Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Science

presented by

MSc ETH Jörgen Mathias Roland Kornfeld
born in Ulm, Germany
Oral-examination date: 28th of July 2017

Connectomic Analyses in the Zebra Finch Brain

Referees: Prof. Dr. Stephan Frings
Prof. Dr. Winfried Denk

Declarations according to § 8 (3) b) and c) of the doctoral degree regulations:

b) I hereby declare that I have written the submitted dissertation myself and in this process have used no other sources or materials than those expressly indicated;

c) I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in this or any other form at any other institution as an examination paper, nor submitted it to any other faculty as a dissertation.

Heidelberg, 31st of May, 2017

_____

Jörgen Kornfeld

Summary

Songbirds (*Passeri*) are one of the few known orders of animals capable of vocal learning. The male zebra finch (*Taeniopygia guttata*), a model organism for the neurobiological basis of speech, performs a highly stereotyped courtship song of about one second in length. Juvenile finches learn their song through imitation of their fathers (or other conspecific males) during overlapping sensory and sensorimotor periods and continue to improve until they reach adulthood, after which the song becomes stable. The finch's vocal organ, the syrinx, is controlled by an interconnected set of brain nuclei that can be subdivided into pathways required for song production and song learning. Of these nuclei, HVC (a premotor cortical area) and Area X (a striatopallidal area) are of particular interest. HVC has been identified as a neuronal sequence generator at the apex of the sensorimotor hierarchy, and Area X has been shown to constitute the center of the reinforcement learning apparatus used during vocal learning. The microcircuits of both nuclei have been analyzed extensively using electrophysiological methods and light microscopy-based neuroanatomy, but a detailed analysis of their synaptic wiring based on high-throughput serial block-face electron microscopy (SBEM) is lacking.

Volume electron microscopy datasets of HVC and Area X were therefore acquired to test well-defined anatomical hypotheses about their synaptic connectivity as predicted by mechanistic circuit models:
1) That the neuronal sequence generator located in HVC is based on direct synaptic connections between the sequence neurons, as predicted by the synfire chain hypothesis.
2) That efference copy and context signal axons in Area X have different synaptic targets, as predicted by a recent reinforcement learning model of the basal-ganglia.
Both hypotheses were found to be consistent with the acquired data, supporting the notion that structural neurobiology, otherwise known as connectomics, enables powerful tests of neurobiological hypotheses.
To help overcome the technological challenges faced during the analysis of large volume EM datasets, a new annotation workflow (GraphWalker) was implemented that allows neurite reconstruction to proceed at least an order of magnitude faster than the current state-of-the-art. Additionally, a machine learning-based computational pipeline (SyConn) was developed and used to derive the first synaptic wiring diagram of the vertebrate basal-ganglia from one of the acquired Area X datasets.

Zusammenfassung

Singvögel (*Passeri*) sind eine der wenigen Tierordnungen die in der Lage sind neuartige Vokalisationen zu erlernen. Der männliche Zebrafink (*Taeniopygia guttata*), ist ein Modellorganismus für die Erforschung der neurobiologischen Grundlagen von Sprache, und erlernt einen hochgradig stereotypen Balzgesang von ungefähr einer Sekunde Länge. Jungfinken imitieren dabei den Gesang ihres Vaters (oder eines anderen konspezifischen männlichen Vogels) in einer sensorischen und sensorimotorischen Lernphase bis zum Erwachsenenalter, wonach der Gesang stabil bleibt. Das Gesangsorgan der Vögel, die Syrinx, wird durch miteinander verschaltete Hirnkerne kontrolliert, die zwei Verarbeitungssträngen zugewiesen werden können: einem Pfad zuständig für die Gesangsproduktion und einem für das Erlernen. Zwei Hirnkerne sind dabei von besonderem Interesse, das prämotorische kortikale HVC und die striatopallidale Area X, da HVC als neuronaler Sequenzgenerator an der Spitze der sensorimotorischen Verarbeitungshierarchie identifiziert wurde, und Area X das Zentrum des Verstärkungslernapparates für das Gesangslernen darstellt. Die Mikroschaltkreise dieser zwei Hirnkerne wurden mittels elektrophysiologischer Methoden und lichtmikroskopie-basierter Neuroanatomie ausgiebig untersucht, allerdings fehlt eine detaillierte Analyse der synaptischen Verschaltung auf der Basis von serieller Hochdurchsatzelektronenmikroskopie der Blockoberfläche (SBEM). Volumen EM Datensätze von HVC und Area X wurden daher aufgenommen um klar definierte anatomische Hypothesen zur synaptischen Verschaltung zu testen, basierend auf den Vorhersagen von mechanistischen Schaltkreismodellen:

1) Dass der in HVC lokalisierte neuronale Sequenzgenerator auf direkten synaptischen Verbindungen zwischen den Schrittmacherzellen basiert, vorhergesagt von der *synfire chain* Hypothese.

2) Dass Efferenzkopie- und Kontextsignal tragende Axone in Area X unterschiedliche synaptische Ziele haben, vorhergesagt durch ein aktuelles Verstärkungslernmodell der Basalganglien.

Die in dieser Arbeit gewonnenen Daten waren konsistent mit beiden Modellen; was zeigt, dass die strukturelle Neurobiologie, auch bekannt als *connectomics*, mächtige Hypothesentests in der Neurobiologie ermöglicht. Um die technologischen Herausforderungen bei der Analyse von großen Volumen EM Datensätzen besser meistern zu können wurde ein neues Verfahren entwickelt (*GraphWalker*), das das Rekonstruieren von Neuriten um mindestens eine Größenordnung beschleunigt. Zudem wurde anhand von einem der Area X Datensätze das erste synaptische Verschaltungsdiagramm der Wirbeltier-Basalganglien erzeugt, wozu ein auf maschinellem Lernen basierendes Computersystem entwickelt wurde (*SyConn*).

Acknowledgments

First, I want to thank my advisor Prof. Dr. Winfried Denk for his intellectual and general support in all my undertakings during the past years. Without his advise, wit and help in the laboratory, this thesis would not have been possible. I also want to thank my close friend and colleague Dr. Fabian Svara for the many scientific and nonscientific discussions which made work productive and entertaining together. I am also grateful to Dr. Stefanie Preuss, Dr. Chintan Trivedi, Dr. Benjamin Titze, Dr. Jonas Binding, Dr. Shawn Mikula, Dr. Johann Bollmann and Sarah Mikula from the Denk department and Bollmann group for many interesting discussions and good times in Heidelberg.

Furthermore, I would like to thank my students, Sven Dorkenwald (Bachelor's thesis), Philipp Schubert (Master's thesis), Marius Killinger (Master's thesis), Larissa Heinrich (Bachelor's and Master's thesis), Julia Buhmann (Master's thesis), Lukas Burchartz (Master's thesis), Lukas Schott (Bachelor's thesis), My-Tien Nguyen (Bachelor's thesis) and Andreas Knecht (Master's thesis). I hope I could pass on some of the rigorous scientific thinking that Winfried tried to teach me.

I would like to thank Joseph (Joe) Scherrer for excellent comments on the manuscript.

I would also like to thank the KNOSSOS team (`knossostool.org`) that took over hands-on programming after Fabian and I had to start working on new projects, namely Norbert Pfeiler, My-Tien Nguyen, Alexander Stepanov, Michal Pronkin, Oren Shatz, Matthias Wegner and Claus Ripp.

I also want to thank my collaborators from New York University, Prof. Dr. Michael Long and Dr. Sam Benezra, and my collaborators from MIT, Prof. Dr. Michale Fee and Michael Stetner, for giving me the opportunity of doing very inspiring science together. I further owe gratitude to the department of Prof. Dr. Manfred Gahr at the MPI of Ornithology in Seewiesen, especially Dr. Moritz Hertel, for giving me access to their zebra finch colony.

I would like to express my gratitude to the outstanding support staff at the MPI for medical Research in Heidelberg, especially Klaus Bauer, Thomas Stather, Chris Roome and Mario Hilpert for IT support, Jürgen Tritthardt and his team for electronics, and Niklas Neef and Martin Lukat of the mechanical workshop.

Additionaly, I want to thank the Boehringer Ingelheim Fonds for financial support and the excellent seminars and retreats which accompanied my doctoral training.

I want to thank the many student assistants who helped with the data annotations. First, the HiWi management team, Hannah Jakobi, Liesa-Marie Schreiber, Anna Biasotto, Annika Gable, Malte Grohmann, Danny Baltissen, Stefan Bassler, Julia Hammerich, Nathalie Tisch, Christian Röhrig, Joris Nassal, Ivo Sonntag, Dimitar Dimitrov and Ji-Hyun Youm.

Second, the many songbird HiWis who annotated neurons and synapses: H. Adam, O. Aderhold, A. Al-Shaboti, T. Albers, M. Alsara, M. Ament, A. Andrade Garcia, E. Atsiatorme, D. Azabdaftari, K. Bach, Y. Badiye, M. Balk, D. Baltissen, F. Bassler, S. Bassler, C. Bauerdick, A. Baumeister, A. Becker, C. Becker, K. Berneiser, J. Bessler, D. Bornhorst, F. Braun, L. Burchartz, D. Dimitrov, J. Düchting, J. Dunder, M. Ederer, L. Ehm, T. Eliguezel, L. Fey, J. Foehr, J. Fuchs, A. Gaubatz, H. Gothe, M. Groß, E. Grün, M. Haan, J. Hammerich, L. Hammes, S. He, H. Hees, J. Hendricks, L. Henkel, C. Hinzmann, F. Höfert, J. Huether, S. Hutzl, R. Janssen, K. Kappler, T. Kehrer, K. Kehrle, J. Kern, K. Kiesl, O. Kornfeld, M. Krause, J. Kupke, C. Lance, R. Lichtenberger, M. Lilow, J. Lin, S. Lorger, M. Maiwald, R. Majidi, R. Mann, G. Marschmann, S. Mayorca de Ituarte, S. Mertens, I. Mintert, M. Müller, O. Müller, P. Müller, M. Nathusius, V. Nguyen Chi, M. Nick, T. Noel, L. Obenauer, D. Parthier, G. Patzer, J. Pellegrino, M. Pohrath, D. Raica, J. Reichert, M. Reitz, N. Rieger, S. Ruder, P. Ruthig, O. Salem, C. Sandhof, C. Schaefer, M. Schiedeck, J. Schlotthauer, C. Schmelas, M. Schoen, L. Schott, M. Schramm, F. Schuebel, A. Schuerholz, J. Schwab, D. Schwarz, C. Serger, J. Serwani, V. Sievers, M. Siggel, Y. Soehngen, A. Sons, D. Talwar, K. Tessmer, T. Toeroek, M. Trache, J. Tytko, P. Varade, V. Viehweger, H. Volpert, L. Volz, J. Volz , D. Wachmann, F. Weichel, M. Weiss, Y. Yu, S. Zaschke, I. Zelienka, T. Zhang, E. Zimmermann, A. de la Porte, L. von Hoessle.

I want to thank my entire family and especially my parents, Dr. Günter and Ursula Kornfeld, who continuously supported me and made sure that I did not get lost in the basement with all those electron microscopes.

Finally, I want to thank my wife Olesya for her continuous support and understanding, her love, and our wonderful time together that allows me to recharge my batteries.

Dedication

This work is dedicated to my teacher Mr. Fella, who introduced me to the world of studying biology.

‘It is not because things are difficult that we do not dare, it is because we do not dare that they are difficult.‘

Lucius Annaeus Seneca

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Structural neurobiology with the electron microscope

1.1.1 From synapses to the first connectome

In 1888 Santiago Ramón y Cajal presented evidence for the "neuron doctrine," the notion that the brain consists of separated cells that interact at synapses (Sherrington, 1906), by studying the cerebellum of birds with Golgi staining (Cajal, 1888). However, "reticularists," proponents of the idea that the brain consists of a cellular syncitium, could not be convinced initially (reviewed in Cimino, 1999). Only electron microscopy (EM) provided sufficient resolution to demonstrate unambiguously that the central nervous system's neurons are separated at their synaptic clefts (Palade and Palay, 1954). However, the high resolution of the electron microscope was both a benefit and a drawback, as it could reveal details never seen before but restricted datasets to small regions. In 1986, John White et al. presented for the first time a wiring diagram (or "connectome") of the complete *Caenorhabditis elegans* hermaphrodite brain, based on serial section transmission electron microscopy (ssTEM) (White et al., 1986). The imaging of the serial sections and the tracing of all 302 neurons required several years of work,

illustrating the difficulties the pioneers of structural neurobiology faced. At this point, volume electron microscopy was a completely manual process, despite first attempts by Leighton in 1981 to automate in-chamber ultramicrotomy and perform alternating cutting and block-face imaging (Leighton, 1981).

1.1.2 Automated volume electron microscopy

The technological progress of the last decades also led to new developments for volume EM, resulting in automation of many of the previously manual steps (see Briggman and Bock, 2012, for a comprehensive review). Apart from electron tomography, which is limited in its field of view (Mahamid and Baumeister, 2012), all currently used volume electron microscopy techniques are based on sample sectioning and subsequent 2D-imaging to construct a 3D-image. Two main techniques can be distinguished: block-face imaging, where the sample is alternatingly ablated and imaged, and serial section imaging, where thin sections are cut, collected and imaged independently. Two inherent advantages of section imaging techniques over block-face imaging are that the imaging can easily be parallelized if several electron microscopes are available, and that section quality can be evaluated before the time-consuming imaging step. The serial sectioning approach with the highest degree of automation is currently the Automatic Tape-collecting Ultra-Microtome (ATUM), an extension of a conventional diamond-knife ultramicrotome that was build by Kenneth Hayworth in his garage and applied to brain tissue in Jeff Lichtman's laboratory (Kasthuri et al., 2015). Thin sections are collected on a tape that is later mounted in short lengths onto silicon wavers for scanning EM (SEM) imaging, since TEM imaging is not compatible with the Kapton tape typically used.

However, despite the mentioned advantages of the serial sectioning approaches, imaging of the heavy metal-stained tissue block-face usually offers better data quality because of the higher z-resolution and smaller number of artifacts. Many problems of the non-block-face methods, such as section wrinkles or difficult image alignment, can be avoided (Wanner et al., 2015). The block-face imaging method that offers the highest resolution (5 x 5 x 5 nm voxel size) is focused ion-beam (FIB) SEM (Knott et al., 2008), in which the top layer of the sample is

milled away by means of targeted Ga^+ ions. The main disadvantage of FIB-SEM for structural neurobiology is currently the limited size of an individual dataset (Wanner et al., 2015), but near-lossless sample partitioning through a "hot-knife" and parallel FIB-SEM imaging of the individual chunks might overcome this limitation (Hayworth et al., 2015).

Diamond knife-based serial block-face electron microscopy (DiK-SBEM, or simply SBEM) makes use of an in-chamber ultramicrotome to alternate cutting and imaging. As mentioned previously, it was devised by Leighton in 1981, but was not successfully demonstrated until 2004 in Winfried Denk's laboratory (Denk and Horstmann, 2004). SBEM does not suffer from the limited field of view of FIB-SEM but is typically used at a lower resolution (10 x 10 x 25 nm voxel size) because of cutting difficulties at high electron doses (Wanner et al., 2015).
In addition to the aforementioned problems, three main factors can be identified that practically limit the acquisition of large volume EM datasets:

1. Sample staining difficulties for samples exceeding several hundreds of micrometers.

2. Cutting or image acquisition process reliability.

3. Insufficient imaging speeds.

While the first problem has been solved for brains up to the size of the entire mouse brain by increasing the penetration limit of the commonly used *en-bloc* heavy-metal fixation and staining (Mikula et al., 2012; Mikula and Denk, 2015), image acquisition speed and reliability have remained unsolved problems for years. However, despite a lack of published datasets, technological solutions for the imaging problem are in sight. Zeiss has developed a novel multi-beam SEM that parallelizes the acquisition of a single micrograph over up to 91 electron beams (Kemen et al., 2015). There are also ongoing attempts to parallelize dataset acquisition over multiple microscopes (Hayworth et al., 2015). Despite these dramatic advances that will be reflected in increasing dataset sizes, data analysis remains a challenge, and dense connectomic analysis of datasets over a volume of $(100\,\mu m)^3$ appears almost insurmountable (Helmstaedter, 2013) using the current manual analysis methods as illustrated in the following paragraph.

1.1.3 The data analysis bottleneck(s)

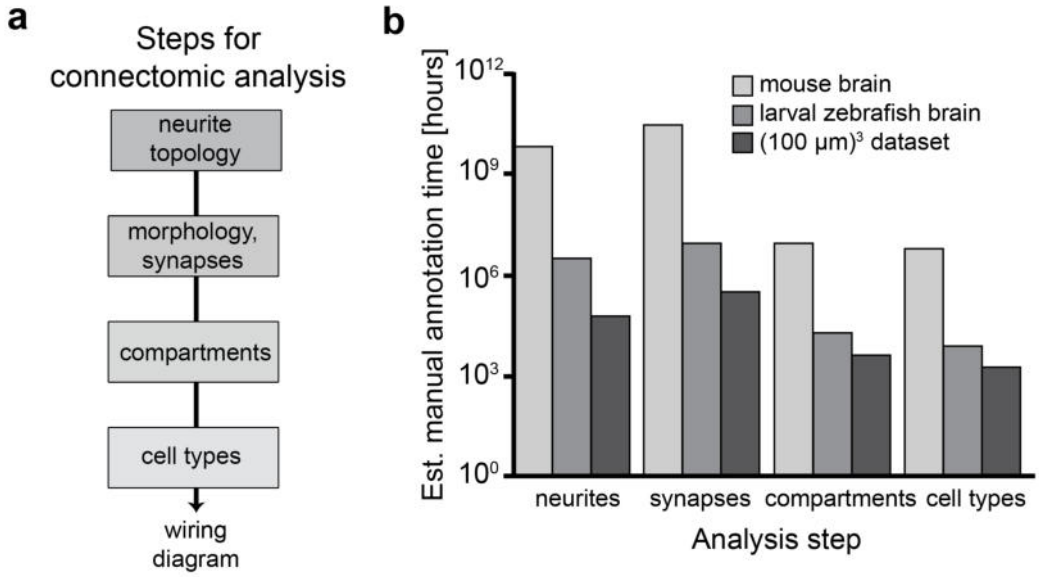

Figure 1.1: *Steps in wiring diagram extraction from volume EM datasets. (a) Necessary analysis steps for a connectome. (b) Estimated manual annotation times for dense (i.e. all neurites in a given volume) analysis. Shown are skeletonization times for the neurite reconstruction step, the fastest method up to date (Berning et al., 2015). Data from (Dorkenwald et al., 2017).*

The *C. elegans* hermaphrodite connectome (White et al., 1986) was reconstructed manually from printed electron micrographs with color pens after it was realized in Sydney Brenner's laboratory that computer graphics were not yet capable enough, although an attempt was made to build a computer system to assist the reconstructions (reviewed in Emmons, 2015). Despite the relatively small size of the *C. elegans* connectome of only 302 neurons and 7,000 connections (White et al., 1986), the reconstruction still took 15 years from project start to publication (Emmons, 2015). In comparison to *C. elegans*, the fruit fly (*Drosophila melanogaster*) brain has about 135,000 neurons (Alivisatos et al., 2012) and the brain of a mouse has about 71,000,000 neurons (Herculano-Houzel et al., 2006), which demonstrates that dense connectomic analysis is not feasible with manual approaches.

Despite many attempts over the last decade to automate neurite reconstructions (Berning et al., 2015; Nunez-Iglesias et al., 2014; Huang and Jain, 2013; Helmstaedter et al., 2013; Beier et al., 2017), there is still no reconstruction method that is faster than manual neuron skeletonization combined with automated neurite "outgrowing" to obtain a full volume representation (Bern-

ing et al., 2015; Helmstaedter et al., 2013). This lack of suitable automated methods can be explained by the extremely low voxel-wise error rates that are required when the neurite reconstruction problem is viewed as a 3D image segmentation problem, where each voxel has to be assigned to a neuron or background. The simplest segmentation algorithm available, connected component analysis, requires that all neurites are separated by boundary voxels, and a single voxel error can falsely merge two neurons over an entire dataset, rendering their reconstructions useless without manual correction.

However, automated neurite reconstruction is not the only unsolved problem in data analysis, and somewhat surprisingly, the dense manual analysis of synapses is about as time consuming as neurite reconstruction (see Fig. 1.1 for estimates of various dataset sizes; Dorkenwald et al., 2017). The high cost of synapse analysis can be explained by the high density of contact sites between neurites in the neuropil, about 15 contacts per μm^3 for extracellular space (ECS) preserved samples (Dorkenwald et al., 2017), all of which require inspection. Steadily growing datasets make it increasingly difficult to manually annotate the compartments (e.g. axons, dendrites, somata and spines) of all contained cells or assign a cell type label to the neurons (Fig. 1.1), illustrating further that structural neurobiology depends critically on progress in automated analysis.

1.1.4 Limits and promises of structural neurobiology

Despite the profound insights that were gained in our understanding of neurons and synapses from studies performed after the initial availability of electron microscopy (see section 1.1.1), breakthroughs in light microscopy (LM) have started to dominate neuroscience research (reviewed in Evanko et al., 2011) and interest in EM-based circuit reconstruction has been fading (Helmstaedter et al., 2008). However, recent advances in automation for the collection of volume EM datasets have triggered interest again (Briggman and Bock, 2012; Denk et al., 2012) and could also help to overcome the single animal and dataset (n = 1) study problem that still plagues the field by enabling intra- and interspecies comparative connectomics (Bumbarger et al., 2013).

Despite these advances, the most obvious limitation of an EM dataset, regardless of its size or resolution, remains the lack of temporal physiological information. Experiments that combine optophysiological measurements such as two-photon imaging of Ca^{2+} indicators (Helmchen and Denk, 2005) with volume electron microscopy of the same cells therefore appear very powerful (reviewed in Helmstaedter et al., 2008). This approach is further supported by the uncertainty of whether biophysical models of cells and their connectivity, as reconstructed from volume electron microscopy datasets, will ever be able to reproduce even short timescale network activity. The question of network activity reproduction leads to the fundamental question of how information is stored in a network of neurons. The structural neurobiologist is tempted to answer this question with the statement that information is simply stored in the synapses; however, it is obvious that such a view is a gross oversimplification that neglects a plethora of evidence about the importance of mechanisms that cannot be identified directly by volume EM, such as:

- Genetic and epigenetic factors influencing information processing of neurons (Levenson and Sweatt, 2005)
- Biophysical factors, such as membrane receptor distributions (Lai and Jan, 2006)
- Volume conductance and direct electric coupling over gap junctions (Agnati and Fuxe, 2014), but see (Pallotto et al., 2015) for evidence of gap junction identification

While this list is certainly not complete, the long-term success of EM connectomics will be determined by the degree to which these and other factors can be systematically correlated with the ultrastructure of a neuron and its wiring. The establishment of such systematic correlations, if they exist, will always depend on physiological ground truth. However, even without the option to simulate neuronal networks bottom-up from volume EM datasets, a possibility which seems still far away, volume EM datasets can be used in exploratory approaches or to test defined anatomical circuit hypotheses. The next paragraphs will introduce the male zebra finch and its remarkable singing behavior as a basis for the description of the anatomical hypotheses studied in this thesis.

1.2 Vocal learning and song production in the zebra finch

1.2.1 Male zebra finches learn their song by imitation

Figure 1.2: Adult male zebra finch. By O. Kornfeld.

Songbirds, and in particular zebra finches, are a prominent model organism to study the neuroscience of learned behaviors (Mooney, 2009). Songbirds are some of only a few nonhuman animal species that produce learned vocalizations; other examples include elephants, bats, pinnipeds, and parrots (Petkov and Jarvis, 2012; Wang et al., 2014). Interest in songbird research strengthened after the discovery that malfunctions in the Forkhead box protein P2 (FoXP2) transcription factor affect both zebra finch song and human speech, hinting at possible similarities in the underlying brain circuits (reviewed in Vargha-Khadem et al., 2005). Developmentally, zebra finch song learning can be subdivided into overlapping phases: an initial sensory period (up to about 35 days post hatching, dph) followed by a sensorimotor period including different qualitative song stages (babbling or subsong and plastic song) which culminate in the adult crystallized song at about 90 to 120 dph (Immelmann, 1969; Price, 1979). The adult song is highly structured and can be subdivided into bouts of stereotyped song motifs of up to 1 s duration, with each motif consisting of several syllables. Some of these syllables can be further partitioned into notes separated by short pauses (Mooney, 2009). While male zebra finches naturally keep their own song motif fixed for their entire life following development, they are still capable of modifying it in conditional auditory feedback (CAF) paradigms (Andalman and Fee, 2009; Tumer and Brainard, 2007) that enable controlled learning settings. Importantly, it seems like the neural mechanism for

the adult modification and juvenile song acquisition are largely overlapping (Ölveczky et al., 2011) and make use of the same neuronal machinery. This neuronal machinery, in its entirety called the song system, will be further discussed in the next sections.

1.2.2 Nuclear organization of the song system in the bird brain

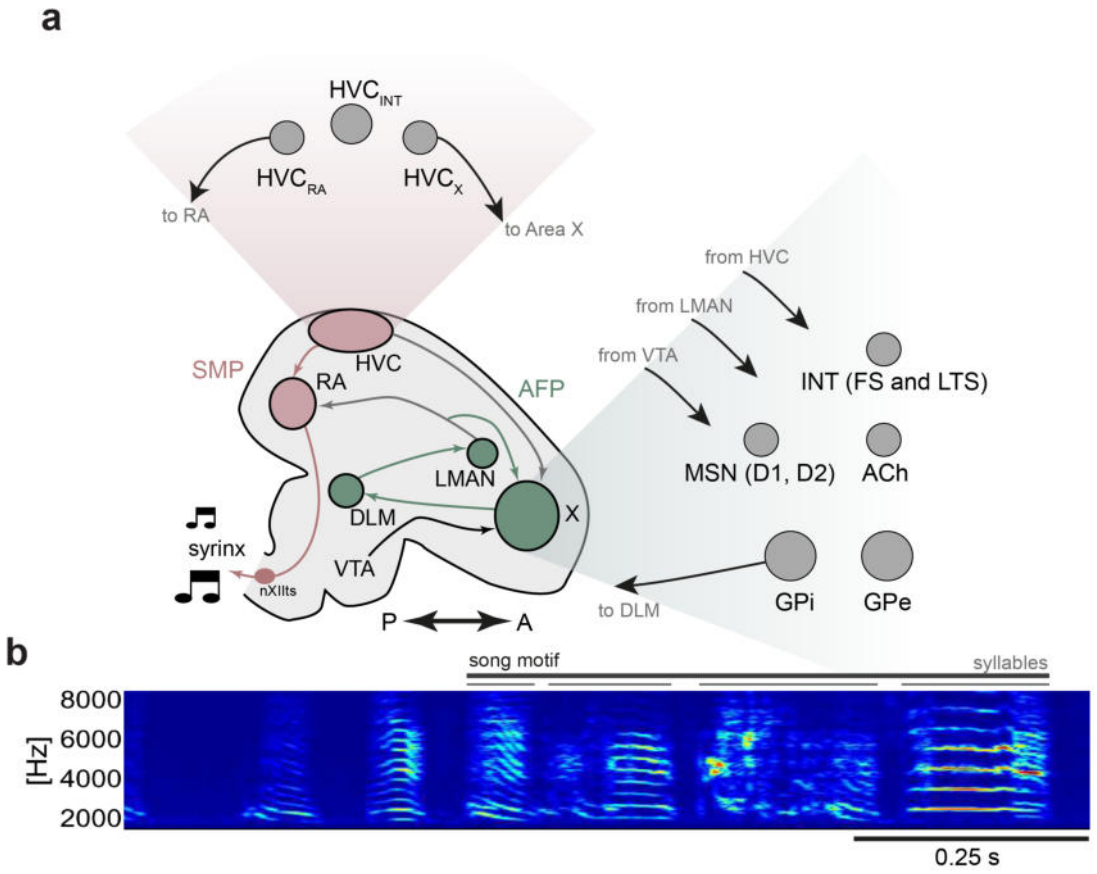

Figure 1.3: *The songbird brain and zebra finch song. (a) Schematic of the two main brain pathways that can be related to singing: the song motor pathway (SMP), responsible for the production of learned song, and the anterior forebrain path (AFP), required for song syllable changes and necessary for juvenile song learning. Also shown are the most important cell types within HVC and Area X, the two brain nuclei studied in this thesis, and main song system afferents and efferents of these areas. INT: Interneuron; FS: Fast-spiking; LTS: low-threshold-spiking; MSN: medium spiny neuron; D1 and D2: dopamine receptor types; ACh: cholinergic neuron; GPi and GPe: pallidal-like neurons of direct and indirect pathway. (b) Spectrogram of zebra finch song, including introductory calls at the beginning, followed by syllables that can be grouped into the stereotyped song motif (recorded by the author in the laboratory of R.H. Hahnloser, ETH, Zurich).*

The parts of the songbird brain that directly relate to singing can be subdivided functionally into areas required for the production of birdsong (song motor pathway, SMT) and areas that

are only required for its acquisition or modification in adulthood (anterior forebrain pathway, AFP), as depicted in Fig. 1.3 (Mooney, 2009). Auditory areas are clearly necessary for the sensory period and refinement phases of song learning. However, song production does not rely acutely on auditory regions, as demonstrated by deafening experiments, although song does deteriorate over longer time-scales without auditory feedback (Lombardino and Nottebohm, 2000). Lesions in the SMT always lead to severely degraded song (Nottebohm et al., 1976). In contrast, lesions in the AFP in adult birds do not alter the singing behavior immediately (Nottebohm et al., 1976; Andalman and Fee, 2009), although such birds have an impaired ability to adapt their song, especially the spectral structure of individual syllables (Tumer and Brainard, 2007; Andalman and Fee, 2009). While the separation of the song system into these two pathways allows the relation of function to macroscale anatomical structures in the songbird brain, it should not be forgotten that this dual pathway picture is certainly an oversimplification, especially given that both pathways diverge from and converge to the same brain areas, namely HVC and the robust nucleus of the archipallium, or RA.

1.2.3 HVC, the cortical song sequence generator

Macro-anatomical and functional characterization

HVC (letter-based proper name, see Reiner et al., 2004) is a brain area of the songbird caudal nidopallium with a volume of approximately 0.4 mm^3 (Kornfeld et al., 2017; Wang et al., 2002) that can be identified in all songbird species (Mooney, 2009). It was first described by Nottebohm et al. in 1976 in canaries and was associated with song production through lesion studies (Nottebohm et al., 1976). It was shown later that electrical stimulation of both HVC and RA elicits complex vocalizations that are part of the bird's own song (Vicario and Simpson, 1995). The importance of HVC was further realized through its placement at the apex of the sensorimotor hierarchy: it receives afferents mostly from auditory regions (NIf, CM), but projects into the primary motor area RA and Area X of the AFP (Mooney, 2009).

The neurons of HVC and their dynamics during singing

Three main neuronal cell types were described for HVC: small RA-projecting cells with spiny dendrites (HVC_{RA}), large basal-ganglia projecting neurons with densely spiny dendrites (HVC_X), and large interneurons with smooth dendrites that do not project out of HVC (HVC_{INT}) (Mooney, 2000; Kubota and Taniguchi, 1998; Dutar et al., 1998). Additionally, a small population of <1 % of cells required for temporal song modifications projects to the caudal mesopallium (CM)(Akutagawa and Konishi, 2010; Roberts et al., 2017), but these cells do not form local axon collaterals (Roberts et al., 2017). HVC_{RA} neurons are by far the largest cell population with about 40,000 cells per hemisphere (Wang et al., 2002). Roughly half of these cells are active during singing (Fee et al., 2004), but only with a single, stereotyped burst of action potentials (Hahnloser et al., 2002). Only about 10,000 HVC_X (Nordeen and Nordeen, 1988) and even fewer HVC_{INT} (Wild et al., 2005) neurons have been found. HVC_X neurons exhibit a sparse, stereotyped time code of up to four AP bursts per song motif (Kozhevnikov and Fee, 2007), while HVC_{INT} neurons are almost continuously active during singing (Hahnloser et al., 2002; Long et al., 2010; Kozhevnikov and Fee, 2007). What is known about the wiring of the described neurons inside HVC? This question, and how the wiring might explain their dynamics, will be discussed in the next section.

Models of HVC's sequence generator

HVC's critical role in song production from a macroanatomical perspective (Nottebohm et al., 1976), together with the discovery of the previously introduced (section 1.2.3) ultra-sparse time code of the HVC_{RA} neurons (Hahnloser et al., 2002; Long et al., 2010), led to the development of many models about how the circuit might operate (Gibb et al., 2009b; Cannon et al., 2015; Armstrong and Abarbanel, 2016; Jin et al., 2007). Two main anatomically relevant models of HVC can be distilled from the ideas of the past years, one putting emphasis on the inhibitory connectivity (Fig. 1.4 a), and the other focusing on structured feed-forward excitatory connectivity (Fig. 1.4 b, c) usually related to Moshe Abeles' synchronous firing chain (synfire chain) model (Abeles et al., 1993; Abeles, 1982). The HVC_X neurons have been found to be

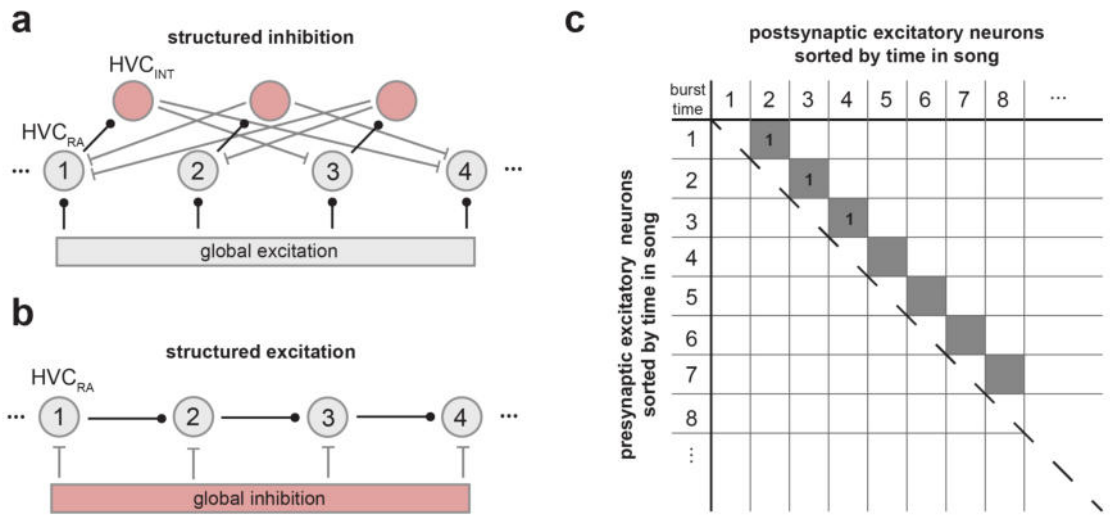

Figure 1.4: *Two extreme views on how the neurons in HVC might be wired for sequence generation. See also Kosche et al. for a more physiological interpretation (Kosche et al., 2015). (a) Structured inhibitory connectivity, with gaps in connectivity, which allows downstream excitatory HVC_{RA} to burst (Gibb et al., 2009b) when driven by unstructured global excitation. Numbers indicate the relative burst times of the HVC_{RA} neurons. Connections from HVC_{RA} neurons onto interneurons "back in time" were left out for clarity. (b) Structured excitatory connections with global inhibition. (c) Excitatory connections of (b), shown in connectivity matrix notation.*

unnecessary for song production in a targeted neuronal death study (Scharff et al., 2000) and are therefore not considered further. It should be noted that there is a substantial amount of controversy about both HVC's role and the involvement of other brain structures in sequence generation (Hamaguchi et al., 2016; Amador et al., 2013; Gibb et al., 2009a) which in part motivates the anatomical investigations of this thesis. These two opposing views are presented for illustrative purposes, but an intermediate model is also plausible in which structured inhibition and excitation cooperate (Kosche et al., 2015). What is the current evidence supporting the different models? Long et al. collected the first data in favor of the structured excitatory connectivity in 2010 by using intracellular microdrive recordings from the HVC of singing zebra finches (Long et al., 2010). The temporally sharp depolarization (about 5 ms) before the action-potential burst, in combination with the observation that cooling slows down singing linearly (Long and Fee, 2008), was interpreted as evidence for converging excitatory synaptic inputs likely emerging from other HVC_{RA} neurons. The feed-forward connectivity pattern has been further supported by comparisons of recurrent vs feed-forward models of HVC and estimates of their compatibility with linear song stretching (Pehlevan et al., 2015). However, no

study has been able to rule out the model that the rapid depolarization is caused by disinhibition, and conflicting and weak evidence for direct connectivity between HVC_{RA} neurons from electrophysiological slice recordings (Kosche et al., 2015; Mooney and Prather, 2005) has spurred doubts. Robust synaptic connectivity between HVC_{RA} and HVC_{INT} has also been demonstrated (Kosche et al., 2015) and supports an intermediate model. One of the objectives of this thesis is therefore to clarify the extent of homotypic HVC_{RA} connectivity and to evaluate whether an excitatory synaptic chain model is still a plausible explanation for the sparse time code of HVC. Interestingly, there is a prominent projection from HVC to the basal-ganglia part of the song system (Area X) that likely makes use of HVC's timing signals for spectral song learning. This feature is the subject of the second topic investigated by this thesis.

1.2.4 Area X, the basal-ganglia element of the song system

Macro-anatomical and functional characterization

Area X is a brain nucleus located in the medial striatum (Wang et al., 2014; Gale and Perkel, 2010) and its volume has been estimated to be about 1.3 mm^3 (Simpson and Vicario, 1991). Two major cortical input areas, HVC and LMAN, project to it, and a single, main efferent projection targets the dorso-lateral medial thalamus (DLM) which projects back to LMAN, making Area X part of a closed topographic cortico-basal-ganglia-thalamo-cortical loop (reviewed in Mooney, 2009; Gale and Perkel, 2010). Lesion studies have shown that it is not required for normal adult zebra finch song production, but is necessary for song learning in juveniles (Sohrabji et al., 1990) and spectral syllable changes in adult birds (Brainard and Doupe, 2000). Of particular significance is one of the aforementioned CAF studies that combined the adult song learning paradigm with transient inactivation of the AFP output (Andalman and Fee, 2009). The pitch of individual harmonic stack syllables was shifted through real time auditory feedback, punishing, for example through unpleasant white noise playback, the syllable renditions with a slight down-pitch variation. Over the course of hours, birds adapted the pitch of the targeted syllable to escape the negative feedback. Inactivation of the AFP output then led to an immediate return to the baseline pitch, illustrating that the motor bias must be

memorized in the AFP (Andalman and Fee, 2009).

The neurons of Area X and their dynamics during singing

Area X was first recognized as a striatal brain area because of its large population of medium-sized neurons with spiny dendrites (MSNs) and the expression of well-known genetic markers (Bottjer and Alexander, 1995). Similar to striatal areas in other vertebrates, the majority of Area X neurons (about 400,000 per hemisphere) are rarely active MSNs (Carrillo and Doupe, 2004; Burek et al., 1991) that fire mostly with sparse bursts during singing (Goldberg and Fee, 2010). Additionally, three different types of interneurons have been identified so far: fast-spiking, parvalbumin-positive neurons (FS); low-threshold-spiking, somatostatin-positive neurons (LTS); and tonically active, cholinergic neurons (ACh) (Goldberg and Fee, 2010; Perkel et al., 2002). Interestingly, Area X also contains about 3,000 cells typically found in other vertebrates in the pallidum (Luo and Perkel, 1999; Farries et al., 2005; Person et al., 2008; Carrillo and Doupe, 2004), large neurons resembling the pallidal cells from the direct and indirect basal-ganglia pathway (Carrillo and Doupe, 2004) that fire at high frequencies (Goldberg et al., 2010). Furthermore, it was recently found that Area X also contains rare glutamatergic neurons, which could be a functional analog to the neurons in the mammalian subthalamic nucleus (STN) (Budzillo et al., 2017). In mammals, the STN is part of the hyper-direct pathway that enables fast cortical control of the GP neurons (Jahanshahi et al., 2015).
The songbird Area X can therefore be considered an amalgamation of the highly conserved vertebrate basal-ganglia structures (Grillner and Robertson, 2016) specialized in its function on spectral song syllable optimization (Ali et al., 2013). As is the case for the mammalian basal-ganglia (Graybiel, 2008), Area X seems to implement reinforcement learning (RL) (Fee and Goldberg, 2011), i.e. learning that is based on trial-and-error and feedback evaluation.

A reinforcement learning model of Area X

In 2011, Fee and Goldberg proposed a novel RL-model for Area X that distinguishes precisely the different roles of efference copy and context signals for learning (Fee and Goldberg, 2011;

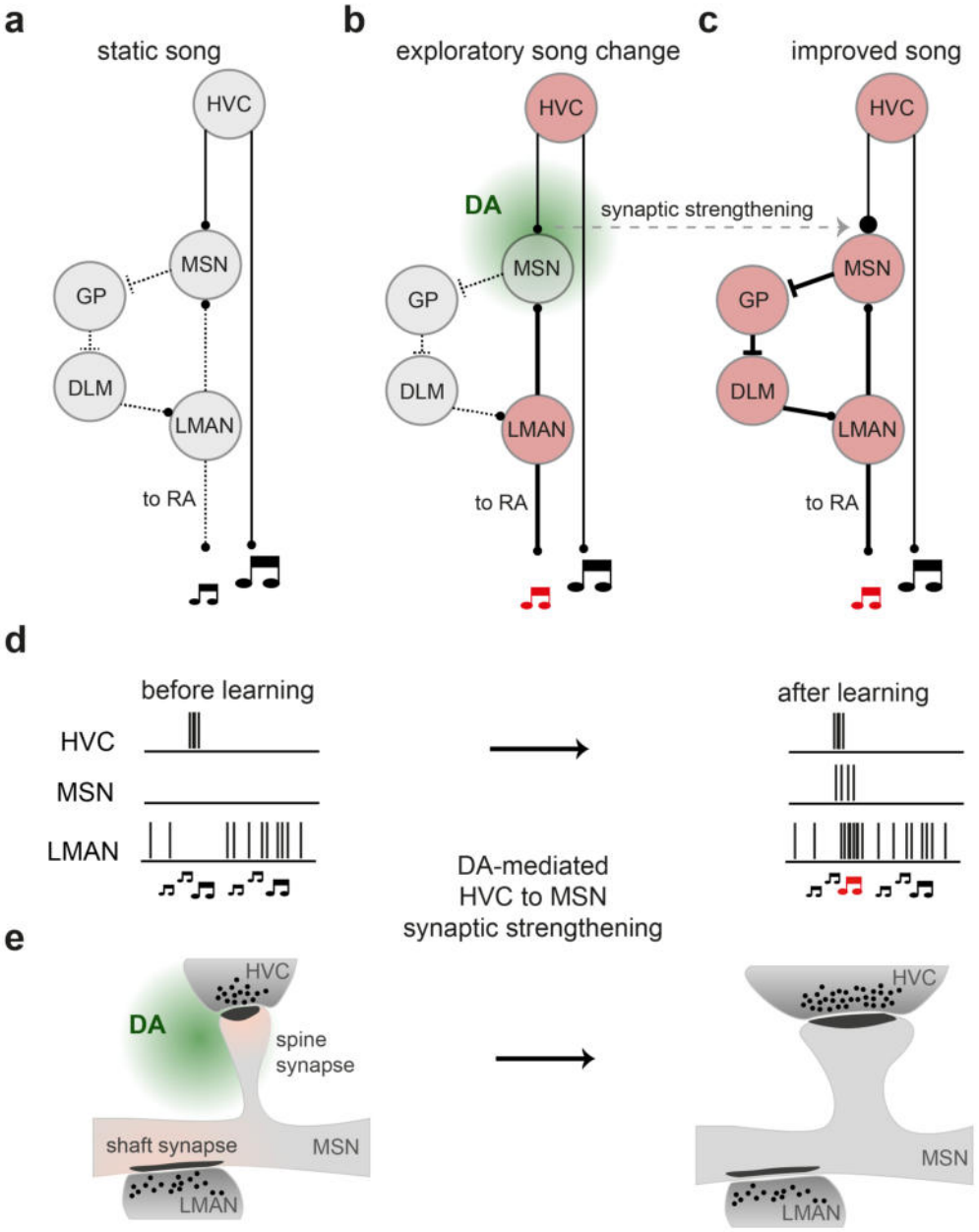

Figure 1.5: *Illustration of the RL model of Area X as proposed by Goldberg and Fee (Fee and Goldberg, 2011). (a) During the female-directed crystallized adult song (baseline), there are no changes made to the spectral syllable composition. (b) During undirected (practice) song, zebra finches inject more variability into their song by means of seemingly randomly bursting LMAN neurons that project to nucleus RA, which also receives the stereotyped time-control signal from HVC. If the injected variability leads to the release of DA in Area X (corresponding to positive song changes, the valence of which is determined in the auditory system), the HVC synapses on MSNs that were previously made eligible through LMAN activity are strengthened. (c) The strengthened HVC to MSN synapses now increase the spiking probability of the postsynaptic MSN to reproduce the previously random change. (d) Schematic neuronal activity before and after learning. (e) Proposed anatomical implementation of the model at the synaptic level.*

Fee, 2014) down to the level of individual synapses impinging on MSNs, which are assumed to be the central information store. In this model, the glutamatergic synapses formed by Area X-projecting HVC neurons are strengthened with a dopamine (DA) dependent plasticity rule (Fig. 1.5) after coincidence detection and DA release (Ding and Perkel, 2004). The basal-ganglia-thalamo-cortical loop formed between Area X, DLM, and LMAN acts as the RL-experimenter that injects variability into the song (by biasing the neuronal activity of the principal neurons inside of RA) (Kao et al., 2005) and in parallel correlates this exploratory activity with both the song context provided by HVC and DA release (Gadagkar et al., 2016). In the case of positive song changes (DA release in Area X), synaptic strengthening occurs and the song will be biased during future renditions in the direction of the previous random exploration. This seems to be implemented with a closed topographic loop (Luo et al., 2001), which would ensure that the directed bias activates the same syringeal muscles (over RA) as the previously exploratory activity. Since the neurons in LMAN transmit the bias to RA, and also the efference copy to Area X, their synapses on the MSNs are unsuitable for synaptic plasticity according to the model: Synaptic strengthening of LMAN's efference copy synapses would lead to uncontrolled positive feedback of neuronal activity in the loop. This is in contrast to HVC's context synapses, which are not part of the loop, and therefore candidates for synaptic strengthening. What does this distinction between efference copy (LMAN) and context (HVC) synapses mean for the individual synapses? Repeated activation of HVC-MSN synapses, made eligible for plasticity by the activity of LMAN-MSN synapses (through depolarization and removal of a Mg-NMDA receptor block) on the same MSN dendrite followed by DA release, would strengthen these synapses. Even without activation of LMAN, the context synapses might then be able to drive the MSNs, with the consequence that they would bias the song as described before. While highly speculative, specific testable anatomical predictions can be inferred from the model. For instance, in order to provide an eligibility trace to many spine synapses that listen to diverse context signals from different HVC axons, context synapses (emerging from HVC_X axons) should target the spines of MSNs, while efference copy synapses (emerging from $LMAN_X$ axons) should target the dendritic shafts of MSNs.

Chapter 2

Objectives

2.1 Anatomical hypotheses tested

2.1.1 HVC_{RA} neurons are predicted to be connected, forming a synaptic chain

Evidence that the ultra-sparse time code of HVC_{RA} neurons (Hahnloser et al., 2002) is the result of direct synaptic connections between these neurons is so far mainly based on indirect observations:

1. HVC_{RA} bursts are accompanied by a sharp depolarization instead of a slow ramp-up (Long et al., 2010).
2. Cooling of HVC slows down song production almost linearly, indicating that the circuit is contained in HVC (Long and Fee, 2008).
3. Feed-forward models of the HVC circuitry are a better explanation for the linear song stretching than recurrent models (Pehlevan et al., 2015).

While these observations support the synaptic chain model of the sequence generating circuit in HVC, there exists no anatomical data confirming these connections. On the contrary, studies

based on pairwise electrophysiology report conflicting connection probabilities (Kosche et al., 2015; Mooney and Prather, 2005). One goal of this thesis is therefore to analyze the homotypic HVC_{RA} neuron connectivity anatomically, using volume EM based on SBEM.

2.1.2 HVC and LMAN axons are predicted to synapse differently onto MSNs in Area X

As illustrated in section 1.2.4, the postulated RL-model of the basal-ganglia by Fee and Goldberg predicts different functional roles for Area X-projecting axons originating from LMAN (efference copy) and HVC (context). The second objective of this thesis is to test whether these postulated functional differences are reflected in the anatomy, and more precisely, whether LMAN axons form synapses with dendritic shafts of MSNs and HVC axons target the dendritic spines of MSNs.

2.2 Development of novel approaches for faster data analysis

As laid out in section 1.1.3, faster data analysis through automation is a prerequisite for the connectomic analysis of larger volume EM datasets of brain tissue. An objective of this thesis was to advance automation and reduce human annotation times in the two most problematic areas of dataset analysis, as shown before:

- Automated neurite reconstruction, i.e. the recovery of neuronal processes from a volume EM dataset.
- Automated identification of ultrastructural features, including the location and type of synapses, to reduce the manual time required for the construction of connectivity matrices further.

Chapter 3

The synaptic anatomy of the sequence generating circuit in HVC

Publication of the results presented in this chapter

This chapter contains text and figures from the following peer-reviewed publication:

> **EM connectomics reveals axonal target variation in a sequence-generating network.**
> Kornfeld, J.*, Benezra, S. E.*, Narayanan, R. T., Svara, F., Egger, R., Oberlaender, M., Denk, W., Long, M.** eLife, 6:e24364 (2017)
> * co-first authors; ** corresponding author

All experiments were performed by J. Kornfeld in the laboratory of W. Denk (MPI for medical Research, Heidelberg; MPI of Neurobiology, Martinsried) unless explicitly stated otherwise. The project was performed in collaboration with the laboratory of M. Long (NYU, NYC).

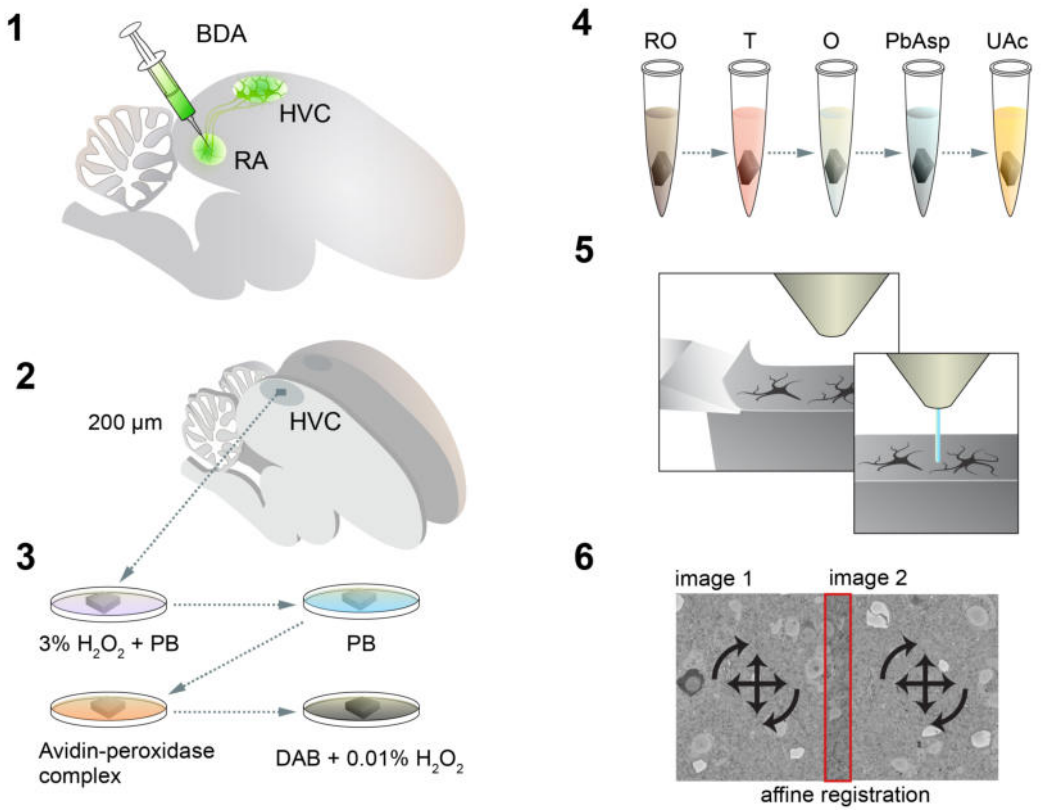

Figure 3.1: *Steps performed to obtain a SBEM dataset with retrogradely labeled HVC_{RA} neurons. After BDA injection (1), the animal was perfused (Cragg, 1980), the brain was extracted and a single 200* µm *thick section containing HVC was used for further processing (2). Standard HRP-DAB staining was then applied to the brain slice (3), and a subsequent post-staining was applied using the ROTO-UAc-PbAsp method for EM (4). SBEM was then performed (5), and the resulting individual image tiles registered through affine transformations (6). Figure adapted from (Kornfeld et al., 2017), originally created by J. Kuhl.*

3.1 Methods

3.1.1 Sample preparation and data acquisition

Retrograde labeling and EM staining

All steps up to the EM staining procedure were performed as part of a collaboration in the laboratory of Michael Long, NYU, New York City, by S. Benezra and G. Kosche. Figure 3.1 illustrates the entire sample preparation process. 200 nL of 3 kDa biotinylated dextrane amine (BDA) was injected with a glass pipette of 30-40 µm tip diameter (Nanoject, Drummond Scientific, USA) into the left-hemisphere RA of a male zebra finch (adult, >90 dph) under

isoflurane anesthesia (1-3 %). A carbostar (Carbostar-1, Kation Scientific, USA) electrode was pointed at 2.3 mm lateral and 1.85 mm posterior to the midsagittal sinus, and the RA-characteristic electrophysiological activity (Long and Fee, 2008) was used as an additional criterion to identify the injection site. The bird was then perfused with an extracellular space (ECS) preserving solution (0.07 M sodium cacodylate (Serva, Germany), 0.14 M sucrose (Sigma-Aldrich, USA), 0.002 M $CaCl_2$ (Sigma-Aldrich) with 2 % paraformaldehyde (PFA) and 2 % glutaraldehyde (GA) (Serva) under high pressure to open the blood-brain barrier (Cragg, 1980). A 200 µm slice of brain tissue containing HVC was cut from the extracted brain and post-fixed in the same solution overnight. Several rinsing steps in 0.15 M cacodylate buffer (CB) were performed and the sample was permeabilized in a 30 % sucrose solution (for cryo-protection) in liquid nitrogen by freeze-thawing (single cycle). Intrinsic peroxidase-activity was inactivated with a 3 % H_2O_2 solution. The sample was then immersed overnight at 4°C in an avidin-peroxidase complex solution to tag it with the peroxidase complex (Vector Labs and Sigma-Aldrich). The slice was then stained in a solution with 2.3 mM diaminobenzidine (DAB, Sigma-Aldrich) with 0.01 % H_2O_2. After several CB rinsing steps, a ROTO stain was applied (Briggman et al., 2011), as described in the following. In between each staining step, the sample was rinsed multiple times, either with CB (after the first Osmium step) or with double-distilled water (all other steps). Staining was started in 2 % OsO_4 (Serva) reduced with 2.5 % Potassium hexacyanoferrate(II) (Sigma-Aldrich)(2 h), followed by immersion in 1 % Thiocarbohydrazide @58°C (Sigma-Aldrich)(1 h), followed by 2 % OsO_4 (2 h), and 1.5 % Uranyl acetate in H_2O @53°C (Serva)(2 h), and 0.02 M Lead aspartate @53°C (Sigma-Aldrich)(2h)(Karnovsky, 1971; Seligman et al., 1966; Walton, 1979). Water was then removed with an ethanol (Electron Microscopy Sciences, USA) series (chilled, 10 min, 15 min, 10 min, 10 min, at 70 %, 100 %, 100 %, 100 %). Propylene oxide (Sigma-Aldrich) was used for epoxy-infiltration before the sample was embedded in epon hard (Serva) and cured at 60°C for 48h (Glauert and Lewis, 1999). The sample was trimmed, glued to an aluminum holder (custom made) and re-embedded into epoxy. Before SBEM, the sample was trimmed again, this time to a pyramidal-shape, and then gold coated for better conductivity. Finally, the surface was smoothed with an ultramicrotome (Leica, Germany).

SBEM acquisition parameters and image registration

A Zeiss UltraPlus field-emission cathode SEM (GEMINI column) equipped with an electrostatic beam blanker and scan generator (Fibics Inc., Canada) was used for SBEM. The microscope was outfitted with a custom in-chamber microtome (Denk and Horstmann, 2004) and back scatter electrons were detected with a custom amplification system (J. Tritthardt) in high-vacuum. Electron landing energy was set to 2 kV, the image scan rate was 5 MHz (effective imaging rate 2.1 MHz), and the electron dose 10.3 electrons per nm^2. Four overlapping image tiles (4400 pixels side length) were acquired per slice at a resolution of 11x11 nm, with a cutting thickness of 29 nm. Global affine transformations were calculated for each image tile for a re-assembly into a continuous volume (Scheffer et al., 2013)[1], and these transformations were applied using custom Python code. The transformed images were then converted to a KNOSSOS readable dataset by cubing into 128x128x128 large raw data blocks using the KNOSSOS cuber[2].

3.1.2 Data analysis

BDA labeling efficiency estimation

The labeling efficiency for HVC_{RA} somata was determined by counting the number of darkly stained cell bodies in the dataset, and dividing it by the number of expected HVC_{RA} cell bodies for the volume of the EM dataset using published cell counts (Wang et al., 2002).

Since different labeling efficiencies for axonal collaterals and somata were suspected, an independent analysis was performed to estimate the labeling efficiency for axons. Three hundred randomly distributed 1 μm^3 boxes were selected, all labeled axons within each were traced, and the path length was compared to the expected HVC_{RA} axon path length for such a volume. This expected path length was calculated by multiplying the total number of HVC_{RA} neurons in HVC with the LM-measured axon length, divided by the total HVC volume.

[1] `https://github.com/billkarsh/Alignment_Projects`
[2] `https://github.com/knossos-project/knossos_python_tools/tree/master/knossos_cuber`

Neuron tracings

Raw image data analysis was performed with KNOSSOS[3] (Kornfeld et al., 2011; Helmstaedter et al., 2011) by either annotators who received at least 10 h of training, or experts who had at least one year experience. All neurites were traced as skeletons, except the ones used for illustrative purposes (see Fig. 3.6 and 3.7), which were volume annotated. The twelve BDA labeled HVC_{RA} neurons were identified at their soma inside the EM volume and traced and proofread in a review step by an expert annotator. Additionally, six BDA labeled axons were traced from seed points (three seeds in myelinated parts, three seeds in unmyelinated parts) and reconstructed with at least double redundancy by annotators. An expert resolved then reconstruction discrepancies manually in a proofreading step, thereby creating a consensus annotation. All synapses were identified on the nine axons by an expert, and an attempt was made to reconstruct a stretch of the postsynaptic partner dendrite for each (regular annotators, 2x redundancy). Dendrites were then proofread by an expert, and dendrites were excluded when only a short stretch could be reconstructed due to dataset artifacts or the dataset boundaries (at least 10 µm required). To identify postsynaptic dendritic reconstructions that belonged to the same cell but were independently created because of the presence of multiple synapses between an axon and its partner, a spatial matching procedure was developed (based on an efficient k-dimensional spatial tree look-up data structure[4]). If two skeleton nodes of different dendritic tracings were closer than 400 nm, they were considered overlapping, and if more than 25 % of nodes overlapped, independent dendritic tracings were assumed to belong to the same neuron.

Synapse annotations

Synapses were annotated by labeling the synaptic cleft in the principal viewing orientation in which the synaptic contact was largest. Synaptic contact areas were then calculated by taking the path length of the cleft annotation as the diameter of a circle. All incoming den-

[3] https://knossostool.org/

[4] `https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.spatial.KDTree.html`

dritic synapses were classified by an expert as symmetric (inhibitory) or asymmetric (excitatory)(Colonnier, 1968; Gray, 1959). If the type of a synapse could not be inferred from a single synapse, the axon was followed until a decision could be made. Synapse-type annotations were validated in two ways. First, twenty synapses randomly sampled from the dendrite of an HVC_{RA} neuron were classified independently by two experts. Second, fifty synapses were sampled from the axons of two HVC_{INT}, one HVC_{RA} and two HVC_X neurons. These cells were identified by their somatic and dendritic appearance. The fifty synapses were then classified in a blind test by an expert annotator.

Neuronal cell type identification by spine density

Neuronal cell types were identified from the traced dendritic stretches (see section 3.1.2) by analysis of their spine density, for which ground truth data was established from LM reconstructions (LM experiments performed by S. Benezra). The spine density was estimated through automatic spine counting (criteria: >1 µm length, spines with >1 incoming synapses from an average of three synapse counts by independent annotators were removed). The synapse counting was necessary to avoid the classification of dendritic interneuron protrusions as spines (see Fig. 3.8 d). The neuron was then assigned to one of the three main classes of HVC by the spine density: HVC_{INT} for less than 0.11 spines per µm path length, HVC_{RA} for values between 0.11 and 0.46, and HVC_X in the case of a higher spine density. The average spine density was used for classification in cases where multiple counts were made for dendrites of the same cell (see section 3.1.2) that were traced from different synapses.

HVC_{RA} axon soma-distance estimation

Since not all labeled HVC_{RA} axons could be traced back to a soma due to the limitations of the EM dataset, a Bayesian method was developed that allowed the estimation of a soma distance for each orphaned axon. For this estimate, the variation of the number of axonal branch points as a function of soma distance ($D_b(r)$) was exploited. The probability distribution over r was calculated using

$$P(r|N,l) \propto P(N,l|r)P_a(r) \tag{3.1}$$

with l being the length of an axonal branch, N the number of branch nodes, and $P_a(r)$ the Bayesian prior. $P_a(r)$ was determined from LM reconstructions (LM experiments performed by S. Benezra) using a Gaussian kernel density estimation (Python scipy.stats.gaussian_kde with the Scott bandwidth selector[5]) and $P(N,l|r)$ then calculated by assuming a Poisson distribution of branch nodes, with a node-count expectation value $\lambda = D_b(r)l$:

$$P(N,l|r) = \frac{D_b(r)l^N e^{-D_b(r)l}}{N!} \tag{3.2}$$

whereby $D_b(r)$ was determined with an exponential fit to LM measurements:

$$D_b(r) = \frac{35.448 \cdot e^{-\frac{r}{43.5\,mm}} + 0.613}{mm} \tag{3.3}$$

Estimation of the fraction of homotypic HVC_{RA} synapses

Two separate approaches were used to estimate the fraction of homotypic synapses.
First, a dendritic perspective was taken, and the number of labeled incoming HVC_{RA} synapses was counted (identified by an expert, scrutinized by another expert) and corrected with the estimated axonal labeling efficiency (see section 3.1.2). Second, an axonic perspective was taken by analyzing the postsynaptic partner type of outgoing synapses of HVC_{RA} axons. For this, the 59 BDA-labeled axons identified for the estimation of the axon labeling efficiency (see section 3.1.2), were traced to the next two synapses (or the end of the dataset), and the post synaptic partner cell type was determined as described in section 3.1.2. The outgoing synapse density per type was then computed from these data, and an estimate of the total number of outgoing homotypic HVC_{RA} synapses calculated by extrapolation to the entire HVC-internal HVC_{RA} axon length, as measured from LM (LM experiments performed by S. Benezra).

[5] `https://docs.scipy.org/doc/scipy-0.19.0/reference/generated/scipy.stats.gaussian_kde.html`

3.2 Results

3.2.1 The j0256 dataset of HVC

To test the hypothesis that HVC_{RA} neurons are synaptically connected anatomically (see objectives section 2.1) as required by the synfire chain model, a SBEM dataset of HVC was acquired. Since small volume EM datasets suffer often from the difficulty of identifying reconstructed neu-

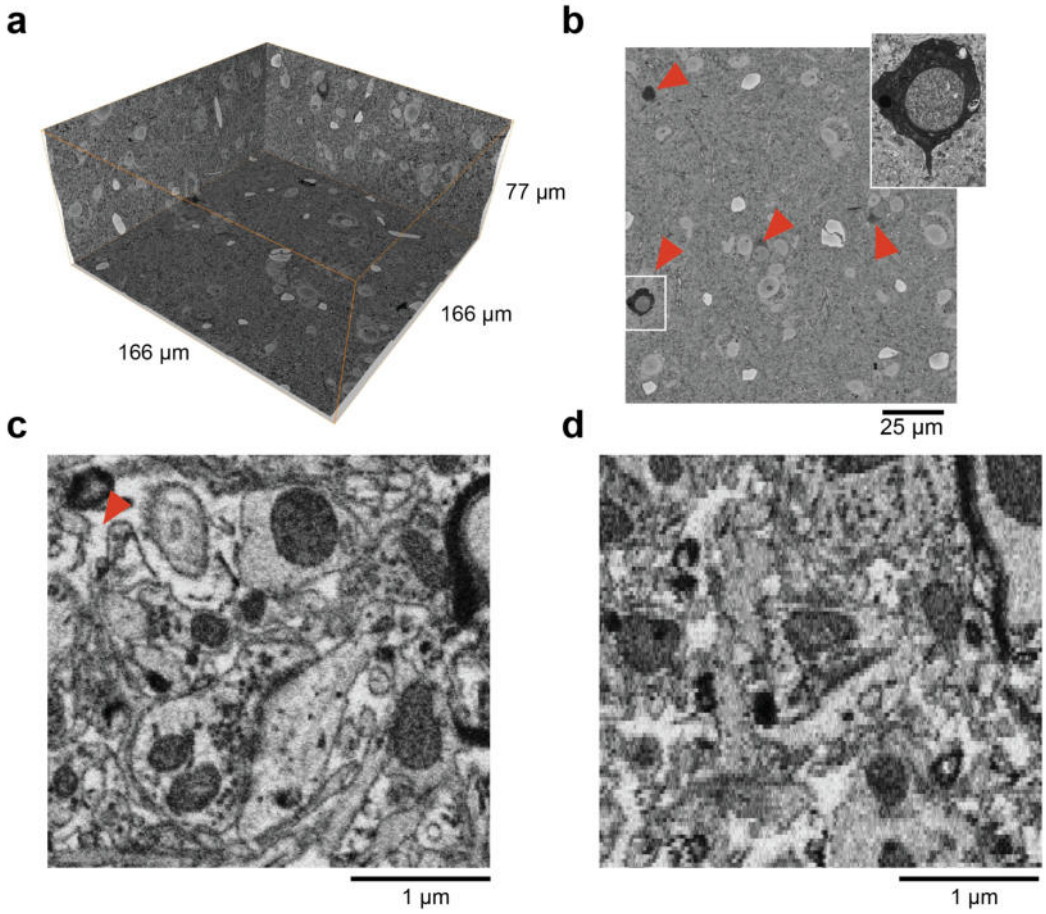

Figure 3.2: *Image quality and dimensions of the acquired j0256 dataset. (a) 3D-rendering indicating dataset size. (b) Overview 2D xy-slice, showing the darkly labeled HVC_{RA} neurons. Red arrows show the BDA-labeled somata visible in this plane (see also inset). (c) xy-slice (EM imaging plane) at full resolution. The red arrow indicates the white ECS, often not visible in conventionally stained EM micrographs. (d) xz-reslice at full resolution. Note the slight jitter from top to bottom, in comparison with c. Figure adapted from (Kornfeld et al., 2017).*

ron fragments and assigning them to previously described cell types, a novel *en-bloc* staining method was first developed that combines the ROTO protocol (reduced Osmium tetroxide, RO; Thiocarbohydrazide, T; Osmium tetroxide, O; followed by Uranyl acetate, UAc; and Lead

aspartate, PbAsp; see section 3.1.1) with retrograde labeling. A retrograde tracer (3 kDa BDA, Reiner et al., 2000) was injected into the brain nucleus RA and was allowed to diffuse to HVC and label a sub-population of HVC_{RA} neurons. An SBEM dataset of 166 x 166 x 77 µm of a central region of HVC was then acquired at a voxel size of 11 x 11 x 29 nm. Darkly stained cells were visible in this dataset, as shown in Fig. 3.2 (for detailed acquisition parameters see section 3.1.1). The dataset contained 563 somata, of which 449 were neuronal and 114 glial cell bodies. 34 of the neuronal cell bodies were labeled with BDA and therefore were confirmed as HVC_{RA} neurons. Based on the morphology of the cell bodies and dendrites (see section 3.1.2), 259 cells were classified as putative HVC_{RA} neurons. In addition, 22 neurons were suspected to be HVC_X cells and 61 were suspected to be HVC_{INT} cells. A large population (n = 73) of neuronal somata could not be further classified because they were located at the boundary of the dataset, making it impossible to analyze their processes, or because of an inconclusive appearance of their soma.

3.2.2 Synaptic afferents of HVC_{RA} neurons

To analyze the synaptic afferents of HVC_{RA} neurons, twelve labeled cells were reconstructed by tracing their processes (skeletonizing) from their somata within the limits of the dataset (total path length: 22.7 mm, mean 1.89 mm, s.d. 0.625 mm), followed by an identification of their incoming synapses. For one of the twelve neurons, all incoming synapses were annotated (Fig. 3.3 a, total of 1,003 synapses) and classified as being asymmetric (excitatory) or symmetric (inhibitory)(Colonnier, 1968; Gray, 1959). 396 asymmetric synapses were found (39.5 %), and 607 symmetric synapses (60.5 %). The synapse type classification was validated in two ways: first, with an independent analysis of twenty randomly selected synapses from the dendrites of an HVC_{RA} neuron by two experts (19 of 20 were assigned the same type), and second, through the blind classification of 50 synapses sampled from axons (8 HVC_{RA}, 11 HVC_X and 31 HVC_{INT}) with known synaptic identity (asymmetric, n = 19; symmetric, n = 31; all correctly classified) (see Fig. 3.3 b for example synapses).

The incoming synapse distribution of the single cell was then compared to the local HVC_{RA}

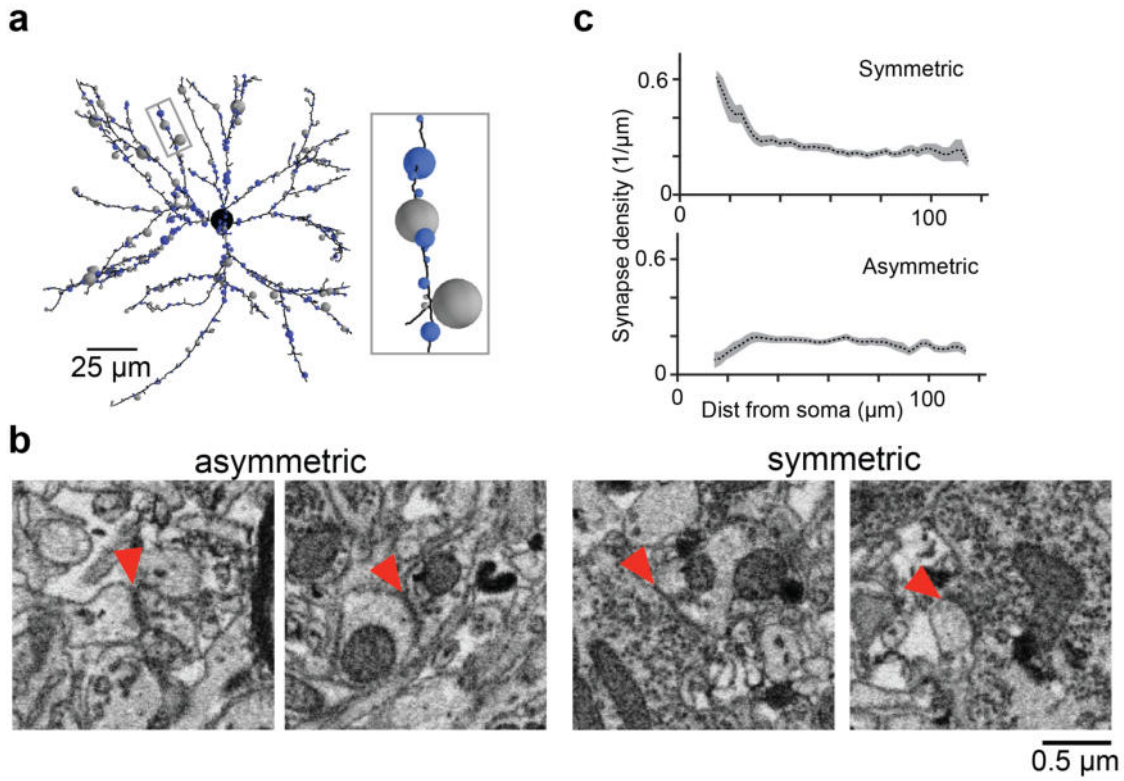

Figure 3.3: *Analysis of the synaptic inputs of HVC_{RA} neurons. (a) Annotated and classified dendritic synapses of a BDA-labeled HVC_{RA} neuron, fully reconstructed in the extent of the j0256 dataset. Found putative excitatory synapses are in gray, inhibitory synapses in blue. (b) Raw EM images illustrating the ultrastructural differences between the asymmetric and symmetric synapse type in this dataset, compare especially the number of synaptic vesicles and synaptic densities. Red arrows: synaptic clefts. (c) Synapse type vs soma distance, gray shading indicates s.e.m. Figure adapted from (Kornfeld et al., 2017).*

neuron population by inspecting synapses from the dendrites of eight additional neurons, and no significant difference could be found ($p = 0.36$, one-way ANOVA). To reduce the annotation workload, short dendritic stretches were randomly sampled from these neurons ($n = 97$, first and third quartile of stretch length 13.3 µm and 21.6 µm) and densely annotated as before. This high amount of inhibitory input synapses (60.5 %) is surprising, given that the fraction of inhibitory synapses that can be found on mammalian cortical neurons is usually less than about 20 % (Beaulieu et al., 1992; Kasthuri et al., 2015; Peters, 2002). Similar to mammalian principal neurons, inhibitory inputs were enriched close to the soma of the HVC_{RA} neurons (Anderson et al., 1994) (Fig. 3.3 c) (68 ± 4 % for less than 40 µm soma distance vs 57 ± 2 % for more than 40 µm; $p < 0.05$, Wilcoxon rank-sum test). While this analysis provides a rough categorization of the synaptics inputs, it could not reveal the exact cell type behind each input,

an important piece of information for the understanding of circuit function.

However, the introduced BDA label allows for the identification of excitatory afferents that originate with high probability from other HVC_{RA} neurons (but see section 3.2.5 for a critical discussion and a control experiment), a prerequisite for the synaptic chain model to generate the temporally precise timing signals for singing. Double-labeled (i.e. pre- and post-synaptically

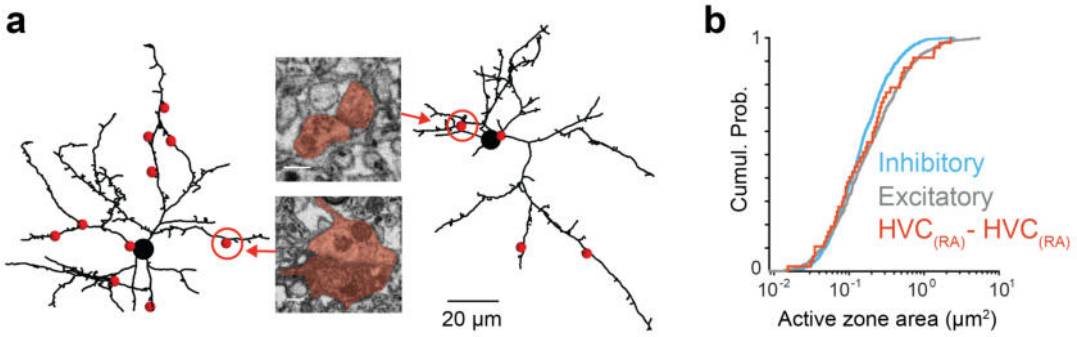

Figure 3.4: *Incoming homotypic synapses on HVC_{RA} neurons. (a) Identified darkly labeled synapses on two of the twelve reconstructed HVC_{RA} dendrites in red. The insets show a small (top) and large (bottom) homotypic synapse. Inset scale bars: 0.25* µm *(b) Cumulative distributions of synapse sizes for the three different synapse types that could be discriminated. Note that the symmetric synapses had fewer large synapses compared to the excitatory synapses and the homotypic synapses. Figure adapted from (Kornfeld et al., 2017).*

stained) homotypic synapses between HVC_{RA} neurons were therefore searched on all twelve HVC_{RA} neuron reconstructions (see Fig. 3.4 a for two examples), and, in total, 44 such synapses were identified on the dendrites, out of an estimated total of 3,817 ± 925 (s.d.) excitatory synapses (calculated from an estimated excitatory synapse density of 0.168 ± 0.014 μm^{-1} s.e.m. and a total searched dendritic path length of about 22.7 mm). The synapse sizes of all types (Fig. 3.4 b) followed a skewed distribution (many small synapses and few large synapses, log-normal like), and the double-labeled synapses were not different in their size distribution from all identified excitatory synapses (median size double-labeled: 0.21 μm^2, asymmetric: 0.17 μm^2, $p > 0.05$, Wilcoxon rank-sum test). The inhibitory synapse distribution appeared to have fewer large synapses, but this remains an anecdotal observation that should be investigated in more detail in future studies. To calculate the overall fraction of excitatory synapses that stem from other HVC_{RA} neurons, a reliable estimate of the labeling efficiency of the retrograde tracer was required. The somatic labeling efficiency was 14 %, as calculated from

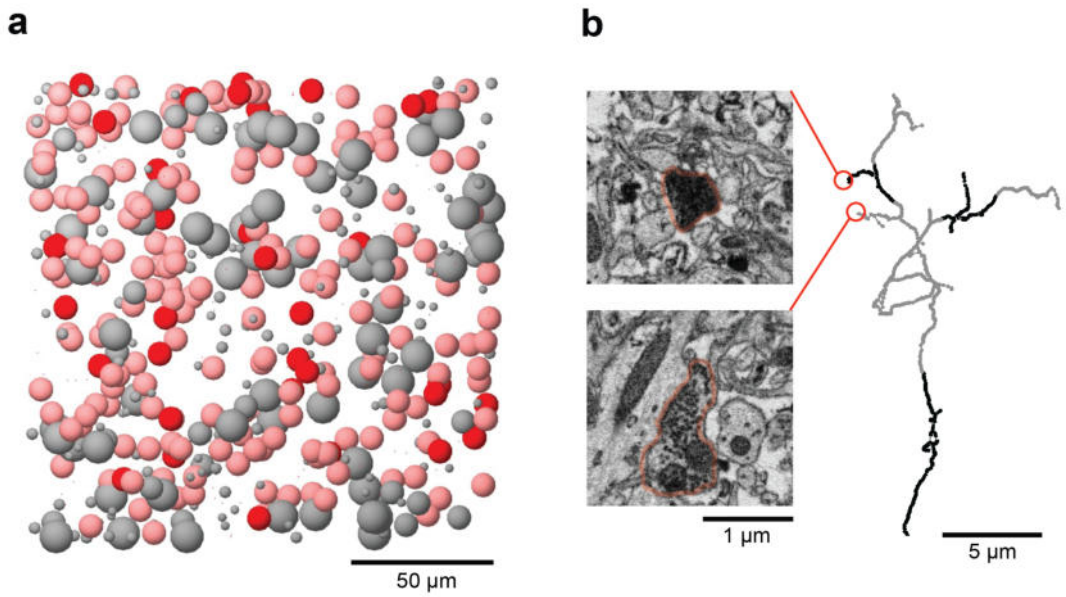

Figure 3.5: *Labeling of the retrograde tracer. (a) Somatic labeling efficiency of the retrograde tracer. Top view of the j0256 dataset, red spheres indicate a labeled neuron, therefore an HVC_{RA} projector, light red a morphologically identifed HVC_{RA} neuron and gray other neurons, HVC_{INT} and HVC_X. Small gray spheres indicate glia. (b) Example of an incompletely labeled axon collateral. The black regions of the reconstruction indicate strong label, the light regions almost invisible label, as illustrated in the insets. Figure adapted from (Kornfeld et al., 2017).*

the number of somata in the dataset that had recognizable staining compared to the number of expected HVC_{RA} neurons (expected value of labeled somata with perfect labeling efficiency for a dataset of this size: 240 ± 28, s.e.m.; based on a total of about 40,000 ± 3,800 (s.e.m.) HVC_{RA} neurons (Wang et al., 2002), and a total HVC volume of 0.35 ± 0.024 mm^3, s.e.m., n = 14, LM measurements performed by S. Benezra) (see Fig. 3.5).

This analysis indicates that only about 8 % of all excitatory synapses could be a part of a synaptic chain for sequence generation. However, the labeling of axonal collaterals appeared to be much less efficient than the somatic and dendritic labeling, as illustrated in Fig. 3.5 b. This observation is not surprising, given that 3 kDa BDA is well known for its retrograde labeling efficiency but not for its strong axonal collateral and terminal labeling (Reiner et al., 2000). The axonal labeling efficiency was therefore quantified independently with a sparse sampling approach to reduce annotation time. 300 cubes of 1 μm^3 size were randomly sampled from the dataset and the total labeled axon path length contained was measured (38.6 µm). The labeling

efficiency could then be calculated from an estimate of the total HVC_{RA} axon path length of the nucleus (585.6 m) and its volume (about 0.35 mm^3, see above), leading to a number of 7.6 ± 1.6 % (s.e.m.), about half the somatic labeling efficiency (14 %). Based on this estimate, about 15 ± 4 % (s.e.m.) of all excitatory synapses of an HVC_{RA} neuron are homotypic and might support a synaptic chain.

To have an independent mean for estimating the amount of homotypic connectivity, a transsynaptic tracing approach was taken and the synaptic targets of labeled (i.e. putative HVC_{RA}) axonal collaterals were analyzed. However, a method to reliably identify the type of an HVC neuron based on its dendritic appearance first had to be established. This identification was necessary because only a small fraction of the postsynaptic dendrites were expected to be labeled by BDA, which would have allowed a direct identification.

3.2.3 Identifying HVC cell types from short dendritic reconstructions

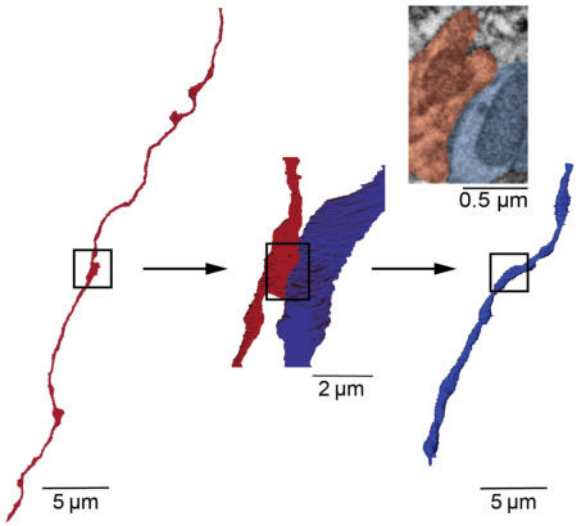

Figure 3.6: *Transsynaptic tracing approach used to analyze the synaptic targets of HVC_{RA} axons. First, all synapses are identified on an axon (left, red), and then all postsynaptic dendrites are traced, either partially or completely, as illustrated for one synapse formed with an aspiny HVC_{INT} dendrite. Figure adapted from (Kornfeld et al., 2017).*

The three main cell types of HVC (HVC_{RA}, HVC_{INT} and HVC_X) can be distinguished by their

dendritic spine density, as reported previously in several studies (Kubota and Taniguchi, 1998; Dutar et al., 1998; Mooney, 2000). To confirm these results, the spine densities of dendritic branches of known cell types were quantified from dendritic LM reconstructions (LM experiments by S. Benzra, see Fig. 3.7 a). As expected, HVC_{INT}, HVC_{RA}, and HVC_X had largely non-overlapping spine density distributions (Fig. 3.7 c). Interneuron dendrites were largely aspinous (0.01 ± 0.01 spines per µm, mean ± s.d.), HVC_X dendrites spiny (0.70 ± 0.13 spines per µm) and HVC_{RA} dendrites in between (0.21 ± 0.07 spines per µm). A simple classification

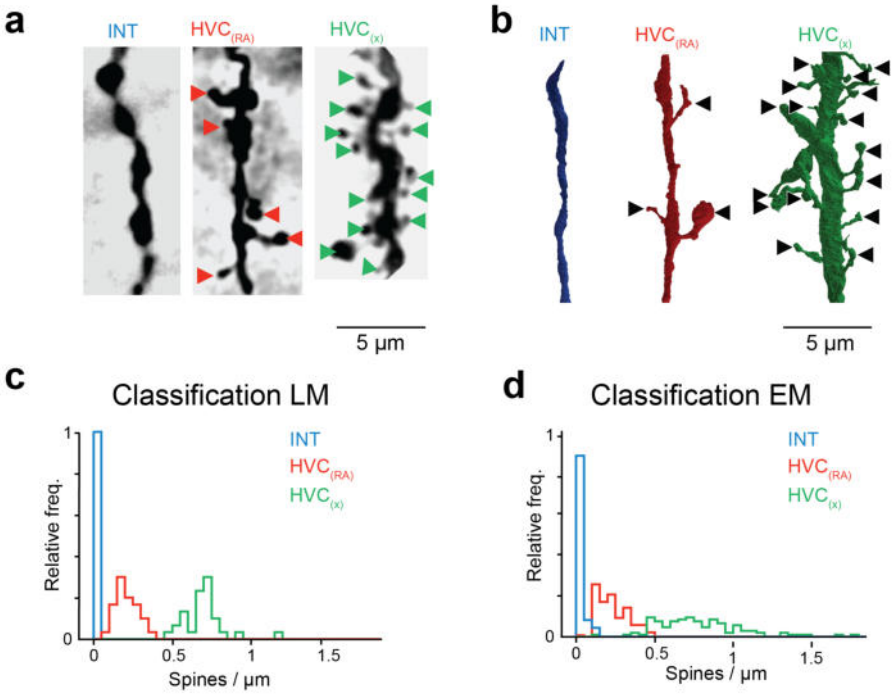

Figure 3.7: *Spine density based classification of short dendritic reconstructions to identify the postsynaptic targets of HVC_{RA} axons. (a) LM micrographs (S. Benezra) showing dendrites from the three main classes. (b) Volume reconstructions of three EM dendrites, colored by the classification method. (c) and (d) show the frequency of measured spine density data for LM ground truth data (c) and unknown EM data (d). Figure adapted from (Kornfeld et al., 2017).*

scheme was then used to identify the EM dendrites (Fig. 3.7 b, d), based on thresholding the spine density (see section 3.1.2). Interestingly, the EM spine density distributions were broader, possibly because of the larger amount of different cells that were incorporated into the analysis (n = 528 for EM dendritic stretches, most likely from different cells, and n = 9 cells for LM dendrites) and the presence of shorter dendritic stretches. The EM analysis revealed further that some interneurons (those that could be clearly identified based on their outgoing synapse type and somatic ultrastructure, see Fig. 3.8 a to c) appeared to have dendritic spines

as well, with one case shown in Fig. 3.8 d. While these spines were similar geometrically, the ultrastructure revealed that they received multiple synaptic contacts at their spine head. This was in contrast to the excitatory cell types, which mainly received just a single synapse. The synapses on all spines identified on the EM reconstructions were therefore counted (three independent annotators), and spines that had more than one synapse on average were excluded for the cell type classification. The EM spine density classification was validated by the correct

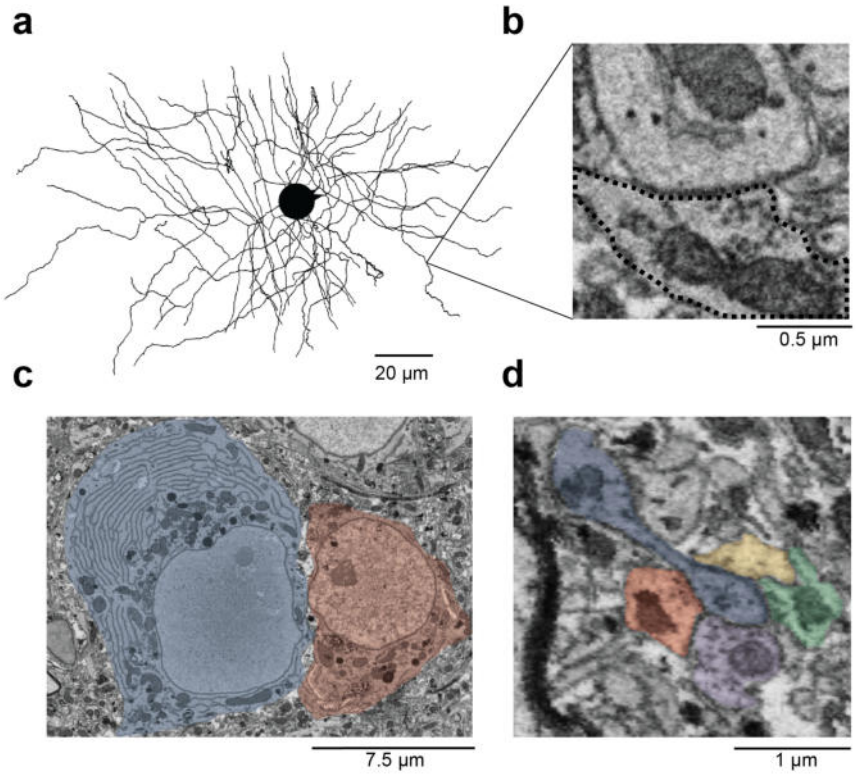

Figure 3.8: *Morphological identification of interneurons in EM. (a) Skeleton reconstruction of a largely aspiny HVC_{INT}, soma approximated with a sphere. (b) Imaging plane EM cut-out of a symmetric synapse formed by the neuron. (c) Different appearance of the soma of an HVC_{INT} on the left side (blue shading) and an HVC_{RA} neuron, right side (red shading). (d) Short dendritic process of an interneuron (blue shading) that receives four synapses and has a spine-like geometry. Figure adapted from (Kornfeld et al., 2017).*

classification of eleven interneuron dendrites that were identified based on other morphological characteristics (soma morphology and symmetric outgoing synapses on their axons, since the dendrites could be traced back to the soma), and the correct classification of 17 out of 18 HVC_{RA} dendrites that were identified by the BDA-label.

3.2.4 Postsynaptic targets of HVC_{RA} axons

The established spine density classification method was then used to identify the synaptic targets of three HVC_{RA} axons that were connected to a soma in the EM dataset (axonal path lengths: 1.37, 0.88 and 0.72 mm). 121 synapses were identified in total, of which 115 were formed on HVC_{INT} and only six on excitatory cell types (four HVC_{RA} and two HVC_X). While this finding is compatible with the high probability of HVC_{RA} to HVC_{INT} connectivity found in an electrophysiological study based on pairwise patch-recordings in slices (Kosche et al., 2015), it is in conflict with the previous estimate (15 %, see 3.2.2) of the fraction of homotypic synapses. By definition, HVC_{RA} neurons must receive on average the same number of homotypic synapses as they form on other HVC_{RA} neurons. When the density of HVC_{RA} to HVC_{RA} synapses from these 121 synapses is extrapolated to the estimate of total HVC_{RA} axonal path length in the brain nucleus, only about 20 homotypic synapses per HVC_{RA} would be expected, about six times fewer than estimated from the dendritic perspective.

The synaptic targets of BDA-labeled axons (putative HVC_{RA} axons) that could not be traced back to their soma were analyzed next (path length: 0.56 ± 0.27 mm, mean $\pm$ s.d.; $n = 6$). Three of these axons were traced from a labeled dendrite and four were partially myelinated. Surprisingly, the number of synapses onto excitatory postsynaptic cells was thirteen times higher for this population, from 5 % to 64.6 % (263 of 407 synapses). The number of putative homotypic synapses increased in a similar way, from 3.3 % to 36.7 % (11-fold, 150 of 407 synapses). Additionally, multiple synapses from a labeled axon to a postsynaptic HVC_{RA} dendrite were found: 17 doubles, 3 triple and 1 quintuple out of 127 pairs. What could explain this striking difference? Axon fragments that cannot be traced back to their soma inside the dataset must be farther away from it (at least on average, but rare cases where a soma could be found directly next to the boundaries of the dataset are possible). This relationship implies that the axon fragment to soma distance should be quantified, requiring a local criterion that can be translated into an estimate of the axon-soma distance. LM reconstructions (S. Benezra) provided a hint: axonal branching is reduced with increased soma distance (see Fig. 3.9). This relationship was also found for the EM reconstructions, where HVC_{RA} axons that could be

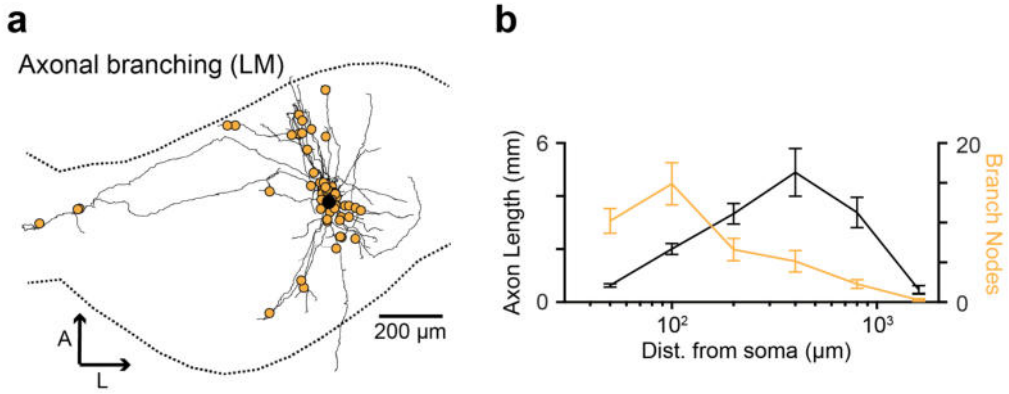

Figure 3.9: *LM ground truth for the estimation of the axon fragment soma distance, data by S. Benezra, Long laboratory. (a) Axon of a HVC_{RA} neuron (soma in black) with all axonal branch points labeled in orange. Dashed lines indicate the outline of HVC. (b) Axon length and branch nodes vs distance from soma. Figure adapted from (Kornfeld et al., 2017).*

followed to the soma had 12.4 ± 3.7 branch points / mm (mean ± s.d.; n = 3), much higher than the orphaned fragments with 4.0 ± 4.3 branch points / mm (mean ± s.d.; n = 6).

Bayesian inference was then used to obtain an estimate and its uncertainty of the fragment soma distance for each orphaned fragment (see section 3.1.2 for details). In short, the axonal length density as a function of soma distance was used as a prior, and the number of branches per axon length was modeled with a Poisson distribution. The soma distance distribution for each axon fragment (including the fragments that could be traced to a soma as a control) was then inferred from its number of branch points and length. The uncertainty of the estimate was given by the width of the distribution, see Fig. 3.10. The soma distance of the three axon fragments that were connected to a soma was estimated to be below 100 µm, validating the approach for the given EM dataset size. It should be noted that a direct analysis on individual axons seemed impossible given the small dataset size and the fact that there were almost no excitatory post synaptic targets, but such an analysis would clearly have been the preferred option.

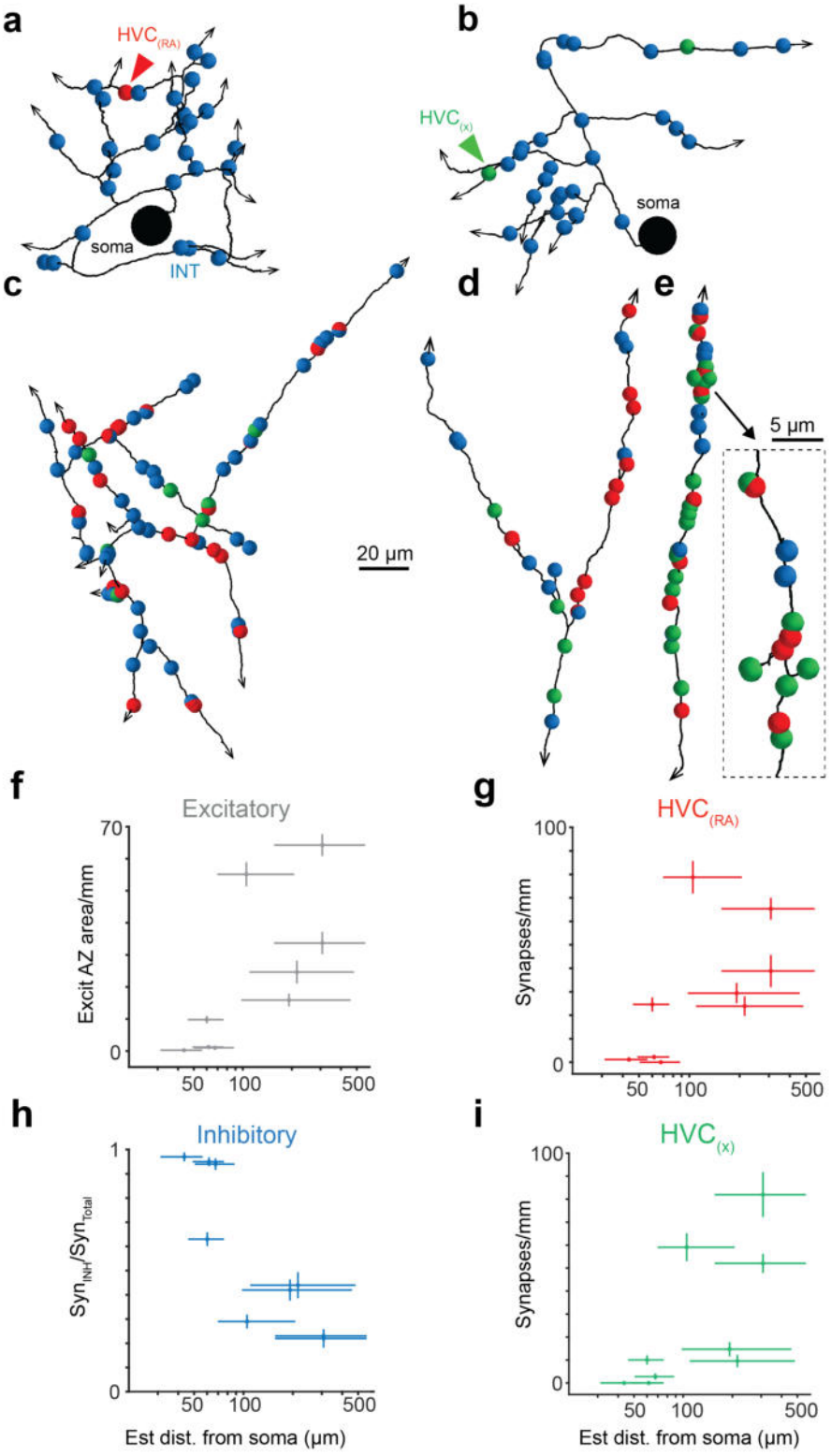

Figure 3.10: *Axonal target variation as a function of soma distance, estimated using Bayesian inference. (a)-(e) Skeleton reconstructions of axons of putative HVC_{RA} neurons, with their outgoing synapses colored by postsynaptic type. (f) Area of active zone onto excitatory neurons, as a function of soma distance, estimated using a Bayesian model; $p < 0.05$, Pearson's correlation. (f)-(i) Error bars on the x-axis indicate the 0.16 quantile and 0.84 quantile of the Bayesian posterior distribution. Y-axis error bars represent the s.e.m. of an assumed underlying Poisson distributed synapse count for (f), (g) and (i), or the s.e.m. of an assumed underlying binomial count distribution (h). Figure adapted from (Kornfeld et al., 2017).*

Since only a few (n = 9) axon fragments were inspected and so could have been affected by selection bias, a separate and independent analysis was performed to estimate the fraction of HVC_{RA} homotypic connectivity from the axon perspective, in comparison with the dendritic perspective (see section 3.2.2). To carry out this analysis, the 59 BDA-labeled axon fragments identified in the 300 member set (see section 3.2.2) were traced until the first two synapses were reached or the tracing terminated at the dataset boundaries. The postsynaptic cell type of these synapses was then determined as described previously (see section 3.2.3). 105 synapses were identified in total, with 65 being targeted interneurons, 22 HVC_X neurons and 18 HVC_{RA} neurons. An estimate was then found for the number of outgoing synapses per HVC_{RA} neuron of 1111 ± 513 (s.d.) (based on a total average axon path length of 14.7 mm, and a synapse density of 75.4 synapses / mm), leading to an estimate of about 191 ± 88 (s.d.) HVC_{RA} homotypic synapses per cell. This number translates to an estimate of 24 ± 4 % (s.e.m.) homotypic synapses as a percentage of all excitatory synapses, slightly larger than the estimate obtained from the dendritic perspective (about 15 %). This difference might be due to the strict acceptance policy for identifying double-labeled synapses for the dendritic estimate, or could be caused by other uncertainties that were not accounted for.

3.2.5 Could the orphaned axons belong to RA_{HVC} neurons?

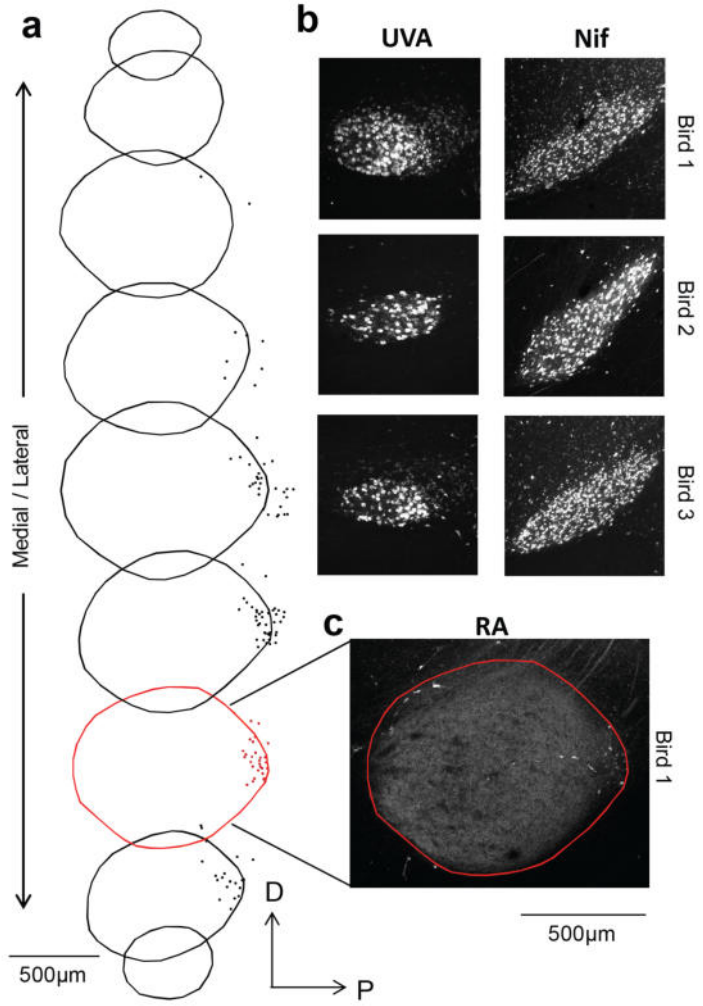

Figure 3.11: *Control experiment to estimate the number of RA_{HVC} neurons, performed by S. Benezra. (a) Borders of nucleus RA, reconstructed from 100 µm thick sagittal sections. Dots represent RA_{HVC} neurons. Only a few cells in RA were labeled (<1 %) by the DiI injection into HVC. (b) Robust labeling in Uva and NIf, in comparison to RA, demonstrating that the tracer is efficient. (c) A confocal micrograph shows the small number of retrogradely labeled neurons in the posterior part of RA. Figure adapted from (Kornfeld et al., 2017).*

Roberts et al. (Roberts et al., 2008) estimated that there is a significant population of RA_{HVC} neurons, and the anterograde projections to HVC might therefore be labeled with BDA after injection of BDA into RA. The number of RA_{HVC} neurons was quantified by DiI injection into HVC (LM experiments performed by S. Benezra) and compared to the number of HVC_{RA} neurons (about 40,000 per hemisphere). Only a small number of RA_{HVC} neurons (125, 163, 171, n = 3 birds) were found, despite heavy labeling of other upstream areas such as NIf and Uva (see Fig. 3.5). Each RA_{HVC} neuron must therefore have an axon of about 4 m path length

inside of HVC to account for the density of labeled axon in the dataset. This length, however, appears unlikely, given that the strongly ramifying HVC_{RA} axons have a path length of only about 0.015 m. It should also be noted that 3 kDa BDA is known to be a better retrograde than anterograde tracer in birds (Reiner et al., 2000), making it even less likely that a substantial fraction of the orphaned axons originates from RA_{HVC} neurons.

3.2.6 Postsynaptic targets of HVC_X and HVC_{INT} neurons

A preliminary analysis was carried out on the synaptic targets of the proximal axons of two putative HVC_X neurons and two axon fragments from HVC_{INT} neurons in a similar way to the previous sections. Due to the manual annotation efforts required for synapse counting on spines (see section 3.1.2), the postsynaptic dendrites were only visually classified, but pure visual classification does not necessarily sacrifice substantial accuracy as can be seen in Fig. 3.7 a and b. Similarly to the proximal HVC_{RA} axons, the proximal HVC_X axons almost exclusively targeted (15 out of 17) smooth dendrites of HVC_{INT}, in addition to two putative HVC_{RA} neurons. HVC_{INT} axons had a very different ultrastructural appearance compared to HVC_{RA} and HVC_X axons, with thick caliber axons and a high density of synaptic vesicles and synaptic sites (see Fig. 3.3). Interestingly, and in stark contrast to the HVC_X and HVC_{RA} neurons, the 20 analyzed inhibitory synapses, sampled from axonal parts proximal to their soma, targeted mainly spiny postsynaptic dendrites (5 HVC_{INT}, 10 HVC_{RA}, 5 HVC_X). This suggests that the axons of inhibitory and excitatory cells follow fundamentally different wiring patterns in HVC.

3.3 Discussion

3.3.1 The role of homotypic HVC_{RA} connectivity

Synaptic chain models of neuronal sequence generation (Abeles, 1982) rely on a wiring diagram with excitatory feed-forward connectivity (in an idealized case with no "noisy" connections that

do not adhere to the temporal wiring scheme), as described in section 1.2.3. This study and previous electrophysiological work (Kosche et al., 2015) demonstrated that the most prominent wiring pattern is the dense recurrent HVC_{RA} and inhibitory interneuron connectivity, and that the pairwise connection probability for homotypic connections is small (<1 %). However, such low pairwise connection probabilities can be deceptive, since the overall fraction of homotypic inputs of all excitatory synapses was estimated here to be almost 25 %, or about 200 per neuron. This observation illustrates the importance of considering the total synaptic input of a neuron instead of focusing on pairwise connection probabilities. However, it should be noted that the remaining 75 % of the incoming HVC_{RA} excitatory synapses could not be characterized in this study. The sensorimotor cortical area NIf, the thalamic nucleus Uva, and other auditory areas are known to project to HVC (reviewed in Mooney, 2009), but to what extent they drive the HVC_{RA} neurons is less clear. Inactivating NIf leads only to transient song degradation (Cardin et al., 2005; Otchy et al., 2015) and auditory inputs were shown to be of little importance (Vallentin and Long, 2015), similar to the contribution of HVC_{X} neurons (Scharff et al., 2000). Despite these cues, and the less clear role of Uva (Hamaguchi et al., 2016; Coleman and Vu, 2005), it remains to be tested whether direct excitatory connections from other HVC_{RA} neurons are the main driver of the temporally precise bursts. It could well be that an individual HVC_{RA} neuron makes use of any excitatory input that contributes to the "right timing" - this would be consistent with the effect of transient NIf inactivation, which leads to song degradation (Otchy et al., 2015). After permanent NIf lesions (Cardin et al., 2005), the lost excitatory synapses would for example be replaced through homeostatic regulation with HVC_{RA} synapses (Otchy et al., 2015). Future experiments should clearly combine functional HVC activity imaging to determine the burst time of individual HVC_{RA} neurons, followed by large scale EM analysis of the same tissue. Such an experiment would provide a better measure of how relevant excitatory feed-forward connectivity between HVC_{RA} neurons is, and whether the temporal succession of activity over a neuron population can be inferred directly from synaptic connectivity.

3.3.2 Hypotheses about the developing HVC

The wiring patterns described here were found in an adult bird with a mature HVC circuit, but it is unclear how they develop. The low number of proximal homotypic synapses in the adult HVC could be the result of genetically encoded dense local recurrent connectivity with inhibitory interneurons: in the beginning of development, HVC_{RA} neurons might form more homotypic synapses on their proximal axonal parts, but these connections could then slowly be pruned through Hebbian plasticity (Hebb, 1949) because of the strong local recurrent inhibition. Alternatively, the observed connectivity patterns might arise from gradients in molecular cues. It would therefore be of interest to perform experiments that elucidate whether the distance-dependent connectivity pattern already exists in young animals.

3.3.3 Functional consequences of the observed wiring

The dense local recurrent inhibition is a feature often ascribed to winner-take-all (WTA) circuits, where the activation of a functionally grouped excitatory neuronal population leads to a suppression of neighboring principal neurons through inhibition. Such circuits were introduced as models for decision making (Hopfield and Tank, 1985), but were more recently also employed as a theoretical framework for neuronal sequence generation (Mostafa and Indiveri, 2014). Interestingly, these models are based on local microcircuits with dense recurrent inhibition that exhibit WTA dynamics, coupled through distal excitatory synapses formed by the locally winning principal neurons (see Fig. 3.12). The synaptic architecture analysis of HVC in this thesis shows that such a connectivity pattern might exist in HVC, but it should be noted that the computations performed by the zebra finch HVC during song production are extremely stereotyped. The assumed local WTA dynamics are therefore expected to always lead to the same outcomes rather than selecting between different sequences. However, this functionality might be different for birds with a more complex syllable repertoire and order (Jin, 2009). The synaptic anatomy reported here is further in agreement with a modeling study performed by Cannon et al., 2015 that is based on the idea of spatio-temporal activity patterns that cycle through HVC. In this model, local inhibitory zones that are the result of the dense recurrent

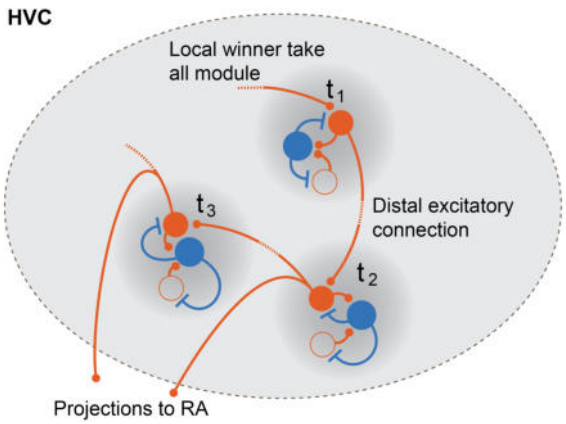

Figure 3.12: *Hypothesized model of HVC function during song production and how inhibitory and excitatory neurons may interact. It is proposed that the burst activity of HVC_{RA} neurons cycles through the brain area, driven by local suppression and distal excitation, consistent with Cannon et al., 2015.*

inhibition sharpen the excitatory pulse that travels through the network.

3.3.4 The neurite identification problem in the j0256 dataset

Despite the usage of a mainly retrograde tracer (3 kDa BDA), the cell type identification of the orphaned HVC_{RA} axons remained indirect and required extensive control experiments (see section 3.2.5) to estimate the probability of confusion with RA_{HVC} projecting axons. This ambiguity was illustrative for the frequent problem of neurite identification in small volume EM datasets, which also had to be solved for the identification of dendritic reconstructions (see section 3.2.3). A larger HVC dataset would have had two significant advantages: it would have allowed to follow all axons back to their somata, instead of performing a Bayesian analysis to estimate the soma distance (section 3.1.2), and additionally would have increased the number of cues (morphological, but also from connectivity) to infer the cell type for a reconstruction. It is therefore concluded that the higher efforts required upfront for the acquisition of a larger dataset might in fact be compensated later by the time saved by avoiding additional analyses required for a smaller dataset.

Chapter 4

An anatomical hypothesis test in the zebra finch basal-ganglia

Publication of the results presented in this chapter

The results of this chapter had not yet been published in a peer-reviewed journal by the submission time of this thesis.

All experiments were performed by J. Kornfeld in the laboratory of W. Denk (MPI for medical Research, Heidelberg; MPI of Neurobiology, Martinsried) unless explicitly stated otherwise. The project was performed in collaboration with the laboratory of M. Fee (MIT, Cambridge).

4.1 Methods

4.1.1 Sample preparation and data acquisition

The same EM sample preparation procedure as described previously for the HVC dataset (see section 3.1.1) was used with some small changes, the most notable being the lack of a retrograde labeling procedure which greatly simplified the protocol. The Area X sample of the j0251 dataset was taken from an adult male zebra finch (>120 dph) that was perfused by M. Stetner in the laboratory of Michale Fee (MIT, Cambridge) using the fixation method described earlier but with 1 % GA and 1 % PFA. The sample for the j0126 dataset was taken from an adult male zebra finch (>120 dph) that was provided by the laboratory of M. Gahr (MPI for Ornithology, Seewiesen).

The same microscope setup (Zeiss UltraPlus, see section 3.1.1) as described previously was used for both datasets that were acquired for this chapter, j0126 and j0251. The different imaging and SBEM parameters are described separately in the following sections.

Acquisition of the j0126 dataset

The j0126 dataset was acquired with a single image tile (10,240 by 10,240 pixels) per microtome slice with a landing energy of 1.6 kV and an electron current of 1 nA at a scan rate of 3.3 MHz. The lateral resolution of each image was 9 x 9 nm and the cutting thickness was set to 20 nm. The individual images were registered with custom MATLAB scripts (K. Briggman and K. Börgens) by globally minimizing cross-correlation shifts between the layers. The registered images were then converted to a cubed 3D image pyramid with a custom Python script[1] for neurite reconstructions and synapse annotations with KNOSSOS, as described previously in section 3.1.1.

[1] `https://github.com/knossos-project/knossos_cuber`

Dataset j0251

The j0251 dataset was acquired with 16 overlapping (4 by 4) and equally sized image tiles (6,600 by 6,600 pixels) per microtome slice with a landing energy of 2 kV and an electron current of 1 nA, at a scan rate of 5 MHz and lateral resolution of 10 x 10 nm. The microtome cutting thickness was set to 25 nm. Registration was performed as for the j256 dataset with affine transformations based on the Karsh alignment framework[2] (Scheffer et al., 2013) and custom Python scripts. Additionally, contrast limited adaptive histogram equalization (CLAHE)[3] was applied (gradient from the top to the bottom of the dataset, from a clip limit of 1.0 to 2.4) to visually compensate for the effects of a degrading signal-to-noise ratio and frequent signal offset shifts that were encountered during the acquisition. The data was then cubed to a KNOSSOS readable format using custom Python scripts, as described for the other datasets.

4.1.2 Neurite reconstructions and synapse annotations

All skeleton and synapse reconstructions were performed with KNOSSOS[4] (Kornfeld et al., 2011; Helmstaedter et al., 2011). Neurites were annotated similarly to the annotations of the HVC j0256 dataset: seed coordinates were provided to at least two independent human annotators. The redundant annotations were then automatically matched by comparing all tracings of the entire dataset, since seeds from different locations could belong to the same neurons (matching parameters as described previously, see section 3.1.2). The matched skeletons were then reviewed by annotators and discrepancies were inspected and resolved, leading to a single consensus skeleton tracing per neurite. Synapses were annotated as described in the previous chapter (see section 3.1.2), in short by tracing of the synaptic junction and extrapolation to a circle for area estimation. All further data analysis was performed using custom Python scripts based on the open-source KNOSSOS utility package[5].

[2] `https://github.com/billkarsh/Alignment_Projects`
[3] `http://docs.opencv.org/trunk/d5/daf/tutorial_py_histogram_equalization.html`
[4] `www.knossostool.org`
[5] `https://github.com/knossos-project/knossos_utils`

4.2 Results

4.2.1 Acquired SBEM datasets of Area X

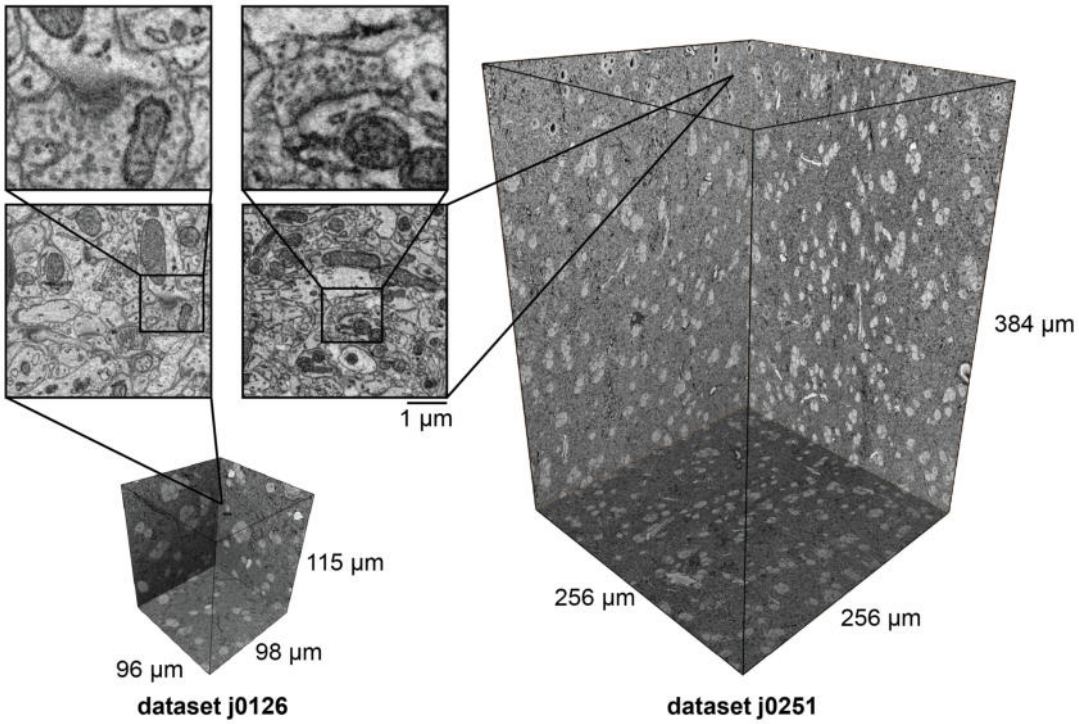

Figure 4.1: *Acquired datasets of the striatopallidal Area X from adult male zebra finches. Left: j0126, which has a total size of about 0.7 TB and contains about 450 somata. Right: j0251, which has a total size of about 10 TB and contains more than 10,000 somata. The insets show the image quality at the location of synapses. Note that individual synaptic vesicles can be discriminated in both datasets.*

To analyze the wiring of the zebra finch basal-ganglia Area X, two SBEM datasets from different adult wild type male zebra finches were acquired. j0126 and j0251 are both from central regions of Area X, and their main difference is dataset size, as illustrated in Fig. 4.1. Both datasets were acquired with ECS preservation (Cragg, 1980) to improve the performance of automated reconstruction algorithms (Pallotto et al., 2015). j0126 offers a slightly better image quality at a voxel size of 9 x 9 x 20 nm, but has only about 4 % of the volume of j0251 (voxel size of 10 x 10 x 25 nm), illustrating the progress made in dataset acquisition size over time. The results of the following sections were derived from the j0126 dataset and are based on manual analysis, since the computational approaches presented in the last chapter of this thesis had not yet been established at the time of analysis. Similarly, a detailed analysis of j0251 cannot

be part of this thesis due to time constraints.

4.2.2 Testing the structural predictions of a model of reinforcement learning

One of the objectives of this thesis was to test the anatomical predictions of the RL model described in the introduction (see 1.2.4). In short, the model predicts that context synapses (from HVC axons) should be found making contact with the spine heads of MSNs, and efference copy synapses (from LMAN axons) should contact the dendritic shafts. To test this hypothesis, skeleton reconstructions were performed of MSNs and axons forming asymmetric synapses in the j0126 dataset. Skeletons and synapses were manually annotated, as described previously (see the methods section from the previous chapter, 3.1).

Morphological classification of excitatory axons in Area X

Context (HVC) and action (LMAN) axons were discriminated based on their morphology, and LM tracing ground truth data was provided by M. Stetner (Fee Laboratory, MIT, Massachusetts) which enabled the visual classification of EM reconstructions. Fig. 4.2 a and b show LM micrographs and their reconstructions, illustrating that HVC axons branch rarely and are mostly straight inside of Area X, while LMAN axons form highly branched focal arborizations. Both axon types were found to synapse onto MSNs (Fig. 4.2 c). Two exemplary EM reconstructions of axons (Fig. 4.2 d, e), transsynaptically traced from MSNs, are shown to illustrate the EM morphologies. To avoid misclassification, the further analyses presented here were restricted to reconstructions that fell clearly into the two described morphological categories.

Axon perspective on the hypothesis test

As a first step, axons that formed asymmetric synapses were skeletonized (traced from afferent asymmetric synapses of MSNs). 11 were selected based on their morphology (5 putative LMAN

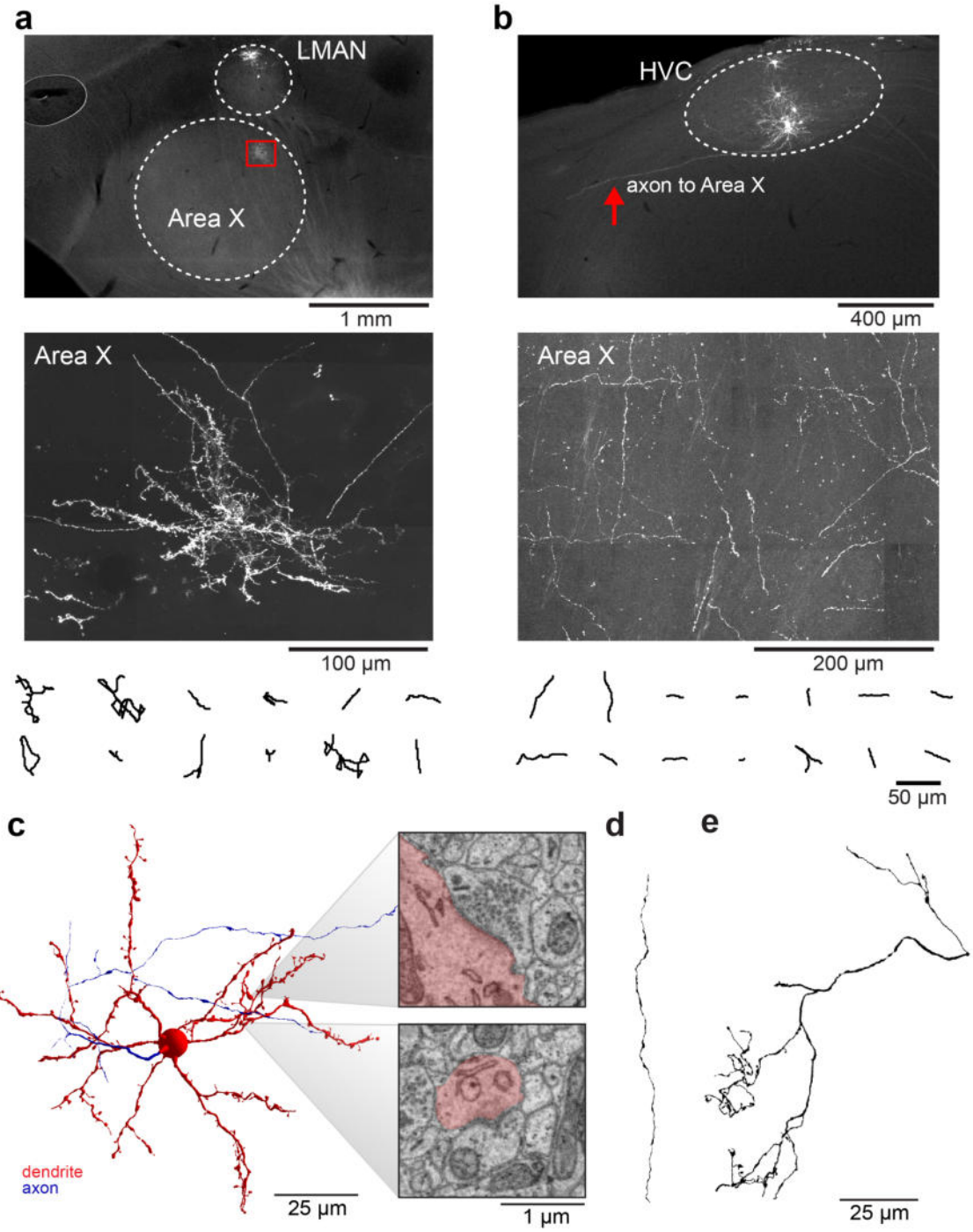

Figure 4.2: *LM ground truth and EM tracings in Area X; LM data produced by M. Stetner, Fee laboratory. (a) Top: A single LMAN-projection neuron was virally labeled with GFP. Middle: Only a small volume is covered by its axon in X, which forms a dense terminal arborization. Bottom: LM tracings of small LMAN axon fragments on the EM dataset size-scale. (b) Top: HVC projection neurons labeled virally with GFP, with one of them projecting to Area X. Middle: Diffuse and largely unbranched axon inside X. Bottom: Reconstructions as in (a). (c) EM MSN tracing with the typical spiny dendrite (skeleton nodes adjusted to neurite thickness, see next chapter). The top inset shows an asymmetric dendritic shaft synapse, the bottom a spine head synapse. (d) HVC-like axon forming asymmetric synapses. (e) LMAN-like axon forming asymmetric synapses.*

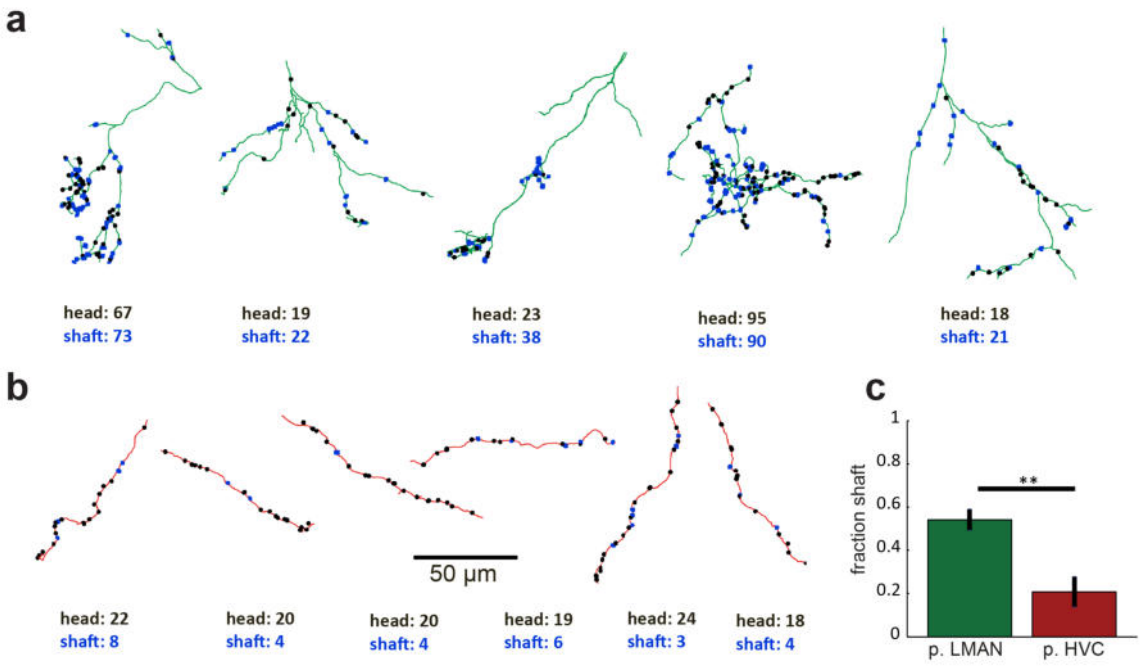

Figure 4.3: *Axons forming asymmetric synapses with LMAN-like or HVC-like morphology target different dendritic compartments. (a) Synapses annotated on 5 LMAN-like axons (branched) and manually classified as spine or shaft synapse. (b) Synapses annotated on 6 HVC-like axons (straight, unbranched) and manually classified as spine or shaft synapse. (c) Quantification of the fraction of shaft synapses from 618 synapses. t-test, $p < 0.01$. Error bars are s.d..*

axons, and 6 putative HVC axons), and all their synapses were manually identified. It was then analyzed for each synapse whether it targeted the postsynaptic neuron on a spine head or the dendritic shaft (see Fig. 4.3). In total, 618 synapses were annotated in this way, whereby most of them were found on p. LMAN axons, often enriched in parts of the axon with many branches. Synapses on p. HVC axons were distributed more evenly, often at regular distances (about 5-10 µm). It was then tested whether LMAN-like axons targeted dendritic shafts more frequently, and a significant difference was found, consistent with the predictions by the RL-model (see section 1.2.4). However, LMAN-like axons showed a high fraction of spine head synapses (about 50 %), and HVC-like axons also formed shaft synapses (about 20 %), an observation for which the model has no explanation. Since these axons were traced from asymmetric synapses of MSNs, a potential selection bias could not be excluded. It is further likely that the axons also form synapses with other cell types than MSNs (despite the other cell types being rare), and synapse counting neglects the functional relevance of synapse sizes (Cossell et al., 2015). A further analysis was therefore devised to address these concerns.

Hypothesis test under the inclusion of synaptic areas

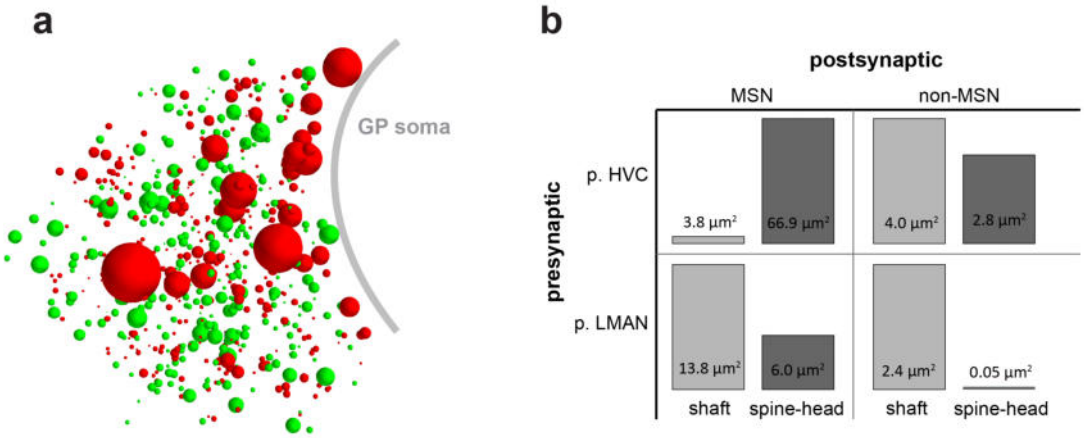

Figure 4.4: *Unbiased analysis of synaptic sizes of LMAN-like and HVC-like axons with MSNs. (a) The central* $(10\,\mu m)^3$ *subregion of j0126 was selected for a dense synapse analysis, and 662 synapses were identified. Symmetric synapses are shown in red, asymmetric in green. The sphere diameter was scaled with the synaptic areas to illustrate the size differences. A large pallidal-like soma is located in the central subvolume of the dataset, attracting massive inhibitory synaptic terminals. (b) "Connectivity matrix" with additional geometric synapse information, and cumulative synaptic areas between the different cell types.*

To eliminate any bias in the choice of the analyzed synapses and axons, the central $(10\,\mu m)^3$ region of the j0126 dataset was selected for a dense manual synapse analysis. In total, 662 synapses were identified, of which 340 were classified as asymmetric and 322 as symmetric (see Fig. 4.4 a). This set of synapses demonstrates an interesting difference when compared to cortical tissue, which contains mainly asymmetric synapses (about 80 % of all synapses, reviewed in DeFelipe and Fariñas, 1992). The pre- and postsynaptic partners were reconstructed, and 260 synapses were identified with a partner that could be classified as pre-synaptically p. HVC or p. LMAN and that belonged post-synaptically to a MSN. The synaptic contact area was then estimated using manual synapse annotations (see section 4.1.2), and a striking dominance of synaptic area was found for HVC-like axons and spine-head synapses of MSNs compared to LMAN-like axons (about 11-fold difference). Similarly, when incorporating synaptic areas, LMAN-like axons showed a stronger preference towards shaft synapses (about 70 %) compared to pure synapse counts (compare to section 4.2.2). These data are again consistent with the predictions arising from the RL-model and demonstrate that synapse counts alone should be judged cautiously.

4.2.3 Neuron reconstructions in the j0251 dataset

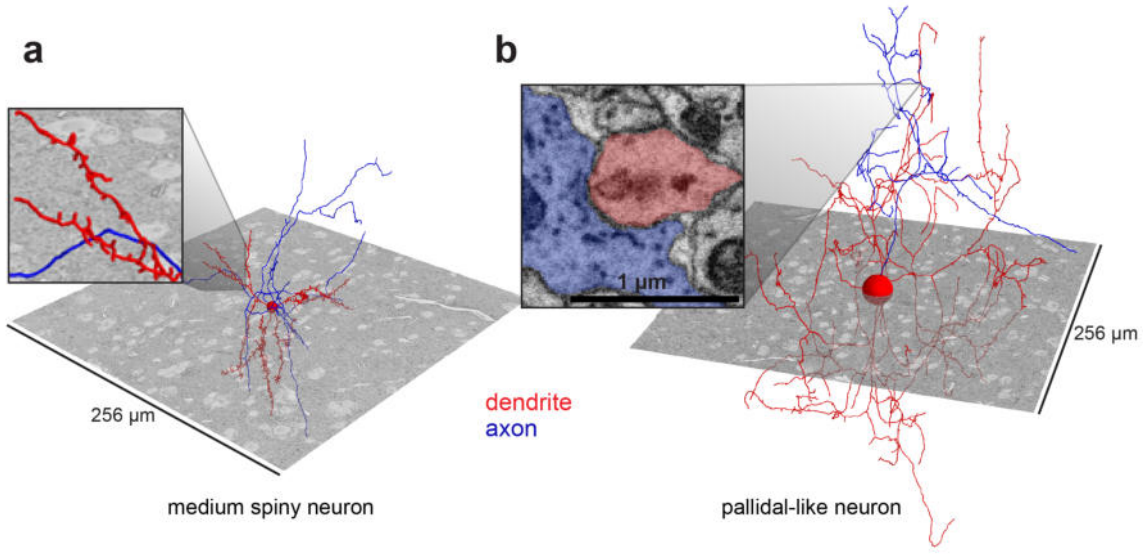

Figure 4.5: *Skeleton reconstructions of neurons in the j0251 dataset of Area X. (a) MSN tracing, where the inset shows dendritic spines. Note that the entire dendritic field of the neuron is included in the dataset. (b) Tracing of a pallidal-like neuron that formed several autapses, one shown in the inset. Even those very large neurons are almost entirely contained in this dataset.*

In total, 30 neurons (10 MSNs and 20 putative pallidal-like neurons) were reconstructed in the j0251 Area X dataset to evaluate traceability and generate ground truth for future automated reconstruction approaches. Fig. 4.5 shows a MSN reconstruction with the typical spines on the dendrite and an aspiny GP neuron. The size of the j0251 dataset enabled near-complete EM-based tracings of MSNs and allowed the reconstruction of large portions of the biggest neurons that can be found in Area X, the GP-like cells, showing the potential of this dataset for basal-ganglia circuit analysis compared to the smaller j0126 image stack. While still preliminary, this reconstruction allowed for the first observation of autapses for a GP neuron (at least to the knowledge of the author, but see Karabelas and Purrura, 1980, for evidence of autapses in substantia nigra pars reticulata neurons), raising the question whether they might be of mechanistical relevance given the high-frequency firing patterns during singing interrupted by short pauses (Goldberg et al., 2010). At the point of the submission of this thesis, no further data analysis could be performed because of the tremendous manual effort required. This limitation motivated to a large degree the work of the next chapter, which presents results regarding how data analysis can be accelerated using machine learning algorithms that are

combined with a proofreading workflow.

4.3 Discussion

4.3.1 An anatomical test of a mechanistic learning model

Similar to the analysis performed in chapter 3, a volume EM dataset was used in combination with specific predictions derived from a mechanistic functional model, without the knowledge of a full wiring diagram. The limiting factor was not the amount of information contained in the datasets, but rather the access to it, due to the manual annotation efforts required. Aside from such technical difficulties, there is a further challenge with the strict hypothesis driven approach: how to generate the anatomical predictions from a functional model? The gap between the mechanistic model of song learning in the basal-ganglia and the anatomy of the region seems large at first, but it was easily bridged with the incorporation of prior knowledge of how physiology and structure are likely related. The model predicts that the context-carrying HVC axons should drive MSNs and that the synaptic strength of these synapses is modulated, while the efference copy signal carrying axons (LMAN) should pave the way for the plasticity of the context synapses. A plethora of evidence points to the importance of dendritic spines in synaptic plasticity, while shaft synapses might have more global effects (reviewed in Yuste and Bonhoeffer, 2001). However, the demonstrated approach also exposes its fundamental limitation: without any prior physiological knowledge or established links between structure and function, it seems difficult to draw conclusions from the static EM data.

4.3.2 The neurite identification problem in the j0126 dataset

The manual analysis of the small j0126 dataset of Area X demonstrated again (as discussed already in section 3.3.4 for HVC) that small EM datasets suffer from a severe neurite identification problem. It was difficult to establish links between short EM reconstruction fragments and assign them to known cell types based on known morphology, and novel ground truth had

to be generated for both HVC (see section 3.2.3) and Area X (see section 4.2.2) to tackle a problem that would hardly exist if all neurites could be traced to their original cell bodies. While the j0251 dataset also contains many neurite fragments that cannot be traced to their cell body, the additional spatial context will likely simplify the morphological identification of neurons substantially.

4.3.3 The challenges and opportunities of large volume EM datasets

The first skeleton reconstructions in the larger Area X dataset (somewhat smaller in volume to the recently published whole zebrafish brain, but at a significantly better resolution (Hildebrand et al., 2017), but see also (Svara, 2017) for a high-resolution zebrafish dataset) revealed new technical difficulties, even for sparse neuron reconstructions. Single tracings of cells could take more than 100 h for a single annotator. This brought the previously used crowd-sourcing approach with student assistants to its limit, since each assistant usually works only a few hours per month. The now somewhat classical reconstruction approach of redundant skeleton annotations followed by the generation of a consensus skeleton, either automatically (Helmstaedter et al., 2011) or manually (Kornfeld et al., 2017), must therefore be accelerated. Speedups could be accomplished using automation, or manual tracing procedures could be modified to use a "shotgun-tracing" approach that focuses on smaller fragments that are later combined, instead of attempting that individual cells are completed by single annotators.

To overcome the aforementioned manual analysis challenges, novel methods were developed to speed up neurite reconstructions and identify synapses automatically. These methods will be presented in the next chapter of this thesis.

Chapter 5

Novel methods for connectomic analysis

Publication of the results presented in this chapter

Parts of this chapter contain text and figures from the following peer-reviewed publication:

> **Automated synaptic connectivity inference for volume electron microscopy.**
> Dorkenwald, S., Schubert, P., Killinger, M., Urban, G., Mikula, S., Svara, F., Kornfeld, J.*
> Nature Methods, 14(4):435–442 (2017)
> * corresponding author

All experiments were performed by J. Kornfeld in the laboratory of W. Denk, unless explicitly stated otherwise. S. Dorkenwald and P. Schubert (student assistants) developed the SyConn software framework together with J. Kornfeld.

5.1 Methods and used data

5.1.1 GraphWalker

j0126 Google segmentation

The j0126 segmentation (referred to as version 3 or v3) used for the proofreading workflow was provided by a group at Google Research (M. Januszewski, P. Li, J. Maitin-Shepard, A. Pope, M. Tyka, T. Blakely and V. Jain) as part of a scientific collaboration. It was generated using convolutional neural networks (CNNs) followed by graph-based active learning of agglomeration (GALA, Nunez-Iglesias et al., 2014), optimized and trained on ground truth data provided by the author of this thesis.

Implementation

The GraphWalker frontend was implemented in Python 2.7 as a plugin for KNOSSOS[1] (Kornfeld et al., 2011; Helmstaedter et al., 2011) in a single Python module that can be distributed to annotators. It uses the PythonQt framework[2] to create a graphical user interface and to integrate the proofreading workflow into KNOSSOS (e.g. by adding keyboard shortcuts for efficient annotation). The currently performed annotation is stored in memory as a networkx graph object (Hagberg et al., 2008) and the undo and redo functionality is implemented with snapshots stored on a stack. The frontend communicates with an extended version of Heidelbrain (originally developed by F. Svara) through a RESTful HTTP API. Heidelbrain is a Python Django application[3] that provides annotator management capabilities, such as time tracking and task management. It was extended with a Neo4j graph database[4] that stores the segmentation supervoxel graph of the analyzed dataset and the created annotations. The Neo4j database is internally accessed via the py2neo connector library from Django[5]. The plugin fetches precom-

[1] `https://knossostool.org/`
[2] `https://sourceforge.net/projects/pythonqt/`
[3] `https://djangoproject.com`
[4] `https://neo4j.com/`
[5] `http://py2neo.org/v3/`

puted supervoxel meshes for display in KNOSSOS from the Google BrainMaps service after Oauth2 authentication with a Google service account.

GraphWalker evaluation

To evaluate GraphWalker, ten neurons that were reconstructed through skeletonization in the j0126 dataset were picked randomly as a ground truth test set. The same neurons were skeletonized independently by an expert annotator to be able to compare the reconstruction speed and accuracy against expert skeletonization. The reconstruction speed was determined by eliminating all skeleton nodes from the ground truth skeletons that were more than 1 µm away from the tested skeleton (combined re-skeletonized supervoxels for GraphWalker reconstructions and expert skeleton for control). The total annotation time of the tested skeleton was then divided by the remaining path length in mm.

The reconstruction accuracy was quantified by measuring two separate quantities:

- The fraction of the recovered path length of the ground truth annotation.
- The number of false mergers shorter than 10 µm and between 10 and 50 µm.

The fraction of recovered path length was computed by eliminating for the test reconstructions all parts of the matching ground truth skeleton that were more than 1 µm away from the reconstruction, and then dividing the remaining path length by the total ground truth skeleton path length. False mergers were visually identified and validated in KNOSSOS by comparing the ground truth skeleton with either the GraphWalker reconstruction or the expert skeleton.

5.1.2 Synaptic connectivity inference of the j0126 dataset

The full Python source code of the latest version of the SyConn software package is available as a Github source code repository[6] and the source code snapshot used for the data presented

[6] `https://github.com/StructuralNeurobiologyLab/SyConn`

in this thesis can be found on the project website[7].

Skeleton annotations

All neurites were skeletonized with the KNOSSOS annotation software by at least two independent annotators, as described in detail previously (see sections 4.1.2 and 3.1.2).

CNN ground truth annotations

KNOSSOS was also used to generate volume annotations for the training of CNNs in subvolumes of the j0126 dataset. The following data categories and volumina were annotated:

- 988 μm^3 of synaptic junctions, vesicle clouds and mitochondria, with one ID each, and everything else as background.
- 941 μm^3 of binary synaptic junction type labels.
- 81 μm^3 of neurite reconstructions, with one ID per neurite, and ECS and membranes as background.
- 603 μm^3 of myelin, with separate IDs for the interior of the axon and the myelin itself, and everything else as background. The annotations were performed on 2-fold downsampled data.

Additionally, coordinates from the centers of the somata in the j0126 dataset were tagged, and all voxels around them in a 1.7 μm radius were used as labels for the prediction of soma locations. The dense neurite reconstructions were converted to the neurite barrier ground truth with a Gaussian gradient magnitude filter to detect neurite boundaries.

[7] `https://structuralneurobiologylab.github.io/SyConn/`

CNN classifiers

Detailed hyperparameters of the employed CNN-architectures are provided in the source code[8] and journal publication supplementary material (Dorkenwald et al., 2017). In short, 6-layered (synaptic junctions, vesicle clouds, mitochondria, barrier) CNNs with <200,000 parameters were used for an initial prediction, followed by a second CNN for a recursion with either 6 layers (synaptic junctions, vesicle clouds, mitochondria, <200,000 parameters) or 10 layers (barrier, about 1,000,000 parameters) that had the original raw data and the prediction from the previous stage as input. Cell bodies, myelin and synapse types were predicted with a single stage CNN (cell bodies: 8 layers, about 260,000 parameters, synapse types: 6 layers, about 190,000 parameters, myelin: 6 layers, about 170,000 parameters).

Post-processing of CNN-predicted probability maps

Thresholding and connected component analysis followed by size filtering (threshold and minimum size determined on the test set) was used to obtain synaptic junction, vesicle cloud and mitochondria objects from the voxel-based probability maps that were generated by the CNNs. Myelinated axons were found by averaging the probability map of the myelin CNN in a local environment (90 x 90 x 100 nm^3) around each skeleton node of an axon. The node was declared to be myelinated if a threshold ($\Theta_{\text{myelin}} = 0.39$) was exceeded. To remove remaining small false positives, a running majority vote was applied along the skeleton (window length 4 μm). The probability maps of the synapse type classifier were used to assign a type prediction to each previously identified synaptic junction object by taking the ratio of the sum of class probabilities within the object, followed by determination of the optimal threshold (see Fig. 5.8) on the test set. Somata were identified by thresholding the soma and blood vessel probability maps ($\Theta_{\text{soma}} = 0.20$, $\Theta_{\text{bv}} = 0.08$), followed by binary closing (10 iterations) of the binarized blood vessel map, which was then used to mask false positive soma predictions.

[8] `https://github.com/StructuralNeurobiologyLab/SyConn`

Assessment of classifier performance

The quality of the extracted connected components (henceforth called objects) of synaptic junctions, vesicle clouds and mitochondria was assessed by visual comparison of the objects with the underlying raw data or independently generated annotations. More precisely, the recall of synaptic junctions and vesicle clouds was estimated by inspecting the locations of consensus synapses that were labeled by three annotators that followed skeletonized axons for predicted objects. Locations were compared in the same way as for mitochondria, where the coordinates were labeled independently and then later compared with the predicted objects. The precision was determined by randomly sampling and inspecting synaptic junction, vesicle cloud, and mitochondria objects from the entire j0126 dataset. Because of the ambiguity of whether a small contact is synaptic or not, false positives were only accepted when two experts agreed that a contact was synaptic. The soma precision and recall were determined in a similar way, predicted soma-locations were inspected for false positives and human annotated soma locations were compared to the predictions. The automated synapse type (symmetric vs asymmetric) prediction quality was tested on manually classified synapses, for which synapses of ambiguous type (defined by a lack of agreement between two experts) were excluded. The identification of myelinated parts of axons was evaluated on short manual tracings from myelinated axons (recall), and skeleton nodes that were predicted to be myelinated (precision) were compared to the raw image data to determine the correctness of the automatic assignment. All F scores were calculated as F1 scores, defined as the harmonic mean of precision and recall (Powers, 2011).

Conversion of skeletons to volume reconstructions

The CNN-predicted barrier probability map was used as input for a raycasting algorithm that sent out rays from a point along the manually traced skeleton while excluding a double-cone shape (20 rays, 18°) that was interpolated to a maximum node distance of 50 nm. The end point for each ray was defined by reaching a threshold ($\Theta_{\text{ray}} = 2.2$) on the cumulated barrier probability that the ray encountered on its way. Additional rays (19 instead of 9) were sent

out for skeleton end nodes, with a slightly adapted angle (135° instead of 30°) to increase the node density at endings of neurites. Outliers were then filtered to remove hull points with little support from others: those with a median neighbor-point distance >1.4 times greater compared to the other points which emerged from the same source skeleton node, and those that had less than 20 other points in their vicinity (220 nm radius).

Assigning objects to neurites

The generated ultrastructural objects (synaptic junctions, vesicle clouds and mitochondria) were assigned to neurite reconstructions by comparing object surface points with hull points from the neurite volume reconstructions using the scalar product (relative side of object surface point to a plane defined by the raycasting vector as normal vector, see section 5.1.2). For synaptic junction objects, it was tested whether they overlap with the neurite hull points.

Contact site detection and classification

Contact sites between neurite reconstructions were identified by measuring proximity between all hull points of all pairs of available neurites with a k-dimensional tree data structure for better computational efficiency. Hull points that were at maximum 60 nm away were considered, and all such hull points that were closer than 300 nm to each other were connected in an undirected graph. Connected component analysis was then performed to identify the hull points of individual contact sites, and the convex hull was used to calculate the contact site area (half of the convex hull surface area was used as estimate of the contact area). The proximity of the so-derived contact sites to the CNN predicted synaptic junction and vesicle cloud objects was then calculated, and contact sites with a predicted synaptic junction (<40 nm, <80 nm for vesicle clouds) were further examined. To avoid the manual tweaking of many heuristic parameters, a RFC was trained on whether contact sites constitute synapses based on various local features, as listed in table 5.1.

Table 5.1: Listing of used RFC features; shortened list of cell type features for brevity; consult the supplementary material of (Dorkenwald et al., 2017) for details. Histogram features were calculated from a breadth-first-search starting at each node and stopping after 6 µm.

RFC	**Features**
Synapse filter	Mean contact area Absolute and relative overlap with synaptic junction (in μm^2) Minimum hull distance
Spines	Mean, s.d. and count of 10 histogram bins of node radii Number and mean size of mitochondria, vesicle clouds, synaptic junctions Distance to next branch node Distance to next end node
Compartments	All spine features plus: Number of spine heads Mean and s.d. of spine head size Mean of spine head probability
Cell types	17 derived features, from predictions of the other RFCs 20 global neural shape descriptors 12 statistics of assigned ultrastructural objects 8 synapse type based features

RFC prediction of neurite compartments and cell types

For the prediction of neurite compartments (axons, dendrites, spines and somata), features (see Table 5.1) were computed for local environments around skeletons nodes, and RFCs were trained on the different classifications tasks. Predictions were made on a skeleton node level, and the results were evaluated on a set of manually annotated nodes. For cell types, the feature set was extended with global neurite properties to improve the classification (see Table 5.1 for an overview).

Connectivity matrix construction

Based on the contact sites that were predicted to be synaptic, a connectivity matrix was constructed, sorted by the automatically inferred cell types. The predicted synapse types were used to assign for each neurite the outgoing synapse type by majority voting according to Dale's rule (Strata and Harvey, 1999). The column of the incoming synapses for the excitatory axon (EA) cell class was omitted, since no dendrites were found for these neurites (external cortical afferents to the basal-ganglia, but see Budzillo et al., 2017 for a rare excitatory cell type located inside of Area X). Individual matrix elements were rendered slightly enlarged (3 by 3 pixels)

for improved visibility.

5.2 Results

5.2.1 GraphWalker: A segmentation proofreading workflow

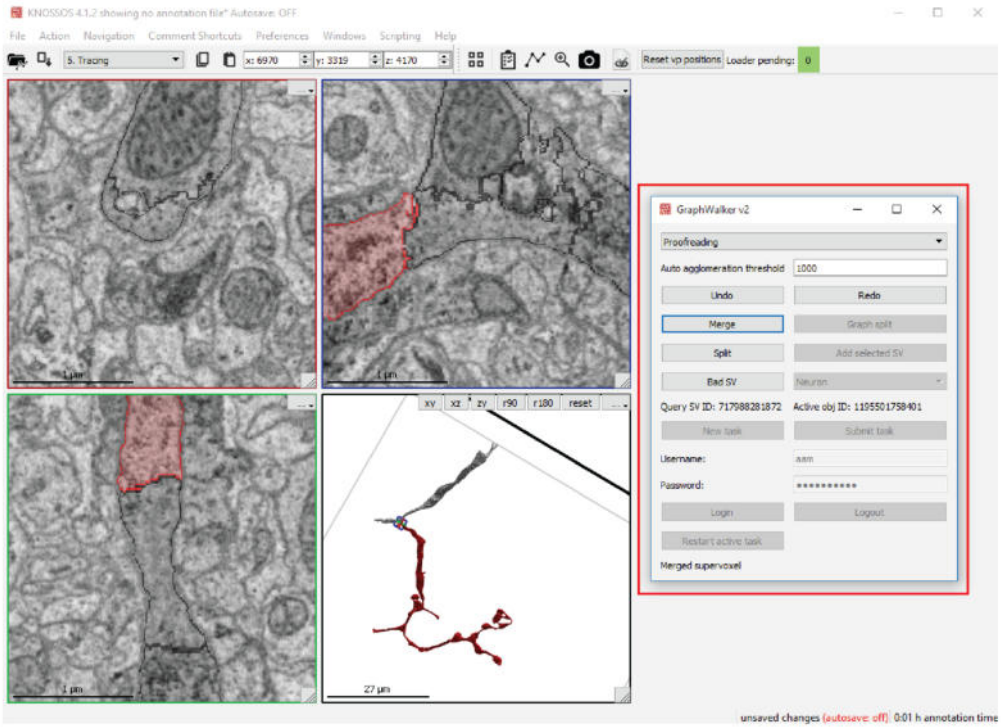

Figure 5.1: *Graphical user interface of GraphWalker plugin. Shown are on the left side the four KNOSSOS data viewports (XY, XZ, YZ and the 3D view), and located on the right, the GraphWalker v2 control panel (red box), currently in proofreading mode. Note that 2D and 3D renderings (supervoxel meshes) of the same neurites are visible to the annotator for decision making.*

KNOSSOS[9] (Kornfeld et al., 2011; Helmstaedter et al., 2011), originally developed for fast skeleton tracing (Helmstaedter et al., 2011), was extended with a Python plugin (GraphWalker) that adds a novel proofreading workflow to KNOSSOS. M. Januszewski contributed important ideas to the development of GraphWalker, as part of a scientific collaboration with Google Research. A short overview about GraphWalker is provided in the next section, followed by

[9] `https://knossostool.org/`

an illustration of the reconstruction principle and an evaluation of the accuracy and speed of neurite reconstructions generated with GraphWalker.

The KNOSSOS graphical user interface (GUI) was extended by a small window (see Fig. 5.1), the GraphWalker front end. It allows the user to control and perform proofreading of an existing automatic segmentation, such as the j0126 v3 segmentation that was supplied by Google. All data shown and analyzed in this chapter are based on this segmentation, but GraphWalker is in principle compatible with all automatic segmentations that provide a supervoxel graph of an EM dataset where probabilities of supervoxel edges are available. The workflow is organized in three phases, reflected by separate operation modes that change the GUI:

1. **Proofreading**, during which the annotator actively grows the reconstruction.
2. **Review**, during which the annotator inspects the reconstruction and can correct errors.
3. **Task management**, during which the annotator requests a new task and submits the task to the GraphWalker back end.

The GraphWalker back end communicates with two separate entities. Firstly, a version of the Heidelbrain annotator management server (programmed originally by F. Svara), which was extended to record the work time of GraphWalker annotations and store them in a graph database[10]. Second, the Google Brainmaps API provides 3D meshes to display in KNOSSOS for version 2 of GraphWalker (version 1 was based on skeletons automatically computed from the segmentation). Fig. 5.2 illustrates the architecture of GraphWalker and shows how the different components interact.

The separation into front- and back end allows GraphWalker to run with moderate hardware requirements on standard laptop computers, with essentially the same hardware requirements as KNOSSOS. Since all data is stored remotely (raw image data, automatic segmentation and the supervoxel graph), the maximum dataset size that can be proofread is mainly limited by the scalability of the back end hardware used. A single virtualized Linux server proved to be

[10] see `https://neo4j.com/`

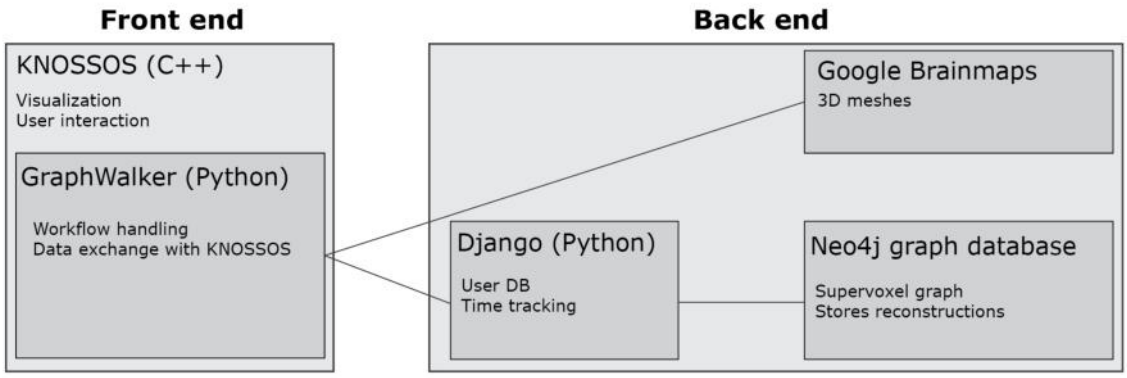

Figure 5.2: *Front- and back end of GraphWalker. The KNOSSOS Python plugin interacts with the Google Brainmaps API and a custom Python Django back end. All data are stored in a highly-scalable graph database.*

more than sufficient for the j0126 dataset with the Google v3 segmentation. All employed back end software components additionally support distribution to multiple compute nodes and can therefore run on scalable cloud computing infrastructure such as Amazon's AWS, making it possible to reconstruct EM datasets with this approach that are limited in size only by the available storage.

Neurite reconstruction principle

During conventional skeletonization, neurites are followed through the dataset and nodes are placed by the annotator to reconstruct the center line of the cells, thereby building a skeleton reconstruction of the neuron step-by-step (Helmstaedter et al., 2011). This approach is fundamentally limited in reconstruction speed by the need to inspect all of the raw data locally. Therefore, GraphWalker takes a different approach that naturally translates advances in automated segmentation into higher reconstruction speeds. Instead of following neurites through the dataset, the annotator extends a neuron reconstruction under visual control by deciding whether a presented automatically segmented cell fragment (i.e. supervoxel) should be added to the already reconstructed part of the cell, as illustrated in Fig. 5.3. Hence, GraphWalker is in principle a binary decision process, that additionally provides feedback to the annotator about the global consequences of each decision. This should be contrasted with independent

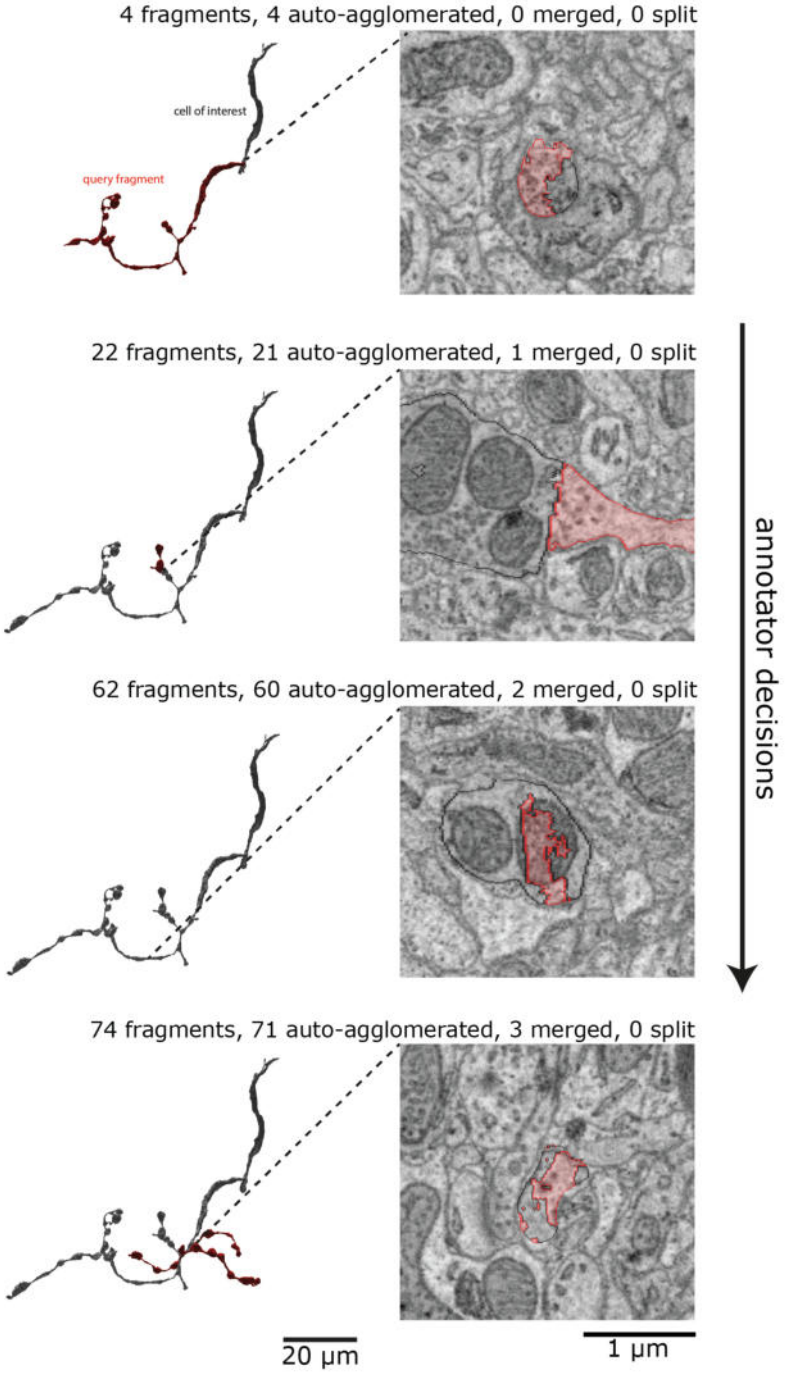

Figure 5.3: *Example GraphWalker reconstruction of a myelinated axon, showing the first four manual decision steps and how the neurite is extended. Left column: 3D visualization of the neurite, with the current query fragment in red and the existing reconstruction of the cell of interest in gray. Note how the auto-agglomeration dramatically reduces the number of decisions by adding small fragments automatically.*

binary decisions that hide the reconstruction progress of a neuron from the annotator. In principle, such an approach could also lead to successful cell reconstructions; however, this was

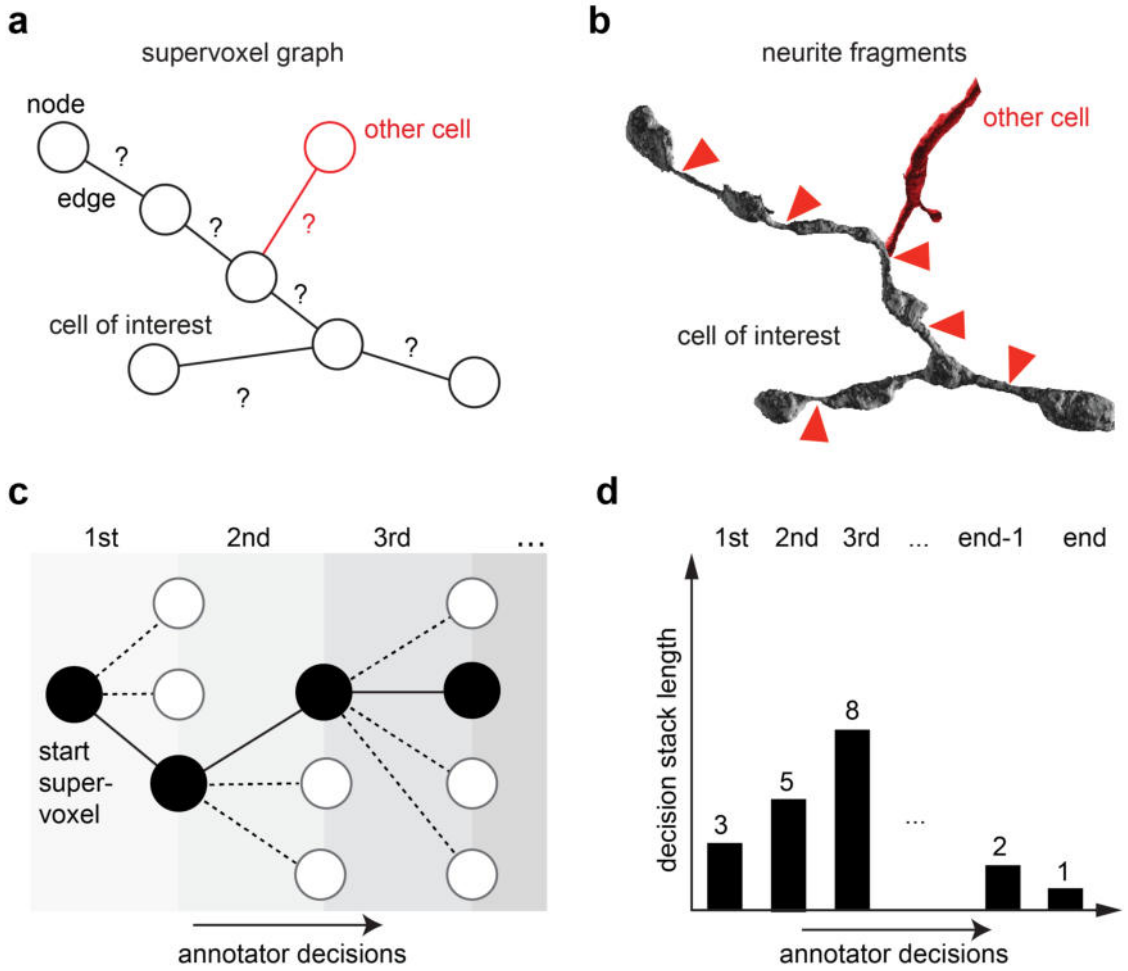

Figure 5.4: *GraphWalker reconstruction principle. (a) Schematic of a supervoxel graph. Nodes represent supervoxels, edges between two supervoxels indicate that they might belong to the same neurite. The black subgraph belongs to the cell of interest that should be reconstructed, but one supervoxel that is connected by an edge belongs to a different cell. (b) Visualization of neurite mesh renderings of the supervoxel graph in (a), with edge locations indicated by red arrows. Note how the red supervoxel deviates morphologically. (c) Schematic of first decisions of an annotator, "walking" along the supervoxel graph. (d) Illustration of GraphWalker decision stack length over the reconstruction process of a cell. Initially, the stack size increases, until the annotator slowly exhausts it.*

found to be problematic due to propagating errors, even when multiple redundant decisions were used (Nguyen, 2014)[11]. Additionally, GraphWalker allows the user to control an auto-agglomeration threshold that adds supervoxels with a high merge probability up to a cumulated size limit (neurite length) without a time-consuming user query, while still allowing the user to undo false decisions introduced this way. This feature was found to be particularly useful since the tested segmentation contained many small neurite fragments that almost always belonged

[11] see `http://brainpuzzler.org/` for the project website

to the neurite of interest. This was a consequence of the intentionally high split-rate to avoid merge errors on the level of individual supervoxels. Fig. 5.4 a, b shows a schematic of how a neurite could be subdivided into fragments that are linked by edges in a supervoxel graph. The annotator then decides for each edge whether it should be added to the cell of interest in consecutive decisions that expand, and eventually exhaust, the edge decision stack (Fig. 5.4 c, d). Algorithm 1 shows the basic principle as pseudo-code, excluding the auto-agglomeration for simplicity.

Algorithm 1 Basic GraphWalker reconstruction algorithm.

```
Require: Initialize a decStack                ▷ The decision stack stores upcoming decisions
procedure GraphWalker(startSV, SVgraph, decStack)
    SVchildren ← query SVgraph with startSV
    push [startSV, SVchildren] onto decStack
    while decStack not empty do               ▷ Repeat until decision stack is exhausted
        currDec ← pop decStack
        querySV ← currDec[1]
        if currDec is doMerge then            ▷ Manual merge decision by user
            SVchildren ← query SVgraph with querySV
            push [querySV, SVchildren] onto decStack
        else                                  ▷ Manual split decision by user, next decision follows
            pass
    return
```

Reconstruction quality

To evaluate the quality of the reconstructions that were created with GraphWalker based on the v3 segmentation, GraphWalker neurites were compared to a test set of ten ground truth skeletons (redundant annotations, followed by a consensus generation), and to independent expert skeletons of the same cells. Fig. 5.5 a shows three of these test neurites (one axon and two medium spiny neurons) in direct comparison with the corresponding ground truth skeletons and the expert skeleton, illustrating that GraphWalker can be used to reconstruct a diverse set of neurites. The two main error categories, added neurite fragments that do not belong to the cell of interest, and missed neurite fragments, were analyzed separately. Fig. 5.5 b is an evaluation of the missed neurite fragments, and shows mixed results for the fraction of ground truth skeleton path length that could be recovered with GraphWalker: test neurite 3 and 10

were completely reconstructed, while the majority of path length for test neurite 2 is missing (see Fig. 5.5 c). Closer inspection revealed that the supervoxel graph was missing the edges required to complete the reconstruction of this thin axon. This problem illustrates the danger of pruning low merge probability edges to reduce the amount of decisions and proofreading work: thin neurites appeared to be particularly affected, an error that is difficult to spot.

The number of false neuron mergers (either due to undiscovered errors during the auto-agglomeration, false manual decisions or false mergers that already existed in the unmodifiable supervoxels) was quantified and is shown in table 5.2. Interestingly, and in contrast to such errors generated by fully automatic reconstruction results, almost all false mergers remained small, i.e. did not merge large parts of a different cell. This can likely be explained by a higher discovery probability for larger incorrectly added fragments by human annotators.

Table 5.2: Qualitative evaluation of the number and length of false mergers (i.e. incorrectly added branches that do not belong to the test neurites). Small mergers: <10 µm; Large mergers: >10 µm and <50 µm.

Test neurite	**Path length** [µm]	**Small mergers**	**Large mergers**
1	527	0	0
2	145	0	0
3	178	1	0
4	111	1	0
5	47	1	0
6	1165	4	0
7	882	3	1
8	764	0	0
9	97	2	0
10	199	1	0
Total	4,115	13	1

It is important to note that not all false merge errors can be discovered in this way, and that errors could remain undiscovered by relying purely on their identification through a suspicious shape in the 3D overview. As an example, tip-to-tip axon mergers can mainly be spotted through local image data inspection, or, in a dataset that includes both somata, by resulting global discrepancies. It remains to be determined to what extent such difficult to discover errors would affect the overall reconstruction quality of a dataset. Additionally, the presented results showed a somewhat unexpected error category: false splits due to missing edges in

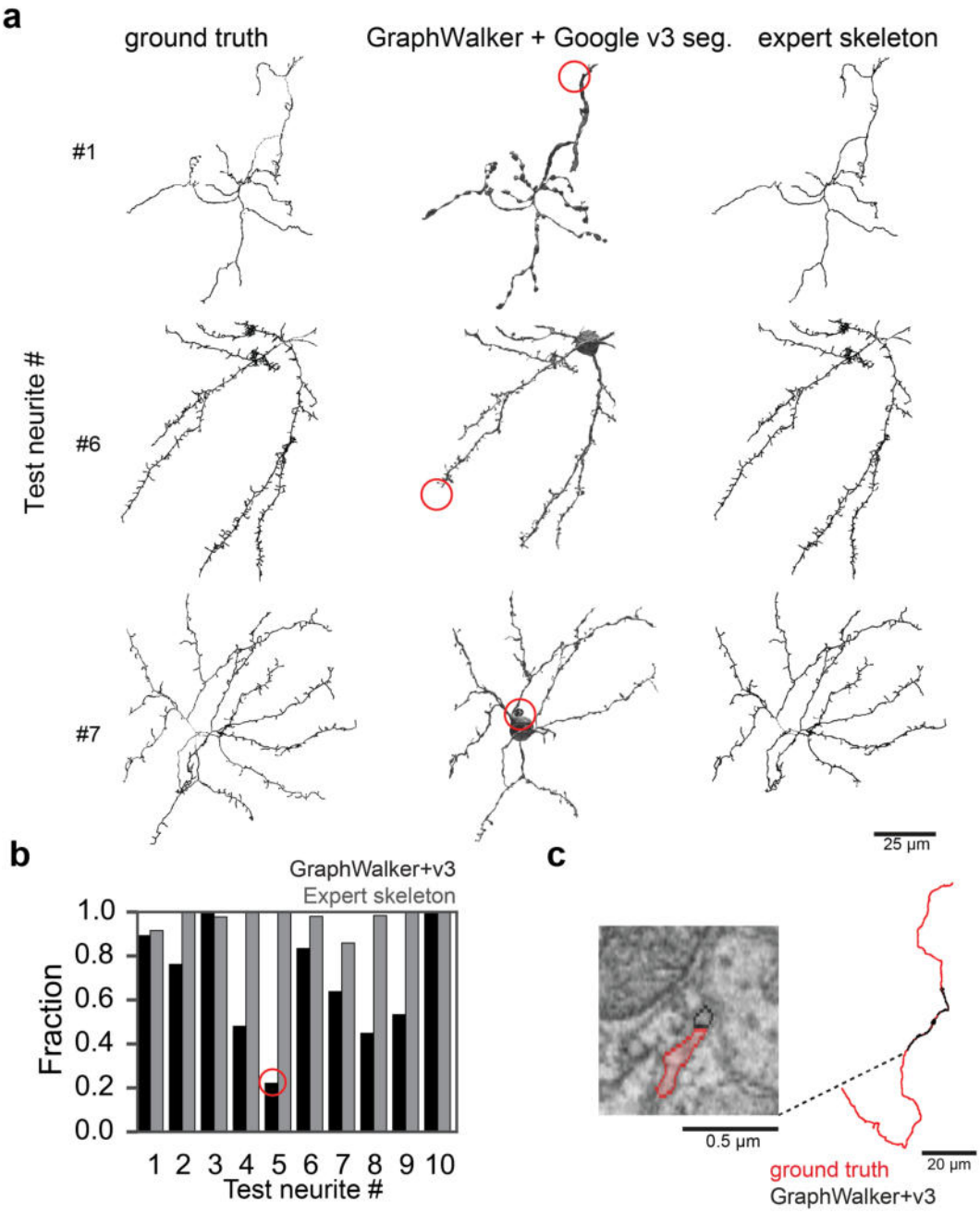

Figure 5.5: *GraphWalker evaluation against ground truth skeletons. (a) Example GraphWalker reconstructions (middle column), compared to ground truth skeletons (left column) and to a single redundancy expert skeletonization (right column). Red circles highlight exemplary reconstruction failures (merger or split). (b) Path length of ground truth skeleton recovered by expert skeletons and GraphWalker reconstructions. (c) Shown is the skeleton reconstruction with the least recovered path length from b (red circle), a thin axon. The inset shows the region (xy-view) in the j0126 dataset with the overlaid supervoxels.*

the supervoxel graph (see Fig. 5.5). These errors are difficult to fix, since they can only be discovered in an active-search manner, instead of the relatively passive approach of deciding on

supervoxel edges that are presented.

Reconstruction speed

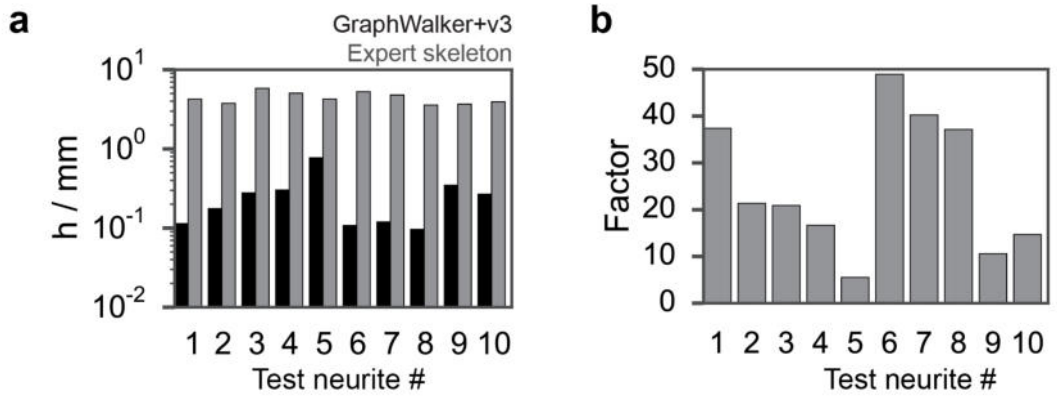

Figure 5.6: *GraphWalker reconstruction speed evaluation. (a) Absolute reconstruction speeds in h per mm path length of ground truth skeleton, for expert skeletonization and GraphWalker reconstruction (performed by an expert) for the 10 test neurites. (b) Relative speed-up compared to expert skeletonization.*

The reconstruction speed was compared to that of an expert skeletonizing the same set of cells, in order to avoid sampling bias: some neurons can be reconstructed faster than others due to easier morphology, and some annotators trace cells faster than others. Fig. 5.6 a shows the achieved speed for expert skeletonization (mean 4.4 h / mm, n = 10, consistent with previous reports, Helmstaedter et al., 2011; Wanner et al., 2016) and GraphWalker (mean 0.26 h/mm, n = 10) as hours per path length of ground truth neurite. This leads to an average speed-up of about 25x over skeletonization (see Fig. 5.6 b), which is, to the knowledge of the author, the first written report of a reconstruction method that outperforms manual skeletonization significantly (Berning et al., 2015). Closer inspection shows that the speed for test neurite 5, the thin axon shown in 5.5 c, was the lowest measured. This can likely be explained by the overhead of starting the task and the short fragment that could be recovered. The two test neurites that achieved the highest speed-ups (6 and 7, see Fig. 5.6) are complex neurons with many spines, demonstrating that the GraphWalker approach can perform well on cells with highly branched neuronal morphologies. Despite these significant advances, it should

be noted, that a reconstruction method must prove itself ultimately on the reconstruction of an entire dataset (such as skeletonization of retina data Helmstaedter et al., 2013), and not only on a small test, rendering the presented results somewhat preliminary. It is also less clear whether redundant annotations would reduce the error rate significantly, given that the presented decisions in the current process are correlated and that edges were missing in the supervoxel graph.

5.2.2 Synaptic connectivity inference from neurite skeletons

Since GraphWalker was a relatively recent development, all of the annotations that were created for this thesis were manual skeleton reconstructions. While these provide morphological information, they are insufficient by themselves to analyze the wiring of a brain circuit due to a lack of synapse information. While synapses can also be annotated manually, as described in the previous chapters, this annotation is as about as time consuming as the neurite tracing itself (see Fig. 1.1). Therefore, a computational framework was developed that can infer a synaptic connectivity matrix automatically by using deep convolutional neural networks (based on the ELEKTRONN library[12] (Killinger, 2016); M. Killinger and G. Urban, supervised by the author of this thesis) and classical feature-based machine learning (random forest classifiers, RFC). As detailed in the introduction (see section 1.1.3, Fig. 1.1), multiple analysis challenges have to be solved for the extraction of a comprehensive wiring diagram from a volume EM dataset of brain tissue. SyConn addresses all but the first step, the skeletonization of neurites, thereby eliminating about half of the manual annotation effort otherwise required for dense connectomic analysis. The first step of SyConn is the conversion of skeleton reconstructions to volume reconstructions, a prerequisite for the reliable association of synapses with skeletons.

Converting skeleton to volume reconstructions

Since there is normally no information about the caliber of axons and dendrites collected during skeletonization (due to the manual effort), synaptic connections cannot easily be identified

[12] `http://elektronn.org/`

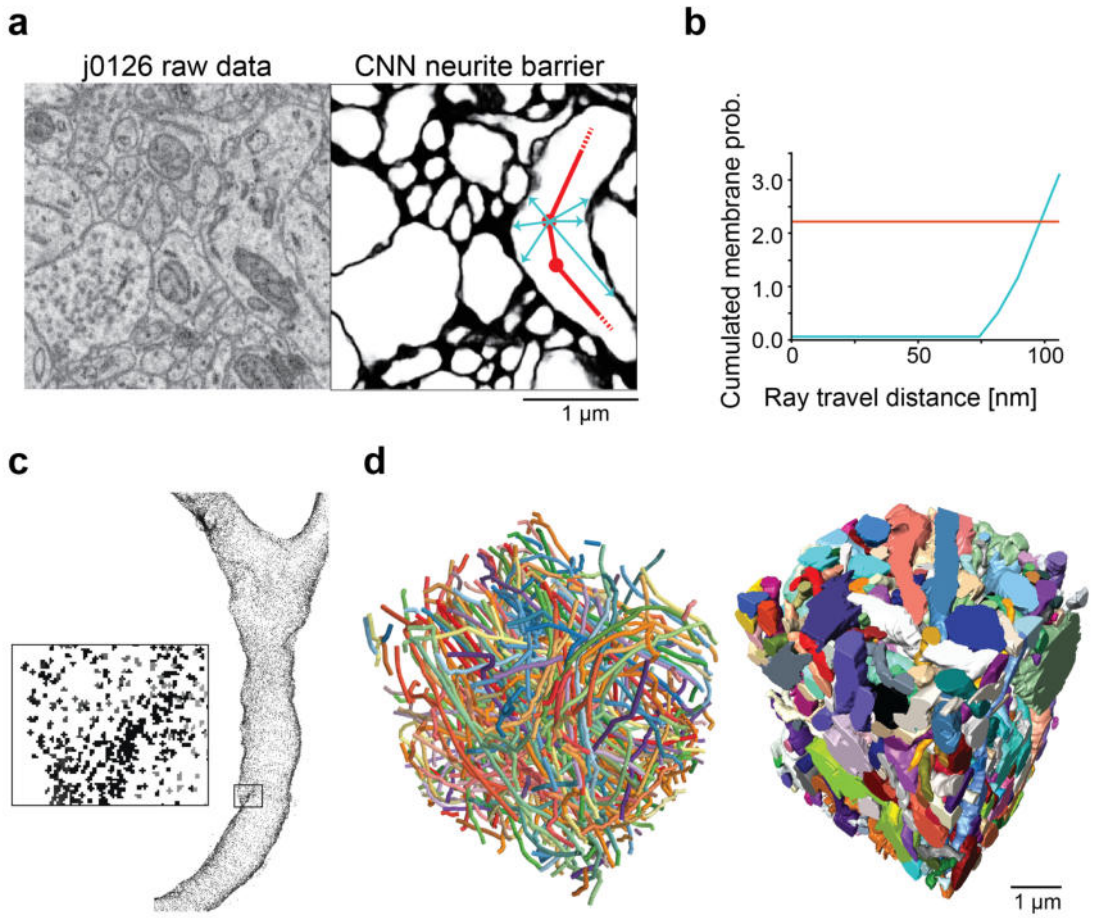

Figure 5.7: *Conversion of skeletons to volume reconstructions. (a) Left: raw image data of j0126 (microscope imaging plane), right: predicted neurite barrier, cell interior in white, membranes and extracellular space in black. (b) Cumulated barrier probability plotted against the traveled path length of a single, example ray. The red line indicates the threshold for ray termination. (c) Exemplary rendering of the detailed morphology of a skeletonized part of a dendrite, combined with raycasting to reconstruct its 3D shape. (d) Dense skeletonization of a small test cube (left), with raycasting applied (right). Figure partially adapted from (Dorkenwald et al., 2017).*

from skeleton proximities in a dense reconstruction setting, as is illustrated with the following thought experiment for which only skeletons shall be available. To identify all synapses through the inspection of skeleton proximity sites, permissive distance thresholds have to be used (ultimately depending on the maximum thickness of the neuronal processes) to avoid false negatives of the nearby neurons. However, this leads to a large number of possible contacts that require inspection (growing with the number of reconstructed neurites), making this approach unfeasible for the dense reconstruction of EM datasets. This number can be dramatically reduced when the exact boundaries of the neurites are known, which is also a prerequisite for the association of automatically identified synapse objects to the neurites. Therefore, the first

step of SyConn consists of a CNN that predicts the interior regions and the barrier regions (membrane regions + ECS, see Fig. 5.7 a) between neurites, which is an established approach (Berning et al., 2015; Pallotto et al., 2015; Helmstaedter et al., 2013; Beier et al., 2017). Rays emitted from the skeleton nodes then terminate in the barrier, essentially sampling the hull of the skeleton reconstruction (Fig. 5.7 b, c). Fig. 5.7 d shows the result of applying this method to a densely skeletonized cube of 125 μm^3 in j0126. Using this approach, 612 manual skeleton reconstructions were converted to volume reconstructions and used for the analyses described in the following.

Automated ultrastructural analysis with CNNs

SBEM datasets from conventionally stained tissue blocks make many ultrastructural details visible (see Fig. 5.7 a), and show essentially all cell biological features (Hughes et al., 2014), in some cases even electric synapses (Pallotto et al., 2015). While these data could be analyzed in principle, ultrastructure was often neglected in existing connectomic studies (Helmstaedter et al., 2013; Wanner et al., 2016), either because a staining procedure with a different focus was used (Helmstaedter et al., 2013) or because analysis was not feasible or seemed not necessary (Wanner et al., 2016). Multiple CNNs were therefore designed and implemented to identify ultrastructure automatically with computer vision, by first generating probability maps for synaptic junctions, vesicle clouds and mitochondria (all three combined in a single, multichannel output CNN; Fig. 5.8), synapse type (Fig. 5.8 c to e) and myelin (Fig. 5.8 f). The synapse type was predicted as either symmetric or asymmetric, based on the observations of Colonnier (Colonnier, 1968) and Gray (Gray, 1959). These probability maps were then used to extract discrete "objects" by thresholding and connected component analysis (except for the synapse type probability map, which was used to assign a type to each synaptic junction object). Having such an object representation made it easier to then combine these data with the neurite volume reconstructions, as described in the following section.

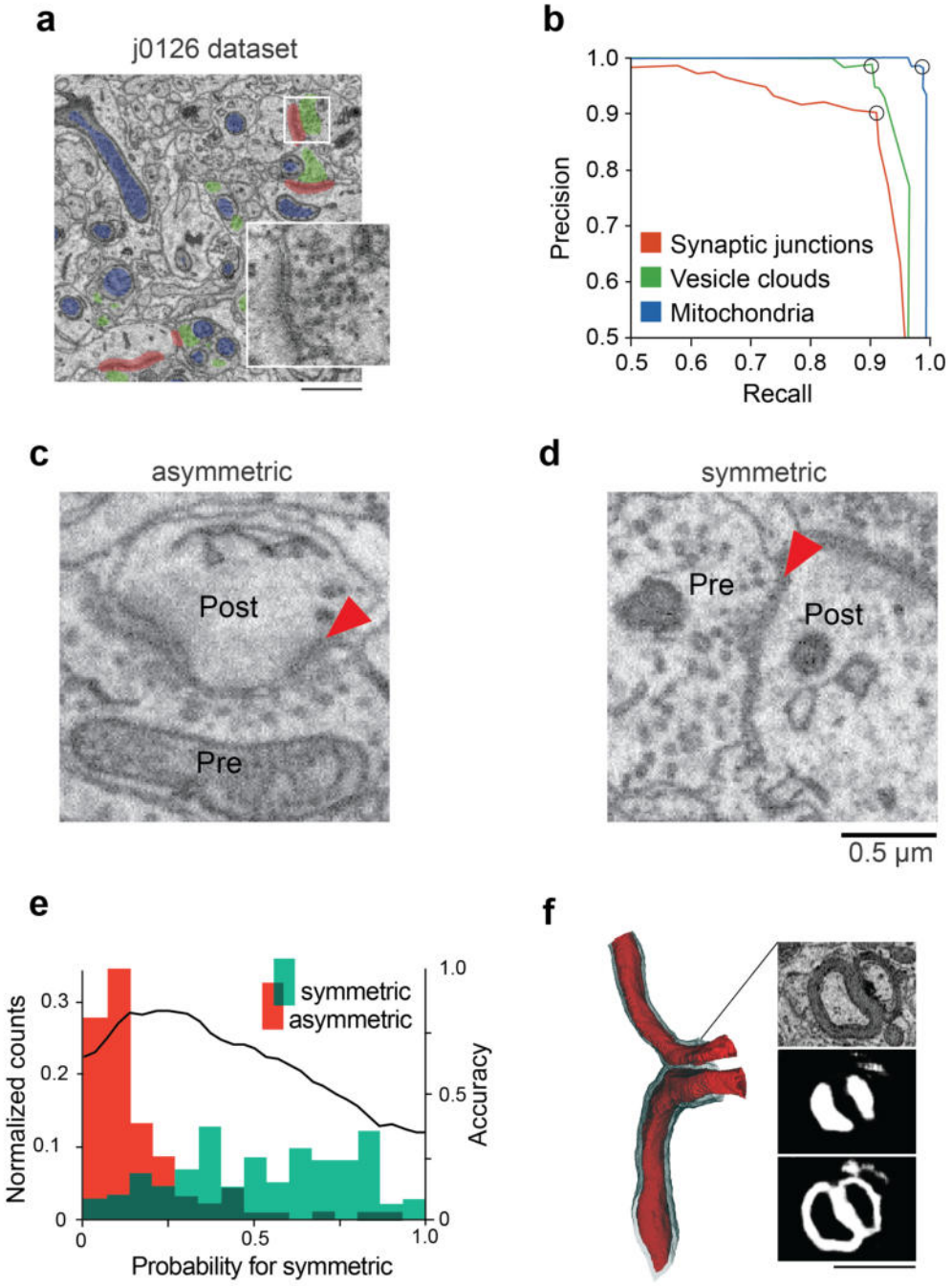

Figure 5.8: *Detection of ultrastructure with CNNs in raw data. (a) EM micrograph with detected mitochondria (blue), synaptic vesicle clouds (green) and synaptic junctions (red). The inset shows an asymmetric synapse that was identified by the classifier. (b) Precision-recall curve for the aforementioned classification tasks, the optimal F1-score is encircled. The small jitter of the curve is the result of the discrete ground truth set. (c,d) High-resolution asymmetric and symmetric synapse example of this dataset. Note that it is easier to identify asymmetric synapses in the j0126 dataset compared to the j0256 of HVC, see Fig. 3.3. (e) Classification of individual synaptic junctions by type. Shown is a normalized histogram of the number of voxels sorted by their inferred probability of being classified as symmetric for a test set of synaptic junction objects. The colors indicate the ground truth labels. The black line shows the synapse-level classification accuracy for a varied threshold. (f) Visualization of two myelinated axons (left), rendered with Amira from the probability maps for the myelin-interior class (middle panel) and the myelin class (bottom panel). The top panel shows the raw image data for comparison. Scale bar in a and f is 1 µm. Figure partially adapted from (Dorkenwald et al., 2017).*

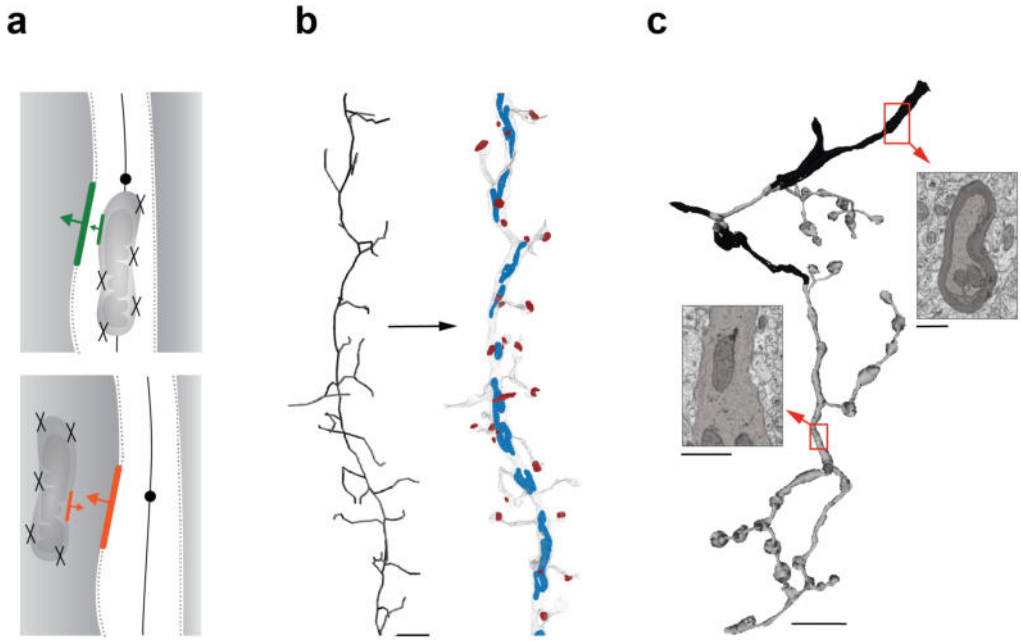

Figure 5.9: *Assignment of detected ultrastructure to reconstructed neurites. (a) Random points of the detected ultrastructural objects were compared to random points sampled from the neurite hull via the sign of the scalar product. (b) Exemplary spiny neurite. Raw skeleton tracing on the left, and enriched neurite with hull on the right. (c) Visualization of myelinated regions (in black) of an axon reconstruction. Scale bar in b 1* µm*, in c 10* µm *and 1* µm *for the micrographs. Figure partially adapted from (Dorkenwald et al., 2017).*

Enriching neurite skeletons with ultrastructural objects

To assign the previously identified objects to reconstructed neurites, random voxel samples from the ultrastructural objects were compared with a set of closest hull points from the neurite (see Fig. 5.9 a). Using this approach (section 5.1.2), excellent object-to-neurite assignment performance was achieved (F1 scores of 0.99, 0.96, and 0.96 for synaptic junctions, vesicle clouds and mitochondria, respectively; tested on a set of 22,167 human generated assignments). All 612 neurite reconstructions of the j0126 dataset were then augmented in this way (Fig. 5.9 b, c). It was realized during inspection of the resulting neurite library that different neurites and their cellular compartments (i.e. axons, dendrites and soma) had striking differences in these properties, opening the door to exploit these features for a further automatic classification approach that targeted the heretofore manually performed identification of neuronal compartments and cell types.

Computational identification of neuronal compartments and cell types

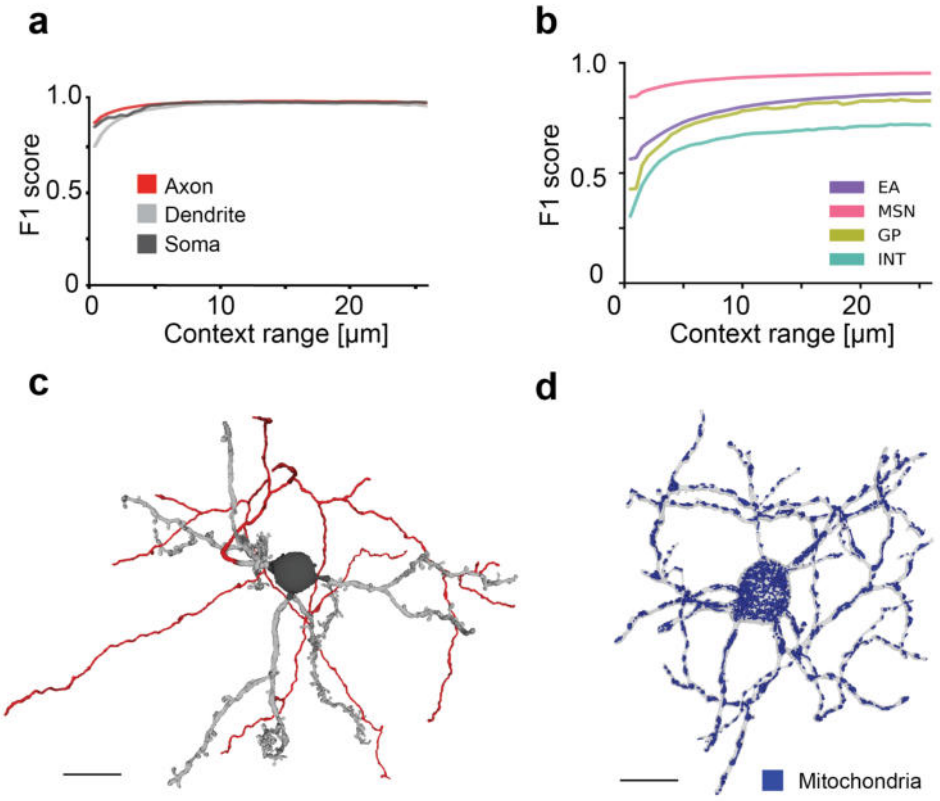

Figure 5.10: *Classification of neurites based on morphology and ultrastructure. (a) Classification performance for neuronal compartments as a function of local neurite context. (b) Same as in a, but for the four main cell types in Area X. (c) Example of an MSN, with different parts color coded by their automatically inferred identity (colors as in a). (d) Pallidal-like neuron with all mitochondria mapped. Scale bar in c and d is 10* µm. *Figure partially adapted from (Dorkenwald et al., 2017).*

The ultrastructural features were then combined with local morphological properties of the neurites (e.g. caliber and variance of caliber, for more details see Table 5.1), and an RFC was trained to label individual compartments of the reconstructed neurons as being part of the axon, dendrite or soma, and on a finer scale, dendritic spine or shaft. This approach could demonstrate that relatively local (see Fig. 5.10 a and b) neurite features are sufficient to infer semantic neurite identity. Fig. 5.10 c shows a reconstruction of an MSN with its compartments automatically labeled, highlighting how strongly the axon and dendrites differ in overall morphology. Different compartments varied also in their ultrastructural properties, as shown in Fig. 5.10 d for mitochondria: for this cell type, the soma is densely filled, while the neurites of the cell appear to have more clustered occurrences. These observations were

then used to train an additional RFC with an extended feature set (Table 5.1) to automatically classify the neurites into cell types. Since the number of reconstructions was limited, and some Area X cell types are rare (Goldberg and Fee, 2010; Goldberg et al., 2010), leading to a lack of example cases, classification was restricted to four main types: excitatory axons (EA), medium spiny neurons (MSNs), interneurons (INT), and pallidal-like neurons (GP). It was possible in many cases to identify the cell type again solely based on a local feature context (Fig. 5.10 b), but global features were used for the best performance (F1 scores MSN 0.97, EA 0.99, GP 0.92, INT 0.82; evaluated using leave-one-out on a manual test set of 456 cells). The lowest performance values were obtained for the rare cell types (GP and INT), suggesting that a lack of ground truth for these cases might affect their classification.

Computation of a synaptic wiring diagram of Area X

The volume reconstructions of the neurites were then used to extract contact sites between cells (precision: 0.92, on n = 2,488 predicted contact sites and recall: 0.92, evaluated on n = 53 manually identified sites; all false negatives were small tip contacts), and these were then further classified as synaptic or non-synaptic with an RFC to remove false positives. This additional RFC was necessary due to the many edge cases that occur in a large and diverse dataset, that would have required extensive manual threshold tweaking if neurites were assigned to identified synapse objects with a simple heuristic. A synaptic connectivity matrix was then constructed, sorted by the predicted cell types (see Fig. 5.11). The synapse detection performance was evaluated on a sample of 2305 manually classified contact sites, and found to be 0.88 precision, and 0.82 recall.

The biological plausibility of the computed connectivity matrix was then assessed by comparing it to known pathways of the zebra finch and other vertebrate basal-ganglia, since the basal-ganglia are an evolutionary highly conserved brain structure (Grillner and Robertson, 2016). The main input pathways to the striatum are the numerous excitatory cortical afferents (Grillner and Robertson, 2016), found for Area X using electrophysiology (Luo and Perkel, 1999; Leblois et al., 2009). The dominance of these connections can also be seen in the connectivity

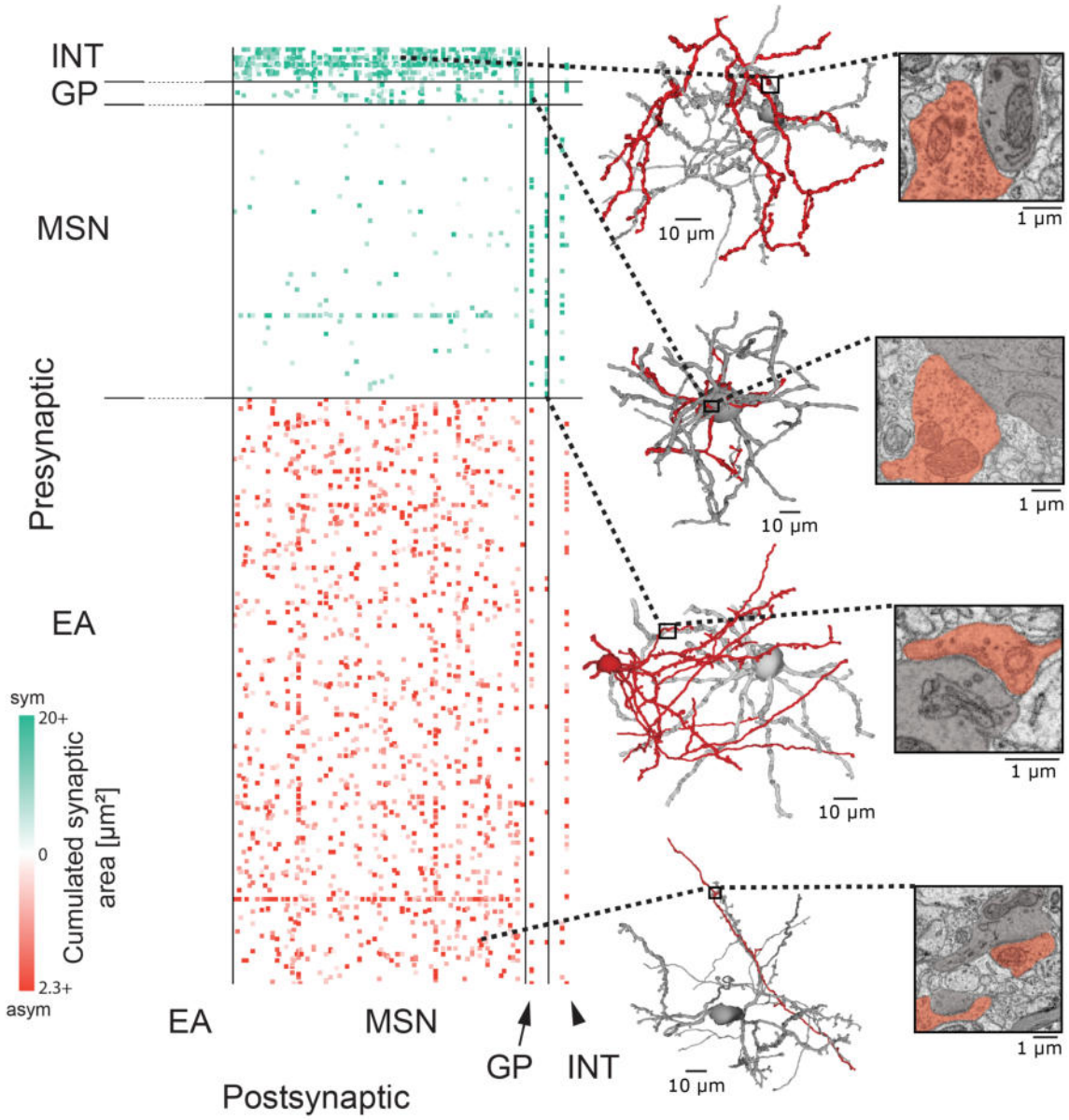

Figure 5.11: *Inferred synaptic wiring of the j0126 dataset, based on 612 skeletonized neurites that were analyzed with SyConn. The insets on the right show EM images of example synapses with pre- and post synaptic neurites that were automatically identified, next to 3D renderings of the outgrown skeleton reconstructions. The axon is shown in red, the postsynaptic cell in gray. Note the diverse ultrastructure of the different synapses. Figure partially adapted from (Dorkenwald et al., 2017).*

matrix, with 1443 EA to MSN synapses, consistent with the electrophysiological data. MSN to MSN connectivity was sparse, and only few synapses were found (127), in agreement with reports from the mammal (Oorschot et al., 2013; Jaeger et al., 1994). This is in stark contrast to the INT to MSN connectivity, as becomes immediately apparent from the matrix: the rows of the identified INT axons are densely populated, and individual INT axons form synapses with

a large set of the reconstructed MSNs. Given the small number of INTs, this predicts a high pairwise INT to MSN connection probability, again consistent with findings in the mammalian striatum (Koós and Tepper, 1999). The second order connectivity implication of this is a strong feed-forward inhibition of MSNs through INTs, since INTs are also driven by EAs (note that the few reconstructions contained almost no INT dendrites), as reported previously (Bennett and Bolam, 1994). Since Area X also contains pallidal-like neurons (Luo and Perkel, 1999; Farries et al., 2005), it becomes possible to follow the entire synaptic pathway through the basal-ganglia with the connectivity matrix in Fig. 5.11. Despite the presence of only two reconstructed GP dendrites, many of the MSNs were found to innervate them, consistent with the convergent innervation from the striatum to the pallidum (Grillner and Robertson, 2016; Alexander and Crutcher, 1990). Interestingly, the two GPs did not receive the same fraction of synapses from other GPs (4.5 % vs 48.2 %, compared to total synaptic input). While highly speculative, given the sample size, this observation could indicate that the two pallidal-like neurons are part of the direct and indirect pathway (Farries et al., 2005). The described findings support the notion that the automated identification of synapses and the construction of a wiring diagram from skeletons is sufficiently advanced not to require additional human proofreading, but it became also clear that analyses based on few (612, compared to an estimated number of 10,000 to 20,000 neurites in the j0126 dataset) reconstructions are problematic, especially for rare cell types.

5.3 Discussion

5.3.1 GraphWalker: A segmentation proofreading workflow

The presented workflow was developed on the basis of an automatic segmentation of the j0126 Area X dataset that was generated by Google Research with CNNs, followed by GALA(Nunez-Iglesias et al., 2014). GraphWalker requires a dataset that is fully segmented into supervoxels, i.e. 3D regions that have a very high probability of belonging to the same object. The relationships of these supervoxels can further be represented in an undirected graph, in which

each supervoxel is a node and nodes that might belong together are connected through edges. A proposal segmentation is then defined by subgraphs. In its current implementation, GraphWalker requires weighted edges, representing estimated supervoxel equivalence probabilities. These weights were only used to order the query supervoxel presentation to the user, and are therefore not essential. However, efficient proofreading (in both reconstruction quality and speed) has multiple requirements towards the underlying automatic segmentation:

1. Automatically segmented neurite fragments should be as correct as possible, with essentially no false mergers on the level of supervoxels.

2. The fragments should be as large as possible, without violating 1).

3. The number of proposed merge decisions should be small, i.e. only fragments that likely belong to the currently reconstructed neuron should be presented.

4. The number of missing edges in the supervoxel graph should be small, i.e. only fragments belonging to the same neuron should be linked by an edge, without violating 3).

These requirements are conflicting, and only an error-free automatic segmentation could satisfy all together. Since this is currently not in sight, it is important to realize that the sole optimization criterion for any automated procedure should be that it allows efficient proofreading (i.e. a reduction of the "nuisance metric"; Arganda-Carreras et al., 2015). This holds true up to the point that proofreading appears completely unnecessary, because remaining error rates are deemed acceptable. The presented results illustrate that the GALA segmentation, together with the applied GraphWalker proofreading procedure, remains problematic. While a significant speed-up was achieved for the path length that was recovered correctly, further experiments are required to demonstrate that the proposed workflow can be scaled to an entire dataset and is able to lead to a dense segmentation that can be considered final.

5.3.2 Synaptic connectivity inference from neurite skeletons

The massive growth of the field of machine learning, triggered by the recent success of deep neural network architectures in solving real world problems (LeCun et al., 2015), also helped the development of the SyConn framework (Dorkenwald et al., 2017): one of the key components are the ultrastructure classifying CNNs that are based on ElektroNN (Killinger et al., 2015; Killinger, 2016), a high-level deep learning library tailored to the needs of 3D image processing that builds on the Theano framework (Theano Team et al., 2016). While increasingly complex CNN models have won essentially all image classification competitions over recent years[13], algorithms such as random forest classifiers still provide excellent performance for feature-vector based classification problems (Breiman, 2001), and both approaches were therefore combined in SyConn. The many different machine learning methods available now can essentially be seen as a toolkit that enables the automated analysis of volume EM datasets: for identification problems in raw image data, algorithms based on an at least partially convolutional neural network architecture are likely the tool of choice, while problems derived from these data, e.g. feature-vector based problems, can still be addressed with more conventional algorithms (such as RFCs or other ensemble methods). SyConn is a demonstration of this principle, and as such deviates from the currently favored end-to-end single model approach often taken in machine learning research today. Instead of attempting to solve all problems with a single, very powerful model, problems were addressed as they came up through the manual combination of many different classifiers, forming together a machinery capable of inferring a wiring diagram from existing neurite reconstructions. However, the biggest challenge in connectomic analysis, the reconstruction of the wires, is not addressed by SyConn. A novel approach based on flood-filling-networks (Januszewski et al., 2016), which outperformed all other methods so far on a small test of j0126, appears to be promising.

[13] see e.g. `https://www.kaggle.com/`

5.3.3 How accurate is accurate enough?

Different strategies were used so far to analyze volume EM datasets, with very different results regarding the annotation accuracy. Two studies are exemplary: The work from Kasthuri in cortex (Kasthuri et al., 2015), in which they labeled structures at extreme detail, with many years of manual work and restricted to a small volume (1,500 μm^3), and a study by Helmstädter (Helmstaedter et al., 2013), in which a piece of retina was analyzed by combining skeleton tracings with machine learning to generate volume reconstruction of neurons. While Kasthuri et al. itemized every synaptic vesicle, Helmstaedter et al. inferred the retinal wiring from the contact surface of cells without the analysis of synapses. The latter is clearly a much less detailed analysis, but also covers a volume that is almost three orders of magnitudes greater (1,200,000 μm^3). The only argument against increased reconstruction accuracy is the added manual or computational effort, and the fact that this effort might not be in relation to the additionally gained knowledge. However, given that it is currently even unclear how much knowledge about a brain can be extracted from a static volume EM dataset, the question should be viewed from a different perspective: which conclusions does the data and accuracy support, or in Karl Popper's terms, is a given accuracy sufficient to rule out a previously formulated hypothesis? To illustrate this rather theoretical point of view, a simulation study of Rajan et al., 2016 can be considered that discusses highly structured vs weakly structured circuits, which were analyzed and compared for sequence generation and memory. If a circuit is only weakly structured, i.e. the exact weight (size) of very few synapses determines its computation(s), a very high reconstruction accuracy will be required to resolve these subtle differences, with the possibility that the imaging modality used (the staining or volume EM acquisition parameters) proves to be insufficient, even at the highest degree of analysis. It would indeed be quite possible that a volume EM dataset cannot answer these question at all, e.g. if the circuit is weakly structured, and all parameters are encoded in biophysical properties that are hidden from the electron microscope. Despite these possible fundamental limitations of the analysis of volume EM datasets, and the open question of the required analysis accuracy, the skeleton-based j0126 connectivity matrix created for this thesis (Dorkenwald et al., 2017) was fully consistent with previously made electrophysiological observations of connectivity in Area X (Farries et al., 2005;

Perkel et al., 2002; Leblois et al., 2009) and confirmed synaptic contacts between cell types that were only assumed to exist from light microscopy data (Farries et al., 2005).

Chapter 6

Conclusions and outlook

6.1 Thesis achievements

This thesis had three main objectives:

1. The investigation of zebra finch HVC with volume EM, with a focus on the local synaptic connectivity of HVC_{RA} neurons that might underly a synaptic chain responsible for neuronal sequence generation.

2. The investigation of zebra finch Area X with volume EM, to test the anatomical predictions of a basal-ganglia reinforcement learning model that explicitly discriminates between context and efference copy signals.

3. The development of novel semi-automatic computational approaches to make the analysis of volume EM datasets more efficient.

A SBEM dataset of a small portion of HVC was acquired and analyzed, and synaptic connections between HVC_{RA} neurons were identified that might form the anatomical substrate of a functional synaptic chain. It was further discovered that HVC_{RA} neurons target mostly interneurons in the region close to their soma, but target other HVC_{RA} neurons at higher

soma-distances, a finding with interesting implications for how neuronal circuits might generate activity sequences.

Two SBEM datasets of the striatopallidal Area X were acquired to test the anatomical predictions of a reinforcement learning model. The connectivity that was found between HVC and LMAN excitatory afferents and MSNs was consistent with the model predictions, supporting the notion of separate efference copy and context signals in reinforcement learning.

Based on the smaller SBEM dataset of Area X, novel methods using state-of-the-art machine learning were developed that will significantly ease (through dramatic cost reductions) the analysis of volume EM datasets in the future. First, a proofreading workflow was developed in collaboration with Google Research, outperforming currently used neurite reconstruction approaches in speed by more than an order of magnitude. Secondly, a synaptic connectivity inference pipeline (SyConn; Dorkenwald et al., 2017) was developed that shows how the daunting task of manual synapse annotations can be fully automated at small remaining error rates. SyConn was then used to generate the first synaptic wiring diagram of the vertebrate basal-ganglia (Area X) based on volume EM.

6.2 Outlook

Despite the presented advances in our understanding of the neuronal sequence generator in HVC, it remains to be determined whether excitatory synaptic feed-forward connectivity underlies the ultra-sparse code observed in $\mathrm{HVC_{RA}}$ neurons (Hahnloser et al., 2002) and what the role is of the other neurons that can be found in HVC. Two-photon based functional $\mathrm{Ca^{2+}}$ imaging should therefore be combined in future experiments with connectomic analysis of the same tissue for a more powerful hypothesis test.

While a first hypothesis test of the anatomical predictions of the reinforcement learning model of Area X was performed, many anatomical features relevant for the model that can be inspected in the existing datasets remain to be explored: Do LMAN axons target inhibitory interneurons to suppress through feed-forward inhibition the activity they evoke in MSNs? Are there further

differences in the targeted MSN compartments? Anatomical hypotheses can now be derived directly from the mechanistic learning model and tested in the existing datasets.

The next step in semi-automated connectomic analysis will be the combination of curated computer-generated neurite reconstructions with the automated synaptic connectivity inference pipeline that was presented in this thesis. This will enable for the first time the dense connectomic analysis of a larger volume EM dataset without investing millions of Euros into manual analyis.

List of abbreviations

ACh . . . . . cholinergic neuron

ANOVA . . . analysis of variance

AFP . . . . . anterior forebrain pathway

ATUM . . . . Automatic Tape-collecting Ultra-Microtome

Area X . . . . Area X of the medial striatum

BDA . . . . . biotinylated dextrane amine

CAF . . . . . conditional auditory feedback

CB . . . . . cacodylate buffer

CLAHE . . . contrast limited adaptive histogram equalization

CM . . . . . caudal mesopallium

CNN . . . . . convolutional neural network

DA . . . . . Dopamine

DAB . . . . . 3,3'-Diaminobenzidine

DLM . . . . . dorso-lateral medial thalamus

DiK-SBEM . diamond knife SBEM

EA . . . . . excitatory axon

ECS . . . . . extracellular space

EM . . . . . electron microscopy

FIB-SEM . . focused ion beam milling SEM

FoXP2 . . . . Forkhead box protein P2

FS . . . . . fast-spiking, parvalbumin-positive neuron

GALA . . . . Graph-based active learning of agglomeration

GA . . . . . glutaraldehyde

GP . . . . . pallidal-like neuron

GPi,e . . . . . pallidal-like neuron of direct, indirect pathway

GUI . . . . . graphical user interface

HRP . . . . . horseradish peroxidase

HVC . . . . . letter-based proper name, see (Reiner et al., 2004)

INT . . . . . interneuron

kDa . . . . . kilo Dalton; molecular weight

kV . . . . . kilovolts

LMAN . . . . lateral magnocellular nucleus of the anterior nidopallium

LM . . . . . light microscopy

LTS . . . . . low-threshold-spiking, somatostatin-positive neuron

Mg . . . . . Magnesium

MSN . . . . . medium spiny neuron

nA . . . . . nanoampere

NIf . . . . . interfacial nucleus of the nidopallium

nm . . . . . nanometer

NMDA . . . . N-methyl-D-aspartate

ns . . . . . nanosecond

PbAsp . . . . Lead aspartate

PFA . . . . . paraformaldehyde

RFC . . . . . random forest classifier

RL . . . . . reinforcement learning

ROTO . . . . reduced osmium, thiocarbohydrazide, osmium staining

SBEM . . . . serial block-face SEM

s.d. . . . . . standard deviation

s.e.m. . . . . . standard error of the mean

SEM . . . . . scanning electron microscopy

SMP . . . . . song motor pathway

STN . . . . . subthalamic nucleus

ssTEM . . . . serial section TEM

TB . . . . . terabytes

TEM . . . . . transmission electron microscopy

UAc . . . . . Uranyl acetate

Uva . . . . . nucleus uvaeformis of the thalamus

µm . . . . . micrometer

vs . . . . . versus

WTA . . . . . winner-take-all

References

Abeles, M. (1982). *Local cortical circuits: An electrophysiological study*, volume 6. Springer-Verlag.

Abeles, M., Bergman, H., Margalit, E., and Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, 70(4):1629–1638.

Agnati, L. F. and Fuxe, K. (2014). Extracellular-vesicle type of volume transmission and tunnelling-nanotube type of wiring transmission add a new dimension to brain neuro-glial networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1652):20130505.

Akutagawa, E. and Konishi, M. (2010). New brain pathways found in the vocal control system of a songbird. *The Journal of Comparative Neurology*, 518(15):3086–3100.

Alexander, G. E. and Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences*, 13(7):266–271.

Ali, F., Otchy, T. M., Pehlevan, C., Fantana, A. L., Burak, Y., and Ölveczky, B. P. (2013). The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron*, 80(2):494–506.

Alivisatos, A. P., Chun, M., Church, G. M., Greenspan, R. J., Roukes, M. L., and Yuste, R. (2012). The brain activity map project and the challenge of functional connectomics. *Neuron*, 74(6):970–974.

Amador, A., Perl, Y. S., Mindlin, G. B., and Margoliash, D. (2013). Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature*, 495(7439):59–64.

Andalman, A. S. and Fee, M. S. (2009). A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences*, 106(30):12518–12523.

Anderson, J. C., Douglas, R. J., Martin, K. A. C., and Nelson, J. C. (1994). Map of the synapses formed with the dendrites of spiny stellate neurons of cat visual cortex. *The Journal of Comparative Neurology*, 341(1):25–38.

Arganda-Carreras, I., Turaga, S. C., Berger, D. R., Cireşan, D., Giusti, A., Gambardella, L. M., Schmidhuber, J., Laptev, D., Dwivedi, S., Buhmann, J. M., Liu, T., Seyedhosseini, M., Tasdizen, T., Kamentsky, L., Burget, R., Uher, V., Tan, X., Sun, C., Pham, T. D., Bas, E., Uzunbas, M. G., Cardona, A., Schindelin, J., and Seung, H. S. (2015). Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in Neuroanatomy*, 9:142.

Armstrong, E. and Abarbanel, H. D. I. (2016). Model of the songbird nucleus HVC as a network of central pattern generators. *Journal of Neurophysiology*, 116(5):2405–2419.

Beaulieu, C., Kisvarday, Z., Somogyi, P., Cynader, M., and Cowey, A. (1992). Quantitative distribution of gaba-immunopositive and-immunonegative neurons and synapses in the monkey striate cortex (area 17). *Cerebral Cortex*, 2(4):295–309.

Beier, T., Pape, C., Rahaman, N., Prange, T., Berg, S., Bock, D. D., Cardona, A., Knott, G. W., Plaza, S. M., Scheffer, L. K., Koethe, U., Kreshuk, A., and Hamprecht, F. A. (2017). Multicut brings automated neurite segmentation closer to human performance. *Nature Methods*, 14(2):101–102.

Bennett, B. and Bolam, J. (1994). Synaptic input and output of parvalbumin-immunoreactive neurons in the neostriatum of the rat. *Neuroscience*, 62(3):707–719.

Berning, M., Boergens, K. M., and Helmstaedter, M. (2015). SegEM: efficient image analysis for High-Resolution connectomics. *Neuron*, 87(6):1193–1206.

Bottjer, S. W. and Alexander, G. (1995). Localization of met-enkephalin and vasoactive intestinal polypeptide in the brains of male zebra finches (part 1 of 2). *Brain, Behavior and Evolution*, 45(3):153–165.

Brainard, M. and Doupe, A. (2000). Interruption of a basal ganglia–forebrain circuit prevents plasticity of learned vocalizations. *Nature*, 404(6779):762–766.

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

Briggman, K. L. and Bock, D. D. (2012). Volume electron microscopy for neuronal circuit reconstruction. *Current Opinion in Neurobiology*, 22(1):154–161.

Briggman, K. L., Helmstaedter, M., and Denk, W. (2011). Wiring specificity in the direction-selectivity circuit of the retina. *Nature*, 471(7337):183–188.

Budzillo, A., Duffy, A., Miller, K. E., Fairhall, A. L., and Perkel, D. J. (2017). Dopaminergic modulation of basal ganglia output through coupled excitation–inhibition. *Proceedings of the National Academy of Sciences*, page 201611146.

Bumbarger, D. J., Riebesell, M., Rödelsperger, C., and Sommer, R. J. (2013). System-wide rewiring underlies behavioral differences in predatory and bacterial-feeding nematodes. *Cell*, 152(1):109–119.

Burek, M. J., Nordeen, K. W., and Nordeen, E. J. (1991). Neuron loss and addition in developing zebra finch song nuclei are independent of auditory experience during song learning. *Journal of Neurobiology*, 22(3):215–223.

Cajal, R. y. S. (1888). Estructura de los centros nerviosos de las aves. *Rev. Trimest. Histol. Norm. Patol.*, 1:1–10.

Cannon, J., Kopell, N., Gardner, T., and Markowitz, J. (2015). Neural sequence generation using spatiotemporal patterns of inhibition. *PLoS Computational Biology*, 11(11):e1004581.

Cardin, J. A., Raksin, J. N., and Schmidt, M. F. (2005). Sensorimotor nucleus NIf is necessary for auditory processing but not vocal motor output in the avian song system. *Journal of Neurophysiology*, 93(4):2157–2166.

Carrillo, G. D. and Doupe, A. J. (2004). Is the songbird area x striatal, pallidal, or both? an anatomical study. *The Journal of Comparative Neurology*, 473(3):415–437.

Cimino, G. (1999). Reticular theory versus neuron theory in the work of camillo golgi. *Physis Riv Int Stor Sci.*, 36(2):431–72.

Coleman, M. J. and Vu, E. T. (2005). Recovery of impaired songs following unilateral but not bilateral lesions of nucleus uvaeformis of adult zebra finches. *Journal of Neurobiology*, 63(1):70–89.

Colonnier, M. (1968). Synaptic patterns on different cell types in the different laminae of the cat visual cortex. an electron microscope study. *Brain Research*, 9(2):268–287.

Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, R., Sader, E. N., Ko, H., Hofer, S. B., and Mrsic-Flogel, T. D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*, 518(7539):399–403.

Cragg, B. (1980). Preservation of extracellular space during fixation of the brain for electron microscopy. *Tissue and Cell*, 12(1):63–72.

DeFelipe, J. and Fariñas, I. (1992). The pyramidal neuron of the cerebral cortex: Morphological and chemical characteristics of the synaptic inputs. *Progress in Neurobiology*, 39(6):563–607.

Denk, W., Briggman, K. L., and Helmstaedter, M. (2012). Structural neurobiology: missing link to a mechanistic understanding of neural computation. *Nature Reviews Neuroscience*, 13(5):351–358.

Denk, W. and Horstmann, H. (2004). Serial Block-Face scanning electron microscopy to reconstruct Three-Dimensional tissue nanostructure. *PLoS Biology*, 2(11):e329.

Ding, L. and Perkel, D. (2004). Long-term potentiation in an avian basal ganglia nucleus essential for vocal learning. *The Journal of Neuroscience*, 24(2):488–494.

Dorkenwald, S., Schubert, P. J., Killinger, M. F., Urban, G., Mikula, S., Svara, F., and Kornfeld, J. (2017). Automated synaptic connectivity inference for volume electron microscopy. *Nature Methods*, 14(4):435–442.

Dutar, P., Vu, H. M., and Perkel, D. J. (1998). Multiple cell types distinguished by physiological, pharmacological, and anatomic properties in nucleus hvc of the adult zebra finch. *Journal of Neurophysiology*, 80(4):1828–1838.

Emmons, S. W. (2015). The beginning of connectomics: a commentary on white et al. (1986) 'the structure of the nervous system of the nematode caenorhabditis elegans'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1666):20140309.

Evanko, D., Heinrichs, A., and Rosenthal, C. (2011). Milestones in light microscopy.

Farries, M., Ding, L., and Perkel, D. (2005). Evidence for direct and indirect pathways through the song system basal ganglia. *The Journal of Comparative Neurology*, 484(1):93–104.

Fee, M., Kozhevnikov, A., and Hahnloser, R. (2004). Neural mechanisms of vocal sequence generation in the songbird. *Annals of the New York Academy of Sciences*, 1016(1):153–170.

Fee, M. S. (2014). The role of efference copy in striatal learning. *Current Opinion in Neurobiology*, 25:194–200.

Fee, M. S. and Goldberg, J. H. (2011). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. *Neuroscience*, 198:152–170.

Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., and Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science*, 354(6317):1278–1282.

Gale, S. and Perkel, D. (2010). Anatomy of a songbird basal ganglia circuit essential for vocal learning and plasticity. *Journal of Chemical Neuroanatomy*, 39(2):124–131.

Gibb, L., Gentner, T., and Abarbanel, H. (2009a). Brain stem feedback in a computational model of birdsong sequencing. *Journal of Neurophysiology*, 102(3):1763–1778.

Gibb, L., Gentner, T., and Abarbanel, H. (2009b). Inhibition and recurrent excitation in a computational model of sparse bursting in song nucleus HVC. *Journal of Neurophysiology*, 102(3):1748–1762.

Glauert, A. M. and Lewis, P. R. (1999). *Biological Specimen Preparation for Transmission Electron Microscopy*. Princeton University Press, 1 edition.

Goldberg, J. H., Adler, A., Bergman, H., and Fee, M. S. (2010). Singing-related neural activity distinguishes two putative pallidal cell types in the songbird basal ganglia: Comparison to the primate internal and external pallidal segments. *The Journal of Neuroscience*, 30(20):7088–7098.

Goldberg, J. H. and Fee, M. S. (2010). Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. *Journal of Neurophysiology*, 103(4):2002–2014.

Gray, G. (1959). Axo-somatic and axo-dendritic synapses of the cerebral cortex. *Journal of Anatomy*, 93(Pt 4):420–433.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31(1):359–387.

Grillner, S. and Robertson, B. (2016). The basal ganglia over 500 million years. *Current Biology*, 26(20):1088–1100.

Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX.

Hahnloser, R., Kozhevnikov, A., and Fee, M. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*, 419(6902):65–70.

Hamaguchi, K., Tanaka, M., and Mooney, R. (2016). A distributed recurrent network contributes to temporally precise vocalizations. *Neuron*, 91(3):680–693.

Hayworth, K. J., Xu, C. S., Lu, Z., Knott, G. W., Fetter, R. D., Tapia, J. C., Lichtman, J. W., and Hess, H. F. (2015). Ultrastructurally smooth thick partitioning and volume stitching for large-scale connectomics. *Nature Methods*, 12(4):319–322.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Psychology Press.

Helmchen, F. and Denk, W. (2005). Deep tissue two-photon microscopy. *Nature Methods*, 2(12):932–940.

Helmstaedter, M. (2013). Cellular-resolution connectomics: challenges of dense neural circuit reconstruction. *Nature Methods*, 10(6):501–507.

Helmstaedter, M., Briggman, K., Turaga, S., Jain, V., Seung, H., and Denk, W. (2013). Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, 500(7461):168–174.

Helmstaedter, M., Briggman, K. L., and Denk, W. (2008). 3d structural imaging of the brain with photons and electrons. *Current Opinion in Neurobiology*, 18(6):633–641.

Helmstaedter, M., Briggman, K. L., and Denk, W. (2011). High-accuracy neurite reconstruction for high-throughput neuroanatomy. *Nature Neuroscience*, 14(8):1081–1088.

Herculano-Houzel, S., Mota, B., and Lent, R. (2006). Cellular scaling rules for rodent brains. *Proceedings of the National Academy of Sciences*, 103(32):12138–12143.

Hildebrand, D. G. C., Cicconet, M., Torres, R. M., Choi, W., Quan, T. M., Moon, J., Wetzel, A. W., Champion, A. S., Graham, B. J., Randlett, O., Plummer, G. S., Portugues, R., Bianco, I. H., Saalfeld, S., Baden, A. D., Lillaney, K., Burns, R., Vogelstein, J. T., Schier, A. F., Lee, W.-C. A., Jeong, W.-K., Lichtman, J. W., and Engert, F. (2017). Whole-brain serial-section electron microscopy in larval zebrafish. *Nature*, 545(7654):345–349.

Hopfield, J. J. and Tank, D. W. (1985). Neural computation of decisions in optimization problems. *Biological cybernetics*, 52(3):141–152.

Huang, G. B. and Jain, V. (2013). Deep and wide multiscale recursive networks for robust image labeling. *arXiv:1310.0354*.

Hughes, L., Hawes, C., Monteith, S., and Vaughan, S. (2014). Serial block face scanning electron microscopy—the future of cell ultrastructure imaging. *Protoplasma*, 251(2):395–401.

Immelmann, K. (1969). Song development in the zebra finch and other estrildid finches. *Bird Vocalizations*, pages 61–74.

Jaeger, D., Kita, H., and Wilson, C. J. (1994). Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. *Journal of Neurophysiology*, 72(5):2555–2558.

Jahanshahi, M., Obeso, I., Rothwell, J. C., and Obeso, J. A. (2015). A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. *Nature Reviews Neuroscience*, 16(12):719–732.

Januszewski, M., Jeremy, M., Li, P., Kornfeld, J., Denk, W., and Jain, V. (2016). Flood-Filling networks. *arXiv:1611.00421*.

Jin, D. Z. (2009). Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus HVC. *Physical Review E*, 80(5):051902.

Jin, D. Z., Ramazanoğlu, F. M., and Seung, H. S. (2007). Intrinsic bursting enhances the robustness of a neural network model of sequence generation by avian brain area HVC. *Journal of Computational Neuroscience*, 23(3):283–299.

Kao, M., Doupe, A., and Brainard, M. (2005). Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature*, 433(7026):638–643.

Karabelas, A. B. and Purrura, D. P. (1980). Evidence for autapses in the substantia nigra. *Brain Research*, 200(2):467–473.

Karnovsky, M. (1971). Use of ferrocyanide-reduced osmium tetroxide in electron microscopy. *The Journal of Cell Biology*, 51:146a.

Kasthuri, N., Hayworth, K. J., Berger, D. R., Schalek, R. L., Conchello, J. A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T. R., Roberts, M., Morgan, J. L., Tapia, J. C., Seung, H. S., Roncal, W. G., Vogelstein, J. T., Burns, R., Sussman, D. L., Priebe, C. E., Pfister, H., and Lichtman, J. W. (2015). Saturated reconstruction of a volume of neocortex. *Cell*, 162(3):648–661.

Kemen, T., Malloy, M., Thiel, B., Mikula, S., Denk, W., Dellemann, G., and Zeidler, D. (2015). Further advancing the throughput of a multi-beam SEM. *SPIE Advanced Lithography*, 9424:94241U–94241U.

Killinger, M. F. (2016). Agent-Based automatic neuron skeletonisation in 3-Dimensional electron microscopy images. Master's thesis. University of Heidelberg.

Killinger, M. F., Urban, G., Dorkenwald, S., Schubert, P., Drawitsch, M., and Kornfeld, J. (2015). ELEKTRONN. http://elektronn.org/, accessed on 05/25/2017.

Knott, G., Marchman, H., Wall, D., and Lich, B. (2008). Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *The Journal of Neuroscience*, 28(12):2959–2964.

Koós, T. and Tepper, J. M. (1999). Inhibitory control of neostriatal projection neurons by GABAergic interneurons. *Nature Neuroscience*, 2(5):467–472.

Kornfeld, J., Benezra, S. E., Narayanan, R. T., Svara, F., Egger, R., Oberlaender, M., Denk, W., and Long, M. A. (2017). Em connectomics reveals axonal target variation in a sequence-generating network. *eLife*, 6:e24364.

Kornfeld, J., Svara, F., Pfeiler, N., Nguyen, M., Pronkin, M., Stepanov, A., Spaar, S., Knecht, A., Ripp, C., Wegner, M., Briggman, K., Helmstädter, M., and Denk, W. (2011). KNOSSOS. https://knossostool.org/, accessed on 05/25/2017; Publication in preparation.

Kosche, G., Vallentin, D., and Long, M. (2015). Interplay of inhibition and excitation shapes a premotor neural sequence. *The Journal of Neuroscience*, 35(3):1217–1227.

Kozhevnikov, A. and Fee, M. (2007). Singing-related activity of identified HVC neurons in the zebra finch. *Journal of Neurophysiology*, 97(6):4271–4283.

Kubota, M. and Taniguchi, I. (1998). Electrophysiological characteristics of classes of neuron in the hvc of the zebra finch. *Journal of Neurophysiology*, 80(2):914–923.

Lai, H. C. and Jan, L. Y. (2006). The distribution and targeting of neuronal voltage-gated ion channels. *Nature Reviews Neuroscience*, 7(7):548–562.

Leblois, A., Bodor, A. L., Person, A. L., and Perkel, D. J. (2009). Millisecond timescale disinhibition mediates fast information transmission through an avian basal ganglia loop. *The Journal of Neuroscience*, 29(49):15420–15433.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

Leighton, S. (1981). SEM images of block faces, cut by a miniature microtome within the SEM - a technical note. *Scanning electron microscopy*, (Pt 2):73–76.

Levenson, J. M. and Sweatt, J. D. (2005). Epigenetic mechanisms in memory formation. *Nature Reviews Neuroscience*, 6(2):108–118.

Lombardino, A. J. and Nottebohm, F. (2000). Age at deafening affects the stability of learned song in adult male zebra finches. *The Journal of Neuroscience*, 20(13):5054–5064.

Long, M. A. and Fee, M. S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*, 456(7219):189–194.

Long, M. A., Jin, D. Z., and Fee, M. S. (2010). Support for a synaptic chain model of neuronal sequence generation. *Nature*, 468(7322):394–399.

Luo, M., Ding, L., and Perkel, D. J. (2001). An avian basal ganglia pathway essential for vocal learning forms a closed topographic loop. *The Journal of Neuroscience*, 21(17):6836–6845.

Luo, M. and Perkel, D. (1999). Long-range GABAergic projection in a circuit essential for vocal learning. *The Journal of Comparative Neurology*, 403(1):68–84.

Mahamid, J. and Baumeister, W. (2012). Cryo-electron tomography: the realization of a vision. *Microscopy and Analysis*, 26:45–48.

Mikula, S., Binding, J., and Denk, W. (2012). Staining and embedding the whole mouse brain for electron microscopy. *Nature Methods*, 9(12):1198–1201.

Mikula, S. and Denk, W. (2015). High-resolution whole-brain staining for electron microscopic circuit reconstruction. *Nature Methods*, 12(6):541–546.

Mooney, R. (2000). Different subthreshold mechanisms underlie song selectivity in identified hvc neurons of the zebra finch. *The Journal of Neuroscience*, 20(14):5420–5436.

Mooney, R. (2009). Neural mechanisms for learned birdsong. *Learning & Memory*, 16(11):655–669.

Mooney, R. and Prather, J. (2005). The HVC microcircuit: The synaptic basis for interactions between song motor and vocal plasticity pathways. *The Journal of Neuroscience*, 25(8):1952–1964.

Mostafa, H. and Indiveri, G. (2014). Sequential activity in asymmetrically coupled winner-take-all circuits. *Neural Computation*, 26(9):1973–2004.

Nguyen, M. (2014). An online micro-jobbing approach for the segmentation of neuronal 3D electron microscopy datasets. Bachelor's thesis. Hochschule Mannheim.

Nordeen, K. and Nordeen, E. (1988). Projection neurons within a vocal motor pathway are born during song learning in zebra finches. *Nature*, 334:149–151.

Nottebohm, F., Stokes, T., and Leonard, C. (1976). Central control of song in the canary, serinus canarius. *The Journal of Comparative Neurology*, 165(4):457–486.

Nunez-Iglesias, J., Kennedy, R., Plaza, S. M., Chakraborty, A., and Katz, W. T. (2014). Graph-based active learning of agglomeration (GALA): a python library to segment 2d and 3d neuroimages. *Frontiers in Neuroinformatics*, 8:34.

Ölveczky, B. P., Otchy, T. M., Goldberg, J. H., Aronov, D., and Fee, M. S. (2011). Changes in the neural control of a complex motor sequence during learning. *Journal of Neurophysiology*, 106(1):386–397.

Oorschot, D. E., Lin, N., Cooper, B. H., Reynolds, J. N., Sun, H., and Wickens, J. R. (2013). Synaptic connectivity between rat striatal spiny projection neurons in vivo: unexpected multiple somatic innervation in the context of overall sparse proximal connectivity. *Basal Ganglia*, 3(2):93–108.

Otchy, T. M., Wolff, S. B. E., Rhee, J. Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S. M. H., and Ölveczky, B. P. (2015). Acute off-target effects of neural circuit manipulations. *Nature*, 528(7582):358–363.

Palade, G. E. and Palay, S. L. (1954). Electron microscope observations of interneuronal and neuromuscular synapses. *Anat. Rec.*, 118:335–336.

Pallotto, M., Watkins, P. V., Fubara, B., Singer, J. H., and Briggman, K. L. (2015). Extracellular space preservation aids the connectomic analysis of neural circuits. *eLife*, 4:e08206.

Pehlevan, C., Ali, F., and Olveczky, B. P. (2015). Flexibility in motor timing constrains the topology and dynamics of pattern generator circuits. *bioRxiv*, page 033472.

Perkel, D., Farries, M., Luo, M., and Ding, L. (2002). Electrophysiological analysis of a songbird basal ganglia circuit essential for vocal plasticity. *Brain Research Bulletin*, 57(3-4):529–532.

Person, A., Gale, S., Farries, M., and Perkel, D. (2008). Organization of the songbird basal ganglia, including area x. *The Journal of Comparative Neurology*, 508(5):840 866.

Peters, A. (2002). Examining neocortical circuits: some background and facts. *Journal of Neurocytology*, 31(3-5):183–193.

Petkov, C. I. and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4:12.

Powers, D. M. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1):37–63.

Price, P. H. (1979). Developmental determinants of structure in zebra finch song. *Journal of Comparative and Physiological Psychology*, 93(2):260.

Rajan, K., Harvey, C. D., and Tank, D. W. (2016). Recurrent network models of sequence generation and memory. *Neuron*, 90(1):128–142.

Reiner, A., Perkel, D. J., Mello, C. V., and Jarvis, E. D. (2004). Songbirds and the revised avian brain nomenclature. *Annals of the New York Academy of Sciences*, 1016(1):77–108.

Reiner, A., Veenman, C., Medina, L., Jiao, Y., Mar, N. D., and Honig, M. G. (2000). Pathway tracing using biotinylated dextran amines. *Journal of Neuroscience Methods*, 103(1):23–37.

Roberts, T. F., Hisey, E., Tanaka, M., Kearney, M. G., Chattree, G., Yang, C. F., Shah, N. M., and Mooney, R. (2017). Identification of a motor-to-auditory pathway important for vocal learning. *Nature Neuroscience*.

Roberts, T. F., Klein, M. E., Kubke, M. F., Wild, J. M., and Mooney, R. (2008). Telencephalic neurons monosynaptically link brainstem and forebrain premotor networks necessary for song. *The Journal of Neuroscience*, 28(13):3479–3489.

Scharff, C., Kirn, J., Grossman, M., Macklis, J., and Nottebohm, F. (2000). Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron*, 25(2):481–492.

Scheffer, L. K., Karsh, B., and Vitaladevun, S. (2013). Automated alignment of imperfect em images for neural reconstruction. *arXiv:1304.6034*.

Seligman, A., Wasserkrug, H., and Hanker, J. (1966). A new staining method (OTO) for enhancing contrast of Lipid-Containing membranes and droplets in osmium Tetroxide-Fixed tissue with osmiophilic thiocarbohydrazide (TCH). *The Journal of Cell Biology*, 30(2):424–432.

Sherrington, C. S. (1906). *The integrative action of the nervous system*. Oxford University Press: H. Milford.

Simpson, H. B. and Vicario, D. S. (1991). Early estrogen treatment of female zebra finches masculinizes the brain pathway for learned vocalizations. *Journal of Neurobiology*, 22(7):777–793.

Sohrabji, F., Nordeen, E., and Nordeen, K. (1990). Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. *Behav Neural Biol*, 53(1):51–63.

Strata, P. and Harvey, R. (1999). Dale's principle. *Brain Research Bulletin*, 50(5-6):349–350.

Svara, F. (2017). *Volume Electron Microscopic Analyses in the Larval Zebrafish*. PhD thesis. University of Heidelberg.

Theano Team, T., Rami, A., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J., Belikov, A., Belopolsky, A., Bengio, Y., Bergeron, A., Bergstra, J., Bisson, V., Snyder, J., Bouchard, N., Nicolas, B., Bouthillier, X., de Brébisson, A., Breuleux, O., Carrier, P., Cho, K., Chorowski, J., Christiano, P., Cooijmans, T., Côté, M., Côté, M.,

Courville, A., Dauphin, Y. N., Delalleau, O., Demouth, J., Desjardins, G., Dieleman, S., Dinh, L., Ducoffe, M., Dumoulin, V., Kahou, S., Erhan, D., Fan, Z., Firat, O., Germain, M., Glorot, X., Goodfellow, I., Graham, M., Gulcehre, C., Hamel, P., Harlouchet, I., Heng, J., Hidasi, B., Honari, S., Jain, A., Jean, S., Jia, K., Korobov, M., Kulkarni, V., Lamb, A., Lamblin, P., Larsen, E., Laurent, C., Lee, S., Lefrancois, S., Lemieux, S., Léonard, N., Lin, Z., Livezey, J. A., Lorenz, C., Lowin, J., Ma, Q., Manzagol, P., Mastropietro, O., T, M. R., Memisevic, R., van Merriënboer, B., Michalski, V., Mirza, M., Orlandi, A., Pal, C., Pascanu, R., Pezeshki, M., Raffel, C., Renshaw, D., Rocklin, M., Romero, A., Roth, M., Sadowski, P., Salvatier, J., Savard, F., Schlüter, J., Schulman, J., Schwartz, G., Serban, I., Serdyuk, D., Shabanian, S., Simon, É., Spieckermann, S., Subramanyam, R. S., Sygnowski, J., Tanguay, J., van Tulder, G., Turian, J., Urban, S., Vincent, P., Visin, F., de Vries, H., David, W., Webb, D. J., Willson, M., Xu, K., Xue, L., Yao, L., Zhang, S., and Zhang, Y. (2016). Theano: A python framework for fast computation of mathematical expressions.

Tumer, E. and Brainard, M. (2007). Performance variability enables adaptive plasticity of crystallized adult birdsong. *Nature*, 450(7173):1240–1244.

Vallentin, D. and Long, M. (2015). Motor origin of precise synaptic inputs onto forebrain neurons driving a skilled behavior. *The Journal of Neuroscience*, 35(1):299–307.

Vargha-Khadem, F., Gadian, D. G., Copp, A., and Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*, 6(2):131–138.

Vicario, D. and Simpson, H. (1995). Electrical stimulation in forebrain nuclei elicits learned vocal patterns in songbirds. *Journal of Neurophysiology*, 73(6):2602–2607.

Walton, J. (1979). Lead asparate, an en bloc contrast stain particularly useful for ultrastructural enzymology. *Journal of Histochemistry & Cytochemistry*, 27(10):1337 –1342.

Wang, N., Hurley, P., Pytte, C., and Kirn, J. (2002). Vocal control neuron incorporation decreases with age in the adult zebra finch. *The Journal of Neuroscience*, 22(24):10864–10870.

Wang, R., Chen, C.-C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., Chakraborty, M., Audet, J.-N., and Jarvis, E. D. (2014). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *The Journal of Comparative Neurology*, 523(6):892–906.

Wanner, A., Kirschmann, M., and Genoud, C. (2015). Challenges of microtome-based serial block-face scanning electron microscopy in neuroscience. *Journal of Microscopy*, 259(2):137–142.

Wanner, A. A., Genoud, C., Masudi, T., Siksou, L., and Friedrich, R. W. (2016). Dense EM-based reconstruction of the interglomerular projectome in the zebrafish olfactory bulb. *Nature Neuroscience*, 19(6):816–825.

White, J., Southgate, E., Thomson, J., and Brenner, S. (1986). The structure of the nervous system of the nematode caenorhabditis elegans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 314(1165):1–340.

Wild, J. M., Williams, M. N., Howie, G. J., and Mooney, R. (2005). Calcium-binding proteins define interneurons in HVC of the zebra finch (taeniopygia guttata). *The Journal of Comparative Neurology*, 483(1):76–90.

Yuste, R. and Bonhoeffer, T. (2001). Morphological changes in dendritic spines associated with Long-Term synaptic plasticity. *Annual Review of Neuroscience*, 24(1):1071–1089.