State-space modeling and electroencephalogram source localization of slow oscillations with applications to the study of general anesthesia, sedation and sleep

by

Gladia Chork Hotan
B.S. Physics, California Institute of Technology
Submitted to the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
September 2020

Author . . . .
Department of Brain and Cognitive Sciences
Aug 14, 2020

Certified by . . . .
Patrick L Purdon
Associate Professor of Anaesthesia, Massachusetts General Hospital
Thesis Supervisor

Certified by . . . .
Emery N Brown
Edward Hood Taplin Professor of Medical Engineering and of
Computational Neuroscience, Massachusetts Institute of Technology
Thesis Supervisor

Accepted by . . . .
Rebecca Saxe
John W. Jarve (1978) Professor of Brain and Cognitive Sciences
Associate Head, Department of Brain and Cognitive Sciences
Affiliate, McGovern Institute for Brain Science

State-space modeling and electroencephalogram source localization of slow oscillations with applications to the study of general anesthesia, sedation and sleep

by

Gladia Chork Hotan

Submitted to the Department of Brain and Cognitive Sciences
on Aug 14, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

General anesthesia, sedation and sleep correspond to distinct physiological states on a spectrum of unconsciousness. Slow oscillations (0.1-1Hz) are a common feature of these unconscious states. It is unclear whether these slow oscillations might have different properties that could relate to mechanistic or behavioral differences observed in these states. In this thesis we develop novel methods to characterize the dynamic properties and spatial relationships of slow oscillations during general anesthesia, sedation, and sleep. First we analyze the electroencephalogram (EEG) power spectrum in each of these states and find that slow oscillation power increases with increasing levels of unconsciousness. Next, we perform source localization analysis to characterize the spatiotemporal relationships among distributed cortical generators for the slow oscillation using canonical coherence analysis. We find that the inherent spatial dispersion of MNE estimates could produce spurious coherence values even when sources were uncorrelated. To improve the accuracy of coherence estimates, we develop an improved source localization method using a state space model for the slow oscillation. This method employs a novel state space representation for oscillatory signals developed by Matsuda and Komaki, combined with an expectation maximization (EM) algorithm to estimate the model parameters in the sensor and source spaces. We demonstrate in simulation studies that this oscillator-EM method improves localization performance as compared to MNE. Finally, we apply the oscillator-EM method to analyze slow oscillations in the propofol, dexmedetomidine and sleep datasets, respectively. We illustrate how the application of this novel state space model and source localization method can elucidate novel properties of slow oscillation dynamics and coherence.

Thesis Supervisor: Patrick L Purdon
Title: Associate Professor of Anaesthesia, Massachusetts General Hospital

Thesis Supervisor: Emery N Brown
Title: Edward Hood Taplin Professor of Medical Engineering and of Computational Neuroscience, Massachusetts Institute of Technology

Acknowledgments

I would like to express my deepest gratitude to my PhD advisor, Dr. Patrick Purdon, for his mentorship and guidance over the years. I would also like to express my heartfelt gratitude to Dr. Emery Brown for giving me the opportunity to work in his laboratory.

I would like to thank my thesis committee members Dr. Emery Brown, Dr. Matthew Wilson and Dr. Mehrdad Jazayeri for providing me with feedback on and insights into my research progress over the years. I would also like to thank Dr. Oluwaseun Johnson-Akeju and Dr. Brandon Westover for providing valuable input on my qualifying examination proposal. In addition, many thanks to Dr. Johnson-Akeju for suggesting that I could add his dexmedetomidine data to my analysis, which has allowed me to make a richer comparison between states of unconsciousness.

I would like to thank my mentors in clinical research: Dr. Farrah Mateen, who has given me numerous valuable opportunities to contribute to statistical analyses for her research projects, and Dr. Westover and Dr. Aaron Struck, who mentored my first paper on clinical decision analysis. Thanks to my mentors, I have learned a great deal about clinical research.

I would like to express my sincere thanks to my labmates Ms. Amanda Beck, Dr. Emily Stephen, Mr. Alex He and Dr. Feng Liu for their help and support discussing source localization, oscillator model fitting and forward modeling. Thank you also to Dr. Laura Lewis, Dr. Seong-Eun Kim and Mr. David Zhou for helping me to get started with EEG data analysis and multitaper spectral analysis. I would also like to thank Dr. Dax Koh and Dr. Siong Thye Goh for useful discussions on linear algebra, probability theory and inference.

I am very grateful to my control course teammates Dr. Jingzhi An, Dr. Sourish Chakravarty, Dr. John Abel, and Ms. Ayan Waite for allowing me the opportunity to work with them on developing and conducting our course, “HST.S56: Introduction to Closed-Loop Control of Physiological Systems”, during MIT’s Independent Activities Period (IAP) in 2019 and 2020. This gave me the chance to delve deeper into the

subject of control in the company of a wonderful team.

I am also thankful to have had the opportunity to independently design and travel to Korea to teach, in 2015 and 2018, a 5-day neuroscience course entitled "Introduction to Neuroscience" to high school students who were passionate to learn about the subject. This opportunity was organized by the Washington Educational Research Institute LLC.

Last but certainly not least, I would like to thank the faculty and speakers who conducted the courses comprising the MIT Sloan Healthcare Certificate which I completed. Thank you to my healthcare lab teammates Mr. Leonardo Escudero, Mr. Anthony Orr and Mr. Akhan Mukhanov who were truly a pleasure to work with on our project about diabetes and hypertension patient outcomes at Boston Medical Center.

Contents

1 Introduction **21**

1.1 General anesthesia, sedation and sleep: States on a continuum of unconsciousness . . . 21

1.2 The neurobiology of the slow oscillation . . . 24

1.2.1 Introduction to the slow oscillation . . . 24

1.2.2 The neuroanatomy of thalamocortical synchrony . . . 25

1.2.3 The thalamus takes input from layer 6 and layer 5 pyramidal cells, which have narrow and extensive axonal ramifications respectively . . . 26

1.2.4 Relay cells are divided into two classes: core cells, which have focused projections to cortex, and matrix cells, which have diffuse projections to cortex . . . 27

1.2.5 Cortical mechanisms of the slow oscillation . . . 27

1.2.6 Thalamic mechanisms of the slow oscillation . . . 31

1.2.7 The delta oscillation . . . 33

1.3 The slow (0.1-1Hz) oscillation as a biomarker of unconsciousness . . . 34

1.3.1 Propofol vs dexmedetomidine . . . 36

1.3.2 Propofol vs sleep . . . 36

1.3.3 Dexmedetomidine vs sleep . . . 37

1.4 Innovation and significance . . . 37

1.5 Thesis overview . . . 39

2 Slow oscillation power and spatial distribution in sensor and source space 41

2.1 Propofol, dexmedetomidine and sleep datasets 41

2.1.1 Propofol dataset 41

2.1.2 Dexmedetomidine dataset 42

2.1.3 Sleep dataset 43

2.2 Conditions of interest: 5 distinct states of unconsciousness 44

2.3 Slow oscillations in EEG sensor space 45

2.3.1 Data cleaning: Removal of bad and bridged channels 45

2.3.2 Multitaper spectral analysis 46

2.3.3 Results: Slow oscillation power and spatial distribution in EEG sensor space 47

2.4 Slow oscillations in source space 48

2.4.1 Overview of source localization techniques 48

2.4.2 MNE source localization 49

2.4.3 Multitaper spectral analysis in source space 50

2.4.4 Results: Slow oscillation spatial distribution in source space . 51

3 Source space functional connectivity in resting-state networks 57

3.1 Resting-state networks of interest 57

3.2 Source-space canonical coherence analysis 60

3.3 Results: Source-space canonical coherence between the Default Mode Network and Salience Network in the slow band 64

3.4 Source-space canonical coherence analysis accounting for the effect of the resolution matrix 65

3.5 Simulation: MNE causes spurious canonical coherence 70

3.6 Summary: Limitations of MNE source localization 73

4 Improving source localization with a state-space oscillator model 81

4.1 The Matsuda-Komaki oscillator model 81

4.2 The oscillator-EM method: Source localization under the oscillator model 82

4.3 Simulation studies: The oscillator-EM method outperforms MNE at source localization 84
4.3.1 Simulation of a single slow oscillation 84
4.3.2 Simulation of a slow and an alpha oscillation in an overlapping region 85
4.3.3 Simulations in different cortical regions 86
4.3.4 Simulations of slow and alpha oscillations with varying relative amplitudes 87
4.4 Oscillator-EM source localization applied to the EEG data 87
4.4.1 Finding neural oscillations using the Matsuda model in EEG sensor space 87
4.4.2 Source localization of real data with oscillator-EM 88
4.5 Source-space canonical coherence in resting state networks under oscillator-EM source localization 89

5 Conclusion **101**
5.1 Thesis summary 101
5.2 Future directions 102

List of Figures

2-1 Power spectral analysis of each condition of interest (dexmedetomidine, NREM2, NREM3, Propofol troughmax, Propofol peakmax) performed on 2 min of 4s non-overlapping data windows using multitaper spectral analysis (3 tapers, 1Hz spectral resolution). (a) Average power spectra for each of the 5 conditions. (b) Average slow band power in each condition. Higher slow band power correlates with deeper states of unconsciousness. 47

2-2 Spatial distribution of average slow band (0.1-1Hz) power in EEG sensor space, Laplacian-referenced. The power spectral density was estimated using multitaper spectral analysis on 2mins of 4s non-overlapping data windows (3 tapers, spectral resolution 1Hz). The slow power is broadly distributed across the scalp. Slow power increases across the whole scalp as the depth of unconsciousness increases (from NREM2 to NREM3; from troughmax to peakmax). The slow power magnitude and distribution under dexmedetomidine sedation resembles NREM2 sleep more closely than NREM3 sleep, possibly indicating that the depth of unconsciousness under dexmedetomidine sedation at this dosage is similar to NREM2 sleep. 53

2-3 Average slow band power spectral density in source space for the propofol subjects. The power spectral density was estimated by performing source localization on the EEG data, then using multitaper spectral analysis (4s non-overlapping windows, 3 tapers, spectral resolution 1Hz) on the source-space timeseries data. For each condition of interest, the spectral estimates for each subject were morphed onto the average brain anatomy for the subjects in that condition. The average slow-band power spectral density was calculated from the morphed source estimates. The slow power density is distributed across the cortex, with some concentration in the anterior and posterior regions. This analysis suggests that propofol-induced slow oscillations are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas. 54

2-4 Average slow band power spectral density in source space for the sleep subjects. The power spectral density was estimated by performing source localization on the EEG data, then using multitaper spectral analysis (4s non-overlapping windows, 3 tapers, spectral resolution 1Hz) on the source-space timeseries data. For each condition of interest, the spectral estimates for each subject were morphed onto the average brain anatomy for the subjects in that condition. The average slow-band spectral density was calculated from the morphed source estimates. The slow power density is distributed across the cortex, with some concentration in the anterior and posterior regions. This analysis suggests that sleep slow oscillations are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas. 55

2-5 Average slow band power spectral density in source space for the dexmedetomidine subjects. The power spectral density was estimated by performing source localization on the EEG data, then using multitaper spectral analysis (4s non-overlapping windows, 3 tapers, spectral resolution 1Hz) on the source-space timeseries data. For each condition of interest, the spectral estimates for each subject were morphed onto the average brain anatomy for the subjects in that condition. The average slow-band spectral density was calculated from the morphed source estimates. The slow power density is distributed across the cortex, with some concentration in the posterior regions. This analysis suggests that dexmedetomidine slow oscillations are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas. 56

3-1 Sources (red points) corresponding to the default mode network in an ico-3 parcellation of the cortex. The regions included in the default mode network for this analysis include the superior frontal, rostral middle frontal, inferior parietal, middle temporal, medial orbitofrontal, precuneus, posterior cingulate, parahippocampal and entorhinal cortices. 75

3-2 Sources (red points) corresponding to the salience network in an ico-3 parcellation of the cortex. The regions included in the salience network for this analysis include the rostral and caudal anterior cingulate cortices. 76

3-3 Canonical coherence (CC) at 0.5Hz of the MNE localized results for one propofol subject. CC within and between the DMN and Sal networks decreases more during peakmax than during troughmax. 76

3-4 Canonical coherence (CC) at 0.5Hz of the MNE localized results for one sleep subject. The CC in the Sal network decreases more prominently than in the DMN as the state of unconsciousness deepens from baseline to NREM2 to NREM3. 77

3-5 Canonical coherence (CC) at 0.5Hz of the MNE localized results for one dexmedetomidine subject. The CC in the Sal network decreases more prominently than in the DMN as the state of unconsciousness deepens from baseline to dexmedetomidine sedation. 77

3-6 Labels defined on the brain for the canonical coherence simulations. The labels are arranged sequentially from anterior to posterior to help characterize changes in canonical coherence as the distance between labels increases. 78

3-7 Example of MNE distributed solution. (a): A simulated signal was generated in label 6 (1Hz oscillation, a single 4s window, white noise elsewhere, signal-to-noise ratio of 3). (b): After MNE source localization, the estimated activity appears distributed across all other labels, covering much of the cerebral cortex. These results show that MNE may not be able to accurately reconstruct the spatial location of cortical source activity. 78

3-8 Median canonical coherence between pairs of cortical labels before (ground truth) and after MNE source localization with a 64, 128 and 256-electrode EEG cap. A slow oscillation (1 Hz) was simulated in each label in turn, with white noise in all other labels (signal-to-noise ratio 3). The simulation duration was 60 4s epochs. The canonical coherence between the label containing the simulation and all other labels was calculated before (ground truth) and after MNE source localization. These simulations were each repeated 20 times to obtain the median canonical coherence. The results are displayed as the rows in each matrix. The coherence after MNE localization is much higher than the ground truth, and is high between pairs of labels that should have low coherence. This is due to the tendency of MNE to spread the solution out over the whole cortex. Recovery of the original location of the signal is intractable even when the EEG electrode density is increased from 64 to 128 to 256 electrodes. 79

3-9 A simulation (1Hz oscillation, 60 4s epochs, white noise with SNR 3, 20 runs of simulation) was placed in a particular label with 64, 128 and 256 electrode setups. Other labels contained white noise. The figures show the median and standard deviation in canonical coherence between the label containing the simulation and all other labels. 1st column: Simulation in label 0. After localization, the coherence increased the most in the labels closer to label 0, and less in the labels which were further from label 0 and more shielded from the spreading of the MNE result by the geometry of the brain. Increasing the number of electrodes from 64 to 128 (or 256) brought the coherence closer to the ground truth in labels 4, 5 and 6. 2nd column: Simulation in label 3. After localization, the coherence has increased in all labels. As the number of electrodes increases from 64 to 128 to 256, the spurious coherence in label 0 is reduced. However, the coherence in all the labels is high, making it difficult to tell which label contained the original signal. . . 80

4-1 Comparing oscillator-EM localization of a simulated slow oscillation with MNE. A 1Hz oscillation travelling at 3.6ms^{-1} was simulated in a series of 4 adjacent sources in the frontal cortex (white noise elsewhere, signal-to-noise ratio 100). The oscillator-EM algorithm appears to recover the amplitude and spatial location of the simulated source with greater accuracy than MNE. 85

4-2 Columns: Simulation of spatially overlapping slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Rows: Results at 5 time snapshots. The oscillator-EM method successfully recovers both oscillations despite their spatial overlap. 91

4-3 Columns: Simulation of slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Rows: Different views from anterior, lateral, and posterior perspectives. Performance of the oscillator-EM method varies with spatial location on the cortex. The oscillations in the frontal cortex (1st row) are localized best. Localization is less precise in a dorsal (2nd row) and posterior (3rd row) region. However, the spatial location recovered is still more precise than with MNE (not shown). 92

4-4 Columns: Simulation of slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Rows: Various lateral regions. Localization performance decreases in regions further away from the scalp sensors (2nd and 3rd rows) and in regions obscured from the scalp by the brain's anatomy (3rd row), as expected. 93

4-5 Columns: Simulation of slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Comparing localization performance in a gyrus versus a sulcus in the dorsal region. The oscillator-EM estimate is more spread out when localizing in a sulcus, as expected since cortical currents within sulcii are more difficult to measure at the scalp with EEG. 94

4-6 Oscillator-EM simulations with a larger slow than alpha amplitude (1 vs 0.5) (a) and vice versa (0.5 vs 1) (b). For the slow oscillation, the source localization is more accurate when the amplitude is larger. The localization of the alpha oscillation remains accurate when the amplitude is reduced. 95

4-7 Histograms of oscillator frequencies identified by fitting the oscillator model to data from all clean EEG channels in a 4s epoch of data for a representative subject from each condition. In each figure, the top panel shows all the frequencies found between 0 and 4Hz (the slow-delta band) and the bottom panel shows all the frequencies found between 7 and 16Hz (the alpha-spindle band). In general, one peak in the slow band and 1 or 2 peaks in the alpha band are found. 96

4-8 Histograms of the canonical coherence (CC) distribution in the 5 conditions under oscillator-EM source localization. The CC is skewed towards values closer to 1.0 in peakmax and NREM3. 97

4-9 Canonical coherence under oscillator-EM source localization in troughmax and peakmax aggregated across propofol subjects. Peakmax shows overall higher coherence than troughmax. 98

4-10 Canonical coherence under oscillator-EM source localization in NREM2 and NREM3 aggregated across sleep subjects. NREM3 shows overall higher coherence than NREM2. 99

4-11 Canonical coherence under oscillator-EM source localization in dexmedetomidine sedation aggregated across dexmedetomidine subjects. The results for dexmedetomidine resemble the results for NREM2 sleep. . . 100

List of Tables

1.1 Properties of core and matrix relay cells 28

Chapter 1

Introduction

1.1 General anesthesia, sedation and sleep: States on a continuum of unconsciousness

General anesthesia, sedation and sleep can be thought of as states on a spectrum of unconsciousness. General anesthesia is a drug-induced, reversible state of unconsciousness, analgesia, amnesia and immobility. [15] The goal of surgical general anesthesia is to place the brain into a precise state of unconsciousness, avoiding both recall of the surgery and over-administration of the anesthetic. On the one hand, recall of a surgery can lead to serious trauma in patients, who may recall the sensations of being cut without pain and feeling locked-in with no outlet to express strong emotions. In severe causes, post-traumatic stress from surgical recall can result in suicide. On the other hand, over-administration of anesthetics may lead to post-operative cognitive dysfunction. Thus, it is important to avoid both underdosing and overdosing of anesthetics.

Sedation is a drug-induced, reversible state of decreased movement and decreased arousal. [15] Under milder sedation, patients may still be responsive to external stimuli. Some sedatives like dexmedetomidine may be used to mimic a state of sleep.

Sleep is a naturally-occurring restorative state during which metabolites are cleared from the brain. During sleep, sensitivity to stimuli is reduced, but the subject can be readily aroused.

The overarching goal of this thesis is to quantify differences in human brain activity during general anesthesia, sedation and sleep so as to advance our understanding of the neural correlates of depth of unconsciousness. To this end, we study EEG data recorded from human subjects under propofol-mediated general anesthesia, sedation with dexmedetomidine, and natural sleep.

General anesthesia under propofol

Propofol is a GABA-A agonist which targets GABA receptors on inhibitory interneurons leading to the cortex, thalamus, striatum, midbrain and pons. In particular, it targets arousal areas in the midbrain and pons. In light doses, propofol induces sedation. In higher doses, such as a 5-10 second bolus administration, propofol induces general anesthesia. The patient becomes unresponsive and their respiration becomes apneic, necessitating respiratory support by intubation. [15]

Propofol binds to GABA receptors on the synapses from inhibitory interneurons onto pyramidal neurons, causing postsynaptic Cl^- channels to remain open. The resulting inward Cl^- current causes pyramidal neurons to hyperpolarize. GABA-A receptors are present throughout the brain, and since small numbers of inhibitory interneurons connect to large numbers of excitatory pyramidal neurons, GABA-A agonists binding to inhibitory interneurons suppress activity throughout the brain. [15]

Each physiological response to GABA-A agonists like propofol is attributable to different possible neural circuit mechanisms. Unconsciousness and sedation are thought to be brought about by the potentiation of GABAergic interneurons in the cortex, reticular thalamic nucleus, and arousal centers in the midbrain and pons. Atonia, a condition in which muscle tone is lost, is thought to arise from GABA-mediated inhibition of the pontine and medullary reticular nuclei. [15]

It is worth noting that anesthesia and analgesia are brought about by different drugs. Propofol induces anesthesia but not analgesia; during surgery, opioid receptor agonists such as fentanyl and remifentanyl are used to produce analgesia. It is typical to administer a cocktail of drugs containing an anesthetic and an analgesic to achieve the twin aims of anesthesia and analgesia. If an anesthetic is administered without an analgesic, the patient loses consciousness but may die of shock from the body's pain response to surgical incisions.

Sedation under dexmedetomidine

Dexmedetomidine is an $\alpha 2$ receptor agonist. It binds to $\alpha 2$ receptors on neurons emanating from the locus coeruleus. These locus coeruleus neurons become hyperpolarized, leading to reducing norepinephrine release. This results in reduced norepinephrine-mediated inhibition of the pre-optic area in the hypothalamus. The disinhibited pre-optic area inhibits ascending arousal centers via GABAergic and galainergic neurons, resulting in a state resembling physiological non-REM (NREM) sleep. [15]

Natural sleep

Sleep is a natural, restorative process observed in virtually all organisms which have brains or prototypical brains (in the form of neural bundles). Sleep promotes the clearance of metabolites such as amyloid beta from the brain [71, 9]. Sleep has also been shown to be necessary for memory consolidation [29], with sleep deprivation hindering the performance of cognitively-demanding tasks [51].

There are four stages of sleep. REM sleep is the lightest state of sleep. During REM sleep, subjects exhibit rapid eye movements, whose number per unit time increases through the night. The body enters a state of skeletal-muscle hypotonia, which prevents movement (and thus injury) from the acting out of dreams. In NREM1 sleep, subjects exhibit slow, rolling eye movements, and muscle tone further decreases. NREM2 sleep is marked by spindles and K-complexes visible on the EEG. These appear intermittently in short bursts of time at frequencies typically ranging from 11-16Hz. NREM3 sleep, the deepest sleep stage, is marked by slow-wave sleep. By this stage, spindles and K-complexes are no longer visible, and a strong slow oscillation is present. [52] Humans transition between the different stages of sleep in the course of a normal night's sleep.

1.2 The neurobiology of the slow oscillation

1.2.1 Introduction to the slow oscillation

In this section, we delve deeper into the neurobiology of the slow oscillation. As we will show in section 1.3, the slow oscillation is a key biomarker of unconsciousness, which is why we chose to focus on it in our analyses. Before discussing its role as a biomarker, we first explain the neurobiology behind how this important oscillation is produced.

The slow oscillation in the brain consists of an alternation between Up states, during which vigorous synaptic activity occurs, and Down states, during which neurons are relatively silent. The oscillation occurs synchronously over large patches of the cortex. The cortical slow oscillation persists after thalamic and callosal lesions, indicating that the cortical network is sufficient to generate the slow oscillation [68]. The slow oscillation can also be produced in cortical slices in vitro. However, subcortical structures are also involved in the slow oscillation, which is observed to occur synchronously between neurons in the cortex, thalamus, striatum and cerebellum. The horizontal axon collaterals of cortical pyramidal cells are likely to be responsible for maintaining the long-range coherence of the slow oscillation.

The emergence of the slow oscillation is correlated with diminished arousal, as in the quiescent waking state, and unconsciousness, such as during natural sleep, sedation or general anesthesia. The functional role of the slow oscillation remains an open question. Thus, we choose to focus on the slow oscillation as a biomarker of depth of unconsciousness. In this section we review current literature on the neural mechanisms that are thought to produce the slow oscillation.

1.2.2 The neuroanatomy of thalamocortical synchrony

The corticothalamic system synchronizes the activities of cortical and thalamic neurons

Majority of the inputs to relay cells and GABAergic cells in the reticular nucleus (RTN) come from the corticothalamic system. Cortical inputs can excite or inhibit thalamic relay cells via the RTN, thus synchronizing high- and low-frequency oscillations respectively.

The substrate which synchronizes populations of cortical and thalamic cells during high-frequency oscillations consists of 1) focused corticothalamic axons arising from layer VI cortical cells, and 2) diffuse corticothalamic axons arising from layer V cortical cells. These interact with 1) specifically projecting core relay cells in the dorsal thalamus and 2) diffusely projecting matrix cells in the dorsal thalamus [38].

Cortical layer 6 provides input to relay neurons in the dorsal thalamus and to the reticular thalamic nucleus (RTN)

Corticothalamic terminals predominate on relay cells in the dorsal thalamus as well as on RTN cells. Cortical layer 6 projects glutamatergically (excitatory) to thalamic relay cells and also projects to GABAergic (inhibitory) cells in the RTN. The cortex can therefore glutamatergically excite thalamic relay cells, or inhibit them via the GABAergic cells in the RTN. Work by Steriade et al. [63] suggests that under normal conditions in the intact brain, the cortex predominantly inhibits the thalamus.

Relay neurons in dorsal thalamus predominantly receive glutamatergic input from layer 6 of the cortex. The RTN also takes input predominantly from cortical layer 6. This is clear from an examination of the rat RTN: Of the synapses received by rat RTN cells, 70% are corticothalamic terminals, 20-25% are thalamocortical collateral synapses, and 15-20% are GABAergic synapses. The RTN primarily inhibits the thalamus. [38]

Corticothalamic stimulation excites RTN cells but inhibits relay cells. When stimulated, corticothalamic fibers induce fast-rising EPSPs in RTN cells. These EPSPs have both NMDA components as well as non-NMDA components, where the latter consist of AMPA-based kainate receptors. Corticothalamic stimulation tends to inhibit relay cells. When thalamic relay cells are hyperpolarized, for example during sleep, the RTN-mediated disynaptic inhibition overcomes the weak excitatory effect of corticothalamic stimulation. In vitro, when relay cells are held at -55mV, weak electrical stimulation of corticothalamic fibers produces a small, short-latency EPSP in relay cells, but this is followed by a deep, prolonged IPSP that lasts up to 100ms. This IPSP consists of both GABA-A and GABA-B-receptor-mediated components. [38]

The synapses of corticothalamic fibers on RTN cells are more powerful than those on relay cells. There are approximately 3 times as many GluR4 receptor subunits at corticothalamic synapses on RTN cells than at corticothalamic synapses on relay cells. This leads to an AMPA-receptor-based difference in synaptic strength between the corticothalamic connections to RTN and the corticothalamic connections to relay cells. Due to the larger number of GluR4 receptors subunits at corticothalamic synapses on RTN, corticothalamic fibers exert a more powerful influence on RTN cells than on relay cells. [28]

1.2.3 The thalamus takes input from layer 6 and layer 5 pyramidal cells, which have narrow and extensive axonal ramifications respectively

Two classes of pyramidal cells project from cortex to thalamus: 1) pyramidal cells with somata located in layer 6, which project from a particular cortical area to its related thalamic nucleus, and 2) pyramidal cells with somata located in layer 5. [38]

The layer 6 pyramidal cells are small and have a narrow dendritic field, influencing a narrow zone of the cortical area in which they lie. However, the terminals of a layer 6 pyramidal cell can influence thalamic relay cells that project to regions of cortex outside the narrow cortical zone in which the pyramidal cell lies. A layer 6 pyramidal cell gives off axon collaterals to the RTN. Its terminations in the dorsal thalamus project to the specific thalamic nucleus associated with the cortical area from which the pyramidal cell receives input. [38]

The layer 5 pyramidal cells have extensive axonal ramifications in the thalamus and cortex. They have thick axons that give off horizontal collaterals which extend for a long distance throughout layers 3 and 4 of the cortex. A layer 5 pyramidal cell projects not only to the thalamus, but also to the tectum, other brainstem regions, and the spinal cord. A layer 5 pyramidal cell does not give off axon collaterals to the RTN. Its terminations in the dorsal thalamus extend into numerous adjacent nuclei, not just the thalamic nucleus from which its parent cortical area receives inputs. [38]

1.2.4 Relay cells are divided into two classes: core cells, which have focused projections to cortex, and matrix cells, which have diffuse projections to cortex

There are two classes of relay cells in the primate thalamus: parvalbumin-immunoreactive core cells and calbindin-immunoreactive matrix cells. Table 1.1 summarizes the characteristics of both types of cell. [38]

1.2.5 Cortical mechanisms of the slow oscillation

The Up state of the slow oscillation is likely initiated by pyramidal cells in cortical layer 5

Up-state initiation refers to the initiation of synchronous activity from a relatively silent network. Up states initiate in the same group of neurons, and engage the rest of the cortical network in the same sequence on each cycle. [49] This was established in various in vitro experiments [45, 23] and various in vivo experiments [43].

Experimental evidence points to the fact that cortical layer 5 is important for initiation of Up states. In vivo, the polarity of the extracellular slow oscillation potential reverses toward the middle cortical layers.[64] In vitro, multiunit activity during the slow oscillation was strongest and earliest in layer 5. [59] When layer 4 is

	Core cells	Matrix cells
Distribution	Core cells are found only in the principal sensory and motor relay nuclei, in certain nuclei of the pulvinar and in some intralaminar nuclei.	Matrix cells form a widely distributed matrix throughout all the nuclei of the dorsal thalamus.
Projections to cortex	Core cells are topographically organized. Adjacent groups of core cells project to adjacent regions of a single area of cortex, terminating in localized zones in the middle layers (deep layer 3 and layer 4).	Diffuse projections. Adjacent matrix cells in a thalamic nucleus may project to two different (although usually adjacent) cortical areas. The axons of matrix cells terminate in superficial layers (1, 2 and upper 3) of the cortex.
Role	Core relay cells send place- and modality-specific information to the cortex.	Matrix relay cells disperse activity across large areas of the cortex.

Table 1.1: Properties of core and matrix relay cells

cut horizontally, the slow oscillation can still be generated in layers 5 and 6, but not in the upper layers. Furthermore, optogenetic manipulation of layer 5 and layer 2/3 pyramidal cells has shown that layer 5 is involved in Up state initiation. [11]

In particular, there is much evidence that layer 5 pyramidal cells initiate the slow oscillation. Many layer 5 pyramidal cells fire rhythmically, exhibiting resonant firing frequencies below 15 Hz in response to both hyperpolarizing and depolarizing current pulses. [2] Low-frequency resonance may initiate the slow oscillation. Layer 5 contains "intrinsic-bursting pyramidal cells", a pyramidal cell subtype that bursts in response to depolarizing current. [21] This pyramidal cell subtype has been implicated in the initiation of cortical epileptiform activity, which some argue is a dysregulated Up state [72]. Recordings from anesthetized and sleeping cats have shown that intrinsic-bursting pyramidal cells fire before all other neuron types before Up state onset. [20] Also, there is a high degree of convergence onto and divergence from layer 5 pyramidal cells, as evidenced by the fact that layer 5 pyramidal cells 1) have wide axonal arborization within layer 5 [18] and 2) have high dendritic spine density [24]. Because of these features, layer 5 pathways cause epileptiform discharges to propagate horizontally. [69]

It is currently thought that there are two possible mechanisms by which the slow oscillation is initiated by pyramidal cells in cortical layer 5. The first is initiation by persistently active, pacemaker-like cortical neurons. The second is stochastic initiation by temporal summation of spontaneous synaptic activity. [49]

There has been debate over the type of neuronal firing that initiates the slow oscillation, with two possible scenarios for the intracortical initiation of the Up state. In the first scenario, Mini-EPSCs [26], which are spontaneous, action potential-independent excitatory synaptic potentials, temporally summate in a critical number of layer 5 pyramidal neurons, causing these neurons to spike. This tips the entire cortical network into the Up state. In the second scenario, Layer 5 pyramidal neurons which fire consistently during the Down state initiate the Up state after the refractory mechanisms from the previous Up state have subsided. There is mixed evidence regarding the second scenario. Most in vivo intracellular and extracellular recording studies done in cats have not identified spiking activity in any cortical layer during the Down state. However, in vivo rodent studies have found spontaneous spiking in layer 5 pyramidal neurons during the Down state. [34] In vitro, pyramidal cells in layer 5 do spontaneously fire during the Down state. [59]

Synaptic activity is probably primarily responsible for the maintenance of the Up state

Following Up state initiation, Up states are sustained for hundreds of milliseconds to a few seconds. The question of whether synaptic or intrinsic membrane properties contribute more to the sustaining of the persistent neural activity that characterizes the Up state remains under debate. There is some role played by intrinsic mechanisms in the persistent activity of the Up state, as seen in the involvement of layer 5 pyramidal cells. However, most evidence supports a synaptic basis for persistent activity during the Up state: recurrent excitatory synaptic activity balanced by synaptic inhibition. [49]

The fact that manipulation of membrane potential does not affect Up state periodicity supports the hypothesis that intrinsic mechanisms are less important for Up state persistence. Injection of current to hyperpolarize or depolarize the cellular membrane potential does not affect Up state duration or period [67], showing that Up state persistence does not depend on membrane voltage.

Instead, there has been substantial evidence suggesting that excitatory synaptic activity is important for Up state persistence. Firstly, membrane potential variance and irregularity of interspike intervals are high in cortical neurons during Up states, which is consistent with excitatory and inhibitory synaptic barrages sustaining the persistent activity. Secondly, when AMPA or NMDA receptor antagonists, which block fast glutamatergic transmission, are applied to cortical slices in vitro, Up states are not observed. [59] Thirdly, blockade of NMDA receptors by ketamine reduced Up state duration in vivo by more than half. [67] Similar observations have been made in slices.

Inhibitory synaptic activity is also important for Up state persistence, since a balance between synaptic excitation and synaptic inhibition in necessary to maintain the membrane potential near spike threshold, where synaptic noise can transiently cause firing.

Thus, synaptic activity, rather than intrinsic membrane properties, is likely to be the major contributing factor to Up state persistence.

Intrinsic membrane properties (the activation of activity-dependent K^+ conductances) are probably primarily responsible for Up state termination

The question of whether Up state termination depends on synaptic or intrinsic mechanisms has also been a subject of debate. In a synaptic explanation for Up state termination, Up state termination may be caused by enhanced activity of inhibitory interneurons or synaptic depression of excitatory synapses near the end of the Up state [70]. Alternatively, in an intrinsic explanation, Up state termination may be caused by the activation of activity-dependent hyperpolarizing conductances.

It has been shown that the activation of activity-dependent K^+ conductances in a neuron (primarily pyramidal cells) during Up states decreases the excitability of the neuron, decreasing its firing rates and leading to Up state termination. On the other hand, while the synchronous initiation of the Up state is linked to increased excitability in a stereotypical subnetwork of neurons, no such stereotypical subnetwork or sequence of activation has been found which may explain synchronous Up state termination. Thus, it is likely that an intrinsic rather than a synaptic mechanism is primarily responsible for Up state termination.[49]

Synaptic inhibition is unlikely to cause the transition to the Down state because 1) there is no increase in inhibitory cell firing when Up states terminate, and 2) inhibitory conductances decrease as the Up state progresses. However, it should be noted that further research is necessary before synaptic inhibition can be ruled out as a mechanism for Up state termination. There are three main pieces of evidence to support this idea. Firstly, slow GABA-B-mediated inhibition may play a role in terminating Up states. In vitro, progressive pharmacological blockade of slow GABA-B mediated inhibition caused a continuous increase in the Up state [44]. Secondly, synaptic inhibition dominates 4-7% of cortical neurons prior to Up state termination [40]. However, it is unclear whether this contributes significantly to Up state termination. Thirdly, “late-firing” inhibitory cells, ie. inhibitory cells that fire towards the end of the Up state, might be important contributors to Up state termination synchrony provided that they have a wide enough axonal spread. Somatostatin (SOM)-positive inhibitory cells are a possible candidate.

All in all, the activation of activity-dependent K+ conductances is the most likely mechanism behind Up state termination. The activation of GABA-B receptors may exert a modulatory influence on Up state duration.

1.2.6 Thalamic mechanisms of the slow oscillation

Thalamocortical projections are mainly involved in Up state initiation and determination of the period of the Up state

The thalamus is thought to be involved in the initiation of the Up state and the determination of the period of the Up state via the following process: [49]

1. After a Down state, thalamocortical neurons (excitatory thalamic relay cells) fire a post-inhibitory rebound spike-burst.

2. An Up state is produced in the cortex.

3. Corticothalamic input depolarizes inhibitory thalamic reticular nucleus (TRN) cells, causing them to fire and inhibit the thalamocortical neurons.

This hypothesis has been supported by evidence that sensory or thalamic stimulation reliably produces Up states in animals [7].

Spontaneous thalamic activity appears to be necessary for the initiation of spontaneous Up states. It has been shown that the severing of thalamocortical axons in mouse barrel cortex in vitro decreases the frequency of the slow oscillation. [57]

Inactivation of the thalamus decreases the frequency of the slow oscillation and reduces the synchronization of the Up states in cortical neurons [41]. This indicates that the thalamocortical dialogue is crucial for synchronization of the slow oscillation, and that the thalamus directly influences the period of the slow oscillation. Interestingly, within 30h after thalamic inactivation, the cortex recovers its normal slow oscillation frequency. This may indicate that the slow oscillation is highly important to cortical networks, since cortical recurrent circuitry is recruited to compensate for the missing thalamic signal for Up state initiation.

It is thought that sensory and non-sensory thalamic nuclei are involved in the initiation and maintenance/termination of the Up state respectively. Core cells form the focused thalamocortical projections from sensory thalamic nuclei, while matrix cells form the diffuse thalamocortical projections from non-sensory thalamic nuclei. During Up state initiation, relay cells in sensory thalamic nuclei are inhibited by the TRN. On the other hand, relay cells in non-sensory thalamic nuclei are dominated by excitation during the Up state, indicating that they may be involved in sustaining and possibly terminating it. Furthermore, since thalamocortical projections from these non-sensory nuclei are diffuse, the non-sensory nuclei might play a role in maintaining the Up state. [49]

Core and matrix cells in the thalamus maintain widespread thalamocortical synchrony

The following outlines a possible process by which core and matrix cells play a complementary role in spreading synchronous activity through large areas of cortex: [38]

1. Within a cortical zone, matrix cells terminate in superficial layers while core cells terminate in middle layers. This forms a coincidence-detection circuit which allows for temporal integration between inputs from the two cell classes. Coincidence could generate synchronous activity within cortical columns and between cortical columns activated by the same stimulus.

2. This activity would in turn be transmitted to the thalamic nuclei corresponding to these columns via layer 6 corticothalamic cells, reinforcing thalamocortical synchrony.

3. The diffuse projections of matrix cells in the thalamic nucleus where the activity originated could be responsible for spreading the oscillation to other cortical areas.

4. Other thalamic nuclei would be recruited via the diffuse projections of layer 5 corticothalamic neurons.

5. The diffuse projections of matrix cells from these other thalamic nuclei would further recruit more cortical areas. It should be noted that the principal targets of layer 5 corticothalamic axons are thalamic nuclei, which are rich in matrix cells.

The widespread collaterals of layer 5 corticothalamic cells within the cortex would also promote the spread of the activity.

1.2.7 The delta oscillation

Closely related to the slow oscillation is the delta oscillation (1-4Hz), which mostly appears during deep slow-wave sleep. Delta oscillations are thought to be generated by 1) intrinsic properties of thalamocortical neurons, and 2) intrinsic properties of layer 5 pyramidal neurons. [49]

The generation of delta oscillations by thalamocortical neurons is associated with h-currents and T-currents. During slow-wave sleep, brainstem cholinergic neurons decrease in activity beyond the spindle range, causing greater hyperpolarization of membrane potentials. [66] At this level of hyperpolarization, interactions between the h-current and the T-current produce delta oscillations. The hyperpolarized membrane potential activates the h-current and de-inactivates the T-current, causing a Ca^2+ spike. The Ca^2+ spike deactivates the h-current and inactivates the T-current, causing repolarization which restarts the cycle. [47]

Steriade et al. demonstrated that the delta oscillation remains after removal of the thalamus, suggesting that there may be a separate cortical generator of the delta oscillation. This generator is thought to function based on mechanisms similar to those which generate Up and Down states in the slow oscillation: the activation of intrinsic bursting pyramidal neurons in cortical layer 5 followed by the activation of long-lasting K+ conductances. [68] Thus, the cortex is thought to be the primary initiator of the slow oscillation, while delta oscillations may originate either in the cortex or via thalamocortical neurons.

Delta oscillations are synchronized by cortical Up states in deep slow-wave sleep. The synchronization is due to the depolarization of corticothalamic cells during the Up state, which produces monosynaptic excitatory input, and disynaptic inhibition from TRN cells. [66] Similar to spindles, the delta oscillation is thought to promote synaptic plasticity. [49]

1.3 The slow (0.1-1Hz) oscillation as a biomarker of unconsciousness

In this section we discuss literature which has studied how the slow oscillation appears as a biomarker of unconsciousness. Steriade et al. [65] conducted experiments in cats, recording intracellularly from reticular thalamic neurons. The cats were anesthetized with urethane, ketamine-xylazine or ketamine-nitrous oxide. The slow oscillation was found in 65% of recorded cells. Action potential spiked were fired during the depolarizing envelopes of the slow oscillation. As the cellular membrane voltage became more hyperpolarized, the frequency of neuronal spiking was reduced and the slow oscillation amplitude increased.

Chauvette et el. [19] experimentally compared the slow oscillation in cats under sleep and ketamine-xylazine anesthesia. They measured field potentials, extracellular unit recordings and intracellular activities of cortical neurons in the cats, recording from a large number of cortical areas including suprasylvian gyrus (associative cortex), ectosylvian gyrus (auditory cortex), postcruciate gyrus (somatosensory cortex), precruciate gyrus (motor cortex), posterior marginal gyrus (primary visual cortex) and frontal gyrus medial prefrontal cortex). They found that the slow oscillation groups neuronal firing, and is larger in amplitude and more periodic under ketamine-xylazine anesthesia than during sleep. Specifically, neuronal firing was grouped by the depth-negative phase of the local field potential (LFP). (The depth-negative voltage was defined relative to a reference electrode placed between the skull and the temporal muscle). Secondly, they found that slow oscillations become present in more cortical areas under ketamine-xylazine anesthesia than during sleep. In slow-wave sleep, the slow oscillation was most prominent in the suprasylvian gyrus. It occurred at irregular intervals and did not occur regularly in all locations. On the other hand, under ketamine-xylazine anesthesia, the slow oscillation was prominent across all recording locations and was highly periodic. Thirdly, from the intracellular recordings, they determined that silent states are prolonged under ketamine-xylazine anesthesia as compared to during sleep. The duration of silent states in several areas and in all areas when pooled together was significantly longer under anesthesia than during sleep. The ratio of time spent in the silent state to the total time of recording was also significantly higher under anesthesia than during sleep. When all recordings were pooled together, the amplitude of slow oscillations throughout the cortex was significantly larger under anesthesia than during sleep, although when individual regions were compared, the slow amplitude was significantly higher under anesthesia only in the somatosensory cortex. Overall, the results suggest that slow oscillations increase in amplitude and become more widespread across the cortex as depth of unconsciousness increases. Within each cortical area, the slow oscillation exhibits greater coherence and periodicity in ketamine-xylazine anesthesia than during sleep.

A study by Lewis et al. [42] identified the slow oscillation in intracranial recordings from epilepsy patients undergoing propofol anesthesia. A microelectrode array in the temporal cortex was used to record LFPs and single units. The slow oscillation appeared at loss of consciousness (LOC) and remained throughout unconsciousness. After LOC, spikes became phase-coupled to the local slow oscillations. The synchrony between slow oscillations across regions in space, as measured by the phase-locking factor, decreased during propofol anesthesia. This signalled a fragmentation of functional connectivity between cortical areas.

1.3.1 Propofol vs dexmedetomidine

There have been few studies comparing EEG activity during propofol anesthesia versus dexmedetomidine sedation in humans. Kasuya et al. [39] compared the BIS index during propofol sedation with the BIS index during dexmedetomidine sedation, but the BIS index omits many important features of the EEG which become evident with spectral analysis. Heard et al. [35] compared the pharmacodynamic responses of children anesthetized with propofol or dexmedetomidine-midazolam, but did not study the changes in their brain states as measurable by EEG.

1.3.2 Propofol vs sleep

While there are numerous studies on propofol and numerous studies on sleep, few have drawn rigorous comparisons between the two using full-head EEG studies and frequency-domain analysis techniques. Murphy et al. [48] conducted 256-electrode EEG studies of subjects undergoing propofol anesthesia as well as during natural sleep. However, they treated propofol-induced unconsciousness as a single state, and

NREM sleep as a single state. It is known that NREM2 sleep differs from NREM3 sleep in that sleep spindles occur during NREM2 sleep but not during NREM3 sleep, and NREM3 sleep is thought to be a deeper state of unconsciousness than NREM2 sleep. Furthermore, Purdon et al. [54] identified via a phase-amplitude modulation analysis two distinct states of unconsciousness under propofol anesthesia – troughmax, a lighter state of unconsciousness that appears during loss and recovery of consciousness, and peakmax, a deeper state of unconsciousness that persists throughout deep unconsciousness. Thus, we investigate in detail the different states of unconsciousness that occur during propofol anesthesia and natural sleep. Our analyses of propofol-induced unconsciousness separately consider the troughmax state during loss of consciousness, the peakmax state during deep unconsciousness, and the troughmax state during recovery of consciousness. Our analyses of sleep separately examine NREM2 sleep and NREM3 sleep.

1.3.3 Dexmedetomidine vs sleep

Huupponen et al. [37] compared time-domain spindle activity during dexmedetomidine sedation and natural sleep, but time-domain plots do not reveal the structure of EEG activity as clearly as frequency-domain analyses do. Furthermore, the study was carried out using a single-channel EEG recorded from the forehead, and thus did not capture the global properties of dexmedetomidine sedation and natural sleep. We extend their work by carrying out frequency-domain comparisons between data recorded from the entire scalp during dexmedetomidine sedation and natural sleep.

1.4 Innovation and significance

In this thesis, we present the first three-way comparison between propofol anesthesia, dexmedetomidine sedation, and sleep, using data recorded from healthy volunteers using the same 64-electrode EEG cap system. Our hypothesis is that sleep, dexmedetomidine sedation and propofol anesthesia place the brain in states which lie on a spectrum of depth of unconsciousness. We use spectral analysis and source localization to make quantitative comparisons between the states of unconsciousness. We use canonical coherence analysis at specific frequencies of interest to elucidate the changes in cortical functional connectivity during propofol anesthesia, dexmedetomidine sedation and natural sleep. We also conduct the first quantitative analysis of how source localization biases the computation of canonical coherence. Finally, we develop a novel source localization method based on a state-space model of neural oscillations which outperforms the standard minimum norm estimate (MNE) source localization technique at spatially localizing neural oscillations. We validate our localization method in simulations and apply it to the EEG data.

Clinical relevance

Current clinical standards for measuring the level of unconsciousness during surgical anesthesia include tests of behavioural responsiveness, checks for brainstem reflexes such as the oculocephalic reflex, and the BIS index. None of these measures provide a rigorous quantification of the depth of unconsciousness. For instance, the BIS index, which is a single number extracted from multiple features of the EEG signal, tends to reflect that young children and elderly patients are not deeply unconscious even when their EEG spectrograms suggest the contrary [22, 3, 55]. This can result in the over-administration of anesthetic drugs, which may lead to temporary cognitive impairments after surgery. We will conduct a precise quantitative characterization of the EEG signatures of propofol-mediated general anesthesia, sedation under dexmedetomidine, and natural sleep. These results will aid in the development of robust metrics of unconsciousness which could be used to ensure that anesthetics and sedatives are administered in appropriate amounts to achieve precise states of unconsciousness ranging from sedation to surgical general anesthesia.

Methodological importance of source localization to neuroscience

While many analyses in the field have examined functional connectivity in the context of coherence between pairs of EEG sensors, connectivity between functional regions in the brain itself is a more fundamental quantity of interest. Source localization allows for neural activity to be inferred non-invasively based on EEG sensor recordings, enhancing our ability to draw scientific conclusions about the brain when invasive recording is not possible. In this thesis, we demonstrate that the standard MNE source localization technique is insufficient to accurately recover the spatial locations of cortical oscillations. We present a source localization method based on a state-space oscillator model that outperforms MNE. We hope that our method will help to improve the ability of the neuroscience community to probe cortical activity using non-invasive scalp recordings, unlocking a deeper understanding of the brain which was not possible before.

1.5 Thesis overview

In Chapter 2 we present the propofol, dexmedetomidine and sleep datasets. We define 5 conditions of interest corresponding to various depths of unconsciousness: troughmax and peakmax in propofol, sedation in dexmedetomidine, and NREM2 and NREM3 in sleep. We show our results for scalp-level frequency-domain quantification of the slow oscillation power in EEG sensor space. We introduce minimum norm estimate (MNE) source localization, and apply it to the sensor-space data to arrive at the spatial distribution of slow oscillation power on the cortex.

In Chapter 3 we introduce canonical coherence analysis, the frequency-domain analogue of canonical correlation analysis. We apply canonical coherence analysis to the source-localized data to examine the functional connectivity within and between two resting state networks of interest – the default mode network (DMN) and salience network (Sal) – under the 5 conditions of interest. We show that slow-band functional connectivity in these networks decreases with increasing depth of unconsciousness. We show that MNE spreads out the source estimate all over the cortex, resulting in spurious canonical coherence that prevents us from accurately estimating the functional connectivity. We explore the possibilities of removing the spurious coherence analytically and numerically, and show that its removal is intractable.

In Chapter 4 we introduce a state-space oscillator model for source localization, which we refer to as the oscillator-EM (expectation maximization) source localization method. We test the oscillator-EM method in simulations and show that it outperforms MNE in terms of accuracy at reconstructing the spatial location of a simulated oscillation. We apply the oscillator-EM method to the real data and recalculate the canonical coherence in the DMN and Sal networks. Under oscillator-EM, the coherence decreases rather than increases with increasing depth of unconsciousness, contrary to expectations and opposite to the results under MNE. We discuss possible reasons for this difference in the results.

In Chapter 5 we conclude with a summary of the thesis and discuss future directions for this work.

Chapter 2

Slow oscillation power and spatial distribution in sensor and source space

2.1 Propofol, dexmedetomidine and sleep datasets

2.1.1 Propofol dataset

The propofol dataset used consists of data from 10 healthy volunteers between 18 and 36 years of age. The subjects were American Society of Anesthesiology Physical Status I with Mallampati Class I airway anatomy. During the study, subjects breathed 30% oxygen by volume. Each subject's heart rate was monitored with an electrocardiogram. Their oxygen saturation was measured through pulse oximetry, respiration and expired carbon dioxide with capnography, and blood pressure through an arterial line. The arterial line was also used for blood sampling. To ensure subject safety, at least three anesthesiologists were present at each study: one was responsible solely for the medical management of the subject during the study, the second controlled the propofol administration, and the third performed blood sampling. When a subject became apneic, the first anesthesiologist assisted breathing with bag/mask ventilation. A phenylephrine infusion was used to maintain mean arterial pressure above a patient-specific level determined from the subject's baseline measurements. [54]

Propofol target effect-site concentrations of 0, 1, 2, 3, 4, and 5 μg/mL were achieved using a computer-controlled infusion. Each concentration level was maintained for 14 min. C_{LOR} was the target effect-site concentration at which the subject stopped responding to the button press stimulus. Emergence was achieved by lowering the target effect-site concentrations to $C_{LOR} = 0.5\mu$g/mL, $C_{LOR} = 1.0\mu$g/mL, $C_{LOR} = 1.5\mu$g/mL, and $C_{LOR} = 0\mu$g/mL, for 14 min each. Structural MRI (Siemens Trio 3 Tesla, T1-weighted magnetization-prepared rapid gradient echo, 1.3-mm slice thickness, 1.3 × 1 mm in-plane resolution, TR/TE = 2530/3.3 ms, 7deg flip angle) and digitized scalp electrode positions (Polhemus FASTRACK 3D) were acquired for each subject prior to each study. [54]

EEGs were recorded using a 64-channel BrainVision Magnetic Resonance Imaging Plus system (Brain Products, Munich, Germany) with a sampling rate of 5000 Hz, resolution 0.5 μV least significant bit (LSB), bandwidth 0.016-1000 Hz. [54]

Subjects were instructed to close their eyes throughout the study. They were tasked to respond to auditory stimuli by button presses to assess their level of conscious behavior. The stimuli consisted of either a verbal stimulus or an auditory click and were presented every 4s in a repeating sequence of click-click verbal-click-click, with a total of 210 stimuli per target effect-site concentration level. Verbal stimuli consisted either of the subject's name or a word, randomized with an equal number of name or word stimuli at each level. Subjects were instructed to press one button if they heard their name and to press the other button if they heard any other stimulus. Button-press stimuli were recorded using a custom-built computer mouse which was strapped to the subject's hand. [54]

2.1.2 Dexmedetomidine dataset

The dexmedetomidine dataset used consists of data from 10 healthy volunteers between 18 and 36 years of age. All subjects provided informed consent and were American Society of Anesthesiology physical status I with Mallampati class I airway anatomy. [4]

A 1 μg/kg loading bolus of dexmedetomidine was administered over 10 min, followed by a 0.7 μg/kg/h infusion for 50 min. During the study, subjects breathed 21% oxygen by volume. Each subject's heart rate was monitored with an electrocardiogram. Their oxygen saturation was measured through pulse oximetry, respiration and expired carbon dioxide through capnography, and blood pressure through a blood pressure cuff. [4]

During induction and emergence, EEGs were recorded using the same 64-channel BrainVision Magnetic Resonance Imaging Plus system (Brain Products, Munich, Germany) at a sampling rate of 1000 Hz, resolution 0.5 μV least significant bit, and bandwidth 0.016 to 1000 Hz. The subjects were instructed to close their eyes throughout the study. [4]

The subjects were presented with auditory stimuli during the study and asked to respond by button presses to assess their level of conscious behavior. The stimuli consisted of the subject's name presented every 2 min. Button-press stimuli were recorded using a custom-built computer mouse which was strapped to the subject's hand. [4]

2.1.3 Sleep dataset

The sleep dataset used consists of data from 10 healthy volunteers between 19 and 32 years of age. The subjects spent two consecutive nights sleeping in the sleep laboratory. Subjects were screened to ensure that they had regular sleep schedules and no histories of sleep disorders, psychiatric problems, neurological disease, tobacco or drug use. One night of home monitoring was performed before the study to exclude subjects with obstructive sleep apnea (using a threshold of AHI<5 and RDI<15) (WatchPAT, Itamar Medical). [53]

EEGs were recorded using the same 64-channel BrainVision MRI Plus system (Brain Products) with a sampling rate of 5000 Hz, resolution 0.5 μV least significant bit (LSB), bandwidth 0.016–1000 Hz. In addition to the EEG cap, the subjects wore standard clinical PSG sensors including PTAF, airflow, abdominal belt, and eye, chin, and limb electrodes. EMG data were bandpass filtered between 10 and 70 Hz with the addition of a notch filter at 60 Hz. Airflow and abdominal belt data were bandpass filtered between 0.1 and 12 Hz. EEG and DC channel data were unfiltered. [53]

2.2 Conditions of interest: 5 distinct states of unconsciousness

Under propofol general anesthesia, we analyze two distinct states of unconsciousness. The first state, “peakmax”, corresponds to surgical general anesthesia, during which a subject cannot be roused by external stimuli. This state is termed peakmax because during this state, the maximum alpha oscillation amplitude occurs at the peaks of the slow oscillation [54]. The second state, “troughmax”, corresponds to induction and emergence from general anesthesia, and is a lighter state of unconsciousness during which the subjects were able to respond intermittently to the auditory stimuli by button-pressing. During troughmax, the maximum alpha amplitude occurs at the troughs of the slow oscillation [54].

Since our dexmedetomidine dataset comprises sedation at a fixed dose of dexmedetomidine, a single state of unconsciousness is produced. We refer to this state as ”sedation”.

From our sleep dataset, we analyze NREM2 and NREM3 sleep, which are the deepest stages of sleep. These stages were chosen so as to compare the deepest stages of sleep with general anesthesia, which is intuitively thought of as a deeper state of unconsciousness than sleep, and with sedation, which has been shown to mimic NREM2 or NREM3 sleep depending on the dose concentration of the sedative [37].

2.3 Slow oscillations in EEG sensor space

2.3.1 Data cleaning: Removal of bad and bridged channels

For each subject, bad channels and electrically bridged channels were identified and excluded from the analyses. Bad channels were identified by visual inspection of the EEG timeseries and spectrograms.

Electrical bridging occurs when EEG electrode gel oozes between electrodes and forms an electrical contact. This results in two or more electrodes having almost identical electrical signals. We chose to remove bridged electrodes from the analyses because the actual spatial location of the source of the bridged signal is unclear. To determine which channels were electrically bridged, we used "eBridge", a tool implemented in Matlab for detecting bridged electrodes. [6] The tool identifies bridged electrodes by calculating electrical distance (ED) between pairs of electrodes:

$$\begin{aligned} P_{i-j}(t) &= P_i(t) - P_j(t) && (2.1) \\ ED_{ij} &= \frac{1}{T}\sum_{t=1}^{T}(P_{i-j}(t) - \overline{P_{i-j}(t)})^2 && (2.2) \end{aligned}$$

where $P_{i-j}(t)$ is the voltage difference between electrode i and electrode j at time t. This calculation is done over all the data epochs. If a pair of channels is bridged, the plot of number of epochs versus electrical distance for the pair of channels will show both a local peak (LP) and a local minimum (LM). The tool searches for an LP with ED ≤ 3 and an LM with ED ≤ 5. The two channels are marked as bridged if their ED in 50% or more of the epochs falls below or at the ED corresponding to the LM. [5]

For each of the 5 states of unconsciousness – NREM2, NREM3, sedation, troughmax and peakmax – the bad and bridged channels for each subject were plotted topographically on the scalp. To obtain accurate estimates of the spatial distribution of power over the whole head and to obtain accurate source localization results, subjects with a large number of bad/bridged channels (approximately 10 or more) or with large regions of adjacent bad/bridged channels were excluded from the analysis. The latter occurred most frequently at the back of the head in the sleep and dexmedetomidine subjects, consistent with the fact that under lighter states of unconsciousness, subjects were more likely to move and thus cause electrical bridging.

After excluding subjects with a large number of bad/bridged channels, a total of 3 subjects for NREM2, 5 subjects for NREM3, 6 subjects for dexmedetomidine sedation, 7 subjects for troughmax, and 9 subjects for peakmax were used in the analyses.

2.3.2 Multitaper spectral analysis

The data from each subject was downsampled to 250Hz. Multitaper spectral analysis was used to estimate power spectra [8]. The multitaper method applies a series of orthogonal tapers to each window of data to reduce the variance and bias of the spectral estimates. These tapers are known as the discrete prolate spheroidal sequences (DPSS).

We used 4s non-overlapping windows. The number of tapers to be used can be calculated from the time half-bandwidth product NW, which is a function of the desired spectral resolution:

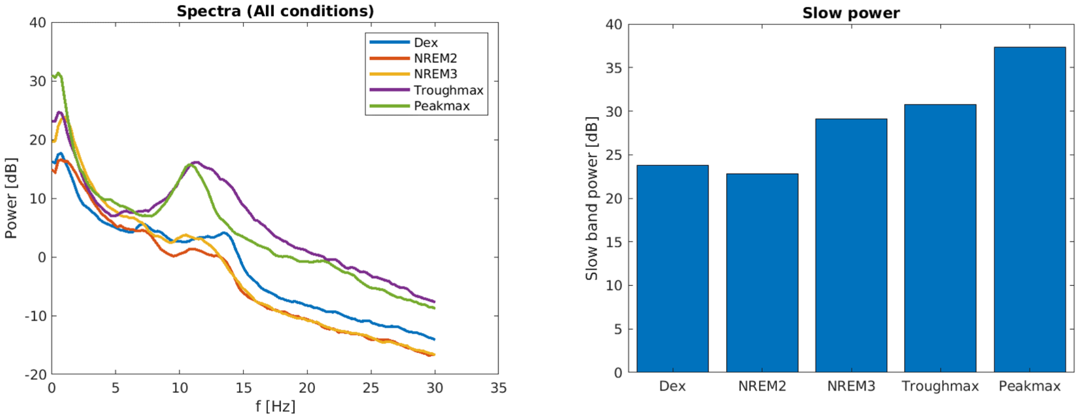
$$NW = N\frac{\Delta f/2}{\mathrm{Fs}} \tag{2.3}$$

$$K = 2 \times NW - 1 \tag{2.4}$$

where N is the length of the data window in samples, $W = \frac{\Delta f/2}{Fs}$ is the half-bandwidth, Δf is the desired spectral resolution in Hz, Fs is the sampling frequency in Hz, and K is the number of tapers. We chose a spectral resolution of 1Hz, and correspondingly used 3 DPSS tapers. The results were averaged across all 64 EEG channels for all subjects in each of the 5 conditions of interest.

2.3.3 Results: Slow oscillation power and spatial distribution in EEG sensor space

The average power spectra and average total slow band power are shown in figure 2-1. The slow band power is the greatest in peakmax, intermediate in NREM3 and troughmax, and smallest in NREM2 and dexmedetomidine (dex) sedation. Thus, higher slow band power is associated with deeper states of unconsciousness. The power spectra of each condition show two main frequency bands of oscillations: the slow-delta in the <4Hz range and the alpha/spindles in the 7-16Hz range.



(a) (b)

Figure 2-1: Power spectral analysis of each condition of interest (dexmedetomidine, NREM2, NREM3, Propofol troughmax, Propofol peakmax) performed on 2 min of 4s non-overlapping data windows using multitaper spectral analysis (3 tapers, 1Hz spectral resolution). (a) Average power spectra for each of the 5 conditions. (b) Average slow band power in each condition. Higher slow band power correlates with deeper states of unconsciousness.

The spatial distribution of the slow oscillation in EEG sensor space is shown in figure 2-2. For this analysis, the power in the 0.1-1Hz range was selected from the multitaper results, averaged across all subjects and all channels, and plotted topographically on the scalp. The plots are scaled such that all plots from each drug type have the same scale. Within each drug type, slow oscillation power at most spatial locations on the scalp increases with increasing depth of unconsciousness. The maximum slow oscillation power is around 12dB for sleep NREM3, whereas in propofol peakmax it is around 20dB, consistent with the fact that peakmax is a deeper state of unconsciousness than NREM3. Thus, average slow oscillation power increases across the whole scalp as depth of unconsciousness increases. The slow power magnitude and distribution under dexmedetomidine sedation resembles NREM2 sleep more closely than NREM3 sleep, possibly indicating that the depth of unconsciousness under dexmedetomidine sedation at this dosage is similar to NREM2 sleep.

While these scalp-level results provide a general overview of the spatial differences in slow power during different states of unconsciousness, an understanding of changes in the brain itself is necessary to make deeper statements about how activity in the various brain regions changes in the different states of unconsciousness. Therefore, we next sought to estimate slow oscillation activity on the cortical surface via source localization.

2.4 Slow oscillations in source space

2.4.1 Overview of source localization techniques

“Source localization” is a term used to describe methods for estimating electrical activity in the brain based on EEG activity measured at the scalp. Two main approaches to source localization exist. The first is to look for a small number of equivalent dipoles which represent a handful of brain regions thought to contribute to the EEG activity. This approach is appropriate when only a small number of active sources is expected, for instance in the case of a stimulus-evoked potential. A number of dipole fitting algorithms fall into this category. The second approach is to assume that dipoles could be present at all possible locations on the cortex, and localize activity to the entire grid of dipoles. This dipole grid is called the source space. The grid-based approach is appropriate when source activity is anticipated to be spread out over the brain. Examples of grid-based approaches include beamforming, spatial scans, and maximum a posteriori (MAP) Bayesian algorithms. Sparse priors such as the Laplacian prior can be used to model distributed but sparse source activity. [56]

We first focus on minimum norm estimate (MNE) source localization, which is an MAP approach that tends to produce distributed source activity. We chose this approach because the slow oscillation is known to be a widespread cortical phenomenon, and because our initial EEG-space analyses indeed showed that the slow oscillation is distributed across the scalp as depth of unconsciousness increased in the 5 conditions of interest.

2.4.2 MNE source localization

We begin by constructing a mesh of current dipoles distributed at regular spacings across the cortical surface. We call these current dipoles "sources". In our study, we used an ico-3 parcellation of the cortex, which assumes a current dipole mesh that contains a total of 1280 sources. Our data was recorded using a 64-electrode EEG cap, so we had measurements from 64 sensors on the scalp. The sensor activity $y(t)$ is related to the true source activity $x_{true}(t)$ by

$$y(t) = Gx_{true}(t) + \eta. \tag{2.5}$$

G is the forward operator or gain matrix which captures the physics of the problem. The forward model is calculated based on MRI scans of the subject's head anatomy. η is the noise in the measurement. Note that $x_{true}(t)$ has dimensions $n_{sources} \times n_{samples}$, G has dimensions $n_{sensors} \times n_{sources}$, and $y(t)$ and η have dimensions $n_{sensors} \times n_{samples}$. Our task is to solve for $x(t)$ given our measurements of $y(t)$.

We divided the data into epochs of length 4s each. Let $y(t)$ be the EEG sensor data in each epoch. We assume that the process is stationary in each epoch, so $y(t) \sim N(\mu, \sigma^2)$.

The source-space activity is recovered from the sensor-space data by solving the inverse problem. Let $x(t)$ be the reconstructed source-space data in an epoch. Then

$$x(t) = My(t) \tag{2.6}$$

where M is the inverse operator. $x(t)$ is normally distributed as $x(t) \sim N(M\mu, M^2\sigma^2)$.

In the minimum-norm estimate (MNE) method, a unique solution for the inverse problem is found by minimizing the L2 norm of the vector of source estimates:

$$\hat{x} = \underset{x}{\operatorname{argmin}}(||y - Gx||^2_{C^{-1}} + ||x||^2_{R^{-1}}), \tag{2.7}$$

where the matrix norm is defined as

$$||x||^2_{A^{-1}} = x^T A^{-1} x. \tag{2.8}$$

The MNE inverse operator can be solved for analytically by matrix differentiation, and is given by

$$M = RG^T(GRG^T + C)^{-1}, \tag{2.9}$$

where C is the noise-covariance matrix and R is the source-covariance matrix. C captures the covariance between EEG sensors while R captures the covariance between sources in source space. In the MNE setup, C and R are diagonal.

2.4.3 Multitaper spectral analysis in source space

We then performed multitaper spectral analysis on the source-space data epochs. In accordance with the multitaper method, the source-space data $\hat{x}$ in each 4s epoch is first tapered with the first three DPSS tapers $\mathscr{S}^{(m)}$, where $m = 1, 2, 3$, to achieve a spectral resolution of 1Hz.

Denote the tapered source-space data by

$$x^{(m)}(t) = x(t)\mathscr{S}^{(m)} \text{ for m=1,2,3.} \tag{2.10}$$

The tapered data is then Fourier-transformed into the frequency domain:

$$X^{(m)}(f) = x^{(m)}(t)\mathscr{F}. \tag{2.11}$$

Note that the time-domain transforms, $\mathscr{S}^{(m)}$ and $\mathscr{F}$, are applied on the right of $x(t)$ because the timeseries are the rows of $x(t)$, not the columns. $x(t)$ has dimensions $n_{sources} \times n_{samples}$.

The average of the tapered data then gives the source-space power spectral density,

$$X(f) = \frac{1}{3} \sum_{m=1}^{3} X^{(m)}(f). \tag{2.12}$$

2.4.4 Results: Slow oscillation spatial distribution in source space

All the analyses involving MNE source localization make use of the MNE software toolbox [25] developed by the MGH Athinoula A. Martinos Center for Biomedical Imaging. For each subject in each of the 5 conditions, MNE source localization was performed in each epoch to obtain time-domain source estimates. The power spectral density in the slow band (0.1-1Hz) was estimated using multitaper spectral analysis on the time-domain source estimates. An average subject brain for each condition was constructed using the built-in morphing functions in the MNE toolbox. This average brain is an anatomical average of each subject's brain. The source-space timeseries data for each subject, each epoch was morphed onto the average subject brain using the morphing functions in the MNE toolbox. The average of the power estimates from all epochs on the average subject brain was computed, giving the overall average slow band power in source space. The source-space spatial distribution of the average slow band power in each of the 5 conditions is shown in figures 2-3, 2-4 and 2-5.

The slow power spectral density is spread over the cortex. For propofol and sleep, there is some concentration of slow power in the anterior and posterior regions. For dexmedetomidine, the slow power density shows some concentration in the posterior regions. This analysis suggests that slow oscillations induced by propofol, dexmedetomidine and sleep are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas.

We next sought to explore how functional connectivity in the slow band changes under the different states of unconsciousness. In chapter 3, we make use of the MNE source localization results to estimate slow-band functional connectivity between brain regions of interest in the source space.

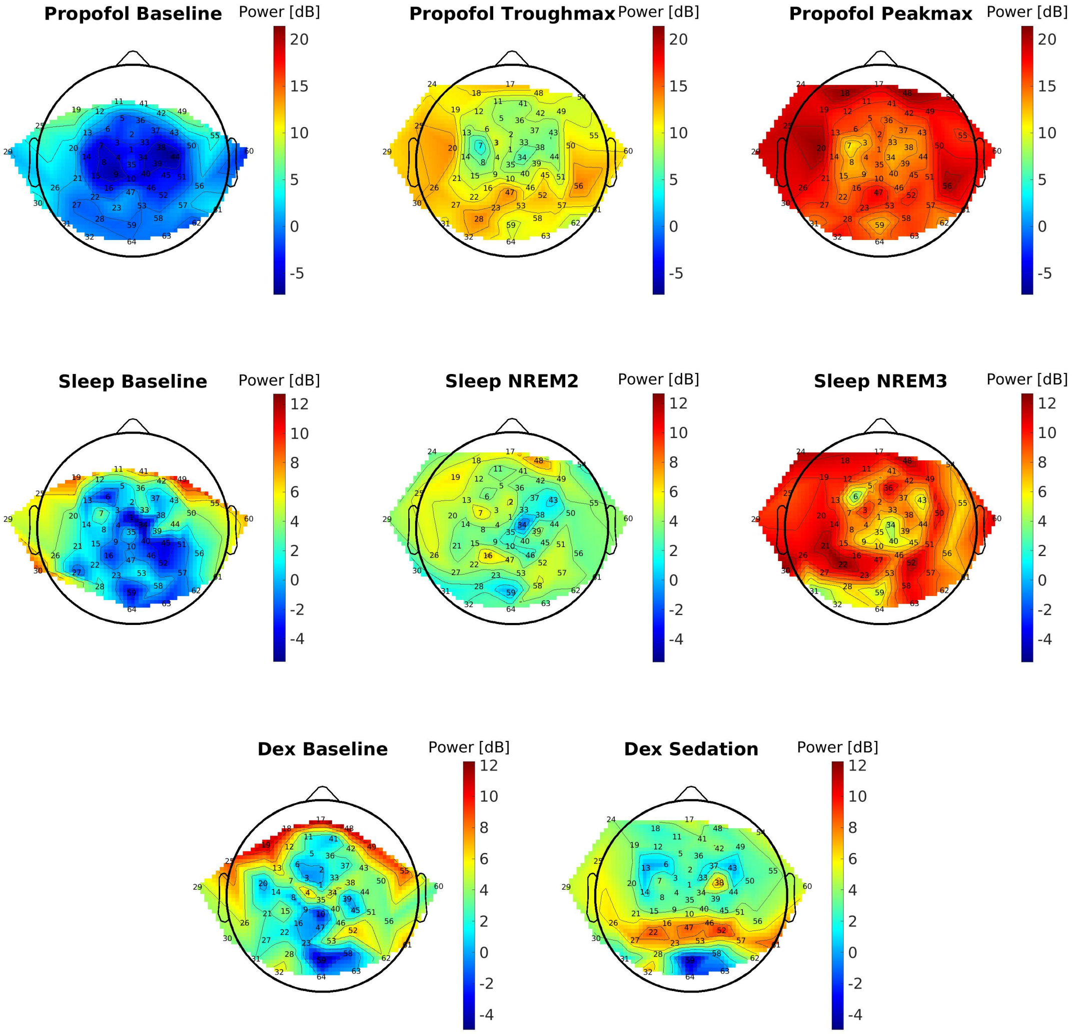
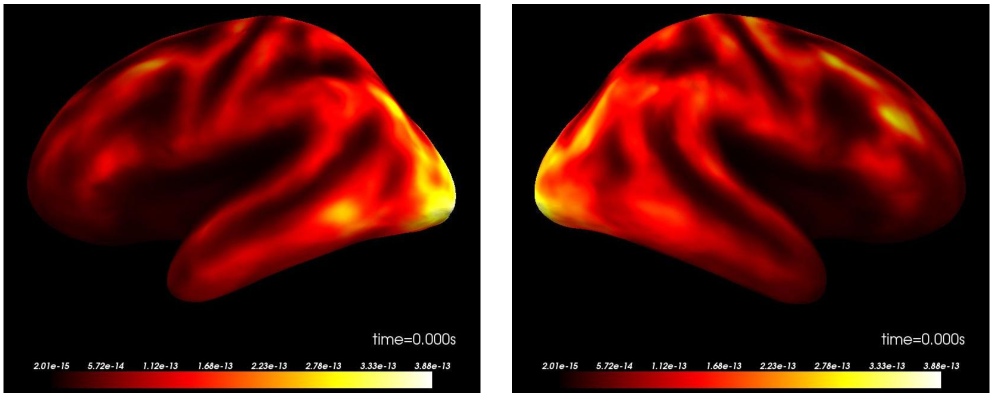
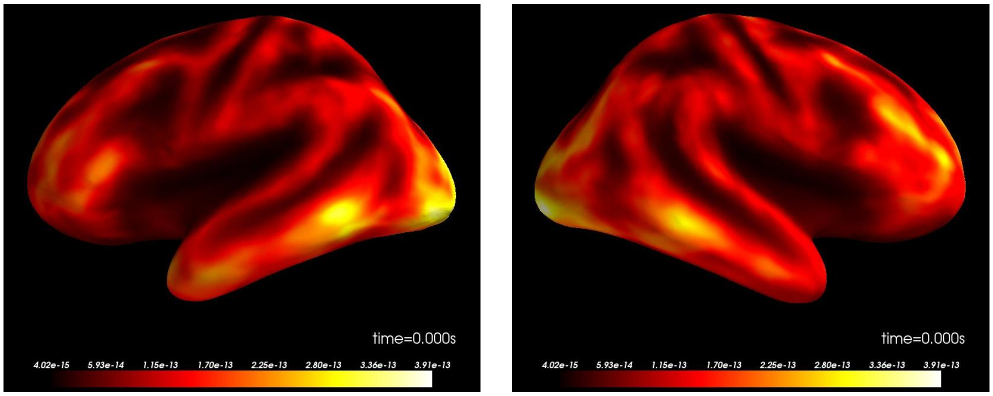


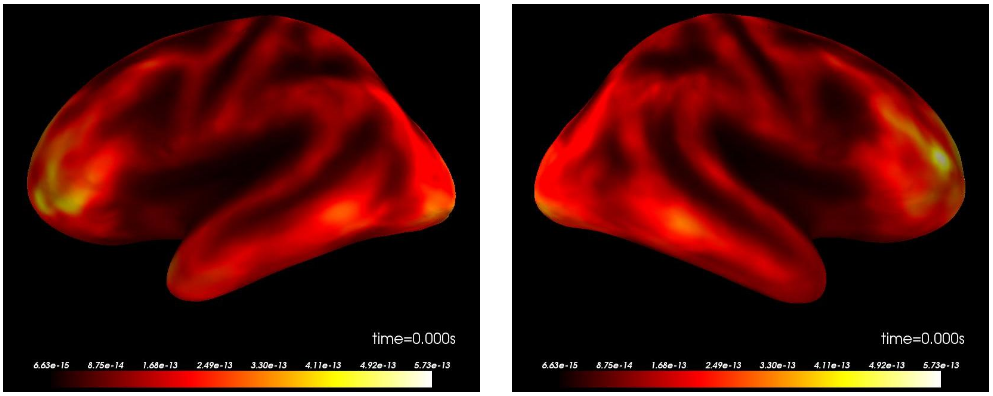
Figure 2-2: Spatial distribution of average slow band (0.1-1Hz) power in EEG sensor space, Laplacian-referenced. The power spectral density was estimated using multitaper spectral analysis on 2mins of 4s non-overlapping data windows (3 tapers, spectral resolution 1Hz). The slow power is broadly distributed across the scalp. Slow power increases across the whole scalp as the depth of unconsciousness increases (from NREM2 to NREM3; from troughmax to peakmax). The slow power magnitude and distribution under dexmedetomidine sedation resembles NREM2 sleep more closely than NREM3 sleep, possibly indicating that the depth of unconsciousness under dexmedetomidine sedation at this dosage is similar to NREM2 sleep.



(a) Propofol baseline

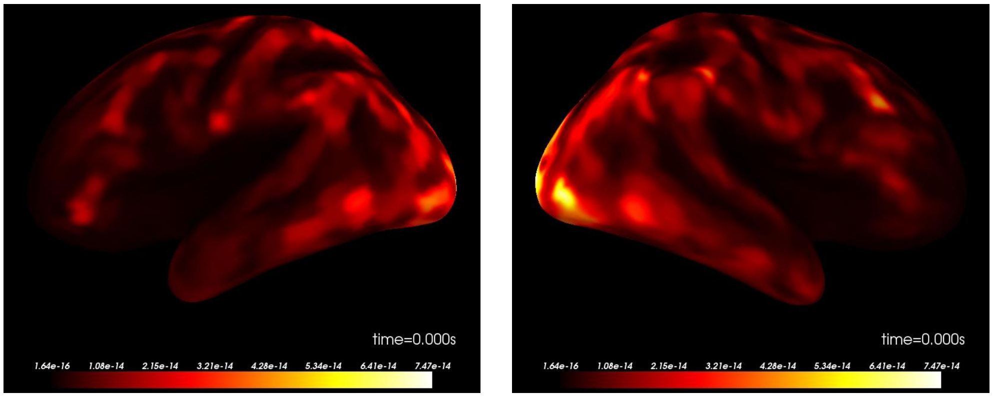


(b) Propofol troughmax

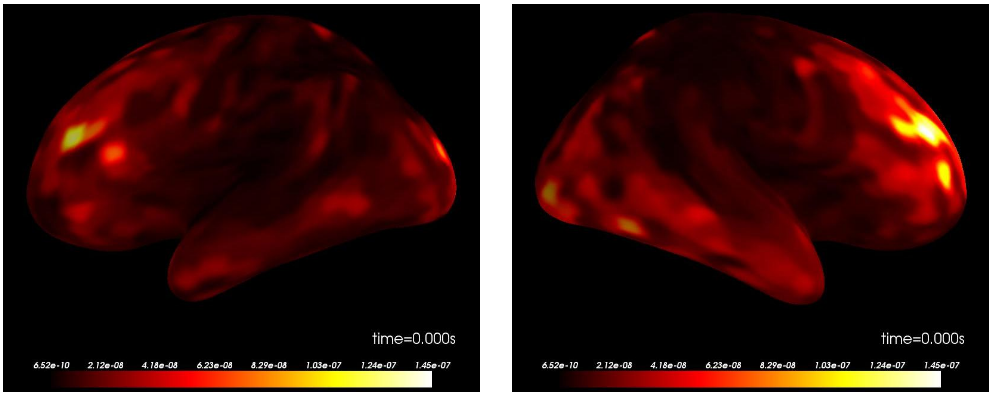


(c) Propofol peakmax

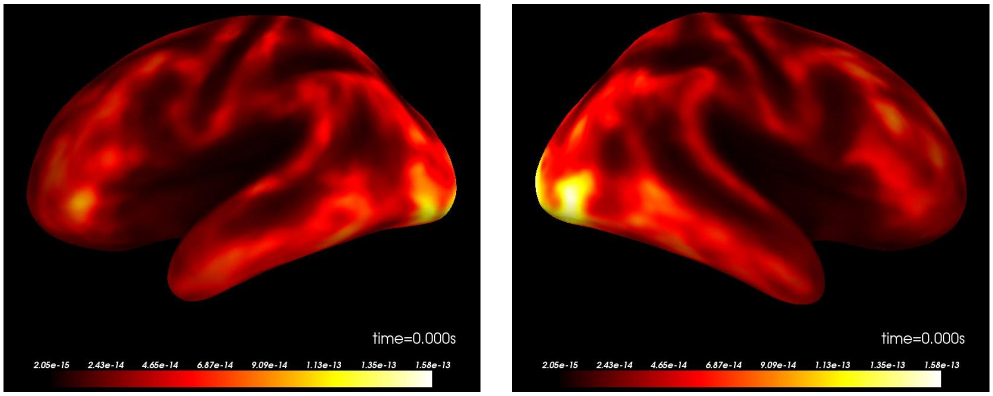
Figure 2-3: Average slow band power spectral density in source space for the propofol subjects. The power spectral density was estimated by performing source localization on the EEG data, then using multitaper spectral analysis (4s non-overlapping windows, 3 tapers, spectral resolution 1Hz) on the source-space timeseries data. For each condition of interest, the spectral estimates for each subject were morphed onto the average brain anatomy for the subjects in that condition. The average slow-band power spectral density was calculated from the morphed source estimates. The slow power density is distributed across the cortex, with some concentration in the anterior and posterior regions. This analysis suggests that propofol-induced slow oscillations are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas.



(a) Sleep baseline

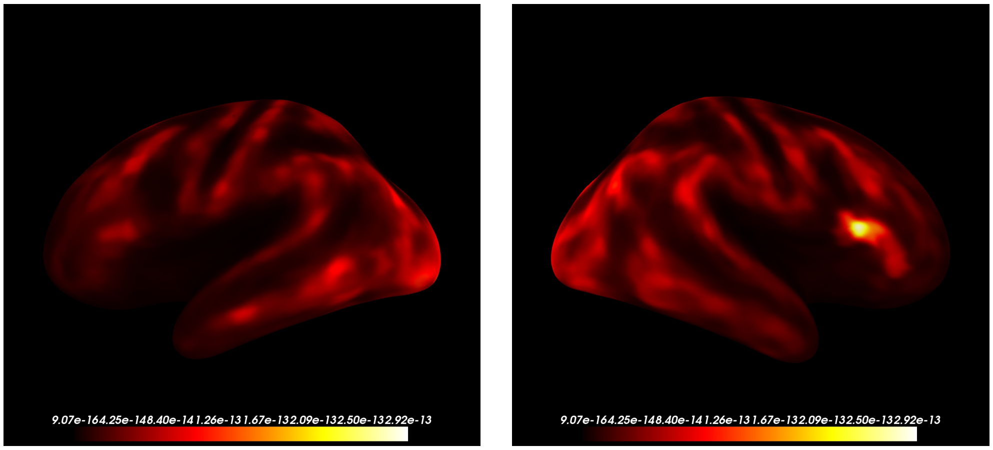


(b) Sleep NREM2

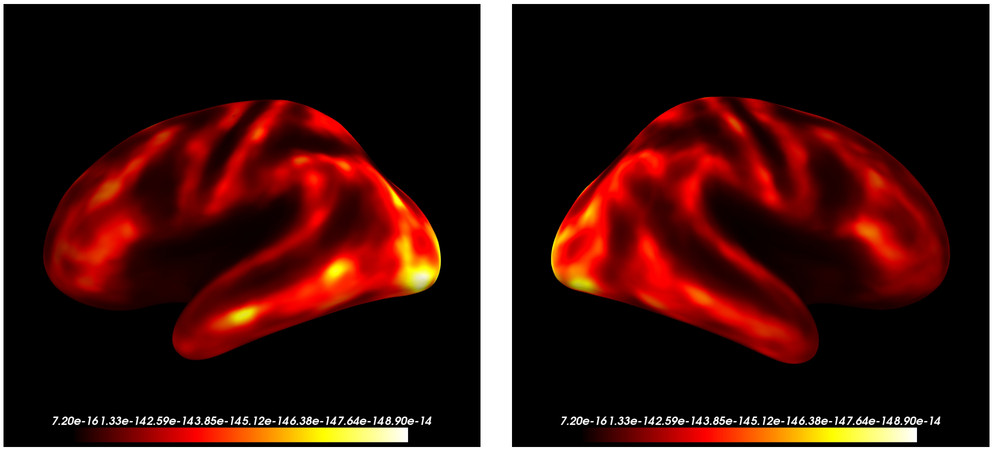


(c) Sleep NREM3

Figure 2-4: Average slow band power spectral density in source space for the sleep subjects. The power spectral density was estimated by performing source localization on the EEG data, then using multitaper spectral analysis (4s non-overlapping windows, 3 tapers, spectral resolution 1Hz) on the source-space timeseries data. For each condition of interest, the spectral estimates for each subject were morphed onto the average brain anatomy for the subjects in that condition. The average slow-band spectral density was calculated from the morphed source estimates. The slow power density is distributed across the cortex, with some concentration in the anterior and posterior regions. This analysis suggests that sleep slow oscillations are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas.



(a) Dexmedetomidine baseline



(b) Dexmedetomidine sedation

Figure 2-5: Average slow band power spectral density in source space for the dexmedetomidine subjects. The power spectral density was estimated by performing source localization on the EEG data, then using multitaper spectral analysis (4s non-overlapping windows, 3 tapers, spectral resolution 1Hz) on the source-space timeseries data. For each condition of interest, the spectral estimates for each subject were morphed onto the average brain anatomy for the subjects in that condition. The average slow-band spectral density was calculated from the morphed source estimates. The slow power density is distributed across the cortex, with some concentration in the posterior regions. This analysis suggests that dexmedetomidine slow oscillations are present throughout the cortex and could alter the functional state and functional connectivity between cortical areas.

Chapter 3

Source space functional connectivity in resting-state networks

3.1 Resting-state networks of interest

Resting-state networks refer to functionally connected brain regions which become active when subjects are not engaged in any particular focused task. Functionally connected regions are regions whose neural activity is coherent. In studies of resting state networks, subjects are typically not assigned any task, but are simply asked to fixate on a cross and remain still, or remain still with their eyes closed. Their neural activity is recorded using fMRI or EEG. fMRI tracks low-frequency BOLD (blood-oxygen level-dependent) signals, while EEG tracks voltages on the scalp, but both types of signals are related to the local field potential. There are two main approaches to analyzing the resting state data. The first is the ROI-based approach, in which neural activity is analyzed in predefined regions of interest. The second is the independent components analysis (ICA) approach, which separates out statistically independent non-Gaussian components of the signal without any prior hypothesis on where regions of interest may be located. Both ROI-based approaches and ICA have been shown to identify similar resting-state functional networks, indicating that the presence of resting state networks is a robust phenomenon. The most commonly-identified resting state networks include the default mode network (DMN), the sensorimotor network, the executive control network, the mesial visual network, the lateral fronto-parietal network, the temporo-parietal network, and the auditory network. [58] Studies have also identified a salience network distinct from the executive control network. [60]

The default mode network (DMN) is arguably the most extensively studied network. Activity within the default mode network increases when the subject is idle or resting, and decreases when the subject is performing cognitively-demanding tasks, such as tasks involving visuo-spatial reasoning or working memory. [58] Anatomical regions involved in the default mode network include the precuneus and posterior cingulate, the lateral parietal cortex, and the mesial prefrontal cortex. Implicated in introspection and awareness, the densely interconnected precuneus/posterior cingulate node is hypothesized to be a center connecting the various parts of the DMN. [16, 17, 33]

Another pair of DMN regions, studied by Buckner et al., consists of the medial temporal hippocampal formation and the dorso-medial prefrontal cortex (dMPFC). The medial temporal hippocampal formation comprises 1) the entorhinal cortex, which plays a role in declarative memories, and 2) the parahippocampal cortex, which is involved in memory encoding and retrieval. These hippocampal regions could play a role in accessing memories which are invoked when conducting mental simulations. The dMPFC is involved in constructing theory of mind. It is hypothesized that the role of the hippocampal formation in default mode activity is to access past experiences, which are then used as analogies to predict future events. The dMPFC region is not functionally correlated with the hippocampal formation, but the two regions are anatomically connected to the posterior cingulate cortex (PCC). Therefore, it is possible that the dMPFC interacts with the hippocampal formation via the PCC when prior episodic memories are being incorporated into mental simulation.[16] Greicius et al. also identified the PCC as part of the DMN, implicating it in episodic retrieval. [31] In a follow-up paper, they demonstrated structural connectivity between PCC and MPFC and PCC and medial temporal lobe using diffusion tensor imaging (DTI), thus providing a structural explanation for the function connectivity between those regions. [32]

Opposite to the DMN, the salience network identified by Seeley et al. is activated during cognitively effortful tasks. The salience network includes the dorsal anterior cingulate cortex (dACC) and orbital frontoinsular cortex (FI), and is connected to subcortical and limbic structures. The frontoinsular cortex activates during working memory tasks and is thought to be involved in interoceptive and autonomic processing. The group found that the lateral and parietal prefrontal cortices were often activated in tandem with the dACC and FI in tasks involving response selectivity, working memory and attention. Furthermore, dACC and FI response has also been observed during pain, uncertainty and other threat to homeostasis, indicating that these regions may respond to personal salience in addition to the demands of cognitive tasks. [60] For the purposes of our study, this makes the salience network an interesting network to study during different states of unconsciousness, since responsiveness to pain and external stimuli diminishes with deeper unconsciousness.

A study by Boveroux et al. showed that resting state network connectivity breaks down under propofol. In their study, subjects were tasked to make keyboard presses in response to beep sounds, and their level of consciousness was assessed by the Ramsay scale. Subjects were considered awake if they scored 2 on the Ramsay scale (clear, strong response to verbal command). They were dosed with propofol, either to induce mild sedation (Ramsay score of 3: Clear but slow response to a verbal instruction to squeeze the experimenter's hand) or to induce deeper unconsciousness (Ramsay score of 5 to 6: No response to verbal commands). It was found that connectivity in the default mode and executive networks decreased during mild sedation and decreased further during deeper unconsciousness. Normal connectivity was restored when consciousness was regained. On the other hand, connectivity within the auditory and visual networks remained unchanged under both types of propofol dosing. There was also no change in connectivity between primary visual and primary auditory cortex. [13] On the other hand, another study by Greicius et al. found that DMN functional connectivity was preserved during midazolam sedation, except in the posterior cingulate cortex. [30] These results suggest that we should expect connectivity to decrease in some resting state networks under sleep, dexmedetomidine sedation and propofol anesthesia, but that the different resting state networks may be differentially affected during these states of unconsciousness.

In our study, we chose to focus on the default mode network and the salience network since their activity is correlated with two different types of processing – introspection and mental simulation in the DMN, and response to salient stimuli in the salience network.

3.2 Source-space canonical coherence analysis

After performing MNE source localization on the EEG data timeseries (Section 2.4.2) and using the multitaper method to convert it into the frequency domain (Section 2.4.3), we performed canonical coherence analysis at the frequency of interest in accordance with the method described in [61]. We used 3 DPSS tapers to obtain independent estimates of the frequency-domain source activity. We chose 0.5Hz as the frequency of interest since it is in the middle of the slow band.

Canonical coherence analysis is the frequency-domain analogue of canonical correlation analysis. We will first explain the goal of canonical correlation analysis.[12] Given two vector spaces, $\mathcal{A}$ and $\mathcal{B}$, the first canonical correlation between $\mathcal{A}$ and $\mathcal{B}$ is given by

$$\begin{aligned}\rho_1 = \cos(\theta_1) = \max\{a_1 \cdot b_1 \quad | \quad & a_1 \in \mathcal{A},\ b_1 \in \mathcal{B}, \\ & \|a_1\| = \|b_1\| = 1\}.\end{aligned} \tag{3.1}$$

The variable θ_1 is known as the first principal angle. Geometrically, the first canonical correlation (ie. the cosine of the first principal angle) is the correlation between a maximally correlated pair of vectors, a and b, drawn from $\mathcal{A}$ and $\mathcal{B}$ respectively. The

kth canonical correlation is given by

$$\begin{aligned} \rho_k = \cos(\theta_k) = \max\{a_k \cdot b_k \quad | \quad & a_k \in \mathcal{A},\ b_k \in \mathcal{B}, \\ & \|a_k\| = \|b_k\| = 1, \\ & a_k \cdot a_i = b_k \cdot b_i = 0, foralli \in \{0, 1, \ldots, k-1\}\} \end{aligned} \tag{3.2}$$

Thus, canonical correlation analysis finds a maximally correlated pair of vectors from $\mathcal{A}$ and $\mathcal{B}$, then finds the next most correlated pair which is orthogonal to the first pair, and so forth.

For the purposes of our study, we apply the canonical correlation method to the frequency-domain source-space activity $X(f)$. For slow-band canonical coherence analysis, we select $f = 0.5$Hz, the center of the slow band, as the frequency of interest. Since $X(f)$ are frequency-domain vectors, the analysis is termed canonical coherence analysis instead of canonical correlation analysis, but the mathematics of the method remain the same. Also note that $X(f)$ are complex.

We seek to find the first canonical coherence between a pair of brain areas of interest. For a given brain area i, the tapered data $X_i^{(m)}(f)$, m=1,2,3 are treated as independent estimates of the frequency-domain source activity. Thus, we concatenate the tapered data in brain region i into a single $(n \times 1)$-dimensional vector $X_i(f)$. Denote the pair of brain regions whose canonical coherence we are computing as i and j respectively, with corresponding source activity $X_i(f)$ and $X_j(f)$.

The variables $X_i(f)$ and $X_j(f)$ are complex random vectors. To understand canonical coherence analysis, we invoke the vector interpretation of random variables. [50] [27] In this vector interpretation, a zero-mean random variable A can be thought of as a vector with length σ_A. The correlation (coherence, in frequency domain) between two zero-mean random variables can be thought of as the cosine of the angle between them. The expectation operator, $\mathbb{E}$, is the analogue of the dot product between vectors. Thus, the vector space $\mathcal{A}$ from which we search for a_1 is the span of the random variables which are the elements of the random vector $X_i(f)$. Likewise, we search for b_1 from the vector space $\mathcal{B}$, where $\mathcal{B}$ is the span of the random variables in the random vector $X_j(f)$. The span of the elements of $X_i(f)$ is defined as the linear combination $a = \sum_{k=1}^{n} \alpha_k^* X_i^{(k)}(f)$ where $X_i^{(k)}(f)$ is the *kth* element of $X_i(f)$. This can be expressed in vector notation as $a = \alpha^\dagger \cdot X_i(f)$. Likewise, the span of the elements of $X_j(f)$ is $b = \sum_{k=1}^{m} \beta_k^* X_j^{(k)}(f) = \beta^\dagger \cdot X_j(f)$, where m is the dimensionality of $X_j(f)$. Thus, for these random vectors, the first canonical coherence is given by:

$$\begin{aligned}
\rho_1 &= \cos(\theta_1) \\
&= \max\{a_1 \cdot b_1^* \mid a_1 \in \mathcal{A},\ b_1 \in \mathcal{B}, \qquad (3.3) \\
&\qquad \|a_1\| = \|b_1\| = 1\} \\
&= \max\{\mathbb{E}[a_1 b_1^*] \mid a_1 \in span(X_i(f)),\ b_1 \in span(X_j(f)), \\
&\qquad \sigma_{a_1} = \sigma_{b_1} = 1\} \qquad (3.4)
\end{aligned}$$

Note that ρ_1 can be interpreted as a coherence because $\rho_1 = \frac{\mathbb{E}[a_1 b_1^*]}{\sigma_{a_1}\sigma_{b_1}}$ since the variances σ_{a_1} and σ_{b_1} are constrained to be 1.

We follow the approach developed in [62]. ρ_1 can be expressed in terms of the cross-spectral matrix:

$$\begin{aligned}
\rho_1 &= \max\{\mathbb{E}[\alpha^\dagger X_i(f) X_j(f)^\dagger \beta] \mid a_1 \in span(X_i(f)),\ b_1 \in span(X_j(f)), \\
&\qquad \sigma_{a_1} = \sigma_{b_1} = 1\} \\
&= \max\{\alpha^\dagger S_{X_i X_j}(f) \beta \mid a_1 \in span(X_i(f)),\ b_1 \in span(X_j(f)), \\
&\qquad \sigma_{a_1} = \sigma_{b_1} = 1\} \qquad (3.5)
\end{aligned}$$

This can be solved by the method of Lagrange multipliers. For convenience, we define

$$u = S_{X_i X_i}(f)^{1/2} \alpha \qquad (3.6)$$

$$v = S_{X_j X_j}(f)^{1/2} \beta. \qquad (3.7)$$

Then

$$\alpha^\dagger S_{X_i X_j}(f) \beta = u^\dagger S_{X_i X_i}(f)^{-1/2} S_{X_i X_j}(f) S_{X_j X_j}(f)^{-1/2} v$$

$$\triangleq \quad u^\dagger J_{X_iX_j}(f)v. \tag{3.8}$$

We seek to find ρ_1. Since ρ_1 can be complex, we modify the problem slightly and instead maximize the magnitude of ρ_1, $\|\rho_1\|^2 = \rho_1\rho_1^*$. Thus, the Lagrange function is

$$\Lambda(\rho_1, u, v) = \rho_1\rho_1^* - \lambda_1(u^\dagger u - 1) - \lambda_2(v^\dagger v - 1). \tag{3.9}$$

We take the gradient of Λ and set the gradient to zero to obtain the extrema of Λ. The solution works out to be a singular value problem. The kth singular value of $J_{X_iX_j}(f)$ works out to be the kth canonical coherence, and u_k and v_k work out to be the left and right singular vectors of $J_{X_iX_j}(f)$ respectively:

$$\begin{aligned}
\rho_k &= u_k^\dagger J_{X_iX_j}(f)v_k \\
&= u_k^\dagger U\Sigma V v_k \\
&= e_k^\dagger \Sigma e_k \\
&= s_k
\end{aligned} \tag{3.10}$$

Therefore, the canonical coherence between a pair of brain regions i and j is computed by calculating

$$J_{X_iX_j}(f) = S_{X_iX_i}(f)^{-1/2} S_{X_iX_j}(f) S_{X_jX_j}(f)^{-1/2}, \tag{3.11}$$

calculating its singular value decomposition

$$J_{X_iX_j}(f) = U\Sigma V^\dagger, \tag{3.12}$$

and then taking the first singular value s_1, which is equal to the first canonical coherence.

3.3 Results: Source-space canonical coherence between the Default Mode Network and Salience Network in the slow band

With reference to the abovementioned literature on the brain regions comprising the default mode network (DMN) and salience (Sal) network, we selected all the source-space anatomical regions belonging to each network for this analysis. The sources in each network are shown in figures 3-1 and 3-2. The DMN consists of the superior frontal, rostal middle frontal, inferior parietal and middle temporal anatomical regions in the Freesurfer [1] anatomical parcellation of the cortex. The Sal network comprises the rostral and caudal anterior cingulate regions.

For each subject, we computed the canonical coherence at $f = 0.5$Hz (the middle of the slow band) between every pair of regions within and between these two networks. The results are shown for one propofol, one sleep and one dexmedetomidine subject in figures 3-3, 3-4 and 3-5 respectively. The results show that DMN-DMN, DMN-Sal and Sal-Sal connectivity decrease more during peakmax than during troughmax, sedation, NREM2 and NREM3, corresponding to the idea that disruption to neural processing is greater under anesthesia than under sedation and sleep. At peakmax, there has been a connectivity decrease in almost all DMN-DMN and Sal-Sal ROI pairs, indicating a disruption to both internal and external processing. By contrast, during NREM2 and NREM3 sleep and during dexmedetomidine sedation, Sal-Sal connections show a clear decrease in coherence, but many DMN-DMN connections maintain a coherence close to the baseline value. This correlates with the idea than external processing (by the Sal) may be more disrupted than internal processing (by the DMN) during sleep and sedation.

However, we note that the minimum coherence observed in all conditions is around 0.68, which is quite high. In the next sections, we show how this minimum coherence is caused by the inaccuracies introduced by source localization.

3.4 Source-space canonical coherence analysis accounting for the effect of the resolution matrix

The standard canonical coherence analysis described in the previous section does not account for the effect of source localization, which introduces spurious coherence between sources. The inverse solution does not reconstruct source activity with perfect accuracy; instead, activity at each source location spreads to other source locations. [36] The relationship between the true source activity x_{true} and the reconstructed source activity x is

$$x(t) = My(t) = MGx_{true}(t) \triangleq Kx_{true}(t) \tag{3.13}$$

where K is the resolution matrix. K captures the amount of activity which is spread from each source to the other source locations. K has dimensions $n_{sources} \times n_{sources}$, and $x(t)$ and $x_{true}(t)$ have dimensions $n_{sources} \times n_{samples}$. In this section, we derive source-space canonical coherence taking into account the effect of the resolution matrix and show that it is intractable to analytically remove the influence of the resolution matrix on the canonical coherence.

To apply canonical coherence analysis in source space, we seek a derivation of canonical coherence which accounts for the spurious coherence introduced by the resolution matrix. Let us begin by tapering and Fourier transforming $x(t)$:

$$X(f) = Kx_{true}(t)\mathscr{S}^{(m)}\mathscr{F} = KX_{true}(f) \tag{3.14}$$

Canonical coherence analysis is performed at a particular frequency of interest, so choosing a particular f of interest, $X(f)$ and $X_{true}(f)$ have dimensions $n_{sources}$ by $(n_{tapers} \times n_{epochs})$. Let $k = n_{tapers} \times n_{epochs}$.

Let a set of sources be called a "label". Assume that we have two non-overlapping labels of interest, label 1 and label 2, containing n_1 and n_2 sources respectively. Let the vectors of true frequency-domain source activity in labels 1 and 2 be $X^{(1)}_{true}(f)$ and

$X^{(2)}_{true}(f)$ respectively.

$$\text{Label 1: } x^{(1)}_{true},\ n_1 \text{ sources}$$
$$\text{Label 2: } x^{(2)}_{true},\ n_2 \text{ sources}$$

Let

$$X_{true}(f) = \begin{bmatrix} X^{(1)}_{true}(f) \\ X^{(2)}_{true}(f) \end{bmatrix}_{n_{sources}\times k} \tag{3.15}$$

Then the resolution matrix acts on $X_{true}(f)$ to produce the reconstructed source activity, $X(f)$, as follows:

$$\begin{aligned} X(f) &= KX_{true}(f) \\ \begin{bmatrix} (X^{(1)}(f))_{n_1\times k} \\ (X^{(2)}(f))_{n_2\times k} \end{bmatrix} &= \begin{bmatrix} (K_{11})_{n_1\times n_1} & (K_{12})_{n_1\times n_2} \\ (K_{21})_{n_2\times n_1} & (K_{22})_{n_2\times n_2} \end{bmatrix} \begin{bmatrix} (X^{(1)}_{true}(f))_{n_1\times k} \\ (X^{(2)}_{true}(f))_{n_2\times k} \end{bmatrix} \\ &= \begin{bmatrix} K_{11}X^{(1)}_{true}(f) + K_{12}X^{(2)}_{true}(f) \\ K_{21}X^{(1)}_{true}(f) + K_{22}X^{(2)}_{true}(f) \end{bmatrix} \end{aligned} \tag{3.16}$$

The first canonical coherence is given by

$$\begin{aligned} \rho_1 = \max\{\mathbb{E}[\alpha^\dagger X^{(1)}(f)X^{(2)}(f)^\dagger\beta] \mid a_1 \in span(X^{(1)}(f)),\ b_1 \in span(X^{(2)}(f)), \\ \sigma_{a_1} = \sigma_{b_1} = 1\} \end{aligned} \tag{3.17}$$

Thus, the quantity to maximize is

$$\begin{aligned} \mathbb{E}\Big[\alpha^\dagger X^{(1)}(f)X^{(2)}(f)^\dagger\beta\Big] &= \mathbb{E}\Big[\alpha^\dagger\Big(K_{11}X^{(1)}_{true}(f) + K_{12}X^{(2)}_{true}(f)\Big)\Big(K_{21}X^{(1)}_{true}(f) + K_{22}X^{(2)}_{true}(f)\Big)^\dagger\beta\Big] \\ &= \mathbb{E}\Big[\alpha^\dagger\Big(K_{11}X^{(1)}_{true}(f)X^{(1)\dagger}_{true}(f)K^\dagger_{21} + K_{11}X^{(1)}_{true}(f)X^{(2)\dagger}_{true}(f)K^\dagger_{22}+ \\ &\qquad K_{12}X^{(2)}_{true}(f)X^{(1)\dagger}_{true}(f)K^\dagger_{21} + K_{12}X^{(2)}_{true}(f)X^{(2)\dagger}_{true}(f)K^\dagger_{22}\Big)\beta\Big] \\ &= \alpha^\dagger\Big(K_{11}\mathbb{E}\Big[X^{(1)}_{true}(f)X^{(1)\dagger}_{true}(f)\Big]K^\dagger_{21}+ \\ &\qquad K_{11}\mathbb{E}\Big[X^{(1)}_{true}(f)X^{(2)\dagger}_{true}(f)\Big]K^\dagger_{22}+ \end{aligned}$$

$$
\begin{aligned}
&\qquad K_{12}\mathbb{E}\Big[X^{(2)}_{true}(f)X^{(1)\dagger}_{true}(f)\Big]K^{\dagger}_{21}+\\
&\qquad K_{12}\mathbb{E}\Big[X^{(2)}_{true}(f)X^{(2)\dagger}_{true}(f)\Big]K^{\dagger}_{22}\Big)\beta\\
&= \alpha^{\dagger}\Big(K_{11}S_{X^{(1)}_{true}(f)X^{(1)}_{true}(f)}K^{\dagger}_{21}+\\
&\qquad K_{11}S_{X^{(1)}_{true}(f)X^{(2)}_{true}(f)}K^{\dagger}_{22}+\\
&\qquad K_{12}S_{X^{(2)}_{true}(f)X^{(1)}_{true}(f)}K^{\dagger}_{21}+\\
&\qquad K_{12}S_{X^{(2)}_{true}(f)X^{(2)}_{true}(f)}K^{\dagger}_{22}\Big)\beta\\
&\triangleq \alpha^{\dagger}\Big(K_{11}S_{11}K^{\dagger}_{21}+K_{11}S_{12}K^{\dagger}_{22}+K_{12}S_{21}K^{\dagger}_{21}+K_{12}S_{22}K^{\dagger}_{22}\Big)\beta
\end{aligned}
\tag{3.18}
$$

by a similar calculation, the constraint $\sigma_{a_1} = 1$ works out to be

$$
\mathbb{E}\Big[\alpha^{\dagger}X^{(1)}(f)X^{(1)\dagger}(f)\alpha\Big] = \alpha^{\dagger}\Big(K_{11}S_{11}K^{\dagger}_{11}+K_{11}S_{12}K^{\dagger}_{12}+K_{12}S_{21}K^{\dagger}_{11}+K_{12}S_{22}K^{\dagger}_{12}\Big)\alpha = 1
\tag{3.19}
$$

and the constraint $\sigma_{b_1} = 1$ is

$$
\mathbb{E}\Big[\beta^{\dagger}X^{(2)}(f)X^{(2)\dagger}(f)\beta\Big] = \beta^{\dagger}\Big(K_{21}S_{11}K^{\dagger}_{21}+K_{21}S_{12}K^{\dagger}_{22}+K_{22}S_{21}K^{\dagger}_{21}+K_{22}S_{22}K^{\dagger}_{22}\Big)\beta = 1
\tag{3.20}
$$

We will solve this by the method of Lagrange multipliers. The objective function to maximize is

$$
C(\alpha,\beta) = \alpha^{\dagger}(K_{11}S_{11}K^{\dagger}_{21}+K_{11}S_{12}K^{\dagger}_{22}+K_{12}S_{21}K^{\dagger}_{21}+K_{12}S_{22}K^{\dagger}_{22})\beta \triangleq \alpha^{\dagger}H\beta
\tag{3.21}
$$

with the constraints

$$
\alpha^{\dagger}\Big(K_{11}S_{11}K^{\dagger}_{11}+K_{11}S_{12}K^{\dagger}_{12}+K_{12}S_{21}K^{\dagger}_{11}+K_{12}S_{22}K^{\dagger}_{12}\Big)\alpha \triangleq \alpha^{\dagger}P\alpha = 1
\tag{3.22}
$$

$$
\beta^{\dagger}\Big(K_{21}S_{11}K^{\dagger}_{21}+K_{21}S_{12}K^{\dagger}_{22}+K_{22}S_{21}K^{\dagger}_{21}+K_{22}S_{22}K^{\dagger}_{22}\Big)\beta \triangleq \beta^{\dagger}Q\beta = 1
\tag{3.23}
$$

Let

$$u = P^{1/2}\alpha \tag{3.24}$$

$$v = Q^{1/2}\beta \tag{3.25}$$

Assume that the inverses $P^{-1/2}$ and $Q^{-1/2}$ exist. In practice, these are numerically computable in the Python numpy package. Then the constraints simplify to

$$\alpha^\dagger P\alpha = u^\dagger u = 1 \tag{3.26}$$

$$\beta^\dagger Q\beta = v^\dagger v = 1 \tag{3.27}$$

And the quantity to maximize becomes

$$\begin{aligned} C(u,v) &= u^\dagger (P^{-1/2})^\dagger H Q^{-1/2} v \\ &= u^\dagger P^{-1/2} H Q^{-1/2} v \qquad (3.28) \\ &\triangleq u^\dagger F v \qquad (3.29) \end{aligned}$$

where (3.28) follows from the fact that $P^\dagger = P$.

Since C is complex, we maximize CC^*.Thus, the Lagrange function is

$$\Lambda(u, v, \lambda_1, \lambda_2) = CC^* - \lambda_1(u^\dagger u - 1) - \lambda_2(v^\dagger v - 1) \tag{3.30}$$

The maximum of CC^* is found by solving

$$\nabla_u CC^* - \lambda_1 \nabla_u (u^\dagger u) = 0 \tag{3.31}$$

$$\nabla_v CC^* - \lambda_2 \nabla_v (v^\dagger v) = 0 \tag{3.32}$$

We choose to differentiate with respect to u^* and v^*, obtaining:

$$\begin{aligned} \frac{\partial}{\partial u^*} CC^* - \lambda_1 \frac{\partial}{\partial u^*} u^\dagger u &= 0 \\ C^* F v - \lambda_1 u &= 0 \qquad (3.33) \end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial v^*}CC^* - \lambda_2 \frac{\partial}{\partial v^*}v^\dagger v &= 0\\ CF^\dagger u - \lambda_2 v &= 0\end{aligned} \tag{3.34}$$

(3.33) and (3.34) form a pair of simultaneous equations for u and v. To avoid trivial solutions, we turn this into a singular value problem. From (3.34):

$$F^\dagger u = \frac{\lambda_2}{C}v \tag{3.35}$$

Multiplying (3.33) by $F^\dagger$, we obtain:

$$F^\dagger F v = \frac{\lambda_1}{C^*}F^\dagger u \tag{3.36}$$

Substituting (3.35) in (3.36):

$$F^\dagger F v = \frac{\lambda_1}{C^*}\frac{\lambda_2}{C}v \tag{3.37}$$

Similarly, from (3.33):

$$Fv = \frac{\lambda_1}{C^*}u \tag{3.38}$$

Multiplying (3.34) by F:

$$FF^\dagger u = \frac{\lambda_2}{C}Fv \tag{3.39}$$

Substituting (3.38) in (3.39):

$$FF^\dagger u = \frac{\lambda_1}{C^*}\frac{\lambda_2}{C}u \tag{3.40}$$

Thus, the Lagrange problem reduces to a singular value decomposition of F:

$$F^\dagger F v = \frac{\lambda_1}{C^*}\frac{\lambda_2}{C}v \tag{3.41}$$

$$FF^\dagger u = \frac{\lambda_1}{C^*}\frac{\lambda_2}{C}u \tag{3.42}$$

The kth canonical coherence is given by the kth singular value of F:

$$\begin{aligned}\rho_k &= u_k^\dagger F v_k \\ &= u_k^\dagger U \Sigma V^\dagger v_k \\ &= s_k \end{aligned} \tag{3.43}$$

To summarize, the resolution matrix modifies the matrix whose singular value decomposition gives the canonical coherences. The relevant matrix is now given by

$$F = P^{-1/2} H Q^{-1/2} \tag{3.44}$$

where

$$H = K_{11}S_{11}K_{21}^\dagger + K_{11}S_{12}K_{22}^\dagger + K_{12}S_{21}K_{21}^\dagger + K_{12}S_{22}K_{22}^\dagger \tag{3.45}$$

$$P = K_{11}S_{11}K_{11}^\dagger + K_{11}S_{12}K_{12}^\dagger + K_{12}S_{21}K_{11}^\dagger + K_{12}S_{22}K_{12}^\dagger \tag{3.46}$$

$$Q = K_{21}S_{11}K_{21}^\dagger + K_{21}S_{12}K_{22}^\dagger + K_{22}S_{21}K_{21}^\dagger + K_{22}S_{22}K_{22}^\dagger \tag{3.47}$$

From the form of the H, P and Q matrices, we see that analytical removal of the impact of the resolution matrix on the canonical coherence is intractable. In the next section, we explore numerical simulations to determine whether there is a possibility of numerically removing the spurious coherence.

3.5 Simulation: MNE causes spurious canonical coherence

In this section, we use simulations to determine the amount of spurious canonical coherence introduced by the MNE method. We show that removal of this spurious coherence by numerical means such as thresholding or bootstrapping is not tractable.

We defined 7 labels on the dorsal cortex for the simulations, as shown in figure 3-6. For each label, we simulated 200 4s epochs of a sine wave at 1 Hz in the label and zero-mean Gaussian white noise elsewhere. The amplitude of the sine wave was chosen to be 1 and the standard deviation of the white noise to be 0.3, corresponding to a signal-to-noise ratio of about 3. Using the forward model computed by MNE, the simulated activity was brought up to sensor space. We used an EEG electrode montage from a Biosemi 64-EEG cap to simulate the sensor space data. The sensor-space data was then inverted back onto the cortex using the MNE inverse solution.

Figure 3-7 shows the degree of spreading out of cortical source activity caused by the MNE inverse solution. In this figure, the simulated activity was originally confined to label 6. However, after MNE inversion, the cortical activity diminishes greatly in amplitude and become spread out into other labels and over the whole cortex.

To systematically investigate the effect of MNE on canonical coherence estimation, we simulated a slow oscillation (1Hz) in each label in turn, with white noise in all other labels (signal-to-noise ratio 3). The simulation duration was 60 4s epochs. The canonical coherence between the label containing the simulation and all other labels was calculated before (ground truth) and after MNE source localization. These simulations were each repeated 20 times to obtain the median canonical coherence. These simulations were conducted with a 64, 128 and 256-electrode EEG cap to determine whether increasing the EEG electrode density improved canonical coherence estimation. Increased electrode density corresponds to denser sampling on the scalp, which might improve the spatial resolution of the source estimates.

The median canonical coherence between pairs of cortical labels before (ground truth) and after MNE source localization with a 64, 128 and 256-electrode EEG cap is displayed as the rows in each matrix in figure 3-8. The coherence after MNE localization is much higher than the ground truth, and is high between pairs of labels that should have low coherence. This is due to the tendency of MNE to spread the solution out over the whole cortex. Recovery of the original location of the signal is intractable even when the EEG electrode density is increased from 64 to 128 to 256 electrodes.

In figure 3-9 we show the median canonical coherence with standard error bars for simulations in two particular labels, labels 0 and 3. When the simulation was placed in label 0, the coherence after MNE localization increased the most in the labels closer to label 0 (labels 1, 2 and 3), and less in labels 4, 5 and 6. This is probably because labels 4, 5 and 6 were further from label 0 and more shielded from the spreading of the MNE result by the curved geometry of the brain. Increasing the number of electrodes from 64 to 128 (or 256) brought the coherence closer to the ground truth in labels 4, 5 and 6, but did not appreciably improve the coherence estimation in labels 1, 2 and 3. When the simulation was placed in label 3, the coherence in all labels after localization became high, making it difficult to tell which label contained the original signal. The spurious coherence in label 0 was reduced as the number of electrodes was increased from 64 to 128 to 256, but there was no appreciable improvement in the accuracy of the coherence estimation in the other labels. This is probably because label 3 was located in a central, flat region of the dorsal cortex, so the MNE solution spread out to all the other labels on the cortex.

The results show that removal of the spurious canonical coherence by numerical methods is not tractable. Even if the average 'baseline' canonical coherence can be removed by thresholding, it may not be possible to tell where the original signal was, especially in flatter areas of the cortex. For instance, when the simulated activity is in label 3, the activity becomes spread almost uniformly into the other labels. Furthermore, attempts to remove the spurious canonical coherence by techniques such as bootstrapping to find coherences significantly above the average value are unlikely to succeed since the overall coherence in multiple brain regions is highly elevated after MNE source localization.

3.6 Summary: Limitations of MNE source localization

We chose MNE to model the slow oscillation because of its tendency to produce distributed source activity, but we have shown that the method introduces a large amount of spurious canonical coherence which is analytically and numerically intractable to remove. Thus, spurious coherence can only be reduced by improving the resolution of source estimates, which corresponds to a resolution matrix that is closer to the identity matrix.

The widely distributed solutions produced by MNE do not appear to match the actual distribution of source activity. Additionally, MNE is known to result in depth bias, which is the phenomenon in which sources located deep in the cortex are incorrectly localized to superficial regions. Gaussian or Laplacian priors can be specified beforehand to reduce the depth bias, as is the case in methods such as weighted MNE, dynamic statistical parametric maps (dSPM), and standard low-resolution electromagnetic tomography (sLORETA). [56] A study by Bradley et al. showed that localization accuracy could be improved by using LORETA with a p-norm of 1.5, or a combination of sLORETA to identify the source with the highest activity and LORETA (p=1.5) to find other active sources. [14] Another approach to improve localization accuracy is to tailor the prior to the dataset, using approaches such as sparse Bayesian learning (SBL) and automatic relevance determination (ARD). These approaches learn the prior from the data, compensating for the fact that the full posterior may not be accurately represented by the posterior mode. [56] Yet another way to encourage sparse solutions is by using an L1-norm in place of the L2-norm.

The various MAP source space methods are nonparametric, since they do not make assumptions on the generative process behind the data. These approaches may obtain the most plausible naive solution when the underlying generative process is unknown. However since we are looking for the slow oscillation in particular, we are able to parametrically model the oscillation we aim to localize. Therefore, for our particular application, reasonable assumptions can be made about the form of the generative process, which we expect will improve source localization performance better than nonparametric methods. In the next chapter, we use a state-space modelling approach to improve the accuracy of source localization.

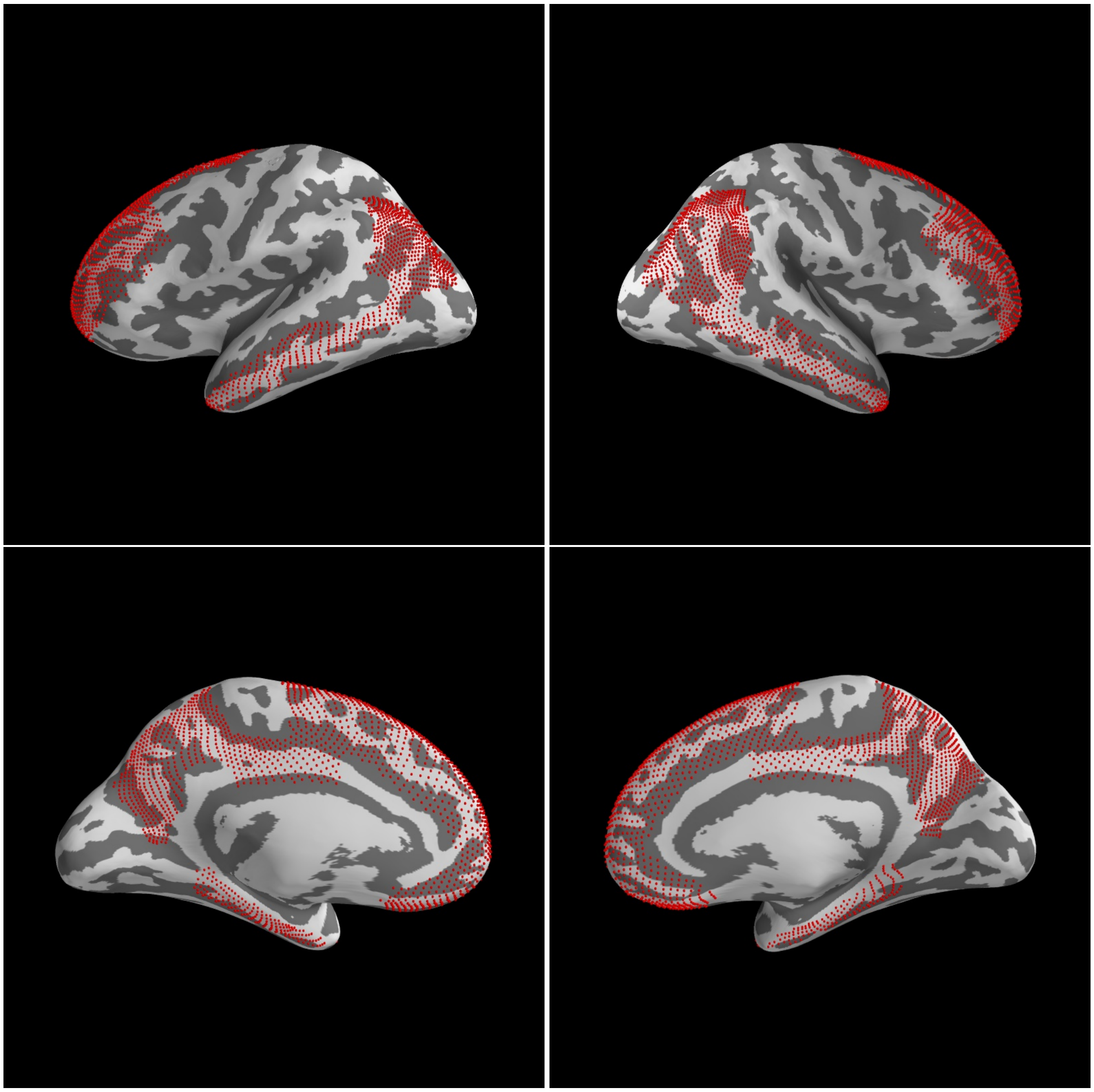


Figure 3-1: Sources (red points) corresponding to the default mode network in an ico-3 parcellation of the cortex. The regions included in the default mode network for this analysis include the superior frontal, rostral middle frontal, inferior parietal, middle temporal, medial orbitofrontal, precuneus, posterior cingulate, parahippocampal and entorhinal cortices.

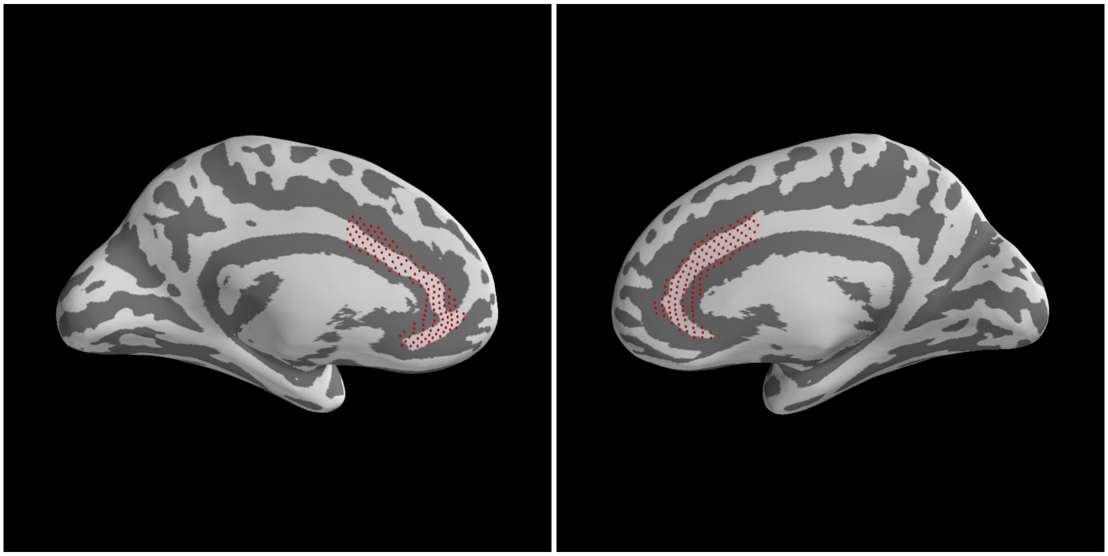


Figure 3-2: Sources (red points) corresponding to the salience network in an ico-3 parcellation of the cortex. The regions included in the salience network for this analysis include the rostral and caudal anterior cingulate cortices.

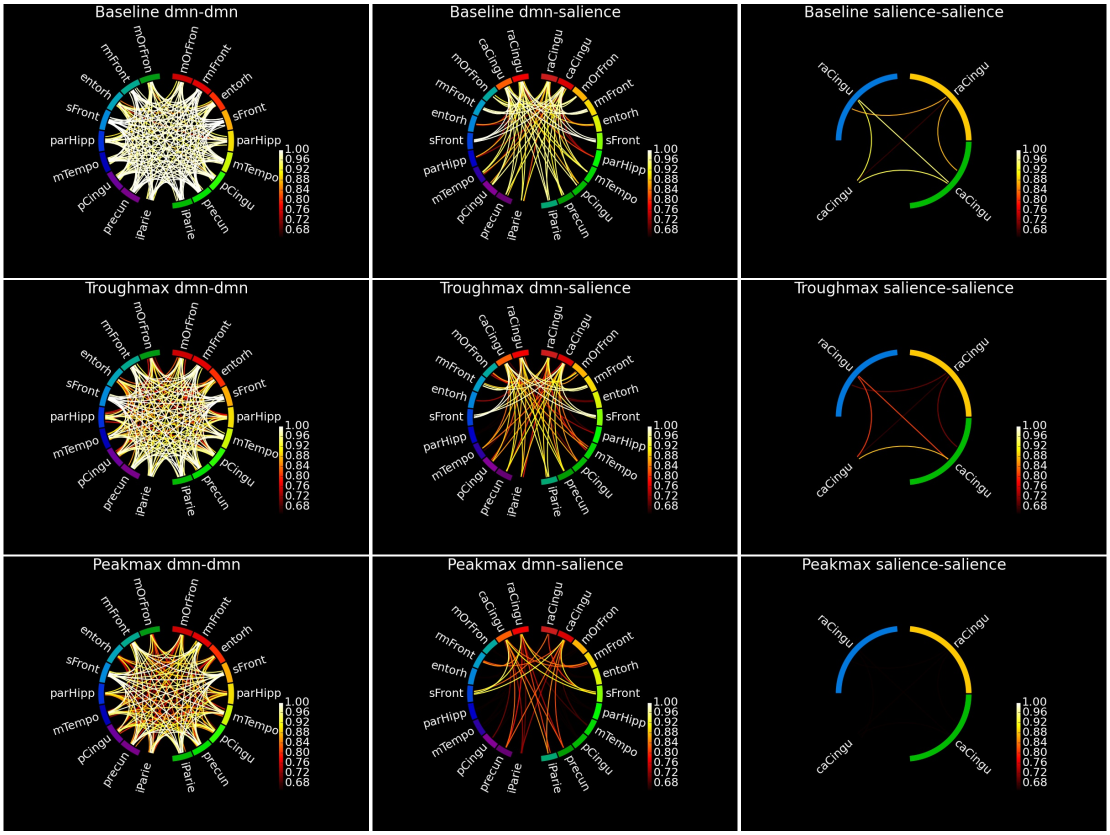


Figure 3-3: Canonical coherence (CC) at 0.5Hz of the MNE localized results for one propofol subject. CC within and between the DMN and Sal networks decreases more during peakmax than during troughmax.

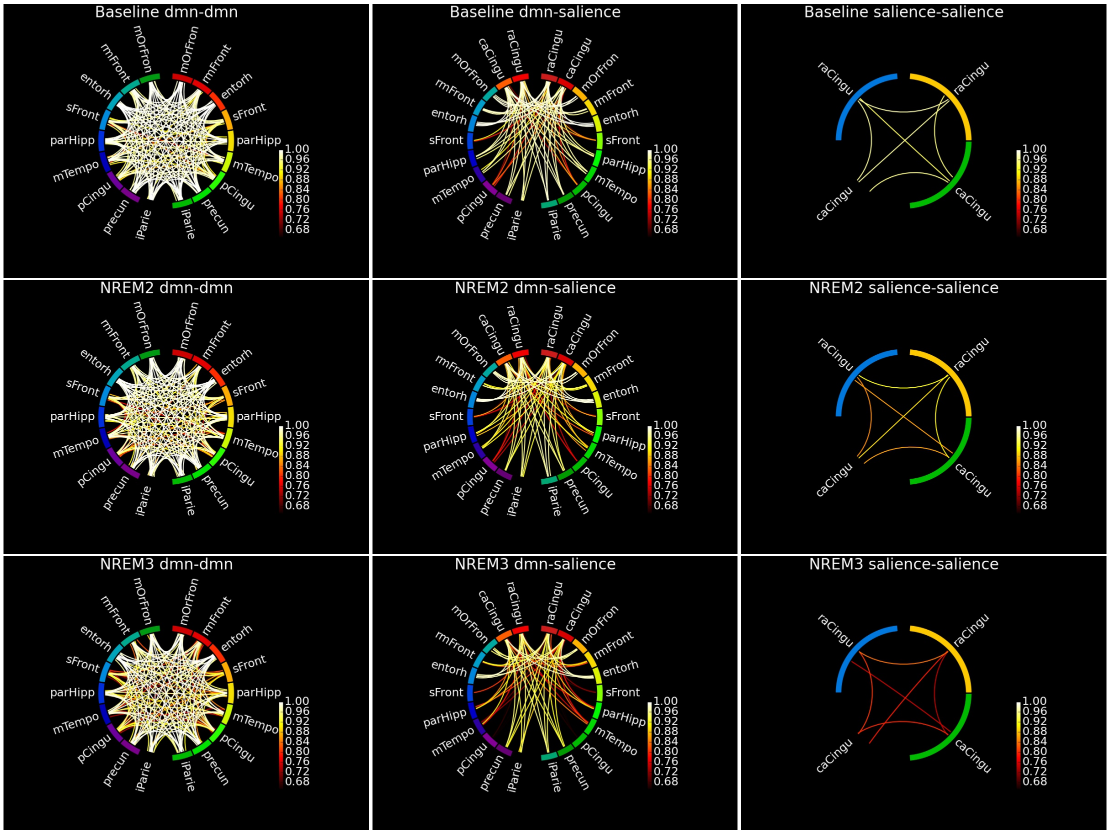


Figure 3-4: Canonical coherence (CC) at 0.5Hz of the MNE localized results for one sleep subject. The CC in the Sal network decreases more prominently than in the DMN as the state of unconsciousness deepens from baseline to NREM2 to NREM3.

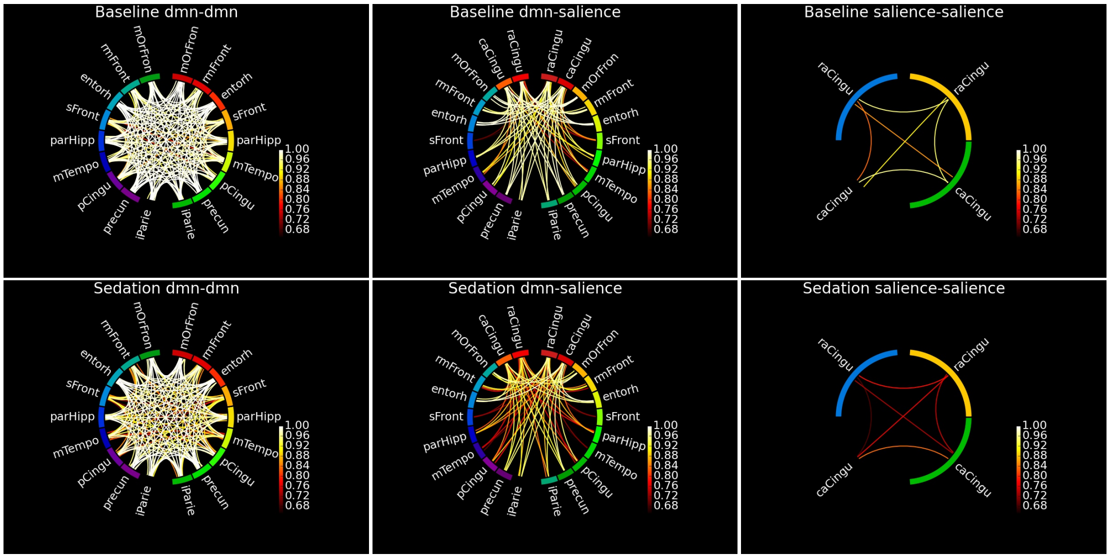


Figure 3-5: Canonical coherence (CC) at 0.5Hz of the MNE localized results for one dexmedetomidine subject. The CC in the Sal network decreases more prominently than in the DMN as the state of unconsciousness deepens from baseline to dexmedetomidine sedation.

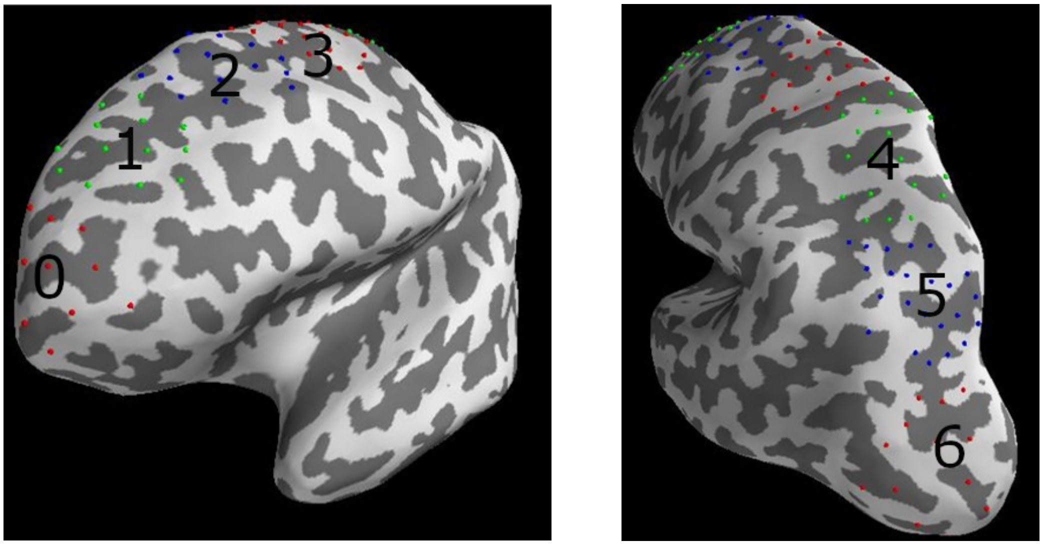
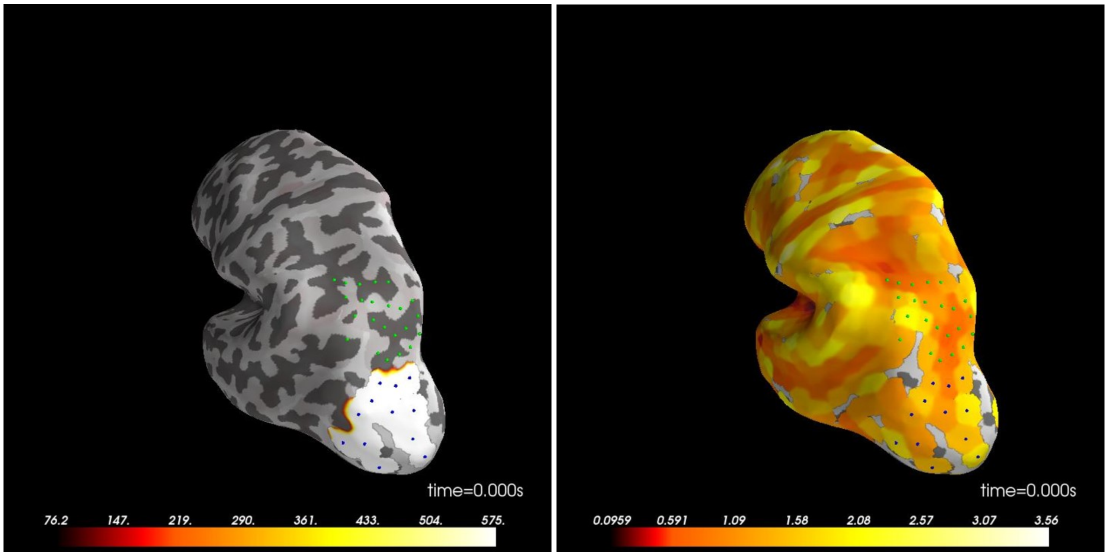


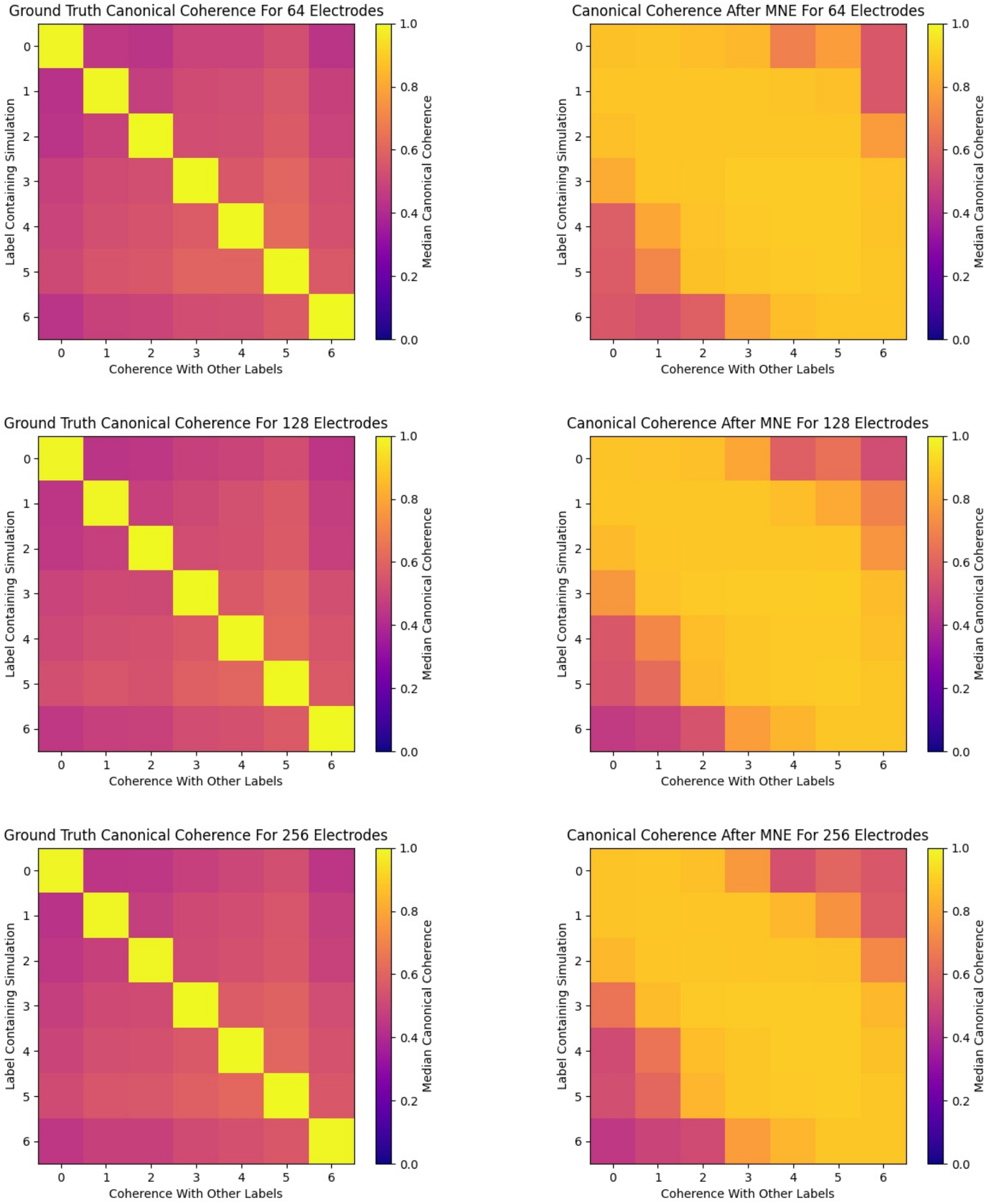
Figure 3-6: Labels defined on the brain for the canonical coherence simulations. The labels are arranged sequentially from anterior to posterior to help characterize changes in canonical coherence as the distance between labels increases.



(a) Before MNE

(b) After MNE

Figure 3-7: Example of MNE distributed solution. (a): A simulated signal was generated in label 6 (1Hz oscillation, a single 4s window, white noise elsewhere, signal-to-noise ratio of 3). (b): After MNE source localization, the estimated activity appears distributed across all other labels, covering much of the cerebral cortex. These results show that MNE may not be able to accurately reconstruct the spatial location of cortical source activity.


Ground Truth Canonical Coherence For 64 Electrodes

Canonical Coherence After MNE For 64 Electrodes

Ground Truth Canonical Coherence For 128 Electrodes

Canonical Coherence After MNE For 128 Electrodes

Ground Truth Canonical Coherence For 256 Electrodes

Canonical Coherence After MNE For 256 Electrodes

Figure 3-8: Median canonical coherence between pairs of cortical labels before (ground truth) and after MNE source localization with a 64, 128 and 256-electrode EEG cap. A slow oscillation (1 Hz) was simulated in each label in turn, with white noise in all other labels (signal-to-noise ratio 3). The simulation duration was 60 4s epochs. The canonical coherence between the label containing the simulation and all other labels was calculated before (ground truth) and after MNE source localization. These simulations were each repeated 20 times to obtain the median canonical coherence. The results are displayed as the rows in each matrix. The coherence after MNE localization is much higher than the ground truth, and is high between pairs of labels that should have low coherence. This is due to the tendency of MNE to spread the solution out over the whole cortex. Recovery of the original location of the signal is intractable even when the EEG electrode density is increased from 64 to 128 to 256 electrodes.

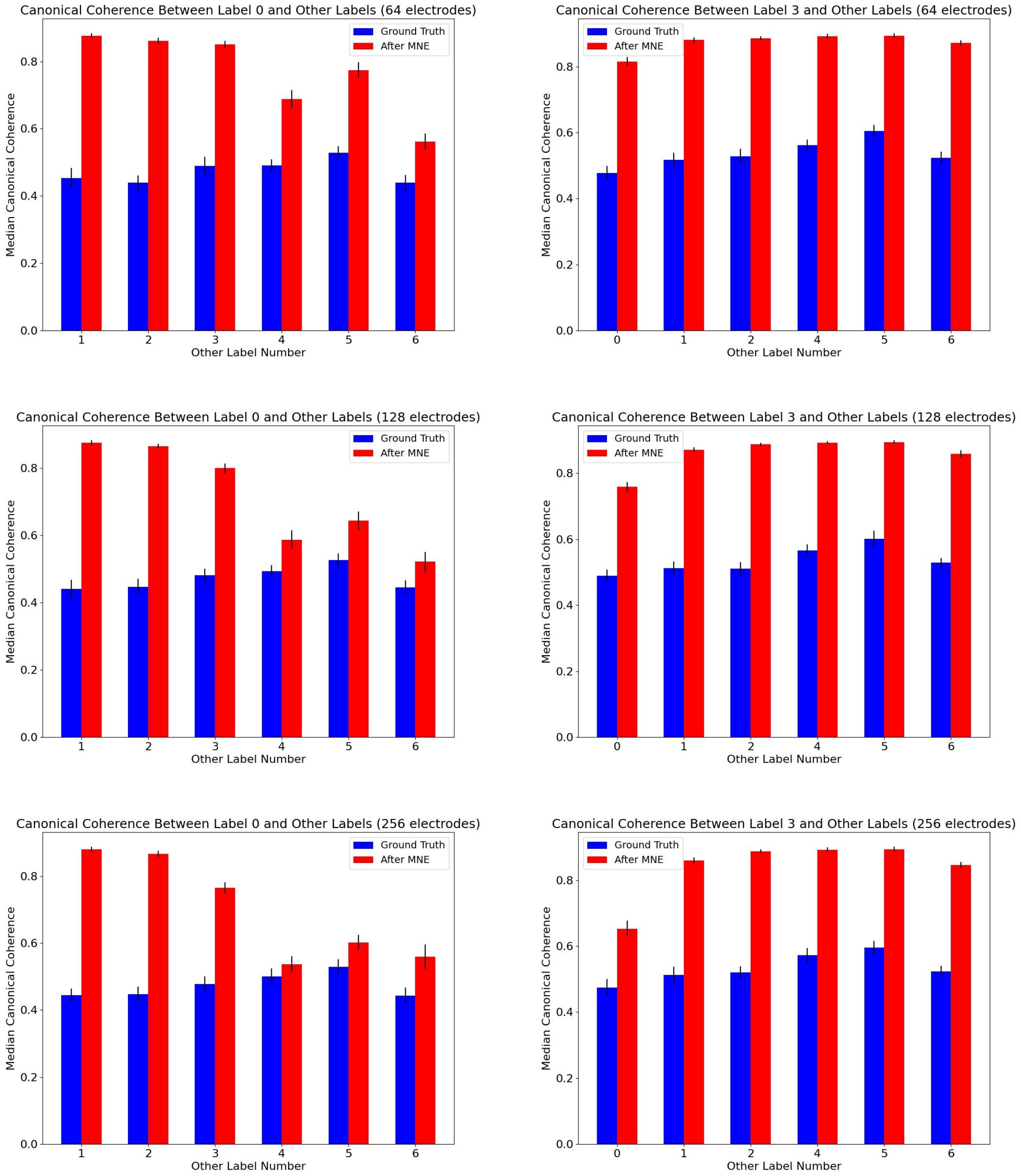


Figure 3-9: A simulation (1Hz oscillation, 60 4s epochs, white noise with SNR 3, 20 runs of simulation) was placed in a particular label with 64, 128 and 256 electrode setups. Other labels contained white noise. The figures show the median and standard deviation in canonical coherence between the label containing the simulation and all other labels. 1st column: Simulation in label 0. After localization, the coherence increased the most in the labels closer to label 0, and less in the labels which were further from label 0 and more shielded from the spreading of the MNE result by the geometry of the brain. Increasing the number of electrodes from 64 to 128 (or 256) brought the coherence closer to the ground truth in labels 4, 5 and 6. 2nd column: Simulation in label 3. After localization, the coherence has increased in all labels. As the number of electrodes increases from 64 to 128 to 256, the spurious coherence in label 0 is reduced. However, the coherence in all the labels is high, making it difficult to tell which label contained the original signal.

Chapter 4

Improving source localization with a state-space oscillator model

4.1 The Matsuda-Komaki oscillator model

To improve source localization performance, we used a parametric method involving a state space model to model the temporal dynamics of neural oscillations. The Matsuda-Komaki oscillator state-space model, as described in [46], has the form

$$\begin{aligned} \begin{pmatrix} x_{t+1,1} \\ x_{t+1,2} \end{pmatrix} &= a \overbrace{\begin{pmatrix} \cos(2\pi f \Delta t) & -\sin(2\pi f \Delta t) \\ \sin(2\pi f \Delta t) & \cos(2\pi f \Delta t) \end{pmatrix}}^{F_f} \begin{pmatrix} x_{t,1} \\ x_{t,2} \end{pmatrix} + \begin{pmatrix} v_{t,1} \\ v_{t,2} \end{pmatrix} && (4.1) \\ y_t &= G x_{t,1} + w_t && (4.2) \end{aligned}$$

where

$$\begin{aligned} \begin{pmatrix} v_{t,1} \\ v_{t,2} \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right) && (4.3) \\ w_t &\sim N(0, \tau^2). && (4.4) \end{aligned}$$

Equation (4.1) is known as the state equation, and equation (4.2) is called the observation equation. The state x cannot be observed directly, but is assumed to evolve according to the state equation. In this formulation, the state equation describes an oscillation in time, corresponding to a neural oscillation. $x_{t,1}$ is the real part of the state at time t and $x_{t,2}$ is the imaginary part. y_t is the observed measurement at time t, related to the hidden state by G. Both the state and observation equation contain zero-mean Gaussian noise, which captures the uncertainty in the state's evolution as well as the noise in the data measurement.

This formulation has been shown to perform well when applied to real data, and has several advantages. Firstly, it accounts for measurement noise, which is always present in real data. Secondly, it does not require bandpassing, which can cause edge artifacts in the signal. Thirdly, the phase of the signal can be readily estimated from the real and imaginary parts of the fitted vector x. [46]

4.2 The oscillator-EM method: Source localization under the oscillator model

For our problem of interest, x is the vector of source-space activity decomposed into real and imaginary parts, and y is the EEG sensor data. G is the forward model, augmented such that every even column is 0, corresponding to not observing the imaginary part of the signal. Since neural data contains several oscillations, one rotation matrix is set up for each oscillation frequency to be fit by the model. The overall state space model is constructed by placing the rotation matrices for each oscillation F_f on the diagonals of a large F matrix. For instance, to fit a model containing a slow and an alpha oscillation, the vector of source activity x is duplicated and concatenated vertically with itself. A large F matrix is constructed to fit the slow and alpha oscillations:

$$\begin{pmatrix} \vec{y} \\ \vec{y} \end{pmatrix} = a \begin{pmatrix} \underline{F}_{slow} & \underline{0} \\ \underline{0} & \underline{F}_{alpha} \end{pmatrix} \begin{pmatrix} \vec{x} \\ \vec{x} \end{pmatrix}. \tag{4.5}$$

Thus, our model has the overall form

$$x_t = Fx_{t-1} + w_t \tag{4.6}$$

$$y_t = Gx_t + v_t \tag{4.7}$$

where

$$v_t \sim N(0, C) \tag{4.8}$$

$$w_t \sim N(0, Q). \tag{4.9}$$

An expectation maximization (EM) algorithm is used to estimate the parameters of the state-space model based on the data. During the E-step, a Kalman filter and fixed interval smoother are used to estimate the state x and calculate the complete data likelihood. During the M-step, parameters of interest such as C, Q and F are estimated by maximizing the complete data likelihood with respect to parameters. The E and M steps are performed iteratively until convergence in the likelihood is reached.

We conducted initial validation studies of the model by simulating activity from the state equation at a frequency of 10Hz. Due to the nonlinear form of F, we used the Newton-Raphson numerical maximization method in the M-step for finding the parameters of F, namely the amplitude a and frequency f. The EM algorithm was able to accurately recover the oscillation amplitude and frequency. We next moved on to source space simulations and demonstrated that this oscillator-EM source localization method outperforms MNE in terms of its accuracy at reconstructing the spatial location of the oscillation(s) (Section 4.3).

After the model had been validated in simulations, we applied it to the real EEG data from the 5 conditions of interest. For the real data, our approach was to first determine the oscillation frequencies present at the scalp level, and then source localize with those identified frequencies plugged into the F matrix. The EM algorithm was then used to estimate C and Q, assuming F was known based on the sensor-space fits (Section 4.4). Finally, based on the oscillator-EM localization results, we recalculated the source-space canonical coherence between resting state networks (Section 4.5).

4.3 Simulation studies: The oscillator-EM method outperforms MNE at source localization

4.3.1 Simulation of a single slow oscillation

We first simulated 4 seconds of a slow oscillation at 1Hz in four rostral sources in the frontal cortex. White noise was simulated in all other sources, with a signal-to-noise ratio of 100. A sampling frequency of $Fs = 100$Hz was used. We projected the simulated activity up to the scalp level using the forward model. We tested the accuracy of the oscillator-EM method by using it to invert the scalp-level sensor space activity back onto the cortex, and comparing the result with the simulated ground truth. We used 10 iterations of the EM algorithm. For comparison, we also used MNE to invert the scalp-level activity. The results are shown in figure 4-1 for one snapshot in time. The oscillator-EM method outperforms MNE at recovering the spatial location of the slow oscillation. As shown in the previous chapter, MNE spreads the solution out over the entire cortex. The highly local simulated slow oscillation is completely overwhelmed by noise in the MNE estimate and cannot be recovered. On the other hand, the oscillator-EM estimate becomes spatially more accurate with each iteration of the EM algorithm. After 10 EM iterations, the location of the simulation oscillation is largely well-recovered. MNE also significantly reduces the estimated current amplitudes on the cortex, whereas oscillator-EM does not.

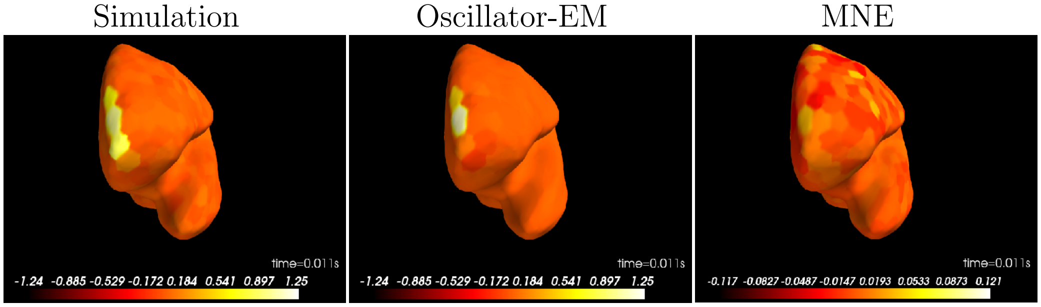


Figure 4-1: Comparing oscillator-EM localization of a simulated slow oscillation with MNE. A 1Hz oscillation travelling at 3.6ms^{-1} was simulated in a series of 4 adjacent sources in the frontal cortex (white noise elsewhere, signal-to-noise ratio 100). The oscillator-EM algorithm appears to recover the amplitude and spatial location of the simulated source with greater accuracy than MNE.

4.3.2 Simulation of a slow and an alpha oscillation in an overlapping region

In practice, real data usually contains multiple oscillations, so we sought to test whether the oscillator-EM method could separate oscillations at two distinct frequencies occurring at the same spatial location. In the same four rostral sources, we simulated a slow and an alpha oscillation of equal amplitude (both the amplitudes were set to 1). We again brought the source-space data up to sensor space using the forward model, and inverted it back onto the cortex using oscillator-EM with 10 iterations of the EM algorithm. This time, the oscillator-EM model used for the fits contained both F_{slow} and F_{alpha} in the F matrix. The results, shown in figure 4-2, demonstrated that oscillator-EM was able to recover both the slow and alpha in the correct spatial location, even though the two oscillations had completely overlapped in space. This is an encouraging result because we expect spatial overlap between oscillations at different frequencies to occur in real data. This simulation has shown that spatial overlap with other frequencies is unlikely to affect localization of the slow oscillation in real data using oscillator-EM.

4.3.3 Simulations in different cortical regions

We next expanded our simulation studies to different regions on the cortex. While the oscillator-EM method worked well in the rostral region, regions further away from the scalp and regions located in cortical sulci are expected to be more difficult to localize in, since activity in those sources is more difficult to observe from the scalp. Thus, we simulated a slow and an alpha oscillation of equal amplitudes (both 1) in various cortical regions.

In figure 4-3 we explore oscillator-EM localization in several dorsal regions. We expect localization to be better for dorsal regions which are close to the scalp. The localization results were the best in the rostral region, reasonably good in a central dorsal region, and not as good in a posterior region. Due to the curvature of the brain, the posterior region is located further away from the EEG sensors on the scalp, so it is reasonable that localization performance is not as good in that region.

In figure 4-4 we choose lateral regions for the simulation. The first two rows show that the oscillator-EM estimate is more spread out for these lateral regions than for the dorsal regions previously tested, probably due to the fact that these lateral regions are located further away from the scalp EEG sensors, and thus the sources in them are less easy to detect. On the whole, however, the performance is still better than MNE because the amount of spreading over the cortical surface is less. The last row of the figure shows that the sources hidden under the curved surface of the frontal lobe cannot be detected at all by the EEG sensors, which is not surprising; however, surprisingly, both the slow and alpha oscillations in the single source on the temporal lobe are well-detected.

In figure 4-5 we compare localization in a dorsal gyrus with localization in a dorsal sulcus. Localization performance is better in the gyrus; in the sulcus simulation, the oscillator-EM estimate is much more spread out over the cortex. This is as expected since the sources in a gyrus are less concealed from the scalp than the sources in a sulcus. Nevertheless, even in the tricky case of localizing in a sulcus, the oscillator-EM estimate spreads out less over the cortex than MNE for both of the oscillations localized.

4.3.4 Simulations of slow and alpha oscillations with varying relative amplitudes

In our final set of simulations, we tested the effect of varying the relative slow and alpha oscillation amplitudes on the oscillator-EM localization performance. These simulations were carried out in the various regions tested in section 4.3.3. The results for the rostral region are shown in figure 4-6. When the slow amplitude is half the alpha amplitude (0.5 is to 1), the localization of the slow becomes less accurate. We observe simular results in the other regions as well; varying the relative oscillator amplitudes may decrease localization performance for one or both oscillations. Nevertheless, the spatial accuracy of the recovered source estimates is still better than that obtained by MNE.

4.4 Oscillator-EM source localization applied to the EEG data

4.4.1 Finding neural oscillations using the Matsuda model in EEG sensor space

After validating the oscillator-EM method in simulations, we applied it to the real EEG data from the 5 conditions of interest. We first used the Matsuda model to fit neural oscillations in sensor space, following the approach developed by Beck et al. [10]. For each subject from each condition, we used one 4s epoch of EEG data. The iterative procedure described in [10] was used to identify the frequencies and amplitudes of oscillations present in each clean EEG channel. The iterative procedure increases the number of oscillations in the model one at a time, and calculates the AIC of each model. The model with the lowest AIC is chosen to be the best model for that electrode.

Based on this procedure, we constructed a histogram of frequencies found for each subject. Examples for a representative subject from each of the 5 conditions are shown in figure 4-7. In each subfigure, the top panel shows all the frequencies found in the 0-4Hz range, and the bottom panel shows all the frequencies found in the 7-16Hz range. From the histograms, we see that a single peak in the slow-delta band (0-4Hz) and 1-2 peaks in the alpha band (7-16Hz) are identified for each subject. In propofol peakmax there is only one alpha frequency with a median of around 10Hz, whereas for the other conditions there are both a "slow" alpha (<10Hz) and a "fast" alpha (>10Hz). In some of the other subjects, there was only 1 alpha frequency in the other conditions. Each subject's histogram was visually inspected. The median of the slow frequencies was used as the identified value of the slow frequency for the subject. If there were two alpha peaks, the median of the peak below 10Hz was taken as the first alpha frequency, and the median of the peak above 10Hz as the second alpha frequency; otherwise, the median of the single alpha peak was used as the alpha frequency.

It is interesting to note from figure 4-7 that in NREM2 sleep, NREM3 sleep and dexmedetomidine sedation, the peak found in the 0-4Hz range was at 1Hz or slightly above 1Hz. This slow frequency is higher than the median slow frequency found in propofol troughmax and peakmax. It is possible that for the dexmedetomidine and sleep data, the oscillator model is explaining the slow and delta oscillations by a single frequency peak instead of two peaks. As we have reviewed in section 1.2.7, the slow and delta oscillations could be generated by different cortical or thalamocortical mechanisms. Further work could be done to determine how well the oscillator model can separate the slow and delta oscillations, given their close proximity in frequency.

4.4.2 Source localization of real data with oscillator-EM

The slow and alpha frequencies identified by the procedure described in section 4.4.1 were input into the F matrix of the state equation in oscillator-EM source-space localization algorithm. For the analysis of real data, we chose to fit the source and noise covariance matrices Q and C, but not the parameters of the oscillators in F.

This is because fitting the oscillator amplitudes and frequencies in source space adds an additional possible source of error and uncertainty. Since the data was collected in EEG sensor space, the EEG-level fits of the oscillator parameters are likely to be the most accurate estimates we can obtain.

For each of the 5 conditions, we selected 90-100 noncontiguous 4s epochs of clean data from the various clean subjects. As far as possible, the number of epochs selected from each clean subject was kept about the same in order to get a good estimate for the data across subjects. In each epoch, we downsampled the data to 100Hz and ran 10 iterations of the oscillator-EM algorithm to arrive at our source estimates for the data.

4.5 Source-space canonical coherence in resting state networks under oscillator-EM source localization

Using the oscillator-EM localization results for the data in each epoch, we calculated the canonical coherence (CC) between and within the default mode network (DMN) and salience network (Sal) at 0.5Hz. Using 3 DPSS tapers with about 100 epochs in each condition, we obtained around 300 independent estimates of the frequency-domain source activity for each condition, and calculated the CC as before (Section 3.2).

The results are shown in figures 4-8, 4-9, 4-10 and 4-11. The histogram of CC values between pairs of brain regions (figure 4-8) shows the opposite trend to the CC results from MNE source localization. The CC values are skewed closer to 1 in peakmax and NREM3, the deeper states of unconsciousness in propofol and sleep respectively, and are lower in the lighter states of unconsciousness. Visualizations of the CC in each condition show that peakmax has overall higher coherence within and between DMN and Sal than troughmax (figure 4-9), NREM3 has overall higher coherence than NREM2 in both networks (figure 4-10), and dexmedetomidine sedation has CC values similar to NREM2 and troughmax. Notably, the minimum value of canonical coherence found in the real data with the oscillator-EM model was about 0.10, much lower than the MNE minimum of 0.68. This likely reflects the reduction in spurious coherence under oscillator-EM.

We hypothesize why the oscillator-EM estimates produce high coherence in the deeper states of unconsciousness, contrary to our expectations. The oscillator-EM model is better able to spatially localize oscillations than MNE, but a 4s window may not be enough time to estimate the phase relationships between oscillations in different parts of the cortex. As can be seen from the simulation results in section 4.3, for instance in figure 4-2, there is a time lag between the ground truth simulation and the recovered slow and alpha oscillations; peaks and troughs in the recovered slow and alpha occur a few time points after they have occurred in the simulated data. This lag is to be expected since the Kalman filter and fixed interval smoother become more accurate after aggregating data from a longer time period. A 4s epoch may not be enough to recover the true phase relationships between the slow oscillation in various regions.

It is encouraging that oscillator-EM works well in simulations, but we plan to move forward with further tests of the accuracy of the oscillator-EM canonical coherence results in real data. One possibility is to conduct simulations in which we test whether oscillator-EM can accurately recover a known ground truth canonical coherence. Another possibility is to select data epochs longer than 4s and see if the oscillator-EM CC results change when more data is used for the estimate. Thirdly, we can test oscillator-EM with simpler meaures of functional connectivity such as the phase-locking factor used in Lewis et al. [42], or the simple pairwise coherence between pairs of single sources randomly selected from within the DMN and Sal regions of interest.

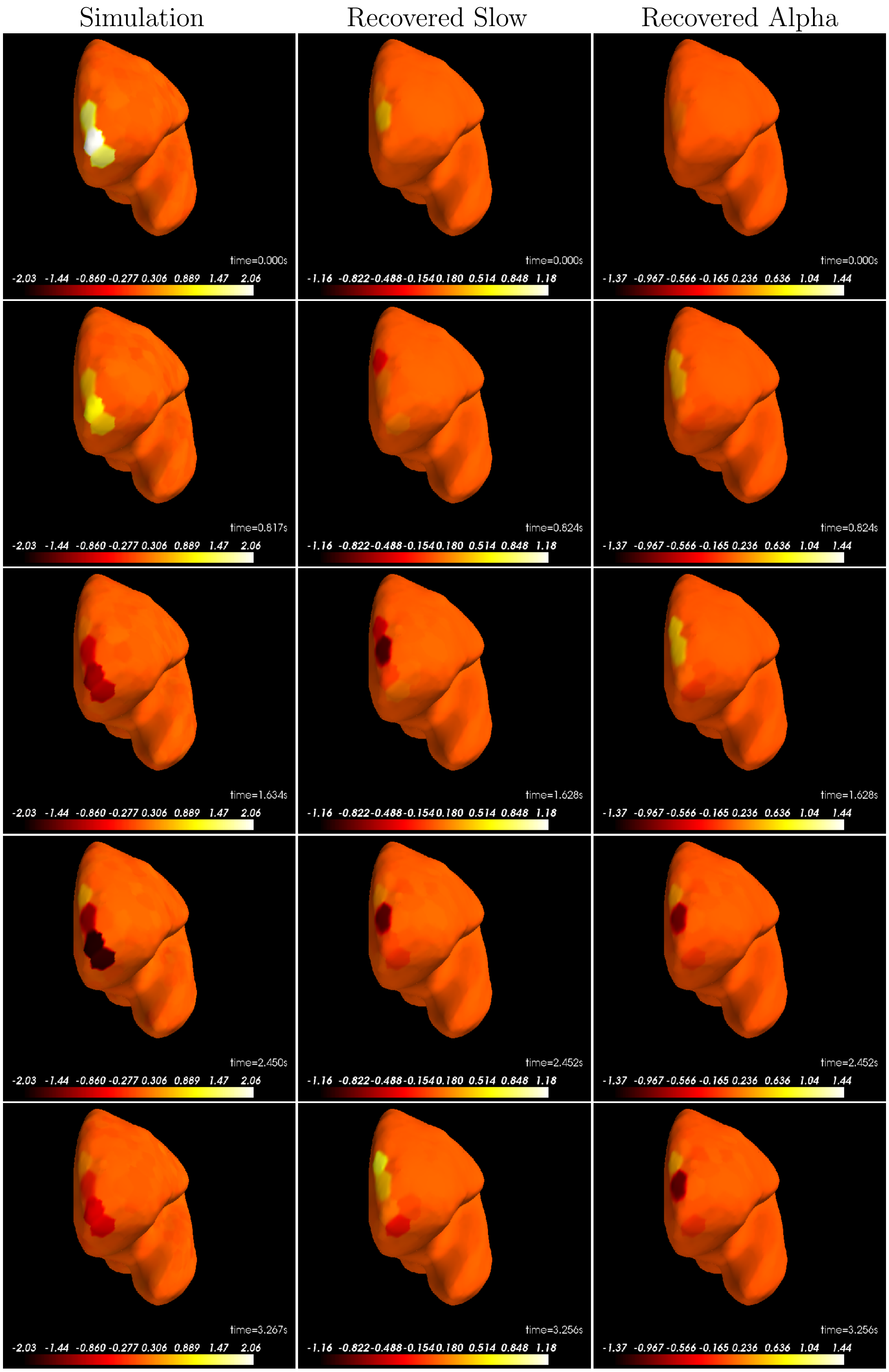


Figure 4-2: Columns: Simulation of spatially overlapping slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Rows: Results at 5 time snapshots. The oscillator-EM method successfully recovers both oscillations despite their spatial overlap.

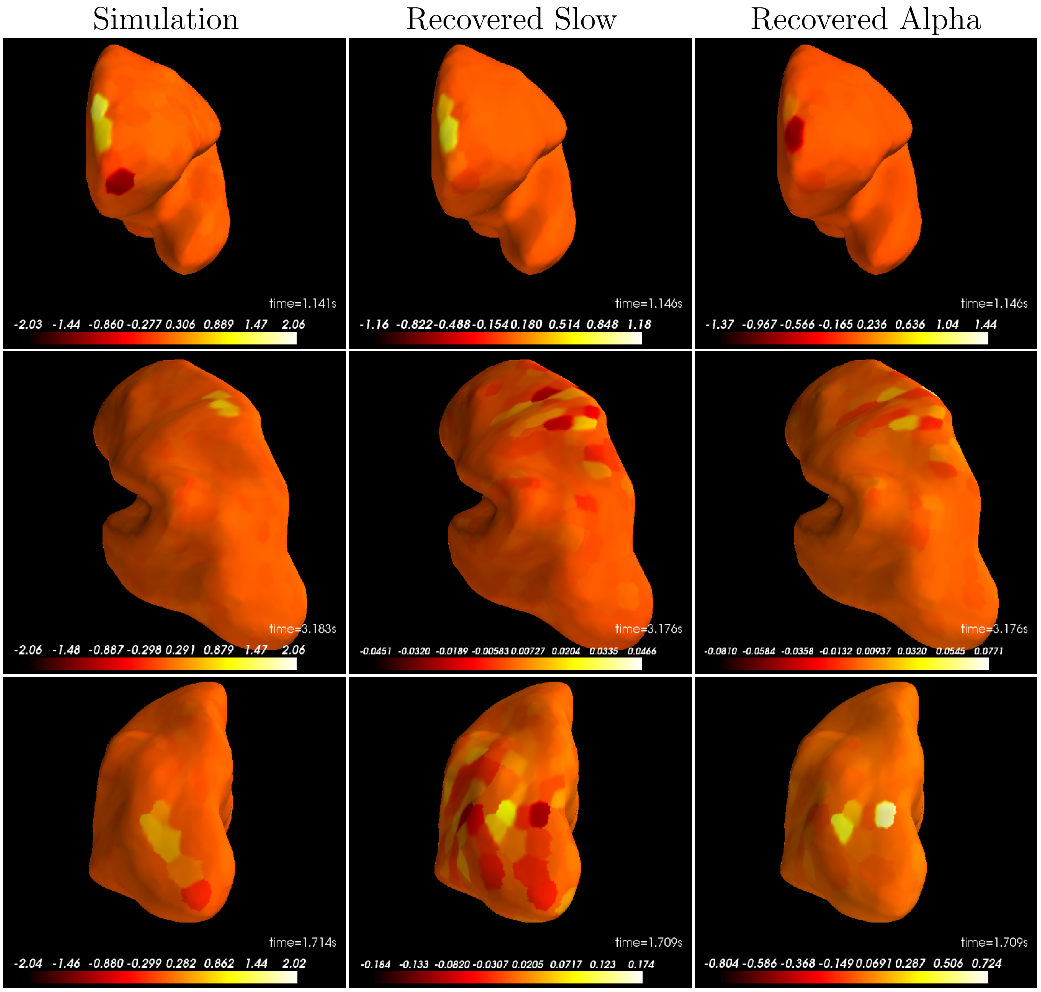


Figure 4-3: Columns: Simulation of slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Rows: Different views from anterior, lateral, and posterior perspectives. Performance of the oscillator-EM method varies with spatial location on the cortex. The oscillations in the frontal cortex (1st row) are localized best. Localization is less precise in a dorsal (2nd row) and posterior (3rd row) region. However, the spatial location recovered is still more precise than with MNE (not shown).

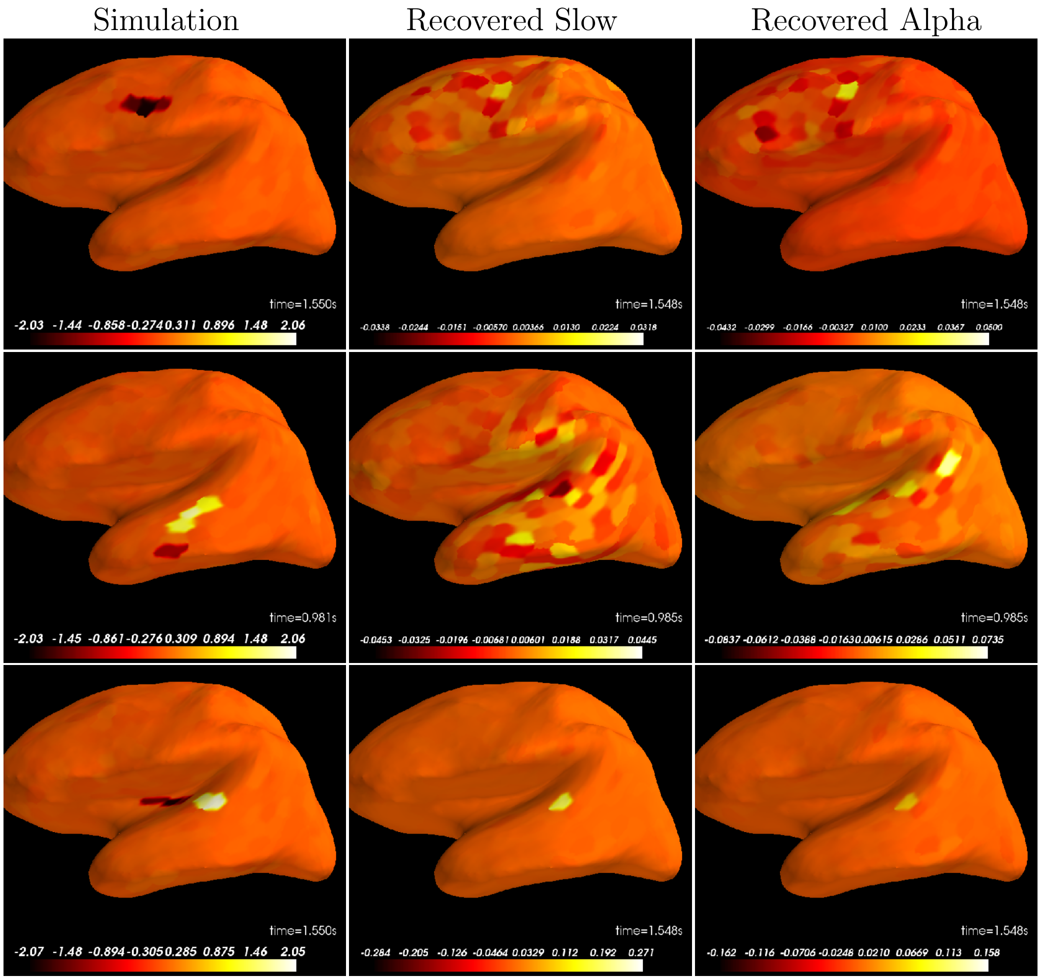
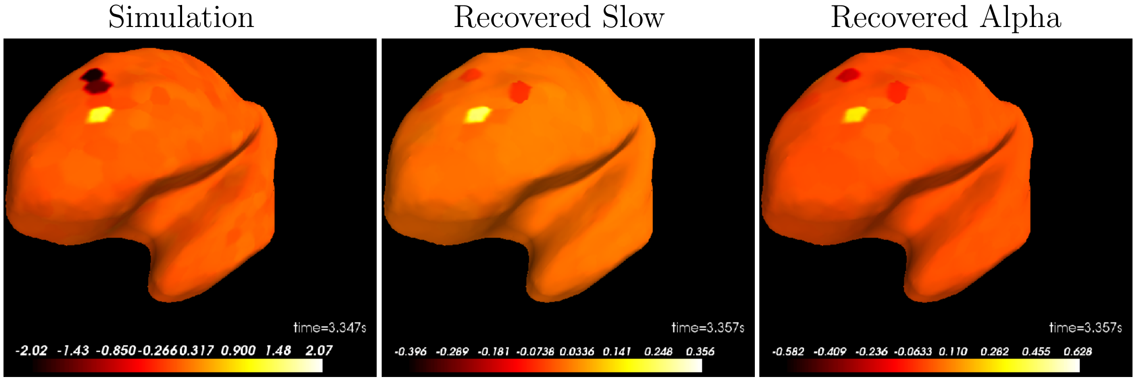
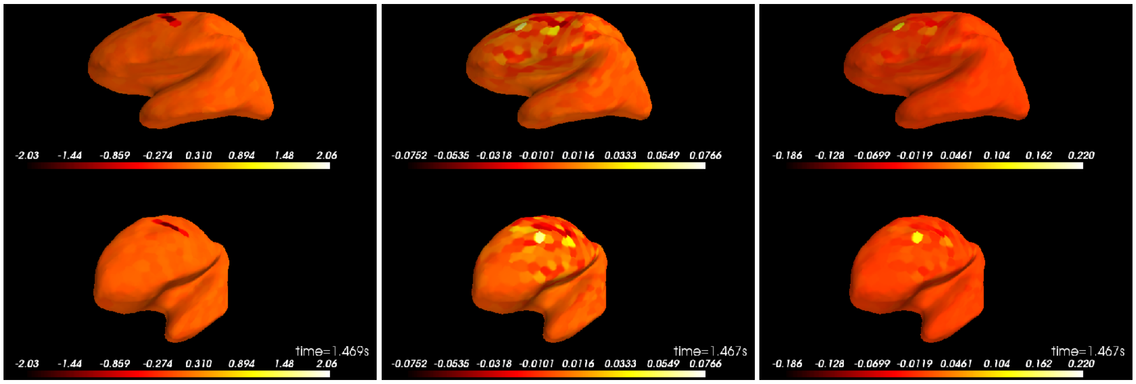


Figure 4-4: Columns: Simulation of slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Rows: Various lateral regions. Localization performance decreases in regions further away from the scalp sensors (2nd and 3rd rows) and in regions obscured from the scalp by the brain's anatomy (3rd row), as expected.

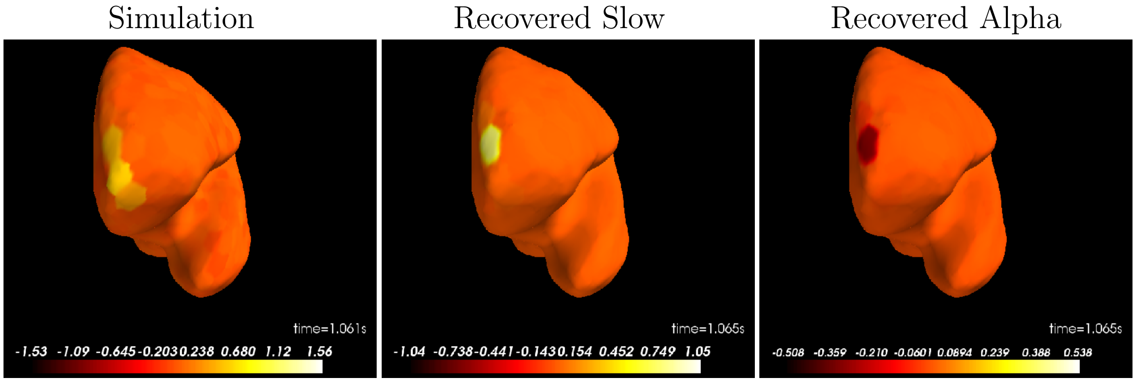


(a) Simulation in a dorsal gyrus

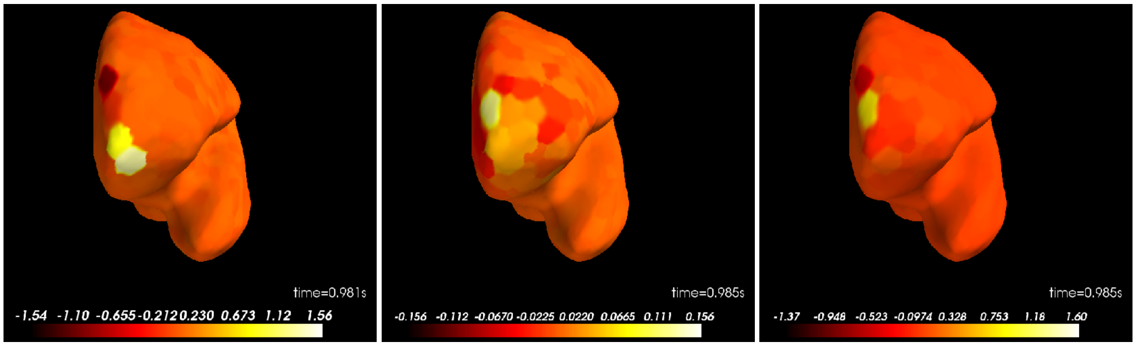


(b) Simulation in a dorsal sulcus

Figure 4-5: Columns: Simulation of slow and alpha oscillations of equal amplitude; Slow and alpha oscillations recovered using oscillator-EM source localization. Comparing localization performance in a gyrus versus a sulcus in the dorsal region. The oscillator-EM estimate is more spread out when localizing in a sulcus, as expected since cortical currents within sulcii are more difficult to measure at the scalp with EEG.

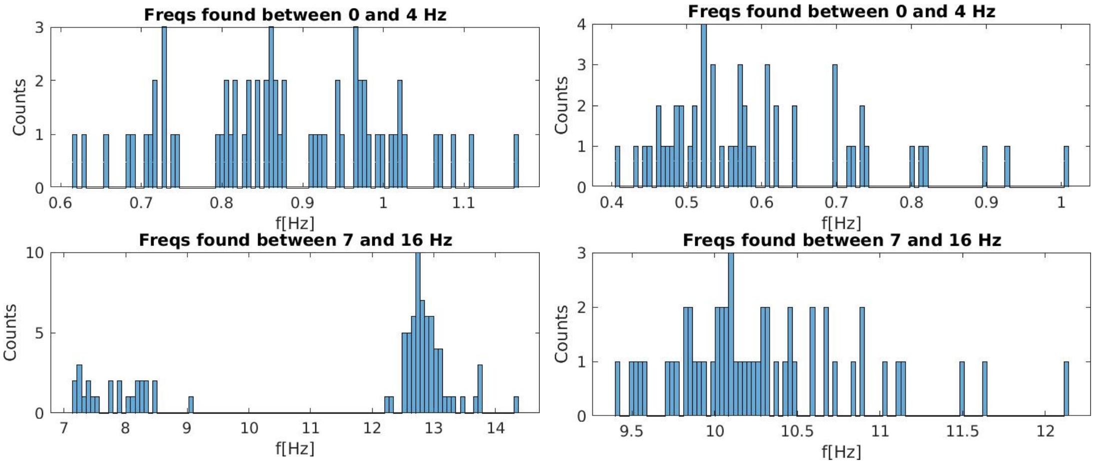


(a) Simulation with slow amplitude 1, alpha amplitude 0.5



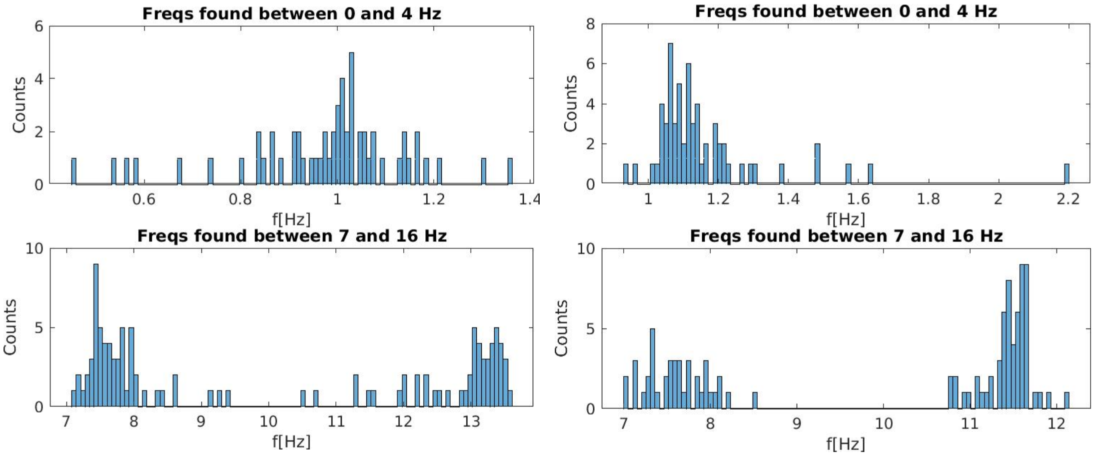
(b) Simulation with slow amplitude 0.5, alpha amplitude 1

Figure 4-6: Oscillator-EM simulations with a larger slow than alpha amplitude (1 vs 0.5) (a) and vice versa (0.5 vs 1) (b). For the slow oscillation, the source localization is more accurate when the amplitude is larger. The localization of the alpha oscillation remains accurate when the amplitude is reduced.



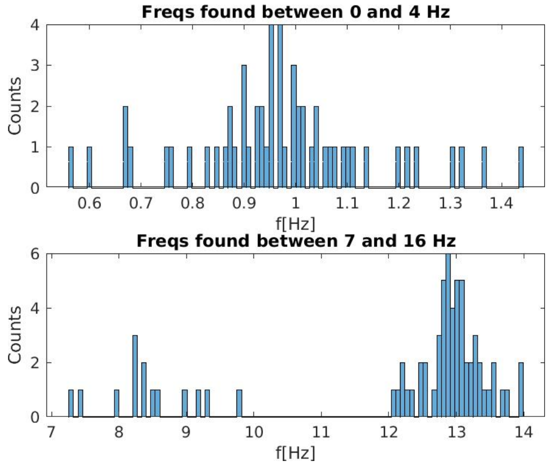
(a) Propofol Troughmax

(b) Propofol Peakmax



(c) NREM2

(d) NREM3



(e) Dexmedetomidine

Figure 4-7: Histograms of oscillator frequencies identified by fitting the oscillator model to data from all clean EEG channels in a 4s epoch of data for a representative subject from each condition. In each figure, the top panel shows all the frequencies found between 0 and 4Hz (the slow-delta band) and the bottom panel shows all the frequencies found between 7 and 16Hz (the alpha-spindle band). In general, one peak in the slow band and 1 or 2 peaks in the alpha band are found.

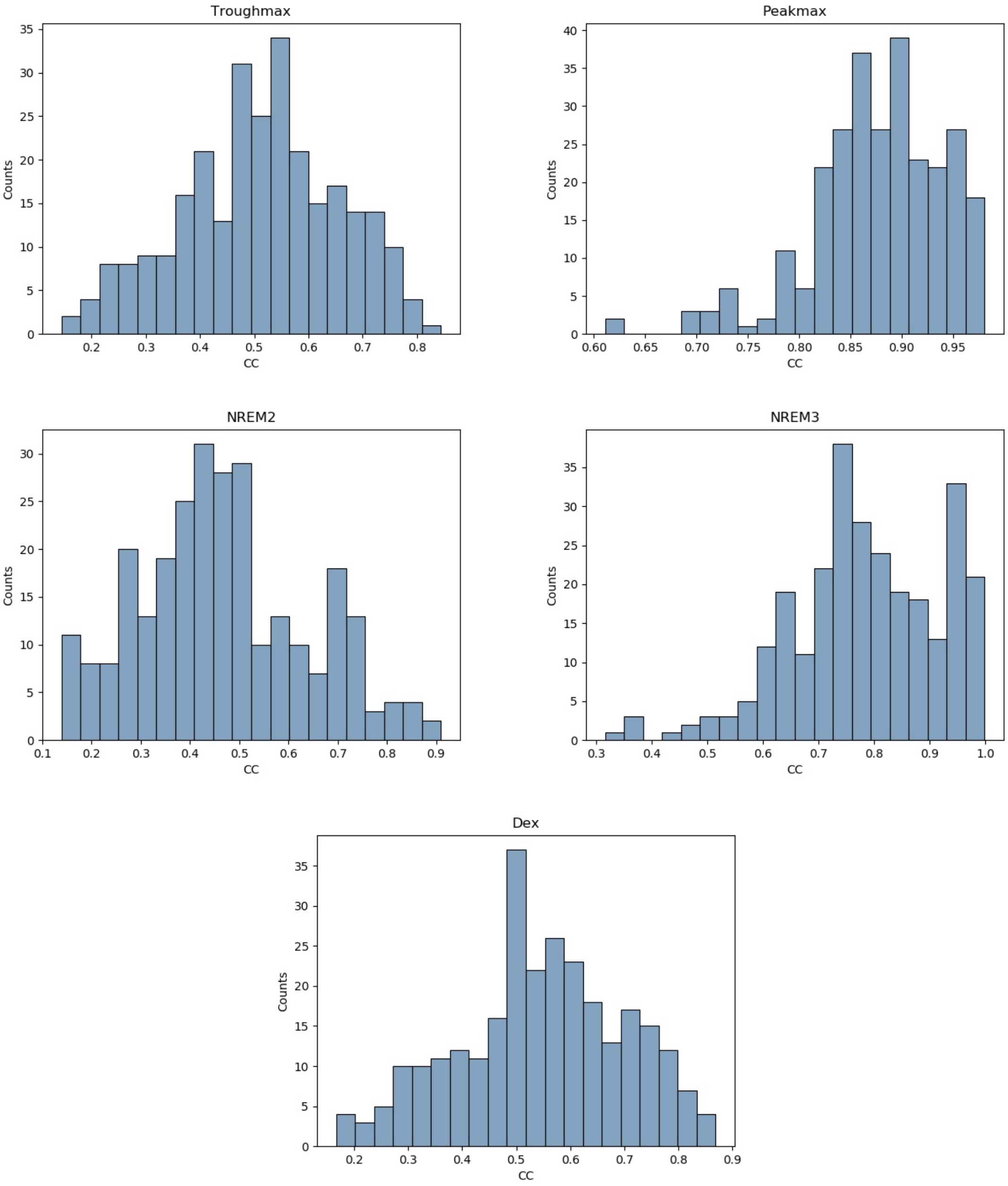


Figure 4-8: Histograms of the canonical coherence (CC) distribution in the 5 conditions under oscillator-EM source localization. The CC is skewed towards values closer to 1.0 in peakmax and NREM3.

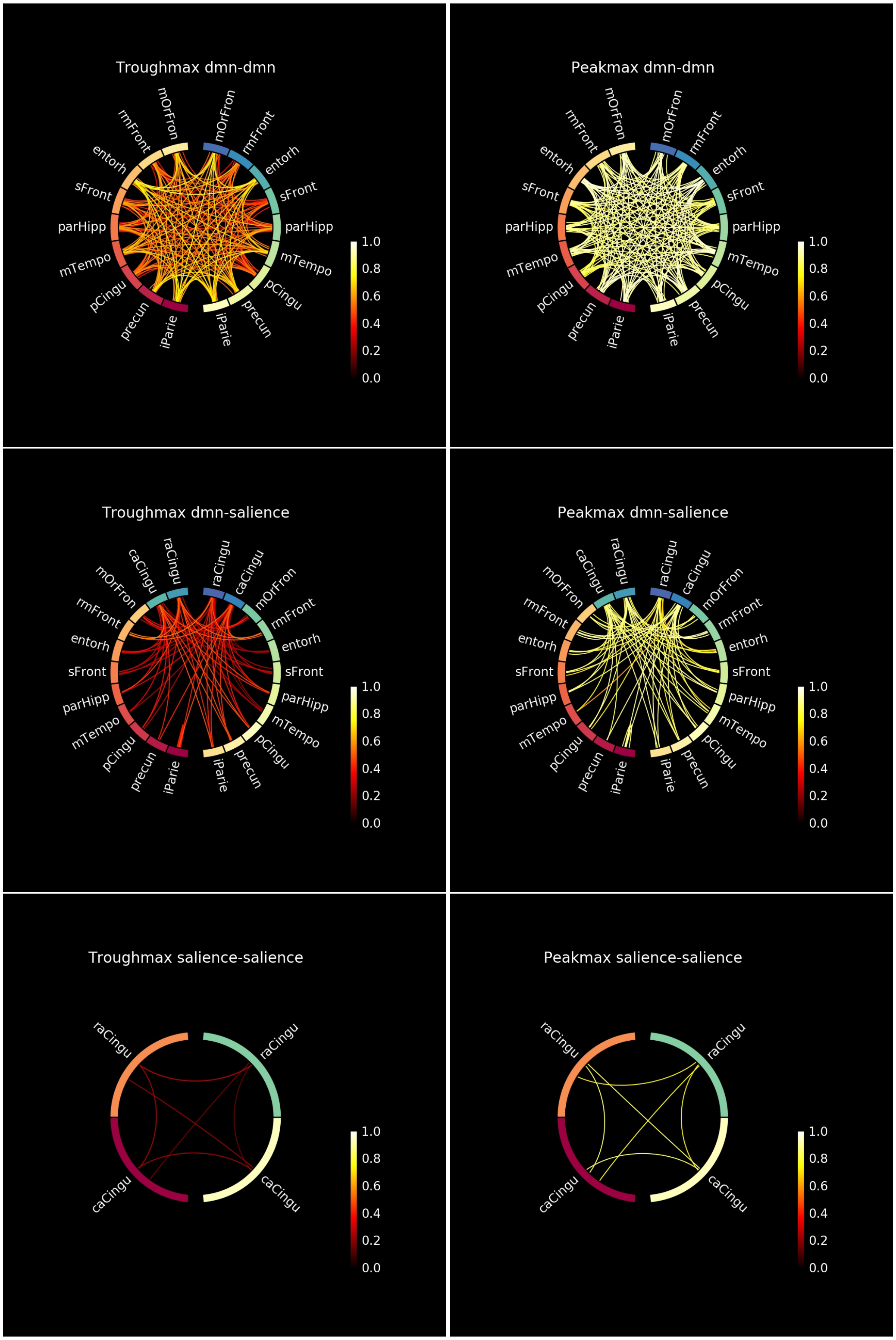


Figure 4-9: Canonical coherence under oscillator-EM source localization in troughmax and peakmax aggregated across propofol subjects. Peakmax shows overall higher coherence than troughmax.

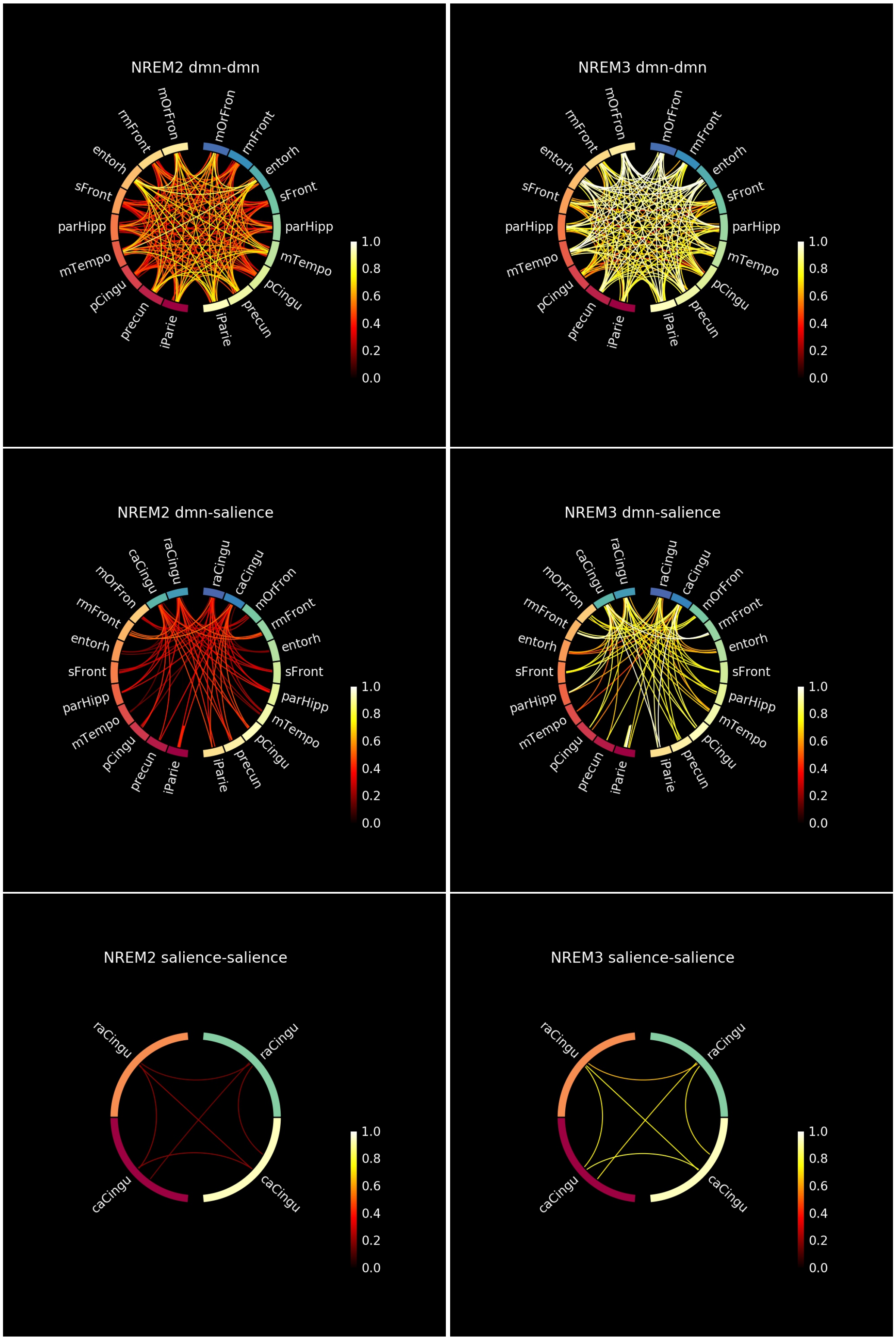


Figure 4-10: Canonical coherence under oscillator-EM source localization in NREM2 and NREM3 aggregated across sleep subjects. NREM3 shows overall higher coherence than NREM2.

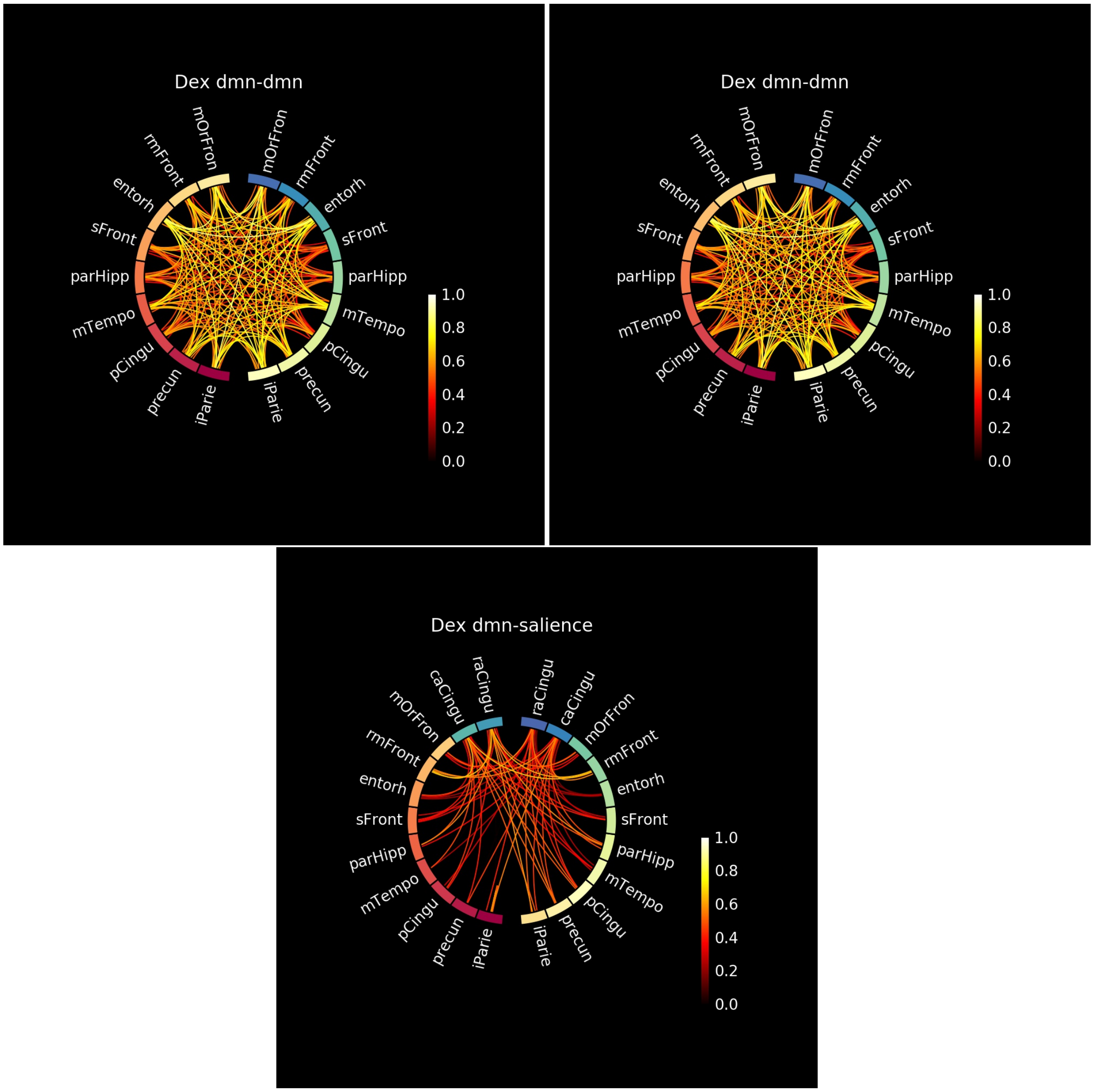


Figure 4-11: Canonical coherence under oscillator-EM source localization in dexmedetomidine sedation aggregated across dexmedetomidine subjects. The results for dexmedetomidine resemble the results for NREM2 sleep.

Chapter 5

Conclusion

5.1 Thesis summary

In this thesis, we have quantified the differences in the slow oscillation under 5 distinct states of unconsciousness: troughmax and peakmax under propofol general anesthesia, sedation under dexmedetomidine, and NREM2 and NREM3 sleep. We showed that slow oscillation power increases from NREM2 to dexmedetomidine sedation, NREM3, troughmax and then peakmax, consistent with the idea that deeper states of unconsciousness produce slow oscillations of greater amplitude. We characterized the spatial distribution of the slow oscillation power in both sensor space and source space, making use of MNE source localization for the latter. We explored the functional connectivity between source-space regions in the DMN and Sal networks under the different states of unconsciousness. Under MNE source localization, the canonical coherence decreases with increasing depth of unconsciousness, correlating with increased disruption to processing.

On a methodological front, we explored the limitations of MNE source localization, demonstrating that the method causes large spurious coherence by spreading out the source estimate solution all over the cortex. After establishing that the removal of the spurious coherence is intractable, we presented oscillator-EM, a novel method for source localization that employs a state-space model of neural oscillations. We showed in simulation that oscillator-EM outperforms MNE. However, the canonical coherence under oscillator-EM is the opposite of the expected result, showing higher coherence in deeper states of unconsciousness. We hypothesize that 1) the 4s epochs used may not be enough to capture coherence relations using oscillator-EM due to the time lag in tracking the oscillations whch we observed in simulations; and 2) that adding spatial smoothing to the Matsuda transition matrix could improve the localization by imposing reasonable values on sources which are difficult to observe from the scalp.

5.2 Future directions

The oscillator-EM method has proven to work well in simulations. More testing is required to ensure that the canonical coherence results obtained from oscillator-EM localization in real data are accurate. We plan to conduct simulations to determine whether oscillator-EM can accurately recover a known coherence. We will also test oscillator-EM on longer data segments to check whether the same coherence results are obtained. Lastly, we will test the oscillator-EM results with simpler measures of functional connectivity such as coherence between pairs of sources and the phase-locking factor.

Bibliography

[1] Freesurfer. `https://surfer.nmr.mgh.harvard.edu/fswiki`, 2020.

[2] A. Agmon and B. W. Connors. Repetitive Burst-Firing Neurons in the Deep Layers of Mouse Somatosensory Cortex. *Neurosci. Lett.*, 99(1-2):137–141, Apr 1989.

[3] O Akeju, KJ Pavone, JA Thum, PG Firth, MB Westover, M Puglia, ES Shank, EN Brown, and PL Purdon. Age-dependency of sevoflurane-induced electroencephalogram dynamics in children. *British journal of anaesthesia*, 115(suppl_1):i66–i76, 2015.

[4] Oluwaseun Akeju, Kara J Pavone, M Brandon Westover, Rafael Vazquez, Michael J Prerau, Priscilla G Harrell, Katharine E Hartnack, James Rhee, Aaron L Sampson, Kathleen Habeeb, et al. A comparison of propofol-and dexmedetomidine-induced electroencephalogram dynamics using spectral and coherence analysis. *Anesthesiology*, 121(5):978, 2014.

[5] Daniel Alschuler. ebridge tutorial. `https://psychophysiology.cpmc.columbia.edu/software/eBridge/tutorial.html`, September 2013.

[6] Daniel M Alschuler, Craig E Tenke, Gerard E Bruder, and Jürgen Kayser. Identifying electrode bridging from electrical distance distributions: a survey of publicly-available EEG data using a new method. *Clinical Neurophysiology*, 125(3):484–490, 2014.

[7] J. Anderson, I. Lampl, I. Reichova, M. Carandini, and D. Ferster. Stimulus Dependence of Two-State Fluctuations of Membrane Potential in Cat Visual Cortex. *Nat. Neurosci.*, 3(6):617–621, Jun 2000.

[8] Behtash Babadi and Emery N Brown. A review of multitaper spectral analysis. *IEEE Transactions on Biomedical Engineering*, 61(5):1555–1564, 2014.

[9] Randall J Bateman, Ling Y Munsell, John C Morris, Robert Swarm, Kevin E Yarasheski, and David M Holtzman. Human amyloid-β synthesis and clearance rates as measured in cerebrospinal fluid in vivo. *Nature medicine*, 12(7):856–861, 2006.

[10] Amanda M Beck, Hugo Soulat, Emily P Stephen, and Patrick L Purdon. State space oscillator models with priors on frequency. Cosyne, 2018.

[11] Riccardo Beltramo, Giulia D'Urso, Marco Dal Maschio, Pasqualina Farisello, Serena Bovetti, Yoanne Clovis, Glenda Lassi, Valter Tucci, Davide De Pietri Tonelli, and Tommaso Fellin. Layer-specific Excitatory Circuits Differentially Control Recurrent Network Dynamics in the Neocortex. *Nat. Neurosci.*, 16(2):227–234, Feb 2013.

[12] Sen's Blog. Subspace learning method. https://sensblogs.wordpress.com/2011/09/06/subspace-learning-method/. Accessed: 2018-04-25.

[13] Pierre Boveroux, Audrey Vanhaudenhuyse, Marie-Aurélie Bruno, Quentin Noirhomme, Séverine Lauwick, André Luxen, Christian Degueldre, Alain Plenevaux, Caroline Schnakers, Christophe Phillips, et al. Breakdown of within- and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 113(5):1038–1053, 2010.

[14] Allison Bradley, Jun Yao, Jules Dewald, and Claus-Peter Richter. Evaluation of Electroencephalography Source Localization Algorithms with Multiple Cortical Sources. *PLoS One*, 11(1):e0147266, Jan 2016.

[15] Emery N Brown, Patrick L Purdon, and Christa J Van Dort. General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual review of neuroscience*, 34:601–628, 2011.

[16] Randy L Buckner, Jessica R Andrews-Hanna, and Daniel L Schacter. The brain's default network: anatomy, function, and relevance to disease. 2008.

[17] Andrea E Cavanna and Michael R Trimble. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129(3):564–583, 2006.

[18] Y. Chagnac-Amitai, H. J. Luhmann, and D. A. Prince. Burst Generating and Regular Spiking Layer 5 Pyramidal Neurons of Rat Neocortex Have Different Morphological Features. *J. Comp. Neurol.*, 296(4):598–613, Jun 1990.

[19] Sylvain Chauvette, Sylvain Crochet, Maxim Volgushev, and Igor Timofeev. Properties of slow oscillation during slow-wave sleep and anesthesia in cats. *Journal of Neuroscience*, 31(42):14998–15008, 2011.

[20] Sylvain Chauvette, Maxim Volgushev, and Igor Timofeev. Origin of Active States in Local Neocortical Networks during Slow Sleep Oscillation. *Cereb. Cortex*, 20(11):2660–2674, Nov 2010.

[21] B. W. Connors, M. J. Gutnick, and D. A. Prince. Electrophysiological Properties of Neocortical Neurons in Vitro. *J. Neurophysiol.*, 48(6):1302–1320, Dec 1982.

[22] Laura Cornelissen, Seong-Eun Kim, Patrick L Purdon, Emery N Brown, and Charles B Berde. Age-dependent electroencephalogram (eeg) patterns during sevoflurane general anesthesia in infants. *Elife*, 4:e06513, 2015.

[23] Rosa Cossart, Dmitriy Aronov, and Rafael Yuste. Attractor dynamics of network up states in the neocortex. *Nature*, 423(6937):283–288, 2003.

[24] J. DeFelipe and I. Fariñas. The Pyramidal Neuron of the Cerebral Cortex: Morphological and Chemical Characteristics of the Synaptic Inputs. *Prog. Neurobiol.*, 39(6):563–607, Dec 1992.

[25] MNE Developers. MNE: MEG + EEG Analysis and Visualization. `https://mne.tools/stable/index.html#`, 2012-2020.

[26] P. Fatt and B. Katz. Spontaneous subthreshold activity at motor nerve endings. *J. Physiol.*, 117(1):109, May 1952.

[27] Robert G Gallager. *Stochastic processes: theory for applications.* Cambridge University Press, 2013.

[28] Peyman Golshani, Xiao-Bo Liu, and Edward G Jones. Differences in quantal amplitude reflect glur4-subunit number at corticothalamic synapses on two populations of thalamic neurons. *Proceedings of the National Academy of Sciences*, 98(7):4172–4177, 2001.

[29] Laurel A Graves, Elizabeth A Heller, Allan I Pack, and Ted Abel. Sleep deprivation selectively impairs memory consolidation for contextual fear conditioning. *Learning & memory*, 10(3):168–176, 2003.

[30] Michael D Greicius, Vesa Kiviniemi, Osmo Tervonen, Vilho Vainionpää, Seppo Alahuhta, Allan L Reiss, and Vinod Menon. Persistent default-mode network connectivity during light sedation. *Human brain mapping*, 29(7):839–847, 2008.

[31] Michael D Greicius, Ben Krasnow, Allan L Reiss, and Vinod Menon. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1):253–258, 2003.

[32] Michael D Greicius, Kaustubh Supekar, Vinod Menon, and Robert F Dougherty. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral cortex*, 19(1):72–78, 2009.

[33] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLoS Biol*, 6(7):e159, 2008.

[34] Andrea Hasenstaub, Robert N. S. Sachdev, and David A. McCormick. State Changes Rapidly Modulate Cortical Neuronal Responsiveness. *J. Neurosci.*, 27(36):9607–9622, Sep 2007.

[35] Christopher Heard, Frederick Burrows, Kristin Johnson, Prashant Joshi, James Houck, and Jerrold Lerman. A comparison of dexmedetomidine-midazolam with propofol for maintenance of anesthesia in children undergoing magnetic resonance imaging. *Anesthesia & Analgesia*, 107(6):1832–1839, 2008.

[36] Tanguy Hedrich, Giovanni Pellegrino, Eliane Kobayashi, Jean-Marc Lina, and Christophe Grova. Comparison of the spatial resolution of source imaging techniques in high-density EEG and MEG. *Neuroimage*, 157:531–544, 2017.

[37] E Huupponen, A Maksimow, P Lapinlampi, M Särkelä, A Saastamoinen, A Snapir, H Scheinin, M Scheinin, P Meriläinen, S-L HIMANEN, et al. Electroencephalogram spindle activity during dexmedetomidine sedation and physiological sleep. *Acta Anaesthesiologica Scandinavica*, 52(2):289–294, 2008.

[38] Edward G Jones. The thalamic matrix and thalamocortical synchrony. *Trends in neurosciences*, 24(10):595–601, 2001.

[39] Yusuke Kasuya, Raghavendra Govinda, Stefan Rauch, Edward J Mascha, Daniel I Sessler, and Alparslan Turan. The correlation between bispectral index and observational sedation scale in volunteers sedated with dexmedetomidine and propofol. *Anesthesia & Analgesia*, 109(6):1811–1815, 2009.

[40] Maxime Lemieux, Sylvain Chauvette, and Igor Timofeev. Neocortical inhibitory activities and long-range afferents contribute to the synchronous onset of silent states of the neocortical slow oscillation. *J. Neurophysiol.*, 113(3):768, Feb 2015.

[41] Maxime Lemieux, Jen-Yung Chen, Peter Lonjers, Maxim Bazhenov, and Igor Timofeev. The Impact of Cortical Deafferentation on the Neocortical Slow Oscillation. *J. Neurosci.*, 34(16):5689–5703, Apr 2014.

[42] Laura D Lewis, Veronica S Weiner, Eran A Mukamel, Jacob A Donoghue, Emad N Eskandar, Joseph R Madsen, William S Anderson, Leigh R Hochberg, Sydney S Cash, Emery N Brown, et al. Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. *Proceedings of the National Academy of Sciences*, 109(49):E3377–E3386, 2012.

[43] Artur Luczak, Peter Barthó, Stephan L. Marguet, György Buzsáki, and Kenneth D. Harris. Sequential structure of neocortical spontaneous activity in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, 104(1):347–352, Jan 2007.

[44] Edward O. Mann, Michael M. Kohl, and Ole Paulsen. Distinct Roles of GABAA and GABAB Receptors in Balancing and Terminating Persistent Cortical Activity. *J. Neurosci.*, 29(23):7513, Jun 2009.

[45] Bu-Qing Mao, Farid Hamzei-Sichani, Dmitriy Aronov, Robert C Froemke, and Rafael Yuste. Dynamics of spontaneous activity in neocortical slices. *Neuron*, 32(5):883–898, 2001.

[46] Takeru Matsuda and Fumiyasu Komaki. Time series decomposition into oscillation components and phase estimation. *Neural computation*, 29(2):332–367, 2017.

[47] D. A. McCormick and H. C. Pape. Properties of a hyperpolarization-activated cation current and its role in rhythmic oscillation in thalamic relay neurones. *J. Physiol.*, 431:291, Dec 1990.

[48] Michael Murphy, Marie-Aurélie Bruno, Brady A Riedner, Pierre Boveroux, Quentin Noirhomme, Eric C Landsness, Jean-Francois Brichant, Christophe Phillips, Marcello Massimini, Steven Laureys, et al. Propofol anesthesia and sleep: a high-density eeg study. *Sleep*, 34(3):283–291, 2011.

[49] Garrett T Neske. The slow oscillation in cortical and thalamic networks: mechanisms and functions. *Frontiers in neural circuits*, 9:88, 2016.

[50] Alan V Oppenheim and George C Verghese. *Signals, systems and inference*. Pearson, 2015.

[51] Svitlana Palchykova, Raphaëlle Winsky-Sommerer, Peter Meerlo, Roland Dürr, and Irene Tobler. Sleep deprivation impairs object recognition in mice. *Neurobiology of learning and memory*, 85(3):263–271, 2006.

[52] Michael J Prerau, Ritchie E Brown, Matt T Bianchi, Jeffrey M Ellenbogen, and Patrick L Purdon. Sleep neurophysiological dynamics through the lens of multitaper spectral analysis. *Physiology*, 32(1):60–92, 2017.

[53] Michael J Prerau, Katie E Hartnack, Gabriel Obregon-Henao, Aaron Sampson, Margaret Merlino, Karen Gannon, Matt T Bianchi, Jeffrey M Ellenbogen, and Patrick L Purdon. Tracking the sleep onset process: an empirical model of behavioral and physiological dynamics. *PLoS computational biology*, 10(10), 2014.

[54] Patrick L Purdon, Eric T Pierce, Eran A Mukamel, Michael J Prerau, John L Walsh, Kin Foon K Wong, Andres F Salazar-Gomez, Priscilla G Harrell, Aaron L Sampson, Aylin Cimenser, et al. Electroencephalogram signatures of loss and recovery of consciousness from propofol. *Proceedings of the National Academy of Sciences*, 110(12):E1142–E1151, 2013.

[55] PL Purdon, KJ Pavone, O Akeju, AC Smith, AL Sampson, J Lee, DW Zhou, K Solt, and EN Brown. The ageing brain: age-dependent changes in the electroencephalogram during propofol and sevoflurane general anaesthesia. *British journal of anaesthesia*, 115(suppl_1):i46–i57, 2015.

[56] Rey R. Ramírez. Source localization. *Scholarpedia*, 3(11):1733, Nov 2008.

[57] Pavlos Rigas and Manuel A. Castro-Alamancos. Thalamocortical Up States: Differential Effects of Intrinsic and Extrinsic Cortical Inputs on Persistent Activity. *J. Neurosci.*, 27(16):4261–4272, Apr 2007.

[58] Cristina Rosazza and Ludovico Minati. Resting-state brain networks: literature review and clinical applications. *Neurological sciences*, 32(5):773–785, 2011.

[59] Maria V. Sanchez-Vives and David A. McCormick. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.*, 3(10):1027–1034, Oct 2000.

[60] William W Seeley, Vinod Menon, Alan F Schatzberg, Jennifer Keller, Gary H Glover, Heather Kenna, Allan L Reiss, and Michael D Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9):2349–2356, 2007.

[61] Emily Stephen, Kyle Lepage, Uri Eden, Peter Brunner, Gerwin Schalk, Jonathan Brumberg, Frank Guenther, and Mark Kramer. Assessing dynamics, spatial scale, and uncertainty in task-related brain network analyses. *Frontiers in Computational Neuroscience*, 8:31, 2014.

[62] Emily Patricia Stephen. *Characterizing dynamically evolving functional networks in humans with application to speech.* PhD thesis, Boston University, 2015.

[63] M Steriade. Corticothalamic resonance, states of vigilance and mentation. *Neuroscience*, 101(2):243–276, 2000.

[64] M. Steriade and F. Amzica. Intracortical and corticothalamic coherency of fast spontaneous oscillations. *Proc. Natl. Acad. Sci. U.S.A.*, 93(6):2533, Mar 1996.

[65] M Steriade, D Contreras, R Curró Dossi, and A Nunez. The slow (< 1 hz) oscillation in reticular thalamic and thalamocortical neurons: scenario of sleep rhythm generation in interacting thalamic and neocortical networks. *Journal of Neuroscience*, 13(8):3284–3299, 1993.

[66] M. Steriade, R. C. Dossi, and A. Nuñez. Network modulation of a slow intrinsic oscillation of cat thalamocortical neurons implicated in sleep delta waves: cortically induced synchronization and brainstem cholinergic suppression. *J. Neurosci.*, 11(10):3200–3217, Oct 1991.

[67] M. Steriade, A. Nunez, and F. Amzica. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *J. Neurosci.*, 13(8):3252–3265, Aug 1993.

[68] Mircea Steriade, Angel Nuñez, and Florin Amzica. Intracellular analysis of relations between the slow (< 1 Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. *Journal of Neuroscience*, 13(8):3266–3283, 1993.

[69] A. E. Telfeian and B. W. Connors. Layer-specific Pathways for the Horizontal Propagation of Epileptiform Discharges in Neocortex. *Epilepsia*, 39(7):700–708, Jul 1998.

[70] A. M. Thomson and J. Deuchars. Temporal and Spatial Properties of Local Circuits in Neocortex. *Trends Neurosci.*, 17(3):119–126, Mar 1994.

[71] Lulu Xie, Hongyi Kang, Qiwu Xu, Michael J Chen, Yonghong Liao, Meenakshisundaram Thiyagarajan, John O'Donnell, Daniel J Christensen, Charles Nicholson, Jeffrey J Iliff, et al. Sleep drives metabolite clearance from the adult brain. *science*, 342(6156):373–377, 2013.

[72] Jokubas Žiburkus, John R. Cressman, and Steven J. Schiff. Seizures as Imbalanced Up States: Excitatory and Inhibitory Conductances During Seizure-Like Events. *J. Neurophysiol.*, 109(5):1296–1306, Mar 2013.