Current-Source Density Estimation from Linear Depth Electrode Recordings

Pascal Kropf

Integrated Program in Neuroscience
McGill University
Montreal, QC, Canada

August 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Amir Shmuel for accepting me in his lab and providing me with continued funding over all those years. Without his extensive efforts and continued support this work could not have been completed.

I would also like to thank my committee members, Dr. Curtis Baker, Dr. Jean-Marc Lina, and Dr. Robert Kearney for all the work and critical feedback they have provided me throughout these years. I would especially like to thank Dr. Curtis Baker for his substantial efforts towards improving my thesis project proposal and steering me along the right direction.

The work presented here would not have been possible without the collaborative action of my labmates and fellow graduate students who have pitched in along the way: Martin Villeneuve, Zeshan Yao, Sujay Neupane, Kuwook Cha, Abdel Baig, Ajay Venkateswaran, Shahab Bakthiari, Alexandre Hutton, Peter O'Connor, Aleksandra Bortel, Clotilde Lecrux, Victor Mocanu, James McNicol, Roland Pilgram, Roberto Sotero, Niladri Mohanti, Sebastian Thomas, and Gabriel Martine-La Boissoniere. A very special thanks goes to Dr. Martin Villeneuve who spent a great amount of time showing me the tricks of the trade and making sure that I never lost sight of the goal.

Thank you also to Ms. Toula Papadopoulos and all the lovely people at the IPN office for their extensive help with administrative duties, and to Animal Care Facility at the Montreal Neurological Institute for their readiness to help with anything, even on short notice.

My journey as a Ph.D. student would not have been nearly as much fun without all the great people that I got the chance to meet along the way. I would therefore like to express my gratitude to Ihab, Ying Syuan, Sujay, Kuwook, Martin, Zeshan, Ajay, Pariya, Shahab, Steven, and many more for all the great discussions as well as their support and understanding.

Finally, I would like to thank my family and Evelin for their never-ending support, critical input, and encouragement. Without their love and patience none of this would ever have happened.

Contribution of authors

This thesis comprises three main chapters, an introduction, a discussion and three appendices. Chapter 3 has previously been published in a peer-reviewed journal (Kropf and Shmuel, 2016). All the work presented herein has been conducted under the supervision of Dr. Amir Shmuel. In all cases, Dr. Shmuel helped conceive the experimental design, contributed to data interpretation, and helped writing the manuscripts.

The introduction, discussion as well as Chapters 3 and 4, and the three appendices were elaborated and written by myself.

Chapter 2 was written in a shared first authorship with Dr. Martin Villeneuve. The data acquisition was performed by Dr. Martin Villeneuve and myself with contributions from Dr. Zeshan Yao. The experimental protocol was elaborated by Dr. Villeneuve and Dr. Shmuel. I provided all the coding related to conducting the experiment. The data analysis and interpretation was shared among Dr. Villeneuve and myself. Dr. Villeneuve performed the histology work and I performed the simulation results.

Table of Contents

Acknowledgements i

Contribution of authors ii

Table of Contents iii

List of Figures vii

List of Tables ix

1 Chapter 1: Introduction 1

1.1 Thesis structure 1

1.2 Electrical Recordings in the Brain 1

1.3 The Forward Problem: Contributions to Extracellular Ionic Current Flow 3

1.4 The Forward Problem: Generation of Field Potentials from Current Sources 5

1.4.1 Volume Conductor Theory 6

1.4.2 Mean-field Approximations of Maxwell‘s equations 8

1.5 The Inverse Problem: Current-Source Density Estimation 11

Preface to Chapter 2 14

2 Chapter 2: Laminar specific, orientation selective current-source density in cat area 18 17

2.1 Introduction 17

2.2 Methods 19

2.2.1 Animal preparation 19

2.2.2 Visual stimulation 21

2.2.3 Optical imaging and analysis of imaging data 21

2.2.4 Neurophysiology 22

2.2.5 CSD analysis 24

2.2.6 Measurements of CSD components 25

2.2.7 Spiking activity 26

2.2.8 Response latency 27

2.2.9 Layer Identification/Histology 27
2.3 Results 31
2.3.1 Laminar specific CSD 37
2.3.2 Validation of the use of standard CSD estimation 41
2.4 Discussion 44
2.4.1 Information transfer between cortical laminae 45
2.4.2 Validation of the standard CSD estimation method 47
Preface to Chapter 3 49
3 Chapter 3: 1-D current-source density (CSD) estimation in inverse theory: a unified framework for higher-order spectral regularization of quadrature and expansion type CSD methods 50
3.1 Introduction 51
3.2 Theory 54
3.2.1 Forward and inverse problem 54
3.2.2 Source parameterization 56
3.2.3 Discretizing the inverse problem 57
3.2.4 Solving the discretized inverse problem using SVD 61
3.2.5 Regularization 63
3.2.6 Choosing a prior 66
3.2.7 Methods for choosing the regularization parameter 67
3.2.8 Solution resolution 68
3.3 Methods 70
3.3.1 “True” CSD function $\boldsymbol{f}(\boldsymbol{z})$, forward model and noise 70
3.3.2 Priors 72
3.3.3 Reconstruction error 73
3.3.4 Comparison of estimation schemes 73
3.3.5 Simulation environment 74
3.4 Results 74

3.4.1 Noise sensitivity 74
3.4.2 Demonstration of the regularization 75
3.4.3 Comparison of regularization parameter selection methods 78
3.4.4 Regularization and resolution 80
3.4.5 Comparison of estimation schemes for each condition for “Sum of Gaussians” profile 81
3.4.6 Comparison of estimation schemes over all conditions for “Sum of Gaussians profile” 87
3.4.7 Effect of conductivity on the stability of the final ranking 91
3.4.8 Validation of final ranking with different CSD depth profiles 93
3.5 Discussion 97
3.5.1 Accuracy of the volume conductor model 99
3.5.2 Accuracy of forward model assumptions 101
Preface to Chapter 4 103
4 Chapter 4: Comparison of 1-D current-source density estimation methods for large diameter sources in a three-layer volume conductor 104
4.1 Introduction 105
4.2 Theory 108
4.2.1 Common Approaches to Standard CSD Estimation 108
4.2.2 Standard CSD as an integral equation 111
4.2.3 3-Layer Medium Parametrization 114
4.3 Results 115
4.3.1 A motivating example 115
4.3.2 Infinite plane versus uniform cylinder CSD estimation 120
4.3.3 Inverse CSD versus Standard CSD 126
4.3.4 Effect of Noise on Estimation 130
4.3.5 Three- versus Two-Layered Medium Parameterization 131
4.4 Discussion 135
4.4.1 How likely are extended sources? 136

4.4.2 Two- versus three-layered medium 137
5 Chapter 5: Discussion 138
5.1 Challenges and Limitations 140
5.1.1 Dealing with multiple time points 141
5.1.2 Dealing with multiple trials 141
5.1.3 Channel-Dependent and/or Non-White Noise 142
5.1.4 Open Challenges 144
5.2 Future work 145
Bibliography 148
Appendices 164
A. Anatomical and functional properties of cat primary visual cortex 164
B. Potential in a three-layered volume conductor 175

List of Figures

Figure 1-1: Schematic of inverse nature of current-source density estimation. 11

Figure 2-1: Functional identification of areas 17 and 18. 22

Figure 2-2: Post-processing of cortical tissue. 28

Figure 2-3: Orientation selectivity of CSD. 30

Figure 2-4: Statistical measures of the changes in the CSD response. 32

Figure 2-5: Modification of CSD response in relation to response strength.. 33

Figure 2-6: Statistical measures of the structural changes in CSD for different contrasts. 36

Figure 2-7: Comparison between CSD patterns elicited by grating stimuli vs. checkerboard stimulus. 37

Figure 2-8: Histology and layer correspondence of electrophysiological data. 38

Figure 2-9: Selective modification of CSD correspondence with layer information transfer. 40

Figure 2-10: Validation of standard CSD method. 42

Figure 3-1: Flowchart describing the spectral regularization framework. 53

Figure 3-2: Simulation of extracellular potential. 55

Figure 3-3: Flowchart describing the estimation scheme comparison process. 69

Figure 3-4: Demonstration of CSD estimation using the rCSD expansion method. 76

Figure 3-5: Demonstration of regularization parameter selection methods. 79

Figure 3-6: Effect of Regularization on Solution Resolution. 81

Figure 3-7: Distribution of labels within the top rank. 85

Figure 3-8: Distribution of error and lambda ratio for top 40 schemes in the final ranking. 90

Figure 3-9: Ranking validation for a conductivity mismatch between the forward and inverse model.. 92

Figure 3-10: Pairwise comparison of the similarity in rankings between different CSD depth profiles. 94

Figure 4-1: Demonstration of boundary correction method at conductivity jumps in recording medium.. 112

Figure 4-2: InfPlane forward model for a source of unit strength positioned at z=1 mm. 114

Figure 4-3: CSD estimation of a localized source overlaid with a laterally extended global activation.. 116

Figure 4-4: Demonstration of effect of source distance on recorded potential. 119

Figure 4-5: Difference between the InfPlane and UniCyl forward models. 122
Figure 4-6: Comparison of errors introduced by mismatch in source diameters. 125
Figure 4-7: Standard CSD estimator applied to the InfPlane and UniCyl forward model. 127
Figure 4-8: Noise sensitivity as a function of noise amplitude. 130
Figure 4-9: Effect of medium parameterization on potential. 132
Figure 4-10: Error introduced by a mismatch between forward and inverse medium parameterization .. 133
Figure A-1: View of the Lateral Geniculate nucleus and the adjacent MIN. 165
Figure A-2: Projections from LGN to cortex separated according to the different pathways. 166
Figure A-3: Parallel projections emanating from cat LGN. 167
Figure A-4: Feed-forward projections from LGN to area 17. 168
Figure A-5: Innervation profile of area 18 and 17. 169
Figure A-6: Model of intra-areal projections with 1 excitatory and 1 inhibitory population per layer. 171
Figure A-7: Laminar origin of cortico-cortical projections in the cat. 172
Figure A-8: Reciprocity of projections between area 17 and 18. 173

List of Tables

Table 3-1: Condition Number and Noise Amplification. 75

Table 3-2: Best ranked estimation scheme for each SNR and source diameter. 82

Table 3-3: Distribution of methods within the Top Rank. 84

Table 3-4: Top 40 estimation schemes over all tested conditions for “Sum of Gaussians” profile. 89

Table 3-5: The top 20 estimation schemes in ranking over all CSD depth profiles.. 96

Abstract

Understanding brain function entails disentangling the activity of different neuronal populations. Owing to advances in microelectrode fabrication technology, neuroscientists are now in position to simultaneously record extracellular potentials from multiple spatial locations in close proximity. In particular, a growing number of studies have been using linear electrode arrays to disentangle laminar processing in cortical and subcortical structures. However, in order to extract useful information about neural processing from such recordings, it is imperative to comprehend how extracellular potentials are related to the underlying neural activity.

Uncovering the relationship between extracellular potentials and neural activity has proven to be particularly challenging for the low-frequency part of the potential, which in intracranial recording is referred to as the local-field potential (LFP). However, the LFP is of particular interest in studying brain function because it is thought to reflect key integrative mechanisms of population activity which are mostly inaccessible to traditional spike recordings.

The difficulty in the interpretation of the LFP stems largely from the fact that electric potentials are a non-local measure of the neural activity due to volume conduction. To address this issue, it has thus become common practice in intracranial extracellular neurophysiology to estimate and analyze current-source density (CSD) instead of directly interpreting the LFP. This is beneficial because, in addition to mitigating the effect of volume conduction, correctly estimated CSD also represents the volume density of net transmembrane currents which generated the measured LFP.

In this thesis, we analyze the process of CSD estimation as it is applied to extracellular recordings from linear, i.e. one-dimensional, electrode arrays inserted perpendicularly to a layered medium. In Chapter 2, we demonstrate the usefulness of CSD estimation in the study of neural processing by providing novel insights into visual orientation processing in cat area 18. In particular, we use standard CSD estimation to show that excitatory activity is selectively relayed from the input layer IV to the superficial layers II/III, with CSD sinks extending more superficially for gratings of the locally preferred orientation and similar orientations.

Although we showed in Chapter 2 the usefulness of CSD estimation, our analysis also highlights known difficulties with the use of the standard CSD estimation method. We thus turn

in Chapter 3 to the study of the methods employed for 1-D current-source density estimation. In particular, we propose a novel estimation framework inspired by linear inverse theory with a special emphasis on higher-order spectral regularization methods. With this framework, we provide a unification of previously published inverse estimation methods under a common format. In addition, we introduce two novel CSD estimation methods. Moreover, the presented framework facilitates dealing with recording noise, a ubiquitous problem in solving inverse problems, by incorporating zeroth- and higher-order regularization methods with multiple regularization parameter selection methods.

Although the framework presented in Chapter 3 successfully unites all previously published inverse methods under a common formulation, it does not provide a clear connection between the inverse estimation methods and the standard CSD method used in Chapter 2. We addressed this connection in Chapter 4, by comparing the 1-D differential (standard) and integral (inverse) formulation of CSD estimation in the presence of laterally extended sources. By expressing the standard method as an inverse problem using the Green's function for the 1-D Poisson equation, we show that the two formulations converge to each other and could be considered similar for lateral source diameters exceeding ~2-5 mm. Moreover, we demonstrate that the standard 5-point CSD estimation is considerably less resistant to recording noise than its regularized integral formulation. We thus conclude that, with the exception of slightly reduced computation time, there appears to be no reason to opt for the standard CSD estimator over inverse estimators.

In summary, the core of this thesis demonstrates the usefulness of 1-D CSD in the study of neural processing and provides novel methods for its estimation. In the process, it also includes a thorough investigation of the stringent assumptions involved in the estimation with a focus on the possible misinterpretations brought about by their violation. The particular challenges and open questions surrounding 1-D CSD estimation are then revisited in the discussion and potential future work for their resolution is proposed.

Résumé

Pour comprendre le fonctionnement du cerveau il faut démêler l'activité des différentes assemblées neuronales. Grâce aux progrès dans la technologie de fabrication de microélectrodes, les neuroscientifiques sont en mesure d'enregistrer simultanément des potentiels extracellulaires à plusieurs emplacements spatialement de plus en plus rapprochés. Notamment, un nombre croissant d'études sur le traitement d'information au niveau mésoscopique (laminaire) dans les structures corticales et sous-corticales utilisent des réseaux de microélectrodes linéaires. Toutefois, afin de pouvoir extraire des informations utiles sur le traitement neuronal à partir de ces enregistrements, il est impératif de comprendre au niveau théorique comment les potentiels extracellulaires sont liés à l'activité neuronale sous-jacente.

Pour la partie à basse fréquence du potentiel, qui dans le cas d'enregistrements électriques intracrâniens est communément appelé potentiel de champ local (PCL), élucider cette relation entre les potentiels extracellulaires et l'activité neuronale s'est avéré être particulièrement difficile. Cependant, le PCL est d'un grand intérêt pour la compréhension du fonctionnement du cerveau, car on croit que ce signal reflète l'intégration d'information au niveau des assemblés neuronales. Et c'est exactement cette intégration d'information qui est inaccessible aux enregistrements de potentiels d'actions traditionnels.

La difficulté principale dans l'interprétation du PCL découle du fait que les potentiels électriques sont une mesure non-locale de l'activité neuronale dû à la conduction de volume. Ainsi, il est devenu pratique courante dans la neurophysiologie extracellulaire intracrânienne d'analyser la densité de courant source (DCS), plutôt que d'interpréter directement le PCL. Cela est bénéfique, car, en plus d'atténuer l'effet de la conduction de volume, la DCS, correctement inférée, représente la densité volumique des courants transmembranaires qui ont généré le PCL mesuré.

Dans cette thèse, nous analysons le processus d'estimation de la DCS tel qu'il est appliqué aux enregistrements extracellulaires linéaires. Nous nous concentrons plus précisément sur les enregistrements effectués à l'aide d'une matrice de microélectrodes insérées perpendiculairement dans le cortex cérébral. Nous démontrons d'abord, dans le deuxième chapitre, l'utilité d'estimer la DCS pour l'étude du traitement d'information neuronale en fournissant des résultats

expérimentaux nouveaux sur le traitement de l'orientation visuelle au niveau du cortex visuel primaire. En particulier, nous utilisons la méthode standard d'estimation de la DCS pour montrer que, dans l'aire 18 du cortex visuel du chat, l'activité neuronale excitatrice est sélectivement relayée aux couches neuronales superficielles (couches II/III) à partir de la couche d'entrée (couche IV). Ceci est démontré par le fait que les puits dans la DCS s'étendent plus superficiellement lorsque les neurones proches du site d'enregistrement sont stimulés avec des grilles orientées parallèles à leur orientation préférée que lorsqu'ils sont stimulés avec l'orientation perpendiculaire.

Bien que dans le deuxième chapitre nous démontrons l'utilité d'estimer la DCS dans l'étude du fonctionnement du cerveau, notre analyse met également en évidence certaines difficultés connues avec l'utilisation de la méthode standard d'estimation de la DCS. Pour cette raison, nous présentons dans le troisième chapitre une étude théorique détaillée des diverses méthodes d'estimation de la DCS. En outre, nous décrivons un cadre d'estimation nouveau, inspiré par la théorie inverse linéaire. Ce cadre met un accent particulier sur les méthodes de régularisation spectrale, et permet une unification de toutes les méthodes d'estimation inverses précédemment publiées sous une formulation commune. Un avantage de ce nouveau cadre est de faciliter grandement le traitement du bruit d'enregistrement. De plus, nous présentons deux nouvelles méthodes d'estimation de la DCS, elles aussi inspirées de la théorie inverse et donc aussi incluses dans notre cadre général.

Bien que ce dernier incorpore avec succès toutes les méthodes inverses précédemment publiées, il ne fournit pas de lien direct et clair entre les différentes méthodes d'estimation basées sur la théorie inverse d'une part, et la méthode standard employée dans le deuxième chapitre d'autre part. Nous analysons en détails ce lien dans le quatrième chapitre pour des sources de courant neuronales qui ont une large étendue latérale. Notamment, nous démontrons que la méthode standard peut être transformée en une méthode inverse en utilisant la fonction de Green de l'équation de Poisson. Avec cette approche, nous démontrons que les solutions des deux types de méthode convergent et que ceux-ci peuvent être considérées comme similaires pour des diamètres de sources supérieures à 2-5mm. De plus, nous démontrons que la méthode d'estimation standard (avec un noyau de cinq points) est beaucoup moins résistante au bruit d'enregistrement que sa formulation intégrale régularisée. Nous concluons donc, qu'à l'exception

d'une légère réduction du temps de calcul, il semble y avoir aucun avantage à utiliser la méthode d’estimation standard et que les estimateurs inverses sont préférables en toute situation.

En résumé, le noyau de cette thèse démontre l'utilité de l’estimation de la DCS dans l'étude du traitement d’information neuronale et fournit de nouvelles méthodes pour son estimation. Cette thèse comprend également une discussion approfondie des hypothèses contenues dans l’estimation de la DCS en mettant un accent particulier sur les conséquences possibles d’une violation de celles-ci. Finalement, la thèse se conclue avec une discussion sur les questions ouvertes autour de l’estimation de la DCS et des pistes de solution sont fournies.

1 Chapter 1: Introduction

1.1 Thesis structure

In the following thesis, I investigate the theory and application of one-dimensional current-source density (CSD) estimation in invasive cortical recordings. In the introduction, I first provide a broad overview of the use of CSD in the study of neural processing. I then proceed to a more detailed description of the theoretical problems encountered in CSD estimation.

Three manuscripts, each focusing on a particular aspect of CSD estimation, form the core of the thesis. In the first manuscript, I show the usefulness of CSD estimation in understanding brain function by presenting novel results on orientation selectivity uncovered by applying CSD estimation to laminar recording in cat primary visual area 18. In the second manuscript, I examine the application of higher-order regularization to CSD estimation. I also provide an elaborate theory section highlighting the inverse nature of the estimation problem. The third manuscript then connects the estimation methods from the second manuscript with the standard estimation method used in the first manuscript by comparing CSD estimation at small and large source diameters. Finally, the last chapter contains a critical discussion of the obtained results, guidelines on the use of CSD in an experimental setting, and an overview of potential future work.

1.2 Electrical Recordings in the Brain

Neurons are electrically excitable cells. Hence, it is not surprising that, next to anatomical studies, electrophysiology has established itself as the gold standard in investigating the functioning of the nervous system. In the most general sense, electrophysiology consists in measuring the effects of ion flow in biological tissue (Plonsey and Barr, 2007). In Neuroscience, this is performed on a wide variety of scales ranging from the study of single ion channel activity all the way to the investigation of whole brain processing as it is done in electroencephalography (EEG) and magnetoencephalography (MEG).

The various invasive electrophysiological techniques can be grouped in two broad categories: intracellular and extracellular measurements. Intracellular measurements are limited to recordings from single cells or even patches of membrane. They consist in measuring

fluctuations in currents or changes in voltage across the cell membrane by creating a conductive path with the interior of the cell either by directly inserting an electrode or by using patch clamp techniques (see e.g. (Li et al., 2004)). This allows for example to measure the voltage changes associated with the firing of an action potential.

In contrast to intracellular measurements, extracellular techniques don't disrupt the integrity of the recorded cells but instead employ electrodes with a conductive path to the extracellular medium embedding the neurons in order to measure broad-band field potentials (Nunez and Srinivasan, 2006). For example, using a high impedance intracranial microelectrode positioned close to a neuronal soma, this allows again the measurement of single-unit action potential signatures (found in the high frequency component ($\gtrsim 500$ Hz) of the recorded field potentials) as well as broad-band potential fluctuations generated by the summed activity of multiple neurons (Brette and Destexhe, 2012). However, it is quite common that, even in such localized extracellular recordings, action potentials from multiple surrounding neurons are picked up on the same electrode. Hence, we then speak of multi-unit activity (MUA) recordings.

A major difference between intra- and extracellular recordings is thus the number of neurons that contribute to the measured signal. In intracellular recordings only a single neuron is observed while extracellular recordings record from populations of neurons. Extracellular recordings are therefore popular experimental tools to study interactions between neural populations between the mesoscopic (few mm^3 of tissue) and the macroscopic (few cm^3 of tissue) scale (Nunez and Srinivasan, 2006). In this thesis, we will be exclusively concerned with 1-D mesoscopic extracellular recordings obtained from linear high impedance microelectrode arrays ($\sim 400\ \mu m^2$ surface area with a resistance of ~1 MΩ).

Although the activity of multiple neurons provides richer information about the functioning of the nervous system, it also complicates the analysis and interpretation of the measurements. This problem is particularly pronounced in the analysis of the low-frequency part ($\lesssim 500$ Hz) of extracellular field potentials because of their non-local nature brought about by a process called volume conduction (see next section). In invasive micro-electrode recordings (i.e. electrodes located in the extracellular space in the brain as used in this thesis), this low-frequency component is called the local field potential (LFP). The low-frequency field potential is also the

only signal that is measured with surface (EcoG) electrodes and with EEG macro-electrodes located on the scalp. Besides being more invasive, a defining characteristic of the LFP is that it samples from more localized neural populations than EcoG or EEG and gives thus access to neural processing at a mesoscopic scale (Nunez and Srinivasan, 2006). However, it is important to note that the exact sampling distance is still highly debated (Katzner et al., 2009, Xing et al., 2009, Kajikawa and Schroeder, 2011, Linden et al., 2011, Leski et al., 2013).

As mentioned above, a major difficulty in interpreting LFPs stems from their non-local nature. In a volume conductor such as the brain, the summed action from currents generated by neurons located hundreds of micrometers away from an electrode will contribute to the measured field potentials due to volume conduction. This represents a major caveat when trying to understand brain function because it makes it extremely difficult to use the recorded signals for disentangling the location and nature of active neural populations. Hence, to try to address this issue, it has become commonplace in extracellular neurophysiology to resort to source localization techniques. As its name implies, source localization simply refers to analysis tools which aim to determine the sources (in our case neural activity) of a measured effect (in our case recorded changes in electric potential). Whenever source localization is performed on LFP recordings, it is referred to as current-source density (CSD) estimation, the study of which is the content of this thesis.

As any source localization problem, CSD estimation is called an inverse problem because it aims to uncover the causal origin of a measured effect. It is the aim of this thesis to discuss CSD estimation in the context of 1-D cortical recordings. However, before being able to discuss the nature of this inversion, it is necessary to closely examine the associated forward problem describing how neural activity is thought to generate recordable field potentials. This step is crucial because the result and interpretation of the source estimation effort critically depends on the proper choice and understanding of the forward model.

1.3 The Forward Problem: Contributions to Extracellular Ionic Current Flow

Neuronal activity is the result of transmembrane ionic currents driven by electrochemical gradients across the cell membrane (Buzsaki et al., 2012, Einevoll et al., 2013). However, extracellular electrophysiology measures changes in electric potentials and not currents directly.

Hence, in order to study neural processing, we need to infer the neural activity from recorded field potentials. But as mentioned in the previous section, this is complicated by the fact that the field potential recorded at a given location is the linear superposition of the potentials generated by transmembrane currents of all excitable membranes within the volume conductor (Plonsey and Barr, 2007). This means that synaptic events, somatic activity, action potentials, and even slow fluctuations in glial cell potential contribute to the extracellular field (Buzsaki et al., 2012). Moreover, further complications arise from the fact that not all neuronal activity is equally represented in the measured signal because currents from multiple sources need to temporally overlap to generate detectable extracellular potentials. Hence, before delving into the mathematical models designed to formalize the generation of extracellular field potentials, it is informative to summarize what is known about the relative contribution from different neural sources.

It is generally assumed that, in physiological situations, post-synaptic activity is the major contributor to the extracellular potential because the slow synaptic currents from different sources most easily overlap in time (Mitzdorf, 1985). However, based on large-scale simulations it has recently been proposed that in addition to post-synaptic currents, active membrane currents also significantly affect the recorded LFP (Reimann et al., 2013). Moreover, even for synaptic currents, there are differences in the contributions from different cell types. Spatially aligned neurons with elongated structures such as pyramidal cells are thought to contribute more heavily to the field potentials than spatially disorganized or radially symmetric neurons such as stellate cells (Riera et al., 2012, Tenke and Kayser, 2012). This is because aligned elongated cells are in a so-called open field configuration in which dipolar current sources generate substantial ionic flow in the extracellular medium and thereby allows effects from neighboring neurons to easily add up (Einevoll et al., 2013). This is in contrast to the closed field configurations in radially symmetric cells where the active source and resulting return currents tend to cancel each other out. Hence, in addition to the temporal alignment between transmembrane currents, geometric factors also need to be considered when estimating the relative contribution of neuronal sources to the extracellular field potentials.

It was already mentioned that action potential signatures visible in the high frequency part of the measured field potentials can only be detected from neurons close to the recording

electrode. Since the largest action potential currents only last about 2 ms, their minimal contribution to the LFP is partly explained by the lack of synchronicity between action potentials from different neurons (Buzsaki et al., 2012). The lack of synchronicity can also result in a destructive interference between the depolarization and after-hyperpolarization phase of action potentials from different neurons. Nevertheless, it has been observed that action potentials exhibit a small bleed through effect that makes them detectable in the higher frequency bands of the LFP (Zanos et al., 2011). Moreover, the synchronous bursting of multiple neurons as well as the ensuing spike after-hyperpolarization can induce similar extracellular currents as coordinated synaptic activity (Buzsaki et al., 2012).

Although the various membrane conductance change mechanisms generating the neuronal currents are quite well understood, there is still considerable uncertainty about how exactly these transmembrane currents lead to the measured extracellular field potentials. This is due to the fact that brain tissue is a heterogeneous medium made of multiple membranes embedded in a highly conductive extracellular fluid filling less than 20 percent of the total volume (Brette and Destexhe, 2012). Hence, predicting the LFP from a set of current sources heavily depends on the approaches taken in modeling the whole medium.

1.4 The Forward Problem: Generation of Field Potentials from Current Sources

The literature contains two major frameworks for modelling the generation of extracellular field potentials: volume conductor theory and mean-field approximations to Maxwell's equations. We will briefly review both of these below since they form the basis of all current-source density estimation methods.

In both frameworks we start with assuming the validity of the electro-quasistatic approximation. This means that we decouple the electric and magnetic variables and thus allow capacitive but not inductive effects. It also assumes that fields propagate instantaneously (i.e. $c \to \infty$). Hence, quasistatics differs from statics in that it allows for time dependence in otherwise time independent laws by presuming that the dynamical system under study proceeds from one state to another as though it was static at every time point (Larsson, 2007). Under these assumptions, the electric field $\vec{E}$ is irrotational (i.e. curl-free) and Faraday's law simplifies to

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} = 0$$

Hence, the electric field can be expressed by the negative gradient of the electric potential φ

$$\vec{E} = -\nabla\varphi \tag{1}$$

From this point on, the two approaches differ from each other in the way they model the sources and the medium.

1.4.1 Volume Conductor Theory

In the volume conductor theory, the neuronal tissue is divided into an intracellular and extracellular space, both of which are three-dimensional conductive media (Plonsey and Barr, 2007). The extracellular current flow giving rise to the extracellular field potential is induced by transmembrane current sources and sinks, where source refers to positive current flowing from the intracellular to the extracellular space and sink to current flowing in the opposite direction. From the point of view of the extracellular space, this can be seen as current appearing (source) or disappearing (sink) from the medium at certain points. Moreover, in the volume conductor model the neuronal elements are treated as core conductors and thus the membranes are viewed as a system of sources and sinks currents (Nicholson, 1973). In this view, the current sources can be approximated by a volume current source density $C(x, y, z, t)$, which can be obtained by averaging the total transmembrane current flux within a small volume of tissue (Nunez and Srinivasan, 2006). $C(x, y, z, t)$ is therefore a scalar field of the total transmembrane current at time t within a small region around point $\vec{x} = (x, y, z)$. Using Kirchoff's current law, the current density vector field $\vec{J}(x, y, z, t)$ can then be related to its source density $C(x, y, z, t)$ by the divergence relation

$$\nabla \cdot \vec{J} = C \tag{2}$$

In a second step, the volume conductor theory approximates the extracellular space as a linear, possibly anisotropic, and purely ohmic (i.e. resistive) medium. These assumptions imply the validity of Ohm's law relating the free current density $\vec{J}$ to the electric field $\vec{E}$

$$\vec{J} = \overleftrightarrow{\sigma}\vec{E} \tag{3}$$

Combining Equations (1), (2), and (3), we find that the current source density C is related to the extracellular field potential φ by

$$\mathrm{C} = -\nabla \cdot \overleftrightarrow{\sigma} \nabla \varphi \tag{4}$$

Equation 4 thus represents the desired general equation relating the sources to the potential assuming the validity of the volume conductor model.

In many practical settings this relationship is further modified by assuming that the conductivity is homogeneous ($\overleftrightarrow{\sigma}(\vec{x}) = \overleftrightarrow{\sigma}$) and isotropic ($\overleftrightarrow{\sigma} = \sigma$) leading to an equivalent of the well-known Poisson equation

$$\mathrm{C} = -\sigma \Delta \varphi = -\sigma \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \varphi$$

However, to better fit recordings from layered media, the homogeneous medium assumption is sometimes slightly modified by employing a piece-wise continuous instead of a homogeneous conductivity. This is often the case in cortical recordings, where it allows to account for the large conductivity jump that occurs between gray matter and the artificial CSF solution which is customarily applied to the brain surface to avoid it drying out (Pettersen et al., 2006). The mathematics of this approach, namely the method of images, and its implications are discussed in more depth in Chapter 3.

It is however clear that even a two-layered medium is still a very coarse approximation to the actual brain geometry. Hence, a few studies have employed a three-layered medium to account for the layered structure of the hippocampus (Gold et al., 2006) or to simulate slice recordings on a micro-electrode array (Ness et al., 2015). However, the existing studies have so far only focused on situations where only the potential in the middle layer is used. In chapter 4, we expand on this approach by exploring the effect of including potentials from all three layers. This allows us to approximate cortical recordings in a slightly more realistic manner by including the conductivity jumps between the surface solution and gray matter as well as between gray and white matter.

Modeling the brain as a volume conductor composed of infinitely extended layers with homogeneous and isotropic conductivity is certainly an acceptable first order approximation. It

does however gloss over many known features of the brain. First, cortex as well as the hippocampus are layered structures in which most neuronal processes are predominantly oriented perpendicular to the layers. This is the case for most pyramidal cells which make up a significant proportion of cortical neurons (Binzegger et al., 2004). Hence, it is reasonable to expect that the vertical conductivity might non-negligibly differ from the lateral conductivity. Under certain conditions, this can sometimes be addressed by appropriately rescaling the spatial axes (Nicholson, 1973). A second salient feature of the brain is that the whole brain as well as most of its components exhibit extensive curvature. This introduces further errors when using a medium approximation made of infinitely extended planar layers. To address these limitations, Goto et al. studied the cortical conductivity profile in Wistar rats using a layered spherical volume conductor. They compared their results to the infinitely extended planar volume conductor and reported significant differences in conductivity across laminae as well as radial/tangential anisotropies (Goto et al., 2010). Moreover, they showed that discounting the curvature as well as approximating the medium with a homogeneous and isotropic volume conductor introduced errors in CSD estimation mainly for charge-unbalanced source distributions. This result is however not easily translated to other species like cats, monkeys or humans because the brains of these species are much more heavily convoluted and are thus only poorly approximated by a spherical volume conductor. To properly account for the effects of complex tissue geometries such as cortical sulci and gyri, it is necessary to model the volume conductor using the more complicated finite element models (FEM) built from anatomical magnetic resonance images (Wagner et al., 2014). However, since this is currently not a scalable approach (i.e. it cannot be easily done for every subject), it remains common practice to use the simplified volume conductors like the infinite planar medium approximation discussed above.

1.4.2 Mean-field Approximations of Maxwell‘s equations

In the volume conductor theory, the medium is assumed to be purely ohmic with conductivity varying only on a relatively macroscopic scale (e.g. conductivity variations across brain structures or across different laminae). This view has a long history dating back to the inception of the standard CSD estimation method by Pitts (Pitts, 1952). It received further support from Mitzdorf who concluded based on experimental evidence and theoretical considerations that any departure from a homogeneous and isotropic conductivity would only

have a minor effect on CSD estimation (Mitzdorf and Singer, 1980). Recently, this assumption has however come under renewed scrutiny. A major point of contention is that a purely resistive volume conductor cannot exhibit a frequency dependent conductivity (Bedard et al., 2004). It is however well known that the temporal power spectrum of the extracellular field potential exhibits a very characteristic $1/f^{\alpha}$ shape (with $\alpha = 1 - 2$), meaning that power at higher frequencies is significantly smaller than at lower frequencies (Buzsaki et al., 2012). Part of this low-pass filtering can be explained by dendritic filtering in passive neuron models or by neural network effects in which, for example, network synchrony and/or correlation could lead to self-organized power law distributions in synaptic activity (Milstein et al., 2009, Linden et al., 2010, Brette and Destexhe, 2012). Although all these effects certainly play a role in shaping the frequency spectrum of the extracellular potential, experimental evidence suggests that neuronal activity cannot always explain the observed frequency scaling (Bedard et al., 2006a). Instead, it has been proposed that the observed frequency filtering could be explained by a frequency-dependent macroscopic conductivity and permittivity resulting from the heterogeneous nature of the brain at microscopic scales (Bedard et al., 2006b, Bedard and Destexhe, 2009).

In the electro-quasistatic approximation, electric fields in biological media can drive current by two mechanisms: ohmic current due to the motion of free charges and capacitive displacement current resulting from the polarization of paired charges (Wagner et al., 2014). The ease with which free charges can flow through a medium in response to an electric field is quantified by the electric conductivity. Similarly, the electric permeability is a measure of the ease of polarization of paired charges (Griffiths, 2013).

In extracellular media, it is generally assumed that the displacement current is insignificant compared to the free charge current (Plonsey and Barr, 2007, Gratiy et al., 2013). But, it has been pointed out that this reasoning might be flawed (Bedard and Destexhe, 2013). Maxwell's theory of electromagnetism requires that the generalized current density ($\vec{J}_g = \vec{J}_f + \vec{J}_d$) composed of the free charge current density ($\vec{J}_f$) and the displacement current density ($\vec{J}_d = \frac{\delta \vec{D}}{\delta t}$, where $\vec{D}$ is the displacement field) must be conserved:

$$\nabla \cdot \vec{J}_g = \nabla \cdot \vec{J}_f + \nabla \cdot \vec{J}_d = 0$$

This relationship must be valid in any medium. However, it is only true in a homogeneous locally neutral medium that the free charge current density must be conserved independently of the displacement current density because local neutrality proscribes charge accumulation (i.e. $\nabla \cdot \vec{J}_f = -\frac{\partial \rho_f}{\partial t} = 0$, where ρ_f is the free charge density). Hence, considering the displacement current as negligible is equivalent to considering the medium as a homogeneous conductor in which charges behave like electrons in a metal. It is however clear that at microscopic, i.e. subcellular scales, the medium is certainly not homogeneous because it is mix of membranes, fluids, organelles and other cell components. Hence, the traditional assumption of the applicability of Kirchhoff's laws might not be valid in neuronal tissue. The inapplicability of Kirchoff's laws could be explainable by the fact that either displacement currents can simply not be ignored or because charges move at non-negligible finite velocities due to inertia and friction (Destexhe and Bedard, 2012). Especially the limited charge mobility could result in a delay between the ionic current induced by the opening of post-synaptic channels and the setting in of the extracellular and return current. Hence, the appearance of a membrane current might not be instantaneously equilibrated by a return current as required for the applicability of Kirchhoff's laws and as assumed in the traditional dipole model of the volume conductor theory. As a consequence, it allows for the possibility of strong monopolar components in the initial, non-stationary, phase of neuronal currents as experimentally observed by Riera et al. (Riera et al., 2012).

To resolve this issue, Bedard and Destexhe have proposed a macroscopic model of LFP generation which incorporates frequency-dependent conductivity and permittivity using a mean-field approach (Bedard and Destexhe, 2009). They show that the frequency-dependence of the macroscopic (i.e. spatially averaged) electrical parameters (conductivity and permittivity) is the result of the microscopically non-neutral nature of the cortical medium (despite the observed macroscopic neutrality). In particular, they predict that ionic diffusion currents are an important physical cause of the frequency dependence of the LFP and are thus a major determinant in the genesis of the LFP.

Although the debate over the exact nature of the medium and thereby over the important factors affecting the forward model is still ongoing, evidence for non-ohmic effects have recently

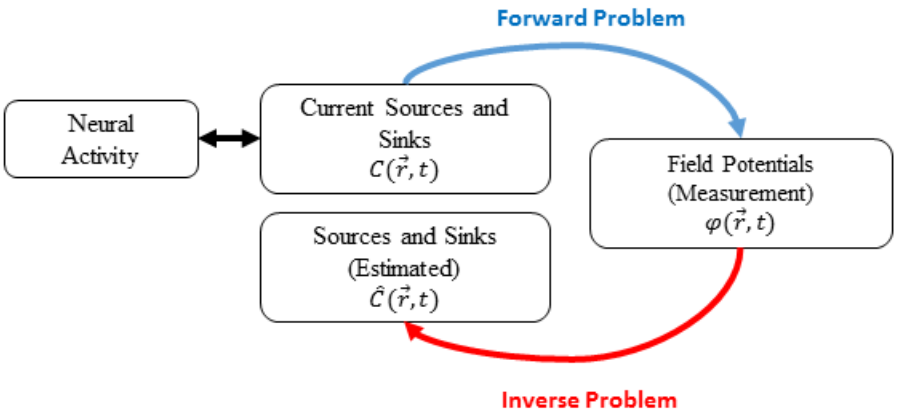

Figure 1-1: Schematic of inverse nature of current-source density estimation.

received interesting experimental support (Gomes et al., 2016). It is thus perfectly conceivable that in the near future the traditional forward model of field potential generation in brain tissue will need to be profoundly revised.

1.5 The Inverse Problem: Current-Source Density Estimation

As mentioned previously, CSD estimation is an inverse problem. The origin of this naming is schematically illustrated in Figure 1-1. The forward problem describes the physical mechanisms by which neural activity generates field potentials in the recording medium which has been discussed in detail in the previous two sections. One might think that once this process is understood, the estimation of neural activity from measured potentials will be a relatively simple and straightforward task. But it turns out that solving the inverse problem is surprisingly difficult. In this section, I provide a short introduction to the reasons behind the difficulties without going into much detail. This topic is then resumed in Chapter 3 where the particular solution of one-dimensional CSD estimation is analyzed in depth. Moreover, the discussion here is limited to CSD estimation in the context of the standard, i.e. resistive, piece-wise homogeneous and isotropic, forward model. The reason behind this limitation being simply the current lack of knowledge and data about the solution to the more complicated forward problems.

In a purely resistive, homogeneous, and isotropic medium, the field potential φ is related to the volumetric source density C by

$$\mathrm{C} = -\sigma\Delta V = -\sigma\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\varphi \tag{5}$$

As mentioned previously, this is arrived at by combining the differential law of charge conservation with Ohm's law and setting $\mathrm{C} = -\frac{\partial \rho}{\partial t}$. As an important side note, it is worth mentioning that this formulation is paradoxical because it is contradictory to Gauss's law since the potential is independent of Poisson's law in a homogeneous medium ($\varepsilon\nabla^2\varphi = -\rho$) and solely determined by the charge conservation law (Bedard and Destexhe, 2011). When including Gauss's law, we find that in order for the source density C to be different from zero, the so-called Maxwell-Wagner time (standing for the spreading time of excess free charge carriers in a material) needs to be non-negligible or the electric parameters need to display strong spatial variations. But both of these conditions violate the assumption of a homogeneous and purely resistive medium. However, despite this inconsistency, the standard CSD model has a long history of useful contributions to the study of brain function (see e.g. (Nicholson and Llinas, 1975, Mitzdorf, 1985, Schroeder et al., 1998, Swadlow et al., 2002, Csicsvari et al., 2003, Jin et al., 2008, Riera et al., 2012)).

The paradoxical nature of the problem aside, it is well known that neither Laplace nor Poisson-like equations such as Eq. 5 uniquely define the potential φ. It can however be shown using the *uniqueness theorem* that specifying the boundary conditions on all surfaces of the medium is sufficient to uniquely define the system (Jackson, 1999). Moreover, using Green's second identity together with Dirichlet boundary conditions at infinity, i.e. setting the potential to zero at infinity, it can be shown that the potential φ produced at position $\vec{r} = (x, y, z)$ by a unit strength point source located at $\overrightarrow{r'} = (x', y', z')$ is given by (Jackson, 1999, Brette and Destexhe, 2012)

$$\varphi(\vec{r}, t) = \frac{1}{4\pi\sigma|\vec{r} - \overrightarrow{r'}|} \tag{6}$$

Analogously, when dealing with a spatially extended source distribution $C(\vec{r}, t)$, the potential at point $\vec{r}$ and time t is given by the volume integral

$$\varphi(\vec{r},t) = \int_V \frac{1}{4\pi\sigma} \frac{C(\overrightarrow{r'},t)}{|\vec{r} - \overrightarrow{r'}|} dV' \tag{7}$$

where the integration limits are chosen to cover the full support of C.

Equation 7 is known as a linear Fredholm integral equation of the first kind and the right-hand side of Equation 6 is sometimes referred to as the data kernel (Wing and Zahrt, 1991). The kernel is what is modified when dealing with non-homogeneous media, while the integral structure of Equation 7 is preserved since Maxwell's equations are linear. It is however possible that a temporal integral needs to be added if the kernel has both spatial and temporal dependency (e.g. when conductivity is frequency dependent).

At this point, the task of CSD estimation is fully defined: We want to estimate $C(\vec{r},t)$ based on a limited number of data samples from $\varphi(\vec{r},t)$. Moreover, in an experimental setting, the data samples might be corrupted by noise. With this in mind, we are now ready to move on to the core of the thesis. We first demonstrate the usefulness of CSD in the study of brain function (Chapter 2). In particular, we will use the standard estimation method to study orientation selectivity in cat early visual cortex. It is important to mention that the traditional CSD estimation method is based on direct numerical estimation of the Laplace operator in Equation 5 without the transformation to an integral equation (Freeman and Nicholson, 1975, Nicholson and Freeman, 1975). In order to address known limitations in the standard CSD method, we therefore perform a detailed analysis of the methodological approaches to solve the CSD estimation problem in a one-dimensional setting (Chapter 3). In particular, we propose a novel estimation framework inspired by linear inverse theory with a special emphasis on higher-order spectral regularization methods. In Chapter 4, we then return to investigate the relationship between the integral and differential formulation of 1-D CSD estimation. Finally, since the focus of this thesis is on CSD estimation, I refrain here from providing an extended introduction into the anatomical and functional properties of cat area 18. Instead, the next section only provides the motivation as well as a short overview of the information necessary to properly contextualize the work presented in the next chapter. The reader interested in a more thorough discussion of the anatomy and functional properties of cat early visual cortex is instead referred to appendix A.

Preface to Chapter 2

Functional integration of activity within and between neuronal populations is essential for rapid and reliable information processing in the brain (Fingelkurts and Kahkonen, 2005). However, the complexity of brain connectivity makes investigation of functional integration a particularly challenging task. Analysis of anatomical connectivity suggests that the brain processes information in a highly parallel and distributed manner (Scannell et al., 1995, Hilgetag et al., 2000). This in turn imposes the question of how the brain integrates information across neuronal populations. We approached this question by investigating orientation selectivity in cat early visual cortex. In this section, I present a summary of how the lamina-specific topology of the anatomical connections of cat area 17 and 18 potentially influences their respective neuronal activity. For greater generality, I introduce both cat visual area 17 and 18 since these are sometimes grouped together to form the primary visual cortex due to their extensive interrelation (Payne and Peters, 2002).

The visual cortex is composed of multiple areas, each containing a distinct functional map of visual space (Payne and Peters, 2002, Frisby and Stone, 2010). For example, both areas 17 and 18 contain a clear retinotopic representation of visual space in addition to a superposed well-described orientation selectivity map (other maps such as ocular dominance, direction selectivity, color and spatial frequency selectivity are also found albeit to a weaker extent). Moreover, the neuronal units constituting these areas are heavily and non-randomly interconnected, both laterally and by inter-laminar connections, leading them to exert an influence onto each other. Using the example of cat area 17 and 18, it was shown that anatomical connections preferably link regions of similar orientation preference (Gilbert and Wiesel, 1989).

Cat visual area 17 and 18 both receive direct innervation from the retina via the lateral geniculate nucleus (LGN) (Payne and Peters, 2002). Specifically, there are three distinct pathways called X, Y, W that project from the retina to the cortex. Anatomical studies have shown that the Y-pathway innervates both area 17 and 18 while the X-pathway probably only innervates area 17 (Payne and Peters, 2002). Additionally, the W-pathway also innervates both areas, as well as higher order visual areas. However, because W-cell responses to visual stimulation are sluggish, I will not discuss this pathway further.

Y- and X- cells are optimally driven by different visual stimuli: The Y-pathway responds stronger to higher temporal but lower spatial frequencies while the X-pathway prefers mid-range to high spatial frequencies but low temporal frequencies. Additionally, both pathways are contrast selective. This makes it possible to preferably activate one area over the other by manipulating the stimulus. Moreover, responses can also be graded by modifying the contrast.

The X-pathway forms the major input of area 17 and innervates extensively layer 4A (upper layer 4) and also layer 6A. On the other hand, Y-cells innervate lower layer 4 in area 17 and the full extent of layer 4 in area 18. Additionally, they also send processes to layer 6 (Payne and Peters, 2002).

Finally, area 17 and 18 extensively project to each other in a lamina-specific manner. Projections from area 17 to area 18 mainly originate in layer 2/3 and terminate in layer 2/3 while projections from 18 to 17 originate in layer 2/3 and 5 and terminate in layer 2/3 and 5 (Symonds and Rosenquist, 1984). The projections between area 17 and 18 have also been shown to preferably connect regions of same retinotopy and similar orientation preference (Salin et al., 1995, Shmuel et al., 2005).

An important item to add here is that it is thought that there are two general classes of neuronal projections, namely driver and modulatory projections (Sherman and Guillery, 1998). The difference between the two, as the name implies, is that driver connections can make their post-synaptic partner neuron fire an action potential while modulatory connections only create sub-threshold effects. It has been proposed that the projection from the LGN to cortex is of the driver kind (Sherman and Guillery, 1998). Recently, studies on the auditory and visual cortex of the mouse have shown evidence for two distinct classes of synapses: Class 1B, a sub-class of driver synapses, and Class 2, assumed to be identical to the modulatory synapses found in the thalamus (Covic and Sherman, 2011). Class 1B is differentiated from Class 1A, which is equated with the previously found driver synapses in the thalamus, in that Class 1B shows graded EPSP amplitude as a function of input intensity while Class 1A shows an all-or-none EPSP amplitude pattern (Covic and Sherman, 2011). It was found that the projections between mouse V1 and V2 contain both Class 1B and Class 2 type of synapses in both directions, i.e. from V1 to V2 as well as from V2 to V1 (De Pasquale and Sherman, 2011). At this point it is not really clear how these

findings can be translated to the cat visual cortex due to the lack of comparable data. However, although it is still unknown whether cat cortico-cortical projections are predominantly driving (Class 1) or modulatory (Class 2), based on the results from mouse visual cortex it seems reasonable to assume that both types of projections could also be found in the cat.

On functional grounds, it is interesting to note that cortico-cortical projections between areas 17 and 18 are sufficiently strong to drive each other without LGN input to one of the areas. By destroying layer A of the LGN which provides the only X-input from the contralateral eye, Malpeli et al. showed that activity in the superficial layers of area 17 could only be silenced if area 18 was inactivated simultaneously with the LGN (Payne and Peters, 2002). This indicates that connections between area 17 and 18 could potentially play an important role in shaping their mutual response. Nonetheless, studies which inactivated either supragranular or infragranular layers of area 18 failed to see a consistent effect on area 17 responses (Alonso et al., 1993a, b, Martinez-Conde et al., 1999). This again imposes the question of the use of these extensive mutual connections.

2 Chapter 2: Laminar specific, orientation selective current-source density in cat area 18

Orientation selectivity in neurons in the visual cortex is the result of a dynamic process comprising a weighted combination of feed-forward, local, and feedback signal integration. Although much effort has been devoted to understanding the emergence of orientation selectivity, the exact mechanisms are still incompletely understood. In particular, it is still unclear how orientation specific signals are selectively relayed across cortical layers. Since current-source density (CSD) is hypothesized to depict synchronized synaptic activity, it should be well suited to further our understanding of the role of intracortical signal processing in the emergence of orientation selectivity. Surprisingly, previous studies reported that the spatiotemporal CSD pattern is virtually invariant to the properties of the visual stimulus. Here we revisit this paradox by comparing 1-D laminar CSD responses in cat visual area 18 to gratings varying in contrast and orientation. Contrary to previous reports, we found clear orientation selective components in the CSD responses. In particular, the primary sink in layer IV extended to supragranular layers when the optimal orientation was presented. Similar spatio-temporal dynamics were also observed in response to a non-orientated stimulus but were absent in the non-optimal orientation. This orientation-specific modification in the extent of the sink was also observed in response to lower contrast stimuli suggesting that the observed effect is not due to a difference in overall neural activation but rather reflects differences in cortical signal processing. Hence, we present evidence that feature selection in early cortical processing involves selective information transfer between cortical laminae.

2.1 Introduction

Lower visual areas comprise functional modules arranged as cortical columns. The functional selectivity shown by neurons in such columns remains approximately unchanged along tracks orthogonal to the cortical surface (Hubel and Wiesel, 1962). In several species, these cortical columns are arranged in complex maps showing gradual changes in functional selectivity across the cortical surface (Swindale, 2004, White and Fitzpatrick, 2007). Several functional properties are organised in cortical columns and maps, including ocular dominance (Hubel and Wiesel, 1968), orientation selectivity (Hubel and Wiesel, 1962, Bonhoeffer and

Grinvald, 1991, Shmuel and Grinvald, 2000) and direction of motion (Albright, 1984, Malonek et al., 1994, Shmuel and Grinvald, 1996, Weliky et al., 1996, Swindale et al., 2003).

Our existing knowledge on the emergence of orientation selectivity and its columnar organisation makes it a classic model system for understanding the cortical processing of sensory input and information transfer between layers. Current models of orientation selectivity differ in the specific contribution of thalamocortical versus intracortical connections. Models based on thalamocortical projections (Hubel and Wiesel, 1962, Palmer and Davis, 1981, Ferster, 1986, 1988, Reid and Alonso, 1995, Chung and Ferster, 1998) propose that orientation selectivity results from the precise spatial arrangement of inputs from the thalamic Lateral Geniculate Nucleus (LGN) onto cortical cells. In contrast, models that emphasize cortical involvement propose that orientation selectivity results from intracortical connections (Ben-Yishai et al., 1995, Somers et al., 1995, Ringach et al., 1997, Sompolinsky and Shapley, 1997, Adorjan et al., 1999, Shapley et al., 2007). It has also been argued that thalamo-cortical inputs may be the basis for orientation preference, but are insufficient for explaining orientation selectivity (Shapley et al., 2007). The nature of the cortical mechanisms that contribute to orientation selectivity remains unclear. In particular, the importance of laminar-specific cortico-cortical synaptic interactions remains elusive.

A major challenge in elucidating laminar-specific synaptic interactions is to collect spatially localized signals resulting from synaptic activity. This is particularly difficult because extracellularly recorded field potentials are volume conducted and are thus a non-local measure of neuronal activity. A potential solution to this problem is to estimate laminar current-source density (CSD) from the recorded local field potentials (LFP). Given certain assumptions, CSD analysis provides the pattern of local transmembrane currents that generate the LFP recorded in the extracellular medium (Nicholson, 1973, Nicholson and Freeman, 1975, Mitzdorf and Singer, 1978). In contrast to the volume-conducted LFP (Kocsis et al., 1999), CSD is locally confined (Mitzdorf, 1985). It therefore enables the localization of synchronized synaptic activity integrated over all processes taking place near the probe.

Intracortical processing is expected to contribute significantly to orientation selectivity (Ben-Yishai et al., 1995, Somers et al., 1995, Ringach et al., 1997, Sompolinsky and Shapley, 1997, Adorjan et al., 1999). In addition, inter-laminar interactions are expected to play a crucial

role in cortical computations. Therefore, one would expect that the spatiotemporal laminar distribution of current sources and sinks would vary with the orientation of grating stimuli. However, although cortical layers differ anatomically and functionally, previous studies concluded that functional parameters of the visual stimuli are reflected only in minor modulations of the CSD (Mitzdorf, 1985). The use of CSD for analyzing intracortical processing underlying orientation selectivity or any other functional selectivity remains questionable, as is the contribution of intracortical processing to orientation selectivity. However, to our knowledge, no analysis of CSD specifically associated with the response to oriented gratings has been reported. Therefore, whether CSD reflects the intracortical processing of orientation selectivity remains unknown.

Here we aim to test whether laminar specific synchronized synaptic activity as estimated by CSD shows orientation selective responses. We further aim at testing the hypothesis that layers 2/3 receive more potent synaptic input when processing the preferred orientation relative to the non-preferred orientation. Cat area 18 has several features that make it suitable for such an investigation. Both areas 17 and 18 receive their main input from layers A and A1 of the LGN in their respective layers IV and VI. However, area 18 receives a more uniform projection comprised of a strong input from the Y pathway and little to no input from the X pathway (Hubel and Wiesel, 1962, Hoffmann et al., 1972, Stone and Dreher, 1973, LeVay and Gilbert, 1976). This makes area 18 a simpler model to study intracortical processing in comparison to area 17. The responses to oriented stimuli in area 18 are organised in columns, i.e. orientation selective cells with similar preferred orientation are present in all layers (Hubel and Wiesel, 1962, Gilbert, 1977). This includes the layers that receive input from the LGN, which is not the case for other species such as the macaque monkey for example (Blasdel and Fitzpatrick, 1984, Hawken and Parker, 1984).

Parts of the results presented here have been previously presented in an abstract form (Kropf et al., 2010).

2.2 Methods

2.2.1 Animal preparation

All procedures were approved by the animal care committees of the Montreal Neurological Institute and McGill University, and were carried out with great care according to

the guidelines of the Canadian Council on Animal Care. Eight adult cats weighing 2.5–3.5 kg were used. Cats were premedicated with acepromazine maleate (Atravet®, Boehringer Ingelheim, 1.0 mg/kg, S.C.) and glycopyrrolate (Sandoz Canada, 0.01 mg/kg, I.M.). Fifteen to thirty minutes after this injection, deep anesthesia was induced by inhalation of 5% of isoflurane (Baxter Medical) mixed with O_2/NO_2 (50:50). The level of isoflurane was gradually lowered and maintained at 2% for the initial stages of the surgery. Blood oxygen saturation and heart rate were monitored using pulse oximetry (Nonin Medical, Inc.), while end tidal CO_2 pressure was monitored with a capnometer (Capnomac Ultima, Datex). Depth of anesthesia was determined by verifying lack of response to clamping the inter-digital web of the posterior paws and by monitoring the heart rate. Following endotracheal intubation and cephalic vein cannulation, muscular relaxation was induced by administering gallamine triethiodide (2%, I.V.; Sigma-Aldrich). Upon cessation of spontaneous breathing, the animal was artificially ventilated with a mixture of O_2/NO_2 (33%/67%) using a respiratory pump (Ugo Basile, 6025). Lidocaine 2% (Lidocaine Neat, Pfizer) was administered at all points of pressure and incision. The cat was then placed in a stereotaxic frame (David Kopf instrument) modified for allowing visual stimulation. End-tidal CO_2 partial pressure was kept between 30 and 38 mm Hg by adjusting the rate and stroke volume of the respiratory pump. The core temperature was maintained at 38°C by a feedback-controlled heating pad (Homeothermic Blanket Systems, Harvard Apparatus). The animals were continuously infused with 5% Dextrose in lactated Ringer's injection solution (Abbott Laboratories) mixed with gallamine triethiodide (50:50, 20 mg/kg/h).

At the end of the surgical procedures, isoflurane was switched to 0.5-1.0% halothane (2-Bromo-2-chloro-1,1,1-trifluoroethane, Sigma-Aldrich), since halothane is less detrimental to neuronal responsiveness (Villeneuve and Casanova, 2003). During imaging and recordings, the animals were infused with 5% Dextrose in lactated Ringer's mixed with gallamine triethiodide (50:50, 20 mg/kg/h), and were ventilated with O_2/NO_2 (33%/66%) mixed with 0.5-1.0% halothane. Pupils were dilated with phenylephrine hydrochloride 2.5% (Mydfrin®, Aventix Animal Health). The eyes were protected using contact lenses of appropriate refractive power, such that they were focused on a tangent screen at a distance of 30 cm. The blind spot of each eye was back projected on the screen and the area centralis was located accordingly (Bishop et al., 1962).

2.2.2 Visual stimulation

All stimuli were written in Matlab 7.10.0.499 (R2010a) using the Psychophysics Toolbox 3.0.8 extensions (Brainard, 1997) run on a Quad-CPU MacBook Pro 6.2 (Mac OS 10.6.7) with a NVIDIA GeForce GT 330M graphic card. The stimuli were presented on a 19 inches Dell 1908FPt monitor set at 60Hz refresh rate. The monitor was positioned at a distance of 30cm from the eyes, covering a field of view of 53º by 64º. The luminance was gamma corrected using Eye-one display II (X-rite). The mean of the maximum luminance of the monitor was set at ~30 cd/m^2.

2.2.3 Optical imaging and analysis of imaging data

Optical imaging was pursued at the beginning of each experiment to functionally identify area 18 and to locate orientation selective modules. Based on atlas and stereotaxic coordinates (Reinoso-Suárez, 1961), a circular chamber with inner diameter of 20 mm was implanted, centered on Horsley-Clarke coordinates AP +4 and ML 0, over portions of areas 17 and 18 (Fig. 2-1A, Ai). The cortex was illuminated uniformly with green light (546 nm, BP 30). Optical imaging was pursued using a VDAQ 3001 system (Optical Imaging Inc., Rehovot, Israel) equipped with a twelve-bit CCD camera (Dalsa 1M60, Canada) and a macro lens (Nikon, AF Micro Nikkor, 60 mm, 1:2.8 D). Drifting oriented sine wave gratings were presented for 8s followed by an inter-stimulus interval of 10s, in which the next oriented grating stimulus was presented in a stationary manner. The parameters of the sine wave gratings were optimized for eliciting responses in area 17 (spatial frequency (SF) 0.5c/deg, TF 2Hz) and area 18 (SF 0.15c/deg, TF 4Hz) (Movshon et al., 1978b). Two oriented gratings, spanning the orientation space at a resolution of 90 degrees were presented. Each of these oriented gratings drifted in two opposite directions orthogonal to their orientation. Orientation maps were computed by summing the raw responses to the same orientation and opposite directions and dividing the result by the sum of responses to the orthogonal orientation. Panels Bi and Bii of Figure 2-1 show orientation modules in area 17 and area 18, respectively. Dark modules indicate regions that were more responsive to horizontal gratings, whereas bright regions indicate a preference for vertical orientation.

To map the responses to stimuli optimized for eliciting activations in area 17 (18), all cortical images obtained during the presentation of the corresponding stimuli were summed and

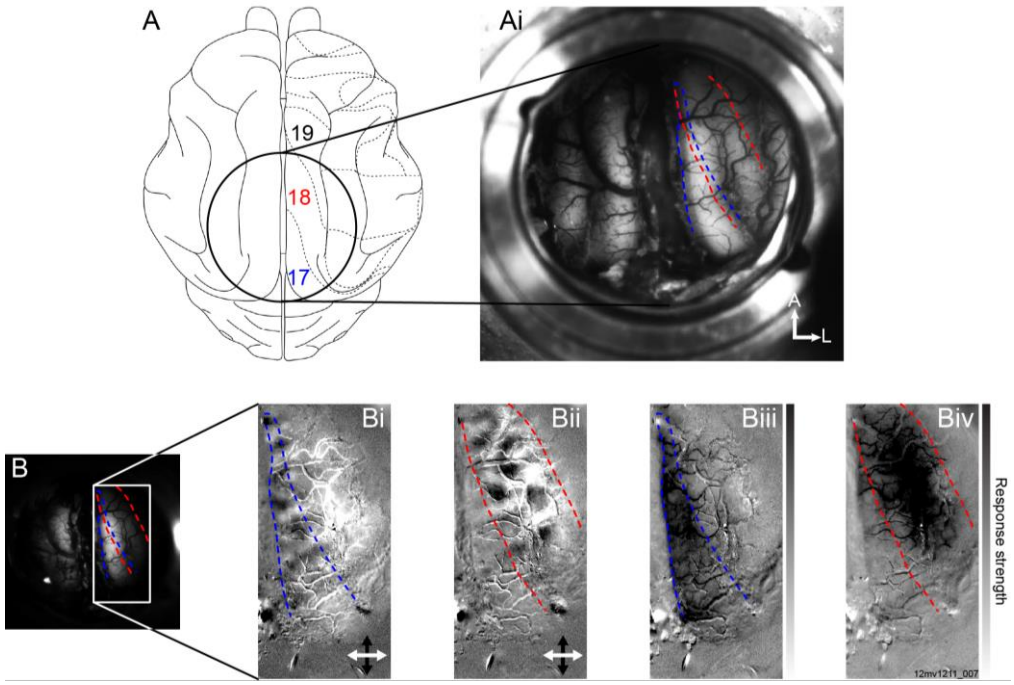

Figure 2-1: Functional identification of areas 17 and 18. Panels A and Ai present the reference atlas coordinates that were used and the corresponding position of the recording optical imaging chamber, respectively. Panel B shows analysis results of optical imaging of intrinsic signals. Panel B presents the cortical image under green light while Panels Bi, Bii, Biii, Biv present differential orientation maps for areas 17 and 18 and activation maps for areas¬ 17 and 18, respectively. Area-specific optimal spatiotemporal parameters were used to elicit responses in each area. Orientation modules are confined to area 17 and area 18 in Panel Bi and Bii, respectively. Similarly, activation maps presented in Panel Biii and Biv illustrate the functional segregation between the two areas. Blue and red dashed curves represent the functional boundaries of area 17 and 18, respectively.

divided by the sum of cortical images obtained during the presentation of a blank gray stimulus. The mappings of areas 17 and 18 are presented in Panels Biii and Biv, respectively. Dark regions indicate locations for which the stimuli elicited stronger response than the blank condition. Functional delineations of area 17 (blue dashed curves) and area 18 (red dashed curves) were made based on their activation maps.

2.2.4 Neurophysiology

Neurons at the centers of orientation modules and those located close to orientation singularities (‘pinwheels’) receive synaptic inputs from neurons with a narrow and wide distribution of orientation preferences, respectively (Yousef et al., 2001, Schummers et al., 2002). In order to study the orientation selectivity of the CSD response, we needed data from columns with a clear orientation preference. We therefore targeted each electrode insertion to the

center of an orientation module in area 18, identified based on the differential analysis of the optical imaging responses described above.

Neurophysiological recordings were performed using multi-channel recording probes (A32, 32 channels, intervals between adjacent contacts $100\ \mu m$, impedance 1-2 MΩ, contact surface area $413\ \mu m^2$, NeuroNexus Technologies) and a Tucker-Davis Technologies RZ2 multi-channel neurophysiology system. The probes were inserted roughly $2.5\ mm$ deep into cortex perpendicular to the cortical surface. The initial screen position was first determined by listening through a speaker to the neuronal activity on multiple channels in response to a computer-generated white/black bar on a black/white background. In order to quantitatively locate the aggregate receptive fields and to center the screen on their centroid, we pursued a reverse correlation analysis of sparse noise stimuli, composed of black and white squares flashed individually for one frame ($16\ ms$) on a gray background. Once the position of the screen was set, we optimized the insertion depth of the probe in order to assure coverage of the gray matter. To this end, flashing ($16\ ms$) full screen checkerboard stimuli with SF of $0.15\ c/deg$ were presented. We then pursued CSD analysis, which guided modification of the insertion depth such that the fast, high-amplitude sink response (assumed to occur in layer 4) was recorded approximately at the middle of the probe.

The angle of the insertion relative to the cortical surface was evaluated by computing the preferred orientation as a function of cortical depth. In cat area 18, preferred orientation is organized in columns, with the preferred orientation remaining approximately unchanged with increasing cortical depth (Hubel and Wiesel, 1962; Shmuel and Grinvald, 1996). The spike rate during the presentation of oriented gratings was used to determine the preferred orientation of all responding channels. We then verified that no systematic gradient in preferred orientation was observed with increasing cortical depth. In order to be included in the main analysis, the SD of the preferred orientations of action potentials recorded across contacts had to be smaller than $22.5°$.

In order to increase the stability of the recordings, in the majority of the recordings (7 out of 8 animals) the exposed cortex was covered with warm ($40°\ C$) 1% agar (Sigma-Aldrich). During recordings in which the position of the probe relative to the surface of cortex was monitored by imaging (see section on histology), the brain was covered with HBSS (Invitrogen)

to prevent dryness. The recordings made with agar and HBSS did not show any differences in their neurophysiology profiles. We therefore included both for further analysis.

Once the position of the screen (centered on the RFs) and the position of the probe (depth and angle) were adjusted, optimal spatio-temporal parameters for oriented gratings were assessed. Multiple contrasts (2.5%, 5%, 10%, 20%, 40%, 80% and 100%) were tested along with eight orientations, spanning the orientation space at a resolution of $22.5°$. Each of these eight oriented gratings drifted in two opposite directions. Stimuli were presented for $1\ s$ with $1\ s$ inter-stimulus interval. The inter-stimulus condition was a gray screen of luminance equal to the mean luminance of the grating stimuli (~30 cd/m^2). Responses to stimuli of the same orientation but opposite-direction were averaged together. In order to compare the cortical responses of oriented gratings to those elicited by non-oriented stimuli, we presented a flashing ($32\ ms$ on, $1\ s$ off) full screen checkerboard pattern of optimal spatial frequency.

The extracellular potentials were recorded at a sampling rate of 24414.0625 Hz. They were band-pass filtered with cut-off frequencies at 1 and 150 Hz using a 10th order FIR filter and down sampled to 1 kHz to obtain the LFP. In addition, the raw signals were separately band-pass filtered between 300 and 3000 Hz to obtain multi-unit activity (MUA). Custom MATLAB scripts (version R2012a; The Mathworks) were used to analyze neurophysiological signals.

2.2.5 CSD analysis

To estimate the CSD, we used the standard method based on a 5-point approximation of the second spatial derivative (Freeman and Nicholson, 1975)*:*

$$\frac{\delta^2\varphi}{\delta z^2} = \frac{\varphi(z+2h) - 2\varphi(z) + \varphi(z-2h)}{4h^2}$$

where h is the interval between electrode contacts ($100\ \mu m$). The one-dimensional CSD was then computed using the following equation:

$$\tilde{I}(z) = -\sigma\frac{\delta^2\varphi}{\delta z^2}$$

where z is the direction perpendicular to the cortical surface, φ the recorded LFPs and σ is the assumed homogeneous cortical conductivity ($0.3\ S/m$) (Pettersen et al., 2006). The method of Vaknin was used to estimate the CSD at the top and bottom 2 electrodes (Vaknin et al., 1988).

2.2.6 Measurements of CSD components

Prior to any quantification, we excluded signals with origin in dysfunctional channels (10 times larger low and high frequency fluctuations in comparison to other channels) which we occasionally observed. We replaced the missing signals by interpolating the recorded LFP using cubic splines prior to estimating the CSD. To evaluate the changes in the CSD profile between different conditions, the amplitudes of the CSD responses were normalized for each condition separately by dividing by the maximum absolute value of the CSD. Since our probe covered 3.1 mm, some channels were in cortex while others were in white matter or outside the brain. Hence, the normalization was performed by taking into account only channels that responded to grating and checkerboard stimuli, i.e. the ones within the gray matter. Where available, the appropriateness of this selection was also confirmed via histological reconstruction.

The cortical thickness in cat area 18 is variable (Beaulieu and Colonnier, 1985, van der Gucht et al., 2001). This made it difficult to align multiple CSD profiles from different recording sessions and animals for the purpose of group analysis. Therefore, pair-wise comparisons of measurements obtained from single insertions were pursued (see results for details).

In order to compare the CSD response across conditions, an analysis time window was set. The CSD response amplitude is higher at the beginning of the response when synchronized activity occurs. Therefore, we focused the analysis on the CSD signal in a time window starting in the range of $20 - 55\ ms$ after stimulus onset and ending $100 - 160\ ms$ after stimulus onset. The starting time and duration of the analysis window was determined for each insertion separately. To avoid biasing the analysis with a manual window selection, we used the following data driven approach: First we averaged the CSD responses across all orientations. We then took the absolute value of the CSD and averaged all the channels within the cortex (W_{CSD}). Subsequently, we high-pass filtered W_{CSD} using a least-square linear filter (cutoff frequency at 5 Hz). We then defined a baseline signal (W_{base}) by taking the 0.25 s prior to stimulation from the filtered W_{CSD}. To determine the window length (Win), we thresholded W_{CSD} in the $200\ ms$ following stimulus onset ($Win = W_{CSD} \geq mean(W_{base}) + 5 * std(W_{base})$). Finally, we added a buffer of $10\ ms$ before and $30\ ms$ after the window to get the final window length. In order to match the analysis of the action potential response to that of the CSD response, the same time window was also used for the analysis of the spiking activity.

For the analysis of the different CSD compartments, the raw CSD was temporally averaged (C_{win}) over the time window (Win) described above. To determine the spatial extent of the sources and sinks, we calculated the mean and standard deviation of the baseline CSD (C_{base} = CSD over 0.25 s prior to stimulation) separately for each channel in cortex. We then computed a threshold vector comprising a threshold for each contact $C_{vec} = |mean[C_{base}]| + 2 \cdot std[C_{base}]$. To obtain a scalar threshold (C_{thresh}) independent of depth, we spatially averaged C_{vec} over the channels within the cortex ($C_{thresh} = mean[C_{vec}]$ over channels within cortex). We also computed the standard deviation of C_{vec} over the channels within the cortex to identify outlier data. Insertions that showed a standard deviation exceeding 0.1 were deemed too noisy and were excluded from further analysis. The spatial extent of the primary sink response was then defined as $sum[C_{win} \leq -C_{thresh}]$ ($C_{win} \geq C_{thresh}$ for the extent of the sources). In other words, we defined any response exceeding twice the baseline fluctuation as a meaningful response. The analysis was also performed with a threshold of 1 standard deviation which did not alter the results presented herein.

2.2.7 Spiking activity

Band-pass filtered data (300-3 kHz) was post-processed for automatic spike detection, sorting and clustering. The spike-sorting pipeline is based on the open-source MATLAB package Wave_Clus (Quiroga et al., 2004). From the filtered traces, signal with amplitude higher (lower, when negative) than a threshold determined as 3 standard deviations of the baseline signal were identified as spikes. A segment of each detected spike (0.75 ms before the peak and 2.25 ms after the peak, 64 data points for each segment) was cut and saved for further processing. All spike segments were then aligned according to their respective peaks following interpolation of the spike shape with cubic splines. To avoid double detections of the same spike, only spikes separated by at least 1.5 ms were included. Feature extraction was performed with wavelet coefficients analysis of the spike events by computing a Kolmogorov Smirnov test of Normality. Finally, superparamagnetic clustering was performed to isolate spike classes.

Upon further inspection, the data showed two different clusters of spikes in each of the recording contacts; one cluster included the majority of the detected spikes. Therefore, this cluster was used for the analysis related to spiking activity. It should be noted that we cannot stipulate that the spike sorting and clustering analysis resulted in single cell activity, but rather an

improved, more specific MUA signal. Simple MUA involves crude thresholding of the signal; thus, every point above the threshold is counted as a spike. The first part of the clustering workflow involved thresholding for spike detection, which was followed by sorting and clustering analysis that removed unreliable detected signal by performing wavelet analysis. Spike sorting and clustering removed a large part of spurious signal that occurs when only thresholding is applied. To quantify and compare the spiking activity with the CSD pattern, we analyzed the spiking activity from the same channels as those used for CSD analysis, thus excluding signals from channels positioned above cortex or in white matter. In addition, we computed the spike rate across the same time window used for quantifying the CSD.

2.2.8 Response latency

In order to locate contacts within specific cortical layers, we computed the latencies of responses to the visual stimuli. Latencies vary significantly between cortical layers in the visual cortex, with the shortest latencies observed in layer IV (Best et al., 1986). Latencies of responses were calculated using the cumulative sum technique (Ellaway, 1978, Vogels and Orban, 1990, Raiguel et al., 1999, Ouellette and Casanova, 2006).The analysis was carried out on the mean peri-stimulus time histograms (PSTH) with a bin resolution of 1msec for each grating orientation. The mean and standard deviation of the spontaneous spike rate over the 150-ms periods preceding stimulus onset was computed separately for each orientation. Next, the mean spontaneous activity was subtracted from the mean PSTH separately for each grating condition. Then, a cumulative sum of the result was calculated. The onset of the response was defined as the first bin after stimulus onset where the response exceeded the mean spontaneous discharge rate by three standard deviations conditioned that it was followed by at least two successively increasing bins. The latencies were then averaged across the different orientations and the standard deviation was computed.

2.2.9 Layer Identification/Histology

In control experiments, probes were dipped in DiI (Invitrogen) prior to insertion into cortex. This fluorescent dye was used to stain the tracks left by the recording probe for optimal registration of neurophysiological responses with neuroanatomical structures (DiCarlo et al., 1996) and to properly locate the recording sites in relation to the optical imaging data. Panels A and Ai of Figure 2-2 shows images taken at the end of the experiment after the skull was

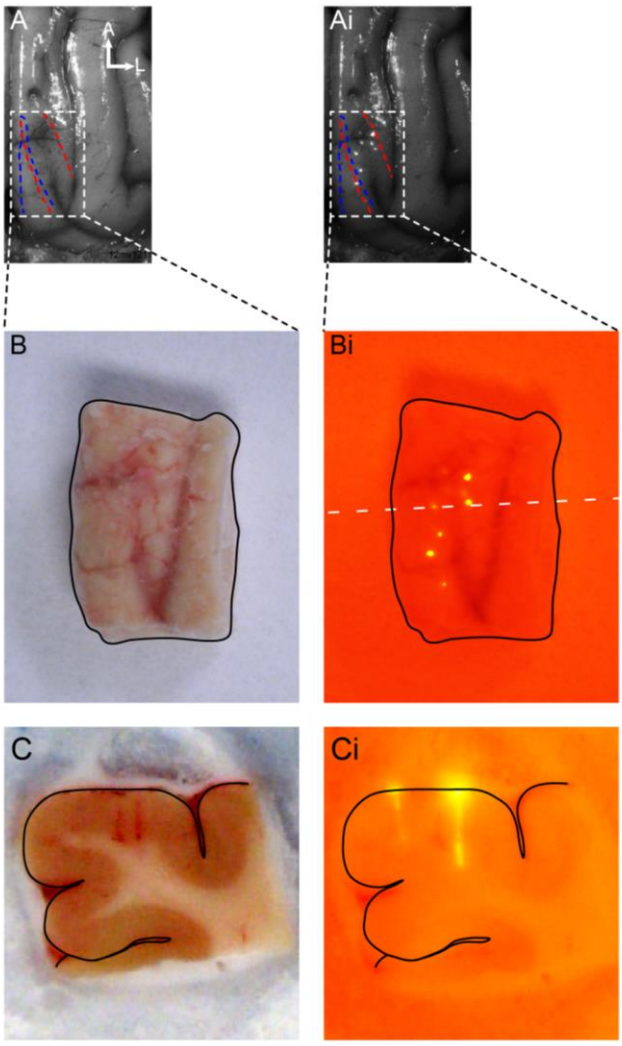

Figure 2-2: Post-processing of cortical tissue. In A, a reference image taken after the skull was removed. In Ai, the same cortical region is now imaged under illumination optimized for detecting DiI fluorescence. The dashed red and blue curves represent the functional localization of areas 17 and 18, respectively, based on the optical imaging data presented in Figure 2-1. In B, the frozen block used for histology presented as a reference. In Bi, the same frozen block is now imaged under illumination optimized for detecting DiI fluorescence. The white dashed line illustrates the plane of cut taken on the microtome of the slice presented in Panel C as reference and Panel Ci under illumination optimized for DiI.

removed (areas 17 and 18 are delineated with blue and red dashed curves respectively). Panel A shows the posterior part of the right hemisphere under green light, while Panel Ai shows the fluorescence produced by the DiI left at each recording insertion (excitation filter 535(BP 50); emission filter 565 (HP)). Fluorescence at each insertion site appears as a bright circular region. Insertions in the transition zone between areas 17 and 18 or in area 17 were excluded from further analysis. Imaging the fluorescence sites at the exact same plane used for the functional

optical imaging allowed precise alignment of electrophysiological recordings to the functionally identified cortical areas. Panels B and Bi present images of a smaller frozen block of tissue taken before histological processing, from the regions shown in Panels A and Ai. Panel B presents the block under white light and Panel Bi presents the same block under illumination and post filter optimized for DiI visualisation. The insertion sites can still be identified. Thus, the fluorescence allows selection of specific brain slices for further staining as shown in Panels C and Ci of Figure 2-2. These panels present images of the brain tissue cut along the dashed white line in Panel Bi. The images were taken while the tissue was on the microtome, under white light and under illumination optimized for DiI fluorescence, respectively. The electrophysiological findings were then interpreted in relation to the histology in Figures 2-8 and 2-9.

To confirm the insertion depth of the probe, we imaged the insertion site and the part of the probe which was above the cortical surface at high resolution. This allowed counting the number of contacts that remained above the cortical surface, and helped determine the position of the surface relative to the probe contacts. In addition, we pursued two electrolytic lesions by means of current (50µAmp for 10s) administered through the bottom most contact (contact 32) and a contact 1 mm more superficially (contact 22). This electrolytic lesion protocol was not conducted in all experiments as it can be highly detrimental to the recording probes.

At the end of the experiments, the animals were euthanized by an intravenous overdose of pentobarbital sodium (Euthanyl, 240 mg/ml/kg). The brain was removed from the skull and a block of tissue from the recorded area was immersed in paraformaldehyde (4%) for 24h, followed by cryo-protection in sucrose (30%) for the successive 24h. Forty micrometer thick serial sections (coronal plane) were cut using a sliding microtome (Leica SM 2000r). Brain slices showing DiI fluorescence were post-processed for cytochrome oxidase (equine heart, Sigma-Aldrich) in order to identify cortical layers. For identification of cortical laminae we used the criteria from Price (1985) who stipulated that cytochrome oxidase primarily stains layer IV and VI. The electrolytic lesions allowed for accurate registration of the recording contacts in relation to the cortical laminae. We confirmed that probe insertion were in area 18 by combining functional information from optical imaging, optimal spatio-temporal parameters used for electrophysiology and cytoarchitectonic standards described by Price (1985).

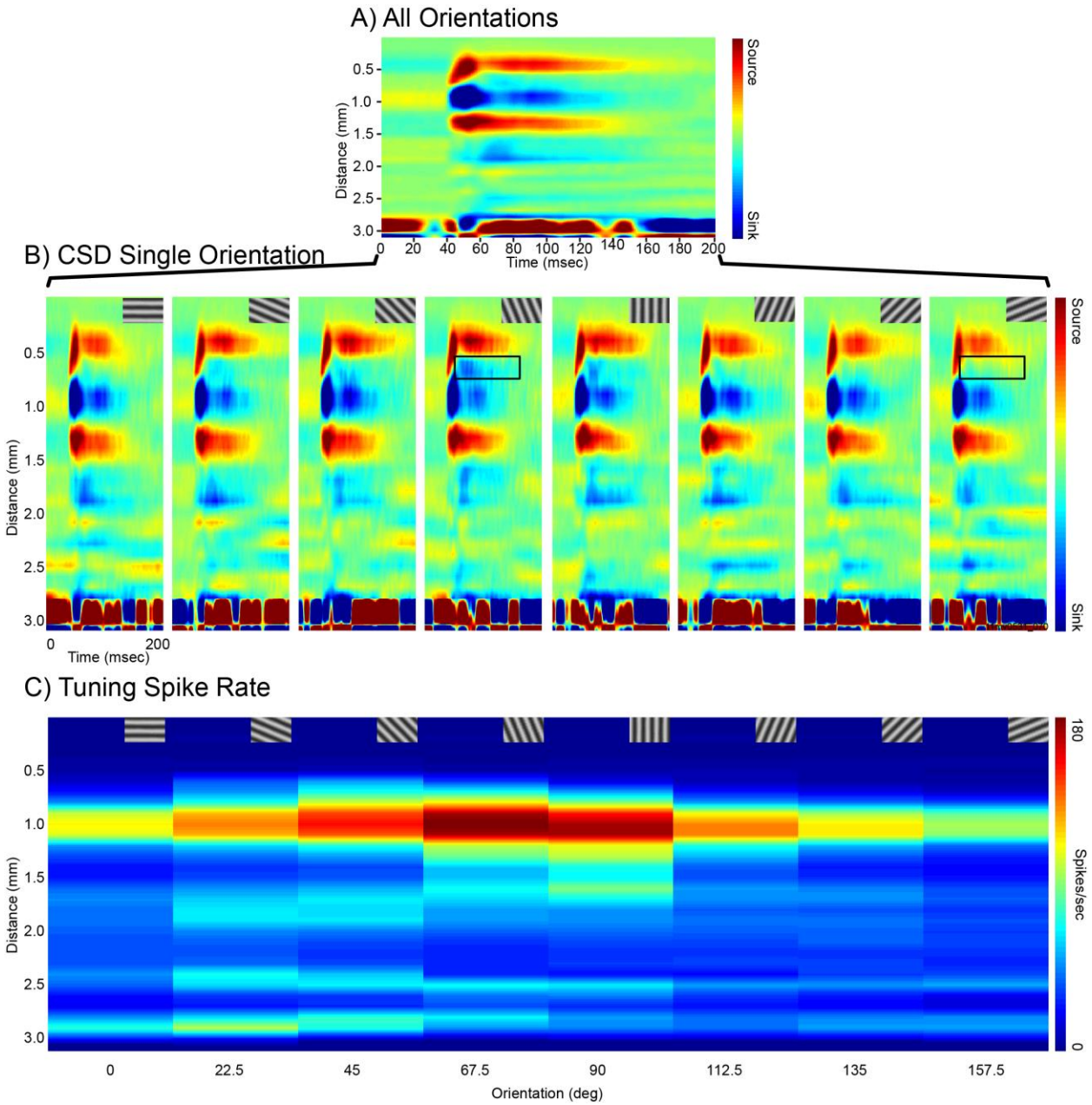

Figure 2-3: Orientation selectivity of CSD. Panel A shows the typical CSD response obtained by averaging the responses to 8 different oriented grating stimuli. Panel B presents the individual responses to each of these grating stimuli. Note that in response to the preferred orientation of this particular column (~67.5º), the primary sink (between coordinates 0.8 and 1.1 mm on the vertical axis) extends towards more superficial contacts. The two dark rectangles show the region where this extension takes place in response to the optimal (67.5º) but not to the non-preferred (157.5º) orientation. Panel C shows the spike rate response obtained from the same insertion, with the preferred orientation (highest spike rate response) at 67.5º.

2.3 Results

Here we show that the lamina resolved cortical CSD response is sensitive to the orientation of grating stimuli. Moreover, we show that this selectivity cannot be explained by differences in the amplitude of neuronal response elicited by the preferred orientation. Instead it is a signature of the neuronal processing of orientation *per se*. We found that this selectivity is specific to sub-compartments of the CSD profile: a specific segment of the CSD response is orientation selective while others are invariant.

Oriented grating stimuli elicited reliable and reproducible CSD responses in cat area 18. Figure 2-3A shows the CSD response obtained in one run, averaged over 20 trials and 8 orientations. The CSD profile presents a clear pattern in response to orientated gratings, which has been replicated in all of our recordings. The overall pattern of the response is very similar to that reported in previous studies in other species or systems (Mitzdorf and Singer, 1978, Mitzdorf, 1985, Happel et al., 2010, Maier et al., 2010, Maier et al., 2011). The CSD response clearly presents a fast, high-amplitude sink (blue; here centered on 1.0 mm along the vertical axis; termed 'primary' in the remaining text), bounded by two adjacent high-amplitude sources (red), one deeper and the other more superficial relative to the primary sink.

The focus of our present study is to investigate whether there are functional differences between CSD responses to different oriented stimuli. Figure 2-3B shows the normalized CSD responses to each of the 8 oriented gratings from the same run presented in Figure 2-3A. Figure 2-3C presents the orientation tuning of the spiking activity for the same insertion. The preferred orientation based on the recorded MUA is approximately 67.5° (Figure 2-3C). Careful examination of Figure 2-3B shows that the spatial extent of the primary sink (in blue) varies across orientations. The orientation-dependent modulation of the CSD pattern is made clearer by comparing the CSD within the two black rectangles on the fourth panel (optimal orientation) and the eighth panel (non-optimal orientation). In particular, the sink extends to more superficial contacts in response to the preferred orientation.

This spatial extension in the sink response was consistent across the majority of our experiments. When observed, it always extended more superficially relative to the fast high-

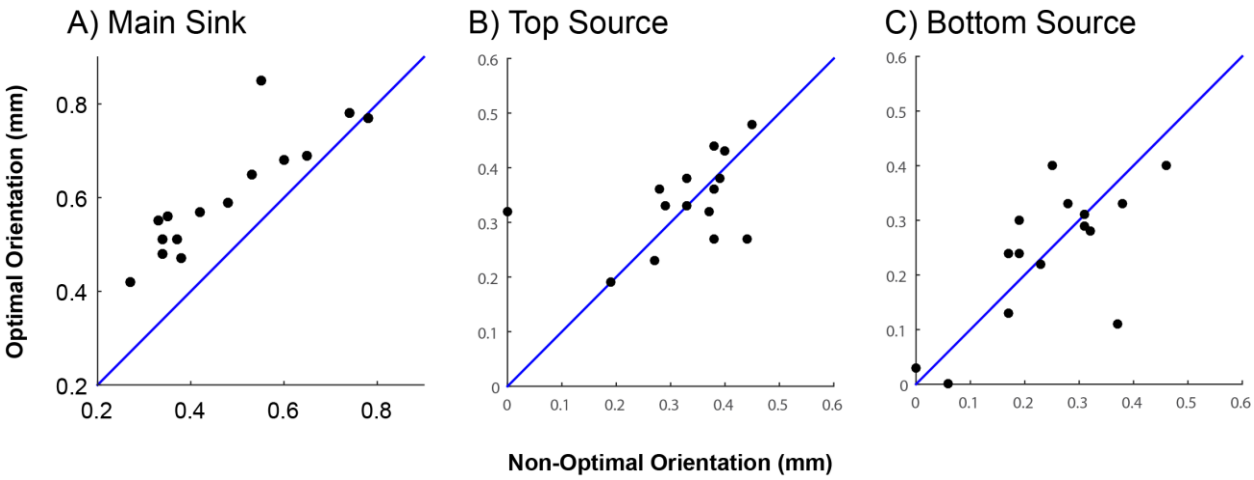

Figure 2-4: Statistical measures of the changes in the CSD response. The length of the different compartments of the CSD are compared in a pair-wise manner (preferred vs. non-preferred orientation) for each insertion. In A, comparison of the lengths of the primary sinks elicited by the preferred and non-preferred orientations. In B, comparison of the length of the source adjacent to- and more superficial relative to the primary sink. In C, comparison of the length of the source adjacent to- and deeper than the primary sink. The length of the primary sink associated with the preferred orientation is significantly different ($p < 0.001$) than that of the non-preferred orientation. In contrast, the superficial and deep sources associated with the preferred orientation are not significantly different than their respective counterparts associated with the non-preferred orientation.

amplitude sink component. In contrast, the two sources (red) surrounding the main sink remained approximately invariant to the orientation of the grating.

In order to confirm these qualitative observations, we conducted a quantitative evaluation of the variations in the CSD responses to different orientations. Figure 2-4 presents three scatter plots of different measurements of the CSD responses to optimal (along the vertical axis) versus non-optimal orientation. All measurements were based on the CSD responses following interpolation, normalization, averaging, and thresholding as described in the methods section. The average primary sink response to the optimal orientation was 0.13 mm longer than that to the non-optimal orientation (Figure 2-4A; main sink optimal orientation= 0.61 ± 0.13 mm (mean±STD), main sink non-optimal orientation= 0.48 ± 0.16mm, computed over 15 penetrations obtained from 10 cats; $p < 0.001$, paired two tailed t-test). Figures 2-4B and 2-4C show that this selective modification in the CSD pattern was specific to the primary sink. In contrast to the modification observed in the primary sink, the superficial source response to the optimal orientation and to the orientation orthogonal to it were not significantly different (Figure 2-4B; average difference 0.014 mm; $p = 0.62$). Similarly, the spatial extent of the deeper source

response to the optimal orientation and to the orientation orthogonal to it were not significantly different (Figure 2-4C; average difference -0.005mm; $p = 0.83$). These findings confirm that the sink responses to the optimal and non-optimal orientations were significantly different, and that this orientation selective CSD response feature is specific to the primary sink component. These findings further demonstrate that the CSD response can show selectivity and specificity in reflecting the spatiotemporal properties of a visual stimulus.

It is however conceivable that the observed differences in the CSD response might be solely explained by the higher response amplitude elicited by the optimal orientation compared to the non-optimal orientation. To test this hypothesis, we compared the CSD responses to oriented gratings of two contrast levels. The rationale behind that approach was that if the pattern of the CSD response merely depends on the amplitude of the response, two sets of grating stimuli presented at different contrasts should present different CSD patterns. On the other hand, if the CSD response pattern reflects intra- cortical processing underlying orientation selectivity rather than the amplitude of the response, CSD responses obtained with high and low contrast gratings should be similar. In particular, we have chosen a contrast level for which the average spiking response to the optimal orientation at lower contrast was similar to the spiking response for the non-optimal orientation at 100% contrast.

Figure 2-5 and 2-6 present the results of comparing the CSD responses to oriented gratings of two contrast levels. The top panel of Figure 2-5A illustrates an example of orientation tuning as a function of cortical depth, reflecting spike rate for eight oriented gratings presented at 100% contrast. The preferred orientation associated with this insertion was 157.5º. The bottom panel of Figure 2-5A shows the corresponding normalized CSD responses. The differences in the pattern of the CSD primary sinks in response to the preferred (157.5º) and non-preferred (67.5º) orientations can be clearly observed. Consistent with our findings from Figure 2-3B, the sink response to the preferred orientation extended more superficially than that to the non-preferred response. A similar layout is used in Figure 2-5B, which presents the responses from the same electrode insertion to oriented gratings presented at 40% contrast. Note the decreased spike rate in gratings presented at 40% contrast (top panel of Figure 2-5B) compared to 100% contrast (top panel of Figure 2-5A). Despite the decrease in response amplitude, the difference in the patterns of the CSD responses to the preferred and non-preferred orientation at 100% contrast (lower

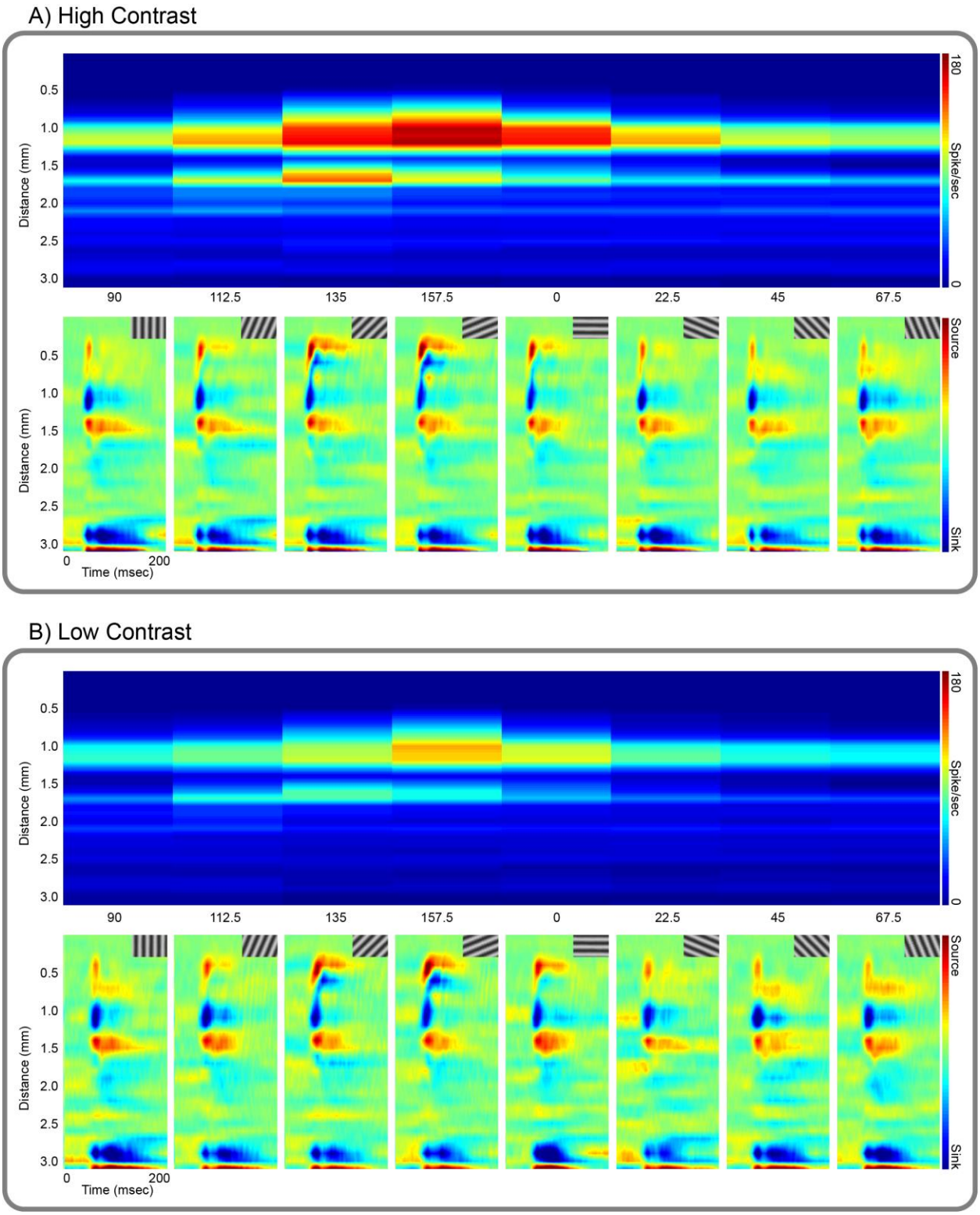

Figure 2-5: Modification of the CSD response pattern in relation to response strength. Comparison of the CSD structure in response to oriented grating stimuli of two different contrasts. (continued on next page)

panel of Figure 2-5A) prevailed under stimulation at 40% contrast (bottom panel of Figure 2-5B). Thus, the CSD pattern reflected the orientation selectivity and not the overall response amplitude. This suggests that the differences in the CSD patterns cannot be explained by the strength of the responses of the column, but are rather reflective of orientation processing *per se*. If the pattern of the CSD was determined by the strength of the response, a shorter sink would have been observed for a grating of the optimal orientation presented at 40% contrast, since it elicited an action-potential response of lower amplitude. This is clearly not the case, since the spatial extent of the CSD sinks elicited by gratings of 40% and 100% contrast were similar, while the spike rate response to the stimuli of lower contrast dropped in comparison to the responses of higher contrast.

These observations are quantified by the statistical analysis presented in Figure 2-6, which shows a scatter plot of the sink extension for the optimal orientation at 40% contrast vs the non-optimal orientation at 100% contrast. The primary sink response to the optimal orientation at 40% contrast was 0.07 mm longer than that to the non-optimal orientation at 100% contrast (main sink optimal orientation at 40% = 0.54 ± 0.19 mm (mean ± STD), main sink non-optimal orientation at 100% = 0.47 ± 0.18mm, computed over 10 penetrations obtained from 5 cats; p = 0.04, paired, one-tailed t-test). The difference in the spike rate was not significantly different between these two conditions (p = 0.77, paired, two-tailed t-test). Thus the difference in the sink pattern observed between the optimal and the non-optimal orientation cannot be explained by differences in response amplitude but instead is related to the underlying intra-cortical processing involved in orientation selectivity.

So far, we have shown that the spatial extent of the sink response to the optimal orientation is larger than that to the non-optimal orientation. However, it is not yet clear whether this response is also larger than that to non-oriented stimuli; or alternatively, that the sink

Figure 2-5 (continued from previous page) The top panel in A presents the response strength (rate of action potential response) to grating stimuli of 100% contrast. The bottom panel in A shows the CSD response to each orientation. The primary sink extends more superficially in response to the preferred orientation (157.5º). Section B presents the responses to oriented grating stimuli of 40% contrast in a format similar to that presented in A. Note the decrease in response strength (top panel) relative to that shown at 100% contrast. Despite this decrease in response amplitude, the primary sink extends more superficially in response to the preferred orientation.

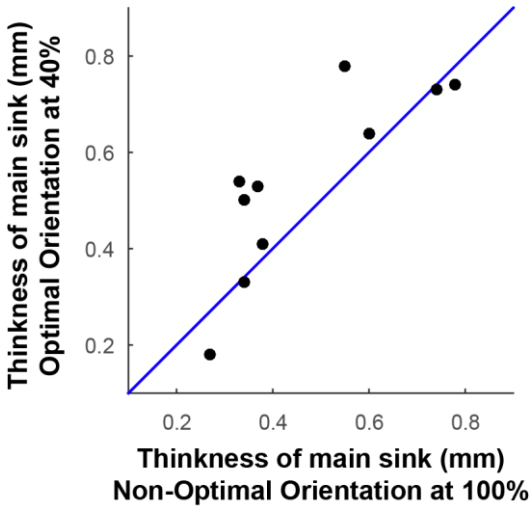

Figure 2-6: Statistical measures of the structural changes in CSD for different contrasts. The distribution of the sink length in response to the optimal orientation at low contrast (40%) along the vertical axis vs. the sink length in response to the non-optimal orientation at high contrast (100%) along the horizontal axis. The primary sink associated with the preferred orientation at low contrast was significantly longer (p= 0.04, one-tailed paired t-test) than that of the non-preferred orientation at high contrast. The spike rates were not significantly different between these conditions (p= 0.77, two-tailed paired t-test). Therefore, the modification of the structure of the CSD quantified in Figure 2-4 cannot be explained by the larger response amplitude elicited by the preferred orientation, but rather reflects the processing involved in orientation selectivity.

response to the non-optimal orientation is smaller than that to a non-oriented stimulus. A third alternative may suggest that a combination of both these mechanisms takes place. In order to address this question, we compared the CSD patterns elicited by the optimal and non-optimal orientations to that obtained in response to a flashing checkerboard of the same spatial frequency. We used the checkerboard stimulus as a reference because it is a more isotropic stimulus which should not activate orientation selective processes to the same extent as oriented gratings.

Figure 2-7 presents a comparison between the extent of the sink responses to the optimal (A) (and non-optimal (B)) orientation vs the response to a checkerboard presented at the same luminance contrast. We found that the primary sink response of the optimal orientation (0.59 ± 0.12) was not significantly different than that to the checkerboard stimulus (0.56 ± 0.18; p = 0.33; Figure 2-7A; computed over 14 penetrations obtained from 10 cats). In contrast, the spatial extent of the primary sink response to the grating of non-preferred orientation (0.46 ± 0.15mm) was significantly shorter than that to the checkerboard stimuli (p=0.03, two-tailed paired t-test;

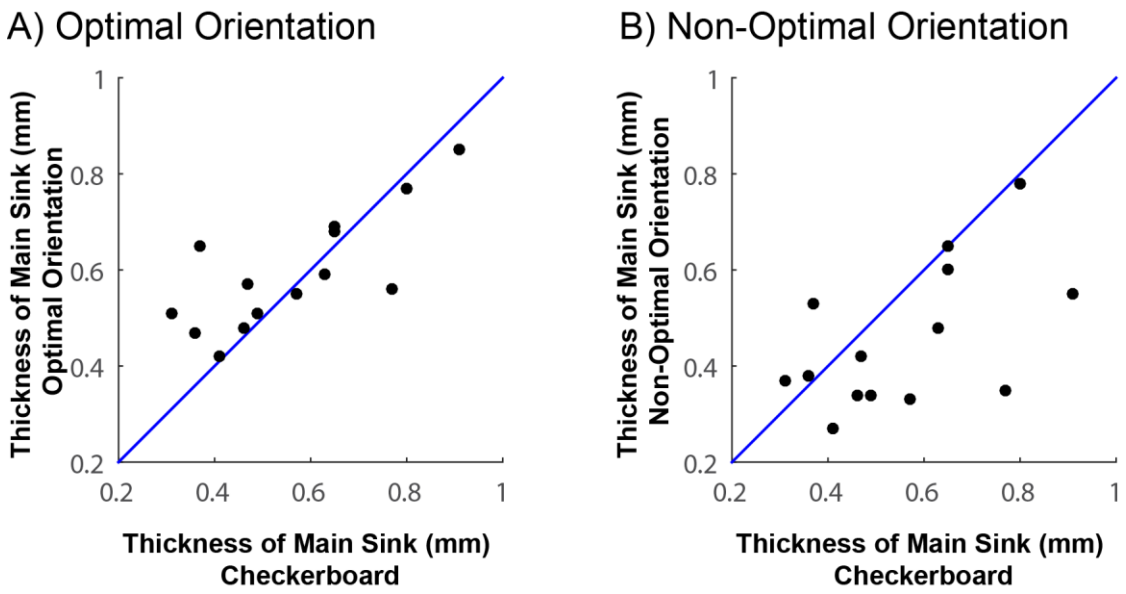

Figure 2-7: Comparison between the CSD patterns elicited by grating stimuli vs. checkerboard stimulus. In A, the length of the primary sink in response to the preferred orientation is shown as a function of the primary sink in response to flashing checkers. The sink elicited by the optimal orientation is not significantly longer than that elicited by the checkerboard condition (paired t-test, p = 0.33). In contrast, the sink elicited by the non-preferred orientation is significantly shorter than that elicited by the checkerboard (paired t-test, p = 0.03).

Figure 2-7B). The source responses to the preferred/non-preferred oriented gratings were not significantly different from those elicited by the checkerboard stimulus; non-preferred orientation vs. checkerboard: p=0.58; Bottom source: preferred orientation vs. checkerboard: p=0.62; non-preferred orientation vs. checkerboard: p=0.59, two-tailed paired t-test). These findings show that relative to a non-oriented stimulus, the sink is shorter when the non-optimal orientation is presented, and is approximately equal when the optimal orientation is presented. Hence, this contributes to a better understanding of the mechanisms involved in orientation selectivity since it supports previous reports which suggested that active intra-cortical processes contribute to orientation selectivity (Crook and Eysel, 1992, Ben-Yishai et al., 1995, Douglas et al., 1995, Somers et al., 1995, Adorjan et al., 1999).

2.3.1 Laminar specific CSD

So far, all the findings were discussed only in relation to differences in the spatial extents of sink and source responses, but nothing has been said about their location within cortex. In order to better understand laminar processing, it is important to identify how these changes relate to information transfer between cortical layers.

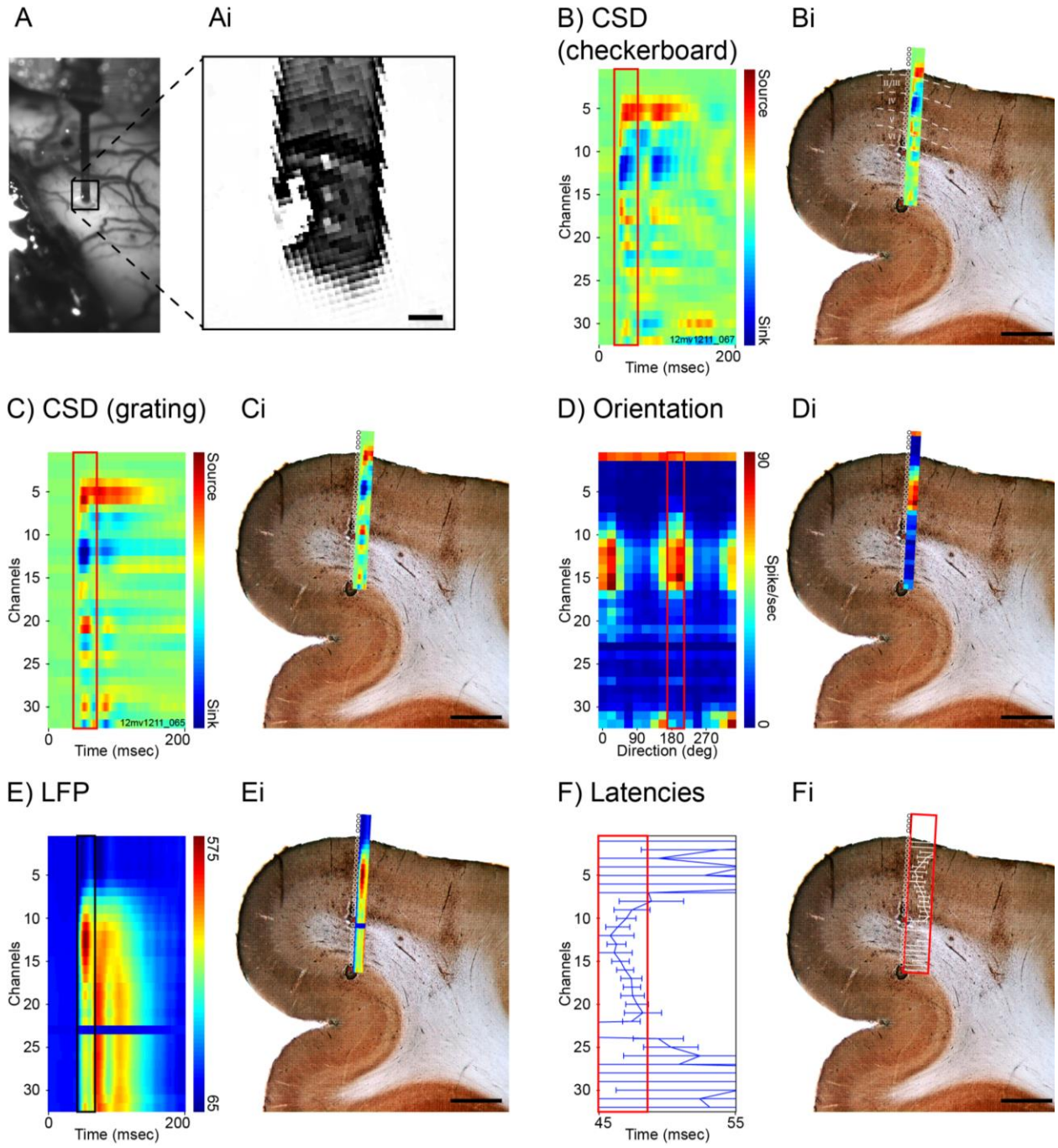

Figure 2-8: Histology and layer correspondence of electrophysiological data. In A, high resolution image of the recording probe near the point where it penetrated cortex. In Ai, a magnification of the part of the probe close to the insertion point, showing contacts just above the surface of cortex (scale bar is 100µm). Panels B, C D, E and F present CSD response to checkerboard, CSD response to grating, orientation selectivity of the action potential response, LFP response, and action potential response latencies, respectively. (continued on next page)

To localize the electrode contacts within cortex, we performed electrolytic lesions by injecting current via two channels (22 and 32; 1mm apart). In addition, we imaged the inserted probe near the site where it penetrated cortex, in order to determine which of the contacts was the closest to but still outside of cortex (Figures 2-8A and 2-8Ai). The exact position of the surface of the brain relative to the contacts could not be precisely determined, because of the presence of the pia, arachnoid and cerebro-spinal fluid. However, given the thickness of these structures (50-150ųm), imaging of the contacts contributed to the localization of the probe relative to the cortical surface. The insertion presented in Figure 2-8 corresponds to the slice presented in Figures 2-2C and 2-2Ci. The tissue was processed for cytochrome oxidase, in order to precisely localize layers IV and VI (Price, 1985). The position of the recording probe relative to histology was established by the electrolytic lesions at channel 22 and 32. Moreover, various electrophysiological properties of the signal (spiking activity, LFP and latencies) were used to validate the probe alignment with histology. In contrast to the data presented so far, the electrophysiological data presented in Figures 2-8 and 2-9 were not interpolated between channels, since the aim was to spatially relate the channel position to cortical laminae.

Figure 2-8, B, C, D, E and F present the spatial correspondence between electrophysiological responses and the cortical layers from which they were recorded. Panel B presents the CSD response to the checkerboard stimulus. The sink with the highest amplitude can be observed in channels 11-13. In Panel Bi (and all other panels depicting histological slices), channels have been positioned relative to the slice by aligning the two contacts (marked in red) used for injecting the lesioning current to the two lesions in the tissue. By aligning and rescaling the probe according to the distance and angle of a line drawn between the two lesions, the neurophysiological responses were superimposed on the histological slice. As a first control, we see that following the alignment of the probe contacts to the electrolytic lesions, the sink with the highest amplitude was localized in layer IV (Figure 2-8Bi). Panel C presents the CSD response averaged across all oriented grating stimuli from one test. It can be seen that the sink response to

Figure 2-8: (continued from previous page) Representative portions of these measures are superimposed on cytochrome oxidase histology in their respective right-side panels. The delineation of cortical layers is shown in Panel Bi. Small bright circles, representing the recording contacts are superimposed on the histological images. In this insertion, channels 22 and 32 were used for electrolytic lesion; here they are illustrated in red. Scale bars in Bi, Ci, Di, E, Fi, 1mm. The data here are from the same insertion presented in Figures 2-1 and 2-2.

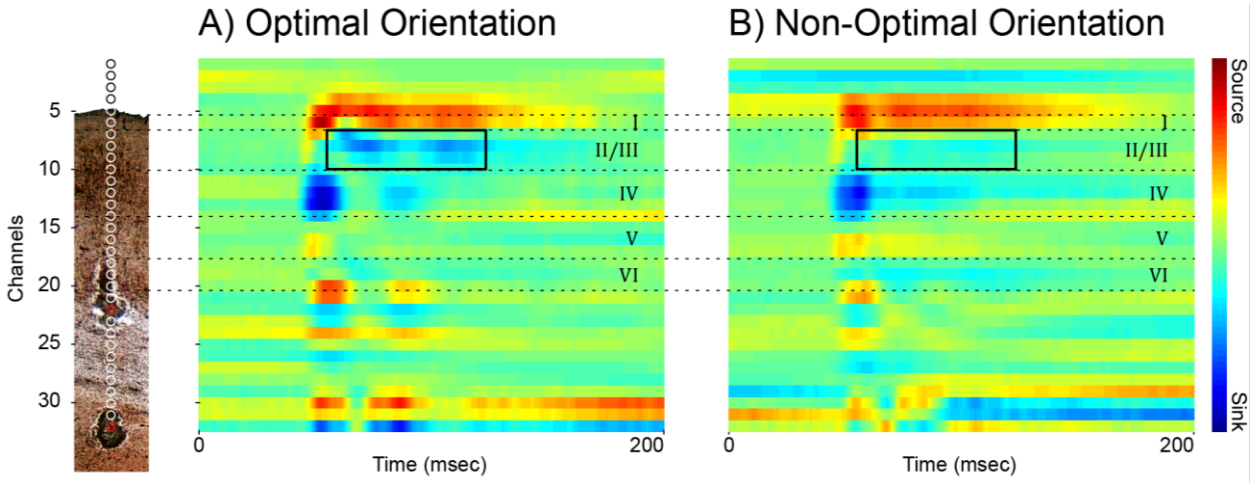

Figure 2-9: Selective modification of CSD correspondence with layer information transfer. Inset, cytochrome oxidase stain histological reconstruction of the same insertion presented in Figure 2-8. The CSD profiles obtained in response to gratings of optimal and non-optimal orientation for this insertion are presented in panel A and B, respectively. Cortical layers are labeled. The change in extent of the main sink can be appreciated when the optimal orientation is presented, extending from layer IV to layer II-III. This transfer of signal does not take place when the non-optimal orientation is presented. This can be seen by inspecting the response in the black rectangles located in layers II-III in both panels.

gratings spatially corresponds to layer IV too. Panel D presents the tuning curves computed according to the spike rate responses in the same penetration. The main action potential response is located at the center of the cortex. Further evidence that the match between the electrophysiological data and histology is accurate is the spiking activity observed in channels 31 and 32. This observation was initially confusing. However, Panel Di shows that these channels are in grey matter on the other side of the gyrus, explaining the action potential responses observed in these channels.

LFP of all the oriented gratings is presented in Panel E with a window of the signal superimposed on the histological slice in Panel Ei. The LFP response of the largest amplitude was centered on layer IV. Panels F and Fi present the mean latency computed across all oriented gratings. As expected, the fastest responses were found in layer IV. Moreover, stable latencies could only be calculated in channels that were inside cortex, thus providing another indication that our alignment was accurate.

These converging findings confirm that the primary sink of the CSD response reflects the activity in layer IV in cat area 18. Therefore, it can be used as a functional marker of the position

of layer IV. The correspondence between electrophysiology data and histology were reproduced in several animals and insertions.

Following the demonstration of the correspondence between the CSD pattern and cortical layers, we can map our findings on orientation selective CSD responses to cortical laminae. The preferred orientation stimulus induced an increase in extent of the primary sink (mean 0.13 mm). This extension was always towards contacts more superficial relative to the primary sink, and took place at a lag relative to the primary sink (Figure 2-3B and 2-5). Given that we associated the primary sink with layer IV, we can specify that following the primary sink in layer IV, supragranular layers II-III show a sink when the preferred orientation is presented, but not when the non-preferred orientation is presented. This result is presented in Figure 2-9, where the CSD responses to the preferred (Figure 2-9A) and non-preferred (Figure 2-9B) orientations are mapped to cortical layers. When the optimal orientation is presented, the sink signal extends towards layer II-III. In response to the non-optimal orientation, such extension either does not take place (Figures 2-3 and 2-5) or shows much smaller amplitude (Figure 2-9B). Hence, we conclude that there is a selective transfer of information from layer IV to layer II-III when the optimal orientation is presented. Our work suggests that this intra-cortical transfer of information is functionally gated and provides an effective means for feature selection.

2.3.2 Validation of the use of standard CSD estimation

A potential objection to the results presented herein is that spatially non-homogeneous activation of cortex with an oriented stimulus violates the infinite plane assumption behind standard CSD estimation. Because, CSD estimation is an inverse problem, its solution always depends on the assumptions built in the chosen forward problem. Indeed, by assuming that the x and y components of the Laplacian are zero, one technically assumes that the sources are infinitely extended in space. This assumption is clearly violated when presenting a stimulus such as orientation which activates cortex in a patchy manner. It is therefore reasonable to ask whether the smaller spatial extent of the sink observed in the non-optimal orientation is simply an artifact from the off-center placement of the recording electrode relative to the center of the non-optimal orientation column. Moreover, inverse CSD methods have been developed exactly to address

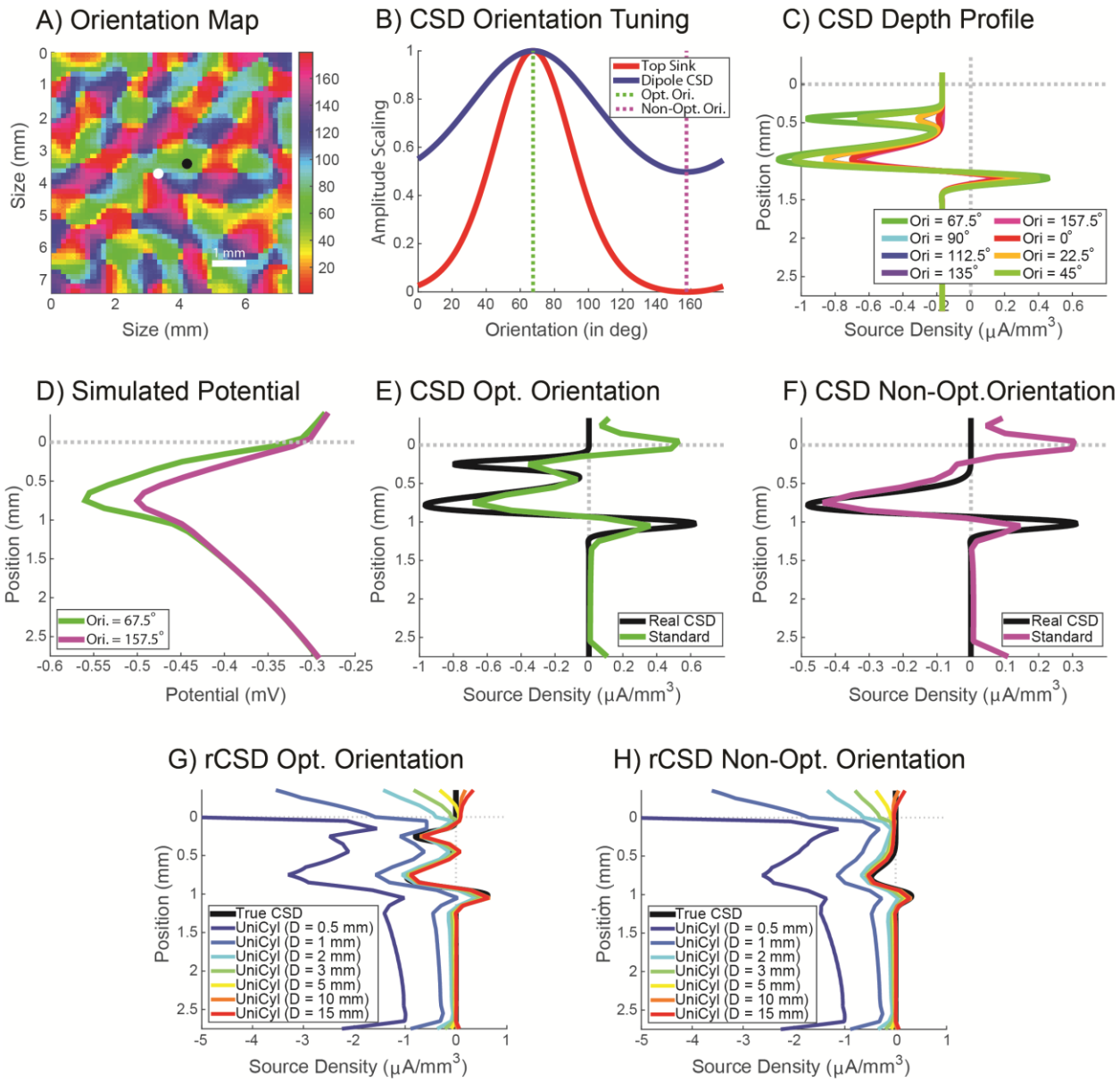

Figure 2-10: Validation of standard CSD method. In A, orientation map simulated using the Topographica toolbox (pixel size 50x50 µm, total width = 7.5 mm). Black and white dots show the simulated electrode positions. The optimal orientation at the location of the black dot electrode is 69.2º. In B, orientation selectivity of the dipolar and monopolar components of the simulated CSD depth profile. The variation of the CSD depth profile is shown in Panel C for 8 orientations. Panel D shows the field potential recorded at the black dot electrode in response to the presentation of gratings closest to the optimal (67.5º) and non-optimal (157.5º) orientation. Panels E and F show the simulated CSD response (black) at the location of the electrode overlaid with the estimation from the standard CSD method in response to the presentation of an optimal and non-optimal grating respectively. Panels G and H show the same as E and F but for CSD estimation based on the inverse CSD method rCSD with various diameters. For the inverse CSD estimation, the sources were assumed to be uniformly distribution in a cylinder of diameter D. We see that in both the optimal (G) and the non-optimal (H) condition, the estimation becomes better with increasing diameter.

issues in CSD estimation with sources of limited lateral extent (Pettersen et al., 2006, Potworowski et al., 2012, Kropf and Shmuel, 2016). It is therefore reasonable to ask why these methods were not employed here.

In order to address this potential shortcoming, we validated our CSD estimation on simulated data. To do so, we have simulated the orientation map shown in Figure 2-10A using the open-source MATLAB toolbox Topographica (Bednar, 2009). The map is discretized at a pixel resolution of 50x50 μm and has a total width and height of 7.5 mm. In other words, the simulation assumes that the orientation preference is constant within a region of 50x50 μm.

To get the LFP response to a presentation of a fullfield grating stimulus, we inserted a simulated electrode in the center of a pixel (shown by the black dot in Figure 2-10A). We then created a hypothetical orientation-selective CSD response aimed at mimicking the result that we observed (see Figure 2-10B). In particular, the CSD is composed of a weakly orientation-selective dipolar component with a sink located around layer 4 and a strong orientation-selective monopolar component in the superficial layer (see Figure 2-10C). The LFP response was simulated by summing the potential generated from each pixel following a stimulation with a fullfield grating of optimal and non-optimal (orthogonal) orientation (discretized in steps of 22.5° like in the experiment). The forward model employed here was made of 2 semi-infinite media with continuous conductivity separated at z = 0 (top conductivity = 1.7 S/m, bottom conductivity = 0.3 S/m). This forward model aims at emulating a recording from a cortical site covered with saline, as used by (Kropf and Shmuel, 2016). The resulting LFP responses are shown in Figure 2-10D.

To validate the use of the standard CSD estimation method, we calculated the CSD for both the optimal and non-optimal condition. Figure 2-10E shows the true CSD depth profile of the pixel containing the simulated electrode (black) in response to the optimal orientation, along with the estimated CSD using the standard 5-point method (the same is shown in Figure 2-10F for the non-optimal orientation). These results clearly demonstrate that the standard CSD estimation method is sensitive to the simulated orientation-selective component. Due to the smoothening effect of the 5-point estimator, the overall amplitude of the components is under-estimated but no artefactual sink is introduced by the off-centered position of the electrode with respect to the non-optimal orientation columns. The only artefactual component observed is the

source estimated at the surface of the brain which likely stems from the discontinuity of conductivity that is not accounted for in the standard CSD estimation.

Although these results indicate that the use of the standard CSD estimation seems warranted, it is still possible that the use of the inverse methods with a cylindrical or Gaussian forward model might provide more accurate estimation. To test this, we also estimated the CSD using the rCSD method with a range of lateral source diameters (Kropf and Shmuel, 2016). Figures 2-10G and H show the resulting CSD estimates for the optimal and non-optimal orientation. It was quite a surprise to us that the larger the diameter, the closer the estimated sources are to the true underlying profile. We therefore conclude that the violation of the infinite lateral source extension in the standard CSD does not appear to be a limitation to the results presented in this work. Finally, in order to make sure that these results are not due to placing the simulated electrode in the center of an orientation domain, we have repeated the same analysis for an electrode placed closer to a pinwheel (see white dot in Figure 2-10A). The orientation selectivity of the standard CSD estimation was preserved (data not shown). The only, rather unsurprising, difference was that the overall amplitude of the estimation was reduced.

2.4 Discussion

CSD analysis is a potent tool for investigating cortical processing since it allows the investigation of the flow of activity within and between cortical layers (Mitzdorf and Singer, 1978). However, its sensitivity and specificity to spatio-temporal properties of visual stimuli has been debated (Mitzdorf, 1985). Here we show for the first time that CSD responses reflect the neuronal processing involved in signalling the orientation of grating stimuli. We specifically compared CSD responses to optimal and non-optimal orientations, and showed that the primary sink is shorter in the non-preferred orientation than in the optimal orientation. Moreover, it is also shorter in the non-optimal orientation compared to a non-oriented (checkerboard) stimulus. Hence, this change in the CSD pattern appears to be linked with a layer specific processing of orientation signals. This is at odds with Mitzdorf (Mitzdorf, 1985) who stipulated that the specific stimulus properties are reflected only in minor modulations of the basic properties of the CSD.

2.4.1 Information transfer between cortical laminae

Features of the visual scene are represented at various levels of cortical processing. This is true for different cortical areas, but also for different cortical layers (Hubel and Wiesel, 1962, Gilbert, 1977, Movshon et al., 1978a, DeAngelis et al., 1995, Ringach et al., 1997, Martinez et al., 2005). Layer IV is often seen as the first step of cortical processing in the visual system since it is the main recipient of thalamic input (Hubel and Wiesel, 1962, LeVay and Gilbert, 1976). Hubel and Wiesel (1962) proposed that the orientation selectivity of simple cells in layer IV is based on the specific configuration of the receptive fields of the presynaptic thalamic relay cells.

Our electrophysiological analysis reflects this synchronized and fast synaptic activity in layer IV, as presented in Figure 2-8. The primary sink of the CSD profile is confined to layer IV, which also coincides with the shortest latencies. The locus of this early synaptic activity does not appear to be modified by the properties of the stimuli, whether it is a checker board, an optimal, or non-optimal oriented grating. As was already shown by Mitzdorf (Mitzdorf, 1985), this CSD response reflects the fast input from the thalamus and does not reflect any spatio-temporal properties of the stimulus. Hence, it reflects the input from the LGN and not the intracortical processes that appear at later stages of the response (Ben-Yishai et al., 1995, Somers et al., 1995).

The second stage of cortical processing is expected to occur in layer II-III, which receives a direct projection from layer IV (Gilbert and Kelly, 1975, Gilbert and Wiesel, 1979, Fitzpatrick, 1996, Callaway, 1998, Hirsch et al., 1998). As revealed by our CSD analysis, this stage of cortical processing appears to reflect the properties of the stimulus, since the primary sink extends from layer IV to layer II-III only when the optimal orientation or a non-oriented stimulus is presented. This functional pathway between cells in layer IV and II-III has been investigated before from many different angles. First, numerous studies that showed dense anatomical connections between layer IV cells and supragranular layers of the cortex (Callaway 1998; Fitzpatrick 1996; Gilbert and Kelly 1975; Gilbert and Wiesel 1979; Hirsch, et al. 1998). Moreover, the functional strength of the connection from layer IV to layer II-III was estimated by fitting a model that captures the effective connectivity between cortical laminae to the CSD response in rat area S1FL (Sotero et al., 2010). In the visual cortex, the functional strength of this connection was demonstrated in vivo using a cross-correlogram analysis (Hubel and Wiesel,

1968, Alonso and Martinez, 1998, Martinez and Alonso, 2001). Strong functional coupling was found between a high proportion of extracellularly recorded simple cells in layer IV and complex cells in layer II-III (50%) of similar orientation preferences. Moreover, it has been shown that this functional coupling between layer IV cells and layer II-III cells arises only when coordinated or nearly synchronous activity among a large population of layer IV cells is produced (Bonhoeffer and Grinvald, 1991, Yu and Ferster, 2013).

In line with these previous findings, our study suggests that the coupling between layer IV cells and layer II-III cells is functionally gated, since information is selectively relayed between the first and second stages of cortical processing. Moreover, this is consistent with a prior study which proposed that the intrinsic circuitry in layer II/III acts in concert with inputs from layer IV to generate the orientation selectivity of layer II/III neurons (Bosking et al., 1997).

But what are the mechanisms that control the opening/closing of this functional gate? Or in other words, why is information relayed to superficial layers in the optimal but not in the non-optimal orientation? It was proposed that facilitatory interactions between converging excitatory inputs establish the foundation of more complex stimulus selectivity. Hirsch et al (2002) suggest that any post-synaptic events generated by a single stream are negligible if not accompanied by the arrival of succeeding inputs. Thus, it is possible that the feed forward stream from layer IV has to be co-activated with a secondary stream to selectively open the gate so that signal is transferred to layers II-III, as proposed by Bosking et al. (1997).

Complementing the feed forward stream is the presence of horizontal connections which are most prominent in the layers II-III (Rockland and Lund, 1982, Gilbert and Wiesel, 1983, Rockland and Lund, 1983, Fitzpatrick, 1996, Bosking et al., 1997, Chisum et al., 2003)**.** These connections are largely viewed as contributing to receptive field surround effects (Rockland and Lund, 1982, Gilbert and Wiesel, 1983)**.** However, these contributions are limited since they link non-overlapping receptive fields under 10% of the time, while the vast majority of the connections link neurons with receptive fields that overlap by 50% (Chisum et al., 2003). Moreover, horizontal connections preferentially link sites with similar orientation preference, suggesting that they might play a role in shaping orientation tuning in layer II-III neurons (Fitzpatrick, 1996, Bosking et al., 1997, Chisum et al., 2003). Therefore, it has been suggested

that the functional contribution of horizontal connections is not restricted to surround effects, but that it also plays a significant role in shaping the responses within the classical receptive field (Crook and Eysel, 1992, Girardin and Martin, 2009). We propose that since horizontal connections provide a powerful amplification of an orientation bias provided by feedforward inputs (Douglas et al., 1995, Chisum et al., 2003, Hang and Dan, 2011), they are possibly responsible for the opening of the proposed functional gate between layer IV and II-III due to a convergence of orientation preference signals. This would explain why some signals but not others propagate from layer IV to II-III in a selective manner. The optimal orientation would activate horizontal intracortical pathways, which in conjunction with feed forward signals from layer IV, would open the gate and allow effective information transfer to layers lI-III. The same might be true for an unoriented stimulus which activates the cortical sheet in an approximately even manner. In contrast, the gate would remain closed in the non-optimal orientation because insufficient reinforcement would be provided by horizontal connections in layer II-III. This interpretation is also supported by computational models that suggest that interactions between orientation columns play a pivotal role in signal processing and influence the speed of information transport, which may have a greater influence than inner-columnar interactions (Stoop et al., 2013).

It is worth noting that when intracortical pathways are silenced by topical application of muscimol (Happel et al., 2010), the resulting CSD profile obtained is very similar to the one obtain with non-optimal orientation stimuli (present data). This supports the concept that intracortical signalling is reduced for non-optimal orientations.

Although we have so far emphasized the contribution of intra-area horizontal connections, one cannot rule out the contribution of axons with origin in area 17 that project to layers II-III of area 18 (Gilbert and Wiesel, 1989). The contribution of this pathway has been shown to be significant in building the functional properties of cells in area 18 (Casanova et al., 1992), and might therefore play a role in their orientation selectivity.

2.4.2 Validation of the standard CSD estimation method

The CSD methods used here is based on the "standard" forward model of electric potentials in biological tissue (Mitzdorf, 1985, Nunez and Srinivasan, 2006, Plonsey and Barr, 2007) which rests on the hypothesis that the extracellular medium is purely resistive, isotropic

and homogeneous. These assumptions have recently come under challenge, and it has been suggested that non-ohmic effects such as ionic diffusion could play a significant role in the spread of current within cortical tissue (Bedard and Destexhe, 2009, 2011, Brette and Destexhe, 2012, Riera et al., 2012, Wagner et al., 2014, Gomes et al., 2016). Although of great relevance and interest, the importance of these possible confounders is still under heavy debate and it is too early to speculate on their effect on the results presented here.

In addition to the assumptions on the medium properties discussed above, the 1-D standard CSD estimation method also assumes that sources are infinite planes perpendicular to the electrode. This assumption has been shown to significantly falsify estimation when sources are spatially localized or off-centered with respect to the recording electrode (Pettersen et al., 2006). To deal with this issue, inverse CSD methods have been developed (Pettersen et al., 2006, Potworowski et al., 2012, Kropf and Shmuel, 2016). However, it has never been tested how responses to a fullfield stimulus (such as orientated gratings), which result in a patchy cortical activation, should be treated. To address this potential shortcoming, we have validated the use of the standard CSD estimation on simulated data. We found that for our fullfield grating stimulus, the standard CSD was appropriate and sensitive to the orientation selective component that we introduced in the simulated signal.

Preface to Chapter 3

In this chapter, we have shown the usefulness of CSD estimation for the study of neural processing. However, our analysis has also highlighted known difficulties with the use of the standard method. Therefore, in the next chapter we conduct an in-depth study of the methods and assumptions behind 1-D CSD estimation. In particular, we propose a novel estimation framework inspired by linear inverse theory with a special emphasis on higher-order spectral regularization methods. With this framework, we provide a unification of previously published inverse estimation methods under a common format and introduce novel estimation tools. In addition, we will see that the presented framework greatly facilitates dealing with recording noise through the incorporation of zeroth- and higher-order regularization methods with multiple regularization parameter selection methods.

3 Chapter 3: 1-D current-source density (CSD) estimation in inverse theory: a unified framework for higher-order spectral regularization of quadrature and expansion type CSD methods

Estimation of current-source density (CSD) from the low-frequency part of extracellular electric potential recordings is an unstable linear inverse problem. To make the estimation possible in an experimental setting where recordings are contaminated with noise, it is necessary to stabilize the inversion. Here we present a unified framework for zero- and higher-order singular-value decomposition (SVD) based spectral regularization of 1-D (linear) CSD estimation from local-field potentials. The framework is based on two general approaches commonly employed for solving inverse problems, quadrature and basis function expansion. We first show that both inverse CSD (iCSD) and kernel CSD (kCSD) fall into the category of basis function expansion methods. We then use these general categories to introduce two new estimation methods, quadrature CSD (qCSD), based on discretizing the CSD integral equation with a chosen quadrature rule, and representer CSD (rCSD), an even-determined basis function expansion method which uses the problem's data kernels (representers) as basis functions. To determine the best candidate methods to use in the analysis of experimental data, we compared the different methods on simulations under three regularization schemes (Tikhonov, tSVD and dSVD), three regularization parameter selection methods (NCP, L-curve and GCV) and seven different a priori spatial smoothness constraints on the CSD distribution. This resulted in a comparison of 531 estimation schemes. We evaluated the estimation schemes according to their source reconstruction accuracy by testing them using different simulated noise levels, lateral source diameters and CSD depth profiles. We found that ranking schemes according to the average error over all tested conditions results in a reproducible ranking, where the top schemes are found to perform well in the majority of tested conditions. However, there is no single best estimation scheme that outperforms all others under all tested conditions. The unified framework we propose expands the set of available estimation methods, provides increased flexibility for 1-D CSD estimation in noisy experimental conditions, and allows for a meaningful comparison between estimation schemes.

3.1 Introduction

Understanding brain function entails disentangling the activity of different neuronal populations. Owing to advances in microelectrode fabrication technology, neuroscientists are now in position to simultaneously record extracellular potentials from multiple spatial locations in close proximity. In particular, a growing number of studies are using linear electrode arrays to disentangle laminar processing in the cortex. In order to extract useful information about laminar processing from such recordings, it is imperative to comprehend how extracellular potentials are related to the underlying neural activity.

With extracellular recordings, it is not possible to measure the neuronal activity directly. Instead we measure electrical potential differences that result from the flow of currents. In particular, the low-frequency part of extracellular potentials, called the local field potential (LFP), is thought to be related to currents generated during synaptic activity (Nunez and Srinivasan, 2006, Einevoll et al., 2013). Direct interpretation of LFP is problematic given that electric potentials are a non-local measure of the neural activity due to volume conduction. Thus, estimating current-source density (CSD), which represents the volume density of net transmembrane currents generating the LFP, has become common practice in extracellular neurophysiology (Freeman and Nicholson, 1975, Mitzdorf, 1985, Schroeder et al., 1998, Bedard and Destexhe, 2011, Tenke and Kayser, 2012).

In a homogeneous, isotropic and purely resistive medium with electrical conductivity σ, current-source density C is related to the extracellular potential φ via the Poisson equation: $C = -\sigma\Delta\varphi$ where Δ is the Laplace operator (Nicholson and Freeman, 1975, Tenke et al., 1993, Brette and Destexhe, 2012). Originally, CSD was estimated by approximating the Laplacian by a discrete second derivative (Freeman and Nicholson, 1975). Moreover, in the analysis of linear (laminar) electrodes inserted perpendicularly to cortical layers, it has been commonplace to simply ignore the lateral, i.e. x and y, derivatives in the Laplacian (e.g. (Mitzdorf, 1985, Schroeder et al., 1998)). This is equivalent to assuming that potentials have minimal curvature in the lateral direction, a situation which can only be achieved with laterally extended current-source density profiles. In recent years, a novel CSD method called iCSD has been developed to circumvent the difficulties associated with spatially localized sources (Pettersen et al., 2006, Leski et al., 2007, Leski et al., 2011). Instead of simply computing the derivative, iCSD assumes

a particular parametric form for the CSD and uses a forward-inverse scheme for estimating CSD amplitude. This method has the advantage of allowing the estimation of CSD at the outermost electrode contacts, something which in the standard method is only doable under strong assumptions (Vaknin et al., 1988, Pettersen et al., 2006). Additionally, it also facilitates the incorporation of slightly more complex medium assumptions like the parametrization of the medium as two semi-infinite media with different electrical conductivity. Recently, a generalized version of iCSD called kCSD (short for kernel CSD) has been introduced (Potworowski et al., 2012). Using results from machine learning, kCSD transforms CSD estimation into an underdetermined problem where uniqueness of the solution is enforced using a minimum-norm requirement. This approach allows greater flexibility in changing the electrode arrangement. This is of particular interest when working on non-regular grids which occurs when the recording electrode has non-functional contacts.

In the present work, we propose a unification of all 1-D current-source density estimation methods in a single framework, by adding elements of linear inverse theory. In particular, we focus on comparing regularization methods based on spectral filtering to deal with recording noise. It is a well-known fact from linear inverse theory that inverse problems with smoothening kernels are ill-posed resulting in a high sensitivity to noise. We start by presenting two general approaches that are common in solving inverse problems, quadrature and basis function expansion. This will allow us to show that both iCSD and kCSD fall into the category of basis function expansion methods. We then use these general categories to introduce two new methods for CSD estimation which we call quadrature CSD (qCSD) and representer CSD (rCSD). As its name implies, quadrature CSD is based upon discretizing the CSD integral equation using a chosen quadrature rule while rCSD is an even-determined basis function expansion method which uses the problem's data kernels (also called representers) as basis functions. We will see that our approach is in line with the approach presented in Riera et al., 2014 (Riera et al., 2014) which focuses on 3-D CSD estimation using a realistic volume conductor model for rat barrel cortex. In a second step, we compare the different methods under various regularization schemes (Tikhonov, tSVD and dSVD), different regularization parameter

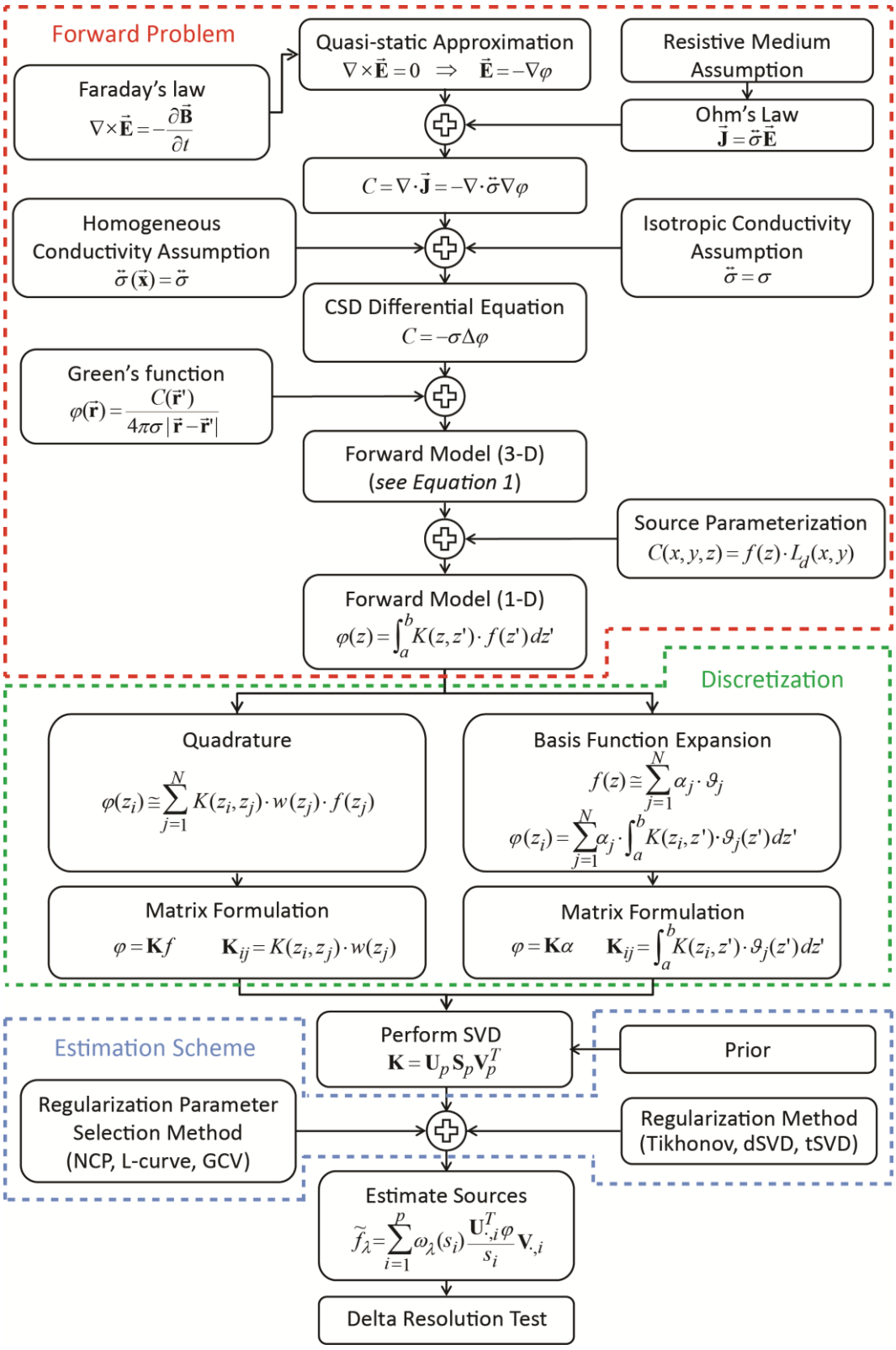

Figure 3-1: Flowchart describing the spectral regularization framework.

selection methods (NCP, L-curve and GCV) along with different a priori spatial smoothness constraints on the CSD distribution. We compare these different estimation schemes using simulations and rank them according to their source reconstruction accuracy under a range of source diameters and noise conditions. We found that no single estimation scheme outperforms all others under all tested conditions. Nevertheless, we found that ranking schemes according to the average error over all tested conditions results in a reproducible and meaningful ranking where the top schemes are found to perform well in the majority of tested conditions.

3.2 Theory

In this section, we describe the different aspects of 1-D CSD estimation from the standpoint of linear inverse theory. We discuss the theoretical aspects in more detail than what is common, in the hope of making it more accessible to newcomers to the field. Figure 3-1 shows a schematic of the steps leading to the final formulation and proposed solution of the problem.

3.2.1 Forward and inverse problem

Before attempting to solve the inverse problem of CSD estimation, it is necessary to have a clear understanding of the associated forward problem. This is crucial because current CSD estimation methods do not allow for the verification of the forward model assumptions from within the model (Bedard and Destexhe, 2011). In the case of current CSD methods, the forward problem is described by the Poisson equation for volume conductors, which is obtained by combining the quasi-static approximation with Ohm's law (Plonsey and Barr, 2007). Moreover, in line with iCSD and kCSD, we describe the recording medium as being composed of 2 semi-infinite media with different but constant electrical conductivity (σ) as graphically depicted in Figure 3-2A. The plane $z = 0$ represents the brain surface and separates the extracellular space (assumed unbounded from below) from the medium above the brain. The upper medium is added to model the effect of substances such as artificial cerebro-spinal fluid (aCSF) or mineral oil on potential generation. It is common in invasive in vivo electrophysiological experiments to add liquids onto the skull opening to avoid drying of the brain surface.

Taken together, the electrostatic assumption as well as the semi-infinite medium parameterization with constant conductivity form the basic set of assumptions defining the forward problem for CSD estimation. Under these circumstances, the potential $\varphi(x, y, z)$

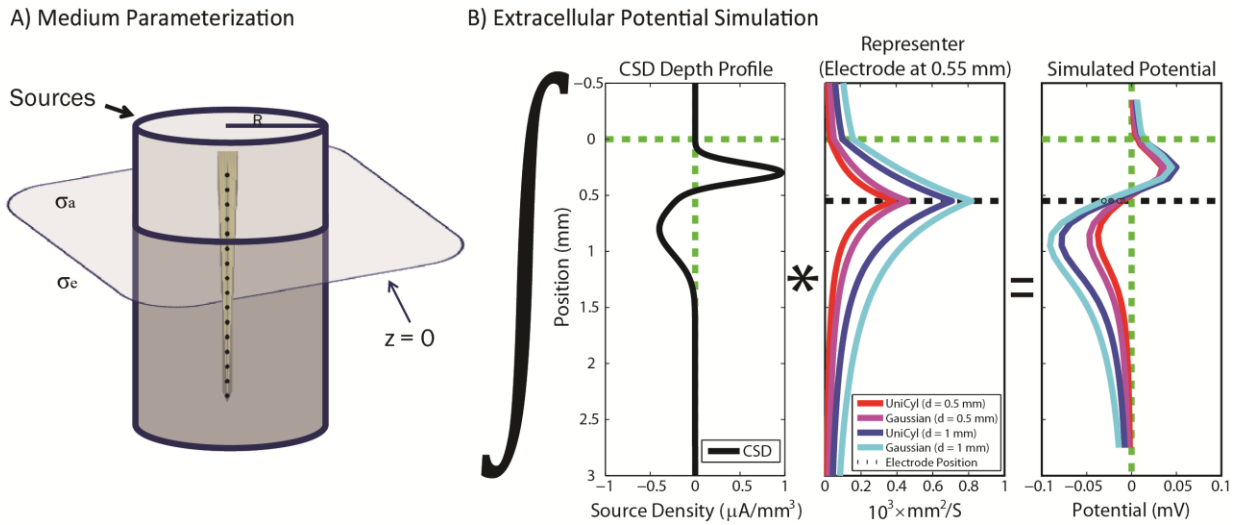

Figure 3-2: Simulation of extracellular potential. (A) Schematic of the medium parameterization and cylindrical lateral source assumption. A 32 channel electrode is located in the center of the cylinder. (B) Effect of lateral source parameterization choice on simulated potential. The value of extracellular potential (dots in last panel) is obtained by integrating the CSD depth profile (first panel) with the corresponding representer (middle panel). The depth profile of the representer is shown for an electrode located at a depth of 0.55 mm. It shows how the potential at this electrode is affected by sources at different depths. Representers for a uniform cylinder (UniCyl) and Gaussian lateral source parameterization with a diameter of 0.5 or 1 mm are shown. To obtain the simulated potential (last panel) as a function of depth, the integration is performed using representer proper to each electrode.

generated at any point in space by a current-source density distribution $C(x,y,z)$ can be found using the method of images (Jackson, 1999)

$$\begin{aligned}\varphi(x,y,z\geq 0) &= \frac{1}{4\pi\sigma_e}\left(\iiint_{z'=0}^{\infty}\frac{C(x',y',z')}{|\vec{x}-\vec{x'}|}+\frac{\sigma_e-\sigma_t}{\sigma_e+\sigma_t}\frac{C(x',y',z')}{|\vec{x}+\vec{x'}|}dV'\right)\\ &+\frac{1}{\sigma_e+\sigma_t}\left(\iiint_{z'=-\infty}^{0}\frac{C(x',y',z')}{|\vec{x}-\vec{x'}|}dV'\right)\\ \varphi(x,y,z<0) &= \frac{1}{4\pi\sigma_t}\left(\iiint_{z'=-\infty}^{0}\frac{C(x',y',z')}{|\vec{x}-\vec{x}'|}+\frac{\sigma_t-\sigma_e}{\sigma_e+\sigma_t}\frac{C(x',y',z')}{|\vec{x}+\vec{x}'|}dV'\right)\\ &+\frac{1}{\sigma_e+\sigma_t}\left(\iiint_{z'=0}^{\infty}\frac{C(x',y',z')}{|\vec{x}-\vec{x}'|}dV'\right)\end{aligned} \quad (1)$$

with $V'=dx'dy'dz'$ and $\vec{x}=(x,y,z)$, while σ_e and σ_t are the electrical conductivity for $z\geq 0$ and $z<0$ respectively. The limits of integration are chosen to cover all current sources. Here,

we slightly expand on the approach used in (Pettersen et al., 2006) because we make it possible to use electrodes located in both media.

The above formulation defines a classic linear inverse problem for $C(x, y, z)$. Here, the particular problem we are interested in solving is: what information about $C(x, y, z)$ can be recovered from 1-D linear (laminar) depth electrode recordings $\varphi(z)$?

3.2.2 Source parameterization

To solve this inverse problem for laminar recordings, the 3-dimensional problem needs to be reduced to 1-dimension. This is required because in laminar recordings all electrode contacts are located on a line approximately orthogonal to the local cortical surface. Therefore, the contacts provide no information about the in-plane potential variation. The dimensionality reduction is achieved by assuming a fixed lateral profile for the function $C(x, y, z)$, i.e.

$$C(x, y, z) = f(z) \cdot L(x, y) \tag{2}$$

It is customary to choose $L(x, y)$ to be either a Cylinder (see (Pettersen et al., 2006)) or a Gaussian function

$$L(x, y) = \begin{cases} 1, & x^2 + y^2 < R^2 \\ 0, & otherwise \end{cases} \quad or \quad L(x, y) = e^{-\frac{(x^2+y^2)}{2R^2}} \tag{3}$$

It is important to note that these parameterizations impose very strong assumptions on the underlying function $C(x, y, z)$.). Their choice stems from a plausibility argument relying on the columnar organization of sensory cortices as main justification. However, to our knowledge, it has never been experimentally verified whether and under which circumstances these assumptions are warranted.

Inserting Equation 2 into Equation 1 and integrating over x and y provides us with the desired 1-dimensional linear inverse problem

$$\varphi(z) = \int_a^b K(z, z') \cdot f(z')\, dz' \tag{4}$$

where $K(z, z')$ is called the kernel of the linear operator. Certain authors also refer to $K(z_i, z)$ as representers or data kernels (Parker, 1994, Aster et al., 2005). If we assume that the electrode is

located at the center of a Cylinder or a Gaussian, i.e. $x' = y' = 0$, then the kernel can be found in closed form using

$$\iint \frac{L(x,y)}{|\vec{x} - \vec{x}'|} dxdy = \begin{cases} 2\pi(\sqrt{(z-z')^2 + R^2} - |z-z'|), & Cylinder \\ 2\pi \frac{\sqrt{2\pi}R}{2} \cdot erfcx\left(\frac{|z-z'|}{\sqrt{2}R}\right), & Gaussian \end{cases} \quad (5)$$

where $erfcx(x) = e^{x^2} \cdot erfc(x)$ and erfc is the complementary error function.

More generally, Equation 4 is an example of a general form of well-studied problems known as Fredholm integral equations of the first kind. Problems of this form are ubiquitous in science and therefore a large body of literature is devoted to their potential solution. Unfortunately, inversion of linear Fredholm integral equations of the first kind is often plagued with issues of uniqueness and/or stability (Wing and Zahrt, 1991, Aster et al., 2005, Hansen and Society for Industrial and Applied Mathematics., 2010).

3.2.3 Discretizing the inverse problem

To make the solving of equation 4 amenable to numerical treatment, the problem needs to be discretized. Since in an experimental setting $\varphi(z)$ is only known at a discrete number of recording positions, we will first discretize the left side of equation 4 by projecting $\varphi(z)$ onto a set of delta functions centered at the recording points, i.e. $\varphi(z_i) =< \varphi(z), \delta(z - z_i) >$ (Wing and Zahrt, 1991). This results in

$$\varphi(z_i) = \int_a^b K(z_i, z') \cdot f(z')\, dz' \quad (6)$$

where z_i is the position of recording electrode i.

Next, we will add the assumption that the function $f(z)$ is a L^2 function on an interval $[a, b]$. It is necessary to make the interval finite because neither the cylindrical nor the Gaussian 1-D CSD kernels are L^2 operators on an infinite interval.

Here, we investigate 2 approaches to discretize the right-hand side of Equation 6. In line with tradition, we will call these approaches qCSD for quadrature CSD and rCSD for representer CSD. Moreover, we also add a variant of kCSD which omits the kernel trick (by 'kernel trick' we refer to a method in machine learning, which allows to operate in a high-dimensional,

implicit feature space by computing the inner products between the images of all pairs of data in the feature space).

3.2.3.1 qCSD

First, we investigate how the problem is solved with simple quadrature methods, i.e. replacing the integral in Eq. 6 by a sum. Using an N-point quadrature formula with associated weights $w(z_j)$, Equation 5 is transformed into (Wing and Zahrt, 1991)

$$\varphi(z_i) \cong \sum_{j=1}^{N} K(z_i, z_j) \cdot w(z_j) \cdot f(z_j) \tag{7}$$

where the sampling points z_j are given by

$$\begin{aligned} z_j &= a + \frac{\Delta z}{2} + (j-1)\Delta z \\ \Delta z &= \frac{b-a}{N} \end{aligned} \tag{8}$$

For M recording electrodes, we see that the original integral equation has been transformed into a matrix equation

$$\begin{aligned} &\varphi = \boldsymbol{K} f \quad \boldsymbol{K} \in M * N \\ &with \\ &\boldsymbol{K}_{ij} = K(z_i, z_j) \cdot w(z_j) \\ &f = [f(z_1), f(z_2), \dots, f(z_N)]^T \\ &\varphi = [\varphi(z_1), \varphi(z_2), \dots, \varphi(z_M)]^T \end{aligned} \tag{9}$$

If the number of quadrature points N is larger than the number of electrodes M, the problem is underdetermined, meaning that there is an infinite number of source functions f that can satisfy the measurements φ.

3.2.3.2 rCSD

A second option for discretizing the right hand side of Equation 6 is to expand $f(z)$ in a finite set of linearly independent functions $\vartheta_j(z)$

$$f(z) \cong \sum_{j=1}^{N} \alpha_j \cdot \vartheta_j(z) \tag{10}$$

Inserting this expansion into Equation 6 results in

$$\varphi(z_i) = \sum_{j=1}^{N} \alpha_j \cdot \int_a^b K(z_i, z') \cdot \vartheta_j(z') dz' \tag{11}$$

Again, we see that for M recording electrodes, the original integral equation has been transformed into a matrix equation

$$\begin{aligned} &\varphi = \boldsymbol{K}\alpha \quad \boldsymbol{K} \in M * N \\ &with \\ &\boldsymbol{K}_{ij} = \int_a^b K(z_i, z') \cdot \vartheta_j(z') dz' \end{aligned} \tag{12}$$

This approach is also referred to as projection or Galerkin method (Wing and Zahrt, 1991).

One can see that this is similar to the approach taken by both iCSD and kCSD. In iCSD, the basis functions $\vartheta_j(z)$ are either delta, step or spline functions and N is chosen to be the same as M, i.e. same number of basis functions as recording electrodes (Pettersen et al., 2006). In contrast, kCSD chooses either Gaussian or Step functions as basis functions. Moreover, kCSD allows the system to be underdetermined by making $N > M$ (Potworowski et al., 2012).

An obvious question arises at this point: What is the best choice of basis function and why? The general answer is that the best choice is given by basis functions that resemble the underlying sources most closely as already hinted at in (Potworowski et al., 2012). This is apparent from Equation 10 since such basis functions will best approximate the underlying function $f(z)$. The difficulty is that under experimental conditions it is not known *a priori* what form $f(z)$ takes. Therefore, we propose to expand $f(z)$ in a set of basis functions naturally given by the problem.

Remember that we assumed that $f(z)$ is an element of the Hilbert space $L^2[a, b]$. As long as $K(z_i, z')$ is bounded (as it is in our case), the forward problem described by Equation 6 can be

expressed as a bounded linear functional. Then, using the Riesz representation theorem, we can rewrite Equation 6 as an inner product (Parker, 1994)

$$\varphi(z_i) = < K(z_i, z'), f(z') > \tag{13}$$

Moreover, because we are working in a Hilbert space, we can employ the decomposition theorem to write

$$f(z) = p(z) + q(z), \quad p \in \mathrm{K}, q \in \mathrm{K}^{\perp} \tag{14}$$

where K is the subspace spanned by the M representers $K(z_i, z')$ and $\mathrm{K}^{\perp}$ is its orthogonal complement.

Since any function in $\mathrm{K}^{\perp}$ is orthogonal to a function in K it follows that

$$\begin{aligned} \varphi(z_i) &= < K(z_i, z'), f(z') > \\ &= < K(z_i, z'), p(z') + q(z') > \\ &= < K(z_i, z'), p(z') > + < K(z_i, z'), q(z') > \\ &= < K(z_i, z'), p(z') > \end{aligned} \tag{15}$$

This means that any part of $f(z)$ which lies in $\mathrm{K}^{\perp}$ has no effect on the recorded potential. Moreover, the existence of $\mathrm{K}^{\perp}$ also implies that there is an infinite number of functions $f(z)$ that satisfy the recorded data perfectly.

In order to single out a particular solution, it is customary to pick the one with the minimum norm (Parker, 1994). In this case, this amounts to set $q(z) = 0$ because

$$\begin{aligned} \|f(z)\|^2 &= < f(z), f(z) > \\ &= < p(z) + q(z), p(z) + q(z) > \\ &= \|p(z)\|^2 + 2 < p(z), q(z) > + \|q(z)\|^2 \\ &= \|p(z)\|^2 + \|q(z)\|^2 \end{aligned} \tag{16}$$

Hence, by setting $q(z) = 0$, the minimum norm solution is given by

$$\begin{aligned} f(z) &= p(z) \\ &= \sum_{j=1}^{N} \alpha_j \cdot K(z_j, z) \end{aligned} \tag{17}$$

with $N = M$. Comparing Equation 17 with Equation 10 shows that this approach leads to choosing the basis functions $\vartheta_j(z) = K(z_j, z)$.

Thereby, inserting Equation 17 into Equation 12 produces

$$\boldsymbol{K}_{ij} = \int_a^b K(z_i, z') \cdot K(z_j, z') dz' =< K(z_i, z'), K(z_j, z') >= \boldsymbol{K}_{ji} \tag{18}$$

A matrix created from the inner products of a set of functions is called a Gram matrix and has a number of interesting properties (Olver and Shakiban, 2006). Most importantly, the linear independence of the representers guarantees that the matrix is full rank which in turn means that it is invertible. Hence, the Gram matrix approach produces a full rank linear system $\varphi = \boldsymbol{K}\alpha$.

3.2.3.3 eCSD

To facilitate the comparison of the expansion methods with the kernel method, we also add a small variation of kCSD. Instead of employing the kernel trick from machine learning to reduce the model estimation dimension, we directly use the expansion into Gaussian basis functions to create an underdetermined inverse problem. This means we simply choose $\vartheta_j(z)$ in Equation 10 to be

$$\vartheta_j(z) = \frac{3}{\sqrt{2\pi}w} e^{-\frac{(z-z_j)^2}{2\cdot(\frac{w}{3})^2}} \tag{19}$$

The number of sources j can be chosen freely in the same way as in kCSD. Note that in kCSD the width of the sources w is called R. We have avoided this choice here to prevent confusion with the lateral source radius in iCSD. Finally, although we describe eCSD as a variation of kCSD, it could equally well be viewed as an underdetermined version of iCSD with a different choice of basis function. However, since we added eCSD specifically to explore the effect of the kernel trick in a 1-D setting, we will refer to it as a variant of kCSD.

3.2.4 Solving the discretized inverse problem using SVD

Both the quadrature (qCSD) and the expansion-type (rCSD, kCSD, eCSD and iCSD) discrete inverse problem can be solved numerically using the singular value decomposition (SVD). As we will see in the next section on Regularization, the SVD approach is particularly

attractive for solving Fredholm integral equations because these problems are often ill-posed and highly sensitive to noise. In the SVD, an M by N matrix $\boldsymbol{K}$ is decomposed into (Aster et al., 2005, Olver and Shakiban, 2006)

$$\boldsymbol{K} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\boldsymbol{T}} \tag{20}$$

where

- $\boldsymbol{U}$ is an M by M orthogonal matrix (i.e. $\boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{U} = \boldsymbol{I}$ where $\boldsymbol{I}$ is the identity matrix) whose columns span the data space R^M
- $\boldsymbol{V}$ is an N by N orthogonal matrix whose columns span the model space R^N
- $\boldsymbol{S}$ is an M by N diagonal matrix with p positive diagonal elements called singular values. The number of non-zero singular values p is equal to the rank of the matrix $\boldsymbol{K}$. Moreover, the singular values are customarily arranged in decreasing order $s_1 \geq s_2 \geq \cdots \geq s_{\min(M,N)} \geq 0$

Because there are only p non-zero singular values, the SVD of $\boldsymbol{K}$ can also be written in the more compact form

$$\boldsymbol{K} = \boldsymbol{U}_{\boldsymbol{p}}\boldsymbol{S}_{\boldsymbol{p}}\boldsymbol{V}_{\boldsymbol{p}}^{\boldsymbol{T}} \tag{21}$$

The compact form simply clarifies that the last $m - p$ columns of $\boldsymbol{U}$ and the last columns $n - p$ of $\boldsymbol{V}$ are multiplied by 0. More importantly, the compact form allows the computation of the Moore-Penrose pseudo-inverse

$$\boldsymbol{K}^{+} = \boldsymbol{V}_{\boldsymbol{p}}\boldsymbol{S}_{\boldsymbol{p}}^{-\boldsymbol{1}}\boldsymbol{U}_{\boldsymbol{p}}^{\boldsymbol{T}} \tag{22}$$

which provides us with the solution to both the qCSD and rCSD-type discrete inverse problem.

For full rank matrices $\boldsymbol{K}$ (like the ones given by iCSD, kCSD and rCSD), there is no difference between the pseudo-inverse and the regular matrix inverse, i.e. $\boldsymbol{K}^{+} = \boldsymbol{K}^{-\boldsymbol{1}}$. This is not so for the underdetermined problems (i.e. $N > M$) like qCSD and eCSD where a regular matrix inverse does not even exist. In the underdetermined CSD methods, the model null space $N(\boldsymbol{K})$ is nontrivial while the data null-space $N(\boldsymbol{K}^{\boldsymbol{T}})$ is trivial as long as the M recording electrodes are at different spatial locations. As a result, there are an infinite number of source distributions $\tilde{f}(z)$ that can fit the recorded data vector $\varphi = [\varphi(z_1), \varphi(z_2), \dots, \varphi(z_M)]^T$ exactly. Hence, an important

question is which solution is found when using the pseudo-inverse? It turns out that the pseudo-inverse solution has the nice property of being unique as well as being the solution with the minimum norm $\left\|\tilde{f}(z)\right\|^2$ (Aster et al., 2005).

As it seems, the problem has now been solved and the estimated current-source density distribution $\tilde{f}(z)$ is uniquely given by

$$\begin{aligned} quadrature\text{:}\ \ \tilde{f} &= \boldsymbol{K}^{+}\varphi \\ basis\ expansion\text{:}\ \ \tilde{\alpha} &= \boldsymbol{K}^{+}\varphi = \boldsymbol{K}^{-1}\varphi \\ \tilde{f} &= \sum_{j=1}^{N} \tilde{\alpha}_j \cdot K\left(z_j, z\right) \end{aligned} \tag{23}$$

Unfortunately, this solution is often useless in practical settings because even under mild noise levels the estimated solution $\tilde{f}(z)$ bears barely any resemblance with the true source distribution $f(z)$. To circumvent this difficulty we turn to so-called regularization techniques.

3.2.5 Regularization

It can be shown that Fredholm integral equations of the first kind with a square integrable kernel are ill-posed (Vogel, 2002). Ill-posed problems are defined in opposition to well-posed problems, which require that the following three properties be given (Parker, 1977):

(i) **Existence of a solution:** For each $\varphi(z)$, there exists a solution $f(z)$ satisfying Equation 6.

(ii) **Uniqueness:** the solution $f(z)$ is unique

(iii) **Stability:** the solution is stable with respect to perturbations in $\varphi(z)$

So how well does our discretized 1-D CSD problem align with these properties? First, as long as the matrix $\boldsymbol{K}$ has full row rank, which is the case if all recording sites are at different locations, any data vector V will be in Range(K) and therefore condition (i) will be satisfied. Second, we already know that underdetermined problems like qCSD or eCSD admit an infinite number of solutions, thereby violating condition (ii). However, we have circumvented this problem by limiting ourselves to the minimum norm solution which we have found to be unique. Third, because discrete inverse problems always have a finite condition number, they cannot

formally be unstable. Instead, they are referred to as numerically ill-posed or ill-conditioned (Aster et al., 2005). This last part is worth investigating more thoroughly because it is a major reason why the CSD inverse problem is difficult to solve.

Let's recall that the solution to the discretized problem is given by

$$\begin{aligned} \tilde{f} &= \boldsymbol{K}^{+}\varphi \\ &= \boldsymbol{V_p S_p^{-1} U_p^T}\varphi \\ &= \sum_{i=1}^{p} \frac{\boldsymbol{U_{\cdot,i}^T}\varphi}{s_i}\boldsymbol{V_{\cdot,i}} \end{aligned} \tag{24}$$

It can be easily seen from the above relationship that small singular values s_i can heavily amplify the contribution of the corresponding model space basis vector $\boldsymbol{V_{\cdot,i}}$ and thereby make it dominate the solution. We have already discussed in the previous section that it is customary to arrange the singular values in a decreasing order. But we have not mentioned yet that the singular values always decay to zero. In fact, the smoother the kernel $K(z_j, z)$ of our Fredholm integral equation, the faster the decay (Hansen and Society for Industrial and Applied Mathematics., 2010). Moreover, the singular vectors can be compared to spectral basis in the sense that they exhibit increasing numbers of zero-crossings (i.e. oscillations) for smaller singular values (Hansen and Society for Industrial and Applied Mathematics., 2010). Thus, the amplification of singular vectors corresponding to small singular values introduces more and more high frequency oscillations into the estimated solution $\tilde{f}$. This explains why even small noise in the data vector φ can have a substantial effect on the solution $\tilde{f}$ as exemplified by

$$\begin{aligned} \varphi &= \boldsymbol{K}f + \varepsilon = \varphi' + \varepsilon \\ \tilde{f} &= \boldsymbol{K}^{+}\varphi = f + \sum_{i=1}^{p} \frac{\boldsymbol{U_{\cdot,i}^T}\varepsilon}{s_i}\boldsymbol{V_{\cdot,i}} \end{aligned} \tag{25}$$

where ε is an additive noise vector and φ' is the noise free data. Considering that

$$f - \tilde{f} = \boldsymbol{K}^{+}(\varphi - \varphi') = \boldsymbol{K}^{+}\varepsilon \tag{26}$$

an upper bound for this noise amplification is given by

$$\frac{\left\|f-\tilde{f}\right\|_2}{\|f\|_2} \leq \frac{s_1}{s_p}\frac{\|\varphi-\varphi'\|_2}{\|\varphi\|_2} \tag{27}$$

where $\|\cdot\|_2$ denotes the regular 2-norm and $cond(\boldsymbol{K}) = s_1/s_p$ is the condition number of $\boldsymbol{K}$. As shown in the Results, it is not rare in 1-D CSD estimation to have condition numbers largely exceeding 1000. This means that unless there are some very lucky cancellation effects, noise at a level of 1 part in 1000 is potentially enough to make the solution practically useless.

A common way to circumvent this difficulty is to filter the singular values by multiplying them by a regularizing function $\omega_\lambda(s_i)$ designed such that $\omega_\lambda(s) \to 0$ for $s \to 0$ (Vogel, 2002). Thereby, the exact solution from Equation 24 is replaced by the regularized solution

$$\tilde{f}_\lambda = \sum_{i=1}^{p} \omega_\lambda(s_i)\frac{\boldsymbol{U}_{\cdot,i}^{\boldsymbol{T}}\varphi}{s_i}\boldsymbol{V}_{\cdot,i} \tag{28}$$

This approach is sometimes referred to as regularization by spectral filtering with the individual entries of the regularization function being called filter factors (Hansen and Society for Industrial and Applied Mathematics., 2010). In the present work, we focus on 3 regularization schemes differing only in their regularizing function $\omega_\lambda(s)$:

- Truncated SVD (TSVD)

$$\omega_\lambda(s) = \begin{cases} 1 & if \quad s^2 > \lambda \\ 0 & if \quad s^2 \leq \lambda \end{cases} \tag{29}$$

- Tikhonov

$$\omega_\lambda(s) = \frac{s^2}{s^2+\lambda^2} \tag{30}$$

- Damped SVD (dSVD)

$$\omega_\lambda(s) = \frac{s}{s+\lambda} \tag{31}$$

It is obvious that all these methods function in the same way: they dampen the effect of small singular values by reducing the contribution of their corresponding singular vector $\boldsymbol{V}$ to the solution. As mentioned above, this process can be likened to a low pass filter because the singular vectors associated with the smaller singular values are increasingly oscillatory. Damped

SVD is similar to Tikhonov regularization but the transition between including large and excluding small singular values is slower.

3.2.6 Choosing a prior

So far we have only considered the case where the minimum norm of the model (in the case of quadrature) or the minimum norm of the coefficients in expansion methods is enforced using SVD. However, it is also possible to enforce the minimum norm of smoothness measures of the model $\left\|\boldsymbol{L}\tilde{f}_\lambda\right\|_2$ ($\|\boldsymbol{L}\tilde{\alpha}_\lambda\|_2$ for expansion methods) using generalized SVD (GSVD) instead of SVD. The matrix $\boldsymbol{L}$ is called a prior and is usually chosen to measure the size or the roughness of the function $\tilde{f}_\lambda$. In the case of quadrature methods, $\boldsymbol{L}$ would then respectively be the identity matrix $\boldsymbol{I}$ or a discrete approximation to a derivative of the function (the first and second derivative are the most common choices). The situation is a bit more difficult with expansion methods because the expansion coefficients α_j do not represent the function $\tilde{f}_\lambda$ but only the contribution of their respective basis function. It is however still possible to define measures of the size and/or the roughness of $\tilde{f}_\lambda$ in terms of its expansion coefficients. By expressing the function f in terms of its expansion as in Equation 10, we find that the norm of its derivative of order d is given by

$$\begin{aligned}
\left\|f^{(\boldsymbol{d})}(z)\right\|_2^2 &= \int_a^b \left(\sum_{i=1}^{N} \alpha_i * \vartheta_i^{(d)}(z')\right)\left(\sum_{j=1}^{N} \alpha_j * \vartheta_j^{(d)}(z')\right) dz' \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j \int_a^b \vartheta_i^{(d)}(z')\vartheta_j^{(d)}(z')dz' \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j \, \Lambda_{ij} \\
&= \alpha^T \boldsymbol{\Lambda} \alpha
\end{aligned} \tag{32}$$

Since $\boldsymbol{\Lambda}$ is composed of the inner products of the basis functions $\vartheta_i^{(d)}$, it is a positive definite matrix (Olver and Shakiban, 2006). This allows us to decompose it into a lower triangular matrix and its transpose $\boldsymbol{\Lambda} = \boldsymbol{L^T L}$ by means of the Cholesky factorization. Thereby, we can rewrite the above relation as

$$\begin{aligned} \left\| f^{(\boldsymbol{d})}(z) \right\|_2^2 &= \alpha^T \boldsymbol{\Lambda} \alpha \\ &= \alpha^T \boldsymbol{L}^{\boldsymbol{T}} \boldsymbol{L} \alpha \\ &= \|\boldsymbol{L}\alpha\|_2^2 \end{aligned} \tag{33}$$

and fulfill the goal of expressing the size of $f^{(\boldsymbol{d})}$ as a function of its expansion coefficients α.

3.2.7 Methods for choosing the regularization parameter

The main difficulty remaining to be solved is to find ways to choose a good value for the regularization parameter λ. The objective of a regularization parameter selection method is to automatically find the value of λ which minimizes the error between the true function f and the regularized solution $\tilde{f}_\lambda$. It is therefore informative to examine the nature of these errors.

Regularization by spectral filtering stabilizes the solution by introducing bias. This implies that there are two types of errors that are traded off against each other. First, there is the bias error introduced by filtering the singular values and second there is the perturbation error resulting from inverting the noise component in the data. Writing the filter factors $\omega_\lambda(s_i)$ as a diagonal matrix $\boldsymbol{W}_\lambda$ allows us to express these as

$$f - \tilde{f}_\lambda = \boldsymbol{V}(\boldsymbol{I} - \boldsymbol{W}_\lambda)\boldsymbol{V}^{\boldsymbol{T}} f - \boldsymbol{V}\boldsymbol{W}_\lambda \boldsymbol{S}^{-\boldsymbol{1}} \boldsymbol{U}^{\boldsymbol{T}} \varepsilon \tag{34}$$

where $\boldsymbol{I}$ is the identity matrix and all the other quantities are the same as in Equation 24 and 25. The contribution of these errors depends monotonically on the value of the regularization parameter λ: If λ is large (oversmoothing), most of the noise will get filtered out but the bias will be increased. If λ is small (undersmoothing), there will be almost no bias but the noise will dominate the solution.

Successful regularization therefore hinges on choosing the regularization parameter in a way to minimize the bias while preventing the perturbation error from blowing-up. In practice, this is far from trivial because neither the true solution f nor the error term ε is known. Therefore, it is necessary to find other measures on the data to choose λ and the literature abounds with regularization parameter selection methods. Here, we have tested three: L-curve, generalized cross-validation (GCV) and the normalized cumulative periodogram (NCP).

The L-curve method aims to balance the residual norm $\left\|\boldsymbol{K}\tilde{f}_\lambda - \varphi\right\|_2$ and a measure of the size of a desirable characteristic of the solution, e.g. $\left\|\boldsymbol{L}\tilde{f}_\lambda\right\|_2$ (respectively $\|\boldsymbol{K}\tilde{\alpha}_\lambda - \varphi\|_2$ and $\|\boldsymbol{L}\tilde{\alpha}_\lambda\|_2$ for expansion methods). The balance is achieved by changing the value of the regularization parameter λ. The L-curve rests on the observation that, away from the optimal solution, the model norm tends to increase rapidly with decreasing λ (undersmoothing) or the residual norm increases rapidly with increasing λ (oversmoothing). Therefore, sweeping through various values of λ optimally generates an L-shaped curve giving the name to the selection method. The regularization parameter is then chosen by identifying the point of maximal positive curvature in the L-curve.

In contrast, GCV selects λ by minimizing the prediction error of the potential φ under the assumption that the additive noise is white. In practice, this is done by separating the data in two sets. An estimate of the solution is found using the first set and then used to predict the values in the second set. The parameter λ for which the error between the predicted values and the values in the second set is minimized is then selected. Here, we have used leave-one-out generalized cross-validation as described in (Hansen, 2007).

Finally, also based on the assumption that the additive noise is white, NCP chooses the regularization parameter λ such that the vector of residuals, $r = \boldsymbol{K}\tilde{f}_\lambda - \varphi$ or ($r = \boldsymbol{K}\tilde{\alpha}_\lambda - \varphi$ for expansion methods), is closest to flat (white residuals). In practice, this is done by choosing λ such that the cumulated sum of the power spectrum of the residual vector (also called the normalized cumulative periodogram) is closest to linear (Hansen and Society for Industrial and Applied Mathematics., 2010).

3.2.8 Solution resolution

We have already mentioned that regularization stabilizes the estimation at the cost of introducing bias. However, this is not the only cost because the filtering of singular values also

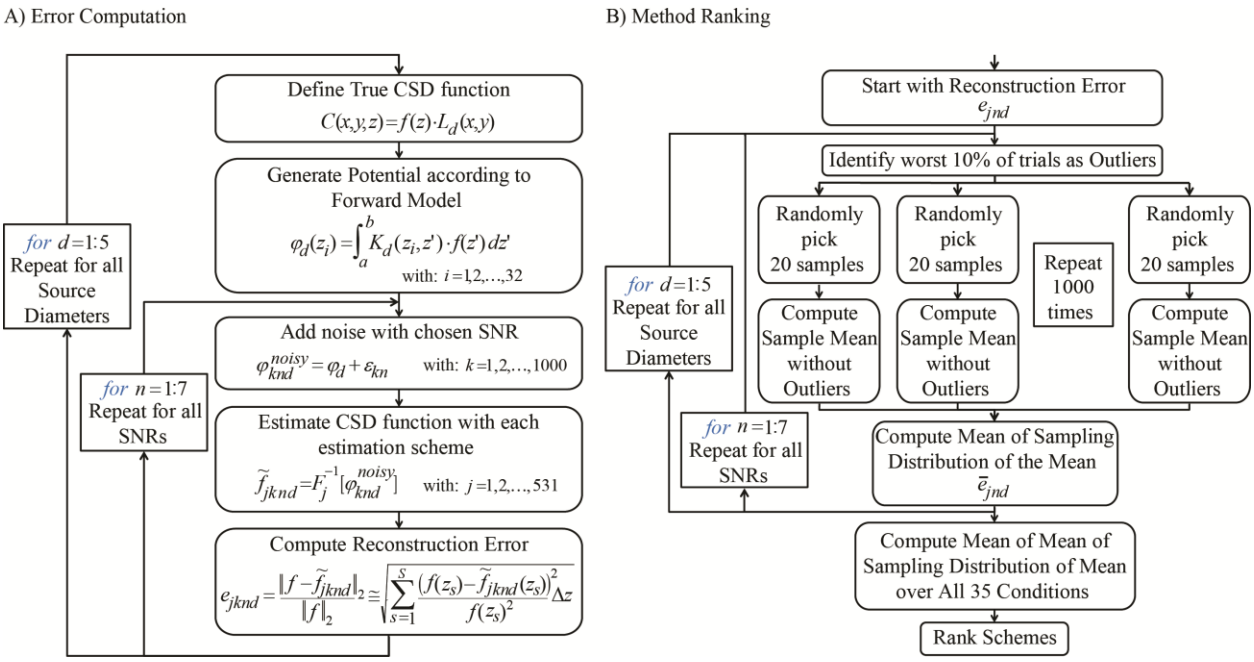

Figure 3-3: Flowchart describing the estimation scheme comparison process. (A) Simulation process and computation of the source reconstruction error. (B) Ranking of estimation schemes according to the mean over all conditions of the mean of the sampling distribution of the mean.

affects the resolution of the solution. This concept is easiest to understand by looking at the forward and inverse estimation process assuming no noise

$$\begin{aligned} \varphi &= \boldsymbol{K}f \\ \tilde{f} &= \boldsymbol{K}^{+}\varphi \\ \Rightarrow \tilde{f} &= \boldsymbol{K}^{+}\boldsymbol{K}f \end{aligned} \tag{35}$$

By identifying the estimation resolution with $\boldsymbol{R}=\boldsymbol{K}^{+}\boldsymbol{K}$, we see that the resolution matrix will not be an identity matrix if $\lambda > 0$ in $\boldsymbol{K}^{+}$. Therefore, whenever a regularized inverse is obtained, it is useful to check the resolution of the estimation to verify that interesting variations in the solution are indeed within the resolution capability of the inversion method.

In practice, the model resolution is usually qualitatively evaluated using a so-called spike-test (we have called it delta test here in order to avoid confusion with a neural action potential). This means simply to replace the model f with a delta function and then to evaluate how well the delta function is resolved using the inversion matrix $\boldsymbol{K}^{+}$ with the estimated λ from the experimental data of interest. The full mathematical treatment of this topic goes beyond the

scope of this work but interested readers are referred to (Menke, 1989, Parker, 1994, Aster et al., 2005) for a detailed discussion.

3.3 Methods

We have compared the accuracy of different CSD methods under various regularization schemes in order to identify the methods that are likely to work best in a noisy experimental setting. We have compared the previously published (1) spline iCSD (Pettersen et al., 2006), (2) kCSD with Gaussian basis functions (Potworowski et al., 2012) and the newly presented (3) eCSD which uses the same Gaussian basis function expansion as kCSD but without the kernel trick, (4) qCSD using Simpson's quadrature rule and (5) rCSD. Each of these CSD methods was tested under 3 regularization schemes (Tikhonov, dSVD, tSVD), 3 regularization coefficient selection methods (l-curve, GCV, NCP) and 7 different priors. Moreover, we have applied the priors in 2 different ways for the expansion methods. We called a particular combination of these methods an estimation scheme. Hence, each scheme is composed of 5 parameters: a CSD method (iCSD, kCSD, eCSD, qCSD, or rCSD), a regularization method (Tikhonov, dSVD, or tSVD), a regularization parameter selection method (NCP, L-curve, GCV), a prior (see below for detailed explanation) and a way of enforcing the prior (model or coefficient). This results in a total of 531 different estimation schemes after removal of impossible or redundant schemes (e.g. qCSD with model prior enforcement). The behavior of each estimation scheme has been evaluated for 5 diameters (0.5, 1, 2, 3 and 5 mm) under 7 noise levels (0, 1, 2, 3, 5 and 10 dB). The reconstruction accuracy of each scheme was evaluated using the process schematically shown in Figure 3-3A. The following describes these steps in more detail.

3.3.1 "True" CSD function $\boldsymbol{f(z)}$, forward model and noise

For the underlying "true" CSD function $f(z)$ we have chosen 8 different CSD depth profiles. As ground profile, we used a sum of Gaussians similar to the one used in (Pettersen et al., 2006)

$$f(z) = \begin{cases} 0, & z \leq 0 \\ e^{-\frac{1}{2}\frac{(z-0.3)^2}{0.08^2}} - \frac{e^{-\frac{1}{2}\frac{(z-0.8)^2}{0.23^2}}}{\sqrt{2\pi}}, & z > 0 \end{cases} \tag{36}$$

The detailed description in the results section is based on this profile. In order to assess the stability and reproducibility of the methods ranking obtained from this "Sum of Gaussian" profile, we repeated the ranking for 7 additional CSD depth profiles:

1, 2 and 3: Small 1 ($\mu = 0.3$ mm; $\sigma = 0.08$ mm), Small 2 ($\mu = 1$ mm; $\sigma = 0.08$ mm) and Large ($\mu = 1$ mm; $\sigma = 0.2$ mm) monopole

$$f(z) = \begin{cases} 0, & z \leq 0 \\ -e^{-\frac{1}{2}\cdot\frac{(z-\mu)^2}{\sigma^2}}, & z > 0 \end{cases} \tag{37}$$

4: Large dipole (the sum of Gaussians negates the need for an additional small dipole)

$$f(z) = \begin{cases} 0, & z \leq 0 \\ e^{-\frac{1}{2}\cdot\frac{(z-0.75)^2}{0.15^2}} - e^{-\frac{1}{2}\cdot\frac{(z-1.5)^2}{0.15^2}}, & z > 0 \end{cases} \tag{38}$$

5 and 6: Small ($\mu = [0.3, 0.6, 0.9]$ mm; $\sigma = 0.08$ mm) and Large ($\mu = ([0.5, 1, 1.5]$; $\sigma = 0.13$ mm) quadrupole. In both cases $A = [0.5, -1, 0.5]$.

$$f(z) = \begin{cases} 0, & z \leq 0 \\ \sum_{i=1}^{3} A_i \cdot e^{-\frac{1}{2}\cdot\frac{(z-\mu_i)^2}{\sigma^2}}, & z > 0 \end{cases} \tag{39}$$

7: CSD depth profile described in Glabska et al. (Glabska et al., 2014). In order to obtain the Glabska profile, we first extracted the CSD depth profile at $x = 0$ from their Figure 3-4B. We then transformed their color coded image into intensity values using a reverse lookup table approach and fitted the obtained depth profile with 7 Gaussian functions. This provided us with a parametric description for the Glabska profile given by

$$f(z) = \begin{cases} 0, & z \leq 0 \ \cup z \geq 2.045 \\ \sum_{i=1}^{7} A_i \cdot e^{-\frac{1}{2}\cdot\frac{(z-\mu_i)^2}{\sigma_i^2}}, & 0 < z < 2.045 \end{cases} \tag{40}$$

with

$$A = [-2.810, 5.925, -14.202, 13.283, 9.637, 6.356, -7.89\ ]$$
$$\mu = [0.021, 0.175, 0.812, 0.84, 1.142, 1.305, 1.680\ ]$$
$$\sigma = [\ 0.062, 0.145, 0.281, 0.148, 0.111, 0.216, 0.104]$$

For the lateral source profile we have used the cylinder assumption (see Eq. 3) with diameters of either 0.5, 1, 2, 3 or 5 mm.

From these CSD distributions, we simulated the potential $\varphi(z_i)$ for 32 electrodes (inter-electrode spacing of 100 μm) according to the procedure described in Figure 3-2B. The choice of this setup is meant to simulate a cortical recording with a linear 32 channel electrode. The medium was parameterized as 2 semi-infinite media separated at $z = 0$ differing only in their electrical conductivity (homogeneous and isotropic conductivity of $\sigma_t = 1.7\ S/m$ for the top medium and $\sigma_e = 0.3\ S/m$ for the bottom medium). The bottom medium conductivity σ_e corresponds to the conductivity generally assumed for extracellular cortical tissue while σ_t is the conductivity of saline (Wagner et al., 2004). The virtual electrode was positioned to have 4 contacts in the top medium and the 5th contact positioned at $z = 0.05\ mm$, i.e. 50 μm below the surface. This setup is a realistic representation of a feasible cortical recording because a 32 channel electrode has a larger span (3.1 mm) than most mammalian cortices. The assumed setup is schematically displayed in Figure 3-2A.

For each level of signal-to-noise ratio (SNR) we then generated 1000 noise realizations of the appropriate amplitude and added them to the noise-free potential φ. The resulting noisy potential $\varphi^{noisy} = \varphi + \varepsilon$ was used as an input to the different source estimation schemes. It should be noted that the 531 estimation schemes were evaluated on the same data, i.e. we did not regenerate different noise for each scheme.

3.3.2 Priors

We have tested the methods under 7 different priors (the symbol used for each prior is shown in square parentheses): 1) minimization of the quadrature/expansion coefficients []; (2) minimization of the model norm [0], i.e. $\boldsymbol{L} = \boldsymbol{I}$ for qCSD or $\boldsymbol{L}$ corresponding to $d = 0$ in Equation 33 for the expansion methods; (3) minimization of the first derivative [1]; (4) minimization of the second derivative [2]; (5) minimization of the combination of the model norm and the first derivative [0 1]; (6) minimization of the combination of the model norm and the second derivative [0 2]; (7) minimization of the combination of the model norm, the first derivative and the second derivative [0 1 2]. In the case of qCSD each of these methods is implemented using the identity matrix $\boldsymbol{L} = \boldsymbol{I}$ for measuring the model norm or with $\boldsymbol{L}$ chosen as a discrete approximation to the d-th derivative. In the case of expansion methods there are actually two ways in which the priors can be applied. In the first case, we can construct $\boldsymbol{L}$ such that it

measures the relevant characteristic, i.e. norm or norm of a derivative of the function f, as shown in Equations 32 and 33. In the second case, we can directly minimize the norm of the coefficients α (or its derivative). For example, minimizing the first derivative would then prioritize all coefficients having the same value instead of directly enforcing a small norm for the first derivative f'. We decided to explore both these options here because calculating the matrix $\boldsymbol{L}$ involves numerical approximation of the integral from Equation 32 which potentially introduces errors/instability into the problem. Thus, we also applied all the priors used in qCSD, i.e. the identity matrix and the discrete derivatives, to the expansion methods. We refer to these priors as the *coefficient* priors in contrast to the *model* priors. Note that this distinction is meaningless for qCSD because it is not a basis function expansion method.

Finally, all priors except the first are implemented using GSVD instead of SVD. We note that prior [] and [0] (for the *coefficient* priors) are enforcing the same condition but using either SVD or GSVD. Hence, these conditions are useful in assessing the effect of numerical differences between the two methods.

3.3.3 Reconstruction error

The source reconstruction error was estimated using a discrete approximation (361 samples regularly spaced in 10 micron steps) to the 2-norm of the error

$$e = \frac{\left\| f - \tilde{f}_\lambda \right\|_2}{\|f\|_2} \cong \sqrt{\sum_{s=1}^{361} \frac{\left(f(z_s) - \tilde{f}_\lambda(z_s) \right)^2}{f(z_s)^2} \Delta z} \tag{41}$$

$$z_1 = -0.6\ mm;\ z_{361} = 3\ mm$$

3.3.4 Comparison of estimation schemes

We compared the 531 estimation schemes under 7 different noise levels and with 5 different source diameters (a total of 35 conditions). In order to rank the estimation schemes, we first created a sampling distribution of the mean estimation error by subsampling the full 1000 trial dataset 1000 times with a sample size of 20 samples without replacement. We then computed the mean of the sampling distribution of the mean and the standard error of the mean and ranked the estimation schemes from lowest to highest mean error. This approach was chosen to answer the following question: “If one records 20 trials of noisy potential, which method will

most likely provide the lowest average CSD estimation error?" Moreover, to avoid excessively contaminating the estimate with outliers, we have discarded the worst 10% of trials for each estimation scheme. Thereby, any method that has a larger number of outliers is implicitly penalized because the remaining outliers will increase the mean error. A schematic of the whole procedure is shown in Figure 3-3B.

3.3.5 Simulation environment

All the simulations were conducted using a custom-made MATLAB (MathWorks) package which we made available for download at http://www.bic.mni.mcgill.ca/~amirs/. This package builds on the freely available regu-tools by Hansen which implements all the SVD and GSVD based regularization methods (Hansen, 2007). In order to handle the large amount of computation required for the comparison we have used a high-performance computing cluster (Guillimin, McGill University) for running the simulations.

3.4 Results

3.4.1 Noise sensitivity

We first evaluated the condition number associated with the naïve (unregularized) CSD estimation to assess the upper bound of noise amplification as defined in Equation 42. Table 3-1A shows the condition number for each estimation scheme and each evaluated source diameter. As expected, increasing the source diameter leads to a larger condition number because of the increased smoothness of the forward kernel K (see Equation 4). This in turn implies that CSD estimation is more difficult at larger source diameters. Moreover, the large condition number suggests that CSD estimation is potentially highly sensitive to noise. In fact, a condition number of 10^6 means that noise at an amplitude of 1 part in a million could be enough to significantly affect the estimation quality. However, it is important to keep in mind that Equation 42 only provides an upper bound on noise amplification. Therefore, to get a better picture of the need for regularization, we have explicitly computed the average noise amplification over trials for the naïve CSD estimation using the following equation

$$NoiseAmp = Mean\left[\frac{\left\|f - \tilde{f}\right\|_2}{\|\varphi - \varphi'\|_2}\right] = mean[A] \cdot \frac{\|f\|_2}{\|\varphi\|_2} \tag{42}$$

A) Condition Number:

	Diam 0.5	Diam 1	Diam 2	Diam 3	Diam 5
rCSD	4.76E+04	3.85E+05	2.62E+06	7.18E+06	2.22E+07
kCSD	2.44E+05	1.94E+06	1.31E+07	3.59E+07	1.11E+08
qCSD	2.50E+02	7.08E+02	1.85E+03	3.05E+03	5.38E+03
iCSD	1.15E+03	1.13E+03	2.08E+03	3.43E+03	6.02E+03
eCSD	4.94E+02	1.39E+03	3.62E+03	5.99E+03	1.05E+04

B) Noise Amplification:

	Diam 0.5	Diam 1	Diam 2	Diam 3	Diam 5
rCSD	63.54	61.48	60.70	60.61	61.12
kCSD	67.23	64.86	63.97	63.82	64.46
qCSD	68.53	66.20	65.29	65.11	65.73
iCSD	64.70	62.65	61.84	61.70	62.28
eCSD	67.23	64.86	63.97	63.82	64.46

Table 3-1: Condition Number and Noise Amplification. (A) Condition number for each CSD methods and each evaluated source diameter. As expected, the condition number increases for increasing source diameter because representers become flatter. (B) Noise amplification for all methods and source diameters computed according to Eq. 38. Note that the noise amplification was averaged over the different SNRs.

Table 3-1B shows the computed noise amplification for all methods and source diameters (the noise amplification was averaged over the different SNRs). Although the picture is not as bleak as suggested by the condition number, we can see that noise still has a considerable effect on estimation accuracy. It is interesting to see that the noise amplification as defined above is independent of the diameter for all methods.

3.4.2 Demonstration of the regularization

Before proceeding to the comparison of the various CSD estimation schemes, it is useful to build some visual intuition about the functioning of regularization. This is especially important in the present case because the large number of compared schemes makes it impossible to show the reconstruction for each case. Moreover, because a central aim of this work is to assess the performance of the estimation under various noise conditions, a visual inspection of every trial is impossible. Therefore, the following section will present the general characteristics of the SVD regularized CSD estimation using rCSD as an example.

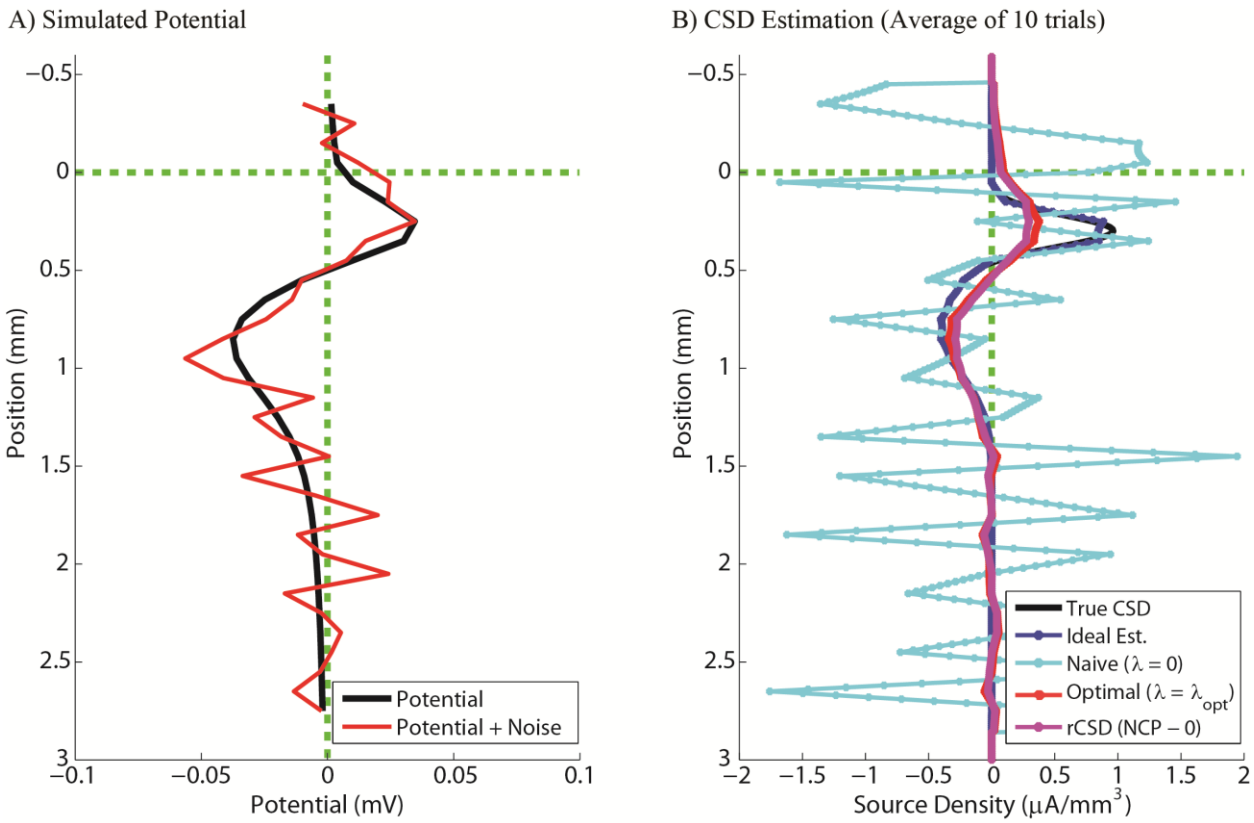

Figure 3-4: Demonstration of CSD estimation using the rCSD expansion method, Tikhonov regularization, NCP regularization parameter selection and a minimum model norm. (A) Simulated potential without noise (black) and with additive Gaussian white noise at an SNR of 3 dB (red). (B) CSD estimation averaged over 10 trials. Ideal (blue) estimation is computed assuming no noise. Optimal (red) is computed by finding the λ that minimizes the reconstruction error. Naive (cyan) estimation corresponds to the source reconstruction obtained without regularization (λ = 0)

In Figure 3-4, we demonstrate the average CSD estimation over 10 trials using the rCSD expansion method, Tikhonov regularization, NCP regularization parameter selection and a minimum model norm. Figure 4a shows the noise-free potential φ (black curve) and the potential with additive Gaussian white noise at a signal-to-noise ratio of 3 dB, simulating experimentally measured potential. The potential has been calculated for an electrode with 32 contacts spaced 100 μm apart using the method demonstrated in Figure 3-2B and the true CSD depth profile named "True CSD" from Figure 3-4B (same as Fig. 3-2B). A cylinder with a radius of 0.25 mm was used for the lateral CSD source profile (see Eq. 3). The electrode was positioned in a way to leave 4 contacts above the brain surface which is shown by the horizontal green dotted line at depth $z = 0$. The 5th electrode is positioned 50 μm below the brain surface. The electrical conductivity above the brain ($z < 0$) was set to 1.7 S/m to emulate the conductivity of artificial

cerebro-spinal fluid (aCSF) and the typical conductivity for gray matter, $0.3\ S/m$, has been chosen for $z \geq 0$ (Wagner et al., 2004, Pettersen et al., 2006). All these parameters were chosen to emulate a laminar cortical recording with aCSF covering the brain surface.

Figure 3-4B shows the average reconstruction over 10 trials using the rCSD expansion method, along with the regularization scheme described above. The ideal estimation (blue curve) was calculated by using a naïve estimation ($\lambda = 0$) on the noisefree potential. This gives a bench mark value for the best possible estimation using the chosen expansion method. It is interesting to note that the ideal estimation is not perfect as shown by the small dent on the positive peak of the CSD curve. This could potentially be explained by the fact that the rapid fluctuation of the CSD profile at the peak does not lay within the span of the expansion functions. However, as can be seen by the actual estimation (magenta curve), this small lack in fidelity is meaningless for the CSD estimation under noise. The naïve estimation (cyan curve) is found using $\lambda = 0$ on the noisy potential. This is the estimation one would get without regularization. As expected, there is substantial noise amplification even after averaging over 10 trials. Moreover, it is interesting to note that the noise amplification seems larger outside of the brain where the conductivity is higher. The optimal lambda estimation (red curve) is calculated by finding the regularization parameter λ which minimizes the error between the ideal CSD and the estimated CSD. This estimation represents the best possible estimation from the noisy potential under the chosen estimation scheme. The optimal estimation depends on the chosen CSD method (in this case rCSD), the regularization method (Tikhonov), the prior and the particular noise realization. However, because the regularization parameter λ_{opt}. is found by minimizing the error with the ideal CSD, it is independent of the particular regularization parameter selection method. In fact, the quality of a regularization parameter selection method can be evaluated by looking at how similar the value of the regularization parameter λ (found using the particular scheme) is to λ_{opt}.

It is clear from Fig. 3-4B that the regularization stabilizes the estimation since the rCSD estimation (magenta curve) is much smoother than the naïve estimation. However, it also exemplifies the cost of the increased stability. Because of the filtering of the right singular vectors $\boldsymbol{V}$ associated with small singular values, the estimation is now biased as seen by the reduced amplitude of the estimated CSD compared to the "True CSD". Furthermore, the

closeness of the rCSD estimation and the optimal estimation suggests that the blind application of the NCP regularization parameter selection works almost optimally in this case.

3.4.3 Comparison of regularization parameter selection methods

A major challenge in regularization is the choice of the regularization parameter λ. We have compared 3 different regularization parameter selection methods: NCP, L-curve and GCV. Figure 3-5A-C shows the process of regularization parameter selection for 1 trial for each of these selection methods. All parameters for the generation of the potential and the source profile are the same as in the previous section. In NCP (Fig. 3-5A), the regularization parameter λ is chosen such that the vector of residuals is closest to flat (white residuals). In practice, this is done by choosing λ such that the cumulated sum of the power spectrum of the residual vector (also called the normalized cumulative periodogram) is closest to linear (Hansen and Society for Industrial and Applied Mathematics., 2010). The changing of shape of the NCP as λ is varied is shown in Fig. 3-5A.

Instead of focusing on the shape of the residuals, the L-curve criterion attempts to balance the residual norm and the model (semi-) norm. The L-curve rests on the assumption that, away from the optimal solution, the model norm will increase rapidly with decreasing λ (undersmoothing) or the residual norm will increase rapidly with increasing λ (oversmoothing). Therefore, sweeping through various values of λ optimally generates an L-shaped curve giving the name to the selection method. The regularization parameter is then chosen by identifying the point of maximal positive curvature in the L-curve. It is not necessarily the case that the optimal λ really lies at the point of maximal curvature as can be seen in Fig. 3-5B, but it generally lays close by. However, especially with very smooth models, the L-curve does not always have a point of positive curvature in which case the L-curve method will fail (Hansen and Society for Industrial and Applied Mathematics., 2010).

Finally, GCV selects λ by minimizing the prediction error of the potential φ under the assumption that the additive noise is white. Here, we have used leave-one-out generalized cross-validation as described in Hansen 2007 (Hansen and Society for Industrial and Applied Mathematics., 2010). The same approach is used in Potworowski et al. (Potworowski et al., 2012). However, it is important to mention that, although the principle is the same, the

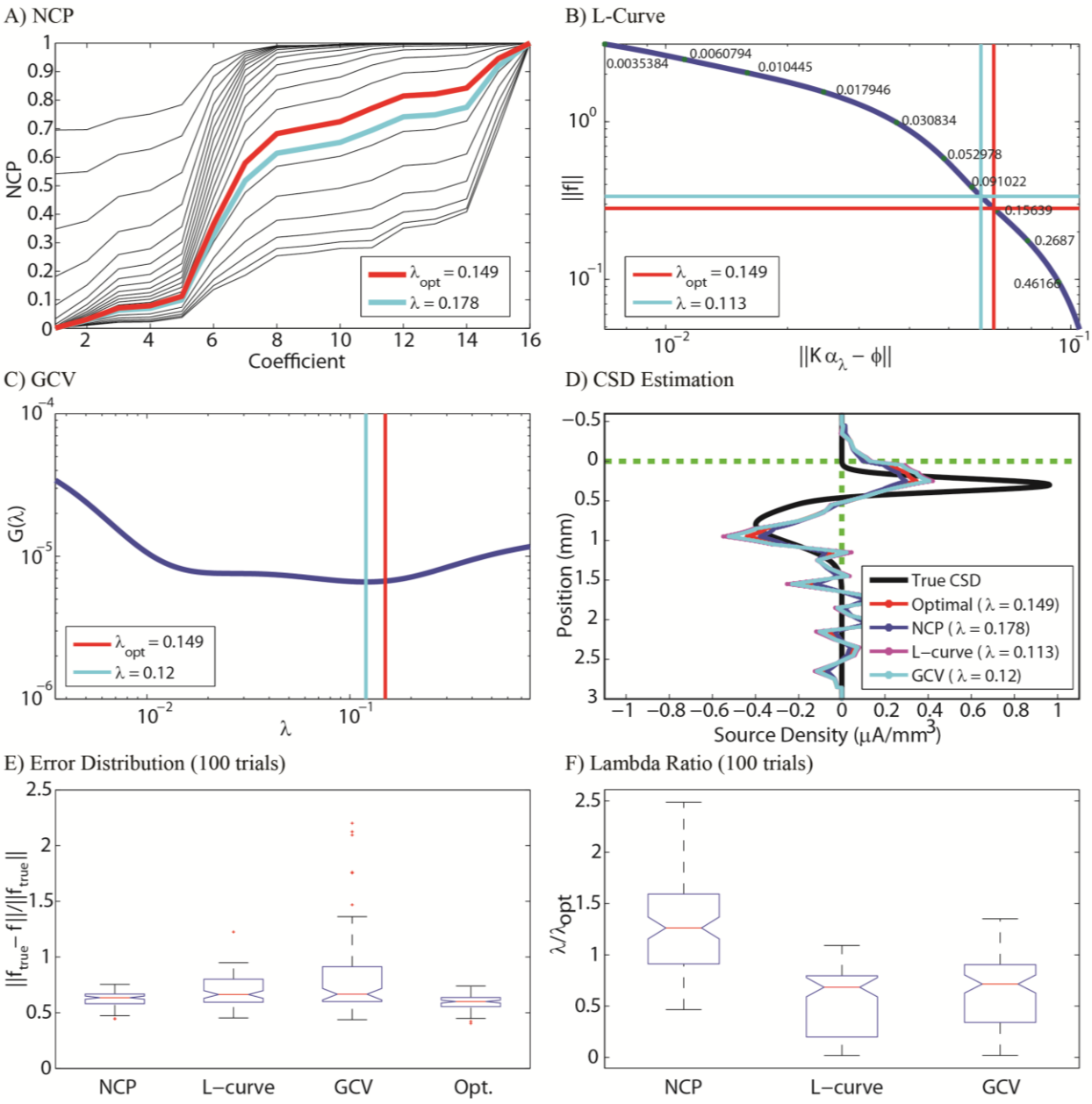

Figure 3-5: Demonstration of regularization parameter selection methods. (A) NCP attempts to find the normalized cumulative periodogram that is closest to linear. The selected curve is shown in cyan and the optimal in red. The optimal is found by minimizing the estimation error as measured by the 2-norm. (B) The l-curve chooses the regularization parameter by identifying the point of maximal positive curvature. (C) In GCV the minimum of the leave-one-out prediction error is used for selecting λ. (D) Overlay of the estimated CSD using the regularization parameter identified using the methods in (A)-(C). (E) Boxplot of the estimation error for 100 trials. (F) Boxplot of the success of each regularization parameter selection method as evaluated by the ratio between the selected and optimal λ. A value of 1 means optimal parameter selection.

implementation of GCV used here differs from theirs. Figure 3-5C shows the value of the generalized cross-validation function as a function of the regularization parameter λ. We see that the chosen regularization parameter is quite close to the optimal. Interestingly, the cross-validation exhibits two local minima suggesting two different solutions where the prediction error is small. This can be problematic and leads to a higher level of outliers.

Fig. 3-5D shows an overlay of the CSD estimation for 1 trial using the three different regularization parameter selection methods. However, Fig. 3-5D does not necessarily provide a representative view of the general situation since it only shows the resulting estimation for a single trial. In order to get a better picture of the quality of each of the selection methods, we have computed the estimation error for 100 trials (outliers removed). The resulting distribution of the error is shown as a boxplot in Fig. 3-5E. We see that for this particular estimation scheme (i.e. rCSD, Tikhonov, minimum model norm prior) NCP performs best followed by the L-curve criterion. In fact, NCP performs almost optimally. In order to show how each of the selection methods compares to the optimal case, we have also investigated how close the regularization parameter lies to the optimal one. Fig. 3-5F shows the ratio between the regularization parameter λ and λ_{opt} for each trial. The closer this ratio is to one, the better the performance of the regularization parameter selection method. As expected from Fig. 3-5E, NCP also performs better in this respect. Moreover, it is interesting to note that NCP tends to overestimate λ under the chosen conditions while both the L-curve and the GCV tend to underestimate it.

3.4.4 Regularization and resolution

It has already been mentioned that regularization stabilizes the solution at the cost of introducing bias. However, the magnitude of the regularization parameter λ also affects the resolution of the solution. We have explored this reduction in resolution by performing a delta-test which consists of looking at the noise-free source reconstruction assuming $f(z) = \delta(z - z'')$. Figure 3-6A shows the CSD estimation for $z'' = 0.55\ mm$. It is clear that there is a marked difference between the unregularized ($\lambda = 0$) and the regularized case. Most importantly, the estimation goes from a wavelet type shape resembling the impulse response of a high pass filter to a smoothening function akin to a low-pass filter.

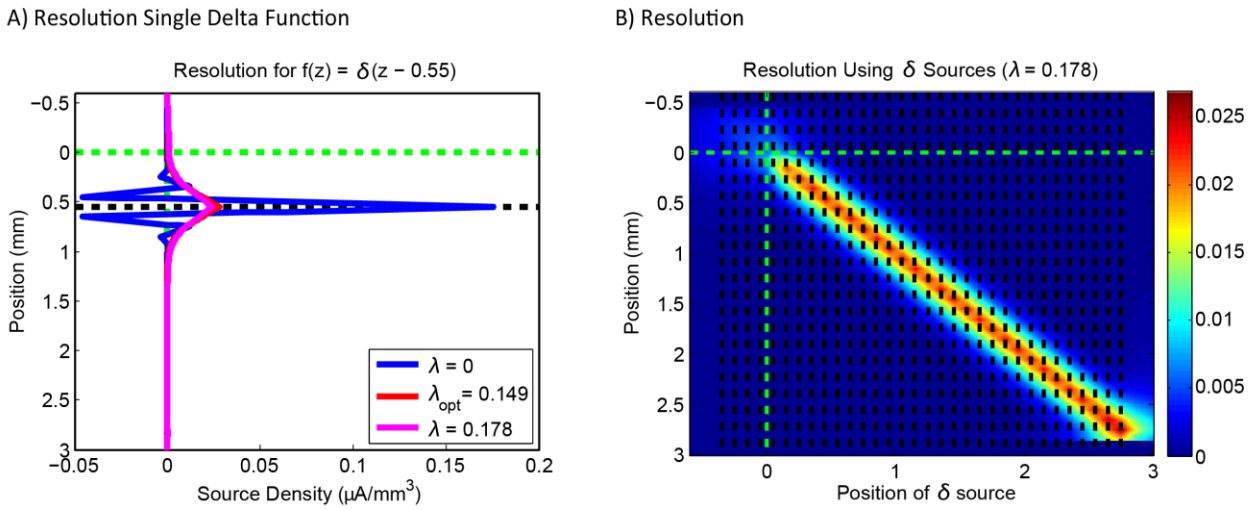

Figure 3-6: Effect of Regularization on Solution Resolution. (A) Resolution of a delta function located at z = 0.55 mm using estimation naive (blue), optimal (red) and NCP-selected (magenta) regularization parameter. The estimation scheme and trial is the same as in Figure 3-5. (B) Pseudo-color image of the resolution of a delta function located at various depths.

Figure 3-6B shows the resolution profiles obtained using a delta-test with the discrete delta functions located at different depths. We see that the conductivity jump at $z = 0$ affects the estimation resolution for $\sim 150\ \mu m$. Moreover, we note that the higher conductivity at $z < 0$ dramatically reduces the estimation. Finally, we also observe a small reduction in resolution towards the bottom-most electrode stemming from the fact that there is no data from below that can be used to constrain the estimate.

The width of the low-pass resolution profile is very useful for a qualitative assessment of estimated CSD profiles because it provides an approximation of the size of the features that can be resolved under the given conditions. In fact, the estimated CSD profile $\tilde{f}(z)$ (assuming no noise) is simply the convolution of the resolution profile with the true CSD profile $f(z)$. This can be nicely approximated by multiplying each column in Figure 3-6B by the corresponding magnitude of the true CSD profile and summing over all rows (the result has to be multiplied by the distance between the delta functions to get the correct magnitude).

3.4.5 Comparison of CSD estimation schemes for each condition for "Sum of Gaussians" profile

So far, we have only demonstrated the working of one estimation scheme under one single condition. This obviously does not give us a very good picture about which estimation scheme

Diam	SNR	CSD Method	Regu. Method	λ Sel. Method	Prior	Prior Method	Mean Error	Standard Error	Mean Noise Amp.
0.5	0	eCSD	Tikhonov	NCP	0	Model - 0	0.6971	0.0189	3.12
	1	kCSD	dSVD	L-curve	0 2	Coeff +	0.6623	0.0212	3.29
	2	kCSD	dSVD	L-curve	0 2	Coeff +	0.6311	0.0163	3.52
	3	iCSD	Tikhonov	L-curve	0 1 2	Coeff +	0.6022	0.0181	3.75
	5	iCSD	Tikhonov	L-curve	0 2	Coeff +	0.5497	0.0157	4.34
	7	iCSD	Tikhonov	L-curve	0 2	Coeff +	0.5085	0.0133	5.07
	10	iCSD	Tikhonov	L-curve	0 2	Coeff +	0.4404	0.0111	6.20
1	0	kCSD	Tikhonov	NCP	0	Model - 0	0.7686	0.0217	1.61
	1	kCSD	Tikhonov	NCP	0	Model - 0	0.7403	0.0201	1.75
	2	iCSD	Tikhonov	L-curve	0 1 2	Coeff +	0.7016	0.0233	1.84
	3	iCSD	Tikhonov	L-curve	0 1 2	Coeff +	0.6804	0.0204	2.02
	5	iCSD	Tikhonov	L-curve	0 2	Coeff +	0.6199	0.0173	2.30
	7	iCSD	Tikhonov	L-curve	0 2	Coeff +	0.5734	0.0145	2.69
	10	rCSD	Tikhonov	L-curve	0 1 2	Model +	0.5109	0.0129	3.40
2	0	rCSD	Tikhonov	NCP	0 2	Model +	0.8489	0.0241	0.91
	1	rCSD	Tikhonov	NCP	0 2	Model +	0.8247	0.0251	1.00
	2	rCSD	Tikhonov	NCP	0 2	Model +	0.7987	0.0236	1.07
	3	rCSD	Tikhonov	NCP	0 2	Model +	0.7627	0.0213	1.16
	5	rCSD	Tikhonov	NCP	0 2	Model +	0.7235	0.0183	1.38
	7	rCSD	Tikhonov	NCP	0 2	Model +	0.6795	0.0159	1.63
	10	iCSD	Tikhonov	L-curve	0 2	Coeff +	0.6175	0.0155	2.08
3	0	kCSD	dSVD	NCP	0 2	Model +	0.8941	0.0179	0.67
	1	iCSD	dSVD	NCP	2	Coeff +	0.8822	0.0191	0.74
	2	rCSD	Tikhonov	NCP	0 2	Model +	0.8685	0.0198	0.81
	3	iCSD	dSVD	NCP	2	Coeff +	0.8559	0.0187	0.89
	5	rCSD	Tikhonov	NCP	0 2	Model +	0.8090	0.0206	1.08
	7	rCSD	Tikhonov	NCP	0 2	Model +	0.7656	0.0186	1.29
	10	rCSD	Tikhonov	NCP	0 2	Model +	0.7165	0.0176	1.68
5	0	kCSD	dSVD	NCP	0 2	Model +	0.9316	0.0144	0.46
	1	iCSD	dSVD	NCP	2	Coeff +	0.9188	0.0166	0.51
	2	iCSD	dSVD	NCP	2	Coeff +	0.9071	0.0169	0.57
	3	kCSD	dSVD	NCP	0 2	Model +	0.8934	0.0181	0.63
	5	iCSD	dSVD	NCP	2	Coeff +	0.8646	0.0184	0.77
	7	iCSD	dSVD	NCP	2	Coeff +	0.8313	0.0181	0.93
	10	kCSD	dSVD	NCP	0 2	Model +	0.7795	0.0178	1.24

Table 3-2: Best ranked estimation scheme for each SNR and source diameter for "Sum of Gaussians" profile. The ranking was obtained according to the procedure described in Figure 3-3B. Note that in this case the last step of averaging over all tested conditions is omitted and each condition is considered independently.

should be preferred in a general situation. For instance, the exact signal-to-noise ratio and the source diameter is usually unknown in an experimental setting. Therefore, we decided to compare the 531 estimation schemes under 7 different noise levels and with 5 different source diameters (a total of 35 conditions). We ranked the estimation schemes based on the mean error of their sampling distribution of the mean computed from subsampling without replacement the dataset 1000 times with a sample size of 20 samples. This approach was chosen to answer the question: "If one records 20 trials of noisy potential, which method will most likely provide the lowest average CSD estimation error?" Moreover, to avoid excessively contaminating the estimate with outliers, we have discarded the worst 10% of trials for each method. Thereby, any method that has a larger number of outliers is implicitly penalized because the remaining outliers will increase the mean error.

Table 3-2 shows the best ranked estimation scheme for each SNR and source diameter combination. As expected from the analysis of the condition number, the mean error increases with increasing diameter. Moreover, unsurprisingly the mean error decreases with increasing SNR. The opposite is true for noise amplification. Moreover, there are some interesting trends in the distribution of the parameters in the top methods. First, regularization involving a prior minimizing the second derivative is dominating the top ranks. In all but 3 conditions the prior involves minimizing the second derivative. Second, there is a clear tendency to favor Tikhonov regularization over dSVD for smaller diameters. A similar tendency is observed for the comparison between the L-curve and NCP regularization parameter selection method. However, the main finding that sticks out from Table 3-2 is that there is no single estimation scheme that completely outperforms the others across all tested conditions. It is therefore necessary to investigate further which types of schemes perform well in the different conditions.

In order to get a better idea of the type of estimation schemes that are found in the top of the ranking in each condition, we chose to consider all schemes whose mean estimation error was within the one-tailed 99% confidence interval of the best scheme's mean error. Each estimation scheme that fulfilled this criterion was considered a good candidate to use in that condition (i.e. $\bar{\varepsilon}_{j,d,n} \leq \bar{\varepsilon}_{j=best,d,n} + 2.54 \cdot SE[\bar{\varepsilon}_{j=best,d,n}]$). The number of estimation schemes that satisfied this selection criterion in each condition is shown in Table 3-3A.

A) # of Methods:

SNR	Diam 0.5	Diam 1	Diam 2	Diam 3	Diam 5
0	47	84	109	98	114
1	48	83	87	100	111
2	38	82	67	103	94
3	40	65	30	90	92
5	31	41	14	82	87
7	27	22	10	65	77
10	20	23	31	55	66

B) Percent Difference:

SNR	Diam 0.5	Diam 1	Diam 2	Diam 3	Diam 5
0	6.96	7.27	7.35	4.76	3.99
1	7.62	7.03	7.51	5.35	4.36
2	6.40	8.58	7.71	5.82	4.66
3	7.64	7.55	6.70	5.88	5.13
5	6.71	6.97	6.19	6.21	5.47
7	6.78	6.40	5.68	6.02	5.53
10	6.24	6.26	6.12	6.06	5.68

Table 3-3: Distribution of methods within the Top Rank. (A) Number of estimation schemes in each condition whose mean estimation error is within the one-tailed 99% confidence interval of the best scheme's mean error in that condition, i.e. $\bar{\varepsilon}_{j,d,n} \leq \bar{\varepsilon}_{j=best,d,n} + 2.54 \cdot SE[\bar{\varepsilon}_{j=best,d,n}]$). (B) Percent difference in error between the best scheme and the last one accepted within the top rank.

Additionally, Table 3-3B shows the percentage difference in mean error between the best scheme and the last accepted as a possible candidate. It is interesting to note that although a fewer number of schemes satisfy the selection criterion at smaller diameters, the percentage difference in mean error is larger. This suggests that the difference among estimation schemes is smaller at larger diameters where CSD estimation is more difficult.

Figure 3-7A displays the number of estimation schemes that satisfy the chosen selection criterion as a function of the number of tested conditions. We see that a small number of schemes (3 – see Table 3-4) satisfy the selection criterion in all conditions. This is an encouraging result for our attempt to identify schemes that are likely to perform well across a range of experimental conditions. To investigate whether certain characteristics dominate the ranking within and across each condition, we counted the number of times the labels identifying each estimation scheme parameter were found within the top ranked schemes (see Fig. 3-7B-F). Each bar within a diameter block corresponds to an SNR value, e.g. the first bar is for SNR = 0 and the last for SNR = 10 dB. It is important to remember that the different labels are not present in equal numbers in the whole set of estimation schemes. For example, because qCSD can only be regularized on the coefficients and not on the model, it makes up only 12% of the schemes instead of 22% for the other CSD methods. This means that if the top ranking was insensitive to

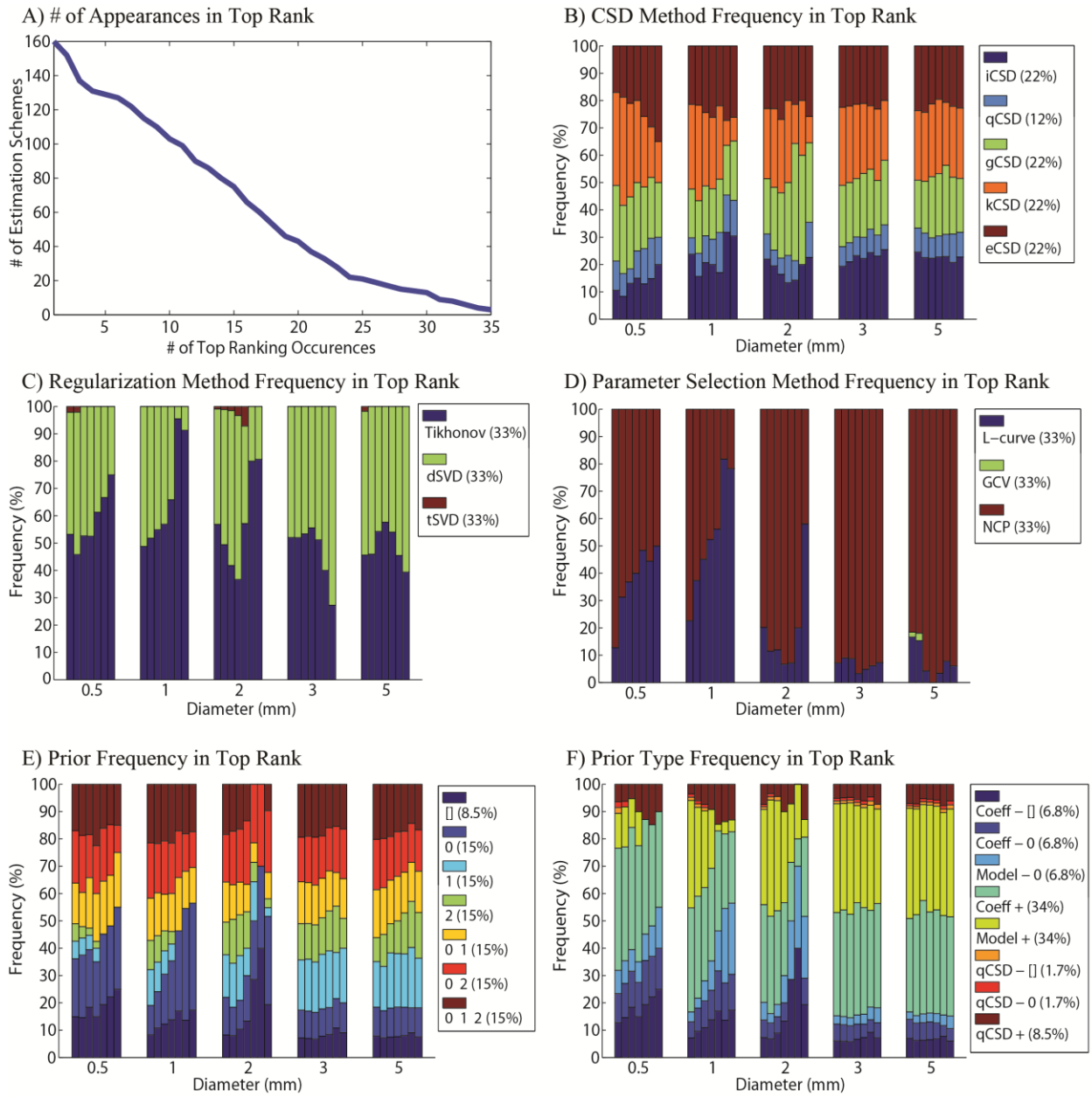

Figure 3-7: Distribution of labels within the top rank. (A) Number of conditions in which a particular estimation scheme was found in the top rank. (B) - (F) Percentage with which the given label appears in the top rank of each tested condition. Note that the number of schemes that were considered in the top rank varies in each condition and is shown in Table 3-3A. Each block corresponds to a diameter and each bar within the block represents a signal-to-noise ratio ([0 1 2 3 5 10] dB). The relative frequency of each label within the 531 estimation schemes is shown as a percentage in each panel legend. The relative frequency is important in order to judge whether a label appears more/less often than expected by chance.

the CSD method, we would expect only 12% of the labels to be qCSD in Figure 3-7B.

Therefore, we show the relative frequency of each label in the full set of schemes in the respective panel legend.

Figure 3-7B shows the relative distribution of CSD method labels. We see that at small diameters, iCSD is represented less frequently than expected by chance thus suggesting that iCSD is not an ideal choice in this situation. Moreover, iCSD appears to perform better at larger SNRs in contrast to kCSD. In fact, there is an interesting trade-off between kCSD and eCSD. At the smallest diameter, it is very clear that kCSD and eCSD evolve in opposite directions with increasing SNR. These variations are much less pronounced at larger diameters where all methods occur roughly proportionally to their relative frequency in the dataset. However, overall, the top ranking scheme identity seems only weakly related to CSD label. This means that the type of CSD method chosen is not a major predictor of estimation accuracy. This is in clear contrast to the relative occurrence of regularization scheme shown in Panel 3-7c. First, we note that tSVD is almost absent from the top ranking. This is not particularly surprising because tSVD is a cruder method than the two others. Second, there is a clear trend for Tikhonov to perform better than dSVD especially at small diameters and large SNRs.

This discrepancy is even more pronounced in the frequency distribution of the regularization parameter selection methods. Again, the top rankings contain almost exclusively L-curve and NCP. Moreover, it is obvious that NCP outperforms the L-curve except for small diameters and large SNRs. It turns out that the absence of GCV is mainly explained by its increased rate of outliers. When increasing the rejected trials to 25%, GCV is much more heavily represented (data not shown).

For the predictive power of the choice of prior, the situation is again split between small and large diameters (see Fig. 3-7E). At large diameters, the methods appear more equal and the priors are more evenly distributed. This is not the case for small diameters where penalizing the model norm (either directly the model or simply the coefficients) seems to be a promising approach. In order to study the choice of priors in more depth, Figure 3-7F also shows the priors broken up according to the way they are applied. Again, the discrepancy between small and large diameters is clearly visible. While penalizing higher order functions seems to be a promising approach for larger diameters regardless of the way it is done, i.e. measuring the model norm or

the norm of the coefficients, it is clearly preferable to penalize the coefficients at smaller diameters. Moreover, at smaller diameters it seems best to implement a minimum norm prior directly. Finally, we also observe that penalizing the coefficients norm using SVD directly (Coeff –[]) is preferable over its GSVD alternative (Coeff – 0) despite them implementing the same prior. The increased algorithmic complexity and numerical approximations of GSVD appears to negatively influence the estimation accuracy.

3.4.6 Comparison of CSD estimation schemes over all conditions for "Sum of Gaussians profile"

Since it is notoriously difficult to know the true source diameter in an experimental setting and one generally does not know the exact signal-to-noise ratio, it is useful to try to select a CSD estimation scheme that performs well across a large set of conditions. We have already observed in Figure 3-7A that there are estimation schemes that consistently appear in the top of the ranking suggesting that it should be possible to create a meaningful ranking across conditions. Therefore, we created an overall ranking of the estimation schemes by ordering the schemes according to the mean over the mean of the sampling distribution of the mean errors of each of the 35 conditions. The top 40 estimation schemes in the overall ranking are shown in Table 3-4. Looking at the characteristics of these top estimation schemes, it stands out that NCP clearly outscores the other regularization parameter selection methods since every scheme in the list relies on it. Next, the results for the preferable regularization methods are slightly favoring dSVD over Tikhonov (tSVD did not make it into the list at all): dSVD (25), Tikhonov (15). In terms of CSD methods, the results are distributed as follows: rCSD (10), kCSD (13), eCSD (11), qCSD (3) and iCSD (3). Although there is a clear preference towards rCSD, kCSD and eCSD, the fact that all methods are represented shows that none of them is definitely worse than the others.

To examine the results for the top estimation schemes in more detail, a distribution of the error (with outliers removed) over the whole dataset is shown as a violin plot in Figure 3-8A. We see that their error distributions are very similar which supports our choice of considering them as similarly good candidates. We also performed a similar analysis on the noise amplification as defined in Eq. 42. Again, the distributions were found to be very similar with a mean noise amplification of ~ 2 (median of ~ 1.5), which constitutes a substantial improvement over the mean noise amplification of ~ 60 for naïve CSD estimation (see Table 3-1B). Finally, Figure 3-8B shows the distribution of the ratio of the estimated regularization coefficient λ and the

Rank	CSD Method	Regularization Method	λ Selection Method	Prior	Prior Method	# Top Ranks	Mean Error	Standard Error	Mean Rank	Std Rank	% Diff.
1	rCSD	Tikhonov	NCP	0	Model - 0	35	0.7581	0.0184	14.89	6.26	0.00
2	kCSD	Tikhonov	NCP	0	Model - 0	35	0.7585	0.0182	16.51	10.98	0.06
3	iCSD	Tikhonov	NCP	[]	Coeff - []	35	0.7590	0.0183	14.49	8.25	0.12
4	rCSD	dSVD	NCP	0 1	Coeff +	31	0.7591	0.0177	15.77	7.84	0.14
5	rCSD	dSVD	NCP	0 2	Coeff +	31	0.7596	0.0177	15.00	7.02	0.19
6	eCSD	Tikhonov	NCP	0	Model - 0	32	0.7603	0.0174	18.86	15.03	0.29
7	kCSD	dSVD	NCP	[]	Coeff - []	33	0.7605	0.0183	16.94	7.84	0.32
8	eCSD	Tikhonov	NCP	[]	Coeff - []	33	0.7605	0.0183	16.97	8.05	0.32
9	rCSD	dSVD	NCP	0 1 2	Coeff +	31	0.7609	0.0176	17.66	8.67	0.37
10	rCSD	dSVD	NCP	[]	Coeff - []	34	0.7615	0.0182	19.63	7.57	0.44
11	rCSD	dSVD	NCP	0	Coeff - 0	32	0.7625	0.0181	21.40	15.95	0.58
12	qCSD	Tikhonov	NCP	0 2	qCSD +	30	0.7628	0.0174	19.17	6.66	0.62
13	qCSD	Tikhonov	NCP	0 1	qCSD +	31	0.7634	0.0168	20.57	9.91	0.71
14	rCSD	dSVD	NCP	1	Coeff +	28	0.7665	0.0181	26.40	12.00	1.10
15	eCSD	Tikhonov	NCP	0	Coeff - 0	27	0.7683	0.0171	28.34	7.82	1.35
16	kCSD	dSVD	NCP	0	Coeff - 0	30	0.7687	0.0170	31.26	9.07	1.39
17	qCSD	Tikhonov	NCP	0 1 2	qCSD +	25	0.7690	0.0169	29.97	8.97	1.44
18	iCSD	dSVD	NCP	2	Coeff +	23	0.7702	0.0171	29.63	30.12	1.60
19	kCSD	dSVD	NCP	0 2	Model +	22	0.7723	0.0171	31.97	28.97	1.88
20	kCSD	dSVD	NCP	2	Model +	22	0.7724	0.0171	32.97	29.06	1.88
21	eCSD	dSVD	NCP	0 2	Model +	23	0.7725	0.0166	32.54	22.23	1.90
22	kCSD	dSVD	NCP	0 1 2	Model +	22	0.7725	0.0171	33.89	28.68	1.90
23	kCSD	dSVD	NCP	0 2	Coeff +	25	0.7728	0.0173	36.31	11.43	1.94
24	eCSD	dSVD	NCP	0 1 2	Model +	21	0.7729	0.0169	32.57	27.19	1.95

(Table 3-4 continued)

Rank	CSD Method	Regularization Method	λ Selection Method	Prior	Prior Method	# Top Ranks	Mean Error	Standard Error	Mean Rank	Std Rank	% Diff.
25	eCSD	dSVD	NCP	2	Model +	23	0.7740	0.0169	33.57	23.54	2.10
26	kCSD	dSVD	NCP	0 1 2	Coeff +	24	0.7755	0.0177	41.34	7.41	2.30
27	eCSD	Tikhonov	NCP	0 1 2	Coeff +	26	0.7763	0.0170	44.46	10.86	2.40
28	eCSD	Tikhonov	NCP	0 2	Coeff +	27	0.7772	0.0174	45.51	13.81	2.52
29	eCSD	Tikhonov	NCP	0 1	Coeff +	25	0.7777	0.0171	46.29	10.98	2.58
30	kCSD	dSVD	NCP	0 1	Coeff +	23	0.7798	0.0171	48.06	16.31	2.86
31	kCSD	dSVD	NCP	1	Coeff +	24	0.7801	0.0179	47.06	11.72	2.90
32	eCSD	dSVD	NCP	0 1	Model +	20	0.7826	0.0162	48.97	32.81	3.24
33	rCSD	Tikhonov	NCP	[]	Coeff - []	23	0.7827	0.0176	55.40	16.43	3.25
34	eCSD	dSVD	NCP	1	Model +	19	0.7830	0.0162	48.97	32.97	3.28
35	kCSD	dSVD	NCP	0 1	Model +	20	0.7831	0.0161	49.40	32.25	3.30
36	kCSD	dSVD	NCP	1	Model +	20	0.7834	0.0161	50.71	31.31	3.34
37	rCSD	Tikhonov	NCP	0	Coeff - 0	20	0.7835	0.0167	58.94	24.65	3.35
38	rCSD	dSVD	NCP	2	Coeff +	19	0.7842	0.0216	54.74	30.42	3.44
39	iCSD	dSVD	NCP	1	Model +	20	0.7845	0.0169	51.94	24.89	3.48
40	kCSD	Tikhonov	NCP	[]	Coeff - []	22	0.7858	0.0174	60.03	14.94	3.65

Table 3-4: Top estimation schemes over all tested conditions for "Sum of Gaussians" profile. The top 40 estimation schemes in ranking over all tested conditions. This overall ranking is obtained by ordering the estimation schemes according to the mean over the mean of the sampling distribution of the mean errors of each of the 35 conditions.

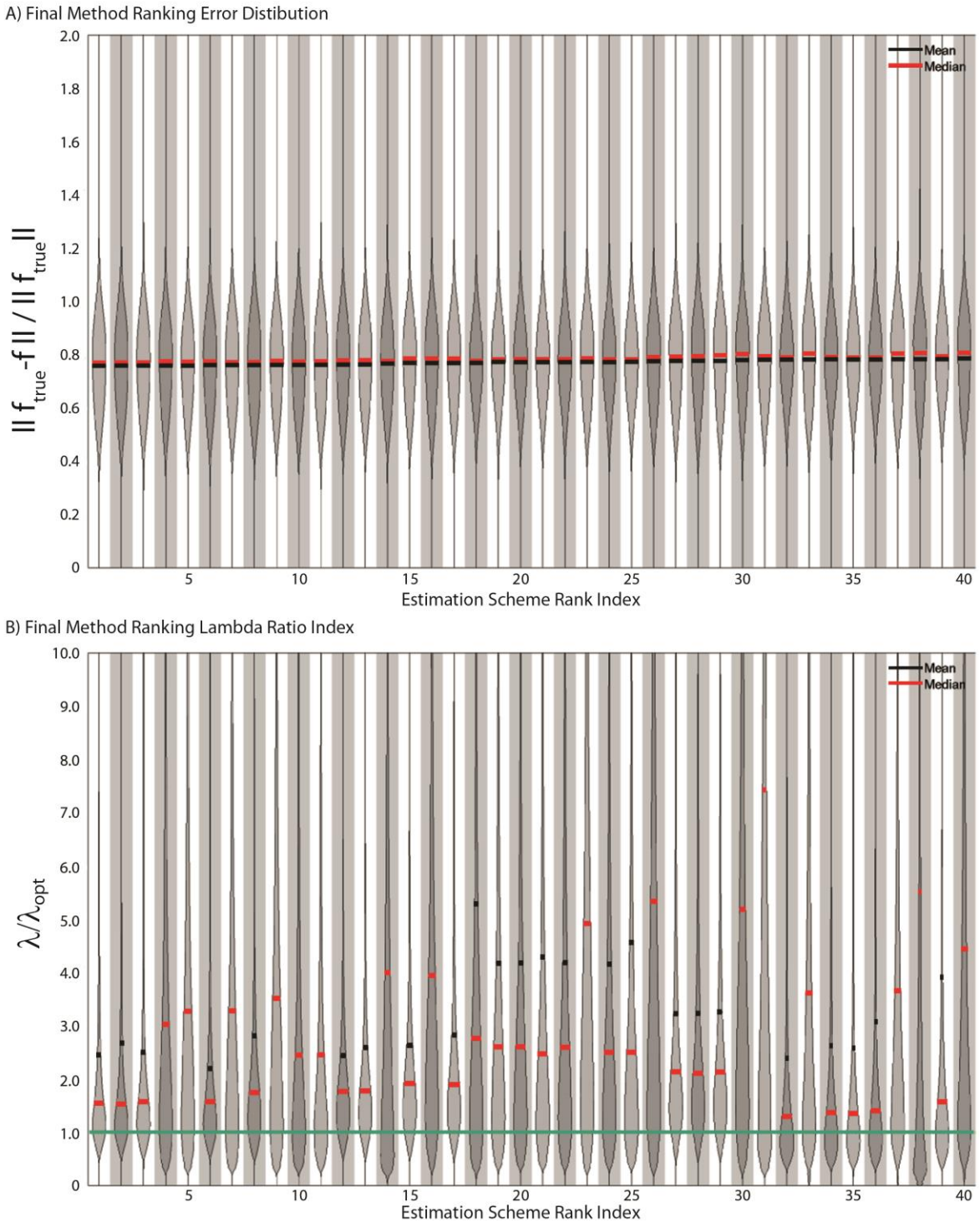

Figure 3-8: Distribution of error (A) and lambda ratio (B) for top 40 schemes in the final ranking. The violin plots are computed over the full dataset but without the worst 10% of trials in each condition. The green line in B marks a ratio of 1, which represents optimal regularization

optimal regularization coefficient λ_{opt} over the whole dataset. A ratio of 1, shown by the green dotted line, represents optimal regularization. In contrast to Fig. 3-8A, there is a much larger heterogeneity between estimation schemes. Although all estimation schemes tend to slightly oversmooth (ratio greater than 1), the oversmoothing tends to be more pronounced for estimation schemes employing dSVD as the regularization method. However, it is interesting to see that the increased spread of λ ratios does not affect the overall distribution of error or noise amplification.

To assess the robustness of the final ranking and make sure that the ranking is not simply fitting to the noise of this particular dataset, we have repeated the whole procedure for another set of 1000 trials. We found that almost exactly the same estimation schemes were found in the final ranking of both datasets (only 1 difference out of 160 schemes). Moreover, we then compared the similarity of the two final rankings using Spearman rank correlation. We found a correlation coefficient of 0.96 when the correlation was computed over the first 160 estimation schemes of the final ranking. Computing the rank correlation over only the top 40 schemes, the schemes displayed in Table 3-4, the correlation coefficient was 0.99. Both correlation coefficients are statistically significant at a significance level of 0.001. Moreover, the average absolute rank difference between the two datasets is 1.56 ± 1.86 when computed over the top 160 schemes, and 0.75 ± 1.13 when computed over the top 40 schemes. Thus, the final ranking was found to be highly reproducible for two realizations using the same "Sum of Gaussians" spatial CSD profile.

3.4.7 Effect of conductivity on the stability of the final ranking

The CSD forward problem used herein depends on 4 general parameters: 1) Lateral Source Profile (in our case the Uniform Cylinder); 2) Lateral Extent of Sources (i.e. the diameter of the cylinder); 3) the electrode positions, and 4) the medium conductivities. Assessing the effect on estimation accuracy of errors in these parameters is a very difficult task and it is beyond the scope of our current work to do this systematically. However, due to the pronounced effect of a discontinuity in conductivity on the amplitude of the potential close to the interface (as demonstrated by Pettersen et al. (Pettersen et al., 2006) and Ness et al. (Ness et al., 2015)), we decided to validate the final ranking assuming a mismatch in conductivity between the forward

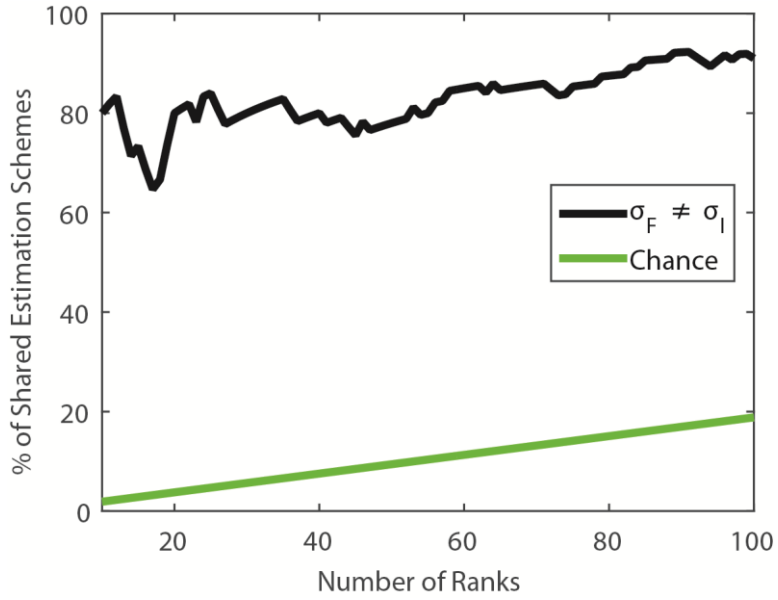

Figure 3-9: Validation of the ranking for a mismatch in conductivity between the forward and inverse model. The potential was simulated using a conductivity of $\sigma_t = 1.7\ S/m$ for the top medium in the forward model but the CSD was estimated with $\sigma_t = 1\ S/m$. The conductivity of the bottom medium was kept constant at $\sigma_e = 0.3\ S/m$ S/m. The black curve shows the percentage of estimation schemes which are found in both rankings within the top N ranks (computed as $f(N) = 100/N \cdot \{\#ranking_1[1{:}N]\} \cap \{ranking_2[1{:}N]\}$). The "chance" curve shows the expected percentage of matches when twice drawing N random samples from a population of 531 elements. It follows a hypergeometric distribution.

and inverse model. In particular, we assumed a conductivity of $\sigma_t = 1.7\ S/m$ for the top medium in the forward model but we estimated the CSD using $\sigma_t = 1\ S/m$. The conductivity of the bottom medium remained $\sigma_e = 0.3\ S/m$ as previously. We then repeated the previously presented analysis using the "Sum of Gaussians" profile and compared the final rankings. We found that the fidelity of the ranking was well preserved. In order to investigate the relationship between the two rankings, we display in Fig. 3-9 the percentage of estimation schemes which are found in both rankings within the top N ranks (computed as $f(N) = 100/N \cdot \#\{ranking_1[1{:}N]\} \cap \{ranking_2[1{:}N]\}$). The "chance" curve shows the expected percentage of matches when twice randomly drawing N samples from a population of 531 elements, which follows a hypergeometric distribution. On average, approximately 80% of the estimation schemes can be found within the same top ranks of both rankings, confirming the similarity between the two rankings.

3.4.8 Validation of final ranking with different CSD depth profiles

The robustness of the final ranking described in the previous section for the "Sum of Gaussians" depth profile does not guarantee that the ranking will turn out the same for different spatial CSD profiles. Discrepancies can be expected because, at least in the case of expansion methods, the estimation accuracy will always depend on the similarity between the spatial profile and the basis functions used in the expansion. This was already pointed out by Potworowski et al. in their discussion about the factors affecting the optimal choice for the width of the Gaussian expansion functions (Potworowski et al., 2012). Hence, it is important to assess how stable the ranking of estimation schemes is to a change in the CSD depth profile. We addressed this issue by repeating the previously presented analysis for 7 additional CSD depth profiles: 1) small monopole close to the surface (Fig. 3-10A); 2) small monopole further away from the surface (Fig. 3-10A); 3) large monopole (Fig. 3-10A); 4) large dipole (Fig. 3-10B); 5) small quadrupole (Fig. 3-10C); 6) large quadrupole (Fig. 10C), and 7) the CSD depth profile from Glabska et al. (Fig. 3-10B) (Glabska et al., 2014). The Glabska profile was added because it provides an interesting additional validation case since it differs from the Gaussian profile on three important points that could affect estimation performance: It is unbalanced, i.e. the sum over the profile is positive, indicating that sources are stronger than sinks. In addition, the amplitudes in Glabska's profile are larger than in the Gaussian profile. Lastly, this profile shows 2 dipole-like structures instead of a single one.

We repeated the same analysis as presented for the "Sum of Gaussians" profile with these new CSD profiles and then performed a pairwise comparison of the overall rankings. When computed over all 531 schemes, the Spearman rank correlation is highly significant ($p << 0.001$) for all pairwise comparisons (smallest rank correlation coefficient $\rho = 0.73$) and remains highly significant under Bonferroni correction for multiple comparisons. However, this is rather unsurprising since we would not expect the rankings to be completely random. Hence, such a null hypothesis is too liberal to serve as a good measure of the usefulness of our rankings. To provide a more focused comparison on the similarities in the top ranks, we show in Figure 3-10D-K the percentage of estimation schemes which are found within the top N ranks in a pairwise manner for all CSD profiles. It is calculated in the same way used for Fig. 3-9 with "chance" again referring to the expected percentage of matches if the rankings were random

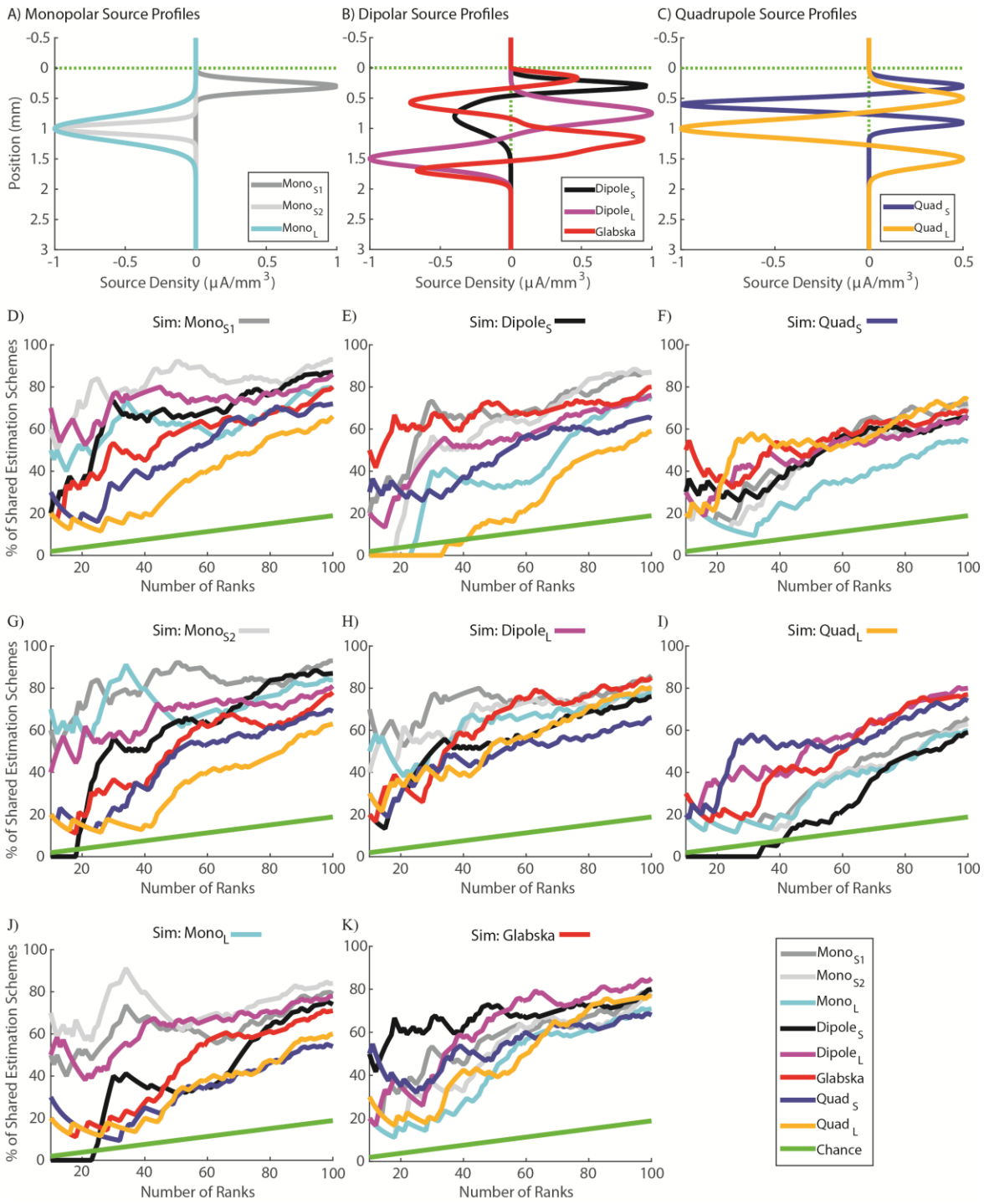

Figure 3-10: Pairwise comparison of the similarity in rankings between different CSD depth profiles. Panels (A)-(C) show the eight tested CSD depth profiles (the amplitude of the Glabska profile was divided by 10 for visualization purposes). Panels (D)-(K) show the percentage of estimation schemes which are found within the top N ranks in a pairwise manner for all CSD profiles. It is calculated in the same way as in Fig. 3-9 with "chance" again referring to the expected percentage of matches if the rankings were random permutations of 531 elements.

permutations of 531 elements (i.e. expected number of identical elements within two independent draws of N elements from a population of 531 elements). We found that there are only 3 situations where the percentage of matches lies below the chance level, namely between the small dipole and 1) the small monopolar source further from the surface, 2) the large monopole and 3) the large quadrupole. Hence, it appears that the presence of a source close to the conductivity jump affects the estimation accuracy of the different estimation schemes in a non-uniform manner. When looking at these discrepancies in more detail, we find that it originates mainly from the higher prevalence of the l-curve among these rankings. For example, the large quadrupole ranking has 37 l-curve estimation schemes within the top 40 ranks while the small dipole has none. When removing the differentiation according to the regularization parameter selection methods, only the comparison between the small dipole and the large quadrupole fails to show a number of matches exceeding the chance level within the top 20 ranks (results not shown). It is however not obvious how this effect comes about. A qualitative analysis of the final rankings suggests that more extended sources with lower spatial frequencies are better captured with the l-curve method. However, this effect is not clearly differentiable from other factors such as the overall extension of the sources or the presence of a source near the discontinuity in conductivity. Moreover, it is probable that these differences affect the estimation differently for small or large diameter sizes or for different SNRs. Hence, although clearly of interest, such an in-depth analysis will have to be deferred to future work. Nevertheless, despite the discrepancies discussed above, Figure 3-10D-K shows clearly that the final ranking for the tested profiles are not independent of each other but rather share common features.

It is important to note however that pairwise similarities don't guarantee that there are estimation schemes that perform well across all tested spatial profiles. To test whether it is possible to come up with a suggestion for optimal estimation schemes, we ranked the estimation schemes according to their average ranking in each of the eight spatial profiles. The results for the top 20 estimation schemes are shown in Table 3-5. The general features already discussed in the final ranking of the "Sum of Gaussians" profile remain valid for this global ranking. First, expansion methods outperform the quadrature methods. Second, there is no preference between the Tikhonov or dSVD regularization method.

Rank	CSD Method	Regu Method	λ Sel. Method	Prior	Prior Method	Mean Rank	Mean Error	% Diff.	Mono S1	Mono S2	Mono L	Dipole S	Dipole L	Glabska	Quad S	Quad L
1	iCSD	Tikhonov	L-curve	0 2	Coeff +	7.4	0.6809	0.00	1	1	2	42	2	8	1	2
2	iCSD	Tikhonov	L-curve	0 1 2	Coeff +	8.4	0.6828	0.28	3	6	1	46	1	6	3	1
3	kCSD	dSVD	NCP	0	Coeff - 0	24.0	0.7066	3.78	15	19	33	16	10	14	28	57
4	eCSD	Tikhonov	L-curve	0 1 2	Coeff +	25.4	0.7030	3.25	10	13	5	75	7	57	9	27
5	eCSD	Tikhonov	L-curve	0 1	Coeff +	27.3	0.7061	3.70	6	7	4	68	6	52	47	28
6	eCSD	Tikhonov	NCP	0	Coeff - 0	28.6	0.7097	4.23	14	22	29	15	28	23	37	61
7	iCSD	Tikhonov	NCP	[]	Coeff - []	29.0	0.7063	3.73	20	37	88	3	30	7	7	40
8	eCSD	Tikhonov	NCP	[]	Coeff - []	29.1	0.7062	3.72	28	44	86	8	18	4	10	35
9	kCSD	dSVD	NCP	[]	Coeff - []	29.1	0.7062	3.72	29	43	87	7	17	5	11	34
10	rCSD	dSVD	NCP	0 2	Coeff +	31.3	0.7101	4.29	7	41	69	5	4	13	43	68
11	kCSD	dSVD	NCP	0 2	Coeff +	31.5	0.7109	4.42	22	20	30	23	5	44	50	58
12	eCSD	dSVD	NCP	0 2	Model +	31.9	0.7125	4.64	8	3	26	21	29	50	35	83
13	eCSD	dSVD	NCP	2	Model +	32.3	0.7125	4.64	5	9	34	25	27	28	54	76
14	kCSD	dSVD	NCP	0 1	Coeff +	33.6	0.7120	4.58	34	17	18	30	32	19	57	62
15	kCSD	Tikhonov	NCP	0	Model - 0	34.0	0.7087	4.08	42	35	85	2	44	3	17	44
16	rCSD	Tikhonov	NCP	0	Model - 0	34.8	0.7103	4.32	43	25	77	1	48	21	15	48
17	eCSD	Tikhonov	NCP	0	Model - 0	34.8	0.7121	4.59	13	23	66	6	71	17	29	53
18	rCSD	dSVD	NCP	0 1 2	Coeff +	34.8	0.7122	4.61	9	47	70	9	8	15	48	72
19	eCSD	dSVD	NCP	0 1 2	Model +	35.8	0.7137	4.82	4	15	28	24	31	41	65	78
20	rCSD	dSVD	NCP	[]	Coeff - []	35.9	0.7102	4.31	27	50	105	10	34	10	8	43

Table 3-5: The top 20 estimation schemes in ranking over all CSD depth profiles. This global ranking is obtained by ordering the estimation schemes according to their average rank over the eight tested CSD depth profiles. The ranks for each of the depth profiles were calculated in the same way as in Table 3-4. They are presented in the eight columns to the right (Dipole S = "Sum of Gaussians" profile).

Finally, NCP is generally preferred over the l-curve for the selection of the regularization parameter and GCV is completely absent from the list. However, in contrast to the "Sum of Gaussians" ranking, four of the top 5 estimation schemes in the global ranking use the l-curve. Although it is clearly not straightforward to predict the accuracy of a particular estimation scheme on a new spatial CSD profile, the consistency of the rankings across the tested subset of profiles suggests that the list in Table 3-5 represents the candidate schemes most likely to achieve good estimation accuracy on a new spatial CSD profile.

3.5 Discussion

In this paper, we have introduced a unified framework for zero and higher-order regularization of 1-D CSD estimation problems. In particular, we have focused on showing how the current (iCSD and kCSD) and the newly presented methods (rCSD, eCSD (variation of kCSD) and qCSD) can be understood as special cases of two approaches for the discretization of linear inverse problems: quadrature and basis function expansion. Moreover, the presented framework greatly facilitates dealing with recording noise, a ubiquitous problem in solving inverse problems, by incorporating zeroth- and higher-order regularization methods with multiple regularization parameter selection methods. To show the utility of the presented framework, we have compared the performance of the different estimation schemes under various noise conditions and source diameters. We show that no single estimation scheme outperforms all others under all tested conditions. Nevertheless, we found that ranking schemes according to the average error over all tested conditions results in a reproducible and meaningful ranking where the top schemes are found to perform well in the majority of tested conditions. We have further validated this ranking procedure by repeating it over a set of spatial CSD profiles. The general features of the ranking are preserved across this comparison as well. However, a qualitative analysis of the distribution of the ranks across the eight spatial profiles reveals a structure that suggests the presence of unidentified features that influence estimation accuracy. The identification of source characteristics (e.g. spatial frequency, extent, and/or position of source) which could be better resolved with certain estimation schemes goes however beyond the scope of our current work. In fact, it is easily conceivable that these features could affect the estimation differently for small or large diameter sizes (or different SNRs) which would require an additional level of detail in the analysis.

Our study also presents a set of additional, smaller contributions: it shows the detrimental effect of noise on the spatial resolution of the solution; it facilitates the use of electrodes from 2 media with different conductivities in the estimation process; and it provides a thorough introduction into the difficulties of solving inverse problems in the hope of making the problem more easily accessible to experimentalists. Finally, our work makes available a ready-to-use open-source MATLAB toolbox containing all the CSD estimation methods discussed, including existing methods and the 2 novel methods we introduce, along with all the regularization methods that we have tested.

An unexpected finding of the present work is the overall performance of the NCP regularization parameter selection method. It is interesting to ponder whether this dominance is likely to translate to the experimental setting. Unfortunately, as discussed by Hansen, there is no simple method to predict which parameter selection method is optimal for a given problem because each inverse problem has its own characteristics and error model (Hansen and Society for Industrial and Applied Mathematics., 2010). By construction, NCP will perform best under white noise. This might explain its success here but could affect its performance in an experimental setting where noise characteristics might be non-white. Although the derivation of GCV also relies on noise being white, it was almost totally absent from the top ranking which is especially surprising given its successful application in Potworowski et al. (Potworowski et al., 2012). A potential explanation for this discrepancy can be found in the higher level of outliers with GCV. Occasional failures of GCV are a well understood phenomenon and tend to occur whenever the minimum is located in the flat part of the GCV function (Hansen and Society for Industrial and Applied Mathematics., 2010). However, in the simulated situation used here the failure rate appears to be further increased by the occasional existence of two local minima (see Fig. 3-5C for example). We have observed a similar situation with the L-curve in certain trials where the L-curve shows two locations of positive curvature. These situations can technically be avoided by setting bounds on the regularization parameter. Although this is quite easy to do in practice, we have not pursued this avenue because the bounds are problem-dependent, thus making this approach impractical in a large scale comparison like here. It has also been suggested that GCV can potentially be improved further by adding a weighting factor to the GCV function (Hansen and Society for Industrial and Applied Mathematics., 2010). Finally,

since the L-curve is the only regularization parameter selection method that is not based on noise being white, it has been advocated for applications where no prior information about noise is available (Vogel, 2002).

A similar problem-dependent optimization of parameters could also be pursued for eCSD, kCSD and qCSD. In both eCSD and kCSD the width and number of basis could be adapted to the problem at hand. Here we chose the basis width (three sigma point) to be 1.5 times the interelectrode distance following the guidelines from Potworowski et al. where a factor of 1-2 was suggested (Potworowski et al., 2012). It is interesting to note that eCSD and kCSD were found to perform equally well suggesting that the kernel trick does not provide any advantages in the 1-D case. However, this should not be interpreted as being generally valid. Especially in higher dimensions, it is quite probable that it provides additional stability to the estimation. The only reason eCSD was included here is to assess the similarity between the expansion and kernel method in the 1-D setting. It remains to be shown whether this similarity also applies in an experimental setting.

A central question arising in any computational study based on simulations is the applicability of the results in an experimental setting. As we have hinted at on various occasions, CSD estimation contains a large amount of built-in assumptions which have the potential to dramatically affect source reconstruction accuracy. Therefore, in order for CSD estimation to become a more streamline technology to be applied to the analysis of intra-cranial electrophysiology data, the theoretical assumptions incorporated into the estimation need to be carefully addressed. In the following we will attempt to critically discuss the major factors that could affect the applicability of the proposed methods to the experimental setting.

3.5.1 Accuracy of the volume conductor model

The volume conductor model employed here relies on the tissue being purely resistive and thus justifying the use of Ohm's law for explaining the relation between the extracellular electric field and neuronal currents. The literature contains both experimental results validating this assumption (Logothetis et al., 2007) and others reporting significant non-ohmic effects (Gabriel et al., 1996a, Gabriel et al., 1996b, c). The interpretation of these contradictory findings is further complicated by the fact that a newer study from Gabriel et al. critically analyses the

usefulness and limitation of their previous measurements. In particular, they point out that their previous work is based on excised tissue instead of in vivo measurements. In addition, they mention that the frequency region below 1 MHz is particularly error prone and hence the results only provide best estimates of the conductivity values (Gabriel et al., 2009). On the other hand, recent experimental findings from Riera et al. find potential problems with the resistive assumption based on an unbalanced ratio of current sinks and sources during evoked neuronal activity (Riera et al., 2012). Similar findings were also reported by a study of the brain tissue's effect on the propagation of neurostimulation fields (Wagner et al., 2014). The authors conclude that "living tissue carries currents through both dipole and ionic mechanisms, in a frequency dependent manner". Although Wagner et al.'s results found ohmic mechanisms to be dominant, permittivities were deemed to be of sufficient magnitude to support significant displacement currents. Finally, based on a theoretical study, Bedard and Destexhe proposed that ionic diffusion could play a significant role in the re-equilibration of extracellular potential following neural activation (Bedard and Destexhe, 2009). This is corroborated by recent experimental evidence which found non-ohmic properties of the extracellular medium around neurons which could possibly be related to ionic diffusion (Gomes et al., 2016). In conclusion, the electric properties of extracellular tissue (e.g. ohmic, capacitive, polarizable or diffusive) are currently under heavy debate and will likely remain debated in the near future. However, it is clear that a modification to the resistive tissue assumption would require a reevaluation of all source estimation methods. In particular, it would entail that the spread of potential would depend on both the spatial and temporal components of neuronal currents instead of only the spatial component as assumed in a purely resistive medium.

A second assumption of the current volume conductor model is that electrical conductivity is homogeneous and isotropic within each medium. This assumption has also been challenged in Goto et al. who found that there is likely a noticeable difference in conductivity between the different cortical layers (Goto et al., 2010). However, in contrast to the non-applicability of the resistive assumption, a non-homogenous conductivity as proposed by Goto et al. could be relatively easily incorporated into the current framework because it only involves changing the Equations describing the induced potential (Eq. 1). The inaccuracies arising from substituting a layered conductivity by a homogeneous conductivity profile are hard to assess

since no data is available on the conductivity in each layer. However, even with current knowledge, it is clear that the assumption of 2 semi-infinite media represents a significant simplification of the actual brain geometry. In fact, simply the presence of the white matter bordering the cortex clearly forms another discontinuity in conductivity analogous to the discontinuity between cortex and the medium above it. To address a similar issue, Ness et al. implemented the method of images in a 3-layer medium (Ness et al., 2015). However, in contrast to our situation, their problem only required to model sources and potentials within the middle layer. Hence, we have refrained from adopting this approach in this work because the method would have to be extended in order to allow the use of electrodes from all three media. To our knowledge, such an extension of the infinite series of image charges has not been presented so far. Alternatively, several papers have discussed an extension to multi-layered media based on the Fourier transform (e.g. (Barrera et al., 1978)). However, the lack of closed form solutions makes these approaches unsuitable for our purposes since both the forward and inverse problem calculation would require a significantly longer computation time. On the other hand, it is unquestionable that even the approximation of the medium by planar surfaces represents a significant simplification of the brain geometry. Addressing this would require employing more elaborate models for medium parametrization such as models based on finite element modeling (e.g. (Ness et al., 2015) or a spherical volume conductor as proposed in Goto et al. (Goto et al., 2010). Although such issues are certainly important for CSD estimation, the comparison of volume conductor models goes beyond the scope of our current work. Hence, we have used the most commonly employed medium parametrization.

3.5.2 Accuracy of forward model assumptions

To our knowledge, the applicability of the cylindrical source assumption has never been experimentally validated. This is most likely due to the fact that such a validation is quite difficult because it requires a way to manipulate the source diameter while at the same time having a good knowledge about the depth variation of current sources. Nevertheless, the choice of source diameter is a major challenge in the application of the current CSD methods and can potentially affect different estimation schemes in different ways. Although we have evaluated the accuracy of estimation under various levels of noise, we have not attempted to systematically assess the robustness of the presented framework to errors in the forward model parameters.

However, because of the pronounced effect of a discontinuity in conductivity on the amplitude of the potential close to the interface, we have performed a single point validation of our ranking approach assuming a mismatch in conductivity between the forward and inverse model. In particular, although we assumed a conductivity of $\sigma_t = 1.7\ S/m$ for the top medium in the forward model, we estimated the CSD using $\sigma_t = 1\ S/m$. The conductivity of the bottom medium remained $\sigma_e = 0.3\ S/m$. A comparison of the final ranking for the "Sum of Gaussians" profile showed that the fidelity of the ranking was well preserved between the matched and unmatched conductivity values. Although encouraging, this should not be understood as ruling out an important effect of conductivity but rather gives us some confidence that the sensitivity to its choice is manageable. Moreover, the validation presented here relies on a single set of values whereas errors in the forward model can take various forms (e.g. error in electrode placement, error in conductivity, inhomogeneity of conductivity, error in source diameter, non-cylindrical source distribution, and/or combination thereof). Therefore, our current analysis does not alleviate the need to perform a more thorough analysis of the errors introduced by mismatches between the forward and inverse model. To our knowledge, a partial assessment of such errors has only been performed on noiseless data so far (Pettersen et al., 2006) but would merit to be further investigated in order to facilitate the application of these methods in an experimental setting.

Preface to Chapter 4

The 1-D CSD estimation framework presented in this chapter provides a successful unification of all previously published inverse methods under a common formulation. But its connection with the standard estimation method used in Chapter 2 remains elusive. In the next chapter, we develop this question by comparing the differential (standard) and integral (inverse) formulation of CSD estimation in the context of laterally extended sources. In particular, we transform the standard method to its integral formulation by using the Green's function for the 1-D Poisson equation. By comparing the two inverse formulations we then show that the standard and inverse estimation methods converge to each other and could be considered similar for lateral source diameters exceeding ~2-5 mm.

4 Chapter 4: Comparison of 1-D current-source density estimation methods for large diameter sources in a three-layer volume conductor

The standard current-source density (CSD) estimation method has a long history in neurophysiology and remains heavily used even nowadays. However, in an experimental setting, discrete estimation of derivatives is afflicted with certain mathematical limitations. In particular, derivative estimation is an unstable operation which heavily amplifies noise. Moreover, the standard CSD estimator is insensitive to jumps in electrical conductivity such as those encountered at transition between cerebro-spinal fluid and gray matter, or between gray and white matter. In addition, 1-D standard CSD estimation is based on an assumption of large lateral source extension, a condition which is not always clearly defined. In order to address some of these caveats, estimation methods based on inverse theory have been introduced. However, these methods have mainly aimed at dealing with restricted source diameters and, up to now, the relationship between inverse estimators and the standard method have not been thoroughly studied. In this work, we have investigated the relationship between 1-D standard and inverse CSD estimation in the presence of laterally extended sources. We have approached this by first showing that the 1-D Poisson equation can be transformed into an integral equation akin to the common inverse formulation. We have then used this formulation to compare the standard and inverse estimators and found that they converge to each other in the presence of lateral source cylinders of uniform density with a diameter exceeding ~2-5 mm. Moreover, errors in the assumed source diameter affects estimation mainly at the edge electrodes, which can be explained by errors in the enforced boundary conditions. In addition, the 5-point standard CSD estimator is more sensitive to noise than the inverse methods. In a second step, we have also investigated the effects on CSD estimation when simplifying a cortical medium parameterization from a three- to a two-layered medium. In particular, we have focused on the impact of ignoring the transition from gray to white matter. We found that the introduced estimation errors were small and mainly limited to the boundary. We thus conclude that (1) there is no particular benefit in using the standard estimation method over the various inverse estimators, and that (2) whereas a three- layered medium describes the conductivity better than a two-layered medium, the difference between these models in 1-D CSD estimation is small for typical cortical conductivities.

4.1 Introduction

Disentangling the interactions between neural populations located in different cortical laminae is an important step for understanding the principles governing neural computation. However, because of their close spatial proximity, it is very difficult with extracellular electrical measurements to localize the laminar origin of neural activity. This difficulty is explained by the fact that, due to volume conduction, extracellular potentials are a non-local measure of the activity generating it. In other words, variations in electrical potential caused by some given neural activity can theoretically be measured throughout the full volume conductor, although in practice, this is obviously limited by the strength of the source and the signal-to-noise ratio of the recording. But, in the case of spatially close recordings, the non-local nature of electrical potential recordings is sufficient to make direct localization quasi impossible.

To circumvent the limitations imposed by volume conduction, it has become common practice in extracellular neurophysiology to estimate and interpret current-source density (CSD), instead of directly interpreting the electrical recordings. This is advantageous because, in addition to mitigating the effect of volume conduction, correctly estimated CSD represents the volume density of net transmembrane currents that generated the measured potential. Although CSD is based on the same approach as source localization in EEG, the term CSD analysis has been principally used for source localization in the context of invasive electrical recording with microelectrodes. In this context, CSD analysis has been extensively applied to the study of the low-frequency part of the electrical potential, called the local field potential (LFP). This focus on the LFP makes sense for two major reasons: First, it is believed that synaptic trans-membrane currents are the major generators of the LFP (Buzsaki et al., 2012). Hence, their analysis and localization provides interesting complementary information to the information provided by the analysis of action potentials. Second, in contrast to action potentials, which for SNR reasons can only be recorded from a few neurons close to the electrode, the LFP gives access to the summed activity of multiple neuronal populations. It is thus mostly reflective of population activity in the mesoscale range (Brette and Destexhe, 2012). Hence, CSD estimation represents a very interesting tool to study the interactions between different neuronal populations.

In a homogeneous, isotropic and purely resistive medium with electrical conductivity σ, current-source density C is related to the extracellular potential φ via the Poisson equation: $C =$

$-\sigma\Delta\varphi$ where Δ is the Laplace operator (Nicholson and Freeman, 1975, Tenke et al., 1993, Brette and Destexhe, 2012). This relationship is obtained by combining the a priori quasi-static approximation of Faraday's law (leading to $\vec{E} = -\vec{\nabla}\varphi$) with Ohm's law ($\vec{J} = \sigma\vec{E}$) and taking the divergence of the result (Plonsey and Barr, 2007). Given the mentioned assumptions, there are two main strategies for estimating the CSD C from the measured potential φ: numerically approximating the Laplace operator by discrete differences or solving the electrostatic linear inverse problem. With both strategies, the estimation depends on the spatial arrangement of the electrodes used in recording the potential. Here, we will focus on CSD estimation from 1-D linear electrodes such as those used in cortical depth recordings.

Direct approximation of the Laplace operator was first introduced by Pitts who used it to study synaptic transmission in the spinal cord (Pitts, 1952). Since then, the method has been employed in a variety of brain structures and in many different animal models (see e.g. (Freeman and Nicholson, 1975, Newman, 1980, Mitzdorf, 1985, Di et al., 1990, Schroeder et al., 1998, Swadlow et al., 2002, Jin et al., 2008, Szymanski et al., 2009, Sotero et al., 2010, Buzsaki et al., 2012)). In particular, it has been extensively used with 1-D multi-channel linear electrodes and on successive recordings from linear tracks. Due to its popularity, the direct approximation of the Laplacian is now commonly called the standard CSD method.

Although the standard CSD method has provided valuable information about brain function, it is often complicated by various experimental and mathematical limitations. In recordings with a multi-channel linear electrode (e.g. an electrode inserted perpendicular to cortical layers), no data is available to estimate the lateral (x and y) derivatives in the Laplacian. Hence, the lateral components of the Laplacian are simply ignored, which is equivalent to assuming that sources are infinitely extended sheets (Mitzdorf, 1985). Moreover, derivative estimation is an unstable operation, which means that it heavily amplifies recording noise. Also, discrete approximation of the derivative cannot be done at edge electrodes without introducing boundary conditions. Finally, standard numerical estimation of the second derivative is insensitive to conductivity jumps within the medium. However, such jumps occur at transitions between different structures such as the border between cerebro-spinal fluid (CSF) and cortex, the border between gray matter and white matter, and to a lesser extent between adjacent cortical layers (Goto et al., 2010).

To address some of these shortcomings, methods to directly solve the electrostatic inverse problem have been developed (Pettersen et al., 2006, Leski et al., 2007, Leski et al., 2011, Potworowski et al., 2012). By assuming particular parametric forms for the lateral source distributions, e.g. a uniform cylinder or a Gaussian distribution, these methods provide accurate estimation with 1-D linear electrodes even in the presence of spatially localized sources as well as naturally providing source amplitude estimates at the edge electrodes. Recently, these methods have also been augmented with a range of regularization methods to stabilize estimation on data corrupted by noise (Kropf and Shmuel, 2016). Moreover, the inverse approach has the advantage of facilitating the incorporation of more complex medium assumptions. In practice, the most common way to do this has been to assume that the volume conductor is made of two semi-infinite media with different but constant conductivity. The forward model can then easily be modified accordingly by using the method of images (Jackson, 1999). It is clear that this is still a major simplification of the actual brain geometry but, at least in cortical recordings, this method is well suited to account for the large conductivity jump from CSF (or saline) to gray matter. This approach does however ignore the presence of a conductivity jump between gray and white matter. It is possible in theory to extend the method of images to account for a multi-layered medium but this has not yet been extensively used (Barrera et al., 1978, Gold et al., 2006, Ness et al., 2015). In addition, to our knowledge, no previous study has investigated the type of errors introduced when using a two-layered medium instead of a three-layered one.

Although the inverse methods were introduced a decade ago (Pettersen et al., 2006), the standard method remains popular (see e.g. (Jin et al., 2008, Kajikawa and Schroeder, 2015)). A possible reason for this is the apparent ease of using the standard CSD method. In contrast to standard CSD, the use of the inverse CSD methods is complicated by the need to specify parameters that are often unknown. In particular, the lateral source distribution and especially its spatial extent is difficult to choose. Although it has been well established that the standard method does not provide adequate estimation for laterally constrained sources (Pettersen et al., 2006), it is not yet fully understood under what conditions it remains applicable. This is especially difficult when dealing with extended but non-uniform activations, which might be encountered for example with cortical recordings from primary visual cortex in response to a

large oriented stimulus. Hence, it is unclear whether standard CSD or inverse CSD is the optimal choice for CSD estimation in the presence of laterally extended sources.

In the present work, we show that the standard CSD method, routinely applied by taking the second derivative of the measured potential along the cortical depth, can also be formalized and written as an inverse problem. We then use this description to compare the estimation accuracy between the standard CSD method and classical inverse methods for large diameter sources. Finally, we investigate the errors introduced by approximating a 3-layered volume conductor with a 2-layered medium. In particular, we investigate the effect of neglecting the conductivity jump between gray and white matter in cortical CSD estimation.

4.2 Theory

As mentioned in the introduction, 1-D standard CSD estimation suffers from 4 major experimental and mathematical limitations:

1. In recordings with linear multi-channel electrodes, no data is available to estimate the lateral derivatives in the Laplacian.
2. Derivative estimation is unstable, which means that recording noise gets heavily amplified.
3. Discrete approximation of the derivative cannot be done at edge electrodes without introducing boundary conditions.
4. Standard numerical estimation of the second derivative is insensitive to conductivity jumps within the medium.

To most of these difficulties, workarounds can be found in the literature. We will shortly discuss the most common ones here. In the following section, we will then propose a way to transform the standard CSD method into an inverse problem which will allow us to study the relationship between the direct and inverse CSD estimation methods.

4.2.1 Common Approaches to Standard CSD Estimation

The lack of data available with linear electrodes is commonly addressed by simply ignoring the lateral, i.e. x and y, derivatives in the Laplacian (e.g. (Mitzdorf, 1985, Schroeder et al., 1998)). In other words, the Poisson equation is simplified from

$$C = -\sigma\Delta\varphi = -\sigma\left(\frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta y^2} + \frac{\delta^2}{\delta z^2}\right)\varphi \quad (1)$$

to

$$C = -\sigma \frac{\delta^2}{\delta z^2} \varphi \tag{2}$$

This assumption is approximately valid when sources are laterally extended so that potentials have minimal curvature in the lateral direction. For this approximation to be exact, the sources would need to be infinitely extended in the lateral direction. However, it has been shown that the approximation introduces minimal error when sources are distributed within a uniform cylinder with a diameter larger than ~5 mm (Pettersen et al., 2006).

The simplest way to numerically estimate the second derivative in Eq. 2 is to approximate it with a centered difference. In this way, $\frac{\delta^2}{\delta z^2}\varphi(z)$ is estimated by replacing it with

$$\frac{\delta^2}{\delta z^2}\varphi(z) \cong (\varphi(z-h) - 2 \cdot \varphi(z) + \varphi(z+h))/h^2$$

where h is the spacing between recording electrodes. However, derivative estimation is unstable and leads to a large amplification of noise. Hence, to minimize its noise sensitivity, it has become commonplace to combine the estimation with a smoothening filter. Multiple options have been proposed by Freeman and Nicholson but the most commonly employed one is the five-point estimator (Freeman and Nicholson, 1975, Rappelsberger et al., 1981)

$$\frac{\delta^2}{\delta z^2}\varphi(z) \cong (\varphi(z-2h) - 2 \cdot \varphi(z) + \varphi(z+2h))/4h^2$$

This 5-point estimation kernel ($[1\ 0\ -2\ 0\ 1]/4$) is obtained by combining the basic three-point estimator $[1\ -2\ 1]$ with a smoothening filter $[1\ 2\ 1]/4$. Using this kernel, the standard CSD estimator can then be written as a $[N-4\ x\ N]$ matrix in which each row contains an appropriately shifted version of the 5-point kernel.

Although the five-point estimator increases estimation accuracy under noisy conditions, it comes at the cost of reducing the effective spatial resolution. In this sense, it is akin to spectral regularization which also comes at the cost of reduced resolution (Kropf and Shmuel, 2016). However, the five-point kernel also amplifies the problem of estimation at edge electrodes because discrete second derivative estimation cannot be obtained at 2 electrodes from the edge

instead of 1 with the three-point estimator. In other words, with a 32-channel electrode, one can only estimate CSD at the 28 inner electrode positions without introducing assumptions on boundary conditions. To alleviate this issue, Vaknin et al. proposed to introduce virtual electrodes at the edge by simply replicating the potential at the edge electrode (Vaknin et al., 1988). This implicitly assumes that the sources are zero outside of the range covered by the electrode. In the terms used in the field of differential equations, this is equivalent to enforcing Neumann boundary conditions of 0 (Hansen and Society for Industrial and Applied Mathematics., 2010). This also has the side effect of reducing the null-space of the estimator to $span\{(1,1,\dots,1)\}$ instead of $span\{(1,1,\dots,1),(1,2,\dots,N)\}$ for the estimator without boundary conditions. In other words, linear trends in the recorded potential will affect the estimation at the edge electrodes with the method of Vaknin but won't affect the inner electrode estimation. It is clear that, in an experimental setting, Vaknin's approximation is only valid if the potential varies minimally at the edge electrodes. It is therefore natural to ask under what conditions this assumption is applicable. Pettersen et al. have proposed that whenever the sources are balanced, i.e. they sum to zero, Vaknin's procedure is asymptotically correct as sources become infinitely extended in the lateral direction (Pettersen et al., 2006). In practice, it suffices that sources are balanced and the radius of the sources is significantly larger than the extent of the probe, i.e. the distance from the first to the last electrode. In Section 3.3 we return to this point and explore the effect of boundary conditions in more detail.

The last major difficulty encountered when employing the standard numerical estimation of the second derivative in an experimental setting is that the estimation is insensitive to conductivity jumps within the medium. However, jumps in conductivity are common in neural tissue such as at transitions from cerebro-spinal fluid (CSD) to gray matter or from gray to white matter. The literature comprises two approaches to solve this issue: In the inverse CSD formulation, it is quite simple to introduce one discontinuity in conductivity by modifying the forward model using the method of images (Jackson, 1999, Pettersen et al., 2006, Kropf and Shmuel, 2016). Moreover, extending the method of images to an infinite series of images sources also allows incorporating additional layers (Barrera et al., 1978, Gold et al., 2006, Ness et al., 2015). We will explore this approach in more detail in Section 2.3. Alternatively, it is also possible to take into account conductivity jumps in the direct numerical approximation of the 1-

D Poisson equation by correcting the estimator for the two electrodes on each side of the boundary (Liu et al., 2000). In particular, if σ_t and σ_e are the conductivities of the top and bottom medium respectively, the three-point estimator of the CSD for the electrodes closest to the boundary are modified as follows (see Fig. 4-1A for a schematic of the described situation): The estimator for C_i is corrected from

$$C_i = \sigma_t[V_{i+1} - 2V_i + V_{i-1}]/h^2$$

to

$$C_i = [\sigma_x(V_{i+1} - V_i) - \sigma_t(V_i - V_{i-1})]/h^2$$

with

$$\sigma_x = \sigma_t\sigma_e h/(\sigma_t\theta + \sigma_e(h - \theta))$$

where θ is the distance from the i-th electrode to the boundary, and h is the inter-electrode separation. Similarly, the three-point estimator for C_{i+1} is corrected from

$$C_i = \sigma_e[V_{i+2} - 2V_{i+1} + V_i]/h^2$$

to

$$C_i = [\sigma_e(V_{i+2} - V_{i+1}) - \sigma_x(V_{i+1} - V_i)]/h^2$$

To our knowledge, such a boundary correction has never been discussed in the context of standard CSD estimation.

4.2.2 Standard CSD as an integral equation

In the previous section, we have discussed two approaches for estimating CSD with extended sources. The first is the standard method which consists in a dimensionality reduction by ignoring the lateral components in the Laplacian (i.e. approximate Eq. 1 with Eq. 2), followed by the numerical approximation of the second derivative (see Eq. 2) using a three- or five-point kernel. The second is the inverse CSD approach with sources of large diameters (e.g. large cylindrical or Gaussian source). Here, we will propose a third approach which consists in solving Eq. 2 using Green's function for the 1-D Poisson equation. This is equivalent to transforming the standard differential CSD equation (Eq. 2) into an integral equation.

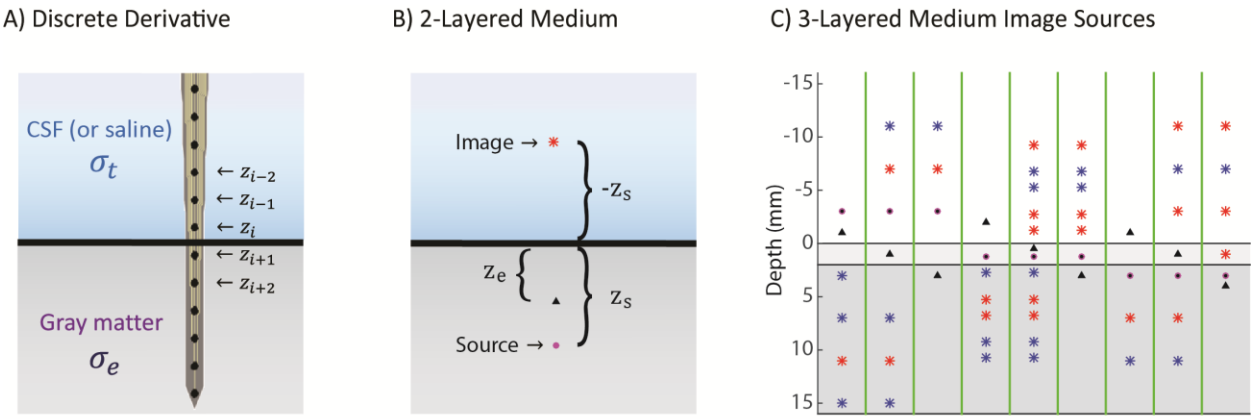

Figure 4-1: Graphic demonstration of the boundary correction method at conductivity jumps in recording medium. A) Electrode position labeling used in the explanation of the boundary correction in second derivative. B) Position of image source in a 2-layered medium. C) First few image sources in the infinite image source series used for approximating potential in a 3-layered medium. The middle layer is assumed to be 2 mm thick. In both Panels B) and C) the position of the electrode is indicated by a triangle and the position of the source by a circle. The blue and red stars denote the positive and negative image sources respectively.

Green's function for a uniform medium with conductivity σ with Dirichlet boundary conditions at $z = L_T$ on top and $z = L_B$ on the bottom is given by

$$K(z_e, z_s) = \frac{-1}{\sigma(L_B - L_T)} \begin{cases} (L_T - z_s)(L_B - z_e), & L_T < z_s \leq z_e \leq L_B \\ (L_T - z_e)(L_B - z_s), & L_T \leq z_e < z_s < L_B \end{cases} \quad (3)$$

where z_s is the position of the source and z_e the position of the electrode (Olver and Shakiban, 2006). As shown by the blue line in Figure 4-2, this forward model is composed of two linear functions decaying to zero at L_T and L_B on each side of a given source position.

Using Equation 3, the potential φ generated by a CSD depth profile $f(z)$ at an electrode located at $z_e = z_i$ is then found using the standard linear integral equation

$$\varphi(z_i) = \int_{L_T}^{L_B} K(z_i, z') \cdot f(z')\, dz' \quad (4)$$

As with the other inverse methods, Equation 3 can be extended to a 2-layered medium using the method of images (Jackson, 1999). Assuming σ_t and σ_e are the conductivities in the top and bottom medium respectively, the method of images will result in the following forward operator (see red line Fig. 4-2 for illustration):

$$K(z_e, z_s \geq 0) = \frac{-1}{\sigma_e(L_B - L_T)}$$

$$\begin{cases} (L_B - z_e)\left((L_T - z_s) + \dfrac{L_T(\sigma_e - \sigma_t)}{L_B\sigma_t - L_T\sigma_e}(L_B - z_s)\right), & 0 \leq z_s \leq z_e \leq L_B \\ (L_B - z_s)\left((L_T - z_e) + \dfrac{L_T(\sigma_e - \sigma_t)}{L_B\sigma_t - L_T\sigma_e}(L_B - z_e)\right), & 0 \leq z_e < z_s < L_B \\ \dfrac{\sigma_e(L_B - L_T)}{L_B\sigma_t - L_T\sigma_e}(L_T - z_e)(L_B - z_s), & L_T \leq z_e < 0 \end{cases} \tag{5}$$

$$K(z_e, z_s < 0) = \frac{-1}{\sigma_t(L_B - L_T)}$$

$$\begin{cases} (L_T - z_s)\left((L_B - z_e) + \dfrac{L_B(\sigma_e - \sigma_t)}{L_B\sigma_t - L_T\sigma_e}(L_T - z_e)\right), & L_T < z_s \leq z_e < 0 \\ (L_T - z_e)\left((L_B - z_s) + \dfrac{L_B(\sigma_e - \sigma_t)}{L_B\sigma_t - L_T\sigma_e}(L_T - z_s)\right), & L_T \leq z_e < z_s < 0 \\ \dfrac{\sigma_t(L_B - L_T)}{L_B\sigma_t - L_T\sigma_e}(L_T - z_s)(L_B - z_e), & 0 \leq z_e \leq L_B \end{cases} \tag{6}$$

These expressions look considerably more complicated than what we are used to get from the image method in the 3-dimensional case for a boundary located at $z = b$

$$K(z_e, z_s \geq b) = \frac{1}{\sigma_e}\begin{cases} \varphi(z_e, z_s) + \dfrac{\sigma_e - \sigma_t}{\sigma_e + \sigma_t}\varphi(z_e, -z_s + 2b), & z_e \geq b \\ \dfrac{2\sigma_e}{\sigma_e + \sigma_t}\varphi(z_e, z_s), & z_e < b \end{cases}$$

where $\varphi(z_e, z_s)$ is the potential at position z_e generated by a source at position z_s in a homogeneous medium of conductivity $\sigma = 1$ (see Fig. 4-1B for the position of the source and its image). But most of the complexity observed in the 1-D case is due to the (possibly) non-symmetric bounding interval around the discontinuity. When setting $L = L_B = -L_T$, we see that we recover the familiar form of the image method. A small peculiarity worth mentioning is that when $z_e < z_s$ then computing the contribution of the image source $\varphi(z_e, -z_s)$ will require using the branch with $z_s < z_e$ since the position of the image source $(-z_s)$ will be smaller than z_e.

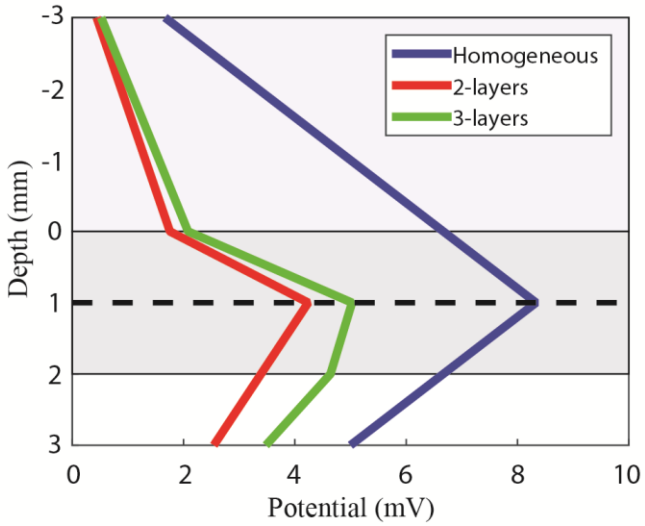

Figure 4-2: InfPlane forward model for a source of unit strength positioned at z=1 mm. The blue, red and green curves show the potential as a function of depth for a homogeneous, 2-layered and 3-layered medium respectively. In the homogeneous medium the conductivity is assumed to be $0.3\ S/m$ (gray matter). In the 2-layered and 3-layered case the conductivity in the top layer is set to $1.7\ S/m$ (CSF conductivity) while in the 3-layered medium the bottom conductivity is set to $0.1\ S/m$ (white matter conductivity). In all three cases $L_B = -L_T = 5\ mm$.

So far, we have only discussed 2-layered media, which can account for a single jump discontinuity in conductivity. A natural extension would be to approximate the volume conductor by a three-layered medium where the two outer media are again taken to be semi-infinite. In the case of cortical recordings, one could think of these three layers as representing CSF (or saline), gray matter, and white matter respectively.

4.2.3 3-Layer Medium Parametrization

The extension of the method of images to a three-layered medium is complicated by the fact that the solution is theoretically composed of an infinite number of image sources (Barrera et al., 1978). This is explained by the fact that the purpose of introducing an image source is to satisfy a boundary condition at the interface between two media. But in a 3-layered medium, every time an image source is placed to satisfy the boundary condition on one interface, it disrupts the boundary condition at the other interface, thus requiring the addition of another image source to counter-balance the first one. It is clear that this leads to a situation akin to an infinite regress.

In practice, the situation is however not as dire as it might look because the image sources have the characteristic of being placed further and further away from the original source (see Fig. 4-1C for an illustration of the image source positions for given source and electrode locations). Moreover, the strength of each image source is always a fraction of the strength of the previous one. Hence, their effect on the potential is continuously diminished. This makes it possible to compute a good approximation to the true solution by truncating the infinite series at a given number N. As described by Barrera et al., another option would be to move to the Fourier domain and evade the issue of the infinite sum altogether (Barrera et al., 1978). However, recovering the potential in space from its description in Fourier space results in an integral expression containing the zero-order Bessel function. Since it can be quite difficult to approximate such an integral numerically, we have preferred to explore the series expansion in the space domain.

In Appendix B we show the resulting equations for approximating the potential for the nine possible combinations of source and electrode locations in a 3-layered medium. To our knowledge, only the case where both the source and electrode are located in the middle layer has been considered previously (Gold et al., 2006, Ness et al., 2015).

4.3 Results

4.3.1 A motivating example

The purpose of this work is to investigate 1-D CSD estimation in presence of laterally extended sources. In particular, we are interested in understanding which CSD estimation method is the best choice when the extension of the sources is large but not exactly known. There are multiple ways in which such a situation can occur in an experimental setting. As a motivating example, we will consider the particular case in which the recorded potential is composed of the superposition of a global and a local activation. Similar situations can be expected when recording from primary sensory cortices in response to a large stimulus containing local features. It also represents a simplification for recordings in response to a stimulus leading to non-homogeneous cortical activation such as the response of the primary visual cortex to oriented gratings for example. For such a situation the potential φ at an electrode located at $z_e = z_i$ is given by

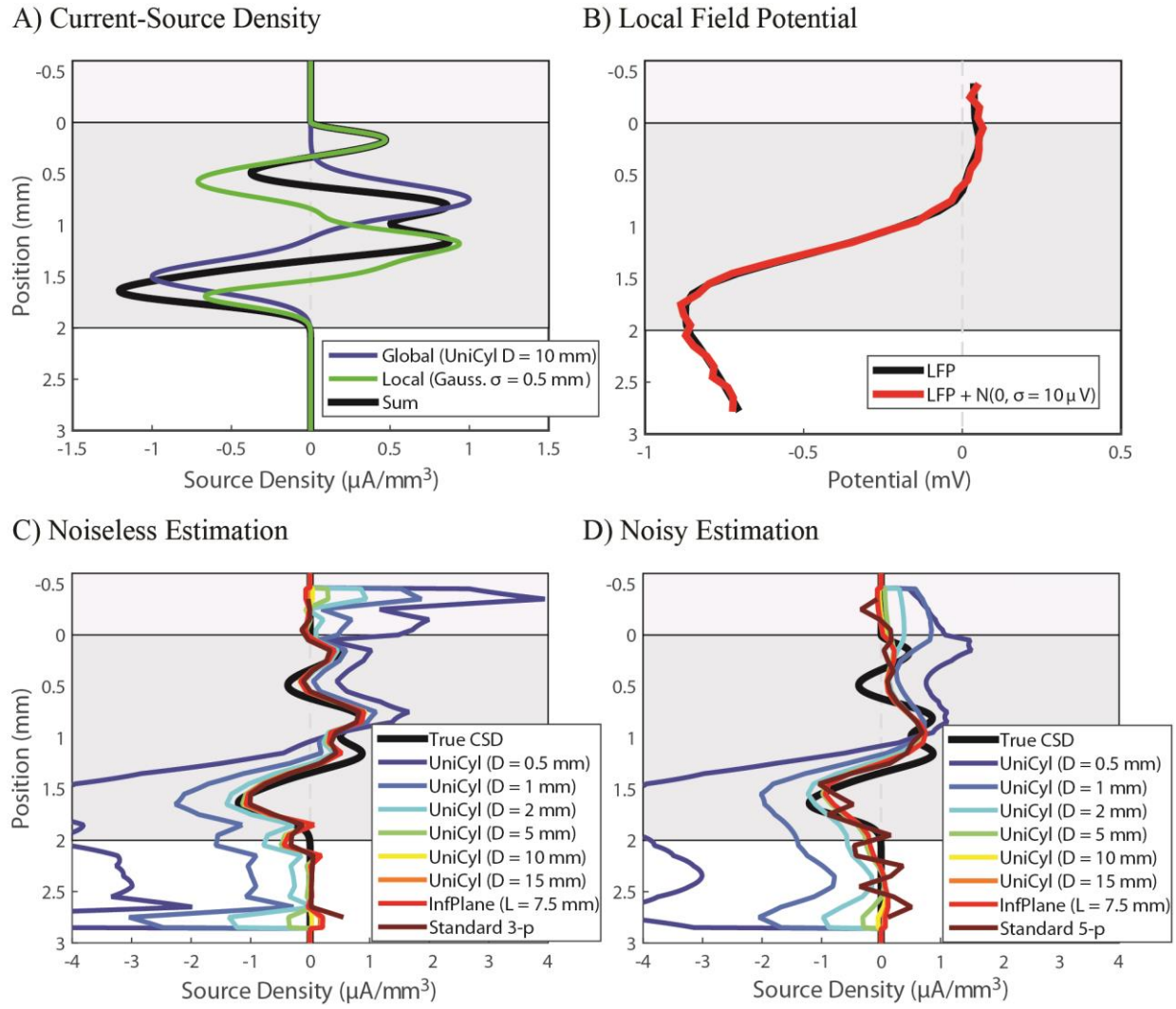

Figure 4-3: CSD estimation of a localized source overlaid with a laterally extended global activation. A) Plot of the CSD depth profiles. The green curve shows the local activation which is assumed to be laterally distributed as a Gaussian with a standard deviation of 0.5 mm. The profile is the same as used in the modelling study by Glabska et al. The blue curve shows the dipolar global activation which is assumed to be uniformly distributed within a cylinder of 10 mm diameter. Both the cylinder and the Gaussian are centered on the recording electrode. The black curve is the sum of the local and global activation and therefore represents the estimation target. B) The noiseless (black) potential generated by the total activation in A). The potential was generated assuming a 3-layered medium with conductivities of $1.7, 0.3, 0.1\ S/m$. The red curve shows the same potential with additive Gaussian white noise with a standard deviation of $10\ \mu V$. C) and D) show the non-regularized and regularized CSD estimation for various assumed source diameters using the noiseless and noisy potential respectively. The sources were estimated assuming a 2-layered medium (the bottom layer was omitted).

$$\varphi(z_i) = \int K_g(z_i, z') \cdot f_g(z') + K_l(z_i, z') \cdot f_l(z')\, dz' \quad (7)$$

where f_g and f_l are the global and local CSD depth profiles respectively and K_g and K_l the associated forward problems.

Given Eq. 7, the question is thus what CSD estimation method will provide the best estimate of the overall activity and how are f_g and f_l represented in the estimated CSD. Currently, most researchers prefer approaching this problem using the standard CSD estimation method instead of the inverse methods despite the limitations of the standard method discussed in Section 2.

In Figure 4-3 we investigate this problem using simulated data. Panel 4-3A shows a large dipolar global CSD depth profile overlaid with a patterned local profile. The local response is the same CSD profile as used in Glabska et al. (Glabska et al., 2014). To simulate the potential in Panel 4-3B, we assumed that the global activation was a uniform cylinder of 10 mm diameter and the local sources were distributed as a Gaussian function with a standard deviation of 0.5 mm. Both the uniform cylinder and the Gaussian were centered on the electrode position. We also included additive zero-mean Gaussian white noise with standard deviation of $10\ \mu V$ to simulate recording noise. The medium was parametrized as a 3-layer setup with aCSF conductivity $\sigma_t = 1.7\ S/m$ in the top medium (blue shading $z < 0$), extracellular gray matter conductivity $\sigma_e = 0.3\ S/m$ for the middle medium (gray shading $0 \leq z \leq 2\ mm$, and white matter conductivity $\sigma_b = 0.1\ S/m$ for the bottom medium (Wagner et al., 2004). The recording points were chosen to simulate a 32 channel electrode with $100\ \mu m$ channel separation. The electrode was positioned with 4 channels in the top medium and the 5th channel located $50\ \mu m$ below the surface. This setup simulates a cortical recording with an electrode spanning the whole cortical depth along with contacts in aCSF and white matter.

Panels 4-3C and 4-3D show the estimated CSD using different inverse model diameters and the standard method for the noiseless (black curve in 4-3B) and noisy (red curve in 4-3B) potential respectively. For the inverse methods, UniCyl refers to an assumed uniform cylinder forward model and the InfPlane is the 1-D Poisson inverse problem presented in Section 2.2. For the InfPlane, the bounding box was measured from the edge of the electrode, i.e. $L_T = 1.2 -$

$(L + 1.55)$ and $L_B = 1.2 + (L + 1.55)$. In all the inverse methods, the rCSD expansion method was used to parametrize the sources (Kropf and Shmuel, 2016). For Panel 4-3D, all inverse methods were regularized using a minimum coefficient norm prior with a regularization coefficient of 0.01.

It is obvious from both Panel 4-3C and 4-3D that small diameter source estimation methods are not well suited for estimating the CSD in this situation. Moreover, the large diameter estimations ($D \geq 5\ mm$) as well as the InfPlane perform very similarly with the main differences being located at the edge of the electrode. These edge effects are particularly apparent in the noiseless estimation where there is no smoothing due to regularization. Although the standard estimation method performs reasonably well, it appears to be more sensitive to noise than the regularized inverse methods with D ≥ 5 mm, even when it is supplemented with the additional smoothing step (i.e. 5-point method).

Interestingly, all large diameter estimation methods as well as the standard method are able to capture the features of the summed CSD despite the large differences in the extent of its components. Hence, we see from this example that the CSD is mainly sensitive to local features even when assuming extended sources. To understand why this is the case and why small source estimation methods perform poorly in this situation it is worth looking at the contributions to the potential from distant sources.

In Figure 4-4A, we show the potential generated by a point source of unit strength ($1\ \mu A$) located at position $\vec{x} = (0,0,0)$ in a uniform medium with conductivity $\sigma = 0.3\ S/m$. It is well established that in a homogeneous and isotropic medium, the potential decays as $1/r$ where r denotes the distance from the source to the recording point (Plonsey and Barr, 2007). The potential is shown for a depth of 3 mm to show the extent to which the potential generated by a point source varies across a 32 channel electrode with an inter-contact separation of $100\ \mu m$. We see that the further we are laterally displaced from the source, the less the potential varies across the electrode. This is highlighted in Panel 4-4B which shows the percentage difference between the potential at depth z compared to the potential at $z = 0$. The fact that the potential varies only

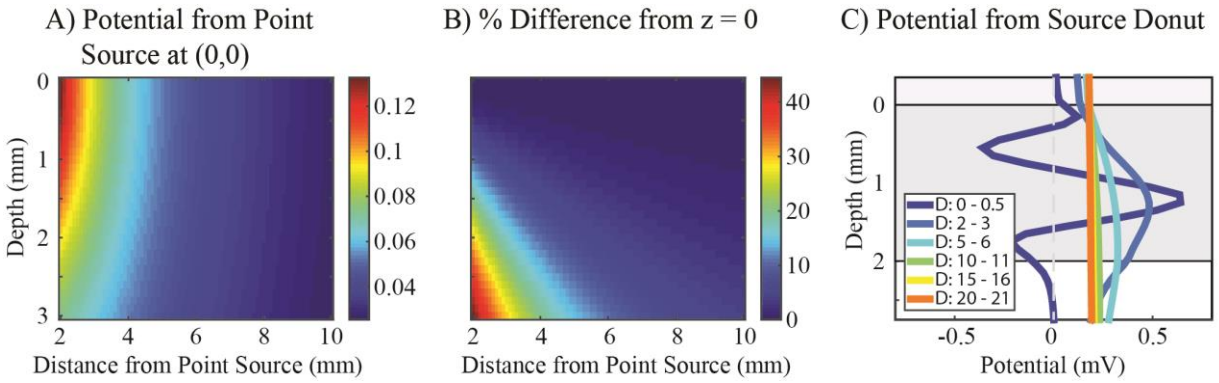

Figure 4-4: Demonstration of effect of source distance on recorded potential. A) Potential generated by a point source located at $(\rho, z) = (0, 0)$ in a homogeneous medium ($\sigma = 0.3\ S/m$) as a function of depth z and lateral distance ρ. The total depth z is chosen to match the typical extent of a cortical electrode. B) Same as A) but shown as the percentage difference from the potential at $z = 0$. C) Potential generated by source donuts of various diameters assuming the local Glabska source profile from Figure 4-3A. It is very apparent that the source structure is very rapidly smoothed out in the potential because of the small percentage variation across the span of the electrode shown in B).

slightly across the electrode the farther it is laterally displaced from the source has an important consequence for CSD estimation. This is further highlighted in Panel 4-4C which shows the potential generated at the center of 1 mm thick donuts of various diameters. The CSD depth profile is the same as the local profile in Figure 4-3A. The major point to note is that the distant sources produce increasingly smoothed potentials until they contribute only approximately linear and then constant functions. Hence, all the structure in the potential is due to the sources close to the electrode. But as mentioned in Section 2.1, constant and linear functions are in the nullspace (also termed kernel) of the standard CSD estimator for the inner contacts (in other words, given constant or linear field potentials, the standard CSD estimation for the inner electrodes is zero). They thus do not affect estimation. This is however not the case for the small diameter inverse estimators which explains why these perform badly in presence of laterally extended sources.

We have seen in Figure 4-3 that large diameter uniform cylindrical (UniCyl) inverse models perform very similarly to the infinite plane (InfPlane) estimator described in Section 2.2. Moreover, we found that both methods were more resistant to noise than the standard CSD method. It is however not yet clear at what diameters these formulations converge to each other. We will thus investigate the relationship between these two formulations in the next section.

4.3.2 Infinite plane versus uniform cylinder CSD estimation

We have shown in Section 2.2 that the standard CSD estimation method can also be expressed as an inverse problem by using the Green's function for the 1-D Poisson equation as the forward operator. The most straightforward way to compare this estimator to the uniform cylindrical inverse formulation is by explicitly comparing their forward models.

In a homogeneous and isotropic volume conductor with conductivity σ, the forward model for cylindrical sources of radius R and uniform lateral density is given by (Pettersen et al., 2006)

$$K_{UniCyl}(z_e, z_s) = \frac{1}{2\sigma}\left(\sqrt{R^2 + (z_s - z_e)^2} - |z_s - z_e|\right) \tag{8}$$

In comparison, the forward model for the inverse formulation of the standard method ($K_{InfPlane}$) with boundaries chosen symmetrically around a point a (i.e. $L_T = a - L$ and $L_B = a + L$), is found from Eq. 3 to be given by

$$K_{InfPlane}(z_e, z_s) = \frac{1}{2\sigma}(L - |z_s - z_e| - \frac{(z_s - a)(z_e - a)}{L}) \tag{9}$$

It is clear that the two formulations share a number of features starting with the term containing the distance between the source and the electrode ($|z_s - z_e|$). Moreover, whenever the radius $R \gg \max(|z_s - z_e|)$, K_{UniCyl} can be approximated by

$$K_{UniCyl}(z_e, z_s) \cong \frac{1}{2\sigma}(R - |z_s - z_e|) \tag{10}$$

If we assume that the sources are all located within the span of the electrode contacts, the constraint $R \gg \max(|z_s - z_e|)$ is equivalent to say that the radius should be larger than the extent of the electrode. Under this condition, comparison of Eq. 8 and 9 shows that the length L of the bounding box in $K_{InfPlane}$ plays a similar role to that of the radius R in K_{UniCyl}. An approximation similar to Eq. 9 can also be found for a Gaussian forward model with a standard deviation of R (see (Kropf and Shmuel, 2016) for the forward model). In this case, the first order McLaurin expansion is given by

$$K_{Gauss}(z_e, z_s) \cong \frac{1}{2\sigma}\left(\frac{\sqrt{2\pi}}{2}R - |z_s - z_e|\right)$$

In this case, the length L of the bounding box of the bounding box is then obviously related to a scaled version of the standard deviation.

The same relationship can be shown to hold for the case of a two-layered medium, for which the image source contribution in the bottom medium (assuming the source is in the bottom medium as well) is given by

$$\frac{1}{2\sigma_e}\frac{(\sigma_e-\sigma_t)\left(1-\frac{a}{L}\right)}{(\sigma_e+\sigma_t)-\frac{a}{L}\cdot(\sigma_e-\sigma_t)}(L-|z_s+z_e|+2*a-\frac{(a-z_s)(z_e-a)}{L})$$

Similarly, the potential in the top medium is given by

$$\frac{1}{(\sigma_e+\sigma_t)-\frac{a}{L}\cdot(\sigma_e-\sigma_t)}(L-|z_s-z_e|-\frac{(z_s-a)(z_e-a)}{L})$$

Since the two formulations clearly converge to each other at large source diameters, the main questions left to answer are: (1) the diameter at which they are sufficiently similar to be exchangeable, and (2) the errors that are introduced when using a wrong diameter.

4.3.2.1 Similarity between discrete uniform cylinder and infinite plane delta estimator

To determine the similarity between the two estimators as a function of the source diameter, we first looked at methods from perturbation theory. Assume that $\boldsymbol{F}_{UniCyl}^{-1}\varphi = f_{UniCyl}$ is the CSD estimation (f_{UniCyl}) obtained with the cylindrical source estimator $\boldsymbol{F}_{UniCyl}^{-1}$ and $\boldsymbol{F}_{InfPlane}^{-1}\varphi = f_{InfPLane}$ the same for the InfPlane estimator (note that the same potential φ is used in both cases). It is then possible to get the following upper bound on the difference in estimation $f_{InfPlane} - f_{UniCyl}$

$$\frac{\left\|f_{InfPlane}-f_{UniCyl}\right\|}{\left\|f_{UniCyl}\right\|} = \frac{\left\|\boldsymbol{E}\,\boldsymbol{F}_{UniCyl}f_{UniCyl}\right\|}{\left\|f_{UniCyl}\right\|} \leq \left\|\boldsymbol{E}\,\boldsymbol{F}_{UniCyl}\right\|$$

where $\boldsymbol{E} = \boldsymbol{F}_{InfPLane}^{-1} - \boldsymbol{F}_{UniCyl}^{-1}$ (Olver and Shakiban, 2006). However, when comparing this upper bound to errors obtained in simulations, it turns out that the upper bound is too generous to serve as a useful measure of the similarity between the estimators. We thus turned to more direct comparison methods.

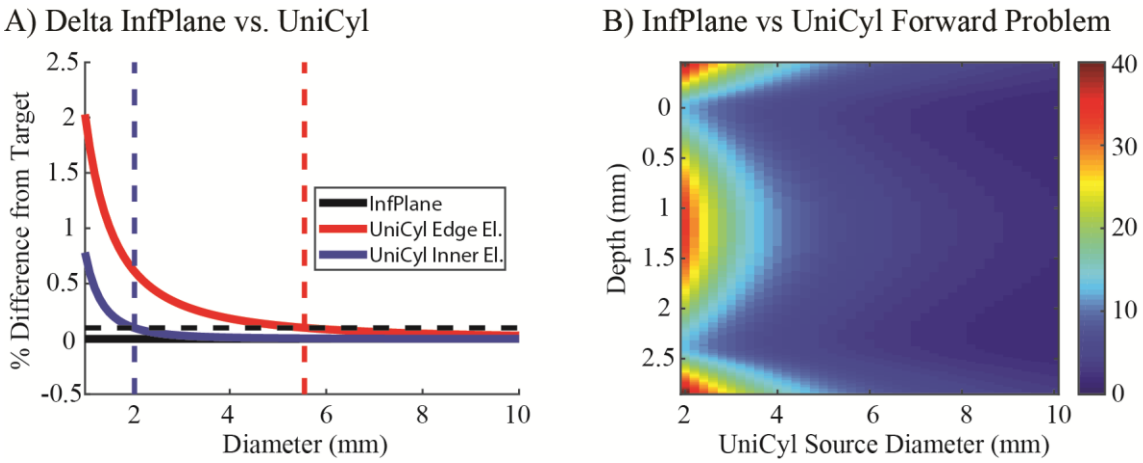

Figure 4-5: Difference between the InfPlane and UniCyl forward models. A) Decay of diagonal elements in *delta* iCSD estimator. The target in this case is a $1 + h/R$ decay for the edge electrode and a constant value of 2 for the inner electrodes. The red and blue dashed lines denote the source diameter at which the percentage difference between the UniCyl model and the target is below 0.1% as shown by the black dashed line. B) Percentage difference between UniCyl and best fitting InfPlane forward models as a function of source diameter and depth. The error is averaged over sources positioned in 0.01 mm steps between [-0.45, 2.85] mm.

Pettersen et al. have shown that, at large source diameters, the *delta* estimator converges to the standard CSD method for the interior electrodes (Pettersen et al., 2006). Interestingly, we found that this also applies to a two-layer medium in which the *delta* estimator converges to Liu et al.'s estimator presented in Section 2.1 (Liu et al., 2000). Moreover, it was shown that the diagonal element of the *delta* estimator at the edge electrodes converged to $1 + {}^{h}/_{R}$ at large diameters with h being the interelectrode separation and R the radius of the cylindrical source disks. Figure 4-5A shows that the same behavior is observed for the InfPlane forward model using the *delta* source parametrization with the box edges set a distance x from the first and last electrode (i.e. $L_T = -z_{e1} - x$ and $L_B = z_{eN} + x$ leading to $a = \frac{z_{e1}+z_{eN}}{2}$ and $L = x + \frac{|z_{eN}-z_{e1}|}{2}$). However, in contrast to the UniCyl forward model, the edge electrodes for the *delta* InfPlane forward model follow a strict $1 + {}^{h}/_{x}$ pattern even at small values of x. Hence, a possible way to identify the radius at which both estimation methods are sufficiently similar to each other is to set $x = R$ and measure the amount by which the edge estimator in $\boldsymbol{F}^{-1}_{UniCyl}$ differs from $1 + {}^{h}/_{R}$. From the red curve in Figure 4-5A, we see that the difference is below 0.1 % at radii larger than ~5.5 mm.

Hitherto, we have only compared the edge electrode estimators but a similar comparison can be done for the inner electrodes. The blue curve in Fig. 4-5A shows the average percentage by which the diagonal elements differ from 2 (excluding the edge estimators). We find that the estimation at the inner electrodes converges even faster than at the edge since the difference is below 0.1 % at radii larger than ~2 mm. Hence, we conclude that for the *delta* estimator, the two formulations can be considered interchangeable for diameters exceeding ~2-5 mm.

4.3.2.2 Similarity between continuous uniform cylinder and infinite plane estimator

A limitation of the comparison approach used in Fig. 4-5A is that it is only applicable to the discrete *delta* source parameterization method. In order to extend the comparison to any source parameterization, we also compared the forward models directly. In a first step, we set out to determine the boundary conditions imposed on the InfPlane forward model (i.e. Eq. 8) that minimize its difference with the UniCyl forward model. We found that choosing $L = R + \frac{1}{(1+h)R} - \frac{3h}{R^2}$ yielded the minimal average error over electrode and source positions (this obviously only applies when the radius is large enough to put the boundary box beyond the edge of the electrode). We also tested this boundary condition on simulations and found that it indeed markedly reduced the source reconstruction at the edge electrodes at small source diameters. In Figure 4-5B, we show the percent difference between the forward model equations 7 and 8 averaged over sources positioned in 0.01 mm steps between $z_s = z_{e1} - h$ and $z_s = z_{eN} + h$. For the purpose of this comparison, we assumed a homogeneous medium with conductivity $\sigma = 0.3\ S/m$. For each source diameter, the percentage error was calculated by dividing the average over source positions of the absolute difference between the UniCyl and InfPlane forward model by the largest value in the UniCyl forward model

$$\varepsilon = 100 \cdot 2\sigma \frac{\overline{\left|K_{UniCyl}(z_e, z_s, R) - K_{InfPlane}(z_e, z_s, L)\right|}}{R}$$

Again, we find that there is a marked drop in the average error at diameters between ~2-5 mm. Hence, the findings from the *delta* source parameterization seem to remain valid for other source parameterizations. It is important to note that the error in Figure 4-5B should not be directly compared to the one from Figure 4-5A because the former compares the forward equation while the latter is based on the inverse estimator.

4.3.2.3 Effect of diameter mismatch between forward and inverse model

So far, we have shown that the two forward model formulations from Eqs. 7 and 8 converge to each other and can be used interchangeably at diameters larger than ~2-5 mm. However, in most experimental settings, it is very difficult to get a good approximation of the activation diameter. It is thus interesting to investigate the errors that are introduced when wrong source diameters are assumed for estimation. Figure 4-6A shows CSD estimation using the rCSD source parameterization and a uniform cylinder (UniCyl) forward model of various diameters. The electric potential data was generated using the local CSD profile from Fig. 4-3A and a uniform cylindrical lateral source profile with a diameter of 2 mm. For the forward problem, the medium was assumed to be 3-layered with the same parameters as in Fig. 4-3. The inverse problem was solved without including the white matter (i.e. using a 2-layered medium). The same setup is used in Figure 4-6B but with a UniCyl lateral source profile with a diameter of 5 mm for the forward problem. No noise was added to the potential and no regularization was used for the estimation shown in Panels 4-6A and B.

For both 2 and 5 mm real source diameter, we find that the estimation is quite insensitive to the assumed source diameter. Moreover, we find that, as predicted from the comparison between the forward models, the estimation is comparable between the UniCyl and InfPlane forward models. In order to determine which of them performs better, we show in Panels 4-6C and 4-6D the subtraction of their respective errors. We find that at the edge electrodes, the UniCyl forward model performs better than the InfPlane, while the InfPlane seems slightly less prone to errors in the middle electrodes. However, from Panels 4-6A and B it is very apparent that the overall error introduced by assuming wrong diameters is small whenever the activation diameter exceeds ~2 mm. This behaviour is also preserved when using noisy potential data. Panels 4-6E and F show the estimation when using the same forward model and source profile used for Panels 4-6A and B but with adding zero-mean Gaussian white noise with a standard deviation of $25\ \mu V$ (the same noise realization is used in both cases). All inverse methods were regularized using a minimum coefficient norm prior with a regularization coefficient of 0.01. Again, the estimations are stable under changes in the assumed source diameters.

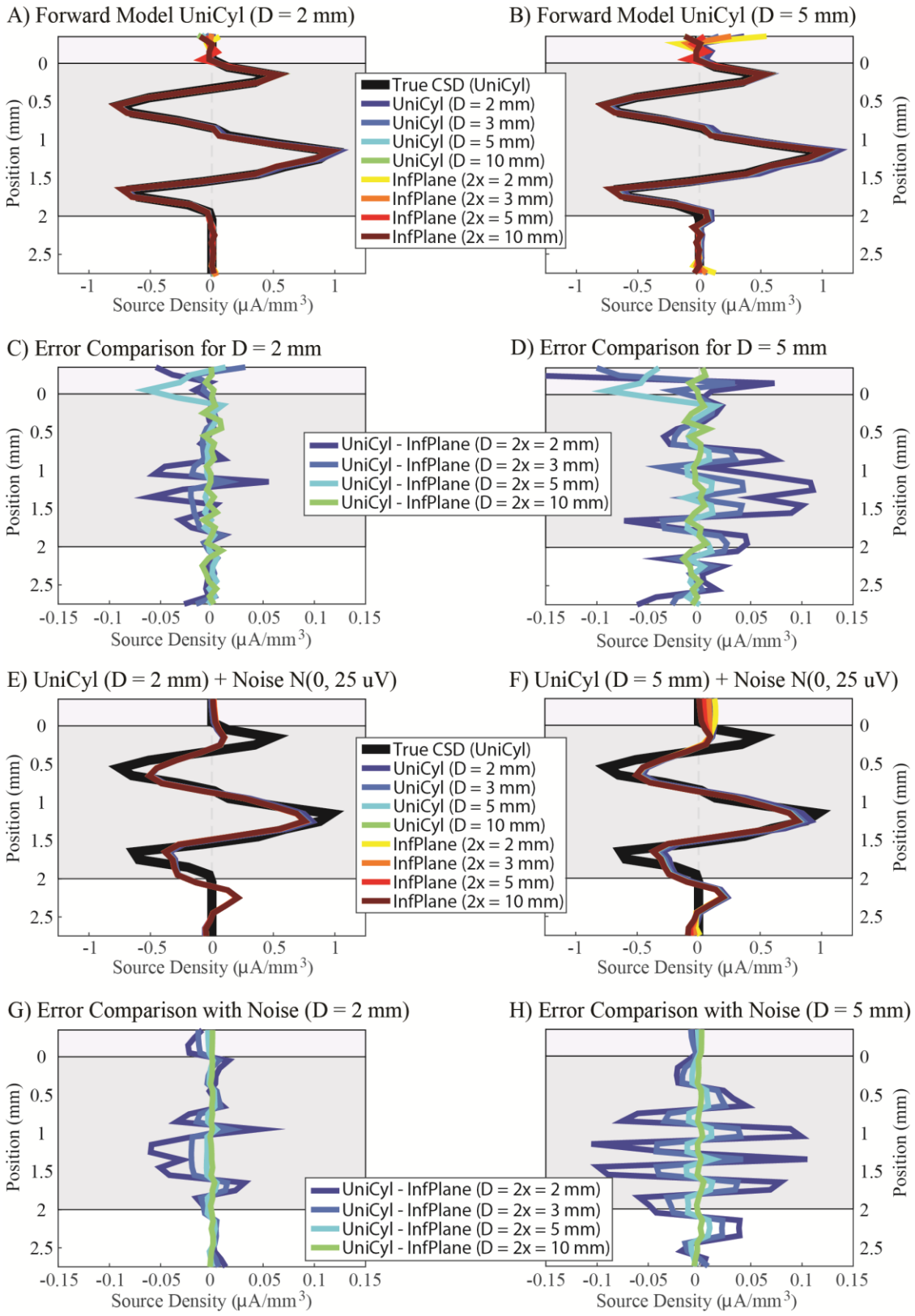

Figure 4-6: Comparison of errors introduced by mismatch in source diameters. (continued on next page)

As mentioned previously, a major appeal of the standard CSD method is its apparent ease of use. However, as we have just shown, the difference between the forward model of the 1-D Poisson equation (InfPane) and the inverse formulation with cylindrical sources (UniCyl) diminishes for large source radii. It thus raises the question of where the parameters of the inverse methods are hidden when using the standard CSD method.

4.3.3 Inverse CSD versus Standard CSD

There are 4 types of parameters that define the forward problem in the inverse CSD methods: the electrode positions, the conductivities, the lateral source distribution (e.g. uniform cylinder), and the lateral extent of the sources (i.e. the radius in the cylindrical source case). In its most common application, the standard CSD method is applied with the assumption of a homogeneous medium. In this case, the requirements on the electrode positions and on the conductivities are the same between the standard and inverse methods. The conductivity will simply scale the estimation and electrode positions only need to be known relative to each other, which means that only the channel separation is of importance. And even if one would use the standard method in a non-homogeneous medium, the same requirements as for the inverse methods would apply for the electrode positions and the conductivities (see Liu et al.'s estimator in Section 2.1). For the lateral source distribution, we have shown above that in the case of a cylinder with radially uniform density (UniCyl) as well as for a Gaussian lateral profile, the forward models of both methods converge to the same formulation. Hence, the only remaining parameter to be found is the radius of the sources.

As mentioned in Section 2.1, reducing Poisson's equation (Eq. 1) from 3-D to 1-D (Eq. 2) theoretically requires working with laterally infinitely extended sources. Therefore, it is tempting to assume that the standard method is simply retrieved in the limit of cylindrical sources with infinite radius. However, examination of the forward model (Eq. 7 or 9) shows that

Figure 4-6: (continued from previous page) The rCSD source parameterization was used for all estimations. A) Source estimation for Glabska CSD profile distributed within a uniform cylinder (UniCyl) of 2 mm diameter. We see that using larger diameters does not significantly affect the estimation. B) Same as A) but for a source diameter of 5 mm. We see that assuming smaller source diameter tends to introduce errors at the edge electrodes. C) and D) Estimation errors using the InfPlane forward model subtracted from the errors obtained with the UniCyl model. E-H) Same as A-D) but with added white Gaussian noise (standard deviation of 25 μV). All estimations in the presence of noise were regularized using a minimum coefficient norm prior with a regularization coefficient of 0.01.

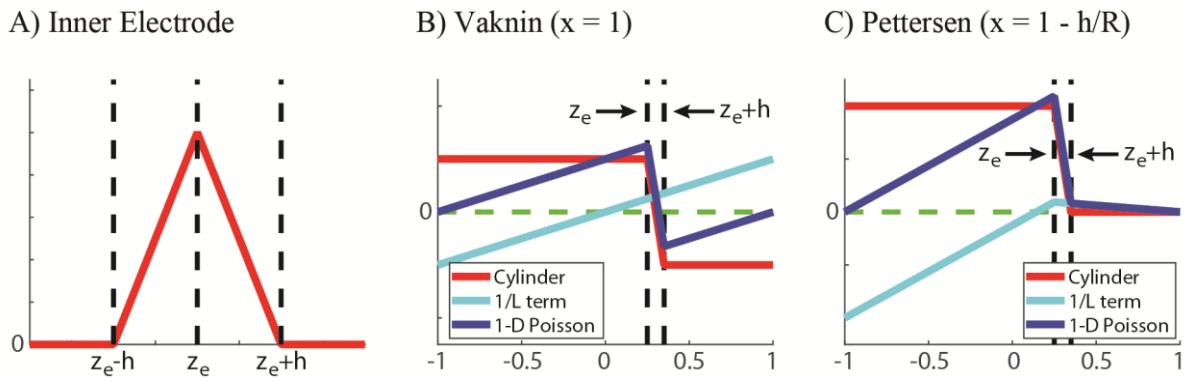

Figure 4-7: Standard CSD estimator applied to the InfPlane and the large source approximation of the UniCyl forward model. A) Inner electrode estimator is found to compute a weighted average of sources between adjacent electrodes. B) and C) Edge estimators assuming Vaknin's and Pettersen's $1 - h/R$ boundary conditions, respectively. The red curve shows the result of applying the standard three-point estimator on the large diameter approximation of the cylindrical forward model. The cyan curve shows the added contribution from the 1/L terms in the InfPlane forward model. The blue curve is the sum of the red and cyan curves: it represents the total source weighting at the top-most electrode for the InfPlane forward model. We see that the sources beyond the electrode span affect the estimation for both boundary conditions. Moreover, sources within the electrode span are not evenly weighted.

this is not possible because the potential diverges in the limit of $R \to \infty$. This is actually a well-known situation in electrostatics and is reflected in the fact that Poisson's equation in 1-D doesn't allow for a Green's function which decays to zero at infinity. Hence, there seems to be no obvious way to relate the radius of the sources to the standard method.

To gain insight into how the source radius affects the standard CSD method, we examined the result of applying the standard method to the large diameter forward equations 8 and 9. For the inner electrodes, this means computing

$$\tilde{C}(z_e) = \sigma \frac{-\varphi(z_e + h) + 2\varphi(z_e) - \varphi(z_e - h)}{h^2}$$
$$= \frac{\sigma}{h^2} \int dz_s \, C(z_s) \cdot (-K(z_e + h, z_s) + 2K(z_e, z_s) - K(z_e - h, z_s))$$

which is obtained by applying the 3-point kernel to Equation 4 assuming an inter-electrode separation of h. The integration boundaries are chosen such that they completely encompass the sources $C(z_s)$. For both the InfPlane forward operator (Eq. 8) as well as the large diameter approximation to the cylindrical source forward model (Eq. 9), this results in

$$\tilde{C}(z_e) = \frac{1}{h^2}\int_{z_e-h}^{z_e+h} dz_s\, C(z_s)\cdot(h - |z_s - z_e|)$$

which is simply a weighted average of the true sources between $z_e - h$ and $z_e + h$ as shown in Figure 4-7A. We note here that, as expected, this estimator is independent of the source radius R and the extent of the bounding box.

As mentioned in Section 2.1, the standard estimation at the edge electrodes is complicated by the fact that it depends on the choice of boundary conditions. For example, the estimator for the top most electrode is given by

$$\tilde{C}(z_{e1}) = \sigma\frac{-\varphi(z_{e1}+h) + (2-x)\varphi(z_{e1})}{h^2}$$

where x depends on the choice of boundary condition (a similar equation applies to the bottom most electrode). Applying this estimator to the forward equations 8 and 9 yields

$$\tilde{C}(z_{e1}) = \frac{1}{2h^2}\begin{pmatrix} (1-x)\int dz_s\, C(z_s)\cdot(R - |z_s - z_{e1}|) + h\int dz_s\, C(z_s) \\ -2\int_{z_e-h}^{z_e+h} dz_s\, C(z_s)\cdot|z_s - z_{e1}| - 2h\int_{z_e+h} dz_s\, C(z_s) \\ +(1-x)b\int dz_s\, C(z_s) + \frac{h-(1-x)(z_e-a)}{L}\int dz_s\, C(z_s)\cdot(z_s - a) \end{pmatrix} \quad (11)$$

The last two terms only apply to Eq. 8. The first one accounts for the difference between L and R (i.e. $b = L - R$) whereas the second term accounts for the $1/L$ term in Eq. 8.

In this formulation, Vaknin's method (i.e. Neumann boundary conditions) is enforced by setting $x = 1$, which simplifies Eq. 11 to

$$\tilde{C}(z_{e1}) = \frac{1}{2h^2}\begin{pmatrix} h\int^{z_e} dz_s\, C(z_s) - \int_{z_e-h}^{z_e+h} dz_s\, C(z_s)\cdot(h - 2|z_s - z_{e1}|) - h\int_{z_e+h} dz_s\, C(z_s) \\ +\frac{h}{L}\int dz_s\, C(z_s)\cdot(z_s - a) \end{pmatrix}$$

The resulting estimator is shown in Figure 4-7B for the cases with and without the $^1/_L$ component and under the assumption that $a = b = 0$. There are a few findings worth pointing out here: First, Vaknin's procedure will provide a correct estimate as long as the sources are balanced, the electrode span encompasses all the sources, and the radius is sufficiently large compared to the electrode span. This is well explained by electrostatics from where we know that the electric field vanishes outside of a region containing a set of balanced infinitely extended sheets, thus leading to a constant potential (Griffiths, 2013). In contrast, all unbalanced sources both within or extending beyond the span of the electrode will affect the estimation at the edge electrodes. This represents a generalization of the result proposed by Pettersen et al. (Pettersen et al., 2006). Moreover, in the case of smaller bounding boxes, the position of the electrode with respect to the center of the box affects the estimation at the edge electrodes because the $^1/_L$ component adds a linear trend with an intercept of 0.

Following the observation of the *delta* iCSD estimator at large diameters, Pettersen et al. also proposed a modified boundary condition corresponding to setting $x = 1 - {}^h/_R$. In this case, Eq. 10 is reduced to

$$\tilde{C}(z_{e1}) = \frac{1}{2h^2}\left(\begin{array}{c} 2h\int^{z_e} dz_s\, C(z_s) - 2\int_{z_e-h}^{z_e+h} dz_s\, C(z_s)\cdot(h-|z_s-z_{e1}|) \\ +\frac{h}{L}\left(1-\frac{z_{e1}-a}{R}\right)\cdot\int dz_s\, C(z_s)\cdot(z_s-a) + \frac{hb}{R}\int dz_s\, C(z_s) \end{array}\right)$$

The resulting estimator is shown in Figure 4-7C again for the cases with and without the $^1/_L$ component. A particularly interesting choice of parameters is $L = R + b$, $a = z_{e1} + b$ and $b = (z_{eN} - z_{e1})/2$. This corresponds to a situation where the bounding box is centered on the electrode and the top and bottom edges of the box are distance R away from the top most and bottom most electrode respectively (as used in Figure 4-3 and 4-6). As mentioned by Pettersen et al. we see that this estimator is indeed insensitive to the sources within the span of the electrode. However, as an extension to Pettersen's previous findings, we show here that the same is not true for unbalanced sources extending beyond the span of the electrode. Also, when adding the $^1/_L$ term, the estimator is again found to be affected by the position of the electrode within the bounding box.

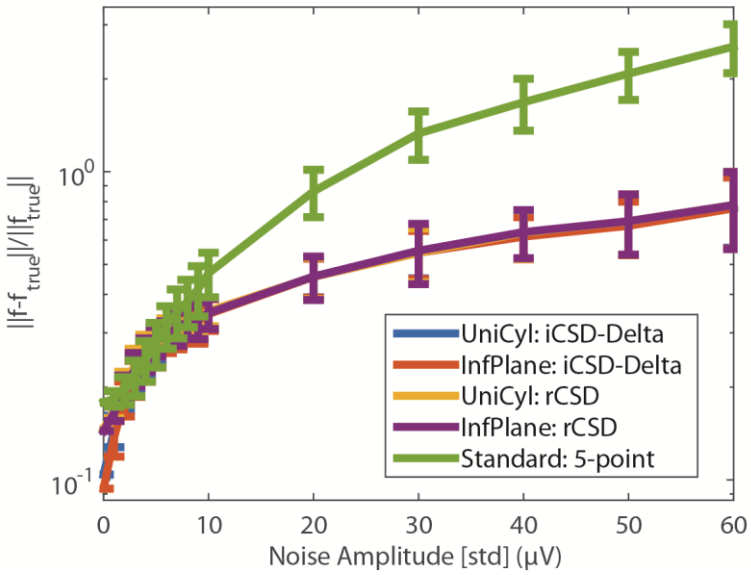

Figure 4-8: Noise sensitivity as a function of noise amplitude. For each noise amplitude, the source estimation error was computed for 100 realizations of zero-mean white Gaussian noise. The same noise realization was used for all estimation methods. It is quite obvious that the 5-point standard estimator (green curve) is less stable under noise than the inverse estimators. Moreover, the InfPlane (red and magenta curves) and UniCyl (blue and orange curves) estimators show highly comparable noise tolerance.

4.3.4 Effect of Noise on Estimation

So far we have only compared the methods assuming a noiseless potential. But in Section 2.1 we have already mentioned that derivative estimation is an unstable operation and that the 5-point estimator has been introduced precisely to deal with the inevitable recording noise in experimental conditions. However, as of now we don't know whether regularized inverse methods or the 5-point standard estimator are more sensitive to recording noise. To address this question, we show in Figure 4-8 the estimation error for the standard and inverse methods as a function of noise level.

To simulate the potential we again used the local activation profile from Figure 4-3 with a uniform cylinder of 10 mm diameter and a 32-channel electrode with 100 μm inter-contact separation. The recording medium was set to have a uniform conductivity of $\sigma = 0.3\ S/m$. We then added white, zero-mean Gaussian noise with a standard deviation equal to the noise level. For every noise level, we created 100 noise realizations of the appropriate amplitude. Figure 4-8 shows the average and standard deviation of the estimation error over these realizations. The

same noise realizations were used for each method. We evaluated the error for both the *delta* iCSD and the rCSD source parameterization assuming a source diameter of 5 mm. This situation was intentionally chosen to favor the standard CSD method because we incorporate the uncertainty of the actual lateral source diameter by choosing the wrong source diameter in the inverse estimator. For the continuous rCSD estimation, the error was evaluated only at the electrode positions to allow a fair comparison with the *delta* iCSD and standard estimation. For all methods the error was only evaluated for the inner electrodes (i.e. electrode contacts 3-30). All inverse methods were regularized using a minimum coefficient norm prior and the regularization coefficient was automatically selected using the normalized cumulative periodogram (NCP) method (Kropf and Shmuel, 2016). We find that all inverse methods outperform the standard 5-point CSD. Moreover, in accordance with the results presented in Section 3.2, we find that the difference between the InfPlane and UniCyl forward model is negligible.

In summary, we have shown that the ease of use of the standard CSD method is largely illusory since all the parameters from the inverse methods are implicitly contained in the standard estimator. In particular, the lateral source extent is affecting the standard estimation at the edge electrodes even at large source diameters. Moreover, we have found that the 5-point standard estimator is more sensitive to noise than the inverse methods. Hence, there appears to be no particular advantage in using the standard estimation method over any of the inverse estimators.

4.3.5 Three- versus Two-Layered Medium Parameterization

The final point that we investigate is the effect on estimation accuracy of simplifying the medium from a three- to a two-layered medium. In particular, we will focus on emulating a cortical recording by choosing electrical conductivity values expected from CSF ($1.7\ S/m$), gray matter ($0.3\ S/m$), and white matter ($0.1\ S/m$).

Figures 4-9A and 4-9B show the potential generated by a unit strength point current source located at various positions in a two- and three-layered medium respectively. For the three-layered medium, the infinite series of image sources was truncated at $N = 20$ as in Ness et al. (Ness et al., 2015). The gray and white lines mark the beginning of the gray and white matter

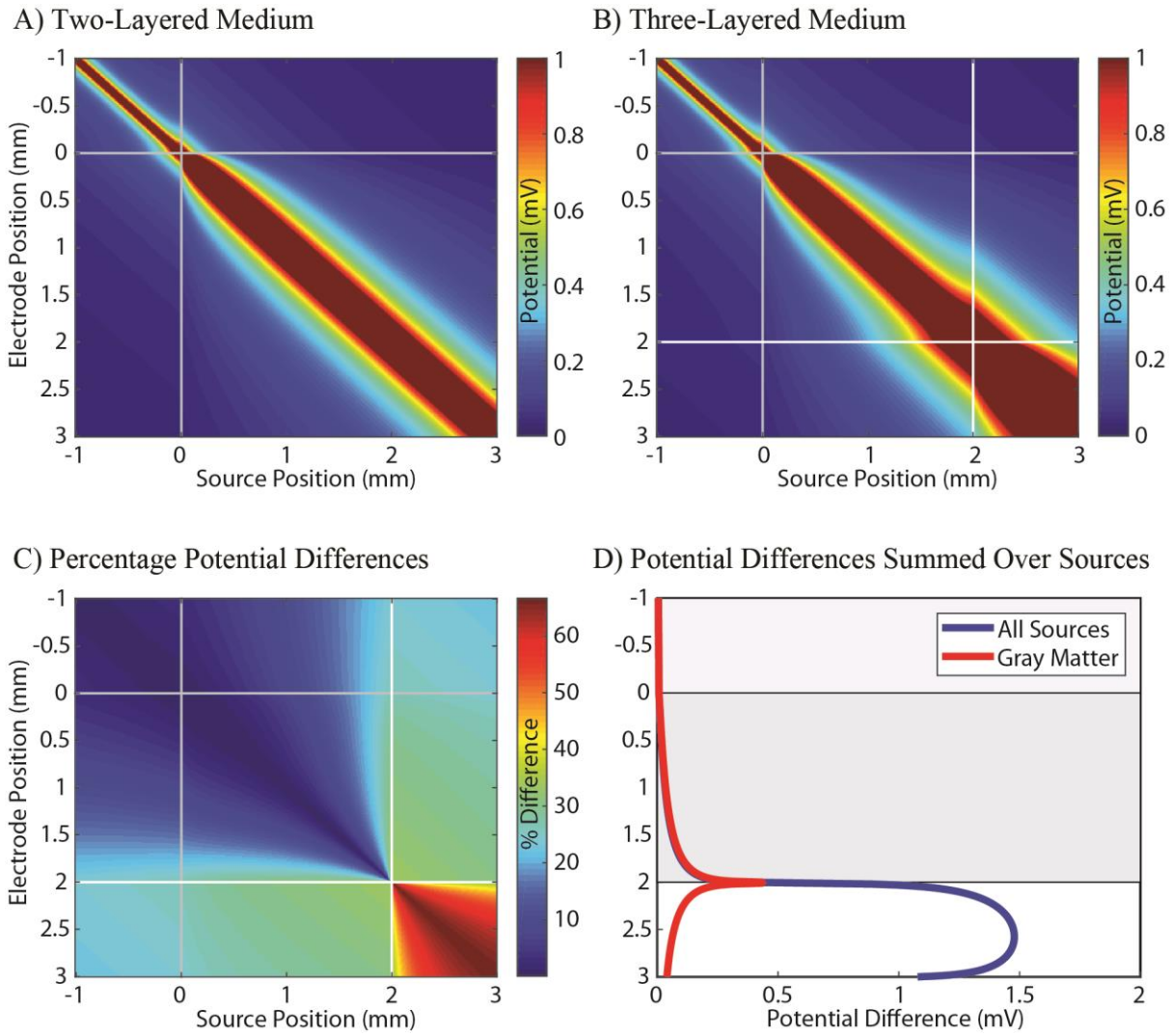

Figure 4-9: Effect of medium parameterization on potential. A) and B) show the potential in depth generated by unit strength point sources located at various depths in a 2- and 3-layered medium, respectively. C) Percentage difference between the 2- and 3-layered medium. As expected, the largest error is found for sources positioned in the bottom layer. D) Potential difference averaged over sources in the gray matter (red curve) or over all sources (blue curve).

respectively. The color scale in the two panels was clipped at a potential of 1 mV to facilitate the comparison. To highlight the discrepancies between the two medium parameterizations, Figure 4-9C shows the percentage difference between the potential in the two media. We see that adding the white matter mainly affects the potential generated by the sources located within the white matter and to a lesser extent the potential generated at the boundary between the gray and white matter by sources located within CSF or gray matter. This border effect is emphasized further in

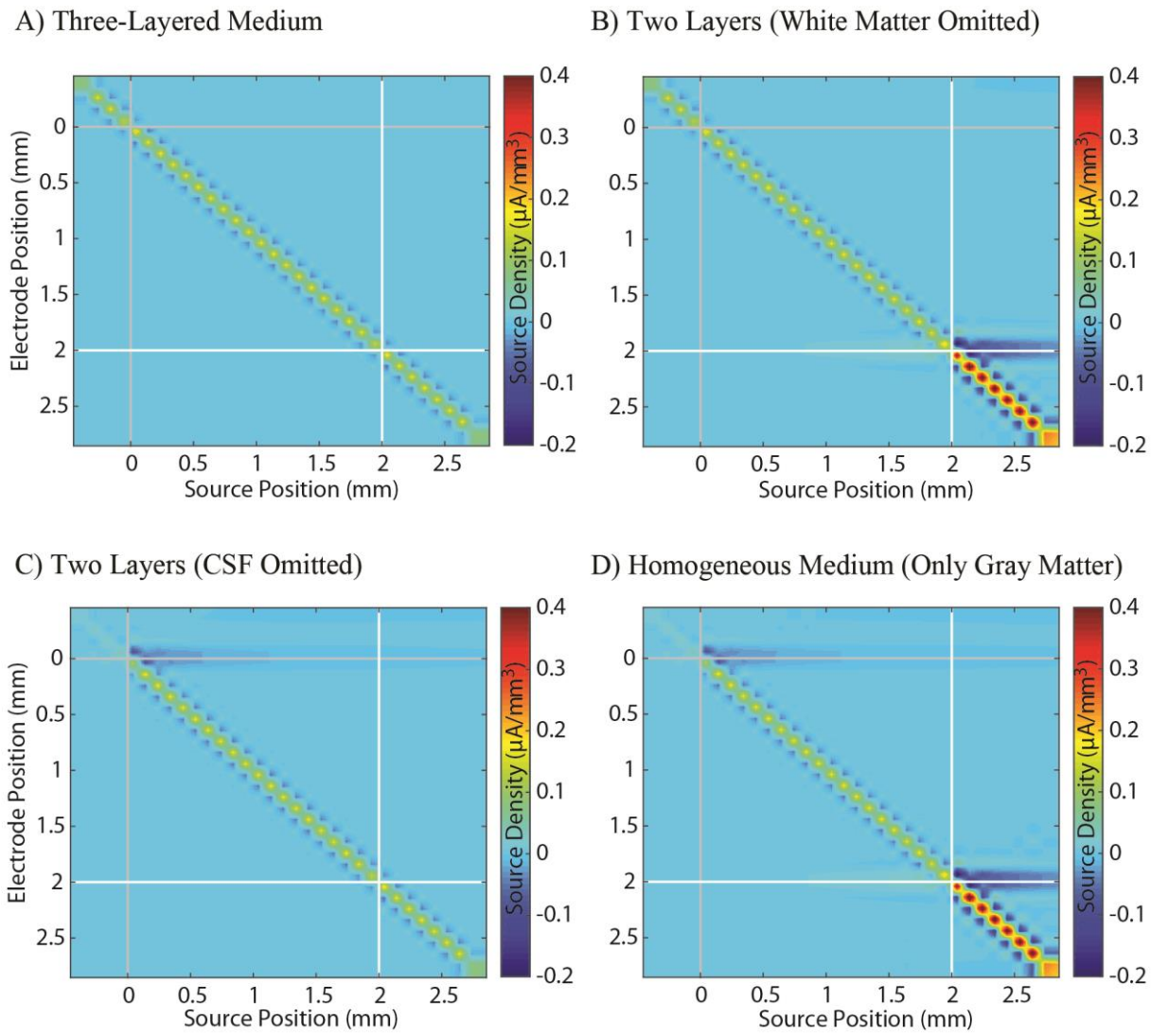

Figure 4-10: Error introduced by a mismatch between forward and inverse medium parameterization assuming unit strength uniform source disks with a diameter of 2 mm. For all panels, the potential is generated using a 3-layered medium with conductivities of $1.7, 0.3, 0.1\ S/m$. In each case the rCSD source parameterization without regularization was used. A) Matched 3-layer estimation. We see that the sources are estimated at the correct positions and with minimal spreading. B) and C) shows the estimations using a 2-layered medium omitting the bottom (white matter) or top (CSF) layer, respectively. In both cases we see a boundary artifact at the omitted interface. Moreover, the sources within the omitted medium are found to be wrongly scaled. D) Estimation using a homogeneous medium. Again, clear boundary artifacts are observed. We also note that the artifacts generated at the boundaries by sources located within the gray matter are of opposite sign and more significant at the top than at the bottom interface. This is explained by the size and polarity of the jump discontinuity in conductivity.

Figure 4-9D which shows the average potential difference over all source positions (blue line) as well as the average over sources within the gray matter (red line). It is very clear that the

discrepancies between the two media disproportionally affect the white matter as well as the boundary between white and gray matter. Hence, it can be expected that the same will be true for the source estimation accuracy.

In Figures 4-3 and 4-6 we already observed that estimating sources assuming a 2-layered forward model when the potential was generated using a 3-layered medium introduces errors at the transition between gray and white matter. This is further highlighted in Figure 4-10 which shows the rCSD estimation of unit strength uniform source discs of 2 mm diameter located at various positions within a three-layered cortical medium. The recording electrode is again assumed to be a 32 channel electrode with $100\ \mu m$ channel separation. The fifth channel is positioned $50\ \mu m$ below the border between CSF to gray matter. The gray matter is 2 mm thick. Panel 4-10A shows the estimation using the correct three-layered forward model, demonstrating that all sources are equally well reconstructed. Moreover, we can observe the boundary effect at the edge of the electrode (i.e. sources located above -0.35 and below 2.75 mm) introduced by the lack of data beyond the edge. In contrast, when the white matter is ignored in the estimation (Fig. 4-10B), a spurious source is estimated at the boundary between gray and white matter. This border effect is most pronounced for sources deeper than $\sim 750\ \mu m$. Moreover, sources within the white matter are amplified in comparison to the gray matter and CSF sources. A similar phenomenon can be observed when instead the CSF layer is ignored (see Fig. 4-10C). However, in this case the source at the border induced by sources within the gray matter is negative and of higher amplitude than in Figure 4-10B. This difference in polarity and amplitude is expected since the jump in conductivity is larger and of opposite polarity at the boundary between CSF and gray matter. Hence, neglecting the gray to white matter boundary introduces less error than omitting the conductivity jump between CSF to gray matter. It is however important to note that the size of the error introduced at the boundary always depends on the spatial profile of the sources. In any given situation, the estimated CSD is given by the summed profile shown here weighted by the source intensity at each depth. Hence, because the strength of the border artifact varies as a function of depth, its summed size will depend on the spatial source profile. Finally, for comparison, Panel 4-10D shows estimation assuming a homogeneous medium with a conductivity of $\sigma = 0.3\ S/m$ (i.e. same conductivity everywhere). As expected, the result is an overlay of the boundary and intensity effects observed in Panel 4-10B and 4-10C.

4.4 Discussion

The standard CSD estimation method has a long history in neurophysiology. However, in an experimental setting, discrete estimation of derivatives is afflicted with certain mathematical limitations. In particular, derivative estimation is an unstable operation which heavily amplifies noise. Moreover, the standard CSD estimator is insensitive to jumps in electrical conductivity such as those encountered at the transition between CSF and gray matter, or between gray and white matter. In addition, 1-D standard CSD estimation is based on an assumption of large lateral source extension, a situation which isn't necessarily fulfilled in all experimental settings. In order to address some of these limitations, estimation methods based on inverse theory have been introduced (Pettersen et al., 2006, Potworowski et al., 2012, Kropf and Shmuel, 2016). However, these methods have mainly aimed at dealing with restricted source diameters and up to now the relationship between inverse estimators and the standard method have not been thoroughly studied.

In this work, we have investigated the relationship between 1-D standard estimation and inverse CSD estimation in the presence of laterally extended sources. We have approached this by first showing that the 1-D Poisson equation can be transformed into an integral equation akin to the common inverse formulation. This novel formulation has allowed us to expand on Pettersen et al.'s findings by comparing the standard estimator to both the discrete (iCSD *delta*) and continuous inverse estimators. In particular, we found that the standard and inverse estimators converge to each other for lateral source cylinders of uniform density with a diameter exceeding ~2-5 mm. The range of this diameter dependence is mainly determined by the type and level of error considered. The smaller end of this range is found when looking at the discrepancies of the inner electrode *delta* estimators whereas the larger error reflects the discrepancy of the delta estimators at the edge electrodes. In the case of the continuous estimators, this range also denotes the region where the discrepancy between the InfPlane and UniCyl forward model drops considerably (Fig. 4-5). In line with these findings, we have also found that errors in the assumed source diameter affects estimation mainly at the edge electrodes, which can be explained by errors in the enforced boundary conditions. The interior electrode estimation was largely unaffected by errors in source diameter as long as the source diameter exceeds the large diameter threshold (~2-5 mm) mentioned above. This is explained by the fact

that sources far away from the electrode contribute approximately linear or constant functions to the measured potential. But as long as the diameter is large enough, linear and constant functions are in the nullspace of the estimator and thus do not contribute to the estimated CSD.

Despite the development of the inverse estimators, the standard CSD method remains popular in the field. This might be due to the fact that it appears easier to use. However, here we have shown that this ease of use is largely illusory because all the parameters required for the inverse methods are implicitly contained in the standard estimator. In particular, the source diameter which is often difficult to estimate in an experimental setting has been shown to mainly affect the edge estimation as mentioned above. Hence, not only does the inverse estimator for the standard method (the InfPlane estimator) converge to the uniform cylinder estimator (UniCyl) at large diameters, but both these estimators are more stable under noisy experimental conditions than the 5-point standard estimator (Fig. 4-8). Thus, except slightly reduced computation time, there appears to be no particular advantage in using the standard estimator over inverse estimators. As we have shown here, in presence of laterally extended sources, the standard method is in fact equivalent to using the inverse methods but without the added benefits of continuous function estimation, flexibility in medium and boundary parameterization, and regularization.

4.4.1 How likely are extended sources?

The 1-D standard CSD estimation method has unquestionably contributed useful information about neural function as attested by the vast body of literature devoted to it (e.g. (Brette and Destexhe, 2012, Buzsaki et al., 2012)). It is thus interesting and important to ask about the reasons behind this success. One of the major assumption limiting its validity is the need for spatially extended sources. This is often interpreted as it being sensitive mainly to laterally extended dynamics. However, we have shown in our motivating example that CSD estimation is in fact sensitive to local dynamics as well. Moreover, we have shown that the estimation is quite insensitive to the real source diameter as long as the overall activation diameter exceeds ~2-5 mm. The success of the standard method could thus be explained by the fact that activation exceeding this threshold is the norm rather than the exception (see e.g. (Kajikawa and Schroeder, 2011)). Moreover, this view is in line with voltage sensitive dye recordings which show that sensory activation often leads to spatially extended waves of activity

within cortex even with spatially localized stimuli (see e.g. (Ferezou et al., 2006, Jin et al., 2008)). This could thus possibly correspond to an overlay of a global activation superposed over a local activation as presented in our motivating example.

4.4.2 Two- versus three-layered medium

We have shown that in the case of cortical recordings, the CSD estimation errors introduced by omitting the conductivity difference between gray and white matter are small. In particular, assuming that the white matter is completely source free, the estimation errors are concentrated around the location of the boundary. The reason why this error is smaller than the one introduced by omitting the discontinuity in conductivity between CSF and gray matter is entirely based on the physics of the situation. The magnitude of the error is directly related to the size of the conductivity jump. Hence, the smaller difference between gray and white matter leads to a smaller induced source density at this boundary. In practice, we have found that this error along with the error introduced by omitting the CSF to gray matter conductivity discontinuity can even provide a useful cue in an experimental setting because it allows for an approximate localization of the gray matter.

5 Chapter 5: Discussion

The aim of this thesis was to investigate the application and methods of estimating current-source density from 1-dimensional electric potential recordings. To do so, we first showed the usefulness of CSD estimation in the study of cortical processing by using the standard CSD approach to study orientation processing in cat cortical area 18 (Chapter 2). I then proceeded to an in-depth analysis of the 1-D CSD estimation problem in the context of linear inverse theory with a special emphasis on higher-order regularization to deal with the inevitable recording noise in experimental settings (Chapter 3). Finally, I extended the inverse formalism to the standard method to investigate the relationship between the two estimation frameworks in the presence of laterally extended sources (Chapter 4). In this last chapter, I will provide a short summary of the results followed by a discussion about the limitations of this work. I then conclude the thesis by proposing some future directions that could address some the discussed challenges.

In Chapter 2 we demonstrated that laminar CSD estimation is a potent tool for investigating cortical processing. Using the standard CSD method, we were able to show for the first time that CSD is sensitive to the neuronal processing involved in signaling the orientation of grating stimuli. Through the comparison of CSD responses elicited by optimally and non-optimally oriented grating stimuli, we found that excitatory activity is relayed from the input layer IV to the superficial layers II/III only in the optimal condition. This finding was replicated with lower contrast gratings suggesting that it was not a by-product of activation strength but rather intrinsically related to the processing of the stimulus. We thus proposed that the coupling between layer IV cells and layer II/III cells is functionally gated. To study the mechanisms governing this putative gate, we also investigated the CSD responses to a checkerboard stimulus and found that the sink extended to superficial layers in a similar manner as in the optimally oriented condition. Hence, we proposed that, in accordance with the study of Bosking et al., activity is only relayed to layer II/III when the feed-forward stream from layer IV is co-activated with a secondary stream (Bosking et al., 1997). In the case of cat area 18, this secondary stream could be generated by intra-cortical horizontal pathways and/or by signals from area 17. As discussed in Appendix A, both types of connections are known to preferably link neurons with matched orientation preference and could thus explain the orientation specificity observed in our

results. Moreover, this interpretation is consistent with the responses to checkerboard stimulus since a non-oriented stimulus is expected to evenly activate the cortical sheet and thus would also generate a sufficiently strong secondary stream to open the putative gate. Although this interpretation is consistent with the observed data, it is important to mention that it entirely depends on the assumption that standard 1-D CSD estimation is an appropriate tool in the context of laterally non-uniform cortical activation. This assumption was however far from trivial since the patchy cortical activation elicited by an oriented stimulus violates the in-plane homogeneity assumption underlying the standard CSD method. In order to address this potential limitation, we have validated the application of standard CSD to a full-field oriented stimulus on simulated data. In particular, we generated an orientation map similar to the one observed in cat area 18 and simulated a potential recording assuming an orientation-specific sink in superficial layers overlaid with a broadly amplitude-modulated dipolar activation of layer IV. We found that in such a situation, the standard CSD was appropriate and sufficiently sensitive to pick up the orientation selective component that we introduced in the simulated signal. We thus concluded that it was unlikely that the orientation-specific CSD responses observed in the experimental setting are artifacts of employing the standard CSD estimation method.

Although we showed that the standard CSD method is useful in the study of cortical orientation processing, the direct derivative estimation procedure used in it does not optimally account for the inverse nature of CSD estimation. Therefore, we decided to conduct an in-depth analysis of the 1-D CSD estimation problem in the context of linear inverse theory with a special emphasis on higher-order regularization. In Chapter 3, we presented a 1-D CSD framework that unifies previous inverse methods (iCSD and kCSD) under a more general umbrella of function expansion methods. We added to that another expansion method which we called rCSD as well as a method based on quadrature (qCSD). We then supplemented these estimation schemes with various spectral regularization method to deal with the inevitable recording noise in experimental settings and tested them using a range of source diameters, noise levels and CSD depth profiles. We found that although the best estimation scheme is problem dependent, a reproducible ranking can be established by ordering the schemes according to the average source reconstruction error over all tested source diameters and noise levels. In addition to providing a ready-to-use toolbox for 1-D inverse CSD estimation, this study also extensively focused on the challenges and

stringent assumptions involved in solving the 1-D CSD inverse problem. This is an important but often underrepresented aspect in discussions about CSD estimation despite the fact that the difficulties and particularities of solving inverse problems are often confounding even to specialists (Wing and Zahrt, 1991).

Although the framework presented in Chapter 3 successfully united all previously published inverse methods under a common formulation, it does not provide a clear connection between the standard and inverse estimation methods. We addressed this shortcoming in Chapter 4 by comparing the standard and inverse CSD estimation with a focus on laterally extended sources. By expressing the standard method as an inverse problem using the Green's function for the 1-D Poisson equation, we show that the two formulations converge to each other and could be considered similar for lateral source diameters exceeding ~2-5 mm. Moreover, we found that errors in the assumed source diameter affected estimation mainly at the edge electrodes, which can be attributed to errors in the enforced boundary conditions. In contrast, the interior electrode estimation was found to be largely unaffected by errors in source diameter as long as the lateral source diameter exceeds the large diameter threshold (~2-5 mm). Nevertheless, despite similar estimation at inner electrodes, the standard 5-point CSD estimation was found to be considerably less resistant to recording noise. We thus concluded that, with the exception of slightly reduced computation time, there appears to be no particular advantage in using the standard estimator over inverse estimators.

5.1 Challenges and Limitations

The validation of the regularization framework presented in Chapter 3 as well as its comparison with the standard CSD method (Chapter 4) have been performed on simulated data. This has the clear advantage of providing clear measures for evaluating the estimation accuracy and thus providing a well-controlled testing environment for the various estimation schemes. This is crucial since, as discussed in the previous section, inverse problems are extremely difficult to solve. However, it also means that certain particularities of dealing with experimental data are left unaddressed. It is thus important to discuss the practical issues and potential limitations that could be encountered when estimating CSD in an experimental context.

5.1.1 Dealing with multiple time points

A first question arising when dealing with experimental data is how to deal with the temporal dimension of the data. In the context of volume conductor theory, the temporal and spatial dimensions are decoupled and thus CSD is estimated separately for each time point. This means that a regularization coefficient needs to be determined for every sample. But we have shown in Chapter 3 that the resolution of the estimation depends on the value of the regularization coefficient. Hence, the interpretation of the data will be complicated if automatic regularization parameter selection methods return markedly different values for different time points due to, for example, changes in the structure of the noise. There are essentially two ways in which this difficulty can be handled. First, instead of defining a regularization coefficient for every time point, one might choose a single one for the entire window of interest. This option is already included in the current version of the toolbox. Alternatively, it is possible to define the regularization problem in a two-dimensional way and include assumptions about the expected temporal dynamics of the signal. Especially if the data is oversampled as is usually the case, it would be perfectly reasonable to impose a temporal smoothness prior analogous to the spatial smoothness priors discussed in Chapter 3.

It is important to note that the points raised above are not unique to dealing with multiple time points but are also encountered when comparing different stimulation conditions. For example, the discussed issues with point-wise automatic regularization coefficient selection would be encountered when using the inverse framework to reanalyze the data from Chapter 2. Since we compare the spatial extent along the cortical depth of particular CSD domains for different orientations and contrasts, it is crucial that the spatial resolution of the estimate remains comparable over conditions. However, the fixed resolution assumption would be violated if the regularization coefficient selection method returned different regularization parameters for different conditions. We would thus need to enforce constraints in the form mentioned above to deal with this data.

5.1.2 Dealing with multiple trials

Another commonly encountered question is whether it is preferable to average the potential over multiple trials and then estimate the CSD or to estimate the CSD for every trial and average the estimates. In the case of the standard CSD method this does not matter because

the estimation procedure is linear. But in the case of the inverse methods the order of operations is important because finding the regularization coefficient introduces a non-linear step. Only for a given regularization coefficient is the problem linear. So which of the two approaches should be preferred?

An argument in support of both strategies can be made. The higher the noise, the more regularization will be required to stabilize the estimate. This means that the average of CSD estimates obtained from different trials will have a larger bias than if a single estimate is found from the averaged potential. This is simply due to the fact that averaging the potential over N trials reduces the noise amplitude by $1/\sqrt{N}$, if we assume that the noise is additive and has zero mean. On the other hand, for statistical purposes it might be interesting to have multiple CSD estimates to get a measure of the variance of the estimate. In this case, it would thus be necessary to avoid averaging the potential. In any case, the benefit of averaging the potential is also heavily dependent on the quality of the data. If the signal is already very clean it is clearly better to estimate every trial separately because only very little regularization will be required.

5.1.3 Channel-Dependent and/or Non-White Noise

The two last experimental challenges that I will discuss here are the potential variation of noise levels across channels and the possibly non-white distribution of noise. Due to the fact that thermal noise varies with the square root of the impedance, variations in the impedance across electrode contacts will lead to different noise levels on each of them (the thermal noise is $\sim 10\ \mu V$ for a contact with an impedance of 1 MΩ and a 10 kHz bandwidth). These variations will in turn introduce systematic errors in both the standard and the regularized CSD methods. There are again two ways to deal with this.

One possibility is to equalize channel impedances prior to data acquisition through a process called electrode activation. With the commonly employed iridium electrodes, it is possible to form a hydrous oxide on the surface of the electrode by running a square-wave cyclic voltammetry protocol while the electrode is dipped into $0.3\ M$ Na2HPO4 (Robblee et al., 1983)(see also tech notes from NeuroNexus). This allows to reduce the channel impedance by increasing the surface area of the contact. By forming different thicknesses of hydrous oxide on each electrode contact, their impedances can be equalized. This approach also has the side benefit that activation increases the charge capacity of the electrode. Hence, if micro-stimulation

or electrolytic lesioning is performed during the experiment, higher levels of current can be delivered without damaging the electrode.

Another possibility is to account for the different noise levels during regularization. We will discuss it in the larger context of how non-white noise affects CSD estimation because the solution is the same in both cases. But in order to answer this question, we first need to define what we consider noise and where a potential non-white component could originate from.

So far, we have only considered the thermal electrode noise which is usually modelled as Gaussian white noise because it results from the random motion of charge carriers in a conductor (Barry et al., 2004). However, neural recordings are also contaminated by line noise and mostly non-white physiological noise such as heart beat artifacts, motion induced by breathing, temporal variations in temperature and blood flow, and many more. Most of these noise sources generate predominantly low frequency noise. This is problematic for regularized CSD estimation because spectral regularization dampens the contribution of the small singular values associated with rapidly oscillating (i.e. high spatial frequency) singular vectors. In other words, low frequency noise is harder to eliminate because it resembles a valid representation of the potential and thus affects estimation accuracy more heavily. A possible way to deal with non-white noise is to prewhiten the data by premultiplying both the potential and the forward matrix with the transposed inverse of the Cholesky factorization of the covariance matrix of the noise (Hansen and Society for Industrial and Applied Mathematics., 2010). This is however not as straightforward as it sounds because it is often quite difficult to get a good estimation of the covariance matrix of the noise. The only exception is when we are confronted with white noise with channel-dependent amplitude as discussed above. In this situation, the inverse covariance matrix is simply a diagonal matrix with the reciprocal of the channel-specific noise variance as entries. Hence, the whole prewhitening process is simply rescaling the problem with the reciprocal of the noise variance. Finally, it is important to note that prewhitening does not lower the obtainable error but only creates a problem with a flat spectrum (Hansen and Society for Industrial and Applied Mathematics., 2010). But this is still useful because it has the advantage of making the problem more accessible to regularization parameter selection method based on the assumption of white noise (e.g. NCP or GCV).

5.1.4 Open Challenges

Up to this point, I have only discussed practical challenges for which possible solutions are known. There are however also important open questions which can significantly influence the validity of current CSD estimation methods. In this section I will focus on three of them: the accuracy of the volume conductor model, the accuracy of the forward model parameters, and the general validity of one-dimensional CSD estimation. However, since these have already been discussed in detail in the discussion of Chapter 3, I will only provide a short summary here and move on to propose potential future work that could address these issues.

The most significant challenge to the work presented in this thesis clearly is the question about the validity of the volume conductor model. As already discussed in the introduction, we have assumed throughout the thesis that the neural tissue can be modelled as purely resistive with piece-wise constant electrical conductivity. This is equivalent to considering the flow of ions in tissue akin to the flow of electrons in a metal. As a consequence, this model dictates that any active sink or source instantly generates a compensatory return current that assures that the total sources are always balanced (i.e. the sum of all sinks and sources is zero). This view has however recently been challenged on experimental grounds by Riera et al. who found monopolar components in current-source density in response to sensory stimulation (Riera et al., 2012). Similar results have been reported by Wagner et al. based on modelling studies where they found non-ohmic effects influencing the propagation of neurostimulation fields in the brain (Wagner et al., 2014). Finally, based on mean-field approximations to Maxwell's equations, Bedard and Destexhe have proposed that ionic diffusion (a non-ohmic effect) could significantly affect current flow in the brain, a view for which they have also provided interesting experimental evidence (Bedard and Destexhe, 2009, 2011, Destexhe and Bedard, 2012, Bedard and Destexhe, 2013, Gomes et al., 2016). These findings could have major implications for all source estimation methods because incorporating non-ohmic effects would significantly alter the problem formulation. In particular, it would not be possible anymore to decouple the spatial and temporal dimensions and it might become preferable to estimate sources in frequency space instead.

The second major open question is the extent to which errors in forward model parameters affect CSD estimation. We have shown in Chapter 3 that already for different spatial CSD

profiles, the optimal estimation methods can vary even when matched forward and inverse model parameters are used. Hence, at the moment we can only guess how errors in the forward model, e.g. error in electrode placement, change in the lateral distribution of sources, or errors in the source diameter, affect estimation accuracy and which estimation scheme is most stable under such perturbations. This problem is particularly difficult to address because the error sources potentially affect each other and the errors are also likely to be state dependent. We have already shown some state-dependency of the error in Chapter 4 where we found that an error in lateral source diameter is increasingly problematic the smaller the true diameter is. Hence, assuming that the current resistive forward model is found to be correct, it is still a complicated task to quantify the sensitivity of regularized CSD estimation to errors in the forward model parameters.

The last open question that I will discuss here is the general applicability of 1-D CSD estimation. To my knowledge, it has never been experimentally verified that either the Gaussian or the Uniform Cylindrical lateral source profiles are valid approximations to the general source distributions in experimental conditions. It has only been shown that the 1-D standard CSD formulation is appropriate in the presence of laterally extended sources (Mitzdorf, 1985). Moreover, we have shown in Chapter 4 that the standard CSD and the inverse models converge to each other at diameters exceeding ~2-5 mm. But it remains to be shown that the dimensionality reduction from a 3-D to a 1-D forward model using a cylinder or Gaussian of fixed lateral extent faithfully represents the sources in an experimental setting.

5.2 Future work

In this last section, I will now outline potential future approaches to deal with the three major open points raised in the previous section, namely testing the accuracy of the volume conductor model, addressing the sensitivity of CSD estimation to errors in the forward model parameters, and assessing the applicability of 1-D CSD estimation.

A particular challenge in finding the correct volume conductor model is that determining the impedance of the medium is heavily dependent on the interface between the tissue and the recording electrode (Logothetis et al., 2007, Gomes et al., 2016). In fact, Bedard and Destexhe have shown that it might be possible to account for the discrepancies in findings about brain tissue impedance only by looking at the method with which the recordings were done (Bedard and Destexhe, 2009). In particular, it is highly probable that the non-ohmic nature of the

extracellular medium only becomes apparent under physiological neuronal activation (Gomes et al., 2016). Hence, I propose as a potential future project that it might be possible to test the importance of non-ohmic effects for current-source density estimation by combining optogenetic neural stimulation with voltage-sensitive dye recordings. In particular, I propose to insert a tapered optical fiber into neural tissue while recording extracellular potential with an electrode and membrane potential using voltage sensitive dyes. This setup would have the advantage of preserving the cellular integrity while eliciting physiologically relevant membrane currents through the activation of light-sensitive ionic channels. Moreover, such an experiment could be conducted in both *in vitro* and *in vivo* conditions. Since it has been shown that voltage-sensitive dye activity is well correlated with membrane potential fluctuations, its recording would provide a meaningful measure of the trans-membrane potential generated in response to optical stimulation (Ferezou et al., 2006). By comparing the optically induced trans-membrane voltage changes with the extracellularly measured potential, we could then estimate the filtering properties of the extracellular space and possibly uncover non-ohmic effects.

To address the second open challenge, I propose to evaluate the sensitivity of CSD estimation to errors in the forward model parameters using a Bayesian framework. Such an approach would have the advantage of allowing an easy incorporation of priors on the form of the forward model into the estimation process. Moreover, given a set of noisy data and priors on the form of the CSD depth profile, a Bayesian framework would also provide the likelihood of different forward models. Hence, this formulation would be ideally suited to evaluate the sensitivity of the estimation to the forward model parameters. Also, expressing CSD estimation in this form would not even require extensive effort because the mathematical formulation of inverse theory in a Bayesian framework has already been well developed (see e.g. (Tarantola, 2005)).

Lastly, I propose that the applicability of 1-D CSD estimation could be tested by comparing it directly to 3-D estimation. Recently, Riera et al. have published a set of technically impressive results using novel 3-D multi-shank electrode arrays (Riera et al., 2014). This novel technology now makes it possible to experimentally verify some of the assumptions behind 1-D CSD estimation. In particular, by comparing the volumetric CSD obtained from a 3-D recording

with the 1-D estimations from the individual electrode shanks, it should be possible to find the best fitting lateral source parameterization to use in the 1-D forward model.

In conclusion, although 1-D CSD estimation has clearly provided valuable contributions to the study of neural processing, there remains important open question to be addressed. In this thesis, I have laid the ground work to address some of these issues and it is my hope that through the potential future work outlined here CSD estimation can be established on even firmer ground.

Bibliography

Adorjan P, Levitt JB, Lund JS, Obermayer K (1999) A model for the intracortical origin of orientation preference and tuning in macaque striate cortex. Visual neuroscience 16:303-318.

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. Journal of neurophysiology 52:1106-1130.

Alonso JM (2002) Neural connections and receptive field properties in the primary visual cortex. The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry 8:443-456.

Alonso JM, Cudeiro J, Perez R, Gonzalez F, Acuna C (1993a) Influence of layer V of area 18 of the cat visual cortex on responses of cells in layer V of area 17 to stimuli of high velocity. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale 93:363-366.

Alonso JM, Cudeiro J, Perez R, Gonzalez F, Acuna C (1993b) Orientational influences of layer V of visual area 18 upon cells in layer V of area 17 in the cat cortex. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale 96:212-220.

Alonso JM, Martinez LM (1998) Functional connectivity between simple cells and complex cells in cat striate cortex. Nat Neurosci 1:395-403.

Anderson JC, da Costa NM, Martin KA (2009) The W cell pathway to cat primary visual cortex. The Journal of comparative neurology 516:20-35.

Aster RC, Thurber CH, Borchers B (2005) Parameter estimation and inverse problems. Amsterdam ; London: Elsevier Academic Press.

Barrera RG, Guzman O, Balaguer B (1978) Point-Charge in a 3-Dielectric Medium with Planar Interfaces. Am J Phys 46:1172-1179.

Barry JR, Lee EA, Messerschmitt DG, Lee EA (2004) Digital communication. Boston: Kluwer Academic Publishers.

Beaulieu C, Colonnier M (1985) A comparison of the number of neurons in individual laminae of cortical areas 17, 18 and posteromedial suprasylvian (PMLS) area in the cat. Brain research 339:166-170.

Bedard C, Destexhe A (2009) Macroscopic models of local field potentials and the apparent 1/f noise in brain activity. Biophys J 96:2589-2603.

Bedard C, Destexhe A (2011) Generalized theory for current-source-density analysis in brain tissue. Phys Rev E Stat Nonlin Soft Matter Phys 84:041909.

Bedard C, Destexhe A (2013) Reply to Gratiy et al. J Neurophysiol 109:1683.

Bedard C, Kroger H, Destexhe A (2004) Modeling extracellular field potentials and the frequency-filtering properties of extracellular space. Biophys J 86:1829-1842.

Bedard C, Kroger H, Destexhe A (2006a) Does the 1/f frequency scaling of brain signals reflect self-organized critical states? Phys Rev Lett 97:118102.

Bedard C, Kroger H, Destexhe A (2006b) Model of low-pass filtering of local field potentials in brain tissue. Phys Rev E Stat Nonlin Soft Matter Phys 73:051911.

Bednar JA (2009) Topographica: Building and Analyzing Map-Level Simulations from Python, C/C++, MATLAB, NEST, or NEURON Components. Front Neuroinform 3:8.

Ben-Yishai R, Bar-Or RL, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. Proceedings of the National Academy of Sciences of the United States of America 92:3844-3848.

Best J, Reuss S, Dinse HR (1986) Lamina-specific differences of visual latencies following photic stimulation in the cat striate cortex. Brain research 385:356-360.

Binzegger T, Douglas RJ, Martin KA (2004) A quantitative map of the circuit of cat primary visual cortex. J Neurosci 24:8441-8453.

Bishop PO, Kozak W, Vakkur GJ (1962) Some quantitative aspects of the cat's eye: axis and plane of reference, visual field co-ordinates and optics. The Journal of physiology 163:466-502.

Blasdel GG, Fitzpatrick D (1984) Physiological organization of layer 4 in macaque striate cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 4:880-895.

Bonhoeffer T, Grinvald A (1991) Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. Nature 353:429-431.

Bosking WH, Zhang Y, Schofield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J Neurosci 17:2112-2127.

Brainard DH (1997) The Psychophysics Toolbox. Spatial vision 10:433-436.

Brette R, Destexhe A (2012) Handbook of neural activity measurement. Cambridge: Cambridge University Press.

Buzsaki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407-420.

Callaway EM (1998) Local circuits in primary visual cortex of the macaque monkey. Annual review of neuroscience 21:47-74.

Cardin JA, Carlen M, Meletis K, Knoblich U, Zhang F, Deisseroth K, Tsai LH, Moore CI (2009) Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature 459:663-667.

Casanova C, Michaud Y, Morin C, McKinley PA, Molotchnikoff S (1992) Visual responsiveness and direction selectivity of cells in area 18 during local reversible inactivation of area 17 in cats. Visual neuroscience 9:581-593.

Chisum HJ, Mooser F, Fitzpatrick D (2003) Emergent properties of layer 2/3 neurons reflect the collinear arrangement of horizontal connections in tree shrew visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 23:2947-2960.

Chung S, Ferster D (1998) Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. Neuron 20:1177-1189.

Covic EN, Sherman SM (2011) Synaptic properties of connections between the primary and secondary auditory cortices in mice. Cereb Cortex 21:2425-2441.

Crook JM, Eysel UT (1992) GABA-induced inactivation of functionally characterized sites in cat visual cortex (area 18): effects on orientation tuning. The Journal of neuroscience : the official journal of the Society for Neuroscience 12:1816-1825.

Csicsvari J, Henze DA, Jamieson B, Harris KD, Sirota A, Bartho P, Wise KD, Buzsaki G (2003) Massively parallel recording of unit and local field potentials with silicon-based electrodes. J Neurophysiol 90:1314-1323.

De Pasquale R, Sherman SM (2011) Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:16494-16506.

DeAngelis GC, Ohzawa I, Freeman RD (1995) Receptive-field dynamics in the central visual pathways. Trends in neurosciences 18:451-458.

Destexhe A, Bedard C (2012) Do neurons generate monopolar current sources? J Neurophysiol 108:953-955.

Di S, Baumgartner C, Barth DS (1990) Laminar analysis of extracellular field potentials in rat vibrissa/barrel cortex. J Neurophysiol 63:832-840.

DiCarlo JJ, Lane JW, Hsiao SS, Johnson KO (1996) Marking microelectrode penetrations with fluorescent dyes. Journal of neuroscience methods 64:75-81.

Douglas RJ, Koch C, Mahowald M, Martin KA, Suarez HH (1995) Recurrent excitation in neocortical circuits. Science (New York, NY) 269:981-985.

Einevoll GT, Kayser C, Logothetis NK, Panzeri S (2013) Modelling and analysis of local field potentials for studying the function of cortical circuits. Nat Rev Neurosci 14:770-785.

Einstein G, Fitzpatrick D (1991) Distribution and morphology of area 17 neurons that project to the cat's extrastriate cortex. The Journal of comparative neurology 303:132-149.

Ellaway PH (1978) Cumulative sum technique and its application to the analysis of peristimulus time histograms. Electroencephalography and clinical neurophysiology 45:302-304.

Ferezou I, Bolea S, Petersen CC (2006) Visualizing the cortical representation of whisker touch: voltage-sensitive dye imaging in freely moving mice. Neuron 50:617-629.

Ferster D (1986) Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 6:1284-1301.

Ferster D (1988) Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 8:1172-1180.

Fingelkurts AA, Kahkonen S (2005) Functional connectivity in the brain--is it an elusive concept? Neuroscience and biobehavioral reviews 28:827-836.

Fitzpatrick D (1996) The functional organization of local circuits in visual cortex: insights from the study of tree shrew striate cortex. Cereb Cortex 6:329-341.

Freeman JA, Nicholson C (1975) Experimental optimization of current source-density technique for anuran cerebellum. J Neurophysiol 38:369-382.

Frisby JP, Stone JV (2010) Seeing : the computational approach to biological vision. Cambridge, Mass. ; London: MIT Press.

Gabriel C, Gabriel S, Corthout E (1996a) The dielectric properties of biological tissues: I. Literature survey. Phys Med Biol 41:2231-2249.

Gabriel C, Peyman A, Grant EH (2009) Electrical conductivity of tissue at frequencies below 1 MHz. Phys Med Biol 54:4863-4878.

Gabriel S, Lau RW, Gabriel C (1996b) The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. Phys Med Biol 41:2251-2269.

Gabriel S, Lau RW, Gabriel C (1996c) The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. Phys Med Biol 41:2271-2293.

Gilbert CD (1977) Laminar differences in receptive field properties of cells in cat primary visual cortex. The Journal of physiology 268:391-421.

Gilbert CD, Kelly JP (1975) The projections of cells in different layers of the cat's visual cortex. The Journal of comparative neurology 163:81-105.

Gilbert CD, Wiesel TN (1979) Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. Nature 280:120-125.

Gilbert CD, Wiesel TN (1983) Clustered intrinsic connections in cat visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 3:1116-1133.

Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 9:2432-2442.

Girardin CC, Martin KA (2009) Inactivation of lateral connections in cat area 17. The European journal of neuroscience 29:2092-2102.

Glabska H, Potworowski J, Leski S, Wojcik DK (2014) Independent components of neural activity carry information on individual populations. PLoS One 9:e105071.

Gold C, Henze DA, Koch C, Buzsaki G (2006) On the origin of the extracellular action potential waveform: A modeling study. J Neurophysiol 95:3113-3128.

Gomes JM, Bedard C, Valtcheva S, Nelson M, Khokhlova V, Pouget P, Venance L, Bal T, Destexhe A (2016) Intracellular Impedance Measurements Reveal Non-ohmic Properties of the Extracellular Medium around Neurons. Biophys J 110:234-246.

Goto T, Hatanaka R, Ogawa T, Sumiyoshi A, Riera J, Kawashima R (2010) An evaluation of the conductivity profile in the somatosensory barrel cortex of Wistar rats. J Neurophysiol 104:3388-3412.

Gratiy SL, Pettersen KH, Einevoll GT, Dale AM (2013) Pitfalls in the interpretation of multielectrode data: on the infeasibility of the neuronal current-source monopoles. J Neurophysiol 109:1681-1682.

Griffiths DJ (2013) Introduction to electrodynamics. Boston: Pearson.

Hang GB, Dan Y (2011) Asymmetric temporal integration of layer 4 and layer 2/3 inputs in visual cortex. Journal of neurophysiology 105:347-355.

Hansen PC (2007) Regularization tools version 4.0 for matlab 7.3. Numerical Algorithms 46:189-194.

Hansen PC, Society for Industrial and Applied Mathematics. (2010) Discrete inverse problems insight and algorithms. In: Fundamentals of algorithms, pp 1 online resource (xii, 213 p.) Philadelphia, Pa.: Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104).

Happel MF, Jeschke M, Ohl FW (2010) Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. The Journal of neuroscience : the official journal of the Society for Neuroscience 30:11114-11127.

Hawken MJ, Parker AJ (1984) Contrast sensitivity and orientation selectivity in lamina IV of the striate cortex of Old World monkeys. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale 54:367-372.

Hilgetag CC, Burns GA, O'Neill MA, Scannell JW, Young MP (2000) Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. Philosophical transactions of the Royal Society of London Series B, Biological sciences 355:91-110.

Hirsch JA, Alonso JM, Reid RC, Martinez LM (1998) Synaptic integration in striate cortical simple cells. The Journal of neuroscience : the official journal of the Society for Neuroscience 18:9517-9528.

Hoffmann KP, Stone J, Sherman SM (1972) Relay of receptive-field properties in dorsal lateral geniculate nucleus of the cat. Journal of neurophysiology 35:518-531.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology 160:106-154.

Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. The Journal of physiology 195:215-243.

Jackson JD (1999) Classical electrodynamics. New York ; Chichester: Wiley.

Jin JZ, Weng C, Yeh CI, Gordon JA, Ruthazer ES, Stryker MP, Swadlow HA, Alonso JM (2008) On and off domains of geniculate afferents in cat primary visual cortex. Nat Neurosci 11:88-94.

Kajikawa Y, Schroeder CE (2011) How local is the local field potential? Neuron 72:847-858.

Kajikawa Y, Schroeder CE (2015) Generation of field potentials and modulation of their dynamics through volume integration of cortical activity. J Neurophysiol 113:339-351.

Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M (2009) Local origin of field potentials in visual cortex. Neuron 61:35-41.

Kocsis B, Bragin A, Buzsaki G (1999) Interdependence of multiple theta generators in the hippocampus: a partial coherence analysis. The Journal of neuroscience : the official journal of the Society for Neuroscience 19:6200-6212.

Kropf P, Shmuel A (2016) 1D Current Source Density (CSD) Estimation in Inverse Theory: A Unified Framework for Higher-Order Spectral Regularization of Quadrature and Expansion-Type CSD Methods. Neural Comput 28:1305-1355.

Kropf PS, Villeneuve MY, Yao ZS, Neupane S, Shmuel A (2010) Visual motion processing across layers in cat area 18. In: Society for Neuroscience, vol. 74.17 San Diego: Online.

Larsson J (2007) Electromagnetics from a quasistatic perspective. Am J Phys 75:230-239.

Leski S, Linden H, Tetzlaff T, Pettersen KH, Einevoll GT (2013) Frequency dependence of signal power and spatial reach of the local field potential. PLoS Comput Biol 9:e1003137.

Leski S, Pettersen KH, Tunstall B, Einevoll GT, Gigg J, Wojcik DK (2011) Inverse current source density method in two dimensions: inferring neural activation from multielectrode recordings. Neuroinformatics 9:401-425.

Leski S, Wojcik DK, Tereszczuk J, Swiejkowski DA, Kublik E, Wrobel A (2007) Inverse current-source density method in 3D: reconstruction fidelity, boundary effects, and influence of distant sources. Neuroinformatics 5:207-222.

LeVay S, Gilbert CD (1976) Laminar patterns of geniculocortical projection in the cat. Brain research 113:1-19.

Li WC, Soffe SR, Roberts A (2004) A direct comparison of whole cell patch and sharp electrodes by simultaneous recording from single spinal neurons in frog tadpoles. J Neurophysiol 92:380-386.

Linden H, Pettersen KH, Einevoll GT (2010) Intrinsic dendritic filtering gives low-pass power spectra of local field potentials. J Comput Neurosci 29:423-444.

Linden H, Tetzlaff T, Potjans TC, Pettersen KH, Grun S, Diesmann M, Einevoll GT (2011) Modeling the spatial reach of the LFP. Neuron 72:859-872.

Liu X-D, Fedkiw RP, Kang M (2000) A boundary condition capturing method for Poisson's equation on irregular domains. Journal of computational Physics 160:151-178.

Logothetis NK, Kayser C, Oeltermann A (2007) In vivo measurement of cortical impedance spectrum in monkeys: implications for signal propagation. Neuron 55:809-823.

Maier A, Adams GK, Aura C, Leopold DA (2010) Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. Frontiers in systems neuroscience 4.

Maier A, Aura CJ, Leopold DA (2011) Infragranular sources of sustained local field potential responses in macaque primary visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:1971-1980.

Malonek D, Tootell RB, Grinvald A (1994) Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT. Proc Biol Sci 258:109-119.

Martinez-Conde S, Cudeiro J, Grieve KL, Rodriguez R, Rivadulla C, Acuna C (1999) Effects of feedback projections from area 18 layers 2/3 to area 17 layers 2/3 in the cat visual cortex. Journal of neurophysiology 82:2667-2675.

Martinez LM, Alonso JM (2001) Construction of complex receptive fields in cat primary visual cortex. Neuron 32:515-525.

Martinez LM, Wang Q, Reid RC, Pillai C, Alonso JM, Sommer FT, Hirsch JA (2005) Receptive field structure varies with layer in the primary visual cortex. Nature neuroscience 8:372-379.

Menke W (1989) Geophysical data analysis : discrete inverse theory. San Diego ; London: Academic Press.

Meyer G, Albus K (1981) Spiny stellates as cells of origin of association fibres from area 17 to area 18 in the cat's neocortex. Brain research 210:335-341.

Milstein J, Mormann F, Fried I, Koch C (2009) Neuronal shot noise and Brownian 1/f2 behavior in the local field potential. PLoS One 4:e4338.

Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol Rev 65:37-100.

Mitzdorf U, Singer W (1978) Prominent excitatory pathways in the cat visual cortex (A 17 and A 18): a current source density analysis of electrically evoked potentials. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale 33:371-394.

Mitzdorf U, Singer W (1980) Monocular activation of visual cortex in normal and monocularly deprived cats: an analysis of evoked potentials. J Physiol 304:203-220.

Movshon JA, Thompson ID, Tolhurst DJ (1978a) Receptive field organization of complex cells in the cat's striate cortex. The Journal of physiology 283:79-99.

Movshon JA, Thompson ID, Tolhurst DJ (1978b) Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. The Journal of physiology 283:101-120.

Ness TV, Chintaluri C, Potworowski J, Leski S, Glabska H, Wojcik DK, Einevoll GT (2015) Modelling and Analysis of Electrical Potentials Recorded in Microelectrode Arrays (MEAs). Neuroinformatics 13:403-426.

Newman EA (1980) Current source-density analysis of the b-wave of frog retina. J Neurophysiol 43:1355-1366.

Nicholson C (1973) Theoretical analysis of field potentials in anisotropic ensembles of neuronal elements. IEEE Trans Biomed Eng 20:278-288.

Nicholson C, Freeman JA (1975) Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. J Neurophysiol 38:356-368.

Nicholson C, Llinas R (1975) Real time current source-density analysis using multi-electrode array in cat cerebellum. Brain Res 100:418-424.

Nunez PL, Srinivasan R (2006) Electric fields of the brain : the neurophysics of EEG. Oxford ; New York: Oxford University Press.

Olver PJ, Shakiban C (2006) Applied linear algebra. Upper Saddle River, NJ: Prentice Hall.

Ouellette BG, Casanova C (2006) Overlapping visual response latency distributions in visual cortices and LP-pulvinar complex of the cat. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale 175:332-341.

Palmer LA, Davis TL (1981) Receptive-field structure in cat striate cortex. Journal of neurophysiology 46:260-276.

Parker RL (1977) Understanding Inverse Theory. Annual Review of Earth and Planetary Sciences 5:35-64.

Parker RL (1994) Geophysical inverse theory. Princeton: Princeton University Press.

Payne BR, Peters A (2002) The cat primary visual cortex. San Diego: Academic Press.

Pettersen KH, Devor A, Ulbert I, Dale AM, Einevoll GT (2006) Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. J Neurosci Methods 154:116-133.

Pitts W (1952) Investigations on synaptic transmission. In: Cybernetics, Trans 9th Conf Josiah Macy, New York, pp 159-162.

Plonsey R, Barr RC (2007) Bioelectricity : A Quantitative Approach. Boston, MA: Springer Science+Business Media LLC.

Potworowski J, Jakuczun W, Leski S, Wojcik D (2012) Kernel current source density method. Neural Comput 24:541-575.

Price DJ (1985) Patterns of cytochrome oxidase activity in areas 17, 18 and 19 of the visual cortex of cats and kittens. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale 58:125-133.

Price DJ, Zumbroich TJ (1989) Postnatal development of corticocortical efferents from area 17 in the cat's visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 9:600-613.

Raiguel SE, Xiao DK, Marcar VL, Orban GA (1999) Response latency of macaque area MT/V5 neurons and its relationship to stimulus parameters. Journal of neurophysiology 82:1944-1956.

Rappelsberger P, Pockberger H, Petsche H (1981) Current source density analysis: methods and application to simultaneously recorded field potentials of the rabbit's visual cortex. Pflugers Arch 389:159-170.

Reid RC, Alonso JM (1995) Specificity of monosynaptic connections from thalamus to visual cortex. Nature 378:281-284.

Reimann MW, Anastassiou CA, Perin R, Hill SL, Markram H, Koch C (2013) A biophysically detailed model of neocortical local field potentials predicts the critical role of active membrane currents. Neuron 79:375-390.

Reinoso-Suárez F (1961) Topographischer Hirnatlas der Katze für experimental-physiologische Untersuchungen. Darmstadt: E. Merck AG.

Riera JJ, Goto T, Kawashima R (2014) A methodology for fast assessments to the electrical activity of barrel fields in vivo: from population inputs to single unit outputs. Front Neural Circuits 8:4.

Riera JJ, Ogawa T, Goto T, Sumiyoshi A, Nonaka H, Evans A, Miyakawa H, Kawashima R (2012) Pitfalls in the dipolar model for the neocortical EEG sources. J Neurophysiol 108:956-975.

Ringach DL, Hawken MJ, Shapley R (1997) Dynamics of orientation tuning in macaque primary visual cortex. Nature 387:281-284.

Robblee LS, Lefko JL, Brummer SB (1983) Activated Ir - an Electrode Suitable for Reversible Charge Injection in Saline Solution. J Electrochem Soc 130:731-733.

Rockland KS, Lund JS (1982) Widespread periodic intrinsic connections in the tree shrew visual cortex. Science (New York, NY) 215:1532-1534.

Rockland KS, Lund JS (1983) Intrinsic laminar lattice connections in primate visual cortex. The Journal of comparative neurology 216:303-318.

Salin PA, Bullier J (1995) Corticocortical connections in the visual system: structure and function. Physiological reviews 75:107-154.

Salin PA, Kennedy H, Bullier J (1995) Spatial reciprocity of connections between areas 17 and 18 in the cat. Canadian journal of physiology and pharmacology 73:1339-1347.

Scannell JW, Blakemore C, Young MP (1995) Analysis of connectivity in the cat cerebral cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 15:1463-1483.

Schroeder CE, Mehta AD, Givre SJ (1998) A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. Cereb Cortex 8:575-592.

Shadlen MN, Movshon JA (1999) Synchrony unbound: a critical evaluation of the temporal binding hypothesis. Neuron 24:67-77, 111-125.

Shapley R, Hawken M, Xing D (2007) The dynamics of visual responses in the primary visual cortex. Prog Brain Res 165:21-32.

Sherman SM, Guillery RW (1998) On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". Proceedings of the National Academy of Sciences of the United States of America 95:7121-7126.

Shmuel A, Grinvald A (1996) Functional organization for direction of motion and its relationship to orientation maps in cat area 18. The Journal of neuroscience : the official journal of the Society for Neuroscience 16:6945-6964.

Shmuel A, Grinvald A (2000) Coexistence of linear zones and pinwheels within orientation maps in cat visual cortex. Proceedings of the National Academy of Sciences of the United States of America 97:5568-5573.

Shmuel A, Korman M, Sterkin A, Harel M, Ullman S, Malach R, Grinvald A (2005) Retinotopic axis specificity and selective clustering of feedback projections from V2 to V1 in the owl monkey. J Neurosci 25:2117-2131.

Somers DC, Nelson SB, Sur M (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. The Journal of neuroscience : the official journal of the Society for Neuroscience 15:5448-5465.

Sompolinsky H, Shapley R (1997) New perspectives on the mechanisms for orientation selectivity. Current opinion in neurobiology 7:514-522.

Sotero RC, Bortel A, Martinez-Cancino R, Neupane S, O'Connor P, Carbonell F, Shmuel A (2010) Anatomically-constrained effective connectivity among layers in

a cortical column modeled and estimated from local field potentials. J Integr Neurosci 9:355-379.

Stone J, Dreher B (1973) Projection of X- and Y-cells of the cat's lateral geniculate nucleus to areas 17 and 18 of visual cortex. Journal of neurophysiology 36:551-567.

Stoop R, Saase V, Wagner C, Stoop B, Stoop R (2013) Beyond scale-free small-world networks: cortical columns for quick brains. Physical review letters 110:108105.

Swadlow HA, Gusev AG, Bezdudnaya T (2002) Activation of a cortical column by a thalamocortical impulse. J Neurosci 22:7766-7773.

Swindale NV (2004) How different feature spaces may be represented in cortical maps. Network 15:217-242.

Swindale NV, Grinvald A, Shmuel A (2003) The spatial pattern of response magnitude and selectivity for orientation and direction in cat visual cortex. Cereb Cortex 13:225-238.

Symonds LL, Rosenquist AC (1984) Laminar origins of visual corticocortical connections in the cat. The Journal of comparative neurology 229:39-47.

Szymanski FD, Garcia-Lazaro JA, Schnupp JW (2009) Current source density profiles of stimulus-specific adaptation in rat auditory cortex. J Neurophysiol 102:1483-1490.

Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Tenke CE, Kayser J (2012) Generator localization by current source density (CSD): implications of volume conduction and field closure at intracranial and scalp resolutions. Clin Neurophysiol 123:2328-2345.

Tenke CE, Schroeder CE, Arezzo JC, Vaughan HG, Jr. (1993) Interpretation of high-resolution current source density profiles: a simulation of sublaminar contributions to the visual evoked potential. Exp Brain Res 94:183-192.

Thomson AM, West DC, Wang Y, Bannister AP (2002) Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2-5 of adult rat and

cat neocortex: triple intracellular recordings and biocytin labelling in vitro. Cereb Cortex 12:936-953.

Vaknin G, DiScenna PG, Teyler TJ (1988) A method for calculating current source density (CSD) analysis without resorting to recording sites outside the sampling volume. J Neurosci Methods 24:131-135.

van der Gucht E, Vandesande F, Arckens L (2001) Neurofilament protein: a selective marker for the architectonic parcellation of the visual cortex in adult cat brain. The Journal of comparative neurology 441:345-368.

Villeneuve MY, Casanova C (2003) On the use of isoflurane versus halothane in the study of visual response properties of single cells in the primary visual cortex. Journal of neuroscience methods 129:19-31.

Vogel CR (2002) Computational methods for inverse problems. Philadelphia: Society for Industrial and Applied Mathematics.

Vogels R, Orban GA (1990) How well do response changes of striate neurons signal differences in orientation: a study in the discriminating monkey. The Journal of neuroscience : the official journal of the Society for Neuroscience 10:3543-3558.

Wagner T, Eden U, Rushmore J, Russo CJ, Dipietro L, Fregni F, Simon S, Rotman S, Pitskel NB, Ramos-Estebanez C, Pascual-Leone A, Grodzinsky AJ, Zahn M, Valero-Cabre A (2014) Impact of brain tissue filtering on neurostimulation fields: a modeling study. Neuroimage 85 Pt 3:1048-1057.

Wagner TA, Zahn M, Grodzinsky AJ, Pascual-Leone A (2004) Three-dimensional head model simulation of transcranial magnetic stimulation. Ieee Transactions on Biomedical Engineering 51:1586-1594.

Weliky M, Bosking WH, Fitzpatrick D (1996) A systematic map of direction preference in primary visual cortex. Nature 379:725-728.

White EL, Keller A (1989) Cortical circuits : synaptic organization of the cerebral cortex--structure, function, and theory. Boston: Birkhäuser.

White LE, Fitzpatrick D (2007) Vision and cortical map development. Neuron 56:327-338.

Wing GM, Zahrt JD (1991) A primer on integral equations of the first kind : the problem of deconvolution and unfolding. Philadelphia: Society for Industrial and Applied Mathematics.

Xing D, Yeh CI, Shapley RM (2009) Spatial spread of the local field potential and its laminar variation in visual cortex. J Neurosci 29:11540-11549.

Yu J, Ferster D (2013) Functional coupling from simple to complex cells in the visually driven cortical circuit. J Neurosci 33:18855-18866.

Zanos TP, Mineault PJ, Pack CC (2011) Removal of spurious correlations between spikes and local field potentials. J Neurophysiol 105:474-486.

Appendices

A. Anatomical and functional properties of cat primary visual cortex

In this section, I summarize the anatomical and functional information of the projections from the retinal ganglion cells up to primary visual cortex in the cat. It includes the projection profile from both area 17 and 18 to higher order areas but does not include information on projections from higher to lower order areas.

From retina to cortex

The cat retina contains three major classes of retinal ganglion cells: α, β, γ, which are found in different relative abundances and give rise to three distinct streams of projection to the cortex. The α cells are the source of the Y-pathway. They account for only 5% of the total number of retinal ganglion cells. The β cells are the most abundant sub-type of retinal ganglion cells (55%) and they give rise to the X-pathway. Together, the X- and Y-pathway form the most reliable visually driven input to cortex. The remaining 40% of retinal ganglion cells are of diverse subtype but they are generally grouped under the umbrella term γ. They form the basis of the W-pathway whose function in visual processing is only poorly studied because they respond inconsistently to visual stimulation.

The X- and Y-pathway can also be distinguished on functional grounds. Due to their smaller receptive fields, β-cells are more selective to higher spatial frequencies (Payne and Peters, 2002). They also respond briskly to visual stimulation; they respond preferably to low temporal frequencies and are highly sensitive to contrast at mid-range spatial frequencies (Payne and Peters, 2002). In contrast, the Y-pathway responds preferably to low spatial but high temporal frequencies and is sensitive to contrast at those lower frequencies (Payne and Peters, 2002). By virtue of their non-linear sub-units, α-cells also respond to high spatial frequencies, but because of their large receptive-fields their sensitivity is associated with low spatial resolution (Payne and Peters, 2002). Under high contrast conditions, α-cells generate the dominant signal up to the limit of α-cell acuity (Payne and Peters, 2002). For low contrast stimuli, both α- and β-cells signal contrast to the brain and the Y and X cells are equally sensitive to their preferred spatial and temporal frequencies (Payne and Peters, 2002). However, β cells are

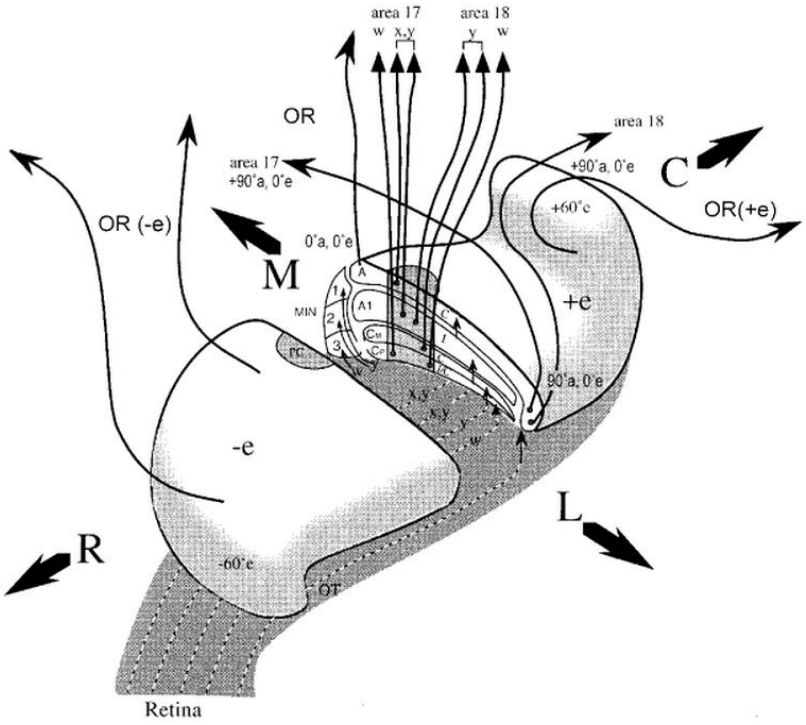

Figure A-1: View of the Lateral Geniculate nucleus and the adjacent MIN. The three parvocellular layers have been combined to a single one labeled C_P and layer C_M is equivalent to layer C1. Angles a and e stand for azimuth and elevation angles in the visual field respectively. OT = optic tract (retina to LGN), OR = optic radiation (LGN to cortex), C = caudal, R = rostral, M = medial, L = lateral. Adopted from (Payne and Peters, 2002)

much more abundant which confers a clear advantage to the X-system in signaling contrast to the brain even at low temporal and spatial frequencies (Payne and Peters, 2002).

From the retina, the major part of the connections project to the lateral geniculate nucleus (LGN) although part of the connections also innervates other sub-thalamic regions such as the medial interlaminar nucleus (MIN), the claustrum, the pulvinar and the superior colliculus.

Before reaching the LGN, the fibers originating in the nasal part of the retina of both eyes cross the midline in order to provide innervations solely from the contralateral visual field. Although the LGN receives input from both eyes, the input from each eye remains segregated in different layers (Payne and Peters, 2002). Figure A-1 shows the laminar pattern of the retinal input to both the LGN and the MIN.

As mentioned previously, we are mainly interested in the X- and Y-pathway because they provide the most reliable visually driven input to the cortex. Before continuing to describe the

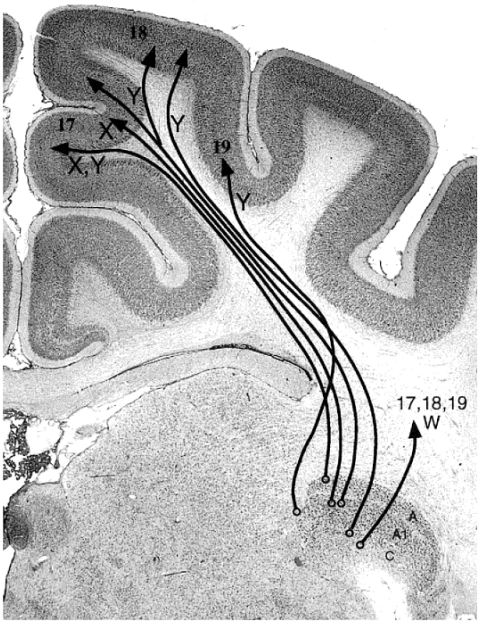

Figure A-2: Projections from LGN to cortex separated according to the different pathways. Adopted from (Payne and Peters, 2002)

innervations of area 17 and 18, there are a few interesting details that should be noted: 1) X-fibers only innervate layer A and A1 of the LGN. This means that the only X-input to cortex from the contralateral eye comes from the layer A of the LGN, which has been used to study the effect of the X-pathway on cortical responses. 2) X-fibers only innervate a single layer in the LGN while single Y-fibers may innervate both cells in layer C_M and in layer A. This amplifies the effect of the Y-pathway even though the α-cells are less abundant in the retina.

The LGN forms the most important relay of visual information between the retina and cortex. While Y-fibers innervate both area 17 and 18, X-fibers project only to area 17. As shown in Figure A-2, W-fibers also innervate area 17 and 18 as well as higher-order visual areas. Additionally, some Y-fibers from the MIN project to area 19 (Payne and Peters, 2002). Y-cells from layers A and A1 of the LGN probably innervate both area 17 and 18 through collateral branches while Y-cells from layer C probably innervate only area 18 (Payne and Peters, 2002).

Based on latency and current-source density studies, Mitzdorf proposed that there are three functionally distinct Y-pathways that project to the primary visual cortex (Mitzdorf, 1985). The one with the fastest conducting fibers innervates only area 18. A slower conducting pathway

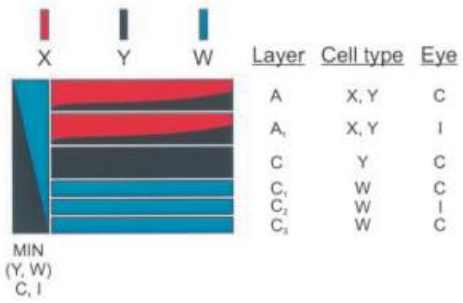

Figure A-3: Parallel projections emanating from cat LGN. Top-down is the proportions of cells found as a function of depth in the layer. Left-right shows the proportions as a function of eccentricity in the visual field with left being the most eccentric and right being the area centralis. We see that layer A and A1 have a bias towards a larger proportion of X cells near the area centralis. Adopted from (Alonso, 2002)

would preferably innervate area 17. Finally, a third also fast conducting pathway would innervate both area 17 and 18 but would exclusively target inhibitory neurons. The latter pathway is particularly interesting because it has been shown in vitro that fast-spiking inhibitory neurons play an important role in the generation of the cortical gamma oscillations (Cardin et al., 2009).

Within cortex

The pathway origin of innervation is mostly preserved in area 17 where the X-fibers preferably terminate in the lower part of layer 4, also called layer 4B, and the Y-fibers terminate in the upper part called layer 4A. This distinction is absent in area 18 since only Y-fibers innervate it. Therefore, layer 4 in area 17 shows a bilaminar structure while layer 4 in area 18 is unilaminar. In both areas, a portion of the LGN-projections terminate in layer 6A, but the layer 6 innervation is much more pronounced in area 17 probably due to a large innervation by X-fibers. However, even in area 17 the proportion of layer 4 and layer 6 innervation is in a ratio of 9:1 (Payne and Peters, 2002). Figure A-3 and Figure A-4 show a summary of the projections from the retina to the LGN and from the LGN to area 17. The projections to area 18 are similar except that the X-projections are replaced by Y. In addition, it has been shown recently that W-cells innervate mostly layer 1, 2/3 and 5 in area 17 (Anderson et al., 2009).

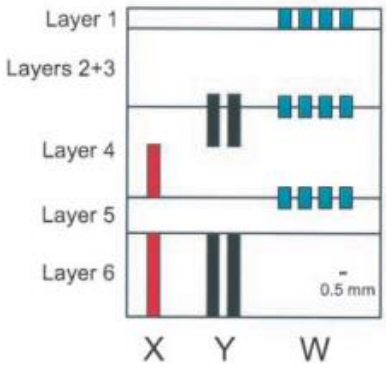

Figure A-4: Feed-forward projections from LGN to area 17. For simplicity the projections of certain X-cells to the entire layer 4 are omitted. Additionally, the graphics does not take into account the strength of the connections, e.g. the projection of X-cells to layer 4 is 9 times stronger than to layer 6. Top-down is the vertical extent of the projection in cortex, e.g. X-cells project to the bottom of layer 4. Left-right corresponds to the lateral spread of the projection. Finally, it has to be considered that the projections of the W pathway are based on relatively sparse data. Adopted from (Alonso, 2002)

Primary visual cortex is the first processing step where the visual input from both eyes is combined. The termination fields from fibers of the two eyes form interdigitated patterns, which gives rise to ocular dominance columns. In contrast to the monkey visual cortex, the ocular dominance columns are more diffuse in the cat. This is mainly due to the fact that innervations of the contralateral eye are spread out to a larger extent (Payne and Peters, 2002).

The connection profile within the areas is quite complicated and would necessitate the discussion of the structure of each lamina. Therefore, I will only describe the general pattern here (see Figure A-5 for details). A more detailed description can be found in (White and Keller, 1989, Payne and Peters, 2002, Binzegger et al., 2004).

Most of the data on intra-areal connectivity comes from anatomical studies on area 17. Due to the proposed similarity between area 17 and 18 and because of the lack of information, we can only assume a similar structure of intra-areal connections for both areas. The right panel of Figure A-5 shows the intrinsic circuitry of spiny cells in area 17. Spiny cells are different from smooth cells in that the former are excitatory (glutamergic) and the latter inhibitory (GABAergic) (White and Keller, 1989). There are two major classes of spiny cells: the spiny stellate and the pyramidal neurons, which are distinguished by the absence (existence) of a pronounced apical dendrite respectively. This distinction is of importance in our case because it

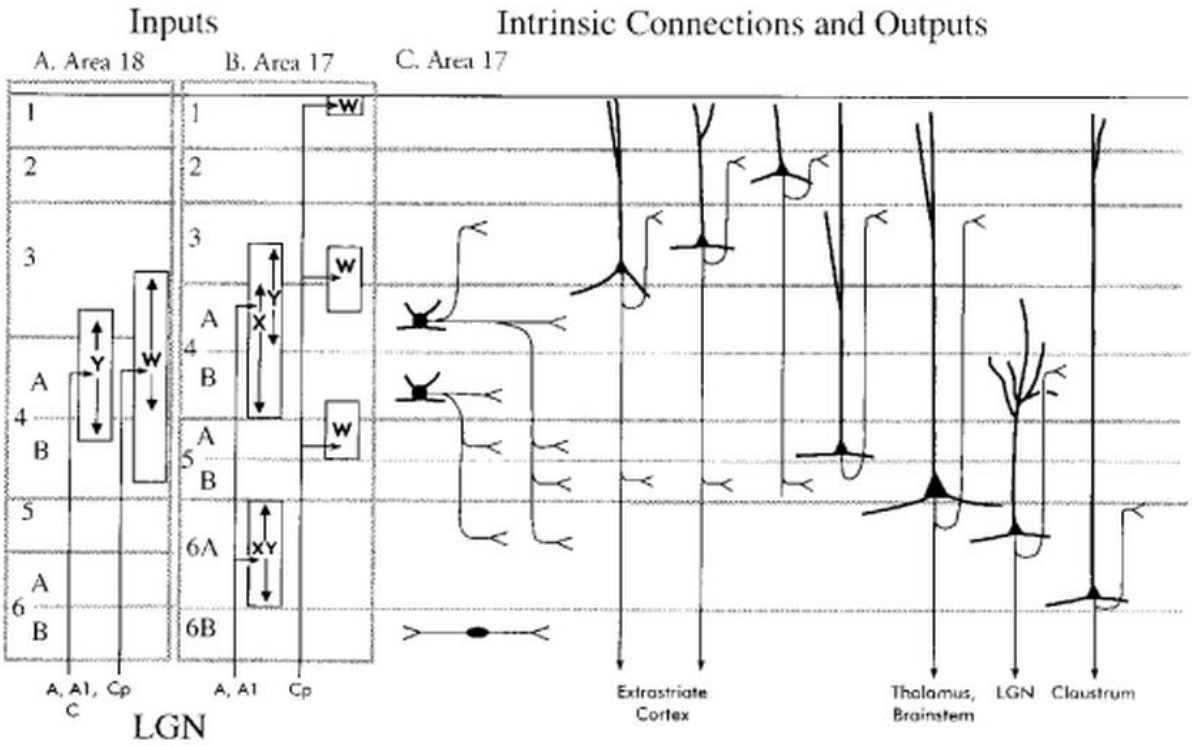

Figure A-5: Innervation profile of area 18 and 17 (Panel A and B respectively) along with the intrinsic circuitry of area 17 (Panel C). Adopted from (Payne and Peters, 2002)

affects the laminar localization of extracellularly measured potentials. Spiny stellates are mainly found in layer 4 where they form the main target of the LGN innervations to cortex. From layer 4, axonal projections go to both supragranular (layer 2/3) and infragranular (layers 5 and 6) layers. As shown in Figure A-5, the visual signal is then further processed within the area and also sent to other sub-cortical and cortical targets.

Although it is of great interest to consider the projection profile in order to explore the possible pathways of signal flow, this approach is limited by the fact that simply knowing the projection profile does not reflect the relative strength of these projections because the number of synaptic contacts might be very different for different projections. Binzegger et al. addressed this problem by making the assumption that the probability that a given neuron *i* makes a synapse with another neuron *j* in a given layer is proportional to amount of axonal membrane of neuron *i* and dendritic membrane of neuron *j* in this layer (Binzegger et al., 2004). This allowed them to come up with a graph of the intra-areal connections in cat area 17 and their respective weights. A major finding of their study was that the whole map of intra-areal cortical circuits is

characterized by "very few 'strong' but many 'weak' excitatory projections, each of which may involve only a few percentage of the total complement of excitatory synapses of a single neuron" (Binzegger et al., 2004). Until now, this is one of the most if not the most complete investigation of intra-areal anatomical connections conducted. Figure A-6 shows a modified version of the intra-areal circuit model omitting contributions from layer 1. The left column corresponds to excitatory populations and the right one to inhibitory populations within the specified layers, while the width of the arrows shows the connection strength. The connection strength is specified as the fraction of the total number of synapses formed by a given connection type. I.e. connections among the left column only involve synapses between excitatory cell pairs (there is an estimated total of $13.6 \cdot 10^{10}$ synapses between excitatory cells in cat area 17). On the other hand, the width of the arrows which go from the left to the right column represent the proportion of synapses between excitatory and inhibitory cell pairs (there is an estimated total of $2.1 \cdot 10^{10}$ synapses between excitatory and inhibitory cell pairs). In other words, the size of the arrows is not given as a percentage of the total number of synapses but only as a proportion of the synapses formed between the different cell types, i.e. excitatory-excitatory (within the left column), excitatory-inhibitory (from left to right), inhibitory-excitatory (from right to left) and inhibitory-inhibitory cell pairs (within the right column).

Finally, after considering the connections that lead to area 17 and 18 as well as the intra-areal circuitry, I will now discuss the connections between these two areas. It has been proposed that cortical neurons receive 3,000–10,000 synaptic inputs, of which 85% are excitatory. Nearly half of the excitatory input to any one neuron comes from nearby neurons that fall within a 100–200 um radius cylinder, arranged as a column. The remaining half can be traced mainly to a local network of horizontal connections originating in the nearby cortex, leaving a minority of inputs from more remote cortical and subcortical structures (Shadlen and Movshon, 1999). In fact, it has been proposed that thalamo-cortical projections only account for 5% of the total synapses of layer 4 (Payne and Peters, 2002). However, the number of synapses cannot necessarily be used as an indicator of the importance of the projection because it may diverge from the strength of the synapses. It is quite likely that the thalamo-cortical synapse is more potent than others and the same might be true for cortico-cortical projections.

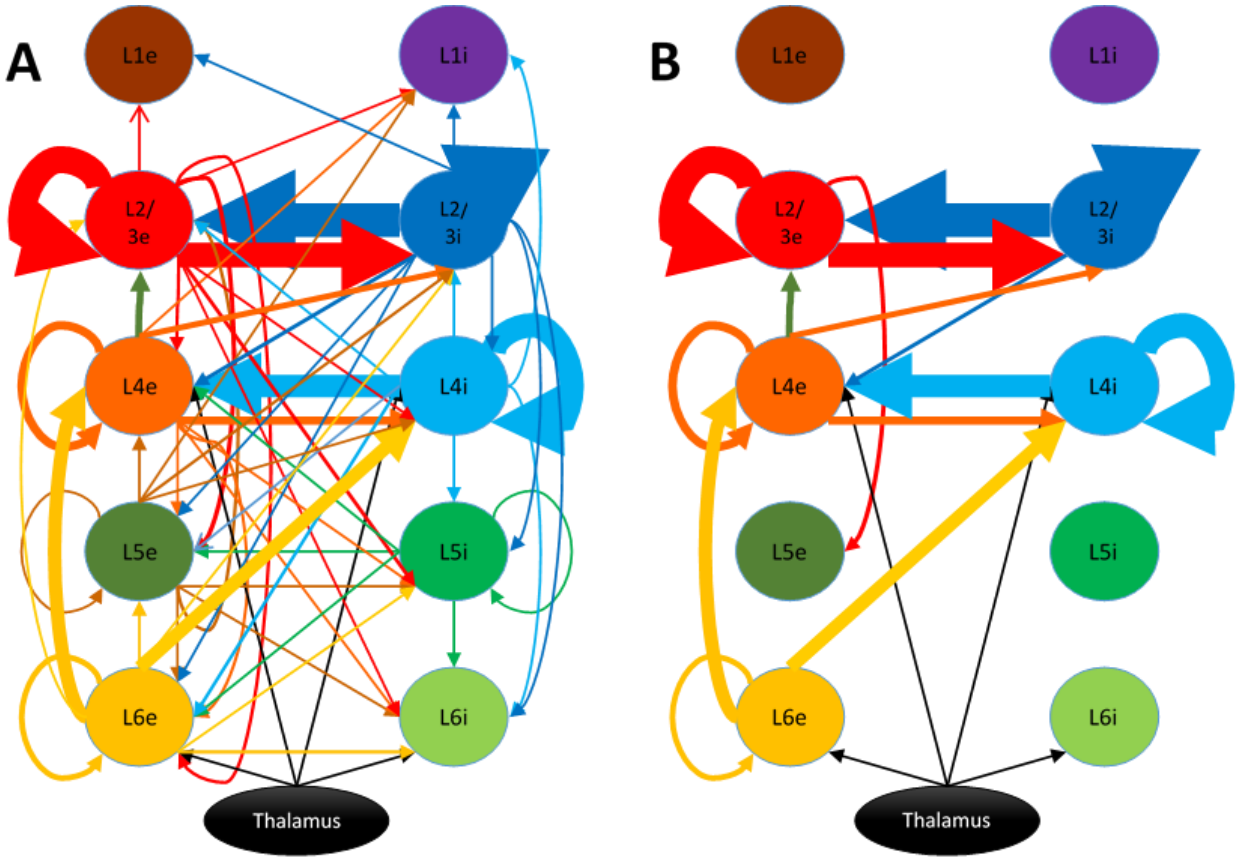

Figure A-6: Model of intra-areal projections using one excitatory and one inhibitory population per layer. This scheme forms the basis of the projected anatomical connection strength between populations. A shows the full model of the connections. B shows only the connections which contain more than 3% of the total number of synapses of the given connection type. This last rule was not applied to the connections from the thalamus to cortex which account for only 0.8% of the connections to excitatory cells and 0.7% of the connections to inhibitory cells. Adapted from (Binzegger et al., 2004, Sotero et al., 2010)

The major projection from area 17 to area 18 originates in the supragranular layers 2/3 of area 17 and terminates in the supragranular layer 2/3 and infragranular layers 5 and 6 of area 18 (Symonds and Rosenquist, 1984, Price and Zumbroich, 1989). Additionally, sparse projections from spiny stellates located in top of layer 4 and from neurons in layer 5 of area 17 have also been found to project to area 18 (Meyer and Albus, 1981, Einstein and Fitzpatrick, 1991). However, no mention of the termination field was made. Conversely, there are two main projections from area 18 to area 17: First, there is a projection from supragranular layer 2/3 of area 18 to supragranular layer 2/3 of area 17 (Symonds and Rosenquist, 1984, Martinez-Conde et al., 1999). However, it is not well studied whether these projections also send collaterals to

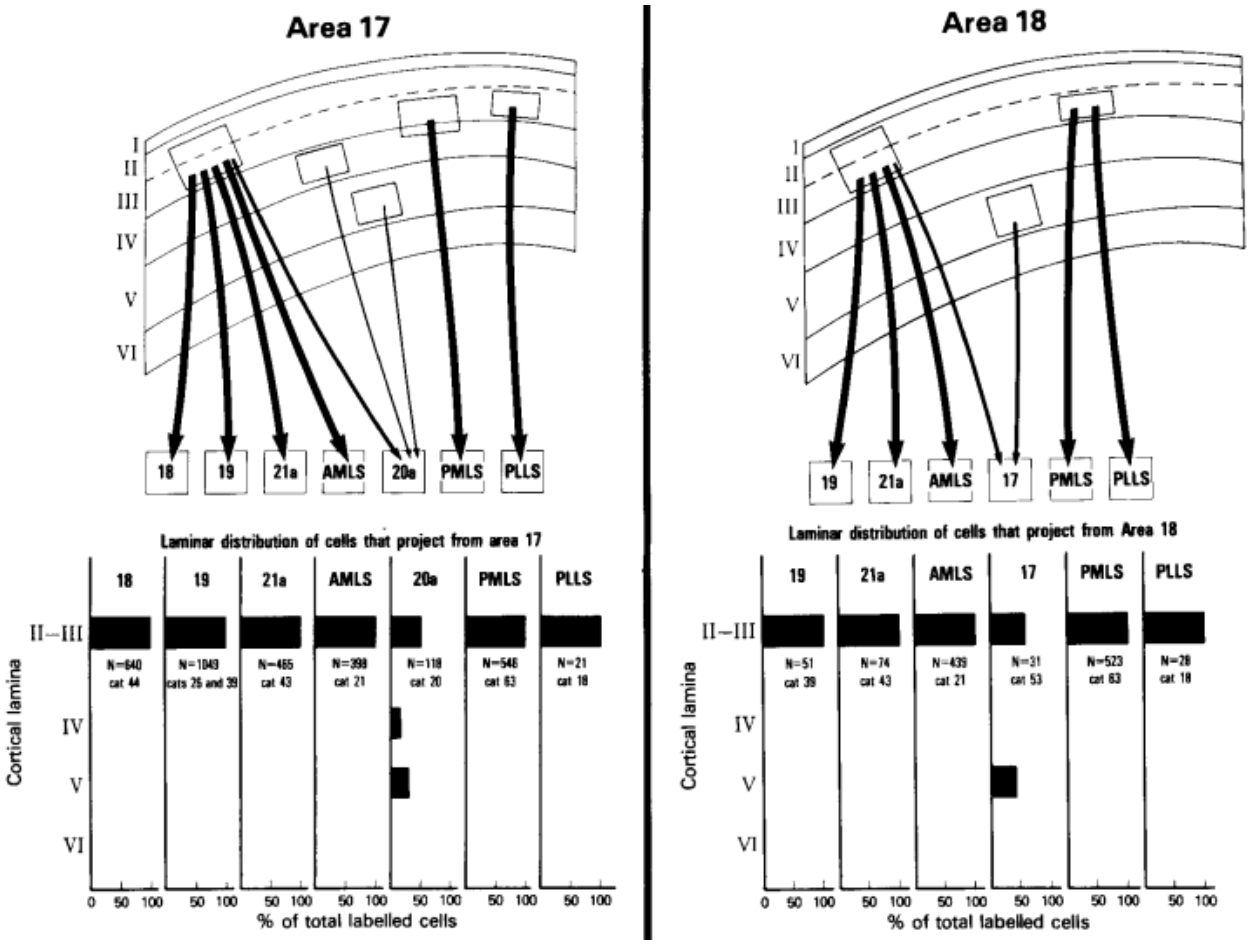

Figure A-7: Laminar origin of cortico-cortical projections in the cat. For further clarity, the extent of the projection originating in each layer is shown in histogram form in the lower part of the figure. Adopted from (Symonds and Rosenquist, 1984)

layers 5 and 6 in area 17 (Martinez-Conde et al., 1999). Secondly, there is a projection originating in infragranular layers 5 and 6 of area 18 and terminating in infragranular layers 5 and 6 of area 17 (Symonds and Rosenquist, 1984). It has also been suggested that many projections from area 18 to area 17 terminate in the top of layer 4 (Salin and Bullier, 1995). Finally, it is generally suggested that cortico-cortical connections in early sensory areas mainly originate from layer 3 while cortical-subcortical projections usually originate in layer 5 and 6 (Thomson et al., 2002). For example, projections from both area 17 and 18 to LGN originate in layer 6 and they are generally reciprocal connections (Payne and Peters, 2002).

Figure A-7 shows a summary of the laminar origin of inter-areal projections originating from area 17 and 18. The thickness of the arrows shows the relative weight of a projection to a

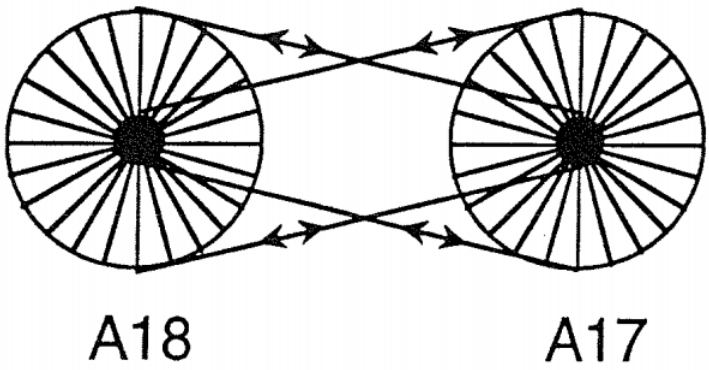

Figure A-8: Reciprocity of projections between area 17 and 18. A point in area 17 sends diverging projections to a wider region in area 18 and also receives converging projections from the same region. The same is true for a point in area 18. Adopted from (Salin et al., 1995)

given area compared to the total strength of the projection to that area. For example, we see that the projection from area 18 to area 17 is quite evenly distributed in origin from layer 2/3 and from layer 5 and 6. The histogram below shows again the weight of the projection but in bar form.

On functional grounds, it is interesting to note that cortico-cortical projections between areas 17 and 18 are strong enough to drive each other without LGN input to one of these areas. By destroying layer A of the LGN which provides the only X-input from the contralateral eye, Malpeli et al. showed that activity in the superficial layers of area 17 could only be silenced if area 18 was inactivated simultaneously with the LGN (Payne and Peters, 2002). This indicates that connections between area 17 and 18 could potentially play an important role in the shaping of their mutual response. However, studies which inactivated either supragranular or infragranular layers of area 18 failed to see a consistent effect on area 17 responses (Alonso et al., 1993a, b, Martinez-Conde et al., 1999).

The projections between area 17 and 18 have also been shown to preferably connect regions of similar retinotopy and similar orientation preference (Gilbert and Wiesel, 1989, Salin et al., 1995). In fact, the connections from area 17 to 18 originating from a given point diverge to connect a wider region in area 18, while also receiving converging connections from the same region. This may sound a little obscure but it should be clarified in Figure A-8. We can see that a point in area 17 (A17) sends diverging projections to a wider region in area 18. This same region in area 18 in turn sends converging projections to this point in area 17 (this is also true for area

18). Therefore, area 17 and 18 are reciprocally connected. The extent of the divergence was found to be about 6 mm with no significant laminar difference (Salin et al., 1995).

B. Potential in a three-layered volume conductor

Here we provide the equations for approximating the potential for the nine possible combinations of source and electrode locations in a 3-layered medium. The notation is the same as previously: namely σ_t, σ_e, and σ_b are the top, middle and bottom conductivities respectively. The electrode and source positions are given by z_e and z_s. Moreover, we set $z = 0$ in the middle of the middle layer which we assume has a thickness D. The top medium is thus located at $z < -a$ and the bottom medium at $z > a$ with $a = D/2$. To facilitate the notation we also use

$$L_{12} = \frac{\sigma_t - \sigma_e}{\sigma_t + \sigma_e} \quad and \quad L_{32} = \frac{\sigma_b - \sigma_e}{\sigma_b + \sigma_e}$$

1. Source in the top layer ($z_s < -a$):
   1.1. Electrode in the top layer ($z_e < -a$):

$$K = \frac{1}{\sigma_t}\left(\varphi(z_e, z_s) + L_{12}\varphi(z_e, -z_s - D) - \frac{4\sigma_t\sigma_e}{(\sigma_t + \sigma_e)^2}\sum_{n=0}^{N}(L_{12}L_{32})^n\varphi(z_e, z_s + 2Dn + D)\right)$$

   1.2. Electrode in the middle layer ($-a \le z_e \le a$):

$$K = \frac{2}{\sigma_t + \sigma_e}\left(\varphi(z_e, z_s) - L_{32}\varphi(z_e, -z_s + D) + \sum_{n=1}^{N}(L_{12}L_{32})^n\big(\varphi(z_e, z_s - 2Dn) - L_{32}\varphi(z_e, -z_s + 2Dn + D)\big)\right)$$

   1.3. Electrode in the bottom layer ($z_e > a$):

$$K = \frac{4\sigma_e}{(\sigma_b + \sigma_e)(\sigma_t + \sigma_e)}\left(\varphi(z_e, z_s) + \sum_{n=1}^{N}(L_{12}L_{32})^n\varphi(z_e, z_s - 2Dn)\right)$$

2. Source in the middle layer ($-a \le z_s \le a$):
   2.1. Electrode in the top layer ($z_e < -a$):

$$K = \frac{2}{\sigma_t + \sigma_e}\left(\varphi(z_e, z_s) - L_{32}\varphi(z_e, -z_s + D) + \sum_{n=1}^{N}(L_{12}L_{32})^n\big(\varphi(z_e, z_s + 2Dn) - L_{32}\varphi(z_e, -z_s + 2Dn + D)\big)\right)$$

2.2. Electrode in the middle layer ($-a \le z_e \le a$):

$$K = \frac{1}{\sigma_e}\left(\varphi(z_e, z_s) - \sum_{n=0}^{N}(L_{12}L_{32})^n\left(L_{12}\varphi(z_e, -z_s - 2Dn - D) + L_{32}\varphi(z_e, -z_s + 2Dn + D)\right)\right.$$
$$\left. + \sum_{n=1}^{N}(L_{12}L_{32})^n\left(\varphi(z_e, z_s + 2Dn) + \varphi(z_e, z_s - 2Dn)\right)\right)$$

2.3. Electrode in the bottom layer ($z_e > a$):

$$K = \frac{2}{\sigma_e + \sigma_b}\left(\varphi(z_e, z_s) - L_{12}\varphi(z_e, -z_s - D)\right.$$
$$\left. + \sum_{n=1}^{N}(L_{12}L_{32})^n\left(\varphi(z_e, z_s + 2Dn) - L_{12}\varphi(z_e, -z_s - 2Dn - D)\right)\right)$$

3. Source in the bottom layer ($z_s > a$):

3.1. Electrode in the top layer ($z_e < -a$):

$$K = \frac{4\sigma_e}{(\sigma_t + \sigma_e)(\sigma_b + \sigma_e)}\left(\varphi(z_e, z_s) + \sum_{n=1}^{N}(L_{12}L_{32})^n\varphi(z_e, z_s + 2Dn)\right)$$

3.2. Electrode in the middle layer ($-a \le z_e \le a$):

$$K = \frac{2}{\sigma_e + \sigma_b}\left(\varphi(z_e, z_s) - L_{12}\varphi(z_e, -z_s - D)\right.$$
$$\left. + \sum_{n=1}^{N}(L_{12}L_{32})^n\left(\varphi(z_e, z_s + 2Dn) - L_{12}\varphi(z_e, -z_s - 2Dn - D)\right)\right)$$

3.3. Electrode in the bottom layer ($z_e > a$):

$$K = \frac{1}{\sigma_b}\left(\varphi(z_e, z_s) + L_{32}\varphi(z_e, -z_s + D) - \frac{4\sigma_b\sigma_e}{(\sigma_b + \sigma_e)^2}\sum_{n=0}^{N}(L_{12}L_{32})^n\varphi(z_e, -z_s - 2Dn - D)\right)$$