

## Brain connectivity in acute and chronic pain

Felix Sebastian Bott

Vollständiger Abdruck der von der TUM School of Medicine and Health der Technischen Universität München zur Erlangung eines  
Doctor of Philosophy (Ph.D.)  
genehmigten Dissertation.

Vorsitz: Prof. Dr. Henning Wackerhage

Betreuer\*in: Prof. Dr. Markus Ploner

Prüfer\*innen der Dissertation:

1. Prof. Dr. Joachim Groß
2. Prof. Dr. Valentin Riedl
3. Prof. Dr. Ulrike Bingel

Die Dissertation wurde am 01.12.2023 bei der TUM School of Medicine and Health der Technischen Universität München eingereicht und durch die TUM School of Medicine and Health am 13.02.2024 angenommen.



# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	<i>Pain</i>	1
1.1.1	Acute Pain	1
1.1.2	Chronic Pain	2
1.2	<i>The cerebral processing of pain</i>	3
1.2.1	Local brain activity in pain	3
1.2.2	Inter-regional brain connectivity in pain and beyond	4
1.3	<i>Research questions and outline</i>	9
1.3.1	Project 1: Brain connectivity in acute pain	9
1.3.2	Project 2: Network dysfunction in chronic pain	9
1.3.3	Publications and code	10
<b>2</b>	<b>Methodological basics</b>	<b>11</b>
2.1	<i>Electroencephalography</i>	11
2.1.1	The physiological origin and significance of EEG	11
2.1.2	Recording devices	12
2.1.3	Source reconstruction	13
2.1.4	Representative time series	15
2.1.5	Connectivity measures	17
2.2	<i>Statistics</i>	27
2.2.1	Bayes factors and model comparisons	27
2.2.2	Specific Bayesian statistical tests	29
<b>3</b>	<b>Project 1: Acute pain</b>	<b>35</b>
3.1	<i>Project-specific methods</i>	35
3.1.1	Participants	35
3.1.2	Paradigm	35
3.1.3	Recordings and preprocessing	37
3.1.4	Evaluating brain measures	37
3.1.5	Statistical analysis	40
3.1.6	Model comparisons	41

3.2	<i>Results</i> .....	43
3.2.1	Behavioral Findings.....	45
3.2.2	Source-level evoked potentials and local oscillatory brain activity.....	45
3.2.3	Inter-regional connectivity .....	49
3.2.4	Model comparisons .....	52
3.2.5	Direct association between neural measures and pain.....	53
3.2.6	Summary .....	55
<b>4</b>	<b>Project 2: Chronic pain</b> .....	<b>57</b>
4.1	<i>Project-specific methods</i> .....	57
4.1.1	Data sets .....	57
4.1.2	Automatic preprocessing.....	62
4.1.3	Evaluating brain measures.....	62
4.1.4	Statistical analysis and machine learning.....	65
4.2	<i>Results</i> .....	72
4.2.1	Primary analyses: Intrinsic brain network function in chronic back pain ....	73
4.2.2	Secondary analyses: Intrinsic brain network function in chronic pain, depression, and age .....	77
4.2.3	Exploratory analyses .....	82
<b>5</b>	<b>Discussion</b> .....	<b>86</b>
5.1	<i>Project 1</i> .....	86
5.2	<i>Project 2</i> .....	88
5.3	<i>Conclusion and outlook</i> .....	93
	<b>Acknowledgements</b> .....	<b>94</b>
	<b>References</b> .....	<b>95</b>
	<b>Publications</b> .....	<b>105</b>
<b>A</b>	<b>Appendix</b> .....	<b>1</b>
A.1	<i>Simulative assessment of representative time series methods</i> .....	1
A.2	<i>Distributions of dependent variables</i> .....	4
A.3	<i>Chronic back pain and activity of intrinsic brain networks</i> .....	5



## List of figures

<b>Figure 2-1:</b> Methods for estimating representative time series. ....	17
<b>Figure 2-2:</b> Effects of amplitude modulation on spectral representations. ....	21
<b>Figure 3-1:</b> Experimental design and possible response patterns. ....	44
<b>Figure 3-2:</b> Pain ratings for different levels of stimulus intensity and expectations. ....	45
<b>Figure 3-3:</b> Noxious stimulus-related brain responses. ....	46
<b>Figure 3-4:</b> Effects of stimulus intensity, expectations, and prediction errors on local oscillatory brain activity and source-level evoked potentials. ....	48
<b>Figure 3-5:</b> Effects of stimulus intensity, expectations, and prediction errors on inter-regional connectivity. ....	50
<b>Figure 3-6:</b> Direction of information flow in selected network connections. ....	51
<b>Figure 3-7:</b> Competition of models predicting connectivity. ....	53
<b>Figure 3-8:</b> Direct association between pain and neural measures. ....	54
<b>Figure 3-9:</b> Synopsis of project 1. ....	56
<b>Figure 4-1:</b> Visualization of the seven Yeo networks. ....	63
<b>Figure 4-2:</b> Visualization of the machine learning pipeline. ....	68
<b>Figure 4-3:</b> Visualization of alternative spatial configurations. ....	70
<b>Figure 4-4:</b> Relation between pain intensity and connectivity in individual pairs of intrinsic brain networks in patients with chronic back pain. ....	74
<b>Figure 4-5:</b> Relation between pain intensity and patterns of inter-network connectivity in patients with chronic back pain. ....	75
<b>Figure 4-6:</b> Relation between pain intensity and intrinsic brain network connectivity in patients with diverse pain conditions. ....	77
<b>Figure 4-7:</b> Relation between depression severity and intrinsic brain network connectivity in patients with diverse pain conditions. ....	79
<b>Figure 4-8:</b> Relation between age and intrinsic brain network connectivity in patients with diverse pain conditions. ....	80
<b>Figure 0-1:</b> Relative explained variance scores achieved by the different methods for the YEO7 network definition. ....	3
<b>Figure 0-2:</b> Distributions of pain intensity, depression scores and age for the individual data sets. ....	4
<b>Figure 0-3:</b> Relation between pain intensity and patterns of intrinsic brain network activity in patients with chronic back pain. ....	5
<b>Figure 0-4:</b> Relation between pain intensity and activity in individual intrinsic brain networks in patients with chronic back pain. ....	5

## List of tables

<b>Table 4-1:</b> Overview of data sets employed in project 2. ....	61
<b>Table 4-2:</b> Summary of analysis strategies. ....	65
<b>Table 4-3:</b> Overview of various analysis variants. ....	72
<b>Table 4-4:</b> Predicting pain intensity using an extended set of candidate models.....	83
<b>Table 4-5:</b> Predicting depression severity using an extended set of candidate models..	84
<b>Table 4-6:</b> Predicting age using an extended set of candidate models.....	85

## Abstract

Pain protects the body from harm. It is an unpleasant yet vital experience that emerges from the integration of sensory and contextual information. However, how the brain implements this integration in healthy individuals and how this process is altered in patients with chronic pain remains unclear. Given the integrative nature of pain, it appears likely that exploring the communication between brain regions, i.e., inter-regional connectivity, could yield fundamental insights into the mechanisms of pain perception in health and disease. Therefore, this work assessed the relation between pain and measures of brain connectivity derived from electroencephalography (EEG), both in healthy individuals and patients with chronic pain.

Expectations are a key contextual influence on pain perception. To assess how expectation effects on acute pain are implemented in the brain and to compare this to sensory effects on pain, the first project analyzed data from an EEG experiment in healthy human individuals. It revealed that sensory information predominantly induced changes of local oscillatory brain activity at alpha, beta, and gamma frequencies, while expectations exclusively shaped inter-regional brain connectivity at alpha frequencies. Moreover, prediction errors, i.e. mismatches between sensory information and expectations, affected connectivity in the gamma band. These findings show how fundamentally different brain mechanisms underlie sensory and expectation effects on pain. Moreover, these findings highlight the relevance of brain connectivity for complex psychological modulations and thus probably also for pathological aberrations of the pain experience.

Building on these insights, a second project investigated the association between brain connectivity and chronic pain. Specifically, we assessed the function of selected intrinsic brain networks, which likely play a role in the pathology of chronic pain: The somatomotor, the frontoparietal, the salience ventral attention, and the default network. We related connectivity among these networks to the intensity of pain using uni- and multivariate strategies. In line with recent recommendations for open science practices, the study was preregistered via the OSF platform (<https://osf.io/qa68n>). The analyses showed that higher levels of pain were associated with increased connectivity at theta frequencies and decreased connectivity at alpha frequencies. Moreover, a multivariate model based on inter-network connectivity at theta, alpha, and beta frequencies predicted pain ratings better than chance. However, neither the results of the univariate nor the multivariate analyses could be replicated consistently in independent data sets. These findings indicate a need for a harmonization of data acquisition procedures with the goal



of reducing variability across data sets and, in this way, obtain clearer and more robust evidence.

Altogether, the current results highlight the importance of brain connectivity for the understanding of pain perception in healthy individuals. Beyond, associations between brain connectivity and the pain experience of individuals with chronic pain likely also exist and were for the first time studied across multiple independent data sets. This assessment revealed that establishing such associations in a robust manner requires important methodological adaptations in future studies. These might include stratifying patients, unifying recording protocols and quality standards across recording sites, and employing longitudinal designs.



# 1 Introduction

## 1.1 Pain

Pain is an aversive experience that signals threat to the body, reinforcing behaviors that protect it from harm. To achieve this, pain integrates both sensory and contextual information. Sensory information originates from physical stimuli, while contextual information encompasses factors such as attention, expectation, motivation, as well as social influences[4]. This understanding of pain aligns with the observation that the perceived intensity of a painful stimulus varies over time, even if the objective stimulus intensity does not. For instance, if an athlete twists their ankle during a competitive run, the pain they feel is likely to be more intense after crossing the finishing line than before. In this example, the contextual modulation of the pain experience helps the athlete to achieve their goal of finishing the race. However, contextual influences do not only attenuate pain but may also exacerbate it both in the short and long term. In this way, pain can persist over an extended period and without an imminent threat to the body. Such longer-lasting, chronic forms of pain have lost their protective function and become a mere burden for the affected individual.

### 1.1.1 Acute Pain

The International Association for the Study of Pain (IASP), defines a noxious stimulus as “a stimulus that is damaging or threatens damage to normal tissue”[5]. Before a noxious stimulus is rendered into a conscious percept by the brain, it is detected, encoded and processed by peripheral and spinal neurons projecting to the brain [6, 7]. This sub- or preconscious process is referred to as nociception. Accordingly, the IASP defines nociception as “the neural process of encoding noxious stimuli”[5].

Nociceptors are sensory neurons that are preferentially activated by noxious stimuli. Mediated by their structural properties, different types of nociceptors exhibit distinct dynamics of noxious information transmission[8, 9]. Nociceptor fibers terminate in the dorsal horn of the spinal cord[10]. The dorsal horn is a first relay station, the output of which projects to the brain. Interestingly, besides the afferent input, the dorsal horn also integrates information of descending fibers originating from the brainstem. Depending on the context, these descending pathways may either facilitate or inhibit the transmission of nociceptive information towards the brain [11, 12], thereby enabling top-down modulations of pain through factors such as expectation or attention.

In the brain, nociceptive information is rendered into a conscious percept which, depending on the context, may or may not qualify as painful[4, 13]. Cortical structures that have prominently been linked to nociception and pain are the somatosensory, the anterior cingulate, the insular, and the prefrontal cortices[14]. The activity of the somatosensory cortex is more closely associated with the nociceptive component of pain, whereas the other supramodal brain regions are thought to mediate cognitive, emotional, and motivational aspects of pain[10].

Taken together, acute pain arises from the interplay of ascending and descending information transmission about threat, leading to behavioral responses that protect the body from harm.

### 1.1.2 Chronic Pain

When pain persists for more than three months it is diagnosed as chronic[15, 16] and loses its protective function. Chronic pain is a debilitating condition with detrimental consequences for the lives and well-being of those affected. In affecting roughly 20% of the global population[17, 18] and being responsible for up to 20% of physician visits[19, 20], chronic pain imposes a substantial burden on healthcare systems[21]. To explain the emergence of chronic pain, several models, which are primarily focused on somatic aspects have been proposed[10, 22-24]. The biopsychosocial model assumes a broader perspective and posits that chronic pain arises from the interplay of biological, psychological, and social factors[24, 25].

Due to its complexity, chronic pain is often difficult to treat. Current treatment approaches include physiotherapy, psychotherapy, and medication[26]. These approaches are combined in multimodal therapies which constitute the current state-of-the-art in chronic pain management[27]. These forms of treatment have been shown to not only improve the patients' quality of life but also significantly reduce their pain[28].

Despite these efforts, there remains a substantial number of patients who cannot be helped adequately, highlighting the crucial need for new means of treating chronic pain. To achieve this, the biological mechanisms underlying complex chronic pain conditions have to be better understood. In this regard, the brain plays a pivotal role as the organ responsible for integrating all factors associated with chronic pain into a conscious percept[7, 14]. Therefore, in order to aid individuals burdened by debilitating chronic pain, we must study the brain.

## 1.2 The cerebral processing of pain

Nociceptive processing and pain perception cannot be attributed to a single brain system[14, 29, 30]. Instead, a diverse set of brain regions are involved, including the thalamus and somatosensory, cingulate, insular, and prefrontal cortices.

### 1.2.1 Local brain activity in pain

Noxious stimuli elicit a complex spatial-spectral-temporal pattern of brain activity. Spatial features of this pattern have been investigated using functional magnetic resonance imaging (fMRI). It was observed that brain regions involved in the processing of nociceptive information include the ventrolateral and medial thalamus, large parts of the insular cortex (INS), the secondary somatosensory cortex (S2), the anterior cingulate (ACC), and the prefrontal cortex (PFC)[14, 31-33]. The spectral-temporal characteristics of the brain's response to noxious stimuli have been studied using electroencephalography (EEG) and magnetoencephalography (MEG). In particular, these modalities have been employed to examine how periodic fluctuations in their signals, so-called brain rhythms or neural oscillations relate to pain. These brain rhythms are typically assessed in four canonical frequency bands which cover frequencies from 4 to 8 Hz (theta), from 8 to <13 Hz (alpha), from 13 to 30 Hz (beta), and >30 Hz (gamma)[34]. The neural responses to both short phasic (in the order of milliseconds) and more prolonged tonic stimuli (in the order of seconds to minutes) have been assessed. Such acute noxious stimuli, both phasic and tonic, induce a transient suppression of activity at alpha and beta frequencies predominantly in the occipital and sensorimotor cortices[35-40]. For phasic stimuli, these spectral changes occur within a post stimulus interval ranging from 300 to 1000 ms[36]. Additionally, phasic stimuli induce increased activity in the gamma band between 150 and 350 ms post stimulus, most markedly observed in the sensorimotor cortex[41-43]. Further, phasic stimuli induce enhanced activity at theta frequencies. These increases can be attributed to three consecutive, large arhythmic deflections of the neural signal which are commonly known as pain-related evoked potentials. These deflections attain their extrema at around 170, 240, and 390 ms after stimulus application and are believed to be generated in the somatosensory cortex, in the parietal operculum (PO) comprising S2 and parts of the INS, and in the cingulate and prefrontal cortex[44, 45].

In addition to the role of local brain activity in experimentally induced acute pain, aberrations in local brain activity in chronic pain have also been studied. Converging lines of evidence indicate that acute and chronic pain are served by distinct mechanisms in the brain. In patients with chronic back pain, it was shown that the intensity of ongoing pain

covaries with gamma oscillations in prefrontal areas and not, as in acute pain, with oscillations at alpha, beta, or gamma frequencies in somatosensory areas[46]. This idea of a shift towards more emotion-related and away from sensory brain regions in chronic pain is corroborated by longitudinal and cross-sectional fMRI observations in patients with subacute and chronic back pain[47]. In particular, these observations implicated emotion-related brain regions such as the ACC, the medial PFC, and the amygdala in chronic back pain.

Evidence from 76 studies investigating the association between chronic pain and changes in brain activity as measured by M/EEG was compiled in a recent systematic review[48]. This review found that the majority of the considered cross-sectional studies reported increased theta and beta-band activity in patients with chronic pain compared to healthy controls. In contrast, correlational studies examining the association between pain intensity and various measures of brain activity did not indicate consistent changes of local oscillatory brain activity in chronic pain. The review also pointed out a high risk of bias in many studies. This was due to small sample sizes, absence of multiple comparison corrections, and insufficient confounder analyses.

In brief, while acute pain was linked robustly to specific patterns of local brain activity, the association between chronic pain and local brain activity appears more diffuse. Considering the complexity of chronic pain conditions, assessing brain measures beyond local brain activity, such as brain connectivity, may be necessary to unveil robust associations which, in the future, could serve as biomarkers for chronic pain.

### 1.2.2 Inter-regional brain connectivity in pain and beyond

Inter-regional connectivity reflects communication between brain regions. One distinguishes between functional and effective connectivity. Functional connectivity describes “statistical dependencies among remote neurophysiological events”[49]. Effective connectivity refers to “the influence one neural system exerts over another”[49]. In this work, we are primarily concerned with functional connectivity among focal brain regions as well as among extended intrinsic brain networks. To clarify, by intrinsic brain networks we mean spatially distributed collections of brain regions sharing common functional properties. A prominent example of an intrinsic brain network is the so-called default network, which is typically observed in task-free states[50].

In M/EEG, connectivity is often quantified in terms of temporal invariance of phase differences between signals (phase-based connectivity) or as the correlation between signal envelopes (amplitude-based connectivity)[51]. More detailed methodological

explanations of different connectivity measures for M/EEG are provided in section 2.1.5. In fMRI, connectivity is commonly but not exclusively defined as the mere correlation between blood flow-related signals in different brain regions[52]. The aim of this section is to illustrate the relevance of inter-regional connectivity for the understanding of the neural substrates of pain and other neuropsychiatric disorders.

As outlined before, pain is a multi-faceted experience shaped by diverse contextual influences and involving many brain regions. It therefore appears plausible that the analysis of the communication between brain regions, i.e., brain connectivity, may unveil fundamental insights into the mechanisms of pain in both health and disease.

### *Brain connectivity in acute pain*

The Tonic Pain Signature (ToPS) is a recently introduced connectivity-based model that predicts tonic pain intensity in healthy individuals with high sensitivity and specificity[53], even when tested in independent data sets. Utilizing measures of dynamic functional connectivity from fMRI time series[54] in a whole-brain setup, the ToPS likely captures both spatial and temporal features of the dynamic pain connectome[55], a concept conceiving pain as being particularly shaped by the dynamics of brain connectivity. Intriguingly, the ToPS outperformed other pain signatures that rely on static local brain activity[31, 56], explained interindividual variability also in subacute and chronic back pain and distinguished patients with chronic back pain from healthy controls with reasonable accuracy. An aggregate view on the ToPS feature weights indicates a positive association between connectivity of the somatomotor-frontoparietal network pair and prolonged pain. The ToPS thus impressively demonstrates how neurobiologically interpretable patterns of functional connectivity can predict experimentally induced pain in healthy individuals and how these findings can be directly translated to clinical pain states. Similar to the ToPS, another recently proposed model called the Resting-state Pain-sensitivity Network (RPN) linked resting-state fMRI-based functional connectivity to pain perception in healthy individuals[57]. In particular, the RPN successfully predicted individual pain sensitivity, underscoring the informative potential of brain connectivity in understanding altered nociception, also a key aspect of many clinical pain conditions.

In addition to fMRI-based efforts, the role of brain connectivity in acute pain processing has also been studied using M/EEG. These recording modalities enable the assessment of oscillatory brain activity and connectivity at shorter temporal scales than fMRI. Based on anatomical considerations[58-60], it has been proposed that oscillatory activity at alpha/beta and gamma frequencies reflects inter-regional feedback and feedforward signaling, respectively. Here, feedforward describes the flow of information from neural

systems responsible for the initial processing of sensory input (hierarchically lower) to less specialized neural systems with a higher degree of integration (hierarchically higher). Feedback refers to the opposite direction of information flow. Within this framework, when pain is primarily shaped by bottom-up processes like stimulus intensity, gamma oscillations in somatosensory regions were proposed to subserve the signaling of nociceptive information toward higher-level brain regions. Conversely, alpha and beta oscillations might mediate top-down modulations of the pain experience[58].

In tonic pain, the subjective experience of pain dissociates quickly from the objective stimulus intensity, implying an increasing influence of top-down modulatory processes[39, 61]. Using a phase-based connectivity measure, tonic pain was shown to be associated with increased communication at alpha frequencies between somatosensory and medial prefrontal areas. While this seems in line with the above-described framework, an analysis of the direction of information flow indicated that connectivity changes occurred in the feedforward direction (i.e., from primary somatosensory to prefrontal cortex) rather than the feedback direction. Irrespective of the theoretical framework, these findings affirm the relevance of somatosensory-prefrontal connectivity in tonic pain processing.

In summary, the discussed studies illustrate the utility of functional connectivity analyses in decoding the cerebral processes of acute pain. Furthermore, the identified relationships between functional connectivity and acute pain can have direct implications for the assessment of clinical pain conditions.

#### *Brain connectivity in chronic pain and beyond*

Several studies have directly examined alterations of brain connectivity in chronic pain[14, 55]. In their seminal paper, Baliki et al. reported that functional connectivity between the medial PFC (mPFC) and the nucleus accumbens (NAc) predicted the transition from subacute to chronic back pain[62]. This connection's implication in top-down pain modulation, a relevant aspect also in chronic pain[63-65], was demonstrated in an experiment showing that self-regulation impacts pain perception via mPFC-NAc connectivity[66]. Further evidence for this connection's involvement in aberrant pain perception stems from rodent studies indicating that the optogenetic activation/silencing of NAc-projections of PFC neurons enhanced/reduced pain-related behaviors[67, 68]. Moreover, a ROI-based, statistical assessment of the feature weights in the previously discussed ToPS revealed that connectivity between the mPFC and the ventral striatum, which encompasses the NAc, is a feature highly predictive of both experimental and clinical pain states[53]. The mounting evidence implicating mPFC-NAc connectivity in



complex modulations of the pain experience thus exemplifies the potential utility of functional connectivity analyses in chronic pain research.

At the intrinsic brain network level, altered network function of the default network (DN) has consistently been associated with chronic pain. Increased connectivity between the DN and the INS was observed in patients with chronic pain relative to healthy controls and correlated with pain intensity on an inter- and intra-individual level[69-71]. Additionally, connectivity between the INS and the mPFC, a central hub of the DN, exhibited associations with pain intensity across various chronic pain conditions[71]. Complementing these fMRI-based findings, an analysis of resting-state MEG recordings of patients with fibromyalgia also linked DN-INS connectivity at various frequencies to chronic pain[72]. In terms of cross-network functional connectivity, the connection between the DN and the salience network (SN) has been linked to chronic pain. One study explicitly investigating this association revealed reduced anticorrelated connectivity between the DN and the SN[73]. Furthermore, a recent meta-analysis comprising 42 chronic pain-related fMRI-studies found reduced connectivity between the SN and the DN in patients with chronic pain when compared to healthy controls or patients with anxiety or major depressive disorder (MDD)[74]. Notably, this study also reports increased connectivity between the SN and the somatomotor network (SMN), although this association is not specific to chronic pain. Lastly, connectivity within the DN or, more precisely, between the DN as a whole and one of its constituent hubs, the mPFC, was shown to be elevated in patients with chronic pain relative to healthy controls[75]. Taken together, these studies indicate an involvement of altered intrinsic brain network function and, in particular, of DN connectivity in chronic pain.

Functional connectivity alterations in chronic pain have also been investigated using non-invasive neurophysiological recording methods, such as M/EEG. One theory suggests that chronic pain arises from an imbalance in brain activity between regions processing nociceptive input and those associated with descending pain suppression[76]. In fact, connectivity alterations at various frequencies, proposed to underly this imbalance, were observed in longitudinal as well as cross-sectional studies[77, 78]. By contrast, a recent systematic review including 76 studies found no conclusive evidence for altered connectivity in chronic pain at any frequency[48]. Nevertheless, the review noted a trend of enhanced theta-band connectivity in chronic pain, with an equal number of studies reporting increased and unchanged theta-band connectivity, and more studies showing increased rather than reduced theta-band connectivity. Notably, the two largest studies included in the review did report an enhanced theta-band connectivity in patients with chronic pain[79, 80]. The larger of these two studies additionally identified increased

gamma-band connectivity and reduced gamma-band global efficiency, the latter indicating changes in whole brain network organization[79, 80].

The reviewed studies provide evidence for an association between brain connectivity and chronic pain more frequently on the basis of fMRI rather than M/EEG recordings. However, the high temporal resolution and, in the case of EEG, the widespread availability, render these neuro recording methods a promising instrument for the investigation of the brain mechanisms of chronic pain. Moreover, recent studies have demonstrated the utility of EEG-based connectivity features for the assessment of other neuropsychiatric disorders, such as MDD or posttraumatic stress disorder (PTSD) [81, 82]. Using a measure of amplitude-based connectivity, Toll et al. showed that combat veterans suffering from PTSD exhibited dysconnectivity at theta frequencies in an extensive network of regions spanning the entire brain[81]. Although these results have not been replicated in independent data sets, they hint at the utility of amplitude-based connectivity patterns for the characterization of neuropsychiatric disorders. Another study by the same group identified two distinct PTSD subtypes by clustering amplitude-based connectivity profiles of patients with PTSD[82]. These clusters were mainly characterized by connectivity at beta frequencies and were replicable not only in PTSD but also in MDD patient cohorts. While baseline clinical severity scores were similar for both subtypes, their responses to various treatments differed significantly.

Overall, the reviewed studies indicate a strong link between brain connectivity and chronic pain, mainly supported by fMRI research. In particular, altered connectivity of intrinsic brain networks appears to play an important role in chronic pain. While evidence primarily stems from fMRI research, studies on neuropsychiatric disorders other than chronic pain convincingly demonstrate the potential also of M/EEG-based connectivity analyses for uncovering brain mechanisms and developing clinically relevant biomarkers of chronic pain

### 1.3 Research questions and outline

This work's objective is to identify neural mechanisms of pain perception and, in this way, inform the development of novel tools for the assessment and treatment of clinical pain conditions. We approached this research objective from two angles. First, we assessed neural mechanisms underlying sensory and contextual influences on acute pain in healthy individuals. Second, we investigated how the brain shapes the pain experience in patients with chronic pain. In both approaches, we relied on neurophysiological data as recorded by EEG. In particular, we wanted to understand whether and how communication between various neural systems, i.e., brain connectivity, is related to acute and chronic pain.

#### 1.3.1 Project 1: Brain connectivity in acute pain

In the first project, to better understand and directly compare how the brain serves sensory and contextual effects on pain, we reanalyzed data from an EEG experiment in which brief painful stimuli were applied to healthy human participants. In this experiment, "sensory and contextual information was modulated by varying stimulus intensity and expectations about upcoming stimulus intensity, respectively"[83]. In light of findings in earlier studies[83-85], we expected that sensory influences on pain would primarily be reflected by local brain activity. The influence of expectations on pain is a higher-level psychological phenomenon that likely relies on the communication between different brain regions. We, therefore, speculated that expectation-related effects on pain are particularly served by inter-regional connectivity. These considerations prompted us to "assess and directly compare how local oscillatory brain activity and inter-regional connectivity in a core network of six brain regions associated with the processing of pain serve the effects of stimulus intensity and expectations on pain"[3].

#### 1.3.2 Project 2: Network dysfunction in chronic pain

In the second project, we investigated brain network dysfunction in chronic pain. We primarily assessed how connectivity among intrinsic brain networks relates to the intensity of pain in patients with chronic pain. Intrinsic brain networks are complex-shaped, spatially distributed networks of brain regions that are synchronously activated during rest or particular tasks[50, 86]. While these networks are typically examined through hemodynamic signals obtained from fMRI, we assessed their involvement in chronic pain using EEG. Leveraging the high temporal resolution of EEG, we hoped to unravel temporal and spectral characteristics of network mechanisms that may elude traditional methods in the field, such as fMRI.

In this project, we adopted the definition of intrinsic brain networks provided by the Yeo atlas[86]. This is the most commonly used definition and divides the brain into seven networks. We primarily focused on four of these networks which figure prominently in the pathology of neuropsychiatric disorders[74, 87] and chronic pain[55, 74]: the somatomotor (a.k.a. pericentral) network, the frontoparietal (a.k.a. lateral frontoparietal/ control/ central executive) network, the salience ventral attention (a.k.a. midcingulo-insular/ salience) network, and the default (a.k.a. default mode/ medial frontoparietal) network. To analyze the function of these networks, we evaluated measures of network activity and inter-network connectivity from resting-state EEG recordings using a tool which we have developed for that purpose. In primary analyses, we related these measures to the reported intensity of the pain experience in patients with chronic back pain. In secondary analyses, we examined cohorts of patients with diverse pain conditions and assessed the specificity of our findings by considering depression severity as an alternative dependent variable. Several studies have reported associations between various neural measures and age[88-90]. To showcase the sensitivity of our methodology, we therefore also considered age as a dependent variable. In exploratory analyses, we employed a more data-driven approach to predict pain, depression, and age from measures of brain activity and brain connectivity. To test the robustness of findings, we implemented an extensive validation strategy involving up to five independent data sets which we acquired from various research groups across the globe.

### 1.3.3 Publications and code

Findings from the first project<sup>1</sup>, concerned with acute experimental pain, have been published in the peer-reviewed journal *Science Advances*[3]. The corresponding analysis code comprising both Matlab and R scripts is available at <https://osf.io/geanc/>. The data, which had been acquired in the context of a different project, are available at <https://osf.io/jw8rv/>. The analysis plan for the second project, concerned with chronic pain, has been preregistered on the Open Science Framework (OSF) platform (<https://osf.io/qa68n>). This dissertation presents results from the second project which have not yet been published. A corresponding manuscript intended for publication is work in progress. Upon publication, the code of this project will be made available at <https://osf.io/qa68n>.

---

<sup>1</sup> This thesis incorporates previously published text. Direct quotations are marked with quotation marks and accompanied by references to the original source. Square brackets indicate deletions.

## 2 Methodological basics

### 2.1 Electroencephalography

Electroencephalography (EEG) is a non-invasive technique for measuring electrical brain activity[91]. Compared to other functional imaging techniques like functional magnetic resonance imaging (fMRI) or positron emission tomography (PET), signals recorded by EEG offer a relatively low spatial but very high temporal resolution[92]. These properties render the method a useful tool for the assessment of system-level dynamic processes in the brain. One of the earliest mentions of EEG dates back to the 1920s and is linked to Hans Berger[93, 94] who, in an effort to decipher physiological mechanisms of telepathy, recognized oscillatory patterns in recordings of scalp potentials. While the research focus diverged from such parapsychological phenomena, EEG has become an indispensable tool in contemporary neuroscience and features of EEG signals have been linked to an array of cognitive and affective states.

#### 2.1.1 The physiological origin and significance of EEG

The human cerebral cortex is estimated to contain up to 16 billion interconnected neurons[95]. By current methods, measuring the individual activity of each of these neurons is infeasible. However, when larger groups of neurons are both structurally aligned and synchronously active, their combined signal can be detected by EEG sensors placed on the scalp.

Cortical pyramidal neurons meet these conditions[96], with their main dendritic branches being oriented perpendicularly to the cortical surface. These neurons receive excitatory as well as inhibitory synaptic inputs. Specifically, various ion channels embedded in the cell's membrane open or close in response to different neurotransmitters being released by presynaptic cells. This results in a net flux of charge across the cell membrane which locally alters the electric field potential. These alterations of electric field potentials are referred to as excitatory and inhibitory postsynaptic potentials, depending on whether they contribute to cell depolarization or hyperpolarization, respectively. When these postsynaptic potentials occur in a spatially and temporally coordinated manner, they add up and their combined magnitude induces potential differences on the scalp that lie within the sensitivity range of EEG systems. Importantly, these postsynaptic potentials are distinct from action potentials, which are rapid and large deflections of transmembrane potentials. Action potentials travel along axons which exhibit a lower degree of structural regularity and thus contribute a lot less to the signals measured by EEG[97].

As previously mentioned, EEG signals display rhythmic fluctuations known as brain rhythms or neural oscillations. A comprehensive account of the neuronal circuitry which generates these oscillations is beyond the scope of this thesis. More relevant to this work is a theory which explains the functional significance rather than the mechanistic cause of neuronal oscillations. The communication-through-coherence (CTC)[98, 99] framework posits that neuronal oscillations subserve selective communication between different neuronal assemblies. Specifically, different phases in the oscillatory cycle are associated with varying levels of neuronal excitability. As a result, a neuronal assembly's readiness to fire action potentials (i.e., send information) and its sensitivity with respect to synaptic input (i.e., integrate incoming information) both increase with higher levels of excitability. Consequently, information transfer preferentially occurs between coherently oscillating pre- and postsynaptic neuronal assemblies. Further, neuronal oscillations and coherence are not hard-wired in the brain but constitute emergent properties of neuronal circuitry, capable of dynamically adapting to situational demands. Thus, according to the CTC framework, adaptive and selective communication in the brain crucially relies on fluctuations of neuronal excitability which, in turn, are reflected in the oscillatory structure of EEG signals.

In summary, the CTC framework provides a plausible mechanistic basis for analyses of local oscillatory activity as well as inter-regional connectivity. It is, however, important to note that while the conceptual ideas behind most connectivity measures harmonize with the CTC framework, they are not rigorously derived from it. Instead, the CTC framework inspires the more abstract notion of communication as statistical relations between time series associated with different brain regions. An overview of some of these measures of connectivity is provided in section 2.1.5.

### 2.1.2 Recording devices

EEG systems commonly used for scientific investigations record signals from a range of 32 to 128 sensors, operating at sampling rates of 1000 Hz or higher, with a resolution of 24 bits in each channel[94]. The signals measured by these devices represent differences in electric field potentials. Specifically, they are defined as the potential difference between a recording and a designated reference sensor. In general, the placement of the reference sensor should be such that it picks up the same electrical artifacts as all other sensors. This theoretically ensures that variations of the electric field potential common to the reference and all other sensors, as is the case for external electrical artifacts, do not affect the measured EEG signal. More formally:

$$\begin{aligned} EEG_{signal} &= (V_{recording} + noise) - (V_{reference} + noise) \\ &= V_{recording} - V_{reference} \end{aligned}$$

While this so-called common mode rejection appears mathematically straightforward, its technical implementation is more involved. Achieving robustness against external influences, requires the amplification of potential differences, typically in the micro-Volt range, using operational amplifiers. These operational amplifiers implement the common mode rejection only imperfectly. Their capacity to reduce the influence of common modes on the EEG signal can, however, be enhanced by integrating an additional ground sensor which is also placed on the scalp[100]. Another source of inaccuracy in EEG measurements results from the variability of contact impedances across sensors. The most straightforward way to minimize this variability, is to reduce the absolute values of contact impedances at all sensors as much as possible by manually applying electrolytic gels. Notably, recent advances in electronics and materials science have facilitated novel, “dry” sensor types which do not rely on the application of such gels and therefore reduce the time and resources needed to conduct EEG[101] experiments.

### 2.1.3 Source reconstruction

The EEG signal recorded by a particular sensor may not provide the most accurate representation of brain activity directly beneath it. This can be attributed to two main factors. First, the electric field generated by a local source in the brain undergoes distortion due to the non-uniform distribution of electrical material properties throughout various tissues like the brain, skull, and scalp. Second, the electric field at a certain site constitutes a superposition of electric fields generated by different sources in the brain. To obtain a better estimate of neural activity at specific sites in the brain, one has to solve the so-called bioelectromagnetic inverse problem. In other words, for various locations in the brain, one has to determine the neural activity that most likely gave rise to the signals observed at the sensors.

Among the most commonly used methods for solving the inverse problem are equivalent dipole fitting, minimum norm estimation, and beamforming[102]. These methods differ in terms of the assumptions that are introduced in order to make the inverse problem solvable. To provide a basic methodological orientation, we present the simplest versions of these methods in the following.

In dipole fitting, one seeks to adjust the location, orientation, and magnitude of a fixed and typically small number of dipoles such that the sensor-level data is explained best[103].

The objective of minimum norm estimation[104, 105] is to identify a linear transformation that maps signals from the sensor-level to the source-level, with sources typically being densely distributed across the entire cortical sheet. This linear transformation, similarly to equivalent dipole fitting, is found by minimizing the distance between the observed and the predicted sensor-level signals. Let the sensor and estimated source-level signals be denoted by  $\mathbf{S} \in \mathbb{R}^{N_s \times n}$  and  $\hat{\mathbf{X}} \in \mathbb{R}^{dN_x \times n}$ , respectively. Herein,  $N_s$ ,  $N_x$ , and  $d$  denote the number of sensors, the number of sources, and the number of source dimensions (usually 3), respectively, while  $n$  corresponds to the number of data samples. Then, the described strategy can formally be expressed as a quadratic optimization problem of the form

$$\hat{\mathbf{X}}_{*t} = \underset{\mathbf{X}}{\operatorname{argmin}} \|\mathbf{L}\mathbf{X}_{*t} - \mathbf{S}_{*t}\| + \lambda \|\mathbf{X}_{*t}\|,$$

where the subscript  $*t$  indicates the  $t$ -th column of the associated matrix and  $\mathbf{L} \in \mathbb{R}^{k_s \times 3k_x}$  represents the so-called lead field matrix which maps source-level to sensor-level signals. The lead field matrix is obtained by simulating the electrical potential at the sensors separately for all sources, given the head's shape and material property distribution. This process is known as solving the forward problem.

The first term of the objective function in the above optimization problem quantifies the discrepancy between observations and predictions. The second term acts as a regularization term for the inverse problem[102, 106]. Without it, the optimization problem would possess infinitely many solutions. Recommendations for appropriate values of the scalar regularization parameter  $\lambda$  can be found elsewhere[107]. The analytical solution to the above optimization problem is[108]

$$\hat{\mathbf{X}}_{*t} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T + \lambda\mathbf{I})^{-1}\mathbf{S}_{*t},$$

where  $\mathbf{I}$  is the identity matrix. We see that the source-level signal is obtained as the product of a matrix  $\mathbf{G}^{\text{MNE}} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T + \lambda\mathbf{I})^{-1}$  and the sensor-level data  $\mathbf{S}_{*t}$ . Interestingly, the matrix  $\mathbf{G}^{\text{MNE}}$  does not depend on the data but solely on the lead field matrix  $\mathbf{L}$  and regularization parameter  $\lambda$ . The literature encompasses different variants of the minimum norm approach that incorporate additional prior knowledge via different weightings of the individual components in the distance and regularization terms[109].

Throughout this work, we utilized linearly constrained minimum variance (LCMV) spatial filters[110] which constitute a special instance of beamforming. As with minimum norm estimation, in beamforming we seek to identify a linear transformation that maps signals from the sensor-level to the source-level, i.e.,

$$\hat{\mathbf{X}}_{*t} = \mathbf{G}^{\text{LCMV}}\mathbf{S}_{*t}.$$



In the context of beamforming, the matrix  $\mathbf{G}^{LCMV}$  is often referred to as the spatial filter. For simplicity, let us assume that the orientation of sources is known, resulting in  $d = 1$ . Then, each source in the brain corresponds to one row of the matrix  $\mathbf{G}^{LCMV}$ . Unlike minimum norm estimation, where all rows of  $\mathbf{G}^{MNE}$  are estimated jointly, beamforming estimates each row of  $\mathbf{G}^{LCMV}$  independently and based on two criteria. To explain the underlying rationale, consider that an ideal spatial filter  $\mathbf{G}^{IDEAL}$  would satisfy  $\mathbf{X}_{*t} = \mathbf{G}^{IDEAL} \mathbf{L} \mathbf{X}_{*t}$ , implying  $\mathbf{G}^{IDEAL} \mathbf{L} = \mathbf{I}$ . Unfortunately, constructing a robust filter which satisfies this condition is infeasible. Instead, LCMV spatial filters provide an approximation by satisfying the equality condition  $\mathbf{G}_{k*}^{LCMV} \mathbf{L}_{*k} = 1, \forall k \in \{1, \dots, N_x\}$ . As before, the subscripts  $*k$  and  $k*$  refer to the  $k$ -th column and  $k$ -th row of the associated matrices, respectively. As the equality condition alone would leave the inverse problem under-determined, LCMV spatial filters, in addition, minimize the variance of the reconstructed source-level time series. The idea is to suppress any signals, including noise, not originating from the source of interest (spatial stop band), while transmitting signals from the source of interest with unit gain (spatial pass band). Assuming zero mean for the rows of  $\mathbf{S}$ , this leads to the constrained optimization problem

$$\begin{aligned} \mathbf{G}^{LCMV} &= \underset{\mathbf{G}}{\operatorname{argmin}} \|\mathbf{X}_{k*}\| \\ \text{subject to } &\mathbf{X}_{k*} = \mathbf{G}_{k*} \mathbf{S} \\ &\mathbf{G}_{k*} \mathbf{L}_{*k} = 1 \end{aligned}$$

for all  $k \in \{1, \dots, N_x\}$ . The solution to this constrained optimization problem is determined independently for each source:

$$\mathbf{G}_{k*}^{LCMV} = (\mathbf{L}_{*k}^{\top} (\mathbf{S} \mathbf{S}^{\top})^{-1} \mathbf{L}_{*k})^{-1} \mathbf{L}_{*k}^{\top} (\mathbf{S} \mathbf{S}^{\top})^{-1}$$

A generalization to the case in which the direction of sources is unknown is obtained in an analogous fashion. In this work we opted for LCMV spatial filters due to their effective noise suppression, their ability to reconstruct activity from deeper sources[111], and the fact that they do not require prior knowledge about the number or distribution of signal-generating sources in the brain.

#### 2.1.4 Representative time series

One research question addressed in this work is concerned with the association between connectivity among intrinsic brain networks and pain intensity. We define as intrinsic brain networks the seven Yeo networks [86] which have intricate and intertwined geometries. Due to the limited spatial resolution of EEG, the source-reconstructed activity of each

individual network is contaminated by the activity of all other networks. To disentangle the activity of individual networks, the following strategy is proposed.

Let's consider networks netA and netB, for which we aim to obtain the representative time series  $\mathbf{r}^A$  and  $\mathbf{r}^B$ , respectively. These representative time series should effectively capture the ground truth activity of netA and netB. In Figure 2-1a, we denote the ground truth activity of netA and netB as  $\alpha$  and  $\beta$ , respectively. However, due to the limited spatial resolution of EEG, we do not have direct access to this ground truth activity; instead, we only have access to a blurred estimate of it, which we denote as the source-reconstructed signals  $\mathbf{A} \in \mathbb{R}^{k_A \times n}$  and  $\mathbf{B} \in \mathbb{R}^{k_B \times n}$  with  $k_A$  and  $k_B$  denoting, respectively, the number of parcels belonging to networks netA and netB and  $n$  being the number of samples in the considered epoch. In the proposed approach, the representative time series  $\mathbf{r}^A$  and  $\mathbf{r}^B$  maximize the explained variance in  $\mathbf{A}$  and  $\mathbf{B}$ , respectively, while being orthogonal. By enforcing orthogonality between  $\mathbf{r}^A$  and  $\mathbf{r}^B$ , we ensure that there is no shared portion of variance between them. Specifically, any variance in, say,  $\mathbf{r}^A$  cannot be explained by  $\mathbf{r}^B$ . As  $\mathbf{r}^B$  is constructed to be representative of the activity in netB, this orthogonalization reduces the contamination of  $\mathbf{r}^A$  by the activity in netB. Formally, this approach can be expressed as the constrained optimization problem

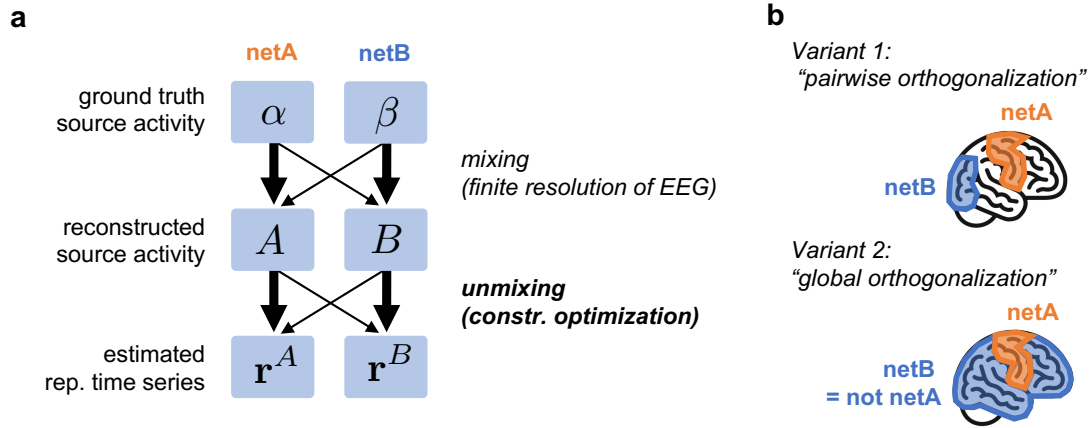
$$\begin{aligned} (\mathbf{r}_1, \mathbf{r}_2) &= \underset{\mathbf{r}_1, \mathbf{r}_2}{\operatorname{argmax}} (v(\mathbf{r}_1, \mathbf{A}) + v(\mathbf{r}_2, \mathbf{B})) \\ \text{subject to } \mathbf{r}_i &= \mathbf{w}_i^\top \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \quad \forall i \in \{1, 2\} \\ \mathbf{r}_i \mathbf{r}_j^\top &= \delta_{ij} \quad \forall i, j \in \{1, 2\}, \end{aligned}$$

where  $\delta$  is the Kronecker delta and, given that  $\mathbf{r}$  has unit length and the mean of the rows of  $\mathbf{X}$  equals zero,  $v(\mathbf{r}, \mathbf{X})$  is the fraction of the variance of matrix  $\mathbf{X}$  explained by vector  $\mathbf{r}$ , i.e.,

$$v(\mathbf{r}, \mathbf{X}) = \frac{\mathbf{r} \mathbf{X}^\top \mathbf{X} \mathbf{r}^\top}{\operatorname{tr}(\mathbf{X}^\top \mathbf{X})}.$$

For the sake of brevity in notation, the representative time series  $\mathbf{r}^A$  and  $\mathbf{r}^B$  have been denoted by  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , respectively, in the above equation. To solve this optimization problem, a standard iterative optimization algorithm, which employs information from both the local gradient and Hessian, is employed[112]. Note that, if we were to remove the orthogonality constraint  $\mathbf{r}_1 \mathbf{r}_2^\top = 0$ , the representative time series would simply correspond to the first principal components of the data matrices  $\mathbf{A}$  and  $\mathbf{B}$ .

Expanding on this idea, we can consider a variant of the method where instead of obtaining a pair of time series for each pair of networks ("*pairwise orthogonalization*"), we estimate a single time series for each network ("*global orthogonalization*"). To achieve this,



**Figure 2-1:** Methods for estimating representative time series. (a) Conceptual diagram of signal mixing due to imperfect source reconstruction (top half) and unmixing procedure (bottom half). (b) Illustration of two variants of the proposed method. Variant 1: For each network pair, one corresponding pair of orthogonal representative time series is determined. Variant 2: For each individual network, one representative time series is determined which is orthogonal to multiple (orthogonal) time series representing activity in all other networks.

we designate netA as the network of interest and define netB as all regions in the brain not contained within netA (cf. Figure 2-1b). Since our focus is on obtaining a single time series only for netA, we can describe the activity of netB using multiple orthogonal components. The associated optimization problem is

$$\mathbf{r}_1 = \underset{\mathbf{r}_1, \dots, \mathbf{r}_{N_c+1}}{\operatorname{argmax}} \left( v(\mathbf{r}_1, \mathbf{A}) + \sum_{i=2}^{N_c+1} v(\mathbf{r}_i, \mathbf{B}) \right)$$

$$\text{subject to } \mathbf{r}_i = \mathbf{w}_i^\top \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \quad \forall i \in \{1, \dots, N_c + 1\}$$

$$\mathbf{r}_i \mathbf{r}_j^\top = \delta_{ij} \quad \forall i, j \in \{1, \dots, N_c + 1\},$$

where  $N_c$  denotes the number of components used to describe the activity of netB and  $\mathbf{r}_1$  is the time series representative of netA. To decide which of the above-described variants should be used as the default, a simulation experiment was conducted. The results (Appendix Figure 0-1) indicate that for the given purpose, the global orthogonalization method with  $N_c = 3$  is optimal. Details of this simulation experiment are provided in appendix A.1.

### 2.1.5 Connectivity measures

Functional imaging methods yield time series that reflect the activity of different brain regions. By quantifying statistical relations between these time series, functional connectivity measures aim to capture inter-regional communication in the brain. There is not a universally agreed-upon rationale which would dictate a unique definition of connectivity. Instead, diverse perspectives on the matter have led to a plethora of different

definitions. A subset of these definitions, which is widely used and relevant to this thesis, will be discussed in the following.

### *Phase-based connectivity measures*

According to the CTC framework[99], neuronal oscillations originating from assemblies of synchronously active neurons are a prerequisite for selective inter-regional communication. Specifically, if neuronal oscillations in different brain regions are coherent, information transfer between the corresponding neuronal assemblies is facilitated. This idea inspired a range of connectivity metrics which quantify the consistency of phase differences between oscillatory signals over time.

The coherence coefficient constitutes the most basic measure of phase consistency between oscillatory signals. Before its definition is presented, some fundamental concepts need to be introduced. First, the Fourier transform  $A(f)$  of a signal  $a(t)$  is

$$A(f) = \int_{-\infty}^{\infty} a(t) \exp(-i2\pi ft) dt.$$

Expressed in polar coordinates, the Fourier transform reads  $A(f) = |A(f)|\exp(i\phi_A(f))$ , with  $|A(f)|$  and  $\phi_A(f)$  representing the frequency-dependent signal amplitude and phase-angle, respectively. The inverse Fourier transform is

$$a(t) = \int_{-\infty}^{\infty} A(f) \exp(i2\pi ft) df = \int_0^{\infty} |A(f)| \cos(\phi_A(f) + 2\pi ft) df.$$

The second equality results from algebraic transformations and demonstrates that the signal  $a(t)$  may be expressed as an amplitude-weighted superposition of sinusoidal components of different frequencies  $f$  and frequency-dependent phase shifts  $\phi(f)$ . As this thesis' focus will lie on frequency-resolved definitions of connectivity, it is convenient to primarily consider the spectral representation of signals and to omit the frequency-dependency in the notation. Using this notational convention, the spectral representation  $A(f)$  at frequency  $f$  of signal  $a(t)$  would simply be referred to as signal  $A$ .

Consider now a second signal, denoted by  $B$ . The cross-spectrum of signals  $A$  and  $B$  is

$$S_{AB} = AB^*,$$

where  $B^*$  denotes the complex conjugate of  $B$ . Note that the phase angle of the cross spectrum  $S_{AB}$  equals the phase angle difference or relative phase of signals  $A$  and  $B$ . The coherence coefficient, like many other connectivity metrics, quantifies the consistency of relative phase across multiple realizations of signals  $A$  and  $B$ . It is therefore convenient to

consider  $A$  and  $B$  as complex-valued random variables. Then, by use of the expectation operator  $\mathbb{E}[\cdot]$ , the complex-valued coherency can be defined as

$$C_{AB} = \frac{\mathbb{E}[S_{AB}]}{\sqrt{\mathbb{E}[A^2] \mathbb{E}[B^2]}},$$

where  $A^2$  and  $B^2$  denote the squared signal amplitudes of  $A$  and  $B$ , respectively. The absolute value or modulus of the coherency constitutes the coherence coefficient:

$$C_{AB}^{\text{coh}} = |C_{AB}|.$$

For fixed signal amplitudes  $|A|$  and  $|B|$ , the above expression is maximized if the variance of the relative phase of signals  $A$  and  $B$  is zero. Conversely, the larger the variability of phase angle differences of signals  $A$  and  $B$ , the smaller the value of the coherence coefficient.

One limitation of the coherence metric is its susceptibility to non-physiological contributions. These arise, e.g., from field spread which leads to the presence of common components in signals measured at different locations in the brain. Since identical signals are always perfectly coherent, the presence of common components artificially inflates the estimate of physiological connectivity. To mitigate this, alternative connectivity metrics have been proposed. One such metric is the imaginary part of coherency[113] (ImC):

$$C_{AB}^{\text{ImC}} = \text{Im}(C_{AB}) = \underbrace{\sin(\angle(C_{AB}))}_{\text{"penalty"}} C_{AB}^{\text{coh}},$$

where  $\angle(C_{AB})$  represents the phase angle of the coherency  $C_{AB}$ . The rationale behind the ImC is based on the insight that coherency resulting from field spread alone would exhibit a zero phase-angle. The imaginary part of coherency  $C_{AB}^{\text{ImC}}$  can be seen as a version of coherence  $C_{AB}^{\text{coh}}$  that penalizes small coherency phase angles as these are likely to result from field spread. This is illustrated by the second equality of the previous equation.

By construction, the ImC is rather insensitive to physiological connectivity occurring at small phase shifts between signals. Moreover, like the coherence coefficient, the ImC is reciprocally linked to the power of the constituent signals. Additional uncorrelated noise in these signals, therefore, reduces the value of the ImC. These properties render the ImC a rather conservative estimate of physiological connectivity. A metric that promises to be more sensitive while being even less susceptible to volume conduction is the phase lag index[114] (PLI). The PLI defines connectivity as the expected sign of the imaginary part of the cross-spectrum:

$$C_{AB}^{\text{PLI}} = \mathbb{E} \left[ \text{sgn} \left( \text{Im}(S_{AB}^i) \right) \right].$$

The sign of the imaginary part of  $S_{AB}$ , which is a binary indicator of the relative phase of signals  $A^i$  and  $B^i$ , is the only information about signals  $A$  and  $B$  that enters the definition of the PLI. It follows that, as compared to the ImC, the influence of the phase difference magnitude is reduced, albeit not eliminated. Moreover, since noise affects the PLI only if it leads to a sign change in the imaginary part of the cross spectrum, the PLI was shown to be less sensitive to noise. A further reduction in noise sensitivity is achieved by a related metric called the weighted PLI[115] (wPLI):

$$C_{AB}^{\text{wPLI}} = \frac{|\mathbb{E}[\text{Im}(S_{AB})]|}{\mathbb{E}[|\text{Im}(S_{AB})|]}.$$

All metrics presented thus far have been defined in terms of expectations of random variables. A common problem among these metrics is that their estimates which are based on limited numbers of realizations, that is trials or epochs, are biased. In practical terms, differences in connectivity between two conditions may thus be due to systematic differences in trial numbers and not due to actual differences in physiological connectivity. For the squared wPLI an estimator has been proposed which reduces this sample size bias. This estimator is simply referred to as the debiased wPLI[115] (dwPLI) and is given by

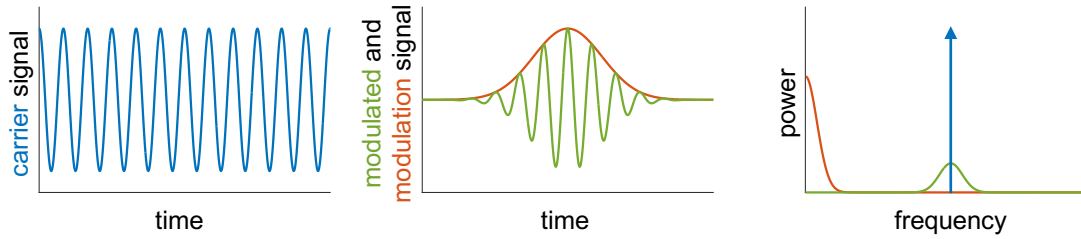
$$\hat{C}_{AB}^{\text{wPLI}} = \frac{\sum_{i=1}^N \sum_{j \neq i}^N \text{Im}(S_{AB}^i) \text{Im}(S_{AB}^j)}{\sum_{i=1}^N \sum_{j \neq i}^N |\text{Im}(S_{AB}^i) \text{Im}(S_{AB}^j)|},$$

where superscripts  $i$  and  $j$  refer to different signal realizations (e.g., trials or epochs),  $N$  denotes the number of realizations, and the hat ( $\hat{\cdot}$ ) in  $\hat{C}_{AB}^{\text{wPLI}}$  serves to distinguish the estimator from its population value. Importantly, this estimator is not completely unbiased. It can be shown that the dwPLI may have a slight positive or negative bias.

### *Amplitude-based connectivity measures*

The amplitude-based connectivity metric presented in this section was first introduced to study brain-wide correlation structures of oscillatory activity akin to the intrinsic resting state networks observed in fMRI[116]. As the name suggests, amplitude-based connectivity metrics assess the relation between time-varying amplitudes of oscillatory signals. As such, they are not immune to spurious contributions due to field spread. The amplitude envelope correlation[116, 117] (AEC) metric mitigates effects of field spread by computing envelope correlations between asymmetrically orthogonalized signals.

The first step in computing the AEC between signals is the transformation of these signals into complex-valued time-frequency representations (cTFRs). For a given frequency and



**Figure 2-2:** Effects of amplitude modulation on spectral representations. Multiplying a sinusoidal carrier signal (blue) with a gaussian modulation signal (red) results in the modulated signal (green). The power spectral density of the modulated signal is wider than that of the carrier signal.

at a given time, the modulus and phase angle of a signal's cTFR correspond to its instantaneous amplitude and instantaneous phase, respectively. There are multiple ways in which to compute cTFRs, but not all are appropriate in every situation. To select the correct approach, the effects of amplitude modulation on the spectral representation of a signal must be understood. To this end, consider the straightforward example depicted in Figure 2-2: Let there be a sinusoidal carrier signal which is represented in the frequency domain by a Dirac impulse (i.e., an infinitely narrow peak) located at some frequency  $f_c$ . Multiplying the carrier signal with a Gaussian modulation signal results in an amplitude-modulated signal, the spectral representation of which takes the form of a Gaussian<sup>2</sup> (i.e. a peak of finite width) centered at frequency  $f_c$ . This illustrates that modulating the amplitude of a band-limited signal leads to a widening of its frequency spectrum. Conversely, amplitude variations of signals within a given frequency band are (at least partially) driven by contributions from outside of that frequency band. Also, the narrower a frequency band of interest, the larger the relative influence of contributions from outside of that frequency band on the amplitude variations of signals within that frequency band.

This implies that, for narrow frequency bands of interest, the broad band signal should not be band-pass filtered prior to the computation of cTFRs. Instead, the cTFR of signal  $a(t)$  for a narrow frequency band centered at frequency  $f_c$  may, e.g., be obtained by convolving the broad-band signal  $a(t)$  with a complex-valued and frequency-dependent Morlet-wavelet[118]  $w_{f_c}(t)$ , i.e.

$$\alpha_{f_c}^{\text{morlet}}(t) = a(t) * w_{f_c}(t).$$

<sup>2</sup> A multiplication in the time domain corresponds to a convolution in the frequency domain. Note that the frequency representation of a Gaussian signal is also a Gaussian and the convolution of a Gaussian with a Dirac impulse yields another Gaussian.

If, by contrast, the frequency band of interest is relatively wide, band-pass filtering is not expected to alter amplitude time courses much and the cTFR of a signal  $a(t)$  may be obtained via band-pass filtering and Hilbert transformation, i.e.

$$a_{f_c}^{\text{hilbert}}(t) = a_{f_c}(t) + i \hat{a}_{f_c}(t),$$

where  $a_{f_c}(t)$  and  $\hat{a}_{f_c}(t)$  denote the band-pass filtered signal and its Hilbert transform, respectively. For notational convenience, the explicit distinction between cTRFs obtained using Morlet wavelets or Hilbert transforms will be omitted from now on.

As noted above, the AEC mitigates effects of volume conduction by means of orthogonalization. Specifically, it is the instantaneous phase angles of signals that are orthogonalized in this approach. Note that field spread leads to the presence of common components in signals measured at different locations in the brain. The phase angle difference among such common components is exactly zero. Conversely, signal pairs exhibiting a phase difference of close to 90 degrees (i.e. orthogonal signals) are thought to be relatively unaffected by field spread<sup>3</sup>. One may conclude that retaining only signal components that exhibit phase differences of exactly 90 degrees, i.e., orthogonal signal components, might mitigate effects of field spread. This orthogonalization can be achieved as follows: For a certain frequency and at a given time, let  $\alpha$  and  $\beta$  be the values of two cTRFs. Then,  $\alpha^{\perp\beta}$ , the modified version of  $\alpha$  that is orthogonal to  $\beta$ , results from solving the optimization problem

$$\begin{aligned} \alpha^{\perp\beta} = & \underset{\alpha'}{\operatorname{argmin}} \|\alpha' - \alpha\|^2 \\ \text{subject to} & \operatorname{Re}(\alpha'\beta^*) = 0 \end{aligned}$$

The constraint enforces orthogonality and the objective function ensures that the modified signal  $\alpha^{\perp\beta}$  is as similar as possible to the original signal  $\alpha$ . Using the Lagrange multiplier method, one obtains the solution

$$\alpha^{\perp\beta} = \alpha - \frac{\operatorname{Re}(\alpha\beta^*)}{|\beta|^2} \beta.$$

Analogously, one may obtain  $\beta^{\perp\alpha}$ , the modified version of  $\beta$  that is orthogonal to  $\alpha$ . This orthogonalization procedure is conducted independently for each point in time and it results in two variants of orthogonalized signal pairs. The AEC is defined as the average

---

<sup>3</sup> This statement relies on the assumption that instantaneous phase differences between physiological signals are bounded to a range between 0 and 90 degrees.



of the power envelope correlations of both signal pairs. When regarding  $\alpha$  and  $\beta$  as complex-valued random variables, the AEC can formally be defined as<sup>4</sup>

$$C_{AB}^{\text{AEC}} = \frac{1}{2} \left( \rho(|\alpha^{\perp\beta}|, |\beta|) + \rho(|\beta^{\perp\alpha}|, |\alpha|) \right),$$

where  $\rho(\cdot, \cdot)$  represents the Pearson correlation. Although the formal derivations of the working principle of AEC and phase-orthogonalization may appear somewhat ad hoc, experimental evidence demonstrates that the AEC provides meaningful insights into the network-level correlation patterns of oscillatory brain activity. Notably, the AEC could capture established phenomena, such as the inter-hemispheric coupling of spontaneous activity in corresponding sensory areas[116].

### *Measures of directed connectivity*

The preceding paragraphs focused on connectivity metrics that quantify undirected relationships between signals. Next, connectivity metrics that discern the direction of information flow will be discussed. A typical starting point in the derivation of such metrics is the description of autoregressive models (ARMs). Let  $\mathbf{x}(t) \in \mathbb{R}^n$  denote the value of an  $n$ -dimensional time series at time  $t$ . The ARM assumes that the current value of a time series can be predicted by past values of that same time series, i.e.,

$$\mathbf{x}(t) = \sum_{\tau=1}^{\infty} \mathbf{a}(\tau) \mathbf{x}(t - \tau) + \mathbf{e}(t)$$

where  $\mathbf{a}(\tau) \in \mathbb{R}^{n \times n}$  is a matrix of coefficients evaluated at time lag  $\tau$  and  $\mathbf{e}(t)$  is an  $n$ -dimensional zero-mean random vector with time-invariant covariance matrix  $\Sigma$ . As it simplifies the derivations of directed connectivity metrics, it will henceforth be assumed that the covariance matrix  $\Sigma$  has a diagonal shape which corresponds to the absence of any (non-causal) instantaneous interactions between signals. The first measure of directed connectivity to be derived on the basis of the ARM, is the time domain definition of Granger causality[119] (tGC). For a pair of time series ( $n = 2$ ), the tGC metric compares the residuals (or prediction errors) from two model configurations. The first configuration does not allow for interactions between components of  $\mathbf{x}$  (only diagonal entries of  $\mathbf{a}$  may deviate from zero), whereas the second configuration does allow for such interactions (all entries of  $\mathbf{a}$  may deviate from zero). Fitting the model based on both the diagonal and full

---

<sup>4</sup> Originally, the AEC has been defined not in terms of amplitude correlations but in terms of correlations between the logarithms of squared amplitudes.

coefficient matrices results in the variance estimates  $\text{Var}[e_i]_{\text{diag}}$  and  $\text{Var}[e_i]_{\text{full}}$ , respectively. Signal  $x_j$  is then said to Granger cause  $x_i$  if the quantity

$$C_{j \rightarrow i}^{\text{tGC}} = \ln \left( \frac{\text{Var}[e_i]_{\text{full}}}{\text{Var}[e_i]_{\text{diag}}} \right)$$

is substantially larger than 0. Note that values below 0 are theoretically infeasible as the model without interaction terms is a special instance of the model including interaction terms. Consequently, the latter will fit the data at least as well as the former.

To derive various spectral definitions of directed connectivity, the autoregressive model is first transformed to the frequency domain:

$$\mathbf{A}(f)\mathbf{X}(f) = \mathbf{E}(f) \Leftrightarrow \mathbf{X}(f) = \mathbf{H}(f)\mathbf{E}(f),$$

where  $\mathbf{A}(f)$ ,  $\mathbf{X}(f)$ , and  $\mathbf{E}(f)$  are the spectral representations of  $\mathbf{a}(t)$ ,  $\mathbf{x}(t)$ , and  $\mathbf{e}(t)$ , respectively.  $\mathbf{H}(f) = \mathbf{A}(f)^{-1}$  denotes the so-called spectral transfer matrix of the system. This result is obtained by rearrangement and subsequent Z-transformation of the ARM. Again, for the sake of notational simplicity, the explicit mention of the frequency dependence will be omitted from now on. Multiplying both sides of the spectral representation of the ARM with their respective Hermitian transposes yields the important decomposition of the cross-spectral matrix:

$$\mathbf{S} = \mathbf{H}\mathbf{\Sigma}\mathbf{H}^H,$$

The diagonal entries of the cross-spectral matrix  $\mathbf{S}$  represent the squared amplitude (i.e., power) of individual signals; the off-diagonal entries correspond to the cross-spectra previously introduced in the context of phase-based connectivity metrics<sup>5</sup>. Note that the above decomposition was derived here on the basis of an ARM. Alternatively, this decomposition can be obtained directly from the cross-spectral matrix  $\mathbf{S}$  via a technique called spectral matrix factorization[120]. While fitting an ARM requires explicit assumptions regarding the model order, spectral matrix factorization is completely non-parametric. In the present work, only the non-parametric strategy was employed. Expanding the above decomposition, while still assuming that  $\mathbf{\Sigma}$  is diagonal and  $n = 2$ , results in

$$S_{ii} = H_{i1}\Sigma_{11}H_{i1}^* + H_{i2}\Sigma_{22}H_{i2}^*$$

It follows that the power of, say, signal  $x_1$  has an intrinsic contribution ( $H_{11}\Sigma_{11}H_{11}^*$ ) and a causal contribution ( $H_{12}\Sigma_{22}H_{12}^*$ ) attributable to the influence of signal  $x_2$ . Therefore, signal

---

<sup>5</sup> The relation  $\mathbf{E}\mathbf{E}^H = \mathbf{\Sigma}$  follows from, e.g., Parseval's theorem

$x_j$  is said to Granger cause  $x_i$  at a given frequency if the total power of signal  $x_i$  at that frequency is substantially larger than its intrinsic contribution, that is, if the quantity

$$C_{j \rightarrow i}^{\text{fGC}} = \ln \left( \frac{S_{ii}}{H_{i1} \Sigma_{11} H_{11}^*} \right)$$

is substantially larger than zero. The quantity  $C_{j \rightarrow i}^{\text{fGC}}$  corresponds to the frequency domain definition of Granger causality[121] (fGC). The tGC and fGC metrics infer directed information flow independently for each connection in a network. In contrast, multivariate measures of connectivity estimate directed information flow within a network by considering all connections jointly. One such metric is directed coherence[122] (DC), which is derived by rewriting the definition of coherence using the decomposition of the cross-spectral matrix provided above. Specifically, the coherence between signals  $x_i$  and  $x_j$  can be written as

$$C_{ij}^{\text{coh}} = \frac{\sum_{k=1}^n H_{ik} \Sigma_{kk} H_{jk}^*}{\sqrt{S_{ii} S_{jj}}}$$

The auto-coherence  $C_{ii}^{\text{coh}}$  of signal  $x_i$  is always one and, hence, the power of signal  $x_i$  can be expressed as

$$S_{ii} = S_{ii} C_{ii}^{\text{coh}} = S_{ii} \sum_{j=1}^n \underbrace{\frac{|H_{ij}|^2 \Sigma_{jj}}{S_{ii}}}_{>0}$$

The component of the transfer matrix  $\mathbf{H}$  that corresponds to signal  $x_j$ 's influence on  $x_i$  is  $H_{ik}$ . Therefore, the relative contribution of signal  $x_j$  to the power of  $x_i$  is linked to the term

$$C_{j \rightarrow i}^{\text{DC}} = \frac{H_{ij} \sqrt{\Sigma_{jj}}}{\sqrt{S_{ii}}} = \frac{H_{ij} \sqrt{\Sigma_{jj}}}{\sqrt{\sum_{k=1}^n H_{ik} \Sigma_{kk} H_{ik}^*}},$$

which constitutes the DC metric. A disadvantage of DC is its limitation to cases where the noise covariance matrix  $\Sigma$  is diagonal. This shortcoming gets addressed by the directed transfer function[123] (DTF) metric which is defined solely in terms of the transfer matrix  $\mathbf{H}$ :

$$C_{j \rightarrow i}^{\text{DTF}} = \frac{H_{ij}}{\sqrt{\sum_{k=1}^n H_{ik} H_{ik}^*}}.$$

This metric quantifies the influence of signal  $x_j$  on signal  $x_i$  as the normalized  $ij$ -th entry of the spectral transfer matrix  $\mathbf{H}$ . The normalization factor  $(\sum_{k=1}^n H_{ik} H_{ik}^*)^{-\frac{1}{2}}$  reflects the combined influence of all signals on signal  $x_i$ . Hence, in simpler terms, the DTF quantifies

the information transmitted from  $x_j$  to  $x_i$  relative to all information received by  $x_i$ . The DTF metric results from the DC metric by setting  $\Sigma_{kk} = 1, \forall k = 1, \dots, n$ .

By following analogous steps as in the derivation of the DTF metric, but using partial coherence instead of ordinary coherence as a starting point, one arrives at the partial directed coherence[124] (PDC) metric. In a setting with  $n > 2$  the partial coherence between signals  $x_i$  and  $x_j$  is defined as the ordinary coherence between the residuals of two models predicting  $x_i$  and  $x_j$  on the basis of all signals other than  $x_i$  and  $x_j$ . Consequently, in the absence of any other signals, i.e., for  $n = 2$ , partial and ordinary coherence are identical. The PDC metric is given by

$$C_{j \rightarrow i}^{\text{PDC}} = \frac{\delta_{ij} - A_{ij}}{\sqrt{\sum_{k=1}^n (\delta_{kj} - A_{kj})(\delta_{kj} - A_{kj}^*)}},$$

where  $\delta_{ij}$  denotes the Kronecker delta. In simpler terms and as opposed to DTF, PDC can be thought of as quantifying the information transmitted from  $x_j$  to  $x_i$  relative to all information sent by  $x_j$ . From the above-mentioned relation between partial and ordinary coherence, it follows also that  $C_{j \rightarrow i}^{\text{PDC}} = C_{j \rightarrow i}^{\text{DTF}}$  for  $n = 2$ . For  $n > 2$ , it was shown by means of simulations that the PDC more accurately represents ground truth connectivity than the DTF.

## 2.2 Statistics

The presence of specific data patterns has traditionally been established using p-values derived from frequentist statistical tests. The degree to which a specific pattern is present in the data is measured by the so-called test statistic. The p-value represents the probability of obtaining a value for the test statistic that is at least as extreme as the actually observed value, under the assumption that the pattern, which is quantified by the t-statistic, is absent in the data. This assumption is commonly referred to as the null hypothesis. While small p-values indicate evidence against the null hypothesis, large p-values must not be interpreted as evidence in favor of it. To see why, note that if the null hypothesis is valid, the distribution of p-values is, by construction, uniform on the interval  $[0,1]$ . Therefore, if the null hypothesis holds, any p-value is as likely to occur as any other. It is only when the null hypothesis is false that the distribution of p-values shifts more towards smaller values. In many research applications it would be of great value to be able to quantify evidence in favor of the null hypothesis. Bayesian statistics offers this possibility. In the following paragraphs, a brief introduction to Bayesian inference and hypothesis testing is provided.

### 2.2.1 Bayes factors and model comparisons

Let  $A$  and  $B$  be two random variables with joint probability distribution  $p(A,B)$ . The conditional probability distributions of  $A$  conditioned on  $B$  and of  $B$  conditioned on  $A$  are given by  $p(A|B) = p(A,B)/p(B)$  and  $p(B|A) = p(A,B)/p(A)$ , respectively. It follows that  $p(A|B)p(B) = p(B|A)p(A)$  or equivalently

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

This relation is known as Bayes' rule and constitutes the core of Bayesian statistics. In the context of hypothesis testing, the variables  $A$  and  $B$  can be identified with the data  $D = (Y,X)$  and hypothesis parameters  $\theta$  (e.g., effect size parameter), respectively. All probability distributions occurring in the formal statement of Bayes' rule are implicitly dependent on a hypothetical model structure  $\mathcal{M}$ . The term model structure is used here to emphasize that  $\mathcal{M}$  accommodates a range of models, each of which corresponds to a particular value of the parameter  $\theta$ . By making the dependence on  $\mathcal{M}$  explicit, one may rewrite the expression to the right of the equality sign of Bayes' rule as  $p(\theta|X,Y,\mathcal{M})$ . This conditional probability distribution is known as the posterior and quantifies the relative likelihood of different parameter values  $\theta$  having given rise to the observed data  $(X,Y)$ , given the model structure  $\mathcal{M}$ . The numerator of the fraction to the right of the equality sign

of Bayes' rule corresponds to the product of the so-called likelihood and prior distribution. The likelihood  $p(X, Y|\theta, \mathcal{M})$  is simply the likelihood of the observed data  $(X, Y)$  under a given model structure  $\mathcal{M}$  and for a specific parameter value  $\theta$ . The prior  $p(\theta|\mathcal{M})$  reflects any a priori knowledge that one might have about the model parameters  $\theta$ . The denominator  $p(D|\mathcal{M})$  is referred to as the model evidence and describes the likelihood of the data given the model structure  $\mathcal{M}$ . It can be expressed as the expected value of the likelihood  $p(X, Y|\theta, \mathcal{M})$  with respect to the prior distribution  $p(\theta|\mathcal{M})$ , i.e.

$$p(D|\mathcal{M}) = \mathbb{E}_{p(\theta|\mathcal{M})}[p(D|\theta, \mathcal{M})] = \int p(D|\theta, \mathcal{M}) p(\theta|\mathcal{M}) d\theta.$$

On the basis of the model evidence and by employing Bayes' rule again, the posterior likelihood of the model structure given the data becomes

$$p(\mathcal{M}|D) = \frac{p(D|\mathcal{M})p(\mathcal{M})}{p(D)}.$$

The ratio of the posterior likelihoods of two competing model structures  $\mathcal{M}_1$  and  $\mathcal{M}_2$  with prior odds  $p(\mathcal{M}_1)/p(\mathcal{M}_2)$  is

$$BF_{10} = \frac{p(\mathcal{M}_1|D)}{p(\mathcal{M}_2|D)} = \frac{p(D|\mathcal{M}_1)p(\mathcal{M}_1)}{p(D|\mathcal{M}_2)p(\mathcal{M}_2)}.$$

Ratios of this kind are known as Bayes factors (BFs). BFs compare the capacity of two models to describe the data at hand. BFs remain informative even if the model structures to be compared have different complexities. This is remarkable as a more complex model usually describes the data at least as well as a less complex model. BFs mitigate this trivial advantage of more complex models by performing the comparison across a range of plausible parameter values and not just for a single optimal parameter value. A complex yet inadequate model typically exhibits a high likelihood only for a small set of parameters values. By contrast, a simpler but more adequate model might exhibit a fairly high likelihood for a large set of parameter values. This simpler model's evidence will, therefore, surpass that of the complex model, despite its maximum likelihood potentially being smaller than that of the complex model.

The concept of Bayes' rule and Bayes factors can be translated into various statistical tests. Therein,  $BF > 3$  and  $BF > 10$  are considered moderate and strong evidence in favor of an effect of interest, whereas  $BF < 1/3$  and  $BF < 0.1$  indicate moderate and strong evidence against an effect of interest, respectively [125]. A selection of Bayesian hypothesis tests, which is relevant for this thesis, is briefly described in the following. First, the derivation of a Bayesian two-sample t-test is roughly outlined. Then some aspects of Bayesian linear (mixed) models and ANOVAs are discussed. The introduction to Bayesian

two-sample t-tests and ANOVAs provides the context for proposing a strategy for comparing sets of logistic regression models.

## 2.2.2 Specific Bayesian statistical tests

### *Bayesian t-test*

Assessing the mean difference between two samples is probably the most frequently encountered statistical task in science. Therefore, this task is used here to exemplify how a statistical question can be translated into a formal test in the Bayesian framework. Let the data  $D$  consist of data points  $X_i$  with  $i = 1, \dots, N_X$  and  $Y_j$  with  $j = 1, \dots, N_Y$ . This data may be described using two normal distributions (model structure  $\mathcal{M}$ ) with means  $\mu_X$  and  $\mu_Y$  and common standard deviation  $\sigma$  (model parameters  $\theta = (\mu_X, \mu_Y, \sigma)$ ). The likelihood of the data given the model structure and parameters is

$$p(D|\theta, \mathcal{M}) = \prod_{i=1}^{N_X} \mathcal{N}(X_i; \mu_X, \sigma) \prod_{j=1}^{N_Y} \mathcal{N}(Y_j; \mu_Y, \sigma).$$

Since the absolute values of  $\mu_X$  and  $\mu_Y$  are not of interest, the likelihood can be reformulated in terms of  $\theta' = (\delta, \mu, \sigma)$  with effect size  $\delta = (\mu_Y - \mu_X)/\sigma$  and grand mean  $\mu = (\mu_X + \mu_Y)/2$  [126]. The re-parametrized likelihood is denoted by  $p'(D|\theta', \mathcal{M})$ . To compute the model evidence, a prior distribution  $p(\theta') = p(\delta, \mu, \sigma)$  over the new model parameters needs to be defined. In ref. [127], the following decomposition of the prior is proposed

$$p(\delta, \mu, \sigma|\mathcal{M}) = p(\delta|\mathcal{M})p(\mu)p(\sigma),$$

where  $p(\delta|\mathcal{M})$  is a generic probability distribution that can be defined flexibly to reflect prior beliefs about the effect size parameter  $\delta$ . The distributions  $p(\mu) \propto 1$  and  $p(\sigma) \propto 1/\sigma$  are improper<sup>6</sup> Jeffreys priors for the location parameter  $\mu$  and scale parameter  $\sigma$ , respectively. Jeffreys priors reflect minimal a priori knowledge. They are constructed to be invariant under reparametrizations of the likelihood. In the present case, this implies that rescaling the data will not affect the resulting Bayes factors. The basic principle of a parametric Bayesian t-test is to compare the evidence of model structures  $\mathcal{M}_0$  and  $\mathcal{M}_1$  reflecting the null and the alternative hypotheses, respectively. Practically, these two model structures only differ with respect to the feasible range of the effect size parameter

---

<sup>6</sup> These priors are described as "improper" as they do not have a finite integral. They can still be employed as long as the resulting posterior does have a finite integral.

$\delta$ . Therefore, the expression for the evidence of both model structures is identical except for the effect size prior  $p(\delta|\mathcal{M})$ . It can be written as

$$\begin{aligned} p(D|\mathcal{M}_i) &= \iiint p'(D|\delta, \mu, \sigma, \mathcal{M}_m) p(\mu)p(\sigma) d\mu d\sigma p(\delta|\mathcal{M}_m) d\delta \\ &= \int T_d(t; \sqrt{N_{\text{eff}}}\delta) p(\delta|\mathcal{M}_m) d\delta, \end{aligned}$$

with  $m \in \{0,1\}$  and  $t$  being the t-statistic of the data  $D$ . In other words, marginalizing the likelihood with respect to parameters  $\mu$  and  $\sigma$  yields an easy-to-evaluate expectation integral of a t-distribution with  $d = N_X + N_Y - 2$  degrees of freedom and effective sample size  $N_{\text{eff}} = (1/N_X + 1/N_Y)^{-1}$ . The null hypothesis postulates the absence of an effect which corresponds to  $\delta = 0$ . The associated prior  $p(\delta|\mathcal{M}_0)$  is thus a Dirac impulse (infinitely narrow peak) at location  $\delta = 0$ . The alternative hypothesis assumes the presence of an effect which corresponds to  $\delta \neq 0$ . Further, the alternative hypothesis should reflect the belief that larger absolute effect sizes are less likely than smaller ones. Therefore, a common choice for the prior  $p(\delta|\mathcal{M}_1)$  of the alternative hypothesis is the centered Cauchy distribution. Finally, with these prior specifications, the Bayes factor  $BF_{10}$  describing the relative evidence for the alternative over the null hypothesis can be written as

$$BF_{10} = \frac{\int T_d(t; \sqrt{N_{\text{eff}}}\delta) \pi(\delta; \gamma) d\delta}{T_d(t; 0)},$$

where  $\pi(\delta; \gamma)$  is the centered Cauchy distribution with scale parameter  $\gamma$ . Note that, irrespective of the data, as  $\gamma \rightarrow 0$ , the Bayes factor  $BF_{10}$  always approaches 1. The default prior used throughout this work is obtained by setting  $\gamma = 1/\sqrt{2}$ . This choice is said to be representative of effect sizes usually observed in the social sciences[125, 128].

### *Bayesian ANOVA and linear models*

Bayesian ANOVAs are best understood from the perspective of linear models. Therefore, linear models and their Bayes factors are briefly outlined in the following. Consider the following representation of a linear model which describes the association between a dependent variable  $y$  and a set of predictor variables  $\{\mathbf{x}_1, \dots, \mathbf{x}_K\}$

$$y = \mu + \sum_{k=1}^K \mathbf{x}_k \boldsymbol{\beta}_k + \epsilon \quad \text{with } \epsilon \sim \mathcal{N}(0, \sigma^2),$$

where  $\mu$  is a bias parameter,  $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_K$  denote coefficient vectors, and  $\sigma^2$  represents the error variance. If the  $k$ -th predictor is continuous, then  $\mathbf{x}_k \in \mathbb{R}$  and  $\boldsymbol{\beta}_k \in \mathbb{R}$  are scalar quantities. If, by contrast, the  $l$ -th predictor is categorical and has  $N_l$  levels, then  $\mathbf{x}_l \in$



$[0,1]^{N_l}$  is a vector quantity representing the binary encoding of the different levels<sup>7</sup>. Accordingly,  $\beta_l \in \mathbb{R}^{N_l}$  is also vector-valued. Many different statistical tests can be derived as special instances of the described linear model. For example, consider the case of only one categorical predictor  $x_1$  with associated coefficient vector  $\beta_1$ . In a frequentist setting, to assess whether the predictor is associated with the dependent variable, one tests the null hypothesis  $\beta_1 = \mathbf{0}$ . To this end, an estimate of the covariance matrix of  $\beta_1$  is required. Depending on how this covariance matrix is estimated, one may compute p-values by comparing a Wald statistic to a  $\chi^2$ -distribution or an F-statistic to an F-distribution. The latter corresponds to performing a conventional one-way ANOVA.

In a Bayesian setting, the relevance of a predictor is established using a different line of argumentation[129]. Rather than computing a p-value associated with a certain predictor, one compares the evidence of models which include that predictor to the evidence of models which do not include it. To make this more concrete, consider the example of an experimental design involving the factors  $F_1$  and  $F_2$ . Some outcome variable  $y$  occurring in this experiment may be described by the model

$$y = \mu + \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{x}_{1 \times 2}\beta_{1 \times 2} + \epsilon \quad \text{with } \epsilon \sim \mathcal{N}(0, \sigma^2),$$

where  $\mathbf{x}_k$  and  $\beta_k$  denote the binary encodings and the parameter vector associated with factor  $F_k$  with  $k \in \{1,2\}$ , respectively. The vectors  $\mathbf{x}_{1 \times 2}$  and  $\beta_{1 \times 2}$  capture the association between  $y$  and the interaction of the two factors. More abstractly, this model can be written as

$$y \sim 1 + F_1 + F_2 + F_1 \times F_2$$

A Bayes factor which quantifies, e.g., the interaction effect of  $F_1$  and  $F_2$  on  $y$  can then be defined as

$$BF_{F_1 \times F_2} = \frac{\mathcal{E}(y \sim 1 + F_1 + F_2 + F_1 \times F_2)}{\mathcal{E}(y \sim 1 + F_1 + F_2)},$$

where  $\mathcal{E}(\cdot)$  denotes the evidence of the model inside the parentheses. To quantify the main effect of, say, factor 1, one could compute a Bayes factor as the ratio of the averaged evidence of models including factor 1 and the averaged evidence of models not including factor 1, i.e.,

---

<sup>7</sup> This type of encoding leads to a linear dependence among the columns of the design matrix. As a consequence, the model parameters are not uniquely identifiable. To solve this problem, additional constraints have to be introduced. For fixed effects, a commonly used constraint ensures that the sum of coefficient vector entries for each predictor equals zero.

$$BF_{F_1} = \frac{\frac{1}{2}[\mathcal{E}(y \sim 1 + F_1 + F_2) + \mathcal{E}(y \sim 1 + F_1)]}{\frac{1}{2}[\mathcal{E}(y \sim 1) + \mathcal{E}(y \sim 1 + F_2)]}.$$

Note that this definition of the Bayes factors is not unique. For example, it could be argued that the average in the numerator of  $BF_{F_1}$  should additionally include the evidence of the full model, i.e., of the model which also includes the interaction term. In the present work, the above-presented Bayes factor definitions are employed as they are said to yield more conservative estimates of the effects of interest[125].

As mentioned previously, to compute a model evidence, prior distributions over the model parameters have to be defined. For hypothesis testing, these prior distributions should introduce as few a priori assumptions as possible[129, 130]. Therefore, one requirement for such a prior is that any inference remains invariant under linear transformations of the continuous data. In an ANOVA setting this means that changing the measurement unit of the dependent variable must not change the resulting Bayes factors. As with the Bayesian t-test, this is achieved by placing Jeffreys priors on the location and scale parameters, i.e.,  $(\mu, \sigma) \sim \frac{1}{\sigma}$ . Further, for categorical predictors, a Gaussian prior with diagonal covariance matrix  $\mathbf{G}$  is placed on the standardized coefficients  $\alpha_k = \beta_k/\sigma$  rather than on the coefficients  $\beta_k$  directly. The diagonal of elements of  $\mathbf{G}$  are not fixed but follow inverse  $\chi^2$  distributions. Thus, the marginal prior on  $\alpha_k$  effectively follows a centered Cauchy distribution. The exact composition of the diagonal of  $\mathbf{G}$  depends on whether predictors are treated as random or fixed. For continuous predictors, the standardized coefficient vectors are assigned default regression priors which equally ensure invariance of Bayes factors with respect to linear transformations of the input.

### *Comparing logistic regression models*

A research question that arose in the context of this work concerned the comparison of the influence of two types of continuous variables on one binary dependent variable. Specifically, it should be assessed whether local brain activity or inter-regional brain connectivity more appropriately described different levels of stimulus intensity, expectations, and prediction errors. To this end, the evidence of models which predict different levels of a categorical dependent variable as a function of a continuous independent variable is computed and compared for different choices of the independent variable. Consider the case in which the dependent variable has two levels. A logistic regression model can be used to describe the probability of observing a certain value of the dependent variable  $Y \in \{0,1\}$  given the value of the continuous independent variable  $X \in \mathbb{R}$  [131], i.e.,

$$p(Y|X; \boldsymbol{\beta}) = \frac{1}{1 + \exp((2Y - 1)(\beta_1 + \beta_2 X))}.$$

Of particular interest here is the case in which the data originate from an experiment with repeated measures design comprising  $N$  individuals. Then, for each individual, both levels of the dependent variable are observed and the associated independent variables can be centered at 0. This implies that for each participant  $i$ , there are two data points  $(X_{i1}, Y_{i1})$  and  $(X_{i2}, Y_{i2})$  with  $X_{i2} = -X_{i1}$  and  $Y_{i1} = 1 - Y_{i2}$ . The fact that two data points of the same individual are not statistically independent has to be accounted for when constructing the likelihood function. Using a conditional logistic regression approach[132], the joint probability of the dependent variables  $Y_{i1}$  and  $Y_{i2}$  given  $X_{i1}$  and  $X_{i2}$  and conditioned on  $Y_{i1} + Y_{i2} = 1$  can be written as

$$\begin{aligned} & p(Y_{i1}, Y_{i2} | X_{i1}, X_{i2}, Y_{i1} + Y_{i2} = 1; \boldsymbol{\beta}) \\ &= \frac{p(Y_{i1} | X_{i1}; \boldsymbol{\beta})p(Y_{i2} | X_{i2}; \boldsymbol{\beta})}{p(Y_{i1} = 1 | X_{i1}; \boldsymbol{\beta})p(Y_{i2} = 0 | X_{i2}; \boldsymbol{\beta}) + p(Y_{i1} = 0 | X_{i1}; \boldsymbol{\beta})p(Y_{i2} = 1 | X_{i2}; \boldsymbol{\beta})} \\ &= \frac{\exp(\beta_2 X_{i1} Y_{i1}) \exp(\beta_2 X_{i2} Y_{i2})}{\exp(\beta_2 X_{i1}) + \exp(\beta_2 X_{i2})}, \end{aligned}$$

which is independent of  $\beta_1$ . The likelihood of the data given  $\beta_2$  then reads

$$p(D | \beta_2) = \prod_{i=1}^N p(Y_{i1}, Y_{i2} | X_{i1}, X_{i2}, Y_{i1} + Y_{i2} = 1; \beta_2)$$

To compute the model evidence associated with this likelihood, a prior for  $\beta_2$  has to be specified. The selection of a suitable prior is guided by the following considerations. First, note that the likelihood asymptotically approaches a value  $> 0$  as  $\beta_2$  tends to infinity. This prohibits the use of any improper priors. Note further that, as  $\beta_2$  approaches 0, the likelihood becomes independent of the predictor  $X$ . Consequently, for  $\beta_2 = 0$ , the Bayes factor arising from the comparison of two distinct predictors always equals one. Therefore, a prior concentrating too much probability mass in the vicinity of  $\beta_2 = 0$  obstructs the view on the effect of interest -akin to a Bayesian t-test in which the BF approaches one as the prior on the effect size  $\delta$  becomes narrower around  $\delta = 0$ . Lastly, it would be convenient if the prior were such that the model evidence is invariant w.r.t. linear transformations of the independent variable  $X$ . This latter property is, however, not essential as any data, prior to computing the model evidence, can be normalized to have a mean of zero and a standard deviation of one. In fact, this normalization is assumed from now on. The described theoretical considerations render, e.g., a zero-mean Gaussian with sufficiently

large standard deviation a suitable candidate for a prior on  $\beta_2$ . The evidence of model  $\mathcal{M}_X$  based on predictor  $X$  then reads

$$\mathcal{E}(\mathcal{M}_X) = \int p(D|\beta_2) \mathcal{N}(\beta_2; 0, \sigma) d\beta_2$$

with  $\sigma \gg 1$  being the standard deviation of the Gaussian prior. This model evidence can be computed for a range of different predictors. Let  $A$  and  $B$  denote two distinct sets of predictors. To assess which set of predictors more appropriately describes the data, the following evidence ratio is computed

$$BF_{A/B} = \frac{\frac{1}{|A|} \sum_{a \in A} \mathcal{E}(\mathcal{M}_a)}{\frac{1}{|B|} \sum_{b \in B} \mathcal{E}(\mathcal{M}_b)}$$

where  $|\cdot|$  denotes the number of elements contained in a set.

## 3 Project 1: Acute pain

### 3.1 Project-specific methods

#### 3.1.1 Participants

“The study was performed in healthy human participants at the university hospital of the Technical University of Munich (TUM). Written informed consent was obtained from all participants prior to the experiment. The Ethics Committee of the Medical Faculty of the TUM approved the study protocol. The study was preregistered at ClinicalTrials.gov (NCT04296968) and conducted in accordance with the latest version of the Declaration of Helsinki. It followed recent guidelines for the analysis and sharing of EEG data [34]. Inclusion criteria were right-handedness and age >18 years. Exclusion criteria were pregnancy, neurological or psychiatric diseases, and regular intake of medication (aside from contraception and thyroidal medication). Severe internal diseases (e.g. diabetes) and skin diseases (e.g. psoriasis, vitiligo), previous surgeries at the head or spine, current or recurrent pain, metal or electronic implants, and any previous side effects associated with thermal stimulation constituted additional exclusion criteria.”[3]

Using the G\*Power toolbox [133], we established that for the current rmANOVA design involving four measurements per individual, a power of 0.95 with an alpha of 0.05, and medium effect sizes of  $f = 0.25$  (corresponding to an  $\eta^2$  of 0.06 [134]) corresponded to  $n=36$  participants. “The original study recruited 58 healthy human participants (29 females, age:  $24.0 \pm 4.3$  y [mean  $\pm$  SD]). Ten participants were excluded due to either the absence of Pain or low pain ratings [ $<10$  on a numerical rating scale from 0 (no pain) to 100 (maximum tolerable pain)] during the familiarization run ( $n = 8$ ), excessive startle responses in response to painful stimulation during the training run ( $n = 1$ ), or technical issues with the response box used during catch trials ( $n = 1$ ). To ensure robust estimates of connectivity values, we here additionally excluded participants with less than 10 trials remaining after the raw data cleaning procedure described below ( $n = 8$ ). The final data set used here thus comprised 40 participants (all right-handed, 21 females, age:  $23.4 \pm 2.9$  y). Average anxiety and depression scores were below clinically relevant cutoff scores of 8/21 [135] on the Hospital Anxiety and Depression Scale [136] (anxiety:  $3.2 \pm 2.2$ ; depression:  $0.9 \pm 1.2$ ).”[3]

#### 3.1.2 Paradigm

“The objective of this analysis was to assess how sensory and contextual modulations are served by local brain activity and inter-regional brain connectivity. The experiment involved

two levels of noxious stimulus intensities (hi and li) and two types of visual cues (HE and LE) resulting in four experimental conditions. The visual cues probabilistically predicted the intensity of the subsequent noxious stimulus. The high expectation (HE) cue was followed by a hi stimulus in 75% of the trials and by a li stimulus in 25% of the trials. Vice versa, the low expectation (LE) cue was followed by a hi stimulus in 25% of the trials and by a li stimulus in 75% of the trials (Figure 3-1a). Figure 3-1b depicts the sequence of events for each trial. After a variable fixation period ranging from 1.5 to 3 s, a visual cue (either blue dot or yellow square) was displayed for 1 s. A brief painful heat stimulus was applied 1.5 s after cue offset. 3 s after the painful stimulus, participants were visually prompted to provide a verbal rating of the perceived pain intensity on a numerical rating scale ranging from 0 (no pain) to 100 (maximum tolerable Pain in the context of the experiment).” To check that participants continuously paid attention to the visual cues, in 10% of the trials (*catch trials*) they were visually prompted to indicate by a button press whether a HE or a LE cue had been presented last. “An average accuracy of  $95.6 \pm 0.1\%$  indicated that participants successfully focused on the task during the entire experiment. Trials were separated by a 3 s period during which a white fixation cross was presented. The experiment consisted of four runs with 40 trials each (hiHE [ $n = 15$ ], hiLE [ $n = 5$ ], liLE [ $n = 15$ ], liHE [ $n = 5$ ]), resulting in total trial numbers of hiHE [ $n = 60$ ], hiLE [ $n = 20$ ], liLE [ $n = 60$ ], liHE [ $n = 20$ ]. Runs were separated by short breaks of  $\sim 3$  min. Pairings of visual cues with stimulus intensities were balanced across participants. Prior to the experiment, the participants were familiarized with the stimulation and the intensity rating procedure by applying a sequence of 10 heat stimuli. Next, participants were informed about the pairing between cues and stimulus intensities and a training run comprising 16 trials was conducted. This was to ascertain that all participants were aware of the pairing and to minimize learning during the main experiment. During the experiment, participants sat in a comfortable chair. They wore protective goggles and listened to white noise on headphones to eliminate effects of ambient sounds. Please see [83] for additional details.”[3]

### Stimulation

“A laser pulse with a wavelength of 1,340 nm, a duration of 4 ms and spot diameter of approximately 7 mm was used to apply painful stimuli to the left hand [137]. For li and hi stimuli, the stimulus intensity was set to 3 and 3.5 J, respectively. These stimulus intensities are known to consistently elicit painful sensations of discriminable intensity [137]. The stimulation site was slightly changed after each stimulus to avoid tissue damage and habituation or sensitization.”[3]

### 3.1.3 Recordings and preprocessing

“Brain activity was recorded using actiCAP snap/ slim with 64 active sensors (EasyCap) placed according to the extended 10-20 system and BrainAmp MR plus amplifiers (Brain Products, Munich, Germany). During the recording, sensors were referenced to FCz and grounded at Fpz. The signals were sampled at 1,000 Hz (0.1- $\mu$ V resolution) and band-pass filtered between 0.016 and 250 Hz while impedances were kept below 20 k $\Omega$ . [] The BrainVision Analyzer software (version 2.1.1.327, Brain Products, Munich, Germany) was used for preprocessing. First, raw signals were low-pass filtered with a cutoff frequency of 225 Hz. After down-sampling to a rate of 500 Hz, a 1 Hz high-pass filter (fourth-order Butterworth) and a band-stop filter between 49 and 51 Hz filter removing line noise were applied. An independent component (IC) analysis based on the extended infomax algorithm was then conducted based on the -4.2 to 3.2 s peri-stimulus time windows of the EEG data. Subsequently, ICs representing artifacts originating from eye movements or muscles were removed from the unfiltered EEG data [138] using visual inspection. Moreover, data segments of 400 ms centered around data samples with amplitudes exceeding  $\pm 100 \mu$ V and data jumps exceeding 30  $\mu$ V were automatically marked for rejection. Finally, the data were inspected visually and remaining artifacts were manually marked for rejection. All signals were re-referenced to the average reference. The cleaned data were exported to Matlab (version R2019b, Mathworks, Natick, MA) and further analyses were performed using FieldTrip [version 20210411 [102]]. Data were segmented into epochs ranging from -4 to 3 s in peri-stimulus time and all trials with marked artifacts or pain ratings of zero were excluded. This resulted in  $49.5 \pm 8.5$ ,  $16.8 \pm 2.8$ ,  $18.0 \pm 1.6$ , and  $52.9 \pm 4.2$  trials per participant in the liLE, liHE, hiLE, and hiHE conditions, respectively. To assure that all analyses for the different trial types were eventually performed on the same number of trials, we matched the numbers of trials.”[3]

### 3.1.4 Evaluating brain measures

#### *Source model*

“To project sensor-level time series to source level, we employed Linearly Constrained Minimum Variance (LCMV) beamformers [139] implemented in FieldTrip [102]. Frequency-specific array-gain LCMV spatial filters for alpha, beta and gamma frequencies were constructed based on a lead field and a frequency-specific covariance matrix. A boundary element approximation of a realistically shaped, three-shell head model was used as the lead field. For each individual and frequency band, the covariance matrix was computed from the band-pass filtered, -1 s to 1 s (peri-stimulus time) concatenated data segments of all (non-rejected) trials. To ensure a robust computation of the inverse of the covariance

matrix we employed Tikhonov regularization as implemented in FieldTrip with a regularization parameter value of 5% of the average sensor power. The fixed orientation of the lead field for every source location was chosen to maximize the spatial filter output. Source-level signals were then obtained by applying the frequency-specific LCMV operator to the corresponding band-pass filtered sensor-level time series.” [3]

#### *Assessment of source-level evoked responses*

To evaluate source-level evoked responses, we first band-pass filtered the sensor-level signals in the 1 to 30 Hz frequency range. We then computed a corresponding frequency-specific spatial filter and projected the sensor level signals to the six source locations of interest. To maximize the signal-to-noise ratio for visualization, the results presented in Figure 3-3a represent the grand average of source-projected signals across all participants and hi trials.

#### *Assessment of source-level time-frequency representations*

“Source-level time-frequency representations were obtained using the following procedure: First, we projected the band-pass filtered sensor-level signals to source space using five frequency-specific LCMV spatial filters (i.e., for frequencies <8 Hz, 8-12 Hz, 13-30 Hz, 30-60 Hz, and 60-100 Hz). For each ROI, we generated TFRs as well as time-courses of alpha, beta, and gamma brain activity. The TFRs are based on Hanning-tapered data. Time courses of brain activity were computed based on moving time windows and using a Slepian multi-taper approach (see below). TFRs and time-courses of brain activity were computed from data segments with widths of 500 ms and 250 ms for frequencies below and above 30 Hz, respectively. Both TFRs and time-courses of brain activity are displayed as percentage change relative to a baseline period ranging from 0.75 to 0.25 s before the stimulus.”[3] To maximize the signal-to-noise ratio for visualization, the results presented in Figure 3-3 represent the grand average of source-TFRs across all participants and hi trials.

#### *Analysis of local brain activity*

“Local oscillatory brain activity was assessed as frequency-specific source power of the 6 ROIs. First, source level timeseries band-pass filtered to the frequency band of interest were obtained using the beamformer described above. For these signals, we computed the power of the frequency in the middle of the frequency band of interest using a Slepian multi-taper approach [140]. The spectral smoothing width was set to one half of the width of the frequency band of interest. In this way, the power value incorporates information of the entire frequency band of interest. We computed source power in the alpha (8-12 Hz), beta (14-30 Hz) and gamma (60-100 Hz) frequency bands for each trial. We then averaged



power values across trials for each condition and subject. To allow for the comparison of the effects on local brain activity to those on brain connectivity, the analysis was primarily performed on a 1 s post-stimulus interval. However, sensor-level findings indicate that the effects of painful stimuli on oscillatory brain activity are usually confined to shorter time windows.”[3] Control analyses based on these shorter time windows are provided in the supplement of the original publication [3].

#### *Control analysis of source-level evoked potential amplitudes*

Source-level evoked potential amplitudes were computed based on the above-described band-pass filtered (1 – 30 Hz) source-level signals. For each individual and ROI, we computed the signal amplitude by averaging the source-projected signal in a ROI-specific time window. We defined the center of each ROI-specific time window as the mean peak latency of intracranial signals measured at that ROI in a different study[2]. The standard deviation of peak latencies reported in that study ranged from 19 to 39 ms. We therefore set the width of the averaging window for all ROIs to 40 ms.

#### *Analysis of inter-regional brain connectivity*

“Connectivity analyses were performed on the 1 s post-stimulus intervals of the source level timeseries of the 6 ROIs. First, we computed the source level cross-spectral density of each participant using a multi-taper approach analogous to the one used for the computation of source power. To assess functional connectivity, we calculated the debiased weighted Phase Lag Index (dwPLI, [115]) based on all trials of each condition and for every subject. We selected the dwPLI measure due to its insensitivity to volume conduction effects. For the assessment of the direction of connectivity, we used an asymmetry score based on bivariate partial directed coherence (PDC, [124]). Specifically, for two ROIs A and B, the bivariate PDC analysis yields two values,  $PDC_{A \rightarrow B}$  and  $PDC_{B \rightarrow A}$ , representing the directed connectivity strength from A to B and from B to A, respectively. We cast these two values into a single asymmetry score,  $(PDC_{A \rightarrow B} - PDC_{B \rightarrow A}) / (PDC_{A \rightarrow B} + PDC_{B \rightarrow A})$ , ranging from -1 to 1. A large absolute value of the asymmetry score indicates a strong asymmetry of directed connectivity. The sign of the asymmetry score reveals the predominant direction of information flow. Direction of connectivity was calculated for connections that had shown intensity, expectation, and/or PE effects in previous analyses. For connections with evidence for an intensity or expectation effect in the Bayesian ANOVA, we included all trial types in the computation of the asymmetry score. For connections with evidence for an interaction effect, we included trials with a mismatch between cue and intensity only.”[3]

### 3.1.5 Statistical analysis

“For each of the four trial types (liLE, hiLE, liHE, hiHE), behavioral and EEG measures were computed based on an identical number of trials. This number was determined as the minimum number of available trials across the four trial types. [...] Building upon previous investigations [141, 142], we made specific predictions about how EEG responses signaling stimulus intensity, expectations, PEs, or combinations thereof are modulated across the four trial types. To formally test these predictions, we performed repeated measures ANOVAs (rmANOVAs) with the independent variables stimulus intensity and expectation. In these rmANOVAs, responses signaling stimulus intensity and expectations would manifest as main effects, whereas responses signaling PEs would manifest as interactions. This applies to definitions of PEs as absolute (unsigned) PE as well as to aversive PE. i.e. a PE occurs only if the stimulus is more painful than expected. To quantify effects and to facilitate interpretation of negative findings, we performed Bayesian rmANOVAs [125].”[3] More detailed explanations of Bayesian rmANOVAs and associated Bayes factors (BFs) can be found in section 2.2. “We considered a neural measure or pain rating as corresponding to the intensity or expectation pattern if there was at least moderate [(i.e.,  $BF > 3$ )] evidence for the corresponding main effect. Accordingly, we considered a neural measure or pain rating as corresponding to the prediction error pattern if the evidence for an interaction effect of intensity and expectation was at least moderate.”[3] To assess the asymmetry of information flow, we tested PDC-asymmetry scores against 0 using a nonparametric Bayesian t-test.

To directly link neural measures to pain scores, we analyzed difference scores of neural measures and pain ratings. These difference scores were computed for each experimental contrast, i.e., for li vs. hi, LE vs. HE, and IPE vs. hPE. In the case of, e.g., the stimulus intensity contrast, we computed the difference between averaged pain ratings in li (i.e., liLE and liHE) and hi conditions (i.e., hiLE and hiHE) for all individuals. Difference scores for the neural measures were computed analogously. We then computed Bayes factors for the correlations between neural measure difference scores and pain rating difference scores, thereby testing whether neural measures and pain ratings are linked on an inter-individual level. To assess whether there is a consistent association on an intra-individual level, we statistically compared the products of neural measure difference scores and pain rating difference scores against zero using a Bayesian one-sample t-test. The rationale for this approach is as follows: If, e.g., an increase in pain for a certain experimental contrast consistently elicits an increase of the neural measure, the product of the two difference scores will always have a positive sign (Figure 3-8). Therefore, the mean of difference score products across individuals will be significantly greater than zero.

Conversely, if the association between difference scores is inconsistent, meaning that an increase in pain entails both increases and decreases of the neural measure, then the mean of difference score products across individuals will not clearly deviate from zero. In essence, we consider there to be an association between pain ratings and a neural measure, if the distribution of the corresponding difference score products significantly deviates from zero.

Lastly, “all parametric Bayesian analyses were conducted using the BayesFactor package in R [130]; for non-parametric Bayesian t-tests we used freely available R code [143].” [3]

### 3.1.6 Model comparisons

“We intended to statistically assess whether an experimental contrast (intensity, expectation, or PE) is associated more strongly with local activity or inter-regional connectivity. To this end, we conducted a Bayesian comparison of power-based and connectivity-based models predicting the levels of intensity, expectation and PE.”[3] Specifically, we computed the Bayesian evidence of conditional logistic models mapping individual power and connectivity values to the probability of observing a certain level of intensity, expectation, or PE. “In the analysis, we consider  $N_{\text{pow}} = 6$  power values and  $N_{\text{conn}} = 15$  connectivity values in each of the  $N_{\text{freq}} = 3$  frequency bands. For each of the three types of experimental contrasts, this resulted in  $N_{\text{freq}} * N_{\text{pow}} = 18$  model evidence values for the power-based models and  $N_{\text{freq}} * N_{\text{conn}} = 45$  model evidence values for the connectivity-based models. The Bayes factor for, e.g., the intensity manipulation reported in the manuscript is the average of the 18 power-based model evidence values divided by the average of the 45 connectivity-based model evidence values. For the factor expectation and the interaction between expectation and intensity, i.e., PE, we proceeded analogously”[3] For methodological details of this model comparison, please refer to section 2.2.2.

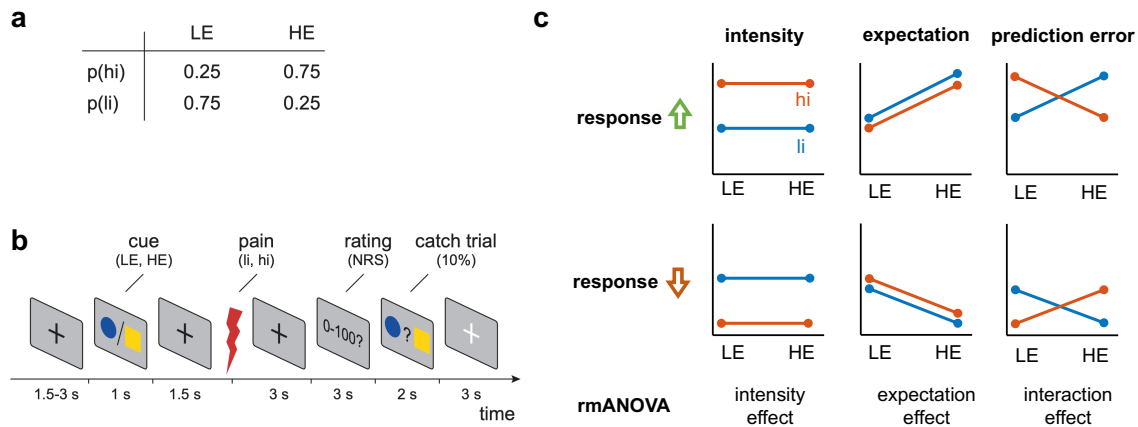
We complemented the above-described Bayesian rmANOVA with a Bayesian model competition approach. Therein, we compared various linear mixed models with regard to their adequacy to describe the data. The considered models differed in terms of the categorical predictor variables they employed. The predictor variables were intensity, expectation, absolute PE, and aversive PE. The participant ID was included as a random effect in all models. For each of these models, we computed a BF by comparing its evidence to that of a null model (random intercept only). The model with the largest BF (i.e., largest absolute model evidence) is considered the winning model. In addition, to

quantify how much better the winning model was in comparison to the other models, we reported the evidence ratio between the winning and the second-best model.

## 3.2 Results

To investigate how the brain serves sensory and contextual influences on pain, we analyzed data from an experiment which employed a probabilistic cueing paradigm[83]. Therein, brief painful heat stimuli were applied to the left hand while stimulus intensity and expectations were modulated independently in a  $2 \times 2$  factorial design. To modulate stimulus intensity, painful stimuli of two different levels (high intensity [hi] and low intensity [li]) were applied. “To modulate expectations, the painful stimuli were preceded by one out of two visual cues, probabilistically indicating the intensity of the subsequent stimulus. The high expectation (HE) cue was followed by a hi stimulus in 75% of the trials and by a li stimulus in 25% of the trials. Conversely, the low expectation (LE) cue was followed by a hi stimulus in 25% of the trials and by a li stimulus in 75% of the trials. The experiment thus comprised four trial types (Figure 3-1a): high intensity, high expectation (hiHE); high intensity, low expectation (hiLE); low intensity, high expectation (liHE); low intensity, low expectation, (liLE). In each trial, after the painful stimulus, the participants were asked to provide a rating of the perceived pain intensity on a numerical rating scale ranging from 0 (no pain) to 100 (maximum tolerable pain). Figure 3-1b shows the sequence of a single trial. The experiment included 160 trials per participant. “[3]

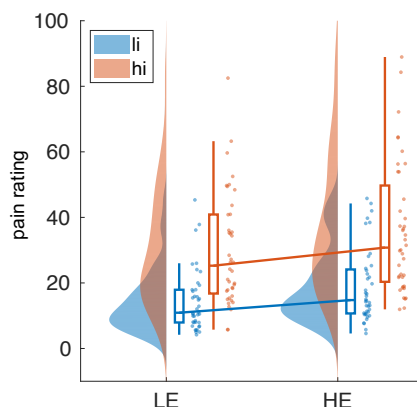
We analyzed local evoked responses, local oscillatory activity, and inter-regional functional connectivity in a network of six brain regions known to play key roles in the cerebral processing of pain [2]. “The brain regions were the contralateral primary somatosensory cortex (S1), the contra- and ipsilateral parietal operculum (cPO, iPO; including the secondary somatosensory cortex and parts of the insular cortex), the anterior cingulate cortex (ACC), and the contra- and ipsilateral prefrontal cortex (cPFC, iPFC). Some of these brain regions are particularly associated with processing of sensory information (S1, cPO, iPO) whereas others are more associated with supramodal cognitive and emotional processes (ACC, cPFC, iPFC) [7, 144]. Coordinates for these six regions of interest (ROIs) were taken from human intracranial recordings which represent the gold standard for electrophysiological brain responses to pain stimuli [2].” [3]. To assess local oscillatory activity and evoked responses, we calculated frequency-specific power and signal peak amplitudes in source space, respectively. “To assess functional connectivity between brain regions, we calculated the debiased weighted phase lag index (dwPLI) [115]. Both local [oscillatory] activity and inter-regional connectivity were assessed in the



**Figure 3-1:** Experimental design and possible response patterns. “(a) Probabilities of different pain stimulus intensities [low intensity (li) and high intensity (hi)] for different levels of expectation [low expectation (LE) and high expectation (HE)]. (b) In each trial, a cue was presented that probabilistically predicted the intensity of a subsequent painful stimulus. Three seconds after the stimulus, a verbal pain rating was obtained from the participants. In 10% of the trials (catch trials), participants were visually prompted to indicate by a button press whether a HE or a LE cue had been presented to ensure that participants continuously paid attention to the visual cues. More details on the experimental design can be found in [3].” [4] (c) “Possible response patterns indicating effects of stimulus intensity, expectations, and (absolute) PEs. Effects of stimulus intensity (low intensity, li; high intensity, hi), expectations (low expectation, LE; high expectation, HE), and prediction errors were tested by means of rmANOVAs. An experimental modulation can lead to either a relative increase (first row) or relative decrease (second row) of oscillatory activity or connectivity.” [4] This figure was adapted by permission from the American Association for the Advancement of Science: *Science Advances*, [4] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023

alpha (8-12 Hz), beta (14-30 Hz), and gamma (60-100 Hz) frequency bands. These frequency bands are known to exhibit changes in oscillatory power in response to brief painful stimuli [35, 38, 41, 43] and play key roles in inter-regional communication in the brain [99]. In addition, to assess the dominant direction of information flow in selected connections and frequency bands, we computed an asymmetry index based on the partial directed coherence measure [124] of directed functional connectivity.” [3]

“To relate neural measures to sensory and expectation effects on pain, we defined different patterns describing the relation between response variables and experimental manipulations [141, 142]. In particular, these patterns characterize how neural phenomena and pain ratings are linked to intensity, expectations, or discrepancies thereof (prediction errors, PEs) across the four trial types (Figure 3-1c). To formally link the data to these patterns, we performed repeated measures analyses of variance (rmANOVAs) [125] with the independent variables intensity and expectation. In these rmANOVAs, features signaling stimulus intensity and expectations would manifest as main effects, whereas features signaling PEs would manifest as interactions without distinguishing between absolute and aversive definitions of PEs. To allow for the interpretation of negative findings, we specifically performed Bayesian rmANOVAs [125].” [3]



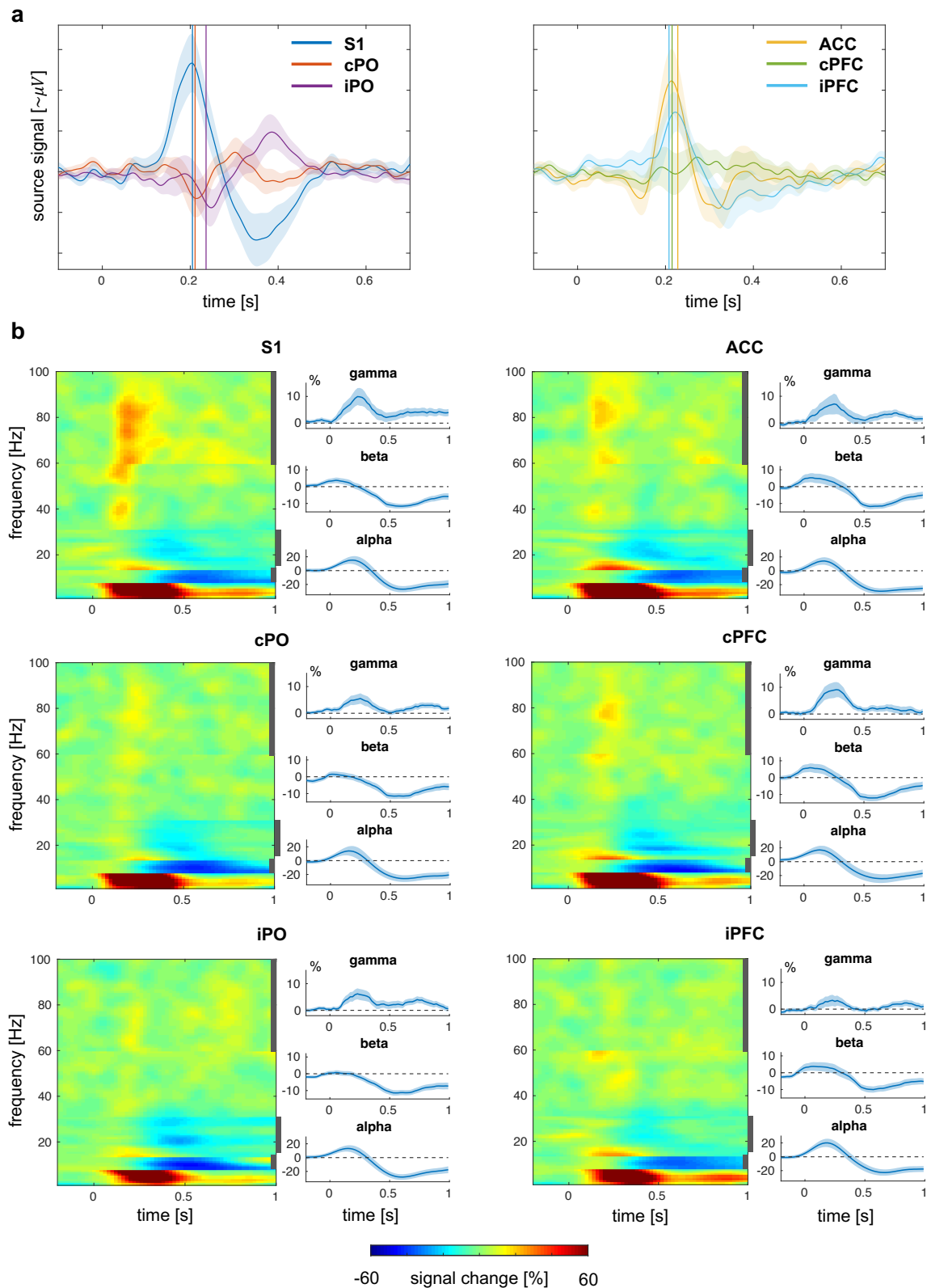
**Figure 3-2:** Pain ratings for different levels of stimulus intensity and expectations. “Rain cloud plot [1] of pain ratings for two levels of stimulus intensity (low intensity, li; high intensity, hi) and expectation (low expectation, LE; high expectation, HE). A Bayesian rmANOVA yielded decisive evidence for main effects of stimulus intensity and expectation ( $BF = 1.1 \cdot 10^{21}$  and  $BF = 5.5 \cdot 10^2$ , respectively). Moreover, there was moderate evidence against an interaction ( $BF = 0.27$ ).” [3] This figure was adapted by permission from the American Association for the Advancement of Science: Science Advances, [3] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023

### 3.2.1 Behavioral Findings

We first assessed how experimental manipulations of noxious stimulus intensity and expectations affected pain perception. Figure 3-2 depicts trial-averaged pain ratings for all participants in the four experimental conditions. A Bayesian repeated measures ANOVA “provided decisive evidence for main effects of intensity ( $BF = 1.1 \cdot 10^{21}$ ) and expectation ( $BF = 5.5 \cdot 10^2$ ) on pain ratings. Specifically, as expected, hi stimuli yielded higher pain ratings than li stimuli, and HE cues yielded higher pain ratings than LE cues. Moreover, there was moderate evidence against an interaction effect of intensity and expectation ( $BF = 0.27$ ). Thus, the results confirmed that stimulus intensity and expectations shaped pain ratings.” [3]

### 3.2.2 Source-level evoked potentials and local oscillatory brain activity

To examine the neural responses elicited by brief noxious stimuli, we considered source-level evoked potentials as well as local oscillatory brain activity in the six ROIs. Figure 3-3a shows source-projected time-domain signals in each ROI. The aperiodic deflections of these signals reflect source-level evoked potentials. The peak latencies of the signals observed here align with those observed in intracranial recordings [2]. In particular, the signal in S1 has the shortest peak latency, and the signal peaks in cPO precede those in iPO. Moreover, it is consistent with expectations that the largest peak amplitude occurs in S1. “Time-frequency representations (TFRs, Figure 3-3b) indicated that noxious stimuli suppressed alpha and beta activity in all ROIs and increased gamma activity



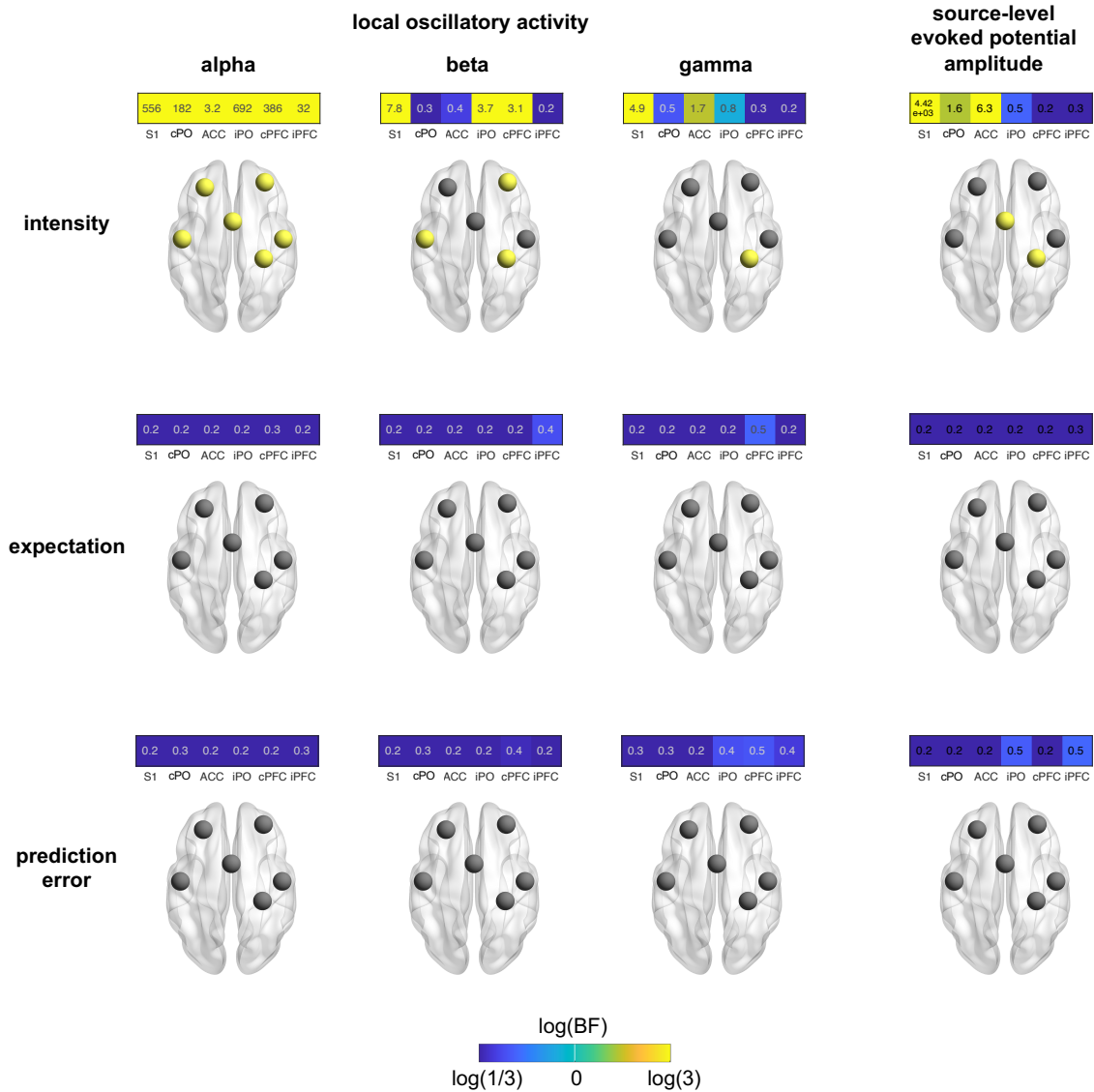
**Figure 3-3:** Noxious stimulus-related brain responses. (a) Source-reconstructed signals in the six ROIs. For better visual clarity, source signals are displayed in two separate diagrams. Vertical lines indicate peak latencies derived from intracranial recordings in a different study[2]. (b) “Time-frequency representations based on hi trials of local oscillatory brain activity in the six ROIs. The first and third columns show concatenated band specific TFRs for all six ROIs. The sharp transitions in the TFRs are due to the employment of frequency band-specific spatial filters. The second and fourth columns show time-courses of brain activity in the alpha, beta and gamma band. Vertical, dark-gray bars in the TFR plots indicate the frequency intervals based on which the time courses of brain activity were computed.” [3] Panel (b) was adapted by permission from the American Association for the Advancement of Science: Science Advances, [3] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023



predominantly in S1.” [3] Increases in signal power at frequencies below 8 Hz can be attributed to above-described evoked potentials. Taken together, the observed patterns of evoked and induced responses to brief noxious stimuli are in good accordance with the literature [2, 35, 36, 145] and, thus, support the validity of the employed source model.

Next, we assessed how local oscillatory activity in the six ROIs was shaped by stimulus intensity and expectations. “We therefore determined the power of brain activity in the six ROIs at alpha, beta, and gamma frequencies averaged across the 1 s post-stimulus interval.” [3] The results of Bayesian *rm*ANOVAs with the factors intensity and expectation are shown in the first three panel columns of Figure 3-4. [...] “We found that stimulus intensity modulated local brain activity at all frequency bands and in all ROIs. Strongest stimulus intensity effects were observed at alpha frequencies where we found moderate to decisive evidence for effects on oscillatory brain activity for all ROIs. In all ROIs, stronger stimuli yielded stronger suppressions of alpha activity [...]. Weaker effects were observed at beta frequencies where we found moderate evidence for an intensity effect on brain activity in S1, iPO, and cPFC. In these ROIs, stronger stimuli yielded stronger suppressions of beta activity. In the gamma frequency band, we observed moderate evidence for an intensity effect on S1 brain activity with stronger stimuli inducing higher amplitudes of gamma activity. In contrast, we found weak to moderate evidence against the effects of expectations or PEs on local brain activity at all frequency bands. [...] In summary, we found that stimulus intensity but not expectations or PEs influenced local oscillatory brain activity in response to brief painful stimuli.” [3]

Similar to measures of local signal power, source-level evoked potentials can be understood as reflecting local brain activity. We therefore performed a control analysis with source-level evoked potential amplitudes instead of local oscillatory activity as dependent variables. To this end, we first computed the peak amplitude at each ROI by averaging the source-projected signal over a time window that was 40 ms wide and centered at the peak latencies reported in [2]. We assessed the effects of stimulus intensity and expectations on peak amplitudes by means of a Bayesian *rm*ANOVA. The pattern of effects mirrored the pattern observed with local oscillatory activity, in that peak amplitudes were exclusively affected by stimulus intensity and not by expectations or PEs (Figure 3-4, last column).

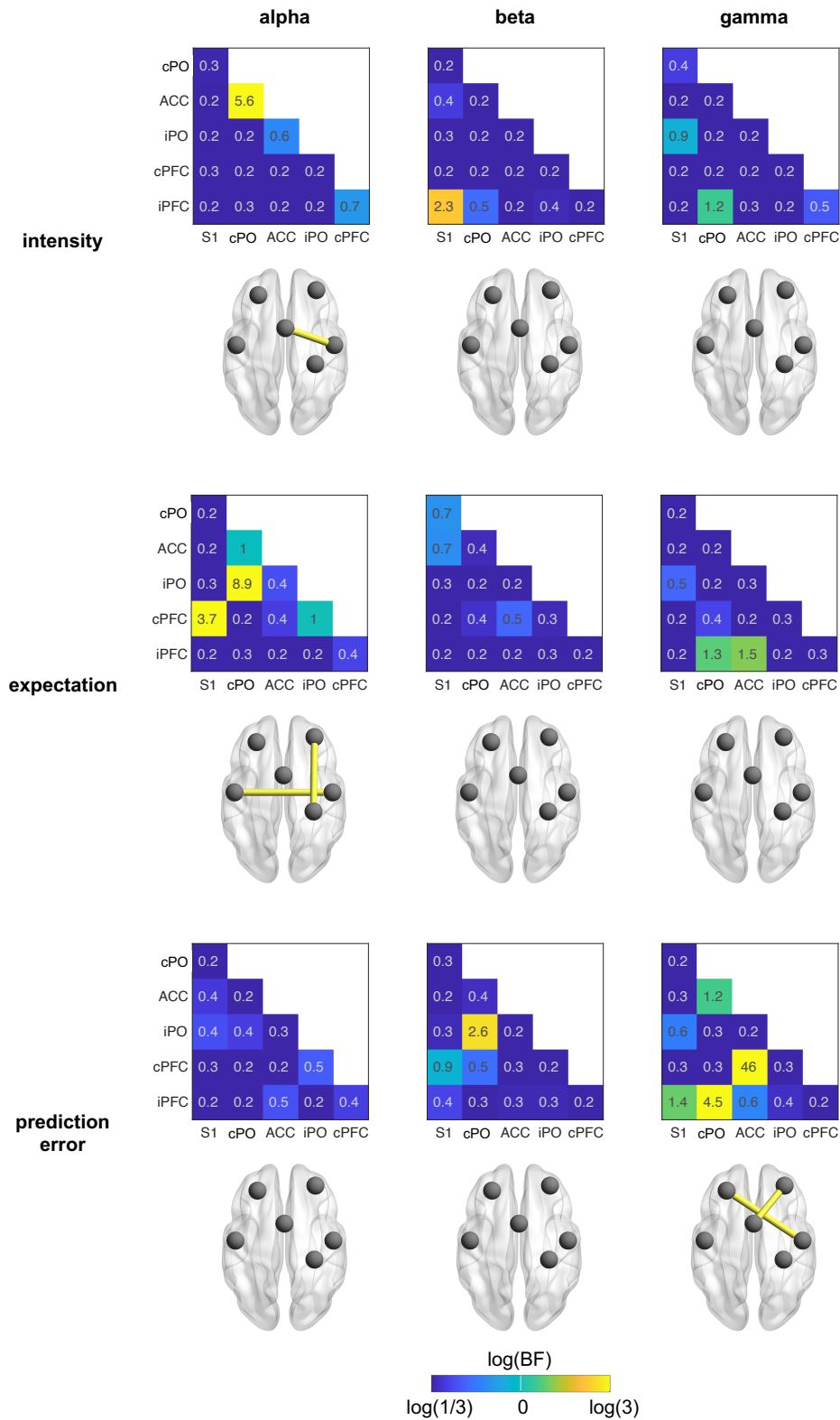


**Figure 3-4:** Effects of stimulus intensity, expectations, and prediction errors on local oscillatory brain activity and source-level evoked potentials. “Effects were assessed by Bayesian *rm*ANOVAs with the factors intensity and expectation. The color of the tiles representing ROIs scales with the log of the Bayes factor. It ranges from blue ( $BF < 1/3$ , at least moderate evidence against an effect) to yellow ( $BF > 3$ , at least moderate evidence for an effect). Brain images display ROIs in yellow which exhibit at least moderate evidence for an effect ( $BF > 3$ ).” [3]. (a) Local oscillatory brain activity. (b) source-level evoked potentials. This figure was adapted by permission from the American Association for the Advancement of Science: *Science Advances*, [3] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023

### 3.2.3 Inter-regional connectivity

#### *Undirected functional connectivity*

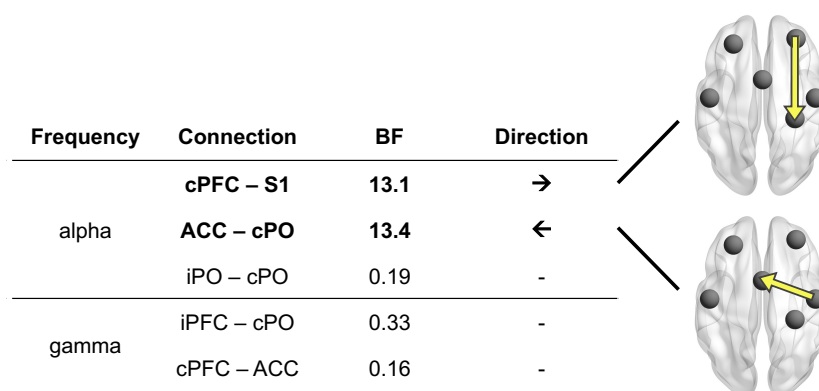
“We next investigated how stimulus intensity and expectations influenced communication in our core network associated with pain processing. We therefore determined pairwise inter-regional connectivity in a network of six ROIs resulting in 15 connectivity values. These analyses were performed separately for the alpha, beta, and gamma frequency bands in the 1 s post-stimulus interval. Figure 3-5 shows the results of Bayesian rmANOVAs. [...] We found moderate evidence for a stimulus intensity effect on the cPO – ACC connection in the alpha band. Here, connectivity was higher in the hi than the li condition. For most other connections and frequency bands, we found weak to moderate evidence against stimulus intensity effects. Effects of expectation were found in the alpha band exclusively. We specifically observed moderate evidence for an expectation effect on the cPFC – S1 and iPO – cPO connections. In these connections, connectivity was lower in the HE than the LE conditions. For most other connections and frequency bands, we found weak to moderate evidence against expectation effects. PE effects were observed in the gamma band exclusively. We found moderate to strong evidence for a PE effect on the cPFC – ACC and iPFC – PO connections. Specifically, the mean connectivity values of mismatch conditions (hiLE, liHE) were lower than those of non-mismatch conditions (liLE, hiHE). In other words, conditions involving a PE exhibited lower connectivity than those without a PE. For most other connections and frequencies, we observed weak to moderate evidence against a PE effect. Taken together, we found that stimulus intensity and expectation influenced connectivity at alpha frequencies whereas PE effects were found at gamma frequencies.” [3]



**Figure 3-5:** Effects of stimulus intensity, expectations, and prediction errors on inter-regional connectivity. “Effects were assessed by Bayesian rmANOVA with the factors intensity and expectation. The color of the heat map tiles scales with the log of the Bayes factor. It ranges from blue (BF < 1/3, at least moderate evidence against an effect) to yellow (BF > 3, at least moderate evidence for an effect). Brain images display connections in yellow which exhibit at least moderate evidence for an effect (BF > 3).” [3] This figure was adapted by permission from the American Association for the Advancement of Science: Science Advances, [3] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023

### Directed functional connectivity

“The previous analyses showed that stimulus intensity, expectations and PEs modulated functional connectivity at alpha and gamma frequencies in a core network associated with pain processing. We were next interested to assess the direction of information flow for connections in which we found at least moderate evidence for intensity, expectation, and/or PE effects. To this end, we calculated an asymmetry score of directed connectivity between pairs of brain regions. The score was based on the bivariate partial directed coherence (PDC, [124]) measure and ranged from -1 to 1. The absolute value and the sign of the score indicate the strength and the direction of asymmetry, respectively. For the cPO-ACC connection, for which intensity effects were observed in the alpha band, we found strong evidence (BF = 13.4) that information flowed from cPO to ACC. For the cPFC-S1 connection, for which expectation effects were observed in the alpha band, we found strong evidence (BF = 13.1) that information flowed from cPFC to S1. For the other connections and frequency bands, we did not find evidence for an asymmetry of information flow. Thus, as summarized in Figure 3-6, for connections showing intensity effects, we found information flow predominantly from sensory to higher-order brain areas. Conversely, for connections displaying expectation effects, we found information flow predominantly from higher-order to sensory brain areas.”[3]



**Figure 3-6:** Direction of information flow in selected network connections. “Using an asymmetry score based on the PDC connectivity metric, we assessed the direction of information flow in connections which exhibited evidence for an effect in the previous connectivity analysis. Brain images depict connections with strong evidence for asymmetric information flow. The arrows indicate the dominant direction of information flow.” [3] This figure was adapted by permission from the American Association for the Advancement of Science: Science Advances, [3] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023

### 3.2.4 Model comparisons

#### *Comparison of connectivity and activity models explaining the levels of stimulus intensity, expectation and PE.*

“The previous analyses indicated that local brain activity and inter-regional connectivity differentially serve sensory and expectation effects on Pain. We specifically observed that stimulus intensity shaped local brain activity more than inter-regional connectivity, while expectations and PEs shaped inter-regional connectivity more than local activity. To address this statistically, we conducted a Bayesian comparison of two types of models predicting the levels of stimulus intensity, expectation and PE. The two types of models differed with respect to the variables they incorporate for their predictions. One type of model incorporated activity variables, the other connectivity variables.”[3] We found decisive evidence that activity-type models predicted stimulus intensity better than connectivity-type models ( $BF_{\text{pow/conn}} > 100$ ). Conversely, there was moderate evidence that connectivity models predicted expectations ( $BF_{\text{conn/pow}} > 3$ ) and PEs ( $BF_{\text{conn/pow}} > 3$ ) better than activity models<sup>8</sup>.

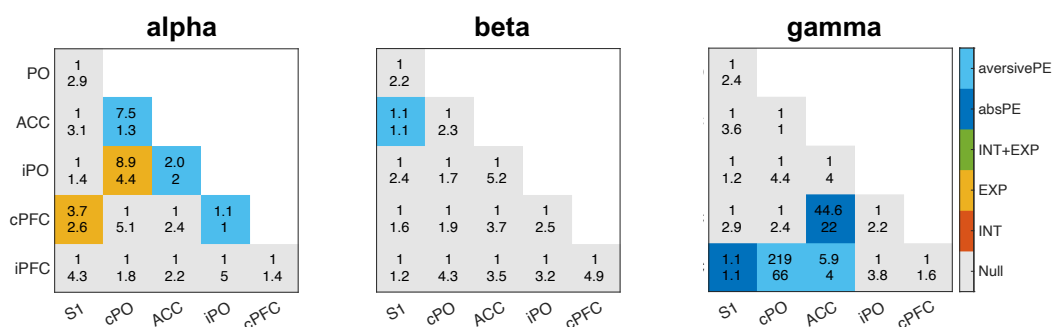
#### *Model competition and type of PE*

As reported above, our data showed that expectations and PEs are linked to connectivity at alpha and gamma frequencies, respectively. To test the robustness of these findings with respect to the choice of statistical method, we complemented the Bayesian rmANOVA of inter-regional connectivity with a Bayesian model competition approach. Moreover, this alternative approach allowed us to determine which type of PE (i.e., aversive or absolute) drove the effects in the gamma band.

The models participating in the Bayesian model competition differed in terms of the categorical predictor variables they employed. Overall, we considered five types of models: intensity, expectation, intensity + expectation, absolute PE, and aversive PE models. To clarify, an aversive PE occurs when a hi stimulus is applied following a LE cue. Absolute PEs arise when there is any mismatch between cue and stimulus intensity (i.e., for cases liHE and hiLE). Figure 3-7 visualizes the outcomes of the model competition. For

---

<sup>8</sup> Bayes factors stated here differ from those reported in the original manuscript. The new Bayes factors have been computed using conditional logistic regression models instead of standard logistic regression models to construct likelihood functions. The reason for this update is that the data are paired, i.e., there are two data points per participant (corresponding to the low and high levels of stimulus intensity, expectations, or prediction errors). Therefore, these data points are statistically independent across but not within participants. This dependency has now been taken into account by using conditional logistic regression models. While the revised models lead to quantitative changes in Bayes factors, the qualitative implications remain unchanged.



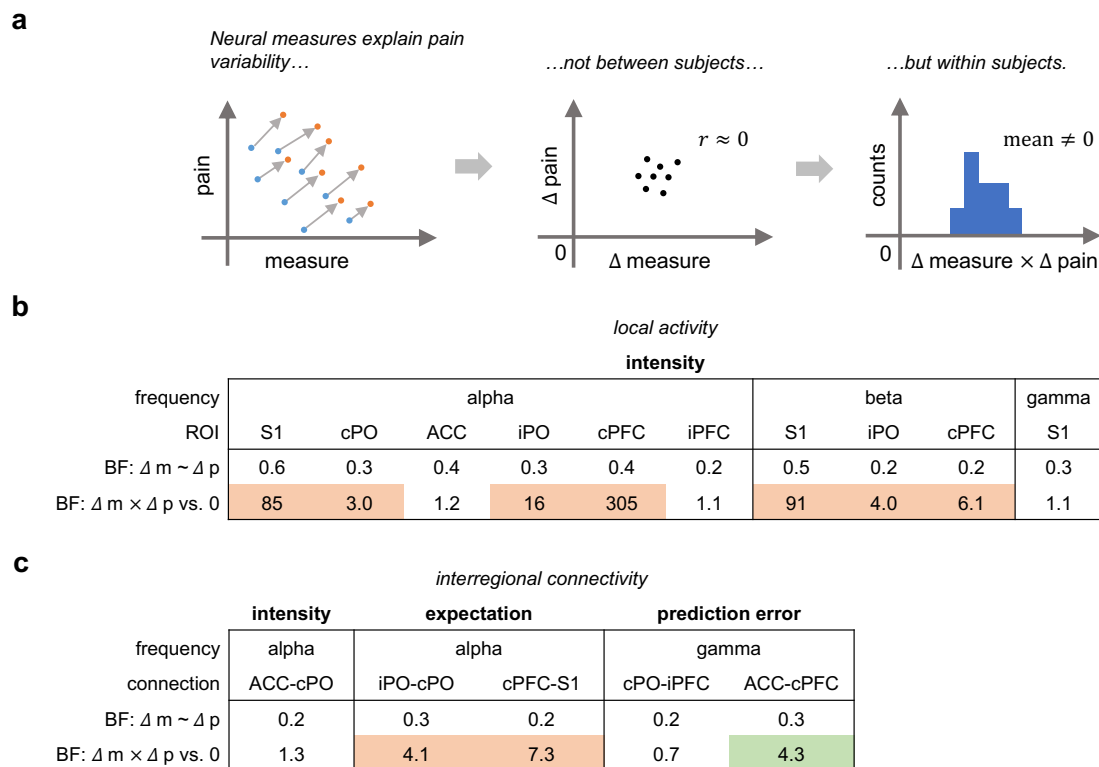
**Figure 3-7:** Competition of models predicting connectivity. The color of the heatmap tiles indicates the model with the largest model evidence, i.e., the winning model. The top number in each tile represents the Bayes factor comparing the winning model to the null model (individual-specific intercept only). The bottom number represents the Bayes factor comparing the winning to the second-best model.

the S1-cPFC and the iPO-cPO connections in the alpha band, the expectation model most adequately described the data. For the ACC-cPO connection the aversive PE model was the best, albeit only marginally better than the expectation model. PE-based models were the best choice to describe connectivity in the gamma band. In particular, the aversive PE model best described connectivity in the cPO-iPFC and ACC-iPFC connections, while the absolute PE model was most adequate to describe connectivity in the ACC-cPFC connection. Taken together, these results are consistent with findings from Bayesian rmANOVAs and offer additional insights into the specific types of PEs that drive the observed PE.

### 3.2.5 Direct association between neural measures and pain

So far, the effects of experimental contrasts on pain scores and neural measures have been investigated independently. To directly link local oscillatory activity and inter-regional connectivity to pain scores, we employed two approaches.

First, we assessed whether individuals showing larger changes in neural measures also exhibited larger changes in pain ratings. To quantify this association for, e.g., the stimulus intensity contrast, we computed the difference between averaged values in li (i.e., liLE and liHE) and hi conditions (i.e., hiLE and hiHE) for pain ratings and neural measures. Subsequently, we computed Bayes factors for the correlations between these difference scores. We restricted these analyses to experimental contrasts and neural measures that had previously yielded evidence for an effect. The results of these analyses are summarized in Figure 3-8. For the intensity contrast, there was no evidence for a relationship between difference scores of neural measures and pain ratings. Likewise, neither expectation nor PE contrasts showed evidence for such relationships in any of the



**Figure 3-8:** Direct association between pain and neural measures. (a) Schematic illustration of the putative data pattern. Blue and orange dots represent data points obtained under two experimental conditions (e.g., li and hi). Gray arrows connect data points of the same individual. Within individuals, an experimental contrast is associated with consistent changes in both pain and neural measures (left graph). At the same time, changes in pain are not associated with changes in the neural measure on the inter-individual level (middle graph). The consistency of the intra-individual association is assessed statistically by testing whether the product of difference scores of neural measures and pain deviates from zero (right graph). (b) Associations between local brain activity and pain. (c) Associations between inter-regional connectivity and pain.

considered ROIs and connections. These results do not imply that pain and the considered neural measures are not directly associated. Rather, the variance of pain rating difference scores across individuals might have been too small to detect such an association. Figure 3-8a illustrates this scenario.

Based on these considerations, we devised a second test to evaluate the relationship between neural measures and pain ratings. Instead of examining the correlation between difference scores, we now tested whether the product of these scores significantly deviated from zero. The rationale behind this approach is illustrated in Figure 3-8 and described in more detail in the methods section. In essence, we consider there to be an association between pain ratings and a neural measure, if the distribution of the corresponding difference score products significantly deviates from zero.

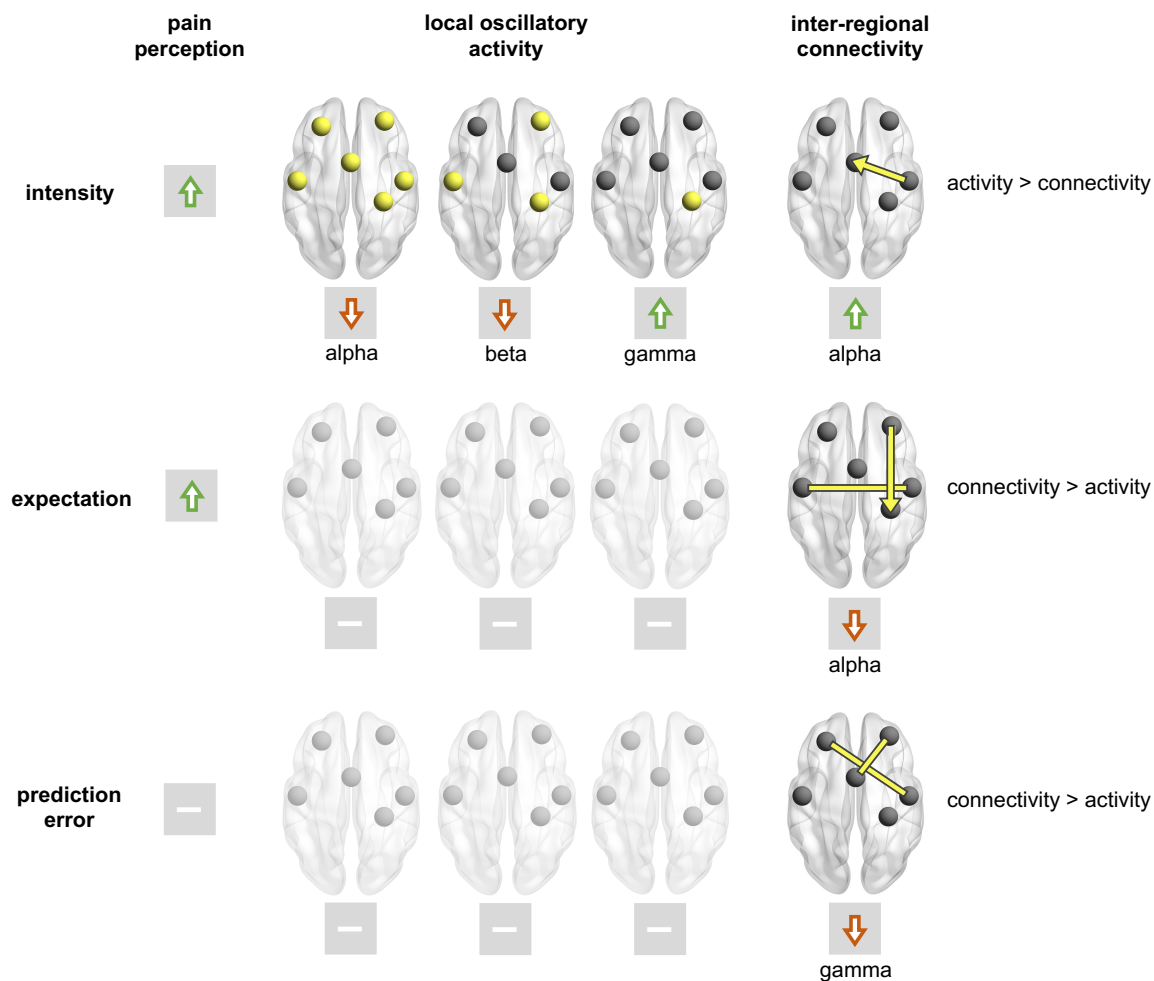
Figure 3-8b and Figure 3-8c depict the results of this analysis. We observed that variations in local oscillatory brain activity and pain ratings, induced by changes in stimulus intensity, are directly linked in most considered ROIs and frequency bands. Moreover,



variations of inter-regional connectivity and pain ratings, driven by changes in expectations, are directly linked in all considered connections at alpha frequencies. For the PE contrast, evidence was found for an association between variations of gamma-band connectivity in the ACC-cPFC connection and pain ratings. For the remaining contrasts and connections/ROIs, a direct association between variations of connectivity and pain was not observed. Overall, these results directly show the differential involvement of distinct neural measures in diverse modulations of the pain experience on a within-subject level.

### 3.2.6 Summary

The main findings are summarized in Figure 3-9. “On the behavioral level, both stimulus intensity and expectation modulated the perception of pain. As expected, both higher stimulus intensities and expectations of stronger stimuli evoked higher pain ratings. In the brain, stimulus intensity effects were predominantly associated with changes of local brain activity. Stronger stimuli yielded stronger responses to brief painful stimuli in alpha, beta, and gamma frequency bands. In contrast, expectation effects on pain were associated with changes of inter-regional functional connectivity but not with changes of local brain activity. We particularly found that expectation effects were associated with top-down connectivity at alpha frequencies from cPFC to S1 and with connectivity between cPO and iPO. PEs were associated with changes of gamma-band connectivity exclusively. Bayesian model comparisons confirmed the differential involvement of local activity and inter-regional connectivity in sensory and expectation effects on pain. Specifically, stimulus intensity has a stronger influence on local brain activity than on inter-regional connectivity. Vice versa, expectations and PEs shape inter-regional connectivity more than local brain activity.” [3]. Frequentist analogs of selected Bayesian statistical tests discussed in this section have also been performed. The results of these frequentist tests are qualitatively similar to those of the Bayesian tests and can be found in the supplement of the original publication[3].



**Figure 3-9:** Synopsis of project 1. “Increases of stimulus intensity led to increases of pain ratings and local brain activity at gamma frequencies as well as to decreases of brain activity at alpha and beta frequencies. Expectations of stronger pain yielded increases of pain ratings and reduced connectivity between cPO and iPO and from cPFC to S1 at alpha frequencies. In contrast, expectations did not modulate local brain activity at any ROI and any frequency band. PEs did not change pain ratings or local brain activity but iPFC-cPO and cPFC-ACC connectivity at gamma frequencies. The last column shows the results of a Bayesian comparison of local brain activity and connectivity models predicting intensity, expectation and prediction errors.” [3] This figure was adapted by permission from the American Association for the Advancement of Science: Science Advances, [3] (Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain, Bott et al.), © 2023

## 4 Project 2: Chronic pain

### 4.1 Project-specific methods

Using fMRI, it was discovered that the brain activity of individuals who are not engaged in a particular task, i.e., resting-state brain activity, does not fluctuate arbitrarily, but rather exhibits distinct spatial patterns. In particular, connectivity profiles of individual brain regions were found to be similar within and distinct between well-defined spatial brain structures. These brain structures are referred to as intrinsic brain networks[50, 86]. In primary analyses, we set out to assess the association between intrinsic brain networks and chronic pain. In secondary and exploratory analyses, we broadened the focus and assessed how connectivity in general relates to chronic pain, depression, and age.

#### 4.1.1 Data sets

We based our analyses on resting-state EEG recordings in patients with chronic pain. To this end, we used EEG recordings from our research group and identified and acquired external EEG data sets. We used the external data sets primarily to test the robustness of our findings. While some of the data sets at our disposal offer resting-state recordings under both eyes closed and eyes open conditions, we restrict our analyses to eyes closed data as these were shown to give rise to more robust results[146]. Given the array of available data, it is convenient to assign unique identifiers to individual data sets. These identifiers are printed in italic font and carry the prefix “*Set\_*”. First, we introduce two data sets that were generated in our research group: *Set\_PainlabDiscovery* and *Set\_Zebhauser2023*. Then, five data sets from external research groups are presented together with the strategy we pursued to obtain them: *Set\_Day2020*, *Set\_Adhia2022*, *Set\_Wager2022*, *Set\_Jarnitsky2022*, and *Set\_Jensen2020*. A summary of all available data sets is provided in Table 4-1.

To counteract any sample size biases, in our analyses, we base computations of neural features in all individuals on a fixed number of two-second-long EEG signal epochs. In our data, to ensure the inclusion of at least 80% of all recordings, this fixed number must not exceed 192. This means that for 20% of individuals there are fewer than 192 epochs available. These individuals are excluded from the analysis.

#### *Internal data sets*

*Set\_PainlabDiscovery* is composed of five data sets that have previously been recorded in our research group to investigate brain dysfunction in patients with chronic pain. These data sets are *Set\_Tiemann2012*[147], *Set\_May2019*[46], *Set\_TaDinh2019*[80], and

*Set\_Heitmann2022*[28]. The first three data sets have been used in combination to compare measures of brain activity, brain connectivity[80], and brain dynamics[148](i.e., microstate analyses[149]) cross-sectionally between patients with chronic pain and healthy controls. *Set\_Heitmann2022* has been used to assess measures of brain activity and brain connectivity longitudinally in a cohort of patients with chronic pain[28].

In all studies of *Set\_PainlabDiscovery*, inclusion criteria for patients consisted in a clinical diagnosis of chronic pain, with pain persisting for at least six months and with an average pain intensity during the four weeks prior to assessment of at least four (two in the case of *Set\_Heitmann2022*) on an 11-point numerical rating scale (NRS) ranging from zero (no pain) to ten (worst imaginable pain). Patients with severe diseases other than chronic pain or those on regular benzodiazepine medication were excluded. In total, *Set\_PainlabDiscovery* comprises data from 148 patients. Nine patients had to be excluded due to not meeting the minimum epoch number requirement. The analyzed cohort consists of 74 patients with chronic back pain (CBP), 33 patients with chronic widespread pain (CWP), 13 patients with primarily neuropathic pain (NP), and 19 patients with pain of other etiologies (OTHER). All data sets were recorded using a passive electrode EEG system with 64 channels (Easycap, Herrsching, Germany) and BrainAmp MR plus amplifier (Brain Products, Munich, Germany). Previous analyses of components of *Set\_PainlabDiscovery* did not yield any evidence which could have informed or biased the analyses of the present study.

*Set\_Zebhauser2023* has not yet been published or analyzed, as data acquisition is currently in progress. The primary goal of this project is to assess pain medication effects on EEG-based measures of brain activity and brain connectivity. Inclusion and exclusion criteria are similar to those described for *Set\_PainlabDiscovery*. At the time of writing this thesis, *Set\_Zebhauser2023* comprised 88 patients, 38 of which were excluded due to insufficient epoch numbers after preprocessing. Among the remaining patients, 21 suffered from CBP, two suffered from CWP, 9 exhibited NP, and 18 could be assigned to the OTHER category. To record EEG, a novel 32-channel system with active dry electrodes (CGX-Quick32r, CGX-systems, San Diego, US) is being used.

### *External data sets*

As detailed in the study preregistration (<https://osf.io/4qmyw/>), we contacted research groups across the globe in a structured data acquisition campaign to inquire about resting-state EEG recordings in patients with chronic pain. To this end, we first created a list of 79 studies which involved resting state EEG recordings. This list was mostly based on a systematic review on EEG and MEG biomarkers of chronic pain[150]. Moreover, the list

comprised studies which had been published in the year 2022. To decide which external research groups to contact, we condensed this list using three criteria:

1. Number of EEG-sensors  $\geq 32$
2. Publication date  $\geq 2013$
3. Number of patients  $\geq 20$

Applying these criteria shortened the list to 18 studies, some of which shared a single data set. We then contacted the corresponding authors of the respective studies to inquire whether they would be willing to share their data. To those who did not reply within a few weeks, we sent a standardized reminder email. To date, we received positive replies from four groups and three of them have already sent us their data; one is work in progress. We additionally included an external data set which comprised less than 32 channels (*Set\_Wager2022*). We included this data set because it was sent to us before we engaged in our systematic inquiries. Further, this data set could be used to test the replicability of findings in EEG setups with fewer channels.

*Set\_Day2020* had been analyzed with respect to local and global sensor-level power changes associated with several non-pharmacological, 8-week interventions in 69 patients with CBP. To be eligible for the study, patients had to exhibit pain in the lower back area for more than three months and they had to rate their average pain in the four-week period prior to assessment with at least a four on an 11-point NRS. Participants with severe comorbidities were excluded from the study. After preprocessing, data from 60 patients recorded prior to interventions could be included in the present study. EEG recordings were obtained using an ANT Neuro EEGO sports system (Medical Imaging Solutions GmbH, Berlin, Germany) with 64 active scalp electrodes (Waveguard cap). In the original study, an analysis of variance revealed that the interventions were linked to both reductions in pain severity scores and reductions in relative power at theta, alpha, and beta frequencies[151].

*Set\_Adhia2022* was recorded in the context of a study which investigated the efficacy of infra-slow neurofeedback training as a treatment for chronic low back pain in 60 patients[152]. Eligibility criteria were analogous to those stated for *Set\_Day2020*. Here, we incorporate baseline data from 57 patients. EEG recordings were obtained using a 64-electrode system with SynAmps-RT amplifier (Compumedics-Neuroscan). In the original study, no analyses relevant to this project were reported.

*Set\_Wager2022* is a data set for which, to date, no analyses have been published. It was recorded in the context of a larger study on the efficacy of the pain reprocessing therapy for the treatment of chronic back pain[153]. Eligibility criteria were very similar to

those described in the two previously presented data sets. *Set\_Wager2022* comprises 19-channel EEG recordings (Evoke system) from 69 patients with CBP. Only one patient had to be excluded due to not meeting the minimum epoch number requirement.

*Set\_Yarnitsky2022* comprises recordings from 133 patients with painful and 47 patients with non-painful diabetic polyneuropathy. In the original study, these data were used to train a machine learning model to distinguish patients with painful diabetic polyneuropathy from those with non-painful diabetic polyneuropathy[154]. This study was part of the larger DOLORisk[155] initiative aiming at identifying risk factors for the development and maintenance of neuropathic pain. Inclusion criteria defined by this initiative were, e.g., a diagnosis of Type 1 or Type 2 diabetes and a clinical diagnosis of peripheral neuropathy or symptoms highly suggestive thereof. For EEG recordings, a 64-channel system with active electrodes was used (ActiCHamp, Brain Products, Munich, Germany). While the trained machine learning model was reported to yield high levels of accuracy, several methodological aspects regarding its validation remained intransparent. In the present study, we included data from 47 patients with painful diabetic polyneuropathy.

*Set\_Jensen2021[156]* is a potentially very informative data set for two reasons. First, it comprises data from as many as 147 patients with different forms of chronic pain. Second, having been recorded with a 128 channel EEG system (GES high-density EEG acquisition system (EGI, Eugene OR), these data promise to have an excellent spatial resolution. However, as the data transfer is still ongoing, *Set\_Jensen2021* could not be incorporated in the analyses discussed in this thesis.

#### *Data curation*

In summary, we currently have two internal and five external resting-state EEG data sets at our disposal. The most important properties of these data sets are summarized in Table 4-1. In addition to EEG recordings, the complete data also comprise meta data tables containing demographic and behavioral information about patients. The meta data variables we used in the present work were

- **participantID:** Unique participant identifier
- **age:** Age of the participant at the time of the recording
- **diagnosis:** Identifier for category of pain diagnosis: Chronic back pain (CBP), chronic widespread pain (CWP), neuropathic pain (NP), postherpetic neuralgia (PHN), polyneuropathy (PNP), Other (OTHER)
- **avgPain:** Average pain over a period of one day to four weeks prior to assessment, rated on an 11-point numerical rating scale ranging from no pain (0)

to worst imaginable pain (10). The time periods to which the ratings refer in the individual studies can be inferred from Appendix Figure 0-2.

- **deprScore:** Depression score. The quantification of depression severity was particularly heterogeneous across studies. Employed metrics included, among others, Beck Depression Inventory and PROMIS questionnaires. Which instrument was used by which study is indicated in Appendix Figure 0-2.

Other meta data variables which were extracted from the studies but not used in the present work are gender (male, female, other), currPain (current pain at time of recording), and neuropathicPain (questionnaire score reflecting the neuropathic pain component). As the data originate from diverse research groups, they are heterogeneous and measures had to be taken to counteract any biases due to systematic differences in questionnaires and/or EEG recording systems. To this end, we computed z-scores within studies for all continuous meta data variables that were extracted, i.e., for age, avgPain and deprScore. Likewise, we computed z-scores within studies for all evaluated brain measures. Moreover, as several brain measures have been linked to age[88-90], we regress out age from both independent and dependent variables in all analyses except those considering age as the dependent variable. As explained in more detail in section 4.1.4, in multivariate analyses, z-scoring and confounder removal were done in a manner to prevent leakage of information from test or validation to training sets in the employed cross-validation procedures.

**Table 4-1:** Overview of data sets employed in project 2. We conducted analyses separately for patients with chronic pain and patients with diverse pain conditions. For each data set, the “CPB” and “diverse pain” columns contain the patient numbers in chronic pain and diverse pain cohorts, respectively. In parentheses is the number of individuals remaining after requiring that at least 192 epochs be available per individual.

		participants (participants with $\geq 192$ epochs)			
	data set identifier	CPB	diverse pain	number of sensors	
Set_PainlabDiscovery	Set_Tiemann2012*	0	20 (20)	64	internal
	Set_May2019	34 (30)	34 (30)		
	Set_TaDinh2019	13 (12)	46 (42)		
	Set_Heitmann2022	33 (32)	48 (47)		
replication sets	Set_Zebhauser2023	30 (21)	88 (61)	32	external
	Set_Day2020	69 (60)	69 (60)	64	
	Set_Adhia2022	60 (57)	60 (57)	64	
	Set_Wager2022	69 (68)	69 (68)	19	
	Set_Yarnitsky2022	0 (0)	133 (52)	64	
	Set_Jensen2021	?	147	128	
		<b>308 (&gt;280)</b>	<b>714 (&gt;437)</b>		

\* For this data set depression scores and age but not average pain ratings are available.

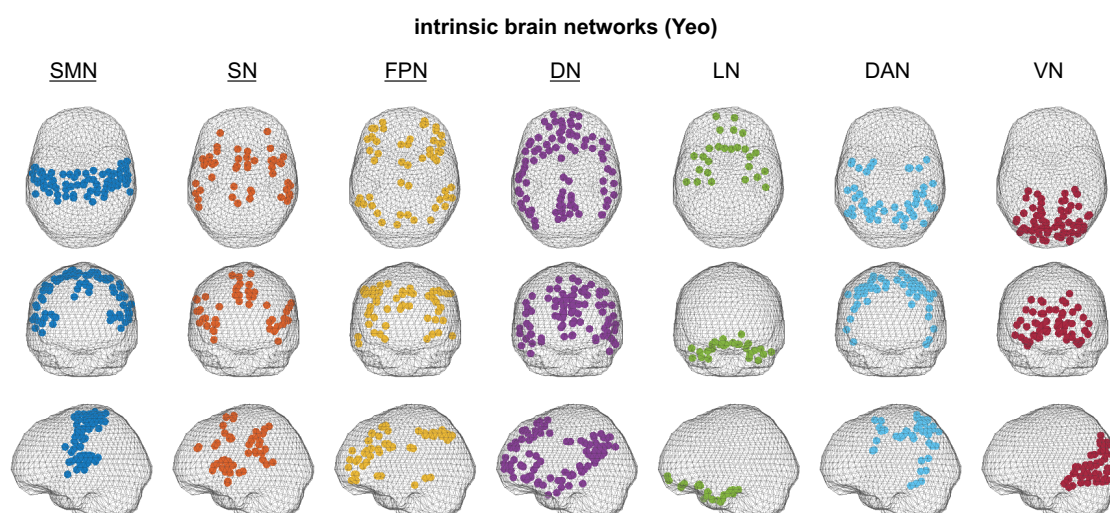
### 4.1.2 Automatic preprocessing

The data we received from external research groups was raw data, i.e., it had not been preprocessed in any way. This was important as we wanted to ensure that preprocessing was done in a unified manner across all internal and external data sets. To preprocess data, we employed an automatic preprocessing pipeline which has initially been proposed by [157] and was adapted in [90] for the use with resting state recordings. This pipeline represents a concatenation of several established functions from the Matlab-based EEGLAB toolbox. It comprises the following steps: The data, which has to be in BIDS format[158], is imported to Matlab and down-sampled to 250 Hz. Then, an algorithm which estimates and subtracts sinusoidal components akin to power line artifacts is applied. Based on a set of exclusion criteria, individual channels are removed from the data prior to re-referencing all channels to the average reference. To further reduce the influence of components not related to actual brain activity, the data are decomposed using independent component analysis (ICA). Independent components (ICs) which have been identified by a machine learning based classifier[159] as corresponding to muscle or eye activity are removed from the data. Compared to its manual counterpart, this procedure generally preserves more components and is, therefore, less invasive. After the removal of artifactual ICs, the previously rejected channels are interpolated. Subsequently, using the artifact subspace reconstruction[160, 161] (ASR) method, time segments which are contaminated by large amplitude artifacts are identified. ASR first decomposes the data of different time segments into their principal components. Time segments are then marked as artifactual if the variance of a principal component exceeds a certain threshold. This threshold, in turn, is defined based on reference data segments which are also automatically identified. Finally, the data are segmented into 2 s epochs with 1 s overlaps. Epochs comprising time segments marked as artifactual are rejected.

### 4.1.3 Evaluating brain measures

In primary analyses, we assess connectivity between and activity within four of the seven Yeo networks (Figure 4-1). Specifically, we include the somatomotor (SMN), the frontoparietal (FPN), the salience ventral attention (SN) and the default network (DN). In exploratory analyses, we explore brain activity and connectivity for spatial configurations other than the four Yeo networks. In all analyses, brain connectivity and brain activity are assessed in theta (4 – 8 Hz), alpha (8 – <13 Hz), and beta (13 – 30 Hz) frequency bands. In the following, the methods for evaluating measures of connectivity and activity in intrinsic brain networks using EEG are briefly described. More detailed explanations of the





**Figure 4-1:** Visualization of the seven Yeo networks. Axial, coronal, and sagittal views of the somatomotor (SMN), the salience ventral attention (SN), the frontoparietal (FPN), the default (DN), the limbic (LN), the dorsal attention (DAN), and the visual network (VN). In primary and secondary analyses, we focus on only four out of these seven networks (indicated by underlined network abbreviation).

methodological aspects of the evaluation of intrinsic brain network connectivity are provided in sections 2.1.3 and 2.1.4.

#### *Source reconstruction*

We project the preprocessed sensor-level data to source space using frequency-band specific linearly constrained minimum variance (LCMV) spatial filters[110]. These spatial filters are constructed as described for project 1 in section 3.1.4, except that sensor-level data covariance matrices are now computed based on epochs derived from continuous recordings rather than from different trials. In primary and secondary analyses, spatial filters were computed for source locations corresponding to the centroids of 400 brain parcels defined by the Schaefer atlas. Exploratory analyses additionally make use of a variant of the Schaefer atlas comprising 100 parcels. Subsequent steps are based on the source-level time series.

#### *Representative time series*

Computing connectivity between intrinsic brain networks using EEG is challenging due to the complex and intertwined geometries of the individual networks. A widely used method for quantifying connectivity between two smaller and simply-shaped geometric entities in the brain, such as parcels or regions, is to calculate the mean of the corresponding entries of a high-resolution connectivity matrix[81, 82]. However, when applied to entire intrinsic brain networks, this approach yields connectivity values that are highly correlated across individuals (pilot assessments yielded average  $r$ -values  $> 0.97$ ). Such spatially indifferent connectivity values are unlikely to adequately capture the actual

underlying network characteristics. The reason for this lack of spatial specificity likely lies in the limited spatial resolution of EEG. At this limited spatial resolution, reconstructed signals at sources in more slender regions of a network are heavily contaminated by contributions originating from surrounding sources, including those belonging to other networks. In other words, many signals associated with a particular network do not appropriately represent that network's activity. Therefore, an alternative way to compute inter-network connectivity is to first identify time series that are believed to be sufficiently representative of the respective networks. The connectivity between networks is then determined by evaluating the connectivity between their representative time series[162].

For simpler geometric entities such as brain regions or parcels, a common approach is to use the first principal component (PC) of the associated source-level signals as the representative time series. This implies the assumption that the contamination from other networks is less reflected in the first than in subsequent PCs. However, for intrinsic brain networks, which have more complex geometries, the first PCs are likely still significantly affected by contributions from outside the respective networks of interest. In this work, we therefore developed a technique that does not remove but aims to mitigate the issue of low spatial resolution. Therein, a network's representative time series is similar to its first PC but additionally satisfies the constraint of being orthogonal to representative time series of other networks. Orthogonality constraints are common in the computation of many connectivity metrics[163]. Two versions of our method exist:

- Pairwise orthogonalization: For a given network pair, an associated pair of orthogonal representative time series is estimated.
- Global orthogonalization: For a given single network, one representative time series is estimated that is orthogonal to a set of time series representing the activity outside of that network. The number  $N_c$  of time series used to represent the activity outside of the network of interest is a parameter of that method.

Through simulation studies (appendix A.1) we found that of the considered methods, the optimal choice is global orthogonalization with  $N_c = 3$ . Compared to the first PCs, the resulting representative time series led to a relative increase of explained variance in ground truth signals of up to 40%. We will, therefore, use this variant for all primary analyses. For exploratory analyses, which explore larger model spaces, both pairwise and global orthogonalization are used.

#### *Intrinsic network connectivity and activity*

As connectivity between two networks, we defined the amplitude envelope correlation between their corresponding representative time series. We opted for an amplitude-based

measure as this is the option that is conceptually most compatible with the fMRI-based approaches that have originally been employed to study intrinsic brain networks[164]. Assessments of network activity will be performed as control analyses. As activity of a network, we define the absolute variance of signals of that network which can be explained by its representative time series. In essence, the activity of a network simply corresponds to the power of its representative time series. Since the distribution of activity values is highly skewed, we consider logarithmically transformed activity values in all statistical and machine learning analyses.

#### 4.1.4 Statistical analysis and machine learning

This project's overarching objective is to examine the relationship between intrinsic brain networks and chronic pain. The multitude of data sets at our disposal provides us with the freedom to approach this research question from different angles. In univariate analyses, we investigated the involvement of individual inter-network connections in chronic pain. Multivariate machine learning analyses aimed to establish a link between patterns of inter-network connectivity and the experience of chronic pain. Further, in secondary and exploratory analyses, we assessed other cohorts as well as different outcome and predictor variables. To be able to quantify evidence against effects, whenever feasible, we employed Bayesian statistics and reported Bayes factors (BFs) along with posterior estimates of correlation coefficients (denoted by "post. r").

We integrated information from diverse data sets primarily using a discovery + replication approach. In the discovery + replication approach, one data set is designated as the discovery data set, while the remaining data sets serve as replication sets. This framework is sensitive to strong effects that are consistent across data sets. Throughout all analyses presented here, *Set\_PainlabDiscovery* is defined as the discovery data set. We complemented the discovery + replication approach with a mega analysis in which

**Table 4-2:** Summary of analysis strategies. In this study we employ univariate and multivariate analyses and integrate data from different sources using a discovery + replication approach. In the discovery + replication approach, one data set is the designated discovery set, while the remaining data sets are exclusively used for replication purposes. We complement the discovery + replication approach with a mega analysis in which all data is analyzed jointly.

	Discovery + Replication	Mega analysis
Univariate	<ul style="list-style-type: none"> <li>In <b>discovery set</b>, one corr. BF for each independent variable</li> <li>No. of <b>replication sets</b> with corr. BF &gt; 3</li> </ul>	In <b>combined set</b> , one corr. BF for each independent variable
Multivariate	<ul style="list-style-type: none"> <li>"In-sample" validation: One permutation-based p-value for <b>discovery set</b></li> <li>"Out-of-sample" validation: No. of <b>replication sets</b> with prediction-outcome corr. BF &gt; 3</li> </ul>	"leave-one-out" cross-validation: Prediction-outcome observation for pooled data and for study sub-sets.

data from all sources are analyzed jointly. The motivation behind this is twofold: First, as more data are incorporated at the same time, more subtle effects, which are present across data sets, may be detected. Second, the hierarchy between discovery and replication sets, which is arbitrary to some degree, is resolved. An overview of the different analysis strategies is provided in Table 4-2.

### *Univariate analyses*

Within the discovery + replication framework, univariate analyses were conducted as follows. For each network pair, we computed the BF of linear correlation across individuals between connectivity values and outcome measures (e.g., avgPain). This was done independently for all data sets. We then visualized the BFs of linear correlation for the discovery data set. For those connections which showed evidence for an effect in the discovery set, we counted the number of replication sets which also yielded at least moderate evidence for an effect ( $BF > 3$ ). Naturally, effects were only counted, if they pointed in the same direction as in the discovery data set. We defined, a priori (preregistration <https://osf.io/4qmyw/>), that if at least half of the replication sets yielded at least moderate evidence for an effect, we would consider this effect “consistent across data sets”.

In univariate mega analyses, all data sets were considered jointly. Accordingly, for each network pair, the analysis yielded one BF of linear correlation. For  $BFs > 3$ , we would consider the corresponding effect as “present in the joint data set”.

### *Multivariate analyses*

The discovery + replication framework is particularly appropriate in the context of machine learning models as these have several degrees of freedom which can be tuned based on the discovery set in the initial discovery stage. The procedure of identifying, fitting and testing the best model, which is inspired by the approaches in [57] and [53], is described in the following and visualized in Figure 4-2.

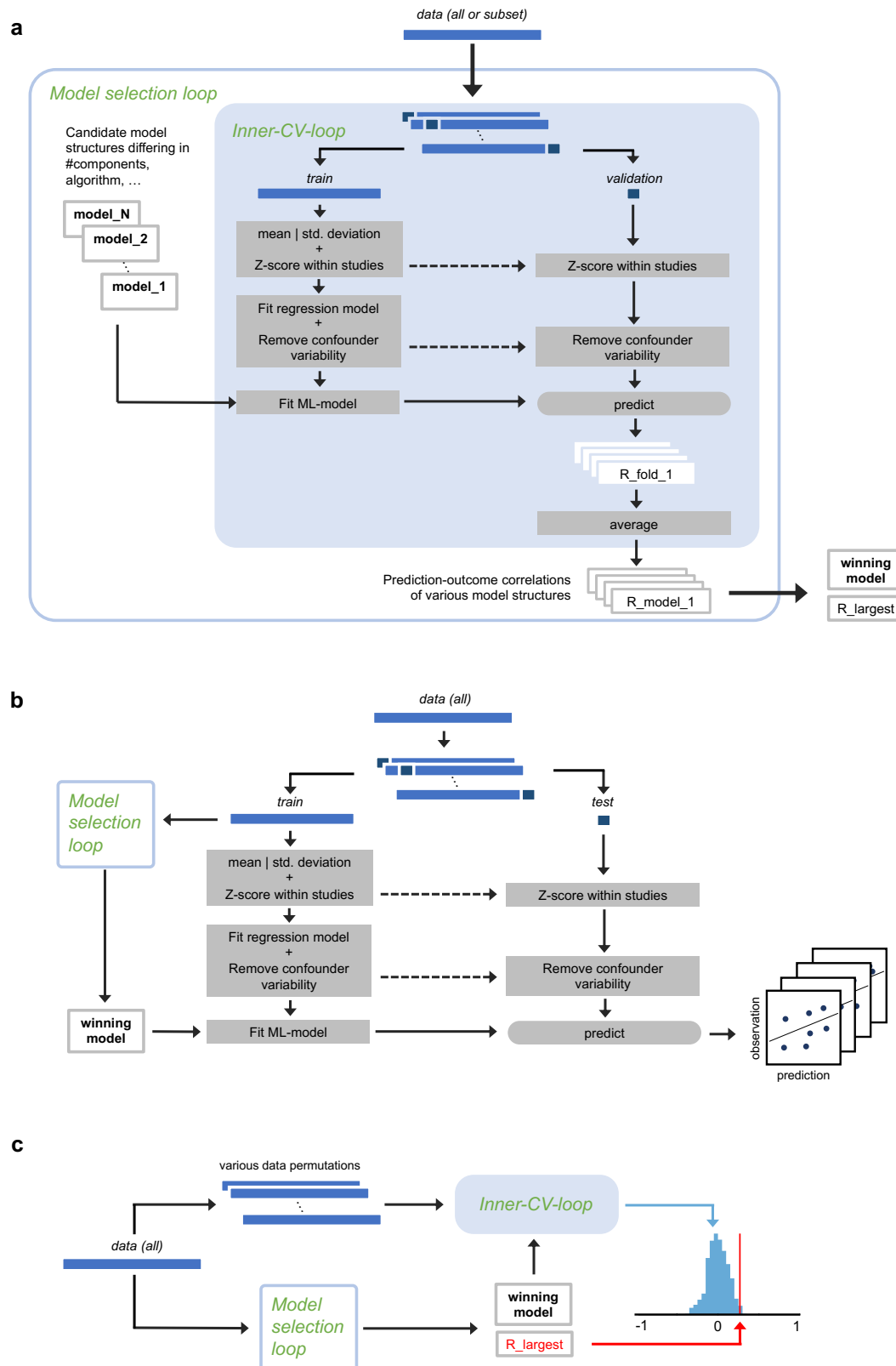
First, we defined several candidate models with distinct attributes. These models differed in their structure (i.e., type of machine learning algorithm) and the number of components they encompassed. Specifically, we explored two model structures: principal component regression and partial least squares regression. These algorithms were selected for their aptitude to deal with highly correlated input, as seen in EEG-based inter-network connectivity values. In both algorithms, the number of included components has to be specified. This number constitutes the second model attribute. In the model discovery stage, we considered component numbers ranging from one to twelve. We opted for a maximum number of twelve as the analysis involves four networks and three frequency

bands. Thus, in primary analyses, our setup involved a total of 24 distinct candidate model structures. In primary and secondary analyses, when average pain and depression scores were defined as dependent variables, the candidate models comprised 18 predictors (6 connectivity values in three frequency bands). For age as the dependent variable, all seven networks were included, resulting in 63 predictors.

Based on the discovery data set, for each candidate model (Figure 4-2a, model selection loop), we computed a cross-validated prediction accuracy as follows (Figure 4-2a, inner-CV-loop).

- **1 Data splitting:** Split the discovery set into a training and a validation sub set in a 9:1 ratio.
- **2 Standardization and confounder removal:** Standardize all independent and dependent variables in both training and validation datasets using mean and standard deviation estimates from the training set. Similarly, fit regression models relating all independent and dependent variables to confounding variables using the training set. Then, apply the regression models to remove confounder-related variability in all independent and dependent variables. This rigorous standardization and confounder removal procedure ensures that there is no information leakage from validation to training sets.
- **3 Model fitting and prediction:** Fit the model using the training set and predict the observations in the validation set.
- **4 Prediction accuracy in current CV-fold:** Compute the correlation coefficient between predictions and observations in the validation set
- Repeat steps one through four  $k = 300$  times (more precisely, repeat steps one through four for the remaining 9 data partitions, then reshuffle and repeat the entire procedure 30 times).
- Compute the prediction accuracy of the model as the average of prediction-observation correlations across all  $k = 300$  validation sets.

For further analyses, the model attributes yielding the highest prediction accuracy were employed. The model possessing these optimal attributes is referred to as the winning model.



**Figure 4-2:** Visualization of the machine learning pipeline. (a) At the discovery stage, different model structures (i.e. machine learning algorithm/ number of included components) are assessed in a “model selection loop” by computing their cross-validated prediction accuracy in the “inner-CV-loop”. (b) The prediction accuracy of models resulting from the developed pipeline is estimated using a leave-one-out cross-validation (CV) procedure. (c) To assess whether a specific model performs better than chance, we employ a permutation-based test.

To assess model performance within the discovery data sets, we performed an “in-sample” validation. First, to estimate the prediction accuracies of models generated by our pipeline, we performed a leave-one-out cross-validation procedure (LOO-CV, Figure 4-2b). This involved selecting and then fitting a winning model structure based on all data points in the discovery set except one. The selected and fitted model was then used to predict the target value of the omitted (test) data point. This process was iterated, leaving out a different data point in each cycle, until all data points were used as test data exactly once. Note that, with this approach, the identified model structure may vary across iterations. Based on the resulting set of prediction-outcome value pairs, we estimated the prediction accuracy by computing spearman’s rho and corresponding p-value. Importantly, in the LOO-CV loop, as in the inner-CV-loop described above, data standardization and confounder removal were performed in a manner to ensure that no information leaks from test to training sets. Next, to statistically evaluate the predictive performance of the winning model which had been identified using the complete discovery data set, we conducted a non-parametric permutation test (Figure 4-2a and c). This involved repeatedly shuffling the response variable and re-calculating the prediction accuracy a large number of times (usually  $2^9$  times). Subsequently, a p-value was derived by comparing the prediction accuracy observed in the unshuffled data to the distribution of prediction accuracies of the shuffled data.

If the permutation-based in-sample validation indicated that the winning model predicted outcomes better than chance, signified by a p-value of less than 0.05, we proceeded to fit this model to the complete discovery dataset. With this refitted winning model, we predicted the observed outcomes in all replication data sets and computed BFs of linear correlation between predictions and observations. We defined, a priori (preregistration <https://osf.io/4qmyw/>), that if the correlation BFs in at least half of the replication sets indicated at least moderate evidence for an effect, we would consider the model to be “generalizing across data sets”.

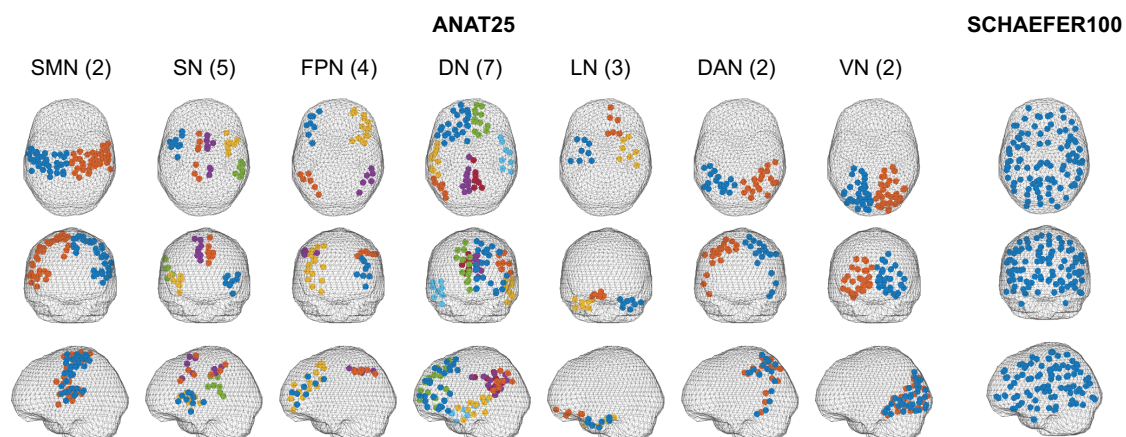
In multivariate mega-analyses, we essentially performed a LOO-CV procedure (Figure 4-2b) on a large data set merging data from all individual studies. We evaluated cross-validated prediction-outcome correlations for the complete set of prediction-outcome value pairs and, in the primary analysis, also for subsets corresponding to each individual study. Specifically, we computed correlation BFs and associated posterior estimates of correlation coefficients. In exploratory analyses, certain model candidates featured a substantially higher number of predictors as compared to primary and secondary analyses. Due to the increased computational demands associated with these larger models and considering our computational resource constraints, all exploratory mega analyses

employed a 10-fold Cross-Validation (10-fold-CV) procedure instead of the computationally more demanding LOO-CV.

### Exploratory analyses

The exploratory analyses were variants of the above-described multivariate discovery + replication and mega analysis strategies. They extended these approaches by augmenting the space of candidate models. In addition to different model structures and numbers of components, the new model space also encompassed several spatial configurations, varied techniques for generating representative time series, and different definitions of amplitude-based connectivity.

Spatial configurations describe the spatial entities between which connectivity is computed. The primary analysis considered only a single spatial configuration -it corresponded to four out of the seven Yeo networks. By contrast, exploratory analyses involve three different spatial configurations. The first configuration, labeled YEO7, corresponds to the complete collection of seven Yeo networks (Figure 4-3). With this configuration the predictive model comprises 63 predictors (21 connectivity values among 7 networks in three frequency bands). The second configuration, labeled ANAT25, involves 25 anatomical brain regions which are defined by grouping the Schaefer parcels according to their anatomical labels rather than according to their network affiliation. Consequently, the predictive model comprises 900 predictors in this configuration. The third configuration, termed SCHAEFER100, defines each individual parcel of the 100-parcel Schaefer atlas as a distinct spatial entity, resulting in 14,850 predictors.



**Figure 4-3:** Visualization of alternative spatial configurations. First seven columns: axial, coronal, and sagittal views of the 25 regions of the ANAT25 spatial configuration. For visualization purposes, the ANAT25-regions are grouped by their Yeo network affiliation; that is, each ANAT25-region belongs to exactly one Yeo network. For example, two out of 25 regions correspond to the left and right halves of the SMN, another five out of 25 regions correspond to five subdomains of the SN, and so on. Different colors are used to distinguish the ANAT25-regions within each Yeo network (due to a limited palette, colors do not distinguish regions across different Yeo networks). Last column: axial, coronal, and sagittal view of the 100 parcel centroids of the SCHAEFER100 spatial configuration.



Visualizations of the ANAT25 and SCHAEFER100 spatial configurations are provided in Figure 4-3.

Moreover, we explored three techniques to compute representative time series. With reference to earlier explanations, these techniques are standard PCA, “pairwise orthogonalization”, and “global orthogonalization” with  $N_c = 3$ . Due to computational resource constraints, “pairwise orthogonalization” was not performed for the spatial configuration SCHAEFER100. We further explored two variants of AEC definitions. The first variant, which is the one originally proposed in [116] and which is also employed in the primary analysis, involves computing correlations between logarithmically transformed squared amplitude envelopes. The second, computes correlations between untransformed amplitude envelopes. In the primary analysis, the maximum number of components explored during model selection was motivated based on the number of Yeo networks and the number of frequency bands included in the analysis. While the number of frequency bands included in the exploratory analysis is still three, the spatial configurations now cover the entire brain, i.e., they comprise all seven Yeo networks. Therefore, the new maximum number of components is set to 21.

## 4.2 Results

In this study, we investigated the role of brain network function in chronic pain. The structure of the results aligns with the description provided in the preregistration (<https://osf.io/4qmyw/>). In primary analyses, we examine the relation between four selected intrinsic brain networks and pain intensity in cohorts of patients with chronic back pain. These analyses comprise both univariate and multivariate methodologies. To test the functional specificity of the primary results, we conduct control analyses in which we use measures of intrinsic network activity as predictors instead of intrinsic network connectivity.

Uni- and multivariate secondary analyses expand the focus to cohorts of patients experiencing diverse forms of chronic pain. Besides average pain, secondary analyses define depression scores and age as alternative dependent variables. Moreover, in exploratory analyses, we pursue a more data driven strategy to predict average pain, depression, and age. Specifically, we enlarge the set of candidate models by including, e.g., different spatial configurations and different methods for computing representative time series and amplitude-based connectivity. An overview of all analyses is provided in Table 4-3.

**Table 4-3:** Overview of various analysis variants.

<i>Primary analysis</i>	<i>Secondary analyses</i>	<i>Exploratory analyses (extended model space)</i>
<p><b>YEO4 connectivity → avg. pain</b> Cohort: CBP Analysis variants:</p> <ul style="list-style-type: none"> <li>• Univariate, disc. + repl.</li> <li>• Univariate, mega</li> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>	<p><b>YEO4 connectivity → avg. pain</b> Cohort: ALL Analysis variants:</p> <ul style="list-style-type: none"> <li>• Univariate, disc. + repl.</li> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>	<p><b>conn./ act. → avg. pain</b> Cohort: CBP/ALL Analysis variants:</p> <ul style="list-style-type: none"> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>
<p><i>Control analysis</i></p> <p><b>YEO4 activity → avg. pain</b> Cohort: CBP Analysis variants:</p> <ul style="list-style-type: none"> <li>• Univariate, disc. + repl.</li> <li>• Univariate, mega</li> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>	<p><b>YEO4 connectivity → depres.</b> Cohort: ALL Analysis variants:</p> <ul style="list-style-type: none"> <li>• Univariate, disc. + repl.</li> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>	<p><b>conn./ act. → depression</b> Cohort: CBP/ALL Analysis variants:</p> <ul style="list-style-type: none"> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>
	<p><b>YEO4 connectivity → age</b> Cohort: ALL Analysis variants:</p> <ul style="list-style-type: none"> <li>• Univariate, disc. + repl.</li> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>	<p><b>conn./ act. → age</b> Cohort: CBP/ALL Analysis variants:</p> <ul style="list-style-type: none"> <li>• Multivariate, disc. + repl.</li> <li>• Multivariate, mega</li> </ul>

#### 4.2.1 Primary analyses: Intrinsic brain network function in chronic back pain

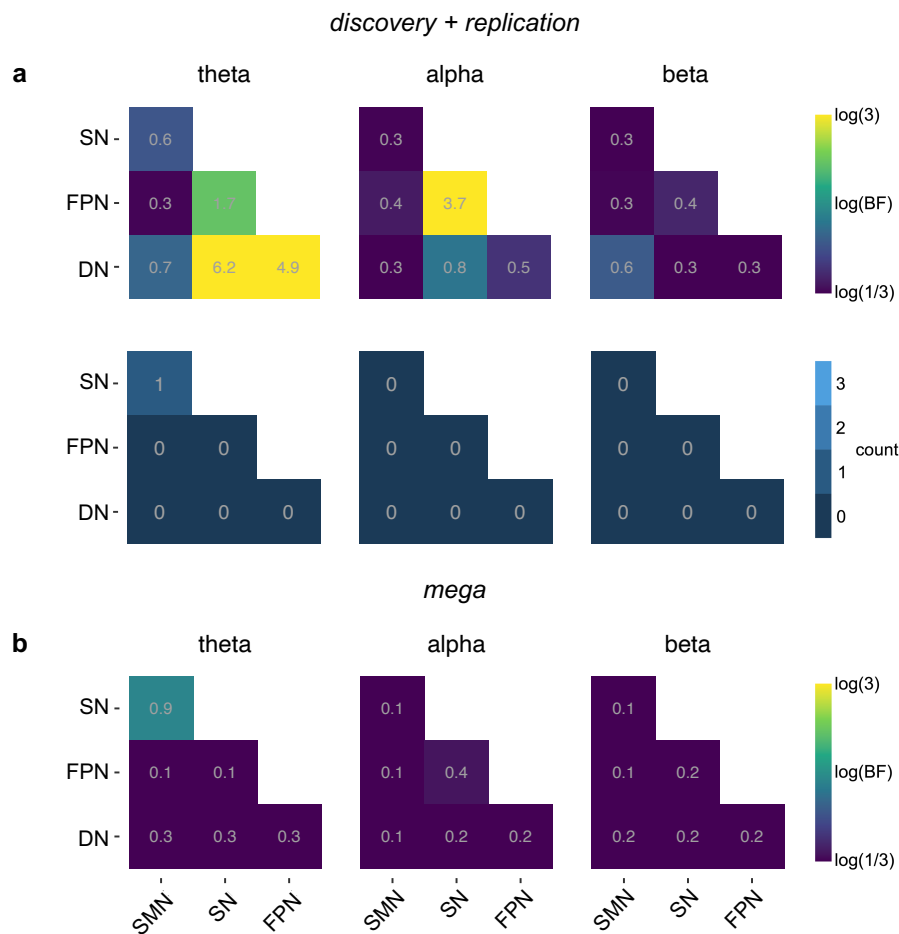
The first set of analyses addressed the question whether inter-network connectivity as measured by EEG relates to pain intensity in cohorts of patients with chronic back pain. Leveraging the high temporal resolution of EEG, we hoped to unravel temporal and spectral characteristics of network mechanisms that may elude traditional methods in the field, such as fMRI. We assessed connectivity among four intrinsic brain networks in the theta, alpha, and beta frequency bands. In particular, we opted to evaluate connectivity among the frontoparietal (FPN), the salience ventral attention (SN), the default (DN), and the somatomotor network (SMN) as these networks have previously been linked to several neuropsychiatric disorders including chronic pain [74, 87]. For each network pair, we computed connectivity based on an associated pair of representative time series, a procedure which is explained in more detail in the project-related as well as the general methods sections of this manuscript.

##### *Univariate analyses*

To begin with, we investigated to which degree inter-network connectivity of individual network pairs could explain the inter-individual variability of pain intensity scores. Figure 4-4a illustrates the results of univariate Bayesian correlation analyses. In the discovery data *Set\_PainlabDiscovery*, we found moderate evidence for a correlation between inter-network connectivity and pain intensity for two network pairs in the theta (DN-SN: BF = 6.2, post.  $r = 0.28$ ; DN-FPN: BF = 4.9, post.  $r = 0.27$ ) and for one network pair in the alpha band (FPN-SN: BF = 3.7, post.  $r = -0.24$ ). Interestingly, pain intensity was associated with hyper and hypo-connectivity in the theta and alpha frequency bands, respectively. These results could not be replicated in the independent data sets comprising patients with chronic back pain. In fact, evidence for an effect was found in only one replication data set. Similar to observations in the discovery data, this effect occurred in the alpha band and involved the SN. Analyzing data from all studies jointly in a mega analysis did not yield any effects either (Figure 4-4b).

To test the functional specificity of the presence and absence of effects, we repeated the analysis with measures of intrinsic network activity as independent variables (Appendix section A.3). No evidence for an association between intrinsic network activity and pain intensity was found -neither in the discovery nor in any of the replication sets. The activity-based mega analysis yielded a similar picture.

In summary, we found evidence of an association between inter-network connectivity and pain intensity in the discovery data set. In contrast, we found mostly evidence against an association between pain intensity and the activity of individual networks. While the

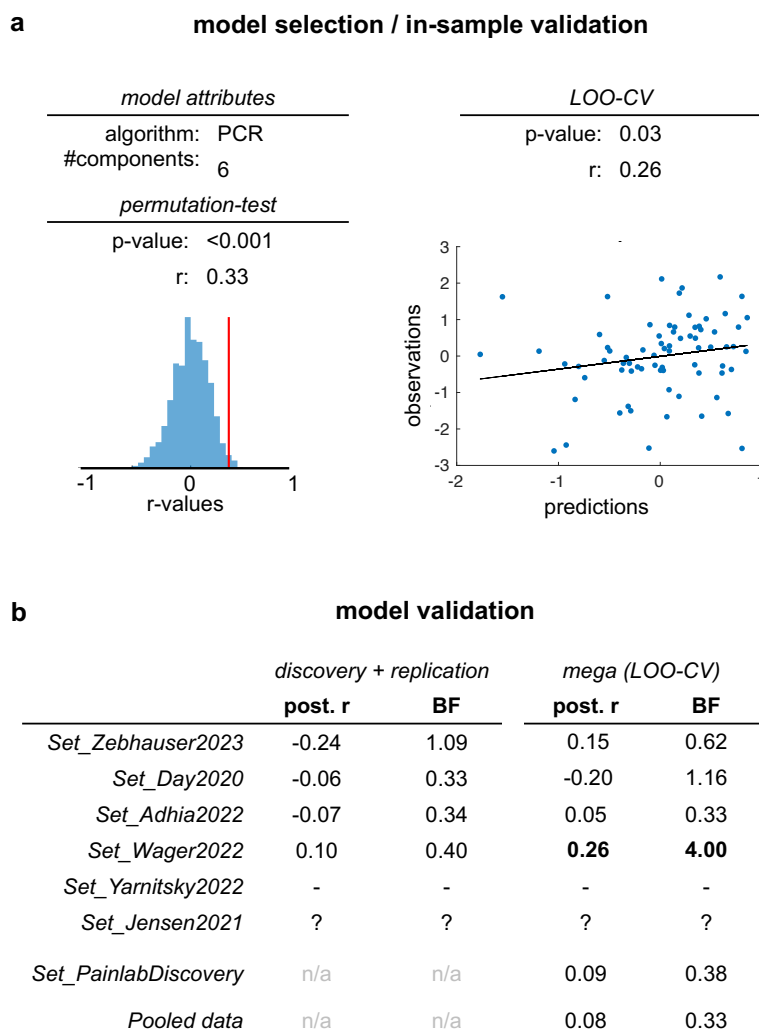


**Figure 4-4:** Relation between pain intensity and connectivity in individual pairs of intrinsic brain networks in patients with chronic back pain. Correlations were assessed using a Bayesian approach. (a) Top panel row: The color of heat map tiles scales with the logarithm of the correlation BF<sub>s</sub> calculated based on the discovery data set. It ranges from blue (BF < 1/3, at least moderate evidence against an effect) to yellow (BF > 3, at least moderate evidence for an effect). Bottom panel row: Visualization of number of replication sets for which the correlation BF was > 3 and the sign of the estimated correlation coefficient was the same as in the discovery data set. (b) Results from the mega analysis. The color coding is the same as in the top panel row of subfigure (a). SN denotes salience ventral attention, FPN frontoparietal, DN default, and SM somatomotor network.

connectivity-based effects in the discovery data set could not be replicated in the independent data sets, the absence of activity-based effects was consistent across data sets.

### Multivariate analyses

Next, we assessed whether pain intensity related to patterns of intrinsic brain network connectivity using a multivariate machine learning approach. As predictors, we used connectivity values among the four preselected intrinsic brain networks in all three frequency bands. We trained several multivariate candidate models on the discovery data *Set\_PainlabDiscovery*. Candidate models differed in their structure (i.e., type of machine learning algorithm) and the number of included components. Using the cross-validated prediction-outcome correlation in the discovery set as a criterion, we nominated a winning



**Figure 4-5:** Relation between pain intensity and patterns of inter-network connectivity in patients with chronic back pain. Panel (a) on the left indicates the attributes of the winning model as well as the results of its permutation-based statistical assessment. The blue histogram and red line in the plot below visualize the prediction-outcome correlations for the permuted and non-permuted data, respectively. Panel (a) on the right shows the results of a LOO-CV assessment of the machine learning pipeline within the discovery data set, that is, the scatter plot of cross-validated prediction-observation pairs as well as the corresponding correlation coefficient and p-value. (b) The left section of the table (“discovery + replication”) shows the prediction-outcome correlations achieved by the winning model in the individual replication data sets. The right section (“mega”) shows the prediction-outcome correlations resulting from a LOO-CV procedure in the combined data set. The LOO-CV procedure results in prediction-observation pairs that can be assessed across studies (“pooled data”) or separately for each individual study.

model and used it to predict pain intensity scores in all replication data sets involving patients with chronic back pain. Figure 4-5 shows the results of this analysis.

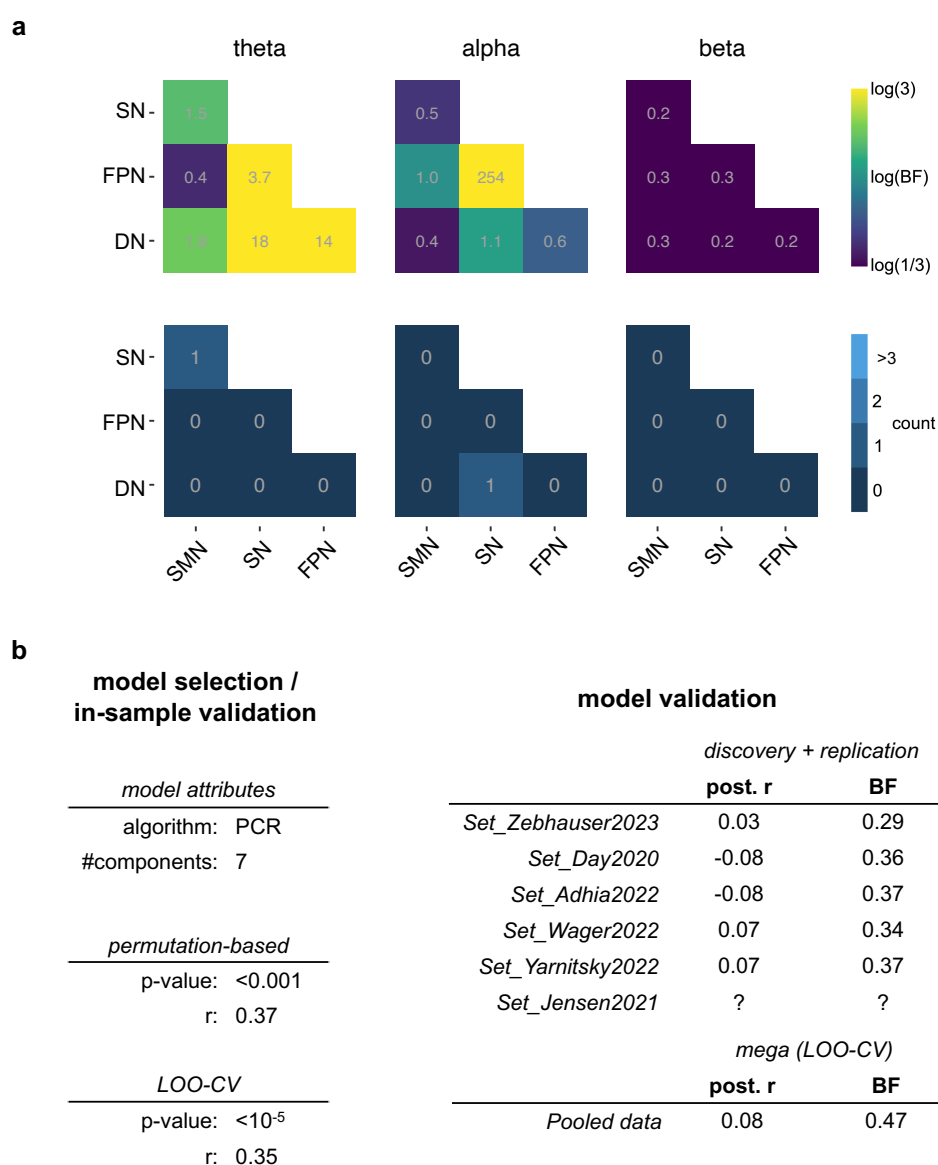
The winning model for the prediction of pain intensity employed principal component regression (PCR) with six components. Its cross-validated prediction-outcome correlation stood at 0.33. To determine whether the winning model predicted pain scores better than chance, we compared it to a null-distribution of prediction-outcome correlations. This null-distribution was generated by repeatedly fitting a model with the same attributes to the discovery data-set, reshuffling the response variable each time. We call this procedure the

permutation-based in-sample validation. The resulting p-value ( $<0.001$ ) indicated that the identified multivariate model predicts pain intensity better than chance (Figure 4-5a, left part), suggesting that pain intensity-related patterns of intrinsic brain-network connectivity do exist in the discovery data set. Additionally, to assess the overall performance of our machine learning pipeline, a leave-one-out cross-validation (LOO-CV) procedure was conducted (Figure 4-5a, left part). This analysis also demonstrated a significant association between predictions and observations of pain ratings (cross-validated  $r = 0.26$ ,  $p < 0.05$ ).

After in-sample validation, the winning model was trained using the complete discovery data set, and used to predict pain intensity ratings in the four replication data sets comprising patients with chronic back pain (Figure 4-5b). In each of these data sets, we determined BFs for prediction-outcome correlations. None of these BFs indicated evidence for an association between predicted and observed pain scores. Furthermore, the posterior estimates of correlation coefficients were not consistently greater than zero.

In addition, to test whether a more subtle patterns that are consistent across data sets could be identified with our machine learning pipeline, we performed LOO-CV procedure on the combined data set. In analogy to the univariate case, we refer to this approach as mega analysis. As with the discovery + replication approach, the mega analysis did not yield evidence for the existence of consistent pain intensity-related patterns of intrinsic network connectivity. Only when the subset of cross-validated predictions corresponding to *Set\_Wager2022* was considered, was there moderate evidence for a positive correlation with observed pain ratings.

As before, we repeated the above-described analysis steps using measures of intrinsic brain network activity in the three frequency bands as predictors (Appendix section A.3). The in-sample validation procedure yielded a permutation-based p-value of 0.25. This suggested that that the winning model utilizing intrinsic brain network activity values as predictors did not significantly outperform chance in predicting pain intensity ratings in the discovery data set. Therefore, testing the model in the replication data sets was not indicated.



**Figure 4-6:** Relation between pain intensity and intrinsic brain network connectivity in patients with diverse pain conditions. (a) Results from univariate analyses. For detailed explanations of the plots, refer to Figure 4-4a. (b) Results from the multivariate analysis. For explanations of the tables, refer to Figure 4-5.

#### 4.2.2 Secondary analyses: Intrinsic brain network function in chronic pain, depression, and age

##### *Different cohort: diverse chronic pain etiologies*

In previous analyses, we focused on patients with chronic back pain in order to minimize variability attributable to differences in pain etiology. Implicit to this approach was the assumption that cerebral pathomechanisms contributing to chronic pain differ substantially across different types of chronic pain. An alternative view suggests that different forms of chronic pain are primarily determined by one shared cerebral pathomechanism. In this scenario, using data from all patients would enhance statistical power and facilitate the detection of these mechanisms. Therefore, in a secondary analysis, we conducted uni-

and multivariate discovery + replication analyses including all patients rather than just those with chronic back pain.

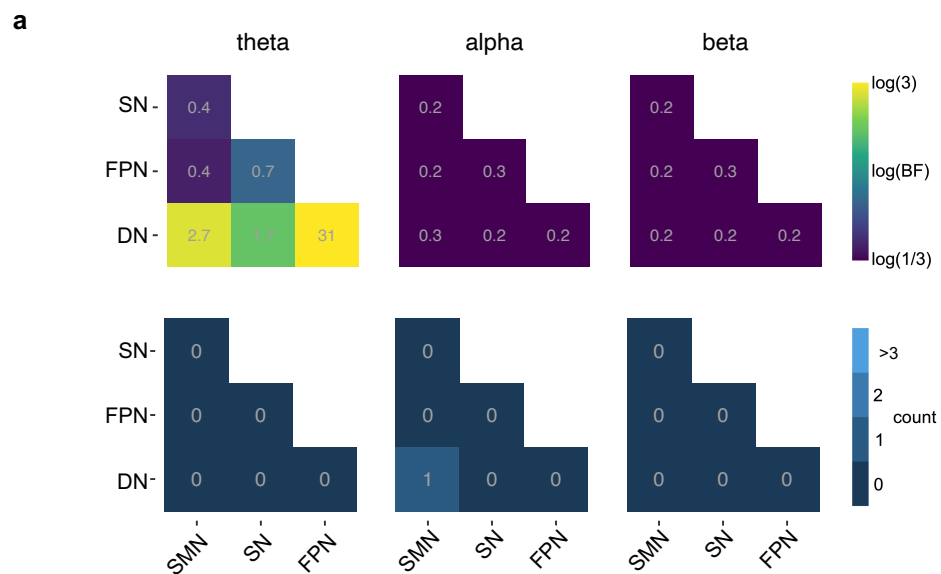
Figure 4-6a depicts the results of univariate analyses. In the discovery data set, we found evidence for an association between inter-network connectivity and pain intensity in four network pairs in the theta and alpha bands. Among these, three network pairs (theta, DN-SN, post.  $r = 0.26$ ; theta, DN-FPN, post.  $r = 0.25$ ; alpha, FPN-SN, post.  $r = -0.32$ ) had also indicated evidence for an effect in the primary analysis involving patients with CBP only. Notably, the strength of evidence in these three network pairs is increased compared to the primary analysis ( $BF > 10$  in all three). In terms of the direction of effects, as before, we observed that pain intensity increased with theta connectivity and decreased with alpha connectivity. In replication data sets, we also found effects in the theta and alpha frequency bands. However, none of these effects spatially coincided with those observed in the discovery data. A multivariate model trained on the diverse pain cohort achieved a significant prediction accuracy in the discovery data set (permutation test:  $r = 0.37$ ,  $p < 0.001$ ; LOO-CV  $r = 0.35$ ,  $p < 10^{-5}$ ). However, this model could not predict pain scores in any of the replication data sets (Figure 4-6b). Similarly, a mega analysis, which involved model-training on the combined data set, did not yield evidence for multivariate pain-related patterns of intrinsic brain network connectivity.

#### *Intrinsic brain network connectivity and depression scores in patients with chronic pain*

Chronic pain commonly co-occurs with a range of neuropsychiatric comorbidities, most notably depression. This prompted us to investigate, in an additional secondary analysis, whether an association exists between intrinsic brain network connectivity and depression severity scores in patients with chronic pain. Furthermore, this analysis would allow us to assess whether the effects found when pain was considered as the dependent variable were specific to pain or overlap with effects seen in depression.

Univariate analyses (Figure 4-7a) in the discovery data set indicated that connectivity between the DN and the FPN in the theta band could explain some of the variability of depression scores ( $BF > 30$ , post.  $r = 0.26$ ). This finding could not be confirmed in the replication data sets. Furthermore, based on the discovery data set, we found evidence against a relation between inter-network connectivity and pain intensity for all network



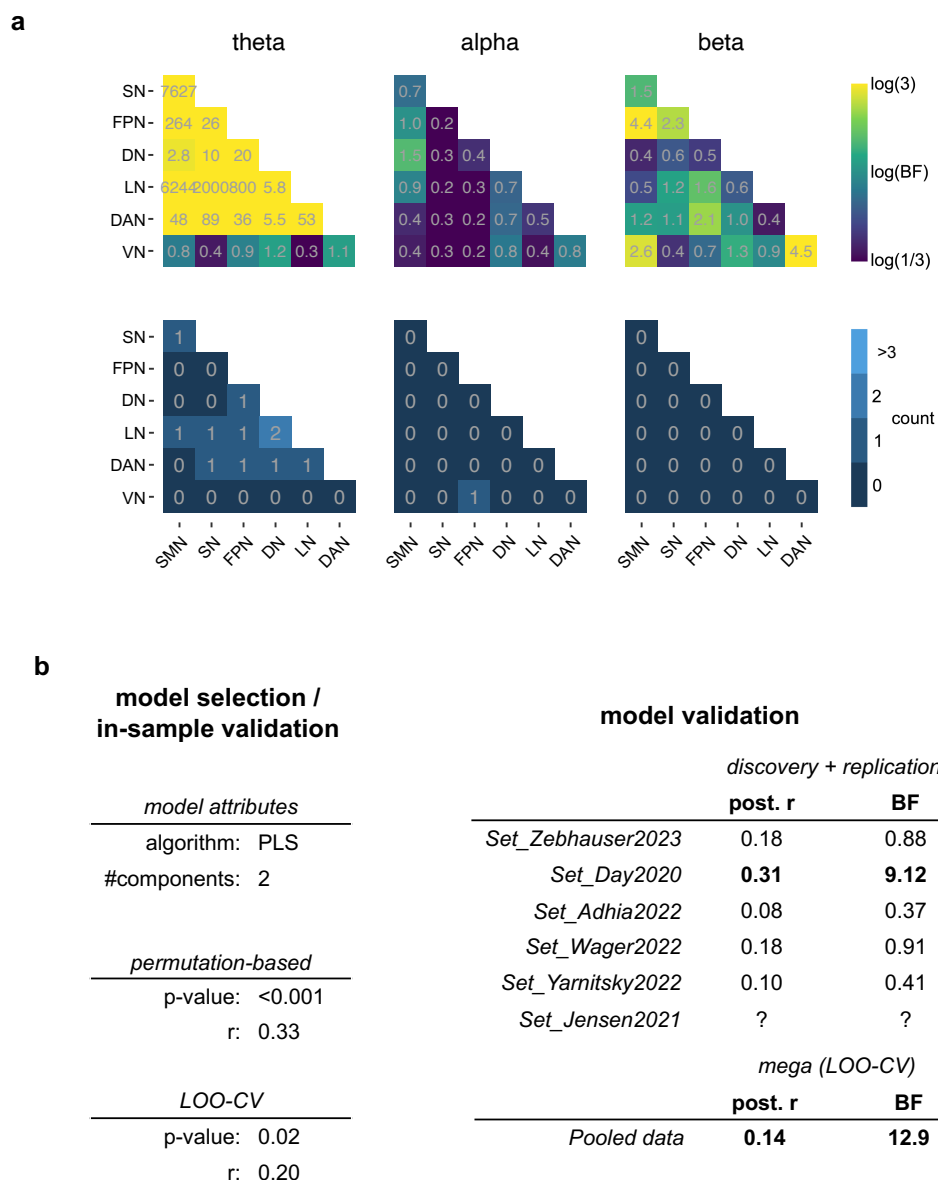


**b**

model selection / in-sample validation		model validation	
<i>model attributes</i>		<i>discovery + replication</i>	
algorithm: PCR		<b>post. r</b>	<b>BF</b>
#components: 5		<i>Set_Zebhauser2023</i>	-0.15    0.63
<i>permutation-based</i>		<i>Set_Day2020</i>	-0.07    0.35
p-value: 0.04		<i>Set_Adhia2022</i>	0.00    0.30
r: 0.19		<i>Set_Wager2022</i>	0.07    0.34
<i>LOO-CV</i>		<i>Set_Yarnitsky2022</i>	-0.02    0.32
p-value: 0.07		<i>Set_Jensen2021</i>	?    ?
r: 0.15		<i>mega (LOO-CV)</i>	
		<b>post. r</b>	<b>BF</b>
		<i>Pooled data</i>	0.05    0.20

**Figure 4-7:** Relation between depression severity and intrinsic brain network connectivity in patients with diverse pain conditions. (a) Results from univariate analyses. For detailed explanations of the plots, refer to Figure 4-4a. (b) Results from the multivariate analysis. For explanations of the tables, refer to Figure 4-5.

pairs in the alpha and beta bands. A multivariate model (Figure 4-7b), which just barely reached significance within the discovery data set (permutation test:  $r = 0.19$ ,  $p = 0.04$ ; LOO-CV  $r = 0.15$ ,  $p = 0.07$ ), did not generalize to any of the replication data sets. Moreover, there was evidence against an association between observed depression scores and the predictions of a model that was trained on the combined data set in the multivariate mega analysis.



**Figure 4-8:** Relation between age and intrinsic brain network connectivity in patients with diverse pain conditions. (a) Results from univariate analyses. For detailed explanations of the plots, refer to Figure 4-4a. (b) Results from the multivariate analysis. For explanations of the tables, refer to Figure 4-5.

### *Intrinsic brain network connectivity and age in patients with chronic pain*

Chronic pain and depression are highly complex phenomena with strong subjective components. To translate these phenomena into measurable quantities and, subsequently, to identify a link between these quantities and neural measures such as intrinsic brain network connectivity, is presumably extremely challenging. In contrast, demographic features such as an individual's age are straightforward concepts and their sensitivity with respect to age has been demonstrated in several imaging studies[88-90]. To showcase the effectiveness of our methodology, we opted to apply it to age prediction in a third secondary analysis. Since previous analyses were concerned with predictions of

neuropsychiatric symptom severity scores, they restricted the focus to only four out of seven Yeo networks. Age is not a measure of neuropsychiatric dysfunction and, therefore, its link with intrinsic brain network connectivity is investigated here in all seven Yeo networks.

Univariate analyses (Figure 4-8a) yielded clear evidence for a relation between connectivity and age. The pattern of effects exhibited both spectral and spatial structure as effects primarily occurred in the theta band and not in network pairs involving the visual network. For all network pairs in the theta band, connectivity values decreased with age. Moreover, Bayesian correlation analyses yielded BFs > 10 in twelve and BFs > 100 in four network pairs ( $-0.37 < \text{post. } r < -0.21$ ) in the theta band. The effect observed in the limbic-default network pair was successfully replicated in two independent data sets (*Set\_Wager2022* and *Set\_Day2020*). Effects seen for an additional nine network pairs could be replicated in one independent data set. Since there is a total of five replication data sets, we cannot quite claim to have found effects that are “consistent across data sets”, as specified in the preregistration. However, in view of the magnitude of BFs and also the spectral and spatial specificity of the results, a relation between theta band connectivity and age seems very likely.

To test the functional specificity of these findings, we repeated the analysis in the discovery data set using intrinsic network activity as predictors. For none of the networks did we find evidence of an association between network activity and age. Out of five replication sets, only *Set\_Wager2022* yielded evidence for a negative association between age and activity in all intrinsic brain networks at theta and gamma frequencies.

The multivariate model trained on the discovery data set clearly indicated the existence of an age-related multivariate pattern of intrinsic brain network connectivity (permutation test:  $r = 0.33$ ,  $p < 0.001$ ; LOO-CV  $r = 0.20$ ,  $p = 0.02$ ). This model achieved positive prediction-outcome correlations in all replication data sets. In none of the replication sets did we see evidence against an association between predicted and actual age. In one replication set, evidence for such an association was even moderate (BF > 3, *Set\_Day2020*). Moreover, there was strong evidence (BF > 10) for an association between actual age and predictions of a model trained on the combined data set in the multivariate mega analysis.

### 4.2.3 Exploratory analyses

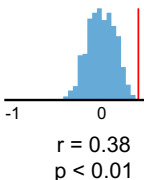
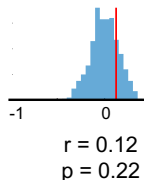
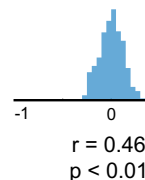
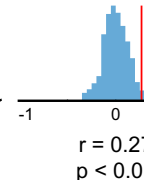
In exploratory analyses we sought to enhance the predictive accuracy of multivariate models by considering a larger model space. At the model identification stage, we incorporate three spatial configurations, component numbers one through 21, up to three different methods for computing representative time series (for one spatial configuration only two variants are computationally feasible), and two definitions of amplitude-based connectivity. A relation between these different feature candidates and observations is established using two distinct machine learning algorithms. This corresponds to a total of 672 candidate models. Conceptually, the use of an enlarged model space represents a shift towards a more data-driven strategy. The structure of exploratory analyses mirrors the one of secondary analyses in that multivariate models are identified which predict average pain, depression scores, and age. Four variants of this analysis were performed: Using either connectivity or activity values as predictors, and including either all patients or only those with CBP.

#### *Predicting average pain*

The results of exploratory analyses with pain intensity as dependent variable are summarized in Table 4-4. Connectivity-based models of pain intensity showed a higher in-sample prediction performance than activity-based models. The largest in-sample prediction accuracy ( $r = 0.46$ ) was attained in the diverse pain cohort by a model employing inter-network connectivity of the seven Yeo networks. This model exhibited positive, albeit very small prediction-outcome correlations in all five replication sets. The second-largest prediction accuracy ( $r = 0.38$ ) was observed for a model which was trained and tested in the CBP cohort and incorporated connectivity among 25 anatomical regions. This model showed positive prediction-outcome correlations in three out of four replication data sets, with moderate evidence for an effect, i.e.,  $BF > 3$ , in *Set\_Wager2022*. The activity-based model trained on the diverse pain cohort, showed a relatively decent in-sample prediction accuracy and exhibited positive prediction-outcome correlations in four out of five replication sets. A  $BF > 3$  signifying moderate evidence for a positive prediction-outcome correlation was observed in *Set\_Wager2022*. For the CBP cohort, the activity-based model failed to predict pain ratings better than chance in the discovery data set.

Since none of the models identified on the basis of the discovery data set generalized consistently to the replication data sets, we additionally performed a multivariate mega analysis. This meant selecting and fitting models based on the combined data set and assessing model performance using a 10-fold cross-validation scheme. The resulting cross-validated prediction-outcome correlations did not indicate improved predictions by

**Table 4-4:** Predicting pain intensity using an extended set of candidate models. Model training and testing was performed separately in cohorts of patients with chronic back pain and in cohorts of patients with diverse pain conditions. Models used either measures of brain activity (within brain regions or networks) or brain connectivity (between brain regions or networks) as predictors.

		model selection							
		CBP		ALL					
		connectivity	activity	connectivity	activity				
model attributes	algorithm:	PLS	PLS	PLS	PLS				
	#components:	6	5	13	4				
	spatial config.:	ANAT25	ANAT25	YEO7	SCHAEFER100				
	rep. method:	pairwise ortho.	pairwise ortho.	pairwise ortho.	pairwise ortho.				
	conn. version:	untransf. env.	-	log. square env.	-				
in-sample validation									
		r = 0.38	r = 0.12	r = 0.46	r = 0.27				
		p < 0.01	p = 0.22	p < 0.01	p < 0.01				
		model validation							
		discovery-replication							
		post. r	BF	post. r	BF	post. r	BF	post. r	BF
	<i>Set_Zebhauser2023</i>	0.21	0.87	0.16	0.69	0.20	1.13	0.01	0.29
	<i>Set_Day2020</i>	-0.13	0.50	0.14	0.56	0.04	0.31	0.22	1.58
	<i>Set_Adhia2022</i>	0.13	0.50	0.08	0.36	0.02	0.30	0.01	0.30
	<i>Set_Wager2022</i>	<b>0.30</b>	<b>9.82</b>	0.19	1.07	0.03	0.28	<b>0.29</b>	<b>7.18</b>
	<i>Set_Yamitsky2022</i>	-	-	-	-	0.02	0.33	-0.17	0.75
	<i>Set_Jensen2021</i>	?	?	?	?	?	?	?	?
		mega (10-fold-CV)							
	<i>Pooled data</i>	-0.04	0.17	<b>0.18</b>	<b>23.1</b>	0.03	0.14	-0.02	0.13

employing the combined data set for model training. Evidence for a correlation between predictions and observations of pain ratings was seen only for an activity-based model. However, considering the incapacity of activity features to predict pain ratings within the discovery data set, the robustness of this finding is questionable.

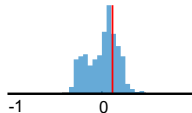
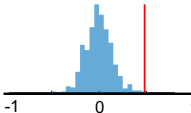
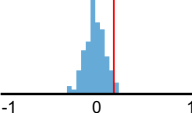
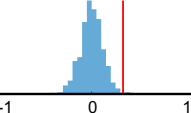
### Predicting depression scores

The results of exploratory analyses with depression as dependent variable are summarized in Table 4-5. In both patient cohorts, the in-sample performance of activity-based models predicting depression scores was better than that of connectivity-based models. The highest in-sample prediction-outcome correlation was observed for the activity-based model in the CBP cohort ( $r = 0.48$ ). As predictors, this model utilized activity values of 25 anatomical regions associated with representative time series derived using the global orthogonalization method. This model could, however, not predict depression

**Table 4-5:** Predicting depression severity using an extended set of candidate models. Model training and testing was performed separately in cohorts of patients with chronic back pain and in cohorts of patients with diverse pain conditions. Models used either measures of brain activity (within brain regions or networks) or brain connectivity (between brain regions or networks) as predictors.

		model selection			
		CBP		ALL	
model attributes		connectivity	activity	connectivity	activity
	algorithm:	PCR	PLS	PCR	PCR
	#components:	1	19	21	11
	spatial config.:	SCHAEFER100	ANAT25	YEO7	YEO7
	rep. method:	global ortho.	global ortho.	global ortho.	pairwise ortho.
	conn. version:	untransf. env.	-	untransf. env.	-

in-sample validation	CBP connectivity		CBP activity		ALL connectivity		ALL activity	
			$r = 0.11$ $p = 0.27$		$r = 0.48$ $p < 0.01$		$r = 0.20$ $p < 0.05$	

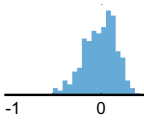
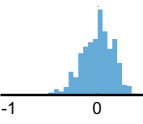
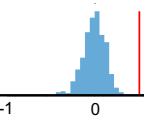
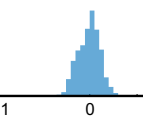
	model validation							
	discovery-replication							
	post. r	BF	post. r	BF	post. r	BF	post. r	BF
Set_Zebhauser2023	0.14	0.63	-0.01	0.47	-0.10	0.40	0.02	0.29
Set_Day2020	0.05	0.31	-0.05	0.32	-0.17	0.71	-0.05	0.32
Set_Adhia2022	-0.05	0.32	0.10	0.41	0.02	0.30	-0.03	0.31
Set_Wager2022	-0.04	0.29	-0.02	0.28	0.12	0.47	-0.10	0.39
Set_Yarnitsky2022	-	-	-	-	0.07	0.36	0.18	0.85
Set_Jensen2021	?	?	?	?	?	?	?	?
	mega (10-fold-CV)							
Pooled data	0.06	0.25	0.08	0.38	-0.05	0.21	0.01	0.11

scores in any of the replication data sets. Mega analyses, employing the combined data set for model training, did not improve predictions in any of the considered analysis variants.

### Predicting age

Lastly, we applied the exploratory multivariate methodology to the prediction of age (Table 4-6). The selected models for all cohorts and predictor types employed the high-resolution SCHAEFER100 spatial configuration comprising 100 brain parcels. Except for the activity-based model in the CBP cohort, all models employed the global orthogonalization method for the determination of representative time series. In-sample prediction-outcome correlations ranged between 0.46 and 0.70. Posterior estimates of prediction-outcome correlations had a positive sign in all replication sets and analysis variants. Interestingly, Set\_Wager2022 exhibited strong evidence for an association between predicted and actual age across all analysis variants. The largest in-sample prediction-outcome

**Table 4-6:** Predicting age using an extended set of candidate models. Model training and testing was performed separately in cohorts of patients with chronic back pain and in cohorts of patients with diverse pain conditions. Models used either measures of brain activity (within brain regions or networks) or brain connectivity (between brain regions or networks) as predictors.

		model selection							
		CBP		ALL					
		connectivity	activity	connectivity	activity				
model attributes	algorithm:	PCR	PCR	PCR	PCR				
	#components:	4	10	5	10				
	spatial config.:	SCHAEFER100	SCHAEFER100	SCHAEFER100	SCHAEFER100				
	rep. method:	global ortho.	PCA	global ortho.	global ortho.				
	conn. version:	log. square env.	-	log. square env.	-				
in-sample validation									
		$r = 0.57$	$r = 0.70$	$r = 0.46$	$r = 0.61$				
		$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$				
		model validation							
		discovery-replication							
		post. r	BF	post. r	BF	post. r	BF	post. r	BF
<i>Set_Zebhauser2023</i>		0.05	0.48	<b>0.36</b>	<b>3.12</b>	0.21	1.34	0.17	0.75
<i>Set_Day2020</i>		0.24	2.26	0.17	0.80	<b>0.29</b>	<b>5.01</b>	0.05	0.32
<i>Set_Adhia2022</i>		0.03	0.31	0.20	1.17	0.07	0.34	<b>0.34</b>	<b>13.5</b>
<i>Set_Wager2022</i>		<b>0.32</b>	<b>13.8</b>	<b>0.36</b>	<b>42.2</b>	<b>0.38</b>	<b>79.2</b>	<b>0.46</b>	<b>1829</b>
<i>Set_Yarnitsky2022</i>		-	-	-	-	0.24	1.79	0.13	0.54
<i>Set_Jensen2021</i>		?	?	?	?	?	?	?	?
		mega (10-fold-CV)							
<i>Pooled data</i>		<b>0.25</b>	<b>&gt;10<sup>3</sup></b>	<b>0.42</b>	<b>&gt;10<sup>11</sup></b>	<b>0.30</b>	<b>&gt;10<sup>8</sup></b>	<b>0.37</b>	<b>&gt;10<sup>13</sup></b>

correlation ( $r = 0.70$ ) was observed for the activity-based model in the CBP cohort. With a value of 0.27, this model also yielded the highest average prediction-outcome correlation across the replication data sets. Mega analyses, using the combined data set for model selection and fitting, further corroborated the existence of pain-related connectivity and activity patterns. While evidence for an association between predicted and actual age was decisive in all analysis variants, higher prediction-outcome correlations were observed for activity-based compared to connectivity-based models. The relatively superior performance of activity-based models is in contrast to what was found in the previous secondary analysis. Notably, the difference between the models identified here and that from the secondary analysis lies in the use of the fine-grained SCHAEFER100 atlas instead of the coarser YEO7 atlas, hinting at a spatially specific age-related pattern of brain activity that is not captured by the considered intrinsic brain networks.

## 5 Discussion

### 5.1 Project 1

In the first project, we investigated how the brain serves sensory and contextual effects on pain. To this end, we analyzed data from an experiment in which noxious stimuli were applied to healthy human participants while independently modulating the levels of stimulus intensity and expectations. “Pain ratings confirmed that stimulus intensity and expectation both influenced pain perception. Analyses of EEG recordings revealed that sensory and expectation effects on pain were served by fundamentally different brain mechanisms. In a core network associated with the processing of pain, sensory information shaped local oscillatory brain activity rather than inter-regional functional connectivity. In contrast, expectation and prediction errors influenced inter-regional functional connectivity but not local oscillatory brain activity”[3].

#### *Sensory and expectation effects on local brain activity and inter-regional brain connectivity*

„We observed that sensory information shapes local oscillatory brain activity more than inter-regional connectivity. The effects of stimulus intensity on local oscillatory activity in various frequency bands are in accordance with previous EEG and MEG studies [35, 36, 84, 145]. [...] We further observed that expectations influenced inter-regional functional connectivity but not local oscillatory brain activity. To the best of our knowledge, expectation effects on functional connectivity have not yet been investigated by neurophysiological recordings”[3]. Some studies have investigated expectation effects on local brain activity with heterogeneous results[84, 85, 165, 166]. “The present findings do not rule out any expectation effects on local brain activity. However, the crucial finding here is not the lack of expectation effects on local oscillatory activity, but that expectation effects on connectivity are stronger than on local oscillatory activity”[3]. Remarkably, the observed spatial-spectral pattern of expectation effects was sparse, indicating their functional, spectral, and spatial specificity. Moreover, aggregate model comparisons which integrated all connections at all frequencies provided direct evidence for expectations being more related to inter-regional connectivity than to local activity.

#### *Results in the context of the predictive coding framework*

“We found that expectation and prediction errors influenced connectivity at alpha/beta and gamma frequencies, respectively. This observation can be interpreted with reference to predictive coding (PC) frameworks of brain function. PC is a general theory used to explain how perception arises from the integration of sensory information and expectations [167]. The framework proposes that the brain maintains an internal model of the environment



which continuously generates predictions about sensory input. Discrepancies between these predictions and the actual sensory evidence, i.e. PEs, serve to adjust the internal model. In this way, the brain allocates its limited resources to events that are behaviorally relevant and useful for updating predictions, i.e., learning. It has been suggested that alpha and beta oscillations serve the signaling of predictions, whereas gamma oscillations have been proposed to signal PEs [59, 141, 168, 169]. The present findings are in good accordance with this framework. They specify that expectation effects on pain might be particularly related to connectivity at alpha frequencies from the prefrontal to the somatosensory cortex. Specifically, expecting less pain was associated with relatively stronger connectivity. This implies that alpha-band connectivity might be mechanistically involved in an active down-regulation of nociceptive input. Prediction errors on the other hand were reflected in reduced gamma connectivity indicating that they are signaled in the brain in terms of a disruption of inter-regional communication which is in line with a recent study on PE signaling in the processing of pain [84] [...]”[3].

In the preceding discussion, PEs are understood as the discrepancies between the intensity of noxious stimuli and the consciously held expectations about that stimulus’ intensity. These consciously held expectations can be conceptualized as high-level predictions. Expanding beyond such high levels of predictions, the PC framework can also be employed to discuss findings on a more basal level. At lower hierarchical levels, any salient stimulus, and in particular any noxious stimulus, represents a deviation from the current perceptual experience and may thus induce a PE. Since it is thought that the feedforward communication of PEs is enabled by neuronal oscillations at gamma frequencies[59, 141, 168, 169], the well-known pattern of enhanced gamma oscillations within the first 350 ms after noxious stimulus application (Figure 3-3, [41-43]) could be interpreted as signaling such lower-level PEs. From this perspective, pain could be understood as reflecting particularly large PEs, a notion which can be extrapolated to clinical pain states. Chronic pain, for instance, could be partially explained by large PEs that occur when the internal model indicates a harmful event but the objective sensory information does not.

#### *Sensory and expectation effects on pain are served by distinct mechanisms*

“The key finding of our study is that sensory and expectation effects on pain are served by distinct brain mechanisms. Previous fMRI studies have already revealed that sensory and contextual effects on pain are associated with different spatial patterns of brain activity[56, 66, 141, 170]. [...] Our results extend these findings by showing that not only the spatial brain activity patterns serving sensory and contextual effects on pain differ but that these effects are served by fundamentally different neurophysiological mechanisms. [...] These

findings might have implications for the understanding, assessment, and treatment of clinical pain conditions. In acute pain, which is predominantly shaped by sensory information, assessing and modulating local oscillatory brain activity might be appropriate. In contrast, in chronic pain, which is often largely detached from sensory information, inter-regional connectivity might be more informative than local activity”[3].

### *Limitations*

“When interpreting our findings, certain limitations should be considered. First, in our paradigm, the effects of expectations on pain perception were weaker than the effects of stimulus intensity. The lack of expectation effects on local brain activity might therefore reflect the weak expectation effects on pain perception, and other paradigms with stronger expectation effects on perception might well modulate local brain activity. However, the central finding of the present study is not the absolute strength of sensory and expectation effects but that the patterns of sensory and expectation effects on local brain oscillations and brain connectivity fundamentally differ. The strength of perceptual effects might well determine the strength of neurophysiological effects but is unlikely to fundamentally change the difference in the patterns of sensory and expectation effects on brain activity and connectivity. [...] Second, to modulate pain, we manipulated participants’ expectations”[3]. It is unclear whether our observations generalize to contextual influences on pain other than expectations. “Third, we applied brief experimental pain stimuli to healthy human participants. It is unclear whether these findings can be translated to other experimental and clinical types of pain”[3].

## 5.2 Project 2

In this study, we used EEG to investigate intrinsic brain network function in chronic pain. In particular, we aimed to explain inter-individual variability of pain intensity using measures of functional connectivity among four a priori selected intrinsic brain networks which likely play important roles in chronic pain. This hypothesis-driven approach bears the potential to offer new insights into the cerebral pathomechanisms of chronic pain. We based our analyses on resting state EEG data from large cohorts of patients with chronic pain and assessed the robustness of our findings in multiple heterogeneous replication data sets. To the best of our knowledge, our study is the first to test the replicability of associations between EEG-features and chronic pain across multiple replication data sets.

### *Intrinsic brain network connectivity and pain intensity in patients with chronic pain*

In the discovery data set, we found associations between the intensity of pain in patients with chronic back pain and connectivity among intrinsic brain networks at theta and alpha

frequencies. Specifically, pain intensity was positively associated with connectivity at theta frequencies in the DN-SN and DN-FPN connections and inversely related to connectivity at alpha frequencies in the FPN-SN connection. An analogous secondary analysis including all patients from the discovery data set, i.e. those with diverse pain conditions, yielded a similar pattern of effects. Although the effects observed in the discovery data set were not replicable in any of the replication data sets, they aligned with existing literature in two aspects: First, a recent systematic review reported a trend towards increased connectivity at theta frequencies in patients with chronic pain compared to healthy controls[48]. It should be noted that the data from one of the studies discussed in the systematic review partially overlapped with our own discovery data set, which could raise concerns about circular reasoning. However, as the conclusions drawn in the systematic review referred to differences between patient and control cohorts and not to pain variability within patient cohorts, we deem the presented line of argumentation admissible. The second aspect in which our findings align with previous observations are alterations of DN connectivity in general[69-72, 75] and of connectivity between the DN and the SN in particular[73, 74]. A recent meta-analysis found fMRI-based DN-SN connectivity to be reduced in patients with chronic pain compared to healthy controls and patients with other neuropsychiatric disorders[74].

We also assessed whether patterns of intrinsic brain network connectivity could be linked to the severity of chronic pain. Using inter-network connectivity values among all four intrinsic brain networks and at all three frequencies collectively, we constructed multivariate models that could account for some variability of pain ratings in both the cohort of patients with CBP and the cohort of patients with diverse pain conditions. However, these models did not generalize to any of the replication data sets for either cohort.

#### *Intrinsic brain network connectivity and depression in patients with chronic pain*

Since chronic pain and depression have a high comorbidity and the specificity of findings to chronic pain is often discussed[7, 74, 171], we also investigated the relationship between intrinsic brain network connectivity and depression severity scores using the entire discovery data set. All analyses were analogous to those of pain intensity in the larger cohort. Univariate analyses yielded strong evidence for a positive association between depression scores and DN-FPN theta-band connectivity. While this association was not observed in any of the replication data sets, a recent meta-analysis did report increased fMRI-based DN-FPN connectivity in MDD compared to healthy controls[74]. In our study, DN-FPN connectivity was positively associated with both depression and chronic pain severity, potentially suggesting an involvement of DN-FPN connectivity in a

broader transdiagnostic mechanism. In contrast, the above-mentioned meta-analysis described alterations in fMRI-based DN-FPN connectivity as specific to depression[74].

#### *Intrinsic brain network connectivity and age in patients with chronic pain*

Many studies have reported a strong dependence of EEG measures on age[88-90], suggesting that age might be predicted more readily from EEG than severity scores of neuropsychiatric disorders. Therefore, to calibrate the trust both in our methodology and in the quality of the data at hand, we assessed how well age could be predicted across data sets. Univariate analyses yielded evidence for an inverse association between inter-network theta-band connectivity and age in the discovery and up to two replication data sets. In line with a recent study in healthy individuals[90], effects were found predominantly in the theta-band and exclusively in connections not involving the visual network. A multivariate model predicted age in the discovery data set better than chance and descriptively exhibited positive prediction-outcome correlations in all replication data sets. Statistically, a Bayesian analysis indicated moderate evidence for an effect in one and inconclusive evidence in the remaining four replication data sets. A model that was selected and fitted on the combined data, i.e., comprising data from all studies, exhibited strong evidence for an association between predicted and observed age. In conclusion, the partial replicability of effects for age as a dependent variable inspires some trust in both the sensitivity of the methodology and the quality of (at least some) data sets. We chose not to employ the irreproducibility of effects as a criterion for the exclusion of individual data sets as none of the effects observed can be considered established.

#### *Exploratory analyses*

In exploratory analyses, we sought to enhance prediction accuracy and replicability using a more data-driven approach. We constructed multivariate models for pain, depression, and age incorporating an extended set of candidate features. When it came to pain and depression, these models did not notably outperform those developed in the primary and secondary analyses. Anecdotally, in the diverse pain cohort, the descriptive average prediction-outcome correlation in the replication data sets increased from  $r = 0.002$  in the initial to at least  $r = 0.062$  in the exploratory analysis. Only when age was considered as the dependent variable, the approach employed in the exploratory analysis did lead to slightly more noticeable improvements in prediction accuracy. In the diverse pain cohort, the average prediction-outcome correlation increased by at least 20% and at least moderate evidence for an effect was observed in two out of five replication data sets, instead of just one in the initial analysis. Moreover, all models identified in mega analyses combining data from all studies yielded decisive evidence for an association between

predicted and observed age. These results indicate that, although predictions of pain intensity and depression scores could not be improved, the exploratory methodology does, potentially, possess a higher sensitivity.

### *Irreproducibility of effects*

In our analyses, we repeatedly observed associations between pain intensity and intrinsic brain network function in the discovery but not in replication data sets. This lack of replicability can be explained by one or even a combination of the following scenarios:

*Scenario 1: Connectivity among intrinsic brain networks, defined by the Yeo[86] atlas, does not shape pain intensity.* The body of literature indicating an association between intrinsic brain network function and neuropsychiatric disorders[55, 74, 87], however, lets it appear more likely that an association between pain intensity and intrinsic brain network connectivity does exist. Maybe this association becomes detectable if intrinsic brain network definitions other than those provided by the Yeo atlas are considered.

*Scenario 2: Connectivity among intrinsic brain networks does shape pain intensity, but this relationship is not observable using EEG.* The relevance of intrinsic brain network function for neuropsychiatric disorders and pain has been primarily established on the basis of fMRI signals[55, 74, 87]. While the transferability of this association to neurophysiological signals as measured by EEG seems plausible[164], it is not guaranteed. For example, the limited spatial resolution inherent to EEG might simply be insufficient to capture the relevant aspects of intrinsic brain network function. Investigations using multimodal approaches that combine the high spatial resolution of fMRI with the high temporal resolution of EEG could assist in guiding the selection of appropriate methodology in the future.

*Scenario 3: Connectivity among intrinsic brain networks does shape pain intensity and this relation can, in principle, be observed using EEG. However, the specific analytical choices here are inadequate.* For example, we specifically opted for an amplitude-based instead of a phase-based connectivity metric due to its conceptual compatibility with fMRI[164]. Maybe a phase-based connectivity metric is, nevertheless, more suitable. Likewise, we chose to aggregate information on the level of networks using representative time-series because, in this way, a higher spatial specificity is attained compared to an approach that operates on a high-fidelity connectivity matrix. Possibly, a different approach for aggregating information on the level of intrinsic brain networks is more appropriate. Adjusting elementary analysis parameters, such as frequency band boundaries, epoch length, or the choice the of machine learning algorithm in multivariate analyses, might also enhance the methodology's sensitivity.

*Scenario 4: Connectivity among intrinsic brain networks does shape pain intensity and this relation is, in principle, reflected in the considered connectivity features. However, high levels of noise in EEG recordings and/or meta data render a detection of this relationship infeasible.* Here, by noise we mean both common measurement noise and variability stemming from unknown confounding factors. In the present scenario, we face two possibilities: First, effects observed in the discovery data set reflect real, pain-related alterations of intrinsic brain network connectivity. In this case, the irreproducibility of effects in replication data sets could be explained by heterogeneous data acquisition procedures and/or simply by the relatively smaller number of patients (i.e., data points). Therefore, to enhance the likelihood of replicating effects, future multi-center studies should adopt standardized protocols and quality standards both for the acquisition of EEG data and for the clinical assessment of patients. The second possibility to consider in the present scenario is that effects observed in the discovery data set do not reflect real pain-related alterations of intrinsic network connectivity. In this case, noise levels should be reduced further by, e.g., including diverse covariates and/or applying a more refined patient stratification. Given the relevance of circadian rhythms for pain perception[172], incorporating the phase of the circadian cycle is a noteworthy option. Moreover, in light of a recent study which identified clinically relevant neuropsychiatric disorder subtypes via patterns of brain connectivity[82], stratifying patients based on such neural patterns might reveal subtype-specific effects. Inter-individual differences in head shape, brain morphology and even brain organization may also contribute to noise. To mitigate the influence of these inter-individual differences, individual-specific head models and/or data from longitudinal rather than cross-sectional study designs could be employed. Lastly, future studies should further increase the number of participants in both discovery and replication cohorts to enhance the likelihood of discovering consistent associations despite high levels of noise.

### *Synopsis*

All in all, our results offer some evidence for a potential link between chronic pain and intrinsic brain network connectivity. Univariate analyses in the discovery data set, but not in any of the replication data sets, indicated that inter-network connectivity in the theta-band, particularly involving the default network, might be linked to chronic pain severity. Multivariate models could explain some variability in pain intensity within the discovery but not within the replication data sets. An analysis of the potential reasons for the lack of replicability of effects suggests that future research should prioritize minimizing the impact of diverse confounding factors while maximizing both sample sizes and uniformity of data acquisition procedures across research centers.

### 5.3 Conclusion and outlook

The analysis of brain connectivity has proven instrumental in uncovering fundamental mechanisms of brain function and advancing the development of clinical tools to address neuropsychiatric disorders. Leveraging the potential of brain connectivity analyses, this thesis assessed the brain mechanisms of sensory and contextual influences on acute pain and investigated aberrant intrinsic brain network function in chronic pain.

In a first project, the thesis assessed and compared the neural mechanisms of sensory and expectation effects on pain in an experimental setting. Evidence from this project indicates that distinct mechanisms serve sensory and contextual influences on pain. In particular, we found that sensory influences on pain primarily shaped local brain activity. By contrast, expectation effects on pain were more closely associated with inter-regional brain connectivity than local brain activity. These findings provide basic science insights into the brain mechanisms of contextual modulations of pain. Moreover, they offer pathways for the development of novel tools for the assessment and treatment of clinical pain conditions.

Considering the importance of expectations in chronic pain, the findings from the first project suggest that brain dysfunction in chronic pain might be particularly reflected in altered brain connectivity. Therefore, in a second project, the thesis explored the relationship between brain connectivity and pain intensity in patients with chronic pain. Specifically, we investigated connectivity among four intrinsic brain networks that were shown to be implicated in the pathology of various neuropsychiatric disorders. To assess the robustness of findings, we set up a validation strategy involving up to five replication data sets. In the discovery data set, both uni- and multivariate analyses unveiled associations between intrinsic brain network connectivity and pain. However, these findings did not consistently replicate in independent data sets. We outlined several scenarios that could explain the replication failures, leading to the conclusion that future multi-center research initiatives should adhere to unified standards for the acquisition of neural as well as behavioral data. Moreover, variability arising from confounding factors should be mitigated by stratifying patients during study planning and incorporating diverse covariates during data analysis. In terms of general scientific principles, the replication failures underscore the critical importance of employing replication strategies using independent data sets, as was done in our study. Following these steps could improve the chances of identifying consistent effects across data sets, potentially leading to valuable mechanistic insights that could advance the fundamental understanding, the clinical assessment and ultimately the treatment of chronic pain.

## Acknowledgements

There is still much to be learned about the brain and its role in pain. With pain being an elementary aspect of the human experience, what will be discovered will be impactful. This potential motivated me to conduct research in this field, and I am thankful to have had the opportunity to do so. In particular, I want to thank my supervisor, Markus Ploner, for granting me this opportunity and for being so stringent and reliable in his guidance and support. Next to my supervisor, I want to thank the members of my thesis advisory committee. I appreciated the frequent discussions with Joachim Gross and his valuable advice on both technical and conceptual matters. In offering a different perspective, Valentin Riedl helped me broaden my scientific horizon.

Throughout my PhD journey, I was privileged to be part of a collaborative and communicative group. I am grateful for Laura's honest humor and open ear, Elisabeth's constructive meticulousness, and Moritz's decisive contributions in the form of data and opinion. Thanks to Vanessa, Özgün, Cristina, Paul, Henrik, and Martin for your kind support, enthusiastic collaboration, refreshing views, and joyful conference experiences.

I want to thank my family for their loyalty and kindness. Toni with her thoughtful patience and our son Ferdi with his open nature are the most important pillars of my optimism. The commitment and active support of the rest of my family during my doctorate, particularly since Ferdi's birth, were invaluable.



## References

1. Allen, M., et al., *Raincloud plots: a multi-platform tool for robust data visualization*. Wellcome Open Res, 2019. **4**: p. 63.
2. Bradley, C., H. Bastuji, and L. Garcia-Larrea, *Evidence-based source modeling of nociceptive cortical responses: A direct comparison of scalp and intracranial activity in humans*. Hum Brain Mapp, 2017. **38**(12): p. 6083-6095.
3. Bott, F.S., et al., *Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain*. Sci Adv, 2023. **9**(16): p. eadd7572.
4. Melzack, R. and K.L. Casey, *Sensory, motivational, and central control determinants of pain: a new conceptual model in pain*, in *The skin senses*, D.R.J. Kenshalo, Editor. 1968, Charles C. Thomas: Springfield, IL. p. 423-439.
5. Pain, I.A.f.t.S.o. *Pain Terms and Definitions*. 2021 [cited 2023 September 26, 2023]; Available from: <https://www.iasp-pain.org/resources/terminology/>.
6. Melzack, R. and K.L. Casey, *Sensory, motivational and central control determinants of chronic pain: A new conceptual model.*, in *The skin senses: Proceedings of the first International Symposium on the Skin Senses, held at the Florida State University in Tallahassee, Florida*, D.R. Kenshalo, Editor. 1968: Springfield. p. 432.
7. Baliki, M.N. and A.V. Apkarian, *Nociception, Pain, Negative Moods, and Behavior Selection*. Neuron, 2015. **87**(3): p. 474-91.
8. Lewis, T. and E.E. Pochin, *The double pain response of the human skin to a single stimulus*. Clinical Science, 1937. **3**: p. 67-76.
9. Bishop, G.H. and W.M. Landau, *Evidence for a double peripheral pathway for pain*. Science, 1958. **128**: p. 712-713.
10. Basbaum, A.I., et al., *Cellular and molecular mechanisms of pain*. Cell, 2009. **139**(2): p. 267-84.
11. Kuner, R. and H. Flor, *Structural plasticity and reorganisation in chronic pain*. Nat Rev Neurosci, 2017. **18**(2): p. 113.
12. Porreca, F., M.H. Ossipov, and G.F. Gebhart, *Chronic pain and medullary descending facilitation*. Trends Neurosci, 2002. **25**(6): p. 319-25.
13. Ploner, M., et al., *Prestimulus functional connectivity determines pain perception in humans*. Proc Natl Acad Sci U S A, 2010. **107**: p. 355-360.
14. Kuner, R. and T. Kuner, *Cellular Circuits in the Brain and Their Modulation in Acute and Chronic Pain*. Physiol Rev, 2021. **101**(1): p. 213-258.
15. Treede, R.-D., et al., *Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11)*. Pain, 2019. **160**(1): p. 19-27.
16. Merskey, H. and N. Bogduk, *Classification of Chronic Pain: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms*. 1994, Seattle: IASP Press.
17. Breivik, H., et al., *Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment*. Eur J Pain, 2006. **10**(4): p. 287-333.
18. Kennedy, J., et al., *Prevalence of persistent pain in the U.S. adult population: new data from the 2010 national health interview survey*. J Pain, 2014. **15**(10): p. 979-84.

19. Mantyselka, P., et al., *Pain as a reason to visit the doctor: a study in Finnish primary health care*. *Pain*, 2001. **89**(2-3): p. 175-180.
20. Koleva, D., et al., *Pain in primary care: an Italian survey*. *Eur J Public Health*, 2005. **15**(5): p. 475-9.
21. Treede, R.D., et al., *A classification of chronic pain for ICD-11*. *Pain*, 2015. **156**(6): p. 1003-7.
22. Hucho, T. and J.D. Levine, *Signaling pathways in sensitization: toward a nociceptor cell biology*. *Neuron*, 2007. **55**(3): p. 365-76.
23. Latremoliere, A. and C.J. Woolf, *Central sensitization: a generator of pain hypersensitivity by central neural plasticity*. *J Pain*, 2009. **10**(9): p. 895-926.
24. Cohen, S.P., L. Vase, and W.M. Hooten, *Chronic pain: an update on burden, best practices, and new advances*. *Lancet*, 2021. **397**(10289): p. 2082-2097.
25. Meints, S.M. and R.R. Edwards, *Evaluating psychosocial contributions to chronic pain outcomes*. *Prog Neuropsychopharmacol Biol Psychiatry*, 2018. **87**(Pt B): p. 168-182.
26. Fitzcharles, M.A., et al., *Nociplastic pain: towards an understanding of prevalent pain conditions*. *Lancet*, 2021. **397**(10289): p. 2098-2110.
27. Kaiser, U., R.D. Treede, and R. Sabatowski, *Multimodal pain therapy in chronic noncancer pain-gold standard or need for further clarification?* *Pain*, 2017. **158**(10): p. 1853-1859.
28. Heitmann, H., et al., *Longitudinal resting-state electroencephalography in chronic pain patients undergoing interdisciplinary multimodal pain therapy*. *PAIN*, 2022.
29. Tan, L.L. and R. Kuner, *Neocortical circuits in pain and pain relief*. *Nat Rev Neurosci*, 2021. **22**(8): p. 458-471.
30. Brodersen, K.H., et al., *Decoding the perception of pain from fMRI using multivariate pattern analysis*. *Neuroimage*, 2012. **63**(3): p. 1162-70.
31. Wager, T.D., et al., *An fMRI-based neurologic signature of physical pain*. *N Engl J Med*, 2013. **368**(15): p. 1388-97.
32. Rainville, P., et al., *Pain affect encoded in human anterior cingulate but not somatosensory cortex*. *Science*, 1997. **277**(5328): p. 968-71.
33. Segerdahl, A.R., et al., *The dorsal posterior insula subserves a fundamental role in human pain*. *Nat Neurosci*, 2015. **18**(4): p. 499-500.
34. Pernet, C., et al., *Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research*. *Nat Neurosci*, 2020. **23**(12): p. 1473-1483.
35. Mouraux, A., J.M. Guerit, and L. Plaghki, *Non-phase locked electroencephalogram (EEG) responses to CO2 laser skin stimulations may reflect central interactions between A-delta- and C-fibre afferent volleys*. *Clin Neurophysiol*, 2003. **114**(4): p. 710-22.
36. Ploner, M., et al., *Pain suppresses spontaneous brain rhythms*. *Cereb Cortex*, 2006. **16**(4): p. 537-40.
37. May, E.S., et al., *Pre- and post-stimulus alpha activity shows differential modulation with spatial attention during the processing of pain*. *Neuroimage*, 2012. **62**(3): p. 1965-74.
38. Hu, L., et al., *Functional features of nociceptive-induced suppression of alpha band electroencephalographic oscillations*. *J Pain*, 2013. **14**: p. 89-99.

39. Nickel, M.M., et al., *Brain oscillations differentially encode noxious stimulus intensity and pain intensity*. Neuroimage, 2017. **148**: p. 141-147.
40. Nir, R.R., et al., *Tonic pain and continuous EEG: prediction of subjective pain perception by alpha-1 power during stimulation and at rest*. Clin Neurophysiol, 2012. **123**(3): p. 605-12.
41. Gross, J., et al., *Gamma oscillations in human primary somatosensory cortex reflect pain perception*. PLoS Biol, 2007. **5**(5): p. e133.
42. Hauck, M., J. Lorenz, and A.K. Engel, *Attention to painful stimulation enhances gamma-band activity and synchronization in human sensorimotor cortex*. J Neurosci, 2007. **27**(35): p. 9270-7.
43. Zhang, Z.G., et al., *Gamma-band oscillations in the primary somatosensory cortex - a direct and obligatory correlate of subjective pain intensity*. J Neurosci, 2012. **32**: p. 7429-7438.
44. Garcia-Larrea, L., M. Frot, and M. Valeriani, *Brain generators of laser-evoked potentials: from dipoles to functional significance*. Neurophysiol Clin, 2003. **33**(6): p. 279-92.
45. Plaghki, L. and A. Mouraux, *EEG and laser stimulation as tools for pain research*. Curr Opin Investig Drugs, 2005. **6**(1): p. 58-64.
46. May, E.S., et al., *Prefrontal gamma oscillations reflect ongoing pain intensity in chronic back pain patients*. Hum Brain Mapp, 2019.
47. Hashmi, J.A., et al., *Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits*. Brain, 2013. **136**(Pt 9): p. 2751-68.
48. Zebhauser, P.T., V.D. Hohn, and M. Ploner, *Resting-state electroencephalography and magnetoencephalography as biomarkers of chronic pain: a systematic review*. Pain, 2023. **164**(6): p. 1200-1221.
49. Friston, K.J., *Functional and effective connectivity: a review*. Brain Connect, 2011. **1**(1): p. 13-36.
50. Uddin, L.Q., B.T.T. Yeo, and R.N. Spreng, *Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks*. Brain Topogr, 2019. **32**(6): p. 926-942.
51. Bastos, A.M. and J.M. Schoffelen, *A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls*. Front Syst Neurosci, 2016. **9**: p. 175.
52. Friston, K., *Causal modelling and brain connectivity in functional magnetic resonance imaging*. PLoS Biol, 2009. **7**(2): p. e33.
53. Lee, J.J., et al., *A neuroimaging biomarker for sustained experimental and clinical pain*. Nat Med, 2021. **27**(1): p. 174-182.
54. Lindquist, M.A., et al., *Evaluating dynamic bivariate correlations in resting-state fMRI: a comparison study and a new approach*. Neuroimage, 2014. **101**: p. 531-46.
55. Kucyi, A. and K.D. Davis, *The dynamic pain connectome*. Trends Neurosci, 2015. **38**(2): p. 86-95.
56. Woo, C.W., et al., *Quantifying cerebral contributions to pain beyond nociception*. Nat Commun, 2017. **8**: p. 14211.

57. Spisak, T., et al., *Pain-free resting-state functional brain connectivity predicts individual pain sensitivity*. Nat Commun, 2020. **11**(1): p. 187.
58. Ploner, M., C. Sorg, and J. Gross, *Brain Rhythms of Pain*. Trends Cogn Sci, 2017. **21**(2): p. 100-110.
59. Michalareas, G., et al., *Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas*. Neuron, 2016. **89**(2): p. 384-97.
60. van Kerkoerle, T., et al., *Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex*. Proc Natl Acad Sci U S A, 2014. **111**(40): p. 14332-41.
61. Schulz, E., et al., *Prefrontal Gamma Oscillations Encode Tonic Pain in Humans*. Cereb Cortex, 2015. **25**: p. 4407-4414.
62. Baliki, M.N., et al., *Corticostriatal functional connectivity predicts transition to chronic back pain*. Nat Neurosci, 2012. **15**(8): p. 1117-9.
63. Vlaeyen, J.W., G. Crombez, and S.J. Linton, *The fear-avoidance model of pain*. Pain, 2016. **157**(8): p. 1588-9.
64. Cormier, S., et al., *Expectations predict chronic pain treatment outcomes*. Pain, 2016. **157**(2): p. 329-38.
65. Peerdeman, K.J., et al., *Relieving patients' pain with expectation interventions: a meta-analysis*. Pain, 2016. **157**(6): p. 1179-91.
66. Woo, C.W., et al., *Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain*. PLoS Biol, 2015. **13**(1): p. e1002036.
67. Lee, M., et al., *Activation of corticostriatal circuitry relieves chronic neuropathic pain*. J Neurosci, 2015. **35**(13): p. 5247-59.
68. Zhou, H., et al., *Inhibition of the Prefrontal Projection to the Nucleus Accumbens Enhances Pain Sensitivity and Affect*. Front Cell Neurosci, 2018. **12**: p. 240.
69. Napadow, V., et al., *Intrinsic brain connectivity in fibromyalgia is associated with chronic pain intensity*. Arthritis Rheum, 2010. **62**(8): p. 2545-55.
70. Napadow, V., et al., *Decreased intrinsic brain connectivity is associated with reduced clinical pain in fibromyalgia*. Arthritis Rheum, 2012. **64**(7): p. 2398-403.
71. Baliki, M.N., et al., *Functional Reorganization of the Default Mode Network across Chronic Pain Conditions*. PLoS One, 2014. **9**(9): p. e106133.
72. Hsiao, F.J., et al., *Altered insula-default mode network connectivity in fibromyalgia: a resting-state magnetoencephalographic study*. J Headache Pain, 2017. **18**(1): p. 89.
73. Hemington, K.S., et al., *Abnormal cross-network functional connectivity in chronic pain and its association with clinical symptoms*. Brain Struct Funct, 2016. **221**(8): p. 4203-4219.
74. Brandl, F., et al., *Common and specific large-scale brain changes in major depressive disorder, anxiety disorders, and chronic pain: a transdiagnostic multimodal meta-analysis of structural and functional MRI studies*. Neuropsychopharmacology, 2022.
75. Loggia, M.L., et al., *Default mode network connectivity encodes clinical pain: an arterial spin labeling study*. Pain, 2013. **154**(1): p. 24-33.
76. Vanneste, S. and D. De Ridder, *Chronic pain as a brain imbalance between pain input and pain suppression*. Brain Communications, 2021. **3**(1).

77. De Ridder, D. and S. Vanneste, *Occipital Nerve Field Transcranial Direct Current Stimulation Normalizes Imbalance Between Pain Detecting and Pain Inhibitory Pathways in Fibromyalgia*. *Neurotherapeutics*, 2017. **14**(2): p. 484-501.
78. Vanneste, S., et al., *Resting state electrical brain activity and connectivity in fibromyalgia*. *PLoS One*, 2017. **12**(6): p. e0178516.
79. Prichep, L.S., et al., *Exploration of the Pathophysiology of Chronic Pain Using Quantitative EEG Source Localization*. *Clin EEG Neurosci*, 2018. **49**(2): p. 103-113.
80. Ta Dinh, S., et al., *Brain dysfunction in chronic pain patients assessed by resting-state electroencephalography*. *Pain*, 2019. **160**(12): p. 2751-2765.
81. Toll, R.T., et al., *An Electroencephalography Connectomic Profile of Posttraumatic Stress Disorder*. *Am J Psychiatry*, 2020. **177**(3): p. 233-243.
82. Zhang, Y., et al., *Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography*. *Nat Biomed Eng*, 2021. **5**: p. 309-323.
83. Nickel, M.M., et al., *Temporal-spectral signaling of sensory information and expectations in the cerebral processing of pain*. *Proc Natl Acad Sci U S A*, 2022. **119**(1).
84. Strube, A., et al., *The temporal and spectral characteristics of expectations and prediction errors in pain and thermoception*. *Elife*, 2021. **10**.
85. Tiemann, L., et al., *Differential neurophysiological correlates of bottom-up and top-down modulations of pain*. *Pain*, 2015. **156**(2): p. 289-96.
86. Yeo, B.T., et al., *The organization of the human cerebral cortex estimated by intrinsic functional connectivity*. *J Neurophysiol*, 2011. **106**(3): p. 1125-65.
87. Menon, V., *Large-scale brain networks and psychopathology: a unifying triple network model*. *Trends Cogn Sci*, 2011. **15**(10): p. 483-506.
88. Cole, J.H. and K. Franke, *Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers*. *Trends Neurosci*, 2017. **40**(12): p. 681-690.
89. Engemann, D.A., et al., *A reusable benchmark of brain-age prediction from M/EEG resting-state signals*. *Neuroimage*, 2022. **262**: p. 119521.
90. Gil Ávila, C., et al., *DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience*. *bioRxiv*, 2023.
91. Buzsáki, G., *Rhythms of the brain*. 2006, Oxford ; New York: Oxford University Press. xiv, 448 p.
92. da Silva, F.L., *EEG and MEG: Relevance to Neuroscience*. *Neuron*, 2013. **80**(5): p. 1112-1128.
93. Cohen, M.X., *Where Does EEG Come From and What Does It Mean?* *Trends Neurosci*, 2017. **40**(4): p. 208-218.
94. Biasucci, A., B. Franceschiello, and M.M. Murray, *Electro-encephalography*. *Current Biology*, 2019. **29**(3): p. R80-R85.
95. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain*. *J Comp Neurol*, 2009. **513**(5): p. 532-41.
96. Murakami, S. and Y. Okada, *Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals*. *J Physiol*, 2006. **575**(Pt 3): p. 925-36.

97. Nunez, P.L. and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. 2nd edition ed. 2006, New York: Oxford University Press.
98. Fries, P., *A mechanism for cognitive dynamics: neuronal communication through neuronal coherence*. Trends Cogn Sci, 2005. **9**(10): p. 474-80.
99. Fries, P., *Rhythms for Cognition: Communication through Coherence*. Neuron, 2015. **88**(1): p. 220-35.
100. Ferree, T.C., et al., *Scalp electrode impedance, infection risk, and EEG data quality*. Clinical Neurophysiology, 2001. **112**(3): p. 536-544.
101. Kam, J.W.Y., et al., *Systematic comparison between a wireless EEG system with dry electrodes and a wired EEG system with wet electrodes*. Neuroimage, 2019. **184**: p. 119-129.
102. Oostenveld, R., et al., *FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data*. Comput Intell Neurosci, 2011. **2011**: p. 156869.
103. Scherg, M., *Fundamentals of dipole source potential analysis*. Auditory evoked magnetic fields and electric potentials. Advances in audiology, 1990. **6**(40-69): p. 25.
104. Hamalainen, M.S. and R.J. Ilmoniemi, *Interpreting Magnetic-Fields of the Brain - Minimum Norm Estimates*. Medical & Biological Engineering & Computing, 1994. **32**(1): p. 35-42.
105. Hauk, O., *Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data*. Neuroimage, 2004. **21**(4): p. 1612-1621.
106. Tikhonov, A.N., *Solution of incorrectly formulated problems and the regularization method*. Sov Dok, 1963. **4**: p. 1035-1038.
107. Lin, F.H., et al., *Spectral spatiotemporal imaging of cortical oscillations and interactions in the human brain*. Neuroimage, 2004. **23**(2): p. 582-95.
108. Grave de Peralta-Menendez, R. and S.L. Gonzalez-Andino, *A critical analysis of linear inverse solutions to the neuroelectromagnetic inverse problem*. IEEE Trans Biomed Eng, 1998. **45**(4): p. 440-8.
109. Pascual-Marqui, R.D., et al., *Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review*. Methods Find Exp Clin Pharmacol, 2002. **24 Suppl C**: p. 91-5.
110. Van Veen, B.D., et al., *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering*. IEEE Trans Biomed Eng, 1997. **44**(9): p. 867-880.
111. Westner, B.U., et al., *A unified view on beamformers for M/EEG source reconstruction*. Neuroimage, 2022. **246**: p. 118789.
112. Byrd, R.H., J.C. Gilbert, and J. Nocedal, *A trust region method based on interior point techniques for nonlinear programming*. Mathematical programming, 2000. **89**: p. 149-185.
113. Nolte, G., et al., *Identifying true brain interaction from EEG data using the imaginary part of coherency*. Clin Neurophysiol, 2004. **115**(10): p. 2292-307.
114. Stam, C.J., G. Nolte, and A. Daffertshofer, *Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources*. Hum Brain Mapp, 2007. **28**(11): p. 1178-93.

115. Vinck, M., et al., *An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias*. Neuroimage, 2011. **55**(4): p. 1548-65.
116. Hipp, J.F., et al., *Large-scale cortical correlation structure of spontaneous oscillatory activity*. Nat Neurosci, 2012. **15**(6): p. 884-90.
117. Siems, M., et al., *Measuring the cortical correlation structure of spontaneous oscillatory activity with EEG and MEG*. Neuroimage, 2016. **129**: p. 345-55.
118. Tallon-Baudry, C., et al., *Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human*. J Neurosci, 1996. **16**(13): p. 4240-9.
119. Granger, C.W.J., *Investigating causal relations by econometric models and cross-spectral methods*. Econometrica, 1969. **37**: p. 424-438.
120. Wilson, G.T., *The factorization of matricial spectral densities*. SIAM Journal on Applied Mathematics, 1972. **23**(4): p. 420-426.
121. Geweke, J., *Measurement of linear dependence and feedback between multiple time series*. Journal of the American statistical association, 1982. **77**(378): p. 304-313.
122. Baccalá, L.A. and K. Sameshima, *Direct coherence: a tool for exploring functional interactions among brain structures*. Methods for neural ensemble recordings, 1998.
123. Kaminski, M.J. and K.J. Blinowska, *A new method of the description of the information flow in the brain structures*. Biol Cybern, 1991. **65**(3): p. 203-10.
124. Baccala, L.A. and K. Sameshima, *Partial directed coherence: a new concept in neural structure determination*. Biol Cybern, 2001. **84**(6): p. 463-74.
125. Keyzers, C., V. Gazzola, and E.-J. Wagenmakers, *Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence*. Nature Neuroscience, 2020. **23**(7): p. 788-799.
126. Gönen, M., et al., *The Bayesian two-sample t test*. The American Statistician, 2005. **59**(3): p. 252-257.
127. Gronau, Q.F., A. Ly, and E.-J. Wagenmakers, *Informed Bayesian t-tests*. The American Statistician, 2019.
128. Rouder, J.N., et al., *Bayesian t tests for accepting and rejecting the null hypothesis*. Psychon Bull Rev, 2009. **16**(2): p. 225-37.
129. Rouder, J.N., et al., *Default Bayes factors for ANOVA designs*. Journal of mathematical psychology, 2012. **56**(5): p. 356-374.
130. Rouder, J.N. and R.D. Morey, *Default Bayes Factors for Model Selection in Regression*. Multivariate Behav Res, 2012. **47**(6): p. 877-903.
131. Bishop, C.M., *Pattern recognition and machine learning*. Information science and statistics. 2006, New York: Springer. xx, 738 p.
132. Gail, M.H., J.H. Lubin, and L.V. Rubinstein, *Likelihood calculations for matched case-control studies and survival studies with tied death times*. Biometrika, 1981: p. 703-707.
133. Faul, F., et al., *G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences*. Behav Res Methods, 2007. **39**(2): p. 175-91.
134. Correll, J., et al., *Avoid Cohen's 'Small', 'Medium', and 'Large' for Power Analysis*. Trends Cogn Sci, 2020. **24**(3): p. 200-207.

135. Zigmond, A.S. and R.P. Snaith, *The hospital anxiety and depression scale*. Acta Psychiatr Scand, 1983. **67**(6): p. 361-70.
136. Bjelland, I., et al., *The validity of the Hospital Anxiety and Depression Scale. An updated literature review*. J Psychosom Res, 2002. **52**(2): p. 69-77.
137. Hu, L. and G.D. Iannetti, *Neural indicators of perceptual variability of pain across species*. Proc Natl Acad Sci U S A, 2019.
138. Jung, T.P., et al., *Removing electroencephalographic artifacts by blind source separation*. Psychophysiology, 2000. **37**(2): p. 163-78.
139. Van Veen, B.D., et al., *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering*. IEEE Trans Biomed Eng, 1997. **44**(9): p. 867-80.
140. Thomson, D.J., *Spectrum Estimation and Harmonic Analysis*. Proceedings of the IEEE, 1982. **70**: p. 1055-1096.
141. Geuter, S., et al., *Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula*. Elife, 2017. **6**.
142. Egner, T., J.M. Monti, and C. Summerfield, *Expectation and surprise determine neural population responses in the ventral visual stream*. J Neurosci, 2010. **30**(49): p. 16601-8.
143. van Doorn, J., et al., *Bayesian rank-based hypothesis testing for the rank sum test, the signed rank test, and Spearman's rho*. Journal of Applied Statistics, 2020. **47**(16): p. 2984-3006.
144. Garcia-Larrea, L. and R. Peyron, *Pain matrices and neuropathic pain matrices: a review*. Pain, 2013. **154 Suppl 1**: p. S29-43.
145. Tiemann, L., et al., *Distinct patterns of brain activity mediate perceptual and motor and autonomic responses to noxious stimuli*. Nat Commun, 2018. **9**(1): p. 4487.
146. Duan, W., et al., *Reproducibility of power spectrum, functional connectivity and network construction in resting-state EEG*. Journal of Neuroscience Methods, 2021. **348**.
147. Tiemann, L., et al., *Behavioral and neuronal investigations of hypervigilance in patients with fibromyalgia syndrome*. PLoS One, 2012. **7**: p. e35068.
148. May, E.S., et al., *Dynamics of brain function in chronic pain patients assessed by microstate analysis of resting-state electroencephalography*. Pain, 2021((May) Department of Neurology, Technical University of Munich (TUM), Germany TUM-Neuroimaging Center, Germany Center for Interdisciplinary Pain Medicine, School of Medicine, TUM, Munich, Germany).
149. Khanna, A., et al., *Microstates in resting-state EEG: current status and future directions*. Neurosci Biobehav Rev, 2015. **49**: p. 105-13.
150. Zebhauser, P.T., V.D. Hohn, and M. Ploner, *Resting-state electroencephalography and magnetoencephalography as biomarkers of chronic pain: a systematic review*. Pain, 2022.
151. Day, M.A., et al., *Change in Brain Oscillations as a Mechanism of Mindfulness-Meditation, Cognitive Therapy, and Mindfulness-Based Cognitive Therapy for Chronic Low Back Pain*. Pain Med, 2021. **22**(8): p. 1804-1813.
152. Adhia, D.B., et al., *Infraslow Neurofeedback Training Alters Effective Connectivity in Individuals with Chronic Low Back Pain: A Secondary Analysis of a Pilot Randomized Placebo-Controlled Study*. Brain Sci, 2022. **12**(11).



153. Ashar, Y.K., et al., *Effect of Pain Reprocessing Therapy vs Placebo and Usual Care for Patients With Chronic Back Pain: A Randomized Clinical Trial*. JAMA Psychiatry, 2022. **79**(1): p. 13-23.
154. Topaz, L.S., et al., *Electroencephalography functional connectivity-A biomarker for painful polyneuropathy*. Eur J Neurol, 2023. **30**(1): p. 204-214.
155. Pascal, M.M.V., et al., *DOLORisk: study protocol for a multi-centre observational study to understand the risk factors and determinants of neuropathic pain*. Wellcome Open Res, 2018. **3**: p. 63.
156. Jensen, M.P., et al., *Pain-related beliefs, cognitive processes, and electroencephalography band power as predictors and mediators of the effects of psychological chronic pain interventions*. Pain, 2021. **162**(7): p. 2036-2050.
157. Pernet, C.R., et al., *From BIDS-Formatted EEG Data to Sensor-Space Group Results: A Fully Reproducible Workflow With EEGLAB and LIMO EEG*. Front Neurosci, 2020. **14**: p. 610388.
158. Pernet, C.R., et al., *EEG-BIDS, an extension to the brain imaging data structure for electroencephalography*. Sci Data, 2019. **6**(1): p. 103.
159. Pion-Tonachini, L., K. Kreutz-Delgado, and S. Makeig, *ICLabel: An automated electroencephalographic independent component classifier, dataset, and website*. Neuroimage, 2019. **198**: p. 181-197.
160. Mullen, T.R., et al., *Real-Time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG*. Ieee Transactions on Biomedical Engineering, 2015. **62**(11): p. 2553-2567.
161. Kothe, C.A.E. and T.-P. Jung, *Artifact removal techniques with signal reconstruction*. 2016, Google Patents.
162. Pellegrini, F., et al., *Identifying good practices for detecting inter-regional linear functional connectivity from EEG*. Neuroimage, 2023. **277**: p. 120218.
163. Bastos, A.M. and J.M. Schoffelen, *A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls*. Front Syst Neurosci, 2015. **9**: p. 175.
164. Engel, A.K., et al., *Intrinsic coupling modes: multiscale interactions in ongoing brain activity*. Neuron, 2013. **80**(4): p. 867-86.
165. Albu, S. and M.W. Meagher, *Expectation of nocebo hyperalgesia affects EEG alpha-activity*. Int J Psychophysiol, 2016. **109**: p. 147-152.
166. Huneke, N.T., et al., *Experimental placebo analgesia changes resting-state alpha oscillations*. PLoS One, 2013. **8**(10): p. e78278.
167. de Lange, F.P., M. Heilbron, and P. Kok, *How Do Expectations Shape Perception?* Trends Cogn Sci, 2018. **22**(9): p. 764-779.
168. Arnal, L.H. and A.L. Giraud, *Cortical oscillations and sensory predictions*. Trends Cogn Sci, 2012. **16**(7): p. 390-8.
169. Bastos, A.M., et al., *Canonical microcircuits for predictive coding*. Neuron, 2012. **76**(4): p. 695-711.
170. Zunhammer, M., et al., *Placebo Effects on the Neurologic Pain Signature: A Meta-analysis of Individual Participant Functional Magnetic Resonance Imaging Data*. JAMA Neurol, 2018. **75**(11): p. 1321-1330.
171. Spisak, T., *Statistical quantification of confounding bias in predictive modelling*. arXiv preprint arXiv:2111.00814, 2021.

172. Daguét, I., et al., *Circadian rhythmicity of pain sensitivity in humans*. *Brain*, 2022. **145**(9): p. 3225-3235.

## Publications

**Bott, F. S.**, Nickel, M. M., Hohn, V. D., May, E. S., Gil Ávila, C., Tiemann, L., Gross, J. and Ploner, M. Local brain oscillations and interregional connectivity differentially serve sensory and expectation effects on pain. *Science Advances*, doi:10.1126/sciadv.add7572 (2023).

Gil Ávila, C., **Bott, F. S.**, Tiemann, L., Hohn, V.D., May, E. S., Nickel, M.M., Zebhauser, P. T., Gross, J. and Ploner, M., 2023. DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience. *Scientific data*, doi:10.1038/s41597-023-02525-0 (2023).

Hohn, V. D., **Bott, F. S.**, May, E. S., Tiemann, L., Fritzen, C., Nickel, M. M., Gil Ávila, C., Ploner, M. How do alpha oscillations shape the perception of pain? – An EEG-based neurofeedback study. *OSF*, doi:10.17605/osf.io/qbkj2 (2022). (stage 1 manuscript with in-principal acceptance at *PLoS Biol*).

**Bott, F. S.** and Gee, M. W. A strong form based moving Kriging collocation method for the numerical solution of partial differential equations with mixed boundary conditions. *International Journal for Numerical Methods in Engineering*, doi:10.1002/nme.6553 (2021)



## A Appendix

### A.1 Simulative assessment of representative time series methods

In the methods section of the main text, several variants of the method for computing representative time series have been proposed. To determine the optimal strategy, a simulation experiment was conducted. In this experiment, I simulated ground truth signals at 400 locations in the brain, denoted as  $\mathbf{G} \in \mathbb{R}^{400 \times n}$ . Simulated ground truth signals differed between all 400 sources. However, signals associated with sources belonging to the same network, possessed a common component, i.e.,

$$\mathbf{G}^A = c_{\text{mix}} \mathbf{G}^{A,\text{common}} + (1 - c_{\text{mix}}) \mathbf{G}^{A,\text{indi}},$$

where  $\mathbf{G}^A$  refers to those lines in  $\mathbf{G}$  that correspond to netA. The rows of  $\mathbf{G}^{A,\text{common}}$  are identical, while the rows of  $\mathbf{G}^{A,\text{indi}}$  differ. The parameter  $c_{\text{mix}}$  controls the degree to which activity in a network is determined by the common component. To obtain the simulated source-reconstructed signals  $\mathbf{S} \in \mathbb{R}^{400 \times n}$ , I applied a spatial blurring filter to the ground truth signals. Specifically, to blur the ground truth signals, I employed the so-called resolution matrix  $\mathbf{M} \in \mathbb{R}^{400 \times 400}$  which results from the composition of the inverse and forward models:

$$\mathbf{S} = \mathbf{M}\mathbf{G} = \mathbf{F}^T(\mathbf{F}\mathbf{F}^T + \lambda\mathbf{I})^{-1}\mathbf{F}\mathbf{G}$$

The forward model is a linear mapping given by the lead field matrix  $\mathbf{F}$  and the inverse model is a linear mapping given by the minimum norm spatial filter  $\mathbf{F}^T(\mathbf{F}\mathbf{F}^T + \lambda\mathbf{I})^{-1}$ . Herein, in accordance with recommendations in the literature, the regularization parameter  $\lambda$  was set to  $\lambda = \text{tr}(\mathbf{F}\mathbf{F}^T)/25$ .

Ultimately, our objective was to compute the amplitude envelope correlation (AEC) between the representative time series. For two signals  $\mathbf{r}^A$  and  $\mathbf{r}^B$ , the AEC is defined as

$$\text{AEC}(\mathbf{r}^A, \mathbf{r}^B) = \frac{1}{2} [\text{corr}(p(\mathbf{r}^A), p(\mathbf{r}^{B\perp A})) + \text{corr}(p(\mathbf{r}^B), p(\mathbf{r}^{A\perp B}))]$$

where  $p(\mathbf{r})$  denotes the envelope of  $\mathbf{r}$ . Further,  $\mathbf{r}^{B\perp A}$  refers to the signal  $\mathbf{r}^B$  phase-orthogonalized w.r.t.  $\mathbf{r}^A$  and  $\mathbf{r}^{A\perp B}$  refers to the signal  $\mathbf{r}^A$  phase-orthogonalized w.r.t.  $\mathbf{r}^B$ . Hence, there are four signals involved in the computation of the AEC:  $\mathbf{r}^A$ ,  $\mathbf{r}^B$ ,  $\mathbf{r}^{A\perp B}$ , and  $\mathbf{r}^{B\perp A}$ .

To quantitatively compare the different methods for estimating representative time series, I compute both the minimum and average of the following values:

- $v(\mathbf{r}^A, \mathbf{G}^{A,\text{common}}) = \text{"fraction of variance of } \mathbf{G}^{A,\text{common}} \text{ explained by } \mathbf{r}^A\text{"}$

- $v(\mathbf{r}^{A\perp B}, \mathbf{G}^{A,\text{common}})$
- $v(\mathbf{r}^B, \mathbf{G}^{B,\text{common}})$
- $v(\mathbf{r}^{B\perp A}, \mathbf{G}^{B,\text{common}})$

Henceforth, the minimum and average of these values are referred to as minimum and average explained variance scores, respectively.

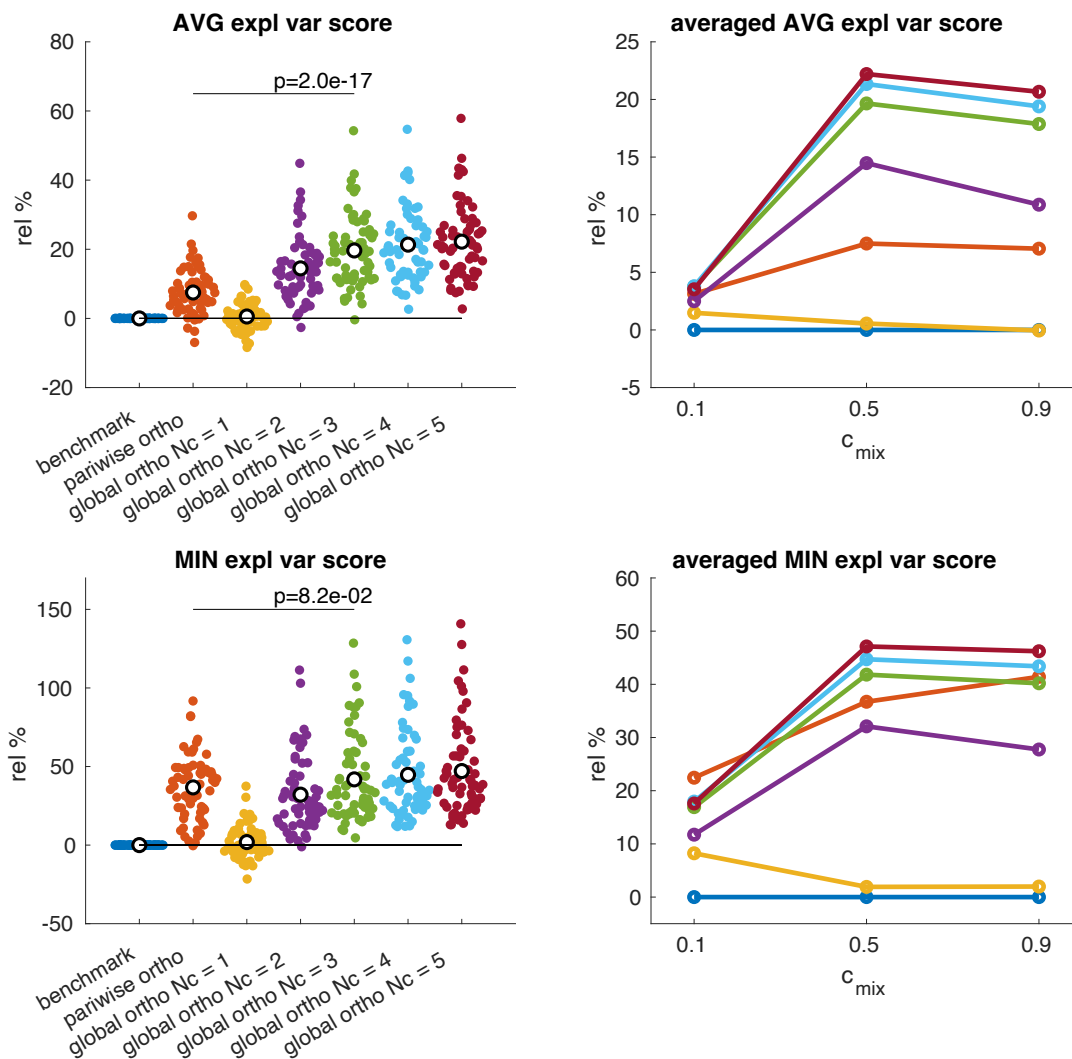
### *Results and conclusion*

I have computed the explained variance scores for several values of the mixing parameter  $c_{\text{mix}}$  (0.1, 0.5, 0.9). Visualizations of the network definitions are provided in Figure 4-1.

Figure 0-1 shows the explained variance scores of the different methods averaged across all connections and for 64 random repetitions of the experiment. The values presented in the figure are relative scores, defined as the ratio between the score obtained from the method of interest and the score obtained from the benchmark method, which involves standard PCA without any orthogonalization. The values in Figure 0-1a and Figure 0-1b correspond to the case  $c_{\text{mix}} = 0.5$ . Scores averaged across repetitions for all tested values of  $c_{\text{mix}}$  are provided in Figure 0-1c and Figure 0-1d. The results demonstrate that both *pairwise* and *global* orthogonalization can significantly enhance the variance of the *common* ground truth signal that is explained by the representative time series. Among the methods tested, the best performing method is the global orthogonalization with  $N_c = 5$ . This method exhibits an average improvement in the average and minimum explained variance scores of roughly 20% and 50%, respectively. The variant of this method, with  $N_c = 3$ , performs similarly well but offers the practically important advantage of being computationally less demanding. Furthermore, the pairwise orthogonalization entails a notable improvement as well, with an increase in average and minimum explained variance scores of roughly 7% and 38%, respectively.

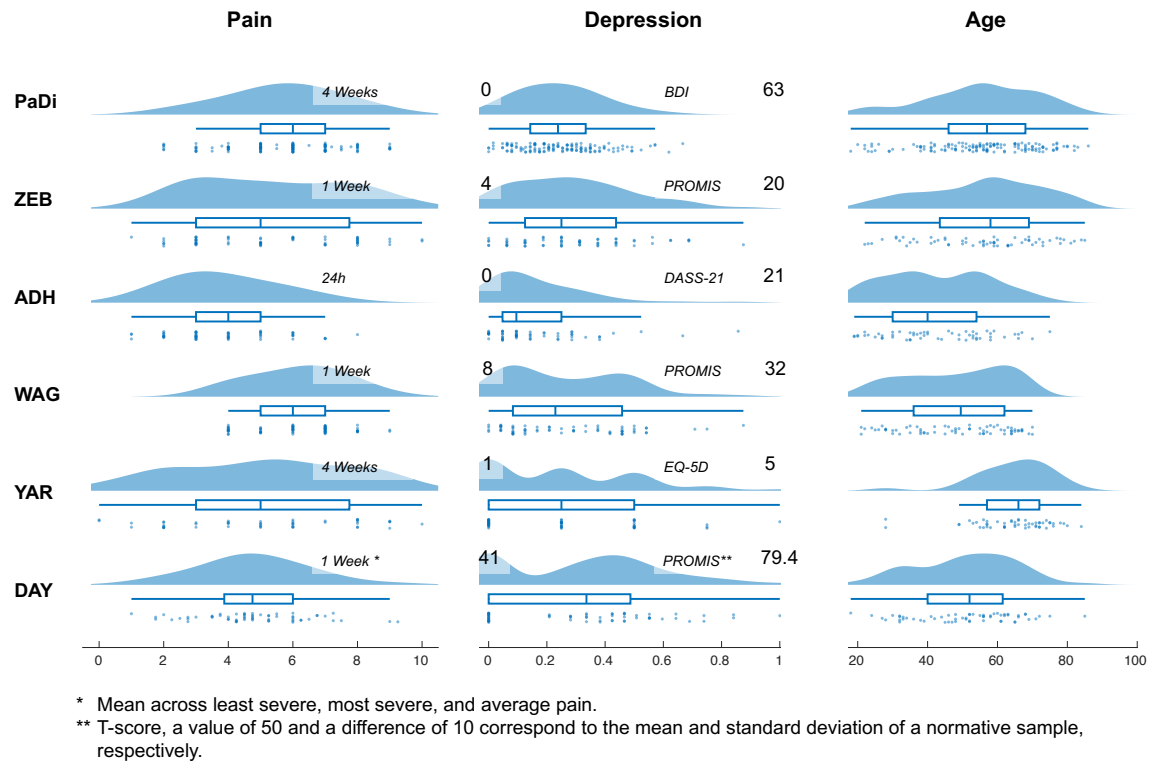
I also evaluated the computation time of the different methods: To compute one pair of orthogonalized time series using the pairwise orthogonalization method, it takes roughly 0.1 s. To compute a representative time series for an individual network using the global orthogonalization approach, it takes about 0.1 s for  $N_c = 1$ , 0.3s  $N_c = 2$ , 1.5s  $N_c = 3$ , 4s  $N_c = 4$ , and 10s  $N_c = 5$ . These times have been obtained for the Yeo 7 network definition.

I conclude that the primary method for extracting representative time series should be *global orthogonalization* with  $N_c = 3$ .



**Figure 0-1:** Relative explained variance scores achieved by the different methods for the YEO7 network definition. Panels a) and b) show results for the case  $c_{mix} = 0.5$  and for 64 random repetitions of the numerical experiment. Black circles indicate the mean of scores across repetitions. Panels c) and d) show averaged explained variance scores across repetitions for three values of the mixing parameter  $c_{mix}$ .

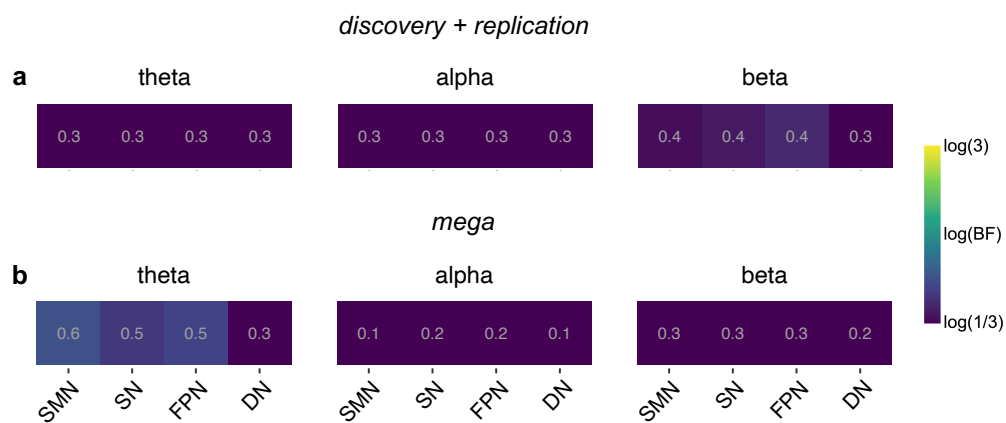
## A.2 Distributions of dependent variables



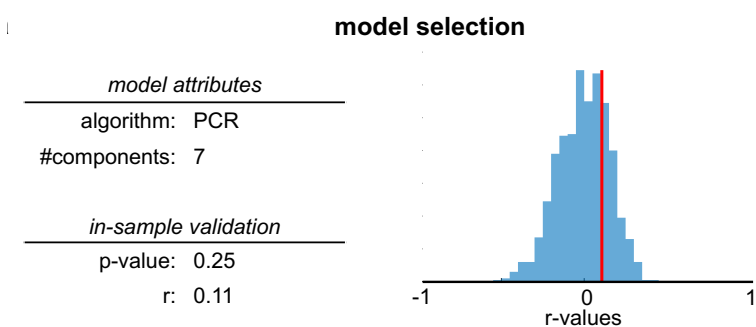
**Figure 0-2:** Distributions of pain intensity, depression scores and age for the individual data sets. For the distributions of pain intensity, it is additionally specified for each data set to which time period (prior to assessment) the ratings refer. For the distributions of depression scores, it is additionally specified which questionnaire was used for the assessment. Moreover, the distributions of depression scores are rescaled to the interval [0, 1]. The original rating scale boundaries are indicated for each data set. The rating scale boundaries do not necessarily correspond to the actual range of the data.



### A.3 Chronic back pain and activity of intrinsic brain networks



**Figure 0-4:** Relation between pain intensity and activity in individual intrinsic brain networks in patients with chronic back pain. The color coding in both subfigures is the same as in Figure 4-4. (a) Results from the discovery+replication approach. (b) Results from the mega analysis.



**Figure 0-3:** Relation between pain intensity and patterns of intrinsic brain network activity in patients with chronic back pain. Explanations of the individual tables can be found in Figure 4-5.