with astrocytes while the changes in inhibitory GABAergic synapses are less pronounced. There is a decrease in the density of large GABAergic synapses, and no changes in their association with astrocytes. Our results suggest that in Fragile X Syndrome astrocytes may mediate at least some of the pathological effects on glutamatergic synapses, while GABAergic synapses are likely influenced by a different mechanism.

4.2 Introduction

Fragile X Syndrome (FXS) is the most common inheritable form of intellectual disability, which affects approximately 1 in 7,000 males and 1 in 1,000 females across all races and ethnic groups (Hunter et al., 2014). FXS patients display a wide spectrum of phenotypes, including moderate to severe intellectual disability, autistic behavior, macroorchidism, predisposition to epileptic seizures, and facial abnormalities (Berry-Kravis, 2002; Fisch et al., 2002; Penagarikano et al., 2007). FXS is caused by the silencing of the FMR1 gene, which encodes the Fragile X Mental Retardation Protein (FMRP) known to play important roles in the translation, trafficking, and targeting of a large number of mRNAs in neurons (Jin et al., 2004; Bagni and Greenough, 2005; Bardoni et al., 2006). FMRP also binds to many proteins which suggests its involvement in a variety of cellular processes, such as genome stability regulation, cell differentiation, and ion channel gating. The unusually long and thin spines found in fixed tissues of both FXS patients (Rudelli et al., 1985; Hinton et al., 1991; Wisniewski et al., 1991; Irwin et al., 2001) and adult FMR1 knock-out (FMR1 KO) mice (Comery et al., 1997; Irwin et al., 2002; Galvez and Greenough, 2005; McKinney et al., 2005; Grossman et al., 2006) are similar to the immature spines observed during development (Dailey and Smith, 1996; Fiala et al., 1998; Zuo et al., 2005). This observation has led to a popular hypothesis that the absence of FMRP in the nervous system causes a defect in spine maturation and pruning, which in turn
alters synaptic connectivity and ultimately results in behavioral defects (Greenough et al., 2001; Churchill et al., 2002; Bagni and Greenough, 2005; Bardoni et al., 2006; Bureau et al., 2008).

While dendritic spine morphology and structural dynamics are good indicators of modifications in synaptic connectivity (Alvarez and Sabatini, 2007; Kasai et al., 2010; Sala and Segal, 2014), they cannot fully represent the molecular diversity of cortical synapses. In neocortex, most synapses fall into two basic categories: excitatory (glutamatergic, ~ 80%) and inhibitory (GABAergic, ~ 15 – 20%) (O’Rourke et al., 2012; Favuzzi and Rico, 2018). Each category comprises multiple types. For example, excitatory synapses can be distinguished based on their presynaptic vesicular glutamate transporter (VGluT) content: cortico-cortical synapses containing VGluT1 and thalamocortical synapses containing VGluT2 (Fremeau et al., 2001; Graziano et al., 2008). However, the synapse type-specific impact of FXS remains little explored (Wang et al., 2014). To add a further layer of complexity, non-neuronal cells may also contribute to FXS pathogenesis.

As the most abundant glial cells in the mammalian brain, astrocytes modulate synaptic structure and function (Verkhratsky and Nedergaard, 2014) and are implicated in many neurodevelopmental diseases (Molofsk et al., 2012). In the mouse brain, astrocytes also express FMRP (Pacey and Doering, 2007), and FMR1 KO mice have fewer hippocampal synapses associated with astrocytes (Jawaid et al., 2018). Interestingly, astrocyte-specific deletion of FMR1 leads to significantly more immature spines in the mouse motor cortex due to overproduction of spines during development (Hodges et al., 2017). Whether such astrocytic contribution varies according to synapse type is not yet known.

To better understand the synapse subtype-specific effects of FXS on the neocortical circuitry, we used immunofluorescent array tomography (IF-AT) to investigate the changes in different synaptic populations and their association with astrocytes
in the adult mouse somatosensory cortex, an area in which a variety of deficits have
been reported for FMR1 KO mice (Pan et al., 2010; Gonçalves et al., 2013; Cynthia X. He, Daniel A. Cantu, Shilpa S. Mantri, William A. Zeiger, Anubhuti Goel,
2017). IF-AT allows for the light level detection of individual synapses within large
volumes of brain tissue, and the ability to apply multiple markers for the identifi-
cation of different synaptic populations (Micheva and Smith, 2007; Micheva et al.,
2010). Synaptic density was quantified using automatic synapse detection methods
previously developed by us (Simhal et al., 2017, 2018).

4.3 Methods

4.3.1 Overview

The methods section is divided into two main components — data generation and
computational analysis. The data generation section specifies the types of mice used,
the antibodies used, and the imaging methodology. The computational analysis
section highlights the methods used to automatically analyze the array tomography
data, including the detection of synapses by their specific type and the detection of
astrocytes.

4.3.2 Data generation

The datasets investigated were obtained from the somatosensory cortex of adult mice
and represent layers one through four. The somatosensory cortex was chosen because
of the well-documented deficits in FMR1 KO mice in this cortical region. We focused
on the superficial cortical layers for which more information is available through live
imaging studies (Pan et al., 2010; Cruz-Martin et al., 2010). The average dataset
volume was $33,897\mu m^3$ with a standard deviation of $9,606\mu m^3$. 
4.3.3 Animals

FMR1 KO mice were obtained from Dr. Stephen T. Warren, Emory University. Thy1-YFP-H mice were purchased from JAX. All mice were backcrossed with C57BL/6 mice more than 10 generations to produce congenic strains. For the current experiments, YFP+ WT males were crossed with YFP- FMR1+/- females, and only male offspring littermates were used for the experiments. WT mice refer to FMR1+/-, and KO mice are FMR1-/-y. Because the YFP expression was highly variable between animals, we did not use it in the analysis. The mice were fourth months old when they were sacrificed. Further details about the mice are in Table 4.1.

Table 4.1: Mice used for the experiments and their condition. ‘WT’ refers to wild-type, ‘KO’ refers to knockout.

<table>
<thead>
<tr>
<th>Mouse specimen #</th>
<th>DOB</th>
<th>Age</th>
<th>KO or WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SS</td>
<td>6/27/17</td>
<td>4 months</td>
<td>WT</td>
</tr>
<tr>
<td>3SS</td>
<td>6/27/17</td>
<td>4 months</td>
<td>KO</td>
</tr>
<tr>
<td>4SS</td>
<td>7/04/17</td>
<td>4 months</td>
<td>WT</td>
</tr>
<tr>
<td>6SS</td>
<td>7/04/17</td>
<td>4 months</td>
<td>KO</td>
</tr>
<tr>
<td>5SS</td>
<td>7/04/17</td>
<td>4 months</td>
<td>WT</td>
</tr>
<tr>
<td>7SS</td>
<td>7/04/17</td>
<td>4 months</td>
<td>KO</td>
</tr>
<tr>
<td>2SS</td>
<td>6/27/17</td>
<td>4 months</td>
<td>WT</td>
</tr>
<tr>
<td>1SS</td>
<td>6/27/17</td>
<td>4 months</td>
<td>KO</td>
</tr>
</tbody>
</table>

4.3.4 Array tomography

The tissue was prepared using standard array tomography protocols (Micheva et al., 2010). Mice were group-housed in the UCSC animal facility, with 12 hour light-dark cycle and access to food and water ad libitum. All procedures were performed in accordance with protocols approved by the Animal Care and Use Committee (IACUC) of UCSC. The mice were anesthetized by halothane inhalation and their brains quickly removed, cut into 2mm slices, and fixed by immersion in 4% paraformaldehyde in phosphate-buffered saline (PBS) for 1 hour at room temperature, then left
in the fixative overnight at 4°C. After rinsing in PBS, the somatosensory cortex was dissected out, quenched in 50mM glycine in PBS for 30 minutes and dehydrated in a series of ethanol washes (50%, 70%, 70%) at 4°C, then infiltrated and embedded in LRWhite resin in gelatin capsules, and polymerized at 50°C for 24 hours.

To prepare ribbons of serial sections, the blocks were trimmed around the tissue to the shape of a trapezoid, and glue (Weldwood Contact Cement diluted with xylene) was applied with a thin paint brush to the leading and trailing edges of the block pyramid. The embedded plastic block was cut on an ultramicrotome (Leica Ultracut EM UC6) into 70nm-thick serial sections, which were mounted on gelatin-coated coverslips.

4.3.5 Immunolabeling

Sections were processed for standard indirect immunofluorescence, as described in Micheva et al. (2010). Antibodies were obtained from commercial sources and are listed in Table 4.2. The sections were incubated in 50 mM glycine in TBS for 5 minutes, followed by blocking solution (0.05% Tween-20 and 0.1%BSA in TBS) for 5 minutes. The primary antibodies were diluted in blocking solution as specified in Table 4.2, and were applied for 2 hours at room temperature or overnight at 4°C. After a 15 minutes wash in TBS, the sections were incubated with Alexa dye-conjugated secondary antibodies, highly cross-adsorbed (Life Technologies), diluted 1:150 in blocking solution for 30 minutes at room temperature. Finally, sections were washed with TBS for 15 minutes, rinsed with distilled water and mounted on glass slides using SlowFade Gold Antifade Mountant with DAPI (Invitrogen). After the sections were imaged, the antibodies were eluted using a solution of 0.2 M NaOH and 0.02% SDS, and new antibodies were reapplied. Several rounds of elution and re-staining were applied to create a high-dimensional immunofluorescent image.
Table 4.2: Antibodies used for the experiments.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Host</th>
<th>Antibody Source</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synapsin</td>
<td>Rabbit</td>
<td>Cell Signaling 5297</td>
<td>1:100</td>
</tr>
<tr>
<td>PSD95</td>
<td>Rabbit</td>
<td>Cell Signaling 3450</td>
<td>1:100</td>
</tr>
<tr>
<td>VGlut1</td>
<td>Guinea pig</td>
<td>Millipore AB5905</td>
<td>1:5000</td>
</tr>
<tr>
<td>VGlut2</td>
<td>Guinea pig</td>
<td>Millipore AB2251</td>
<td>1:5000</td>
</tr>
<tr>
<td>GAD2</td>
<td>Rabbit</td>
<td>Cell Signaling 5843</td>
<td>1:100</td>
</tr>
<tr>
<td>Gephyrin</td>
<td>Mouse</td>
<td>NeuroMab 75-443</td>
<td>1:100</td>
</tr>
<tr>
<td>Glutamine synthetase</td>
<td>Mouse</td>
<td>BD Biosciences 610517</td>
<td>1:25</td>
</tr>
</tbody>
</table>

4.3.6 Imaging method

The immunostained ribbons of sections were imaged on an automated epifluorescence microscope (Zeiss AxioImager Z1) using a 63x Plan-Achromat 1.4 NA oil objective. To define the position list for the automated imaging, a custom Python-based graphical user interface, MosaicPlanner (obtained from https://code.google.com/archive/p/smithlabsoftware/), was used to automatically find corresponding locations across the serial sections. Images from different imaging sessions were registered using a DAPI stain present in the mounting medium. The images from the serial sections were also aligned using the DAPI signal. Both image registration and alignment were performed with the MultiStackReg plugin in FIJI (Schindelin et al., 2012).

4.3.7 Computational analysis

A main goal of this analysis is to examine the effects of the lack of FRM1 gene on the synaptic composition of the somatosensory cortex. This requires the ability to quantify synapses by their proteometric composition and their adjacency to an astrocytic process. To achieve this, we took existing methods and expanded their scope to meet the computational challenges posed by these experiments, including developing a method for detecting astrocyte processes adjacent to synapses.
4.3.8 Synapse detection

For the present purposes we define ‘synapse type’ as a specific combination of synaptic proteins. For example, a GABAergic (inhibitory) synapse type is defined by the presence of the general presynaptic marker synapsin, the postsynaptic marker of inhibitory synapses gephyrin, and the presynaptic marker of inhibitory synapses GAD. A glutamatergic (excitatory) synapse is defined by the presence of the general presynaptic marker synapsin and the postsynaptic marker for excitatory synapses PSD-95. A glutamatergic synapse with VGlut2 and adjacent to an astrocytic process is defined by the presence of synapsin, VGlut2, PSD-95, and GS, a marker for astrocytes.

Detecting synapses by their proteometric composition is the first step of the computational pipeline. In order to quantitatively analyze large array tomography volumes, it is vital to find an appropriate synapse detection technique. The majority of published synapse detection methods use traditional machine learning approaches (Busse and Smith, 2013; Kreshuk et al., 2014; Collman et al., 2015; Bass et al., 2017). These approaches all consist of a few common steps to detect synapses. First, for each synapse type, a large number of synapses are manually identified and labeled in the array tomography data. Next, a classifier (such as a support vector machine or convolutional neural network) is trained with these manual annotations. Lastly, the entire dataset is appropriately parcellated and potential synapses are labeled by the classifier. While this method works well for certain questions in synapse biology, the difficulty in manually labeling different synapse types in immunofluorescent data renders it ineffective for our applications.

The probabilistic synapse detection method introduced in Simhal et al. (2017), is a synapse type focused approach which does not require any training data, making it a viable option for exploring synapses imaged via array tomography. ‘Synapse
type' focused means the user specifies the proteometric composition and the relative spatial arrangement along with the size of the synaptic markers prior to running the probabilistic synapse detection method. The combination of a user-defined synapse type and marker size is called a ‘query,’ as highlighted in Figure 4.1.

A query to detect a glutamatergic synapse would look like the following: a PSD-95 punctum of a minimum of $2px \times 2px \times 2slices$ (which is $0.2\mu m \times 0.2\mu m \times 0.14\mu m$ for our data) adjacent to a synapsin punctum of the same size. Adjacency in this case means that the puncta of the two different antibody markers do not occupy the same space but instead are juxtaposed from each other. Since PSD-95 is a postsynaptic protein and synapsin is presynaptic protein, this simple glutamatergic synapse query follows the known biological model for a glutamatergic synapses. In Simhal et al. (2017), the query comprises only of presynaptic and postsynaptic makers. In this work, we have expanded the query to comprise of presynaptic, postsynaptic, and astrocytic markers. This combination of markers is often referred to as a ‘tripartite synapse’ (Araque et al., 1999). As the left side of Figure 4.1 shows, the tripartite synapse model assumes that the proteometric markers for each ‘subclass’ (presynaptic, postsynaptic, astrocytic) lie adjacent relative to each other.

In summary, the probabilistic synapse detector is a method of detecting specific synapse types. Instead of requiring the user to manually annotate multiple instances of a synapse type to train a machine learning classifier, the query-based approach asks the user to define a synapse by specifying basic characteristics, that is, the requisite markers, requisite punctum volume for each marker (which depends on the microscope known resolution), and their relative spatial arrangement.

Once the query has been established, the probabilistic synapse detection method follows the query to automatically detect synapses matching the query. Figure 4.2 shows an example pipeline going from the raw input data to the result probability map. The details of the synapse detection method are described in Simhal et al.
What is a query? The cartoon on the left side shows the relative spatial arrangement of the different fluorescent markers used to detect an excitatory synapse expressing VGluT1, next to an astrocyte process. This visual description of a synapse is translated into a query, shown to the right of the large blue arrow. A query is a user-defined description of what the synapse type of interest should ‘look’ like. In this case, the presynaptic protein markers - synapsin and VGluT1, are expected to colocalize (occupy the same 3D space) with each other. Furthermore, the presynaptic, postsynaptic, and astrocyte markers (as a group) are all expected to be next to each other. The top right portion of the figure shows three 1.5 × 1.5 \( \mu m \) cutouts of different marker combinations of what the query looks like in the data. The first cutout shows the synapsin and PSD-05 punctum overlaid; the second cutout includes the GS punctum and the third cutout includes the VGluT1 punctum.

(2017) and its applications to antibody characterization are studied in Simhal et al. (2018). Both report extensive validation, rendering the tool to be exploited to address the novel biological questions in this work.

4.3.9 Synapse type definitions

For the analysis presented in this work, we used the queries listed in Table 4.3. A synapse of a particular type is defined as having all the relevant markers and that the markers are of the specified size. For this study, we require the markers to span one or more slices, depending on the desired synapse size, and have a minimum \( x, y \) size of 0.2\( \mu m \times 0.2\mu m \). The one exception is the definition of VGluT2 synapses where the VGluT2 marker is required to span two or more adjacent slices. This is due to the
Figure 4.2: Probabilistic synapse detection pipeline The first column shows the raw PSD-95 and synapsin data. The second column shows the output of the synapse detection method, where the value at each pixel is the probability that pixel belongs to the specified definition of a synapse. The third column shows the detections (in white) overlaid upon the superposition of the PSD-95 and synapse data. For this visualization, the definition of a synapse was the adjacency of a PSD-95 and synapsin punctum of a minimum size of $0.2 \mu m \times 0.2 \mu m \times 0.07 \mu m$ which corresponds to $2px \times 2px \times 1slice$.

Properties of the VGluT2 antibody, which in addition to the expected robust label of a synapse subpopulation, also gives higher, randomly distributed, background signal.

Each synapse type was further subdivided in three sizes: small, medium, and large. A small synapse is defined as spanning one slice, a medium synapse is defined as spanning two slices, and a large synapse is defined as spanning three or more slices. To calculate the synapse density of synapses that only span a single slice, a query where synapses span two or more slices is subtracted from a query where synapses span one or more slices. In the same vein, to calculate the synapse density of synapses that span only two slices, a query where synapses span three or more slices is subtracted from a query where the synapses span two or more slices. 'All
synapses’ of a type are defined as having punctum markers that span one or more slices.

Table 4.3: Queries used for this analysis.

<table>
<thead>
<tr>
<th>Query Name</th>
<th>Presynaptic Markers</th>
<th>Postsynaptic Markers</th>
<th>Astrocyte Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory 1</td>
<td>Synapsin</td>
<td>PSD-95</td>
<td>None</td>
</tr>
<tr>
<td>Excitatory 2</td>
<td>Synapsin, VGluT1</td>
<td>PSD-95</td>
<td>None</td>
</tr>
<tr>
<td>Excitatory 3</td>
<td>Synapsin, VGluT2</td>
<td>PSD-95</td>
<td>None</td>
</tr>
<tr>
<td>Excitatory 4</td>
<td>Synapsin, VGluT1, VGluT2</td>
<td>PSD-95</td>
<td>None</td>
</tr>
<tr>
<td>Inhibitory 1</td>
<td>Synapsin, GAD</td>
<td>Gephyrin</td>
<td>None</td>
</tr>
<tr>
<td>Excitatory-astrocyte 1</td>
<td>Synapsin</td>
<td>PSD-95</td>
<td>GS</td>
</tr>
<tr>
<td>Excitatory-astrocyte 2</td>
<td>Synapsin, VGluT1</td>
<td>PSD-95</td>
<td>GS</td>
</tr>
<tr>
<td>Excitatory-astrocyte 3</td>
<td>Synapsin, VGluT2</td>
<td>PSD-95</td>
<td>GS</td>
</tr>
<tr>
<td>Excitatory-astrocyte 4</td>
<td>Synapsin, VGluT1, VGluT2</td>
<td>PSD-95</td>
<td>GS</td>
</tr>
<tr>
<td>Inhibitory-astrocyte 1</td>
<td>Synapsin, GAD</td>
<td>Gephyrin</td>
<td>GS</td>
</tr>
</tbody>
</table>

4.3.10 Volume calculation

Synapse density is calculated as the number of synapses detected in a dataset over the volume of the dataset. The volume of the dataset is defined by the volume of the neuropil (i.e., excluding the volume occupied by cell nuclei and large blood vessels, which can vary significantly between areas). When choosing the areas to image, we avoided large blood vessels. To calculate the volume of cell nuclei, the nuclear stain DAPI was converted into probability space using the methods outlined in Simhal et al. (2018). Briefly, the value at each pixel in probability space is the probability it belongs to the foreground, with a range of 0 to 1. To do so, the background noise is modeled and the foreground probability is one minus the probability the pixel belongs
to the background. The background noise is modeled as a Gaussian, for which the mean and variance is calculated from the raw data itself. Once the probability map is calculated, it is thresholded (t=0.6, chosen by observation) and cleaned up by a sequence of morphological operations. The code is available for download on the project’s website. In summary, the volume of the neuropil was obtained by subtracting the volume of the DAPI stained nuclei from the total imaged volume.

4.3.11 Statistical analysis

Statistical analysis to determine significance between the two populations was done via a two-tailed unpaired t-test.

4.3.12 Data and code availability

The code and raw data are available for download at github.com/aksimhal/fxs.

4.4 Results

We used immunofluorescent array tomography (IF-AT) to quantify the synaptic density, composition and glial involvement in layers 1 through 4 of the somatosensory cortex of adult FMR1 knockout (KO) mice and wild-type (WT) mice. IF-AT is based on digital reconstruction of images acquired from arrays of serial ultrathin sections (70 nm) attached to coverslips, immunofluorescently labeled and imaged under a fluorescence microscope. The use of ultrathin sections allows the light level detection of individual synapses, while the possibility of applying multiple immunofluorescent markers (10 or more) enables the identification of different synaptic populations (Micheva and Smith, 2007). Synaptic density was quantified using automatic synapse detection methods previously developed by us (Simhal et al., 2017, 2018).
4.4.1 Overview of the datasets and detected synapses

Volumes of approximately $140 \times 400 \times 3\mu m$ spanning layers 1 through 4 of the somatosensory cortex of FMR1 KO mice and WT mice were imaged, as shown in Figure 4.3. We detected an average of 200,000 synapses in each volume for a total of approximately 1.6 million synapses across all eight datasets. Excitatory synapses were identified by the presence of immunofluorescent signals from both synapsin, a presynaptic protein, and PSD-95, a protein of the postsynaptic scaffold of excitatory synapses. Excitatory synapses were further subdivided depending on their vesicular glutamate transporters into VGluT1 positive, generally thought to be of intracortical origin, and VGluT2 positive, belonging predominantly to thalamocortical inputs Fremeau et al. (2001); Kaneko and Fujiyama (2002); Graziano et al. (2008). Inhibitory synapses were identified by the presence of the general presynaptic marker synapsin and the presynaptic marker for GABAergic synapses, glutamic acid decarboxylase (GAD), together with the postsynaptic marker gephyrin. Astrocytes, including their processes, were detected using an antibody against glutamine synthetase (GS) (Martinez-Hernandez et al., 1977; Anlauf and Derouiche, 2013), which allowed the identification of the fraction of synapses that are immediately adjacent to astrocytic processes. In addition to identifying synapses based on combinations of different markers, synapses were also analyzed based on their size, with small synapses identified as having the relevant markers on only one slice, medium synapses — on two consecutive slices, and large synapses — on three or more consecutive slices. Visual inspection of the datasets did not uncover any obvious differences in immunofluorescence intensity and pattern for any of the markers between the KO and WT mice. The cortical thickness was also comparable between the two conditions ($0.88 \pm 0.03 mm$ for KO vs. $0.89 \pm 0.02 mm$ for WT, $p = 0.71$).
4.4.2 Single channel analysis

Quantification of the density of the puncta of the different synaptic markers did not reveal any statistically significant differences ($p > 0.05$). There was a tendency for a decrease in the number of small VGlurT2 puncta ($p = 0.065$) and large PSD95 puncta ($p = 0.065$). On the other hand, medium PSD-95 puncta tended to increase ($p = 0.056$), as shown in the middle panel of Figure 4.3.

The only significant change that was detected was in the density of GS puncta of different sizes: in FMR1 KO mice there were more small puncta ($p = 0.01$), as well as a tendency for a decrease in the number of large GS puncta ($p = 0.07$), as shown in the bottom panel of Figure 4.3.

Even though we did not detect any significant changes in the densities of puncta of the different synaptic markers, this does not necessarily preclude changes in the synaptic populations of FMR1 KO mice. The synaptic proteins assessed in our study are indeed highly enriched at synapses, but they are also found at extra-synaptic sites, and thus their presence does not necessarily equate to the presence of a synapse.
**Figure 4.3: Overview of the datasets.**

A, Immunofluorescent array tomography of wild-type and FMR1 knockout mouse somatosensory cortex. The left panel shows the imaged area in each sample which consists of four tiles spanning cortical layers 1 through 4. DAPI staining of nuclei, volume reconstruction of 30 serial sections, 70 nm each. The right panel shows the immunofluorescence for synaptic (synapsin, VGluT1, VGluT2 and GAD2) and glial (GS) markers in wild-type and knockout mouse somatosensory cortex layer 4, volume reconstruction of 10 serial sections, 70 nm each.

B, Summary of single channel punctum density changes between wild-type and knockout mice with all layers averaged.

C, This plot shows the density distribution of GS puncta by size. ‘Small’ puncta span one slice, ‘medium’ puncta span two slices, and ‘large’ puncta span three or more slices.
4.4.3 Overall synapse densities

A much more accurate detection of synapses is achieved by using combinations of synaptic markers, ideally at least one presynaptic and one postsynaptic marker, as used by our synapse detection algorithm (Simhal et al., 2017, 2018). Indeed, using such combinations of synaptic markers, the detected synapse densities and distributions in WT mice are consistent with previous estimates as shown in Figure 4.4. The overwhelming majority of cortical synapses are known to be either excitatory glutamatergic or inhibitory GABAergic synapses (DeFelipe et al., 2002). Thus, the total density of synapses was estimated by the sum of the densities of the detected glutamatergic (synapsin + PSD95 markers) and GABAergic synapses (synapsin + GAD + gephyrin markers) resulting in approximately 1.94 synapses per $\mu m^3$ of embedded tissue. Because tissue dehydration and embedding with our protocol results in approximately 23% linear shrinkage, or 54% volumetric shrinkage (Busse and Smith, 2013), this equals to 0.9 synapses per $\mu m^3$ of unprocessed tissue, very similar to the reported synapse density in mouse cortex (Schüz and Palm, 1989).

The relative contributions of inhibitory and excitatory synapse types that we detect is consistent with the known synapse composition of mouse cortex. The percent of inhibitory GABAergic synapses in our detections varies between 12% in layer 1 to 20% in layer 4 (Figure 4.4F). EM counts in mouse somatosensory cortex show that inhibitory synapses consist of 11% of the synapses in layer 1 (Cali et al., 2018) and 18% of the synapses in layer 4 (Knott et al., 2002).

The layer distribution of VGluT2 synapses is consistent with their known preference for layer 4 (Figure 4.4G). VGluT2 is known to label thalamocortical synapses which target mostly layer 4 and lower layer 2/3 (Koralek et al., 1988; Meyer et al., 2018). Thalamocortical synapses, identified either by degeneration techniques, anterograde transport of lectin, or VGluT2 immunostaining, have been shown to com-
prise approximately 20% of glutamatergic synapses in layer 4 of mouse somatosensory cortex (White, 1978; Keller et al., 1985; Bopp et al., 2017) and we indeed see 21% in layer 4.

To further verify the accuracy of our detections, we used a different way to calculate the density of excitatory glutamatergic synapses. These synapses can be subdivided into two major populations depending on the vesicular glutamate transporters (VGluTs) at the presynaptic site, with the majority of excitatory synapses containing VGluT1, and a smaller population, mostly concentrated in layer 4, VGluT2. In addition, some synapses express both VGluT1 and VGluT2 (Nakamura et al., 2008). Thus, the density of glutamatergic synapses should be approximately equal to the densities of VGluT1 and VGluT2 synapses minus VGluT1/VGluT2 synapses to prevent double counting of the same synapse. This was indeed the case, as seen in Table 4.4.

Table 4.4: Density distribution of excitatory synapses across layers as calculated by two methods.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Glut</th>
<th>VGluT1+VGluT2-VGluT1/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>1.617</td>
<td>1.629</td>
</tr>
<tr>
<td>L2/3</td>
<td>1.700</td>
<td>1.750</td>
</tr>
<tr>
<td>L4</td>
<td>1.574</td>
<td>1.585</td>
</tr>
</tbody>
</table>
Figure 4.4: Overview of wild-type synapse density distributions. A-D, The colored images show sample cutouts of array tomography data. The white circles represent a possible synapse. Note that a synapse is defined as the localization of multiple relevant synaptic markers. E, The distribution of synapse types across different layers. Layers with significant differences for a synapse type are marked. While there is no significant difference in layer densities for glutamatergic synapses overall and for glutamatergic synapses with VGluT1, there is a significant difference between layers for glutamatergic synapses with VGluT2, with both VGluT1 and VGluT2, and for inhibitory synapses. F, Fraction of inhibitory synapses in layers 1 through 4 of mouse somatosensory cortex. G, Fraction of VGluT2 synapses in layers 1 through 4 of mouse somatosensory cortex.
4.4.4 Changes in synaptic densities in FMR1 KO mice

Next, we compared the densities of the different synapse populations in the WT mice to the FMR1 KO mice. Even though the individual synaptic marker puncta did not show any statistically significant differences in the two conditions, there were wide-ranging changes in synaptic densities. These changes were dependent on the synapse type, size, as well as cortical layer (Figure 4.5). There was an increase of small glutamatergic VGluT1 synapses in layer 4 accompanied by a decrease in large VGluT1 synapses in layers 1 and 4. VGluT2 synapses, on the other hand, showed a rather consistent decrease in density, regardless of their size, with the majority of changes occurring in layers 1 and 2/3, but not layer 4 (except for large VGluT1/2 synapses). Large inhibitory synapses decreased across all layers examined without detected changes in small and medium size inhibitory synapses.

These changes in density of the various synaptic populations resulted in a decrease in the Excitation-Inhibition ratio in the FMR1 KO mice: $6.20 \pm 0.56$ for the KO vs. $5.24 \pm 0.32$ in WT (average for layers 1 through 4; $p = 0.04$) (Figure 4.6G).

Because the strength of excitatory synapses is known to be proportional to the size of the postsynaptic density, we also analyzed the changes specifically at the postsynaptic side (Figure 4.6A-F). There was a statistically significant increase in the density of glutamatergic and specifically VGluT1 synapses with small PSD-95 puncta (spanning only 1 or 2 slices), and a significant decrease in the density of the VGluT2 synapses with large PSD-95 puncta (spanning 6 or more slices). No changes in the densities of inhibitory synapses depending on the size of gephyrin puncta were detected.
Figure 4.5: Changes in synapse densities across layers. A, Summary table showing the synapse types that have a significant increase or decrease in synapse density between wild-type and knockout mice. B-D, Differences in synapse density for different synapse types for three expected synapse sizes.
**Figure 4.6:** Postsynaptic marker size distributions. A, Summary table showing the differences in the size distributions of PSD-95 puncta associated with synapses by type. The main takeaway is that for excitatory synapses overall, there is an increase in the number of small PSD-95 puncta in FMR1 knockout mice while there is a decrease in the number of very large PSD-95 puncta associated with synapses containing VGluT2. B-E, Plots showing the distribution of PSD-95 puncta for both wild-type and knockout mice. F, Size distribution of gephyrin puncta associated with inhibitory synapses. G, Plot shows the significant increase in the ratio of excitatory to inhibitory synapses in knockout mice.
4.4.5 Involvement of glia (astrocytes)

Astrocytes are intimately involved in synaptic function and their processes are found adjacent to many synaptic clefts in the neocortex. Because astrocytes in the mouse also express FMRP (Pacey and Doering, 2007) and are suspected to have a role in FXS pathogenesis (Hodges et al., 2017), we analyzed the potential changes in astrocytic involvement at synapses in the FMR1 KO mice. Astrocytes were detected using antibodies to glutamine synthetase, an enzyme known to be expressed predominantly by this cell type (Martínez-Hernández et al., 1977; Anlauf and Derouiche, 2013) and specifically found at the peripheral astrocytic processes that contact synapses (Reichenbach et al., 2010). Figure 4.7A,B show an example of the data.

In wild-type mice, we found that the majority of glutamatergic synapses (72±2%) are adjacent to astrocytic processes as detected by immunolabel to glutamine synthetase. This is very similar to previous EM estimates in mouse somatosensory cortex layer 4, where 68% of glutamatergic synapses on dendritic spines were in contact with astrocytic processes at the bouton-spine interface (Genoud et al., 2006). Interestingly, we observed significant differences in the astrocytic association of the different synapse types, as shown in Figure 4.7C, D. Thus, compared to the majority glutamatergic synapses containing VGluT1 (61±2% astrocytic association), significantly less VGluT2 synapses (46±1%, p < 0.001) and GAD synapses (29±1%, p < 0.001) were adjacent to astrocytes. There are no previous estimates about the glial association of VGluT2 or inhibitory synapses.

Comparison of WT with KO mice revealed a number of significant changes in astrocytic involvement at synapses. Consistent with the detected overall changes in synaptic density, there were significant decreases in the densities of synapses adjacent to astrocytes, for almost all synapse types and sizes as shown in Figure 4.8A. Only layer 4 small glutamatergic synapses adjacent to astrocytes increased in density in KO
mice. While the density of synapses adjacent to astrocytes is very much influenced by the changes in overall synaptic density, the fraction of synapses adjacent to astrocytes reflects the actual changes in glial involvement in Fragile X. A significant decrease in the fraction of synapses adjacent to astrocytes was detected for glutamatergic and VGluT1 synapses of all sizes, except small VGluT1 synapses (layers 1 - 4 combined) as shown in Figure 4.8B. VGluT2 synapses (medium size) and GAD (medium and large) also tended to decrease. When analyzed by cortical layer, there were significant decreases in the fraction of astrocytic association for large glutamatergic synapses in layers 1 and 2/3, large VGluT1 synapses in layer 1, small and medium VGluT2 synapses in layers 2/3, and medium GAD synapses in layer 1. Interestingly, no changes in the astrocytic association for any of the synapse types and sizes were detected in layer 4.
4.5 Discussion

Using immunofluorescent array tomography and automatic probabilistic synapse detection methods we showed wide-ranging changes of synapses and their association with astrocytes in the somatosensory cortex of adult FMR1 knock-out mice, a Fragile X mouse model. Overall, there is a significant decrease in the density of excitatory glutamatergic synapses and their association with astrocytes, without significant changes in inhibitory GABAergic synapses. However, the changes vary greatly, and are at times in opposite directions, depending on synapse type, size, as well as cortical layer. The changes in supragranular layers (layers 1 and 2/3) reflect the overall decrease in the density of glutamatergic synapses, both VGluT1 and VGluT2 type. Meanwhile in the granular layer (layer 4) there is a significant increase in the density of glutamatergic synapses, mostly due to an increase in small VGluT1 synapses, and no significant change in VGluT2 synapses. No changes in the astrocytic association of synapses are seen in layer 4, while in the supragranular layers a significantly lower fraction of glutamatergic VGluT1 synapses are associated with astrocytes. As for the inhibitory GABAergic synapses, the only change detected is an overall decrease in the density of large synapses, and a decrease in the astrocytic association of medium-sized synapses in layer 1. Overall, these changes result in an increased Excitation/Inhibition ratio in FMR1 knock-out mice. Thus, the absence of FMRP markedly alters the neocortical synaptic circuitry by both changing the relative contributions of synapses of different types, and the astrocytic involvement at synapses.

Our results are consistent with reported data in adult FMR1 knock-out mice suggesting cortical circuit changes such as increase in smaller size synapses and decrease in larger size synapses (as evidenced by the spine size in Ishii et al. (2018); Jawaid et al. (2018)), as well as decreased association of astrocytes with hippocampal synapses (Jawaid et al., 2018). We have extended these observations, by showing that
these changes are not uniform, but depend on the synapse type, as well as cortical layer. A previous study had indeed shown that layers 4 and 5 synapses of different types in mouse somatosensory cortex exhibit various deficits in FMR1 knock-out mice (Wang et al., 2014) and we have now characterized the synapse type specific changes in the supragranular layers as well. Finally, we are showing for the first time that the changes in astrocytic involvement at synapses specifically affect excitatory glutamatergic synapses, and to a much lesser extent inhibitory GABAergic synapses in supragranular layers of mouse somatosensory cortex, with no detectable effect on layer 4 synapses.

Previous studies of Fragile X syndrome have mostly focused on the alterations occurring in the synaptic circuitry, but we now showed that there are specific deficits in cortical astrocytes and their interactions with synapses. Interestingly, these changes affect almost exclusively glutamatergic synapses. Indeed, it has been shown that glutamatergic but not GABAergic neurons, critically depend on the presence of glia to establish synaptic transmission (Turko et al., 2018). It thus appears that in Fragile X Syndrome astrocytes may mediate at least some of the pathological effects on glutamatergic synapses, while GABAergic synapses are likely influenced by a different mechanism.

Overall, our study reveals complex, synapse-type and layer specific changes in the neocortical circuitry of FMR1 knock-out mice. Indeed, because the product of the FMR1 gene, the Fragile X Mental Retardation Protein (FMRP) interacts with a large number of mRNAs, as well as proteins, its absence is expected to cause profound deficits at multiple cellular and circuit levels. As shown by our study, the ability to dissect the circuit deficits by synapse type, as well as astrocytic involvement, will be crucial for understanding how these changes affect circuit function, and ultimately define targets for therapeutic treatment and prevention.
Figure 4.7: Astrocytic synapse densities in wild-type mice. A, The image on the left is a portion of the wild-type synapse data; the image on the right is a zoomed version of the left image with two glutamatergic synapses adjacent to astrocytic processes circled in white. Last row, left: the density of synapses associated with an astrocytic process decreases dramatically for all synapse types across all layers. C, There are significant differences in the fraction of synapses of different types that are associated with astrocytic processes. D, Between layer differences in the astrocytic association of different synapse types.
Figure 4.8: Summary of the astrocytic synapse density differences in the knockout mice. A, Changes in the density of synapses adjacent to astrocytic processes. The density decreases between the wild-type and knock-out mice for Layers 1-3 and the difference is especially pronounced for medium and large synapses. B, Changes in the fraction of synapses adjacent to astrocytic processes.
Conclusion

The previous three chapters detail our contributions towards creating and validating new tools to explore synapse anatomy via array tomography. In Chapter 2, we detail a method for synapse detection in light microscopy data that is easily adaptable and expandable for detecting specific synapse types and requires no manual annotations for training. In Chapter 3, we detail a method for quantitatively characterizing new antibodies for array tomography. In Chapter 4, we expand the probabilistic synapse detection method for tripartite synapses and explore the differences in synapses between wild-type and FMR1 knockout mice. This analysis lead to the discovery of several new effects of the FMR1 gene on astrocytic synapse density.

Alongside our efforts to create new tools for exploring array tomography data is an effort to disseminate the tools to the broader neuroscience community. To that end, we open-sourced the probabilistic synapse detection method as both a python and MATLAB tool and created a website, https://aksimhal.github.io/SynapseAnalysis/SynapseDetection. This website contains both the code and examples needed to install and run the method. The SACT has its own website (https://aksimhal.github.io/SynapseAnalysis/) and also includes the code and
examples needed to install and run the tool. Furthermore, we traveled to a series of universities to share these tools with our collaborators and show them how to use it and adapt it for their own research. These tools are now installed in labs at the following universities, aside from Duke University: University of California (UC) San Francisco, UC Davis, Columbia University, Johns Hopkins University, and the University of North Carolina Chapel Hill.

As the next steps for this project, we have a series of proposals that will both further validate and expand the frameworks presented in this dissertation. First, we plan on expanding the synapse detection method to detect synapses that are not only adjacent to an astrocytic process but also determine whether they originate or terminate on a YFP filled neuron. This would allow neuroscientists to study synapses by their subtype and group them by which neuron they originate or terminate from. This would provide a new framework for clustering synapses and lead to new insights about the impact of neurodegenerative diseases, which have been known to affect the size, shape, and density of dendrites. We already have unpublished work which demonstrates use of the probabilistic synapse detector in determining whether or not a synapse terminates upon a YFP filled dendrite for a few datasets, but we have yet to collect enough YFP data to test this method at scale. The Fragile X study in Chapter 4 highlights a framework for comparing synapses for two classes. For our next steps, we hope to expand this framework and conduct more studies comparing synapse populations between mice and human samples. Currently, very few treatments that are developed using mice models models translate to human models. Finally, we hope to develop FIJI plugins for both the probabilistic synapse detector and synaptic antibody characterization tool to make it easier for researchers to use.

Synaptomics is still a relatively young and difficult field of study. Every imaging modality appears to have as many drawbacks as benefits. Antibodies are notoriously
Finicky. Datasets are too large to be analyzed or even visualized properly on a ‘traditional’ laptop. Synapses are hard to both define and identify, and every publication has a varying definition of a synapse. Furthermore, synapses do not exist in a vacuum and thoroughly studying synapses requires also studying the ‘neighborhood’ in which they reside, including astrocytes and mitochondria. However, studying synaptomics remains an extremely promising avenue for better understanding synaptopathies, which devastate the lives of many individuals around the world. We hope our computational tools and research serve a small but useful contribution to the community.
Bibliography


transporters 1 and 2 at single axon terminals during postnatal development of mouse neocortex: a quantitative analysis with correlation coefficient,” European Journal of Neuroscience, 26, 3054–3067.


to micro-scale connectome: array tomography and mGRASP,” Frontiers in neuro-
anatomy, 9, 78.

ics of perisynaptic glia,” Brain Research Reviews, 63, 11–25.

Rhodes, K. J. and Trimmer, J. S. (2006), “Antibodies as valuable neuroscience re-
search tools versus reagents of mass distraction,” Journal of Neuroscience, 26,
8017–8020.

Roncal, W. G., Pekala, M., Kaynig-Fittkau, V., Kleissas, D. M., Vogelstein, J. T.,
“VESICLE: Volumetric evaluation of synaptic interfaces using computer vision at

Rudelli, R. D., Brown, W. T., Wisniewski, K., Jenkins, E. C., Laure-Kamionowska,

Sala, C. and Segal, M. (2014), “Dendritic Spines: The Locus of Structural and
Functional Plasticity,” Physiological Reviews, 94, 141–188.

of Comparative Neurology, 493, 477–478.


Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J.,
open-source platform for biological-image analysis.” Nature methods, 9, 676–82.

cortex of the mouse,” Journal of Comparative Neurology, 286, 442–455.

Shanks, N., Greek, R., and Greek, J. (2009), “Are animal models predictive for
humans?” Philosophy, Ethics, and Humanities in Medicine, 4, 2.

Spring Harbor perspectives in biology, 3, a005678.


Biography

Anish K. Simhal grew up in Virginia and attended high school at James W. Robinson Secondary School. He attended the University of Virginia in Charlottesville, VA and graduated with a Bachelors of Science in Electrical Engineering with a minor in physics. He graduated with a PhD in Electrical Engineering from Duke University in Durham, NC. At Duke, his research focused on creating computational tools to detect and characterize synapses. During his time in Durham, Anish coached the Durham Boys and Girls Club Little League team in the Durham Bulls Youth Athletic League for three seasons from 2016 to 2018. In 2018, Anish founded a local nonprofit, Bull City Classrooms, to help Durham Public School teachers in their classrooms.