UNIVERSITY OF CALIFORNIA, SAN DIEGO

Characterizing the Spatial Density Functions of Neural Arbors

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in
Neurosciences with a specialization in
Computational Neuroscience

by

Corinne Michelle Teeter

Committee in charge:

Professor Charles Stevens, Chair
Professor William Kristan, Co-Chair
Professor Edward Callaway
Professor Lawrence Saul
Professor Virginia de Sa

2010

The dissertation of Corinne Michelle Teeter is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

Co-Chair

_____

Chair

University of California, San Diego

2010

DEDICATION

I dedicate this thesis to my parents.
Without their unconditional and unwavering support,
I would have never achieved such heights.

EPIGRAPH

We shall not cease from exploration and the end of all our exploring will be to arrive where we started... and know the place for the first time.

T. S. Eliot

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

adf	Arbor density function
AUC	Area under the curve
DSCAM	Down syndrome cell adhesion molecule
IPL	Inner plexiform layer
PCA	Principal components analysis
RGC	Retinal Ganglion cell
RF^{TM}	Random Forest^{TM}
ROC	Receiver operator characteristic

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This thesis would not have been possible without the generous ideas, time, support and encouragement of my advisor Chuck Stevens. I also thank the members of my thesis committe: Ed Calloway, Virginia DeSa, William Kristan, Lawrence Saul, for their time, input and support. I would like to extend a special thanks to William Kristan and Andrew McCulloch for all of their advice, guidance and support over the years. I also thank Anirvan Gosh for his support and for being an excellent program chair.

I thank Joseph Snider, Thomas Nowotny, Julie Haas, Rafael Levi, and Marcelo Reyes who provided scientific mentorship during their postdoctoral positions.

Philip Meier, thank you for many conversations that have lead to the advancement of my scientific ideas. In addition, thank you for the technical and computational guidance and support you have provided.

Thank you to the members of my thesis laboratory for valuable scientific conversations: Bayle Shanks, Nikoosh Carlo, and Trygve Bakken. Also, thank you Hoang Nhan for answering my histology questions.

For their contribution to my development as a scientist, I thank my past mentors: Oscar Vilches, Julia Heiman, Steven Van Sciver, Anna Hayes, William Wilson, John George, Al Selverston, Henry Abarbanel, and Gert Cauwenburghs.

I extend a special thanks to Vitaly Effimov who inspired me to study physics and therefore has played a lasting and sacred role in my life. In addition to being a successful theoretical physicist, Vitally is a very kind and supportive teacher who cares a great deal about teaching and students. His class demonstrations and personal stories about his past life experiences as a student and physicist were memorable. Despite the large number of students he taught (classes with 100+ students), he somehow, without asking, knew my name. I am forever grateful for the respect he shows his students, his thoughtfulness as a teacher, and his inspiration.

I thank all the people who contributed to the data used in this thesis and to those who were involved in the effort to ensure its public access in the interest of ad-

vancing science. Without them, this project would not have been possible. A special thanks to those who took the time to answer my often long and detailed questions about their data: Sharon Furtak, Guillermo Gonzalez-Burgos, and Patrick Hof, Bob Jacobs, Zoltan Kisvarday, Craig McDonald, Dirk Schubert, Gordon Shepherd, Armen Stepanyants, Jochen Staiger, and Rafael Yuste. I also thank Robbin Gibb who did not contribute data to the website, but answered my questions concerning a histology protocol she helped develop which was referenced by several data contributers.

Finally, I thank all my friends and family for their support. Your contribution to my thesis has been every bit as important as the scientific support I have received.

Thank you all again.

VITA

2010	Ph. D. in Neurosciences, with a specialization in Computational Neuroscience University of California, San Diego
2009	Complex Systems Summer School Santa Fe Institute, Santa Fe, NM
2007	Okinawa Computational Neuroscience Summer Course Okinawa, Japan
2004–2006	Integrative Graduate Education and Research Traineeship (IGERT) scholar
2003–2004	Post Baccalaureate Research in Biophysics Los Alamos National Laboratory
2002–2003	Post Baccalaureate Research in Theoretical Astrophysics Los Alamos National Laboratory
2002	B. S. in Physics University of Washington
2002	B. S. in Psychology University of Washington
2001	Psi Chi honor society
1997–2002	Dean's List, University of Washington

PUBLICATIONS

The contents of this thesis are currently in preparation for peer reviewed publication by authors Teeter and Stevens.

Nieto, Hayes, Wilson, Teeter, Stanbro. "Detection of Antineutrinos for Non-Proliferation". Nuclear Science and Engineering. 2005.

Teeter, Hayes, Wilson. "Antineutrino Energy Spectra From Nuclear Reactors." Los Alamos National Laboratory Theoretical Division Special Feature. 2003.

ABSTRACTS

Teeter, Stevens. "Differentiating Neural Arbors Using a Random Forest." Women in Machine Learning (WiML) Conference. Vancouver, British Columbia, December 2010. Poster presentation.

Teeter, Stevens. "Neural Arbor Classification Using a Minimal Set of Parameters." Society for Neuroscience Conference. San Diego, California, November 2010. Poster presentation.

Teeter. "How Would You Describe a Neuron? Insights from Neural Scaling Laws." Autonomous Neurodynamics Conference. San Diego, California. July 2009. Oral presentation.

Teeter, Stevens. "Characterizing Neurons Using Density Distribution Functions." Howard Hughes Medical Institute Interfaces Meeting. Chevy Chase, Maryland, September 2008. Poster presentation.

Teeter, Stevens. "Characterizing Neural Arbor Morphology Using Density Distribution Functions." Computational and Systems Neuroscience (COSYNE) Conference, Salk Lake City, Utah, February 2008. Poster presentation.

Teeter, Stevens. "Classification of Axonal and Dendritic Arbors Using Dimensionality Reduction Techniques." Society for Neuroscience Conference. San Diego, California, November 2007. Poster presentation.

ABSTRACT OF THE DISSERTATION

Characterizing the Spatial Density Functions of Neural Arbors

by

Corinne Michelle Teeter
Doctor of Philosophy in Neurosciences
with a specialization in Computational Neuroscience
University of California, San Diego, 2010
Professor Charles F. Stevens, Chair
Professor William B. Kristan, Co-Chair

Recently, it has been proposed that a universal function describes the way in which all arbors (axons and dendrites) spread their branches over space. Data from fish retinal ganglion cells as well as cortical and hippocampal arbors from mouse, rat, cat, monkey and human provide evidence that all arbor density functions (adf) can be described by a Gaussian function truncated at approximately two standard deviations. A Gaussian density function implies that there is a minimal set of parameters needed to describe an adf: two or three standard deviations (depending on the dimensionality of the arbor) and an amplitude. However, the parameters needed to completely describe an adf could be further constrained by a scaling law found between the product of the standard deviations and the amplitude of the function.

In the following document, I examine the scaling law relationship in order to determine the minimal set of parameters needed to describe an adf. First, I find that the flat, two-dimensional arbors of fish retinal ganglion cells require only two out of the three fundamental parameters to completely describe their density functions. Second, the three-dimensional, volume filling, cortical arbors require four fundamental parameters: three standard deviations and the total length of an arbor (which corresponds to the amplitude of the function). Next, I characterize the shape of arbors in

the context of the fundamental parameters. I show that the parameter distributions of the fish retinal ganglion cells are largely homogenous. In general, axons are bigger and less dense than dendrites; however, they are similarly shaped. The parameter distributions of these two arbor types overlap and, therefore, can only be differentiated from one another probabilistically based on their adfs. Despite artifacts in the cortical arbor data, different types of arbors (apical dendrites, non-apical dendrites, and axons) can generally be differentiated based on their adfs. In addition, within arbor type, there is evidence of different neuron classes (such as interneurons and pyramidal cells). How well different types and classes of arbors can be differentiated is quantified using the Random Forest^{TM} supervised learning algorithm.

Chapter 1

Introduction

The brain is a extraordinarily complicated circuit containing approximately 10^{10} neurons with approximately 10^{14} connections between them (Nagarajan and Stevens, 2008). To make things more complicated the components of this circuit are not homogenous, but instead, there are many different classes of neurons.

How does one go about unraveling the secrets of the brain? To understand the function of any circuit both the components and the connections must be identified and understood. This is a rather daunting proposition when considering the number of neurons in the brain. However, there must be underlying organizational principles that govern brain structure and allow the network to function.

What are the structural principles of the brain? It is clear that neural morphology plays a critical role in brain function. The morphology of neurons can mediate the functional properties of neurons. In addition, because neurons actually reach out and make physical contact with one another in order to communicate, morphology will determine how these components can connect and therefore constrain network computations. Because of the vital role structure plays in the brain, this document will focus on the organizational principles of morphology.

1.1 How are network connections specified?

So what are the structural principles that allow the brain to function? Principles of structure must begin at the level of neural growth laws. How neurons grow will affect neuron structure. Neural structure will both influence neuron function and the connections between neurons thus constraining network computation. How neurons 'reach out' and find each other in a labyrinth of densely packed biological media will eventually affect network function.

How are the connections of the brain specified? It is probable that long range connections that connect distant parts of the brain are determined genetically as it seems unreasonable that neurons would be able to find their way without some sort of programing. However, it is doubtful that the local connections are specified genetically. The brain is constantly changing and adapting and connections are shaped by experience. This plasticity in the brain gained attention with the famous Hebb postulate which identified that activity can invoke changes in synapses so that the activity of one neuron could more strongly influence the activity in another (Hebb, 2002) (originally published in 1949). Since then, at least three different types of plasticity have been discovered. The seemingly simplest and fastest form of plasticity involves the change in synaptic strength which can happen within a minute (Bliss and Collingridge, 1993; Gustafsson et al., 1990; Madison et al., 1991; Malinow et al., 2000). The next least drastic structural change involves the creation and destruction of synaptic spines which happens on a time scale of minutes to hours (Dailey and Smith, 1996; Engert and Bonhoeffer, 1999; B and Markram, 2006; Lendvai et al., 2000; Trachtenberg et al., 2002). Finally, the remodeling of neural branches can take place on the order of days (Yamahachi et al., 2009; Lee et al., 2005).

Even if the brain were perfectly static, it seems unlikely that the genome could encode the location of every synapse and therefore the location of every piece of dendrite and axon from every neuron. It seems probable that the genome specifies general rules or strategies for connectivity and then connections form within the context of these general rules.

So what are the general rules that specify local connections? Because of the intimate relationship between neural form and connectivity, many studies have focused on deriving connectivity from morphology (Chklovskii, 2004; Hellwig, 2000; Kalisman et al., 2003, 2005; Liley and Wright, 1994; Sholl, 1953; Stepanyants et al., 2002, 2004; Uttley, 1955). Many investigators have provided evidence that neurons spread themselves out randomly over space. In 1953, Sholl and Uttley wrote, "There is reason to believe that many of the complexities of the cerebral cortex are due to an organization that is fundamentally of a statistical nature and the parameters of which can only be described in terms of probabilities" (Sholl and Uttley, 1953). In 1954, Uttley published a paper showing that, theoretically, a network of randomly connected neurons could in fact learn simple classification tasks, thereby showing a proof of principle that randomly connecting neurons could indeed perform 'brain like' tasks (Uttley, 1954). Then in 1955, Uttley estimated the number of connections that would be expected based on random connectivity between neurons (Uttley, 1955). Since then, many researchers have provided evidence that neurons largely connect randomly. (Braitenberg and Schuz, 1998; Hellwig, 2000; Kalisman et al., 2003, 2005; Liley and Wright, 1994). However, there is some evidence that interneuron growth targets certain neurons while pyramidal cells do not (Kisvarday et al., 2002; Song et al., 2005; Stepanyants et al., 2004). In this case, it is possible that the interneurons spread out randomly over space and then use a different mechanism such as chemical attractants to find their target when in close proximity.

1.2 The brain as an optimization problem

Optimization theory can add rationale regarding why arbors might spread out randomly over space. Optimization principles have long been invoked to explain organizational structure in the brain. The brain is almost certainly a complicated optimization problem where networks of neurons have to communicate both locally and across different regions. The network is highly constrained by space as neurons

take up physical volume and extend out to make physical contact with one another. In addition, neurons have metabolic requirements and are therefore constrained by metabolic resources. Somehow the brain has developed an organizational structure that balances its need for communication and computation while minimizing space and metabolic requirements

As is nicely summarized in Chklovskii (2004), brain optimization theory started with Ramon y Cajal's laws of economy of space, time and matter (y Cajal et al., 2002) (originally published in 1899). In these laws, Dr. Cajal hypothesizes that network functionality is maximized while wiring cost is minimized. Since then, optimization principles of wire minimization have been employed to explain many phenomenon found in the brain. These include: why the brain is located in the head, why the neocortex folds in a species specific pattern, why gray and white matter are segregated in the cerebral cortex, why there are separate visual areas, why the number of areas and neuron density scale with brain size, how cortical areas in mammals and the ganglia of C. elegans are arranged with respect to one another, why topographic maps exist, why the visual cortex has ocular dominance patterns and orientation maps, why axonal and dendritic arbors have particular dimensions and branching angles, and why axons and dendrites occupy a certain fraction of the gray matter. Please see Chklovskii (2004) for an excellent list of references.

Optimization principles have also been employed to explain how a neuron might distribute its arbors over space (Cherniak et al., 1999; Chen et al., 2006; Chklovskii, 2000; Wen and Chklovskii, 2008; Wen et al., 2009). It is likely that a neuron would maximize its possible connections within the circuit but minimize overall wire length or volume. In addition, the distance of each piece of branch to the soma would likely be minimized as it is costly to 'pump' resources out to the neuron extremities. With this type of morphological strategy it is unlikely that the location of every piece of neurite matters, but instead the neuron spreads itself out over space in such a way as to maximize its connections and minimize metabolic needs.

1.3 Scalable architecture

Scalable architecture is an interesting phenomena that can provide yet further insight into the organizational principles of the brain. Scalable architecture refers to the ability of a system to 'scale up' or be capable of processing more information by increasing the number of components in the network, but maintain the same basic architectural structure of system. Scalable architecture seems to be a general feature of mammalian brain evolution as the neocortical and subcortical structures grow disproportionately (Finlay and Darlington, 1995). For example, if the thalamus of a monkey and a bat were scaled to the same size, the cortex of the monkey would be much larger than the cortex of a bat. An example, of scalable architecture observable within an organism's lifespan is found in the fish brain. The fish brain continues to grow with the size of the fish throughout its life (Lee and Stevens, 2007). During this process the neurons continue to get larger, but not at the same rate as the brain. Instead, the existing neurons grow, while new neurons are created and integrated into the circuit (Lee and Stevens, 2007). This is an extraordinary feat! The network continues to grow and integrate new neurons into the circuit without interrupting normal function.

There must be a mechanism that allows this growth and integration. Such a phenomena implies, at the very least, some sort of scaling law for neural structure.

In fact, recently, scaling laws have been discovered for neural arbors. Neural arbors (axons and dendrites) are the tree like processes that grow from the cell body of a neuron in order to sample (dendrites) or disseminate (axons) information over a region of the brain. Scaling laws are observed as a power law relationship between two variables (Note: a power law is linear in logarithmic space). Wen et al. (2009) found that there is a power law relationship between the total dendritic length and the radius of an arbor. Snider et al. (2010) found that there is scaling relationship between the the total arbor length and the area or volume covered by an arbor.

1.4 A universal function for neural arbors

Scalable architecture has been extensively studied in physics and is associated with many interesting phenomena, one of which is universality. Perhaps unsurprisingly, a universal function that describes the way neural arbors spread themselves over space has recently been uncovered. The general way in which an arbor distributes itself over its territory can be described by an arbor density function (adf); a function that is proportional to the probability of finding a piece of arbor at every point in space (Figure 1A). Snider et al. (2010) showed compelling evidence that there is a universal function that describes the spatial density function of all neural arbors. In addition, they go on to show that the universal function is likely a Gaussian truncated at approximately two standard deviations. This is unsurprising as a Gaussian is a famous scale invariant function.

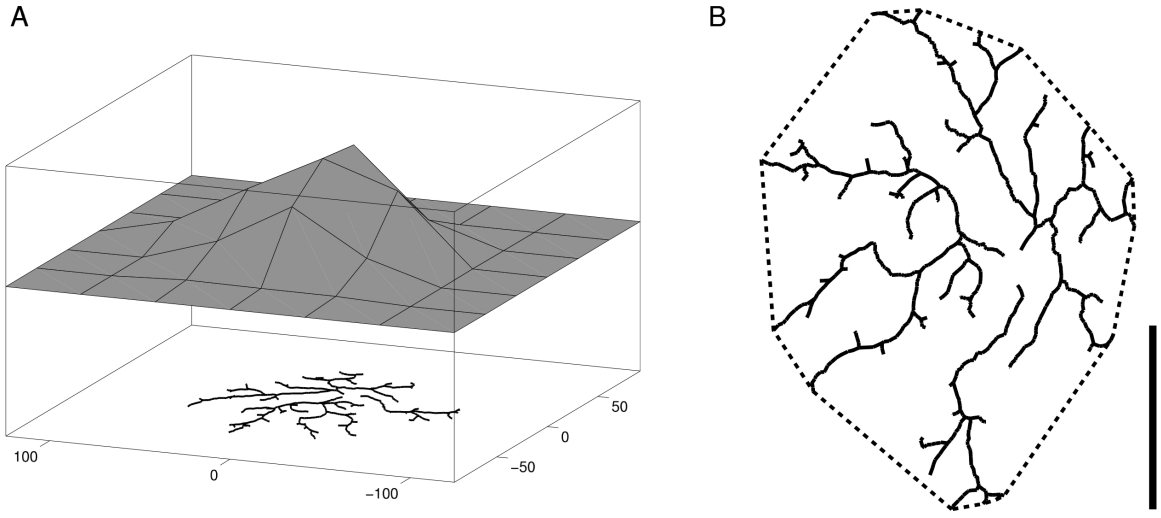


Figure 1.1: Arbor density function and convex hull of a retinal-ganglion cell axonal arbor. A. Axon reconstructed with the arbor density function plotted above. Amplitude of density function is proportional to the probability of finding a piece of arbor at each point in space. B. Same arbor as in A. The area enclosed within the dashed lines is defined as the convex hull area. Scale bar is 50 microns.

1.5 Structure and function: neuron classification based on morphology

It has long been recognized that there are many different classes of neurons in the brain such as pyramidal cells, stellate cells, as well as many different kinds of interneurons. The architectural principle that a structure must have a certain form to perform a certain function (if a wheel is not round it can not roll) has prompted over a century of research attempting to classify neurons based on morphology. Neuron classification started soon after Camillo Golgi discovered the revolutionary Golgi stain in 1873 and for the first time the neurons of the brain could be visualized. In 1899, Ramon y Cajal published his famous book, Comparative study of the sensory areas of the human cortex, where he classifies different types of neurons based on their morphology.

Many years later, numerous studies have shown that morphology does indeed influence the activity of neurons (Larkman and Mason, 1990; Mason and Larkman, 1990; Mel, 1994; Mainen and Sejnowski, 1996; Jaffe and Carnevale, 1999; Vetter et al., 2001; Gulledge et al., 2005). This is strong evidence for differing neural structure. A universal function that describes arbor structure seems incompatible with the evidence that structural variations achieve different neural activity. Can these seemingly conflicting ideas be reconciled?

1.6 Reconciling a universal function with different neural classes

How do different classes of neurons fit into the framework of scaling laws and a universal arbor density function? Having different morphological classes of neurons seems contradictory to a universal structural law for all neural arbors. However, several topics should be considered when pondering morphological classes versus a universal function.

First, the universal Gaussian function describes neural arbors not neurons as a whole. The neural arbors grow out of the soma to a region of space where they will distribute their neurites. It is possible that all arbors are universal but are located at different locations relative to the soma and this is what constitutes different classes of neurons. Addressing this topic will be very important to characterizing different classes of neurons; however, it is outside the scope of current research and could be the topic of future research.

The differing location of arbors on different neurons does not completely escape the conundrum of differing neural structure. Arbors have also been classified into different types based on their structure and functional properties: apical dendrites, non-apical dendrites, and axons. How can these different morphological types of arbors be accounted for with a universal function?

It is possible that different branching laws from different types of arbors could yield a Gaussian spatial density function with the appropriate scaling laws. Although the connection between branching laws and the spatial density function is a very intriguing line of research it is difficult to address. It is impossible to exhaustively check all possible branching laws that could yield correct experimental values. Perhaps this could be addressed using optimization theory, but this is outside the scope of the current research.

What will be discussed in this document are the parameters that define the spatial density functions of arbors. Arbors may have different parameters within the context of a universal law and it is possible that different types of arbors could be differentiated by these parameters. If the spatial density functions of all arbors are Gaussian, it implies that relatively few parameters are needed to specify the adfs. These parameters are the same parameters needed to define a Gaussian function: a standard deviation for each dimension and an amplitude. If an arbor is two dimensional (flat like a pancake) such as retinal ganglion cells, the arbors territory covers an area. In this case, two standard deviations along with an amplitude will be required to describe the arbor. If an arbor is volume filling, such as cortical arbors,

three standard deviations and an amplitude will be required. This means that even if all arbors are Gaussian, different types of arbors might be distinguished by having different parameters or different combinations of parameters.

Different combinations of standard deviations will define the geometrical shape of an arbor. For two-dimensional arbors, the ratio of the standard deviations will describe eccentricity: how circular or ovular an arbor is. For a three dimensional arbor, the standard deviations will determine how spherical or 'egg-like' an arbor is. Interestingly, neither Wen et al. (2009) nor Snider et al. (2010) considered arbor shape. Wen and colleagues recognized a scaling law between the total length of an arbor and the radius of an arbor but assumed that the arbors were spherical. I.e they assumed one radius could completely describe an arbor. Arbors are certainly not always spherical which will be addressed in the current research. Snider et al. (2010) went as far as to define a scaling law between the total length of an arbor (amplitude of the Gaussian) and the product of the standard deviations that define the Gaussian, but did not consider the standard deviations separately to define arbor shape.

The final parameter that could vary is the amplitude of the Gaussian function. The amplitude corresponds to the total length of the arbor and will describe how dense an arbor is. The scaling laws found by Wen et al. (2009) and Snider et al. (2010) define the power law between parameters that describe arbor size and the total length of an arbor. If these relationships were exact (i.e. there is no scatter in the relationships) it would imply that the size of the arbor would perfectly define the amplitude of the Gaussian function. However, this is not the case. Instead, there is scatter in the relationships. Therefore it is also possible that different types of arbors have different densities.

A universal function could explain the difficulties researchers have had when classifying neurons based on their structure. Morphological classification has proven to be challenging due to the apparent complexity of neuron architecture. In an attempt to overcome this complexity, many different morphological traits have been

used as a basis for characterization and sophisticated quantitative methods have been applied to these detailed descriptions of arbor structure (Badea and Nathans, 2004; Kong et al., 2005; Tsiola et al., 2003; Uylings and van Pelt, 2002). However, it is difficult to know what morphological features are required to classify different types of neurons. Theoretically there are an infinite number of parameters one could use in an attempt to differentiate different types of cells (Kong et al., 2005). Overall, classification based on somewhat arbitrary characteristics, using sophisticated clustering methods has been unsatisfying. These methods either do not adequately characterize known types of neurons or the functional relevance of the classification parameters of resulting classes is unknown. Interestingly, Kong et al. (2005) showed that the parameters chosen for classification are often correlated which clouds the analysis. They show that the best classification is obtained by using relatively few, independent parameters. In addition, the parameters they find that optimize classification turn out to be parameters that are similar to the fundamental parameters suggested by the universal function. In fact, the power law relationship is visible in their data but they do not describe it.

Perhaps morphological classification has been difficult because the arbors of neurons are very similar and arbor types exist on a continuum with no clear cut defining types.

1.7 In this thesis

In the following manuscript, I aim to create a more complete description of the spatial distribution strategies of neural arbors and describe the differences between types of arbors. I will begin by considering the simpler two-dimensional retinal ganglion cells of fish and then discuss more complicated three-dimensional cortical arbors. I will show the power law relationship between the product of the standard deviations and and the total length of an arbor. I will show that the product of the standard deviations is proportional to the area or volume covered by an arbor

(depending on the dimensionality of the arbor), and as a result, there is a power law between the total length of an arbor and the area or volume covered by an arbor.

Next, I will describe the scatter around the regression line. Because the power law relationship characterizes area or volume versus length, the scatter in the regression line will correspond to the range in average density for an arbor of a given size. Defining the scatter around the line will have a dual purpose. First, it will provide evidence for or against different density arbor types. Second, it will define the minimal number of parameters needed to describe the spatial density function of an arbor. If density does not contain any information about arbor type, the number of parameters needed to describe an adf will be reduced from three to two. For example, if the 2D spatial density function of a fish retinal ganglion cell is a Gaussian, three standard deviations will be required to define it: two standard deviations and an amplitude. However if there is no information in the scatter around the power law relationship, only two parameters will be needed to define an adf, the third will be specified by the power law equation. Similarly for 3D cortical arbors, if there is no information about arbor type contained in arbor density, the number of parameters needed to define a cortical arbor will be reduced from three to two standard deviations and a total length.

For fish, I will find that both the axons and dendrites of fish are distributed around the regression line as Gaussians and that the minimal set of parameters needed to completely describe the arbors can be reduced to two parameters. I will find that axons and dendrites are shaped similarly and can not be differentiated based on arbor shape. Axons are generally larger than dendrites but the arbor size distributions overlap and therefore axons and dendrites can only be differentiated probabilistically. In addition, there is no evidence to support different classes of arbors within the arbor types (which would be revealed as clusters in the data).

For cortical arbors I will show that the axons and the two types of dendrites (apical and non-apical) are distributed differently around the volume versus total length power law. In addition, different types of arbors from different classes

of neurons have different relative densities (on average, the axons of interneurons are more dense than the rest of the arbors). Therefore, the density of arbors is indeed varied by some process. Thus, four parameters are needed to describe cortical adfs: three standard deviations and a total length (which corresponds to the amplitude of the function). I will provide a description of arbor shape and quantify the distributions of the arbor parameters. In addition, I will identify differences in density and eccentricity between arbors of interneurons and pyramidal cells

Finally, I will show that different types of arbors are located in different regions of the parameter space and quantify how well they can be distinguished from one another using a supervised learning technique. Then, I will show that arbors of the same type but from different predefined classes of neurons have some differentiability but can not easily be distinguished.

Chapter 2

Results: Fish Retinal Ganglion Cells

I begin my investigation by considering both the axons and dendrites of fish retinal ganglion cells. The fish data provide the simplest possible example with which to begin my analysis in order to deduce a complete description of arbor structure. The arborizations of these cells are flat, pancake like, two-dimensional processes. The dendrites reside in different layers, tiling the front of the retina. The axons project through the optic nerve into the tectum where again they make flat two-dimensional arborizations. Because these arbors are flat, they essentially cover an area as opposed to a volume. Therefore, the Gaussian function that describes each arbor density function (Figure 1.1) will have three parameters associated with it: an amplitude which corresponds to the total length of an arbor and two orthogonal standard deviations, σ_x and σ_y. σ_x and σ_y are calculated from the arbor reconstruction data using principal components analysis (please see the Methods section).

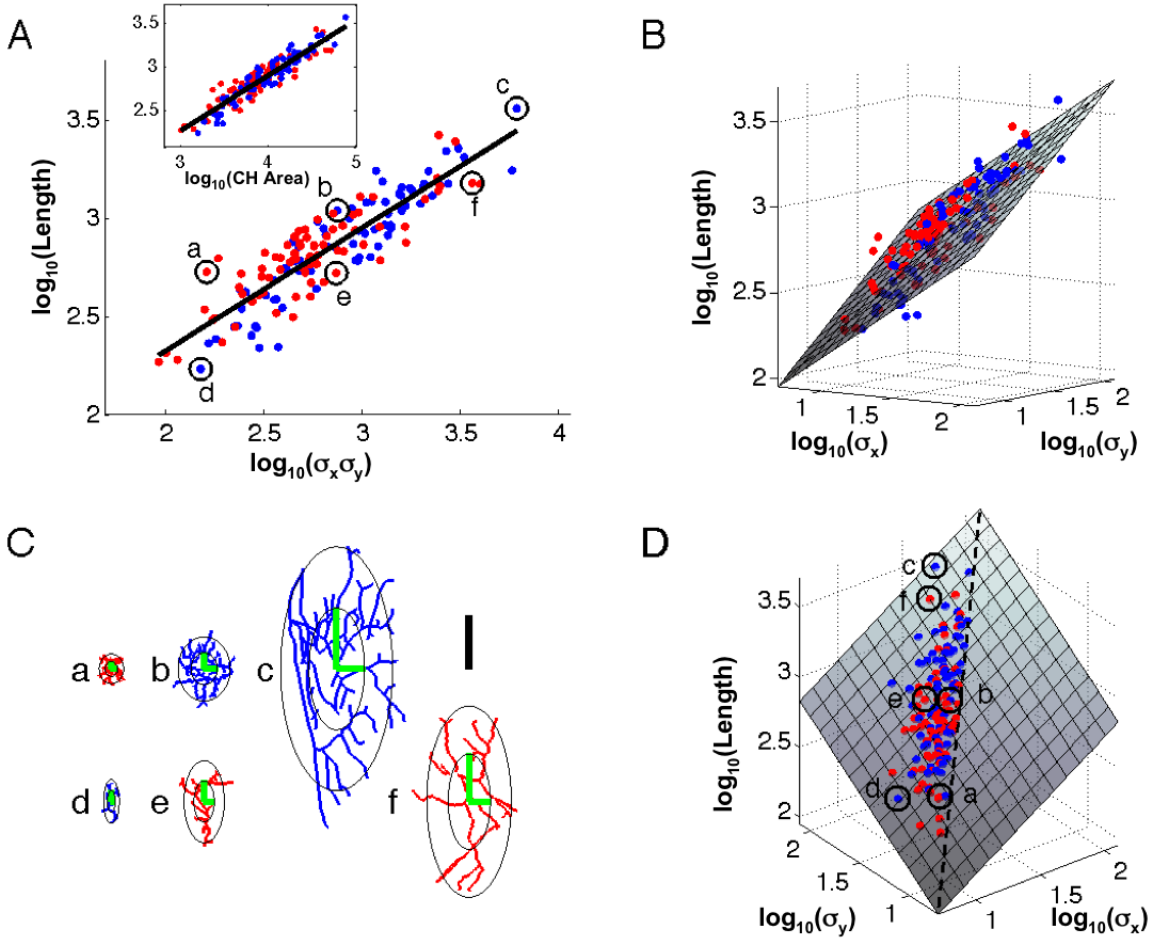
2.1 Power law between arbor size and length

Snider et al., (2010) discovered that there is a power law relationship (which is linear in logarithmic space) between the total length of an arbor and the product

of the standard deviations (Figure 2.1A). The area an arbor covers is commonly described by its convex hull (Wassle and Boycott, 1991) and is defined as the area within a convex polygon enclosing the outermost extremities of an arbor (Figure 1.1B).

Figure 2A, of the Snider et. al., 2010, paper shows that the product of the standard deviations is proportional to arbor convex hull area. I do not replot their figure; however, Figure 2.1A shows that the power law relationship between the convex hull area and the total length of an arbor, is nearly identical to the plot of $log(\sigma_x \sigma_y)$ versus $log(L)$. The y intercepts of the linear regressions differ but the slopes of both plots are the same. This implies that in linear space, area and $log(\sigma_x \sigma_y)$ are proportional. Figure 2.1C illustrates how area is characterized by the two standard deviations, σ_x and σ_y. Six arbors are plotted along with their standard deviations σ_x and σ_y, of the spatial data in two principal, orthogonal directions. The relationship between the standard deviations and the area is demonstrated by the ellipses characterized by the standard equation, $x_2/r_x^2 + y_2/r_y^2 = 1$. Here r_x and r_y are the radii of the ellipse. The inner ellipses are plotted by setting $r_x = \sigma_x$, and $r_y = \sigma_y$. The outer ellipses are plotted by setting $r_x = 2\sigma_x$, and $r_y = 2\sigma_y$.

Figure 2.1: Relationships between three parameters that describe retinal ganglion cell arbors. Both axons and dendrites plotted on all plots. Axons are blue. Dendrites are red. A. A plot of the logarithm of the standard deviation product, $\sigma_x\sigma_y$, versus the logarithm of the total length, the summation of all the segment lengths, of the arbor reveals a power law relationship. The line is a linear regression fit to both the axonal and dendritic data. Slope =0.63 $\pm$ 0.05 , Intercept=1.07, $\pm$ 0.14 , r^2=0.80. Inset: plot of arbor convex hull area versus total arbor length shows that plot is very similar to the larger plot. Slope =0.63 $\pm$ 0.04, Intercept= 0.37 $\pm$ 0.15, r^2=0.87. (all confidence intervals at 95%). This means the convex hull is proportional to $\sigma_x\sigma_y$ (explicit relationship not shown here, see Snider et al. (2010), Figure 2A). Circled data points correspond to the neurons shown in C. B. Three dimensional plot of $log_{10}(\sigma_x)$, $log_{10}(\sigma_y)$ versus $log_{10}(Length)$. Plotting $log_{10}(\sigma_x)$ and $log_{10}(\sigma_y)$ separately has the effect of expanding the linear regression line shown in A, into a plane. C. Neurons corresponding to the data points circled in A and D. Orthogonal, principal, standard deviations obtained by applying principal components analysis to each arbor are plotted in green. The ellipses are characterized by the equation, $x^2/r_x^2+y^2/r_y^2 = 1$. Inner ellipses: $r_x = \sigma_x$, and $r_y = \sigma_y$. Outer ellipses: $r_x = 2\sigma_x$, and $r_y = 2\sigma_y$. Notice the clear relationship between σ_x, and σ_y and the area covered by the ellipses. Scale bar is 100 um. D. Same plot as in B but now viewed from an angle looking down on the plane. The dashed line illustrates where σ_x equals σ_y. All the data are located to the left of the line as σ_y is defined to be the larger principal standard deviation. Notice that eccentricity is now represented in this plot. Data points that lie closer to the dashed lines represent neurons that are more circular whereas data points that lie farther away from the dashed line represent neurons that are more ovular. Circled data points correspond to the neurons shown in C.


A
B
C
D
log10(Length)
log10(σxσy)
log10(CH Area)
log10(σx)
log10(σy)
a
b
c
d
e
f

It appears that the same power law characterizes both the axons and dendrites of retinal ganglion cells as a single power law is fitted to both types of arbors. This suggests that whatever process that governs the creation of the dendrites also governs the creation of the axons.

2.2 Defining average arbor density

Next, I will define the average density of a neuron and deduce whether density could contain information about arbor type. To do this I will define the scatter around the $log(\sigma_x\sigma_y)$ versus the $log(L)$ regression line. I do this because the product of the standard deviations is a proxy for the area covered by an arbor. Average density is defined by the length divided by the area. Therefore, the scatter around the regression line in Figure 2.1A corresponds to a difference in average density. Points above the line correspond to higher density arbors as they have more length per area than points below the line. Because the slope of the regression line is less than 1, on average arbors get less dense as they get bigger. This relationship is plotted explicitly in Figure 2.1.

The constant scatter around the line in log space indicates a constant relative density of the arbors: the density relative to the mean average density is constant in linear space. That is, the variation in the density is a constant percentage of the mean average density at that size. Figure 2.3D shows histograms of the distance of the data from the regression line (representing relative average density), commonly known as the residuals of the regression line shown in Figure 2.1A. Histograms of axons and dendrites are plotted separately in blue and red respectively. A Lillifors test for Gaussianity reports that both the axons and dendrites are not significantly different from a normal distribution with 95% confidence. A Kolmogrov-Smirnov test reports that these two distributions are not significantly different from each other with 95% confidence. Therefore, I see no evidence to suggest that the brain is intentionally creating different density classes of fish retinal ganglion cell arbors. In addition, I

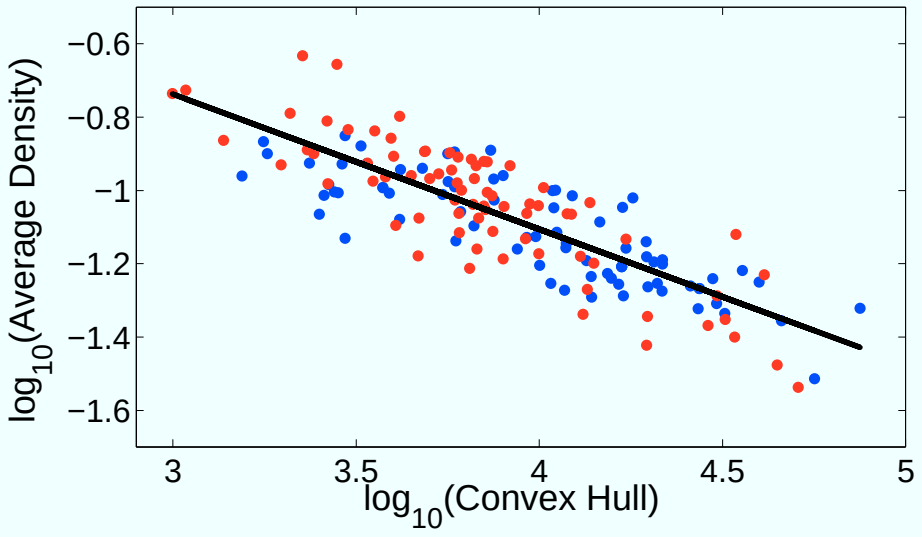


Figure 2.1: Average density decreases with arbor area (as defined by the convex hull) in fish retinal ganglion cells. Axons in blue, dendrites in red. Average density is the total arbor length divided by the convex hull area. The slope of the regression line is -0.37 $\pm$ 0.04, the intercept is 0.37 $\pm$ 0.15, $r^2 = 0.70$.

do not see a difference in density distribution strategies for axons or dendrites; it appears that the size length density creation mechanism is the same for both axons and dendrites.

2.3 Only two parameters are needed to define the spatial density function of fish retinal ganglion cells

The lack of information content in the density implies that the spatial density function of fish retinal ganglion cells can be completely described by just two parameters. The Snider et al. (2010) paper shows evidence that the spatial distribution function is a Gaussian. If this is true only three parameters are needed to define the shape of the arbor: two standard deviations, σ_x, and σ_y, and an amplitude. In the arbor case the amplitude can be represented by the total length of an arbor. Since

the scatter around the regression line is constant, corresponding to a constant relative average density of arbors, the amplitude of the Gaussian function can be described by a density range for an arbor of a certain area. This means that only two parameters are needed to describe the density function: the other will be defined by the equation of the regression line shown in Figure 2.1, $log(L) = 0.63log(\sigma_x\sigma_y) + 1.07$.

2.4 Defining arbor shape

Next, I consider arbor shape. The formulation of Snider et al. (2010) provides a convenient context in which to define arbor shape. The relationship between total length and the product of the standard deviations describes how length and area scale together; however, it does not describe the geometry of an arbor. For example, an arbor could be very circular or very ovular and have the same area. As shown in Figure 2.1 the standard deviations are proportional to the radii of an oval and therefore will describe the geometry of the arbor.

Eccentricity, defined as the ratio of σ_x to σ_y, will define how ovular or circular an arbor is (Figure 2.1). A perfectly circular arbor will have an eccentricity value of 1 and progress toward more ovular as the number decreases to 0. The $\sigma_x\sigma_y$ versus length plot does not characterize arbor geometry. However, conveniently, plotting L against σ_x and σ_y separately has the mathematical effect of expanding the regression line in Figure 2.1A into a plane shown in B and D. The distance of a data point from the plane will be the same as the distance of the same point from the log(L) versus $log(\sigma_x\sigma_y)$ regression line. That is, the data residuals will be the same for the plane and the regression line. Now, eccentricity is also described in the plot by the way the data are distributed across the plane. Figure 2.1D is a top view of the plane plotted with the axonal and dendritic data. The dashed line corresponds to $\sigma_x = \sigma_y$ on the plane. All data lie on the left side of the dashed line due to the definition $\sigma_y \geq \sigma_x$. Circled data points correspond to the arbors plotted in C. Notice that ovular arbors such as c,d,e and f, fall farther away from the dashed $\sigma_x = \sigma_y$ line and more circular

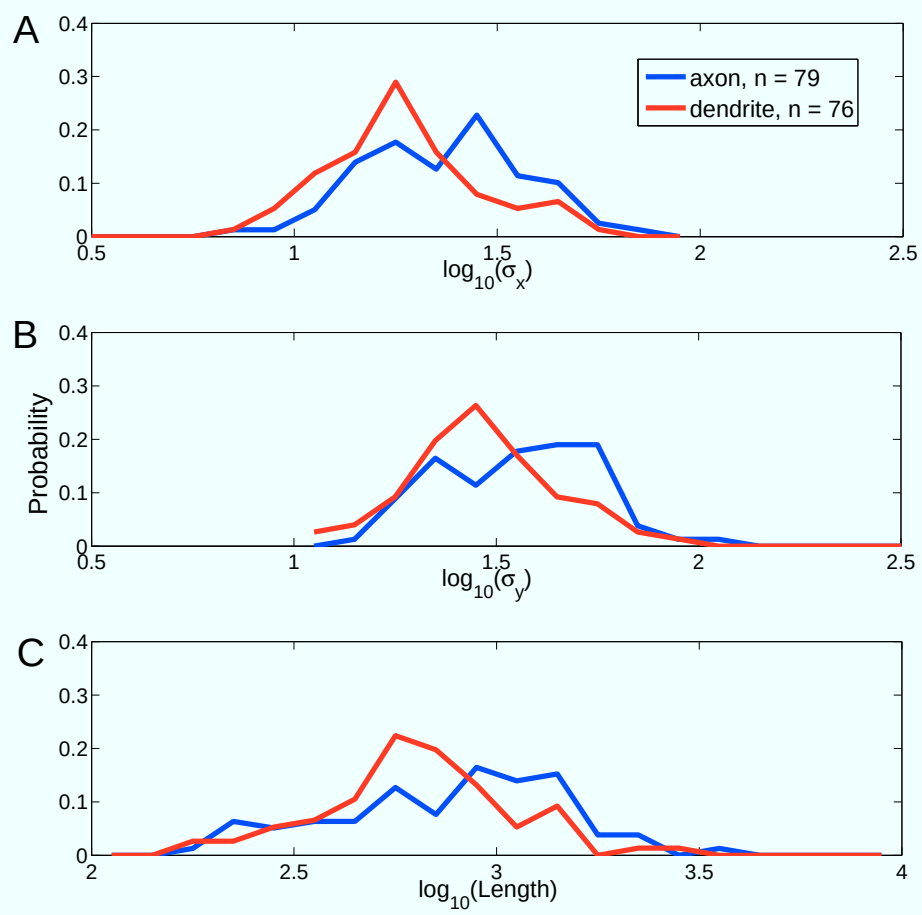


Figure 2.2: Normalized histograms of σ_x, σ_y, and length of fish retinal ganglion cells in logarithmic space. All distributions are significantly different from one another. In all cases the parameters of axons are generally larger. None of the distributions are significantly different from a normal distribution. Kolmogrov-Smirnov and Lillifors statistical tests were used. All tests performed at 95% confidence.

arbors such as a and b are closer to the dashed line.

2.5 No correlations between relative density and eccentricity or area

Looking at Figure 2.1D, it appears that the data are distributed above and below the plane uniformly and that the plane adequately characterizes the data. However, this may not have been the case. For example, it could be that ovular arbors

are less dense and circular arbors are more dense. Looking at the neurons illustrated in Figure 2.1C, it appears that this could be the case since the ovular neurons tend to fall below the regression line in A. Simply characterizing the distribution of the data around the regression line would not have identified this type of relationship. However, no relationship was found between eccentricity and the residuals (Figure 2.3) for either axons or dendrites (95% confidence). Therefore, the plane adequately describes the retinal ganglion cell arbors.

In addition, it is possible that a relationship between area and eccentricity could exist. This also would have gone unnoticed in the $\log(\sigma_x\sigma_y)$ versus log(L) plot. However, it is likely that the relationship would be visible on the plane in Figure 2.1 B and D. For example, if larger arbors were more eccentric they would be farther away from the $\sigma_x = \sigma_y$ line. To confirm a lack of relationship, correlations between eccentricity and area were examined in Figure 2.3B; no relationships were found.

2.6 Summary

It appears that the spatial distribution strategies for fish retinal ganglion cell axons and dendrites are very similar. They have similar relative average density distributions (Figure 2.3D) and eccentricity distributions (Figure 2.3C). There are no correlations between eccentricity and the area an arbor covers or relative average density (Figure 2.3 A & B) for either axons or dendrites. Their differentiating characteristic is size. Both the standard deviations, and the total length of an arbor are generally bigger for the axonal arbors Figure 2.2. Therefore, the area (which is proportional to the product of the standard deviations) covered by an axon is also generally larger Figure 2.3E). Although the relative average densities for axons and dendrites are not significantly different from one another, in general the average density of axons is less than dendrites because bigger arbors are less dense (Figure 2.1).

In addition, because there is no information about arbor type contained

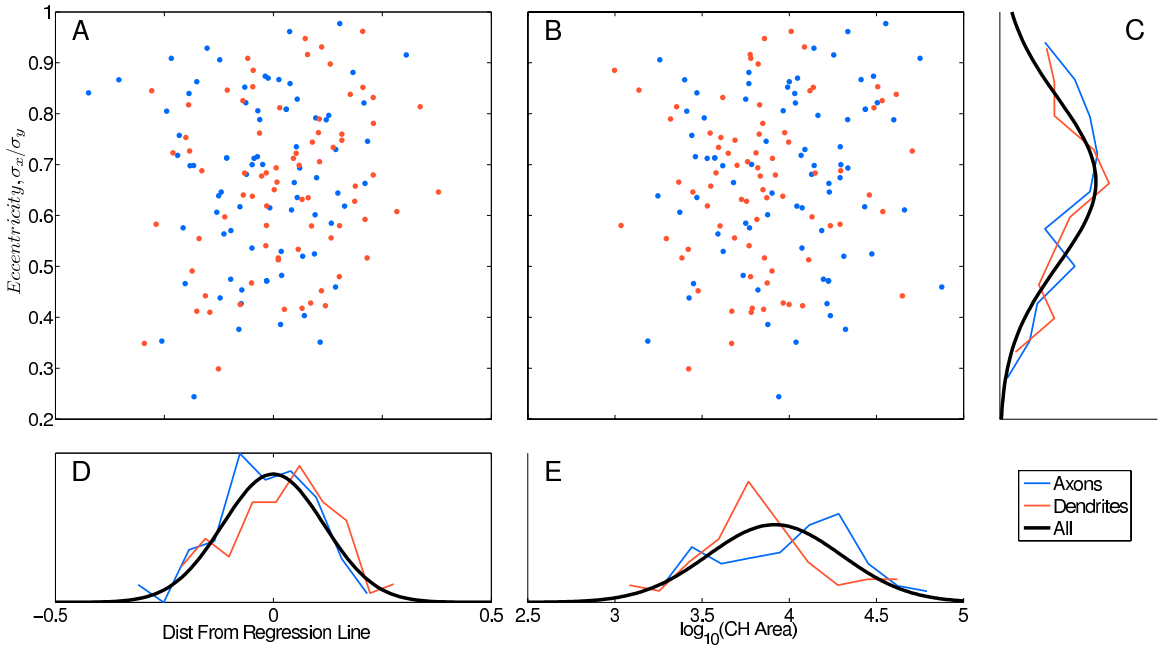


Figure 2.3: Arbor area distributions are the only differentiating factor between axons and dendrites. A & B scatter plots of eccentricity (defined as σ_x/σ_y) versus relative average density (represented as distance from the regression line plotted in Figure 2.1), and $\log_{10}$ of the arbor area respectively. C, D, & E. Normalized histograms of the data are plotted. In all plots axons are in blue, n = 79, and dendrites are in red, n = 76. Black lines correspond to Gaussians fits of combined axonal and dendritic data. Axis assignments correspond to all plots across rows and columns. A. Distance from the regression line (which characterizes the density of an arbor relative to the mean density of an arbor of a given size) and eccentricity are not linearly correlated revealing that different shaped arbors are not scattered around the regression line shown in Figure 2.1 in any systematic way. Axons: p=0.7160. Dendrites: p=0.0734. B. There is no linear correlation between convex hull area and eccentricity revealing that larger neurons are not more circular or have some other such relationship. Axons: p= 0.4729. Dendrites: p=0.3636. Lillifors tests were used to determine if distributions are significantly different from Gaussians. None of the distributions in this image for axons and dendrites separately or combined were found to be significantly different from a normal distribution except for the $\log_{10}$(area) in E. The axonal and dendritic distributions of relative density (D) and eccentricity (C) were not found to be significantly different from one another. The arbor area distributions of axons and dendrites shown in E are significantly different. Although the axonal distribution is not shown to be significantly different from a Gaussian, area might differentiate two different types of axons corresponding to the two peaks in the distribution. It is also interesting that the dendritic distribution is significantly different from a Gaussian and fills in the space between the two axonal peaks. All statistical tests performed at 95% confidence.

in the relative average density, the minimum number of parameters needed to characterize the adf of a fish retinal ganglion cell is two out of the three fundamental parameters. The remaining parameter will be determined within the relative average density range by the scaling equation log(Length)= 0.63*log($\sigma_x\sigma_y$)+0.37

In conclusion, the axonal and dendritic spatial distribution strategies of fish retinal ganglion cells are very similar. The spatial arbor density function of both axons and dendrites can be described by two parameters and knowledge of the power law relationship between the total length of a neuron and the area an arbor covers. In general axons are bigger and less dense, but, although size and average density do help determine arbor type, axon and dendrites do not independently cover the parameter space. Therefore dendrites and axons can only be differentiated probabilistically.

Chapter 3

Results: Cortical Arbors From Rat, Cat, Monkey & Human

Six hundred and forty-six neurons with a total of 1406 arbors from the cortex of rat, cat, monkey, and human were included in this analysis. The arbors of the neurons were categorized into three types by the original investigators: apical dendrites, all other types of dendrites, and axons. In addition, information regarding neuron class was provided. These arbors are a subset of the data available from the Neuromorpho.org database. Many arbors in the data base were deemed inappropriate for this study as they were subject to many experimental and tracing artifacts discussed in the Methods. Only arbors which had minimal artifacts (relatively little slicing artifacts and no obvious tracing errors) and whose data could be corrected by a shrinkage factor (if needed) were used in this analysis. Despite this quality control, errors in the data set remain due to slicing artifacts (which could not be totally eliminated), inaccurate shrinkage corrections, and potentially unnoticed tracing artifacts. Assessments of how artifacts might affect the analyses below will be included as necessary. Details of the data selection along with a table of the included neurons can be found in the Appendix.

3.1 Scaling relationship between arbor size and length

As with fish, there is a power law relationship between the product of the standard deviations and the total length of an arbor. Because these arbors are three-dimensional volume filling arbors, three standard deviations will describe the spatial arbor density function: σ_x, σ_y, and σ_z. The linear relationship between the $\log(\sigma_x\sigma_y\sigma_z)$ versus log(L) is plotted in Figure 3.1A and a single regression line is fitted to all three arbor types. The product of the standard deviations is proportional to the volume an arbor covers as defined by the convex hull (Figure 3.1). This means that the plot of the product of the standard deviations versus the total length shown in Figure 3.2A also represents a plot of arbor volume versus total length (Figure 3.2A, inset). Figure 3.2A, shows these plots are almost identical.

3.2 Defining average arbor density

3.2.1 The distribution of density around the power law

I begin by considering the distribution of the data around the power law relationship between the product of the standard deviations and the total length. Remember that this power law is equivalent to a power law between arbor volume and total length (Figure 3.1A). Here, the average density of an arbor will be the total length divided by the volume. Therefore, the scatter around this line corresponds to a variation in average density relative to an arbor of a certain volume. If the scatter around this line is constant (corresponding to a constant relative average density) and the different arbor types have the same distribution around the line, the minimal set of parameters needed to describe the spatial adfs of cortical arbors can potentially be reduced to three parameters. As with fish, because the slope of the power law is less than one, the average density decreases with arbor size. This relationship between size and average density is shown in Figure 3.2B.

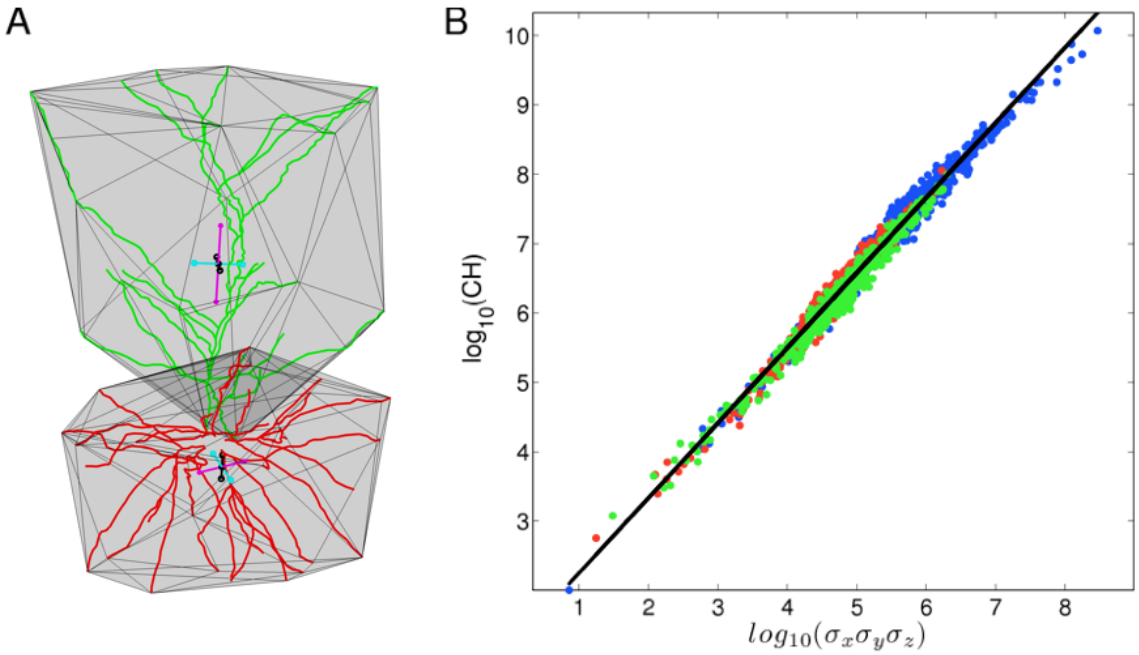


Figure 3.1: The product of the standard deviations is proportional to the volume covered by an arbor. Red are non-apical dendrites, green are apical dendrites, and blue are axons. A. Pyramidal cell with separate arbors delineated. The volume of an arbor is defined as the volume within its convex hull (in gray). Cross bars represent the principal standard deviations found using PCA. Center of colored crossbars denote the center of mass of each arbor. The length of the crossbars (center of the cross bar out to the end) is one standard deviation. The largest of the standard deviations is colored magenta, cyan is the second, and black is the smallest. In this case the magenta standard deviation will be defined as σ_z, as it is most directly pointing toward the pial surface. The largest of the remaining two is defined as σ_y (cyan) and the smallest is σ_x (black). B. There is a proportionality relationship between the product of the standard deviations and the convex hull volume. The regression line has a slope of 1.08 $\pm$ 0.01, an intercept of 1.17 $\pm$ 0.04, and an $r^2 = 0.98$. On a logarithmic plot a slope of one denotes proportionality. The proportionality constant is determined by the intercept, *int*. $CH = e^{int}(\sigma_x\sigma_y\sigma_z)^{slope} = e^{1.17}(\sigma_x\sigma_y\sigma_z)^{1.08} = 3.22(\sigma_x\sigma_y\sigma_z)^{1.08}$. Where, CH, is the convex hull volume, $\sigma_x\sigma_y\sigma_z$, is the product of the principal standard deviations. The slope and intercept are obtained from the regression line.

In contrast to the fish, it appears that information concerning arbor type is present in the density of an arbor. All three of the relative density distributions are significantly different from one another (Kolmogrov-Smirnov, 95% confidence) (Figure 3.4H). In addition, the relative density distributions are not distributed as Gaussians for any of the arbor types. The relative average density is not constant along the entire length of the regression line (Figure 3.2). Instead it appears that scatter 'fattens' as the arbors get larger. Note however, that larger arbors are axons and the increase in relative density along the regression line appears to be the result of an increase in the relative density of the axons (this will be pursued later in the document). This can be seen in Figure 3.4H where the distribution of the relative densities (represented as the distance form the regression line) of axons is wider than the apical and non-apical dendrite distributions.

3.2.2 The effect of slicing and shrinkage artifacts

At this point, evidence suggests that there is information about different types of arbors contained in the density. However, the differences in the density distributions could be due to artifacts created by different laboratories or systematic differences in the data collection of different types of arbors. For example, axons in the cortex are much larger than other arbors and can span thousands of microns. The neurons in this data set are stained and traced after the brains have been cut into slices and therefore, most times, the edges of the neurons are sliced off. Because axons are larger, systematic slicing artifacts will be more pronounced for axons as more of their arbors will be sliced off. This artifact will certainly affect the distributions of all the axon parameters and most likely will affect the other arbors too. It is feasible that the scatter around the line or the differences in scatter between different types of arbors could be due to these artifacts.

In order to assess how much slicing and shrinkage artifacts might affect the data, completely reconstructed data provided by Eysel were manually sliced down to 200 um (a typical section thickness). After manual slicing, the arbors were shrunk

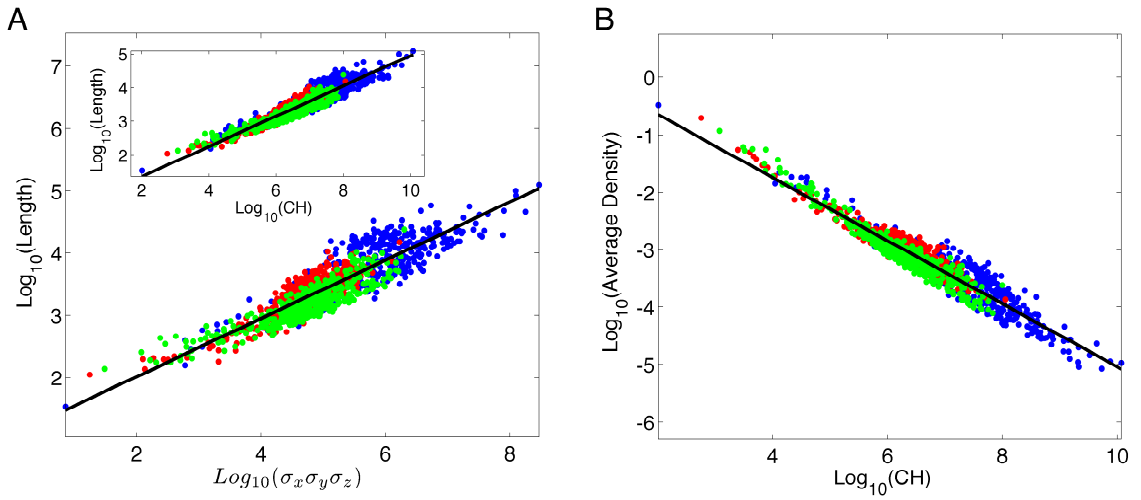


Figure 3.2: Scaling law between length and volume corresponds to a decrease in average density with arbor size. Apical dendrites are in green, non-apical dendrites are in red, axons are in blue. A. There is a power law relationship between the product of the standard deviations, $\sigma_x\sigma_y\sigma_z$, and the total length of the arbor. The regression line is fit to the combined, apical dendrite, non-apical dendrite and axon data. Slope = 0.47 ± 0.01, Intercept = 1.08 ± 0.06, r^2=0.79. Because the product of the standard deviations is proportional to the volume covered by an arbor as defined by the convex hull (CH) (Figure 3.1), the inset plot of the power law between the convex hull and total length looks nearly identical. Slope = 0.45 ± 0.01, Intercept = 0.454 ± 0.06, r^2=0.88. B. Because the power law relationship shown in A has an exponent less than 1 (the exponent of the power law corresponds to the slope of the linear regression), the average density, length divided by convex hull volume, decreases as the arbor gets larger. Slope = -0.55 ± 0.01, Intercept = 0.45 ± 0.06, r^2=0.91. All intervals are reported at 95% confidence.

by 25%. Although shrinkage corrections were applied to the data, the necessary corrections were difficult to assess. A value of 25% was chosen because it is probably similar to the shrinkage error that could remain in the data after corrections were applied. Figure 3.3 shows the completely reconstructed data, the sliced data, and the sliced and shrunken data. Slicing has the effect of shifting the data down the regression line to the left as both the area and the length of the neuron is reduced. This artifact affected axons more than the apical and non-apical arbors because they are much larger and therefore highly sliced. However, the slicing did not noticeably shift the arbors above or below the line. Shrinkage will increase the density of the

neuron as it will compact the same amount of length into a smaller area. This should have the effect of shifting the data to the left on the plot. In Figure 3.3, this minimal shift can be noticed, but, it does not appear that it could account for all of the scatter in the line (especially the cluster of dense axons). Therefore the scatter around the regression line most likely reflects a real relative average density range for arbors of a given size.

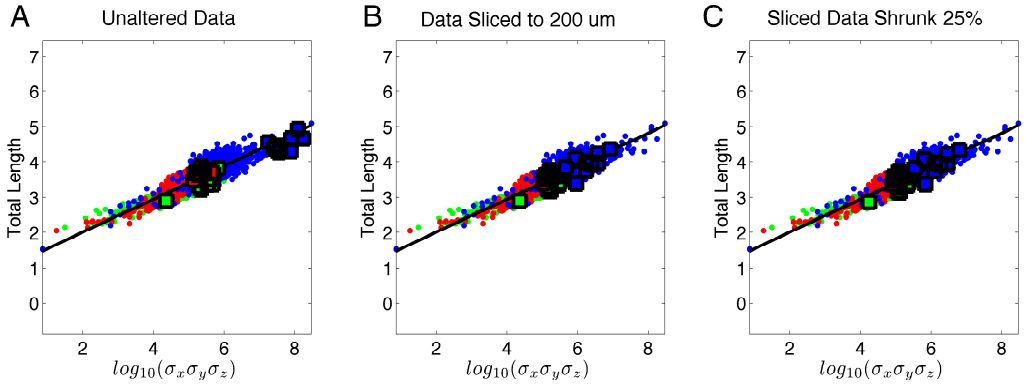
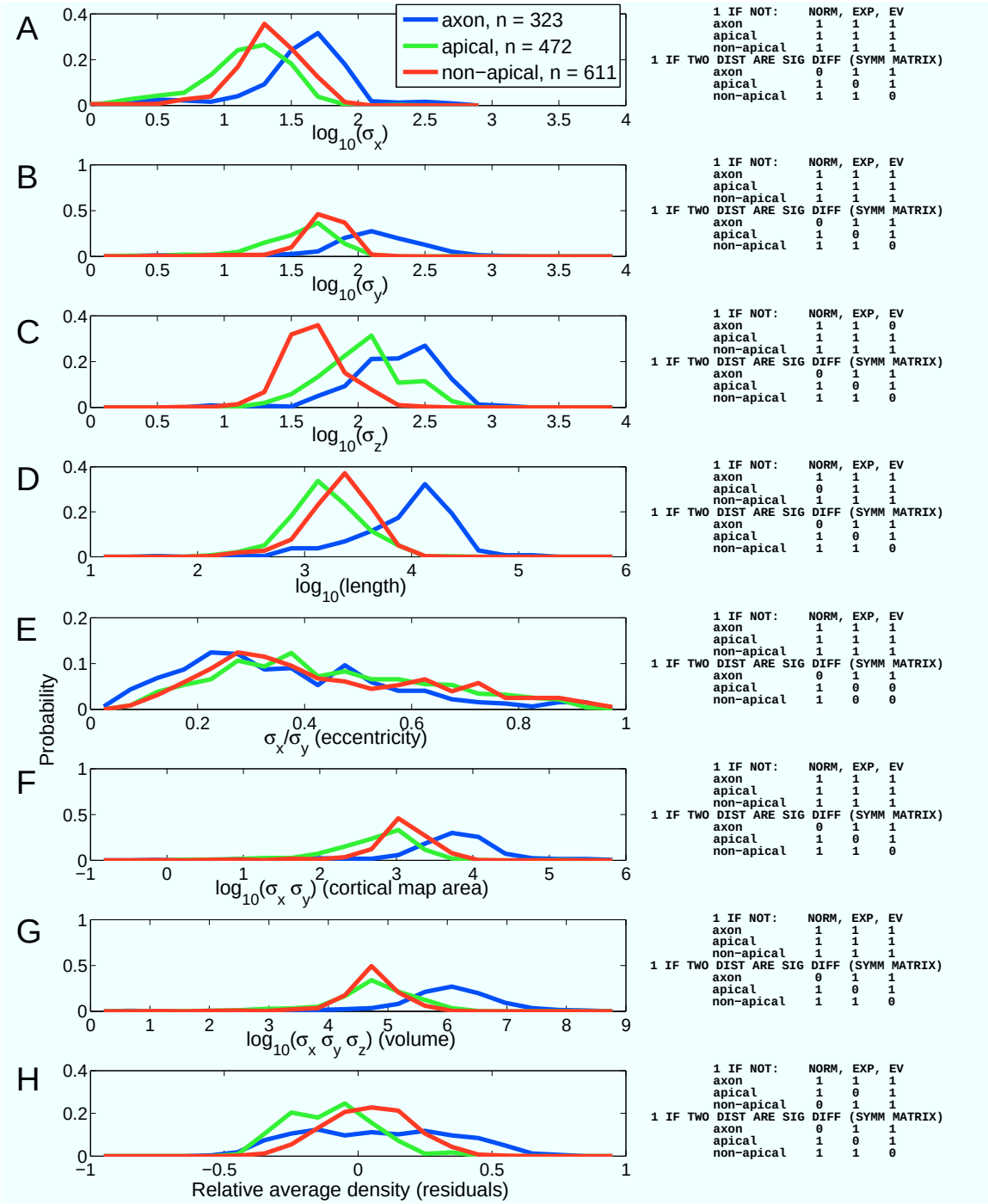


Figure 3.3: Effects of slicing and shrinkage on fully reconstructed cat pyramidal cell data provided by the Eysel lab. Axons are blue, non apical dendrites are red, and apical dendrites are green. Dots represent all the data, squares represent the data of fully reconstructed arbors. A. Unaltered data. B. Fully reconstructed arbors manually sliced to 200 um. This artifact shifts the data down the regression line by reducing both the product of the standard deviations and the total length. C. Sliced neurons then shrunk by 25%. Artifact has the effect of very slightly shifting the data left by reducing the area containing a specified length.

Figure 3.4: Normalized parameter histograms of apical dendrites, non-apical dendrites and axons of arbors from rat, cat, monkey and human. Apical dendrites are in green, non-apical dendrites are in red, axons are in blue. All significance tests are done at 95% confidence. Lilliefors tests were used to test if each individual distribution is significantly different from a normal, exponential, or extreme value distribution. Kolmogrov-Smirnov tests were used to determine if the distributions are significantly different from one another. Matrices to the right show the results of these significance tests. A value of 1 denotes a significant difference, 0 denotes no significant difference found. Top matrices report whether a distribution is significantly different from a normal distribution (NORM), exponential distribution (EXP), or extreme value distribution (EV). Lower distributions report whether apical dendrites, non-apical dendrites, and axons are different from one another. The matrices are symmetrical and the columns of this matrix are in the same order as the rows: axon, non-apical dendrites, apical dendrites. A-C. Logarithmic distributions of the standard deviations σ_x, σ_y, and σ_z in the principle axes coordinate system. σ_z is the standard deviation along the principle axes most pointing toward the pial surface. The remaining two standard deviations describe arbor coverage over the cortical map. σ_y is defined to be larger than σ_x. D. Logarithmic distribution of the total length of the arbors. E. Eccentricity distribution characterizing how ovular or circular the arbors are over the cortical map. An eccentricity value of one corresponds to circular. Arbors get increasingly ovular as eccentricity values progress from one toward zero. F. The product of σ_x, and σ_y represents the coverage area of an arbor over the cortical map. G. The product of the three standard deviations is proportional to the volume an arbor covers (Figure 3.1). H. The relative average density of an arbor for a given size and is represented by the distance from the regression line shown in Figure 3.2. This measure is commonly known as the residuals. Zero is the mean density for an arbor of a given size and corresponds to data points which lie on the regression line. Numbers greater than 0 correspond to arbors that are more dense relative to the mean density for an arbor of a given size. These data lie above the regression line. Numbers less than 0 correspond to arbors that are less dense relative to the mean density for an arbor of a given size. These data lie below the regression line.


A
B
C
D
E
F
G
H
axon, n = 323
apical, n = 472
non−apical, n = 611
log10(σx)
log10(σy)
log10(σz)
log10(length)
σx/σy (eccentricity)
log10(σx σy) (cortical map area)
log10(σx σy σz) (volume)
Relative average density (residuals)
Probability
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 1 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 1 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 0
apical 1 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 0 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 1 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 0
non-apical 1 0 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 1 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 1 1 1
non-apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
apical 1 0 1
non-apical 0 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
apical 1 0 1
non-apical 1 1 0

3.2.3 Density differences between pyramidal cells and interneurons

There are differences in the relative density distributions between interneurons and pyramidal cells (compare Figure 3.5H and Figure 3.6H). Results show that the widening of the axon density distribution of the axon arbors is due to a cluster of interneuron axons that sit above the regression line (Figure 3.7). These arbors are in a cluster and are generally more dense than the rest of the arbors. Only the axons of the interneurons have a relatively high density: the non-apical dendrites of the interneurons do not appear to be different than the non-apical dendrites of the pyramidal cells. Although the relative average density distribution of pyramidal cell basal dendrites is technically significantly different from the interneuron dendrite distribution (Kolmogrov-Shmirnov test, 95% confidence), the arbors from both classes of neurons appear to be located in the same general location on the plot in Figure 3.7. Based on the current data and the knowledge that artifacts do exist, I am not confident in concluding that the relative density of non-apical dendrites of interneurons and pyramidal cells are different.

As a side note: one might have noticed that there are apical arbors in this interneuron data set (Figure 3.7 A & C). Apical neurons are not generally associated with interneurons. These 23 apical interneuron arbors all come from the Markram data set and are associated with 5 Basket cells, 1 Bipolar cell, 5 Bitufted cells, 1 Chandelier cell, 8 Martinotti cells and 3 Stellate cells. These arbors usually originate from the bottom of the neuron and travel down across layers toward layer VI (like an apical dendrite in the opposite direction). I am not sure if these arbors are mislabeled, if they are supposed to be a different interneuron arbor type, or if they are intentionally labeled as apical due to their morphological similarity to apical dendrites. I left them in the data set 'as is'. Looking at (Figure 3.7), I am not convinced that these interneuron apical dendrites could be differentiated from the pyramidal apical arbors.

Figure 3.5: Normalized parameter histograms of interneurons. Apical dendrites are in green, non-apical dendrites are in red, axons are in blue. All significance tests are done at 95% confidence. Lilliefors tests were used to test if each individual distribution is significantly different from a normal, exponential, or extreme value distribution. Kolmogrov-Smirnov tests were used to determine if the distributions are significantly different from one another. Matrices to the right show the results of these significance tests. A value of 1 denotes a significant difference, 0 denotes no significant difference found. Top matrices report whether a distribution is significantly different from a normal distribution (NORM), exponential distribution (EXP), or extreme value distribution (EV). Lower distributions report whether apical dendrites, non-apical dendrites, and axons are different from one another. The matrices are symmetrical and the columns of this matrix are in the same order as the rows: axon, non-apical dendrites, apical dendrites. A-C. Logarithmic distributions of the standard deviations σ_x, σ_y, and σ_z in the principle axes coordinate system. σ_z is the standard deviation along the principle axes most pointing toward the pial surface. The remaining two standard deviations describe arbor coverage over the cortical map. σ_y is defined to be larger than σ_x. D. Logarithmic distribution of the total length of the arbors. E. Eccentricity distribution characterizing how ovular or circular the arbors are over the cortical map. An eccentricity value of one corresponds to circular. Arbors get increasingly ovular as eccentricity values progress from one toward zero. F. The product of σ_x, and σ_y represents the coverage area of an arbor over the cortical map. G. The product of the three standard deviations is proportional to the volume an arbor covers (Figure 3.1). H. The relative average density of an arbor for a given size and is represented by the distance from the regression line shown in Figure 3.2. This measure is commonly known as the residuals. Zero is the mean density for an arbor of a given size and corresponds to data points which lie on the regression line. Numbers greater than 0 correspond to arbors that are more dense relative to the mean density for an arbor of a given size. These data lie above the regression line. Numbers less than 0 correspond to arbors that are less dense relative to the mean density for an arbor of a given size. These data lie below the regression line.

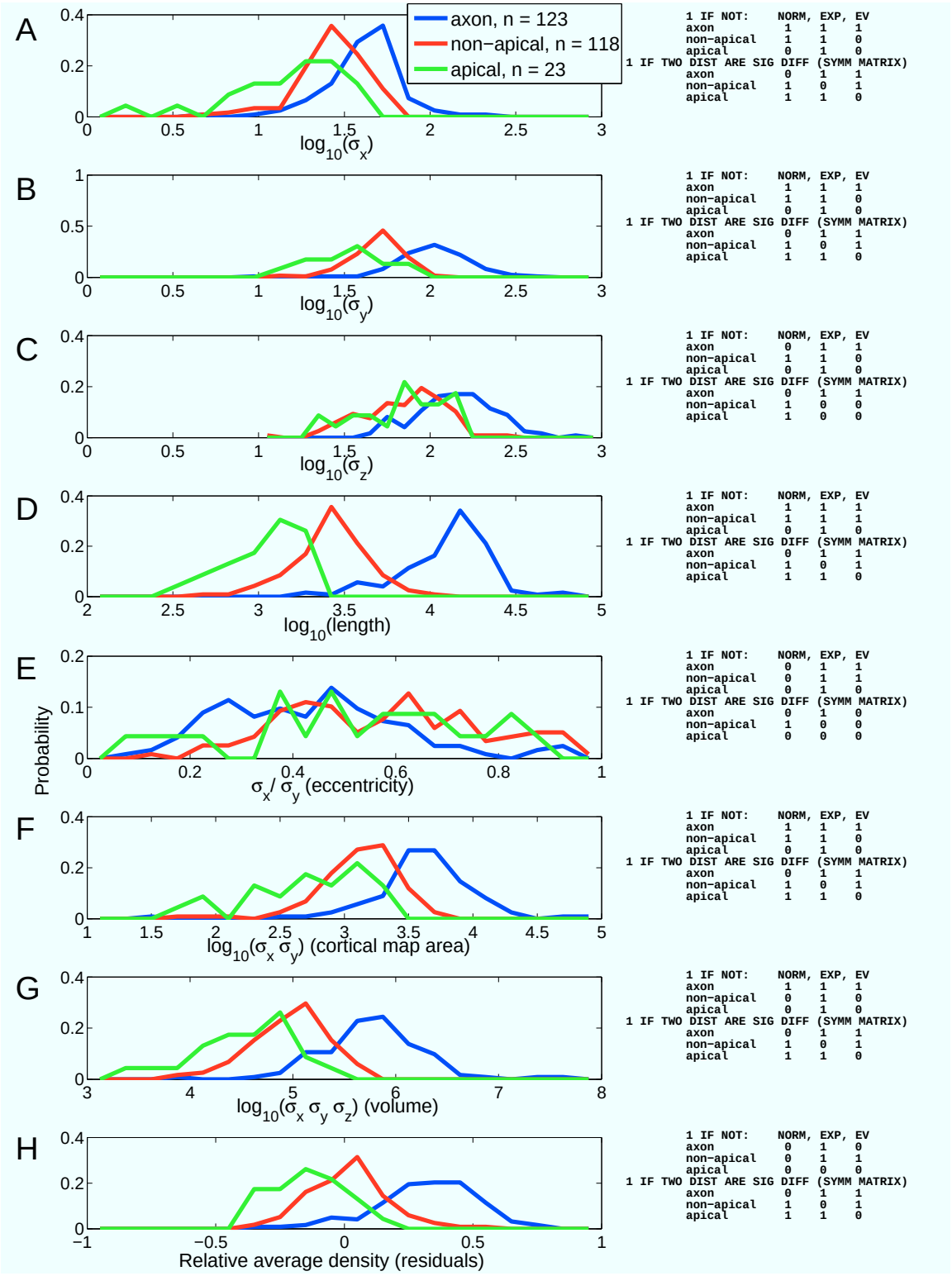
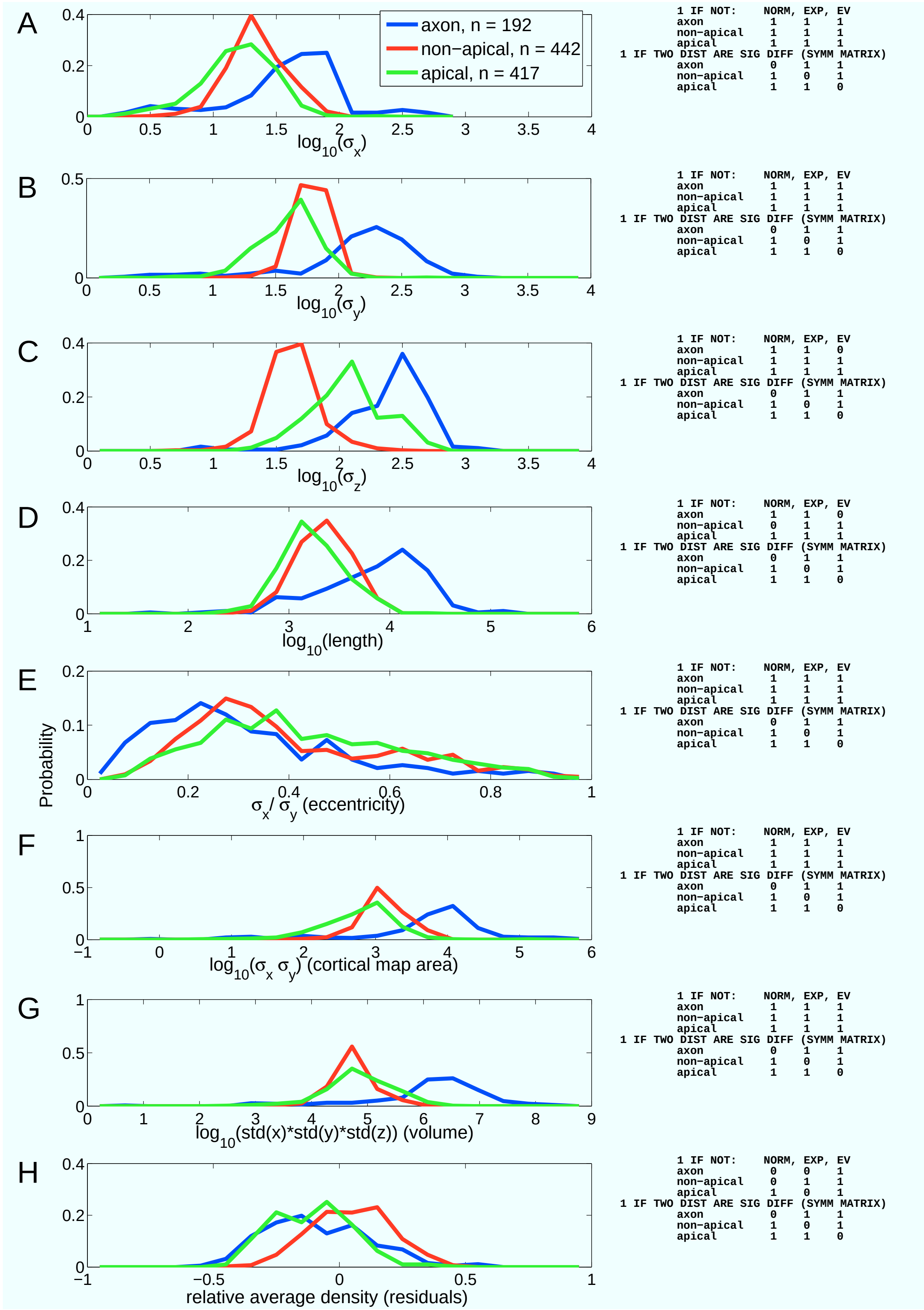

A
axon, n = 123
non−apical, n = 118
apical, n = 23
log10(σx)
B
log10(σy)
C
log10(σz)
D
log10(length)
E
Probability
σx/ σy (eccentricity)
F
log10(σx σy) (cortical map area)
G
log10(σx σy σz) (volume)
H
Relative average density (residuals)
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non-apical 1 1 0
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 1
apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non-apical 1 1 0
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 1
apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 0 1 1
non-apical 1 1 0
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 0
apical 1 0 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non-apical 1 1 1
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 1
apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 0 1 1
non-apical 0 1 1
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 0
non-apical 1 0 0
apical 0 0 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non-apical 1 1 0
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 1
apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non-apical 0 1 0
apical 0 1 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 1
apical 1 1 0
1 IF NOT: NORM, EXP, EV
axon 0 1 0
non-apical 0 1 1
apical 0 0 0
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non-apical 1 0 1
apical 1 1 0

Figure 3.6: Normalized parameter histograms of pyramidal cells. Apical dendrites are in green, non-apical dendrites are in red, axons are in blue. All significance tests are done at 95% confidence. Lilliefors tests were used to test if each individual distribution is significantly different from a normal, exponential, or extreme value distribution. Kolmogrov-Smirnov tests were used to determine if the distributions are significantly different from one another. Matrices to the right show the results of these significance tests. A value of 1 denotes a significant difference, 0 denotes no significant difference found. Top matrices report whether a distribution is significantly different from a normal distribution (NORM), exponential distribution (EXP), or extreme value distribution (EV). Lower distributions report whether apical dendrites, non-apical dendrites, and axons are different from one another. The matrices are symmetrical and the columns of this matrix are in the same order as the rows: axon, non-apical dendrites, apical dendrites. A-C. Logarithmic distributions of the standard deviations σ_x, σ_y, and σ_z in the principle axes coordinate system. σ_z is the standard deviation along the principle axes most pointing toward the pial surface. The remaining two standard deviations describe arbor coverage over the cortical map. σ_y is defined to be larger than σ_x. D. Logarithmic distribution of the total length of the arbors. E. Eccentricity distribution characterizing how ovular or circular the arbors are over the cortical map. An eccentricity value of one corresponds to circular. Arbors get increasingly ovular as eccentricity values progress from one toward zero. F. The product of σ_x, and σ_y represents the coverage area of an arbor over the cortical map. G. The product of the three standard deviations is proportional to the volume an arbor covers (Figure 3.1). H. The relative average density of an arbor for a given size and is represented by the distance from the regression line shown in Figure 3.2. This measure is commonly known as the residuals. Zero is the mean density for an arbor of a given size and corresponds to data points which lie on the regression line. Numbers greater than 0 correspond to arbors that are more dense relative to the mean density for an arbor of a given size. These data lie above the regression line. Numbers less than 0 correspond to arbors that are less dense relative to the mean density for an arbor of a given size. These data lie below the regression line.


A
axon, n = 192
non−apical, n = 442
apical, n = 417
log10(σx)
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non−apical 1 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
B
log10(σy)
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non−apical 1 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
C
log10(σz)
1 IF NOT: NORM, EXP, EV
axon 1 1 0
non−apical 1 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
D
log10(length)
1 IF NOT: NORM, EXP, EV
axon 1 1 0
non−apical 0 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
E
Probability
σx / σy (eccentricity)
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non−apical 1 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
F
log10(σx σy) (cortical map area)
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non−apical 1 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
G
log10(std(x)*std(y)*std(z)) (volume)
1 IF NOT: NORM, EXP, EV
axon 1 1 1
non−apical 1 1 1
apical 1 1 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0
H
relative average density (residuals)
1 IF NOT: NORM, EXP, EV
axon 0 0 1
non−apical 0 1 1
apical 1 0 1
1 IF TWO DIST ARE SIG DIFF (SYMM MATRIX)
axon 0 1 1
non−apical 1 0 1
apical 1 1 0

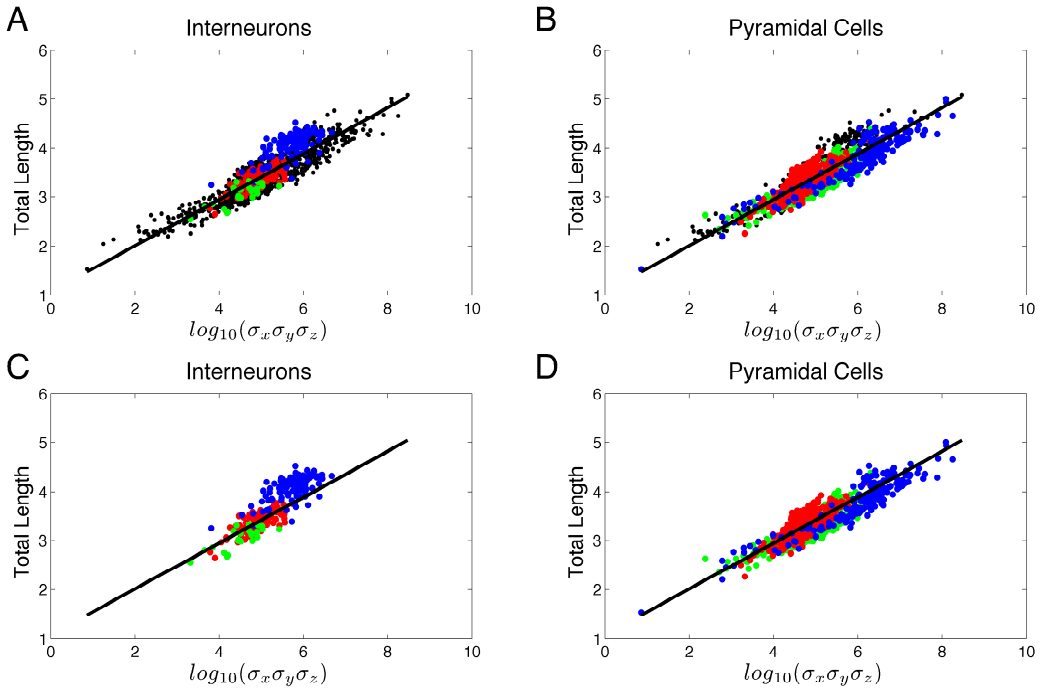


Figure 3.7: Interneurons and pyramidal cells are distributed around the regression line differently. Axons are blue, non apical dendrites are red, and apical dendrites are green. A. Interneurons in color are plotted over entire data set in black. B. Pyramidal cell in color are plotted over entire data set in black. C. Interneurons plotted alone. D. Pyramidal cells plotted alone.

3.2.4 Density summary

Unlike fish, there are density differences between different types of arbors. The relative average density distributions of different types of arbors are significantly different. In addition, different relative density strategies have been identified for arbors from different classes of neurons; the axons of interneurons are relatively more dense and clustered than other arbors. Therefore, unlike the fish, density cannot be neglected when characterizing or differentiating arbors. The minimum set of parameters needed to characterize the spatial density distribution function cannot be reduced from four to three parameters due to the power law.

3.3 Arbor shape as defined by the standard deviations

Now, I will explore the possibility that different types of arbors can be differentiated based on their shape. The relationship between the standard deviations will describe the arbor shape. The shape of cortical neurons can be described just as it was for fish, except fish required the consideration of just two standard deviations; cortical neurons require the consideration of three standard deviations, σ_x, σ_y, and σ_z. The orientation of cortical neurons in the brain is important as information flows in at least two directions: across layers, and across the cortical map. The principal standard deviations were oriented to preserve this biological meaning. The standard deviation that points most toward the pial surface and therefore represents how information is spread across cortical layers is defined as σ_z. The two remaining standard deviations will describe how a neuron spreads information across the cortical map. No information is available to orient σ_x and σ_y with respect to one another and so, as with fish, σ_y is defined as greater than σ_x. Please see the Methods section for further details on orientation.

3.3.1 Most arbors are eccentric

As with fish, eccentricity is defined as the ratio of σ_x to σ_y and will describe how the arbor distributes itself across the cortical map. The different arbor types have a similar range in eccentricities with values between perfectly circular, $\sigma_x/\sigma_y = 1$, to very eccentric with a σ_x/σ_y value of approximately 0.1. The eccentricity distributions of all the types of arbors are skewed toward ovular with peaks located at approximately 0.3 (Figure 3.4). This means that many of the arbors are approximately three times wider in one direction than in the other direction. The apical and non-apical dendrite eccentricity distributions are not significantly different; however, the axon distribution is significantly different from the other two arbor types (Kolmogrov-Smirnov test, 95% confidence). Because axons are much larger

than dendrites, and are much larger than the slice thicknesses, it is likely that the difference in the axon eccentricity distribution from the dendrites is an artifact caused by the axons being more sliced than the other two dendritic arbor types.

The eccentricity of interneurons and pyramidal cells could be different

It appears that the skew in the eccentricity distributions is largely due to the pyramidal cells as the eccentricity distributions of interneurons are not as skewed. Figure 3.6 shows that the pyramidal arbor eccentricity distributions are similar to the previously shown skewed distributions (Figure 3.4) with peaks located at approximately 0.3. In contrast to the pyramidal arbors, the interneuron arbor eccentricities are not found to be significantly different from Gaussian distributions and have means and standard deviations around 0.5: dendrites: 0.57 ± 0.19, apical dendrites: 0.52 ± 0.23, and axons: 0.43 ± 0.18 (Figure 3.7). All statistical tests are done at 95% confidence using a Kolmogrov-Smirnov test. These arbors are still eccentric but generally have a ratio of length to width of 2:1 as opposed the the 3:1 ratio of pyramidal cells.

Can artifacts account for the differences between pyramidal and interneuron eccentricities?

Unfortunately, the eccentricity distributions will be affected by slicing and shrinkage artifacts. Slicing could either make originally large circular arbors more eccentric or it could make originally eccentric arbors more circular. Differences between the interneuron and pyramidal cell eccentricity distributions could be due to artifacts. It is possible that the pyramidal cells and interneurons were subject to different experimental protocols resulting in different artifacts. In an attempt to assess whether the skewed distributions of the pyramidal cells are due to some sort of systematic artifact, I did two assays. First, I looked at the neurons of both very circular arbors ($\sigma_x/\sigma_y > 0.8$) and very eccentric arbors ($\sigma_x/\sigma_y < 0.2$) of the pyramidal cells to look for artifacts that may be causing the distribution to be overly skewed toward eccentric arbors. For example, maybe all the highly eccentric arbors are obviously sliced and

the skewed distribution can be attributed to slicing artifacts. Second, I looked at the eccentricity distributions of completely reconstructed data in the Eysel and Hirsch data sets to see if the distributions are the same as the larger data set. Both the Hirsch and Eysel data were collected from cat cortex. The Hirsch data consists of pyramidal and interneruons. The Eysel data consists of solely pyramidal cells.

The following observations of highly eccentric and highly circular pyramidal cell arbors suggest that artifacts are not the cause of the skewed pyramidal distributions.

1. Often a highly eccentric arbor was part of a neuron that also had a circular arbor providing evidence that slicing or shrinkage was not an issue in that neuron.
2. Many of the circular arbors were noticeably sliced on one side, therefore slicing is not necessarily causing the arbors to be more eccentric: slicing may also make some of the originally eccentric arbors more circular.
3. Many of both very eccentric and very circular arbors were reconstructed by the same laboratories. Therefore different artifacts generated by different laboratories is unlikely to be the cause.

The unsliced data set provides contradictory evidence against skewed pyramidal cell eccentricity distributions. Figure 3.8 shows the eccentricity distributions of the unsliced arbors. Although 20 unsliced arbors is too few arbors with which to obtain reliable eccentricity distributions, the pyramidal cell distributions do not appear to be as skewed. In addition, there are no highly eccentric arbors present in the unsliced data. However, since this data is solely from cat and the data set only contains 20 neurons, it is hard to know whether this data is representative of all arbors. The arbor data from the 9 fully reconstructed interneurons appear to match the larger data set. However, the number of unsliced interneurons is too small to perform any convincing statistical test to confirm.

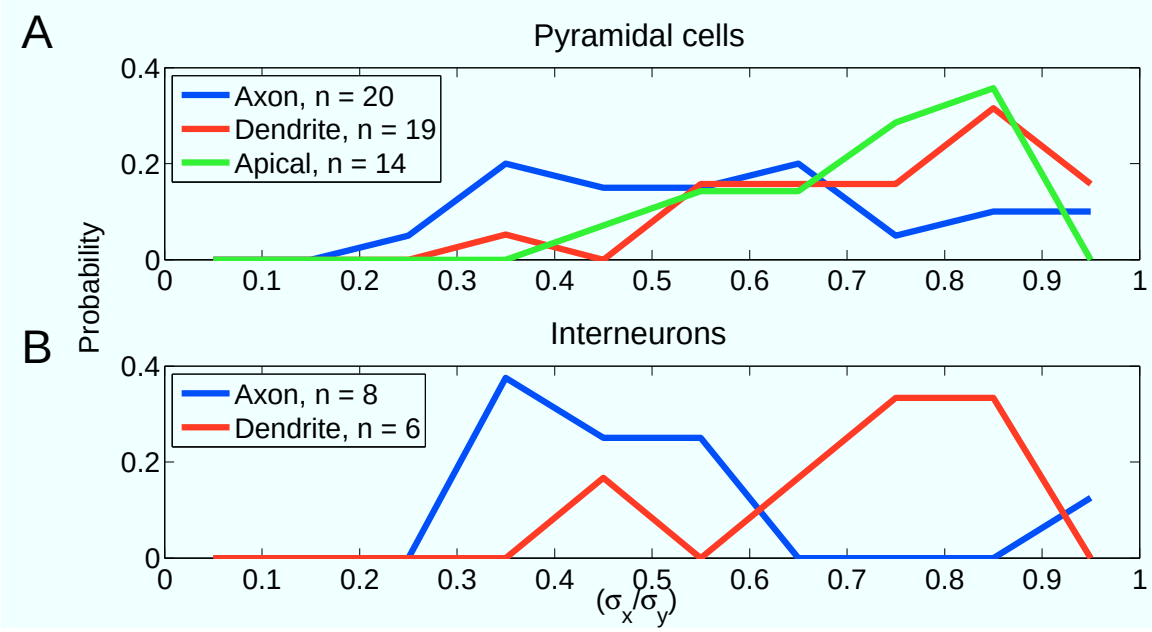


Figure 3.8: Normalized eccentricity distributions of completely reconstructed pyramidal cells and interneurons (filled in vivo, no sectioning artifacts). Eccentricity describes the aspect ratio of the arbor coverage over the cortical map (how circular or ovular an arbor is). Eccentricity is defined as the short axis standard deviation, σ_x, divided by the long axis standard deviation, σ_y, in the the non-pial axes directions A. Pyramidal cell data provided by Eysel and Hirsch. B. Interneuron data provided by Hirsch. Axons are blue, non-apical dendrites are red and apical dendrites are green. In the legend, n, denotes the number of arbors in the data set. Note that the dramatic difference between the eccentricity distributions of the interneurons shown in Figure 3.5E and the pyramidal cells shown in Figure 3.6E is not present here in the fully reconstructed data.

Eccentricity summary

Despite the artifacts in the data, I think it is safe to conclude that arbors are often eccentric. There is evidence to suggest that the eccentricities of interneurons and pyramidal cells are different. The distributions from pyramidal cells are highly skewed toward eccentric with a peak corresponding to a length to width aspect ratio of approximately 3:1. While the interneurons are still eccentric, the distributions are not significantly different from a Gaussian with a peak corresponding to an aspect ratio of approximately 2:1. However, a very small data set containing fully reconstructed arbors solely from the cortex of cat provides contrary evidence. Therefore, whether

pyramidal cells and interneurons have different eccentricity distributions can not be resolved with the current data. If a difference does indeed exist, it is intriguing to ponder the possible reasons and repercussions of this difference.

3.4 Correlations between volume, eccentricity and density

Just as I did for fish, I looked for relationships between the arbor shapes, sizes and arbor density. In the cortical data, no correlations between relative average density (distance from the regression line) and eccentricity or area were found for any of the arbor distributions. However, there are possible correlations between area and eccentricity (Figure 3.9). As axons get larger they tend to be more eccentric. The opposite trend appears to be true for all dendrites (apical and non apical): as dendrites get larger they tend to be more circular. Here, once again, I run into the problem of whether or not these correlations are due to artifacts such as slicing and shrinkage which will affect the area and eccentricities. The two opposite trends might be explained by the effect slicing or shrinkage has on arbors of different sizes. It is probable that as the axon arbors get bigger they are more sliced causing the bigger area = bigger eccentricity trend. However, with the dendrites, bigger arbor areas would probably be correlated with thicker slices because they have less area sliced off. Because the bigger arbors are less sliced they would be less eccentric.

Again, I will look to see if the same trends exist for non-sliced arbors from the Hirsch and Eysel data sets. Figure 3.9 shows the unsliced data overlaid on the entire data set. It appears that no trends exist in the unsliced apical and non apical dendrite data. It is less clear for the axon data. It appears there could be two classes of unsliced axon data: a small group and a large group. If indeed these are two separate groups of data it appears that both of these groups follow the same trend as the artifact prone data: larger axonal arbors are more eccentric. However it is possible that these two apparent groups are a result of the small data set and the

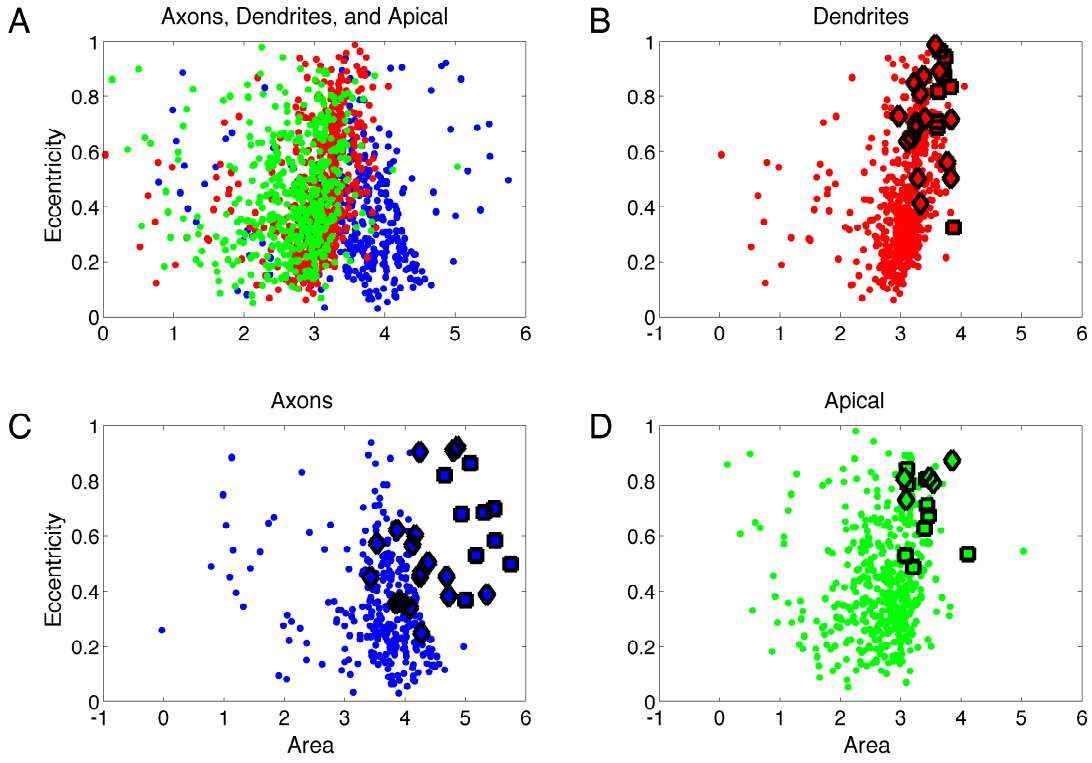


Figure 3.9: Possible correlations between cross-sectional eccentricity and area. Eccentricity is defined as σ_x/σ_y and area is defined as $log_{10}(\sigma_x\sigma_y)$, where σ_x and σ_y are the standard deviations in the x and y directions (non-pial directions) respectively. σ_y is defines to be greater than σ_x. Axons are blue, non apical dendrites are red, and apical dendrites are green. Completely reconstructed data of Eysel are in squares and Hirsch in diamonds (neurons were filled in vivo and do not have sectioning artifacts). A. All data plotted. Correlations appear to exist; however, this data is subject to slicing and shrinkage artifacts. B. Non-apical dendrite data are plotted. There is a trend of larger arbors being more circular in the larger data set. However, this trend is not apparent in the non-sliced data. C. Axon data are plotted. There is a slight trend of larger arbors being more eccentric apparent in the data subject to artifacts. It is unclear whether the non-sliced data follows the same trend. There may be a larger and a smaller class of non-sliced axons following the same trend. The larger class mostly consists of the Eysel data and the smaller class the Hirsch data. D. Apical dendrites are plotted. There is a trend of larger arbors being more circular in the data subject to artifacts that is not apparent in the non-sliced data. Without more completely reconstructed data, it is hard to resolve whether these trends are real or artifact.

data should be considered one group. If this is the case it would appear that there is no relationship between eccentricity and area for the completely reconstructed data. Again, evaluating whether these area versus eccentricity correlations are real will require a larger set of completely reconstructed arbor data.

3.5 Differentiating axons, non-apical, and apical dendrites using the Random Forests^{TM} supervised learning method

Through out this document, I have shown many distributions of the arbors parameters, many of which are statistically different between different arbor types and many of which are not. However, I have not considered all of the parameters at the same time. Figure 3.10 shows the data plotted with all the data dimensions in one graph. As with fish, plotting the standard deviations separately has the effect of expanding the regression line obtained from the plot of the logarithm of the product of the standard deviations verses the logarithm of the length. Because the fish have only two standard deviations to consider, the regression line expands into a plane and the scatter around the plane can easily be visualized in 3D space. However, cortical neurons have three standard deviations; therefore, the regression will expand into a three dimensional space and I will have four dimensions to consider.

Although four dimensional data are not quite as convenient to visualize, one can plot the standard deviations on the three Cartesian axes and represent the distance from the 3D regression space (scatter around the regression line corresponding to relative average density) as color intensity (Figure 3.10). Plotting the data in this way organizes the data and makes it easy to identify relationships or interesting phenomena. Distance between data points on the plot can be used to quantitatively describe how similar or different certain arbors are from one another. Figure 3.10 illustrates that different types of arbors are generally located in different regions of

the parameter space. However, the data are distributed continuously and the different regions do overlap. Thus, different types of arbors can only be differentiated probabalistically. There are no obvious clusters within the data insinuating distinct classes or types of arbors (Figure 3.10). Note that there are regions in the data space where no data points exist on the planes where σ_x=σ_z and σ_y=σ_z. These data-void planes are artifacts created by the use of principal components analysis (PCA) and should be ignored (Please see Methods Section for more information).

I have shown the distributions of each single parameter both to quantify their distributions and show how they differ from one another (Figure 3.4). However, to quantify how differentiable the three arbor populations are, I need a method that will take into account all four parameters at one time. Supervised learning techniques from the machine learning field are ideal for this type of problem. There are many different types of supervised learning techniques but they all aim find a method to most accurately differentiate predefined classes of data. They can take into account non-linearities, correlations between parameters, and ignore artifacts which are irrelevant to classification (such as the artifacts introduced by using PCA).

Of the many different types of supervised learning algorithms available, I chose the Random Forest^{TM} technique (Breiman, 2001a,b) due to its general high performance in comparison to other supervised learning methods. The Random Forest^{TM} algorithm was so named because it works by creating a forest of classification trees. For each data point, each tree votes for the class it thinks each data point belongs. The many tree voting procedure minimizes errors that could be created by using only one classification tree. Please see the methods section for more information on how this algorithm works.

3.5.1 The receiver operator characteristic (ROC)

To evaluate how well the Random Forest^{TM} algorithm can differentiate arbor types, I will use the receiver operator characteristics (ROC). The ROC curve is a plot of false positives versus true positives as the class discrimination threshold is

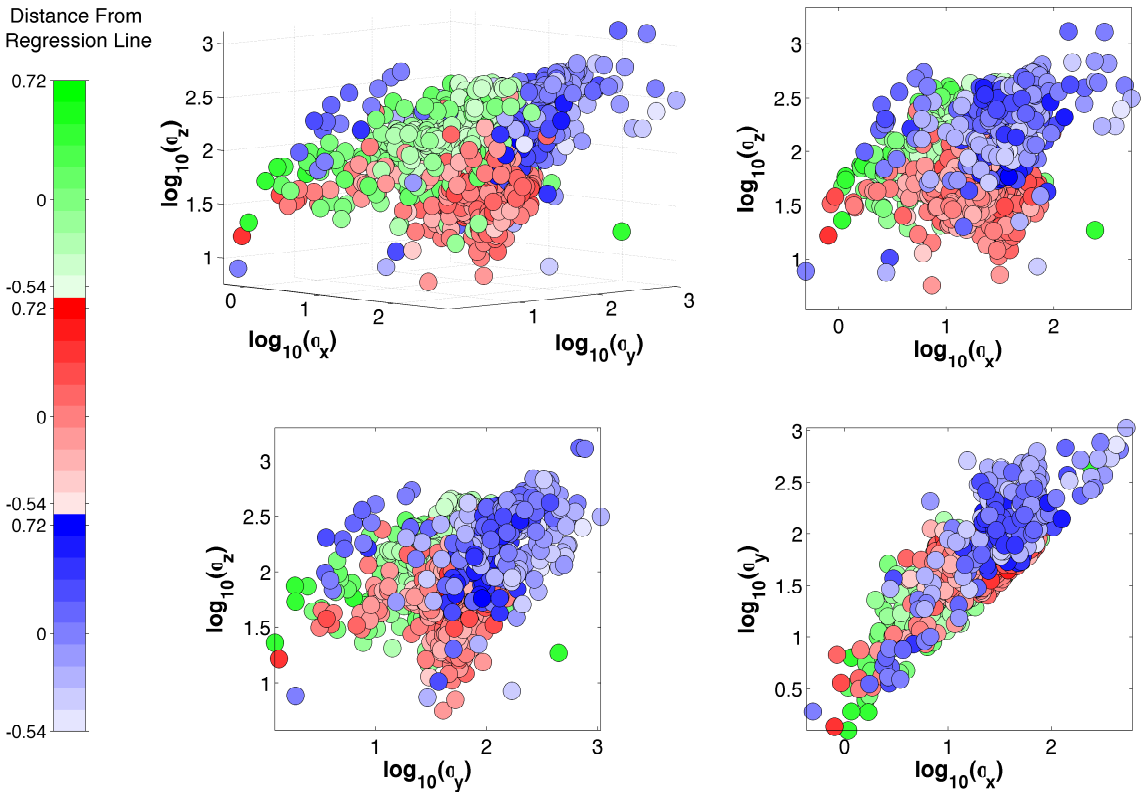


Figure 3.10: Apical dendrites, non-apical dendrites and axons reside in different regions of the parameter space. Four different views of the same four-dimensional arbor data of apical dendrites (green), non-apical dendrites (red), and axons (blue) from rat, cat, monkey and human. $log_{10}(\sigma_x)$, $log_{10}(\sigma_y)$, and $log_{10}(\sigma_z)$ are plotted on the three major axes. The forth dimension is the distance from the regression line (residuals) in Figure 3.2A and is represented by color shade. This value characterizes the average density of an arbor relative to the mean average density of many arbors of a specific size: the darker the color shade the higher the relative density. Zero is the mean average density for an arbor of a given size and corresponds to data points which lie on the regression line. Numbers greater than 0 correspond to arbors that are more dense relative to the mean density for an arbor of a specific size. These data lie above the regression line. Numbers less than 0 correspond to arbors that are less dense relative to the mean density for an arbor of a given size. These data lie below the regression line. The lines which cut through the data where $log_{10}(\sigma_x)$=$log_{10}(\sigma_z)$ and $log_{10}(\sigma_y)$=$log_{10}(\sigma_z)$ are artifacts from the use of principal components analysis (PCA). The artifact is caused because PCA will always find the maximum possible variance and very rarely will the value of the next largest orthogonal variance be exactly the same size as the first. Hence, the lines where the standard deviations are equal are bare of data.

varied (Figure 3.11 and Figure 3.13). ROC curves are commonly used to assess what classification thresholds are needed to give optimal classifications. The area under the ROC curve (AUC) evaluates the power of a classification method (Ling et al., 2003). Although percent correct might seem like the most intuitive measure of goodness, it is in fact threshold dependent and misleading in many cases. For example, say there are two groups of data: group 1 has 900 data points and group 2 has 100 data points. Assume these two groups are drawn from the same distribution and are thus indistinguishable. Any supervised learning algorithm aims to achieve the highest number of correct classifications and therefore will classify all the data points as group 1. The naive investigator would look at the resultant 90% percent correct and conclude that the two groups are quite distinguishable, when in fact, they are not. To correct this issue, when using a Random Forrest^{TM} one can change the threshold of votes required to vote for one class or another. In this case the threshold could be set to require 90% of the trees to vote for group 1 instead of a simple majority vote. The ROC curve plots the false positives versus true positives while altering these thresholds and thus provides a summary of overall performance.

3.5.2 Area under the curve (AUC)

Using the AUC to quantify how well an algorithm performs is very convenient because it eliminates the issue of choosing classification threshold values as mentioned in the previous section. The value of the AUC can range from 0.5 (concluding a random chance classification) to 1 (perfect classification). Note, that a less than chance classification (AUC < 0.5) actually means that a classification schema has been found: reversing the classification criteria would yield a classification accuracy higher than chance. Conveniently, it is possible to calculate the AUC without actually integrating the area under the ROC curve (Hand and Till, 2001). Because the AUC can be calculated independently of the actual ROC, and our purpose is solely to quantify how well different types of arbors can be differentiated (not to choose classification thresholds) the actual ROC curves are irrelevant for our analy-

sis. However, in order to provide intuition, I illustrate the ROC curves along with the values of the AUC.

3.5.3 Axons, apical dendrites & non-apical dendrites are mostly differentiable and continuously cover the range in parameter space

The most straightforward way to assess arbor differentiation is to do multiple binary classifications: each single arbor type is evaluated against every other arbor type. Figure 3.11, illustrates the ROC curves and AUC values for each pair of arbor types obtained from the Random Forests^{TM} algorithm. Indeed, the three arbor types are quite differentiable as evident from the large AUC values (Figure 3.11). Notice in Figure 3.10, that although the different types of arbors reside in generally different areas of the parameter space, all the arbor types are clustered together to form a continuous blob. There are no regions of the parameter space void of data (except for the σ_x=σ_z and σ_y=σ_z artifacts created by the use of PCA).

3.5.4 Summary

I conclude that different apical dendrite, non-apical dendrite, and axon types of arbors are generally restricted to different parts of the parameter space but completely cover the range of the parameter space. This means that different types of arbors generally have different spatial distribution strategies which are described by spatial adfs specified with different parameters. However, there is overlap between different types of arbors, therefore, different arbor types can only be differentiated probabilistically.

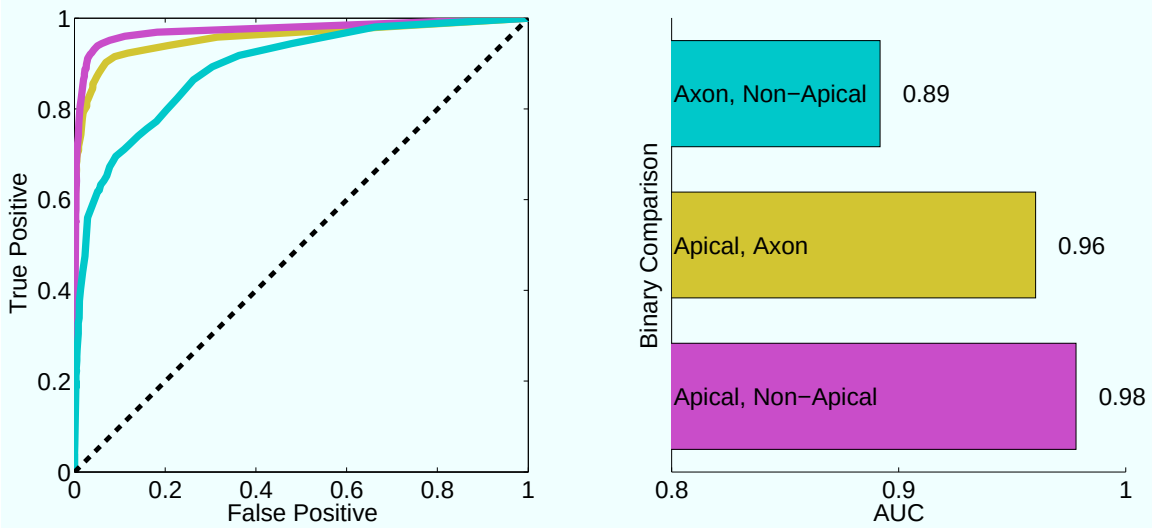


Figure 3.11: Axons, apical dendrites, and non-apical dendrites are largely differentiable. Left: Receiver operator characteristic (ROC) curves for binary classifications between axons, non-apical dendrites and apical dendrites obtained from the Random Forest^{TM} supervised learning method. Colors correspond to the labeled bar graph to the right. The closer the area under the curve (AUC) is to 1, the better the algorithm is at differentiating the two classes. Each curve is the average of results from running the Random Forest^{TM} algorithm 50 times (see Methods section). The dashed, black unity line denotes chance classification. Right: AUC values for ROC curves. AUC values and standard deviations are as follows: Axons versus non-apical dendrites, 0.8919 $\pm$ 0.0009, axons versus apical dendrites, 0.9604 $\pm$ 0.0010, apical dendrites versus non-apical dendrites, 0.9781 $\pm$ 0.0009.

3.6 Are the same type of arbors from different classes of neurons differentiable?

Previously, I showed that different arbor types are somewhat differentiable by the way they spatially distribute themselves regardless of species or neuron type. Next, I consider how well the same type of arbors from different classes of neurons can be differentiated. Do the same type of arbors from different classes of neurons have different spatial distribution strategies or are they the same? It is possible that the arbors of the same type but from different classes of neurons are undifferentiable and thus arbor types are the same for all classes of neurons. For example, perhaps all non-apical dendrites are the same, i.e. the basal dendrites of pyramidal cells are the

same as the dendrites of basket cell interneurons (although this is unlikely based on my previous discussion concerning the differing eccentricities of pyramidal cells and interneurons). Or, it could be the case that within the general arbor type, arbors from different classes of neurons reside in different regions of the parameter space (just as the different arbor types generally reside in different regions of the greater parameter space).

Ideally, it would be possible to utilize the plethora of data available on the Neuromorpho database to analyze many different types of arbors across species. However, different types of neurons from different species were often collected from different laboratories using different experimental protocols. As a result, variation between populations arise from systematic artifacts introduced by the acquisition of data from different laboratories. It is unclear how much variation is due to the artifacts or due to actual differences between the arbors. To circumvent this problem, I considered data provided by one laboratory. Out of the reduced data set, Henry Markram provided the largest number of neurons of various neuron class from rats of the same age range (Wang et al., 2002).

Non-apical dendrites from preclassified groups of neurons with a sufficiently large number of cells were selected for this analysis. The data consisted 30 Basket Cells, 15 Bitufted Cells, 30 Martiniotti Cells, and 72 pyramidal cells. All but the pyramidal cells are inhibitory interneurons and are found in layers II-VI. In the past, these different classes have been identified by morphology, electrophysiology, and gene expression profiles. Although morphological differences between dendritic arbors have been mentioned, these interneurons are most often distinguished by their axons, which target different regions of the cortex (different layers and within versus between multiple cortical columns) and neurons (soma and proximal dendrites versus distal dendrites) (Markram et al., 2004). Here however, I do not consider axons because there is not enough high quality axonal data available to adequately access differences. Instead, I will investigate the non-apical dendritic arbors so that the slicing artifacts which heavily affect the axonal arbors will not be an issue.

Details of the four different classes of neuron I will be considering follow:

Basket cells are interneurons which target the soma and proximal dendrites of both pyramidal neurons and interneurons (Wang et al., 2002; Markram et al., 2004). Wang et al. (2002) define three types of basket cells based on their axonal and dendritic morphologies: large basket cells (LBCs), small basket cells (SBCs) and nest basket cells (NBCs). Morphological classification of these subtypes was mostly attributed to differences between their axons. However, the authors do report differences between cell types in dendritic segment length, branch order, and the average length of the dendritic tree. Because of the limited data available, I will not investigate differences between these three subclasses of basket cells, but instead will lump them in to one class.

The dendritic processes of bitufted and martinotti cells emerge from the top and bottom poles of the soma. These cells are mainly differentiated by their axons. The bitufted axons spread horizontally and can expand past the boundaries of a cortical column. Martinotti axons project vertically toward layer I and can then spread over a distance of millimeters across many cortical columns. However, Martinotti cells are reported to have a more elaborate dendritic tree than most interneurons (Markram et al., 2004).

Finally, pyramidal cells are excitatory neurons. These neurons are characterized by their pyramidal shaped soma and the tall apical dendrite that extends up though the layers toward the pial surface. The basal dendrites of these cells originate from the lower side of the neuron and usually reside in a single layer.

Figure 3.12, shows the non-apical dendrites from the four different classes of neurons in the four-dimensional parameter space. Figure 3.13 shows the ROC curves and AUC values obtained from the Random Forrest^{TM} algorithm. Although the different groups are somewhat discernable, none of the groups are as differentiable as the general axon, apical, and non-apical dendrite classifications. This suggests that different classes of neurons may use slightly different spatial distribution strategies for their arbors but that the same arbor types between different classes of neurons

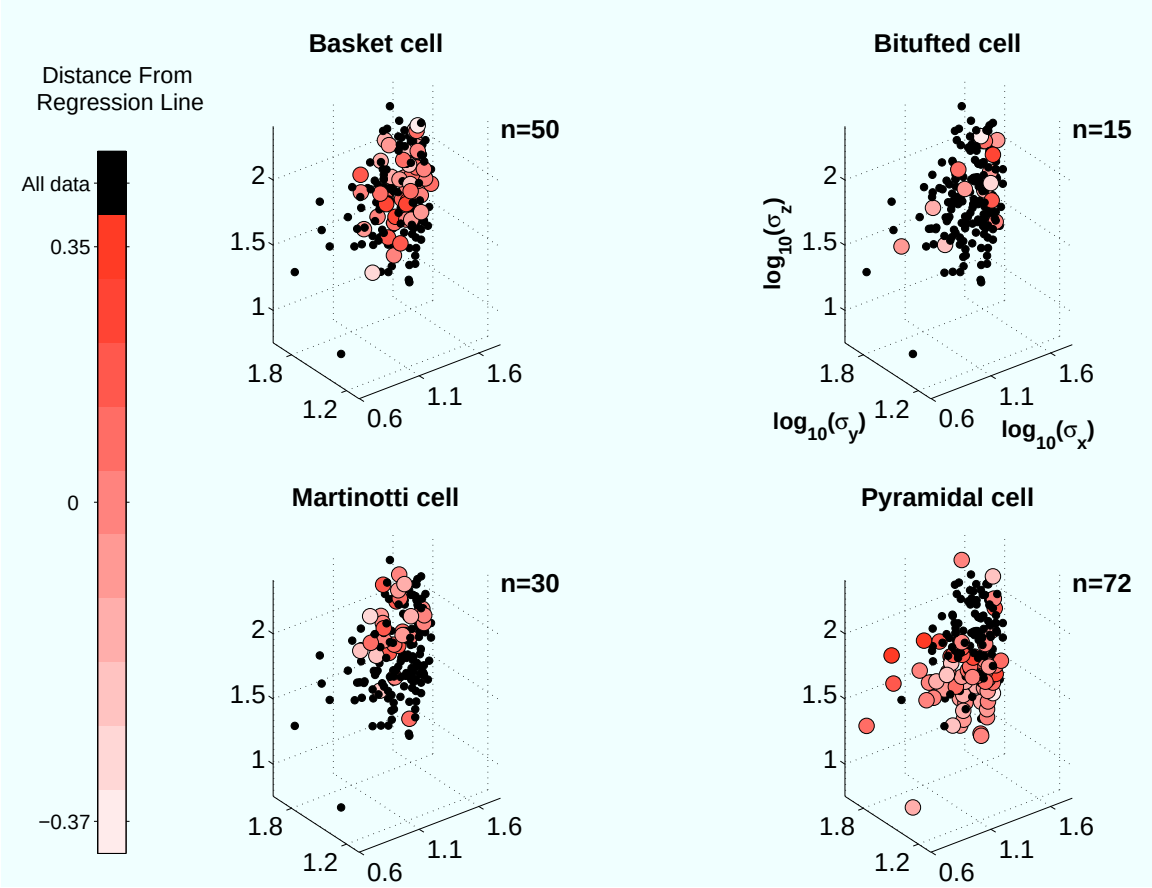


Figure 3.12: Basket, bitufted, martinotti and pyramidal non-apical dendrites from Markram data set plotted in four-dimensional parameter space. Black dots correspond to all the cell data points. The red data represent the data of the corresponding cell non-apical dendrites. σ_x, σ_y, and σ_z are plotted on the three major axes. The shade of the red circles represents the average density of an arbor relative to the mean average density of many arbors of a specific size. The relative density is characterized by the distance from the regression line (residuals) shown in Figure 3.2A (the darker the color shade, the larger the relative density). Zero is the mean density for an arbor of a specific size and corresponds to data points which lie on the regression line. Numbers greater than 0 correspond to arbors that are more dense relative to the mean density for an arbor of a given size. These data lie above the regression line. Numbers less than 0 correspond to arbors that are less dense relative to the mean density for an arbor of a given size. These data lie below the regression line.

are similar.

Notice that the three largest values of the AUCs are between the pyramidal and interneuron classes. Figure 3.12 shows that the basal arbors are smaller than the interneuron arbors: all of the standard deviations of the pyramidal basal dendrites are smaller on average than the interneuron dendrites. This is evident as the basal dendrites are clustered in the lower region of the plot. Therefore, it seems that the spatial density distribution strategies between interneurons and pyramidal cells differ more than the spatial density distribution strategies within different classes of interneurons.

3.6.1 Cortical summary

In conclusion, I have provided evidence that the relative density of cortical arbors contains information concerning the spatial distribution strategies of different arbors. The distributions of the relative average density are significantly different for the three types of arbors. In addition, different density strategies have been identified for different classes of cells: interneuron axons have a higher relative density than the pyramidal cells. Therefore, the minimal set of parameters needed to describe an adf cannot be reduced to three parameters but instead requires four parameters: the three standard deviations and the total length of the arbor. The eccentricity distributions between pyramidal cells and interneurons may be different and there may be correlations between area and eccentricity; however, conclusions are hard to reach due to the slicing and shrinkage artifacts present in the data.

Apical dendrites, non-apical dendrites and axons largely have different spatial distribution strategies as they cover different regions of the parameter space and are generally differentiable based on the minimal set of parameters. However, these distributions do overlap and therefore different arbor types can only be distinguished probabilistically.

The same types of arbors from different classes of neurons show some differentiability but cover highly overlapping regions of the parameter space. Therefore,

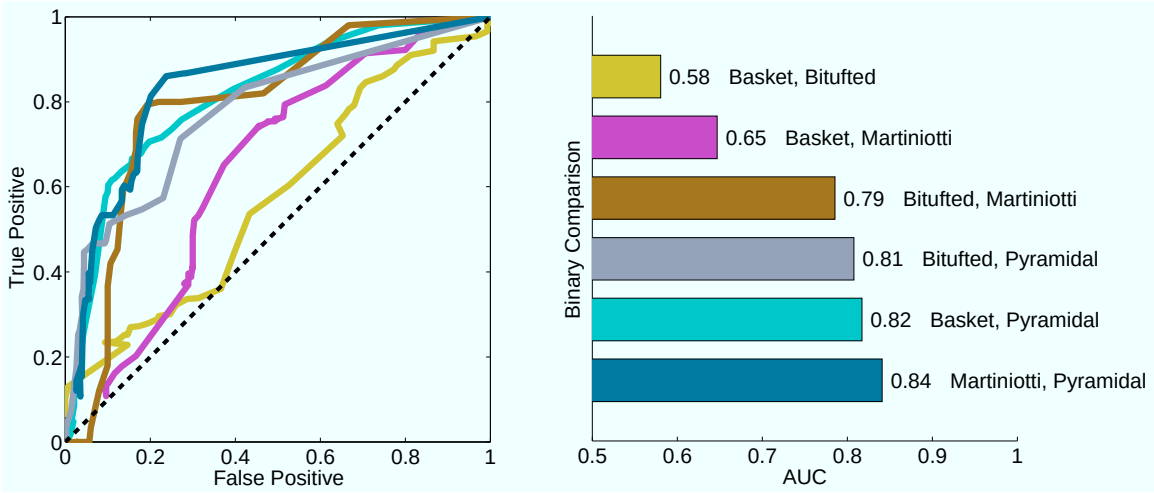


Figure 3.13: Non-apical dendrites of pyramidal cells are more differentiable from interneurons than different classes of interneurons from one another. Left: ROC curves for classification of non-apical arbors of different classes of neurons (from the Markram data set). ROC curves for every pair of arbor types from Random Forest^{TM} supervised learning method. Neuron classes considered were basket, bitufted, martinotti and pyramidal cells. Colors correspond to the colors in the bar graph to the right. Pyramidal cells are excitatory neurons. Basket, bitufted, and martinotti cells are inhibitory interneurons. The dashed black unity line denotes chance classification. Right: AUC values for the ROC curves. AUC values and standard deviations are as follows: basket versus bitufted, 0.5807 $\pm$ 0.0125, basket versus martiniotti, 0.6470 $\pm$ 0.0090, bitufted versus martiniotti, 0.7857 $\pm$0.0108, bitufted versus pyramidal, 0.8080 $\pm$0.0083, basket versus pyramidal, 0.8174 $\pm$0.0050, martiniotti versus pyramidal, 0.8411 $\pm$ 0.0040. The three smallest AUC values are between different classes of interneurons. The largest three AUC values are between pyramidal cells and various classes of interneurons. Note that all of these values are less than the AUC values defining how differentiable different types of arbors are in Figure 3.11. Therefore, non-apical dendrites of different neuron classes are less differentiable than different arbor types.

it is likely that within an arbor type there are somewhat different spatial distribution strategies for different neurons; however, these different classes are not nearly as differentiable as the different arbor types. The highest differentiability occurs between the basal dendrites of pyramidal cells and the different classes of interneuron dendrites.

Chapter 4

Conclusions

4.1 Summary of fish retinal ganglion cells

1. Both axons and dendrites can be described by the same scaling law relationship between the product of the standard deviations and the total length of an arbor. Because the product of the standard deviations is proportional to the area an arbor covers, there is also a power law relationship between the area an arbor covers and the total length of an arbor (Figure 2.1A).

2. As all arbors get larger, their average density decreases (Figure 2.1).

3. The relative average density (scatter around the regression line fit to the $log(L)$ versus $log(\sigma_x \sigma_y)$ power law relationship) is constant (Figure 2.1A).

4. The relative density of both axons and dendrites are distributed as Gaussians around the regression line (Figure 2.3D). In addition, the distributions of axons and dendrites are not significantly different from one another. This means that the minimal set of parameters needed to describe the spatial density function of these arbors can be reduced to two of the three parameters: the two orthogonal standard deviations and the total length of the arbor. The power law relationship will determine the remaining parameter.

5. The eccentricity distributions of axons and dendrites are not significantly different from one another and are not significantly different from a Gaussian (Figure 2.3C)

6. There are no correlations between eccentricity and area or between eccentricity and relative density (Figure 2.3).

7. There is no evidence of clusters within the data that would suggest different subclasses of arbors.

8. Axons and dendrites can only be differentiated from one another probabilistically based on their spatial density functions. Axons are generally bigger than dendrites. On average, axons have larger total lengths and standard deviations (Figure 2.2), have larger coverage areas (Figure 2.3E), and are less dense (Figure 2.1). Axons and dendrites have the same eccentricity and relative average density distributions (Figure 2.3 C & D).

4.2 Summary of cortical data

1. There is a power law relationship between the total length of an arbor and the product of the standard deviations that describe the spatial arbor density function (Figure 3.2A). The product of the standard deviations is proportional to the volume an arbor covers (Figure 3.1). Therefore, there is a corresponding scaling law between the total length of an arbor and the volume an arbor covers (Figure 3.2A).

2. As all arbors get larger their average density decreases (Figure 3.2B).

3. The relative average density is not constant along the $log(L)$ versus $log(\sigma_x\sigma_y\sigma_y)$ power law relationship. There is structure along the length of this line hinting at different types of arbors (Figure 3.2A).

4. The distributions of the relative densities (scatter around the power law) is not the same for apical, non-apical, and axonal arbors (Figure 3.4H). This means that the set

of minimal parameters needed to describe all arbors can not be reduced to three from four parameters. Information concerning arbor type is contained in arbor density. All four parameters (three orthogonal standard deviations and a total length) are needed to describe the spatial density function of cortical arbors.

5. The relative density distributions of each of the three different types of arbors are non Gaussian (Figure 3.4H).

6. Axons of interneurons are relatively more dense than the rest of the arbors (both arbors from the same class of neuron and those of the pyramidal cells) (Figure 3.7 and compare Figure 3.5H with Figure 3.6H).

7. Most arbors are eccentric (Figure 3.4E). There may be a difference between the eccentricity distributions of interneurons and pyramidal cells (compare Figure 3.5E with Figure 3.6E); however, slicing and shrinkage artifacts will affect these distributions. Artifacts in highly eccentric and highly circular arbors appear similar suggesting that the eccentricity differences between the two cell classes may not be due to artifacts. However, data from a few fully reconstructed cat arbors provide evidence that the eccentricities of pyramidal cells might not be as skewed as the larger data set suggests (Figure 3.8). More fully reconstructed data will be required to resolve this issue.

8. There are no correlations between relative average density and eccentricity or area (not pictured).

9. There are possible correlations between area and eccentricity and these correlations are different for different types of arbors. As axons get larger, they tend to be more eccentric. As apical and non-apical dendrites get larger they tend to be more circular (Figure 3.9). For the dendrites this relationship does not seem to hold true in the fully reconstructed data of cat. The fully reconstructed axonal data is uninterpretable as it is unclear whether there are two clusters with the same correlations or just one cluster without correlations.

10. The three different types of arbors (apical dendrites, non-apical dendrites, and axons) generally occupy different parts of the parameter space (Figure 3.10). However, there are not sharp boundaries between these different types. The Random Forest^{TM} classification algorithm demonstrates that these arbors are mostly differentiable (Figure 3.11). Therefore, different types of arbors can be generally differentiated by their spatial arbor density function, but because the adf parameter distributions overlap (Figure 3.4), the arbor types can only be differentiated probabilistically.

11. The same types of arbors from different classes of neurons (the dendrites of interneuron basket, bitufted, and martinotti cells and the basal dendrites of pyramidal cells) are not as differentiable as the different types of arbors (Figure 3.13). Unlike different types of arbors, different classes of arbors within an arbor type do not largely cover different regions of the parameter space (Figure 3.12).

12. The basal dendrites of pyramidal cells are more differentiable from the dendrites of the interneurons (Figure 3.13) than the different classes of interneuron dendrites are from one another.

Chapter 5

Discussion

The research in this thesis involves characterizing the spatial distribution of arbor branches in the context of a universal Gaussian density function. If the spatial density function of an arbor is shaped as a Gaussian a minimal set of parameters will be needed to describe an adf. These parameters will be the same parameters needed to describe a Gaussian function: the standard deviations (two for a flat 2D arbor, and three for a volume filling 3D arbor) and an amplitude, which, in this case, is equal to the total length of an arbor. In the following text I will refer to these parameters as the 'fundamental parameters'. In addition, a scaling law between the total length of the arbor (corresponding to the amplitude of the function) and the area or volume covered by an arbor (corresponding to the product of the standard deviations which define the Gaussian spatial adf) further constrains the parameters that will describe the adfs. In this document, I have defined the fundamental parameter distributions of different types/classes of arbors and have explored how well different types/classes of arbors can be differentiated based on the fundamental parameters.

5.1 Retinal ganglion cells

I began by considering the retinal ganglion cells (RGCs) of zebrafish and goldfish. These arbors are flat and tile the back of the retina. Their adfs can be

described by two standard deviations and the total length of an arbor.

In the current work, the fundamental parameters of fish RGCs were distributed homogeneously. Both axons and dendrites can be fitted by the same scaling law relating the arbor coverage area to the total length of an arbor (Figure 2.1A). The scatter of the data around best fit regression line (corresponding to the average density of an arbor relative to the mean density of an arbor of a given size) is distributed evenly along the regression line. The axonal and dendritic relative density distributions are distributed as Gaussians around the line and are not significantly different from one another (Figure 2.3D). In general, axons are larger and less dense than dendrites. On average, all of the fundamental parameters are larger for axons (Figure 2.1) as well as the cortical map coverage area (Figure 2.3E). The average density decreases with arbor coverage area as the result of the length/area scaling law exponent being less than 1 (Figure 3.2). Axons and dendrites have the same distributions of eccentricity with peaks corresponding to an aspect ratio of approximately 2:1 (Figure 2.3 C). There was no visual evidence of clustering to suggest different classes of arbors within the axonal and dendritic arbor types based on their spatial distribution strategies.

5.1.1 Classification

Because retinal ganglion cells transmit information from the retina to the brain and perform visual computations, numerous studies have been devoted to their characterization and classification. Previous studies have classified RGCs using parameters related to parameters used in this study including arbor coverage area, density, and length (Akaishi et al., 1995; Amthor et al., 1983; Badea and Nathans, 2004; Bailes et al., 2006; Dacey and Petersen, 1992; Dunn-Meynell and Sharma, 1986; Hitchcock and Easter, 1986; Huxlin and Goodchild, 1997; Kolb et al., 1981; Kong et al., 2005; Sun et al., 2002). Specific cell types differ between species and studies identify anywhere between 4 and 20+ different classes of RGCs. RGC size has been linked with functional cell classes. In the macaque, the parasol and midget cells,

so named because of their difference in dendritic arbor coverage area (Dacey and Petersen, 1992; Squire, 2003), have different functional properties (midget cells are sensitive to color and parasol cells detect contrast) and project to different layers of the lateral geniculate nucleus (Squire, 2003).

In the current data, I do not see any evidence to support different cells types on the basis of area or length (or density which is defined as length divided by area). It is possible that I did not have enough data to resolve clusters (79 axons, and 76 dendrites). Or, it is likely that arbor length and coverage area (and consequentially density) alone can not differentiate different types of arbors. Kong et al. (2005) classify the RGCs of mouse based on area, density, and stratification depth. In Figure 2B, of the Kong et al. (2005) paper, area versus density is plotted. The data of Kong and colleagues has a power law relationship and closely matches the data of the current research. For easy comparison, I have plotted the current area versus density data in linear space in Figure 5.1 (the logarithmic version of this data is plotted in Figure 2.1). It does not appear that the different clusters found by Kong et al. (2005) can be differentiated solely based on area and length. In addition, upon visual inspection of the different classes of arbors found in the literature, different classifications do not appear obvious based on area or density. Viewing RGC arbors, it appears that there is a continuum of arbor coverage area and that larger arbors are less dense on average, just as the scaling law between area and density demonstrates (Sun et al. (2002), Figure 1; Kong et al. (2005), Figure 16; Hitchcock and Easter (1986), Figure 2; Akaishi et al. (1995), Figure 11; Huxlin and Goodchild (1997), Figures 3, 7, 8, 10, & 11; Badea and Nathans (2004), Figures 13C, 14B, 15, 16G, 17E, 18D; Mangrum et al. (2003), Figures 3, 4, & 5; Dunn-Meynell and Sharma (1986), Figures 6, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21; Amthor et al. (1983), Figures 2, 3; Kolb et al. (1981) Figures 8, 10, 11, 12, 13, 14, 15, 16; Dacey (1993), Figure 13).

Most studies use additional parameters in their classification schema such as the arbor stratification depth in the inner plexiform layer (IPL). Stratification

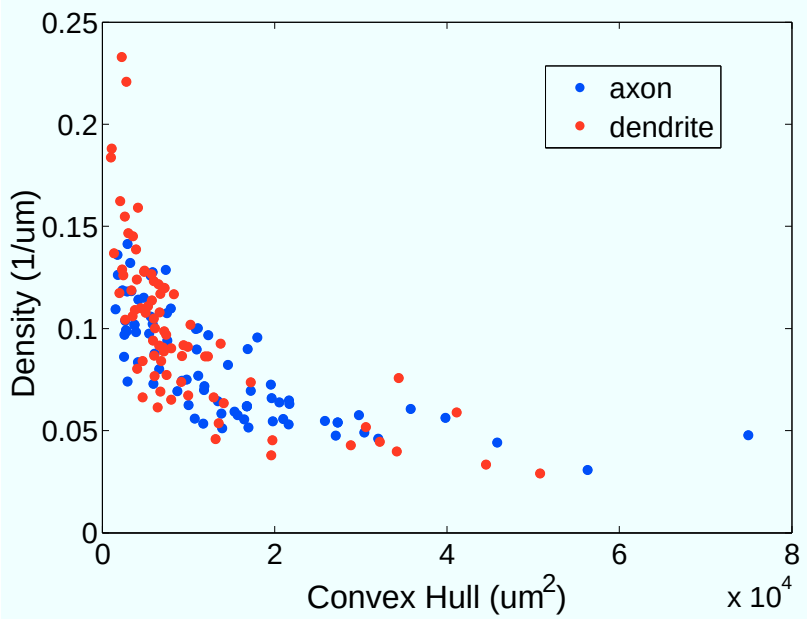


Figure 5.1: Power law relationship between area and density of fish retinal ganglion cells shown in linear space. Same plot as shown in Figure 2.1 but here it is plotted in linear space for comparison with Kong et al. (2005), Figure 2B.

depth has been shown to have function significance in the retina. For example 'on' and 'off' center surround cells are found in different regions of the IPL (Famiglietti and Kolb, 1976; Famiglietti et al., 1977; Nelson et al., 1978). In addition, it appears that different cell classes can be differentiated by whether or not they form a non-overlapping mosaic tiling in the retina (Dacey, 1993; Vlgyi et al., 2009; Wassle and Boycott, 1991).

Arbor geometrical shape, i.e. eccentricity, was one focus of the current research and has not yet been mentioned. There is clearly a range in eccentricities of RGC arbors (current document, Figure 2.3C; Lui et al., 2009); however, eccentricity is not used as a basis for morphological classification. I found no mention of eccentricity in the RGC classification literature. (Note: in the retinal literature eccentricity is commonly used to refer to a neuron's distance from the fovea. Here, I use eccentricity to describe how ovular or circular the arbor coverage area is.) This is likely because eccentricity does not help differentiate different types of retinal ganglion cells. The

current data support a lack of information contained in RGC eccentricity. In Figure 2.3C, I showed that the eccentricity distributions of axons and dendrites were distributed as Gaussians and not significantly different from one another. I saw no reason to suspect that different types of arbors could be differentiated by eccentricity. In addition, it has previously been found that the geometric shape of RGCs are irregular and that their eccentricity is the result of optimizing the RGC mosaic by maximizing RGC retinal coverage but minimizing overlap (Wassle et al., 1981; Liu et al., 2009). Therefore, it is unlikely that arbor eccentricity is related to different classes of neurons.

5.2 The relationship between a universal function and branching laws

It seems, there must be some relationship between the arbor branching statistics and the adfs; after all, an adf describes the way in which an arbor spreads its branches over space. However, it is hard to deduce the specifics this relationship. It is possible that different branching rules could create the same arbor density functions.

One characteristic of a Gaussian function is that the higher order statistics (the statistics are also referred to as moments: mean standard deviation, kurtosis, etc.) are completely characterized by the mean and the standard deviations (the 0th and the 2nd moments). In the Snider et al. (2010) paper, the authors showed that the higher order arbor moments (up to 20 orders) could be described by the 0th and 2nd order moments. In fact, Snider and colleagues found that only one moment was needed to describe the higher order moments because a power law relationship constrained the relationship between the 0th (total arbor length) and 2nd (arbor area or volume) order moments. Larger moments resolve higher frequency components of the arbors and at some point approach the resolution of the individual branches. Moments high enough to resolve the arbor branches are not considered in the Snider et al. (2010) paper because the authors were interested in the basic shape of adfs.

Snider and colleagues considered up to 20 moments which corresponds to a resolution of approximately 1/20 of the arbor diameter: too course to resolve properties of individual arbor branches.

Theoretically, if all of the moments of a single branching arbor were known, the entire arbor branching structure could be reconstructed perfectly. However, it is impossible to fully reconstruct the branching structure of a single arbor from the moments of an arbor density function which only generally describes the way in which an ensemble of neurons spread their branches over space. The tree-like structure of an arbor can be viewed as a single sample taken from an adf. Therefore, a potentially infinite number of branching structures could be drawn from the exact same density function. This is why all arbors do not look exactly the same.

Interestingly, research conducted by Kong et al. (2005) provides evidence that the branching statistics of an arbor do not provide useful information for the classification of retinal ganglion cells. Parameters equivalent to the 0th and 2nd order moments provided the best classification results. Kong and colleagues attempted to classify neurons using many different measured parameters. Many of these parameters were concerned with the specifics of branching such as the number of branch points, dendritic segment lengths, branch angles, and segment tortuosity. Originally they measured 42 parameters many of which they found to be uninformative due to redundencies. They whittled the data set down to 26 parameters (Kong et al. (2005), Table 1). Many of these parameters were correlated with one another and they found that using correlated parameters diluted the clustering results.

In the end, Kong and colleagues chose just three parameters which they thought would be maximally informative to cluster the data. They choose the arbor density, area, and stratification depth within the inner plexiform layer (IPL). Interestingly, two of three of these parameters are made up of the fundamental parameters that would specify a Gaussian spatial adf. The final parameter, stratification depth specifies the location of the arborization.

Other studies provide evidence for universal properties of neural branches.

Binzegger et al. (2005), found find that the axons of cat neocortical and thalamic neurons are topologically self-similar and that a simple random branching model can predict the bifurcation ratio, length ratio and collateral length distribution of the axonal arbors. Wen et al. (2009) found that smaller sections of an arbor are statistically similar to a scaled-down version of the whole arbor (self similarity).

A minimal set of parameters supports a simple design principle for arbor structure where the location of the arbor is specified along with the standard deviations defining the arbor size and shape. The question of whether the arbor density functions are being regulated or whether they are the result of some set of branching rules will be fundamental to understanding how the brain develops, establishes connections and processes information. It is likely that transcription factors play a role in arbor size and density (Parrish et al., 2007). A mechanism such as the Down syndrome-related cell adhesion molecule (DSCAM) is also likely to be involved in the distribution of arbors over space. DSCAM has been shown to be necessary for self avoidance and to obtain regular RGC mosaics in the retina (Fuerst et al., 2008, 2009; Hattori et al., 2008; Parrish et al., 2007).

5.3 Cortical arbors

Cortical neurons are not as homogenous as retinal ganglion cells. All of the arbor types (apical dendrites, non-apical dendrites, and axons) can be fitted to a single power law; however, there is more structure in the data suggesting different morphological arbor types or classes (Figure 3.2). In the following text, and thoughout the entire document (although previously, it hasn't been explicitly mentioned), I use the word 'type' to refer to the different apical dendrite, non-apical dendrite, and axon arbor types. I use the word 'class' to refer to different classes of neurons. There may be different arbor classes within an arbor type which correspond to different classes of neurons. For example, perhaps the axons of interneuron and pyramidal cells can be differentiated from one another. If this is the case, there would be two classes of

axonal arbors which would correspond to the interneuron and pyramidal cell neuron classes.

As discussed below (and in the Cortical Results chapter), although the data can generally be described by a scaling law (which constrains the volume an arbor covers for a given length), there is scatter in the relationship that corresponds to the relative density of an arbor. There is information concerning both arbor type and class contained in the relative density of an arbor. This means that unlike the fish where the relative density was homogenous, we cannot reduce the minimal set of parameters needed to describe a cortical arbor from four to three parameters. Instead all three standard deviations and the total arbor length will be needed to describe the density function of an arbor.

It should be noted that there are many artifacts in this data set that will affect the adfs. These artifacts include slicing, shrinkage, and tracing errors. I made efforts to reduce the impact of these artifacts on the results by eliminating arbors with substantial errors and applying shrinkage corrections when necessary. However, errors remain in the data and the results within this document are certainly affected by these artifacts.

5.3.1 Arbor type

The non-apical dendrite, apical dendrite, and axonal arbor types have long been differentiated from one another based on morphology and functional properties. Dendrites receive information from the network. After the dendrites and soma integrate the information from many neurons, the axons distribute information to the network in the form of a discrete action potential. Apical dendrites are a special type of dendrite which are present on pyramidal neurons. These arbors usually grow from the top of the pyramidal neurons and project themselves across many layers of the cortex toward the pial surface. Pyramidal cells also have a basal dendrite which grows from the bottom of the pyramidal neurons and generally does not project across many layers. The basal dendrites are considered non-apical dendrites along with the

dendrites of stellate cells and interneurons.

Unlike the retinal ganglion cells, different types of arbors are generally distributed around the power law differently (Figure 3.4H). The scatter around the best fit power law corresponds to the relative average density of an arbor. The relative density distributions are not the same for any of the arbor types. This implies that average density carries information about arbor type and suggests that different types of arbors may have a different control mechanism for relative density. The relative average density of non-apical dendrites is distributed normally around the best fit line, but the apical dendrite and axonal distributions are non-Gaussian. The range of axon relative density is larger for axons than it is for dendrites. The relative density of non-apical neurons is normally distributed. Apical dendrites are relatively less dense than the other arbors.

In general, axonal arbors cover a larger volume than do dendritic arbors. On average, each of the fundamental parameters are larger for axons (Figure 3.4A-D). They have a larger volume (Figure 3.4G) and also span a larger region of the cortical map (Figure 3.4D). There is some evidence to suggest that axons cover a slightly more ovular region of the cortical map. However, the difference is slight and I suspect is due to slicing artifacts that will affect the axons more than the dendrites due to their larger volume.

It appears that in general, non-apical dendrites have a larger cortical coverage area than apical dendrites (Figure 3.4F). This is because both standard deviations in the direction of the cortical map are larger for the non-apical dendrites (Figure 3.4A & B). Note that the dendrites of interneurons are lumped together with the basal dendrites of pyramidal cells. The larger coverage area of the non-apical dendrites is most likely due the the larger spread of the interneurons over the cortical map (compare Figure 3.6F and Figure 3.7F). However, the standard deviation that describes how an arbor distributes its information across layers is much larger for apical arbors. Clearly this makes sense as apical arbors are known to reach up toward the pial surface and spread their arbors across the different brain layers.

When taking all four fundamental parameters into account, the three different arbor types can be differentiated from one another fairly well (Figure 3.11). However, the data from different types of arbors do overlap (Figure 3.10) showing that the spatial density distribution strategies are not completely independent from one another and that arbors can only be distinguished probabilistically.

5.3.2 Arbor class

Since the invention of the Golgi stain, investigators have been attempting to categorize neurons into different classes. Although the intricacies of how neuron classes or subclasses are identified and defined are still not agreed upon (Bota and Swanson, 2007), certainly different functional classes of neurons exist within the brain.

In the current data, there is evidence to support different classes of arbors corresponding to different classes of neurons within the same arbor type. Differences were found between two major cell classes: the interneurons and the pyramidal cells. Pyramidal cells are excitatory neurons and usually interneurons are inhibitory. The axons of interneurons are relatively more dense than the axons of pyramidal cells (compare Figure 3.7H & Figure 3.6H). In addition, it appears that interneurons cover a less ovular area of the cortical map than pyramidal cells do. It is possible that these differences are caused by artifacts. However, the eccentricity distributions of pyramidal cells and interneurons are substantially different (compare Figure 3.7E & Figure 3.6E) and it is reasonable to suspect that these differences may be real.

The eccentricity distribution of pyramidal cells is skewed toward eccentric arbors, with a peak corresponding to an aspect ratio of approximately 3:1. Why are pyramidal cells so eccentric? Is there a functional reason for this shape? It is possible that the high eccentricity is due to a mechanism much like in the retina where arbors are shaped to best fill space (Liu et al., 2009). However, this conclusion would bring up many questions. First, this would assume that there are mosaics of different classes of pyramidal cells. In the retina, there are overlapping mosaics of different classes of RGCs. These different classes look very similar to one another (Wassle and Boycott

(1991), Figure 17 A, B, & C); however, functionally RGCs of a certain class only connect with other RGCs of the same class (Vlgyi et al., 2009). This situation would be different in the cortex; although there could be overlapping mosaics of different classes of pyramidal cells, pyramidal cells do connect with their overlapping neighbors. However, perhaps a similar mechanism is used to ensure homogenous pyramidal cell coverage of the cortex.

But why would pyramidal cells be more eccentric than interneurons? In the retina, evidence suggests that eccentricity is related to the regularity of the mosaic. Bistratified cells have more irregular mosaics than parasol cells. Bistratified cells also have more eccentric coverage areas and a greater variance in their eccentricities than parasol cells. Liu et al. (2009), find that the difference in eccentricities between these two types of RGCs is as expected to optimize retinal coverage. Perhaps a similar mechanism is at work in the cortex which could explain the differences in eccentricities of interneurons and pyramidal cells? If this is the case, it would suggest that the initial grids of pyramidal neurons are more irregular than interneurons.

Finally, the Random Forest^{TM} algorithm was used to quantify how well different non-basal arbors could be differentiated from one another when considering their fundamental parameters simultaneously. These arbors were not nearly as differentiable as the general arbor types. The ability to differentiate some of the interneuron classes was not much better than chance. The best classification results were consistently between the pyramidal basal dendrites and the dendrites of the interneurons. This suggests that the different functional excitatory versus inhibitory arbors have the most different spatial information distribution strategies and that the dendrites between different classes of interneurons are similar in their distribution strategies.

Chapter 6

Methods

6.1 Fish: data collection and reconstruction

As described in Lee and Stevens (2007), Goldfish (Carassius auratus) were obtained from local pet stores and zebra fish (Danio rerio) were maintained in the Salk Institute animal facility. Fish were anesthetized with 0.1% trican methanesulfonate, and perfused through an intracardiac catheter constructed from a glass micropipette. For perfusion, phosphate buffered saline (0.8M PBS), followed by 4% paraformaldehyde was used. For DiI (Molecular Probes, Eugene, OR) staining of the retinal ganglion cell dendritic arbors and the retinotectal axonal arbors, the brain was removed from the fish, fixed in 4% paraformaldehyde for 2 hours, and a small crystal of DiI was inserted into the nick in the optic nerve. The brain was maintained in fixative and scanned with a confocal microscope (LSM510: Carl Zeiss, Thornwood, NY). Arbors in the confocal stack were reconstructed with Neurolucida (Microbrightfield, Williston, VT).

6.2 Cortical data

6.2.1 Data download and selection

All non-cultured cortical data available on March 12, 2009 were downloaded from the website http://neuromorpho.org/. The data were collected from 14 different laboratories and consisted of 6,290 neural arbors from 3,094 neurons of mouse, rat, cat, monkey, and human species. In addition, 19 neurons with 41 arbors were provided by Judith Hirsch (Hirsch et al., 1998, 2002).

Many arbors were excluded from this analysis due to the unrealistic nature of the arbor traces. Often an arbor was supplied in the data base but was severely truncated, consisting of only a few segments. Many artifacts existed in the data such as slicing artifacts, tracing errors, and incorrectly aligned branches. Arbors containing drastic errors were eliminated. Shrinkage was a common artifact and was corrected for as covered in a following section. Table 6.1 provides a summary of the data used and the applied shrinkage corrections.

6.2.2 Included neurons

Whether neurons were used in this study was based on a qualitative judgement call. Details of arbor selection and inclusion are discussed in the following sections. Briefly, neurons that were included in the study had no obvious tracing artifacts. Many neurons in the data set appeared as though they were not stitched together correctly or they had strange artifacts in the section direction (the direction that focuses down through the slice if one is looking at the section through a microscope) that distorted this dimension of the reconstructions: these were not used. Virtually all the neurons in the data set were filled in sections except for the Eysel and Hirsch data. As a result, the edges of most of the neurons were sliced off on one or both sides. I will refer to these artifacts as slicing or sectioning artifacts. How much of any given neuron was sliced off depended upon the thickness of the section, the size of the neuron and where the neuron was located in the section. Because of

the prevalence of slicing artifacts in the data, it was impossible to exclude all sliced arbors. However, neurons that were 'very sliced' were eliminated from the data set. What constituted 'very sliced' was a qualitative judgement which corresponded to the removal of more than approximately 1/4 of the spatial volume of the arbor.

6.2.3 Basic experimental data collection methodology

Although the exact experimental protocols and experimental aims vary between laboratories and research projects, a general explanation of the basic histology methods follow:

If experiments such as electrophysiology were performed with live tissue, before histology, the animal is sacrificed and the brain is excised and submerged in chilled artificial cerebral-spinal fluid (ACSF). Then the brain is usually cut into blocks. At this point the blocks may be cut into smaller sections that are between 200 and 400 μm in thickness using a vibratome. Then the electrophysiology or other such experiments are conducted after which the cells are usually injected with a dye such as biocytin. Finally, sections are fixed in a substance such as paraformaldehyde and histology protocols are applied as described below.

If there are no live tissue experiments needed, often the animal has received injections of tracers and is sacrificed by perfusion. The brain is then excised and fixed using a substance such as paraformaldehyde or emersed in Golgi-Cox solution if a Golgi stain is being utilized. Then, as mentioned above, the brain is usually first cut into blocks and then sliced into sections between 200 and 400 μm in thickness. Perhaps dyes are injected at this point. Then the tissue may be embedded with a material such as celloiden or paraffin. At this point, the tissue may be cut into even smaller sections of 50 μm. If the tissue is not embedded, the tissue must then be frozen and cut on a cyrostat or a frozen microtome. Then the sections (thick or thin) are mounted for viewing under a microscope. This can be done by whole mounting or by coverslipping. Usually the tissue undergoes a series of rinses in water and alcohols (referred to as dehydration) and is subjected to any materials necessary

for the staining procedure. If the tissue is whole mounted, it is put into a liquid solution such as alcohol and covered with a slide. If the tissue is coverslipped, it is pressed onto a slide. The tissue may or may not be air dried using a hot plate. A slide (coverslip) is then glued to the top of the tissue.

6.2.4 Neuron reconstruction

After the neurons are mounted on slides they are traced under a microscope using a computer program such as Neurolucida. It is difficult to trace the neurons though thick sections because the view becomes 'muddy' deep within the section. Therefore, neurons that were traced in thick sections most likely have errors associated with difficult viewing. This is why after dye injection some investigators resection the tissue down to thinner sections such as 50 um. This makes tracing easier; however, now the reconstructions have to be 'restitched' back together. This is most likely the source of many of the 'stitching' errors present in the neuron traces.

6.2.5 Artifacts

Artifacts were prevalent in the neural reconstructions. The following is an explanation of common artifacts found. Many neurons were eliminated based on these artifacts. 1. Incomplete filling. In this case, the neuron was not completely filled with dye and therefore could not be fully reconstructed. 2. Tracing past the boundaries of the section due to over-focusing. The z-coordinate (Figure 6.10) location of each piece of neurite is identified by scrolling through the section under a microscope. The location where the dye is the brightest or most intense is denoted the z-coordinate of the neurite at that position. Often it is difficult to discern where the dye is most intense. Therefore, the z-coordinates are especially prone to errors. An over-focusing artifact occurs when an experimenter traces the neurites past the boundries of the section as if the neurites where extending up or down out of the slice. 3. Incomplete tracing due to 'muddiness' when tracing through thick sections. 4. Mismatched stitching. In these cases, the pieces of the arbors were not stitched back together correctly. 5.

Slicing. The sides of the neurons were sliced off. 6. Shrinkage. Depending on the histology process used, tissue can experience shrinkage in the slice direction (Figure 6.10) of up to 80%. 7. Truncation. Neurons were truncated because the investigators stopped tracing the neurons after they extended past a certain distance from the soma.

6.2.6 Shrinkage correction

Most of the data were acquired from brain sections and were subject to shrinkage especially in the slice direction (dependent on the method used to collect the data). Some laboratories corrected their data for shrinkage, but other laboratories did not. When shrinkage was not corrected, I did my best to make the most realistic possible correction to the data. This endeavor was quite difficult as there seems to be quite a bit of confusion in the field surrounding shrinkage. Difficulty assessing an appropriate correction factor arose due to: 1. unreported shrinkage in referenced papers, 2. possible variation in section shrinkage within an experiment, 3. tracing artifacts including disconnected neurons, tracing with incorrect magnification, and tracing past the section edge by over-focasing, 4. unknown effects of tissue distortion, and 5. mislabeling of article references or original section thickness. Often choosing a correction came down to a judgement call and undoubtably artifacts remain.

A common example of a situation where the necessary corrections were unclear follows. Let's say a group of neurons all reference a paper that reports a section thickness of 100 μm and a shrinkage of up to 75%. Looking at the data, most of the neurons look shrunken but some do not. Or perhaps a few look a bit thicker than they should be. This is most likely due to incorrect magnification calibration of the Neurolucida computer tracing system. However, now, instead of fixing all of the data from the study in a uniform way, a judgement call must be made concerning which neurons need to be corrected. Although choosing which neurons need to be corrected is not usually straight forward, assume that somehow a decision is made concerning which neurons to leave unaltered and the remaining data are corrected by a factor

consistent with a 75% shrinkage. With this correction, many of the neurons look more reasonable but many of the neurons are thicker than the reported 100 μm. One might wonder if there is variable shrinkage between different sections. Some experimenters think that there is variable shrinkage between sections, others do not. It may depend on the histology protocol.

In order to discuss another method by which to estimate shrinkage, I need to define the term 'neuron thickness'. I will use the term neuron thickness to refer to the extent of the neuron (including all arbors) in the direction perpendicular to the section. If a neuron is sliced off on both the top and bottom sides of the section, the neuron thickness is representative of the section thickness. Because I can calculate the neuron thickness from the data, I then know the thickness of the section after it has undergone shrinkage. In this case the neuron can then be expanded to reported section thickness. However, often it is difficult to assess whether the neurons are sliced on both sides. If they are not sliced on both sides and the neuron thickness is used as a proxy for section thickness there will be an overestimate of the shrinkage. For example, a section may have a thickness of 100 μm but the neuron is only sliced off on one side and extends down 50 μm into the section. If the section shrank by 50% the neuron is now 25 μm wide (assuming the tissue shrank uniformly). To estimate the percent shrinkage using the width of the neuron:

$$\%\ shrinkage = \left(1 - \frac{shrunken\ tissue\ thickness}{reported\ section\ thickness}\right) * 100 \tag{6.1}$$

In this case the percent shrinkage would be:

$$correct\ \%\ shrinkage = \left(1 - \frac{50\mu m}{100\mu m}\right) * 100 = 50\% \tag{6.2}$$

.

However if the shrunken tissue thickness is replace by the neuron thickness an overestimate of the shrinkage is calculated:

$$estimated\ \%\ shrinkage = \left(1 - \frac{25\mu m}{100\mu m}\right) * 100 = 75\% \tag{6.3}$$

.

In addition, the data are suseptable to tracing errors. Sometimes it is difficult to obtain the correct coordinates describing the depth within the section using Neurolucida. As a result, it is likely that at times, pieces of the reconstructed neurons extend farther than the neurons do in reality (Figure 6.8). This creates a problem adjusting all the neurons to the reported section thickness. For example, if a branch that was incorrectly traced far past the limits of the section is used as a boundry of the neuron then the neuron would be smashed into a space much smaller than it should be (for example, say a piece of neuron was traced 100 μm past this section such as is possibly the case in Figure 6.8). In addition, it is possible that there is some tissue distortion. Occasionally the tissue appears skewed sideways in the section (Figure 6.9). Also, at times, the branches of the neurons are not stitched together correctly (although these neurons will be eliminated from the data set). These issues make it difficult to decided whether to correct the neurons to their section thickness or to use one universal correction factor for a data set even though there may be variation between sections.

One could consider every individual neuron and estimate what caveats are associated with each in order to make cell specific shrinkage corrections. However, it is difficult to gauge specific artifacts and adding experimenter bias to data is undesirable.

The general corrections made to the data are summarized in Table 6.1. The reasoning regarding the corrections of each data set follows in the, "Laboratory specific details of artifacts, shrinkage corrections, and data exclusion" section.

6.2.7 Data exclusion

Many arbors were eliminated from the set. Common reasons for exclusion included severe slicing artifacts, incomplete tracings, tracing alignment artifacts, and unrealistic two-dimensioal arbors (z-coordinants all equal to 0).

The entire data sets from some laboratories were not suitable for this study and therefore eliminated from the data set. Table 6.1 contains a summary of which

Table 6.1: Summary of data usage and manipulation. NA stands for not applicable. The Hirsch neuron provided online was eliminated. The correction for the Staiger neurons was not applied to 8 outliers. Please see the text for details.

Archive Name	Shrinkage Correction	Used	Reason excluded
Allman	none	Yes	
Brown	1.4815	No	Platform tracing
Dendritica	none	Yes	
Destexhe	none	Yes	
Eysel	none	Yes	
Gonzalez-Burgos	NA	No	Section misalignment
Hirsch	none	Yes	
Jacobs	NA	No	120 um sections (highly sliced)
Lewis	NA	No	Truncated at 100 μm
Markram	none	Yes	
Smith	1.4815	Yes	
Staiger	4.0	Yes	
Svoboda	variable	Yes	
Wearne Hof	none	Yes	
Yuste	NA	No	Z-axis inconsistencies

laboratories were excluded. Briefly, the Brown laboratory data were traced in 'platforms' (Figure 6.3). The Gonzales-Burgos neurons often showed problems 'stitching' the sections back together. The Jacobs laboratory data were obtained in 120 um sections and significant portions of the neurons were truncated due to slicing artifacts. The Lewis lab was excluded because the investigators stopped tracing the neurons after the neurites extended past 100 μm from the soma. Finally, the Yuste laboratory neurons were excluded due to large artifacts in the z-axis. Details concerning the exclusions can be found below in the, "Laboratory specific details of artifacts, shrinkage corrections, and data exclusion," section.

6.2.8 Laboratory specific details of artifacts, shrinkage corrections, and data exclusion

Allman

The paper associated with this data set does not mention shrinkage (Watson et al., 2006). Many of the neurons in this data set are close to, or larger than the reported section thickness of 200 μm (Figure 6.1). Therefore, I will assume that there is either little or no shrinkage in the data, or that it was already corrected.

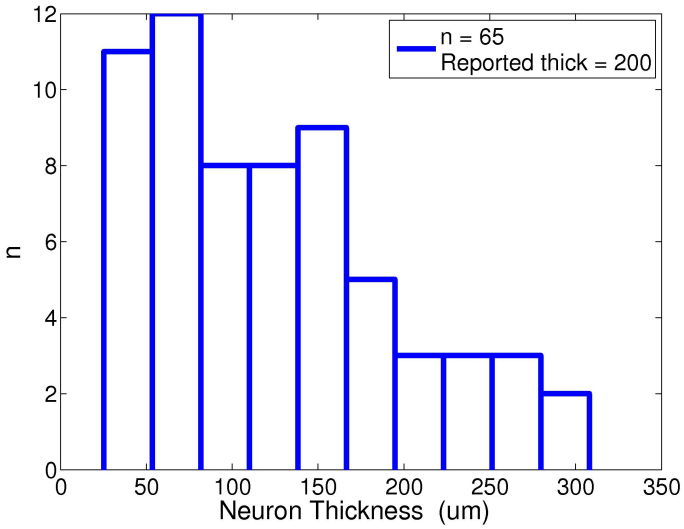


Figure 6.1: Neuron thickness of unaltered Allman data. Some neurons are thicker than the reported section thickness of 200 um. Perhaps this is due to tracing the arbors beyond the boundaries of the section as a result of over-focasing.

Brown

These data were taken for a study done in rats of postnatal day 0 (P0) to adulthood (P45) (Furtak et al., 2007). Therefore, many of the neurons were small and not fully developed. One neuron was much thicker than the others (Figure 6.2). This neuron, P42-DEV384, appeared to have a tracing artifact and was eliminated from the data set. Eighteen of the neurons were completely flat (had a thickness of 0 μm). It is hypothesized that there was a problem with the data transfer to the

Neuromorpho.org website (conversations with Dr. Sharon Furtak). These neurons were eliminated from the data set.

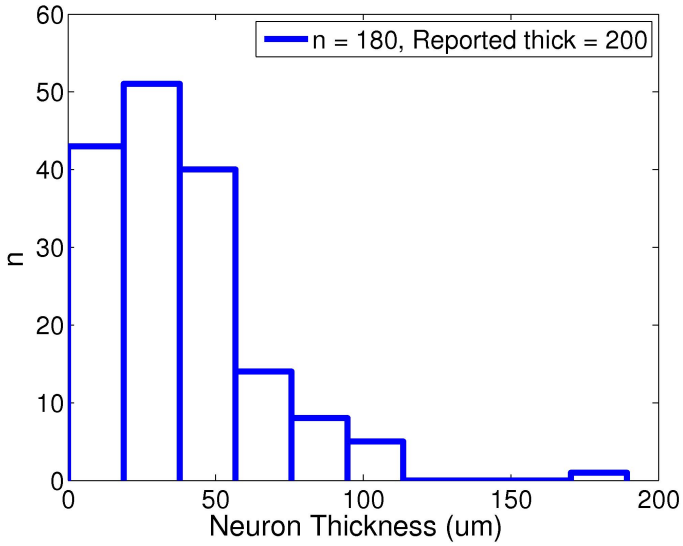


Figure 6.2: Neuron thickness of unaltered data from the Brown archive. These data are associated with the Furtak et al. (2007) paper. The reported section thickness is 200 um.

These data were collected using the protocol of Gibb and Kolb (1998). Dr. Furtak is no longer at Yale where the slides are kept and therefore could not measure the section thicknesses to assess shrinkage. However, the data submitted by the Smith laboratory reported using the same protocol. The Brown data were therefore corrected by the same reasoning used for estimating the Smith correction factor below. The end result corrected the section thickness by a factor of 1.4815. Sharon Furtak agreed that this was a reasonable correction for the data (personal communication).

However, these traces had problems with z-axis coordinates (in the reference frame of the section). They often look platform-like when viewed from the side (see Figure 6.3 as and example). As a result, I decided not to use these data.

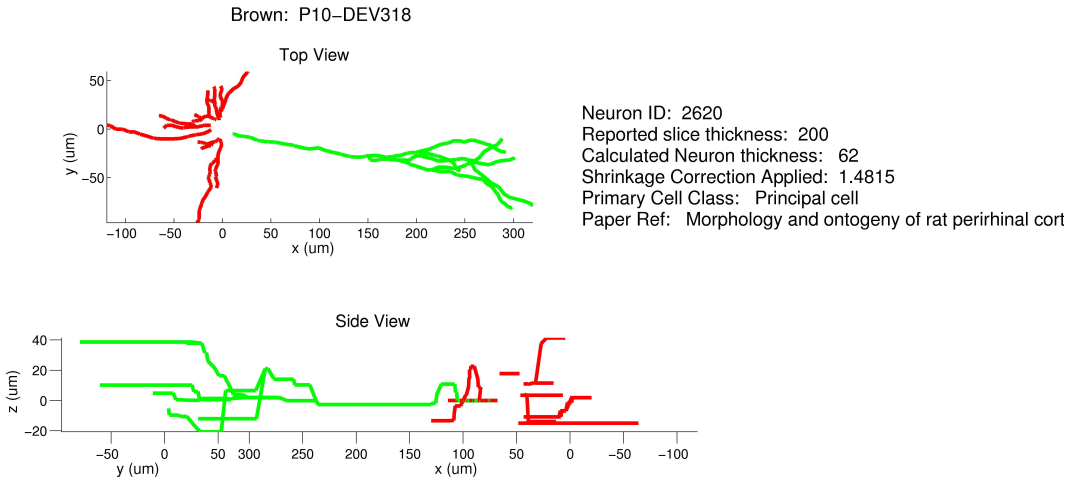


Figure 6.3: Example of 'platform' tracing of Brown data. Both images are 3D reconstructions rotated to show different views. The x, y, and z axes are in the reference frame of the section.

Dendritica

These data reference two papers: Stuart and Spruston (1998) and Vetter et al. (2001). Four of the neurons in this data set are not currently available on neuromorpho.org (503c, 602c,822, 913c). How I acquired these traces is unknown. It is likely that they were available on the website previously but are no longer posted. These neurons reference the Vetter et al. (2001) paper, appear artifact free, and do not look shrunken, therefore, I left them in the data set.

Two neurons which reference the Vetter et al. (2001) are listed as neurons from cat (j8_L23pc and j7_L4stellate). There is no mention of cat neurons in the Vetter et al. (2001) paper, therefore, their origin is unknown. However, these neurons also appear artifact-free and do not look subject to shrinkage. Therefore, I left them in the data set.

Two neurons are listed on the neuromorpho.org website as Stuart and Spruston (1998) (p21 and p22). However, the Vetter et al. (2001) paper says that three neurons are taken from the Stuart and Spruston (1998) paper. I believe the third

Stuart and Spruston (1998) neuron is file 'p18' (due to the naming conventions of the files). This neuron's paper reference was probably mislabeled. These three neurons look a bit flattened however the Stuart and Spruston (1998) paper mentions that there is minimal shrinkage. Therefore, I will leave these neurons 'as is'.

The three neuron files remaining (par, pds and pjs) are layer 5 rat pyramidal neurons. The Vetter et al. (2001) paper reports that it filled two rat neocortical layer 5 pyramidal neurons, obtained three from G. Stuart and N. Spruston, and one from D. Smetters. Perhaps two of these remaining three files are the neurons they traced themselves and one is the neuron from D. Smetters. Although, this seems a bit unlikely due to the consistency of the naming convention of these three neurons. However, the 'pds' neuron does not look shrunken and the remaining two look either flattened or severely sliced on one side. The histology was not reported in the Vetter et al. (2001) paper. I will leave these two neurons 'as is' as the appropriate correction is unclear and the effect of just two neurons on the results will be minimal.

Destexhe

There are two neurons in this data set which reference the paper Contreras et al. (1997). The paper mentions that the neurons were sectioned at 80 μm. However, these sections must have been 'stitched' back together as the neurons are much thicker than 80 μm and the traces look very nice. There is no mention of shrinkage in the paper. The neurons look realistic and I will leave them 'as is'.

Eysel

These are beautiful traces reconstructing the entire cortical depth in cat. The neuromorpho.org website reports that these data correspond to a paper by Kisvarday and Eysel (1992). However, in conversations with Zoltan Kisvarday it was determined that these data do not correspond to the listed paper which reports a dry mounting histology method (which would consequently cause a great deal of shrinkage). Instead these neurons were resin embedded to minimize shrinkage. In

addition any shrinkage was corrected (unlike in the 1992 paper). Dr. Kisvarday recommended Buzs et al. (2006) as a reference for these data.

Gonzalez-Burgos

It was difficult to assess the shrinkage in this data set due to the small number of neurons referencing four different papers: Gonzalez-Burgos et al. (2004), Gonzalez-Burgos et al. (2005), Zaitsev et al. (2005), and Krimer et al. (2005).

Dr. Gonzalez-Burgos mentioned that these data may not be very accurate. Most of the data were resectioned at 50 μm and were provided to Dr. Ascolli (for the neuromorpho.org compilation) 'unstitched' back together. It is unknown whether Dr. Ascolli's team attempted to repair these data. Many of the neurons do not look properly stitched together (Dr. Gonzalez-Burgos, personal communication). In addition to the obvious problems that would be caused by using incorrect traces in our analysis, incorrect stitching adds many complications to the calculation of shrinkage. Therefore, these data are excluded from our analysis. However, for completeness more about the shrinkage follows.

All of the referenced papers report fixing in paraformaldehyde and incubation with a avidin-biotin peroxidase staining kit. However, each paper reports different steps in the histology protocol. These steps may or may not have led to different shrinkage artifacts. It is probably safe to assume that the neurons with a reported section thickness of 50 μm were sliced on both sides. There are only two of these in the data set and their neuron thicknesses are 39 and 44 μm. This corresponds to a maximum of 22% shrinkage. Dr. Gonzalez-Burgos remarked that he believes that neuroanatomists report a 30% shrinkage in volume with the type of method he uses (personal communication). This would probably correspond to approximately a 20% shrinkage in section thickness. However, as in other data sets it is possible the section thickness was misreported and this in not an accurate shrinkage estimate.

Hirsch

One neuron was provided on the neuromorpho.org website. This neuron referenced Hirsch et al. (2002) which in turn references the histology in Hirsch (1995). The Hirsch (1995) paper reports a section thickness of 200 and 350 μm for the electrophysiology. The paper reports reslicing to between 70 and 100 μm for histology. These sections must have been 'stitched' back together even across electrophysiology sections as they are approximately 800 μm thick. This neuron may have been shrunken as the histology protocol reports dry mounting which generally causes significant shrinkage. The side view of this neuron in Figure 6.4 appears flattened in comparison to the top view. I removed this neuron from the analysis because I could not deduce the experimental details or caveats associated with this trace.

Nineteen additional neurons from cat cortex were provided by Dr. Hirsch. These neurons look very nice and are included in the analysis.

Jacobs

These neurons report 120 μm section thickness cut on a vibratome. However, many of the neuron are thicker than the reported section thickness (Figure 6.5). There was no mention of shrinkage or shrinkage expansion in the referenced papers (Jacobs et al., 1997, 2001; Travis et al., 2005). Dr. Jacobs believes that, "Tracings that are close to the 120 thickness (say up to 150 or so) are probably the result of the tracer over focusing–that is, focusing beyond the actual thickness of the section when tracing. This is usually the result of following one branch of the cell out too far in a particular direction, and it thus makes it seem like the section is thicker than it is. This would increase dendritic length slightly, only for those branches traced in this manner. And/or, it could simply be variation in the section thickness coming off of the vibratome. Tracings that are absurdly thick (such as 10-3-2, 10-3-4, 10-5-5A) are the result of tracer error. In looking at these more extreme cases, it appears the tracer may have traced part way, stopped, and then started again, but the origin of the dendrite no longer lines up with the soma (0 on the z plane) and instead began

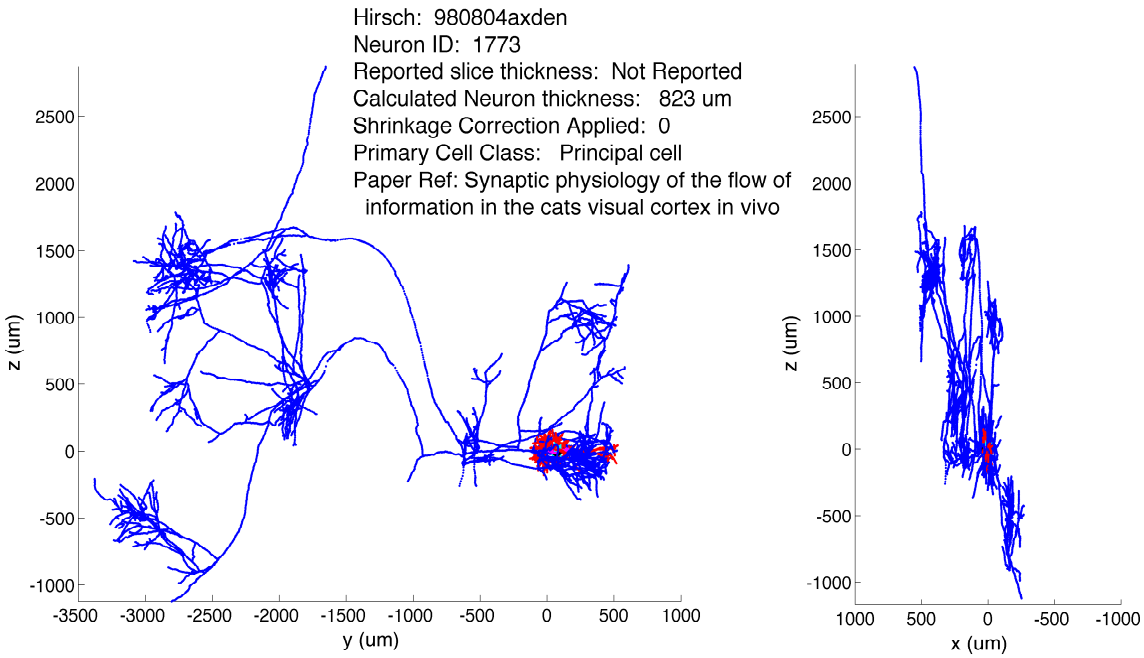


Figure 6.4: Cat cortical neuron (980804axden) traced by Judith Hirsch availiable on neuromorpho.org. The neuron is much thicker than the reported section thickness in the referenced paper. The referenced histology method (dry mounting) should yield shrinkage; however, shrinkage is not mentioned in paper. Despite the neuron thickness (approximately 824 μm) this neuron may be shrunken: notice in the side view that the axon tufts are quite flattened. Both images are 3D reconstructions rotated to show different views. The x, y, and z axes are the coordinates provided by Neuromorpho.org.

the dendrite "out in space" as it were. This makes the section appear much thicker than it actually is, though it probably doesn't significantly affect the overall length of the dendrites. This is something I did not notice when checking over the cells because it only shows up after rotating the cell on the z axis." (personal communication).

These data were excluded, not because of the above mentioned issues which are fairly common in the data sets, but instead because it appears that the neurons are severely sliced on each side of the section. These sections were only 120 μm thick and it was obvious that the neurons were originally much thicker.

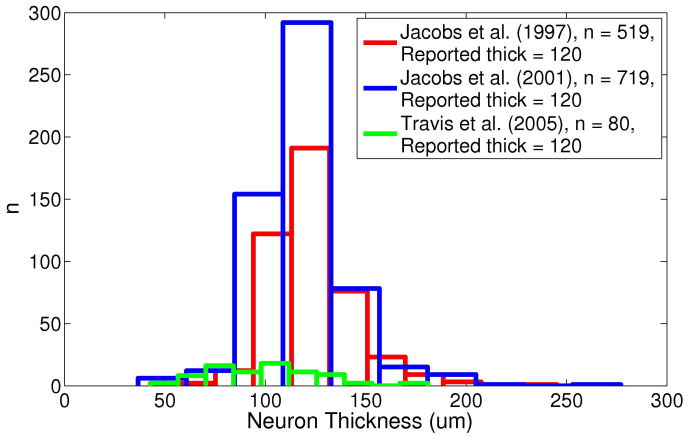


Figure 6.5: Neuron thickness of unaltered data from the Jacobs data set.

Markram

These data look good and references Wang et al. (2002). The paper reports that shrinkage was corrected.

Smith

All of the data in this set refers to Bergstrom et al. (2008), except one neuron which references McDonald et al. (2005). Both papers report following the Golgi staining protocol of Gibb and Kolb (1998). Neither paper mentions shrinkage. Dr. Craig McDonald, who said that by no means did he do it officially, he measured thicknesses of between 130-140 um for a few of the sections after shrinkage. He also said that he thought section shrinkage was fairly consistent and also that he would not assume neurons were cut on both sides of the section (personal communication). When I performed a shrinkage correction by multiplying the appropriate axis by a correction factor of 135/200=1.4815, the data looks suitable. I.e., some of the neurons are close to 200 um thick but none are larger. It can be seen how this would be the case from the histogram of the neuron thickness of the uncorrected data (Figure 6.6). In addition, although the Gibb and Kolb (1998) reference did not report shrinkage for their protocol, Dr. Gibb believes shrinkage is on the order of 30%, maybe up to 40% (personal communication). This is on the order of the measured shrinkage.

Therefore, I corrected these data by a factor of 1.4815.

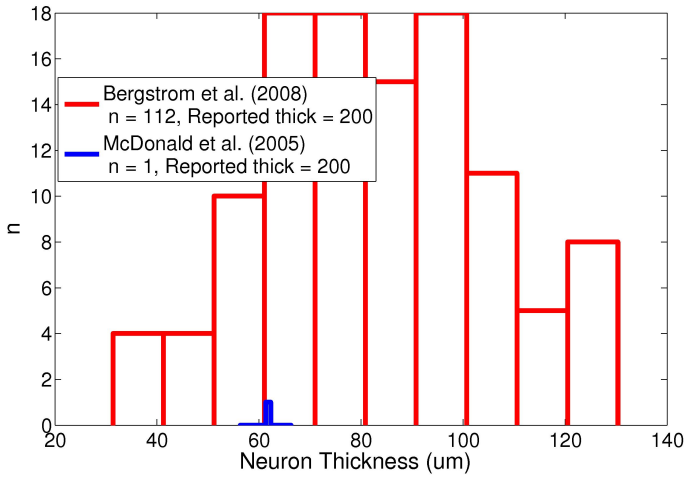


Figure 6.6: Neuron thickness of unaltered Smith data.

Staiger

It appears that this data set was subject to most all the possible issues listed at the beginning of this chapter. In summary, it is highly probable that the neuron references are mislabeled or neurons in smaller sections may have been 'stitched' back together. I think that all of these neurons (except 8 outliers) should be corrected by a 75% shrinkage factor despite the fact that this often results in neurons which are thicker than the section thickness reported on the website or in the referenced papers. The reasoning follows.

Figure 6.7 shows a histogram of the neuron thickness of the data. The histogram lists the section thickness reported by the referenced papers. The section thicknesses reported in the papers match the neuron thickness reported on the website except for the Schubert et al. (2003) paper which reports 300 um on the website, and one neuron from the Staiger et al. (2004) paper which reports a different thickness on the website than the rest of the neurons which reference the same paper. There appears to be eight outliers that are thicker than the rest of the data (i.e. larger than 100 um). Indeed these neurons look less 'squished' than the rest. It is unknown

whether these 'unsquished' neurons were already shrinkage corrected, traced on an incorrect magnification which essentially canceled out the effect of shrinkage, or there was an error in the data transfer to neuromorpho.org.

These data referenced three papers: Schubert et al. (2003) (which reports following the histology protocol of Angulo et al. (1999)), Staiger et al. (2004), and Schubert et al. (2006) which reports following the Staiger et al. (2004) histology method. Essentially two different protocols are used. The first few steps of these protocols are the same. However, the Angulo et al. (1999) protocol describes reslicing to 50 um on a vibratome, dehydrating the tissue using ethanol, and then flat embeds the tissue using Durcupan AMC. The Staiger et al. (2004) paper reports that the sections were cyroprotected and resliced to 100 um on a freezing microtome which were then subject to cytocrome oxidase histochemistry. They were air-dried, mounted on glass slides, immersed in xylene and coverslipped in Depex. The Schubert et al. (2003) paper reports a slight difference from the (Staiger et al., 2004) only in that it uses and additional silver/gold intensification following the protocol of (Bender et al., 2003) to tinct the barrels. In all the other cases, the barrel field was either visualized by cytochrome oxidase histochemistry or the barrel pattern of the photograph of the native section was transferred manually into the reconstruction.

The Staiger et al. (2004) paper reports a z-axis (direction of the section thickness) shrinkage of up 75%. I believe most of the shrinkage is the result of the cytocrome oxidase histochemistry or the air-drying. Therefore, the Angulo et al. (1999) protocol might not be subject to this much shrinkage (shrinkage is not reported in this paper). Looking at the histogram of the neuron thickness of the Schubert et al. (2003) paper, the neurons are 40 μm after shrinkage as opposed to the reported section thickness of 50 μm. This would correspond to approximately a 20% shrinkage. However the neurons look much more flattened than 20%. When corrected by a 75% shrinkage they look appropriate. Therefore, it is highly possible these neurons should reference a paper which corresponds to a 75% shrinkage. Or it is possible they shrank approximately 75% despite the difference in protocol, and were then 'stitched' back

together for the reconstructions.

As for the remaining two data sets, in the histogram one can see that many of the Staiger et al. (2004) neurons are around 60 μm thick. If these are corrected by 75% (as reported) they would be approximately 240 μm thick, which is much larger than the reported 100 μm section thickness. My guess is that these neurons should either reference the Schubert et al. (2006) paper which reports a 300 μm section thickness, or, the reported 100 μm sections were 'stitched' back together for the reconstructions. These neurons look appropriate when corrected by a 75% shrinkage. The Schubert et al. (2006) referenced neurons look appropriate when corrected by the reported 75% shrinkage (excluding the outliers mentioned above). None of these neurons are thicker than reported section thickness of 300 μm when corrected by a 75% shrinkage correction factor.

To conclude, all but the eight outliers were corrected by a 75% correction factor which is equivalent to multiplying the section thickness coordinates by a factor of four. The eight outliers were left 'as is'. Dr. Staiger agreed that my reasoning was probably correct although he could not confirm my suspicions as he was unsure of the collection details of each neuron (personal communication).

Svoboda

These data experienced expansion presumably due to the DMSO-based method used for mounting these sections (personal communication with Dr. Gordon Shepherd; section expansion is not discussed in the paper). The section expansion was measured for each section. The shrinkage coefficients of all but four neurons (File names: $g0445-7\_cellA$, $g0701P1$, $g0703P1$, $g0762\_b$) were provided to me by Dr. Armen Stepanyants and are listed in Table 6.2 . Correcting the section thickness by these values made the neuron thicknesses very close to the reported section width of 300 μm. The four correction factors that were not provided correspond to neurons that were not used in their studies most likely due to bad dye labeling (personal communication with Dr. Armen Stepanyants). I removed $g0445-7\_cellA$ and $g0762\_b$

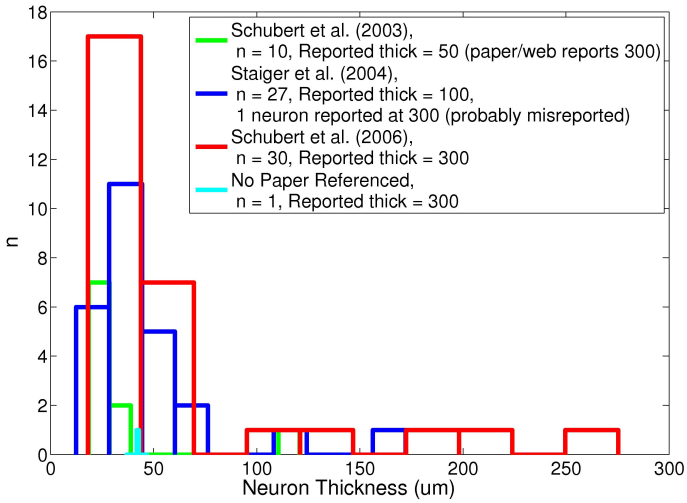


Figure 6.7: Neuron thickness of unaltered Staiger data. Each neuron referenced one of 3 papers: Schubert et al. (2003), Staiger et al. (2004), or Schubert et al. (2006). One neuron did not reference any paper. Eight outliers are thicker that 100 μm and correspond to the files named: 030213DS2, 040112DS1-lower, 040112DS1-upper, 040112DS2, 040113DS3, DS1_160501, DS2_120601_wL, IF1_300300. The neuromorpho.org website reports a section thickness of 300 μm for the Schubert et al. (2003) paper. However, this paper references using the histology method of Angulo et al. (1999) which reports reslicing to 50 μm. The Staiger et al. (2004) reference reports an original section thickness of 500 μm and then reports reslicing down to 100 μm. On the neuromorpho.org website all of these neurons have a reported section thickness of 100 μm except one which reports 300 μm. The Schubert et al. (2006) paper reports an original section thickness of 300 μm but reports to follow the histology protocol of Staiger et al. (2004). It is unclear whether or not these neurons were resliced to 100 μm. All of the data in this set (excluding the outliers) were corrected by a 75% shrinkage factor. The outliers were left 'as is'. Please see the text for the reasoning behind this correction decision.

from the data set because they were not sliced on both sides and looked too expanded when altered to be 300 μm thick. I did not see a marked difference between $g0701P1$ and $g0703P1$ and the other neurons in the data set, therefore, I did not exclude them from the study. The axons of these neurons looked underdeveloped however many axons of other neurons also looked underdeveloped. Therefore, the dendritic arbors of these neurons will remain in the study while the axons of these neurons will be removed as will incomplete axon traces from other neurons.

Table 6.2: Shrinkage correction factors for Svoboda data files. Multiplying the z-coordinates of the data provided on the Neuromorpho website will correct the neurons for expansion artifacts.

File Name	Correction	File Name	Correction
I03192A	0.66458	g0700P1	0.75994
I0319_3A	0.65595	g0701P1	0.76947
I0319_4A	0.57442	g0702P1	0.72564
I0319_5A	0.73174	g0703P1	0.85312
I0319_5B	0.74019	g0704P1	0.81248
I03464	0.70264	g0705P1	0.71705
I03471	0.69238	g0706P1	0.87283
I03472	0.68231	g0707P1	0.90818
I03474	0.67838	g0709P1	0.82271
I03481	0.755	g0715_b	0.81453
I03482	0.70517	g0716_b	0.82332
I03485	0.66742	g0717_b	0.79453
I03538	0.76705	g0718_b	0.73174
I03539	0.77071	g0719_b	0.73174
I03541	0.73174	g0720_b	0.73174
I03542	0.70284	g0721_b	0.73174
I03544	0.78296	g0722_b	0.73174
I03555	0.72228	g0724P1	0.73174
I03564	0.71721	g0746_b	0.75982
I03566	0.63585	g0758_b	0.73174
I03569	0.64412	g0759_b2	0.73174
g0435P1	0.73174	g0760_b	0.75609
g0436P1	0.8036	g0761_b	0.815
g0437P1	0.73174	g0763_b	0.82974
g0611P1	0.73174	gR001P1	0.60459
g0612P1	0.73174	gR002P1	0.67209
g0684P1	0.73174	gR003P1	0.62488
g0685P1	0.73174	gR004P1	0.73174
g0692P1	0.71109	i03535	0.72977
g0697P1	0.78623	i03537	0.73174
g0698P1	0.82122	i03557	0.78143
g0699P1	0.7425	i03558	0.75586

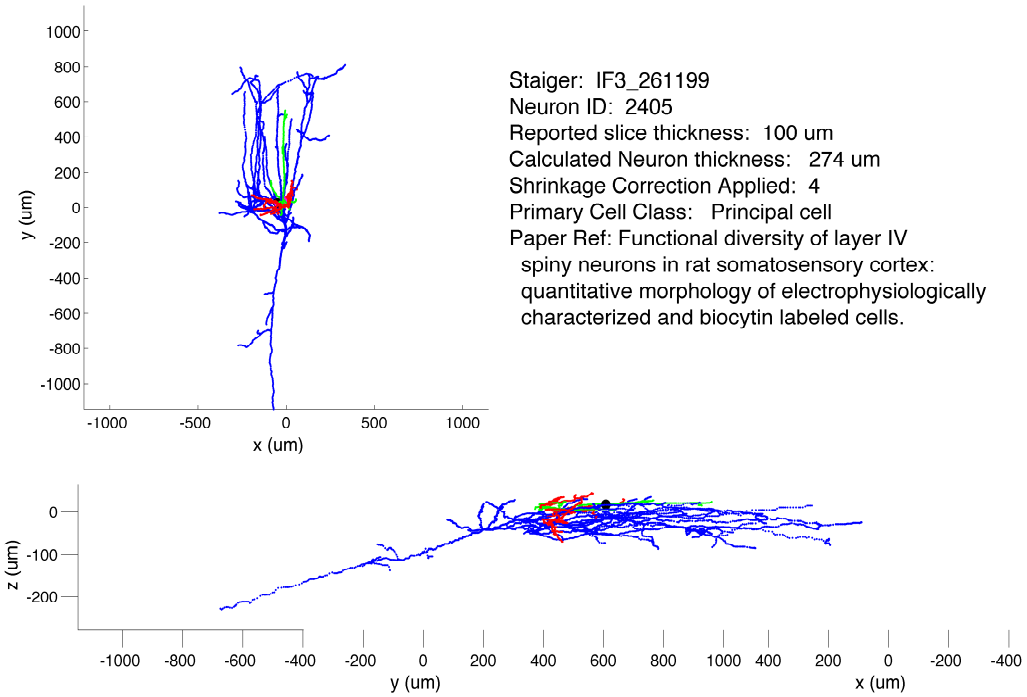


Figure 6.8: Example of a Staiger neuron with a correction corresponding to 75% shrinkage. After correction, the neuron is thicker than the reported section thickness of 100 μm. However, it does not look like this neuron should be shrunk more. It is highly probable that the section thickness was misreported. In addition, the solitary process that extends off to the left, may have been 'over traced,' a common artifact leading to tracings that extend beyond the section edge caused by over focusing.

Yuste

These data were not used due to problems with the z-axis coordinates (personal communication with Dr. Yuste).

Wearne_Hof

These data reference Duan et al. (2002). The paper does not reference shrinkage. Dr. Hof remarked there was little shrinkage: probably between 10 and 15%, 20% max (personal communication). Therefore, I did not perform any shrinkage corrections on these neurons.

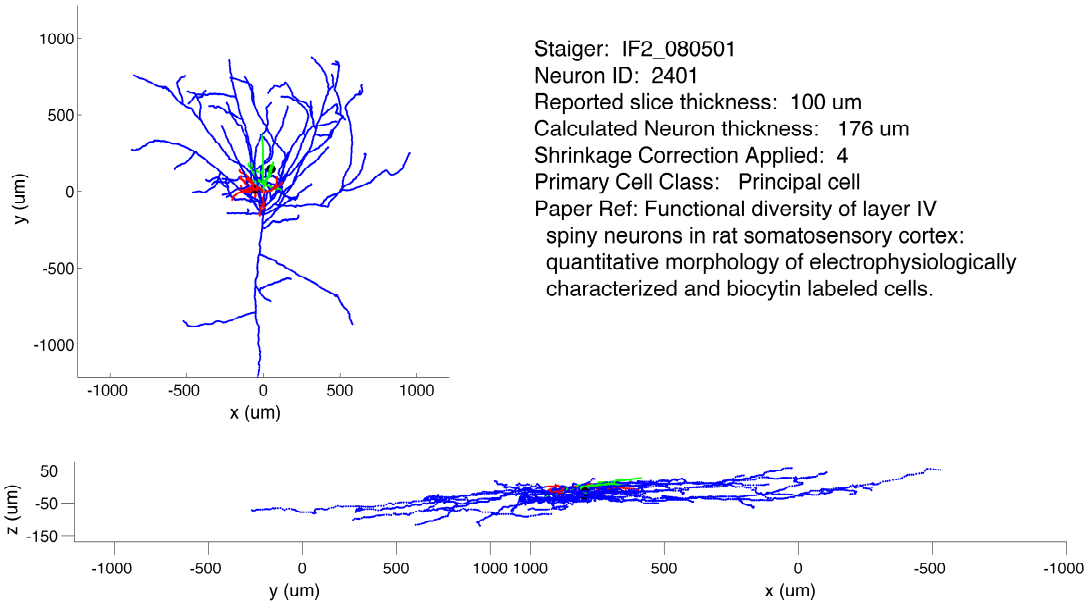


Figure 6.9: Staiger neuron looks a bit diagnol in the section. This could be due to tissue distortion artifacts.

6.3 Finding the principal standard deviations using principal components analysis (PCA)

Principal orthogonal axes along with the variance in the principal directions, σ_x^2, σ_y^2, and σ_z^2 were obtained using weighted principal components analysis (PCA). PCA is a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. PCA was used because maximally independent axes are needed in order to describe the adfs as Gaussians (the different dimensions of a Gaussian are independent from one another). Otherwise, I could have used some other frame of reference to orient the arbors such as the axis pointing directly toward the pial surface as opposed to the principal axis pointing most toward the pial axis.

PCA is achieved by finding the eigenvectors of the covariance matrix. In this case the covariance matrix is weighted by the corresponding arbor segment weight. Using matrix algebra, if X, Y, and Z are column vectors of the x, y, and z coordinates of the segment centers, and W is a column vector of the segment weights (segment lengths) then the covariance of x,y and z is:

$$Cov(x, y, z) = ([X, Y, Z]^T [X, Y, Z]W) / \sum w_j \tag{6.4}$$

The eigenvalues which correspond to the orthogonal variances were then found using the Matlab function 'eigs'.

In the case of the fish data, the principal components are calculated using the same method except that there is no 'Z' component in equation 6.4.

6.3.1 Artifacts created by PCA

PCA is designed to find the the direction of the highest variance. Very rarely are data perfectly symmetrical. Even if data are created using a Gaussian function, usually the variance in orthogonal directions is not equal. In the case of neural arbors this means that the standard deviations will very rarely be equal. In the case of the fish retinal ganglion cells the smaller standard deviation is always defined as σ_x, and the larger standard deviation is σ_y. Therefore when σ_x is plotted versus σ_y, the data always lie on one side of the $\sigma_x = \sigma_y$ unity line (Figure 2.1): the data never touch the unity line. The maximum value of σ_x/σ_y=0.977.

In the case of cortical arbors the standard deviations are chosen in the direction the matches the orientation of the neurons in the brain (see the section on orienting the data below). σ_z is defined as the standard deviation pointing toward the pial surface. Then, the remaining standard deviations are defined as σ_x, and σ_y, where $\sigma_x < \sigma_y$. Because, σ_z is not always greater than or less than the other values, and because PCA very rarely yields instances where the standard deviations are exactly equal, there are planes void of data in the σ_x, σ_y, σ_z parameter space were $\sigma_x = \sigma_z$, and $\sigma_y = \sigma_z$ (Figure 3.10).

6.4 Orienting the data

6.4.1 Fish retinal ganglion cell arbors

Retinal ganglion cells are flat, two-dimensional arbors that tile the front of the retina. Therefore, they will only need two orthogonal standard deviations, σ_x, and σ_y, to describe their adfs. There is no information by which to orient these standard deviations, so, σ_y, was assigned to be the larger of the two standard deviations.

6.4.2 Cortical arbors

The cortical arbors are three dimensional arbors and therefore will have three standard deviations to describe their adfs. So that information concerning the orientation of the arbor within cortex is conserved, principal axes were assigned to relevant directions within the brain. Information transfer seems to happen in two relevant directions with in the cortex: within layer, and across layer. Finding the principal axes and calculating the standard deviations along those axes was done using PCA and is described in the previous section. After the principal axes were found, the z-principal axis was assigned as the principal axis which projects through the layers toward the pial surface. Note that in histology experiments the z-axis refers to the axis which moves through the section if you are looking down on the section through a microscope (Figure 6.10).

In general, sections are not mounted so that the pial surface is pointing in a particular direction. The investigators usually cut the sections so that they view the neurons from the side as opposed to looking down on them from the pial surface. Therefore, usually the pial surface is somewhere in the xy-direction in the section frame of reference. Because the exact direction of the pial surface is unknown, the direction had to be inferred using various methods described below. When possible, the process was automated; otherwise, it was done by eye and as such, was susceptible to experimenter bias. At times, the assignment was difficult and there are bound to be errors.

For pyramidal cells where both an apical and basal dendrite were traced the orientation process was automated. The reference axis used to assign principal arbor axes was the line connecting the center of mass of the apical and basal arbors (For an example see Figure 6.10). The principal axis with the smallest angle in reference to this line was assigned to be the z-axis. That is to say, the axis which was most parallel to the reference line was assigned to be the z-axis.

In some data sets it appeared that the the pial surface was always pointing in a certain direction. For example, if all the pyramidal cells consistently point in one direction then that direction can be used to orient the other cells in that data set. In this case, when the arbor was not a pyramidal cell the standard deviation pointing in the direction closest to parallel with the proposed axis was chosen as σ_z. This was done for the Markram data. For the Markram data, in the cases where a neuron was not a pyramidal cell with both the apical and basal dendrites present, the principal axis most parallel to the [0,1,0] vector (in the section frame of reference) was chosen for the z-axis.

In all other cases, I used visual guidance to determine which axis should be assigned to the z-axis. Often, especially with neurons such as interneurons that are largely spherical and do not provide any visual clue to their orientation, this task was quite difficult and there are bound to be errors in the axes assignments. However, in these instances, the variances in each direction are similar and thus largely interchangeable. Therefore, these errors most likely will not significantly affect the outcome of the study.

After the z-axis was assigned, the x and y principal axes were assigned to the remaining axes and will describe the way in which an arbor samples or distributes information over the neural map. Because no information is provided concerning the orientation of the neurons within the cortical map (and I don't know what a meaningful orientation would be), just as with the fish, the x-axis was assigned as the axis corresponding to the smaller of the two remaining standard deviations.

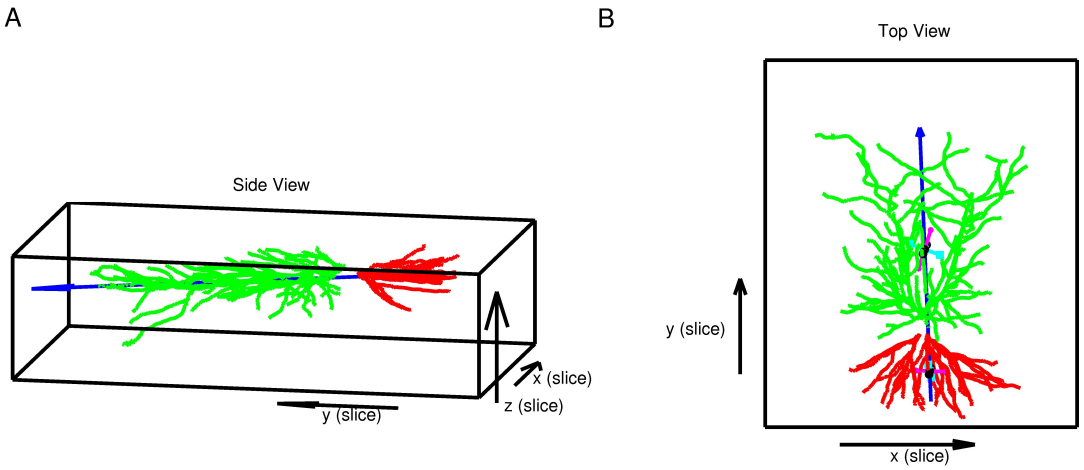


Figure 6.10: Example of neuron orientation. A. Side view showing a pyramidal cell within the section. B. View looking down on the section as you would though a microscope. The black arrows at the bottom corners of the figure denote the cartesian xyz-axis assignments used for histology (which differ from the axes assignments I used in this study). The z-axis in the section frame of reference is defined to point toward the top surface of the section and is the axis most susceptible to shrinkage. The shapes in the center of the arbors in B are located at the center of mass of the arbor and the extending lines represent the principal (which assumes orthogonal) standard deviations of the spatial coordinates of the arbor segments. Magenta, cyan, and black represents the largest, middle and smallest standard deviations respectively. The blue line connects the center of masses of the apical and basal dendrites. In this study the neurons were oriented to reflect their orientation within cortex. The standard deviation which is most pointing toward the pial surface is assigned to be the z-standard deviation, σ_z. Although the pial surface is usually in a direction perpendicular to the xy-plane, the specific direction is unknown which makes orienting the neurons difficult. In the situations where a neuron has both a apical arbor and a basal arbor, the direction of the pial surface is taken to be in the direction of a line connecting the center of masses of the apical and basal dendrites (blue line). The principal axis found by PCA that is closest to parallel with this reference line is assigned to be the z-axis in the principal axis frame of reference. There is no information to orient the remaining tsswo principal standard deviations. Therefore, the largest of the remaining standard deviations is assigned to be the y-standard deviation, σ_y, and the smallest is the x-standard deviation, σ_x. In some cases there were horizontal pyramidal cells. In these cases the standard deviations were assigned in the appropriate directions.

6.5 Arbor differentiation via Random Forests^{TM} supervised learning

The Random Forest^{TM} (RF^{TM}) algorithm (Breiman, 2001a) was implemented in Matlab (Jaiantilal, 2009). For a clear explanation of how RF^{TM} work please see Dr. Breiman and Dr. Cutler's website (Breiman and Cutler, 2007). Briefly, RF^{TM} uses a sample of the data to grow many classification trees. Each tree is created by randomly selecting N data points, with replacement, from the data set. If there are M input variables ($\sigma_x, \sigma_y, \sigma_z$, and L in our case), a number m<<M is specified so that at each branching point in the tree (called a node), m out of the M variables are randomly chosen and the best split on these m variables is used to split the node. The default m is the $\sqrt{M}$. Not all of the data will be used to create each tree. The classification of each data point is achieved by putting each data point 'down' each tree that it did not help create (and consequently there is no need for a 'hold out' test set). Each tree 'votes' for the class it thinks the data point belongs. Then depending on what classification thresholds are set, each data point is assigned to a class based on the tree voting. For example, if a majority wins threshold is set, if the majority of the trees vote that a data point is in class 1, it will be assigned to class 1. Other thresholds may be desirable depending on a particular data set. I do not have to choose thresholds because I use a method to calculate the AUC (a measure of classification ability) which is independent of thresholds, as discussed in the following section.

In order to find the best implementation of the algorithm we optimized two parameters for each binary classification schema. The first, was how many parameters (of the 4 fundamental parameters) we included in the algorithm. For example, if all the information for classification was contained in just σ_x, then including more superfluous variables could dilute the results. Conveniently, the algorithm will rank the variables in order of importance for correct classification. Therefore, my strategy for optimizing the number of parameters to include in the algorithm was to first

rank the parameters in order of their importance for classification. This was achieved by running the RF^{TM} 100 times (using the default m=$\sqrt{M}$=2) and the variable importance was in essence averaged over these 100 trials for each binary classification schema.

Then in order to asses how many of the parameters should be used, I ran the model in four cases. In the first case, only the most important parameter was used. In the second case, the first and the second most important parameters were used. This pattern was continued until all four parameters were included in the algorithm. For example, in rank order of importance, the variables for differentiating non-apical dendrites from apical dendrites are σ_z, L, σ_y, σ_x. For the first run of the RF^{TM} algorithm only σ_z was included, and the AUC was found to access classification capability. During the second run both σ_z and L were included and again the AUC was calculated. Then σ_z, L, and σ_y were used and then all of the parameters were included. Because the algorithm has an element of randomness to it, the results vary from run to run. Therefore in reality the algorithm was not run just once for each added parameter but instead the algorithm was repeated 50 times with each added parameter. The reported AUCs are the average of the AUCs found for each case.

The second optimization task is to assess how many variables, m, should be used to split the nodes. It is reported that the algorithm may be sensitive to this parameter (Breiman and Cutler, 2007). Because I do not have many parameters to deal with, it is possible to test all potential values of m for each binary classification schema and every included parameter case. Therefore, during each case where a parameter is added, all possible values of m were tried. If only one parameter was being used in the algorithm it is only possible to split on one variable. If two parameters are being used it is possible to split on one or two variables and so on and so forth. In summary, the AUCs reported and ROCs graphed in the Results section are the best found values for each binary classification schema from testing each parameter case with each possible value of m, 50 times.

6.6 Calculation of area under the receiver operator characteristic (ROC)

Although it seems counter intuitive, the area under the ROC curve (AUC) can be calculated independently of the ROC curve (Hand and Till, 2001). This is very convenient as it eliminates the need for integration of the curve and the necessity of having to pick a series of classification thresholds that will make the curve smooth. This independent calculation utilizes all the data and therefore is the most accurate method of calculating the AUC. Ling et al. (2003) provide an example of using this method and also show that it is superior to using a percent correct classification evaluation criteria. The AUC is calculated by the following equation:

$$AUC = \frac{(\sum r_i) - n_0(n_0 + 1)/2}{n_0 n_1} \tag{6.5}$$

Where n_0 and n_1 are the numbers of positive and negative examples respectively. r_i is the rank of the i_{th} positive example in the ranked list.

References

Akaishi, Y., Uchiyama, H., Ito, H., and Shimizu, Y., 1995: A morphological study of the retinal ganglion cells of the afghan pika (Ochotona rufescens). *Neuroscience Research*, **22**(1), 1–12.

Amthor, F. R., Oyster, C. W., and Takahashi, E. S., 1983: Quantitative morphology of rabbit retinal ganglion cells. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **217**(1208), 341–355.

Angulo, M. C., Staiger, J. F., Rossier, J., and Audinat, E., 1999: Developmental synaptic changes increase the range of integrative capabilities of an identified excitatory neocortical connection. *J. Neurosci.*, **19**(5), 1566–1576.

Badea, T. C., and Nathans, J., 2004: Quantitative analysis of neuronal morphologies in the mouse retina visualized by using a genetically directed reporter. *The Journal of Comparative Neurology*, **480**(4), 331–351.

Bailes, H. J., Trezise, A. E., and Collin, S. P., 2006: The number, morphology, and distribution of retinal ganglion cells and optic axons in the australian lungfish neoceratodus forsteri (Krefft 1870). *Visual Neuroscience*, **23**(02).

Bender, K. J., Rangel, J., and Feldman, D. E., 2003: Development of columnar topography in the excitatory layer 4 to layer projection in rat barrel cortex. *Journal of Neuroscience*, **23**(25), 8759.

Bergstrom, H. C., McDonald, C. G., French, H. T., and Smith, R. F., 2008: Continuous nicotine administration produces selective, age-dependent structural alteration of pyramidal neurons from prelimbic cortex. *Synapse*, **62**(1), 31–9.

Binzegger, T., Douglas, R. J., and Martin, K. A., 2005: Axons in cat visual cortex are topologically self-similar. *Cerebral cortex*, **15**(2), 152.

Bliss, T. V. P., and Collingridge, G. L., 1993: A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, **361**(6407), 31–39.

Bota, M., and Swanson, L. W., 2007: The neuron classification problem. *Brain Research Reviews*, **56**(1), 79–88.

Braitenberg, V., and Schuz, A., 1998: *Cortex: statistics and geometry of neuronal connectivity*. Springer Berlin.

Breiman, L., 2001a: Random forests. *Machine Learning*, **45**(1), 5–32.

Breiman, L., 2001b: Random forests. *Mach. Learn.*, **45**(1), 5–32.

Breiman, L., and Cutler, A., 2007: Random forests - classification description. http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

Buzs, P., Kovcs, K., Ferecsk, A. S., Budd, J. M. L., Eysel, U. T., and Kisvrday, Z. F., 2006: Model-based analysis of excitatory lateral connections in the visual cortex. *The Journal of Comparative Neurology*, **499**(6), 861–881.

B, J. L., and Markram, H., 2006: Spontaneous and evoked synaptic rewiring in the neonatal neocortex. *Proceedings of the National Academy of Sciences*, **103**(35), 13214 –13219.

Chen, B. L., Hall, D. H., and Chklovskii, D. B., 2006: Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(12), 4723–4728.

Cherniak, C., Changizi, M., and Kang, D. W., 1999: Large-scale optimization of neuron arbors. *Physical Review E*, **59**(5), 6001.

Chklovskii, D. B., 2000: Optimal sizes of dendritic and axonal arbors in a topographic projection. *J Neurophysiol*, **83**(4), 2113–2119.

Chklovskii, D. B., 2004: Synaptic connectivity and neuronal morphology two sides of the same coin. *Neuron*, **43**(5), 609617.

Contreras, D., Destexhe, A., and Steriade, M., 1997: Intracellular and computational characterization of the intracortical inhibitory control of synchronized thalamic inputs in vivo. *Journal of neurophysiology*, **78**(1), 335.

Dacey, D., 1993: The mosaic of midget ganglion cells in the human retina. *J. Neurosci.*, **13**(12), 5334–5355.

Dacey, D. M., and Petersen, M. R., 1992: Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences of the United States of America*, **89**(20), 9666–9670.

Dailey, M. E., and Smith, S. J., 1996: The dynamics of dendritic structure in developing hippocampal slices. *J. Neurosci.*, **16**(9), 2983–2994.

Duan, H., Wearne, S. L., Morrison, J. H., and Hof, P. R., 2002: Quantitative analysis of the dendritic morphology of corticocortical projection neurons in the macaque monkey association cortex. *Neuroscience*, **114**(2), 349359.

Dunn-Meynell, A. A., and Sharma, S. C., 1986: The visual system of the channel catfish (Ictalurus punctatus). i. retinal ganglion cell morphology. *The Journal of Comparative Neurology*, **247**(1), 32–55.

Engert, F., and Bonhoeffer, T., 1999: Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature*, **399**(6731), 66–70.

Famiglietti, E., Kaneko, A., and Tachibana, M., 1977: Neuronal architecture of on and off pathways to ganglion cells in carp retina. *Science*, **198**(4323), 1267–1269.

Famiglietti, E., and Kolb, H., 1976: Structural basis for ON-and OFF-center responses in retinal ganglion cells. *Science*, **194**(4261), 193–195.

Finlay, B., and Darlington, R., 1995: Linked regularities in the development and evolution of mammalian brains. *Science*, **268**(5217), 1578–1584.

Fuerst, P. G., Bruce, F., Tian, M., Wei, W., Elstrott, J., Feller, M. B., Erskine, L., Singer, J. H., and Burgess, R. W., 2009: DSCAM and DSCAML1 function in Self-Avoidance in multiple cell types in the developing mouse retina. *Neuron*, **64**(4), 484–497.

Fuerst, P. G., Koizumi, A., Masland, R. H., and Burgess, R. W., 2008: Neurite arborization and mosaic spacing in the mouse retina require DSCAM. *Nature*, **451**(7177), 470–474.

Furtak, S. C., Moyer, J. R., and Brown, T. H., 2007: Morphology and ontogeny of rat perirhinal cortical neurons. *J Comp Neurol*, **505**(5), 493–510.

Gibb, R., and Kolb, B., 1998: A method for vibratome sectioning of Golgi-Cox stained whole rat brain. *J Neurosci Methods*, **79**(1), 1–4.

Gonzalez-Burgos, G., Krimer, L. S., Povysheva, N. V., Barrionuevo, G., and Lewis, D. A., 2005: Functional properties of fast spiking interneurons and their synaptic connections with pyramidal cells in primate dorsolateral prefrontal cortex. *J Neurophysiol*, **93**(2), 942–953.

Gonzalez-Burgos, G., Krimer, L. S., Urban, N. N., Barrionuevo, G., and Lewis, D. A., 2004: Synaptic efficacy during repetitive activation of excitatory inputs in primate dorsolateral prefrontal cortex. *Cereb. Cortex*, **14**(5), 530–542.

Gulledge, A. T., Kampa, B. M., and Stuart, G. J., 2005: Synaptic integration in dendritic trees. *Journal of Neurobiology*, **64**(1), 75–90.

Gustafsson, B., Wigstrm, H., Zimmer, J., and Storm-Mathisen, O. O. J., 1990: Chapter 16 long-term potentiation in the hippocampal CA1 region: its induction and early temporal development. In *Understanding the Brain Through the Hippocampus the Hippocampal Region as a Model for Studying Brain Structure and Function*, volume Volume 83, 223–232. Elsevier. ISBN 0079-6123.

Hand, D. J., and Till, R. J., 2001: A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning*, **45**(2), 171–186.

Hattori, D., Millard, S. S., Wojtowicz, W. M., and Zipursky, S. L., 2008: Dscam-Mediated cell recognition regulates neural circuit formation. *Annual Review of Cell and Developmental Biology*, **24**(1), 597–620.

Hebb, D. O., 2002: *The organization of behavior: a neuropsychological theory*. Routledge.

Hellwig, B., 2000: A quantitative analysis of the local connectivity between pyramidal neurons in layers of the rat visual cortex. *Biological Cybernetics*, **82**(2), 111121.

Hirsch, J. A., 1995: Synaptic integration in layer IV of the ferret striate cortex. *Journal of Physiology*, **483**(1), 183200.

Hirsch, J. A., Alonso, J., Reid, R. C., and Martinez, L. M., 1998: Synaptic integration in striate cortical simple cells. *J. Neurosci.*, **18**(22), 9517–9528.

Hirsch, J. A., Martinez, L. M., Alonso, J. M., Desai, K., Pillai, C., and Pierre, C., 2002: Synaptic physiology of the flow of information in the cat's visual cortex in vivo. *The Journal of Physiology*, **540**(1), 335.

Hitchcock, P., and Easter, S., 1986: Retinal ganglion cells in goldfish: a qualitative classification into four morphological types, and a quantitative study of the development of one of them. *J. Neurosci.*, **6**(4), 1037–1050.

Huxlin, K. R., and Goodchild, A. K., 1997: Retinal ganglion cells in the albino rat: Revised morphological classification. *The Journal of Comparative Neurology*, **385**(2), 309–323.

Jacobs, B., Driscoll, L., and Schall, M., 1997: Life-span dendritic and spine changes in areas 10 and 18 of human cortex: a quantitative golgi study. *The Journal of Comparative Neurology*, **386**(4), 661680.

Jacobs, B., Schall, M., Prather, M., Kapler, E., Driscoll, L., Baca, S., Jacobs, J., Ford, K., Wainwright, M., and Treml, M., 2001: Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cereb Cortex*, **11**(6), 558–71.

Jaffe, D. B., and Carnevale, N. T., 1999: Passive normalization of synaptic integration influenced by dendritic architecture. *J Neurophysiol*, **82**(6), 3268–3285.

Jaiantilal, A., 2009: Classification and regression by randomForest-matlab.

Kalisman, N., Silberberg, G., and Markram, H., 2003: Deriving physical connectivity from neuronal morphology. *Biological Cybernetics*, **88**(3), 210–218.

Kalisman, N., Silberberg, G., and Markram, H., 2005: The neocortical microcircuit as a tabula rasa. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(3), 880 –885.

Kisvarday, Z. F., and Eysel, U. T., 1992: Cellular organization of reciprocal patchy networks in layer III of cat visual cortex (area 17). *Neuroscience*, **46**(2), 275.

Kisvarday, Z. F., Ferecsk, A. S., Kovcs, K., Buzs, P., Budd, J. M. L., and Eysel, U. T., 2002: One axon-multiple functions: Specificity of lateral inhibitory connections by large basket cells. *Journal of Neurocytology*, **31**(3), 255–264.

Kolb, H., Nelson, R., and Mariani, A., 1981: Amacrine cells, bipolar cells and ganglion cells of the cat retina: A golgi study. *Vision Research*, **21**(7), 1081–1114.

Kong, J. H., Fish, D. R., Rockhill, R. L., and Masland, R. H., 2005: Diversity of ganglion cells in the mouse retina: Unsupervised morphological classification and its limits. *Journal of Comparative Neurology*, **489**(3), 293310.

Krimer, L. S., Zaitsev, A. V., Czanner, G., Kroner, S., Gonzalez-Burgos, G., Povysheva, N. V., Iyengar, S., Barrionuevo, G., and Lewis, D. A., 2005: Cluster analysis-based physiological classification and morphological properties of inhibitory neurons in layers 2-3 of monkey dorsolateral prefrontal cortex. *Journal of neurophysiology*, **94**(5), 3009.

Larkman, A., and Mason, A., 1990: Correlations between morphology and electrophysiology of pyramidal neurons in slices of rat visual cortex. i. establishment of cell classes. *J. Neurosci.*, **10**(5), 1407–1414.

Lee, S., and Stevens, C. F., 2007: General design principle for scalable neural circuits in a vertebrate retina. *Proc Natl Acad Sci U S A*, **104**(31), 12931–5.

Lee, W. A., Huang, H., Feng, G., Sanes, J. R., Brown, E. N., So, P. T., and Nedivi, E., 2005: Dynamic remodeling of dendritic arbors in GABAergic interneurons of adult visual cortex. *PLoS Biol*, **4**(2), e29.

Lendvai, B., Stern, E. A., Chen, B., and Svoboda, K., 2000: Experience-dependent plasticity of dendritic spines in the developing rat barrel cortex in vivo. *Nature*, **404**(6780), 876–881.

Liley, D. T. J., and Wright, J. J., 1994: Intracortical connectivity of pyramidal and stellate cells: estimates of synaptic densities and coupling symmetry. *Network: Computation in Neural Systems*, **5**(2), 175.

Ling, C. X., Huang, J., and Zhang, H., 2003: AUC: a statistically consistent and more discriminating measure than accuracy. In *Proceedings of the 18th international joint conference on Artificial intelligence*, 519–524. Morgan Kaufmann Publishers Inc., Acapulco, Mexico.

Liu, Y. S., Stevens, C. F., and Sharpee, T. O., 2009: Predictable irregularities in retinal receptive fields. *Proceedings of the National Academy of Sciences*, **106**(38), 16499 –16504.

Madison, D. V., Malenka, R. C., and Nicoll, R. A., 1991: Mechanisms underlying Long-Term potentiation of synaptic transmission. *Annual Review of Neuroscience*, **14**(1), 379–397.

Mainen, Z. F., and Sejnowski, T. J., 1996: Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature*, **382**(6589), 363366.

Malinow, R., Mainen, Z. F., and Hayashi, Y., 2000: LTP mechanisms: from silence to four-lane traffic. *Current Opinion in Neurobiology*, **10**(3), 352–357.

Mangrum, W. I., Dowling, J. E., and Cohen, E. D., 2003: A morphological classification of ganglion cells in the zebrafish retina. *Visual Neuroscience*, **19**(06), 767779.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C., 2004: Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci*, **5**(10), 793–807.

Mason, A., and Larkman, A., 1990: Correlations between morphology and electrophysiology of pyramidal neurons in slices of rat visual cortex. II. electrophysiology. *J. Neurosci.*, **10**(5), 1415–1428.

McDonald, C. G., Dailey, V. K., Bergstrom, H. C., Wheeler, T. L., Eppolito, A. K., Smith, L. N., and Smith, R. F., 2005: Periadolescent nicotine administration produces enduring changes in dendritic morphology of medium spiny neurons from nucleus accumbens. *Neurosci Lett*, **385**(2), 163–7.

Mel, B. W., 1994: Information processing in dendritic trees. *Neural Computation*, **6**(6), 1031–1085.

Nagarajan, N., and Stevens, C. F., 2008: How does the speed of thought compare for brains and digital computers? *Current biology*, **18**(17).

Nelson, R., Famiglietti, E. V., and Kolb, H., 1978: Intracellular staining reveals different levels of stratification for on- and off-center ganglion cells in cat retina. *J Neurophysiol*, **41**(2), 472–483.

Parrish, J. Z., Emoto, K., Kim, M. D., and Jan, Y. N., 2007: Mechanisms that regulate establishment, maintenance, and remodeling of dendritic fields. *Annu. Rev. Neurosci.*

Schubert, D., Kotter, R., Luhmann, H. J., and Staiger, J. F., 2006: Morphology, electrophysiology and functional input connectivity of pyramidal neurons characterizes a genuine layer va in the primary somatosensory cortex. *Cereb Cortex*, **16**(2), 223–36.

Schubert, D., Kotter, R., Zilles, K., Luhmann, H. J., and Staiger, J. F., 2003: Cell type-specific circuits of cortical layer IV spiny neurons. *J Neurosci*, **23**(7), 2961–70.

Sholl, D. A., 1953: Dendritic organization in the neurons of the visual and motor cortices of the cat. *Journal of Anatomy*, **87**(Pt 4), 387–406.1.

Sholl, D. A., and Uttley, A. M., 1953: Pattern discrimination and the visual cortex. *Nature*, **171**(4348), 387–388.

Snider, J., Pillai, A., and Stevens, C. F., 2010: A universal property of axonal and dendritic arbors. *Neuron*, **66**(1), 45–56.

Song, S., Sjstrm, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B., 2005: Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol*, **3**(3), e68.

Squire, L. R., 2003: *Fundamental neuroscience*. Academic Press.

Staiger, J. F., Flagmeyer, I., Schubert, D., Zilles, K., Kotter, R., and Luhmann, H. J., 2004: Functional diversity of layer IV spiny neurons in rat somatosensory cortex: quantitative morphology of electrophysiologically characterized and biocytin labeled cells. *Cereb Cortex*, **14**(6), 690–701.

Stepanyants, A., Hof, P. R., and Chklovskii, D. B., 2002: Geometry and structural plasticity of synaptic connectivity. *Neuron*, **34**(2), 275288.

Stepanyants, A., Tams, G., and Chklovskii, D. B., 2004: Class-specific features of neuronal wiring. *Neuron*, **43**(2), 251259.

Stuart, G., and Spruston, N., 1998: Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. *Journal of Neuroscience*, **18**(10), 3501.

Sun, W., Li, N., and He, S., 2002: Large-scale morphological survey of mouse retinal ganglion cells. *The Journal of Comparative Neurology*, **451**(2), 115–126.

Trachtenberg, J. T., Chen, B. E., Knott, G. W., Feng, G., Sanes, J. R., Welker, E., and Svoboda, K., 2002: Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature*, **420**(6917), 788–794.

Travis, K., Ford, K., and Jacobs, B., 2005: Regional dendritic variation in neonatal human cortex: a quantitative golgi study. *Dev Neurosci*, **27**(5), 277–87.

Tsiola, A., Hamzei-Sichani, F., Peterlin, Z., and Yuste, R., 2003: Quantitative morphologic classification of layer 5 neurons from mouse primary visual cortex. *The Journal of Comparative Neurology*, **461**(4), 415428.

Uttley, A., 1954: The classification of signals in the nervous system. *Electroencephalography and Clinical Neurophysiology*, **6**, 479–494.

Uttley, A. M., 1955: The probability of neural connexions. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 229240.

Uylings, H. B., and van Pelt, J., 2002: Measures for quantifying dendritic arborizations. *Network: Computation in Neural Systems*, **13**(3), 397414.

Vetter, P., Roth, A., and Hausser, M., 2001: Propagation of action potentials in dendrites depends on dendritic morphology. *Journal of Neurophysiology*, **85**(2), 926.

Vlgyi, B., Chheda, S., and Bloomfield, S. A., 2009: Tracer coupling patterns of the ganglion cell subtypes in the mouse retina. *The Journal of Comparative Neurology*, **512**(5), 664–687.

Wang, Y., Gupta, A., Toledo-Rodriguez, M., Wu, C. Z., and Markram, H., 2002: Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cerebral Cortex*, **12**(4), 395.

Wassle, H., and Boycott, B. B., 1991: Functional architecture of the mammalian retina. *Physiol Rev*, **71**(2), 447–80.

Wassle, H., Peichl, L., and Boycott, B. B., 1981: Dendritic territories of cat retinal ganglion cells. *Nature*, **292**(5821), 344–345.

Watson, K. K., Jones, T. K., and Allman, J. M., 2006: Dendritic architecture of the von economo neurons. *Neuroscience*, **141**(3), 11071112.

Wen, Q., and Chklovskii, D. B., 2008: A cost-benefit analysis of neuronal morphology. *Journal of neurophysiology*, **99**(5), 2320.

Wen, Q., Stepanyants, A., Elston, G. N., Grosberg, A. Y., and Chklovskii, D. B., 2009: Maximization of the connectivity repertoire as a statistical principle governing the shapes of dendritic arbors. *Proceedings of the National Academy of Sciences*, **106**(30), 12536 –12541.

y Cajal, S. R., 1899: *Comparative study of the sensory areas of the human cortex*.

y Cajal, S. R., Pasik, P., and Pasik, T., 2002: *Texture of the nervous system of man and the vertebrates*. Springer.

Yamahachi, H., Marik, S. A., McManus, J. N., Denk, W., and Gilbert, C. D., 2009: Rapid axonal sprouting and pruning accompany functional reorganization in primary visual cortex. *Neuron*, **64**(5), 719–729.

Zaitsev, A. V., Gonzalez-Burgos, G., Povysheva, N. V., Kroner, S., Lewis, D. A., and Krimer, L. S., 2005: Localization of calcium-binding proteins in physiologically and morphologically characterized interneurons of monkey dorsolateral prefrontal cortex. *Cerebral Cortex*, **15**(8), 1178.