This thesis brings together two formerly distinct topics: The research about neural foundations of interaction and machine-mediated interaction, particularly using BMIs. Both have been active fields of research for years and there are countless contributions on these topics. To my knowledge, however, studying machine-mediated interaction settings as a special case of human interaction is new to neuro interaction research. In particular the use of BMIs, which are the most direct connection between humans and machines, as a technique in neuro interaction research is unprecedented.

This thesis aims to pave a way for this new approach. The first experiment (HExMInE) aims to verify the general feasibility of the approach, in particular to evaluate whether or not neural correlates of interaction (in particular hyper-connectivity) still occur when interaction is machine-mediated. The second experiment (iCusss) then is intended to showcase the potential of this approach in a fully featured (BMI) machine-mediated interaction experiment.

The thorough exploration of this approach's potential is undoubtedly by far too ambitious for a single thesis. The same holds for the (thorough) investigation of the impact of machine-mediation on human interaction and its neural correlates. As this research advances, though, the results will undoubtedly shed light on important aspects of human interaction and human-machine interaction and can be expected to have major impact on the design of future machine-mediation technology. Therefore, this new approach addresses highly relevant research goals.

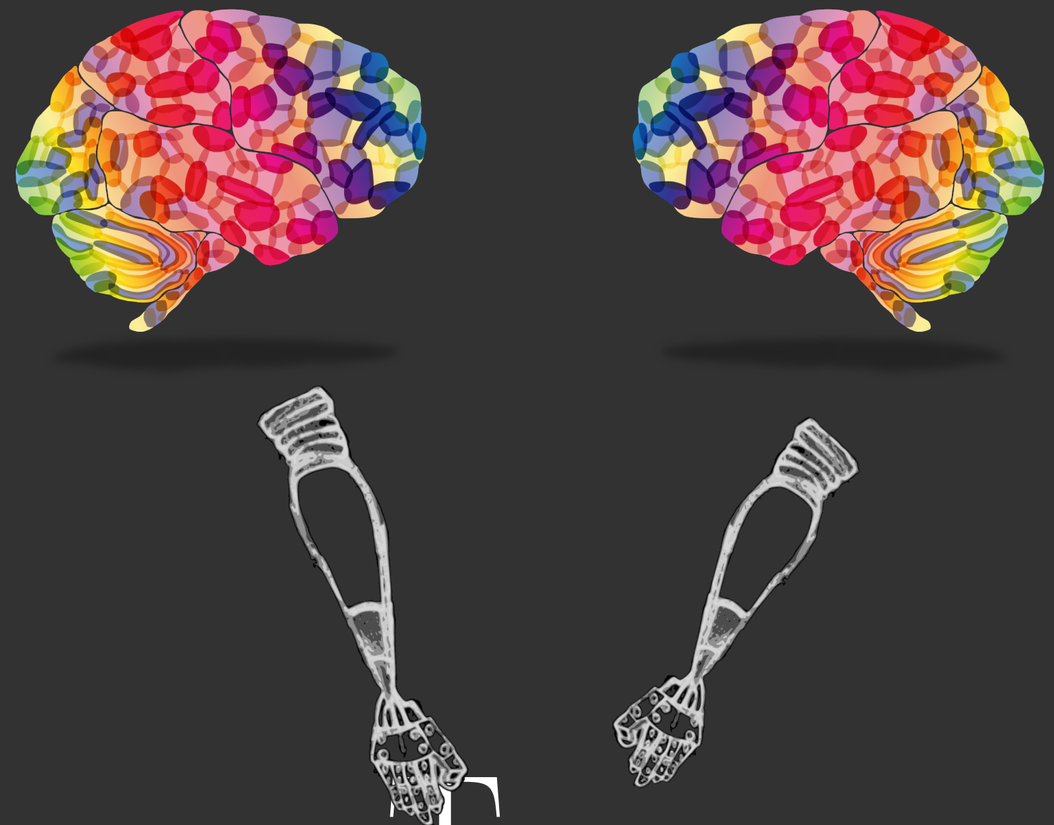
Besides these rather general aspects, both experiments address specific research questions, relevant for neuro interaction research: In the HExMInE experiment I compare neural connectivity during interaction and during solo action of one participant. In the iCusss experiment I compare neural connectivity during cooperation with independent, concurrent action. The main method for this evaluation is hyper-scanning and -analysis aiming for neural connectivity – within a participant and across participants. This connectivity analysis is conducted on different frequency bands, addressing several of the standard neural rhythms in human EEG and thereby contributing to the interpretation of their roles. These roles have been another active research topic for years and still evidences for new facets regarding their interpretation/function is being accumulated.

Additionally, different scientifically relevant topics are addressed as a side-effect when pursuing my main research goals: For the HExMInE experiment a new type of training for the mental strategy of Motor Imagery, often employed in BMIs, is tested. For the two experiments two different robots which are diametrical in many aspects are employed. In particular one is highly anthropomorphic while the other is the exact opposite. Up to now, there are relatively few publications on BMI usage of more than one participant simultaneously. This thesis contributes indirectly to the field of multi-user BMIs by demonstrating its feasibility even for very complex settings. In the course of the PhD project the development of the UBiCI BMI software framework was advanced. And finally, I employ two different BAPs for the BMI control, both of which are correlated with some function vital for interaction (P300 $\leftrightarrow$ attention and ERD $\leftrightarrow$ motor co-representation) with the intention to allow space for interesting side effects to occur in the neural recordings.

This thesis deviates from two different, well explored paths of research at once, converging in and pioneering a brand new direction of neuro interaction research. I hope this path will lead research to the neural foundations of human interaction from a new, different angle, allowing an illumination of new aspects of what Schilbach et al called "the dark matter of social neuroscience".

More Than One

Nils Hachmeister – Dissertation



More Than One

Observing Human Brains in Uncommon Interaction Scenarios

Nils Hachmeister

Dissertation zur Erlangung des akademischen Grades
Doktor der Ingenieurswissenschaften (Dr.-Ing.)
Vorgelegt dem Promotionsausschuss der Technischen Fakultät der Universität Bielefeld

Gedruckt auf alterungsbeständigem Papier °° ISO 9706

More Than One

Observing Human Brains in Uncommon Interaction Scenarios

Nils Hachmeister 2018

Dissertation zur Erlangung des akademischen Grades
Doktor der Ingenieurswissenschaften (Dr.-Ing.)
Vorgelegt dem Promotionsausschuss der Technischen Fakultät der Universität Bielefeld

1. Gutachter: Prof. Dr. Helge Ritter
2. Gutachter: Prof. Dr. Stefan Kopp

To my kids, who picked me up when experiments went wrong.

To my wife, who kept me level when experiments went right.

You have power over your mind - not outside events.
Realise this, and you will find strength.

Marcus Aurelius

Contents

List of Figures

1. Introduction

Interaction – this word evokes associations like communication, synergies, cooperation, reciprocal action and influence. Interaction happens in various contexts: In chemistry molecules, in physics elementary particles or cosmic basic forces and in statistics independent and dependent variables interact. Musicians interact when they are improvising. In Biology different organisms interact. Whenever two or more entities exhibit an influence on one another, we speak of an interaction. The term interaction is ubiquitous in sciences. One predominant interpretation in everyday language, is the interaction between humans. This type of interaction is often particularly complex and difficult to predict, analyse and model, but it is also of high relevance to each human's everyday life.

Human-human interaction can have various forms. The ability of humans for cooperation, including communication on different levels simultaneously, is one of the key abilities of the human species. In particular, the ability to cooperate to achieve a common goal has presumably contributed much to the survival of the human species and to its ascension.

Human interaction has been a focus of research for a very long time. Interaction was studied on a behavioural level and, recently, by recording neural activity during human interaction, aiming for the neural mechanisms from which this rich social behaviour emerges. Those mechanisms have been regarded as "the dark matter of social neuroscience" (Schilbach et al., 2013) and this thesis aims to contribute to the illumination of this dark matter by studying neural recordings during a special form of human interaction, namely when interaction is machine-mediated.

I will start with some general thoughts about interaction and related topics. First, I will dwell a little on (human) *interaction* itself in section 1.1, below. Next I will talk about the term *actions* as the components of an interaction and, subsequently, some terms vital to the experimental settings of this thesis in section 1.2 on page 2. Then I will discuss some general ideas pursued in neuro interaction research in section 1.3 on page 3. I will conclude this chapter by discussing the term of *mediation*, *machine-mediation* and *Brain Machine Interfaces (BMIs)*, as one way to implement machine-mediation, in section 1.4 on page 3.

1.1. Thoughts About Interaction

Let me start with the following definition of (human) interaction:

> *An interaction is a mutual process between more than one individual, consisting of a set of actions carried out by these individuals. For each individual involved, also called interactant, the actions he/she takes could potentially influence some other interactant during the course of the interaction. An interaction has a well-defined beginning and ending and thus a finite duration.*

Hence, interaction emerges from a series of discrete units: The interactants' actions. I would like to stress here, that the definition demands a mutual effect. I.e. every interactant needs to be able to affect the other interactants at some point during the interaction. Pure receivers would not be considered part of the interaction.

Furthermore a "set of actions" is demanded, which makes no assertion about the temporal ordering. It would be possible that only one interactant can perform an action at a time (this might be called sequential interaction) or that each interactant might perform actions at any point in time (which might be called concurrent or parallel interaction) or any other mode in-between.

Interactants always pursue a goal in the interaction e.g. the construction of an object, information interchange or social manipulation of other interactants. Each action taken during the interaction serves this goal. Either directly or indirectly by serving a sub-goal. The interactants can pursue individual goals or a common goal. If a common goal is pursued, two very basic forms of human interaction can be distinguished: cooperation and competition. The criterion for this distinction is basically whether all or just one interactant benefits from achieving the goal.

Of these two forms of interaction, I would consider cooperation to be scientifically more interesting. The ability to cooperate has been developed by relatively few species during the course of evolution. For those species which have this ability, it has proven to be an immense advantage in evolutionary competition. And no other species has brought this ability as far as humans did. From a neuro interaction research perspective, cooperating partners can be expected to align and adopt to one another to a degree not to be expected during competition.

Particularly when cooperating, individuals often rely on their rich, multi-modal communication capabilities. Indeed, their unrivalled and most sophisticated ability to communicate with one another might be the central feature why human cooperation outperforms that of other species. However, the study designs of this thesis do not allow direct communication among the participants, in order to control a multi-layered process and with the intention to gradually release this constraint in future research. Hence, I do not want to comment too much on this topic.

Due to its high relevance and intriguing nature, interaction and its special form cooperation have been a field of active research for a long time. Most of the time this research was conducted at a behavioural level and this kind of research yielded many remarkable results and will, and should, be continued. However, in the last decades more and more sophisticated recording techniques to track neural activity in different situation, conditions and states have become available. Among other applications, these techniques allow us to tack the neural activity of an individual while he/she is engaged in interaction and thereby to track down the neural foundations of one of our own key-abilities.

Recently, the notion that human interaction cannot be fully understood from recordings of single participants has received much attention in the community (Dumas, 2011; Sebanz et al., 2006; Schilbach et al., 2013). Its supporters claim that humans engaged in an interaction form a tightly interwoven system which is more than the sum of its parts and which, thus, cannot be understood if not considered as a whole. An adequate analogy might be a clockwork which is, at least, much harder to understand if one just examines the individual gears and feathers rather than observing the interplay of these parts. Consequentially, researchers have started to record the neural activity of all participants engaged in an interaction (rather than just one), to synchronise the recorded data and to analyse these data sets as a whole.

In this thesis I contribute to this research by recoding and analysing neural activity from (all) participants engaged in an interaction under special constraints: Machine-mediated interaction in a shared space.

1.2. Action and Interaction

At several occasions I used the term "action" in the above section, as a discrete, elementary block of an interaction. I will now try a definition of the term action:

> *An action is the finite process of an individual influencing the external world in order to achieve a goal.*

Any action needs a goal whose implementation it serves. Obviously, goals can be defined hierarchically. E.g. the goal of placing a banana in a bowl could be divided into the sub-goals to picking the banana and placing it. This results in an equivalent hierarchy for actions.

Some aspects of the definition of interaction can be re-encounter here. The individual from the definition of interaction occurs in the definition of an action in two places: Once, of course, as the individual which performs an action. Furthermore the definition of interaction demands that actions must (potentially) have an influence on other individuals. Other individuals in an interaction are, from the perspective of the actor, part of the external world which is being influenced by an action.[1]

Interaction can take many different forms: Some random examples might be, a conversation between two people, some people repairing a car or a parent playing blocks with a child. In many cases the actions taken by the interactant are implemented in a confined space. This leads to the next definition I want propose:

> *If more than one individual implements actions within a confined, physical space, these individual act in a shared space. This space needs to be sufficiently constricted that the actions of each individual are likely to occasionally have an effect on the outcome of other individuals' actions*

[1] The aphorism from Marcus Aurelius I cited in the beginning of this thesis, states that we have no power over outside events, meaning the external world. While this is certainly true we can still influence the external world, though.

While interaction does not require a shared space (e.g. a conversation), individuals acting in a shared space necessarily interact. In a shared space, actions potentially have effects on the outcome of other individuals' actions and thus, indirectly, on other individuals.

1.3. Neural Correlates of Interaction

This thesis aims to contribute to illuminate the neural foundations of human interaction. By analysing neural recordings made during interaction, I hope to find neural correlates of this interaction. But what can we expect to find? In the end, neural recordings are multi-variant signals. Different approaches exist to estimate the degree to which two or more signals are similar. One of the most renown such estimates is coherence. Various others exist and some of the more popular estimates are discussed in section 5.6 on page 47.

Some of these are undirected just describing commonalities between the signals. Others can also identify a temporal delay in these commonalities and thereby identify which signal exercised an influence on which of the other signals. Or, put differently, they identify the information flow between the different signals.

Such estimates have been applied to neural recordings for decades. The brain consists of groups of neurons on different scales. These groups of neurons (which can easily encompass several millions of neurons) influence one another. They interact. And from this interaction emerges human behaviour and capabilities. Identifying the information flow between these neural groups yielded significant scientific insights.

Recently these techniques have been applied to the synchronised recordings of multiple participants aiming to identify neural information flow between participants, e.g. Astolfi et al. (2011b); Saito et al. (2010); Dumas (2011); Dodel et al. (2011); Sänger et al. (2013); Lindenberger et al. (2009); Babiloni et al. (2007c) and many more. I will adopt this approach and take it into a new direction by combining it with Brain-Machine Interfacing (BMI) technology to study machine-mediated human interaction.

1.4. Mediation, Machine-Mediation and BMIs

I already mentioned that this thesis will study a special form of interaction: machine-mediated interaction. Let me, therefore, start by discussing the term mediation. The term mediation has two distinct common meaning: Mediation can be a social process with the aim of solving a conflict and the act of a medium relaying something between entities (particularly individuals). For this thesis, when talking about mediation I always refer to the latter meaning of the term. Let me try the following definition of this type of mediation, linking to the previous definitions:

> *If the influence one individual exercises on another individual is relayed by a medium, this is called mediation.*

The medium which relays the influence is arbitrary. A letter can be a medium as well as television. The term machine-mediation consequentially refers to mediation for which the medium is a machine. A machine here can be any technical device: from a video chat over a telepresence robot to an elaborate robotic system designed to relay actions remotely (e.g. the TAO system, compare section 3.1 on page 23).

Machine-mediation in practice always has certain implications. Especially when it comes to communication that is being mediated, machine-mediation imposes some limitations. Despite such limitations, machine-mediation of communication has become very common nowadays. Telephone, mobile communication, text chat and video-telephony are commonplace. Telepresence robots are mostly still research projects, but no longer a toy application. All of these, however, limit the natural communication capabilities of humans (e.g. the lack of an overtone in text chat or the lack of facial expression and body language on the phone). Communication modalities which are unavailable or hampered make communication between humans error prone, which in turn brought about behavioural adaptations of humans using these technologies to prevent or compensate for such errors (e.g. use of phonetic alphabet and the use of the word "over" in CB-radio transmission or the use of emojis in text chat). Furthermore, some forms of machine-mediated communication limit the amount of information which can be transferred at a time (e.g. in text chat hardly anybody can type as fast as he/she can talk). Again adaptations occurred (e.g. by omitting less important information or by the use of abbreviations as *cu, fyi* or *imho*).

The term of machine-mediation encompasses settings such as video calls and normal phone calls. These are what I would call "transparent mediators". Such a machine mediator is intended to transmit one or

several of the humans' natural modalities for interaction in a maximal unobtrusive (transparent) manner. The mediating machine is intended *not* to play a role on its own in the interaction. Phone calls and video calls would be such transparent mediators, as they transmit voice (and live image) while the mediator itself recedes into the background.

I would like to propose a second concept, narrower than machine-mediation: A triadic human-machine-human setting. In such a triadic setting, which is always a machine-mediation setting, the machine needs to play a role in the interaction, i.e. the machine is *not* a transparent mediator. I assume that triadic settings might be an utile concept for (neural) interaction research, as humans tend to anthropomorphise machines in an interaction. This might even be more the case when the machine mediates another interactant's actions, intentions, communication, etc. Furthermore, a triadic setting allows a degree of control about the information flow between participants to a level way beyond what would be possible with a transparent mediator, particularly when participants are engaged in remote interaction (compare section 7.2.1 in page 80).

The term "information flow" is meant to encompass any type of information transfer between participants, volitional and involitional, verbal and non-verbal, by gesture, body-posture, spatial positioning and even mediated by actions. I will refer to such a behavioural type of information flow between interacting individuals as the interaction level information flow, particularly in contrast to a neural information flow which can be found only by adequate mathematical analysis of neural recordings.

In the studies of this thesis I will aim to achieve a high level of control over the interaction level information flow and, therefore, stick to settings imposing rather strong limitations to it. This might be relaxed step by step in future research (compare section 9.3 on page 115). The experiments which are presented in this thesis induce limitations on two levels:

First, no direct communication is possible. The mutual intentions of the participants are encoded in the actions they perform. That is all the information interchange available to the participants.

Second, all actions were executed via a special device called Brain Machine Interface (BMI), which accesses the user's intentions directly at the brain and which brings its very own challenges and opportunities. For the use of a BMI the individual needs to volitionally produce a certain brain activity. Neural recordings are analysed by a dedicated software autonomously and are translated into actions to be executed by a machine. These actions can be as simple as displaying a text on a computer screen or as complex and elaborate as a bipedal, humanoid robot walking or a robotic hand picking and placing an object.

The reason for the use of BMI machine control for the machine-mediation are three-fold: i.) It is the most direct link of a human to a machine possible, which I consider advantageous for my human-machine-human triadic interaction settings. ii.) I aimed for a medium which is unfamiliar to virtually all participants and whose usage requires participants to adopt to the medium iii.) The brain activity patterns (BAPs) employed in BMIs are often correlated with some neural function relevant for human interaction. In this thesis I use P300 which is highly correlated to attention/expectation and motor-ERDs which are correlated with motor-planning, motor imagery and motor co-representation of others' actions. Hence, I use neural correlates relevant for human interaction to implement the machine-mediation of that interaction. I will introduce these BAPs in some detail in section 2.2.5 on page 17.

In summary, I embark to study the neural foundations of human interaction and, in particular, cooperation by studying a special case of it, namely, machine-mediated interaction. To implement the machine-mediation, I use the most direct link between a human and a machine possible. Now that I have established some basic terms and sketched the idea of this thesis, in the next chapter I will go over some fundamental knowledge required to conduct this kind of research.

2. Foundations

Here I will discuss the foundations upon which this thesis is based. For those familiar with neurology, neural recordings and BMIs this chapter will contain little new information. I will start with generally accepted knowledge about the brain and its functions in section 2.1 below.

BMIs are a technique, central for this thesis. I will describe this technique and give an overview of the field of BMIs, prerequisites for constructing a BMI and the central components of a BMI in section 2.2 on page 8.

2.1. The Human Brain in a Nutshell

In this section I will give some details about the human brain. However, this thesis is from the field of computational neuro science and I will, thus, limit this section to a brief overview over aspects relevant for this thesis. The human brain is a complex thing, whose thorough discussion would fill many books. And many of these still need to be written.

Anyhow, in section 2.1.1 I will briefly describe the most elemental functional unit of the brain: A neuron. I will then quickly move on to the brain and its architecture, dividing it into different units, regions and layers in section 2.1.2 on page 6. The different regions of the brain are known to interact and exchange information. This is discussed in section 2.1.3 on page 7. How such information flow can be technically measured will not be a topic yet, but will be discussed later in this thesis (in section 5.6 on page 47). Finally, in recent years, information flow between brains of individual engaged in some type of interaction has been demonstrated. This is the central approach, which this thesis aims to advance one more step. This and the methods used to demonstrate this information flow between brains are discussed in section 2.1.4 on page 7.

Having established how EEG signals are generated in the brain I will then discuss some components which can typically be found in the human EEG in section 2.2.5 on page 15.

2.1.1. From Neurons to Neural Signal

Central to this thesis is the technique of Electroencephalography or EEG. This method basically measures the electrical field of the brain, which is subject to constant change, giving away quite some information about the brain's (or more precisely the cortex's) activity. The most elemental functional unit producing this activity is a single nerve cell or neuron. When neurons are stimulated, they can undergo a process of depolarisation and repolarization which produces a (change in a) small electrical dipole. This behaviour is called spiking of firing. The process of how the electrical fields of vast groups of neurons sums up to the electrical field measured with an EEG is complex and the subject of active research and discussion. Part of the problem is that even the term neuron is a broad cover term, as up to today more than 10,000 different types of neurons have been identified in the human brain.

Anyhow, all these types of neurons share some characteristics in their structure and behaviour. Figure 2.1 shows a sketch of a typical neuron. Four main parts can be identified: The neuron's body (soma) is the widest part and contains the cell nucleus. The axon is a large cylindrical continuation of the neurons body. It can be as long as one meter or as short as a fraction of a millimetre in humans. The axon splits towards its ending into sever branches, the axon terminals. These terminals are attached to other neurons. The parts of another neurons to which the axon terminals are attached are called dendrites and are surrounding the cell body in a web-like structure. The interface between one neuron's axon terminal and another neuron's dendrite is called a synapse. The (modulation of the) sensitivity of synapses is a primary means to the brains ability to adapt/learn. Neurons are also interconnected by pure dendritic connections (Sanei, 2013). However, as the dendrites are much shorter than the axon, dendritic connections are strictly local.

The synapse allows interaction between neurons. When a neuron fires, that impulse travels along the axon. When it reaches the axon's terminals, a messenger substance is released into the synapse which gives an impulse to the succeeding neuron. This impulse can have an inhibitory (decreasing the stimulation level) or and excitatory (increasing the stimulation level) effect on the receiving neuron. When the level of stimulation

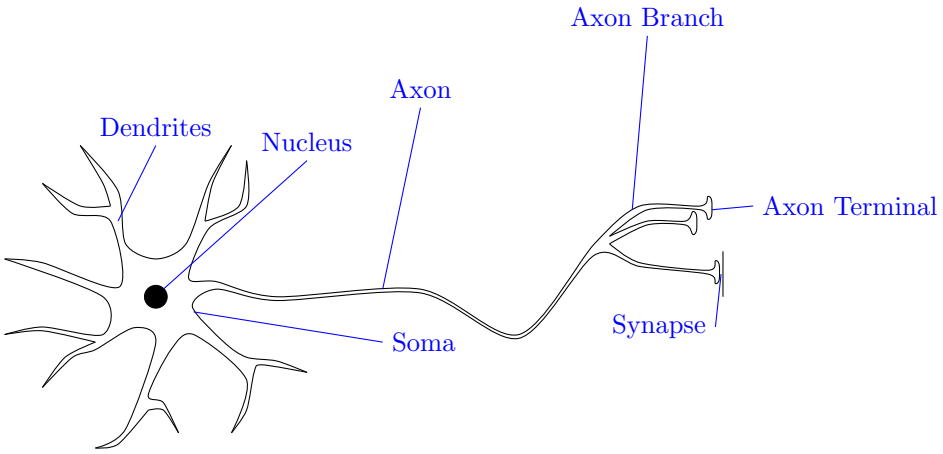


Figure 2.1.: A schematic representation of a neuron (vastly simplified) identifying important parts of the neuron. Figure adapted from Sanei (2013).

of a neuron reaches a certain threshold, the neuron fires. After firing, there is a phase called refractory period during which a neuron cannot fire again, independently of its stimulation level.

In malls and public places one occasionally finds water plays which consist of a series of water containers. Water flows into different containers. When a container is filled to a certain level it tilts, releasing its water to subsequent containers before returning to an upright position. This is often used as an analogy to neurons firing and interacting.

While this analogy can be quite helpful to get an idea of how neurons interact, it has several sever shortcomings. For one, the stimulation level of a neuron declines over time, when no stimuli arrive from other neurons. This is as if someone had drilled small holes into the bottoms the containers of the water play. Second, in the water play there are no inhibitory interactions. When a container tilts it always contributes to other containers filling and eventually tilting. Third, due to gravity, there are no recurrent interactions in the water play while networks of neurons within the human brain are highly recurrent.

The electrical potential of the dipole a neuron produces, lies in the range of several dozens millivolt. As the structure generating this potential is very tiny the resulting electrical field can only be measured in the direct neighbourhood of the neuron. To measure neural activity from further away, particularly outside the skull as in EEG, millions of neurons must be engaged in the same firing pattern. And even than these neurons must be arranged in a way, that their respective electric field can sum up (open-field) rather than eliminating one another (closed-field).

The human brain is surrounded by different layers of tissue, bone and skull. The way from the individual firing of sets of neurons to a signal measurable by EEG devices is a field of active study. One aspect is the interaction of the different neuro-electrical fields involved. Starting at the level of individual neurons, ascending in several steps to the level of cortical areas whose activity can indeed be accessed by EEG. Another aspect is the influence of the different layers of tissue the signal has to traverse before it reaches the EEG sensor (electrode). Third, electrical fields generally propagate omni-directionally. As a consequence the signals of different sources of neural activity mix at each of the electrodes (Volume Conduction).

The complexity of these processes is one of the main reason why, for many questions of computational neuro science we still lack a solid ground truth. There are various computational models which try to simulate this process based on bio-physiological plausibility assumptions. Some of these are more sophisticated than others. In the end, however, none of these models can provide us with a ground truth, because we are lacking a ground truth to verify them. There is no way one can pull oneself out of a swamp.

2.1.2. Brain Structure

The brain is not one homogeneous unit – not structurally/anatomically and much less functionally. It is roughly separated in three main layers (from outer to inner): The cerebrum, the cerebellum and the brain stem. The cerebrum itself has a surface layer which is highly convoluted, as in, it has ridges and depressions (sculci). This surface layer is called the cerebral cortex (often simply called cortex). When performing EEG recordings virtually all of the recorded signal is generated by the cortex. To record activity taking place

deeper within the brain, other recording techniques are required (such as fMRI, fNIRS, etc.).

When looking at the brain's evolution one can observe a tendency that deeper layers tend to be older in terms of evolution. As a consequence, subconscious and unconscious functions of the brain tend to be localised in the deeper layers, whereas conscious perception, thinking, motor control and so forth are all at least in parts located in the cortex. This makes EEG signals, although they only encompass the activity of a relatively small fraction of the brain (speaking in terms as mass or volume), so valuable for diagnostic and research purposes. The cortex itself is again divided into different regions which perform specialised task and are loosely coupled.

In recent years, systems have been developed, which allow individuals to learn to influence certain cortex areas. This can, in some cases be used to train the usage of a BMI (e.g. ERD-based BMIs, see section 6.3 on page 63). Therefore, a feedback-loop is established: Brain activity is measured, certain parameters of this activity are derived (e.g. the level of activity of a certain cortical area) and then presented in some suitable form to the participant. The individual is then instructed to try to influence the presented values using the certain mental strategy. The feedback-loop of participant - recording - analysis - presentation allows the participant to learn to influence brain activity in a way impossible without such techniques. Recently first therapeutic usages of such systems have been proposed, e.g. by Egetemeir et al. (2011).

2.1.3. Neural Information Flow

Information is propagated throughout the brain over synaptic connections (see section 2.1.1 on page 5). This type of information propagation can be as fast as $100m/s$ (Sanei, 2013). I discussed in section 2.1.2 on page 6 that the brain is divided into many different functional and structural units. This is true in the scale of cortical regions as the primary motor cortex, but also in smaller scales of several thousands highly interconnected neurons which are connected to other neuron-populations with relatively few connections.[1] Considering the potential length of a neuron's axon (up to one meter in humans) neuron population and cortical areas could potentially influence one another over virtually any distance in the brain.

It is commonly assumed today, that the understanding of the brain's functions heavily depends on the understanding of the interplay of the different brain areas. This is done on many levels, from recording single neurons, over recording of neuron populations in few square centimetres of cortex (inter-cranial recordings) up to macro-scale recording methods as EEG.

When measuring EEG signals one records a, potentially non-linear, mixture of signals from different neural sources. The signal generated by any given neural source will contribute, in theory, to all and practically to at least several of the signals recorded with different electrodes. This effect is called Volume Conduction. In volume conduction, signals propagate at the speed of light and considering a.) the rather small distances between any given neuronal source and any electrode and b.) the sampling rates at which EEG is typically recorded (seldom above $1024Hz$) volume conduction can be considered to happen instantaneously. Volume conduction is a physical effect and not related to brain function.

If between two channels in an EEG recording an information flow can be demonstrated, one has to verify whether or not this information flow is synaptic in nature or to be attributed to volume conduction. Particularly, if it can be demonstrated that a propagation of information has a non-zero delay, volume conduction can be ruled out.

A number of computational methods for analysing neural information flow have been developed and I will discuss some of the more prominent methods in section 5.6 on page 47.

2.1.4. Hyper-Information Flow

The study of the interaction of areas in a human's brain has been conducted for several decades now. More recently researches started to apply these methods to synchronised data, recorded from *all* participants engaged in an interaction.

The ability of humans to interact and cooperate is unprecedented in nature and engineering. When humans interact or even more cooperate to solve a common task, information flows between these individuals over several modalities: the interaction level information flow . Although the neural activity of the different individuals can, of course, not influences one another directly, the influence is there – mediated by body language, facial expression, orientation in the room, the space in which the interaction takes place and many more.

[1] In mathematical network analysis, this is a typical attribute of small-world networks.

Could such coupling of neural activity between participants be such tight that a connectivity can be shown with methods previously used to identify connectivity between neural sources within the same brain, given that the delay over synaptic connections is usually in the sub-second range? In recent years, various studies have demonstrated that this is indeed the case.

E.g. Sänger et al. (2013) showed this for pairs of guitar players engaged in a duet. This might, however, be a very favourable scenario: Music is a process which is strictly paced and requires actions timed as precise as a few dozens of milliseconds. Furthermore, the music itself could potentially sever as a pace-maker for the inter-individual synchronisation of neural sources. However, even for less strictly paced processes such as pilots during a flight (Astolfi et al., 2012) or a game of cards (Babiloni et al., 2007b) this type of synchronisation has been shown. Connectivity between neural signals recorded from different brains is commonly referred to as hyper-connectivity.

A prerequisite for these types of study is, however, that EEG recordings of the different participants are synchronised with millisecond precision. Furthermore, as any form of interaction usually requires movement, one has to account for motor artefacts (induced by the electrical field generated by the muscles on activation).

The recording of neuronal activity from more than one person simultaneously in the same setting is referred to as Hyper-Scanning. The joint analysis is usually referred to as Hyper-Analysis.

2.2. Brain Machine Interfacing

In this section I am going to discuss the technique of Brain Machine Interfacing, which plays a central role in the studies of this thesis. However, this method is not the subject of study itself. I will start with a definition of the term Brain Machine Interface (BMI) in section 2.2.1 below. Then I will give an overview over different types of BMIs in section 2.2.2 on page 9. Afterwards I will discuss the general structure of BMIs and introduce essential components of any BMI in section 2.2.3 on page 10. Finally I will talk about renown activities patterns found in human brain activity recordings, which can be used for BMIs and about their advantages and disadvantages.

2.2.1. Definition of a BMI

Although Brain Machine Interfaces (BMIs) are not in the centre of focus of this thesis, they play an important role and pervade the entire thesis. What is a BMI anyway? An early definition is given by Wolpaw et al. (2002).

> "A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles."

Wolpaw et al actually refer to Brain *Computer* Interface (BCI) rather than Brain Machine Interface. What is the difference between a Brain Machine Interface and a Brain Computer Interface? These terms are used almost synonymously in the literature, really. The term "machine" is somewhat wider than the term "computer", hence one could argue that BMI is a wider term than BCI. Using a wider notion of the term computer, encompassing devices such as robots and, generally, any device containing a computational unit, the terms BMI and BCI almost coincide. For this thesis I will generally use the term Brain Machine Interface as I belief it is less ambiguous.

Central in the definition of Wolpaw et al. (2002) is that the "normal output pathways of peripheral nerves and muscles" are circumvented. If one circumvents the normal pathways and still wants to bring messages or commands the individual intends to send volitionally to the external world, one needs to access these messages and commands before they reach the peripheral nerves and muscles.

A second point to note about this definition is that the individual in this definition has an active role. It "sends" the "messages and commands" implying that it does so volitionally. This is supported by the system being characterised as a "communication system". In 2002 this basically encompassed all such systems presented so far. Meanwhile, a variety of systems have been proposed which access brain activity which cannot not be actively controlled by the individual, but still represents its mental or emotional state. These interfaces would not be included in Wolpaw's definition. Today, a modern definition of BMIs needs to include systems which are not actively controlled by the participant (passive BMIs).

Based on this discussion and the definition by Wolpaw et al. (2002), I propose a new definition of the term Brain Machine Interface:

> *A BMI is a technical system which allows brain activity of an individual to influences the external world in a systematic way while circumventing the brain's natural output pathways of peripheral nerves and muscles.*

When dropping terms like "send", "communication", "message" and "command" I allowed passive BMIs to fit into my definition. However, the type of influence became somewhat arbitrary. For instance, one could correlate the activation of a standard LED to the raw, per-sample values of an EEG and call that a BMI. Although surly influenced by brain activity, the LEDs behaviour would be chaotic and lacking any true meaning and utility. Hence, by adding the demand that the influence of the brain activity on the external world shall be systematic, I rule out such systems, while still allowing for a wide range of couplings between the brain activity and its influence on the external world.

2.2.2. Taxonomy

Many different BMIs have been proposed in the past decades, varying in neural recording technique, the BAP(s) employed, the device controlled and more. Beside this, different methods for real-time data classification have been proposed, evaluated and compared. Consequentially researchers began to group BMIs into different categories:

When a BMI depends on the cooperation of the individual this BMI is called active. I.e. the individual actively tries to evoke a certain brain activity, either by attending to stimuli or by performing some mental strategy. The aim of the individual is to communicate their intentions. A classical example would be a spelling system in which the individual can choose one out of a matrix of letters and symbols by concentrating on it (Farwell and Donchin, 1988).

In contrast, passive BMIs do not require any actions on part of the individual. These BMIs monitor brain activity that is not volitionally influenced, but still reflect the user's state. An example would be a system, which monitors sleepiness of the individual (by analysing its α-rhythm, compare 2.2.5 on page 15) and then influence lighting of the room or music selection based on the results. These systems allow the brain activity to influence the real world in a systematic way. They are, thus, covered by my definition of a BMI. But they do not constitute a type of communication in which the individual sends messages nor commands and, thus, these systems would not be covered by Wolpaw's et al. definition of BCI.

Another distinction commonly applied to BMIs divides systems into dependent and independent BMIs. An independent BMI requires no physical action of the participant. One example would be a BMI which classifies the brain response to imagined movements of the individual (as long as the movements are not actually executed). A passive BMI is actually always independent as no active participation, neither mental nor physical, is required from the individual.

A dependent BMIs, consequentially, requires that the individual is able to perform physical movements to use the BMI. An example would be a SSVEP-based BMI, which requires individuals to fixate certain stimuli and therefore eye-movements. The fact that dependent BMIs require physical movements does not contradict my definition of a BMI. The brain activity influences the real world and that brain activity is accessed directly at the brain, regardless of the need of physical movement to evoke the brain activity. Here it is, of course, important that the physical movement are not actually used to infer the individual's choice (as, for instance, an eye tracking system would), but that this information is derived from the brain activity recordings only.

Furthermore, one can attribute BMIs either to be synchronous or asynchronous. Synchronous BMIs require some time stamps. It will classify data time-locked to these. E.g. When a participant is asked to focus on one out of a set of stimuli in order to choose the next letter in a spelling system, this would be a synchronous BMI. In contrast an asynchronous BMI does not use temporal cues, but classifies the data continuously. The previously used example of a BMI controlling lighting in relation to participant's sleepiness would be asynchronous.

Adding to the above, I would like to propose two other categorisations:

The first is Choice BMIs in contrast to Continuous BMIs. To my knowledge it has not yet been proposed as such, although the general distinction is probably familiar to anyone in the field of BMIs. The difference lies in the type of the result. If the result is one out of a set of discrete options this would constitute a choice

BMI. Hence, the influence on the external world that could be exhibited with a choice BMI is discrete as well. The previously mentioned spelling BMI system is a choice BMI.

Whereas a continuous BMI outputs a continuous value/influences the real world in some continuous manner. If the BMI dims the lighting of a room seamlessly, depending on the individual's level of sleepiness, this would be a continuous BMI.

Second, I would like to suggest an usage-dependent categorisation of BMIs which again, to my knowledge, has not yet been suggested so far:

Communication: The BMI allows to share information. This can happen pure virtually e.g. by sending Twitter messages or by displaying letters on a screen.[2] But also the utterance of spelled words and phrases using a speech synthesiser could be imagined.

Representation: This category encompasses all types of BMIs aiming to physically represent their users. Although information is typically shared with these types of BMIs as well, their purpose goes above and beyond pure communication. By their design they are supposed to be an interaction partner for others, physically representing the BMI-operator. Examples would be the robots performing gestures and facial expressions, a telepresence robot or even a virtual face on a screen combining verbal (synthesised) communication with facial expressions.

Manipulation: Any BMI allowing the user to manipulate physical objects would fall into that category. This could be fetching objects, picking and placing objects and so forth.

Navigation: When the BMI's primary purpose is navigation of some kind, it would fall into this category. This could be a BMI-controlled wheelchair, the navigation through a virtual mace or the navigation of a camera mounted on wheels, allowing the user, for instance, to visit a museum. This category is, in some aspects, pretty close to the "Representation" category. The main difference here is whether or not the controlled device is meant to represent the operator in interaction with others or if its primary function is navigation.

Creativity/Gaming: Any BMI whose purpose is to allow for the expression of creativity or to play some type of game would fall into this category. Examples would be brain painting devices (Münßinger et al., 2010), devices which allow to compose music via a BMI[3] or the BMI control of a car racing game[4].

In some cases BMIs might plausibly fit into more than one of these categories, especially when two or more purposes are deliberately combined into one single system. However, I believe that any BMI should fit into at least one of these categories and that this categorisation might prove utile.

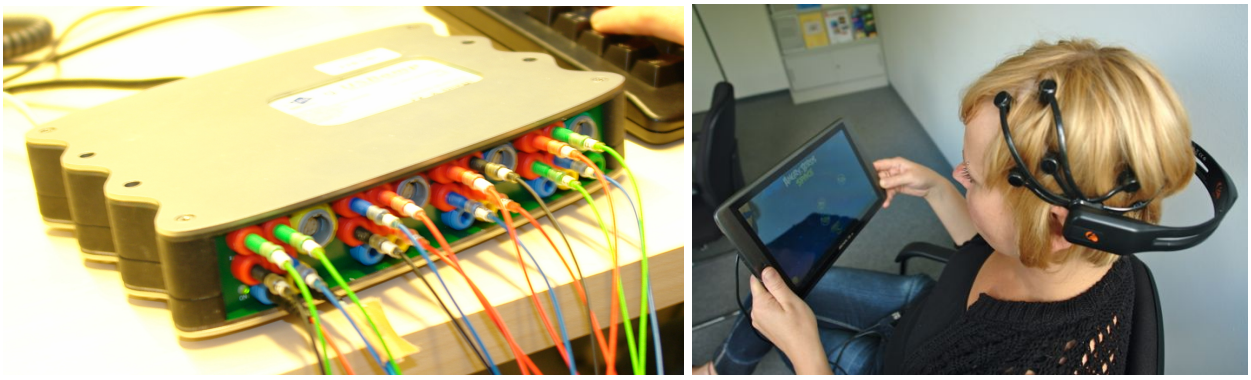
2.2.3. Components of a BMI

Closing in on the inner working of a generic BMI, I will next list its essential components. To my knowledge Wolpaw et al. (2002) were the first to identify and list such components. In the following I will give an overview over the vital components of a BMI. Figure 2.2 shows a generic scheme of these components.

Data Acquisition: As a BMI must not rely on peripheral nerves and muscle, it needs to scan the activity of the brain one way or the other. There is a set of different techniques for neural recordings and most have already been employed for BMIs. EEG is the most popular among these techniques. Many BMIs also require time stamps on the data recorded - in some cases with millisecond precision. More details on data recording in the context of BMIs can be found on page 11.

Stimulus Presentation: For active BMIs some type of Stimulus Presentation and/or Feedback is often required. While not all BMIs relay on a stimulus presentation to evoke BAPs, it is virtually always necessary to inform the participant about the general state of the BMI. Additionally one generally wants to inform the participant about the outcome of the latest classification/decision. For passive BMIs a feedback or stimulus presentation is not mandatory. I will give some more details on stimulus presentations on page 13.

[2] This has been a student project in our group.
[3] Another student project in our research group.
[4] This was a demo system of our research group, actually combining BMI and eye tracking control.



(a) A GUSBAmp EEG device as it has been used for the studies of this thesis.

(b) The Epoc EEG device from EMotive. Wireless and with saline electrodes.

Figure 2.3.: Two different EEG devices: One for clinical studies, the other mobile and low-cost.

Classification: Recording brain activity results in a multi-variant data stream in which the desired information is embedded/encoded. The extraction of this information and, thereby, the translation from raw neurological data to machine commands in real-time is a challenging process, regardless of the neural recording technology employed. For EEG, the main challenges are two-fold: First, the information is encoded in different recording channels due to volume conduction. Second, the EEG recordings have a disadvantageous signal-to-noise ratio, whereby noise encompasses both: external noise from electrical fields e.g. emitted by electrical devices in the surrounding and biological noise by muscle activity and irrelevant/"background" brain activity. Some more discussion about real-time classification of EEG data can be found on page 14.

Device: In the end, the extracted information shall be used in some way to influence the external world. Examples for this influence can be as simple as a display on a computer screen or as complex as a pick-and-place action executed by a robot. What ever it is, there is some physical device involved in implementing the designated influence on the external world. The role of the device controlled by a BMI is discussed in some more detail on page 15.

Of course, some concrete BMIs might need additional components depending on their design and setting, but these are the essential BMI-components.

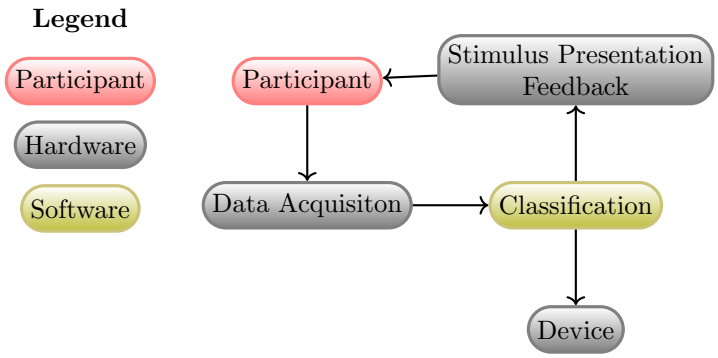


Figure 2.2.: The general components of a BMI system. The classiciation is typically done by an elaborate digital signal processing pipeline and, thus, in software.

Data Acquisition

Brain activity recording can be done, nowadays, in a variety of ways, each having its own properties, advantages and disadvantages. I will first discuss some of the common categories and features relevant when comparing recording techniques. Afterwards I will briefly present several common neural recording techniques and discuss their properties along these lines.

Recording techniques can first and foremost be divided into two categories: Those, which do require surgical intervention before recording (invasive) and those which do not (non-invasive). Another distinguishing feature is whether or not a recording techniques suffers from volume conductions, i.e. whether the activity of different neural sources mix into the signals recorded with different sensors. Furthermore different techniques suffer from noise to different degrees and from different sources of noise.

Next, recording techniques offer very different resolutions, both temporal and spatial. A good spatial resolution allows to locate the source of a given activity pretty accurately, which is of great help when

interpreting the brain activity. A high temporal resolution is important to access all types of fast paced brain responses e.g. P300 and brain rhythms (compare section 2.2.5 on page 15.

Another aspect, especially relevant in the practical application, is the degree of obtrusiveness and, thereby, the degree of freedom the participant retains during recording. In human centred sciences, the influence the method of data acquisition has on the participants and, thereby, on the recorded data is crucial. E.g. camera recordings of participants during the experiment will already have an influence on the participant. For neural recordings, this is even more relevant, as there is no neural recording technique which is unobtrusive. However, some techniques are more obtrusive than others or in different ways. When performing behavioural studies including neural recording, the degree of freedom of movement a recoding technique leaves the participant might be of particular importance.

Another distinguishing feature is, which brain areas can be accessed with the recording technique. Some techniques, particularly EEG, are capable to access only the activity which is produces on the surface of the brain (the cortex), while others can also access activity of the deep brain. As virtually all of the high level brain functions in the human brain are located in the cortex, recording techniques which cannot access the deep brain activity are still well suited for many BMI applications and neuro interaction/cognition research.

Finally, I want to introduce a category of recording techniques which record Blood Oxygen Level Dependent (BOLD) signals. When a certain area of the brain becomes more active, its energy consumption increases. As oxygen from the bloodstream is required to produce this energy, this affects the oxygen level in the blood in that brain region. BOLD signals depend on a metabolic process and, thus, have a low temporal resolution. However, some recording techniques (particularly fMRI) offer a superior spatial resolution.

Next I will discuss some of the most common neural recording techniques, concluding with a broader discussion of the technique of EEG which is being used throughout this thesis.

One prominent invasive recording technique is the Electro-Cortiogram (ECoG), which records the electrical activity of the brain (just as EEG), but the electrodes are surgically implanted into the skull and placed directly on the cortex. This is done with an electrode array which is only a couple of square centimetres large and thus provides data with a high spatial and temporal resolution, but which is limited to a pretty small area of the cortex. This technique is commonly used in medicine on patients which suffer from sever epileptic seizures, in order to identify the source of the seizures, in preparation for later surgical removal of the brain tissue from which the seizures originate.

fMRI stands for functional Magneto Resonance Imaging. A fMRI scanner produces a strong, oscillating magnetic field and measures how the body and in particular the brain tissue interacts with this magnetic field. For brain scanning it is of special interest that a fMRI scanner can give an estimation of the oxygen level in the blood. As discussed before these BOLD signals suffer from a poor temporal resolution, but fMRI scanners provide an excellent spacial resolution. Depending on the device in question, as small as one cubic centimetre and below. They can access any part of the brain including deep brain regions. However, fMRI scanners are very expensive both in acquisition and maintenance. They require specialised personal (a radiologist) and the patient needs to remain motionless during recording.

fNIRS is a technique in which light in the near-infrared spectrum is emitted into the skull. Light in that range can traverse the skull and interacts with the blood oxygen. NIRS, therefore, record a BOLD signal. It does not suffer from volume conduction (as EEG does) nor from muscle artefacts and has a very good spatial resolution. It is generally more expensive than EEG, but still by orders of magnitudes cheaper than fMRI. It is rather obtrusive, as rather bulky probes need to be attached to the participants head.

When neurons fire, they produce (a change in) a dipole, both electrical and magnetic. The sum of these magnetic dipoles can be measured in the magnetic field at the outside of the skull, similar to how EEG measures the electric field. This technique is called Magneto Encephalogram (MEG). The MEG has an excellent spatial and temporal resolution. However, for quantitative measurements of a magnetic field, superconducting components are required. Additionally, magnetic fields are ubiquitous in our world and MEG recordings suffer from these external artefacts even worse than EEG does. Therefore, MEG recordings can only be conducted in a magnetically shielded room. These two factors make MEG recordings rather expensive. MEG devices usually refrain the participant from walking around or moving the head, but the participant could potentially move arms and legs freely.

By far the most used recording technique for brain activity, not only in context of BMIs, is EEG. As discussed in section 2.1.1 on page 5 the sum of the electric field of millions of neurons in the cortex can be sufficiently strong to be measured at the outside of the skull. However, the electric field measured by EEG is pretty small (electric potential between electrodes rarely rises above a level of some dozens of microvolt).

Changes in the electrical field propagate fast and thus the EEG has an excellent temporal resolution. However, EEG recordings are very prone to volume conduction and suffer from a poor spatial resolution as

compared to basically all other recording techniques presented here. It is also the recording technique with the lowest costs. Actually, in recent years several EEG devices in the low-cost segment have been introduced, making this technique affordable even for home use. Although these low-cost devices offer an inferior signal quality, devices which cost a few hundred Euro (see figure 2.3b) haven been proven to indeed record neural activity (rather than arbitrary noise). Advancing this development, recently an EEG system reduced in size and with a miniaturised electrode array have been proposed (Debener et al., 2015). These devices allow for an unprecedented degree of unobtrusive EEG recording and providing virtually full freedom of movement, although, when analysing this data, the impact of muscle artefacts induced by movements always needs to be considered.

Due to the relatively weak electrical field of the brain, EEG is always prone to noise, not only from external sources (e.g. power lines, electrical devices in the area), but also from the participants muscular activity and background brain activity. When activated, each muscle generate an electrical field, which interferes with the neural recording. Of course the further away the muscle is, the lower is its impact on the recording. This decrease it cubic in nature and, hence, the muscles of the legs and lower body, although relatively strong, have a much smaller impact on the EEG recording than the relatively weak muscles which move the eye-ball and close and open the eye lid. Actually, ocular artefacts and those of the yaw-muscles are the strongest source of muscle artefacts found in EEG recording.

Noise in an EEG recording can be reduced. Some EEG recordings are conducted in an electrically shielded room. Although this is not necessary to obtain reasonable data, it surly improves the signal quality. If one does not want to use such a room for some reasons or no such room is available, one can remove most of the artefacts from electrical devices and power lines by frequency filtering. In Europe the power grid operates with an alternating current at a frequency of $50Hz$. When removing all portions of the recorded signal in the range $48 - 52Hz$ one removes most of this noise. Much of human neural activity takes place in a frequency spectrum from $1 - 40Hz$ and is, thus, unaffected by such filtering. But some portions of the upper γ band extend to $50Hz$ and beyond. Hence, these signals would be removed alongside with the artefacts and thus for analyses including these frequency ranges, electrical shielding is basically mandatory.

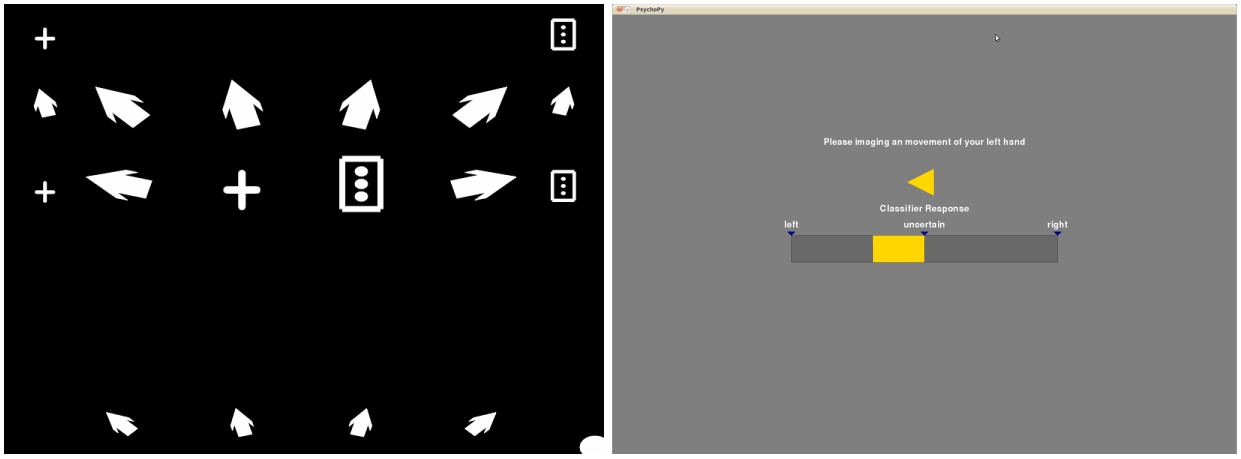
Regarding ocular artefacts, there are four common countermeasures: First, one can minimised the need for saccades and eye blinking. Depending on the experimental design, this might, however, not be possible or even desirable. Second, one can quite reliably identify ocular artefacts by visual inspection of the data and reject the parts of the data containing ocular artefacts. Third, one can record the activity of the ocular muscles (called Electro-Oculogram, EOG) alongside the EEG and subtract the contribution of the EOG to the EEG recording. Finally, one can transfer the data to source space using a technique such an Independent Component Analysis (ICA, see section 5.3 on page 42) identify and remove the component containing the ocular artefacts and, eventually, transfer the data back to sensor space.

Stimulus Presentation/Feedback

At least for active BMIs the experimenter usually wants to inform the participant about the outcome of the classification (feedback) and eventually present a series of stimuli, to which the brain response could give away the participants intentions. The required form of a stimulus is, however, strongly depended on the BAP employed for the BMI. Some BAPs (such as ERDs, see section 2.2.5 on page 15) do not require any form of stimulation of the participant, others require visual stimulation (e.g. CVEPs or SSVEP) and yet others have been shown to respond to different modalities of stimulation (e.g. P300 have been shown to occur for visual, auditive and even haptic stimulation). Further constraints, such as the oddball paradigm required to evoke a P300 (compare with section 2.2.5 on page 17), might apply. In order to align the occurrence of stimuli with the neural recording, the timing of stimuli needs to be recorded, typically with milli-second-precision.

In contrast to stimulus presentations, feedback components for a BMI have much less strict constraints. Any human interpretable form is usually adequate. For systems requiring stimulus presentation and feedback, those two are often combined, which is why they are described in common here, too. Furthermore, such a modality is, in more complex control scenarios, often used to inform the participants about events and general systems status (e.g. success or failure of an autonomous robotic action).

In section 2.2.2 I defined the categories of continuous and decision BMIs. This division carries over to feedback systems. Depending on whether the BMI has a continuous or a discrete output, different feedbacks are adequate. For choice BMIs a typical feedback would be, to display a word or icon representing the option chosen. For systems including a stimulus presentation in which each stimulus represents one option, often the stimulus is highlighted in some way, as feedback. Figure 2.4a shows such a stimulus presentation. After the stimulus presentation has finished and the classification system has returned a results, the corresponding



(a) A stimulus presentation for a CVEP-based BMI. From Riechmann (2014).

(b) The feedback used for the HExMInE experiment of this thesis.

Figure 2.4.: Examples for stimulus presentation and feedback.

icon would be highlighted to communicate that results to the participant.

For continuous BMIs the required feedback will need to represent continuous values. A typical example is given in figure 2.4b. The (continuous) output of the associated classification system is zero-centred. As a feedback, a coloured bar extends to either side depending on the sign of the result and to a degree depending on the result's absolute value.

Classification

Once the data has been recorded, it needs to be analysed. I claimed that, for a BMI, the relation between brain activity and its influence on the external world needs to be systematic (compare 2.2.1 on page 8). BMI systems generally even aim for a relation that is not only systematic, but also utile. It is important to note that a systematic relation does not necessarily imply a deterministic relation. All established techniques for neural data analysis are prone to misclassification and, therefore, occasional misinterpretation the participant's intentions.

The analysis of the neural data for a BMI has to happen in real-time. However, we have a rather soft definition of the term real-time here. For a BMI we expect that the influence on the external world is manifested such that the individual can (for an active BMI) still experience the system's response as being related to the individuals intentions. Hence, a time span of several seconds from the BAP happening to the implementation of the desired influence on the external world would be acceptable. This demand on a BMI is often also referred to as online data analysis, in contrast to offline analysis where data is stored for analysis days, month or even years later.

The actual classification, that is the reduction of neural data to one out of a set of classes or to one continuous value, is actually only the last step in the process of online data analysis. Although the details of the data analysis differ between BMIs, some general steps can be identified in the majority of all BMIs.

First, some type of filtering is applied to the data. This is virtually always a frequency filtering to a defined band in which the targeted BAP takes place. Sometimes, additionally, some spatial filtering or even filtering of some source components is applied.

Second, some type of feature extraction is applied. That is, the data is transposed in a way, which makes the targeted brain activity or the difference in brain activity between two or more conditions (e.g. imagining left or right hand movement) easier to discriminate. As the data in question is fairly high dimensional, feature extraction usually also aims to reduce the number of dimension.

Last, the data is actually classified. A large variety of linear and non-linear classification techniques have been proposed over the years. Classification can be two-class, multi-class or the reduction to a continuous value.

Device

In the end, the aim of a BMI is to influence the external world. What defines the device of a BMI is that it takes the classification results and, based on these, implements some effect on the external world. E.g. for a spelling BMI the device could be a computer screen or Twitter, if the spelled texts are posted there.[5] For other BMIs, the device could be a wheelchair navigating a room (Carlson and Del R. Millan, 2013). Or a robot bringing objects to the participant, pointing towards objects to identify them for interaction partners, expressing emotions by mimicking facial expressions (Hachmeister et al., 2011) or serving as a telepresence system which is operated via a BMI (Tonin et al., 2011). The device is closely linked to the categorisation of BMIs by usage-scenario, which I suggested in section 2.2.2 on page 9.

2.2.4. Measures of Quality for BMIs

BMIs have been optimised regarding different metrics depending on the scenario in which they are to be used. The most common measure to assess the quality of a BMI is the so called Information Transfer Rate (ITR). The ITR measures the amount of information that can be transferred over a BMI, including the time needed to recover from errors. This is usually done theoretically (rather than experimentally) and denominated in bits per minute. Kaper and Ritter (2004) proposes the formula:

$$ITR = \frac{60}{tc}\left(log_2(N) + p \cdot log_2(p) + (1-p) \cdot log_2 \frac{1-p}{N-1}\right) \tag{2.1}$$

Where N is the number of possible choices (in a choice BMI), p the classification rate (i.e. probability of a correct recognition) and tc the time needed for a single choice in seconds.

Depending on the scenario the amount of information which can be transferred in a given time frame is not the central or not the only optimisation criterion. E.g. the usage-scenario of a BMI might call for fast responses from the participant or a BMI might be optimised to produce a low number of false-positives (usually on cost of a higher number of false-negatives) or vice versa.

2.2.5. The B in BMI

I already stated that for brain activity to be accessible to EEG recordings, three main prerequisites must be fulfilled: i.) The activity must take place in the cortex and ii.) a vast number of neurons must be engaged in that activity and iii.) these neurons must be arranged in an open-field.

Hence, EEG will miss most of the details of the brain activity, but there are still numerous BAPs which have been identified in EEG recordings ever since Berger (1938) introduced EEG as a method almost 80 years ago. However, not all of these are eligible for BMI use. For active BMIs, participants need to be capable of volitionally influencing the targeted brain activity in some way. For passive BMIs this constraint does not apply, however, for an utile BMI one will usually only target brain activity representing some mental state which is also consciously perceived (or otherwise the participant would perceive the system's behaviour as chaotic).

I will go over some of the most prominent brain activity patterns and rhythms, now, many of which have already been employed in BMI-usage. The first type of activity patterns identified in an EEG were brain rhythms, rhythmic activity which takes place within defined frequency ranges and which have been studied extensively. Each of these rhythms has been associated with certain mental states or tasks. Although definition of frequency ranges for the different rhythms are similar throughout the literature, they may differ in details. The following list is adopted from Sanei (2013).

δ-rhythm ($0.5 - 4Hz$): This rhythm occurs in healthy persons usually in deep sleep states. Their presence in wake-state persons can be an indication for brain trauma. They can be found in frontal regions.

θ-rhythm ($4 - 7.5Hz$): This can be found in light-sleep states or in sleepy wake-state persons. It has also been associated with creativity and meditation.

α-rhythm ($8 - 13Hz$): Is associated with sleepiness. It is amplified when the participant closes his/her eyes and diminished when he/she engages in mental activity. It vanishes when one falls asleep. The interpretation of α-rhythm as an idle pattern of the visual cortex has recently been challenged by different researchers. It is found primarily in occipital and parietal regions.

[5] This was another student project in our group, actually.

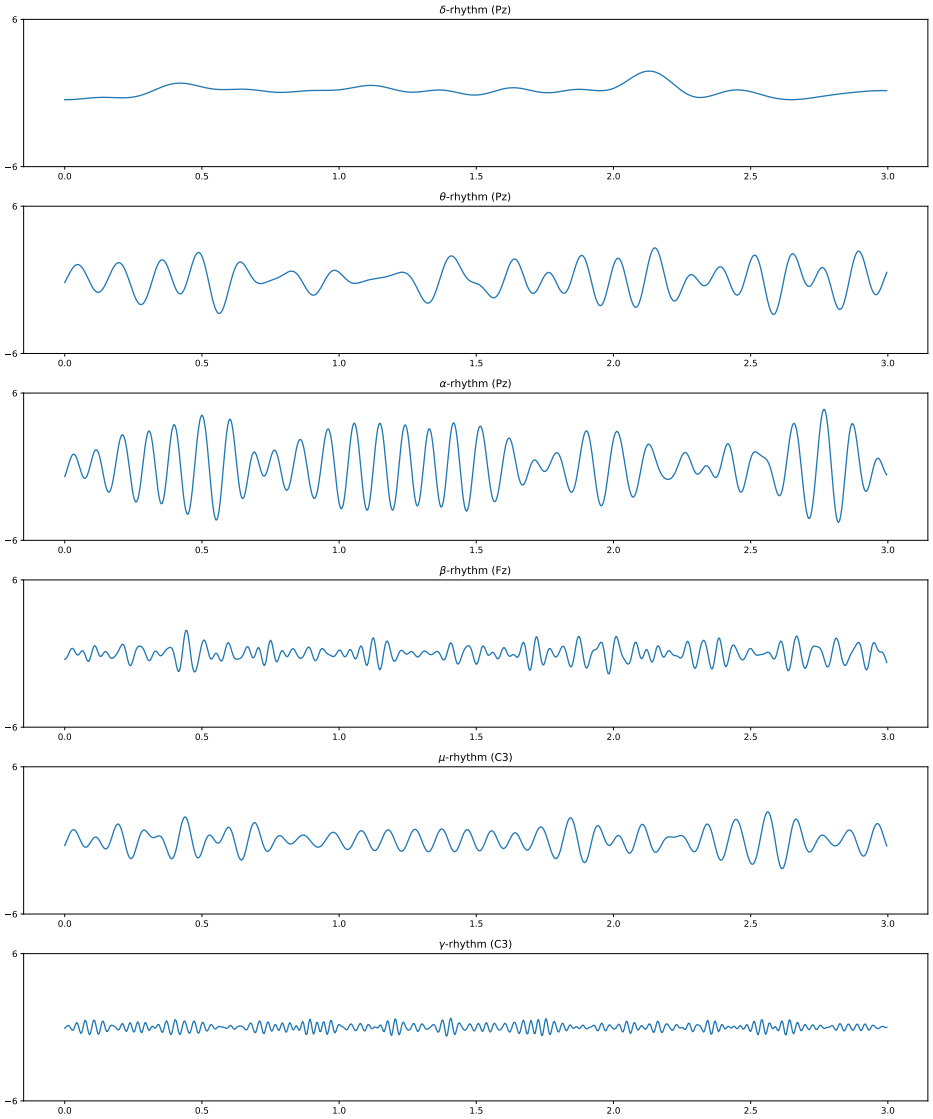


Figure 2.5.: The same data segment (but different channels) filtered with regard to the different brain rhythms.

β-rhythm ($14 - 26Hz$): Has different interpretations. It has been associated with problem solving and active thinking. It is mostly encountered in frontal and central regions. The central part of the rhythm declines when engaged in motor activity.

γ-rhythm ($> 30Hz$): Is associated with concentration, learning and meditation. Some body parts' motor activity seems to be related to this band as well (e.g. index fingers and tongue).

μ-rhythm ($8 - 12Hz$): Is associated with the idle state of motor neurons. The suppression of the μ-rhythms has been associated with planning and execution of motor task and also with motor observation. This has been interpreted as a first indication for the presence of mirror neurons in humans, previously found in some monkey species.

The frequency band of μ- and α-rhythms largely overlap, raising the question why they are deemed distinct rhythms: The distinction between these rhythms is largely grounded in locations they occur in ($\alpha \rightarrow$ parietal, $\mu \rightarrow$ somatosensory cortex) and in the states in which they appear/are suppressed.

For the use within an active (EEG-based) BMI, a BAP not only needs to be accessible with an EEG. There must be a way for the participant to volitionally control or influence the brain activity pattern. Various such BAPs have been proposed in the past decades and I will briefly go over the most prominent of these:

P300: May well be the most commonly used BAP in the context of BMIs. It appears as a response to a stimulus the individual has been expecting and which is embedded into a series of stimuli irrelevant to the individual. As it will be one of the two BAP used for this thesis, I will discuss it in some more detail in section 2.2.5 on page 17.

ERD: When synchronisation between neurons engaged in a brain rhythm is lost (usually in response to some event), the rhythm's contribution to the EEG signals is diminished up to the point at which it completely vanishes. This effect is called Event-Related Desynchronisation (ERD). When (millions of open-field) neurons (re-)gain synchronisation this results in an increase in amplitude of the corresponding rhythm in the EEG and this is accordingly called Event-Related (Re-)Synchronisation (ERS). Motor-correlated ERDs, as one special form of ERDs, are the second BAP used in this thesis and I will discuss them in more detail in section 2.2.5 on page 18.

SSVEP: The Steady-State Visually Evoked Potentials are a response in the visual cortex to some stimulus flickering at a certain frequency. This response can be found within the corresponding frequency band in the EEG. If, for instance, I present two different stimuli, one flickering at $5Hz$ and one flickering at $7Hz$ I would assess the band power of two narrow frequency bands around these two frequencies. If the band power in one of the bands suddenly rises, I can infer that the individual has been looking at the corresponding stimulus. Because the individual needs to fixate one stimulus or the other, SSVEP-based BMIs are always dependent BMIs.

CVEP: The Codebook Visually Evoked Potentials have originally been proposed by Sutter (1992). They also rely on the response of the visual cortex to a certain flickering stimulus. However, the stimulus does not flicker with a fixed frequency, but in a random pattern. If several stimuli flicker, governed by circularly shifted versions of that same pattern, one can infer the stimulus to which the individual attended (by shifting the EEG data in the reverse direction). This BAP is remarkable in so far, as it has not further been exploited for more than 20 years and only recently researches have started evaluating it. Riechmann (2014) wrote a dissertation evaluating the applicability of CVEPs for different real world BMI control tasks. Furthermore, CVEP-based BMIs have by far the highest ITR and response time of all established BMI-BAPs.

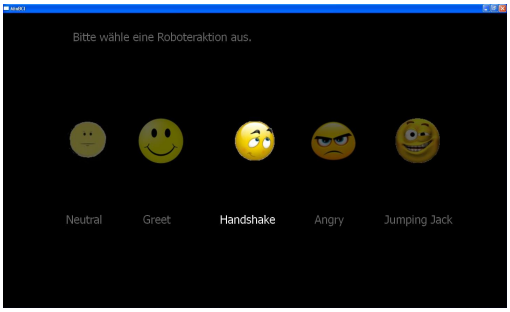
ERP: Event-Related-Potentials are actually a category of different BAPs. They "are voltage fluctuations in the EEG induced within the brain, as a sum of a large number of action potentials (APs) that are time locked to sensory, motor or cognitive events." (Sanei, 2013). ERPs are usually characterised by their amplitude, latency and spatial distribution over the scalp. They typically have an amplitude in the range from 1 to 30 μV (Sanei, 2013) and are therefore significantly weaker than much of the other electrical activity of the brain, i.e. various brain rhythms (see section 2.2.5 on page 15). To make them visible, commonly frequency filtering and averaging of several data segments time-locked to known occurrences of the ERP in questions are performed. ERPs are commonly divided into exogenous ERPs and endogenous ERPs (Sanei, 2013). Both types of ERP can be a response to an external stimulus. Exogenous ERPs are, however, a more direct response and occur inevitably (like a reflex of the brain) and they depend in their properties (amplitude, shape, scalp distribution, ...) on the properties of the stimulus (CVEPs would be an example). Endogenous ERPs are independent in their properties from the properties of the stimulus and depend more on the mental state of the participant. Also they can be diminished or suppressed due to some mental condition, i.e. they do not occur inevitably. As a rule of thumb, exogenous ERPs occur in the first $100ms$ after the stimulus/event while endogenous ERPs generally occur later than that.

In the following I am going into some more details on the two BAPs used for the BMIs in this thesis.

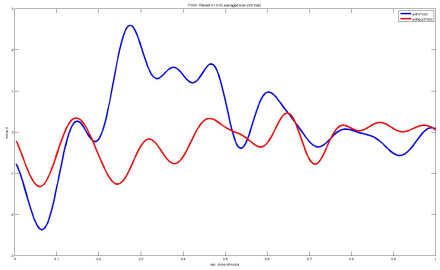
P300

The P300 brain response is an ERP. The P300 occurs about 300 milliseconds after the event that triggered it, is strongly positive (thus the name P300) and originates at a central-parietal location spreading over virtually the entire cortex.

What makes the P300 so valuable for BMI usage is that the conditions under which it occurs are quite sharply defined and that it occurs with a high reliability. When a.) a series of stimuli are presented to an individual, b.) most of these stimuli are irrelevant to the individual, c.) the individual is expecting one specific out of these stimuli, this is called oddball-paradigm. After each of the stimuli certain characteristic



(a) The stimulus presentation from the example. Five different smileys which are highlighted repeatedly in random order.



(b) The average brain response of a sample participant to relevant stimuli (blue) and irrelevant stimuli (red). The clear peak in the blue curve is the P300 brain response to the occurrence of the relevant stimulus.

Figure 2.6.: Stimulation and characteristic P300 brain response.

brain responses, particularly ERPs, can be measures. E.g. a strong negative peak about 200 ms after each stimulus (the N200). If the expected, relevant stimulus finally appears there is, however, one brain response, the P300, which does not occur for the irrelevant stimuli. Thus, with proper analysis of the EEG data, the relevant stimulus can be inferred.

An example should help clarifying the oddball paradigm and its relation to the P300: For an experiment during my master thesis I set up a stimulus presentation featuring five different smiley faces displaying different facial expressions. The idea was to select one of these facial expressions (via a BMI) which would than be displayed by an iCub humanoid robot. The participant in this experiment would eventually choose one of the five facial expressions and start concentrating on the corresponding smiley. The smileys would than be briefly highlighted (flash) in a random order. For each of theses flashes a N200 could be expected in the EEG, but there should only be a P300 response to flashes of the chosen smiley (relevant stimulus). For the other smileys no P300 would occur (irrelevant stimuli). With a suitable data analysis and classification the P300 response can be identified automatically, which allows inferring the selected smiley.

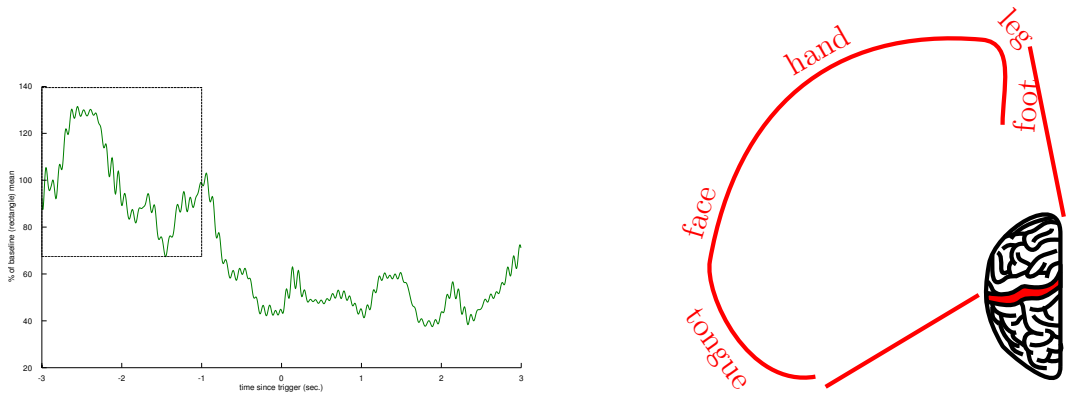
The P300 response is a rather strong potential, however, it is still deeply embedded in the residual brain activity (and noise) when inspecting a raw EEG signal. The P300 cannot be identified visually in raw EEG data for most participants. At least not by non-experts. Up to date, data processing/classification algorithms can, when presented with data segment of which about 50% contain P300 and the others not, classify P300 from non-P300 segments with an accuracy of about 80%, strongly depending on the individual. This can be improved by repeating the stimulation several times and accumulating the results. With such a system classification accuracies beyond 99% have been achieved. The data processing and classification and the accumulation of classification results used for this thesis are discussed in section 7.5.2 on page 86.

Event-Related Desynchronisation

When millions of cortical neurons arranged in an open-field fire at the same rate and in a synchronised manner this can be measured as a rhythm in the EEG. When some event breaks the synchrony in neuron-firing, this can be observed as a drop in band power in the corresponding frequency band. This is called an Event-Related Desynchronisation (ERD). There are several mental techniques which allow individuals to volitionally trigger such ERDs. Hence, some types of ERDs are eligible for the use within a BMI.

An ERD is visualised in figure 2.7a. The y-axis gives the values in relation to the mean of the time window depicted (which is assumed to be before the onset of the ERD). The later values drop to $50-60\%$ of those in the reference time frame. As these are squared value (compare 6.5.1 on page 74), this is not to be interpreted as a drop of $40-50\%$ in band power caused by the ERD.

The ERDs most used in BMIs are probably motor-ERDs, i.e. ERDs that occur in the primary motor cortex. The primary motor cortex is located in the human brain in a relatively thin, vertical stripe on both sides, central on the fronto-parietal axis. The motor cortex is segmented and each segment is associated with a different part of the body. The size of each of these segments depends on the complexity of the control of the corresponding body part. The control of the arms is relatively simple, while the control of the hand, feed and tongue is highly complex and consequentially occupies large segments of the motor cortex. It has even been reported, that persons which regularly perform extremely complex task with their hands (e.g.



(a) Event-Related Desynchronisation volitionally triggered by a participant at the C5 electrode. The countdown after which the participants were instructed to perform motor imagery expires at 0 on the time-axis. Already a second before this one can see a clear drop in the squared and averaged signal. Visualised using a method proposed by Pfurtscheller and Lopes da Silva (1999), compare 6.5.1 on page 74.

(b) A schematic of the organisation of the primary motor cortex. The black arc represents the (motor/sensor) cortex as it stretches from the top to the bottom of the brain's surface. The different body parts are depicted over the sections, which are associated with them. Tongue, hand and feet have the largest representation on the motor cortex. Figure inspired by BCI2000 project.

Figure 2.7.: Visualisation of an ERD and schematic of the organisation of the primary motor cortex.

violinists) reorganise their motor cortex allocating even more space (and neural capacity) to the control of the hands (Schwenkreis et al., 2007). Control is organised contra laterally, meaning the control of the left body part (left hand, left foot, etc.) is done in the right hemisphere and vice versa.

The segmentation of the motor cortex is often illustrated in a figure called the motor homunculus. In such a figure, body-parts are depicted in a size corresponding to the size of the associated motor cortex area. This results in an extremely disarranged figure with large hand, feet and tongue and small arms and legs. Another representation (figure 2.7b) depicts the body parts at the side of a section of the motor cortex.

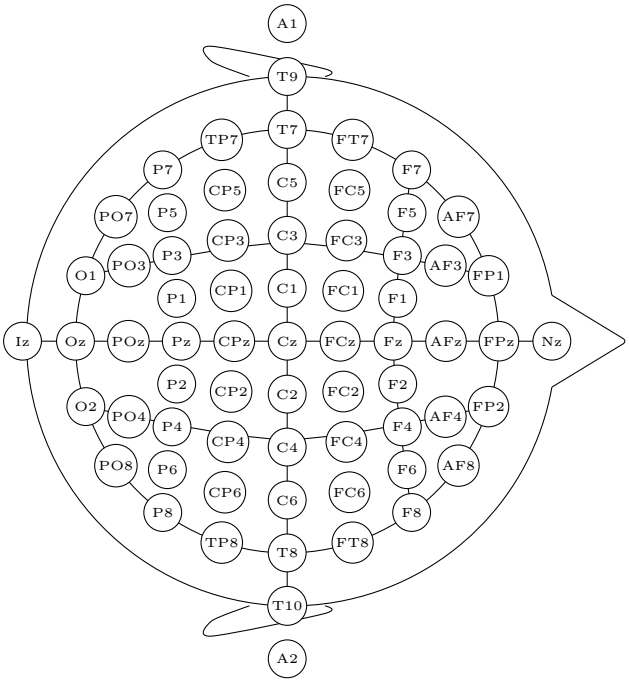
When idle, the neurons of the motor cortex fire in a synchronised manner at a rate of $\sim 10Hz$. This can be measured in EEG and is commonly referred to as the μ-rhythm (compare section 2.2.5 on page 15). Of course, the larger the portion of the cortex tasked with the control of that (currently) idle body part is, the larger is its contribution to the μ-rhythm. When the individual intends to make a movement with a body part (say the left hand), the neurons start to fire in different patterns, as they pick up their task to plan and control the motor action. They desynchronise. For body parts with a large representation in the motor cortex, this results in a drop in band power of the μ-rhythm, detectable with sophisticated EEG data analysis. In my example, the initiation of a movement of the left hand would lead to a desynchronisation of motor neurons in the right motor cortex which would result in a measurable drop in band power in the μ-band, localised approximately in the middle of the right motor cortex. This desynchronisation would persist for the entire duration of the movement. After the movement has come to an end, neurons modulating this movement would go back to their idle state. They would resynchronise resulting in a (localised) increase in band power of the μ-band (ERS).

However, actually measuring ERDs that occur during the execution of a movement is difficult, as the muscle activation produces an electrical field, that would be present as an artefact/noise in the EEG recording and which would occlude most of the μ-band activity (and the brain activity in general). Luckily ERDs in the motor cortex occur also in a series of other states.

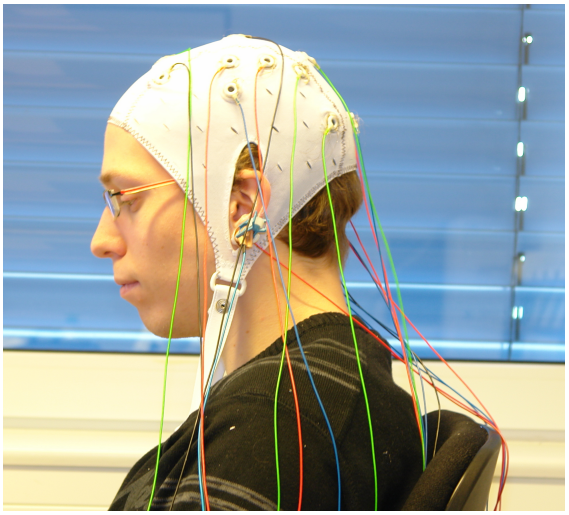
First, the ERD precedes the actual movement, i.e. the ERD already occurs several dozens of millisecond before the muscle activation. During that short time window it would theoretically be possible to record an ERD. I used that in the training process for my first experiment, see section 6.3.1 on page 63.

Second, the ERD also occurs when observing a movement performs by someone else. When a person watches someone else perform a hand movement, the first person's motor cortex would be activated. An ERD would occur.

In monkeys certain types of neurons exist, which have been shown, in single cell recordings, to respond to motor activity and to motor observation. These neurons are called mirror neurons and have been associated



(a) A schematic of the 10-20 channel configuration.



(b) The type of cap used for electrode placement throughout this thesis.

Figure 2.8.: Standard Positioning of EEG electrodes and an electrode placement cap for easy fixation of EEG electrodes.

with co-representation of others actions and understanding other's action's intentions. That the μ-band also responds to movement execution and observation, has been interpreted by some researches as a first indication that there exist mirror neurons also the human brain.

Third, the ERD also occurs when imaging a movement. This mental technique is called Motor Imagery (MI) . Volitionally evoking ERDs using Motor Imagery which can be recognised by a BMI, however, needs training on part of the participant and in any case takes a great deal of concentration. The process of training MI is described in detail in section 6.3 on page 63.

ERDs, even less than P300, can be spotted in raw EEG data by mere visual inspection. Automated classification need specialised data processing and classification algorithms which need to be parametrised to the characteristics of the neural data of each participant. The classification and data processing used for my experiments are described in section 6.4.3 on page 68.

2.2.6. Electrode Positioning - the 10-20 System

Various schemata exist, how electrodes for EEG recordings are to be placed on the scalp. Most commonly used is probably the so called 10-20 system which was proposed by Jasper (1958). Figure 2.8a shows a schematic of the 10-20 system.

It uses two landmarks: The nasion, the base of the nasal bone (identified by the label Nz in the 10-20 notation) and the inion, the notable bone bulk at the back of the human skull (Iz). With these two landmarks two paths are defined: The first between nasion and inion all the way centrally over the scalp and the second from nasion to inion following the scalp laterally on either side. Along these axes electrode positions are distributed in equal distances every 10% or 20 %, depending on how tightly electrodes shall be placed (thus, the name 10-20 system). Starting with the electrode positions on these two initial paths, a net of electrodes spreads covering the entire scalp equidistantly is extrapolated.

The positions are named according to the broad cortical regions: frontal (F...), central (C...), parietal (P...), occipital (O...), temporal (T...). The electrodes along central sculcus of the brain share the last letter z (for zero). The remaining, lateral electrodes are enumerated. Left side electrodes with odd numbers starting with one directly left of the z electrodes and with even number on the right scalp side accordingly. A1 and A2 identify the earlobes, which are considered to be unaffected by electrical activity of the brain and are often used a references and ground during EEG recording.

The absolute distances between electrodes obviously vary between participants as they depend on the two path lengths between Nz and Iz. Measuring these positions can be a tedious and time-consuming process. Additionally the electrodes need to be fixated on the scalp with a special paste, which allows for easy removal of the electrodes after the recording. Therefore, electrodes have the tendency to come off during the recording.

For these reasons electrode placement caps have become very popular. These elastic caps are available in various sizes and are strapped either to the chin or to a breast belt. The caps are equipped with special clamps at the positions at which electrodes are to be placed. The elasticity of the cap ensures a good approximation of the 10-20 system positions, as long as the cap itself is placed adequately. The clamps allow for a steady fixation of the electrodes without the need for a fixation paste.[6]

So much for the foundations required for the rest of this thesis. This chapter did of course not cover the topics treated in length. It was my endeavour to be brief and still equip those foreign to neuro-research and BMIs for the rest of the thesis.[7]

[6] Which is not to say that no paste was required, but the sole purpose of the paste is to ensure a good electrical connection between electrode and scalp (i.e. a low resistance). Therefore, the paste is electrolytic and has a peeling-like effect which allows removal of dead skin scurf with cotton sticks.

[7] Considering I am a computer scientist by trade, it was not too difficult to take that perspective.

3. Related Work

This thesis draws from several other research fields. It uses techniques of Brain-Machine Interfacing (BMI) and hyper-scanning. It investigates the neural foundations of human interaction and therein the special case of machine-mediated interaction.

In this chapter I will present several publications which have had considerable influence on this work and/or are corresponding to different aspects of it. I will start by discussing some publications which include machine-mediation, with and without BMI-usage, in section 3.1, below.

One form of human interaction is cooperation and I consider this form of interaction to be scientifically particularly interesting. I will describe some recent hyper-scanning publications focussing on cooperation in section 3.2 on page 24.

In many hyper-scanning experiment participants have asymmetric roles, i.e. they are not equal partners but have different tasks in the interaction. From this asymmetric role assignment a certain pre-dominant direction of the information flow between the participants (interaction level information flow) can often be postulated. E.g. when one participants observers the results of actions of another participant, an interaction level information flow from the initiator of the actions to the observer can be assumed. This might have an impact on neural connectivity and a few studies could actually show such impact. These will be discussed in section 3.3 on page 27.

Finally, several theoretical contributions on neural connectivity estimation exist. Those publications discuss connectivity estimation in general, not hyper-analysis in particular.[1] I will present two publications discussing connectivity estimation on two different levels – i.) comparing connectivity estimators and ii.) comparing transformations on which connectivity estimation is based – in section 3.4 on page 28.

3.1. Human-Machine-Human Interaction

One idea which is at the heart of this thesis is the use of a triadic human-machine-human interaction setting as a mean to study human-human interaction. Machine-mediated interaction is not a new phenomenon: Online-chats, email, telepresence systems, video-calls and even normal phone calls are forms of machine-mediated interaction. Most of these are transparent mediators, though, (compare section 1.4 on page 3).

Riedenklau et al. (2012a), the developer of the TAO robotic platform[2], examines a remote machine-mediated interaction scenario including non-transparent mediators. The Tangible Active Object (TAO) robotic system is intended as an enhanced (in particular tangible) user interface. The TAO robotic system is described in some more detail in section 6.2 on page 62.

In the experiment, two special desks (so called T-Desk) are connected to allow for remote interaction, more precisely remote collaborative planning of the interior furnishing of a virtual flat. The furniture is represented by the cube-shaped TAO robots, which can be moved and rotated by the participants, resulting in translation and rotation of the virtual furniture. More than that, the TAO robots are actuated, such that a translation or rotation applied to one TAO by participant A on one T-Desk is repeated by the corresponding TAO in front of participant B on the second T-Desk, autonomously.

I consider this study interesting, as the mediating role of the TAO robots is particularly central and prominent. They reflect the partners actions and intentions and the entire interaction flows through these machines, i.e. the mediators are not transparent. This is a good example for a triadic human-machine-human setting.

Another aspect of this study has caught my attention: In section 1.2 on page 2 I defined the term of a shared space. However, Riedenklau et al. (2012a) blurs this definition. At first glance, there is no shared space for the interaction in this study. The virtual space of the flat does, by definition, not count and the participants are remote, acting on two different desks. However, the actions performed on one desk are transmitted to the other Desk. Hence, actions of the two participants could potentially interfere on a

[1] To my knowledge the applicability of connectivity estimators used for within-participant analysis, for the analysis of hyper-scanning data has never been challenged or discussed.

[2] Which I also use in the first study of my thesis.

physical level, even more so as the TAO system can be set to upkeep certain constraints with respect to the relative position and rotation of different TAOs. The possibility for interference (or congruence) of different participants' action on a physical level is a key-feature of a shared space. In a certain sense[3] the systems makes two distinct spaces coincide, forming a shared space for the interaction.

Including a BMI in machine-mediated interaction allows for a much more direct and close link between each of the humans and the mediating machine. As Finke et al. (2012) state: "Correlates of brain activity are the most direct and unbiased output signals that can be acquired from a human."

A predecessor work of this thesis, Hachmeister et al. (2011), already involved BMI-based machine-mediated interaction. Participants could use a BMI to control the iCub Humanoid robot. In a first step the participant could control the hip-joint rotation of the robot. As soon as he/she was satisfied with the orientation, e.g. when he/she had oriented the robot towards another human in the room, the participant could select a facial expression which the robot displayed. In an alternative version, the iCub could point towards one out of five objects, rather than displaying a facial expression.

This system pointed out a possible direction for the development of BMI based rehabilitation. The first BMIs in this field were spelling devices, allowing the user a slow, written form of communication (Wolpaw et al., 2002). However, the limitations of written-only communication has not only been an issue of extensive research, but we all experience these each and every day (compare section 1.4 on page 3).

The idea of this BMI-system was to go beyond pure spelling devices and to include non-verbal forms of communication. A humanoid robot could sever as a surrogate body. This is already a setting pretty similar to the triadic settings I aim to explore. But the machine-mediation in this setting is a "one-way" or asymmetric machine-mediation. The (handicapped) user of this system interacts with a (non-handicapped) other person. The interaction direction from the user to the partner is indeed machine-mediated (even already including a BMI system). The other direction (from the partner to the user) is thought to be direct, i.e. without any mediation (neither by a machine nor other). The partner responds to the performed actions and displayed facial expression directly.

In Finke et al. (2012) a system is described which features a full/duplex BMI-based machine-mediation. Two users were connected to two independent BMI systems. Hence, in this study hyper-scanning was conducted. In Finke et al. (2012), however, no hyper-analysis was performed. Each participant steered a NAO humanoid robot. Both robots were located in a shared space alongside some paper cubes as objects for manipulation.

Two types of interaction were realised: A cooperative scenario, in which one participant had to deliver a cube to the middle of the shared space, where the other participant had to pick it up and return it to his/her home zone. The scenario was intended as an abstraction of a shop-scenario. The shop assistant delivers goods at a predefined hand-over position (counter) where the customer collects them. The second scenario was a competition in which two cubes were placed in the middle of the shared space and each participant intended to be the first to deliver a cube to his/her home zone.

This study has several parallels to my thesis. For one hyper-scanning (but no hyper-analysis) was performed. Furthermore, the interaction took place in a well-defined shared space, in which the actions of participants could potentially interfere. Finally, the participants could only see the shared space and the robots by a live video stream. In so far, this scenario was a teleoperation scenario similar to this thesis's final study (compare section 8 on page 93). Finke et al. (2012) focused on a general evaluation of the feasibility of a multi-user BMI and a BMI-based human-machine-human interaction. They thereby laid foundations upon which this thesis depends, as I embark to combine BMI-based human-machine-human interaction with hyper-analysis to contribute to the investigation of the neural foundations of human interaction.

3.2. Cooperative Behaviour

Cooperation is an important form of interaction, not only because it is excised frequently in everyday life, but also because it is a key-ability which only relatively few species developed and none of these has developed this ability as far as humans did.

A natural first step when aiming at the neural foundations of cooperation is to compare neural data during cooperative and during non-cooperative episodes. This is what De Vico Fallani et al. (2010) have done: They compared neural connectivity during cooperation and during defection. The authors used a very simple, controllable and widely studied scenario, namely the Prisoner's Dilemma. In each iteration of the game participants can either choose to cooperate or to defect. The mutual decisions are only revealed to the

[3] And neglecting still persisting issues with latency and the TAOs' rather slow positioning.

participants after both participants have made their choice. If both participants defected, both would get a small reward, if both participants cooperated they would get a larger reward. If one participant cooperated while the other defected the defecting participant would get a huge reward, while the other would get no reward at all. This game has been extensively studied in various contexts. Three main strategies each player can choose are commonly considered: Either cooperate regardless of the partners past decision (C), defect regardless of the partner's past choices (D) or tit-for-tat strategy where the player always adapts the choice of the partner from the previous trial.

The authors could show many statistically significant neural connections, both within and between participants (hyper-connections). On a per-participant level, the authors could show differences in connectivity especially in α-band. They also reported a high variability in connectivity between pairs of participants. I made similar observation during my experiments and for the same reason many hyper-analysis studies report their results on a per-participants basis and in a descriptive manner. One approach to counter this is, not to compare (hyper-)connections, but to take these connections as a graph and then to apply some standard metrics from the field of graph theory (such as in-/out-degree) to identify some overarching pattern. Regarding this approach De Vico Fallani et al. (2010) state that "... averages and standard deviations of graph metrics computed over the 26 couples do not allow for the characterization of the typical hyper-brain network associated with a specific strategy."

Employing more complex graph measures (namely graph efficiency, divisibility and modularity) the authors could show that hyper-brain networks while both participants have chosen the defect strategy, feature a lower efficiency and a higher divisibility and modularity as compared to the other strategies. Further pursuing these results they trained a Multi-Layer Perceptron to discriminate DD from non-DD networks[4] on the basis of these three graph measures. They applied this for θ-, α-, β- and γ- band separately and achieved classification accuracies way beyond chance level on all frequency bands.

This way the authors could show that hyper-connectivity during non-cooperative behaviour has distinct properties. However, because the Prisoner's Dilemma knows more strategy profiles than just cooperation and non-cooperation, the distinctiveness of hyper-connectivity during cooperation does not follow automatically.

The Prisoner's Dilemma is very popular in scientific studies because of its simplicity and because it is well studied. However, regarding interaction research, it lacks any relevance for everyday life. Conversely, in Astolfi et al. (2011a) the authors examine the cooperative behaviour of a professional pilot and his/her first officer, employing hyper-scanning. This is a particularly interesting domain to study cooperation, as a normal flight has phases with a high need for cooperation, in particular starting and landing, and phases in which almost no cooperation is needed, namely mid-flight. This allows to compare neural data in these two conditions with one another. In this study, while starting the aircraft, the pilot was responsible for steering the plane, while the first officer attended to the various instruments. During landing, roles were exchanged.

Of course, it is not possible for safety reasons, to perform neural recordings from the pilots during a real flight, but it is very common for professional pilots to receive simulator training regularly and these simulators offer a high degree of realism. Consequentially recordings where made in a flight simulator.

The methods for neural recording used in this study are actually pretty similar to those used throughout this thesis. Recording was done with two 16-channel EEG devices at 256 Hz. The connectivity estimation was done using Partial Directed Coherence (PDC) based on a Multi-Variant Autoregressive Model (MVar). The connectivity estimation was done on θ, α, β and γ band.

However, the connectivity estimation was done on electrodes (sensor space) while I did the connectivity estimation on ICA components (source space, compare section 5.3 on page 42) for reasons I will discuss in section 5.6 on page 47. Furthermore, Astolfi et al use a different method of statistical evaluation.

Astolfi et al could show a highly increased number of significant hyper-connections in α-band during starting phase as compared to mid-flight. The connections involved mainly frontal and parietal regions and were directed from the first officer (who was assisting the captain during this phase by controlling the instruments) to the captain (responsible for controlling the aircraft). Astolfi et al found this in line with intuition as "the temporal delay between the activity of the two subjects is at the basis of [...] Granger based estimators like PDC." (Astolfi et al., 2011a)

I cannot follow this notion, as in my understanding the first officer would see the results of the actions taken by the captain with a time delay. Hence, the first officer would have an information later than the captain. Therefore, it would be my expectation that the hyper-connections would be directed from captain to first officer.

It would have been interesting to know, whether or not the general direction of the neural information flow

[4]Meaning networks which established while both participants defected (D)

was reversed during landing phase (during which responsibilities of captain and first officer were reversed). Unfortunately, although Astolfi et al confirm that they had a high number of significant hyper-connections during landing, they make no statement about their direction.

In a later study Astolfi et al. (2014) showcased to which degree connectivity is actually shaped by the belief of the individual and, simultaneously, refuted suspicions whether hyper-connectivity really is a correlate of social interaction or has different causes.

The authors developed a small, game-like program. On a screen a bar with a ball on it was visible. Each end of the bar could be lifted by pressing a corresponding button. It was the task of the participants to keep the ball in balance on the bar, reach the upper end of the screen with it and meanwhile avoid collision of the ball with obstacles (rectangles) on the screen. The setting is depicted in figure 3.1. The participants played this game in three conditions:

SOLO The participants were paired, but in this condition participants played the game all by themselves. Each participant could control both ends of the bar.

JOINT The participants played the game as a team. Each participant of a pair controlled only one end of the bar and participants needed to cooperate to achieve the goal.

PC Participants were told, the second end of the bar was controlled by a computer program. In fact this condition was technically congruent with the JOINT condition, i.e. each participant still controlled one end of the (same) bar.

Astolfi et al could show a significant difference in hyper-connectivity between SOLO and JOINT conditions. And they could show significant differences between the JOINT and PC conditions. They could not show any significant differences between SOLO and PC condition. Put differently, the only difference between the JOINT and the PC condition was that participants were told, they were cooperating with a machine or with their partner, respectively. In both cases they actually cooperated with their partner. Still, the belief that in the PC condition they did not cooperate with their partner, decreased hyper-connectivity to a point at which it was indistinguishable from that found during SOLO condition.

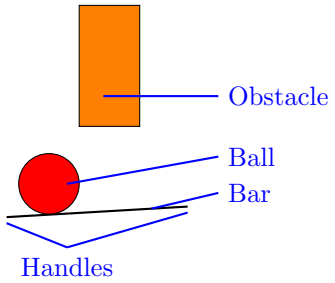


Figure 3.1.: This figure was redrawn from memory, based on the poster-presentation of Astolfi et al at the EMBC2014. In the 2D virtual scene the two ends of a bar can be controlled by the participants. The ball will roll on the bar according to the bars slope (as if a gravity force was applied). The participants' task was to reach the upper screen border, not let the ball drop off the bar and still avoid contact of the ball with obstacles.

This study is particularly remarkable because it showcases to what extend a human's belief determines higher brain function. This is particularly true regarding hyper-connectivity. And it demonstrates that hyper-connectivity is indeed correlated with social interaction. If hyper-scanning was de-correlated with social interaction, no difference would have been found between JOINT and PC condition.

An interesting modification of the described experiment would be, to tell only one participant that he/she was actually playing with his/her partner and to tell the other he/she was playing with a machine (let us name this condition PC'). As cooperation is a mutual process, it would be my expectation to not find any increase in hyper-connectivity in PC' condition compared to SOLO condition, however, it would be worthwhile to verify this hypothesis experimentally.

Comparisons of hyper-scanning studies need to be treated with caution, generally, due to the lack of standardise experimental protocol and the vast number of independent variables involved in social situations. Yet the results of the above studies are a promising basis for my own research. They showed that one can identify a significant difference in hyper-connectivity between phases of cooperation and phases without cooperation/phases of non-cooperation. In my second study I will complement these results by evaluating a similar question in a scenario where interaction is remote and machine-mediated.

3.3. Asymmetric Roles

During the analysis of the data recorded for this thesis, I developed the hypothesis that hyper-connectivity might be sensitive to assignment of roles and/or interaction level information flow. That is not to say that if a dominant direction of interaction level information flow can be postulated between participants, an ipsi-directional neural information flow can be expected. But when participants have different roles in the interaction (asymmetric role assignment) and/or a dominant direction of the interaction level information flow can be postulated, hyper-connectivity is fostered (regardless of its direction).

Many asymmetric role assignments actually imply some type of interaction level information flow. In Astolfi et al. (2011a) the first officer could see the results of the actions taken by the pilot on the instruments, hence an interaction level information flow from pilot to first officer could be assumed. Many other hyper-scanning studies published in the past decade have a similar asymmetric role assignment.

One of the first attempts to really tackle the impact of different roles or the lack thereof was Dumas et al. (2010). They placed to participants in two different rooms, transmitted two live video streams of their hands mutually and recorded their neural activity using EEG (hyper-scanning).

They instructed their participants to perform meaningless gestures. During some phases, the participants were assigned the role of a model and an imitator, meaning that one partner was instructed to imitate the other's free hand movements. During other phases participants were free to either imitate their partner or to perform independent movements. The video recording was later used to annotate the data during free phases as to whether participants were in synchrony (behaviourally). The role assignment of model and imitator (regardless if taken up by the participants volitionally or by instruction from the experimenter) implies an interaction level information flow from model to imitator.

The connectivity was analysed using the Phase Locking Value (PLV) proposed by Lachaux et al. (1999). Only hyper-connections were considered. This is similar to the data analysis of the first study of this thesis.

Dumas et al found significant phase synchronisation between centro-parietal regions on the right side of both participants in α/μ-band. They also state that this band has been associated with the mirror neuron system. The authors could also show neural synchronisation in β and γ-band. However, the α/μ-band inter-brain synchrony was found to be the most expressive activity, when it comes to the discrimination of behavioural synchrony or asynchrony.

PLV is a symmetric estimate, meaning that it can identify neural connectivity but cannot pin-point a direction of an information flow. Hence, the question whether a dominant directionality in hyper-connections existed and whether or not this was corresponding to the interaction level information flow (from model to imitator) cannot be answered based on the publication.

In Dumas et al. (2012a) the previous study design is advanced to tackle questions of agency and perception of body ownership. However, this study is methodically and thematically only remotely related to this thesis. Similarly in Dumas et al. (2012b) the potential of modelling human brain activity as a method to examine neural (hyper-)connectivity based on the same study design is presented, but again only marginally relevant for this thesis.

Another hyper-scanning study in which participants have been assigned different roles is Sänger et al. (2013). Here participants are assigned different roles as leader and follower in a guitar duet. Obviously this role assignment again implies an interaction level information flow from duet-leader to -follower. And indeed Sänger et al could show a directionality in the hyper-brain network from leader to follower at certain points during the play in α- and β-band.

I would consider the production of music a particularly well suited domain for this type of analysis (compare section 2.1.4 on page 7). The production of music is a cooperative process which requires presumably an exceptionally high degree of neural synchrony, further strengthened by the continuous presence of auditory cues/"pacemakers". As a consequence, usually rather subtle effects might be amplified – the contrast in the neural connectivity analysis is increased, speaking figuratively.

In general there is quite some evidence for an impact of asymmetric role assignment. However, all these studies also feature an interaction level information flow. Having an interaction level information flow means having a sender and a receiver of this information flow, thus two different roles. Having two different roles does, however, not necessarily imply an interaction level information flow. To differentiate whether it is the asymmetric role assignment or the interaction level information flow which fosters hyper-connectivity, a hyper-scanning study featuring an asymmetric role assignment but no interaction level information flow would be needed. To my knowledge no such study exists.

3.4. Hyper-Analysis

Now that I have discussed different aspects of machine-mediation, interaction, hyper-scanning and hyper-connectivity I want to conclude this chapter giving some insight in the more theoretical aspects of this research field. Various connectivity estimates have been proposed to measure within-participant connectivity and most of these estimates have later been applied in hyper-analysis, too. Properties, advantages, disadvantages and the validity of the results obtained with different connectivity estimates have ever since been a subject of active discussion. Various studies applying different estimates on either real, neural data or simulated data have been published (Burgess, 2013; Astolfi et al., 2006; Nolte et al., 2008; Kuś et al., 2004). The results of these studies are often contradicting one another. This is particularly an issue for EEG data, as its signals suffer from volume conduction. The details of this mixing process are unknown and can be assumed to be individually different, which makes it hard to model.

We are facing a predicament here: Studies on real neural data are always lacking a ground truth. We cannot objectively verify the results obtained with different estimates. Depending on the conditions under which the data has been recorded, we can in some cases make an educated guess about the existence of some neural connections, but we can never know.

On the other hand, using simulation data we can precisely determine which connectivity is present in the data at any time. One central point of debate, when modelling neural data for such analyses is the modelling of signal mixing. Various of the above mentioned studies use different models of this mixing process, some of which are more inspired/rooted in human physiology than others. In the end, however, it is impossible to verify any such model by comparison with the original – the human brain. We are lacking a ground truth.

Hence, the discussion about validity obtained with different connectivity estimates persists and there will be no conclusion for this discussion in the foreseeable future. However, while this discussion is far from over, it is still not irrelevant. I will pick two seminal contributions to the debate, which are related to my own work in several ways and discuss these here:

For the hyper-analysis during the first study of my thesis I used the Phase Locking Value (PLV) and the Phase Locking Statistics (PLS) as they were proposed in Lachaux et al. (1999). PLV is a bi-variant, symmetric connectivity estimate. It is introduced in some detail in section 5.6.2 on page 50. To actually compute any phase locking we need an estimate of the phase of the signals recorded. Basically all phase-true time-frequency transforms can be used for this. Two such transforms, Wavelet and Hilbert Transformation, have commonly been used as a basis for PLV computations. This caused Le Van Quyen et al. (2001) to compare PLV computations based on each of the two transforms.

They applied both methods on simulation data and on real neural data. This testing data encompassed single neuron recordings, inter-cranial EEG and classical EEG. For the first set of simulation data they used two coupled Hindmarsh-Rose model neurons. A second simulation data set was generated using a model introduced by Ermentrout and Kopell. While the PLV estimates based on Hilbert and Wavelet Transformation respectively showed some small differences, no systematic disagreement was found. Results were very much alike.

Furthermore, the authors applied both methods to inter-cranial recordings during epileptic seizure and EEG data sets of participants during a cognitive task. Again the authors found small differences between the two methods, but no systematic disagreement.

Due to these results the authors state regarding the two methods that "... one can safely conclude that they are fundamentally equivalent for the study of neuro electrical signals." This is reassuring, as both transforms have been widely used for PLV estimation and allows researchers to choose freely in future research.

Haufe et al. (2012) published one of the most thorough and critical assessments of connectivity analysis in recent years. Haufe et al assesses three different connectivity measures, three different approaches for statistical evaluation of connectivity results and three different methods for source estimation on the basis of simulated neural data. And en passant the authors also examine the impact of the choice of reference electrode and signal-to-noise ratio.

The authors identify one of the main problems for connectivity estimation in the effects of volume conduction. They argue that, while most connectivity estimates are based on the assumption that the cause precedes the effect (Granger Causality, compare section 5.4 on page 44) this is not necessarily the only and maybe not even the dominating factor influencing the results of a given connectivity estimate.

To clarify this point, they suggest to classify asymmetries found during the comparison of two time series into weak and strong asymmetries. Strong asymmetries are actually time-lagged asymmetries, while weak asymmetries are general interactions, in particular they are based on uni-variant properties of the time series compared. Consequentially the authors define a criterion for weak and strong asymmetries based on

the cross-covariance matrix C by stating that a process contains weak asymmetries if not all elements on the main diagonal are identical (i.e. $C_{i,i} \neq C_{j,j}$ for some i, j). Consequentially, a process contains strong asymmetries if the matrix is not symmetric (i.e. $C_{i,j} \neq C_{j,i}$ for some i, j). Note that there can (and for EEG data will) be both, weak and strong, asymmetries in a single process.

Of course weak asymmetries do not represent an information flow and are therefore (in most cases) irrelevant. Haufe et al propose a stunningly simple statistical approach to separate strong from weak asymmetries: For strong asymmetries the inversion of the time axis should lead to a flipping of the connectivity estimation (cause and effect change roles). The authors call this statistical approach time inversion.

The authors then compare three connectivity estimates: Phase Slope Index (PSI – compare section 5.6.5 page 53), Partial Directed Coherence (PDC – compare section 5.6.3 page 52) and Granger Causality based on MVar models (compare section 5.5 on page 45). Additionally they included different numbers of electrodes into their analysis.

The results of these experiments are discussed at length in the publication. Here I will highlight certain aspects of greater relevance to this thesis. First and foremost, the authors could reconstruct the correct information flow using PSI, regardless of the statistical method used. This was not true for Granger Causality nor PDC. On bi-variant data the correct information-flow could be reconstructed with PDC and Granger Causality in conjunction with certain statistical methods. Using time-inversion this was also possible with 19 channels, but not with 59 (the maximum number).

Feeding larger numbers of channels into the two multi-variant estimates generally seemed to lead to more spurious connectivity, blurring the actual connectivity pattern. The authors hypothesise that this might be an effect of over-fitting, as the number of weights in the MVar matrices grows quadratically with the number of channels.

Furthermore, the authors found that a change of reference of the electrical signal leads to a "rotation" of the identified information now connecting parietal and frontal regions on contra-lateral sides, rather than left-right direction (the ground truth). A reference free version of the data could be obtained by using a scalp Laplacian. However, this method does not counteract the effects of volume conduction. Source reconstruction or Blind Source Separation (BSS) are also reference-free and better suited to obtain interpretable connectivity estimations.

The publication of Haufe et al is one of the most critical simulation studies in the literature. Following its results and arguments the results of large portions of the connectivity studies of the past decades needed to be critically revised. Some question regarding the methodology of Haufe et al's publication remain, however.

For one, they are advocating the use of bi-variant connectivity estimates (PSI and bi-variant PDC and Granger Causality). Their test data is remarkably well suited for these estimates, as it encompasses exactly two neural sources. The authors address this issue in their discussion but the fact remains that they tested bi-variant estimates "in their comfort zone", hardly a fair comparison of multi- and bi-variant estimates.

The main disadvantage of bi-variant estimates is that they cannot identify indirect connections as such. Let us assume channel A only has an influence on channel B which in turn influences channel C. A bi-variant estimate will probably identify a direct connection between A and C, although there is none. Haufe et al address this point by stating: "Note that multivariant methods are commonly employed based on the consideration that the inclusion of more time series helps to rule out indirect connectivity between channels that are caused by a common confounder. However, that argument does not apply to EEG data, where all causal confounder contribute to all channels due to source mixing."(Haufe et al., 2012). I cannot agree with this argument, however. While it is certainly true that all neural sources will contribute to the signals measured at each of the sensors, this argument is of highly theoretical nature as the contribution of a given neural source might be very small for distant sensors, down to a degree where it is completely occluded by noise. In the PSI-based (bi-variant) results of this thesis I occasionally encounter seemingly significant connections which might be explained by an indirect connection mediated by a third channel and I cannot safely conclude whether or not the direct connection actually exists (as well).

Haufe et al also comment on the previously described predicament between real EEG data lacking a ground truth and simulation data suffering from a lack of proof of bio-physiological validity. They advocate a thorough verification of any connectivity estimate on simulation data before it is applied to EEG data. And the authors are certainly correct stating that: "... any connectivity estimation should achieve reliable performance on appropriately designed artificial data ..."(Haufe et al., 2012). However, they leave open what constitutes "appropriately designed artificial data". The validity of any given forward model could be debated and any such model might be refined indefinitely, lacking a possibility for verification. I do not generally deny the utility of simulation studies. The most interesting and path-breaking results of this study alone are well worth its effort. However, in my opinion, the point from which on a connectivity estimate has

been sufficiently tested on simulation data, such that it might be applied to real EEG data, is a proverbial line in the sand. There is no objective criterion to decide when a connectivity estimate has been thoroughly tested. The authors "encourage attempts on developing a generalized quantitative evaluations scheme for EEG-based connectivity analysis ..."(Haufe et al., 2012). The existence of such a standardised test for newly developed connectivity estimates would be reassuring for all researchers in the field. However, for as long as neural connectivity estimation is still cutting edge science, I would advocate to use the connectivity measures at hand on neural data to obtain new insights (and, of course, also test them on simulation data), while being aware of the risk that some of these insights might ultimately turn out to be in-valid.

Regardless of theses critics, I think Haufe et al. (2012) published one of the most remarkable and pioneering studies in the field in recent years and I want to adopt certain aspects for the hyper-analyses performed in this thesis (compare section 5.6 on page 47).

The two studies presented here showcase the ongoing discussion about the properties and validity of estimates of neural connectivity. In particular the two studies lead this discussion on two different levels: Haufe et al compare different connectivity estimates. These estimates are usually not based on raw data but on some other estimate derived from the raw data, e.g. a time-frequency transformation (PLV), a MVar model (PDC, DTF) or a Cross Spectrum (PSI). To obtain these, in many cases, different methods are available again and the fitness of each of these methods for the application on neural data can (and is) once more debated, as it was the case for Hilbert and Wavelet Transform as a basis for PLV in Le Van Quyen et al. (2001)

After careful consideration, I decided to use PLV for the first study and (taking into account the publication of Haufe et al. (2012), which was only being published when I conducted my first study) a combination of two complementary estimate for the second study: PDC and PSI. Other methods could have been used alternatively (e.g. DTF instead of PDC) and there is, in my opinion, no clear evidence from the literature that either is superior.

Hyper-scanning, machine-mediation and BMIs. My thesis is located between these three landmarks and in this chapter I aimed to illuminate different aspects relevant for this thesis by citing literature from these fields. In the next chapter I will try to further localise my own research within this terrain.

4. Contribution of this thesis

Having laid out the surrounding in which my own research resides, I will now outline what distinguishes this thesis from other research and why it is a novel contribution to this field of science.

This thesis brings together two formerly distinct topics: The research about neural foundations of interaction and machine-mediated interaction, particularly using BMIs. Both have been active fields of research for years and there are countless contributions on these topics. To my knowledge, however, studying machine-mediated interaction settings as a special case of human interaction is new to neuro interaction research. In particular the use of BMIs, which are the most direct connection between humans and machines, as a technique in neuro interaction research is unprecedented.

This thesis aims to pave a way for this new approach. The first experiment (HExMInE) aims to verify the general feasibility of the approach, in particular to evaluate whether or not neural correlates of interaction (in particular hyper-connectivity) still occur when interaction is machine-mediated. The second experiment (iCusss) then is intended to showcase the potential of this approach in a fully featured (BMI) machine-mediated interaction experiment.

The thorough exploration of this approach's potential is undoubtedly by far too ambitious for a single thesis. The same holds for the (thorough) investigation of the impact of machine-mediation on human interaction and its neural correlates. As this research advances, though, the results will undoubtedly shed light on important aspects of human interaction and human-machine interaction and can be expected to have major impact on the design of future machine-mediation technology. Therefore, this new approach addresses highly relevant research goals.

Besides these rather general aspects, both experiments address specific research questions, relevant for neuro interaction research: In the HExMInE experiment I compare neural connectivity during interaction and during solo action of one participant. In the iCusss experiment I compare neural connectivity during cooperation with independent, concurrent action. The main method for this evaluation is hyper-scanning and -analysis aiming for neural connectivity – within a participant and across participants. This connectivity analysis is conducted on different frequency bands, addressing several of the standard neural rhythms in human EEG (compare section 2.2.5 on page 15) and thereby contributing to the interpretation of their roles. These roles have been another active research topic for years and still evidences for new facets regarding their interpretation/function is being accumulated.

Additionally, different scientifically relevant topics are addressed as a side-effect when pursuing my main research goals: For the HExMInE experiment a new type of training for the mental strategy of Motor Imagery, often employed in BMIs, is tested. For the two experiments two different robots which are diametrical in many aspects are employed. In particular one is highly anthropomorphic while the other is the exact opposite. Up to now, there are relatively few publications on BMI usage of more than one participant simultaneously. This thesis contributes indirectly to the field of multi-user BMIs by demonstrating its feasibility even for very complex settings. In the course of the PhD project the development of the UBiCI BMI software framework was advanced. And finally, I employ two different BAPs for the BMI control, both of which are correlated with some function vital for interaction (P300 $\leftrightarrow$ attention and ERD $\leftrightarrow$ motor co-representation) with the intention to allow space for interesting side effects to occur in the neural recordings.

This thesis deviates from two different, well explored paths of research at once, converging in and pioneering a brand new direction of neuro interaction research. I hope this path will lead research to the neural foundations of human interaction from a new, different angle, allowing an illumination of new aspects of what Schilbach et al called "the dark matter of social neuroscience" (Schilbach et al., 2013).

5. Methods

In this chapter I am going over i.) basic signal properties and ii.) processing of multi-variant signals/time series in general and electro-physiological signal in particular. The focus will be on the identification of commonalities of signals and on questions of mutual influence of signals (connectivity).

I start with some discussion of basic properties of signal, i.e. amplitude, frequency and phase and the topic of noise in recordings of physical signals in section 5.1, below.

After that, I will discuss the mathematical methods used for the online classification of EEG signals performed within the BMI systems of this thesis in section 5.2 on page 38. Namely I will introduce the Principal Component Analysis on page 39, the Common Spatial Patterns Analysis on page 40 and the Fisher Discriminant Analysis on page 41.

Time-frequency transforms are not the only transformations commonly applied to signals. Another class of transformations aims to transfer signals from sensor into source space. In some systems multiple sources generate signals and these signals contribute differently to the signal recorded at the sensor (mixing, volume conduction). When multiple sensors record the signals from the same sources with different mixing (as it is the case in EEG) various methods exist, which aim to de-mix these signals, i.e. transferring the signals from sensor into source space. For EEG recordings encompassing relatively few channels, as for this thesis, a transformation into source space can be done, but an exact localisation of the sources in the cortex cannot reasonably be achieved. This makes a special class of source estimators particularly interesting: Blind Source Separation. The most prominent and widely used member of this family is the Independent Component Analysis. I will use this i.) for artefact removal throughout this thesis and b.) for the analysis of experiment data of the iCusss experiment and discuss it as a method in section 5.3 on page 42.

Next I will discuss interdependence between signals: connectivity. First, I will introduce the concept of Granger Causality in section 5.4 on page 44 which gives us a mathematical foundation upon which the mutual influence of two channels on one another can be evaluated. Then a linear, multi-variant modelling approach which can be used to compute a variety of measures characterising the signals, including different estimates for Granger Causality, will be the topic of section 5.5 on page 45. Then I will give a broader overview over different connectivity estimates by discussing a taxonomy of connectivity estimates in section 5.6 on page 47 followed by the individual discussion of some selected estimates. In section 5.7 on page 53 I will present methods for statistical evaluation of the results of a connectivity estimation.

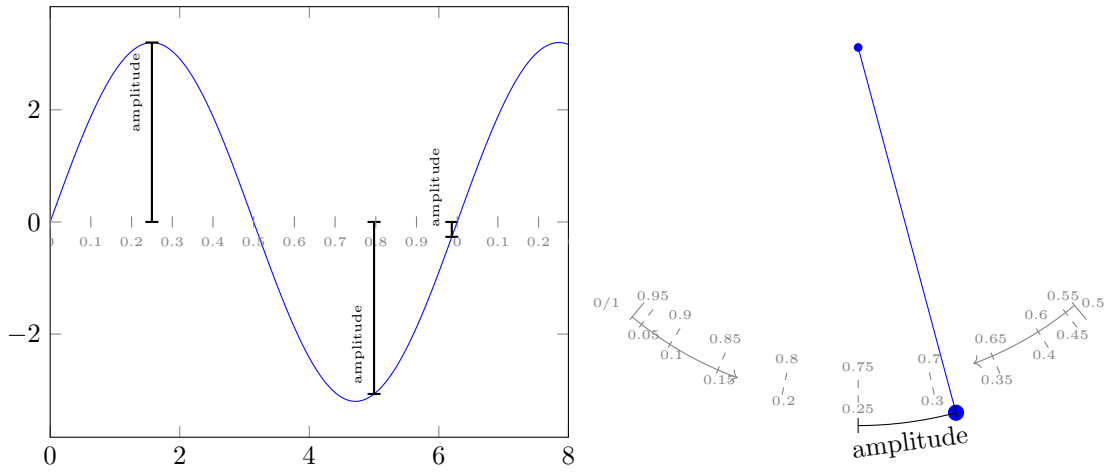
Starting with my Master thesis, I was involved in the development of a software framework for online classification of EEG data and BMIs. This framework was developed further during the course of my PhD and was also used for the construction of the BMIs for this thesis. I will conclude this chapter by going over its most important features and structures in section 5.8 on page 55.

5.1. Basics of Digital Signal Processing

Any (real-valued) variable which changes over time can be regarded as a signal. Often it is additionally assumed that a signal is zero-centred. If that is not the case, e.g. if the signal has a trend to grow or decline indefinitely, de-trending methods can be applied. A special class of signals are periodic signal. These are signals which repeat their values after some finite period. Each sine- or cosine-function and any combination of such signals is a periodic signal. Real, physically recorded signals are seldom periodic. Still many terms important for digital signal processing are best explained at periodic, even sinusoidal, signals first and expanded to general signals later.

When recording and processing physical signals one central issue which needs to be considered is the presence of noise in such signals. I will briefly discuss this issue in section 5.1.2 on page 35.

Physical signals are not generally assumed to have a single or even a discrete number of frequencies. Instead they have a spectral density which describes, which frequency bands are more prominent in the signal and which are less. For any frequency band the (generalised) phase, amplitude, power and other parameters can be estimated using time-frequency transformations. I will discuss time-frequency transformations in general and two popular methods for time-frequency transformation in particular in section 5.1.3 on page 36.



(a) Amplitude and phase in a basic sine function.

(b) Amplitude and phase in a pendulum.

Figure 5.1.: Amplitude (black) and phase (grey) in the examples of a sine function and a swinging pendulum. For better readability the phase is given in the interval $[0, 1]$ here.

5.1.1. Signal Properties

Three of the most basic terms to describe a (sinusoidal) signal are its *frequency*, *amplitude* and *phase*.

The amplitude describes the degree to which the signal is perpetuated from 0. Basically it is the value of the signal at that time.

A (sinusoidal) signal's frequency is the number of iterations it undergoes per time unit. A sine which completes three and a half iterations in one second would have a frequency of $3.5Hz$.

A sinusoidal signal repeats itself after a fixed time span t. The phase of such a signal basically describes how much of its period the signal has already finished. It could be described by a value $p \in [0, 1]$ where 0 would be the start of a repetition and 1 would be the end of the repetition (at time t). As the signal's phase has a circular value range, similar to that of a geometric angle, it is also common to express it by a value $p \in [0, 2 \cdot \pi]$ (radiant).

For a sinusoidal signal the frequency does not change, while the phase and the amplitude are functions over time. A classical analogy to explain these terms is a swinging, undamped pendulum. Here the amplitude would be equivalent to the angle by which the the pendulum is perturbed from vertical orientation. The phase would be the point of its entire movement at which the pendulum is at a given point in time. The frequency (in Hz) would be the number of complete swings (back and forth) is completes per second.

The amplitude and phase of such a signal can be described in a comprehensive way when using complex numbers. In such a representation the phase of a signal would be the imaginary part of the complex number, while the amplitude would be the real part. Plotting a sine or cosine would then result in a zero-centred circle on the complex plane.

From this representation we can derive another common characteristic of a signal: its power. The amplitude is a fast-changing property. For a sinusoidal signal it would be convenient to have a characteristic which describes how much the signal is perpetuated from zero at its maximum. This is equivalent to the signals power and it can be computed by taking the Euclidean Norm of the complex number (the length of the vector on the complex plane). For a sinusoidal signal this is constant, because a sinusoidal signal describes a (zero-centred) circle on the complex plane. The power of the signal is the length of the complex number which is the (constant) radius of that circle.

But how can these basic terms of frequency, phase, amplitude and power be applied to non sinusoidal signals, particularly to physical measurements? The basic idea is to use some type of (complex) model function, called kernel, and to convolute this with the signal. The kernel could be a sine or cosine, but many other different kernel function with different properties of the resulting estimates exist. The phase and amplitude are then directly derived from that of the kernel. In Figure 5.2 I filtered a piece of EEG data to the α-band and computed phase, amplitude and power estimates based on a Wavelet kernel (compare section 5.1.3 on page 37).

5.1.2. Noise

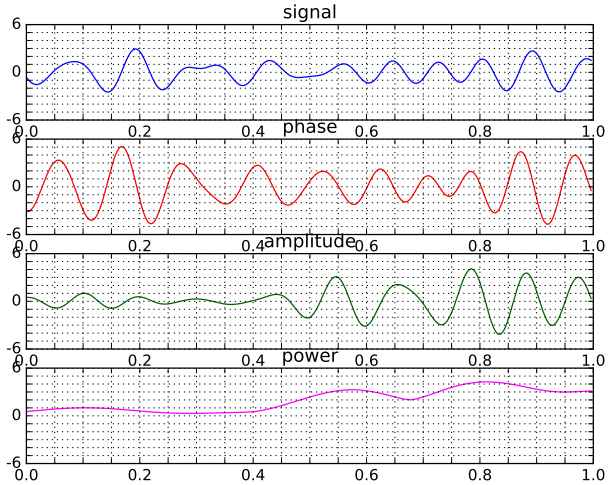


Figure 5.2.: A one second epoch of EEG data filtered to α-band and the phase, amplitude and spectral power estimation based on a Wavelet Transformation using a complex Morlet Wavelet.

When measuring signals in real world scenarios, one rarely has only one signal source. Signals from different sources mix. Some of these sources might be desired others are regarded as noise. For instance when measuring EEG signals many individual sources within the brain generate signals and all these signals mix before reaching the sensor/electrode. Each activation of muscles generates electrical signals. These are regarded noise in EEG and signal in EMG. The categorisation between signals and noise is subjective and depends on the task / analysis at hand. Outside sources such as computers, lights and even power lines present in the lab, generate an electrical signal which also mixes with the neural signals.

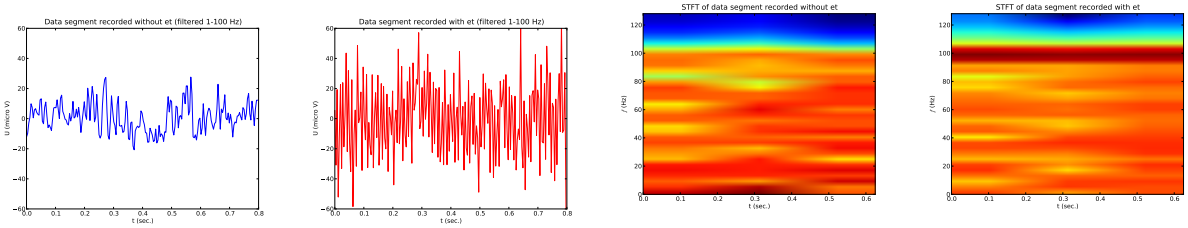
Some of these signals are stronger than others. The brain generally produces signals which are rather weak and which are further diminished by tissue and the skull before they reach the electrodes. The power lines and electrical devices generate a (relatively) strong signal. The ratio of the power of the target signal and the power of the noise is regarded to as the Signal-to-Noise-Ratio (SNR). This is a measure especially critical for EEG recordings as for EEG the SNR is very unfavourable. Luckily there are a number of ways to improve the SNR. The most sophisticated way is of course to avoid noise wherever possible. This means removing electrical devices and conducting experiments in a shielded room. I did not take such measures a.) because in our lab we generally want to push towards commercial and home use of BMI systems and in a home use scenario such restrictions are not suitable. b.) For my experimental settings involving one or more robots such restrictions are particularly impracticable.

There are computational methods to improve the SNR in recorded data. The most basic method is frequency filtering, which can, actually, be performed either in hard- or in software. In Europe the power grid operates with an alternating current of $50Hz$. A hardware filter within EEG recording devices removing the band from $48Hz$ to $52Hz$ will, therefore, already remove most of the noise from the power lines and many electrical devices. Although there is some neural activity taking place in that band, most of the relevant neural activity, in particular most of the rhythms identified in EEG (compare section 2.2.5 on page 15) take place in the frequency range $< 45Hz$.

In EEG the noise produced by the muscles moving the eye-ball and opening and closing the eye-lid are particularly prominent. Therefore, the electrical activity of these muscles is sometimes recorded (EOG) alongside of EEG and subtracted from the EEG signal.

Finally one can use a method called Independent Component Analysis (ICA) to remove most of the muscle artefacts, especially from eye movement and eye blinking. I will discuss the ICA in section 5.3 on page 42.

In preparation of the first experiment (see section 6 on page 61) I encountered a source of noise usually not considered during BMI experiments. For that experiment I used eye trackers in addition to the EEG recordings. The idea was to find correlations between the participants' gaze and their neural activity. However, when using the eye tracker I found the EEG data was heavily contaminated with noise way beyond the usual extend. After I realised the presence of the eye tracker caused that noise I started to shield the eye trackers with aluminium foil, with no visible effect. In the end a co-worker pointed me to the possibility that the noise might not be transmitted "over air" as an electrical field emitted by the eye tracking device, but over the ground wire of the power grid. Luckily the SMI eye trackers I used, could run on battery. And indeed the heavy contamination vanished the moment I pulled the eye trackers' plug. Figure 5.3 shows this effect. Thus, this special form of noise was generated by the eye tracking system, which induced this noise into he ground wire of the power grid. The EEG device which was connected to the same power line (and ground) then picked up this noise. Indeed there are hi-fi devices of the upper class segment which first correct/filter their power input to a $50Hz$ sine before this power then supplies the amplifier, aiming for an enhancement of the audio quality.



(a) A segment of data with no eye-tracker.

(b) A segment of data with eye-tracker.

(c) A spectrogram of the same data segment without eye-tracker.

(d) A spectrogram of the same data segment with eye-tracker.

Figure 5.3.: The impact of the SMI eye-tracker when it was plugged to the same power line as the Gugler EEG Device. I recorded $800ms$ of data with and without the eye-tracker. The difference is easily visible in raw data. The spectrogram pin-points the mayor contribution slightly below $100Hz$. This part of the frequency spectrum was filtered out before classification. Still cross-validation of the recorded data confirmed that the eye-tracker-induced noise was disruptive to classification accuracy. For this test, I did not apply the canonical notch-filtering around $50Hz$ as I wanted to capture as much noise as possible to identify the mayor source of noise.

Another source of noise, is to mention when it comes to digital signal processing: When digitalising a signal this means that the signal is sampled in discrete time steps. When choosing the sampling rate one has to keep in mind that only signals up to half of the sampling rate are properly represented and can be accessed e.g. by software frequency filters or other analytical procedures. This is called the Nyquist frequency . For all of my experiments I chose a sampling rate of $256Hz$, which allows me to access the frequency spectrum up to $128Hz$. It is, however, important to note that signals above this frequency still contribute to the recorded signal and enter the recording as noise which cannot easily be removed, in particular not by means of digital frequency filtering. This process is often called aliasing.

5.1.3. Time-Frequency Transform

In section 5.1.1 on page 34 I talked about how phase, amplitude and spectral power density of a real-world signal can be estimated by convolution with a kernel function.

It is actually pretty common to compute such convolutions with the kernel function for different frequencies. This is commonly referred to as a time-frequency transformation of the signal. It is basically the computational equivalent to a prism splitting the light into different colours.

If the kernel function is real, the resulting time frequency transform is real. This splits the signal into its spectrum, i.e. the activity in different frequency bands. Throughout this thesis the discrete frequency points considered in an analysis are denoted by a ω. Using a complex kernel function, one can obtain a complex time-frequency transform which encodes the signal's amplitude and phase as a complex number as shown in section 5.1.1 on page 34.

Different time-frequency transformations exist. Their main difference is the kernel they use. The Fast Fourier Transformation uses sine and cosine, Hilbert Transformation uses Hilbert kernel or Cauchy kernel (depending on its definition) and for the Wavelet Transformation different kernel functions, called Wavelets have been proposed and evaluated. The different transformations have different properties, depending on their kernel function. It is important to keep in mind that the estimations of amplitude and phase are dependent on the kernel used. Hence, comparing time-frequency transformations obtained with different kernels will most probably not yield any meaningful results.

Taking the euclidean norm of each entry of a complex time-frequency transformation (compare spectral power, above) results in the spectral density of the signal, representing how much each frequency contributes to the overall signal at a given time.

If for a given time-frequency transformation there exists an inverse transformation, this can be used for frequency filtering: One computes the time-frequency transform, sets the entries corresponding to undesired frequencies to 0 and applies the reverse transform.

Fourier Transformation

The Fourier Transformation is probably the most used time-frequency transform. It tries to approximate a given function $f(x)$ with a linear combination of a set of sine and cosine functions (its kernel-function):

$$1, cos(x), sin(x), cos(2x), sin(2x), ..., cos(nx), sin(nx) \tag{5.1}$$

$$f(x) \approx g_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n}(a_k cos(kx) + b_k sin(kx)) \tag{5.2}$$

The coefficients a_k, b_k are chosen such that the quadratic error is minimised. Two versions exist: If the signal is given as a continuous function, the quadratic error is computed as an integral, whereas if the signal is only represented by discrete samples (the common case in digital signal processing) the error to minimise is the sum of the squared errors over all samples. This is also called Discrete Fourier Transformation. The coefficients can be computed with the formulas

$$a_k^* = \frac{2}{n}\sum_{j=0}^{n} f(x_j)cos(k_j), k = 0, 1, 2, ... \tag{5.3}$$

$$b_k^* = \frac{2}{n}\sum_{j=0}^{n} f(x_j)sin(k_j), k = 1, 2, 3, ... \tag{5.4}$$

The main issue here is, that the runtime for the computation of the coefficients a_k, b_k grows quadratically with n (the order of the transformation). However, for the special case that n is a power of 2 and that the transformation is actually a complex transformation (complex sine and cosine), a special algorithm called Fast Fourier Transformation can be applied. This is based on the fact that from a Fourier Transformation of order 2N the transformation of order N can be efficiently computed. The FFT has runtime $O(n \cdot log_2(n))$ (rather than $O(n^2)$) (Schwarz and Köckler, 2006). FFT gives exactly one (complex) estimate per frequency point representing the complete time-span. The result is, hence, a complex vector.

The FFT samples the frequency domain from $0Hz$ to the Nyquist frequency equidistantly. Hence, the frequency resolution depends on the length of the data epoch. E.g. using a sampling rate of $256Hz$ the Nyquist frequency would be $128Hz$. For an epoch of 1024 samples length 1024 samples are distributed equidistantly over the interval $[0Hz, 128Hz]$ resulting in a frequency resolution of $\frac{1}{8}Hz$.

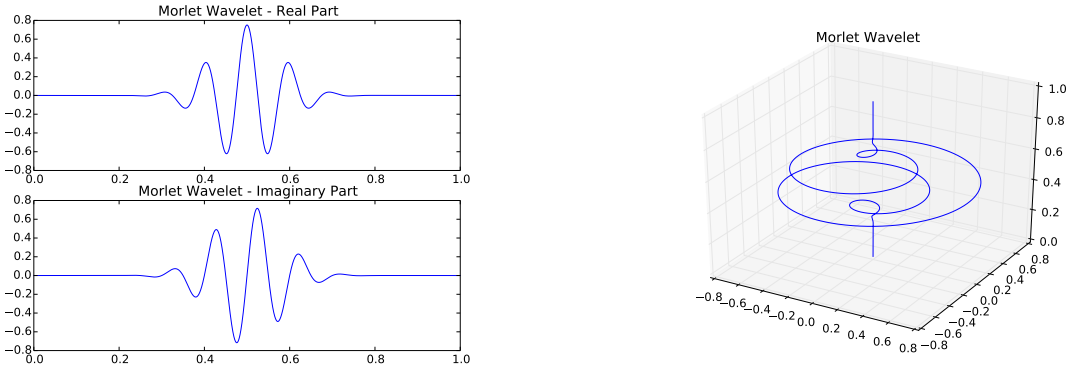
FFT per se has no time-domain, i.e. returns only one estimate per frequency point. To obtain an actual time-frequency transformation, windowing is applied, i.e. FFT is applied to overlapping epochs (windows). This is called Short Time Fast Fourier Transformation (STFT). The result is a 2D complex matrix (time $\times$ frequency).

Applying FFT to EEG data has some restrictions. FFT is mostly used in EEG and BMI context for frequency filtering and estimation of the power spectrum. Both can reasonably be done, although better choices exist. For other applications FFT on EEG data is infeasible. In particular the phase estimate of FFT on EEG data is invalid and should generally not be used. The reason is, that FFT assumes that the spectral power density of the signal does not change throughout the entire epoch (as for FFT's kernel functions). This is generally not true for EEG signals and this violation of FFT's prerequisites lead to a false phase estimation.

I already stated, that FFT is often used for (software) frequency filtering. For this it is important to note that FFT as an operator can be revered: iFFT (compare section 5.1.3 on page 36). For EEG data (and many other types of real world data) FFT has, however, the tendency to produce an over-swinging towards the data epoch's ends. This can easily be countered, by using a longer epoch and subsequently cropping the ends of the epoch.

Wavelet Transformation

Another commonly used transformation is the so called Wavelet Transformation. This is basically a convolution of the raw signal with a kernel function called the Wavelet to which different scaling factors are applied to obtain different frequencies. Sanei (2013) compares it with "a mathematical microscope with properties that do not depend on the magnification" (Sanei, 2013). There are a variety of wavelets with



(a) The complex morlet wavelet separated in real and imaginary component.

(b) The complex morlet wavelet plotted in 3D, where x,y are real and imaginary part, respectively.

Figure 5.4.: The complex morlet wavelet used for the analysis in this thesis.

different properties. They usually converge to 0 towards both ends. This is important as it allows a Wavelet Transform to properly represent signals with a non-static spectrum.

When the wavelet is complex the resulting transform is complex. Another important property of many wavelets is, that when equidistant scaling factors are applied, the resulting frequency spectrum has a non-equidistant (logarithmic) distribution of frequency sampling points. In general, higher frequencies are sampled more tightly. As a consequence of this, the time resolution is lower. Considering that for lower frequencies the localisation in time is generally more diffuse, this is actually a far more adequate representation of the frequency spectrum as compared with an equal sampling of the frequency space.

Although the Wavelet used for convolution with the signal is not fixed and many different Wavelets with different properties have been proposed over the years, in the analysis of EEG signals the complex version of the Morlet Wavelet has become very popular (Sanei, 2013) :

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2} + i2\pi b_0 t} \tag{5.5}$$

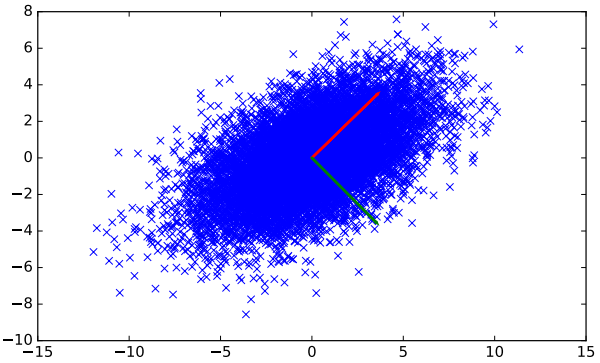
Where $b_0 > 0$ is constant. The Morlet wavelet is depicted in figure 5.4. Similar to the FFT a discrete version (DWT) and a continuous version (CWT) of the Wavelet Transformation exists. Also like for FFT a method for the reconstruction of the signal from a frequency transform with Wavelet Transformation (inverse operator) exists.

5.2. Automated BAP Recognition

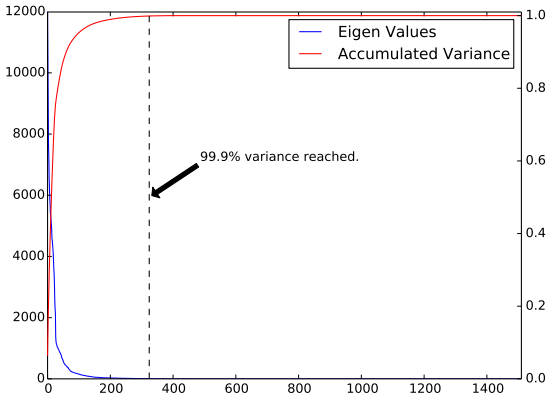
BMIs generally employ some type of BAP. By adequate data analysis BMIs try to infer i.) whether or not the BAP is present, ii.) which of several BAPs is present or iii.) at which location (on the cortex) the BAP occurred.

In the clinical application of EEG, usually neurologists inspect the EEG data visually and use their extensive experience and the amazing pattern recognition capabilities of the human brain to analyse the data. For BMI applications, classification needs to be done in the range of a few seconds, at best. Additionally some of the activity patterns are not visible in single epoch data, not even by experts. Hence, although automated EEG signal classification is pretty uncommon in clinical application, it is the basis for any BMI system.

In most cases the EEG signal classification in BMIs is a three stage process: First, frequency filtering is applied, filtering to a preferably narrow band, in which the target BAP can be found. This is often done using FFT or Wavelet Transformation (see section 5.1.3 on page 36). Second, the data is rearranged and the number of dimensions is reduced in order to make the distinguishing features more salient. This is called Feature Extraction. Finally, the data is classified. In the past 30 years many different methods for feature extraction and classification have been tested, some of which have advantages for some BAPs but may be completely infeasible for others.



(a) The PCA applied to a 2D point cloud with a normal distribution. The first component (red) points in the direction of the maximal variance. The second component is orthogonal to the first component. In the 2D case this means it is fully determined. For high dimensional cases it would point into the direction of maximal variance in the $(n-1)$-d subspace orthogonal to the first component.



(b) The eigenvalues/variance from a real (P300) data set. Shown are only the first 1500 of originally 4096 values. The variance captured by the different principal components (blue) drops steeply. The percentage of the total variance captured by the first n components, accordingly, increases steeply. The first 324 dimensions capture already more than 99.9% of the data variance. Hence the data's dimensionality can be reduced to 21.5% while maintaining 99.9% of its variance/structure.

Figure 5.5.: The direction of the principal components for a simple example and a typical decay in variance captured by the principal components.

In this thesis the BMI itself and the classification of BAPs is not the subject of study itself and, therefore, I chose pretty much standard regarding the mathematical methods. These are Principal Component Analysis, which I will introduce in section 5.2.1 below, and Common Spatial Pattern Analysis, to be discussed in section 5.2.2 on page 40 for feature extraction and the Fisher Discriminant Analysis, which will be covered in section 5.2.3 on page 41 for classification.

For basically all BMIs these mathematical methods need to be parametrised per participant. Before the BMI can actually be used (labelled) training data needs to be collected, usually by confronting the participant with the same system, but instructing the participant which actions to take in each iteration (rather than letting him/her choose freely). This (labelled) data is then used to compute matrices and other parameters used by the methods employed in the classification system.

5.2.1. Principal Component Analysis

The Principal Component Analysis (PCA) is basically a basis transformation and subsequent projection of data on a subspace of the original space. Figure 5.5a shows a 2D point cloud and the new (orthogonal) basis identified with the PCA. When dealing with high dimensional data one often finds that most of the variance of the data can be captured within relatively few dimensions. The remaining dimensions can then be discarded without loosing much information. This is the basic idea of many compression algorithms e.g. used for image compression.

I will follow the argumentation given in Bishop (2006) for a formal introduction of PCA. Let D be the data matrix and $C = D^T \cdot D$ be the covariance matrix of the data. If we wanted to reduce the data to just one dimension using PCA, we choose a normalised vector $\vec{w}$ such that the variance of the projected

$$S = \vec{w} \cdot C \cdot \vec{w}^T \tag{5.6}$$

data is maximised.[1] For our case the fact that $\vec{w}$ is normalised ($\vec{w} \cdot \vec{w}^T = 1$) leads to (using a Lagrange-

[1] This is similar to the Linear Discriminant Analysis (see section 5.2.3 on page 41), but for (LDA) one tries to choose $\vec{w}$ to maximise a different criterion (class separability, not variance).

Multiplier)

$$S \cdot \vec{w} = \lambda \cdot \vec{w} \tag{5.7}$$

And thus $\vec{w}$ needs to be an eigenvector of S. Additionally it follows that

$$\vec{w} \cdot S \cdot \vec{w}^T = \lambda \tag{5.8}$$

meaning that the variance of the projected data is given by the eigenvalue λ. Hence, to find $\vec{w}$ all we need to do is compute the eigenvectors and eigenvalues of the covariance matrix of the data C and find the eigenvector $\vec{w}$ with the largest eigenvalue. If we want to have a M-dimensional subspace rather than just one dimension we just choose the M eigenvectors with the largest M largest eigenvalues.

For applications on real data in many cases one can observe a sever drop in variance/eigenvalues, if we sort them by decreasing magnitude. Or, phrased differently, only relatively few dimensions already capture most of the data's variance. In practice that means that one usually defines a threshold (say 99.9%) for the fraction of the data's variance one wants to retain. Figure 5.5b shows this for an exemplary EEG data set.

By projecting the original data onto the sub-space spanned by the first M eigenvectors, one obtains a good approximation of the data. This way data encompassing thousands of dimensions can be reduced to a few hundred dimensions or below in many real world applications .

For P300 classification PCA is mostly used as follows: Say we have n epochs of length s (in samples) with c channels.

During Parametrisation:

1. We linearise each sample of the training data resulting in an one-dimensional vector of size $s \cdot c$. If each segment is one second long recorded with 16 channels at a sampling rate of $256Hz$ we would end up with vectors of length $1 \cdot 256 \cdot 16 = 4096$.
2. The PCA then identifies a set of 4096 orthogonal vectors, the eigenvectors, which form a new basis for the vector space. It also computes the corresponding eigenvalues and sorts eigenvectors and eigenvalues by decreasing eigenvalue.
3. The PCA now computes how many dimensions are needed to retain the desired amount of variance in the data and generates a matrix of the first k basis vectors.

During Usage: The data recorded during the actual usage of the BMI is linearised the same way. After multiplication with the (pruned) eigenvector matrix the dimensionality of the linearised data is greatly reduced.

Reducing the dimensionality of the data in BMI has little to do with memory requirements of the EEG data. Many machine learning algorithms need more training data the more parameters/weights are available to them. And the number of parameters available depends on the size of their input data. Hence, reducing the dimensionality of the data makes parametrisation of the actual classification algorithm (LDA, in our case) much easier. Phrased differently, we extract the features of the data and do not burden the LDA with irrelevant portions of the signal. As a consequence, with fewer input dimensions the chance of over fitting of the FDA is reduced.

5.2.2. Common Spatial Pattern

The Common Spatial Pattern Analysis (CSP) is a mathematical method which is, again, a basis transformation. Other than the PCA, which is agnostic of the actual classification problem, the CSP needs training data which represents two classes (labelled data). Assuming we have c channels, the CSP uses this data to first compute a matrix of size $c \times c$. The $c \times t$ data epoch is multiplied with the matrix. In the resulting data the first channel has the property that its variance is maximised when the data epoch is similar to the data of class 1 from the training data and minimised when the data is similar to class 2. The last channel is just the other way round: Data similar to the class 2 of the training data leads to large variance in that channel and data similar to class 1 leads to minimal variance in that last channel. Towards the middle channels these trends are diminished.

Hence, when using CSP one usually eliminates the middle columns of the CSP matrix as the corresponding channels carry little information (ratio between class-variances ~ 1) regarding the classification problem.

How many channels are to be deleted depends on the data properties. For ERD classification (the problem for which I will use the CSP later on) it is assumed that more than six residual channels have little benefit for the classification, independent of the number of original channels.

The CSP matrix is computed as follows:

$$J(w) = \frac{W^T D_1^T D_1 W}{W^T D_2^T D_2 W} = \frac{W^T C_1 W}{W^T C_2 W} \tag{5.9}$$

Where W is the CSP matrix, D_1 and D_2 are the data of class one and two, respectively, and C_1 and C_2 are (consequentially) the covariance matrices for the data of the two classes. The computation of W can be reduced to the generalised Eigenvalue equation

$$C_1 W = \lambda C_2 W \tag{5.10}$$

$$C_2^{-1} C_1 W = \lambda W \tag{5.11}$$

The CSP matrix is then composed of the eigen-vectors of this problem and the ratios between the class-variances are the eigenvalues. When sorting the eigenvectors according to eigenvalues one gets the structure described previously (Ramoser et al., 2000; Sanei, 2013).

5.2.3. Fisher Discriminant Analysis

The Fisher Discriminant Analysis is a linear classifier. There are two ways to interpret linear classifiers:

(a) The classifier spans a hyperplane through the data space. Any data point which is on the one side of the hyperplane is classified as class one and all others are classified as class two.

(b) The analysis defines a line through the data space (the normal of the aforementioned hyperplane) and projects all data points onto that line. It further defines a threshold. Any data point whose one dimensional representation is below this threshold is classified as one and all others are classified as two.

Those two views are equivalent, however, following the approach presented in Bishop (2006) I will argue from the second perspective: The high dimensional input data is reduces in its dimensionality to just one dimension by orthogonal projection.

During parametrisation this one dimension must be chosen such that it allows for an optimal separability of the projected (training) data, i.e. the direction of the dimension must be chosen such the overlap of the two classes' projected data points in minimal. To achieve this, two goals need to be pursued:

For one we want the projected means of the classes to be far apart. Assuming that $\vec{m_1}, \vec{m_2}$ are the means of the two classes before projection, let $\vec{w}$ be the normalised vector with which we multiply the data for projection into 1d space. Let m_1, m_2 be the means of the projected data. Hence we want

$$m_1 - m_2 = \vec{w} \cdot (\vec{m_1} - \vec{m_2}) \tag{5.12}$$

to be large. However, just choosing $\vec{w}$ such that it maximises the distance of the projected means does not guarantee optimal separability. Choosing $\vec{w}$ such that $m_1 - m_2$ is maximal might lead to a high variance in one or both of the projected data classes, i.e. there might be a choice for $\vec{w}$ for which these variances were considerably smaller up to a point where the projected data might be better separable than for a $\vec{w}$ exclusively driven by $m_1 - m_2$. Let s_1, s_2 be the variances of the projected data of the two classes. We want to have a large distance between class means and a low overall within-class variance. There are many ways how to weight these two goal against one another. The Fisher Discriminant aims to maximise the following therm:

$$J(w) = \frac{(m_1 - m_2)^2}{s_1^2 + s_2^2} = \frac{\vec{w} S_B \vec{w}^T}{\vec{w} S_W \vec{w}^T} \tag{5.13}$$

Where S_B is the between-class variance and S_W is the within-class variance. This can be rewritten such that

$$\vec{w} \propto S_W^{-1} \cdot (\vec{m_1} - \vec{m_2}) \tag{5.14}$$

Now that we have fixed a direction for the 1d projection, all that needs to be done is to determine a threshold. This can be done by considering thresholds between each two projected data points, determine the number of misclassification given each threshold and thereby find the optimal threshold for the training data.

5.2.4. Combining Classification Results

It is common practice for P300-based BMIs to repeat the stimulus presentation several times and to combine the classification results of these iterations to increase the classification accuracy. Typically, real values which are assumed to represent the classification confidence are summed per option the participant can choose from. These sums are also called the score of the options. For FDA classifiers the distance of a given data sample to the dividing hyper-plane is often used as a classification confidence. When one score exceeds a certain threshold that option is returned as a result.

A more sophisticated approach was proposed by Lenhardt et al. (2008). The basic idea is to scale the scores linearly to a (common) range from 0 to 1. Then all these scaled scores are summed up (over all options). When this over-all sum is lower than a certain threshold the item/option with the highest score is returned as the final result for the corresponding decision request.

The idea behind this is as follows: Choosing an option when it has a high/the highest score/exceeds a threshold might easily lead to misclassification, if other options have a score which is only slightly lower. Hence, Lenhardt et al. (2008) also aim for the other scores to be relatively low. In the proposed method the "winning score" would (always) receive the scaled value 1. If the other scores are relatively low these would be scaled to values $\epsilon_1 ... \epsilon_{n-1}$ close to 0, with n the number of options available .

$$sl = 1 + \sum_{i=1}^{n-1} \epsilon_i \gtrsim 1 \tag{5.15}$$

When, on the other hand, several other options receive high scores as well, the sum will be larger

$$sl < sh = 1 + \sum_{i=1}^{c} \gamma_i + \sum_{i=1}^{n-c-1} \epsilon_i \gg 1 \tag{5.16}$$

where c is the number of options whose scaled scores $(\gamma_1 ... \gamma_c)$ are close to 1.

Hence, if the overall sum of (scaled) scores is close to 1, the method would terminate the decision and pronounce the option with the highest score the outcome of the decision. If the overall sum is much larger than 1 the method would notify the stimulus presentation that it had not yet decided for any of the options and the stimulus presentation would have each of the icons representing the options flash once more, allowing processing pipeline to have more data to decide on. The exact threshold below which the method would render a decision is determined on a per-participant basis on the training data by determining the average (scaled) score at which the classification was correct for the first time over all decisions.

5.3. Sensor and Source Space

Data can be projected/transformed into several spaces. One class of transformations assumes that several sources contribute to signals which are recorded by multiple sensors at multiple (spatial) locations. As the sources (and the signals they produce) are usually unknown, a set of hypothetical sources is assumed and these transformations try to attribute parts/components of the recorded signals to these hypothetical sources. This is often referred to as a transformation from sensor into source space or as source separation.

A common analogy is a room full of people chatting (a party) in which we place a series of microphones. The audio-stream from the different microphones are the data in sensor space. Applying a transformation into source space means to aim at de-mixing the conversations and to obtain audio streams such that each audio stream only contains the utterances of one person in the room.

The mixing of the signals may have various properties. If the mixing of signals can be described by multiplying a fixed matrix this is referred to as linear mixing. If not only the sources and their signals, but

also the mixing process are unknown, the problem of reconstructing the sources is commonly referred to as Blind Source Separation (BSS). For EEG signals the mixing process and its statistical properties are subject to ongoing debate, mostly because researches are lacking a ground truth for evaluation of models for mixing.

Another important property of the mixing process is the delay with which the signals from different sources reach the different sensors. For EEG signals an instantaneous mixing is generally assumed (volume conduction has zero-delay). Additionally a linear mixing is often assumed:

$$r(t) = M \cdot s(t) + v(t) \tag{5.17}$$

Where $r(t) \in \mathbb{R}^k$ is a multi-variant time series representing the recordings at the k sensors/electrodes, $s(t) \in \mathbb{R}^l$ is the multi-variant times series generated by the l sources which contribute to the signals recorded, $M \in \mathbb{R}^{k \times l}$ is the mixing matrix and $v(t) \in \mathbb{R}^k$ is a multi-variant noise (one per sensor – usually assumed to be an independent Gaussian process).

(Linear) BSS means identifying a matrix W which reconstructs $s(t)$ in absences of any knowledge about M (or $v(t)$). If $l = k$ then W would need to be the inverse of M. For EEG signals, as for most other real world BSS applications $l \gg k$, meaning the problem of of BSS is ill-posed.

Common examples for such source space transformations are LORETA source localisation, Dipole Source Localisation and Independent Component Analysis. LORETA and Dipole Source Localisation are not BSS methods, because they use assumptions about the sources and their location. LORETA and Dipole Source Localisation also try to localise the sources they identify in 3D-space, based on the known 3D positions of the sensors. For a well grounded 3D localisation of sources of EEG signals, one needs a high spatial resolution, i.e. many electrodes. 3D source localisation based on a 16 channel data set as for my experiments would be dubious, a best.

Hence, I work with the Independent Component Analysis which is a classical BSS. This is a linear approach meaning that it can be expressed as a basis transformation. The transformation tries to identify parts of the signal which occur with high temporal correlation in different channels and subsumes these into one channel of the transformed data. Those are called independent component (IC). For instance in EEG recordings one always finds activity generated by muscle activity of eye movements. These are actually rather prominent in the EEG. This activity would be found in multiple (if not all) channels simultaneously, although with different power. The ICA can identify this as one component as it always occurs in different channels with high temporal correlation. Actually ICA is often used to remove ocular artefacts from EEG recordings. Of course the activity of any given neural sources would occur with high temporal correlation as well and should hence be subsumed to one component.

There are a number of different algorithms for estimating a good de-mixing matrix for the ICA. The algorithm commonly used for EEG recordings is the so called Infomax algorithm proposed in Bell and Sejnowski (1995). It uses gradient learning to find a good de-mixing that minimises the mutual common information in the output.

In any case, the ICA can only isolate a number of components equal to the number of channels of the original data. My data has 16 channels and the ICA can, therefore, isolate 16 different components. There are, however, certainly much more than 16 (neural) sources in the EEG (the BSS problem is ill-posed). Hence, the ICA is forced to mix different sources into one component. This is especially problematic when performing artefact removal using ICA, because one can never be certain that a component identified to contain an artefact not also contains some neural activity as well.

There are, however, ways to inspect the ICs and to find certain hints to whether or not a given component contains neural activity or muscle activity. One important index is the spectral power density of the component. Components containing neural activity usually show a clear peak in one band usually associated with neural activity (e.g. α-band, $8 - 13Hz$). Figure 5.6 shows a component properties plot for such a component, as it can be generated using the EEGLAB toolkit. Components containing artefacts usually have a monotone spectral power distribution.

Another index is the spatial distribution of the components contributors. If the component is strongly dominated by frontal electrodes this is usually an index that this component contains ocular artefacts. The plotting the activity over time is another important index to identify ocular components. This type of components typically shows a spike-like behaviour (short bursts of activity with longer periods of inactivity in-between) which is not typically found in neural sources.

The EEGLAB toolbox for MATLAB not only provides an ICA implementation optimized for EEG data analysis, but also some great tools for component inspections. Some add-ons to the EEGLAB toolbox

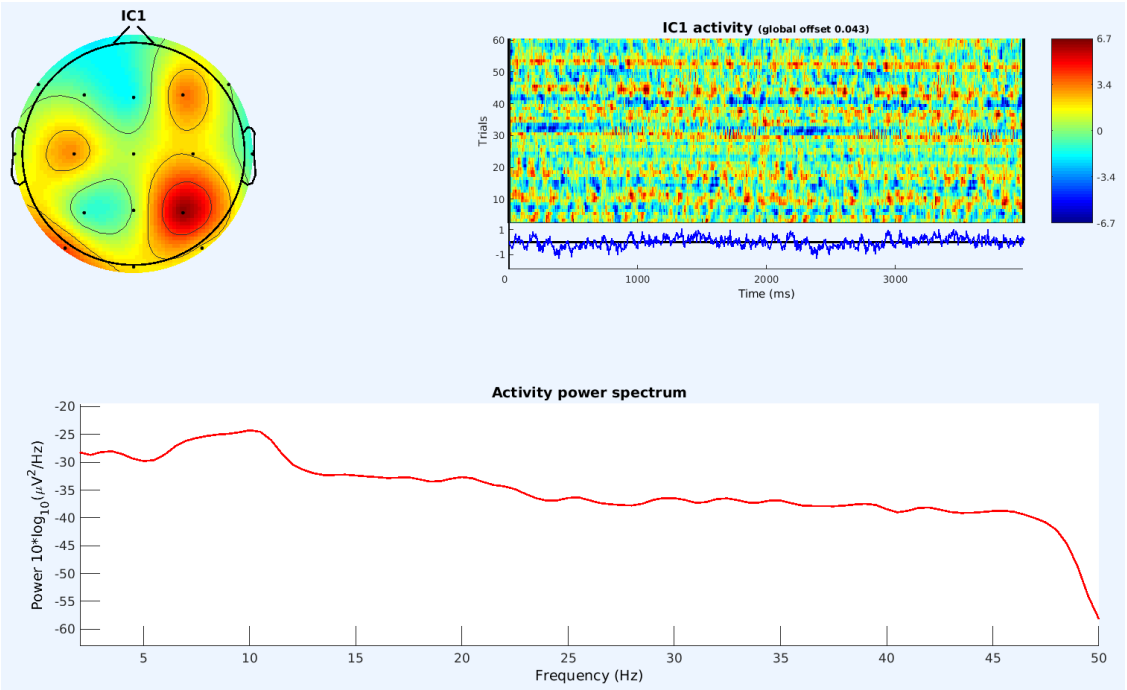


Figure 5.6.: The properties of an Independent Component identified by ICA. The peak in the spectral density in a frequency range relevant for EEG is a clear indication that this is a neural component (rather than a muscle artefact).

also offer suggestions for the categorisation of components in artefacts and neural activity based on different mathematical measures. For this thesis I employed the ADJUST toolbox which uses Spatial Average Difference, Temporal Kurtosis, Maximum Epoch Variance and Generic Discontinuities Spatial Features as parameters to identify artefact ICs (Mognon et al., 2011a).

Reducing the role of ICA to artefact removal would, however, be a mistake. The ability of ICA to isolate contributors of an EEG signal is of great help when studying the interaction of sources within a brain. Connectivity estimation (and other analyses) cannot only be conducted in sensor space (electrodes) but also in source spaces (on ICs). ICA has, however, the downside that the ICs of different participants are hardly comparable, limiting the comparability of results from different pairs of participants and, thereby, the possibilities to formulate results valid across participant pairs. Another property of ICA is that it is a reference-free representation of the data. As discussed in section 3.4 page 28, Haufe et al. (2012) demonstrated the effect of the choice of reference. This, ultimately, convinced me to conduct the analyses of the second experiment in source space (on ICs), despite the mentioned drawbacks.

5.4. Granger Causality

One central point in the data analysis performed for this thesis is to identify the flow of information between sources of brain activity a.) within one brain b.) between the brains of participant. It is renown that the human brain consists of different, loosely coupled areas which interact/exchange information. If we could fully describe the information flow between brain areas, we would have in our hands the key to unlock the secrets of the human brain. However, already the definition of a neural information flow is already non-trivial. A concept commonly employed in computational neuroscience to define information flow originates in economics: The Granger Causality (Granger, 1969).

Given a multichannel time series D_t with c channels and t time step. Let further be P a predictor which tries to predict the values of the time series for the next time step, based on the time series history: $P(D_{\grave{t}}) = \hat{D}_t \quad \grave{t} \in [1, t-1]$. If for two channels $i, j \in [1, c]$ there is a predictor for which the prediction of j is generally better if the predictor knows the history of all channels, rather than all channels except i, we say that i Granger-causes j. In formula: i Granger-causes j if:

$$P(D_{\grave{i}}) \quad \grave{t} \in [1, t-1] \quad \text{Prediction including all channels} \tag{5.18}$$

$$P(D^{\grave{i}}_{\grave{t}}) \quad \grave{t} \in [1, t-1] \quad \text{Prediction excluding } \ i \tag{5.19}$$

$$\mid P(D_{\grave{i}}) - D_t \mid < \mid P(D^{\grave{i}}_{\grave{t}}) - D_t \mid \quad \text{Comparing prediction errors} \tag{5.20}$$

Where $D^{\grave{i}}_{\grave{t}}$ is the time series $D_{\grave{t}}$ without the channel i. This does not need to hold for any t but in general.

The general idea behind this is that the cause precedes the effect. If there is an influence which is exhibited by i on j, it needs to be in the history of i. Thus, a good predictor should improve if it knows the history of i compared to its prediction without knowledge about i.

Here I understand Granger Causality as a general concept which deliberately leaves the choice of the predictor open. There is a variety of publications which understand Granger Causality linked to Multi-Variant Autoregressive Models (see section 5.5 below) as the predictor. I think this is a needless limitation of an otherwise very powerful concept. I admit, however, that leaving the choice of the predictor open makes it hard, if not impossible, to disprove Granger Causality. Proving that there cannot be any predictor which would make better prediction knowing the history of i will in most cases not be possible.

5.5. Multi-Variant Autoregressive Linear Models

A Multi-Variant Autoregressive Model (MVar, discussed in depth in Lütkepohl (2005)) for a c-variant times series consists of a set of $c \times c$-matrices $M_i \quad i \in [1, m]$. These matrices are then used to predict the values of a given time series based on the time series history. To predict the values of the times series at time t one multiplies the matrices with the values of the time series during the last m steps:

$$\hat{D}_t = v + \sum_{i=1}^{m} M_i \cdot D_{t-i} \tag{5.21}$$

$$\sigma_t = \hat{D}_t - D_t \tag{5.22}$$

v is the mean of the different channels and σ is called the residual of the model. For a good model this residual should be an uniformly distributed (i.e. "white") random process. A non-white residual indicates, that the model failed to capture some of the data's structure. Sometimes increasing the model order m can help in such situations. There are various algorithms which can be used to determine the M_i matrices.

The prediction of time series is only one application for MVar models. The matrices M_i, when adequately chosen, capture a lot of information about the structure of the time series and allow to derive various estimates for parameters of the time series. We can reformulate the MVar model as:

$$D_t = v + \sum_{i=1}^{m} M_i \cdot D_{t-i} + \sigma_t \tag{5.23}$$

With D_t being the real, rather than the predicted value of the times series at time t. The linear nature of these models brings certain limitations. I.e. the model assumes the data is stationary, which is generally not the case for EEG data. To overcome this, local stationarity is used: When EEG data is divided into sufficiently small epochs (and detrended) the data is (locally) stationary. Hence, one way to obtain a valid MVar representation of EEG data is to epoch the data into a series of (eventually overlapping) short windows and to compute a new MVar model for every window. The downside of this approach is, that this dramatically cuts down the training data available for each model. Furthermore, the stationary of each window and the whiteness of each models' residual needs to be tested mathematically.

Another approach to overcome this obstacle is to add a forgetting term to the model or (similar) to apply a Kalman filter. Both allow the model to adapt to the changes that cause the in-stationarity. The downside, especially of the Kalman Filter is, that it dramatically increases computation time for a model.

To compute a valid MVar model, a certain amount of data is needed. As a rule of thumb, for EEG data for each weight in the matrices one should not have less than ten training data samples. Given that we have c channels, a sampling rate of sHz and a model order of m (m matrices)

$$\text{we have} \quad w = c^2 \cdot m \qquad \text{weights in the matrices} \tag{5.24}$$

$$\text{and would need} \quad p \geq c^2 \cdot m \cdot 10 \qquad \text{samples of training data} \tag{5.25}$$

$$\text{or a window length of} \quad pt = \frac{p}{s} sec. \geq \frac{c^2 \cdot m \cdot 10}{s} sec. \tag{5.26}$$

For my recordings I had 16 channels and a sampling rate of $256Hz$. Assuming a (quite typical) model order of 3 I would end up at

$$p \geq 16^2 \cdot 3 \cdot 10 = 7,680 \qquad \text{samples} \tag{5.27}$$

$$pt \geq \frac{7,680}{256} sec. = 30 sec. \qquad \text{window length} \tag{5.28}$$

and, even worse, for hyper-analysis I need to treat two such data sets as one, doubling the number of channels and, thus, suffering a quadratic increase in weights. Hence,

$$p \geq 32^2 \cdot 3 \cdot 10 = 30,720 \qquad \text{samples} \tag{5.29}$$

$$pt \geq \frac{30,720}{256} sec. = 120 sec. \qquad \text{window length} \tag{5.30}$$

Both is well beyond any window length for which local data stationarity could reasonably be assumed. Hyper-analysis is particularly problematic in that regard as the need for training data grows quadratically with the number of channels. Luckily there are methods to reduce the need for training data.

One such method was presented by Ding et al. (2000). The basic assumption is that several repetitions of data which share the same properties are available. Let us assume we have n repetitions in an experiment. For each repetition we have a data epoch of length l *sec.* of data. We would still apply windowing to the data, choosing the window sufficiently small that the assumption of local stationarity is reasonable (e.g. 0.35 *sec.*). However, we would assume that the corresponding windows from different epochs are sufficiently similar, such that we can use them to compute a common model to describe them. This divides minimum window length by the number of repetitions available in the data.

$$pt \geq \frac{c^2 \cdot m \cdot 10}{s \cdot n} sec. \qquad \text{window length} \tag{5.31}$$

This reduces the need for training data linearly. In the iCusss experiment of this thesis there are ca. 70 such repetitions available in each data set. This means we would still need a window length of $120 sec./70 \approx 1.714$ *sec.* which is still too large to assume local stationarity.

Another, very basic approach is to reduce the number of channels. The need for training data grow quadratically with the number of channels. And it shrinks the same way. The risk in this is the common drive phenomenon: We might accidentally remove one channel which is the driver of two other channels. In that case connectivity estimates usually (falsely) identify a connection between the two receivers of the information flow, while they are really only influenced by the same sender and do not influence one another.

In this thesis I used MVar-based methods only on basis of independent components (rather than channels). In that case I can leave out any component containing muscle or ocular artefacts. To further reduce the number of channels, one can compute pair-wise partial coherence on the components. For any components which shows little coherence with any other component it can be assumed that this component will not be part of any information flow network. This way one can further reduce the number of channels used for MVar-based analyses, while controlling the risk to accidentally remove parts of the connectivity network.

There is a great EEGLAB plug-in implementing MVar model computation, validation and computation of various connectivity estimates based on these models. This is called Source Information Flow Toolbox (SIFT) (Mullen, 2010; Delorme et al., 2011). All MVar based connectivity estimation done during the PhD project was performed using SIFT.

5.5.1. Auxiliary Quantities

I already stated that the MVar matrices, if properly chosen, capture a lot of information about the data they model. While this is intuitively plausible, it is not quite obvious whether and how this information could be put to use.

From the MVar model I can derive a number of auxiliary quantities of which the most important one is the so called transfer matrix. Let me start with the model from equation 5.23.

$$D_t = v + \sum_{i=1}^{m} M_i \cdot D_{t-i} + \sigma_t \tag{5.32}$$

The data is assumed to be zero-mean and we can, thus, drop v. Furthermore, a white residual σ_t is assumed. We can express D_t as $I \cdot D_t$ (I being the identity matrix). Hence we can transfer the above equation to:

$$\sigma_t = -\sum_{k=0}^{m} M_i \cdot D_t \tag{5.33}$$

where $M_0 = -I$. If we now obtain a spectral representation of both sides using z-transformation we get:

$$U(\omega) = A(\omega) \cdot D(\omega) \tag{5.34}$$

$$\text{where} \qquad A(\omega) = \sum_{k=0}^{m} -M_k \cdot e^{-i2\pi\omega k} \tag{5.35}$$

Rearranging the previous equation for $D(\omega)$ we get:

$$D(\omega) = A(\omega)^{-1} U(\omega) \tag{5.36}$$

Mullen (2010) explains the three matrices as:

$\boldsymbol{U(\omega)}$ "... is a matrix of random sinusoidal shocks ..."(Mullen, 2010).

$\boldsymbol{D(\omega)}$ "... is the [...] spectral matrix of the multi-variant process, ... "(Mullen, 2010).

$\boldsymbol{A(\omega)^{-1} = H(\omega)}$ "... is the transfer matrix of the system."(Mullen, 2010).

As I cannot put it any better, I cite him here.

Mullen et al further argue that it is actually the transfer matrix $H(\omega)$ that transfers the random shocks of $U(\omega)$ into a structured spectrum and that, therefore, we can assume that it contains much information about the internal structure of the modelled data.

Additionally, we can define the spectral density matrix $S(\omega)$:

$$S(\omega) = D(\omega) \cdot \overline{D(\omega)} = H(\omega)\Sigma\overline{H(\omega)} \tag{5.37}$$

From these three quantities, $S(\omega)$, $H(\omega)$ and Σ, we can derive a whole lot of information about the system modelled. In this thesis I will, however, only discuss the two most prominent connectivity measures derived from these quantities: PDC in section 5.6.3 on page 52 and DTF in section 5.6.4 on page 52. Before presenting these complex estimates, however, I will go over some basics of connectivity estimation, first.

5.6. Estimates of Signal Connectivity

Nowadays the human brain is commonly regarded as an assembly of coupled, interacting neural groups. Consequentially, the need for mathematical methods to analyse and quantify such interaction (neural connectivity) grew. This development and discussion is an ongoing process and is by far not yet concluded

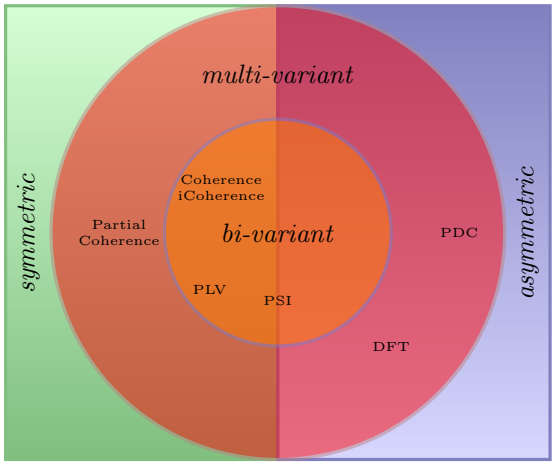


Figure 5.7.: A Taxonomy of different estimates of signal interaction. Estimates are commonly categorised as symmetric/asymmetric and as bi-variant/multi-variant. PSI is a somewhat special case, as it is neither symmetric nor truly asymmetric, but anti-symmetric.

(Haufe et al., 2012). Or as Plomp et al. (2014) put it "... a 'gold standard' has yet to emerge.". Here I will try a taxonomy of different connectivity estimates that have found broader acceptance.

Connectivity estimates are usually classified along two different lines: Whether they are symmetric or asymmetric and whether they are bi-variant or multi-variant

For two time series X and Y, a symmetric estimate $s_{XY}(\omega)$ has no sense of directionality. Whether one considers the time series X, Y or Y, X is all the same for a symmetric estimate: $s_{XY}(\omega) = s_{YX}(\omega)$. Probably the most prominent symmetric estimate in this context is the Coherence of two time series.

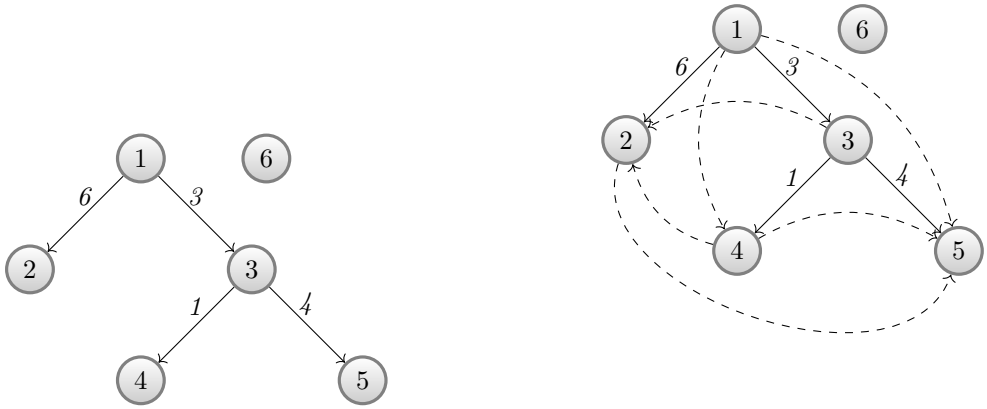
An asymmetric estimate $a_{XY}(\omega)$ takes into account the relation of the two time series as sender and receiver (or cause and effect, to put it in Granger terminology). Generally $a_{XY}(\omega) \neq a_{YX}(\omega)$. This gives the estimate the power to express the direction of an interaction. It allows an estimate to express whether time series X drives time series Y or vice versa. Even more, a truly asymmetric estimate could even express that X drives Y and simultaneously that Y drives X (a feedback loop).

One estimate presented in this thesis does not really fit into that taxonomy: For the Phase Slope Index (PSI) it always holds that $PSI_{XY}(\omega) = -PSI_{YX}(\omega)$. It is therefore not symmetric, but it is not truly asymmetric neither. Particularly it could not identify a feedback loop as such, but only return the "net-influence". In Haufe et al. (2012) PSI is classified as anti-symmetric, which is, in my opinion, a pretty accurate declaration.

A second terminology, which is closely tied to the classification of connectivity estimates as symmetric and asymmetric, is that of functional and effective connectivity. Functional connectivity is the "temporal correlations between spatially remote neurophysiology events" (Friston et al., 1993b) while effective connectivity is "the influence that one neural system exerts over another either directly or indirectly"(Friston et al., 1993a). So it can be said that a symmetric estimate can find functional connectivity but it cannot distinguish functional and effective connectivity. An asymmetric estimate can distinguish these two. Lee et al. (2003) gives a very comprehensive discussion about the these terms, ending by stating that "the distinction between functional and effective connectivity [...] emphasizes the shift from a description of what the brain does to a theory of how it does it."(Lee et al., 2003).

The second categorisation of connectivity estimates is whether an estimate takes into account two channels (bi-variant) or an arbitrary number of channels (multi-variant). Theoretically multi-variant estimates should return a more sophisticated representation of the connectivity structure in the data. I would like to pick up here an allegory I once heard in a lecture given by Tim Mullen to clarify the theoretical argument in favour of multi-variant estimates: A waiter sees a man chasing down the street. Shortly thereafter he sees a second man chasing the first man. What the waiter does not see is the bus a little bit further down the street which both men are actually trying to catch.

A bi-variant estimate takes two channels and tries to identify a relation between them (the two running man). For instance channels two and three from figure 5.8a. It cannot see a third channel, channel one from the figure or the bus in the allegory, as it is by definition bi-variant. There is no connection between channels two and three or the two running man. The influence is actually exhibited by channel one (or the



(a) The real dependencies in this example: One drives two and three, three drives four and five. Six is isolated. Each edge gives the delay with which the influence is propagated. This is, of course, an artificial example to make a theoretical argument.

(b) The dependencies a bi-variant estimate could identify: Additional to the actual dependencies (solid edges) it would identify connections between any two channels expect six which is isolated.

Figure 5.8.: An example to illustrate which dependencies would be found by a bi-variant (asymmetric) estimate. Adapted from a lecture by T. Mullen. Lecture slide including the example at `http://sccn.ucsd.edu/mediawiki/images/a/ab/SIFT_Lecture.pdf`

bus). The influences it exhibits on the two, both have a delay and as long as this delay is not equal (both man do not run beside one another) the waiter and the bi-variant estimate will identify a causality which is not actually there.

Figure 5.8b shows which impacts this can have: A relatively simple structure in the ground truth is inflated to a complex setting by the inability of a bi-variant estimate to distinguish certain interdependencies. It would neither realise that the influence of channel one on four and five is indirect nor would it realise that two, four and five have common ancestors.

Whether or not the theoretical ability of multi-variant estimates to identify such connectivity structures come into effect when those applied to neural data has, however, been heavily disputed Haufe et al. (2012).

For the rest of this section I am going to discuss a variety of connectivity estimates. This starts with Coherence and Imaginary Coherence in section 5.6.1 below. Then I will discuss the concept of phase locking and an estimate for this (PLV) in section 5.6.2 on page 50. Then I will come around to the first multi-variant estimate which is the Partial Coherence and its asymmetric advancement Partial Directed Coherence (PDC) in section 5.6.3 on page 52. A second multi-variant estimate is the Direct Transfer Function (DTF) which is introduced in section 5.6.4 on page 52. Finally I will come to an estimate which has been developed only recently and which has become very popular, despite its bi-variant and not quite asymmetric nature: The Phase Slope Index (PSI) developed by Nolte et al. is introduced as the last connectivity estimate in this thesis in section 5.6.5 on page 53.

5.6.1. Cross Correlation, Coherence and Imaginary Coherence

A classical estimate to test two time series for commonalities is Coherence. I will describe it here, taking a short detour over the Cross Correlation. Coherence has some sever shortcomings when it comes to "real world" data and hyper-analysis, which is why several adaptations of it, namely Imaginary Coherence, Partial Coherence and Directed Partial Coherence have been proposed in order to overcome these. I will discuss these later in this section.

Assuming we have the signals/time series X and Y each with N samples. The two time series need to be zero-mean and unit-variance. If this is not the case it can easily be achieved as a preprocessing step. The most basic estimate for commonalities between two signals is the Cross Correlation. It is defined as a function of a time lag $\tau \in [0, N-1]$. It is basically the mean of the products of the samples with τ distance in time.

$$CC_{XY}(\tau) = \frac{1}{N-\tau} \sum_{n=1}^{N-\tau} X_{n+\tau} \cdot Y_n \tag{5.38}$$

If the signals correlate, i.e. if both signals take high and low values with a delay of τ, this average takes a value close to 1. If the signals are linearly independent, the values will tend to average out and the Cross Correlation is close to 0. If the signals take opposing values the Cross Correlation will be around -1. The Cross Correlation is a bi-variant and symmetric estimate.

Another important concept (although not an estimate for signal alignment in its own right) is the Cross Spectrum. The signals are first transformed into a time-frequency domain before they are transferred into a common spectral representation:

$$C_{XY}(\omega) = E(F_X(\omega) \cdot \overline{F_Y(\omega)}) \tag{5.39}$$

Where $F_X(\omega)$ and $F_Y(\omega)$ are the time-frequency transforms of X_n and Y_n (usually computed with FFT, see section 5.1.3 on page 37) for the frequencies ω.

The cross spectrum offers a way to compute the power spectrum of a signal:

$$P_X(\omega) = \mid C_{XX}(\omega) \mid \qquad P_Y(\omega) = \mid C_{YY}(\omega) \mid \tag{5.40}$$

The Complex Coherence finally is the amplitude of the cross spectrum normalise by the two signals power spectra:

$$CCoh_{XY}(\omega) = \frac{C_{XY}(\omega)}{P_X(\omega) \cdot P_Y(\omega)} \tag{5.41}$$

while the classical coherence is computed as:

$$Coh_{XY}(\omega) = \frac{\mid C_{XY}(\omega) \mid^2}{P_X(\omega) \cdot P_Y(\omega)} \tag{5.42}$$

One of the main drawback of coherence when it comes to EEG signal analysis is that it is sensitive to phase and amplitude of the signals. I already stated that communities between two signals can (in case of EEG) have two different sources: i.) It is a product of volume conduction. ii.) It is a product of actual information transmission between two different brain regions. Commonalities caused by volume conduction have a zero phase difference between the two signals, whereas signal commonalities caused by actual neural connectivity, do generally have a non-zero phase difference. That is why as a first step to improve the expressiveness of coherence in regard to neurological information flow, imaginary coherence was introduced (Nolte et al., 2004).

$$iCoh_{XY}(\omega) = \Im(CCoh_{XY}(\omega)) \tag{5.43}$$

5.6.2. Phase Locking

Phase Locking describes a phenomenon which can be observed in experimental data from experiments which had several repetitions of the same task/stimulation/event. When for a certain frequency and time, time-locked to the task/stimulus/event, the distance in phase between two times series is similar over repetitions, this phenomena is referred to as phase locking.

To check for phase locking we need to perform three steps:

1. Compute the phase for the two signals for all desired points in time and frequency and for all repetitions.
2. Compute the difference in phase between the two signals for each point in time and frequency and for all repetitions.

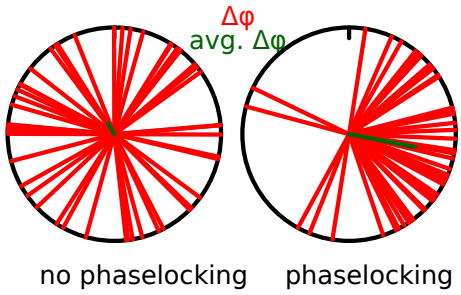


Figure 5.9.: The average complex vector (green) when averaging a number of phases (red) expressed as vectors on the complex number plane's unit circle. Left when phases are uniformly distributed and right when phases share a bias.

3. Compute the average phase difference over all repetitions for each point in time and frequency.

The computation of the average phase (difference) is slightly more involved than classical mean computation, due to the circular value range of the phase. One pretty elegant solution is to transform the phase into a complex number on the unit-circle where the phase is encoded in the vectors angle. This can easily be done using the e function:

$$PLV_{XY} = \frac{1}{N} \cdot | \sum_{k=1}^{N} e^{i \cdot \Delta\phi_{XY}^{k}} | \tag{5.44}$$

Where $\Delta\phi_{XY}^{k}$ is the difference in phase between the two signals in repetition k given as a real numbers. N is the number of repetitions. The resulting complex number's angle is the average phase, but even more interesting the absolute value of the result represent how biased the phases deltas are. This is the Phase Locking Value (PLV).

When the $\Delta\phi_{XY}^{k}$ are uniformly distributed (white) the vectors point in every direction and tend to eliminate one another in averaging. The resulting complex number has a low absolute (euclidean) value. When there is a bias in the phase delta (and therefore in the vectors' orientations) the averaging will come up with a complex number with a larger absolute value.

This is depicted in figure 5.9. All phases were expressed as complex numbers with an absolute value of 1. The average of such a group of vectors is a vector with a length in $[0, 1]$. The length would be 1 if and only if all the original vectors would point into the same direction (i.e. all phases were exactly equal). For PLV this would mean perfect phase locking. The length would be 0 if all vectors would eliminate one another. For an uniformly distributed, finite set of vectors, the average vector will have a small length which is asymptotically 0 for a growing number of (uniformly distributed) vectors.

Lachaux et al. (1999) first presented this estimate. It is a symmetric and bi-variant estimate and completely dependent on the signals' phase. Le Van Quyen et al. (2001) compared PLV using Hilbert and Wavelet Transformation for estimation of phase of the two signals. Their result was that both deliver comparable results (compare section 3.4 on page 28). I picked Wavelet Transformation with a complex Morlet Wavelet.

I already mentioned that the PLV will not generally be 0 even if there is no phase locking between the time series what so ever. Furthermore, PLV will tend to be smaller for larger numbers repetitions (if no phase locking is present in the time series). The question which PLV results can be considered significant phase locking is therefore non-trivial.

Lachaux et al. (1999) suggest a statistical method for assessing the PLVs obtained. The idea is that if we could obtain data which has the same statistical properties as the original data in all regards, but definitely cannot have any phase locking, we could compute PLVs on that data and compare them to the original PLV values.

This can actually easily be done: We shuffle the repetitions for one of the two time series and by that pair epochs from different repetitions. There can be no systematic phase locking in this surrogate data set while it still shares virtually all other properties with the original data. We now compute surrogate PLVs on that surrogate data and compare these with the original PLVs.

This is repeated many times (Lacheaux et al. propose 200 shuffles) and it is counted how often the original PLV was larger than the surrogate PLV. The rate of surrogate PLVs which were larger than the original

PLV is called the Phase Locking Statistics (PLS) and can be treated similarly as a p-value from statistical t-testing. Particularly if the surrogate PLV was larger in no more than 5% (or whatever significance threshold one wants to pick) this can be considered significant phase locking. However, for larger numbers of tests some type of multiple-comparison correction needs to be applied.

5.6.3. Partial Coherence and Partial Directed Coherence

As already mentioned the Coherence is a bi-variant estimate. It, therefore, cannot reconstruct indirect dependencies and common drivers.[2] Partial Coherence is an advancement of classical Coherence aiming to eliminate that shortcoming. It is a multi-variant estimate and its basic idea is that for each channel pair i and j only that coherence is taken into account which cannot be explained as a linear combination any of the other channels. As a result, both, indirect dependencies and common drivers can theoretically be reconstructed by Partial Coherence.

It can be computed most easily from a MVar model as:

$$PCoh_{XY}(\omega) = \frac{S_{XY}(\omega)^{-1}}{\sqrt{S_{XX}(\omega)^{-1} S_{YY}(\omega)^{-1}}} \tag{5.45}$$

$S(\omega)$ is the spectral density matrix as defined in section 5.5.1 on page 47.

The Partial Coherence is a symmetric estimate and as such cannot analyse the directionality of the influence (can infer functional but not effective connectivity). For this reason the Partial Directed Coherence (PDC) has been developed:

$$PDC_{XY}(\omega) = \left| \frac{A_{XY}(\omega)}{\sqrt{\sum_{Z \in C} \mid A_{ZY}(\omega) \mid^2}} \right|^2 \tag{5.46}$$

Where C is the set of all time series involved in the multi-variant analysis. The $PDC_{XY}(\omega)$ is from the interval $[0, 1]$ and it is normalised by outflow:

$$\sum_{Z \in C} PDC_{XZ}(\omega) = 1 \tag{5.47}$$

In other words, the sum over PDC-values between channel X and any other channel is always 1.

5.6.4. Direct Transfer Function

The Direct Transfer Function is a second multi-variant, asymmetric estimator for information flow. It was first introduced by Kaminski and Blinowska (1991). The Direct Transfer Function (DTF) can be computed from a MVar model:

$$DTF_{XY}(\omega) = \sqrt{\frac{\mid H_{XY}(\omega) \mid^2}{\sum_{Z \in C} \mid H_{XZ}(\omega) \mid^2}} \tag{5.48}$$

$$\tag{5.49}$$

$H(\omega)$ is the transfer matrix (see section 5.5.1 on page 47) of the MVar system. The DTF is normalised such that

$$\sum_{Z \in C} DTF_{ZY}(\omega) = 1 \tag{5.50}$$

In other words the DTF is normalised such that the sum of all contributions to the output is 1. Which is just the complement of the normalisation done for PDC.

In the literature the discussion about advantages and disadvantages of PDC vs. DTF has been particularly keen. For my point of view, there is currently no clear argument favouring any of the two over the other.

[2] Remember the example with the two man chasing a bus?!?

5.6.5. Phase Slope Index

The general idea of the Phase Slope Index, first proposed by Nolte et al. (2008), is that an information flow between time series X and Y will results in a slope in the phase of the cross spectrum. The sign of that phase slope reflects the directionality of the information flow.

The Phase Slope Index is computed as

$$\widetilde{\Psi}_{XY}(\omega) = \Im(\sum_{i=1}^{|\omega|-1} \overline{C_{XY}(\omega_i)} C_{XY}(\omega_{i+1})) \tag{5.51}$$

where $C_{XY}(\omega)$ is the complex coherency as in equation 5.41 (page 50). Finally the PSI is usually normalised by its variance.

$$PSI_{XY}(\omega) = \frac{\widetilde{\Psi}_{XY}(\omega)}{Var(\widetilde{\Psi}_{XY}(\omega))} \tag{5.52}$$

The variance of $\widetilde{\Psi}$ is unknown. Hence, Nolte et al. (2008) suggest to estimate it using a Jackknife approach. For $PSI_{XY}(\omega)$ Nolte et al. (2008) al state that any value larger than 2 can be considered significant.

It is important to note that $PSI_{XY}(\omega) = -PSI_{YX}(\omega)$. Hence, $PSI_{XY}(\omega) > 0$ indicate an information flow from X to Y and $PSI_{XY}(\omega) < 0$ indicate an information flow from Y to X. Therefore, PSI is not symmetric but not really asymmetric (particularly $PSI_{XY}(\omega)$ and $PSI_{YX}(\omega)$ are not independent), neither. It is fit to estimate the directionality of an information flow (effective connectivity), however, for recurrent inter-dependencies it would present only a "net information flow". Theoretically speaking a loop between X and Y in which both channels influence one another to the same degree would have a $PSI_{XY}(\omega) \approx PSI_{YX}(\omega) \approx 0$. This issue is address in Nolte et al. (2008) by stating: "in complex systems [...] asymmetries in detection power may as well arise due to other factors, specifically independent background activity...". I would interpret this sentence in a way that for neural and, in particular, EEG recordings it is too ambitious to hope to identify such closed loops reliably.

Finally, PSI is computed on a segments of arbitrary length resulting in a single value, i.e. it collapses the time dimension. Having only one result value and thereby loosing the time dimension can be a drawback, but can also be an advantage for certain types of analyses. Theoretically one could obtain PSI as a function of time using windowing techniques. However, in an email-conversation with Guido Nolte, who proposed this estimate, he discouraged this because "the idea for PSI was to be applied on spontaneous EEG/MEG [...]".

I assume that the problem if indirect connectivity persists for bi-variant connectivity estimates, regardless of the claim of Haufe et al. (2012) on the contrary, as discussed in section 3.4 on page 28. I still decided to (also) use PSI for analysis, due to the convincing results in Haufe et al's study in its favour, particularly "Since PSI inherently implements the ideas of anitsymmetrization and time inversion testing,..." (Haufe et al., 2012). Both concepts Haufe et al. (2012) showed to be extremely effective in improving the reconstruction of the information flow on their simulation data.

5.7. Statistics

Many of the previously introduces estimates are relative, i.e. their absolute values bear little intrinsic meaning. Only in comparison with the same estimate computed on some suitable comparison data they exhibit any meaning. This comparison data can be data from another condition within the experiment. Another method which is often used, is to compute the values on surrogate data which is (with respect to its statistical properties) similar to the original data, but has by its design no connectivity.

In the following I will discuss statistical methods which are often used in the context of connectivity analysis. In section 5.6.2 on page 50 I already explained a method for statistical evaluation of PLVs (PLS) which can be easily adapted for other estimates. In section 5.7.1 on page 54 I will discuss phase randomisation as another method to generate surrogate data.

In many cases connectivity analysis requires a large number of statistical tests. Using a canonical threshold for the p-value of a statistical test of 0.05 will result many type-I errors. There are different methods to compensate for this. In section 5.7.2 on page 54 I will present a method aiming to control the rate of type-I errors at a certain level.

5.7.1. Phase Randomisation

One popular method for generating surrogate data is to randomize the data's phase. For that, first a (complex) time-frequency estimation needs to be computed (e.g. using FFT, see section 5.1.3 on page 37), then the imaginary part is uniformly randomised and finally the reverse transformation is applied.

This is apparently only liable if the connectivity estimation operates exclusively on the signals phase.

5.7.2. False Discovery Rate

When working with neural data, one often is in the situation to have to perform a huge number of statistical tests in a row (e.g. for each voxel of a fMRI scan or for each combination of different channels when estimating connectivity on EEG data). Let us assume we conducted N tests and obtain an estimate for the likelihood that the 0-hypothesis is true (p-value). Let us further assume that on the given data set the 0-hypothesis is actually true for all of the tests. When one applies the canonical 0.05 threshold there would be an expected (a-priori) number of $N \cdot 0.05$ false rejections of the 0-hypothesis. E.g. for a hyper-scanning data set with 22 channels (a typical size for my later analyses) $22 \cdot 22 = 484$ connections would have to be tested. I would still get an expected number of $484 \cdot 0.05 = 24.2$ connections for which the 0-hypothesis would be (falsely) rejected.

In Bennett et al. (2009) this problem is illustrated even more saliently: Bennett et al put a dead salmon into a fMRI scanner and presented the salmon with a series of pictures. After statistical evaluation with a rather rigours threshold of 0.001 they could still show the salmon's neural reaction to the stimuli.[3] Of course the dead salmon reacted to the stimuli by no means (neither neural nor otherwise). The sheer number of statistical tests involved (8064 voxels were tested) lead to 16 significant voxels (expected would have been $\sim$ 8).

This shows that we need to correct for this effect in one way or the other. A standard method in the field is the family-wise error rate (FWER). This controls the chance that there are any type-I errors (or false discovery), i.e. a falsely rejection of the 0-hypothesis. FWER as it was suggested by Friston et al. (1994) was one of two algorithms Bennett et al suggested to compensate for the effect of multiple statistical testing.

As mentioned, the expected number of type-I errors grows with the number of statistical tests. Thus, a method guaranteeing a certain probability that no type-I errors occurred, needs to be stricter for a growing number of tests. This yields more and more type-II errors (0-hypothesis being falsely accepted). For larger numbers of tests it gets pretty hard to obtain any significant results using FWER. In Bennett et al. (2009), FWER always accepted the 0-hypothesis (recognised the dead salmon to be dead). This is, however, hardly surprising, considering what was said about how conservative FWER is on large sets of tests.

For these reasons Benjamini and Hochberg (1995) suggested a different approach: Instead of controlling the chance to have any type-I error, they suggested a method to control for the faction of type-I errors (false discoveries) among all rejections of the 0-hypothesis (discoveries). To achieve this, they define an adaptive threshold on p-values, which is computed based on all p-values. The method guarantees that the expected value for the faction of type-I errors among all rejections of the 0-hypothesis is lower or equal a predefined α. The proportion of type-I errors among all rejections of the 0-hypothesis Q is an unobservable random variable. Another unobservable random variable V the number of type-I errors and S the number of correct rejections of the 0-hypothesis. We can express Q as:

$$Q = \frac{V}{V+S} \quad \text{and} \tag{5.53}$$

$$FDR = E(Q) = E\left(\frac{V}{V+S}\right) \tag{5.54}$$

Up to now we defined what we consider to be the FDR. But to retrieve a threshold for significance from it, we need to actually control the FDR. This is achieved by ordering the p-values $p_1 \leq p_2 \leq ... \leq p_m$. Then we define a k as the largest i for which:

$$p_i \leq \frac{i}{m} \cdot \alpha \tag{5.55}$$

[3] This paper is, obviously, a sarcastic comment and the journal does not actually exist. However, the paper can be found and has been widely discussed in the community, e.g. by Lyon (2017).

Hence, controlling the FDR means applying a threshold:

$$T_{FDR} = p_k \tag{5.56}$$

and reject the 0-hypothesis for all tests belonging to the p-values $p_i \leq T_{FDR}$. For the proof that this indeed controls FDR at α consult Benjamini and Hochberg (1995)!

Bennett et al also applied FDR to their salmon data and FDR also accepted the 0-hypothesis for all of the tests. This was to be expected due to a very important property of FDR: For a data set for which all 0-hypotheses are true, FDR is equivalent to FWER. In different words, for the dead salmon, FDR is no more likely to commit type-I errors than FWER is. For this thesis I use $\alpha = 0.2$. Hence, an expected rate of type-I errors of $\leq 20\%$ can be tolerated.

Benjamini and Hochberg state that for a set $S = T_1 \cup T_2$ of p-values $T_{FDR}(S) \sim T_{FDR}(T_1) \sim T_{FDR}(T_2)$, i.e. splitting the p-values into two sets and applying FDR separately for both sets will not systematically alter the result. This, of course, only holds true as long as the p-values in T_1 and T_2 follows the same distribution. If there are parts of S for which the distribution of p-values can be expected to be different than for others it might actually be advisable to apply FDR to those sets separately. If the assumption of distinct p-value distributions is false, the result will not be systematically altered. If the assumption is, however, true the results will be far more valid when FDR is applied separately.

I will give an example to clarify why the application of the FDR controlling method on p-values from different distributions is dangerous: In chapter 8 I need to test connections between neural components found in the EEG data of two participants. These connections can be divided into three different groups, for which it can be assumed that the distribution of p-values is different:

Autocorrelation: Connections of a channel with itself. These are generally pretty strong. The future of a neural component is basically always highly influenced by its own past.

Within-Participant: A connection of a component with a different component of the same participant. These connections are generally weaker than auto-correlative connections, but still stronger than hyper-connections.

Hyper-connection: These connections are generally much weaker than within-participant connections, simply because they form between two different brains.

FDR always first discards those discoveries with a high p-value. Applying FDR to all of the p-values of the three described types of connection would always recognise the auto-correlative connection (rightfully) as significant, would eventually recognise the within-participant connections as significant and would hardly ever recognise the hyper-connections as significant. Which in no way means that hyper-connections cannot form or could not be detected. For the connectivity analyses of the second experiment of this thesis, I always applied FDR separately to these three groups of connections.[4]

5.8. The UBiCI Software Framework

Disclaimer: Everything in this section refers to UBiCI version 0.0.1, which has been used for all studies which are part of this thesis. At the time of writing of the thesis a new version was under development, which introduced quite substantial changes. Other changes might be introduced in the later development. While some general concepts will probably persist, many details described here might not be true for current versions of the UBiCI.

During my Master thesis at CITEC, which I wrote as a joint thesis together wit Hannes Riechmann, we started the development of an online BMI software framework. We continued developing this framework over the course of our respective PhD studies and today the UBiCI (short for University of Bielefeld Brain Computer Interface) software framework is an elaborate tool for recording, storing and online classification of EEG data.[5] Here I will introduce some general concepts and ideas of the UBiCI.

[4] More precisely, I neglected the auto-correlative connections entirely, as they can be expected to be always significant.
[5] Recent developments even allow for the online classification of other types of data (e.g. eye tracking data).

5.8.1. Components and Connections

The UBiCI framework encapsulates each working step within a software unit called a "component". Data can enter and leave a component and the component can manipulate the data. I will discuss different data types used in UBiCI in section 5.8.3 on page 57. Each component is, by design, agnostic of source and receiver of incoming and outgoing data. Notwithstanding some components have restrictions about their incoming data. One can technically connect a component encapsulating the driver of an EEG device to a component computing the matrices for a FDA. As the computation of FDA matrices requires labelled data – which the driver cannot offer – starting the training will result in an error. A malfunctioning component should, by the UBiCI design guidelines, not cause the application to crash, but leave verbose error descriptions in the log.[6] Possible functions of a component are:

- Fetching data from the driver of an EEG device and wrap the data into an UBiCI EEGData object.
- Store data to a file on persistent memory.
- Read data from a file on persistent memory.
- Performing mathematical operations (CSP, PCA, FDA, ...) on data.
- Send data over a network stream.
- Receive data over a network stream.
- Collect different classification results to render a final decision.
- Present stimuli to a participant and record their timing.
- Translate the results of the classification into commands for a robot.
- Control the course of an experiment.
- Many, many others.

Components are interconnected using signal-slot connections as they are defined in the Qt-Framework. In fact all of the UBiCI framework is heavily dependent on Qt. When a component wants to send data (usually because it has finished its task on that data) it emits the corresponding signal. Different signals are defined for the main data types. They all start with the word `result` followed by the name of the data type (e.g. `resultEEGData` or `resultEpochInfo`). A component should only define such signals in its header that it will emit eventually.

When a component can process data of a certain type it should define as slot for the given data type. These slots start with the word `process` followed by the name of the data type (e.g. `processEEGData` or `processEpochInfo`).

There can be components which are sinks or sources of data of a given type, meaning not every component providing a `process...` slot for a given data type must necessarily also provide the corresponding `result...` signal and vice versa. A component reading EEG data from hard drive has a `resultEEGData` signal, but it lacks a `processEEGData` slot.

As already hinted, the components themselves do not establish any connections between one another. A set of components is encapsulated by a "module" and it is the task of the module to instantiate the components it needs and to establish connections between them. I will cover modules in some more depth in section 5.8.4 on page 57.

5.8.2. Deployments and Configuration

The file type which describes a module, it's components and their connections is called a deployment (extension `.dpl`). A deployment has a task it is designed to fulfil. E.g. classifying EEG data for P300 potentials. When an user conducts a study using P300 classification she/he will use the same deployment for each and every of the participants of the study.

In contrast to that is the configuration file (extension `.conf`). The vast majority of the components in the UBiCI has at least some parameters which influence the component behaviour (thresholds, files containing matrices, modes of operation, etc.). The entirety of these parameters forms the configuration. In contrast to the deployment which is the same for all participants in a study, at least some parameters will always be adapted per participant (e.g. the HD file to which the data is stored).

[6] The current development status of many components does not yet measure up with this guideline, though.

Deployment and Configuration file can both include other deployment and configuration files, respectively. Additionally, configuration files can use `#define` statements to perform string replacement as a preprocessor step. Together these two file types contain all information the UBiCI needs to conduct an experiment.

5.8.3. Data Types

The UBiCI knows a number of data types. Any connection between two components in a given deployment has a "connection type" which identifies the data type which can be transferred over that connection. For instance the CSP-component defines a signal `resultEEGData` and the channel variance component defines a `processEEGData` slot. These two can therefore be connected within a deployment with a connection of the `EEGData` type.

I will described the three most important data types in UBiCI and their meaning.

`EEGData:` This data types contains data arranged in a matrix like layout. It knows a number of channels and a number of samples this EEGData object contains. It further knows some meta information of the data, in particular at which time the data was recorded (millisecond precision) and with which sampling rate the data was recorded. The EEGData class implements a set of common operations which can be performed on EEGData objects such as joining, slicing or (partial) deletion. These operations are optimised for speed on the expense of a higher than necessary memory consumption.

`EpochInfo:` An EpochInfo is designed to collect all information about a data epoch which either already exists or which shall be generated. In particular an EpochInfo holds a time stamp of a certain event that has transpired. A typical example would be a flash on a computer screen that has been triggered by the software as part of a stimulus presentation to a participant. Usually one wants to identify an epoch in the EEG data which coincides with the flash and analyse/classify the brain's response to the stimulus. Several components within UBiCI use the information provided by an EpochInfo to perform this task. The EpochInfo can also hold some meta information such as, which of several items on the screen flashed. And it holds two labels: The training label which can be used to label the data "from outside" when some kind of ground truth is available (e.g. during the recording of P300 training data). And the classification label to which a classifiers (e.g. the FDA-component) stores the result of the classification for use by later components, in particular for use by a device control.

`DecisionRequest:` In basically all control scenario the overall system will time and again come into a state at which new input from the participant/a new decision by the participant is required. E.g. a BMI controlling a telepresence robot recognises corridors crossing and needs to know which way to go. In such a situation the component controlling the robot will generate a decision request which in turn e.g. triggers a stimulus presentation and the generation and classification of one or multiple EEG data epochs. In the end the classification will come up with a result which will be stored in the decision request. This result might actually be the composite results of several classification outcomes. The decision request will then be returned to the component controlling the robot which translates it into robotic actions. For many offline scenarios, such as the analysis of training data, decision request are theoretically not required, but might still be needed to properly operate components which have primarily been developed for online use.

Each of these data types can (and has) been sub-classed, allowing it to store additional data needed for a specific scenario or BAP. For instance for the classification of P300s the EpochInfo not only needs to hold the information when a flash occurred on the computer screen but also which items on the screen flashed. Therefore there is a FlashEpochInfo which is a direct subclass of the EpochInfo class and which hosts this (additional) information.

Should the function of a component depend on that it is fed with instances of a certain subclass of a data type it needs to check that on every incoming data item itself. The deployment/the module has no way of enforcing such constraint. If a component receives an instance of a data type it cannot handle, it should pass the data on unmodified (if the corresponding signal is defined) and maybe place a warning in the logs.

5.8.4. Modules

When a deployment is loaded from a file, this file is passed to a class called `Module`. The module takes care of setting up all the components and their connections. It makes use of the `UBiCIFactory` which is a singleton

class and which is basically the UBiCI implementation of the factory-software design pattern found in many software framework, i.e. only the UBiCIFactory is intended in generate instances of components.

A deployment can contain a reference to another deployment. If this happens, the deployments are generated recursively, bottom-up, each in an own instance of the Module-class. From the view of the parent Module the sub-module behaves like a component which can receive and send any of the UBiCI data types. This allows re-usability of deployments.

5.8.5. Temporary and Step Connections

Especially during offline analysis of data, some steps need to be performed in a sequential order. E.g. first all data has to be loaded from HD before the PCA matrix can be computed. And only thereafter the FDA matrix can be computed.

For such situations the Module class knows two interwoven mechanisms: First, there is a special connection type `step` which can be received by components implementing a slot called `stepIn` and which is usually connected with its module. Such components also need to implement a signal `stepOut` which is then connected to the Module via a `step_return` connection. The idea is that the top-level Module (and only that) emits the step signal after setting up all components and connections. A component that receives the signal (and it should be only one component, really) then starts some task. Upon completion of the task the component emits the `stepReturn` signal notifying the module about its task's termination.

The second part of the mechanism are temporary connections. A mundane connection between two UBiCI components persists virtually for the entire program execution. From its creation during the module setup process to its destruction when its module is destroyed.

A temporary connection has a list of integers. Initially a module does not create any temporary connections. Right before emitting the `step` signal for the first time, it will create all those temporary connections which have a 0 in their list[7]. Then it will emit the `step` signal. When the component receiving that `step`-signal has completed its task and the module has received the `stepReturn` signal the module destroys all connections which have a 0 in their list and creates all connections with a one in their list (this can lead to a connection being destroyed and recreated shortly thereafter). Then the module emits a `step` signal and so forth. When the step with the number equal to the highest integer in any temporary connection's list finished, the program terminates.

This mechanism allows:

(a) To perform several consecutive steps by adding exactly one temporary `step`-connection per step.

(b) To redirect data between components as needed for different steps. For instance, the data from a data buffer component should first be passed to a PCA matrix computation component and at a later step to a FDA matrix computation component.

Note that a module does not take into account any temporary connections from its sub-modules. Rather a (sub-)module whose `stepIn` slot is triggered will start the process of emitting `step` signals itself and only after all its steps have been finished it will notify the super-module via `stepReturn` (just as if it was a normal component). This way a module can be executed several times during a program run. This is for instance very useful when doing a cross-folding of data.

An example deployment for computing PCA and FDA matrices together with a detailed explanation can be found in the appendix A on page 121. The explanation of the general concepts in this section is pretty abstract and, thus, this concrete example might be helpful in understanding the step-mechanism and temporary connections.

5.8.6. UBiCI Plug-ins/Extensibility

The UBiCI framework is a versatile tool. The downside of its growing versatility was, however, that it had a growing number of software dependencies that needed to be fulfilled before one could use it. Even worse, although the UBiCI framework was supposed to be a multi-platform tool, some drivers for external devices (EEG data recording or (robotic) device control) were not. Therefore, at some point, the framework was split into several libraries. All components depending on a certain external library were packed into one UBiCI library. E.g. all components depending on OpenCV functions were packed into `libubici_computation`. This also lead to a division by tasks.

[7] It is common in computer science to start enumerations with 0, rather than 1.

Each of these libraries is fit for dynamical loading at runtime (they are plug-ins). If a given deployment needs a certain library to operate it can declare this in a special line within the deployment file and the software will ensure the library is loaded before the deployment is realised in a module.

The plug-in mechanism allows for an almost arbitrary extensibility of the UBiCI. Anyone with some C++ and Qt programming skills can set up a new plug-in and define his/her own components and data types (and unit tests).

5.8.7. Comparison with Existing Software

Several solutions for BMI software already existed prior to the UBiCI. I will briefly describe some of the more prominent solutions in the field and compare it with UBiCI:

OpenVibe: OpenVibe is probably the BMI software whose focus is most alike the one of UBiCI. Both aim for maximal flexibility on the cost, that users require some knowledge about underlying algorithms and techniques. OpenVibe is developed by Inria Institute, Fance and its partners. Compared with UBiCI, UBiCI is better extensible and better capable of multi-modal data processing. It might be slightly more flexible but in its current state requires more programmers expert knowledge to use. OpenVibe has a LUA interface, UBiCI has a Python interface. Both have a Matlab interface. UBiCI is also compatible with Android systems. OpenVibe has many more supported recording devices, more visualisations and is generally more advanced in its development (fewer bugs, more advanced GUI interfaces, better documentation etc.).

EEGLAB: The EEGLAB software is a very elaborate MATLAB toolbox which is designed for sophisticated offline analysis of EEG data. However, there are two plug-ins which offer online classification and Brain Machine Interface construction based on EEGLAB: ERICA and BCILAB. ERICA allows synchronization of data sources on different machines over network with precision better than two ms and data preprocessing. BCILAB allows for online classification and stimulus presentation. The advantage of this system is, that it allows access to the vast amount of mathematical methods for data preprocessing and classification implemented either in MATLAB or in EEGLAB. The downside it that the closed source nature of MATLAB might impede extensibility of the framework, especially when the extensions should be in C/C++ or other languages than MATLAB. This can be an obstacle especially when including control of robotic systems. Additionally EEGLAB, ERICA and BCILAB, although they are free and open source, all rely on the commercial MATLAB software.

BCI2000: The BCI2000 system is developed in the Schalk Lab in Alberny, New York. This software suit has a different focus than UBiCI in such regard, as it allows to setup a BMI system with minimal expert knowledge on the expense of flexibility and extensibility. It is mainly developed for Windows OS, which is the only OS for which there are pre-build binaries available. It is reported to compile and pass standard tests on Mac OS and Linux, too. But the featured OS is Windows, which is in-line with addressing a non-expert audience.

For our requirements the OpenVibe framework would be probably the most suited from the three competitors presented here. However, the plug-in structure of UBiCI better meets our requirements for extensibility of the framework. Furthermore in 2010, when the UBiCI project was started, the OpenVibe project was a whole lot less advanced, than it is today.

6. Can Hyper-Connectivity Occur in Machine-Mediated Interaction?

In this chapter I will describe a study that aims to evaluate whether or not neural hyper-connectivity can be shown during interaction which is machine-mediated (via a BMI). I called the experiment I conducted for this study *Hyper-scanning Experiment for Machine-Mediated Interaction Estimation* (HExMInE). This study is, to my knowledge, the first, aiming to evaluate whether or not hyper-connectivity, previously demonstrated in different direct interaction settings (e.g. guitar playing, pilots in flight simulator, etc. compare section 3 on page 23), still occurs when interaction is machine-mediated. As this study aims to tackle a rather fundamental question, I aimed to have as few independent variables in my experimental design as possible. Hence, the experimental setting is relatively plain as compared to the much more involved setting in the second study of this thesis (iCusss, described in section 8 on page 93).

I first describe the experimental idea (section 6.1 below), the robotic system used (section 6.2 on page 62), the design of the training phase necessary to operate the BMI (section 6.3 on page 63). Then I will go over the BMI system and experimental setup (section 6.4 on page 66) and finally present and discuss the experiment's results (sections 6.5 on page 74).

Additionally I will have a short excursion, discussing some possible implications of the choice the robotic system in such experiments in section 6.2.1 on page 62.

6.1. Experimental Idea/Questions

Quite some studies on EEG hyper-scanning and more specifically hyper-scanning during human-human interaction have been published in the last few years, e.g. Babiloni et al. (2007b); Lindenberger et al. (2009); De Vico Fallani et al. (2010); Dumas et al. (2010); Astolfi et al. (2012); Astolfi et al. (2004); Astolfi et al. (2014, 2010b, 2011b) and more. With this study I aimed to contribute to that corpus by evaluating the question whether or not the hyper-connectivity described in these publication still occurs when interaction is machine-mediated.

There is no such thing as a simple experimental design including the interaction of two persons via BMI-controlled robots. For this experiment I wanted to get close to this, however. As the feasibility of hyper-scanning experiments has not been shown for machine-mediated interaction before, one aim of the experimental design of this study was to limit the number of independent variables.

The first step in that direction was the choice of a preferably plain robot which poses a minor distraction, at most. I decided to use a small cube-shaped robot called Tangible Active Object (TAO) (Riedenklau et al., 2012b) which has been developed as part of the research towards tangible user interfaces conducted at CITEC. The TAO system, including the robots and a special desk (T-Desk), was mainly developed by E. Riedenklau and he was of great aid in developing an interfaces between the TAO system and the UBiCI software framework. I will describe the TAO system in some detail in section 6.2 on page 62.

The second step was choosing a rather simple, i.e. easy to understand/unambiguous task. The operation of an active BMI always demands some form of mental activity on part of the participant, be that imagination of movements, fulfilment of mental tasks or attending to certain stimuli. Thus, having an easy to understand task predominantly means having a rather obvious/intuitive mapping from the mental strategy to robotic actions. The robotic actions I settled with were, directing the (own) TAO towards the left- or the right-hand side of the T-Desk, respectively. The mental task I chose, was, correspondingly, imagination of movements of the left or right hand, respectively. This motor imagery (MI) triggers ERDs in the contra-lateral primary motor cortex which can be measured in an EEG and detected automatically with reasonable accuracy (see section 2.2.5 on page 17 for details on MI and ERDs). Steering a robot to the left/right by imagining movements of the left/right hand might well be the most intuitive mapping from mental strategy to robotic action possible.

Participants received a cue which hand to use before each trial. The knowledge about that cue was not used in the classification system in any way, but the cues were given such that different conditions emerged

(see section 6.4.1 on page 66). I wanted to facilitate some degree of interaction level information flow between the participants. Hence, I delivered the cue for each trial to the participants via headphones (i.e. audio) such that none of the participants could know their partner's task for that trial and had to infer it from the TAO's movement.

6.2. The TAO Robots

The TAO robotic system consists of a series of cube-shaped robots and a special desk upon which these robots can navigate (the T-Desk). It is shown in figure 6.1. These robots have two miniature chain drives for actuation, an Arduino micro-controller and a XBee wireless communication module. Due to their modular design they could be equipped with various additional sensors or a display. However, I used none of these extensions.

The TAOs have a 3D printed housing of $5cm$ length. At their bottom a total of 13 IR-LEDs are located. These are used for localisation of the different TAOs on the T-Desk. Seven of them are used to pinpoint the TAOs orientation and six encode the TAOs id. The robots were equipped with a LiPo-battery which lasted for about 40 minutes of continuous operation.

For the T-Desk's surface a semitransparent material has been used, which allows to track each TAO in the image of a camera placed at the bottom of the T-Desk facing the table surface from below. This camera has been equipped with an IR-filter, such that in the otherwise black camera image only the bright IR-LEDs of the TAOs were visible. Additionally a short range beamer was mounted to the T-Desk, which allowed projection of arbitrary content on the semitransparent surface.

The TAOs could execute commands such as to drive forward or backward and rotate, however, could not correct for errors in the trajectory autonomously. Therefore a software framework running on a computer continuously evaluated the camera image and send commands to the TAOs via XBee communications. This framework could execute commands such as "navigate TAO A to position x, y". The software suit features a modular design and is easily extensible. The communication with the software and between different software components was implemented using XCF, a XML-based communications middle-ware developed at Bielefeld University.

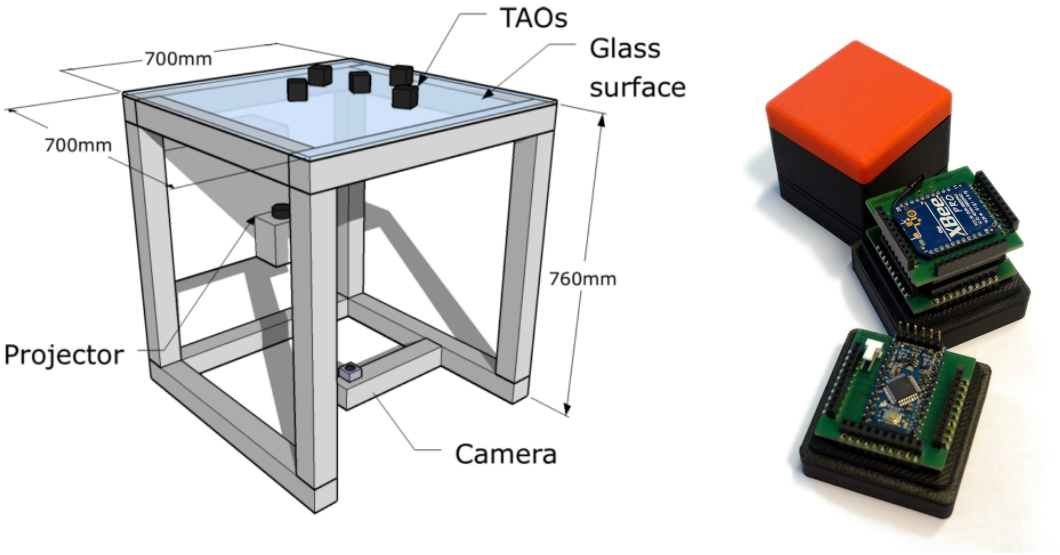
6.2.1. Excursion: Choice of Robots and the Implications

Robots exist in a rich variety for forms. It is reasonable to assume that different factors of the robots form might have an impact on the neural data. This is particularly reasonable when robots are anthropomorphic, i.e. take, at least in parts, human like form. For experiments aiming for the neural foundations of human social interaction, these implications might be even more important. Several studies demonstrated the relevance of appearance, smoothness of motion and spatial behaviour of a robot for its acceptance by a human. Saygin and Ishiguro (2010) for example showed that the so called uncanny valley, i.e. a drop in the perception of human-likeness of robots, has a neural correlate which can be shown in fMRI data.

For the HExMInE I deliberately chose a robot which a.) poses a minimal distraction to the participants and b.) is far from any anthropomorphic shape, thus not tempting the participants to attribute some human intentions, feelings or other to it.

For the later experiment (described in section 8 on page 93) I chose differently. The robotic hands used for that experiment resemble human hands in many regards. Furthermore, the participants saw their own actions and that of their partner being carried out by the same robotic hands from a view position similar to where a head would be located in relation to the robotic arms, i.e. from an I-perspective. This might suggest a feeling of sharing the same body.

This contrast between the two studies of this thesis is intentional. As I approach from a plain experimental design to a more involved setting from the first to the second study, I also move from total "non-anthropomorphism" to high anthropomorphism.



(a) A scheme of the TAO system. From Riedenklau et al. (2012b)

(b) The TAO cube-shaped robot. Housing, XBee communication module and Arduino. From Riedenklau et al. (2011)

Figure 6.1.: The TAO robotic system. A sketch of the entire system and a photo of the real robots.

6.3. Training

Virtually all present BMIs need some labelled training data to parametrise their classification system for that specific participant (training). Some brain activity patterns such as ERDs, however, also require the participant to acquire a skill in invoking the employed BAP, while other BAPs occur without training of the participant (such as P300). The participant training for ERDs should ideally take place on different days with a pause of not more than two days between two training sessions. During training, parameters of the classification are constantly recomputed, allowing participant and classifier to adapt to one another.

I designed a training process for the participants with three stages:

1. EMG-based data collection for classifier initialisation.
2. Screen-based MI-training.
3. Training with TAO.

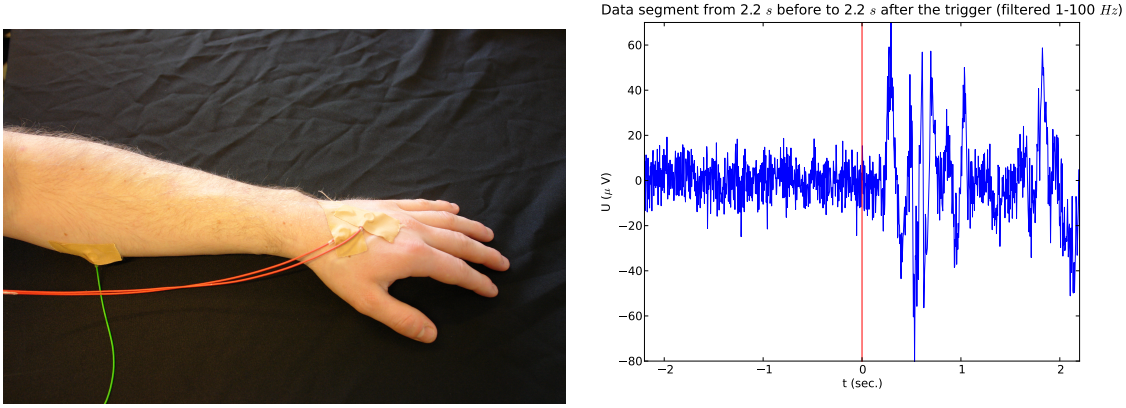
I will go over these stages in the next sections.

All these trainings had in common that 200 trials were recorded. The participants received instructions on what hand to use for MI before each trial, such that I could label the recorded data for parametrisation of the classification system. One-hundred trials used left-hand and one-hundred used right-hand MI, in blocks of ten (ten left-hand, ten-right hand, ten-left hand, ten right-hand, ...).

6.3.1. Collecting ERD Data Using EMG

The classical training process for the use of a MI-based BMI usually starts with a randomly parametrised classification system. Participant and classification system need to adapt to one another over several days of training. I aimed to speed-up this process by using a reasonable, i.e. non-random, initial parametrisation of the classification system.

One idea would be, to use ERD data from previous recordings with other participants for the initial parametrisation. However, pasted experiments aiming for cross-participant BMI-classification were unsuccessful. Hence, it was unclear whether or not such an initialisation would actually lead to a speed-up in the training process.



(a) Fixation of electrodes: Reference and ground on the back of the hand and the recording electrode on the lower arm.

(b) Muscle activation is marked with a heavy increase in EMG-activity.

Figure 6.2.: The onset of a movement can be identified by recording and analysing the muscle activity of the participants (EMG). A rapid increase in the standard deviation of the EMG marks the muscle activation.

It is known from the literature that a.) ERDs are not only triggered during MI but also during motor execution (and observation of movements) and b.) that the motor cortex starts movement planning shortly before a movement is actually executed (Blankertz et al., 2001). This motor pre-planning already triggers an ERD which lasts during the entire movement execution until the movement has come to an end, but is occluded by motor artefacts after movement onset. Hence, there should be a short time window (in the range of a few hundred milliseconds) during which an ERD would occur, while the muscles were still resting and, hence, the EEG signal would not suffer from any contamination from muscle activity. I hoped this motor pre-planning ERD would resemble the ERDs triggered by MI to a degree that it could be used for a reasonable initialisation of the classifier.

Hence, if I could accurately identify the onset of a movement, in the first training session I could ...

1. ... ask the participants to execute hand movements rather than imagining them ...

2. ... collect data from right before the movement onset and ...

3. ... use that data to initialise the classifier for the second session.

To identify the movement onset, during the very first session with a given participant, I did an Electromyogram (EMG) (see figure 6.2a) of the two lower arms additionally to the EEG and instructed the participants to execute movements with either the left or the right hand[1]. Muscles for hand actuation are located in the lower arm, hence the EMG was recorded there, rather than on the hands themselves. The hands hardly house any muscles at all, which makes them suitable place for reference (and ground) electrode. In a later offline analysis, I identified the onset of the movement by analysing the standard deviation of the EMG (see figure 6.2b). A rapid increase of the standard deviation marked the start of the movement, see figure 6.2b.

The participants were instructed which hand to move before each trial via a screen display. This screen showed a text stating which hand to use and a triangle pointing the respective direction (either left or right). Additionally this triangle was colour-coded: Yellow colour represented usage of the left hand and blue colour usage of the right hand (see figure 6.3). In contrast to the later stages of the training, no feedback was provided to the participants.

This allowed me to identify the onset of a movement with high accuracy and to collect EEG data from the participant which I could be confident it would a.) contain ERDs and b.) be free from muscle artefacts stemming from the hand-movements. This data should therefore allow for a reasonable initialisation of the classifier used during the second session.

[1] Rather than only imagining hand movements as for the later sessions.

6.3.2. Screen-based training

Following the EMG data recording, the participants underwent a series of sessions during which they only imagined movements of their hands (rather than executing them). As for the EMG data recording, at the beginning of each trial the participants were instructed which hand to use (for MI) by a screen display.

Furthermore the participants where instructed to experiment with different (imaginary) movements, as it was our experience from earlier MI experiments that participants achieved different results imagining different movements and that the movement which worked best was different for different participants. During these sessions feedback was provided on a screen after each trial. Figure 6.3 shows this feedback. A bar extended into the direction the classification identified (to the left side for left hand MI and vice versa). The length to which it extended depended on the (scaled) confidence of the classification result (this constitutes a continuous feedback, compare section 2.2.3 on page 13). I will discuss classifier confidences and their scaling in sections 6.4.3 (page 68) in more detail. The bar was again colour coded with the same colour-hand mapping used for the triangle (see above).

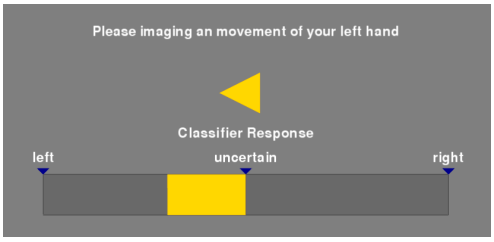


Figure 6.3.: The instructions and the feedback presented to the participant on a screen. Directions were colour coded: Left - yellow and right - blue

For the first screen-based training session (second overall session) the classifier was initialised using the data from the EMG-based session (see section 6.3.1) and the classifier was kept static, forcing the participant to adapt to the classifier rather than allowing for a mutual adoption process. During the later screen-based training sessions the classifier was initialised using the data from the respective previous training and was recomputed every 20 repetitions. This way, during the first screen-based training the participant could learn to volitionally trigger ERDs because the classifier could already provide reasonable feedback. During that second and further screen-based trainings the mutual adoption of classifier and participant could take place as in the classic MI-training paradigm.

The screen-based training war repeated until participants had a stable, reproducible control over the system, but not more than three times. Participants who had no stable control after this amount of sessions were excluded.

For some participants the EMG-based initialisation of the classifier worked well. They already established good control over the system towards the end of the first screen-based session. For others it did not have the desired effect and I had to exclude an unusual high number of participants who would not gain (stable) control. A more systematic evaluation of this approach might reveal more evidence.

6.3.3. Training with TAOs

For the final training session the screen-based feedback was removed. Instead the TAO system was introduced to the participants. In fact this training resembled all aspects of the experiment, including the possibility of TAOs to change direction mid-trial and the projection on the T-Desk (see section 6.4.3 on page 68, figure 6.4 on page 66 and figure 6.8 on page 70). Exactly three aspects of this training were different compared to the experiment:

1. No second participant was available and hence the second TAO did not move at any time (although it was physically present).

2. During this training the classifier was still retrained every 20 trials. During the experiment it was kept static.

3. During this training, instructions were given by a display (projected onto the T-Desk) and instructions changed in blocks of ten left, ten right. During the experiment, instructions were given via earphones and in randomised order (see section 6.4.1).

All participants who had completed the second stage of the training also showed stable control over the TAO system.

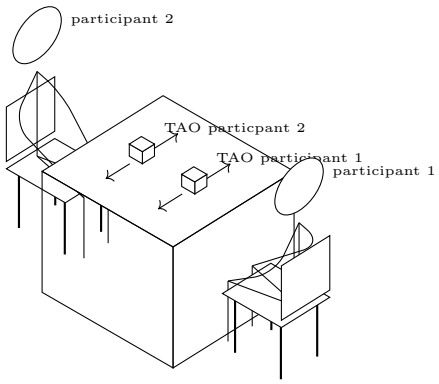


Figure 6.4.: The setup of the HExMInE experiment. The participants were seated in opposition to one another, equipped with earphones, eyetrackers and EEG devices. Two TAO robots were placed on the T-Desk which moved according to the participants' motor imagery towards the left or the right T-Desk border. The TAOs were programmed to move along two axes both oriented parallel to the participants' left-right axis. These axes split the T-Desk approximately in three equal parts. A countdown marking the start of the trial was projected onto the tables surface.

6.4. Experiment Setup and Conduction

In the following, the experimental setup will be described. First, from a participant's perspective: Figure 6.4 shows how participants and the robotic system were arranged and in section 6.4.1 below which instructions where given to the participants and how these were given. Second, the technical implementation is described. Namely, the data recording in section 6.4.2 on page 67, the data processing in section 6.4.3 on page 68 and the interface between TAO system and the UBiCI BMI framework in section 6.4.4 on page 72.

The experiment took place on a different day than the training sessions. To re-familiarize the participants with the system and to get some current data for classifier parametrisation, a shortened training (100 trials is total) was conducted directly before the experiment. This was similar to the screen-based training described above, except for the fact that both experiment partners were present and performing the same task on two machines next to one another. Finally, it should be noted that during the experiment no recalculation of the classifier matrices was performed.

6.4.1. Cues and Conditions

At the beginning of each trial the direction in which to steer the TAO was indicated to the participants via earphones.[2] When both participants had received their cue, a countdown from three to zero was projected onto the T-Desk. When the countdown expired, the participants were supposed to start their MI. One second later I had acquired sufficient data for a (first) classification and subsequently the TAOs started their movement (compare section 6.4.3).

In order to compare joint action with solo action, during some trials one participant should be passive. Consequentially, there were three (rather than two) possible cues, one of which each participant received at the beginning of each trial.

Left-hand MI: Imagine a movement of the left hand to which the TAO should respond by moving towards the left side of the T-Desk (from that participant's perspective).

Right-hand MI: Imagine a movement of the right hand to which the TAO should respond by moving towards the right side of the T-Desk (from that participant's perspective).

Observe: Do not perform any motor imagery but watch the partner's performance, i.e. his/her TAO's movement. During that trial the BMI control for the observer's TAO was switched off, because not issuing any commands to the system can be a mentally demanding task in its own right, i.e. the participant would not be passive.

[2] The knowledge about the participants' cues was not used in the classification system in any way.

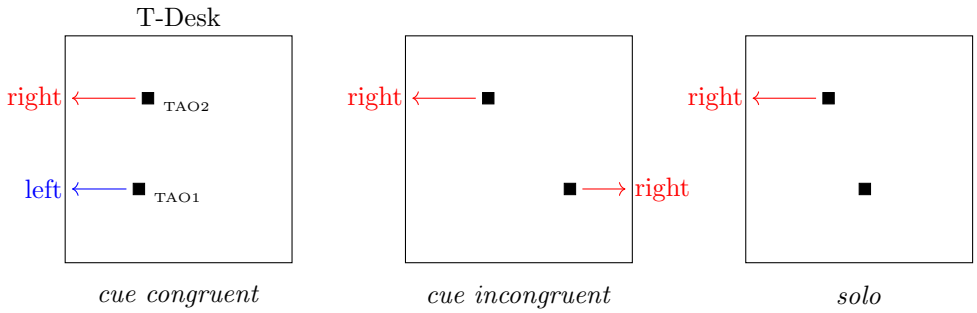


Figure 6.5.: The three cue-based conditions. Both TAOs shall be steered towards the same desk's end (cue congruent), both TAOs shall be steered towards different ends of the desk (cue incongruent) and only one TAO shall be steered at all (solo).

These three cues were combined in a way that it yielded three different conditions which are depicted in figure 6.5:

Cue congruent: Cues were given such that both TAOs would move towards the same side of the T-Desk. Hence, one participant had to be cued to move his/her TAO towards the left border while the other participants was cued to steer his/her TAO towards the right border (from the respective participant's perspectives). There were 50 such trials during the experiment.

Cue incongruent: Cues were given such that the TAOs would move towards different sides of the T-Desk. Hence, both participant would need to receive the same cue (either both left-hand MI or both right-hand MI). There were 50 such trials during the experiment.

Solo: One participant was cued to not perform MI and observe, while the other participant was instructed to perform MI. There were 50 such trials during the experiment. In 25 trials participants one was in the role of the observer and in the other 25 trials participants two was in that role. The cues for the active participant were divided between *left-hand* and *right-hand MI* equally except for one trial difference (one cannot divide 25 by two).

The order of these trials has been randomised once and was than kept fix for all of the participant pairs. There were no trials during which none of the participants had to perform MI (i.e. both would have received the observe cue).

6.4.2. Data Recording

Per participant one GUSBAmp EEG devices from GTec has been used for data recording (one device was of version one one of version 2). These offer 16 channel and hardware bandpass and notch frequency filtering. Ag/AgCl electrodes have been used and were placed on the scalp using an EasyCap electrode placement cap. An electrolyte EEG paste was applied between electrode and scalp.

I did not want to make any assumptions where in the cortex relevant activity, i.e. hyper-connectivity might occur. Hence, the electrodes have been positions to cover most of the cortex. The only exception from that guideline was that two electrodes have been placed over each motor cortex side (C4, C2, C3, C5) to achieve a stable ERD classification. For an optimal ERD classification accuracy, it would have been advantageous to concentrate more electrodes over the hand-related areas of the motor cortex. However, it has been shown in the past, that two electrodes are sufficient to achieve stable (though not optimal) classification. The placement of the electrodes is depicted in figure 6.6 on page 68.

The reference- and the ground-electrode were placed on the mastoids. There is no brain tissue located on the other skull side on that altitude and therefore the mastoids are electrically neutral or, more precisely, electrically independent from neural activity.

When placing the electrodes, impedances were kept below $10k\Omega$. Impedances were measured using a standard UFI model 1089 mk III impedance gauge.

EEG data was recorded using a different computer than the one that was being used to display the

countdown. These computers had their clocks synchronised (using NTP)[3] Allowing them to operate with the same time reference.

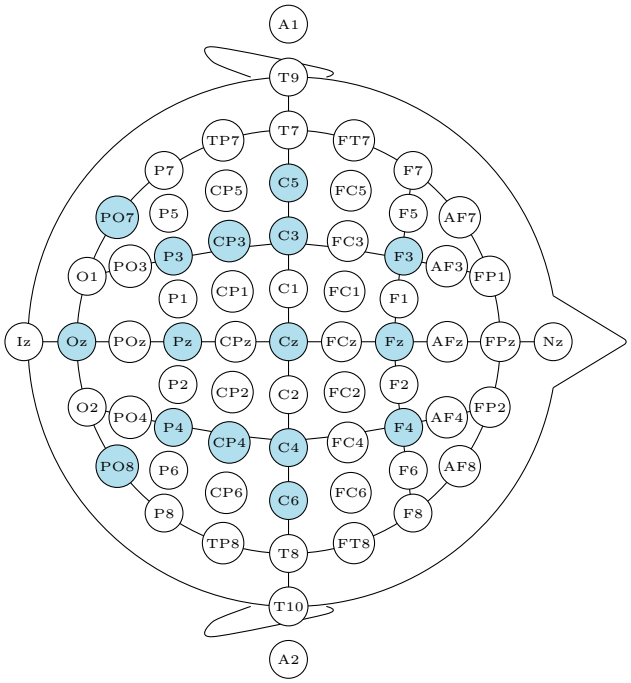


Figure 6.6.: The positions of the electrodes during the HExMInE experiment. For an optimal hyper-scanning analysis the entire cortex was covered equally with a slight bias towards the motor cortex to allow for a better motor-ERD classification.

Additionally to the EEG electrodes, participants were equipped with eye trackers (SMI iViewX) and with earphones (see section 6.4.1). The idea of using the eye trackers was, that the gaze data collected would eventually give timing cues for the later offline analysis of the experiment data. However, inspection of the eye tracking data revealed that participants mostly fixated a random spot and focused on the motor imagery task. This is in-line with the reports of the participants after the experiment. Thus, the eye-tracking data was not used in the later data analysis.

6.4.3. Online Data Processing

Figure 6.7 on page 69 gives an overview of how online classification of the EEG data works. Nodes from that graph represent individual steps in that process in accordance with the UBiCI concept (compare with section 5.8 on page 55). The nodes will be treated within this section roughly in a top-down order. Note that classification was performed for each participant separately and that thus one such deployment existed for each participants, each parametrised individually for the respective participant.

GetData: The top three nodes receive data over network and pass it on. Two of the nodes receive meta information: i.) Meta information[4] describing the current trial. For this experiment the information contained is rather plain and it serves organisation purposes mostly. ii.) Meta information[5] about a data segment/epoch which needs to be generated, in particular at which point in time the countdown presented to the participants expired and, thus, when participants started to perform MI. The third node receives the actual EEG data.

DataStorageCenter: All of this information is first passed to a node which stores it, before any filtering, segmentation and so forth has been applied. I thus have full information for later offline analysis.

TimeToSample: EEG data and the timing information are then passed to a component which synchronises that information. The EEG data stream is furnished with one times stamp every 32 samples (the number of samples read from the EEG device's buffer en block). I need to identify the sample in the EEG data stream which was recorded at the time the countdown expired. The TimeToSample component does this by interpolation of time stamps. It then adds the number of the closest sample (counted from system start) to the epoch info, before passing it on.

We will leave out the ArtificialEpochInfoGenerator for the moment and assume the epoch info connection would directly lead to the segmentation. I will come back to the ArtificialEpochInfoGenerator at the end of this section and explain its purpose.

TriggeredSegmentation: Two types of information enter the TriggeredSegmentation component: Continuous EEG data and epoch infos. EEG data enters the component as a continuous stream, but only individual data segments/epochs leave it. For each epoch info that enters the segmentation component, an epoch is generated and passed on. The length of the epoch to be generated by the segmentation is given

[3] Within the Bielefeld University network, which includes a NTP reference clock, the jitter between the university machines' clocks is usually below one millisecond. Since I record with $256Hz$, which is to say roughly one sample every four milliseconds, this degree of synchronisation is sufficient for my purposes.

[4] An UBiCI decision request object.

[5] An UBiCI epoch info

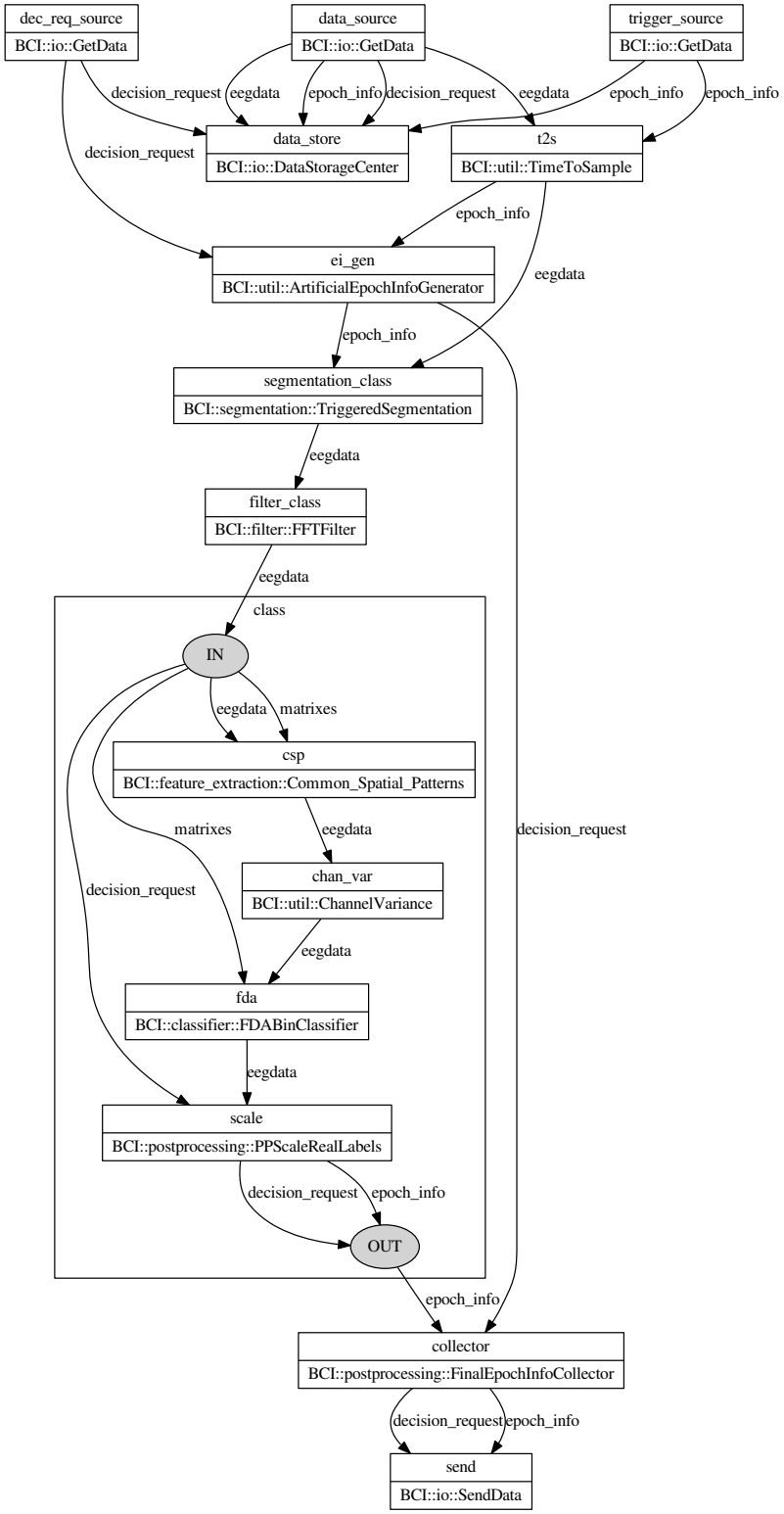


Figure 6.7.: The UBiCI deployment used for classifying the EEG data online. Each node represents an UBiCI component (see section 5.8 on page 55). The upper line gives the component name within the deployment while the lower line gives the (fully qualified) component's type (i.e. which task it performs within the deployment). The edges represent connections over which (meta-)data of the given type is transmitted.

Figure 6.8.: A screen-shot from the video of one participant's eye-tracker during the HExMInE experiment. Participants are seated in opposition to one another and each participant has his/her own TAO to control. During the TAO training the second chair was empty and the TAO on the far end did not move.

by the epoch info which triggered its generation. The segmentation can be configured to include an offset between the time stamp of the epoch info and the start of the epoch. However, epochs are always time-locked to that time stamp. The segments generated for different epoch infos may overlap. The components buffers and copies data as needed. The segments leaving the segmentation component keep a pointer to the epoch info which triggered their creation. This is why no epoch information connections exist beyond this point: epoch infos ride as hitch-hikers with their epoch. It is important to note, that from here on the EEG data no longer is a continuous data stream, but only segments which are time locked to epoch infos!

FFTFilter: The next step is a FFT-based frequency filter (STFT, actually). See section 5.1.3 on page 37 for details on the mathematics of FFT. Filtering is done as a bandpass filtering using a high pass and a low pass filter. Frequencies have been adapted to work for each participant, but aim for the $\mu-$ and β-band.

The next components are being encapsulated by an UBiCI-module named *class*. These components do the feature extraction and classification. They are used in that configuration in different other deployments, too. Keeping them within their own (sub-)module allows re-usability of this part of the deployment. It has no impact on the function and only negligible impact on computation time.

CommonSpatialPatterns: As described in section 5.2.2 on page 40 the CSP is a linear method which is trained to transform the data, such that the two classes can be distinguished more easily. It transforms the data such that for data samples belonging to class one (e.g. left-hand MI) the variances in the first few channels is maximised and minimised in the last few channels . For class two (e.g. right-hand MI) the variance is minimised in the first few channels and maximised in the last few channels. A consequence of this behaviour is, that the inner channels usually contribute little in terms of distinguishability of the data. Hence, the inner channels are dropped when applying the CSP. The exact number of channels that should be retained is depended on the individual participant and determined on the training data (usually four or six).

ChannelVariance: Given the data characteristics introduced by CSP, described in the previous paragraph, the logical next step is, of course, to compute the variance on each channel separately, thereby collapsing the time dimension. We now only have a data vector left with the number of entries equal to the number of channels after CSP (usually four or six).

FDABinClassifier: The actual (binary) classification is now done in the last but one component: The Fisher Discriminant Analysis. It is a linear classifier. For details on the method please refer to section 5.2.3 on page 41. Other then the previous components, it does not actually change the EEG data object it receives, but sets the *classification label* field of the epoch info attached to the EEG data object. This label contains a binary classification result (either 0 or 1) and a so called classifier confidence which is a real value describing the distance of the data point to the dividing hyper-plane spanned by the FDA. The sign represents on which side of the hyper-plane the data point is located and is therefore redundant with the (binary) classification result.

Different BMI systems in the literature (e.g. Lenhardt et al. (2008)) use per-option sums over this distance to improve the reliability of BMI classification, particularly for P300 BMIs. This approach implicitly assumes that data points which are further away from the dividing hyper-plane, are less likely to be misclassifications.

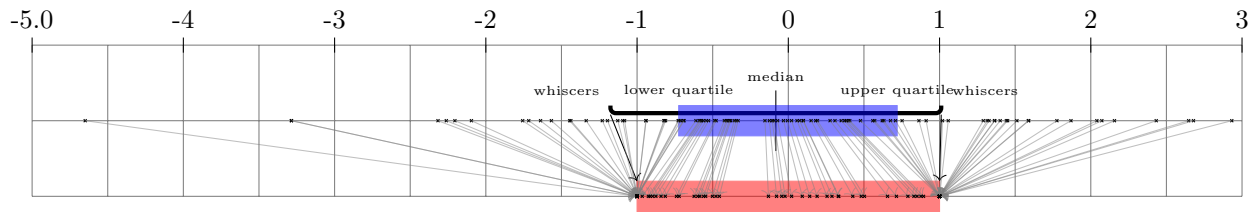


Figure 6.9.: Confidence values from the classifier are collected, quartiles (blue) are computed and new values are scaled into a given interval (red) based on that quartiles. The quartiles are updated continuously.

This assumption could be debated and has, to my knowledge, never been mathematically substantiated. However, during practical applications, approaches based on this assumption have been shown to make BMIs more robust and reliable.

PPScaleRealLabels: The TAO robot of a given participant should react to MI of hand movements. However, if the (absolute) confidence is rather small, the robot should remains stationary, representing the system's uncertainty.

During the screen-based training a bar should extend depending on the classifier confidence (see section 6.3 on page 63). The degree to which the bar could extend was, of course, limited (if only be the borders of the screen).

There are no theoretical boundaries to the values of the confidence. Even worse, the distribution of confidence values can change between different recordings depending on a series of factors: Impedances, degree of focus of the participant, recording hardware used, electrical noise and so forth. This constitutes a problem for both, feedback during screen-based training and uncertainty interval during TAO control.

The PPScaleRealLabels component aims to scale these values into an interval $[-1, 1]$. -1 and 1 would then represent the maximal extension of the bar during training. The sub-interval $[-0.1, 0.1]$ would be interpreted as uncertainty in classification, causing the TAO to stop its movement.

I needed to ensure that the distribution of scaled values was such that scaled data points would occur in the entire interval $[-1, 1]$ in practice. If not so, (during training) the bar might eventually only ever extend to a fraction of its potential range or (when controlling a TAO) all classification results might fall into the $[-0.1, 0.1]$ interval after scaling and the TAO would never actually move.

The bottom line here is that a fixed interval in the space of FDA confidences which would then be linearly projected onto the $[-1, 1]$ interval is not a feasible solution. I needed a method which scales the confidences based on the actual distribution of (past) confidence values. However, it should be insensitive to outliers which could occur and potentially hinder the component from using the entire $[-1, 1]$ interval.

To achieve this, the so called box plots from statistics were used as an inspiration. The box plot defines any data point whose distance to the data's median is larger than 1.5 times the inter-quartile range to be an outlier. This idea was adapted: The interval in the confidence space $[m - 0.7 \cdot iqr, m + 0.7 \cdot iqr]$ with m as past confidence values' median and iqr as the inter quartile range was mapped to the $[-1, 1]$ interval.[6] Every point within the $[m-0.7 \cdot iqr, m+0.7 \cdot iqr]$ interval was scaled to the $[-1, 1]$ interval linearly. Every point outside this (source-)interval (the outliers) are projected to the borders of the target interval (-1 and 1 respectively). Figure 6.9 illustrates this process. It was intended to modify this behaviour such that the mapping between values was not linear but would follow a continuously differentiable function, i.e. the mapping function would converge towards $[-1, 1]$ when approaching $-\infty$ and ∞, however, as a rather cosmetic improvement, this has, to date, not been implemented.

ArtificialEpochInforGenerator: Finally, I come back to the ArtificialEpochInfoGenerator. In order to create a system that reacts dynamically to the participant, the TAO robot should be able to change direction during the trial. Therefore the classifier should "reconsider" its choice every half a second.

When an epoch info reaches the ArtificialEpochInfoGenerator, this generates a series of epoch infos. It multiplies the epoch info it received, so to speak. All these epoch infos share the time stamp (i.e. the point in time when the countdown expired) but they have different length set that will be used by the segmentation. For this experiment the lengths started with one second and then increased up to 6.5 seconds in steps of 0.5 seconds. All these epoch infos then entered the segmentation, where epochs with the desired length were produced as soon a sufficient EEG data was available to the segmentation. Figure 6.10 depicts which epochs

[6] We actually started with a factor of 1.5 as in the original box plot diagrams. The factor was then adapted empirically until the entire range $[-1, 1]$ was exploited regularly.

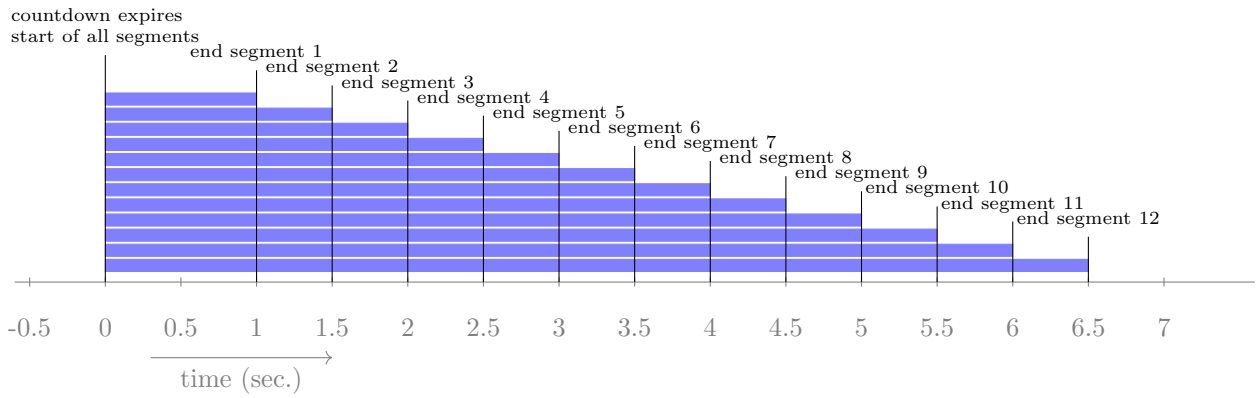


Figure 6.10.: The segments resulting from the use of the artificial trigger generator: All time-locked to the expiration of the countdown with increasing length from 1 to 6.5 seconds with 0.5 seconds step size.

are produces and how the overlap. Hence, every half a second one epoch left the segmentation which was half a second longer then the preceding epoch, travelled down the deployment until the FDA classifier would classify it and was then translated to the corresponding robotic movement.

In the end the classifier would re-evaluate the entire data that had arrived since the countdown's expiration every half a second and the TAO robot could eventually start or stop moving or even change direction midway. During the training with the TAO, most participants proved capable to use this to counter an initially wrong classification.

FinalEpochInfoCollector: As mentioned above the decision request represents a decision to make via the BMI. This object contains all the information about that decision, in particular its final outcome. The FinalEpochInfoCollector stores a copy of each epoch info which passes it and associates them with their decision requests. When for any given decision request it had collected 12 epoch infos (and therefore all classification results), it copies the classification label from the latest epoch info (which is based on the most data) to the decision request and send the decision request onwards.[7]

The final component in the deployment is again a networking component which sends the intermediate (epoch infos) and final (decision request) classification results to another computer on which these results are translated to commands to the TAOs. How this is done, i.e. the interface between the UBiCI framework and the TAO system is described in the next section.

6.4.4. Robotic Interface

The C++-API for operating TAOs provides three main classes:

TAO Scene: Is a GUI interface which displays the position and orientation of any TAO recognised by the system.

TAO Serial Control: This class provides a series of function which allow to steer the TAO on a basic level (forward, backward and rotation).

TAO Planner Interface: This class allows to give high level commands for the TAOs, namely a target position and orientation. The Planner would then plan a path towards that position. When the TAO is on its way, the TAO Planner monitors the TAO's position and corrects the navigation if needed. At the time this experiment was conducted it could, however, not do any obstacle avoidance (including other TAOs). During this experiment this problem should, however, hardly occur.

The TAO code heavily depends on Qt. The TAO Scene is a subclass of the QGraphicScene class and the TAOs are QGraphicsItems. The TAO Serial Control is a QObject just as the Planner Interface. This means they inherit the Qt Property mechanism and they can send and receive Qt Signals. It is indeed possible to set some properties of the Planner Interface, call the method `driveByProperties` and the planner will read these properties and execute the command. The properties are:

[7] Regarding decision requests, you may notice the decision request connections within the sub-module. As no decision request connection enters the sub-module and none of the components within the sub-module can produce decision requests these connections are dead. They are, however, used when the sub-module is used within other deployments.

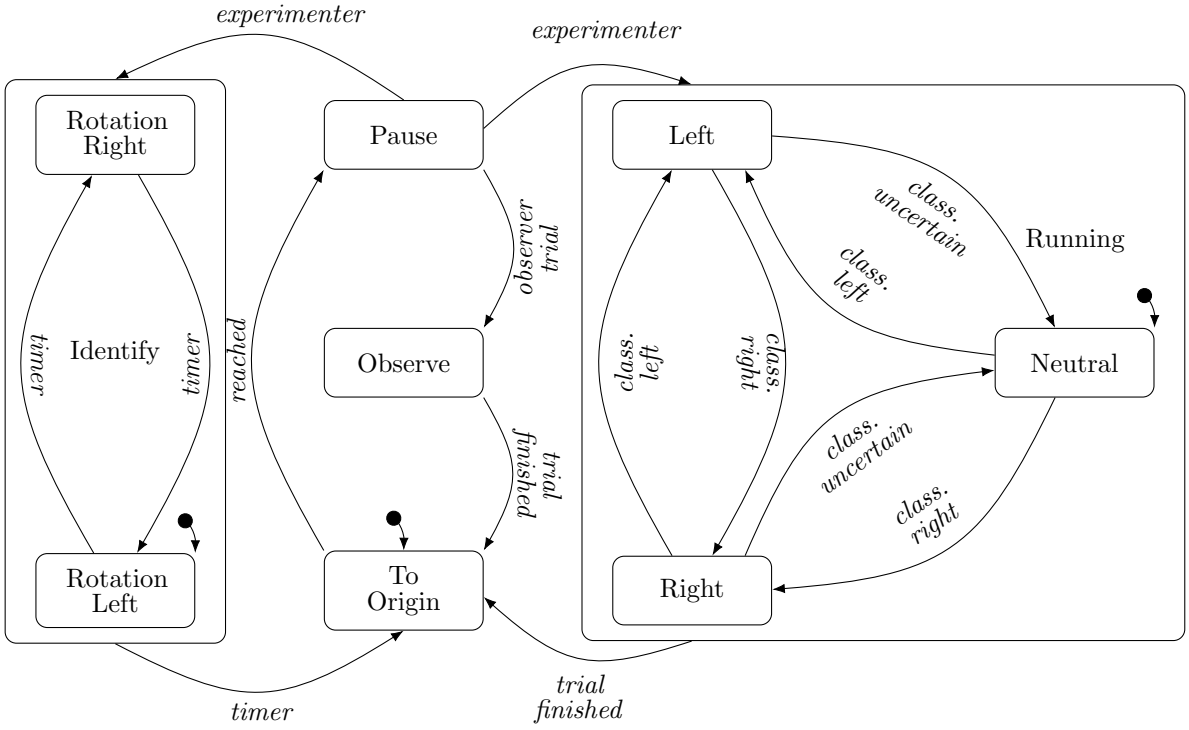


Figure 6.11.: The HSM which governed the experiment. On entering different states, properties were set on the Planner Interface and the `driveByProperties`-method had been called. One of these HSMs was generated for each of the participants.

taoId The communication with the TAOs is done over a (wireless) serial BUS (XBee). Hence, any command send over that BUS reaches all of the TAOs. Each TAO has a fixed id, though, used to identify individual TAOs.

toX The x coordinate of the TAO's target position on the T-Desk.

toY The y coordinate of the TAO's target position on the T-Desk.

toAngle The orientation the TAO should take when it has reached the target position.

The speed of the TAO (a PWM value) could also be set. For this experiment the speed was kept at a constant, low value. Using a higher value, the TAOs would eventually reach the border of the T-Desk too soon and would be passive for most of the trial.

To control the experiment two identical hierarchical state machines (HSMs) have been used. It is good practice for complex robotic systems (and beyond), to have a central software unit to govern such aspects of control flow. HSMs have proven extremely utile in that regard. They are easily interpretable, deterministic and can model virtually any control flow typically required in robotic systems (and experiments) in a comprehensive way.

The HSM used for this experiment is depicted in figure 6.11. It was implemented in the Qt HSM Framework. The initial state is the To Origin state. On entering that state a navigation to the central position on that TAO's motion-axis is triggered with the previously described mechanism. The TAOs could be placed anywhere on the T-Desk and would first drive to their respective home positions. Then the HSM would switch to Pause state. When the operator started a trial, both participants' HSMs switched to Running state, in which they start with the Neutral sub-state. When an epoch info with an absolute confidence larger than 0.1 arrived, the state switched to left or right depending on the label (or, equivalently, the confidences sign). On entering the Left state, properties for the respective TAO were set such that it would drive to a position 10% from the left T-Desk border and the `driveByProperties` method is called. On entering the Right state, things were done accordingly. On entering the Neutral state, the id of the TAO controlled by

the respective participant is set as a property of the Planner Interface, `stopByProperty` is called and the TAO stops.

This way the three sub-states of the Running state switch between one another (independently for each participant) until the end of the trial. Then the control switches to the To Origin state. When that navigation is finished, the control switches back to Pause, waiting for the operator to start the next trial.

When a participant received the instruction to passively observe, his/her HSM entered the Observe state rather than the Running state. After the trial had ended, control was switched to To Origin state (it might as well be switched to Pause state).[8]

The identify state is of little relevance for the actual experiment. I needed a way to ensure, before the experiment was started, that the TAOs were not accidentally exchanged. When the operator switches the state machine to that state, the TAO associated with that machine rotates to the left and to the right again and again for a fixed period, while also triggering an audio output to the corresponding participant. After a fixed time period the To Origin State was entered, in order to realign the TAO with its motion-axis (left-right).

6.5. HExMInE Results

The results of the HExMInE are two-fold: First, I observed a classification accuracy on chance level and consequential a loss of control for almost all participants. In section 6.5.1 below, I investigate the reasons for this. Second, I discuss the results of the connectivity analysis (section 6.5.2 on page 75) which, despite the disappointing classification, still stress the feasibility of the investigation of neural connectivity in machine-mediated interaction scenarios.

6.5.1. Out of Control

The first result of the experiment was that for all but one participant the control they previously established over the system vanished completely. I.e. only one participant had a classification accuracy significantly above 0.5 (which is chance level). Investigating the reasons for this, I applied a method suggested by Pfurtscheller et al.

ERD is essentially a drop in power of a defined frequency band and at a defined scalp location. The relative power of such a band is usually too small for this drop to be observable by visual inspection. Even when filtering the data to that narrow band, the drop is still not prominent enough to be visible in single epochs. For some brain activity patterns (such as P300) averaging over several epochs is a common method to overcome this obstacle. Other than P300 the ERD does not have a fixed sign in raw data, though, it is a change in a property of the rhythm. Averaging over several epochs the contributions of that rhythm in the different epochs will tend to cancel out, rather than making any differences in the band power more salient.

To circumvent this and simultaneously further increase the salience of differences in the magnitude of values, Pfurtscheller and Lopes da Silva (1999) proposed to first square the single epochs and then average over all epochs. After squaring, the signals are all positive and cannot cancel out one another. Additionally, differences in band power will become more salient due to squaring. Furthermore, Pfurtscheller et al suggested to define a reference time window which lies before the (presumed) onset of the ERD and compute the mean for that window. All values in the times series should now be express as percentage values of that mean. Thus, a period for which the curve lies consistently below 100% can indicate an ERD.

I applied this method to a.) the data recorded in the re-training described in section 6.4 on page 66 and b.) the experiment data. It is important to note that only minutes have passed between the two recordings: After re-training participants left their respective computers, have been positioned at two sides of the T-Desk, the experiment software was started and the experiment began.

However, as figure 6.12 shows, the ERD was far less pronounced during the experiment. This can be shown, to different degrees, for any of the seven participants who lost control during the experiment. One possible explanation could be mental fatigue of the participant after the re-training. This is, however, not very plausible. That re-training was like the screen-based training the participants completed during earlier sessions, except for the partner's presence and that it took only half of the time. That all participants had suffered from "premature" mental fatigue on the day of the experiment seems improbable.

[8] Remember that control should be switched off while a participant is meant to be a passive observer. Otherwise I would not need an Observe state and would switch to Running state instead.

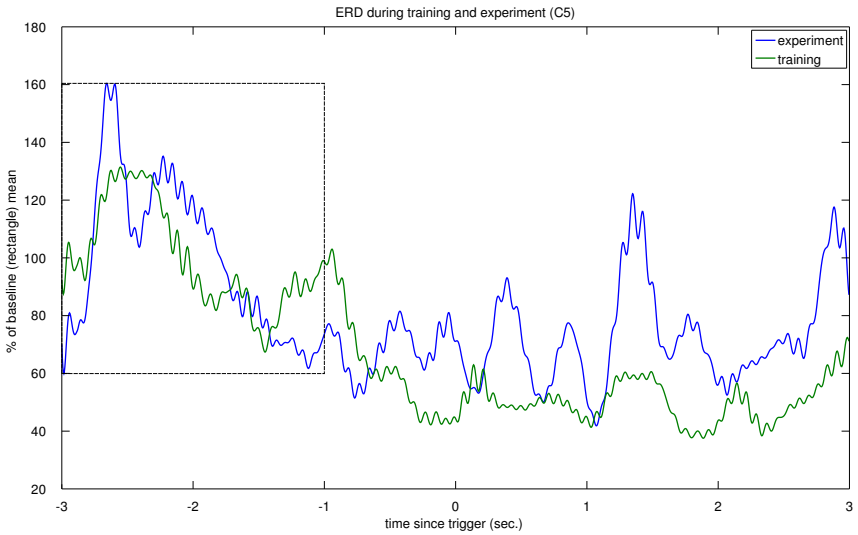


Figure 6.12.: The decrease in the magnitude of the ERD between the retraining and the experiment for one exemplary participant. The rectangle on the left denotes the reference time window used.

The only real difference between the (re-)training and the experiment was that the partners were seated in opposition to one another and had a shared space to operate their robots in. I can only assume that this resulted in a distracting effect causing the decrease in ERD. This hypothesis is, however, hard to prove experimentally and furthermore scientifically not very thrilling. Hence, I refrained from further investigating this issue.

As a side remark it might be added, however, that the ERD did not vanish *completely* during the experiment. Neither is the data recorded during the experiment in-separable. Applying a five-fold cross-validation on the experimental data revealed good (although not excellent) classification rates for all participants. When I performed a cross-validation on the re-training and the experiment data in a common data set, the results were once again basically chance level. A cross-validation on the re-training data alone again reveal solid classification rates. The data from the two recordings seems to be incompatible as far as the classifier is concerned.

A second side remark might be, that all participant pairs met for the first time on the day of the experiment with one exception. One pair of participants had known each other for quite some time before the experiment. It is interesting to note that the one participant who retained control during the experiment is exactly from that pair. Bonnet et al. (2013) already formulated the hypothesis that participant familiarity might have an effect in hyper-scanning experiments. It would be worthwhile to explore this aspect further. Furthermore the second participant from that pair reported afterwards that he suffered from a headache on the day of the experiment. Hence, I cannot rule out the possibility that he would also have retained control, if he had not suffered from that headache.

6.5.2. Phase Locking Analysis

The analysis for phase locking was the first venture into the field of signal interactions undertaken during the PhD project. It is a comparatively simple approach which still yields meaningful results and is widely used in the literature, e.g. by Dumas et al. (2010); Naeem et al. (2012); Yun et al. (2012); Supp et al. (2007). It is an estimate of the signals phase and should not suffer from correlations in amplitude as for instance Coherence does. It is, however, a symmetric estimate and as such not suited to give any information about the interaction's direction, i.e. effective connectivity.

Before applying the PLV to the data, it has been preprocessed in three steps: First, data has been segmented into epochs of six seconds length, centred at the countdowns expiration (i.e. the start of the MI of both participants). Second, (ocular) artefacts have been removed using an independent component analysis (ICA, see section 5.3 on page 42). Third, all segment have been inspected visually and segments with heavy residual contamination have been removed.

Then the PLV has been computed on two different frequency bands: θ-band $(4-7Hz)$ and μ-band $(9-13Hz)$ and I only considered hyper-connections between corresponding electrodes (e.g. C4 from participant

one with C4 from participant two). As this study is intended to evaluate whether hyper-connectivity still occurs in machine-mediated setting (the feasibility of the machine-mediated interaction approach) I neglected within participant connections for this analysis.

Data has been filtered to the according band. Then a Wavelet Transformation using a complex Morlet wavelet has been applied to obtain a time-frequency representation of the signal. Afterwards the phase locking value has been computed as described by Lachaux et al. (1999) (see section 5.6.2 on page 50).

Even for non-correlated data, some PLV larger than 0 is to be expected and the magnitude at which values become significant is not a-priori clear. Hence, PLV per se can only be used as a relative estimate. If we observe an increase or decrease of PLV over time this can be meaningful. But even then, such an assertion lacks any statistical undergirding

Lachaux et al. (1999) et al propose a method for statistical testing: PLS (see section 5.6.2 on page 50). PLS returns values which can be regarded as an estimate for the probability that the 0-hypothesis is true (p-value).

I perform a large number of statistical tests. The six seconds of a segment contain 256 samples each. I treat 16 channel pairs, I have four pairs of participants and I have two frequency bands. Hence, I have $6 \cdot 256 \cdot 16 \cdot 4 \cdot 2 = 196,608$ statistical tests. Assuming the canonical threshold of 0.05 for significance, we could expect $9,830$ "discoveries" even if the 0-hypothesis was true for every test. However, these false discoveries would be expected to occur randomly in the series of PLS values. For any larger time period for which the PLS values consistently undercut the 0.05 threshold the probability for type I errors steeply declines. Based on that assumption I applied a ten sample temporal smoothing on the series of PLS vales. Any insulated significant PLS (the form false discoveries can be expected to take) would average out and only those portions of the PLS which were significant with some temporal consistency would remain.

Next I want to define what constitutes a *stable* phase locking for me: If the PLS indicates a significant phase locking (after temporal smoothing) for a longer time frame, I refer to this as stable phase locking. However, I deliberately left open what is a "larger time frame", for the moment.

It is in the nature of signals their features tend to change fast for higher frequency bands. Hence, for the term stable phase locking to be reasonably applicable for arbitrary frequency bands, the length of the time frame will have to depend on the frequency band considered. I propose the following definition: Let $\omega = [\omega_{low}, \omega_{high}]$ be the frequency band for which the $PLS(\omega)$ time series has been computed. Any time interval $t = [t_1, t_2]$ with length

$$t_2 - t_1 \geq \frac{1}{2} \cdot \frac{1}{\left(\frac{\omega_{low}+\omega_{high}}{2}\right)} = \frac{1}{\omega_{low} + \omega_{high}} \tag{6.1}$$

for which $PLS_{t_i}(\omega) \leq 0.05, \forall t_i \in t$ this would be considered stable phase locking. Put differently, we compute half the period length of the centre frequency for the frequency band ω. Any phase locking that persists for longer than that is considered stable phase locking. For instance the θ band $(4-7Hz)$ has a centre frequency of $5.5Hz$. Thus, any phase locking which reaches significance for longer than $0.5 \cdot \frac{1}{5.5Hz} \approx 90.9ms$ would be considered stable phase locking.

Finally, I aimed to evaluate the phase locking divided by conditions. The cues given to the participants via headphones were designed to form three different conditions: cue congruent, cue incongruent and solo (see section 6.4.1 on page 66). However, given that a.) each participant knew only his/her own cue, b.) the only possibility to infer the partner's cue was by observing his/her TAO's behaviour and c.) there was no relation between the cue given and the TAO's behaviour (the participants had no control over their respective TAO's) cue congruent and cue incongruent conditions have lost relevance and grouping the data accordingly for analysis is pointless.

I decided to group the data based on the actual TAO behaviour. During one trial the accumulated data had been re-classified 12 times resulting in 11 command given to the TAO in intervals of $0.5sec.$ I wanted to define conditions based on a "majority vote" of the commands. If a TAO would have, for instance, received four commands to drive to the left, then one command to stop (inconclusive classification) and six commands to drive to the right I would label this trial as "movement right" for that TAO/participant. Assuming an ideal driving behaviour of the TAOs, this TAO would end up on the corresponding, in my example the right, half of the T-Desk. Based on that I defined a data grouping, analogously to the cue-conditions defined on page 66:

movement congruent: If the trial was labelled differently for each TAO/participant (one "movement left" and one "movement right") I assigned that trail to the movement congruent data group, as both TAOs would be assumed to end their movement on the same half of the T-Desk.

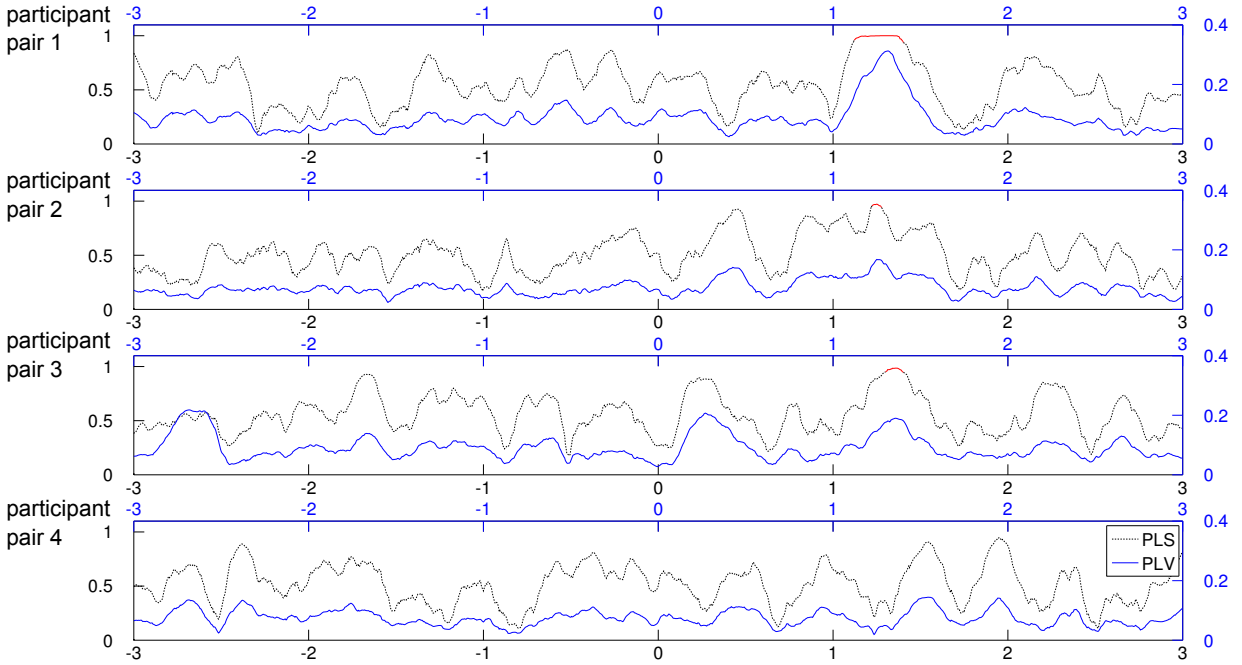


Figure 6.14.: PLV is depicted in blue and PLS in black with red parts for significant phase locking. An increase of PLV which appears to be time-locked to the onset of the TAO movement can be observed for all four pairs. For three of the pairs this reaches significant. PLS is expressed here as the rate of surrogate PLV being smaller than the original PLV (approximating the probability that the 0-hypothesis is false).

movement incongruent: If the trial was labelled the same for both TAOs/participants (either both "movement left" or both "movement right") I assigned that trial to the movement incongruent data group, because the TAOs would end their movement on different halves of the T-Desk.

solo: I keep the definition of solo as it was defined above: All trials for which one of the participants received a "observe" cue are assigned to the solo condition. As the control was switched off for the corresponding TAO, that TAO did not move and hence this definition is still grounded in the TAO's behaviour.

In theory it could happen that there was no majority in the 11 commands given (e.g. five commands "left", one command "stop" and five commands "right"). In that case the trial could not unambiguously be labelled neither "movement left" nor "movement right" making this definition of conditions incomplete. In practice this case never occurred in any of my data sets.

I applied PLV as described above to a.) the entire data sets b.) the data groups defined above individually. I found that for one pair of participants 95 out of 100 trials which were not movement solo, were movement congruent. Computing PLV/PLS on only five trials (the movement incongruent trials) would be infeasible. Hence I excluded this pair from the analysis divided by movement congruent/incongruent condition.

Each data set contained at least 145 epochs (I excluded some epochs due to heavy contamination). Computing PLV and PLS on these complete data sets, I identified a series of statistically significant phase locking. The spatial and temporal distribution of the phase locking were diverse between participant pairs. However, some observations can still be made:

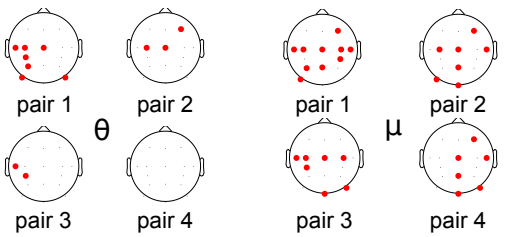


Figure 6.13.: The channels with stable phase locking in μ and θ band for each of the four participant pairs.

Figure 6.13 shows the channels for which stable phase locking could be shown at least once during the six seconds. Generally μ band showed stable phase locking more often than θ band. Stable phase locking seems

to appear less frequently in frontal regions. Apart from that, no clear spatial pattern can be discerned in the occurrence of stable phase locking.

When looking at the temporal occurrence of phase locking, one can note that it appears mostly after the countdown's expiration (i.e. during the phase when the participants performed motor imagery). Any phase locking before the countdown's expiration is spurious and insulated. But most remarkably I found a significant phase locking which occurred consistently for all but one pair[9]. That phase locking develops around electrode positions C3,C5,CP3 (depending on participant pair) about $1.2sec.$ after the countdown's expiration in θ-band. This timing is remarkable especially because $1sec.$ after the countdown's expiration the TAOs start moving. Figure 6.14 depicts this phase locking for the four participant pairs. A similar peak could be identified at the corresponding electrodes of the right hemisphere (C4,C6,CP4), but was less pronounced and did not reach significance.

Inspecting the phase locking by data groups I found that this phase locking is even more pronounced in either movement congruent or movement incongruent data group while it is greatly diminished in the other, for all three pairs (one was excluded from this analysis, see above). The phase locking found when analysing all epochs in common needs to be attributed to one of the two data groups. Strangely, to which of the two groups it needs to be attributed seems to depend on the pair of participants.

For solo trials virtually no phase locking could be observed in θ-band while in μ-band quite some phase locking could be shown. As the μ rhythm has been associated with mirror neuron function, the hyper-connectivity present in solo condition might be an indication for co-representation of the actor's actions by the observer.

6.6. Discussion

The loss of control for all participants was most disappointing after much work has been invested into the acquisition and training of the four pairs of participants. I have a reasonable explanation for this: The presence of a (foreign) partner in plain sight imposed a distraction during a mentally demanding task. Although this hypothesis cannot be proven scientifically on the present data, it is still reinforced by the findings in that data. To scientifically prove this hypothesis a whole new experiment would have to be designed, participants would have to be acquired and trained and the experiment would have to be conducted. But even then, proving that a partner imposes a distraction which diminishes the ERD would be a somewhat trivial finding, not worth the effort which would have to be invested.

It is, however, most remarkable that still hyper-connectivity could be demonstrated. Participants, of course, only knew about their own poor TAO responsiveness. They had no way learning that their partner suffered from similar problems. Hence they apparently still paid close attention to their partner's TAO's actions. This is stressed by prominent phase locking which appears to be time-locked to the start of the TAO movement. Interesting enough this effect is larger, when the participants where both active and diminished in the solo condition. Joint action seems to be important for this type of connectivity.

Furthermore, from the experience made during the experiment, the hypothesis has been derived that familiarity among partners might play a role in this type of settings. If this effect is limited to the impact of distraction by the partner, this would again be a mostly trivial finding. Bonnet et al. (2013) in their hyper-scanning experiment had a large ratio of partners which knew each other in advance and the authors also raised the questions whether this might have an impact. If this could be proven in future studies, it would be a most interesting result which would have implications for all further research and our understanding of human interaction. Investigating this further, maybe in an experimental design that allows to correlate neural and behavioural data, might yield some results highly relevant for research about human-human cooperation.

Finally, the most central research question of this study could be answered positively: Hyper-connectivity, previously shown in direct human-human interaction, still occurs when that interaction is machine-mediated. This is the prerequisite for conducting the more involved machine-mediated human-human interaction study described next.

[9] And even for that pair an increase in PLV can be observed, although it does not reach significance.

7. Experiment on Machine-Mediated Cooperation

I could show that hyper-connectivity, previously described during direct interaction between humans, also occurs when interaction is machine-mediated. I thereby demonstrated that studying the neural foundations of human interaction in such settings is generally feasible. In a second step, I now want to show-case the scientific potential of this approach by tackling a specific research question, which I will formulate and discuss in section 7.1, below. I occasionally hinted that this second experiment will be more complex and elaborate than the first experiment of this thesis. Therefore, this chapter will describe the experiment design and the BMI system developed, while the data analyses and their results will be the topic of the next chapter on page 93.

In section 7.2 (of this chapter) on page 80 I will formulate some rather general constraints regarding the experiment. I will then describe how a participant can issue a command via BMI in this experiment in section 7.3 on page 83. On the system side, the control flow for the same action is less trivial, particularly because control has to flow back and forth between the BMI and the robotic system seamlessly. I will describe this control flow and its organisation in section 7.4 on page 83.

After defining the goals I want to achieve, in the development of the system, I will then describe the concrete implementation in section 7.5 on page page 85. In this process some minor adaptations have been made to the existing robotic system to achieve optimal performance for the planned experiment. These will be described in section 7.6 on page 91.

The resulting BMI/robotic system is pretty versatile and might be of use beyond this experiment and this thesis. I will, therefore, discuss limitations and chances of the implemented BMI solution and offer some ideas for future advancements and research with that BMI system in section 7.7 on page 92.

7.1. Research Questions and Experimental Ideas

It can be assumed that a variety of factors influence the emergence and the intensity of neural hyper-connectivity. It has been shown by various studies that hyper-connectivity emerges during human interaction (compare chapter 3 on page 23). The first experiment of this thesis demonstrated that hyper-connectivity (still) occurs in machine-mediated settings. I already stated that I consider cooperation the scientifically most interesting form of interaction. In this experiment I will compare connectivity during tasks which require cooperation and tasks which allow for independent actions.[1]

Here I am aiming for a triadic human-machine-human setting for which interaction takes place in a shared space. More specific I want the machine (robot) to be the only channel of communication available to the participants (compare section 1.4 on page 3).

The specific research question this study aims to answer is:

What differences exist in hyper-connectivity between cooperation and independent pursuit of goals, when all interaction is mediated by a machine?

This calls for an experimental setting in which participants need to repeat a similar task a series of times, some iterations of that task inducing cooperation and others omitting it. The completion of each task should be an interaction in its own right, i.e. it should consist of several actions from all participants.

Furthermore, I want the cooperative tasks and the non-cooperative tasks to be as similar as possible in all aspects, expect the degree of required cooperation. And I shuffled the order of tasks for each experiment to control for any effects that might emerge from the task order chosen.

[1] Since the partners can see each others actions, they still influence one another. This would, by the definition in section 1.1 on page 1, still constitute an interaction.

7.2. The Experimental Design

From the above discussion stem two consequences for the experimental design: i.) The only communication between participants should be the robot executing the participants' actions and thereby mediating their mutual intentions and ii.) I need tasks which can be solved independently and tasks which require cooperation.

7.2.1. Meeting the Prerequisites for Hyper-Connectivity

The first of these requirements has some implications regarding the emergence of hyper-connectivity and our ability to computationally detect it, which I will discuss in this section.

The conditions under which hyper-connectivity occurs are, to date, not yet clearly characterised. However, some basic necessities can be deduced without which it is outright unreasonable to expect hyper-connectivity to emerge. The most basic of these necessities I would refer to as a "common reference" for the participants.

Basically hyper-connectivity is an alignment of phases between signals recorded from different brains. This does not emerge out of nothing.[2] Participants need some base of information about their partner's state and actions and the state of the over-all interaction. This information needs to be non-static or put differently it needs to be continuously updated. This is what I would call a "common reference". In normal every-day interaction the common reference is formed implicitly and it is pretty rich (including a shared space in which a task is solved, gestures, facial expression, speech, auditory cues about the partner's action and state and so forth).

The common reference can be seen as a channel over which information is exchanged between partners, volitionally and involitionally I already stated that I want to limit and control this information exchange. Only the mutual actions should be transmitted, mediated by the robot, from which the partner's intentions might be deduced. The best way to cut off all other levels of communication is to spatially separated participants from one another, i.e. to design a remote interaction experiment. If the participants are situated in different rooms three options exist regarding the location of the shared space in which the interaction takes place/in which the robot executes the participants' actions: It can be situated in either participant's room or in a third room. As I aimed for a symmetric experiment, i.e. both participants should have the same interface and information at their disposal as well as similar sensory input, I chose the latter option. As a result the participants teleoperated the robot.

The richest, most current and expressive common reference I could think of under the given constraints, was to project a live video stream of the shared space on two monitors. However, transferring this stream over IP network – kind of the standard solution – would induce an unpredictable and almost unmeasurable jitter between the two video streams.

The common reference as it is defined here, therefore, has a property which is for a natural common reference never actually relevant: synchrony. Could a low synchrony of the common reference prohibit the emergence of hyper-connectivity? I will discuss this issue using an example:

Provided we would present the same video stream of the shared space to two interacting, remote participants as the common reference in a hyper-scanning experiment similar to this experiment. But for participant two this stream would have a defined, fixed delay of exactly $300ms$ as compared to participant one. The participants have no means to realise this (neither on a conscious nor on a sub-conscious level). Neural hyper-connectivity might, thus, still form, but we would have to regard to the same delay in our analysis to detect it. Having a non-constant and (even more important) unknown jitter, as it is the case for a video stream transmitted over a TCP network, would make it impossible to us to computationally detect hyper-connectivity (although it would presumptively still emerge).

To provide a common reference and to guarantee a degree of synchrony way beyond the time scale of relevant neural activity, I connected two monitors with the same GPU (of the same computer). I used an extra-long DVI-cable and placed one of the monitors in the next room. Then I set the graphics card to clone the image on both screen. This way the jitter between the images should be bound by the frame rate of the two screens and, even more importantly, it should be mostly constant, at least much more constant than IP network delays are. Hence, at a frame rate of $120Hz$ the jitter should never be more than $1/120sec. \approx 8.3msec.$ As I can only observe neural processes up to a frequency of $\sim 40Hz$ (corresponding to a period of $25msec.$) this degree of synchrony should be sufficient. I decided that this synchronised video stream would be the common reference provided to the participants.

[2] As a fan of science fiction literature, the idea of telepathy intrigues me, but as a scientist I work on the basis of the assumption that telepathy does not happen between humans.

7.2.2. Choice of Task

The second requirement was to find tasks which induce a need for cooperation and other tasks which make cooperation pointless. These two types of tasks should in each other regard be as similar as possible. Furthermore, the participants should solve tasks in a shared space, needed to perform several actions (a whole interaction) to solve a task and interaction should be mediated by the robot. Additionally, as the BMI teleoperation of a robot will be perceived as complex and alien by most participants, it would be desirable if the tasks' domain would be familiar to participants from everyday life.

A variety of robots were available in CITEC to implement the machine-mediation, each with its own capabilities. I already discussed the impact of the choice of the robot to use in section 6.2.1 on page 62. In the first study (section 6 on page 61) I chose a robot with a plain shape which impose little distraction and, more importantly, was not anthropomorphic in any way. In this experiment I decided I wanted to go the other way, namely to employ a robot that is very human-like and allows for an identification of the participants with the machine. Among the robotic systems available at CITEC is one system whose core components are the so called Shadow Hands, one of the most human-like robotic pair of hands on the market. The main capability of that robotic system, is to perform pick-and-place tasks with high versatility regarding object shapes and grasp types.

Object manipulation (i.e. moving, handling and arranging objects) and, in particular, pick-and-place actions are very common in everyday life and highly familiar to humans. Actually object manipulation is a core ability of humans. Object manipulation could easily take place in a shared space, designing task such that they require several actions and that some require cooperation and other do not, should not be too hard, neither. Realising that I had a robotic system at my disposal whose core ability nicely fits the requirements for the task's domain formulated earlier, the questions which robotic system and which task domain to use, was settled.

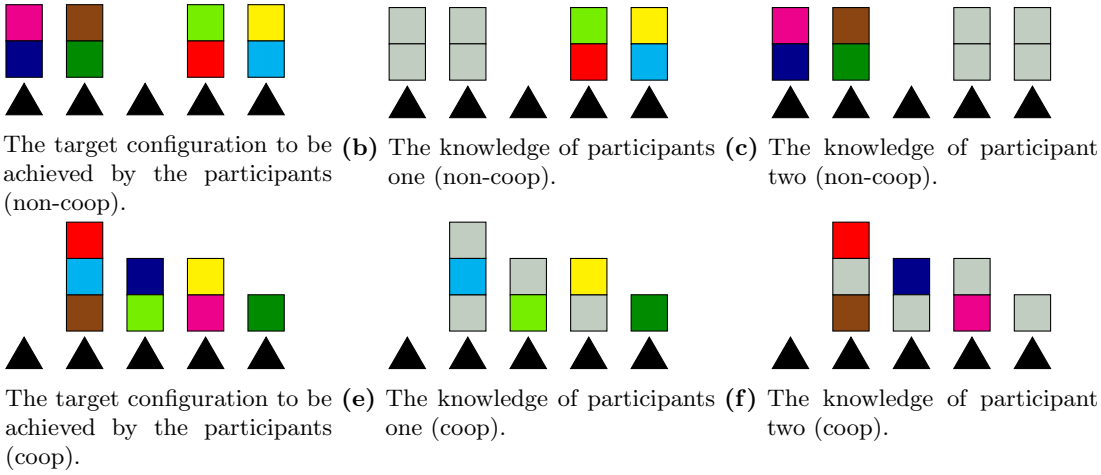
I will describe the robotic system in section 7.5.1 on page 85. Operating these hands using a BMI, seeing the action carried out by the robot from an I-perspective in the video stream and seeing how the partner operates the same set of hands might actually induce a feeling of sharing the same body. This is intentional and a clear contrast to the first experiment.

Pick-and-place tasks are probably the most common object manipulation tasks and everyone experiences/executes countless of these tasks each day. It is generally consent in the BMI community that robot control via (EEG-based) BMI is best organised when participants give high-level commands and let the robot/machine execute these commands with a high degree of autonomy (Finke et al., 2012; Bell et al., 2008; Lotte et al., 2010). The description of a pick-and-place task in a series of high-level commands comes pretty naturally: Selecting the object to grasp and selecting the position to place the object. Low-level decisions such as the approach-vector for the pick- and the place-action, which fingers to use for the grasp, grasping force and much more shall and can be handled by the robotic system autonomously.

Finally, I decided to have several instances of a single, abstract object type, in particular differently coloured cubes. These cubes were to be placed on five predefined target positions. Each task demanded for another configuration of cubes.

The coloured cubes resemble common building blocks for children, hence they are not strange or alien objects to participants and suggest a game-like character of the experiment, potentially motivating the participants. Furthermore, the cubes allowed to include the stacking of objects (cubes) in the task design. This way I was be able to define tasks which enforce cooperation between participants and others which need no cooperation, easily:

This was achieved by providing only partial information to each of the participants. In the target cube-configurations/tasks-description handed to the participants, half of the cubes were greyed out such that the participants had partial, complementary knowledge of the target configuration. Tasks which required cooperation were designed such that each participant had to place some of the cubes he/she had knowledge about ("was responsible for") on top of cubes his/her partner had knowledge about. This way the partners were dependent on one another and needed to cooperate to fulfil the task. For cooperation-free tasks each participant had to build his/her own structure, without dependence on the partner. Two examples (one cooperative and one non-cooperative) for the task descriptions handed to the participants are given in figure 7.1. The complete list of target configurations is given in appendix B on page 125. Furthermore, I decided that participants should not correct mistakes they made eventually and, consequentially, excluded cubes which had already been placed from the stimulus presentation (for the duration of that task). I intended to let the participants deal with situations in which a cube was placed in error and, thus, the given target configuration could no longer be achieved, as they saw fit.



(a) The target configuration to be achieved by the participants (non-coop).

(b) The knowledge of participants one (non-coop).

(c) The knowledge of participant two (non-coop).

(d) The target configuration to be achieved by the participants (coop).

(e) The knowledge of participants one (coop).

(f) The knowledge of participant two (coop).

Figure 7.1.: The approach to control the degree of cooperation between participants. A target configuration is defined, but both participants only have partial, complementary knowledge about the target configuration. The upper row shows a target configuration with no need for cooperation and the lower row a configuration with a high need for cooperation.

In summary, the tasks given to the participants were configurations of (up to) eight differently coloured cubes placed on five target positions, including cube stacking. Each participant was able to pick arbitrary cubes and place them on arbitrary target positions using the P300-based BMI. A need for cooperation was induced by stacking cubes which different participants had knowledge about. The experiment was named *Interactive Cube-Stacking in Shared Space (iCusss)*.

7.2.3. Sharing the Robot

The robotic arms cannot, currently, avoid obstacles, neither objects/cubes nor the other arm. Thus, one cannot use both arms at the same time safely. Hence, the robot can only execute one pick-and-place task at a time which makes it an exclusive, shared resource which needs to be managed. Several options exist:

1. I could enforce turn taking in the use of the robot. The robot would execute a command from the first participant and then only accept a command from the second participant before accepting any new commands from the first participant, and so forth.

2. When the robot is busy executing one participant's command and a new command arrives (from either participant) that task is placed in a queue and the tasks in that queue are processes in a first-come-first-server order.

3. Commands are accepted (and queued) in an arbitrary order, but a new command from any participant is only accepted if the previous command from that participant has been finished.

The first option would be easy to implement and intuitive, however, I found it rather rigid and inflexible, opposing the idea of a vivid interaction.

The second option would be most flexible. The downside would be that, if participants are significantly faster in issuing commands using the BMI than the robot is in executing them, the participants would produce a backlog of commands waiting to be processed. This would decouple the commands from the real world execution. In other words, there would still be a shared spaces technically, but the queue with the commands would be what is relevant to the participants, not the shared space.

I ultimately decided to implement the third option in that way, that after a participant had issued a command, the experiment for that participant paused until the command given had actually been executed by the robot. Consequentially, if a command is completed at a time at which the robot is still busy executing another command, that new command would be delayed until the robot is available again. Of course, when a command is delayed, a display message would inform the participant.

In an ideal case, in which participants would always succeed in issuing commands in a similar (small) number of stimulus presentation repetitions, this mode of execution would result in the participants taking

turns in giving commands. However, it is generally not the case, that both participants have an equally high level of control over the system. Given that participants might take significantly different time for issuing a command with the BMI (see section 7.5.2 on page 86) it can be expected that one participant would finish two commands in the time in which the other participant finishes one, occasionally. This is the most deregulated management of the shared resource, that is the robot, possible, given its exclusiveness.

After the experiments I learned from participants' reports that they occasionally tried to not issue any command while waiting for their partner to place a certain cube. Here participants volitionally deviated from turn-taking in the robot's usage. This was done either by not paying attention to the stimulus presentation or by selecting a so called dummy object volitionally (see, section 7.5.3 on page 88).

I belief that brief pauses between the different actions are desirable for my study. Not only have we made the experience that such pauses are beneficial to the degree of control the participants exhibit over the system. While making decisions using the BMI, the participants need to direct their visual attention towards the stimulus presentation. The synchronised video stream serving as the common reference for the participants would still be visible in the peripheral field of view, but could, during the flash-sequences in the stimulus presentation not be in the centre of attention. Pauses between commands would allow participants to recentre their attention towards the common reference.

7.3. Control Flow – Participant

Before closing in onto the technical detail of the implementation of the robotic system, I want to describe a pick-and-place action performed with this system from two sides: First, in this section, I will describe the participants view. Then, in the next section on page 83, I will discuss the same action from a system's perspective.

From a participant's point of view, two decisions had to be made using the BMI: i.) which cube to pick ii.) at which of the target positions to place the cube. I decided that only after the second decision had been met, the robot would actually start executing the task (rather than already picking the selected cube and only then to poll the participant regarding the target position). During the execution the system would evaluate whether the target positions was occupied and, if necessary, stack the cubes autonomously.

The two selections to be made (cube and target positions) were both "one out of many" decisions. Regarding the taxonomy of BMIs discussed in section 2.2.2 on page 9 this BMI would, hence, be an active, dependent Choice BMI. For decisions with more than two options, (visually) evoked potentials are especially well suited, because each option can be associated with one stimulus and the participant can just concentrate on the stimulus that represents the desired option.

In summary the participants will perceive the BMI control having three steps:

1. All cubes available for grasping, i.e. all cubes which have not yet been moved during this task, are involved into a P300 stimulus presentation. The participant selects the cube to pick by concentrating on it.

2. Markers (triangles) indicating the predefined target positions are involved in a P300 stimulus presentation. The participant selects the position onto which he/she wants to place the cube previously selected, by concentrating on the corresponding marker.

   (a) If the robot is still executing the latest task of the partner, the experiment is paused for the participant and he/she can watch the execution of the partner's task.

3. Subsequently the robot will eventually start the execution. It will pick the cube, determine at which altitude to approach the target position (stacking? at which level?) and then place the cube. During this phase the participant is idle and can watch the execution.

This control flow is actually pretty comprehensive and clear. From the system's side it is, however, somewhat more involved.

7.4. Control Flow – Machine

Thinking the same action from the system's perspective several additional aspects need to be considered/occurred during the development:

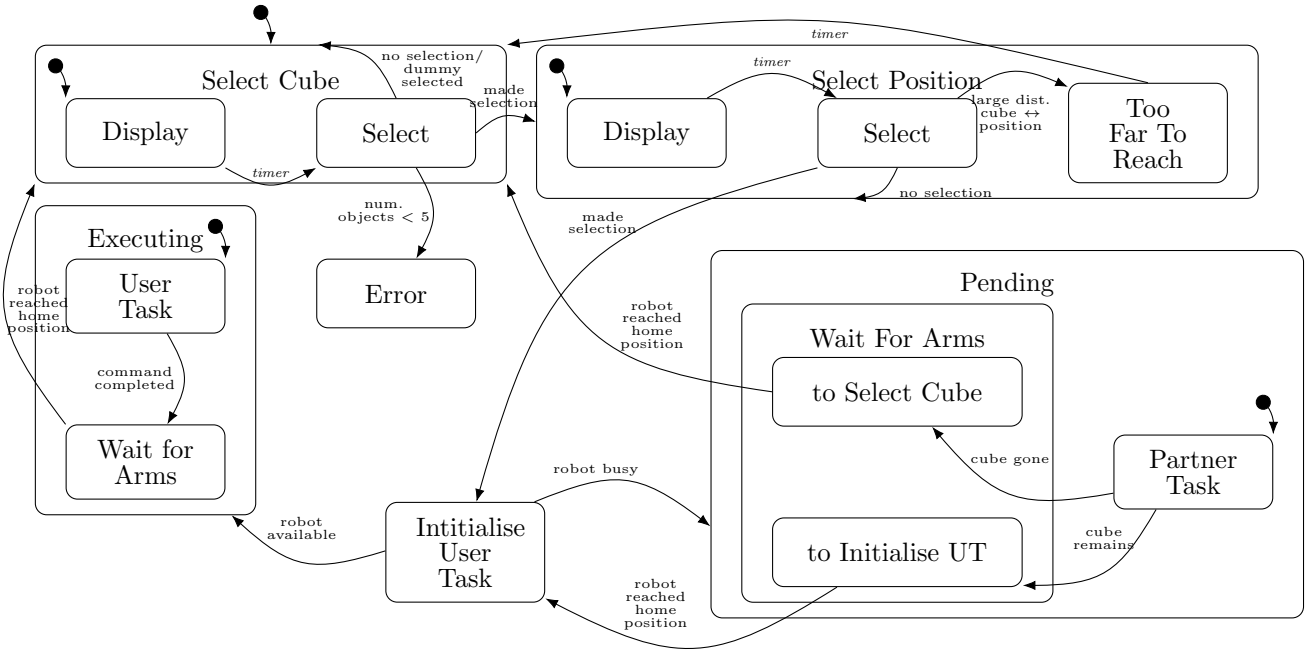


Figure 7.2.: The BMI Control HSM. For the hyper-scanning experiment two of these HSM were instantiated in an encompassing concurrent state.

- Before each of the two choices (cube and position) a text message should be displayed to the user, notifying him/her which decision he/she is supposed to make. There should be a brief pause (three seconds) to given the participant sufficient time to read the message.
- When the BMI classification was inconclusive after ten repetitions of the stimulus presentation the BMI would cancel this decision making and the HSM should react by starting over with the same decision (either for a cube or a target position).
- The same should happen when a dummy object was selected for grasping. I will describe dummy objects and their use in section 7.5.3 on page 88.
- Four dummy objects existed in my experiment. Hence, when only four objects remained for grasping, these four objects had to be dummy objects and, thus, the task was completed. Consequentially, participants should no longer be polled for further decision.
- There are certain movements the robot is physically not capable to perform. In particular, it cannot pick a cube from the one end of the desk and place it on the far end of the desk.
- When a command given by one participant is delayed because the robot is still busy with a command from the partner, both commands (the active and the pending command) might have targeted that same cube. But each cube should only be picked and placed once.

In section 6.4.4 on page 72 I already expressed my preference for HSMs to control complex robotic settings and experiments. Strictly speaking the description of the HSM is already part of the implementation and should be located in the next section. However, it is also the best way to describe the control flow from the machine perspective, which is why I placed it here.

The robotic system here comes with a software which reads HSMs described in a XML-format from a file and executes them. Such a HSM can react to incoming messages, send messages and execute snippets of Python code. One such HSM already managed the control flow of pick-and-place tasks in the pre-existing robotic system.

The HSM I developed for the system is depicted in figure 7.2. This HSM would be duplicated during a hyper-scanning experiment: One instance for each participant. I will now describe how the HSM, and therefore the control flow, was organised by going through the states of the HSM:

Select Cube: In this state a message was displayed to the participant that next a cube is to be selected. After three seconds that message disappears and the stimulus presentation starts highlighting objects in random order. If the classification could not identify the users intention after ten repetitions of the stimulus presentation or if the selected object was labelled as "dummy" by the vision system the decision was repeated. Otherwise I went on with the selection of a target position. If there are less than five objects left, the system enters a terminal state. It can only leave this with operator interaction. It is assumed that the four remaining objects are dummy objects and that the task has, thus, been completed.[3]

Select Position: Again a message was displayed and after three seconds the triangles indicating the potential target positions were highlighted. If the resulting pick-and-place command could not be executed by the robot, the system displayed an error message and initiated a whole new command. After ten repetitions without a valid classification the decision for a target position was started over.

Initialise User Task: This state was practically a transient state. In this state a command was send to the robotic system including the results from the two preceding BMI decisions. The robotic system responded to this directly either by accepting or rejecting the command. The system actually only rejected a command when it was busy (executing a command given by the other participant). The BMI Control HSM responded to this by switching to the Pending state. If the command was accepted the HSM switched to Executing state. What ever the case, the Initialise User Task state was left almost instantly.

Executing: While in this state, the robot executed the task given. When it was completed, the arms needed to go to their home position before I could start a stimulus presentation for the next decision. If I had started it right away, the arms would still have been in the field of view of the 3D camera and the arms would have been segmented into various meaningless objects which would have been included into the stimulus presentation. Furthermore, cubes might have been occluded by the arms and would therefore not have been included into the stimulus presentation.

Pending: This state was entered when at the time the Initialise User Task State had been entered, the robot was busy. The HSM waited until the robot had finished the task it was busy with. Then it evaluated whether the task just finished, targeted the same cube as the task that was still pending. If that was the case, the pending task was discarded and a new decision was triggered. Otherwise the task was re-initialised. In both cases I needed to wait for the arms to reach home position. The pending state basically implements the required synchronisation between the two users.

The HSM allows me to unambiguously model the control flow for the experiment in a clear and comprehensive way, while addressing all issues listed at the beginning of this section.

7.5. Implementation

Now that I have outlined what experiment I want to conduct and which capabilities the BMI/robotic system needs, I am going to describe how this was implemented. I will start by describing the original state of the robotic system in the section below. Next I will describe the different software processes involved in the final system and how they communicated in section 7.5.2 on page 86.

Then I will start to describe the BMI control loop, starting with the stimulus presentation in section 7.5.3 on page 88, briefly going over the P300 classification system in section 7.5.4 on page 90 and finally describing how the classification results were translated into robotic action in section 7.5.5 on page 91.

At this point I want to thank Guillaume Walck, André Ückermann, Christof Elbrechter, Martin Meier and Robert Haschke for their collaboration when interfacing the robotics system with the BMI system and execution of experiments.

7.5.1. The Shadow Hand System at AGNI

To get started I will briefly go over the original robotic system, as it was available at the neuroinformatics group, Bielefeld University, at the time the development started. Figure 7.3 shows a photograph of the system.

[3] The state was named Error for historic reasons and that name had not been changed when the semantic usage of the state changed during development of the HSM.

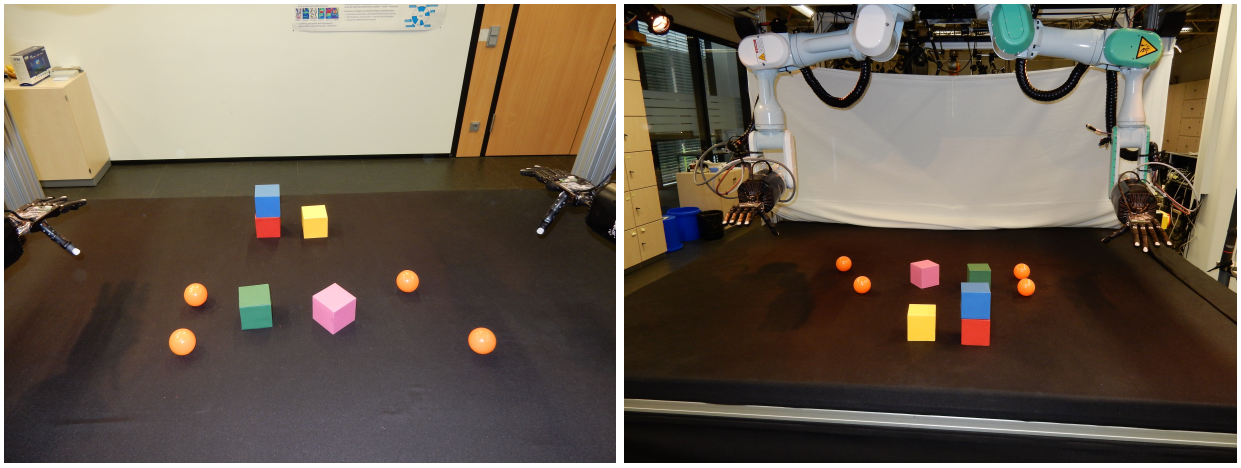


Figure 7.3.: The robotic setting with the two hands, the desk and the cubes used during the experiment. The cubes in the scene were used for the interaction, while the orange balls were the so called dummy objects, discussed in section 7.5.3 on page 88.

The system consists of two five-fingered anthropomorphic hands from Shadow Robotic Company, London. These are mounted to two PA-10 (Mitsubishi) industrial robot arms. The hands are mounted at head height above a desk which was mounted at hip height, allowing manipulation of objects placed on the desk. A KINECT RGB-D camera is mounted on top of the scenery, approximately at the position at which a head would be in relation to the arms and hands. Finally a set of speakers and a wireless microphone allow for verbal communication with the system.

The computer vision uses the information from the 3D point cloud (from the KINECT) to identify individual objects in the scene. When an object is to be grasped, the 3D point cloud segment of the object is fitted with a super-quadrics model and the robotic software plans the trajectory and the grasp according to the shape and a selection heuristic (Ückermann et al., 2014).

The colour information from a colour camera mounted beside the KINECT can be used to compare each object identified with a database of object-prototypes and thereby to label each object with a meaningful name, which can also be used in the verbal communication with the system.

The system is capable of understanding and executing verbal commands such as "Put one red apple into the basket!". If there are several such objects, e.g. several red apples, the system can ask questions for disambiguation.

There are certain limitations to the system, which shall be overcome in the future, some of which where important to keep in mind for the BMI (and for the experimental) design. I already mentioned that one cannot use both arms simultaneously. Another important fact is, that the vision does not recognise (and filter) the robotic hands. This is of particular importance for the stimulus presentation of the BMI system, as I do not want to include the robotic arms or parts of them into the stimulus presentation (compare section 7.5.3 on page 88). Another point is that there are limitations regarding how far each arm and hand can reach out and there is currently no way to perform a handover of an object from one hand to the other. This implies that it is impossible to pick an object from one side of the desk and place it on the other end of the desk. Another technical restriction was that no more than three cubes could be stacked on the central three target positions and no more than two cubes on the outer two target positions. Finally, the robot cannot align neighbouring cube as precisely as a human can and, therefore, the target positions had to be defined such that a small gap was left between any two piles of cubes.

Regardless of these limitations much great research is being conducted with this system at the neuroinformatics group, e.g. Elbrechter et al. (2011); Maycock et al. (2010); Twardon and Ritter (2015); Steffen et al. (2007, 2010), substantially advancing our understanding of manual object manipulation including handling of deformable objects, tactile sensing, slip detection, tactile surveying, folding of paper and much more.

7.5.2. Components of the System

Both, the pre-existing BMI system and the pre-existing robotic system require several processes which need to communicate. Interfacing the robotic system with the BMI involved two main challenges: Enforcing a

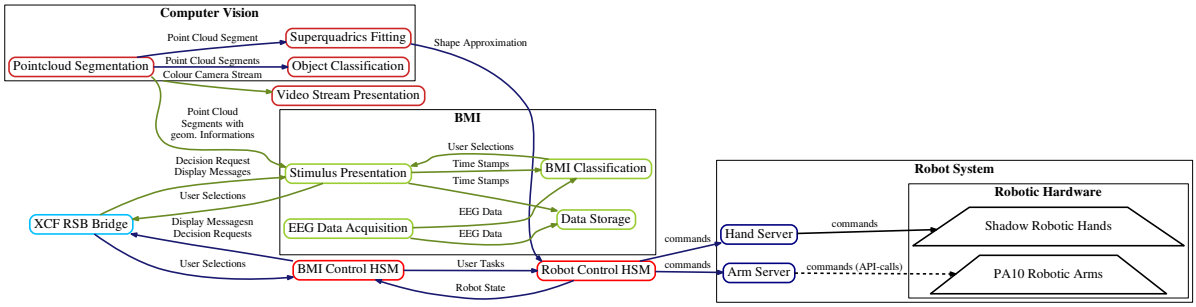


Figure 7.4.: The system architecture for the robot BMI control. UBiCI processes and RSB communications are depicted in olive green, robot control processes and ROS communications are depicted in black, HSM processes are depicted in light red, vision processes are depicted in dark red, XCF communications are depicted in dark blue. The XCF RSB bridge is coloured separately as it constitutes an interface component. Communication is encoded in edge-colour: dark-blue is XCF, olive is RSB and black is ROS.

common control flow, with no deadlocks and undefined states. The solution for this was, again, the usage of a HSM, as discussed in section 7.4 on page 83. Even more basal, communication between the processes of the different systems needed to be established.

The different processes of the final system and their communication is depicted in figure 7.4. Because the system is based on two different pre-existing systems, a variety of different technologies need to interact seamlessly. This includes three different network communication protocols/middlewares: RSB (Wienke and Wrede, 2011), XCF (Fritsch et al., 2005) and ROS (Quigley et al., 2009). The communication needs to be translated back and forth for the system to run. Some details about these middlewares are given in Appendix C on page 127

I will start with the components of the vision system. The components belonging to the vision-system are depicted in dark red in figure 7.4. It consists of three different main processes:

Point Cloud Segmentation: This program receives a point cloud from the 3D camera and tries to segment this into a set of compound segments, representing objects, in a multi-stage process including detection of normals, edges, surfaces and their relations to one another. The system can reliably discriminate/segment objects of almost arbitrary shapes. The main limitation is the resolution of the 3D camera: When the objects are too tiny they can no longer be adequately perceived and discriminated (Ückermann et al., 2014).

Object Classification: Each segment identified by the segmentation is compared to a set of prototypes of known objects from a data base. This is used to assign a label to different point cloud segments/objects (such as “apple” or “basket”).

Superquadric Fitting: When one of the objects identified by the point cloud segmentation is to be grasped by a robotic hand, this is fitted with a super-quadric. This serves as a computationally well suited representation of the point cloud segment and also gives a reasonable hypothesis of the object’s shape on the camera-averted side.

The robotic system consists of two processes tasked with controlling hardware and a control process governing the overall systems behaviour.

Robot Control HSM: A HSM has been used for the over-all control flow of the robotic system, making the different processes of the robotic and the vision system interact, yielding the overall robotic system’s behaviour.

Hand Server: The hand server is one of the parts the HSM controls. It is basically an interface translating the rather high-level commands given by the Robot Control HSM to low-level commands used to control the hands.

Arm Server: The arm server does just about the same thing for the arms, that the hand server does for the hands.

The **BMI Control HSM** is the HSM discussed in section 7.4 on page 83. The rest of the BMI system consists of four different processes. These acquire, classify and store data and present stimuli:

EEG Data Acquisition: This process encapsulates the driver for the GUSBAmp EEG recoding device. It collects packages of 32 samples at a time, wraps them into an UBiCI EEGData object (see section 5.8.3 on page 57) and attaches a millisecond precision time stamp to them before sending them over a RSB connection to other processes.

Stimulus Presentation: This process receives a continuous stream of 3D point cloud objects from the computer vision. This allows it to present a more or less real-time 3D image of the scene[4] and, when needed, to freeze that image and perform a P300 stimulus presentation on it, using each identified object as one option the participant can choose from. It passes the timestamps at which these flashes occurred to the classification and to the data storage. It also receives from the BMI Control HSM the cues when to start the stimulus presentation and when to display one of the predefined messages to the participant. Finally it also receives the result of the classification process, displays this result to the participant (feedback) and then passes it on to the BMI Control HSM.

BMI Classification: This process receives the stream of EEG data recorded from the participant as well as the time stamps at which the stimulus presentation rendered flashes of objects or potential target positions as part of a P300 stimulus presentation. It classifies the brain responses to these stimuli, collects the classification results and finally tries to infer the users choice based on that data.

Data Storage: This process has the exclusive task to store the recorded EEG data and the time stamps from the stimulus presentation before any segmentation, filtering, feature extraction whatsoever is applied. The only treatment the data has undergone is the (hardware) filtering done within the EEG device: A $0.1Hz$ high-pass filter for de-trending and $48-52Hz$ notch filter to remove electrical noise from power lines and electrical devices.

The **XCF RSB Bridge** is really only that: A program which translates back and forth messages between XCF and RSB.

The distribution of tasks over different processes makes the system more robust, allows load distribution over different machines and allows for a remote experiment and teleoperation of the robot.

7.5.3. From Vision System to Stimulus Presentation

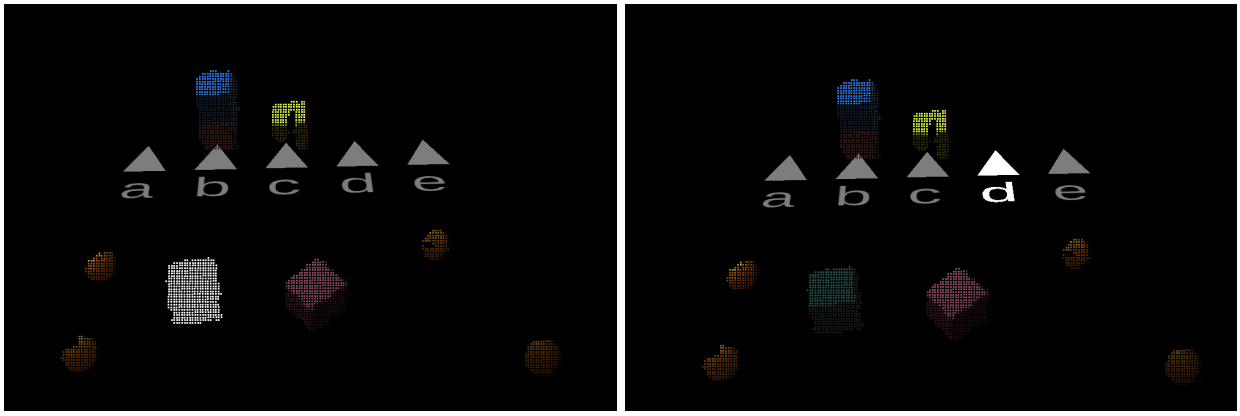
The stimulus presentation should be dynamic and intuitive. Dynamic in so far as it adapts the stimuli in number, shape, colour etc. to the options the current situation offers. And intuitive in so far as the meaning of the stimuli should be unambiguous.

As I stated, the robotic system includes a performant computer vision system. Particularly for the selection of cubes, using the results of the existing computer vision for the stimulus presentation offers many advantages. A stimulus presentation based on the computer vision segmentation would automatically be adaptive, i.e. offer only those cubes available on the desk. Sorting out those cubes which had already been place was as easy as defining a y-threshold in the 3D space. And it would be intuitive as well, because I could highlight the 3D representations of the actual cubes in their respective colours and at their respective positions within a 3D scene. Mapping between cubes in the stimulus presentation and in the live video stream should be highly intuitive for the participants.

The existing computer vision is based on the Image Component Library (ICL), which is being developed at Bielefeld University.[5] The ICL offers components to visualise the depth and RGB information from the KINECT and to modify this visualisation (e.g. to highlight certain parts of it). To implement this I needed three different information from the computer vision system:

1. The image of the 3D camera (as a 3D point cloud).
2. The results of the object segmentation. In particular a structure assigning a label to each point of the point cloud representing to which object/segment this was assigned.
3. A RGB-colour that corresponds to each point of the 3D point cloud.

[4] Not to be confused with the live video stream of the shared space used as a common reference.
[5] See iclcv.org



(a) The stimulus presentation with a cube highlighted. **(b)** The stimulus presentation with a target position highlighted.

Figure 7.5.: The iCusss stimulus presentation. Basically it is a 3D point cloud. The points are tinted in the colour of the corresponding pixel of the RBG camera. All parts of the point cloud not belonging to the objects recognised are discarded, particularly all points belonging to the table. Each of the objects can be highlighted during a P300 stimulus presentation. The same holds for each of the markers for target positions projected into the scene.

Then I visualised the 3D point cloud and tinted each 3D point in the colour of the corresponding pixel of the colour camera. The result would be a virtual 3D scene consisting of coloured points (see figure 7.5). I assumed that in that virtual 3D scene the objects would yet be well recognisable, particularly that it would be unambiguous to map between the task descriptions handed to the participants and the cubes on the screen. To verify this assumption, I prepared a scene with the eight coloured cubes, put a task description as I used them for the later experiment printed on paper beside the screen and asked ten colleges to match the cubes in the 3D scene with the cubes they could see in the task description. All ten did so without hesitation or any ambiguities.

Visual stimulus presentation for P300 usually highlights the different options for a short period (flashing). In preparation of a stimulus presentation using a given point cloud, versions of the point cloud where generated in which one of the segments/objects was coloured in white. To let one of the objects flash during the stimulus presentation, the software would first display the original 3D point cloud and then switch for very short time periods to the corresponding white-object version before going back to the original point cloud. The resulting P300 stimulus presentation will be highly dynamic without any hard-coded limitations regarding object number, shape, colour etc.

The P300 brain response is highly dependent on the so called oddball paradigm (see section 2.2.5 on page 17). If there are too few irrelevant stimuli the oddball paradigm is no longer fulfilled and the P300 is diminished and vanishes eventually. From our experience, having fewer than five stimuli/options will result in a significantly decreased classification accuracy. However, the stimulus presentation developed for this system does not ensure this constraint is satisfied. When only three different objects are present in the scene it will operate with these three objects. It is up the operator to ensure, that sufficient objects are in the scene at any time.

To help with that, the stimulus presentation knows one special class of objects: the “dummy” objects. In section 7.5.1 on page 85 I already mentioned that the vision system has a database of different objects it is familiar with. When one of the objects in the scene is recognised as belonging to a class labelled “dummy”, the P300 stimulus presentation will operate on all other objects in the scene for as long as there are at least five of those non-dummy objects. When there are less than five non-dummy objects in the scene the stimulus presentation will involve as many of the dummy objects into the stimulus presentation as necessary to have five different stimuli/options or all of them if there are not enough objects at all. Phrased differently, by placing four of the dummy object in the scene, the operator can ensure that there are always sufficient objects for a reliable P300 classification, even when there is only one non-dummy object (left). I later settled with small orange balls in contrast to the larger coloured cubes (none of which was orange), see figures 7.3 and 7.5. If these were selected during the experiment, the selection was repeated, thus, they were never grasped.

Having thoroughly discussed how the stimulus presentation allows for selection of objects, the selection of target positions is pretty straight forward. The target position had no physical representation on the desk, hence I used virtual triangles (2D objects projected into the 3D scene) pointing towards the far end of the desk as position markers. The robot aimed to centre each cube on that far vertex of the triangle. It is generally assumed that P300 classification is more stable when the (visual) stimuli have some distinguishing feature. For that reason, I enumerated the position markers using letters. Highlighting these position markers is as simple as switching the colour for a single object in the 3D scene from grey to bright white and back again.

For both stimulus presentations the selected option (object to pick/position for placement) was highlighted afterwards as a feedback.

7.5.4. P300 Classification

Next I want to take a closer look into the Classification Process. For those familiar with P300-based BMIs, this is pretty much a standard setup when classifying P300 data. Figure 7.6 shows the UBiCI deployment used for the P300 classification.

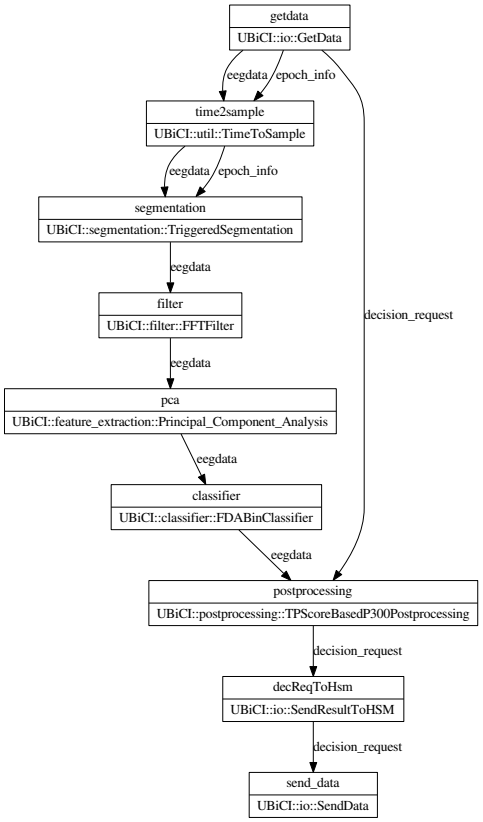


Figure 7.6.: The classification process. Data is segmented, frequency filtered, a feature extraction is applied, the data is classified as either P300 or non-P300 and, finally, results are collected over several repetitions to increase accuracy.

The EEG data, the decision requests and the time stamps/epoch infos (compare section 5.8 on page 55) are received over RSB and enter the classification process from the GetData component. They enter the TimeToSample component where timestamps from the stimulus presentation and EEG data are synchronised. Then they are send to the segmentation where most EEG data is discarded and only data epochs time-locked to the timestamps are passed on to the frequency filtering via STFT (FFT). This process is the same as for the previous experiment and is covered in some more detail in section 6.4.3 on page 68.

The data at this time has 16 channels and each segment has 256 samples. This data is then linearised resulting in vectors of length 4096. The number of dimensions is then reduced by the Principal Component Analysis (see section 5.2.1 on page 39) resulting in a vector of as few as hundred dimensions. The exact number of dimensions used is determined on a per participant basis at time of the computation of the classification matrices, compare section 5.2.1 on page 39.

This data is then classified into two classes, either as containing a P300 or not containing a P300 brain response. The EEG data is not only annotated with a classification label (0 or 1) but also with a confidence (which is basically the distance of the projected data point to the threshold, compare section 6.4.3 on page 68).

The classification results over the repetitions are then collected in the TPScoreBasedP300-Postprocessing component, which implements the approach to combine the classification results from different repetitions of the stimulus presentation to enhance classification accuracy, presented in section 5.2.4 on page 42. This postprocessing was configured to always trigger at least three repetitions of the stimulus presentation to have a reasonable basis for evaluation and not to be fooled by possible outliers. On the other hand the postprocessing would never have more than ten repetitions. If after that many repetitions no score has surpassed the other scores, it would presumably be difficult for any score to achieve that in the future. Therefore, the postprocessing would notify the stimulus presentation (and that in turn the BMI Control HSM and the participant) that the decision was inconclusive. The BMI Control HSM would

then start over the same decision from scratch. This postprocessing method has been evaluated in depth in Lenhardt et al. (2008). For my experiment it worked extremely well. Decisions were rendered in average after three to four repetitions (with three being the fixed min) and misclassifications were sparse.

The final two components notify the stimulus presentation and the BMI Control HSM about the (interim) results of the P300 classification.

7.5.5. From Classification Result to Robotic Action

Robotic commands can be transferred to the robot control software via the XML-based XCF protocol. The XCF message that needs to be sent for starting a robotic action is a `user_task`. This `user_task` contains information about the object to be grasped and the target position in XML-tags `object` and `target` which are children of the `user_task` tag.

Below an exemplary XCF message for starting a robotic action is printed. A more detailed description of its attributes and their meaning can be found in appendix D on page 129.

```
<user_task armPref="targetpos" ... type="put" ... xcfIP:serial="0" ...>
  <object cube_class="browncube" pointingProb="0" probability="0" shapePreference="b" timestamp="1429891262804974">
    <position x="299.6297302246094" y="429.5333251953125" z="66.08498382568359"/>
    <aabb xmax="354.6943359375" xmin="248.6115570068359" ymax="464.1672058105469" ymin="376.7919006347656"
          zmax="87.20902252197266" zmin="24.20656394958496"/>
    <size points="417" x="703.8733520507812" y="588.1016235351562" z="166.0480346679688"/>
  </object>
  <target pointingProb="0" position_num="4" probability="0" relation="at" timestamp="1429884978">
    <position x="280" y="150" z="29"/>
    <euler pitch="0" roll="0" yaw="0"/>
    <approach angle="-1.57" x="0" y="14" z="3"/>
    <aabb xmax="316.5" xmin="243.5" ymax="186.5" ymin="113.5" zmax="65.5" zmin="-7.5"/>
    <size x="73" y="73" z="73"/>
  </target>
  <STATUS origin="Handler" value="initiated"/>

</user_task>
```

For the predefined target positions the values are fixed. On system start one `target` XML-tag containing the information for each of the five target positions is pre-generated and during the experiment copied into the `user_task` tag as needed. The information about the objects are send by the computer vision alongside the point clouds to the stimulus presentation. This in turn sends them along-side the epoch info generated for each flash taking place during stimulus presentation. The `user_task` XCF message is generated and send over XML by the BMI Control HSM when it enters the Initialise User Task state.

7.6. Adaptations of the Robotic System

There are some adaptations to the robotics system we made in order to ensure optimal performance of the robotic system during the experiment. Some of those modifications would be undesirable for a robotics study as they induce domain knowledge not suitable for an objective evaluation of the robotic system. For my BMI study these changes can well be accepted, as the robotic system is not to be evaluated here.

Grasp Preference: Usually the robotic system decides autonomously which type of grasp (power grasp, two finger grasp, etc.) it uses for a given object. However, the validity of this choice very much depends on the validity of the super-quadrics model which in turn is sometimes hampered by the low resolution of the 3D camera. For the cubes used during the experiment a three-finger pinch grasp was optimal, which is why the use of this grasp was enforced throughout the experiment.

Three Finger Pinch Grasp: Actually the three finger pinch grasp was not in the repertoire of the grasps, the system could perform. It was newly introduced when I decided for cubes as the objects to be manipulated during the experiment.

Shape Preference: We also enforced that the super-quadrics model would use the correct archetype to fit the cubes (type “box”). Again the free choice of the super-quadrics arch-type might be hampered by the low camera resolution.

Arm Preference: We introduced a way to influence the choice of the arm which picks the cube. This can now take into account the reachability of the target position. This would be dispensable once the system reliably can perform a hand-over of objects from one hand to the other.

Publish World Belief: The vision system previously had no way to publish the results of its 3D point cloud segmentation and the comparison of the segment with the data base. This was introduced.

For my purposes the reliability of the robotic system was top priority. Hence, modification which would be unacceptable for a robotics study were even desirable for this study.

7.7. Limitations of the System/Future Work

The resulting system is already extremely flexible and versatile and I would advocate its use for further research in the fields of BMI and machine-mediated interaction. The system might also be an interesting platform for research in rehabilitation robotics and beyond. It would probably make for an impressive demo-system for official guests and the press at CITEC, too.

Any system can, however, be improved and advanced and here I want to share some ideas on that:

Positioning in Relation to Objects: The user of the system cannot yet place the grasped object in some relation to a second (target) object (in front of, inside, etc.). This is already possible using the verbal input of the robotic system and could be implemented for the BMI input in the future as well. One possibility would be, for the choice of a target position, to include five markers for any object recognised: One each, for placing in front, behind, left, right and on top/into the respective object. For scenes with many objects this might lead to an abundance of possible choices, not only increasing the chance for a misclassification but primarily elongating the duration of a P300 decision. This might be avoided by a three-stage decision process: Choosing i.) the object to be grasped ii.) the object in relation to which the grasped object shall be positioned iii.) the relation between grasped and target object (in front of, left of, ...).

Arbitrary Positioning: Instead of having just a set of predefined target position one could use a matrix like stimulus presentation to allow for an almost arbitrary selection of the target position. Similar to the renown P300 spelling matrices, e.g. Farwell and Donchin (1988).

Different Actions: Besides picking and placing objects, the robotic system can perform a few other action with a given object: The robotic hand can point at the object and it can grasp and offer it to the robot operator. When selecting an object via the BMI a second choice could be offered by the BMI for one of the three actions to perform with that object. As long as this choice is being made via P300, one needs to find a way to maintain the oddball paradigm, however. Maybe in the future the system will be capable of performing further actions on objects.

The BMI robotic system has been developed keeping it extensible, but with a clear focus on the experiment. Its capabilities are limited to those necessary for the planned experiment. However, it would require relatively little effort to extend the robotic BMI to be even more versatile.

8. (Hyper-)Connectivity Dependence on Cooperation

Having established the experimental design and the system for the iCusss experiment, next I will describe the conduction of the experiment, the analyses of the resulting data and their results in this chapter. I will describe the data acquisition in section 8.1 and the structure of the resulting data set in section 8.2 below.

The analytic methods used for the analyses on this data have been described in chapter 5 on page 33. How these are applied to the data is described here: The preprocessing performed before the actual analyses is described in section 8.3 on page 94. I analysed the data with regard to two different groupings of the data by two different criterion. These are described in section 8.4 on page 95. I applied two connectivity estimates with complementary properties: PSI which is bi-variant and collapses time dimension and PDC which is multi-variant and maintains time dimension. I will describe how these estimates were applied to the data in section 8.5 on page 96.

Having established how the data was analysed, I will then come to the results of the analyses in section 8.6 on page 97 and discuss these results in section 8.7 on page 110.

8.1. EEG Recording

The recording of the EEG data faces similar requirements and is done in a similar way as for the first experiment (see section 6.4.2 on page 67).

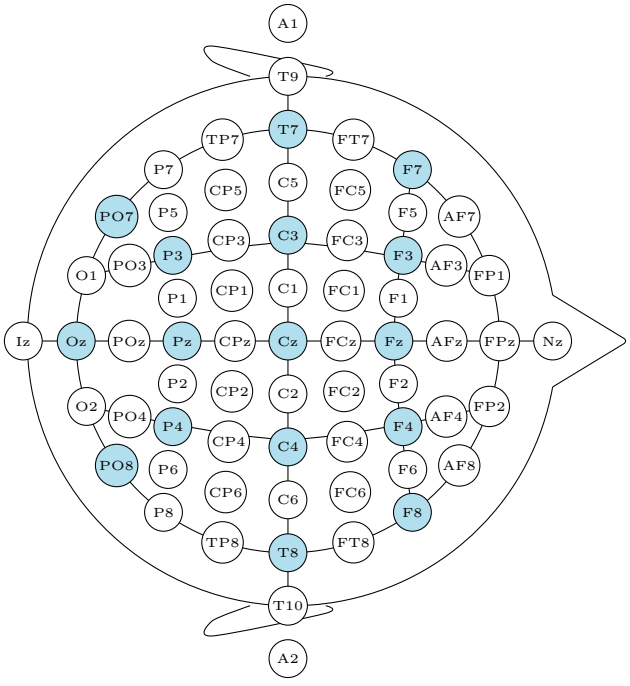


Figure 8.1.: The position of the electrodes during the iCusss experiment. The aim was to cover the entire cortex, as I wanted to avoid any assumptions about loci relevance for the later analysis.

The main difference is in the positioning of the electrodes. When no assumptions can be made about the loci of relevant neural activity, the positioning of the electrodes should cover as large portions of the cortex as possible, meaning that electrodes should be distributed equidistantly.

Contrary to the first experiment, in which I employed a brain activity pattern which is strongly spatially focused (ERD), the P300 brain activity pattern used this time spreads almost over the entire cortex. Furthermore it is more prominent than motor ERDs. Hence, a partial concentration of electrodes at certain loci (which I had during the first experiment) is not necessary this time. I, therefore, used an approximately equidistant distribution of electrodes all over the cortex. The electrode positioning which was chosen is given in figure 8.1. Ground and reference electrode were placed on the mastoids on both sides.

Again, impedances were kept below $10k\Omega$, but impedances were measured using the impedance measurement mode of the GUSBAmp EEG Device.

8.2. Data Sets

I conducted the experiment a total of 11 times. Two data sets had to be excluded from analysis, due to chance level classification of P300 BAPs.[1] Although no exact classification accuracy can be determined due

[1] Based on cross-validation of the training data.

to the lack of a ground truth, the cube configurations achieved by the participants as well as their oral reports, both confirm that for the remaining experiments all participants had reasonable to excellent control over the system.

Each of the nine remaining pairs of participants completed the 11 tasks. The first task (depicted in figure B.3a on page 126) was intended as a test run. With respect to the complexity of the system, I saw fit to start with such a dry run during which the participants could verify whether or not they understood the instructions correctly and after which they had the chance to ask questions. This task was consequentially excluded from analysis.

A variety of time markers were introduced into the data sets:

`baseline`: When preparations of the experiment were completed, the screens in front of the participants were switched off and the participants were instructed to remain seated in a relaxed manner for some baseline data recording. The baseline data recording was performed for two minutes starting when the operator decided (by visual inspection) that the initial oscillation in the data caused by the hardware frequency filters in both devices had vanished. Markers in two second intervals were introduced into the data offline.[2]

`cube - <xy>`: When one of the participants had to select a cube for grasping, the available cubes were highlighted in quick succession (flashing, compare section 7.5.3 on page 88). Two different types of makers identify the points in the data at which the stimulus presentation of either of the participants (identified by his/her initials, here xy) highlighted a cube.

`position - <xy>`: When a participant had to select a position to place the previously selected cube, the five predefined, triangular position markers were highlighted. The time markers in the data are analogous to those for the cube selection.

`newTask`: When a new task (configuration of cubes to achieve) was given to the participants this marker was set. When all cubes where placed in their starting positions and both participants confirmed they were ready to proceed, the hierarchical state machine controlling the experiment (see section 7.4 on page 83) was switched manually to the Select Cube state. This is the point in time at which the marker was placed.

`userTask:started`: When a participant had successfully selected a cube and a position, the robot would eventually start executing the task. The point in time at which the robotic system confirmed it would now start with the execution (when the <`user_task`> XCF-message received the <`STATUS`> `accepted`, compare appendix D on page 129) was marked. Only such <`user_tasks` > which were successfully completed (<`STATUS`>of the user task reached `completed`, eventually), were included.

`userTask:finished`: The point in time at which the robot had finished executing a command it had previously started was marked (when the <`user_task`> XCF-message received the <`STATUS`> `completed`, compare appendix D on page 129).

8.3. Data Preprocessing

After the experiments, the data sets of all participant were subjected to an ICA (see section 5.3 on page 42), individually. ICs containing muscle artefacts, in particular ocular artefacts, were identified by visual inspection aided by the ADJUST toolbox (Mognon et al., 2011b). The artefact ICs were **not yet** rejected!

Having completed the ICA, the two 16-channel data sets of each experiment were synchronised and joint to a single 32-channel data set. The ICA matrices were arranged such that the components identified in the individual ICA runs were maintained, i.e. the lower left and upper right parts (which would combine channels from different participants into one component) were (compare figure 8.2).

Next I needed to define on which data epochs I wanted to do a connectivity estimation. I listed the events for which markers were recorded in the data in section 8.2 above.

The time frame which I suppose to be most interesting when aiming for neural correlates of human interaction, is the execution of the task by the robot. This is the only time frame during which participants can try to infer their partner's intentions which are being mediated by the robot. Hence, this is the only time frame for which we know that some interaction level information flows over the channel that is the

[2] These were actually the only markers which were introduced after the experiment had ended.

common reference (compare section 7.2.1 on page 80). This time frame is, therefore, surely the most relevant for hyper-analysis. The PSI-based analyses is done on data epochs of two seconds length which are aligned with the `userTask:started` event markers. I could have used the entire timespan from the beginning to the end of each of the robotic actions (`user_task:started` to `user_task:finished`), but I wanted to avoid to have data segments of different length in my analysis.[3] For PDC analysis, for which no comparison with baseline data was conducted, I used a time frame of six seconds, which should encompass most of the robotic movement.

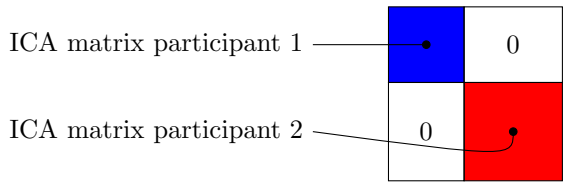


Figure 8.2.: The ICA matrix for the hyper-scanning data set as it was constructed from the two individual ICA matrices (red and blue).

Before any connectivity estimates are applied to the data it is, in a final step, subjected to detrending and normalisation. Particularly the normalisation would not be necessary for PSI connectivity estimation, but it does not hamper the results neither and it is strictly required for PDC connectivity estimation.

8.4. Data Groupings

First, I compare the epochs recorded during the experiment (i.e. all epochs from the last ten, but excluding the initial task, compare section 8.2 on page 93) with epochs generated from the baseline period at the beginning of the experiment. This way I can identify those connections which change in the transition from baseline period to the experiment.

There are several statistical methods aiming to evaluate whether a connectivity estimation is actually significant. Some of these methods test against some type of surrogate distribution (e.g. phase randomisation) or use some fixed threshold. For instance a PSI value larger than 2 can be deemed significant (Nolte et al., 2008; Sanei, 2013). The assertion of these tests is slightly, but notably, different from that of my approach. I do not test whether a connectivity estimate is significant (per se), I rather test whether a connectivity estimate is significantly different between baseline and experiment data. The different areas of a human brain are never completely decoupled. Hence, the question whether there is a significant change in a connectivity estimate from this baseline to the experiment condition is more relevant (or at least more strict) than the question whether a connectivity estimate is significant per se.

Apart from that, there are different aspects regarding which this data can be analysed. After careful consideration I decided to pick two data groupings, I considered to be most promising.

First, I have two different conditions by design: There are tasks (cube configurations to achieve) which require cooperation to be completed and others for which each participants solves his/her part of the task on his/her own (see section 7.2 on page 80 for details and appendix B on page 125 for the tasks divided by these condition). Any data epoch which was recorded while a cooperative task was carried out, was labelled as "coop" and all other epochs were labelled as "non-coop".[4]

Second, I divided epochs by the roles of the participants. Each epoch is aligned with the start of a robotic action and each robot action was initiated by one of the participants. That participant already knew what the robot would do, while the other participant could learn this only from observation of the robot. All epochs for which participant one was the initiator of that robotic action were grouped as one condition and all epochs for which participant two was the initiator were grouped in a second condition. I named these data groups "init1" and "init2".[5] These data groups (init1 and init2) can be (semantically) interpreted in two ways: As roles (action initiator and action observer) and as two different directions of an interaction level information flow (from initiator to observer).

The robotic actions are the only means for information interchange available to the participants in my remote setting. When one participant initiates a robotic action, that participant is the only one who is capable of sending information (volitionally or not) at that time and any information send, is being mediated by the robot. Sorting the trials according to the initiator of the robotic task, hence, means sorting the epochs according to the direction of the interaction level information flow and that participants have different roles

[3] Including baseline epochs for the comparison of experiment and baseline data, see below.
[4] Obviously, all data segments recorded while the initial task was carried out, were labelled neither coop nor non-coop, but were excluded from analysis.
[5] Again, excluding epochs from the initial task.

in the different data groups (sender and receiver). However, the interaction level information flow postulated, not necessarily needs to translate to neural information flow, i.e. I cannot expect to find a neural information flow from initiator/sender to observer/receiver just because I can assume an interaction level information flow in that direction.

8.5. Connectivity Analysis and Statistics

I perform the analyses on four frequency bands: θ, α, β and γ (see section 2.2.5 on page 15). These are the frequency bands which can be considered most relevant for neural interaction research. Various other studies have chosen the same set of frequency bands, e.g. Babiloni et al. (2007a); Astolfi et al. (2009, 2010b, 2011a); De Vico Fallani et al. (2010); Yuan et al. (2010); Dumas et al. (2010).

A variety of connectivity estimates exist. Some of these are introduced in section 5.6 starting on page 47. In the end I decided to use the Phase Slope Index (PSI) and Partial Directed Coherence (PDC) as estimates for neural connectivity (see section 5.6.5 on page 53 for PSI and section 5.6.3 on page 52 for PDC). The results of the PDC analysis are given in section 8.6.3 on page 107, towards the end of this chapter. The results of the PSI analysis are presented in the next section on page 97.

More precisely I used the Renormalised Partial Directed Coherence (RPDC, Schelter et al. (2009)) whose main advantage for my analyses is that it is scale free, i.e. not dependent on the unit of the input data. It should be noted, however, that RPDC does not share the property of normal PDC to be normalised by output (compare section 5.6.3 on page 52).

The properties of PSI and PDC are complementary in many regards: PDC perseveres the time dimension, PSI collapses it. PDC is multi-variant, PSI is bi-variant. PDC represents recurrent connectivity, PSI gives one net-information flow for any channel pair. Due to the high need for training data for the MVar model on which PDC is based, I could not conduct any PDC analyses comparing different conditions but only estimated the PDC on all experiment epochs, using phase randomisation for statistical testing (compare section 5.5 on page 45 for MVar models, section 5.6.3 on page 52 for PDC, section 5.7.1 on page 54 for phase randomisation and section 8.6.3 on page 107 for the results of the PDC analysis). Furthermore, PSI and PDC differ in their approach: PDC is basically an advanced form of the classical Coherence and PSI is a direct result of considerations about the propagation of signals' phases.

Hence, the comparison of different conditions can only be done using PSI on my data. To compare two groups of data (e.g. coop/non-coop or experiment and baseline data) one simply subtracts the corresponding PSI values. I will refer to this a differential PSI throughout this thesis. The resulting values still need to be statistically substantiated, though. It has been argued that for PSI a value larger than 2 could generally be considered significant (Nolte et al., 2008; Sanei, 2013). This could be adapted to test the significance of the difference between two PSI values.

I prefer an alternative significance test, which, in my opinion, gives a more direct answer to the question: "Does connectivity change between two groups?". After computing a differential PSI value, I repeatedly, randomly reassigning the condition labels and then recomputed the differential PSI value. When the differential PSI value on the original data grouping is larger than the vast majority of the differential PSI values on the randomised data grouping, we can consider the original differential PSI value to be significant.[6] I used 2,000 random data grouping for all analyses.

I was deliberately unspecific, speaking of a "vast majority" before. The rate of differential PSI values on randomised groupings that is larger than that on the original grouping, can be interpreted as something similar to the p-value of other statistical tests such as t-test. For those, in most cases one would apply a threshold of 0.05, meaning that no more than five percent of the random grouping differential PSI values may be larger than the original differential PSI value, if that original differential PSI value shall be deemed significant.

However, there is a total of $9 \times 4 = 36$ (number of experiments times number of frequency bands considered) connectivity tables for each of the three analyses listed below. The different data-sets have a total of $5,151$ possible connections (number of components to the square for each data set). Hence, for each of the three analyses I had to conduct $186,156$ differential PSI computations. The sheer number of statistical comparisons made, calls for a correction for multiple comparisons. I decided to use the approach of controlling the false-discovery rate (compare section 5.7.2 on page 54). I control the expected rate of false discoveries (FDR) among all discoveries such that it does not exceed 20%. This is done by computing a new threshold below

[6] This method is following a suggestion by G. Nolte during the above mentioned email conversation. I thank him for this insightful discussion.

which p-values are deemed significant, on the basis of all p-values. The threshold acquired by controlling the FDR basically replaces the canonical 0.05 threshold.

I conducted various differential PSI analyses:

baseline vs. experiment I analyse which connections are different during the experiment as compared to the baseline period recorded before the experiment began (see the listing of events in the data in section 8.3 on page 94). In the appendix E starting on page 131 these results are listed.

coop vs. non-coop The results of the differential PSI analysis comparing epochs during cooperative tasks with epochs during non-cooperative tasks can be found in appendix F on page 167 and following pages.

init1 vs. init2 The results of the differential PSI analysis between epochs with different initiators of the robotic action can be found in appendix G on page 203 and the following pages.

The most relevant and interesting of these results will be discussed in the next sections.

8.6. iCusss Results

Now that I have established which data exists, how this data is to be prepared and which analyses are to be conducted on that data, I will now present the results. I will begin with the differential PSI analyses. First, in section 8.6.1, below, over all experiments. And, second, in section 8.6.2 starting on page 102, describing some observations made on a per-participant basis. Finally, I will discuss the results of the PDC analysis in section 8.6.3 starting on page 107.

8.6.1. General Observations on the Differential Connectivity Analysis

There is no obvious method how to draw such general conclusions from connectivity analyses on neural data, i.e. results which go beyond treatment of individual participants (or pairs of participants as the case may be). Therefore, many publications in this field remain on a rather descriptive, per-participant level. Some studies employ graph analysis in order to obtain some general truth. Testing different graph analysis approaches did not yield any additional insights on my data. Additionally, as I only have 16 channels (and therefore less than 16 non-artefact ICA components) per participant, the resulting graphs are pretty small and the applicability of the commonly used graph measures could be debated.

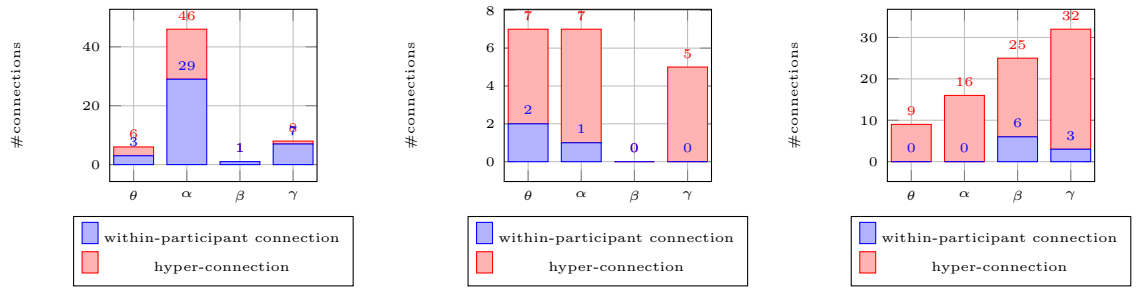
One very basic approach, is counting the number of connections for which significant differences have been found. For all parts of the analyses, dozens of connections undercut the canonical 0.05 threshold. However, controlling the FDR to a max. of 20% discards many of the potentials discoveries. The distribution of the remaining significant findings over frequency bands and analyses still offers interesting insights.

A very general observation is, that the number of significant connections is subject to high variability between pairs of participants. While in many cases significant differences are found for no or few IC-pairs at all, in other cases up to 22 significant differences were found for a single pair of participants and a single frequency band.

Figure 8.3 depicts the number of significant connections found. Examining these plots we learn that contrasting baseline with experiment data, differences were found predominantly (but not exclusively) among within-participant connections. Hyper-connections with significant differences between baseline and experiment data concentrate on α-band, mostly. Contrasting coop and non-coop or init1 and init2 data, the vast majority of connectivity differences found in these analyses affect hyper-connections.

Furthermore, in total much fewer significant differences were found contrasting coop and non-coop data than for the other two analyses. This is remarkable. I designed the tasks of this experiment around these two conditions, expecting that the degree of cooperation needed to solve a task would be highly relevant for the emergence of neural connectivity.

It is interesting to note that many studies from the literature imply some sort of distribution of roles between participants, e.g. pilot and first officer (Astolfi et al., 2012), first and second player (Astolfi et al., 2010b), model and imitator (Dumas et al., 2010), leader and follower in a guitar duet (Sänger et al., 2013)). In my analyses, the init1/init2 set data grouping, which implies roles as initiator and observer, yielded many more significant differences in connectivity than the coop/non-coop data grouping, which lacks such an implication. From any of the asymmetric role assignment listed (others' studies as well as my init1/init2 data grouping) a dominant direction of the interaction level information flow can be derived. Hence, the effects of asymmetric roles cannot be told from the effects of a dominant direction of interaction level



(a) Contrasting experiment and baseline data.

(b) Contrasting cooperative and non-cooperative trials.

(c) Contrasting epochs for which either participant one or two initiated the robotic action.

Figure 8.3.: The number of significant differences found in the three PSI analyses conducted.

information flow. A careful study design dedicated to this research question is required, particularly because an interaction level information flow always has a sender and a receiver and thus implies assignment of different roles.

In the the following I will examine the results per frequency band considered in more detail, relating the results to research from the literature. I will cover the four bands from low to high frequency, consequentially starting with θ-band.

θ-band

θ-rhythm definitely plays a functional role in machine-mediated human-human interaction. I found significant changes in both, within and hyper-connections, when contrasting baseline with experiment data and we find significant differences in the two other analyses, mainly affecting hyper-connections. These observations are not easily explained with the classical interpretation of θ-rhythm as being associated with sleep, meditation and creativity. The setting does not ask for a much creativity. The high level of concentration required to operate the BMI actually comes with a high level of tension, while meditation is a state of relaxed concentration. Much less do these classical interpretations of θ explain its active role in hyper-connectivity. Several recent studies, however, also found that θ-rhythm plays a role in hyper-scanning settings, e.g. Babiloni et al. (2007b); Sänger et al. (2012); Astolfi et al. (2014, 2011a, 2010b, 2009); Dumas et al. (2010); De Vico Fallani et al. (2010). The role of the combined θ/α-band in human-human interaction is also highlighted by Kawasaki et al. (2013).

Babiloni et al. (2007b) reported significant differences in θ band, but these were in activation (meaning change in spectral power) not connectivity. Later Astolfi et al. (2010b) reported a number of significant within-participant connections in θ-band in a hyper-scanning setting. They analysed the in- and out-degree of the nodes of the resulting graph, but did not suggest a semantic interpretation of θ's role in the hyper-scanning setting. Again later, Astolfi et al. (2011a) reported a synchronisation in θ-band which they related to "an increase of the resources employed for the information processing by the cortex."(Astolfi et al., 2011a).

These findings from the literature are being supplement by my own results by indicating that θ plays a role in human interaction, even when this is machine-mediated. Furthermore they indicate that θ (hyper-) connectivity changes depending on the degree of cooperation needed to solve a task and depending on the role assignment between participants (initiator/observer) or, equivalently, the dominant direction of interaction level information flow. It should be noted, that significant differences in θ-band between init1 and init2 epochs affect exclusively hyper-connections.

α-band

I found a huge number of significant differences in connectivity in α-band. This is especially true when contrasting baseline and experiment data, where three quarter of all significant changes happen in α-band, a good third of these affecting hyper-connections. This predominant role of α was, to my knowledge, not previously reported in such clarity. I can, currently, only speculate about the reasons for this role of α. It could possibly be an effect of the machine meditation, but to my knowledge no other study performed a comparable analysis in a non-machine mediated setting. Hence, this cannot be confirmed, currently.

Contrasting coop with non-coop data I could still find quite some significant changes. Although α no longer plays a dominant role in this analysis, it is also not inferior to the other bands analysed, neither. And the vast majority of the affected connections are hyper-connections. Finally, contrasting init1 and init2 data I again found quite some changes in α-band, all of which affected hyper-connections.

A vast corpus of literature can be found on the interpretation of the α-rhythm. One very basic interpretation is that α is an idle rhythm of the visual cortex. Tognoli et al. (2007) compared a baseline condition with relatively little visual information to process with an interaction condition requiring much more visual processing. Consequentially they reported a drop in α (and μ) band power which they attributed to α's role as an idle rhythm of the visual cortex.

That α's role in social situations goes beyond this, has been demonstrated by Astolfi et al. (2009). Similarly to Tognoli et al they compared spectral band power of the α-rhythm between a baseline condition and social interaction. However, the complexity of the scene and therefore the degree of visual processing needed, was comparable between the two conditions. Astolfi et al reported an increase in α (and θ) band power, clearly hinting at a role of α which goes beyond that of a pure idle rhythm of the visual cortex.

My analyses go beyond these studies, insofar as they are based on connectivity estimation rather than α's spectral power. My results show that α-band connectivity changes profoundly between the experiment and the baseline period. This is further evidence for a role of α beyond that of a pure idle rhythm.

Other studies confirm changes in α-connectivity and hyper-connectivity during human interaction. Babiloni et al. (2007a) could show a varying degree of within-participant connectivity depending on the strategy in the Prisoner's Dilemma (and therefore on the degree of cooperation) chosen by the participant. De Vico Fallani et al. (2010) confirmed these results and also could (to a certain degree) predict the strategies of the participants based on the neural activity. Astolfi et al. (2012) showed that pairs of professional pilots develop a series of hyper-connections in the α-band during starting or landing, which vanish during normal flight.

Several other studies reported differences in α-connectivity between different conditions, e.g. Babiloni et al. (2007c); Sänger et al. (2013); Astolfi et al. (2010b, 2011a, 2010a); Dumas et al. (2012a, 2010), further strengthening evidence for the functional role of α in social situations. These combined results stress the important, functional role α-rhythm plays for the neural foundations of human interaction.

My results further under-pin the functional role of α, adding to its interpretation in several aspects. First, the results of the baseline vs. experiment analysis show that α still has a functional role when interaction is machine-mediated. Even more so, considering that a good third of the affected connections are hyper-connections. This suggests an integrative functional role of α in a team of humans. This conclusion is further reinforced, by the dependence of α-hyper-connectivity on the need for cooperation (coop/non-coop) and the direction of the interaction level information flow (init1/init2).

During baseline, participants were seated in the same way as during the experiment, they knew about the presence of a partner in another room – the situation had not physically changed between baseline and experiment condition. What had changed is that a.) participants were now engaged into an activity and b.) participants knew/felt they were engaged in that task together. Apparently this mental state or sensation changes the neural connectivity profoundly.

That a sensation or the belief to be in a common activity can fundamentally change hyper-connectivity has most clearly be shown in Astolfi et al. (2014). Participants had to solve the same task twice with a partner. But the fact that they were told during one iteration, they would solve the task with a computer (rather than with the partner) lead to fundamental changes in hyper-connectivity.

The results furthermore suggest that the degree of cooperation has an impact on the neural connectivity of the participants in α-band. Hence α connectivity is apparently not only modulated by the (perceived) engagement of in a common activity but also depending on whether participants (belief) to cooperate or not. Comparing the results of the experiment vs. baseline analysis to the coop vs. non-coop analysis I can, however, state that the effect of engagement in a common activity is dominant over the effect of different degrees of cooperation.

β-band

In β-rhythm almost no increase in connectivity could be shown, contrasting baseline and experiment data. Contrasting coop and non-coop data, again no significant differences were found in β-band. This is remarkable, since the β-rhythm is classically associated with problem-solving and active thinking, two fields of mental activity which I would deem to be highly relevant in the setting. Also different studies from the literature demonstrated a functional role of β-band during social situations, namely Sänger et al. (2013);

Dumas et al. (2010); De Vico Fallani et al. (2010). Astolfi et al. (2009) could also be named here, but the analysis of this study is based on band power, rather than connectivity, limiting the comparability with my own results.

Contrasting init1 and init2 data I found quite many significant differences, mostly, but not exclusively, affecting hyper-connection. The β-band is actually the band with the most significant changes in within-participant connections contrasting init1/init2 among all frequency bands considered. The first two analyses (baseline/experiment and coop/non-coop) could have raised suspicions whether β-activity plays a functional role in my setting at all. The last analysis resolved these concerns. β plays a functional role. But why could (almost) no significant differences in β band be found for the first two analyses? To understand that I will evaluate different hypotheses for factors for the emergence of β-(hyper-)connectivity. It is my general assumption that several factors might bring forth this type of connectivity.

One possible factor can be derived from β's association with motor-planning. Eventually β connectivity emerges when actions are (physically) executed.

In Sänger et al. (2013) a duet of guitar players showed an increased β connectivity while playing. The task in that study required complex finger and hand movements to be executed and coordinated while in the iCusss study the participants executed no action themselves.

In Dumas et al. (2010) participants were assigned the roles of a model and an imitator of hand movements. Participants showed neural synchronisation in periods during which they synchronised their activity. Again, participants had to execute hand movements themselves (in contrast to the iCusss study). However, the movements were probably less complex and with a much diminished degree for coordination compared to guitar playing.

Regarding my own results, assuming the hypothesis above was true, the fact that participants observed the robotic hands executing their commands from an I-perspective, a-priori, might or might not have lead to the same emergence of β-connectivity as the other two studies. Regardless, such β-connectivity needed to be present in all experiment/baseline analyses. Hence, it can only be deduced that, if actual motor execution is one factor yielding β-hyper-connectivity, the same is not true when actions are carried out by a robot and, therefore, during machine-mediation.

The study by De Vico Fallani et al. (2010) shows that motor execution cannot be the only factor yielding β-hyper-connectivity: The participants played iterations of the Prisoner's Dilemma. The analysed data period for which β-connectivity was found, was recorded while the results of the iteration were displayed and *after* the participants communicated their choices via keyboard, i.e. during that period participants were idle. The β-connectivity demonstrated by this study, therefore, must have been evoked by some other factor.

It is interesting to note that in Sänger et al. (2013) and Dumas et al. (2010), participants were assigned distinct roles (leader and follower). As I already stated coop/non-coop data grouping implies a symmetric role assignment while init1/init2 implies an asymmetric role assignment. And indeed, the only data grouping which yielded significant differences in β-band is init1/init2. Another potential factor for the emergence of β-connectivity could, therefore, be an asymmetric role assignment or (equivalent for the cited studies) an interaction level information flow with a predominant direction.

The same hypothesis can be reached from another line of argumentation. The β-rhythm has recently often been associated with the mirror neuron system (see section 2.2.5 on page 15), and, therefore, with co-representation. Both, Astolfi et al. (2009) and Dumas et al. (2010) had one participant execute actions which were observed by the partner and which were relevant for the partner. Hence, it can be assumed that the partner co-represented the action on a neurological level. The same can be said for my own study. But while for coop/non-coop data grouping the assignment of the role of the observer who needed to interpret his/her partner's actions was balanced, the init1/init2 grouping basically sorts the data according to this criterion. However, although the two lines of argumentation (asymmetric roles/motor co-representation) are quite distinct and operate on different semantic levels and would represent two different factors, on the given data (my study and the cited literature) the two factors are indistinguishable.

Non of the different, hypothetical factors for the emergence of β-connectivity discussed here would, however, explain the emergence of β-connectivity reported by De Vico Fallani et al. (2010). The reported difference is between a condition in which both participants follow a strategy of defection in Prisoner's Dilemma and the other possible strategies, hence, whether or not the participants cooperated or not. In so far these results directly contradict my own results, as I could not find significant differences between coop and non-coop condition in β-band in my data.

However, the notion of cooperation in my own study and that of De Vico Fallani et al. (2010) differs substantially. During the non-coop condition of my study, participants solved their (partial) tasks independently.

The mutual defection of both participants in the Prisoner's Dilemma actually implies, participants mutually hinder one another. Furthermore the Prisoner's Dilemma knows three standard strategy profiles/degrees of cooperation, while my study knows only two degrees of cooperativeness. Finally, the differences reported by De Vico Fallani et al. (2010) have been found using a Perceptron on high-level graph measures on the connectivity networks. Those graph measures would have limited validity to my own data (see above) and I can, therefore, not conduct the same analysis.

The questions which factors influence the emergence of β-connectivity cannot be fully explored in this thesis. However, due to their uniqueness, I consider the results of this thesis to be highly relevant. All these potential factors are, of course, speculative. What ever the case, the discussion of the above paragraph might hint towards potential future lines of research regarding β's role in these types of setting.

γ-band

Before I move on to the discussion of the γ-band I want to pass a cautious note: It has been debated whether or not results from the EEG γ-band can be considered valid when recording is done outside of an electrically shielded room. In Europe the power grid operates with an AC of $50Hz$. My analysis of γ activity has an upper limit of $40Hz$. It has been debated if electrical fields from power lines and electrical devices in the surrounding could systematically affect neural recordings of $\leq 40Hz$. Despite this ongoing discussion various recent studies include results from this frequency range without reporting the use of electrical shielding, e.g. Babiloni et al. (2007a) with an upper bound of $40Hz$, Dumas et al. (2010) with $\leq 48Hz$, Astolfi et al. (2010b) with $\leq 40Hz$ (although the "room was tested previously for the absence of particular electrical noise."(Astolfi et al., 2010b).), Astolfi et al. (2009) $\leq 40Hz$. Regardless, it is surely worth to keep in mind that objections regarding the use of this frequency range exist and cannot easily be rebutted.

The γ-rhythm is associated with concentration and learning. Both mental processes, which can be supposed to be predominant during the experiment. I could show significant differences contrasting baseline and experiment data. It is interesting to note that most findings about the γ band in the literature also applied similarly to the β-band. This is in contrast to my own results, as I found no significant difference between baseline and experiment in the β at all, while I found some significant results for the γ-band. I can currently only speculate on the reasons. Almost all of the significant differences in γ-band concern within-participant connections. Hence, the differential PSI analysis between baseline and experiment seems to suggest that γ plays an integrative role for different brain areas, but its integrative potential for the pair as a whole is limited.

The differential PSI analysis between coop and non-coop refutes this assumption by revealing various significant changes exclusively affecting hyper-connections. Hence, γ-connectivity changes depending on the need for cooperation (or the absence of such a need) which would indicate a strong integrative role of γ across participants. This is underpinned by the results contrasting init1 and init2 data. Again the vast majority of the affected connections are hyper-connections.

For init1/init2 analysis the γ-band is actually the frequency band with the second-most significant changes affecting within-connections and the most significant changes affecting hyper-connections. The high relevance this band receives in this analysis might best be explained with the degree of concentration needed. The use of a P300-based BMI demands a vast degree of concentration. The initiator of the task has just completed that process and, although presumptively remaining focused, now at least he/she reorients his/her focus towards the video stream and can relax from peak concentration during the P300 stimulus presentation.

Some studies confirm the functional role of γ in social situations. In Astolfi et al. (2010a) a consistent connectivity across all participants could be shown in β and γ-band during a card game. De Vico Fallani et al. (2010) also identified β and γ as most relevant for their post-hoc analysis. In Astolfi et al. (2009) a significant activity of the prefrontal cortex was found, mostly independent from condition. For Astolfi et al. (2010a) these findings "suggest that the right prefrontal cortical areas interested are related to the stress of the task performed against the other player."(Astolfi et al., 2010a).

So much about the observations on the distribution of significant results between conditions and frequency bands. The results from this analysis and the cited literature demonstrated significant, functional roles for all of the considered frequency bands. This does, however, not make these results arbitrary, it much more shows that social interaction is a holistic process, involving a wide range of neural functions, resources, brain areas and rhythms.

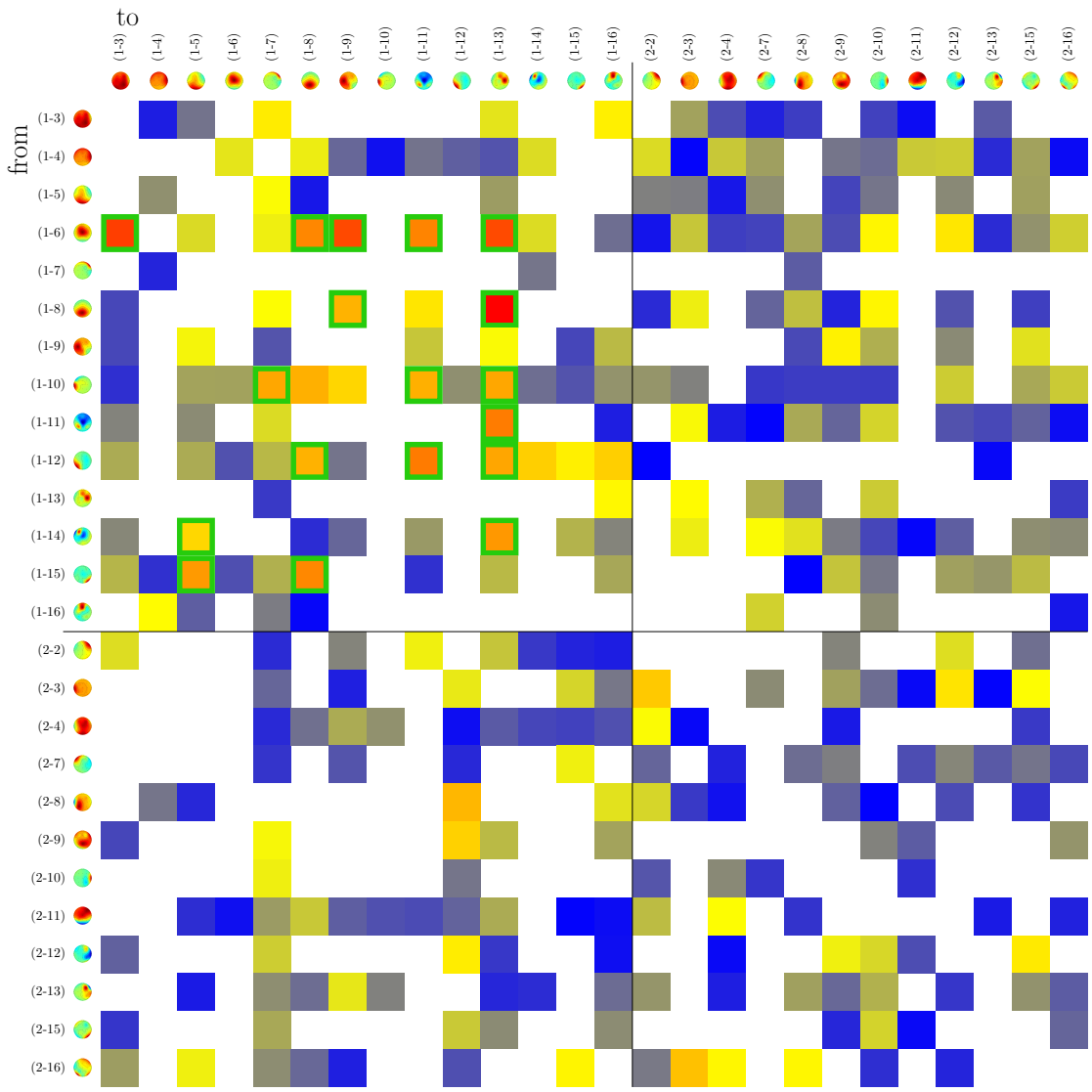


Figure 8.4.: An exemplary differential connectivity analysis between baseline and experiment period. This is the differential analysis of the α-band of experiment nine.

8.6.2. Individual Assessment of the Differential Connectivity Analysis

Formulating observations, conclusions and hypotheses based on the entirety of the nine experiments is important. However, for these types of experiments an examination of the results on a per-participant level can be also worthwhile. Participants are different. This is particularly true when dealing with neural recordings. Observations on an individual level, which have been made repeatedly in different settings and studies might lead to new insights. These might not always apply to every individual, but sometimes only to certain groups of individuals. Furthermore the examination of individual analyses allows us to go into details and illuminate fine-grained aspects of the connectivity analysis.

For the analysis contrasting baseline and experiment data, α-rhythms played a major role. Investigating this phenomenon further I will now examining the α-band analysis of experiment nine, which showed the most connections with significant changes in α-band of all experiments.

Figure 8.4 shows the connectivity table of that analysis. Each cell represents one pair of ICs, i.e. one potential connection. The number of rows and columns (and therefore the number of cells) is dependent on the number of independent components which remained after the removal of those components containing artefacts (i.e. muscular artefacts, compare section 5.3 on page 42). Here a total of 26 components, 14 for participant 1 and 12 for participant two, remained. Before each row the sending component associated with that row is depicted. Above each column the receiving component for that column is depicted. The

components are the same for rows and columns because any two components could potentially interact in arbitrary direction. The components are labelled with two numbers. The first number (1 or 2) denotes the participant to which this component belongs. The second identifies the IC for that participant. Components are enumerated in the order given by the ICA, i.e. components are ordered by decreasing projected variance (compare online documentation of EEGLAB, SCCN)[7]. This enumeration was done before components identified as artefacts have been remove and as a result the enumeration of the components depicted in the connectivity tables is not consecutive.

The components are from two different participants. In the table, they are ordered such that the first k of n components all belong to participant one and the last $n - k$ components belong to participant two. This is true for rows as well as for columns. The components belonging to different participants are separated by a horizontal and a vertical line. This divides the graph into four quadrants. The upper left quadrant contains all within-participant connections of participant one. Similarly the lower right quadrant contains the within-participant connections of participant two. The upper right quadrant contains connection from components of participant one to components of participant two and the lower right quadrant connections from components of participant two to components of participant one – thus, these quadrants contain the hyper-connections.

The colour of each cell indicates the differential PSI as a colour gradient, ranging from blue for the lowest non-negative PSI values to bright red for the highest PSI values. A little more than half of the cells are empty. In particular I omitted all connections of components with themselves (auto-regression connections). Demonstrating information flow from one neural source to itself yields no new insights, as such an information flow is virtually always present and pretty dominant. Furthermore, PSI is anti-symmetric, i.e. $PSI_{XY}(\omega) = -PSI_{YX}(\omega)$. Phrased differently, PSI can identify an information flow between two components, but a recurrent information flow from X to Y and back to X would be averaged out. Thus, a negative PSI value for the connection from X to Y actually indicates a (net-)information flow from Y to X. Hence, I omitted all negative entries, as the actual information flow for those entries is directed in the reverse direction. Consequentially, in the table, for any pair of components $i, j,\ i \neq j$ exactly one of the two entries (i, j) and (j, i) in the matrix is filled.

Finally all connections for which the changes identified were deemed significant in statistical testing including multiple comparison correction (FDR) are marked with a green frame. For the analysis depicted in figure 8.4 this means that any connection for which statistical testing detected a significant difference between baseline data as compared to the experiment data of experiment nine is marked with a green frame.

In figure 8.4 a total of 18 connections with a significant change can be identified, all of which are within-participant connections of participant one. It is, however, very likely that I committed a series type-II errors, i.e. not all effective connections the participants established during the experiment have been recognised.

Not only are all significant differences affecting connections of participant one, but also there is a clear bias to higher PSI values in that quadrant in general. Furthermore, a slight bias among hyper-connections may be identified, favouring participant one as the sender (the upper right quadrant is populated more densely than the lower left quadrant). Considering that for none of these connections the changes are is significant, however, and that most of the values are rather lower, this might be coincidental.

For most analyses the presentation as a table is comprehensive. For analyses with many significant connections these often form a non-trivial network and other questions become relevant, e.g. which components are sources in the network (influence many others but are not influenced themselves), which components are sinks (are influenced, but do not exhibit any influence themselves) and which serve as network hubs (relay information between different network parts). To address these types of questions a representation of the significant connections as a directed graph is advantageous.

Figure 8.5 shows the connectivity network for the α-band analysis of experiment nine (the same analysis as the connectivity table in figure 8.4). Twelve of fourteen components which have not been rejected as artefacts are involved in this network. Because all connections are within-participant connections of participant one the entire network describes the change in neural connectivity within that participant during the experiment as compared to the baseline period. It can be assumed, that a similar degree of change took place in all of the participants, only that the stochastic methods involved in the analysis failed to identify them as significant (type-II errors).

One can observe that component (1-6), a central component, influences many other components (namely five). For two of these connections, (1-6) to (1-13) and (1-6) to (1-9) we need to be cautious, however: The influence of (1-6) on (1-9) and (1-13) might be mediated over (1-8) (or (1-11)) – partially or entirely. In

[7] `https://sccn.ucsd.edu/EEGLAB/`

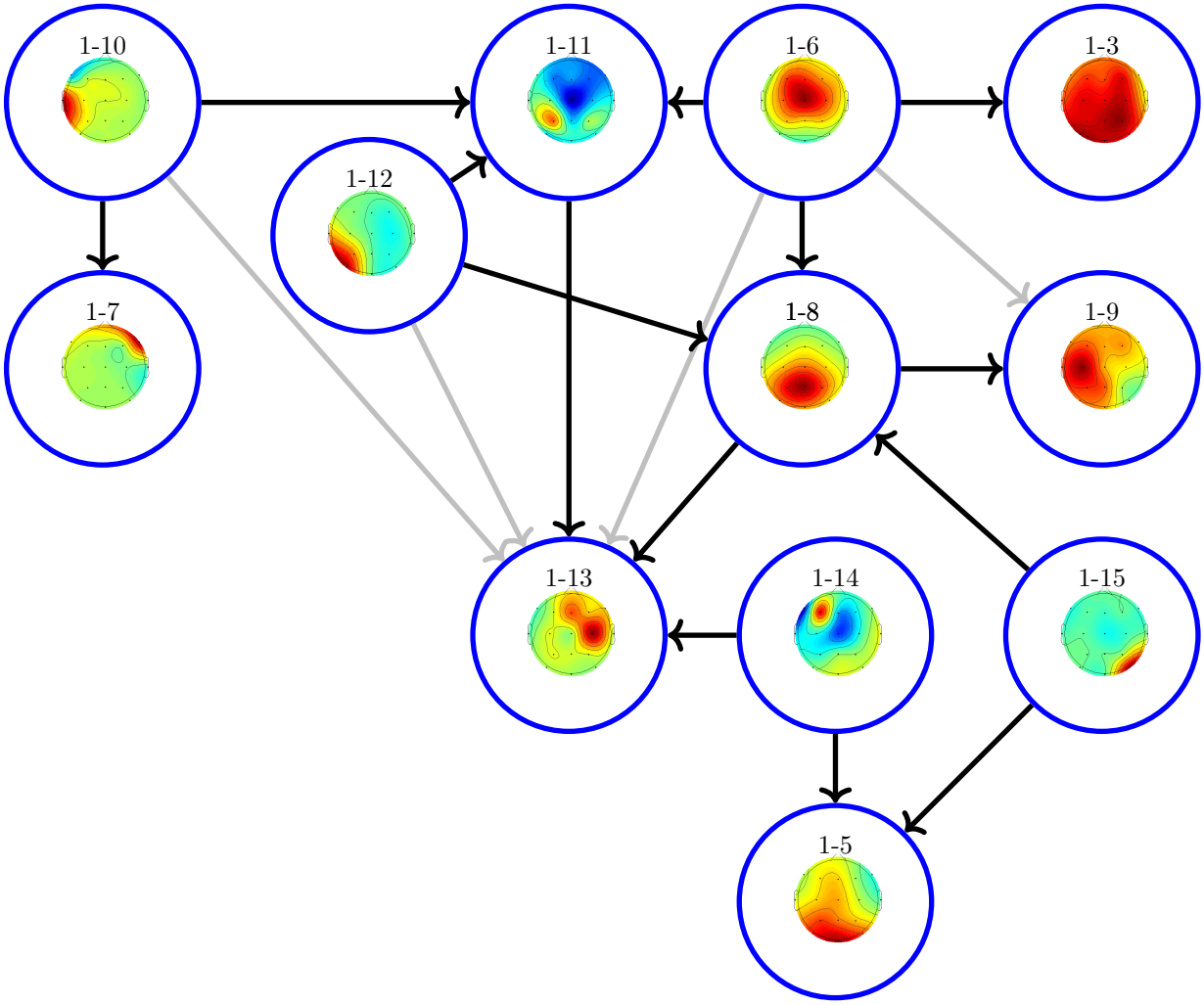


Figure 8.5.: The connectivity network of connections for which a significant change in connectivity in α-band could be shown contrasting experiment and baseline data for experiment nine. Grey lines are connections which might be explained by an indirect connection mediated by another component and only identified as a direct connection as well, because PSI is a bi-variant connectivity estimate (compare section 5.6 on page 47).

such a case, a bi-variant connectivity estimate (such as PSI) would still identify a (direct) connection from (1-6) to (1-9)/(1-13) (compare figure 5.8 on page 49). (1-6) might thus have a total of three to five outgoing direct connection. Another such case is the connection from (1-12) to (1-13). This influence might or might not be mediated over (1-8) and/or (1-11). The last connection we cannot be certain that it exists as a direct connection is (1-10) to (1-13), as the influence might be mediated by (1-11). In the end, component (1-13), which is the recipient of a total of six significant connections identified by PSI, might actually only have three direct incoming connections.

Regardless of these suspicions, one can identify that component (1-8) serves as a hub. It is a parietal component, which could potentially relay the influence components (1-6), (1-12) and (1-15) exhibit on it to components (1-9) and (1-13) (whether or not a direct influence between the respective components is identified as significant or not).

Another hub in this network is (1-11). Regardless whether or not (1-6) and (1-12) also have a direct influence on (1-13) the parietal-lateral component (1-11) is fit to mediate this influence to (1-13).

Although we cannot be certain about three of its ingoing connections, component (1-13) is the major information sink in the network. It is a fronto-lateral component and it has at least three incoming connections and no outgoing connections. A total of seven components could potentially influence this component either directly or indirectly.

Component (1-6), on the other hand, is a major information source, propagating information to three or four other components directly and to up to five component in total (either directly or indirectly).

The longest (shortest) path in the network as it has been identified by PSI[8] has a length of two. Four connections with significant differences identified by PSI might be explained as a mediated influence over a third component and might not actually exist as a direct connections. When being conservative and rejecting those connections, the longest (shortest) path length is still two. Short (average) path length while maintaining a relatively low total number of connections (18 of 132 potential connections) is a property of networks commonly referred to as small-worldness. It has been described in a variety of different network types ranging from social to neural networks (Rubinov and Sporns, 2010). Small-world networks are also attributed to be rather robust. This is underpinned by the fact that removing the four edges we are suspicious about, does not change the longest, shortest path length.

Finally, I will examine a second individual analysis more closely, namely the analysis of the α-band connectivity of experiment six when contrasting init1 and init2 data. The connectivity table can be found in figure 8.6.

Here, no bias can be found among hyper-connections. The difference in PSI values for init1 and init2 data among hyper-connections is positive for both quadrants in approximately equal shares. This is affirmed by the fact that significant differences were found for hyper-connections in both directions: from participants one to participant two and vice versa.

Among the connections for which significant differences have been shown, component (1-14) is remarkable in that way that it acts as a sender for six significant hyper-connections. It is a rather narrow fronto-central component.

For many of these individual analyses, significant differences can only be found for hyper-connections in one direction (either from participant one to participant two or vice versa). For these individual analyses a representation as a graph would be pretty pointless, as it would results in pairs of connected nodes. No three nodes in such a graph would be connected. However, for the present individual analysis this is not true. We have significant differences for connections in both directions resulting a non-trivial graph, discussing which is worthwhile.

Figure 8.7 depicts this graph, however, I removed all connected pairs (two connected component which are otherwise isolated). These contribute little to the overall layout of the graph (in terms of path-length, small-worldness, etc.).

All of the significant connections were hyper-connections, hence the graph is bipartite. As mentioned, component (1-14) is the source for many different hyper-connections. It influences the components (2-7) and (2-3), which focus on frontal regions and (2-13) which has a broad parietal-lateral focus. The last two components on which (1-14) has an influence are the lateral component (2-15) and the fronto-central component (2-6) which are both being influenced by one other component. (2-15) is being influenced by (1-8) which is strongly lateral and to which (2-15) is contra-lateral. (2-6) is being influenced by (1-6) which is almost an exact mirror image of (1-8).

Even further (2-15) is the starting point for kind of a daisy-chain connection over (1-3) (central), (2-12)

[8] That is the longest path, out of all shortest paths between any two connected nodes.

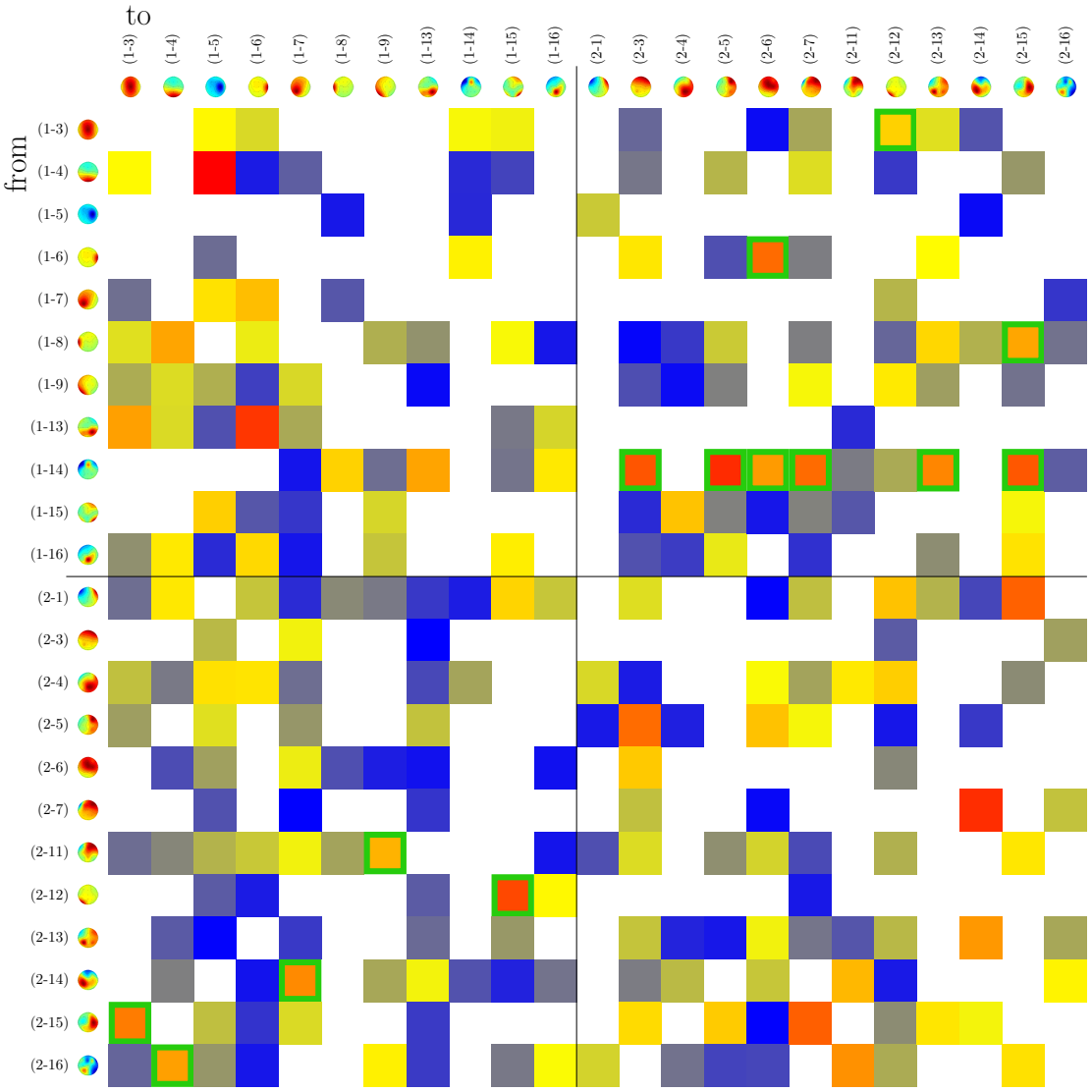


Figure 8.6.: An exemplary differential connectivity analysis between epochs during which the robotic action was initiated by different participants. This is the differential analysis of the α-band of experiment six.

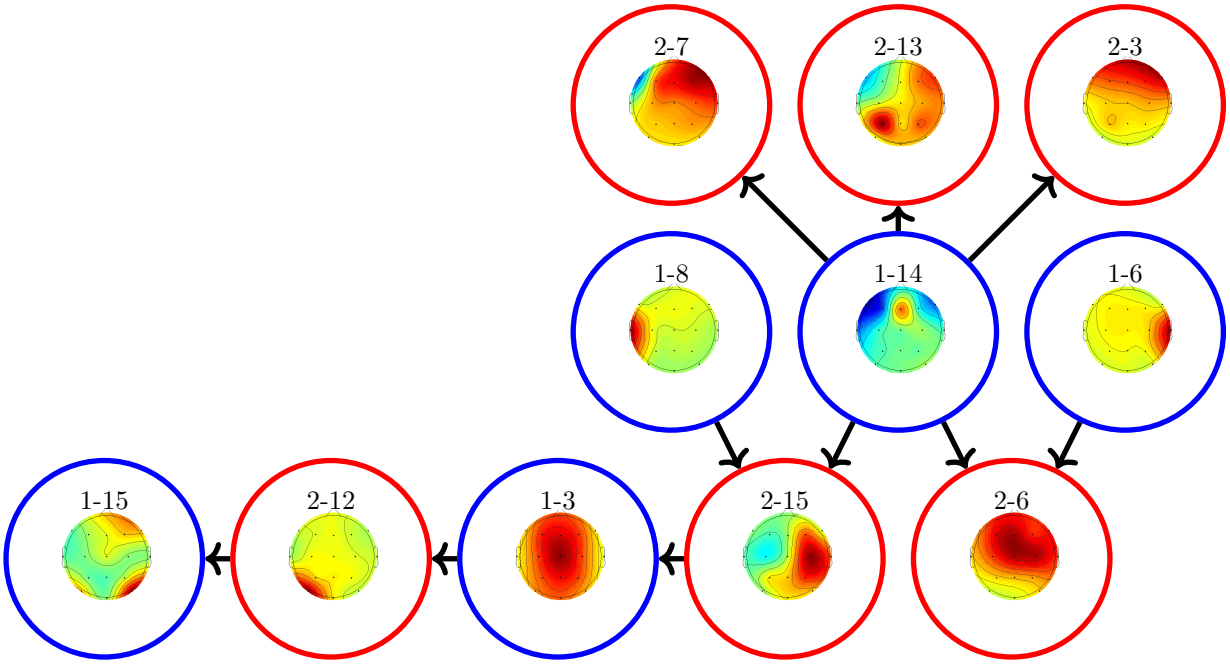


Figure 8.7.: Non-isolated connections with significant differences between init1 and init2 data of experiment six α-band as a graph. The coloured circles indicate the participants to which these components belong.

(occipital-lateral) to (1-15) (occipital-lateral with an additional fronto-lateral focus). (2-15) could be viewed as a hub relaying the activity of (1-8) and (1-14) to the daisy-chain.

From a network perspective the role of (1-14) is even more central. It has a direct or indirect influence on eight of ten of the other components. (2-6) is one sink of the network. (1-15) is the sink of the daisy-chain connection and therefore (in-)directly influenced by four other components. More than any other component in the network.

The longest (shortest) path length is four, from (1-14) to (1-15). This network cannot be considered a small-world network. A path length of four in a graph with 11 nodes is not exactly short. This is remarkable as neural networks are usually small-world networks, but considering the relatively low number of node in the network this might be an effect of chance.

It is highly probable that a.) this network is even larger (i.e. involving more nodes and more connections), but the analysis failed to recognise the others as significant and b.) that similar graphs have formed during the other experiments as well, but again could not be analytically verified (type-II errors). Committing potentially many type-II errors by avoiding type-I errors is actually a pretty standard trade-off in science. On the other hand, this graph encompasses ten connections. Considering that I controlled the FDR to an α-value of 0.2 this means that the expected number of type-I errors (falsely significant changes) for this graph is lower or equal 2 (by virtue of FDR, see section 5.7.2 on page 54).

8.6.3. Connectivity Over Time

I described the drawbacks a bi-variant connectivity estimate such as PSI theoretically has, previously in section 5.6 on page 47 and in the last section I highlighted how they occur on the basis of real data. Considering to use a second, multi-variant estimate such as the Partial Directed Coherence PDC (see section 5.6.3 on page 52) to compensate this, I realised that PSI and PDC are actually complementary in many aspects (compare section 8.5 on page 96).

One of the main advantages of PDC is that it is multi-variant and should therefore represent indirect connectivity correctly. The multi-variance of PDC comes, speaking figuratively, with a price-tag. As discussed in section 5.5 on page 45 the generation of a MVar model requires a number of data-points per weight in the MVar-matrices and the number of the latter grows quadratically with the number of channels/components involved. The number of data-points is on the other hand limited by the EEG signal's local stationarity. This need for training-data can be reduced by a method called AMVar (Ding et al., 2000; Mullen, 2010). This is based on averaging over epochs, meaning that all repetitions of the data (epochs aligned with the

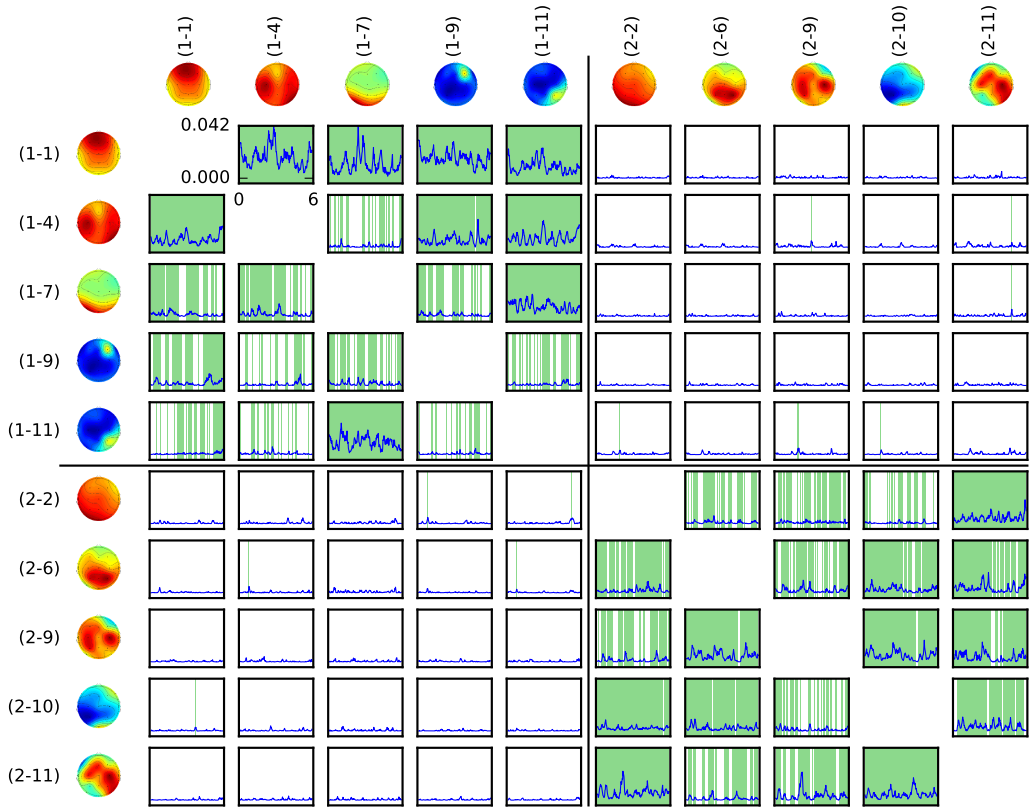


Figure 8.8.: An exemplary PDC analysis (β-band of experiment seven). Components are arranged and named as for the differential PSI analysis. Periods which were deemed significant based on a phase randomisation surrogate distribution and after controlling FDR are marked in green.

onset of the robotic action) are included in the estimation of a MVar model. As a consequence, I cannot apply group randomisation for statistical testing, as I did for PSI differential analyses. As an alternative, I applied phase randomisation to generate a surrogate distribution, again using 2,000 repetitions and again controlling FDR for an $\alpha = 0.2$ for multiple comparison correction.

Even with these measures to reduce the need for MVar training data, I had to reduce the number of components considered in the PDC analysis to five (in one case four) per participant. This opens way to the "common drive" phenomenon (two components are falsely identified to be interdependent, while they are really both dependent on a third component not included into the analysis, compare section 5.6 on page 47). To minimise this risk, components were selected by computing partial coherence an all component pairs and then selecting the components which showed the most coherence with others, following a suggestion of Mullen (2010).[9]

Despite these drawbacks, I still consider the application of PDC on my data worthwhile, as it offers a view on the connectivity which is in many aspects complementary to the PSI results. I computed the PDC on data segments of six seconds starting with the onset of the robotic movement.

Figure 8.8 shows an example for such a PDC analysis. The components are again ordered by participant and the plot is divided into the same four quadrants. The main diagonal (auto-regression) is left out for the same reasons as for the PSI analyses. Each entry contains a plot of the PDC over the course of six seconds. All plots in the table have the same scaling in y-direction. The PDC is usually normalised such that it lies in a value range of $[0, 1]$, however, for this analysis I used RPDC for which this is not the case. In each plot, time series which were deemed significant based on phase randomisation surrogate distribution and

[9] Of course, the components containing artefacts were not included and, thus, the components chosen for the PDC analyses are subsets of the components used in differential PSI analyses.

after controlling the FDR are marked in green. FDR has been applied to the within-participant and hyper-connections separately, as we have to assume that the results of the phase randomisation follow different distributions between these two groups (compare section 5.7.2 page 54).

Several observations can be made, examining this plot: First, the PDC value for hyper-connections is much smaller and gets significant only for very short periods. The within-participant connections on the other hand, have much higher values and are significant for much longer periods, sometimes even for the entire six seconds. This is in-line with my observations on baseline vs. experiment PSI differential analysis for which hyper-connections were significant much more seldom, too. I might, hence, expect to find more significant differences in hyper-connections when contrasting coop and non-coop or init1 with init2 data using PDC. Unfortunately it would be impossible to obtain a stable MVar model using only (roughly) half of the total amount of data (which would be necessary to obtain a PDC value e.g. only for coop or init1 data).

It should be noted that baseline vs. experiment differential PSI analysis contrasts experiment connectivity with baseline connectivity while *this* analysis contrasts experiment connectivity with that from phase randomisation surrogate distribution. Hence, the comparison of the two analyses has to be treated with caution. I assume that a connectivity estimate (either PSI or PDC) will be deemed significant more often when contrasted with a phase randomisation surrogate distribution as compared to contrasting with baseline data. The baseline data will (especially for within-participant connections) already contain some base-level of connectivity, simply because the different areas of the brain are never completely decoupled. A connectivity estimate contrasted with this data, will only be deemed significant if it surpasses this base-connectivity. Whereas the phase randomisation will, by design, destroy any connectivity in the data. In other words, contrasting with baseline data yields a stronger statement than contrasting with phase randomisation data does – not necessarily for every instance, but in general.

One of the advantages of PDC can be seen when examining e.g. components (1-1) and (1-4): PDC here identifies a significant information flow from (1-1) to (1-4) and a significant (although much weaker) information flow from (1-4) to (1-1). In a network of neural sources such constellations are not only possible, but presumably also rather common. PSI could, due to its anti-symmetry property, not identify such recurrent information flow.

We can also observe that the PDC values are subject to a great variability a.) when comparing within-participant components and b.) even more so when comparing within-participant components of a participant with those of another participant or even with hyper-connections. Within one plot one can observe the tendency, that significant PDC values are generally larger than non-significant values. Or, in other words, it is usually peaks in the PDC time curve, which are significant. This is hardly surprising, but does not hold when comparing values from different tables:

Going over all of the PDC plot tables in appendix H starting on page 239 and comparing different plot tables (different participants and bands) the range of PDC values varies greatly. One analysis came up with a maximum PDC value 0.012 another with a maximum of 0.781. One might expect that the results from analyses with lower PDC values are generally less often significant. But no general trend can be observed that for experiments for which lower PDC values were found, these PDC values were deemed significant less often.

In experiment eight similarities can be found between the PDC curves of different connections. For connections which share the same sender (plots in the same row) this only means that this sender undergoes phases in which it exhibits more influence on the receiving components and phases in which its influence is diminished. For other connections with a similar connectivity time course, the reasons for this are subject to speculation.

I assume, however, that these might be indications of a component excluded from the analysis, which acts as a common driver. Figure 8.9 show the scheme for this theory. Component A drives components B,C and D. These components are independent from one another. Component A is not included into the analyses, components B,C and D are. Hence, the PDC would falsely assume an inter-dependence between B,C, and D depending on the latency of the influence A exhibits on each of the three components.

If I now assume that component A (also) undergoes phases during which its influences on B,C and D is higher or diminished, i.e. that A's influence on the three components is not uncorrelated, the PDC value for these virtual dependencies would have a time course which is dependent on the time course of the influence A exhibits on the three dependent components.

Other, unknown, effects might also lead to similarities in the PDC time course, hence, I cannot be sure that these effects are actually caused by a common driver. But it would be an explanation for an intriguing observation, which only occurs in one analysis.

8.7. Discussion

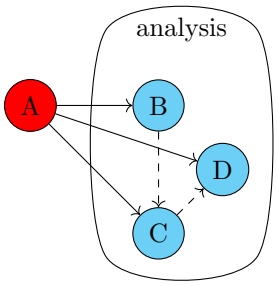


Figure 8.9.: One component (A, the common driver) exhibits an influence on three other components. If it was included into the analysis, its influence would be separated out in the estimation of the connectivity of components B,C and D among one another. Because it is not included, the three components seemingly exhibit an influence on one another and that influence is stronger when the influence of the common driver is stronger (and vice versa). Hence, the PDC time course for these connections is similar.

My primary objective with this study was, to showcase the potential of (BMI) machine-mediated scenarios for neuro interaction research. This study encompassed a complex setting and an elaborate connectivity analysis yielding highly relevant and intriguing results. Such studies will always have a large number of parameters or independent variables one can modify or needs to control. This can be a benefit as well as an obstacle. Regardless, the resulting data sets offer a rich pool of information. On the data sets of this study alone, one could probably conduct a variety of further analyses still yielding significant insights. The true potential of this approach is, however, yet to be determined as future studies explore this new research direction.

The research question of this study was whether or not the need for cooperation among humans has an impact on neural information flow. I could show, that this is indeed the case. I could demonstrate significant changes in θ-, α- and γ-band connectivity. Almost no changes could be shown in β-band. Furthermore I could show that these changes predominantly affect hyper-connections.

I aimed to evaluate whether neural connectivity evolves at all, if participants are interacting remotely. I was, based on the previous findings, confident that this would be the case and the results of the differential PSI analysis confirmed my assumption. I could show a total of 61 changes in connectivity, a good third of them affecting hyper-connections. I was, however, surprised by the prominent role α-rhythms played in this regard. Further research will be required to confirm this prominent role and pinpoint the conditions under which α assumes it.

The differential PSI analysis allows to compare different conditions in a statistically meaningful and computationally workable way. The representation as a graph yields additional insights. However, a high-resolution EEG, allowing for volume source localisation would make the results from different experiments more comparable and the evaluation of these analyses even more expressive.

From a technical point of view, the use of the Shadow Hand robotic system as a basis for the iCusss study worked reasonably well. I knew in advance the robot would occasionally fail to grasp a cube or might not be able to complete a placing action. But these problems occurred rather sparsely. Participants found the robotic system exciting and, thereby, motivating. Beyond this, the reasons for my decision to use the robotic system (good control over degree of cooperation, familiarity with the domain, game-like character, etc.) have proven true.

The PDC analysis has important advantages. In particular, it allows for an inspection of connectivity over time. This is particularly appealing as one of the main advantages of EEG as a neural recording technique is its superior temporal resolution. PDC allows to inspect the neural connectivity at a sub-second temporal resolution in a meaningful way. The PDC analysis and the PSI analysis have properties, which are highly complementary. However, the PDC analysis has high constraints regarding the amount of data, and thereby the number of repetitions needed in order to obtain a stable model. For hyper-scanning data, the additional problem occurs that, by design, such data has double (or more) the amount of channels/components than a single participant study and that the need for training data for the estimation of the MVar model grows quadratically with the number of channels. I would still advocate the use of PSI and PDC as two complementary analyses. However, for future studies one needs to ensure that sufficient training data is available even if the repetitions are divided into conditions. This way randomisation of data groupings and therefore a statistically meaningful, differential PDC analysis would be possible. This would make the results of PSI and PDC much more comparable and one could much more relate the analyses' results with one another.

The data set probably allows for many further analyses, many of which might not even be from the field of hyper-scanning and neural interaction research. From this experiment several other research direction might

be pursued, such as the impact of different properties of a robotic system involved, if a gradually varying cooperation can be correlated with connectivity estimates on the data, if hyper-scanning with more than two participants could yield any new insights or how a parallel cooperation (in contrast to only one participant acting at a time) might influence the connectivity. Furthermore different hypotheses about the role of the different frequency bands might be verified with proper study designs.

In summary, I could achieve my main objectives with this study. But I also learned a series of lesson for future studies in this direction. This is probably only natural, when pursuing a whole new direction of research.

9. Discussion

The research presented in this thesis seeks new ways in neural interaction research. By combining the formerly distinct research topics of BMIs and hyper-scanning, a new approach to research on the neural foundations of human interaction has been laid out. The feasibility of the approach has been demonstrated in the first study. The potential of this approach has been shown-cased in the second study. The experiments conducted in the context of this study, resulted in highly complex data sets which allow for a wide range of analyses – even more than those conducted for this thesis. I illuminated the role of different brain rhythms and the importance and role of hyper-connectivity, as differences between different sets of conditions predominantly affected hyper-connections. This thesis could, however, not even remotely exhaust the potential of this new approach.

In this chapter I want to conclude with some general thoughts and remarks about my results and the techniques used and give some ideas for possible future directions. In section 9.1, below, I will discuss the potential of machine-meditated interaction and the approach to combine BMIs and hyper-scanning. More generally the role of hyper-scanning for computational neuroscience will be the topic of section 9.2 on page 114. Finally in section section 9.3 on page 115 I will sketch some promising future projects.

9.1. Interaction, Connectivity and Machine Mediation

This thesis contributed to the field of neural interaction research, whereby its main methodical novelty is in studying machine-mediated interaction. Previous studies showed hyper-connectivity in a variety of direct interaction settings, e.g. Astolfi et al. (2011a) studied interacting pilots, Babiloni et al. (2007b) studied interaction during a card game, Sänger et al. (2013) studied guitar duets and Dodel et al. (2011) studied team performance in a combat simulation. Dumas et al. (2010) is to my knowledge the only hyper-scanning study featuring a form of machine-mediation, however, with a transparent mediator and the mediation was a means to an end, not the object of study.

Pursuing the approach of machine-mediation in hyper-scanning, this thesis first yielded two very basic results: It showed (among others) that a.) hyper-connectivity still occurs when interaction is machine-mediated (particularly including non-transparent, robotic mediators). This confirms the general feasibility of the approach and allows us to follow a new and different path to approach the neural foundations of human-human interaction. And b.) that hyper-connectivity still emerges when participants are remote to one another and to the place at which their actions are executed. This allows future experiments to precisely control the interaction level information flow between participants.

Connectivity, and particularly hyper-connectivity, is therefore robust to machine-mediation and remoteness of action/interaction. Furthermore, the first experiment showed pretty strong hyper-connectivity despite the limited interactivity of the task, which was even diminished by the lack of control of the participants over the BMI. The fact that connectivity is robust against effects which can be assumed to influence the degree of interaction, might raise suspicions whether or not the emergence of hyper-connectivity is linked to interaction at all. The results of the second experiment refute these suspicions, as I could show significant differences in hyper-connectivity between coop and non-coop condition. These significant changes are, however, relatively few in numbers, namely I found 16 significant changes in hyper-connections over nine pairs of participants and four frequency bands (compare section 8.6.1 on page 97). And particularly not as many as for the second data grouping (init1/init2), for which I found 73 significant changes affecting hyper-connections.

I could show that cooperation has an impact on the emergence of hyper-connectivity, but the results also showed that it is unlikely to be the only factor. Particularly as my own results show that other data groupings affect hyper-connectivity to a much higher degree. This second data grouping hints to assignment of roles and/or the direction of interaction level information flow as other factors.

That knowledge (or belief) about the interaction (partner) plays a mayor role in the emergence of hyper-connectivity was also the outcome of the study by Astolfi et al. (2014) (see section 3.2 on page 26). Connectivity between two participants was established as the participants were engaged in the same task, but was diminished when participants believed that their partner was a machine rather then a human. The results

from this study and my own results complement nicely in certain aspects: In both, my and Astolfi et al's, studies a certain level of hyper-connectivity was retained when the participants only perform the same task simultaneously (but did not cooperate). Effects such as co-representation and theory of mind could play a role in this phenomenon. A different/alternative factor influencing the emergence of hyper-connectivity might be whether the two partners operate in a shared space. The sheer presence of a partner, however, does not seem to suffice for the emergence of hyper-connectivity or else I would have found hyper-connectivity in the solo-condition of my first experiment.

A concrete comparison between connectivity in interaction with and without machine-mediation is beyond the scope of this thesis. This would require a series of carefully designed experiments encompassing an interactive task which is meaningful with and without machine mediation. However, from the experiences from this thesis, I would assume that the differences in (hyper-)connectivity are rather subtle, not to be detected on the level of significant connections and their numbers, but maybe only on the level of connectivity graph analysis (compare section 8.6.1 on page 97).

9.2. Hyper-scanning in Interaction Research: Its Role and Perspectives

The huge chance of neural recordings is to observe human-human interaction "where it happens". Without neural recording techniques, interaction has been studied on the basis of observing one or more of its modalities (gesture, facial expression, speech, behavioural measure of any kind) maybe even integrating observations of several modalities. Theoretically, all this information is integrated at one place: The interactant's brain. More than that, some foundations of human interaction which happen "internally" and are never actually expressed in some kind of behaviour might only be accessible by the use of neural recordings.

Recently the notion that all information about how the interaction emerges can be found in an interactant's brain has been challenged (Dumas, 2011; Sebanz et al., 2006; Schilbach et al., 2013). Researchers argued that the interaction between humans is a holistic process which can only be understood by monitoring (and analysing) the brain activity of all participants involved. Hyper-scanning was born. If we could fully record the brain activity of all interactants we would have, theoretically, all information about how humans implement interaction.

That is, of course, pure theory. In practice ...

a ... we do not have any means to obtain such a full description of the brain activity. Any recording technique only either describes the activity of populations of millions of neurons at once, i.e. are rather coarse or can only monitor pretty small portions of the brain. We either have a detail-poor image of the entire cortex or a detailed image of small portions of the cortex. Figuratively speaking this could be called the Uncertainty Principle of Computational Neuroscience. Other than for Heisenberg's Uncertainty Principle, neuroscientists may, however, hope that this uncertainty might be overcome some day.

b ... the information is deeply encoded in a high-dimension data set. Without sophisticated data analysis the knowledge gain from such a recording would be minimal, even if we had a complete recording.

Therefore, both approaches, neural hyper-scanning and behavioural analysis, have a function in our endeavour to understand how humans achieve their most remarkable skill with such ease and grace. Future research might emphasise even more the correlation of behavioural data with neural recordings.

Additionally it is widely recognised that significant differences exist in the neural activity between humans. E.g. although a P300 can be shown in virtually every healthy human, the form, amplitude and timing may differ substantially. Even more, these features change within the same subject over time. The neural activity shaping human interaction is subject to similar variability. For those reasons, many connectivity studies remain on a descriptive, per-participant level. Some approaches to overcome these limitations have recently been proposed. From counting significant connections (as I did), over graph analysis measures (e.g. Astolfi et al. (2014) or Sänger et al. (2012)) up to approaches to identify the connectivity fingerprint of a certain condition (often employing techniques of machine learning and pattern recognition, e.g. De Vico Fallani et al. (2010)). Some of these have been applied to the results presented in this thesis, but only counting the number of significant connections revealed new in-sights.

Computational neuroscience and particularly neural interaction research is still in the state of a proto-science, where methods and terms have not yet been fully established and are subject to constant debate. This state can, however, only be over-come by continued work on this field, even when many results will

need to be revised once a systematic has been established. I, therefore, strongly advocate the utility of this research.

I would like to stress one particular aspect of this lack of a "gold standard": Hyper-scanning and connectivity analyses generally result in rather complex data. This complexity calls for a most sophisticated visualisation. In my opinion the question how this data can be visualised/plotted in a comprehensive and expressive manner is of a similar importance as questions about feasibility of connectivity estimates, forward modelling, ground truth and others. Yet it has received pretty little attention in the scientific discourse, so far.

In the analyses of the iCusss experiment data I tried to apply two different connectivity estimates whose features are quite complementary: PSI and PDC. This worked pretty well, I could perform analyses featuring a time course of connectivity values and, by collapsing the time dimension, show other significant result, e.g. the prominent role of α-rhythm. However, the data hunger of PDC (and other MVar based estimates) is always an obstacle. For experiments which allow for a rather huge number of repetitions/trials the results of the two estimate might be related to a much higher degree than it was possible for this thesis.

9.3. Future Work

Many experiments have been conducted in the field of interaction research. Interaction situations can take various forms, resulting in a huge number of variables one needs to control in an experimental design and subsequently in an exponentially huge space of different conditions under which interaction can be studied. Theoretically most interaction experiment conducted would profit from additional hyper-scanning as it accesses the observed interaction where it happens. For maximal utility, behavioural and neuro-computational results should, of course, not stand side-by-side, but be correlated. Hyper-scanning might not always be feasible as some forms of interaction might interfere with the recording technique. We have, however, seen improvements in recent years in the field of EEG recording regarding obtrusiveness and restrictiveness.

Generally speaking, I would advocate not to limit hyper-scanning to neural interaction research. Finding neural correlates for human-human interaction is of highest relevance for sure, but the potential of hyper-scanning for social neuroscience probably goes above and beyond it.

Speaking more concretely, what could be next step in machine-mediated interaction research relying on hyper-scanning? One approach could be, to no longer restrict machine-mediation to the level of machines representing mutual interests and intentions (as it was done throughout this thesis). One of the advantages of triadic human-machine-human settings is that they allow to control the interaction level information flow rather precisely. Such research could and should probably be embedded in and correlated with the results of comparable non-hyper-scanning studies.

Another interesting aspect is the impact of the type of interaction. In my experiments interaction was never competitive. Although I consider cooperation to be the scientifically more interesting research subject, competitive task solving should not be completely disregarded. One approach here could be to design two types of tasks: One which requires to infer the competitors intentions to win and another for which concentrating on ones own efforts would suffice. Differences in connectivity between these conditions could further shape our understanding of the neural foundations of co-representation, theory of mind and interaction. Finally, an experiment directly comparing (hyper-)connectivity during competition and cooperation might also prove insightful.

A third approach would be to try to achieve different levels of involvement/motivation in the participants. The use of game-like tasks is a common method to induce motivation in participants in scientific studies. A task which can be adapted to have varying degrees of "game-likeness" seems challenging to design, but might yet be possible. Inducing time-pressure would be another common method to induce motivation. Rewarding the participants in some way relevant for them would be a third way. Comparing connectivity in conditions inducing different levels of motivation could further advance our understanding of human interaction.

Another interesting research direction would be team-building. By monitoring the development of neural connectivity and hyper-connectivity between participants who cooperate on a complex task on different days, we might gain new insights of how interaction partners adapt to one another. Will hyper-connectivity increase, the more adapted the participants are to one another and the better they function as a team? Or will hyper-connectivity actually decline the more participants are confident in their mutual performance and the less they need to anticipate their partner's behaviour? These could be questions of great scientific relevance.

Finally, it would be most relevant to compare connectivity during machine-mediated and direct interaction. This would require, of course, a task which can be meaningfully pursued under both conditions.

I would like to conclude by pointing out three aspects which, in my opinion, should be tackled next social neuro science:

- Which are the main factors influencing the emergence of hyper-connectivity?
- How can results of connectivity estimation be visualised comprehensively and expressively?
- How could the correlation of connectivity estimation with behavioural data be improved?

9.4. Conclusion

This thesis demonstrated the feasibility of studying machine-mediated interaction with methods of hyper-scanning and hyper-analysis. Using this approach I could narrow down the conditions under which hyper-connectivity emerges. I could show a dominant role of α-connectivity and I could show the intriguing role of β-connectivity, which seemed insignificant during my analyses at first, but then became highly relevant when examining certain data groupings. The combination of BMI techniques with hyper-scanning and -analysis, is not only appealing, but has worked exceptionally well for me.

Schilbach et al. (2013) called the neural foundations of social interaction the "dark matter of social neuroscience" and I think this is a well-fitting comparison. The nature of actual dark matter in physics remains largely unknown. It cannot be accessed with standard methods of astrophysics and is therefore hard to fathom. What we do learn about it, however, changes our understanding of the universe per se. Analogous assertions can be made about the neural foundations of social interaction. And as astrophysicists never cease to further illuminate the nature of dark matter, so do neuro-scientists.

With this thesis I have laid out a new, alternative path to approach this dark matter. Following this path should allow us to illuminate new aspects of it. It is one more step in pursuing one of the fundamental questions of mankind.

The search for "the dark matter of social neuroscience" (Schilbach et al., 2013) continues.

Index

A. Example for UBiCI Step-Mechanism

The step mechanism of the UBiCI software framework introduced in section 5.8.5 on page 58 allows to follow a sequence of steps needed to achieve a goal. In contrast to the "normal" online mode which runs in a more asynchronous fashion, this mechanism is most often used for offline analysis of EEG data. One very common application for this is the computation of matrices needed for feature extraction and classification of P300 data (PCA and FDA matrices, see sections 5.2.1 and 5.2.3 on pages 39 and 41). For a better understanding of the step mechanism I will explain how this is used to compute the required matrices. Please compare to figure A.1 when reading this section!

Two UBiCI modules are involved here. The top level module has a very simple task: Load the continuous data set and generate epochs aligned with the stimuli presented during the recording of the data. The sub-module does the actual matrix computation. The step mechanism is initially only executed for the top-level module.

Initially only those connections lacking a number in their label are created. Before the first step (step number 0) is triggered, all connections of the top-level module marked with a 0 are created. This is only one: The one connecting the module's `step`-signal to the DataReplayCenter. When the connection is established the module's `step`-signal is emitted for the first time. This causes the DataReplayCenter to send all EEG data, all epoch infos and all decision requests of the data set it is configured with (given in the configuration file, see section 5.8 on page 55). These run through the TimeToSample component which translates the time stamps of the epoch infos to sample numbers within the EEG data stream and trough the segmentation component, which generates data epochs aligned with the epoch infos based on its configuration (e.g. starting $150ms$ before the time stamp of the epoch info and ending $1150ms$ after the time stamp). After the segmentation we no longer have a continuous stream of EEG data, but a series of (possibly overlapping) segments with their respective epoch infos attached to them. Then these data epochs are frequency filtered using a FFT filter. For P300 a typical filter setting would be to filter for the $1 - 12Hz$ band. Additionally the filter could be set to crop the first and last $150ms$ in order to get rid of the filter artefacts introduced by FFT filtering (compare section 5.1.3 on page 37). Now we have an epoch of one second length which is perfectly aligned with the occurrence of the flash on the screen. Then the segments enter the sub-module and there the EEGDataBuffer who stores them and does not send them anywhere for now. The decision requests are intended for use within an online system and are not required for the training. Because there is no outgoing connection for them from the DataReplayCenter they just vanish.

When the DataReplayCenter has finished sending all data of the recorded session it sends a `step_return` to the module. This causes the module to end step 0 and initiate step 1: All connections which are temporary for step 0 (only one: the `step`-connection from the module to the DataReplayCenter) are removed and all connections which are temporary for step one are created. This is, again, just one, which connects the module with its sub-module via a `step`-connection. Then the module again emits the `step`-signal. Hence, the sub-module now receives a `step` signal. This causes the sub-module to initiate its own step mechanism and it will only send a `step_return` to the top-level module when all the step of the sub-module have been finished. When that happens the top-level module has no more steps to perform and the program terminates.

I will now discuss what happens during the execution of the sub-module's step mechanism: First, all temporary connections for step 0 are created. These are the `step`-connection from the module to the EEGDataBuffer and the connection from the data buffer to the PCAMatrixGenerator. Then the `step`-signal is emitted by the module and received by the EEGDataBuffer. This causes the buffer to emit a copy of all data it contains. This is received by the PCAMatrixGenerator due to the (temporary) `eegdata`-connection connecting the buffer with the PCAMatrixGenerator. The matrix generator stores that data for the moment.

When the buffer is finished it sends a `step_return` to the module which then destroys all temporary connections for step 0 (the two already mentioned). This ends step 0. Then the module creates all step one temporary connections. There is just one, connecting the module's `step`-signal to the PCAMatrixGenerator. When the `step`-signal is triggered this causes the PCAMartixGenerator to start computing a PCA matrix (and the corresponding Eigenvalues). When this is done the result is emitted via a `matrices`-connection. The matrices are received by the PCA component which will update its matrices for use on any incoming

data. Furthermore, the matrices are passed to the top-level module where they reach the MatrixWriter component, which (as the name suggest) writes the matrices to a file. Then the PCAMatrixGenerator emits a `step_return`-signal. The temporary connections for step one are destroyed and the temporary connections step two are created. These are a `step`-connection to the EEGDataBuffer and an `eegdata`-connection from the buffer to the PCA and an `eegdata`-connection from the PCA to the FDAMatrixCalculator. When the `step`-signal is emitted, the buffer again emits a copy of all the data it contains. But this time that data does not reach the PCAMatrixCalculator, but the PCA, which now applies the previously computed PCA matrices to that data. The data then is passed on and reaches the FDAMatrixCalculator which stores it for now.[1] When the buffer is finished (and has communicated that via a `step_return`) the module initiates step three (I will skip description of temporary connection destruction and setup from now on). This is actually nothing else, but a `step`-signal which is received by the FDAMatrixCalculator, which then starts to compute the FDA matrices using all data it has received. When it is done, it again send the matrices over the `matrices`-connections to the FDA and the MatrixWriter. In step four the data of the buffer passes through the PCA and the FDA to a TPScoreBasedP300Postprocessing and a component called TPMeanComputation. The TPScoreBasedP300Postprocessing is a component which combines several P300 classification results to improve the overall classification accuracy. The method is described in section 5.2.4 on page 42. The TPMeanComputation uses the results from that to optimise a threshold used by the TPScoreBasedP300Postprocessing for later online runs. This computation is initiated in step 5, which is the last step of the sub-module. The sub-module now informs the top-level module that it is finished via a `step_return`. The top-level module could now initiate further steps, including sending another `step`-signal to the sub-module causing the sub-module to start over. However, in this deployment there are no further steps in the top-level module and, thus, the program terminates.

[1] There are two components which compute different matrices in that deployment. There is no particular reason that one is called a MatrixGenerator and the other a MatrixCalculator. They are actually very alike in terms of their behaviour.

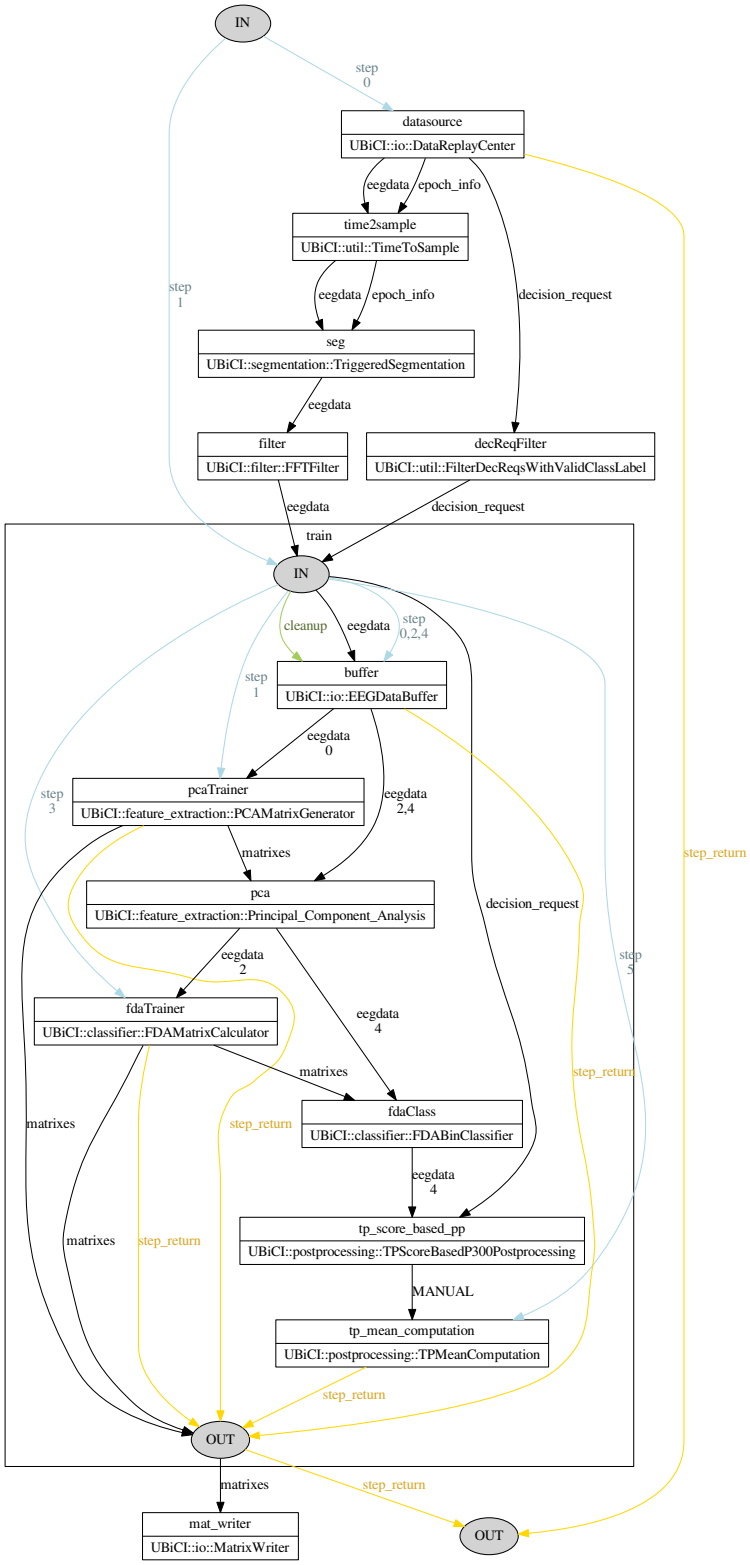
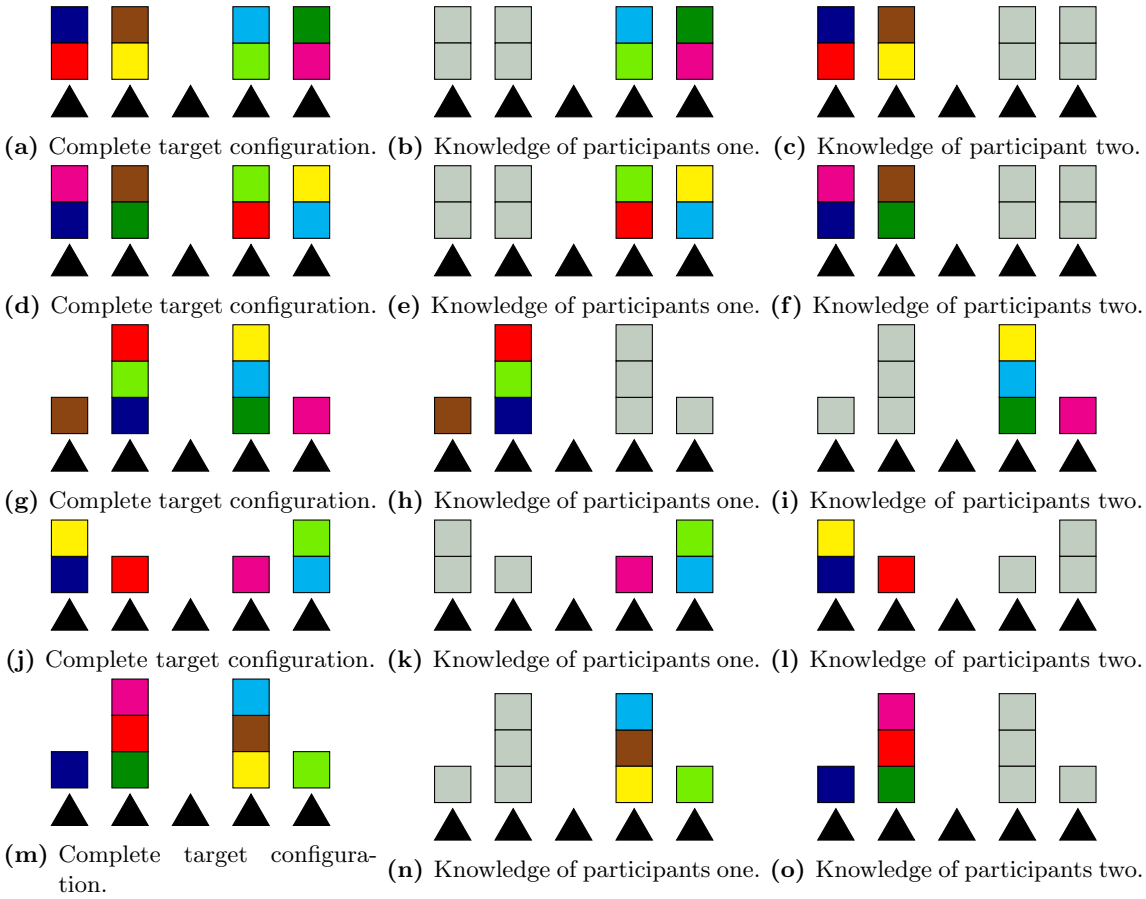


Figure A.1.: The UBiCI deployment used to compute P300 classification matrices. Blue connections are `step` connection, yellow connections are `step_return` connections. Connections with numbers are temporary connections only present during the execution of the steps enumerated. The IN and OUT nodes actually represent their respective modules (connections including them are connection with the module).

B. Target Configuration for Cube-Stacking

Here the tasks the participants had to fulfil during the iCusss experiment are depicted. Participants had to arrange (up to) eight cube in different, pre-defined target configurations. They had partial, complementary knowledge about the target configuration. Some of the tasks were designed to induce a need for cooperation, while others were designed to make cooperation pointless.



(a) Complete target configuration. **(b)** Knowledge of participants one. **(c)** Knowledge of participant two.

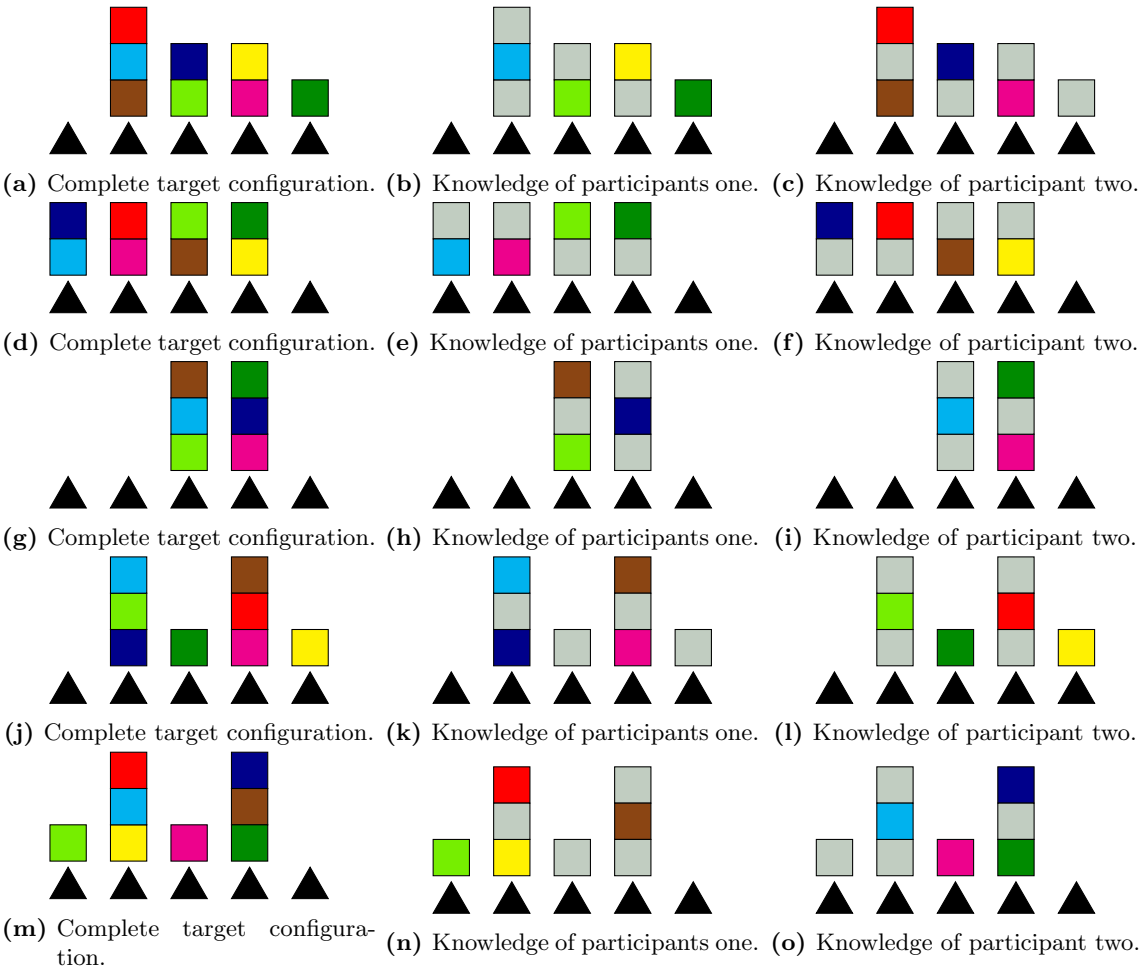
(d) Complete target configuration. **(e)** Knowledge of participants one. **(f)** Knowledge of participants two.

(g) Complete target configuration. **(h)** Knowledge of participants one. **(i)** Knowledge of participants two.

(j) Complete target configuration. **(k)** Knowledge of participants one. **(l)** Knowledge of participants two.

(m) Complete target configuration. **(n)** Knowledge of participants one. **(o)** Knowledge of participants two.

Figure B.1.: The target configurations which required no cooperation among the partners.



(a) Complete target configuration. **(b)** Knowledge of participants one. **(c)** Knowledge of participant two.

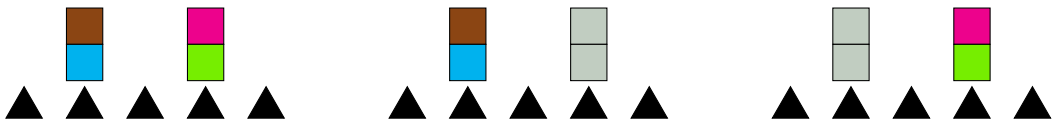
(d) Complete target configuration. **(e)** Knowledge of participants one. **(f)** Knowledge of participant two.

(g) Complete target configuration. **(h)** Knowledge of participants one. **(i)** Knowledge of participant two.

(j) Complete target configuration. **(k)** Knowledge of participants one. **(l)** Knowledge of participant two.

(m) Complete target configuration. **(n)** Knowledge of participants one. **(o)** Knowledge of participant two.

Figure B.2.: The target configurations which require cooperation among the partners.



(a) Complete target configuration. **(b)** Knowledge of participants one. **(c)** Knowledge of participant two.

Figure B.3.: The first target configuration. The data recorded during that trial was not part of the analysis.

C. Network Communication in the iCusss Experiment System

As stated in section 7.5.2 on page 86 three different network communication protocols/middlewares were involved in the iCusss experiment: RSB (Wienke and Wrede, 2011), XCF (Fritsch et al., 2005) and ROS (Quigley et al., 2009). I will briefly describe these here:

RSB: An event-driven middle-ware which is developed at Bielefeld University. The communication is based on informers, each publishing information to a certain scope. Scopes are organised hierarchically using a slash as a delimiter. For instance the scope `/UBiCI/EEG/` is a sub-scope of the `/UBiCI/`-scope. A listener can register to a scope and will receive any information published to that scope and all of its sub-scopes. RSB is easy to use and increases the robustness of the system especially because processes can be started in an arbitrary order and it is even possible to restart processes in the running system.[1] It is, however, still under development and some features planned for later releases would greatly improve the middle-wares utility.

XCF: XCF is a XML-based communication protocol. It has also been developed at Bielefeld University, but is no longer maintained. However, many systems developed at Bielefeld University still use it.

ROS: ROS is widely used for robotic systems. It ss actually more than just a communication platform, but the experiments of this thesis use only its communication capabilities. Shadow company uses it for its products and contributes to its development.

[1] RSB ensures that all connections are re-established. Whether the processes using those connections can handle the interruption in communication is, of course, up to the programmer.

D. Discussion of user_task XCF-Message

In section 7.5.5 on page 91 I described that the classification results from the P300 classification used during the iCusss experiment were translated to `user_task` XCF messages, which triggered the execution of the task described by these messages.

Such a `user_task` looked like the following

```
<user_task armPref="targetpos" ... type="put" ... xcfIP:serial="0" ...>
  <object cube_class="browncube" ... probability="0" shapePreference="b" timestamp="1234">
    <position x="299.62" y="429.53" z="66.08"/>
    <aabb xmax="354.69" xmin="248.61" ymax="464.16" ymin="376.79"
          zmax="87.20" zmin="24.20"/>
    <size points="417" x="703.87" y="588.10" z="166.04"/>
  </object>
  <target pointingProb="0" position_num="4" probability="0" relation="at" timestamp="1235">
    <position x="280" y="150" z="29"/>
    <euler pitch="0" roll="0" yaw="0"/>
    <approach angle="-1.57" x="0" y="14" z="3"/>
    <aabb xmax="316.5" xmin="243.5" ymax="186.5" ymin="113.5" zmax="65.5" zmin="-7.5"/>
    <size x="73" y="73" z="73"/>
  </target>
  <STATUS origin="Handler" value="initiated"/>

</user_task>
```

I will now cover the most important entries in this structure:

armPref: Was actually introduced newly for this system. I already mentioned, that the hands were not yet capable of performing a hand-over, i.e. grasping an object with one hand, then hand it over to the other hand which then would place it. When both hands are capable of grasping a certain object there are different possible heuristics to decided which arm to actually use. For the system and the planned experiment, we wanted to have as much of the robot's working space available as possible. Therefore, we decided to make the choice depended on which hand can best reach the target position. Other values for this attribute would lead to the use of different heuristics and thus to other system behaviour.

cube_class: A classifier tried to identify each object isolated by the 3D point cloud segmentation, based on proto-types of the objects which were given in a database. The result of this classification is given in this attribute.

shapePreference: Was also newly introduced when it was decided that I wanted to use cubes for the experiment conducted with this system. As the robotic system and its performance are not studied in this thesis, a bit of domain knowledge was introduced into the robotic system (which for a robotics-centred study one would want to avoid) by hinting the super-quadrics fitting which shape the object has. In this case `b` for box shape.

object::aabb: The bounding box of the object. This bounding box is aligned with the axes of the coordinate system rather than being minimally fitted around the object.

object::size: The size of the object, first in points in the point cloud delivered by the 3D camera and then in millimetres. The cube are 73 mm in each dimension. However, the 3D camera could eventually only see parts of each cube's surface due to occlusion or the low 3D camera resolution. Hence, the given size does no resemble a cube. This is why superquadrics fitting is that important: It allows to give a reasonable guess about what the system cannot see of a given object's shape.

position_num: The predefined target positions are enumerated. This is the number of the position. This information is currently only used by the BMI, but not by the robotic system.

relation: The robotic system is capable of placing objects in relation to other objects (such as "in front of an apple" or "in a basket") rather than at an absolute position. This attribute describes the desired relation of the object to the target. As the target is a position (given in 3D coordinates) the relation is `at`.

target::euler pitch: The rotation to perform with the object. This was always 0 for all values for this system. I do not want to rotate the object, currently. If this is to be introduced one day, this could be done using a Motor Imagery based-BMI (see section 2.2.5 on page 17) which would make the overall system a Hybrid-BMI.

target::approach angle: This was also newly introduced for this system. I already covered that the robotic system does not yet have collision avoidance. It could potentially decided to approach the target position from any direction. I wanted to be able to pile objects one on top of another. If there would be a pile of objects at position A and the next object would have to be placed at position B (next to A) the robot could potentially decide for a trajectory which would lead through the pile at position A destroying that pile. By giving a defined approach angle this can be avoided. When collision avoidance is implemented in the robotic system this attribute can be removed.

target::position/aabb/size: I already mentioned that the robotic system would be capable to place an object in relation to another object rather than at a given target position. Hence, the <`target`>-tag needs to be capable of holding information to represent such an object.

STATUS: When generating the `user_task` I want to initiate the described action. Hence, I set the `STATUS` to `initiated`. The robotic system would update the status accordingly later on. Possible other statuses would be `accepted`, `rejected`, `completed` or `failed`.

E. Differential PSI: Robotic Action vs. Baseline

These tables show the results of the differential PSI analysis contrasting experiment with baseline data. For details on the interpretation of these plots, please refer to section 8.6.2 on page 102.

E.1. Experiment One

Three significant changes affecting connections in each of α- and θ-band and one in γ-band. All significant changes in θ-band are hyper-connections. All other connections are within-participant connections.

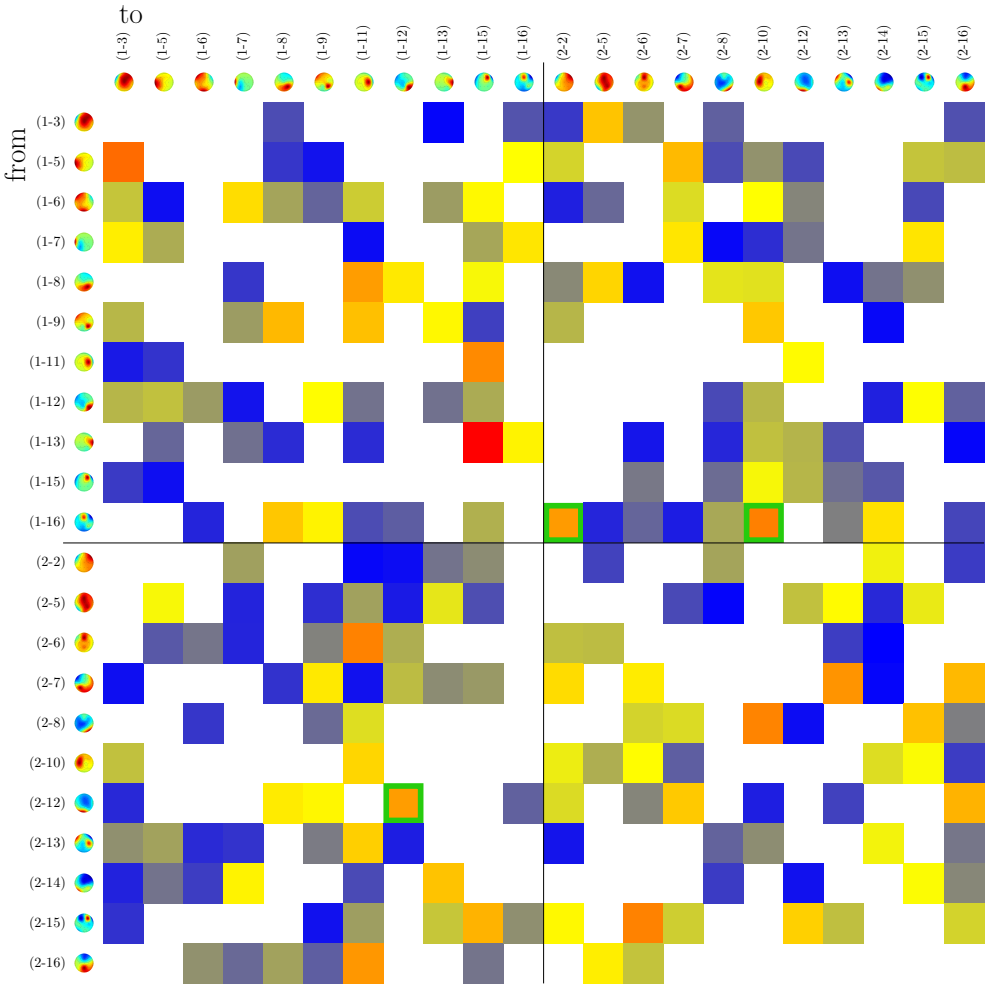


Figure E.1.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment one and θ-band. Three hyper-connections with significant changes were found. One connects a lateral component of participant two with a parietal component of participant one. The others connect a narrow frontal component of participant one with a fronto-central component and a central component extending along the sagittal axis of participant two. Hence, all three connections are hyper-connections.

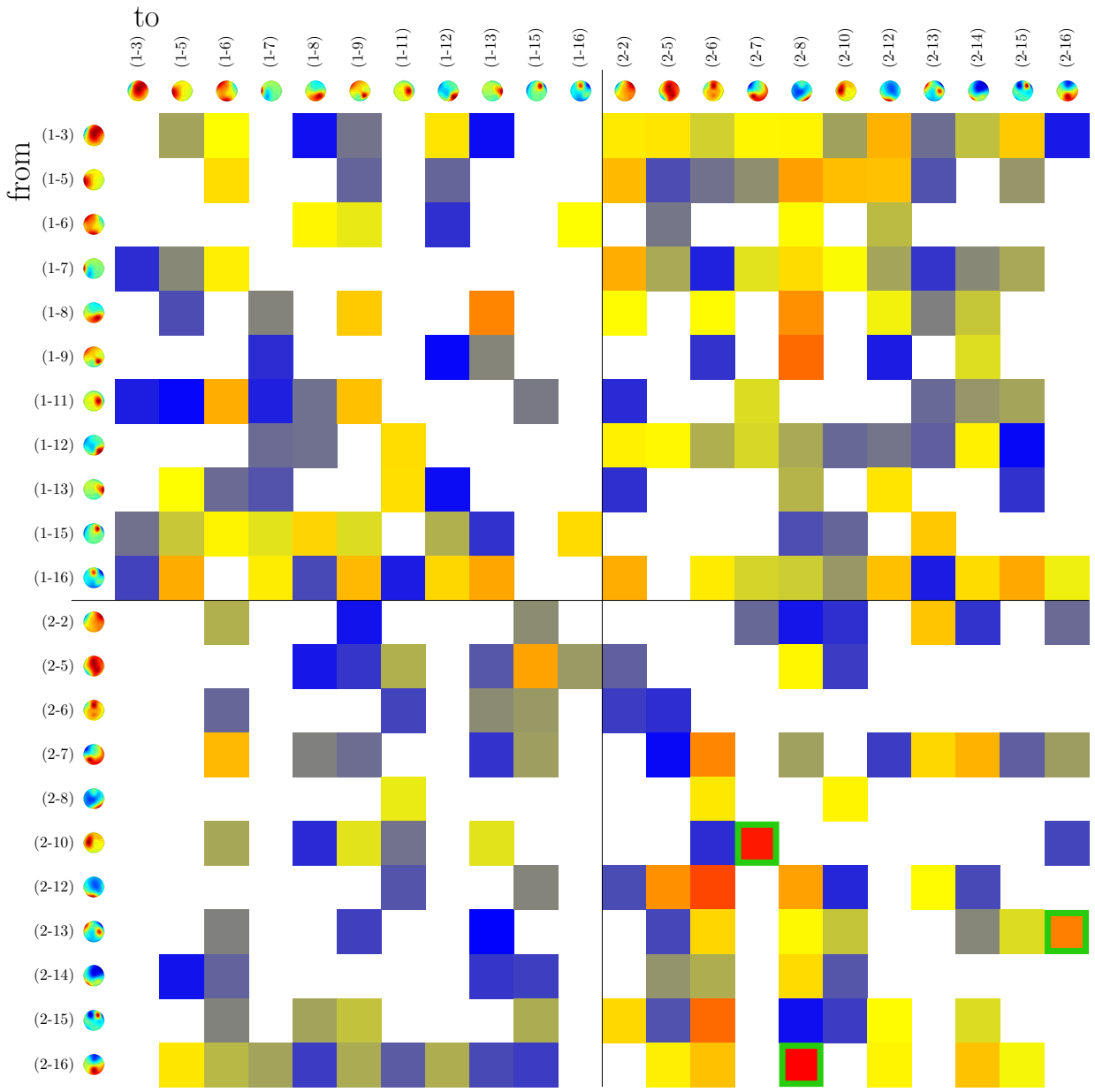


Figure E.2.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment one and α-band. Significant changes for three within-participant connections have been found for participant two. One is connecting a fronto-central component with a broader component roughly running in sagittal direction from frontal to occipital regions. The second connects an occipital with a fronto-central component. The third is connecting a component with lateral foci in the region of the primary motor cortex on both sides with a very broad occipital-lateral component.

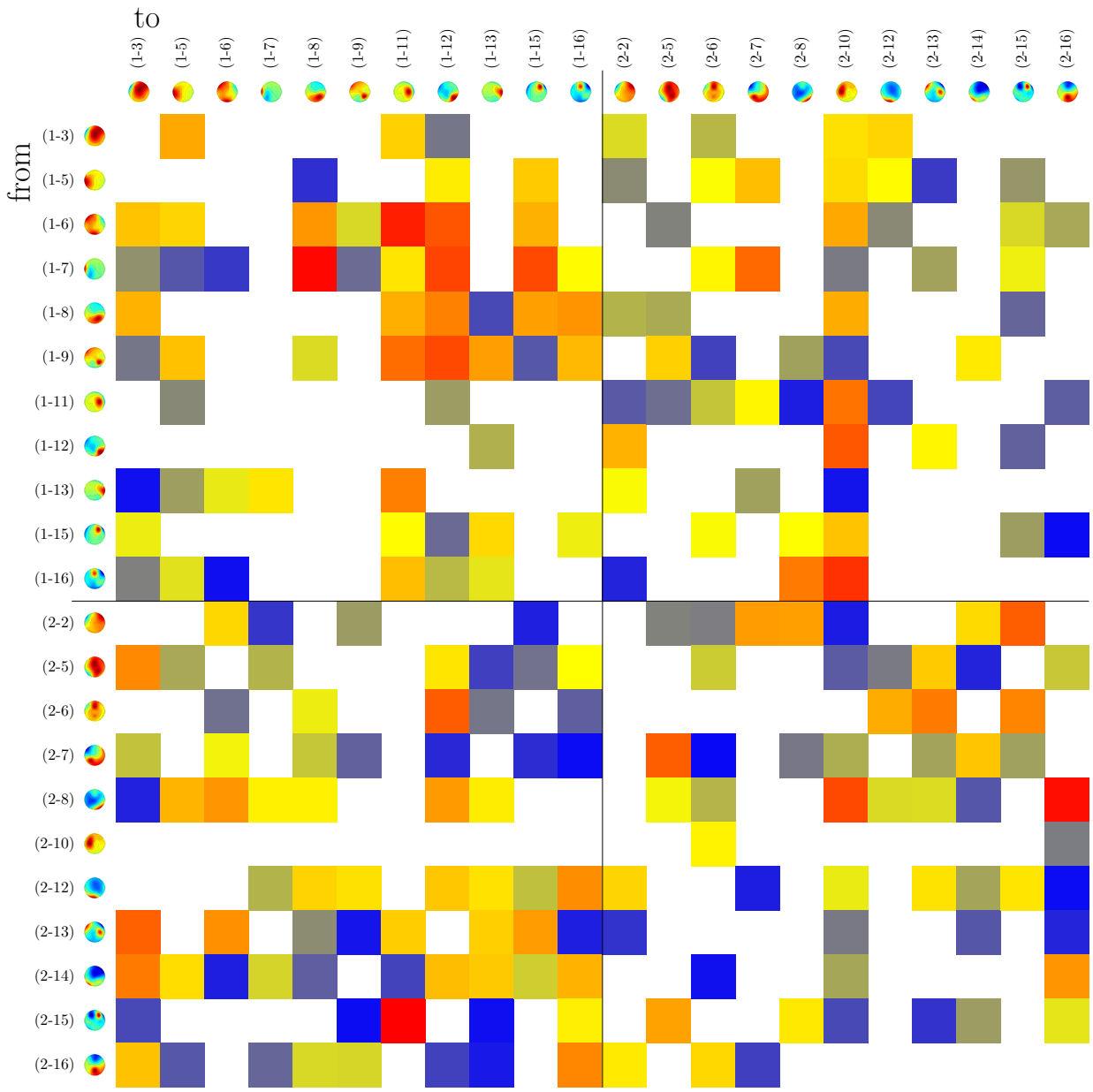


Figure E.3.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment one and β-band. No significant changes in connectivity or remarkable distribution of connectivity values.

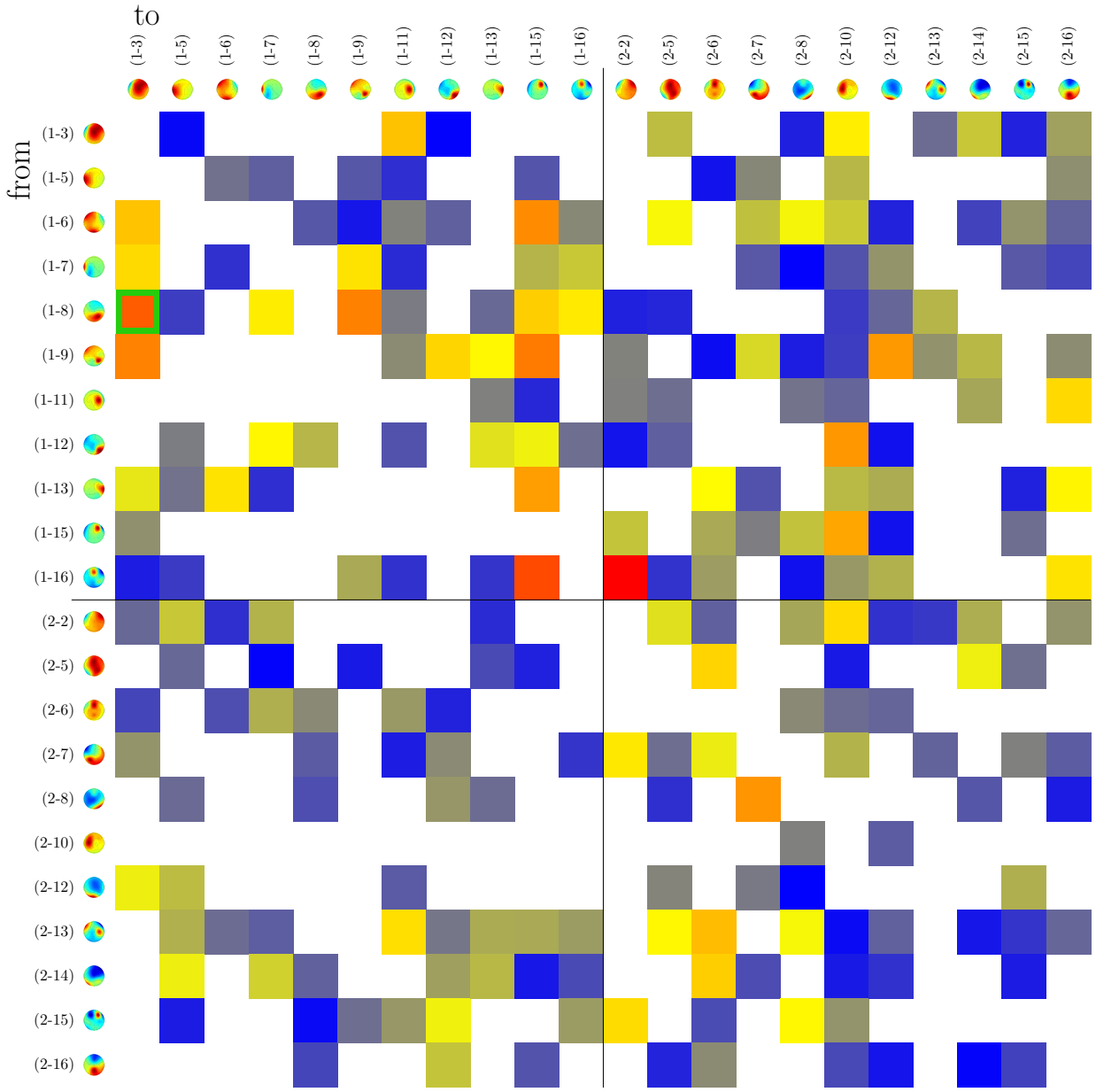


Figure E.4.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment one and γ-band. One significant change affecting a within-participant connection from a parietal to a fronto-central component in participant one.

E.2. Experiment Two

Sixteen significant within-participant connections, all in α band have been found.

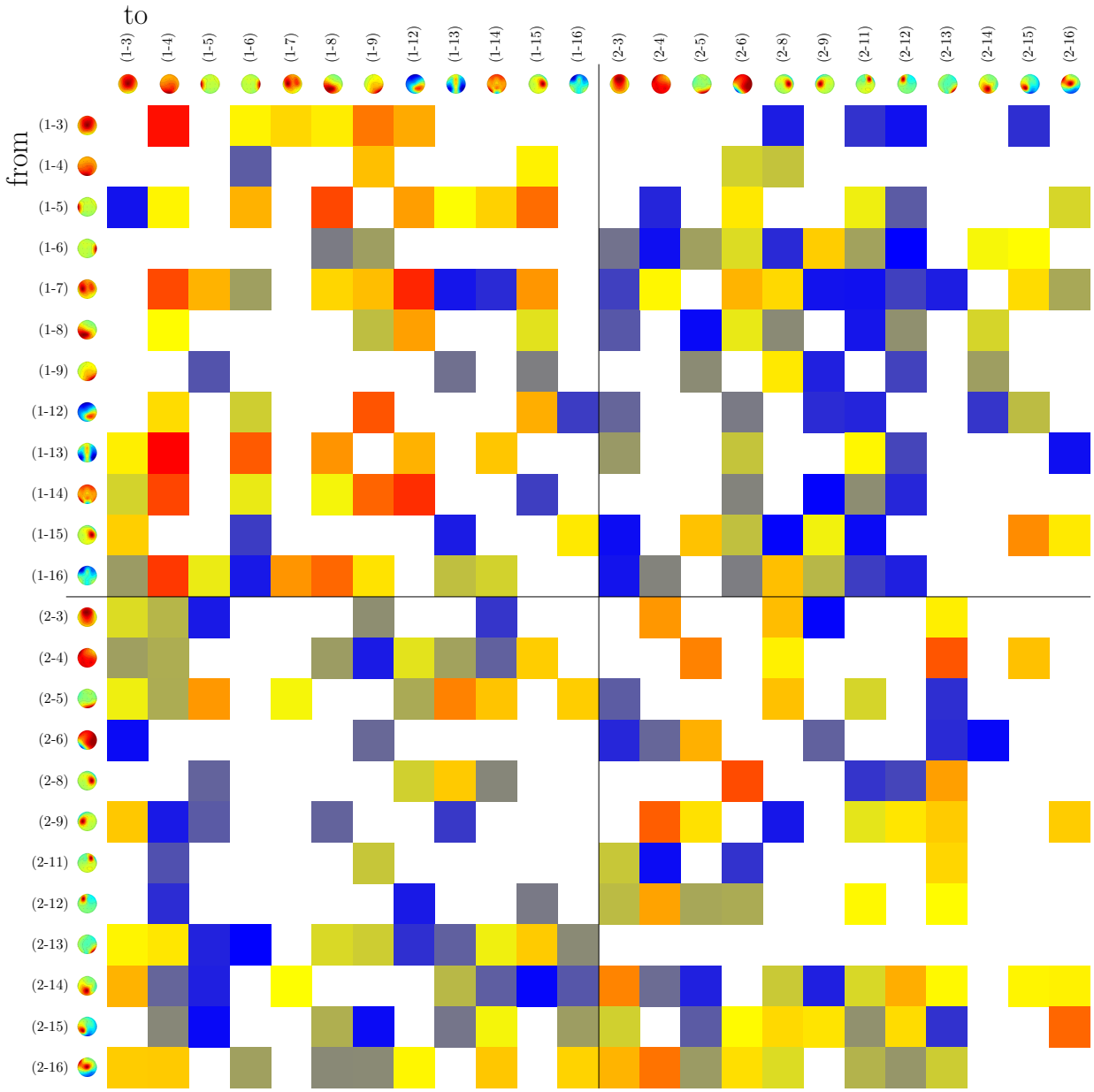


Figure E.5.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment two and θ-band. No significant changes in connectivity or remarkable patterns.

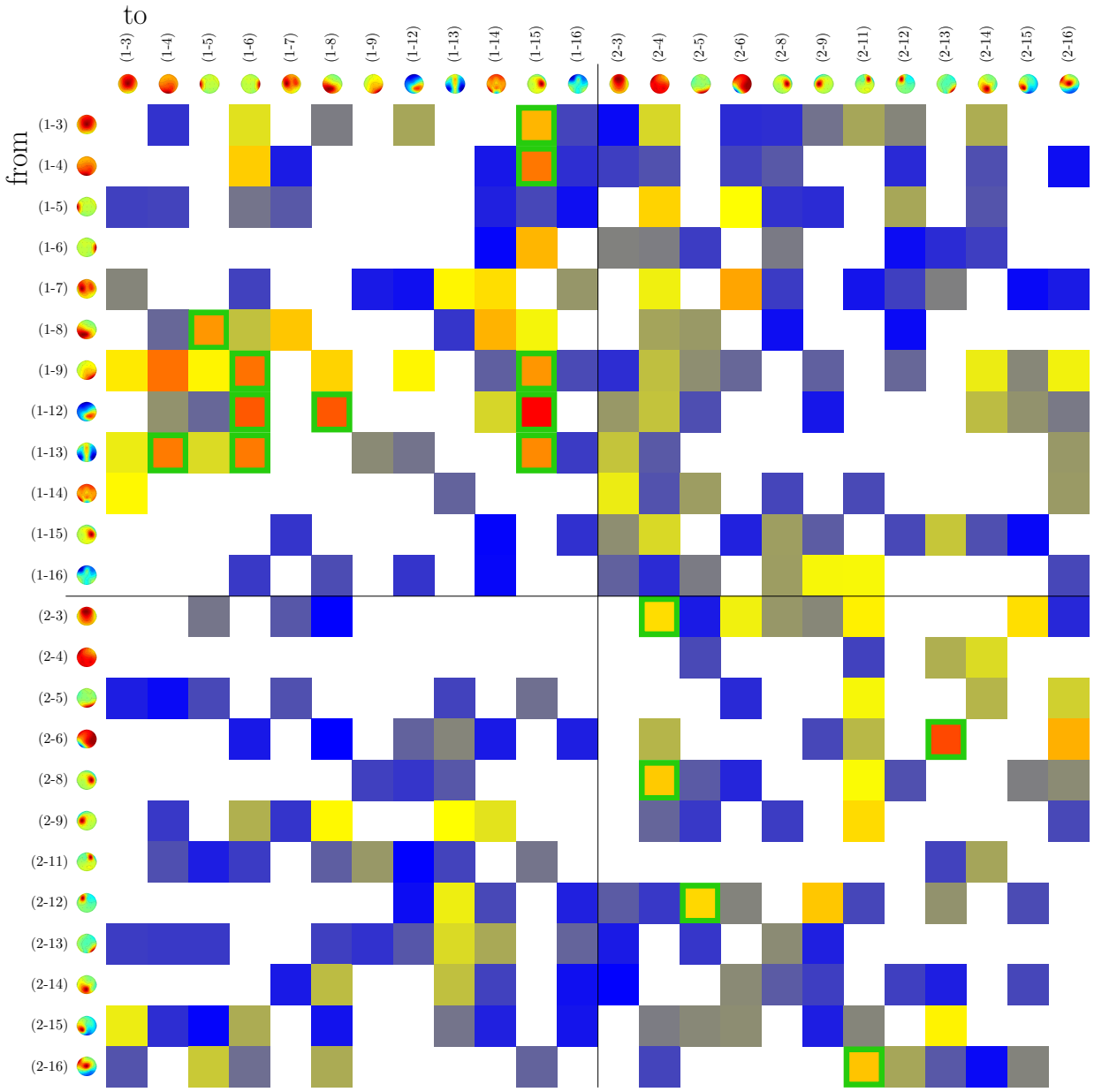


Figure E.6.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment two and α-band. Significant changes were found for 11 within-participant connections for participant one and five significant changes affecting within-participant connections of participant two. Remarkable is especially a lateral component of participant one which acts as the recipient for almost half of the significant connectivity changes of that participant.

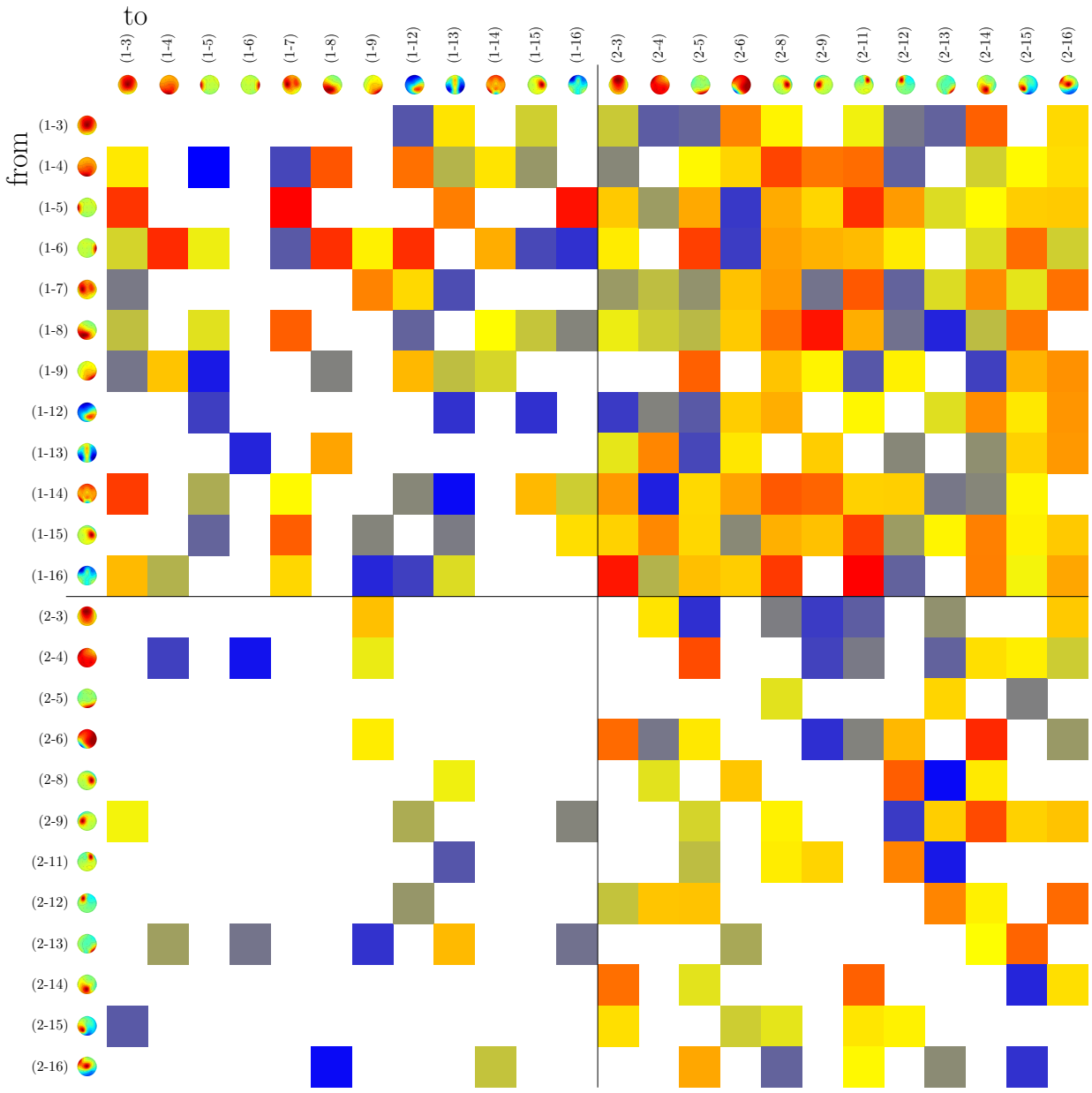


Figure E.7.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment two and β-band. No significant changes in connectivity, but a clear bias among hyper-connections favouring participant one in the role of a sender.

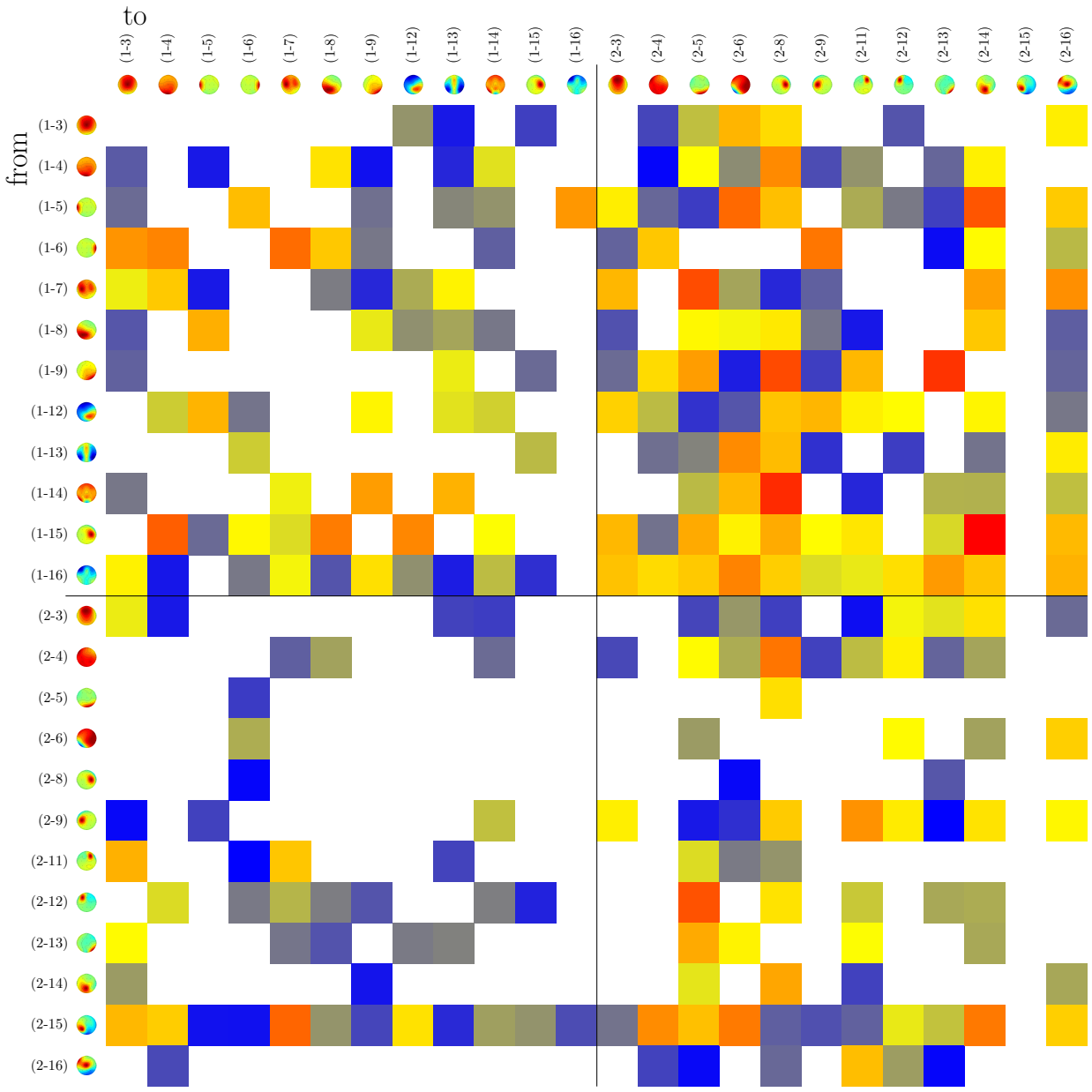


Figure E.8.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment two and γ-band. No significant changes in connectivity and no remarkable biases.

E.3. Experiment Three

Three significant changes affecting within-participant connections were found in α-band and one significant change in a hyper-connection in γ-band.

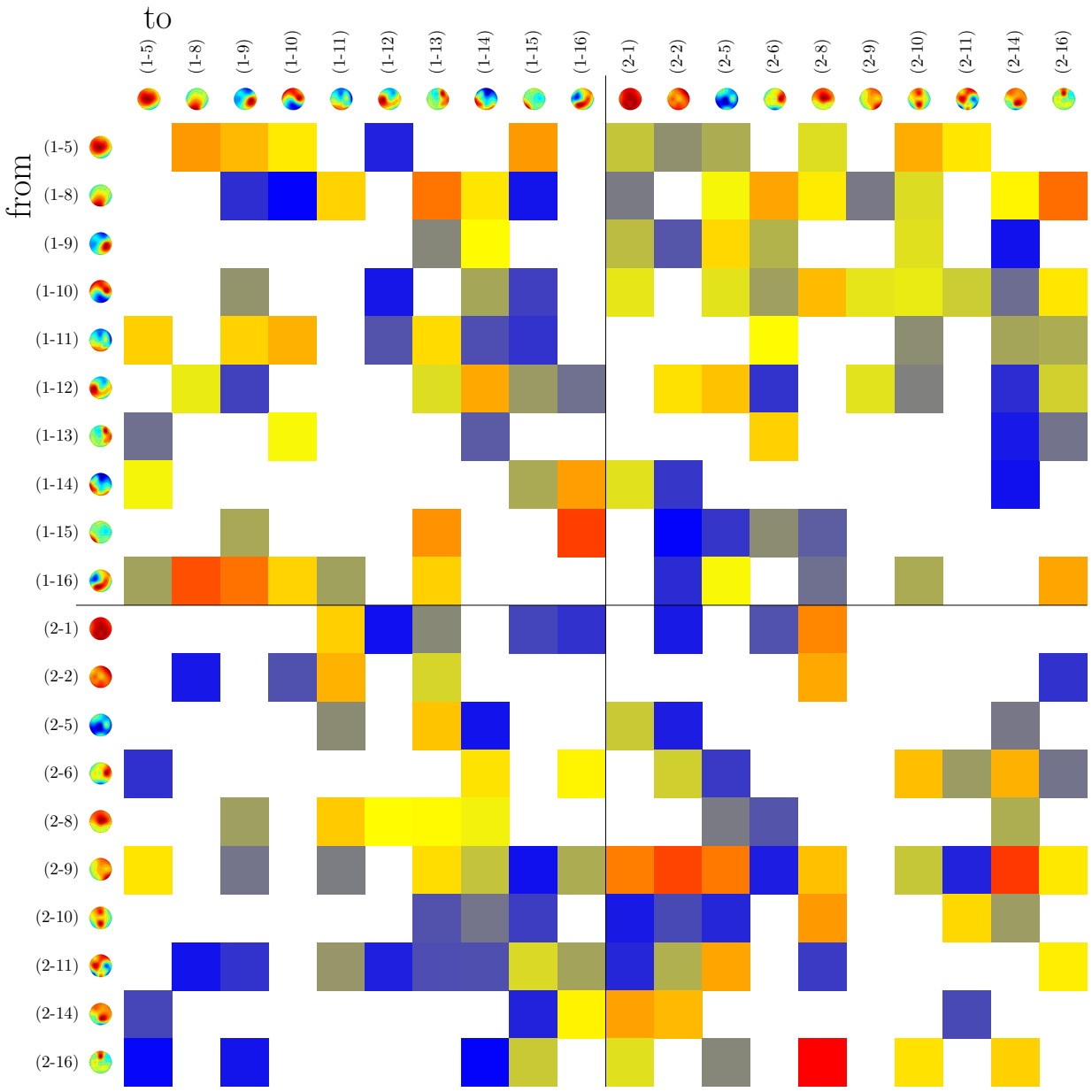


Figure E.9.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment three and θ-band. No changes in connectivity or remarkable patterns in the connectivity values.

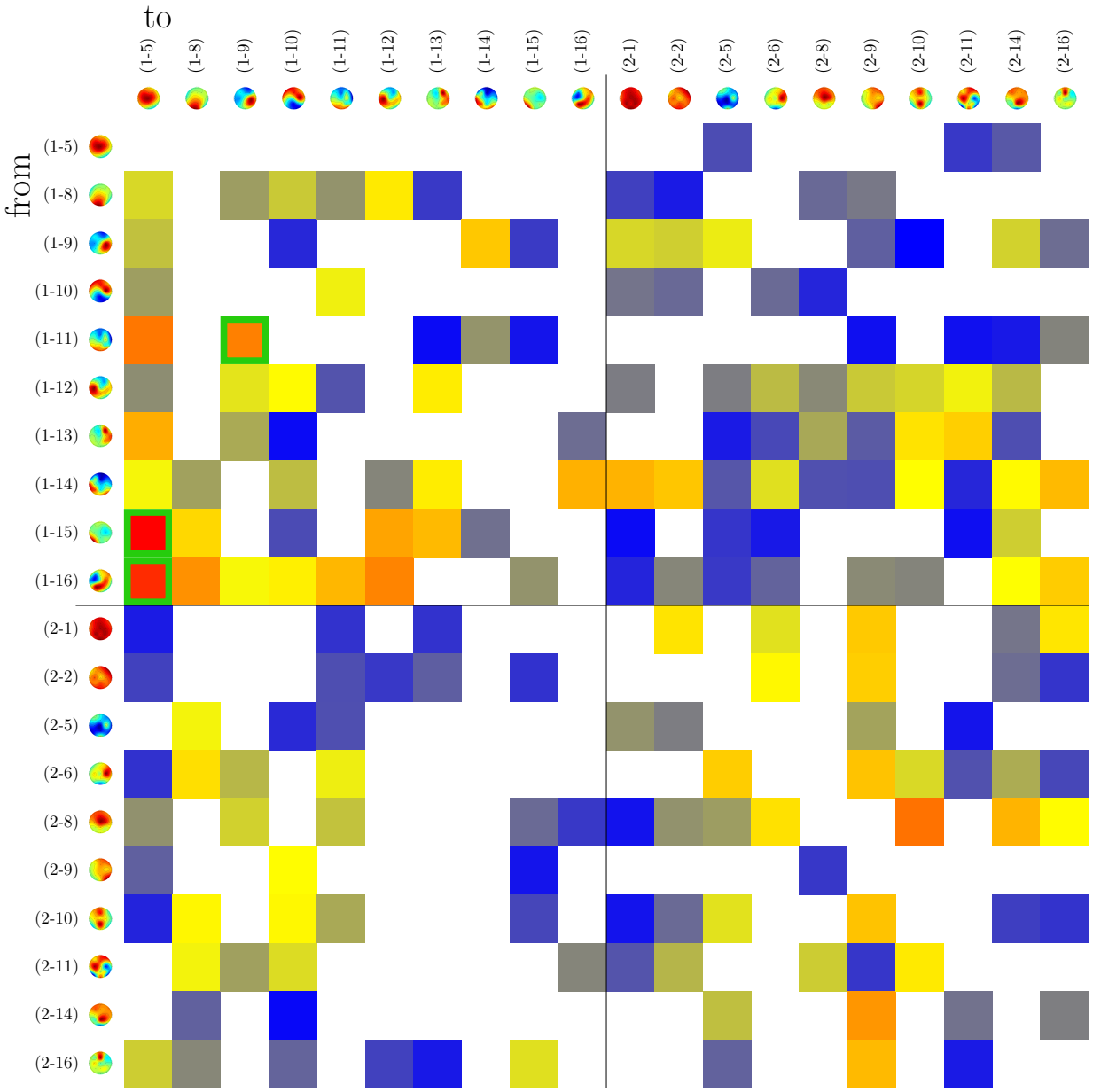


Figure E.10.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment three and α-band. Three within-participant connections of participant one underwent significant changes. A broader central component is being influenced by an occipital-lateral component and a component whose focus extends from a parietal-lateral region to a fronto-contra-lateral region. Furthermore, an occipital component influences a lateral component.

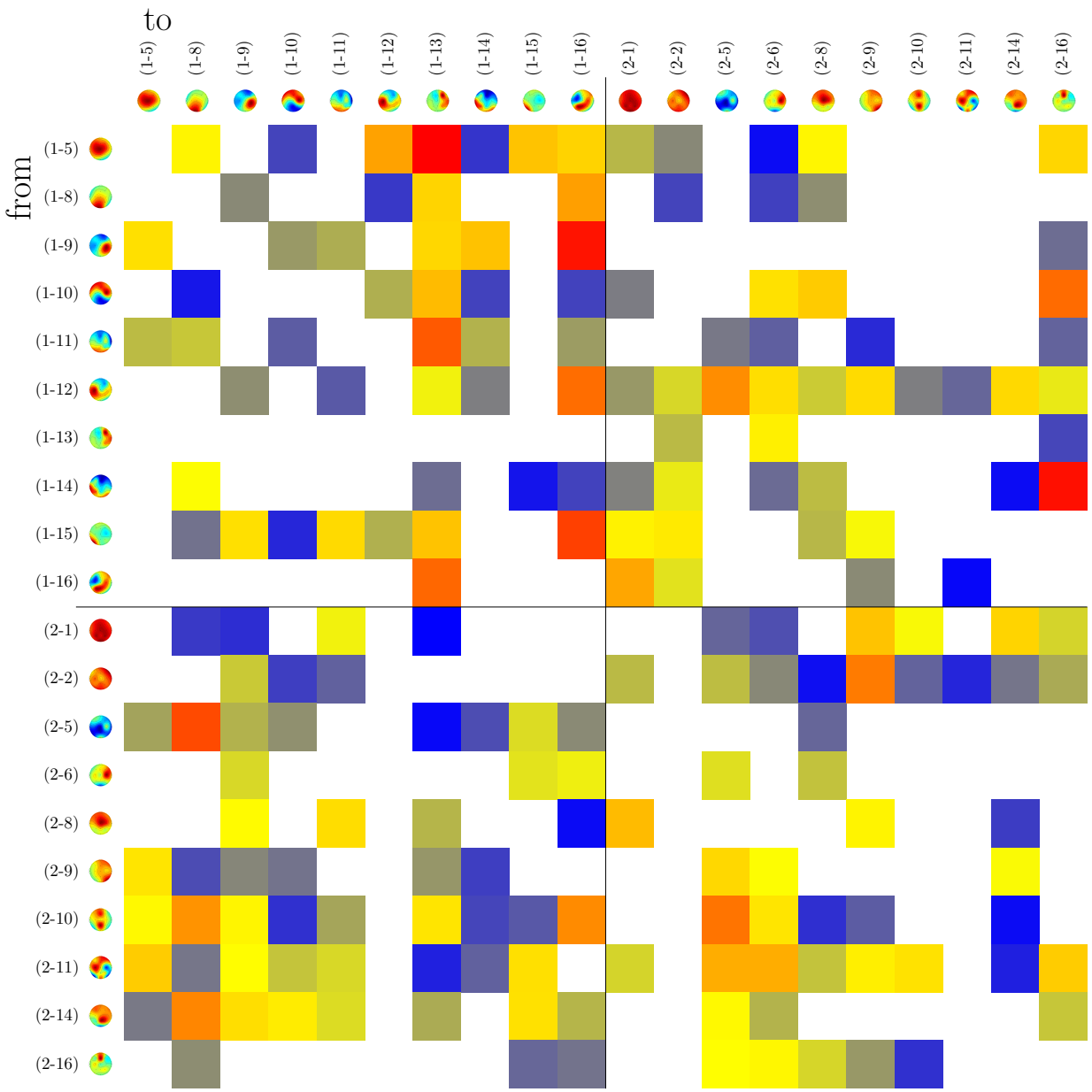


Figure E.11.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment three and β-band. No significant changes of connectivity or remarkable patterns.

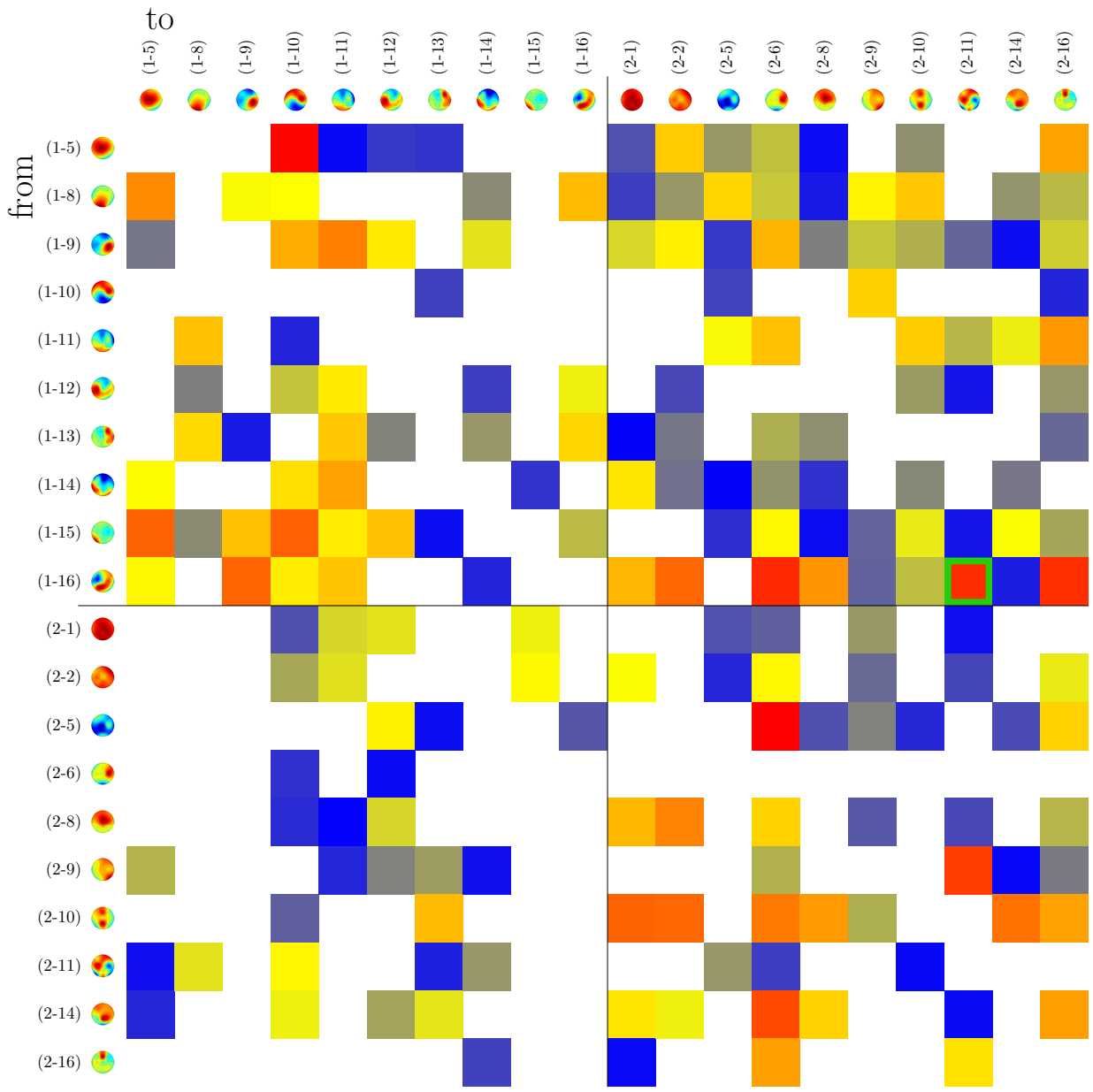


Figure E.12.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment three and γ-band. One significant change was found affecting a hyper-connection which connects a component of participant one encompassing parietal, lateral and frontal regions to a component of participant two focusing on frontal and central regions.

E.4. Experiment Four

One significant change in a within-participant connection was found in β-band.

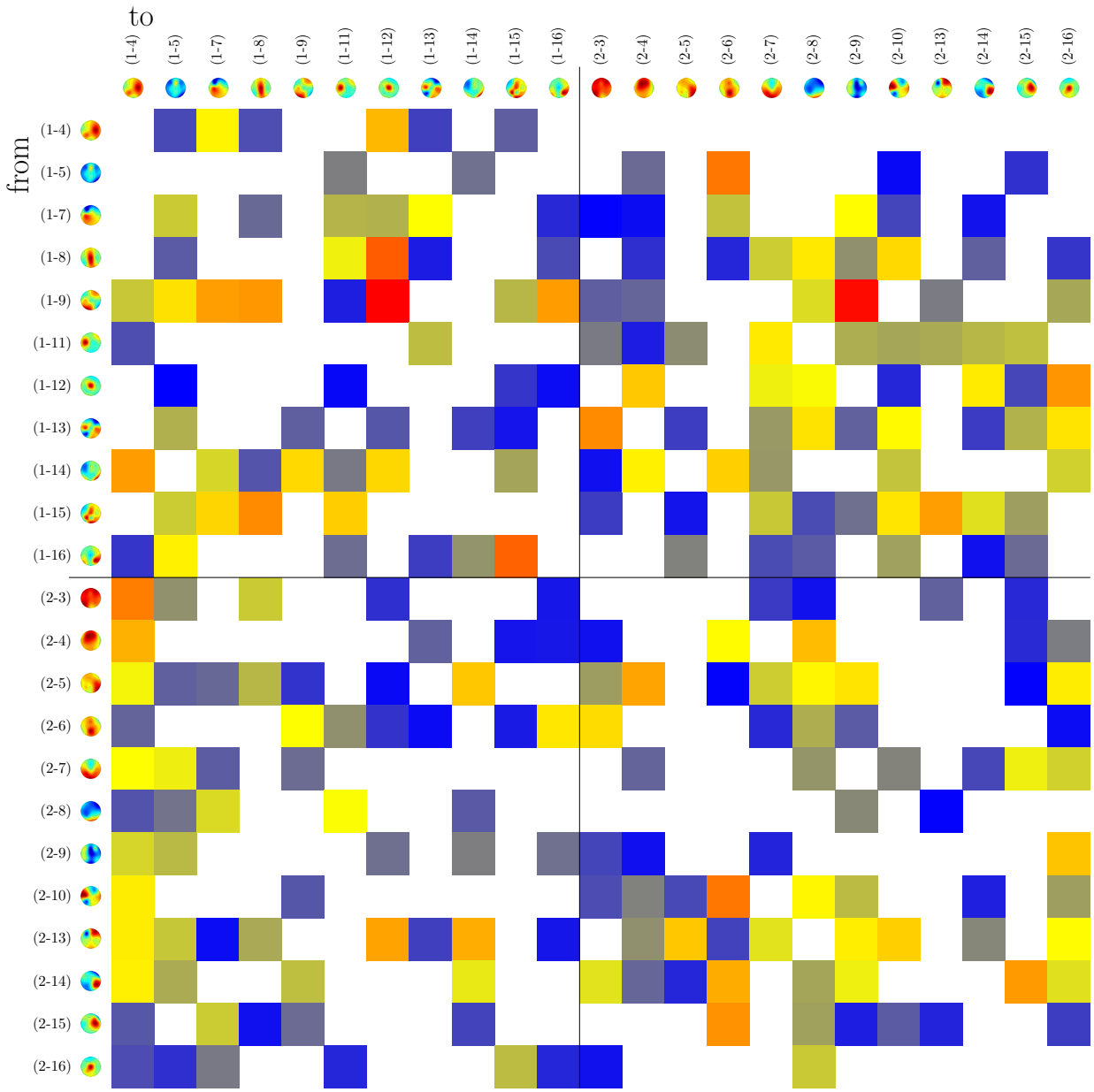


Figure E.13.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment four and θ-band. No significant changes in connectivity or remarkable patterns.

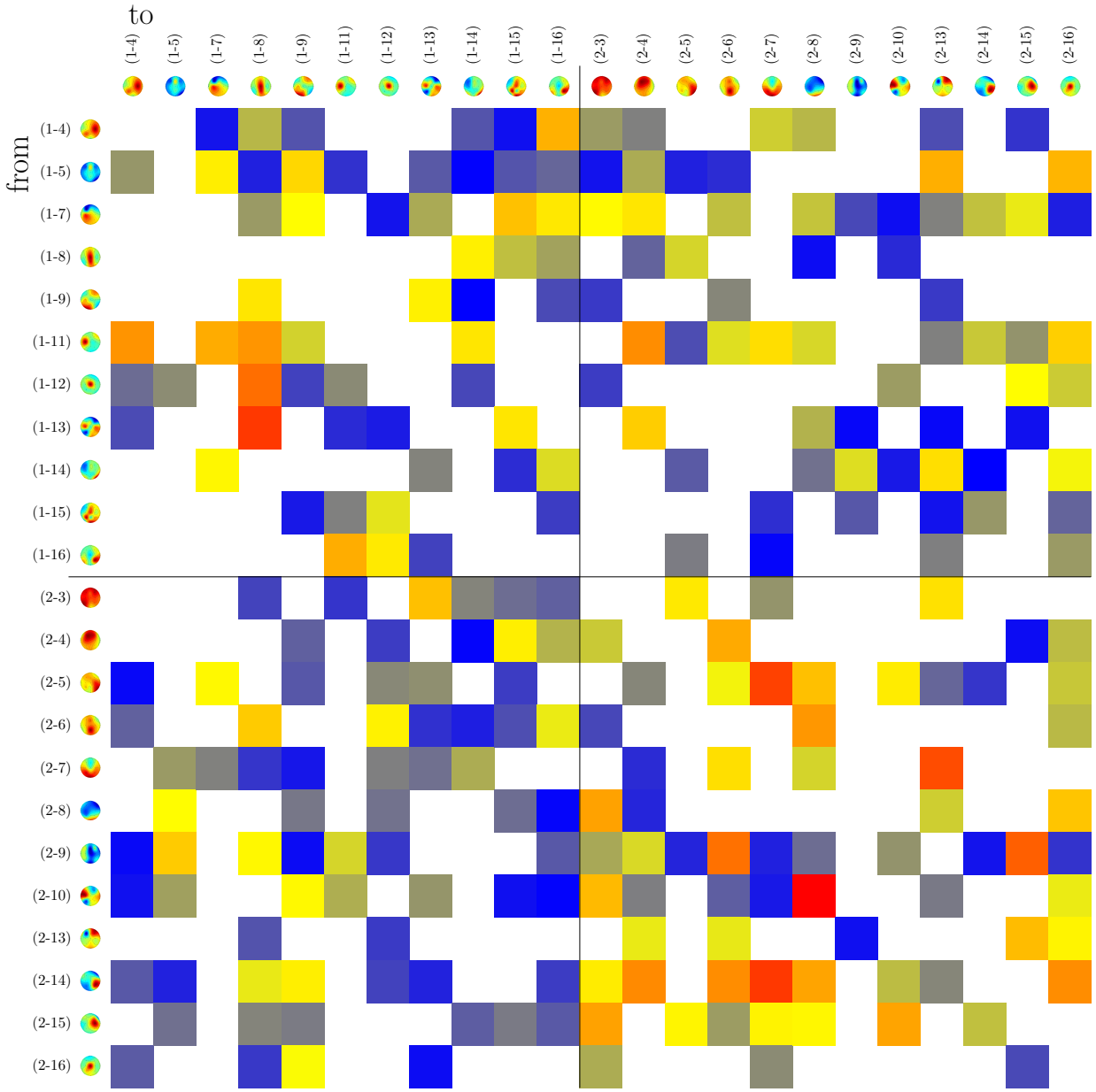


Figure E.14.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment four and α-band. No significant changes in connectivity or remarkable patterns.

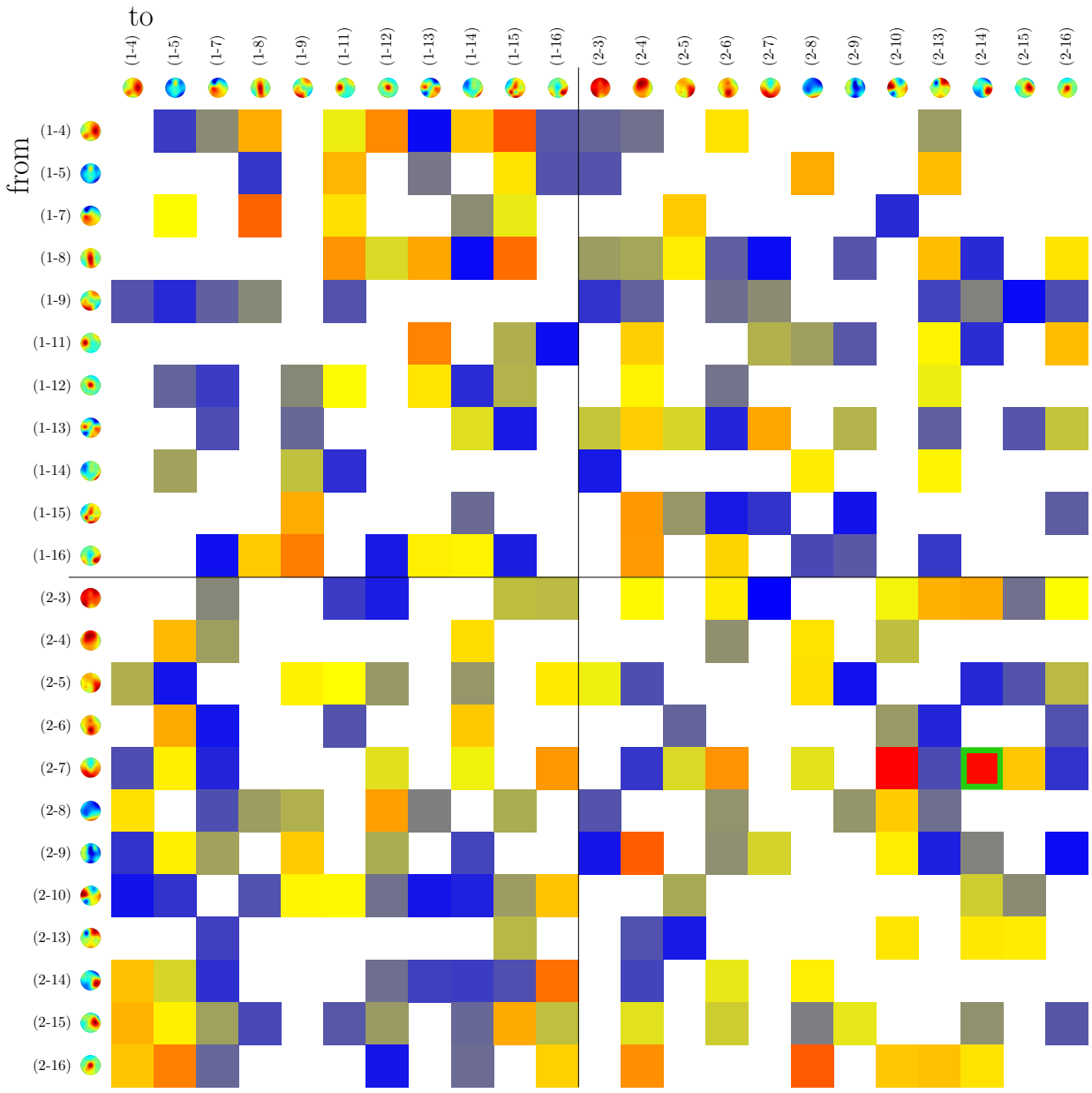


Figure E.15.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment four and β-band. One significant changes in a connection between an occipital component and a parietal-lateral component of participant two was found.

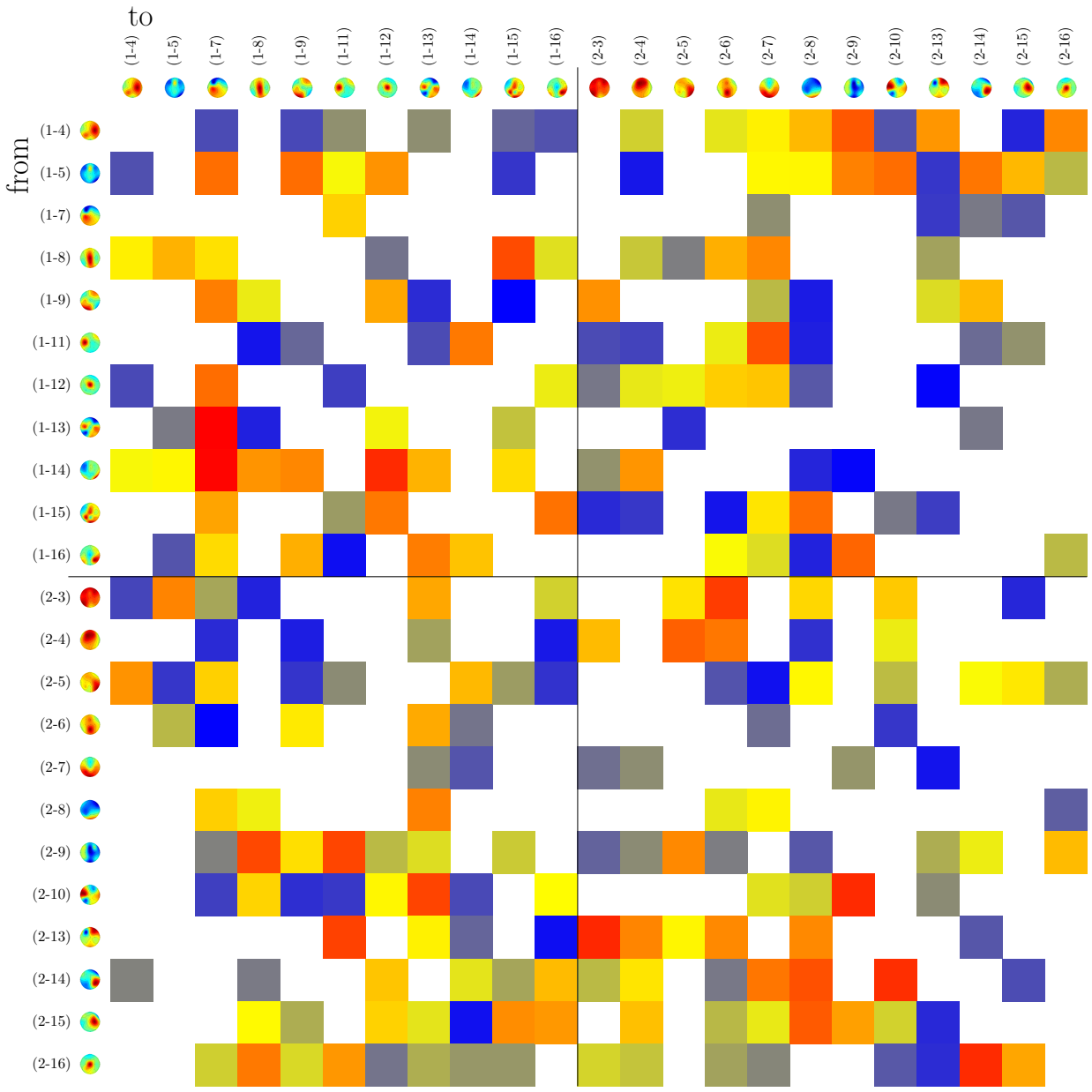


Figure E.16.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment four and γ-band. No significant changes in connectivity or remarkable patterns.

E.5. Experiment Five

One significant change in a within-participant connection in the β-band was found.

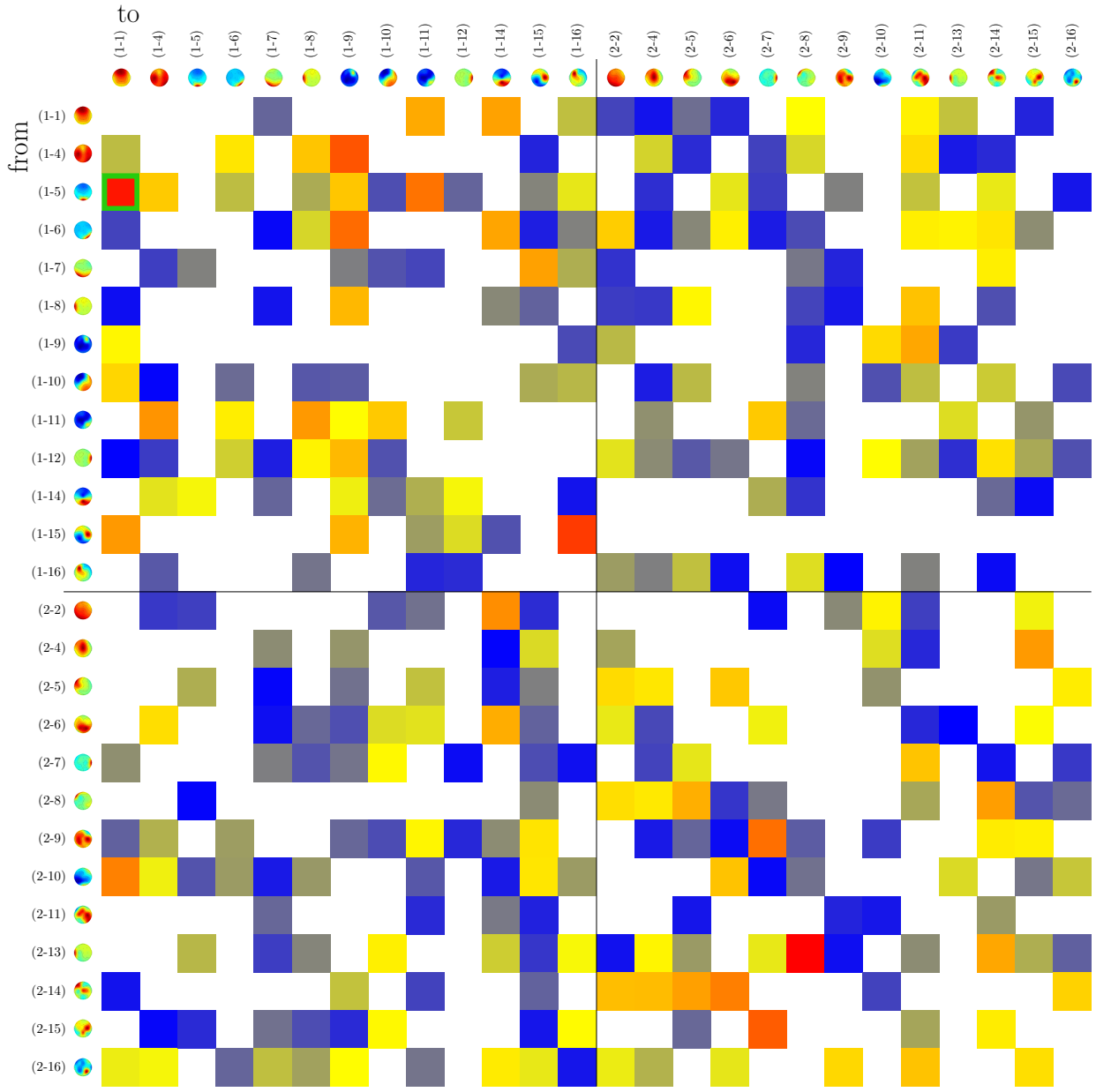


Figure E.17.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment five and θ-band. One significant change in a connection was found for participant one, connecting a narrow occipital component with a broad frontal component.

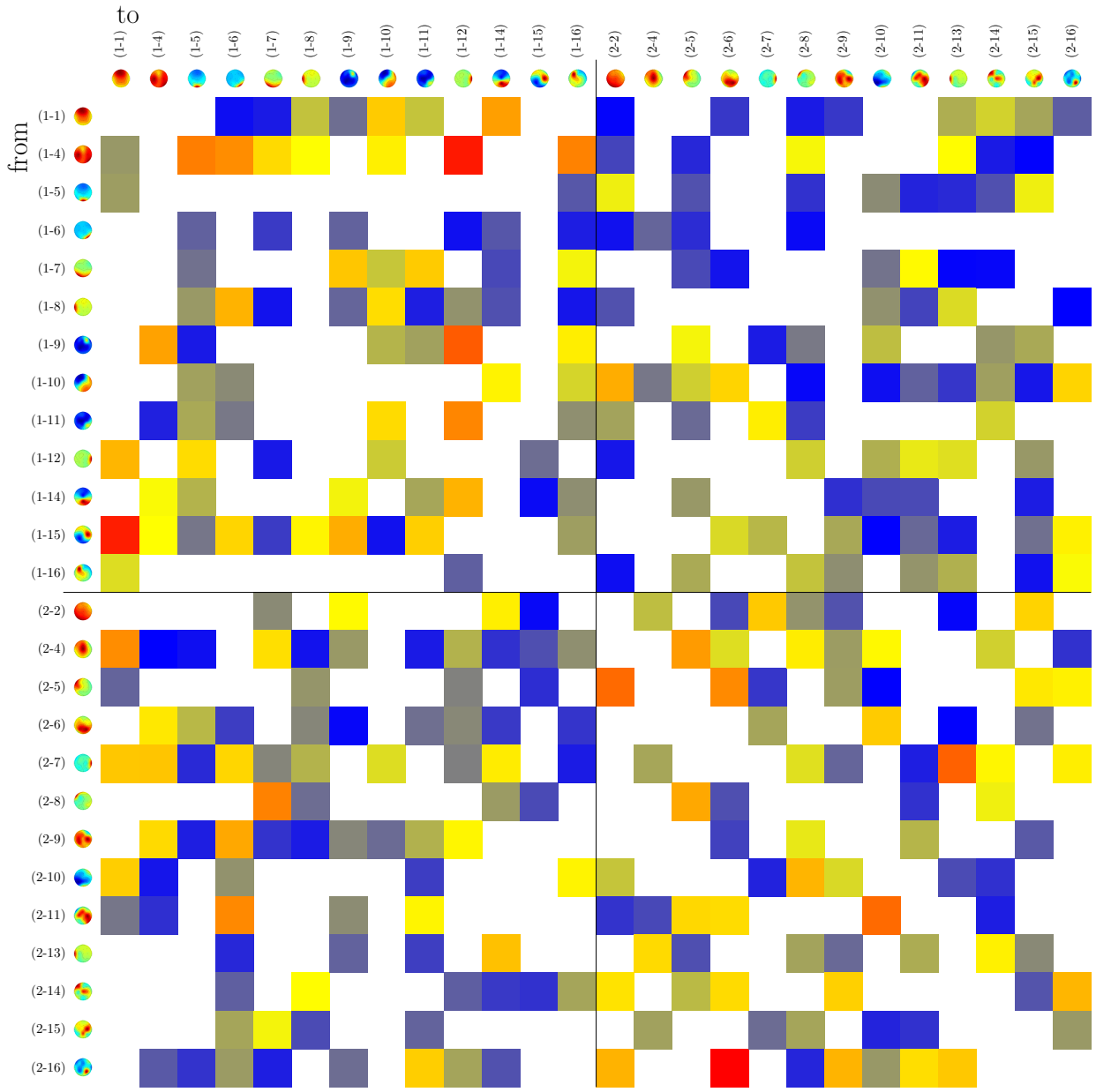


Figure E.18.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment five and α-band. No significant changes in connectivity or remarkable patterns.

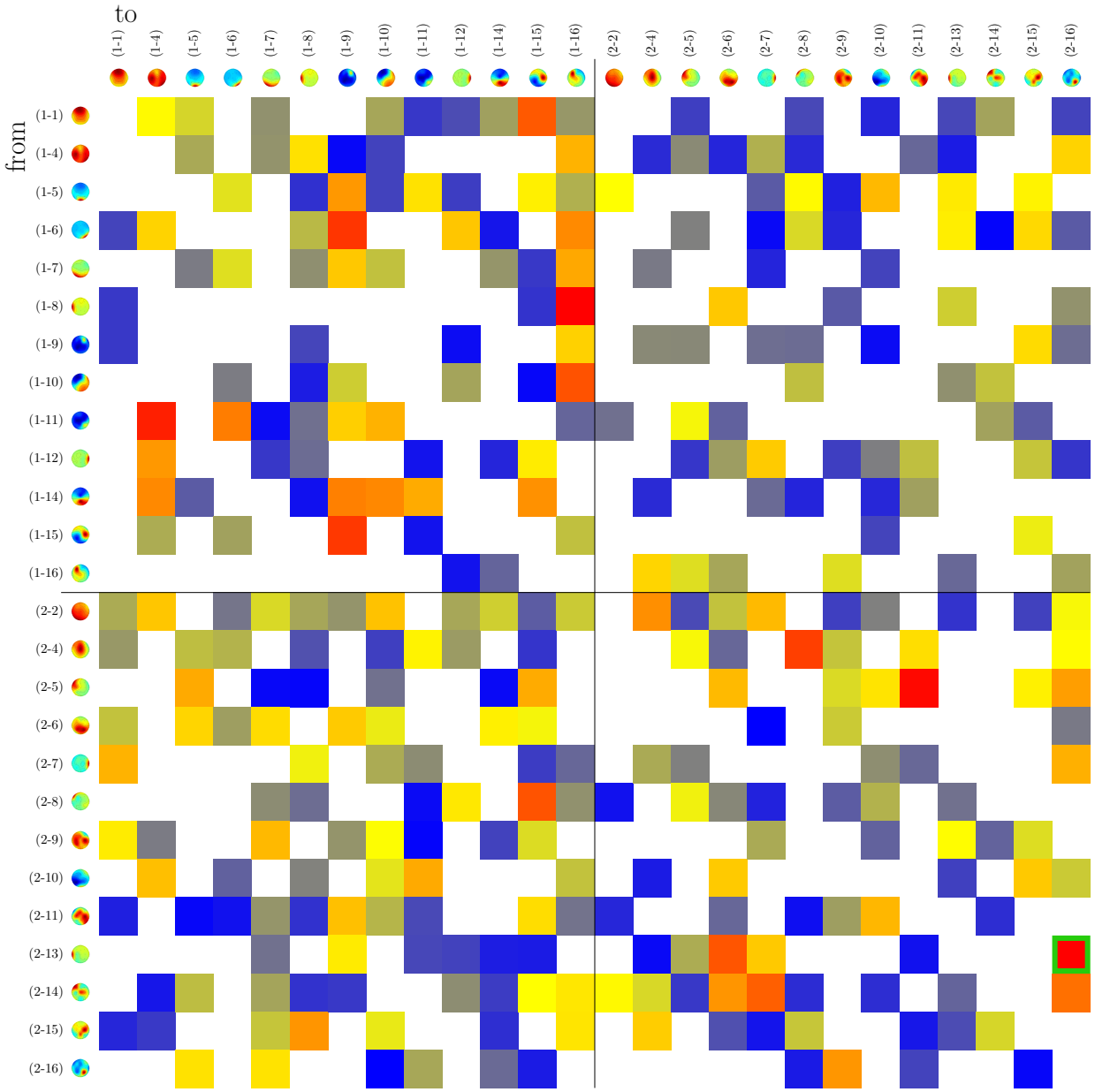


Figure E.19.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment five and β-band. One significant change was found for a within-participant connection of participant two, connecting a strongly lateral component with a narrow parietal-lateral component.

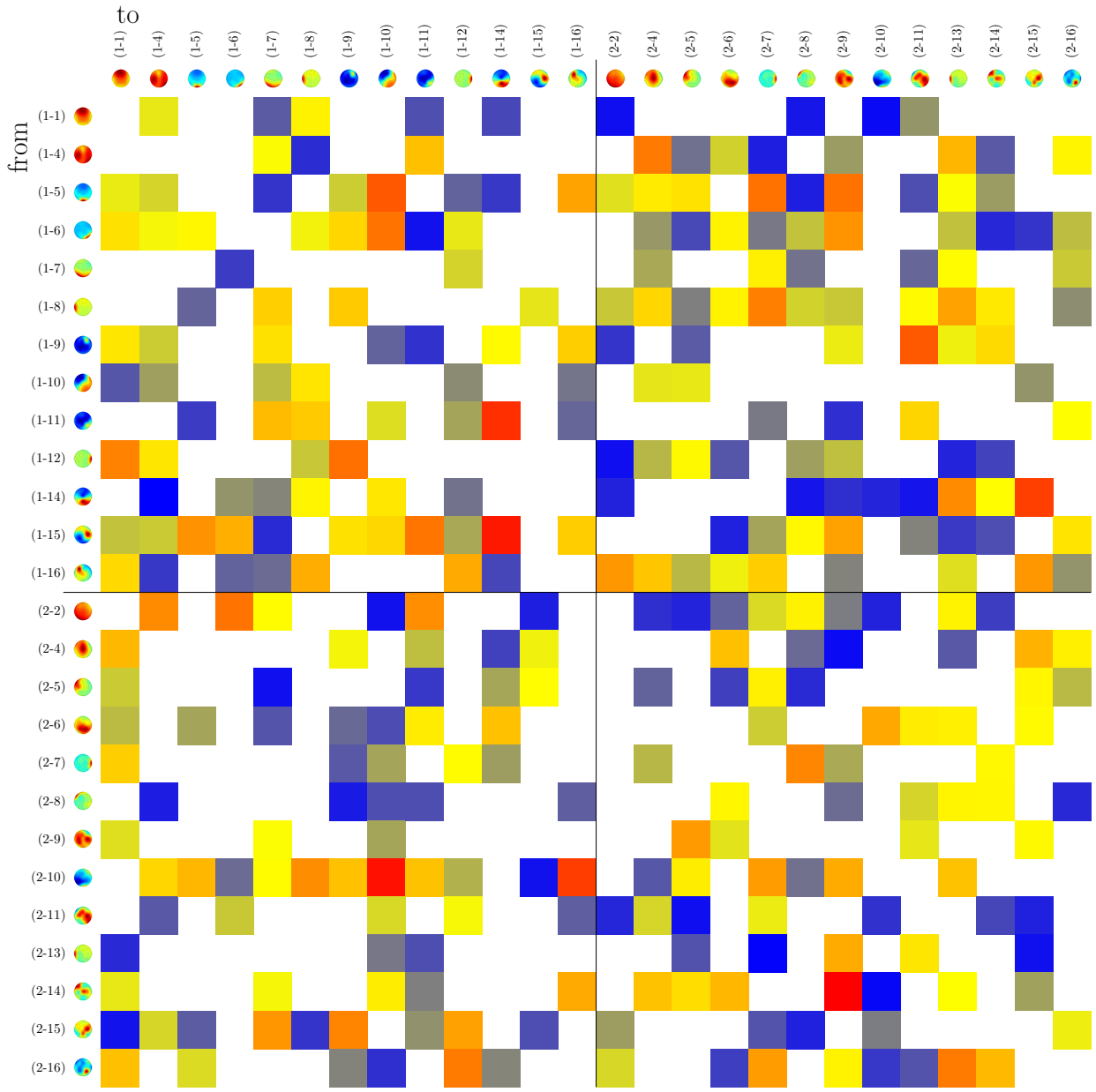


Figure E.20.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment five and γ-band. No significant changes in connectivity or remarkable patterns.

E.6. Experiment Six

Significant changes were found in five within-participant connections and one hyper-connection in α-band as well as in four within-participant connections in γ-band.

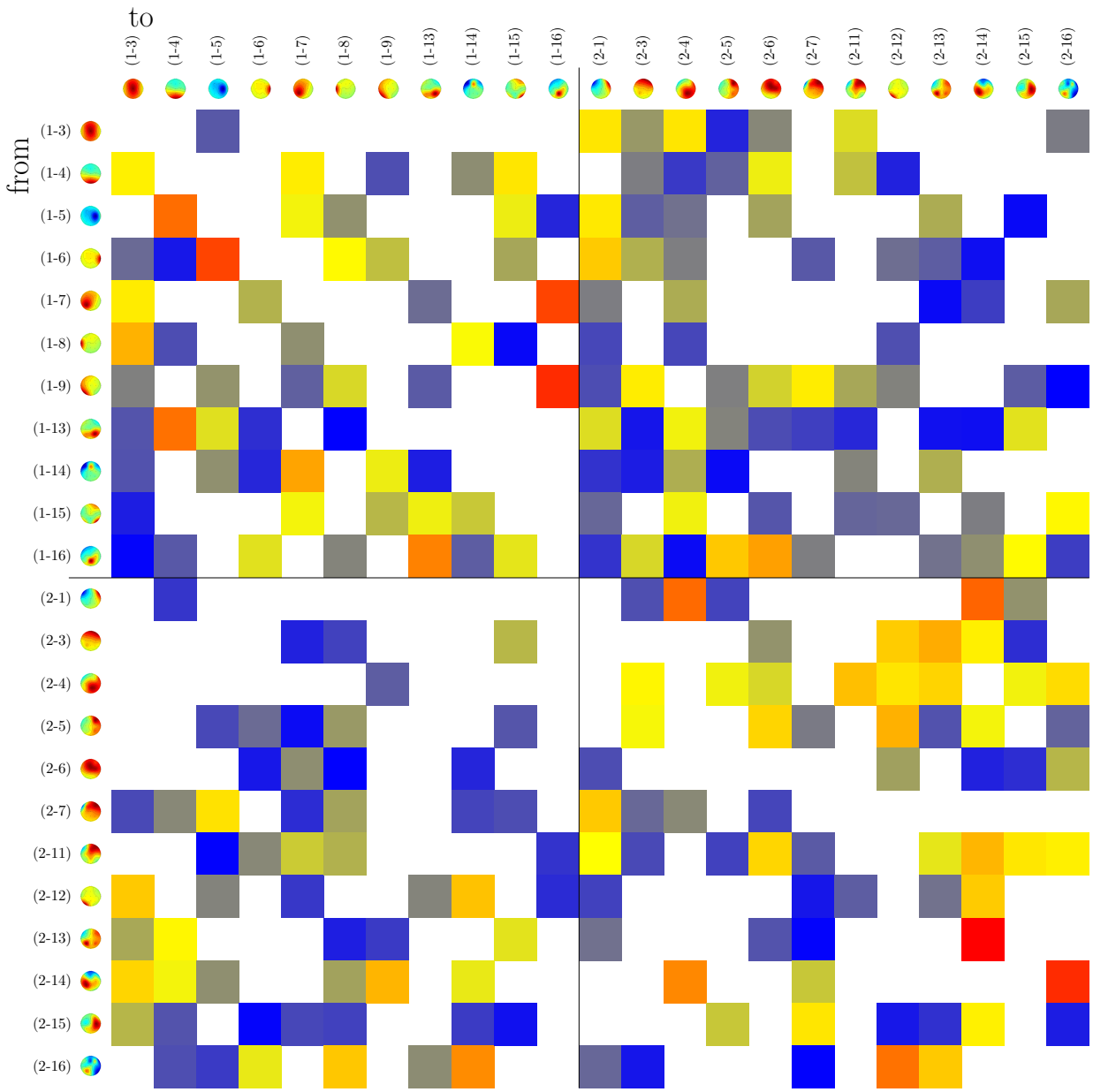


Figure E.21.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment six and θ-band. No significant changes in connectivity or remarkable patterns were found.

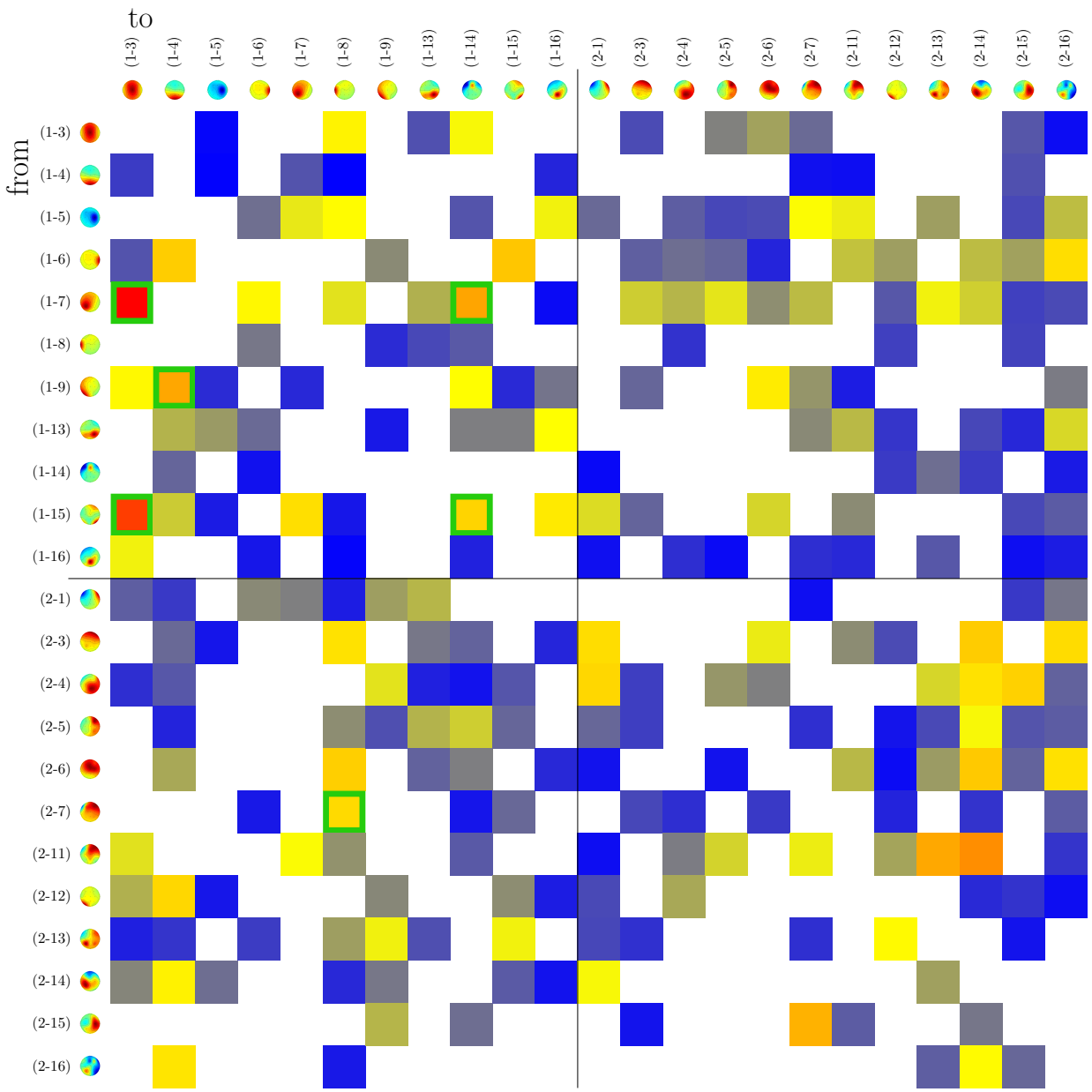


Figure E.22.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment six and α-band. Five significant changes of within-participant connections and one connection from a fronto-lateral component of participant two to a strongly lateral component of participant one were found.

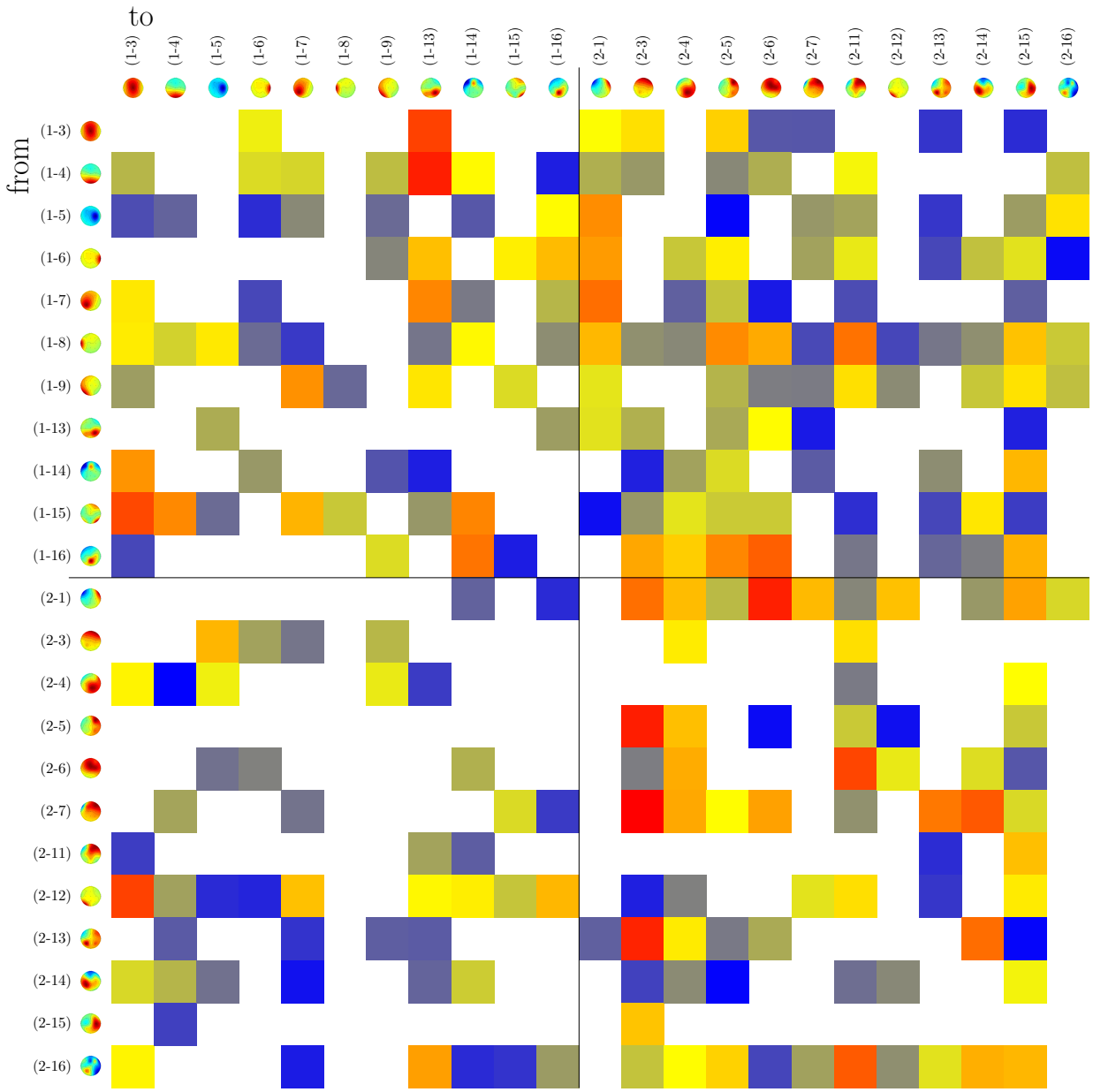


Figure E.23.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment six and β-band. No significant changes in connectivity or remarkable patterns were found.

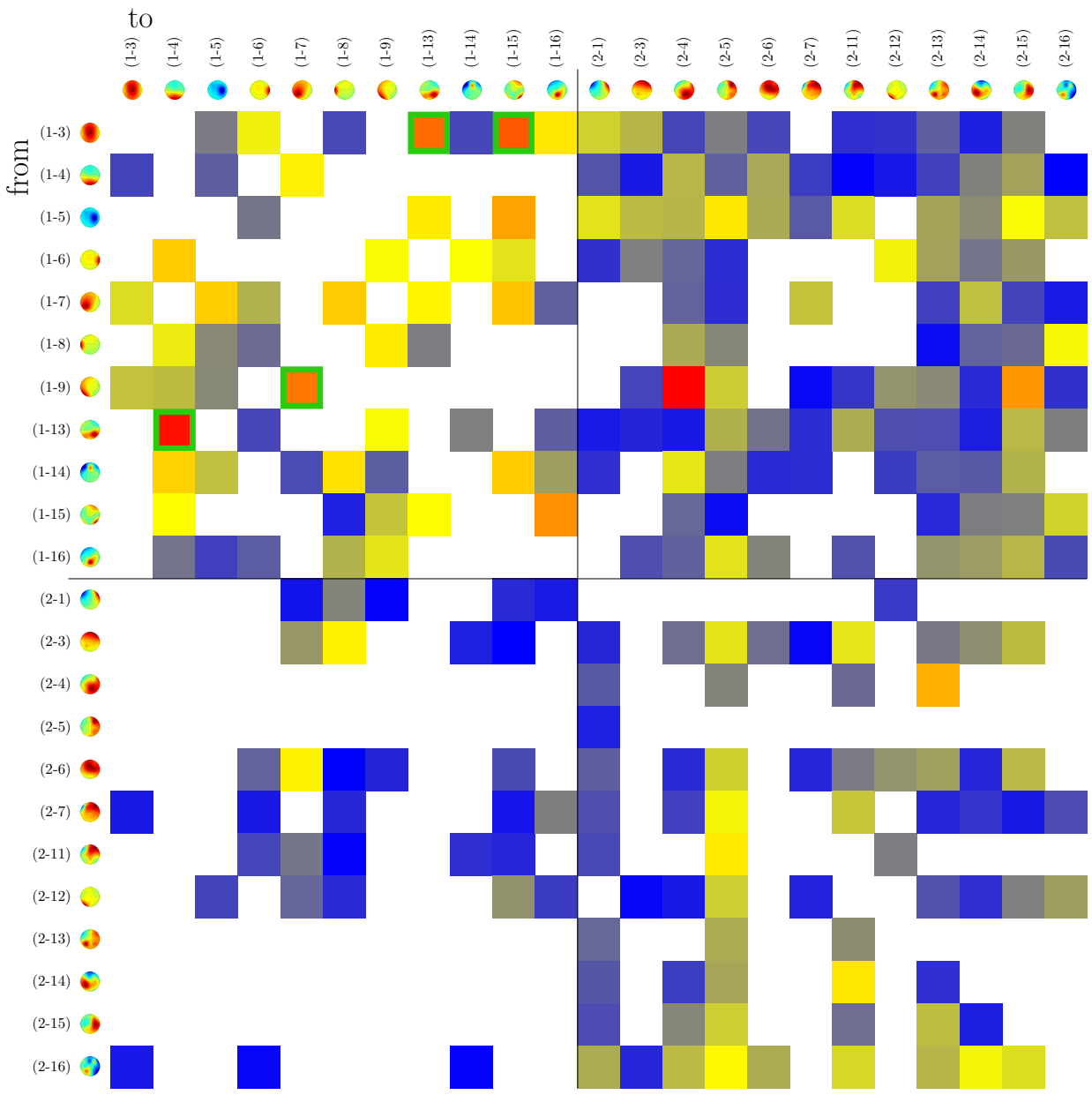


Figure E.24.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment six and γ-band. Four significant changes in within-participant connections were found for participant one. Additionally a bias favouring participant one as the sender of hyper-connections can be observed.

E.7. Experiment Seven

No significant changes in connectivity could be shown. But there seems to be a bias among hyper-connections in β- and γ-band favouring participant two as the sender.

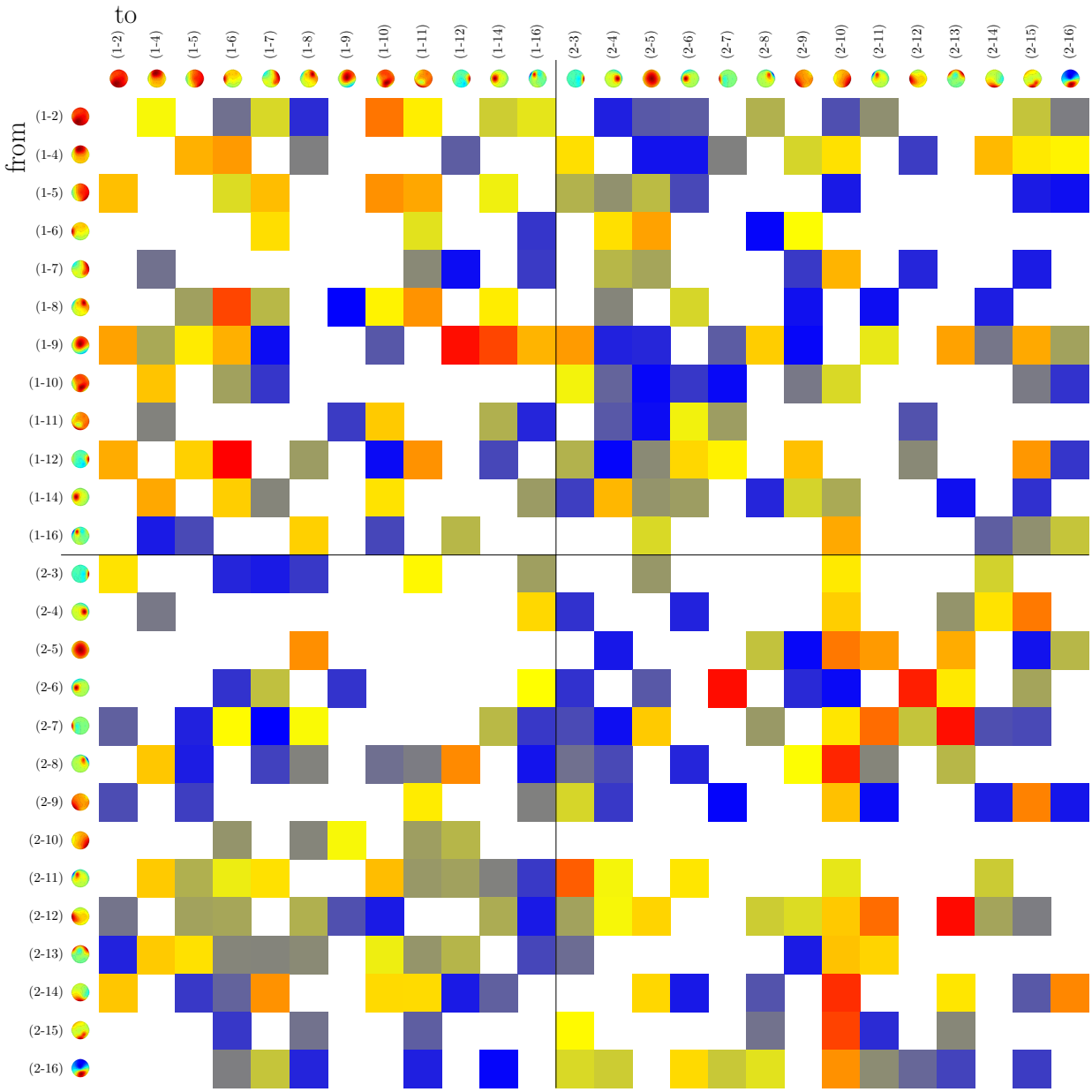


Figure E.25.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment seven and θ-band. No significant changes in connectivity were found, nor remarkable patterns.

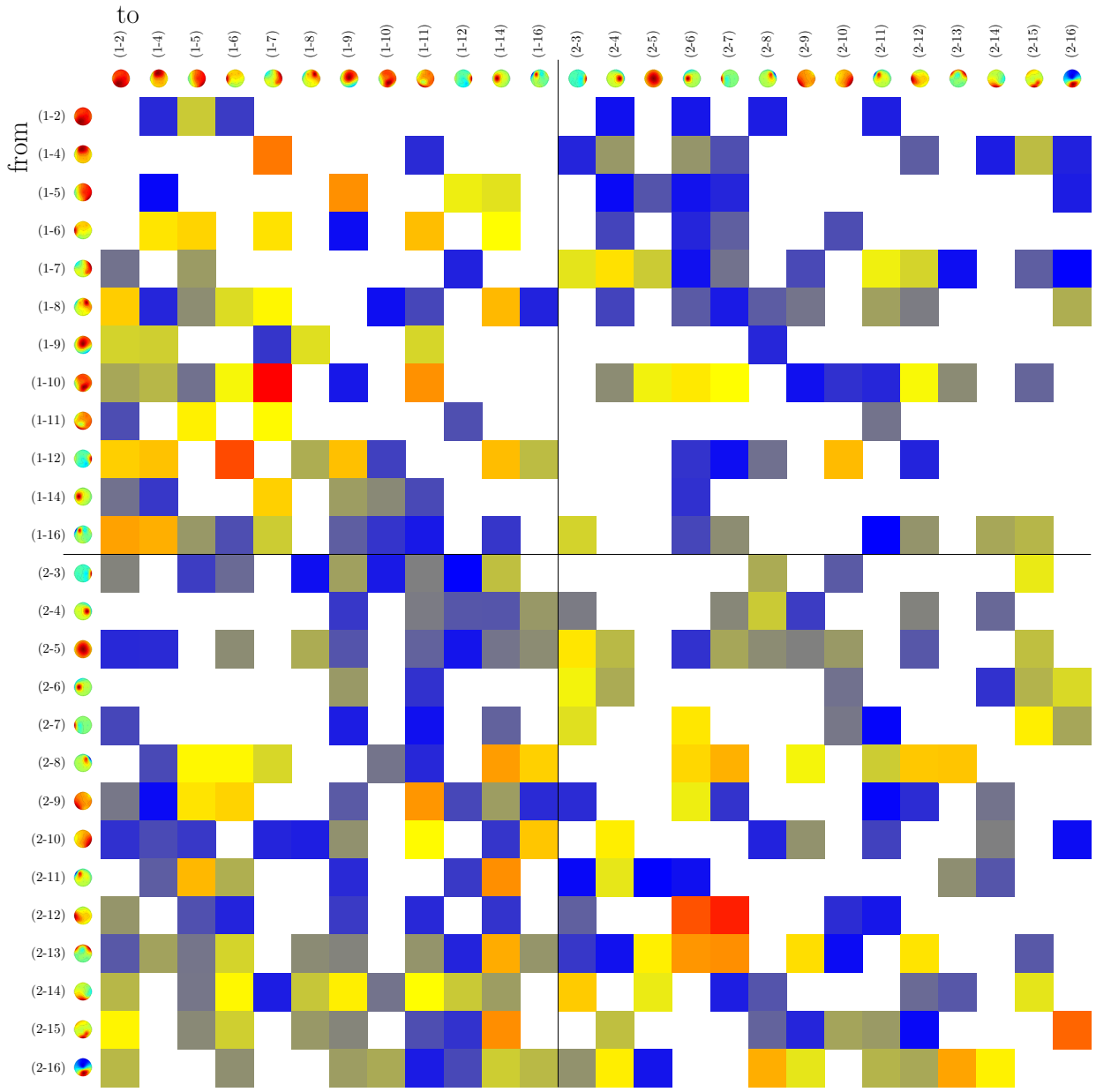


Figure E.26.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment seven and α-band. No significant changes in connectivity, nor remarkable patterns were found.

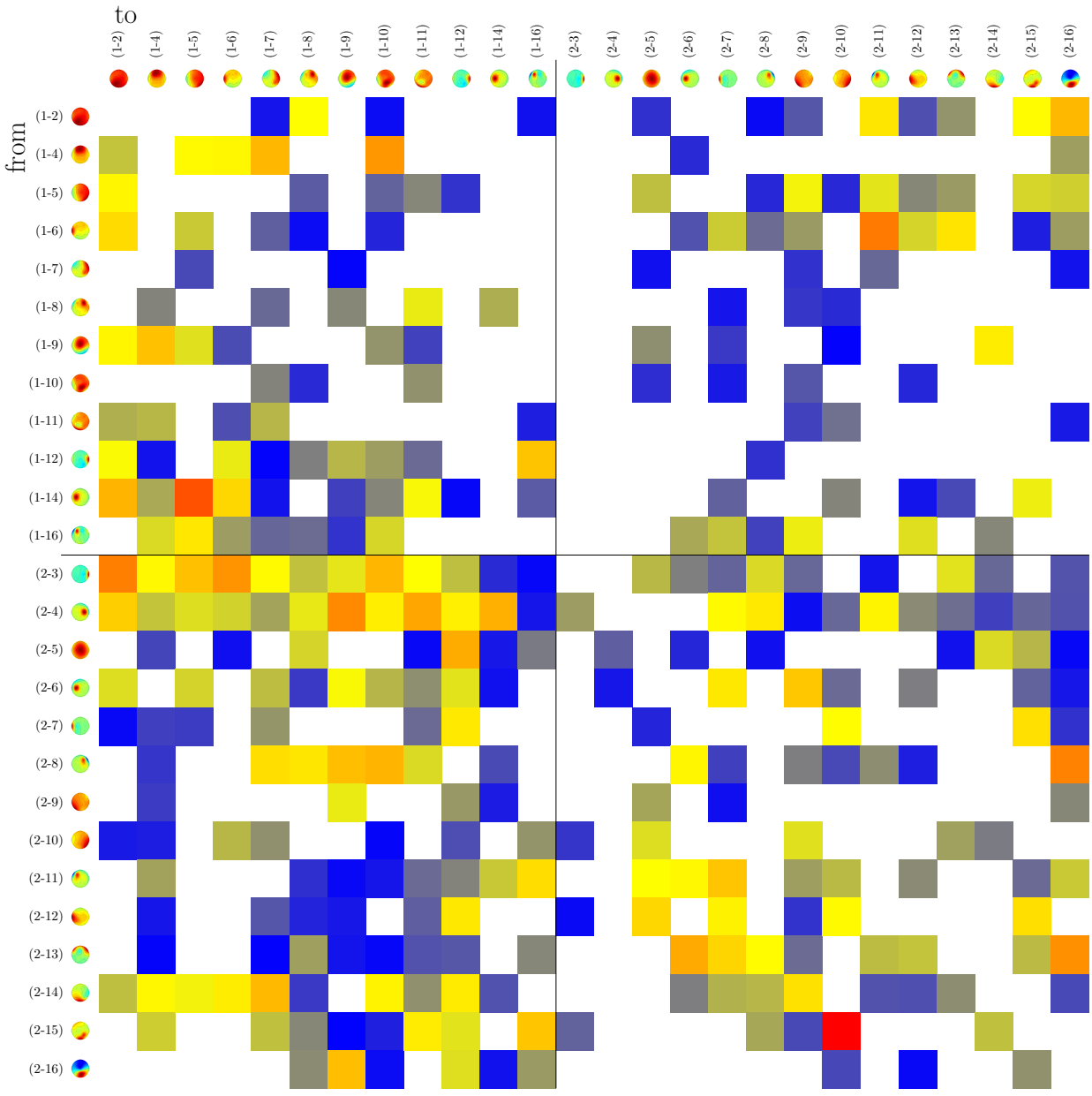


Figure E.27.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment seven and β-band. No significant changes in connectivity were found, although there seems to be a slight bias among hyper-connections favouring participant two in the role of the sender.

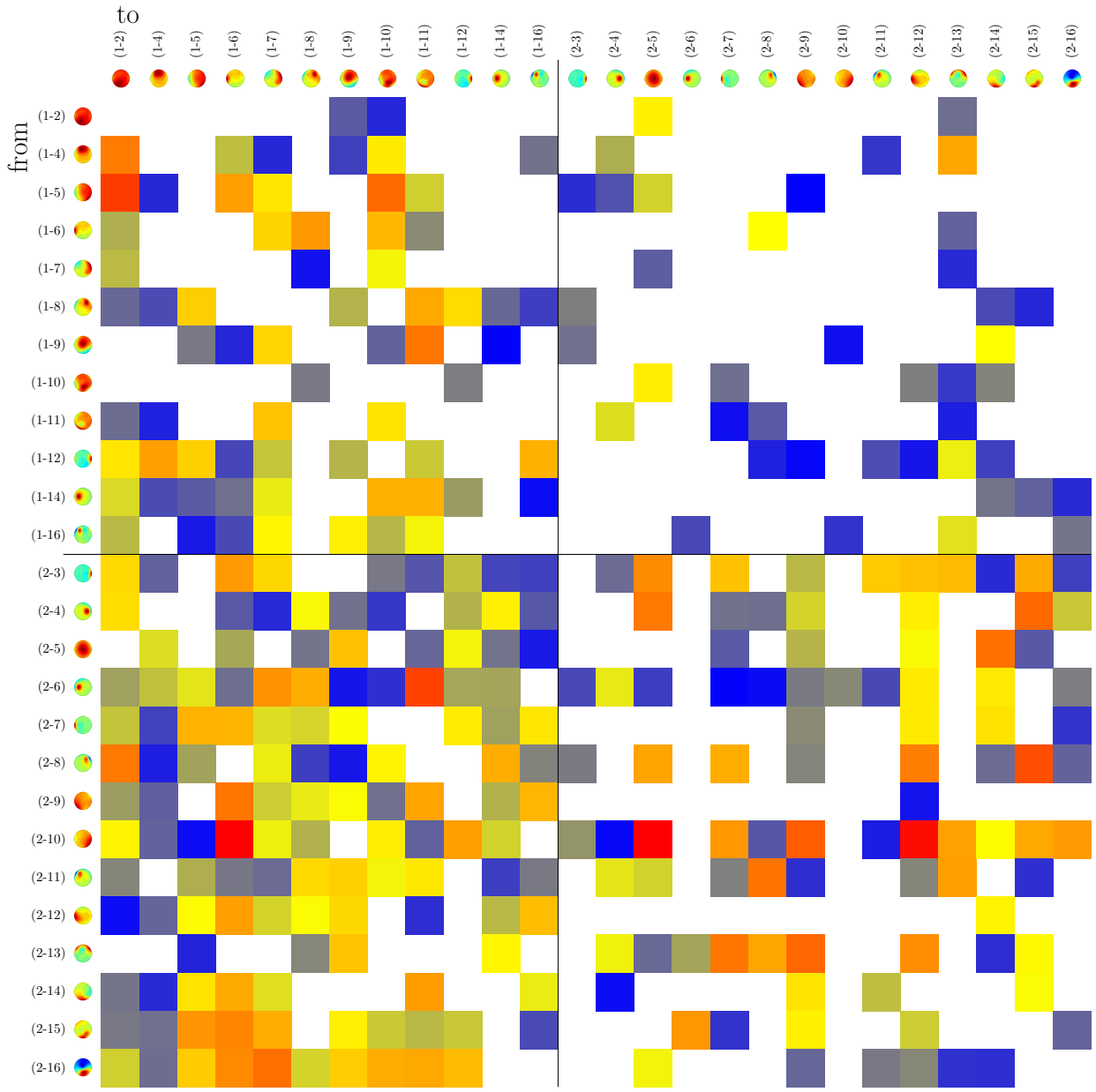


Figure E.28.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment seven and γ-band. No significant changes in connectivity were found, but a distinctive bias favouring participant two as the sender for hyper-connections can be observed.

E.8. Experiment Eight

One significant change for a within-participant connection was found in θ-band.

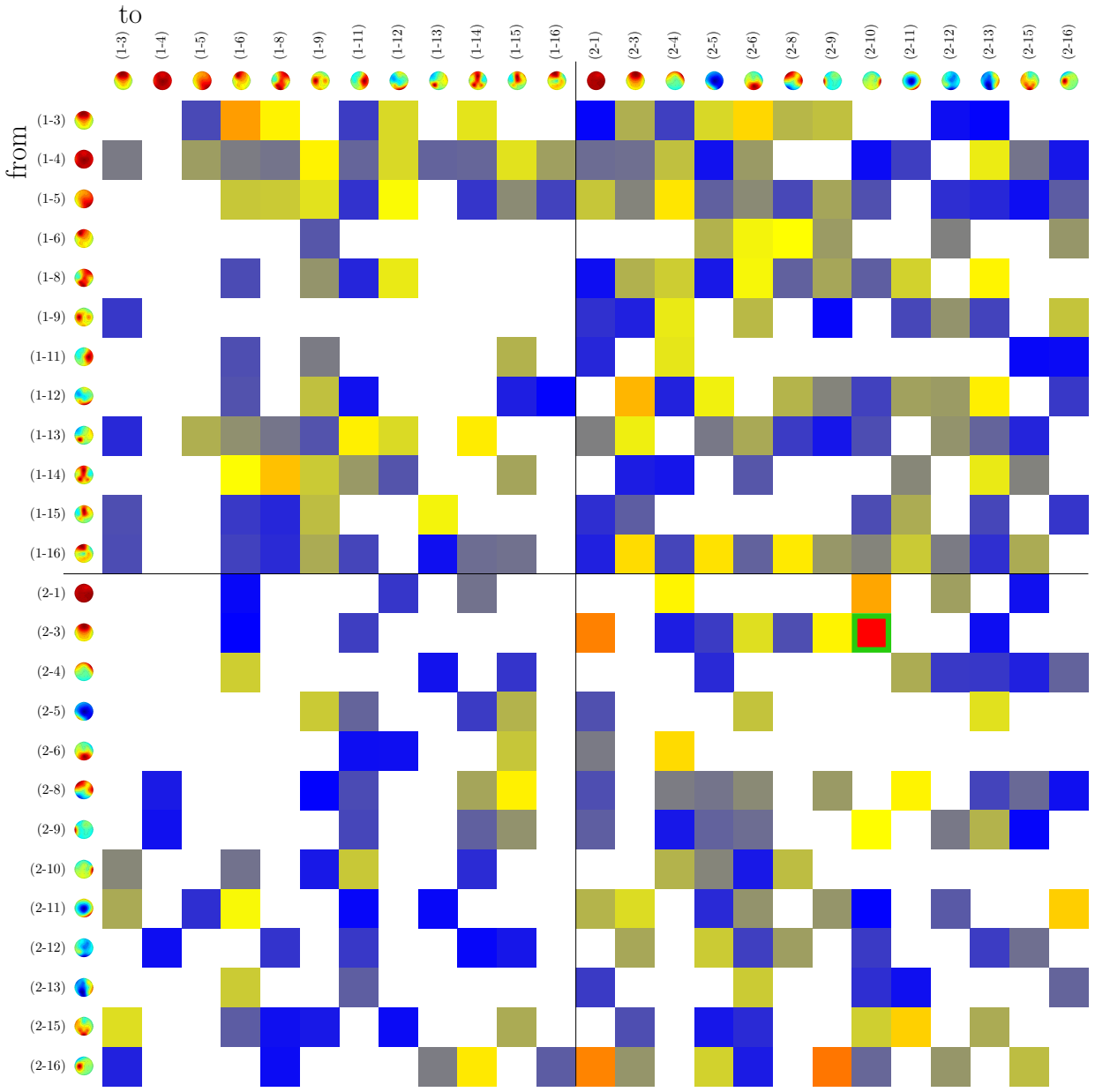


Figure E.29.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment eight and θ-band. One significant change in a connection from a frontal to a strongly lateral component of participant two was found.

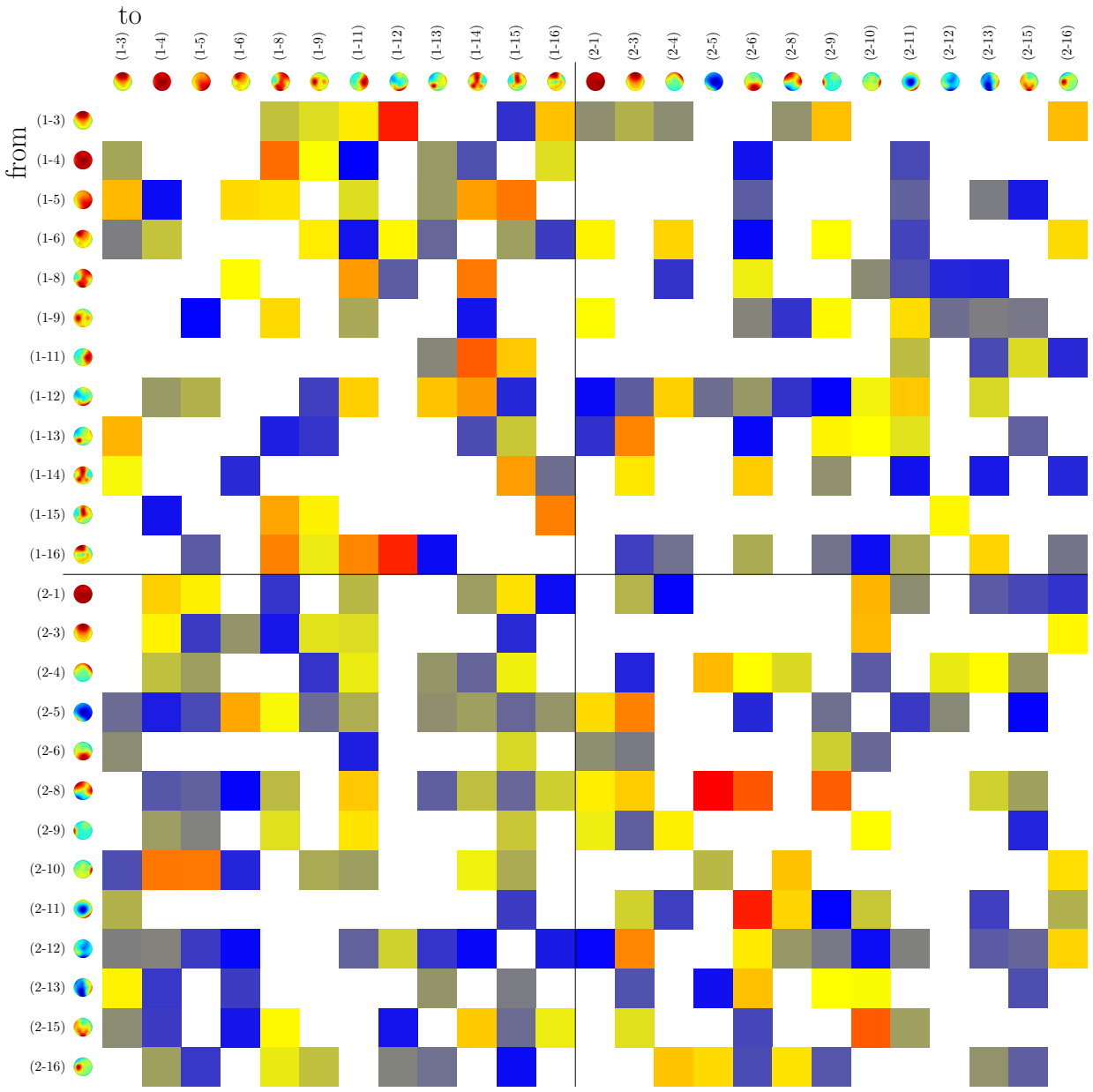


Figure E.30.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment eight and α-band. No significant changes in connectivity, nor remarkable patterns were found.

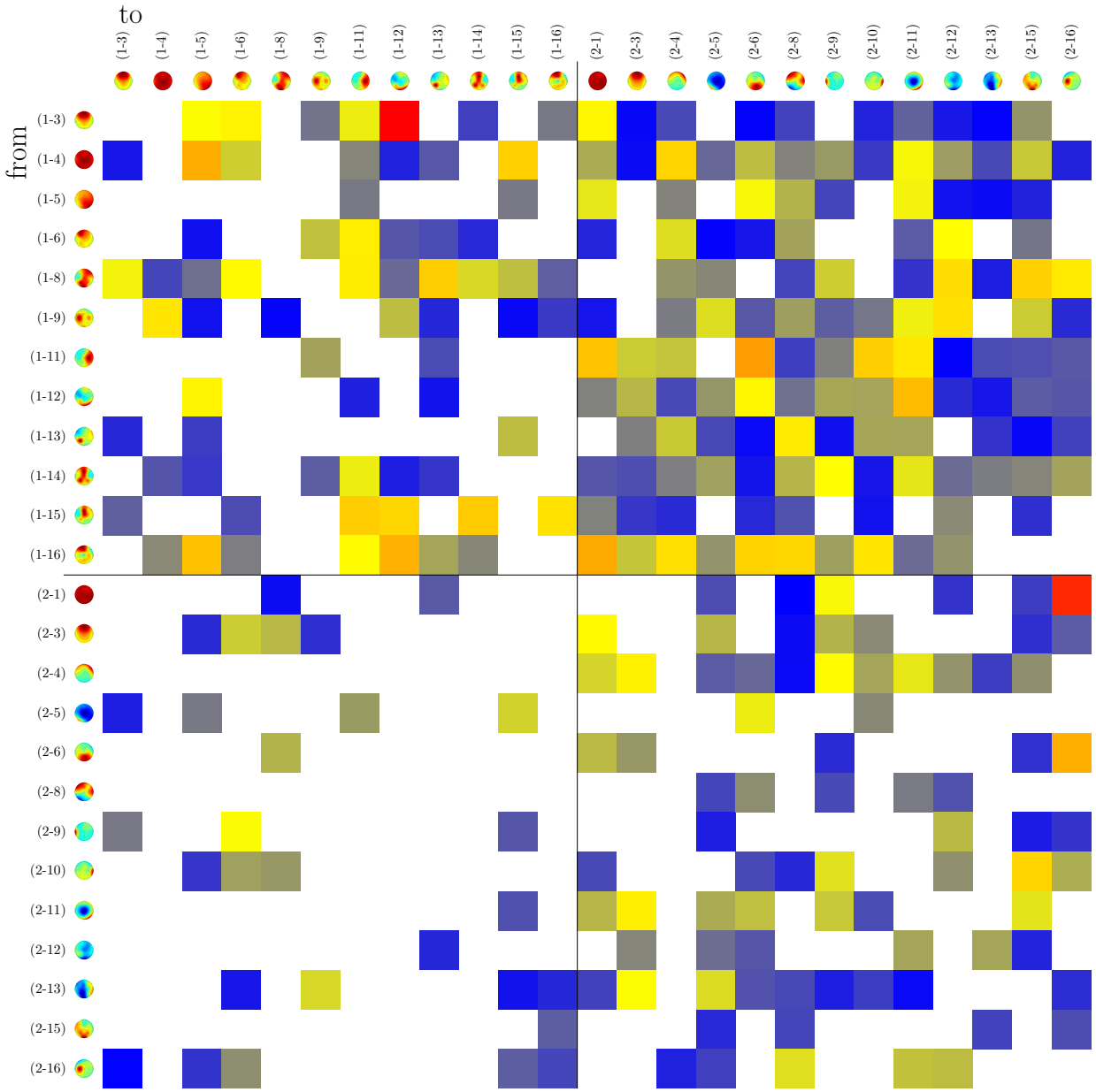


Figure E.31.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment eight and β-band. No significant changes in connectivity were found. Among hyper-connections participant one is indicated as the sender much more often than as the receiver. However, since all these connections are rather weak (have low values) this is far from conclusive.

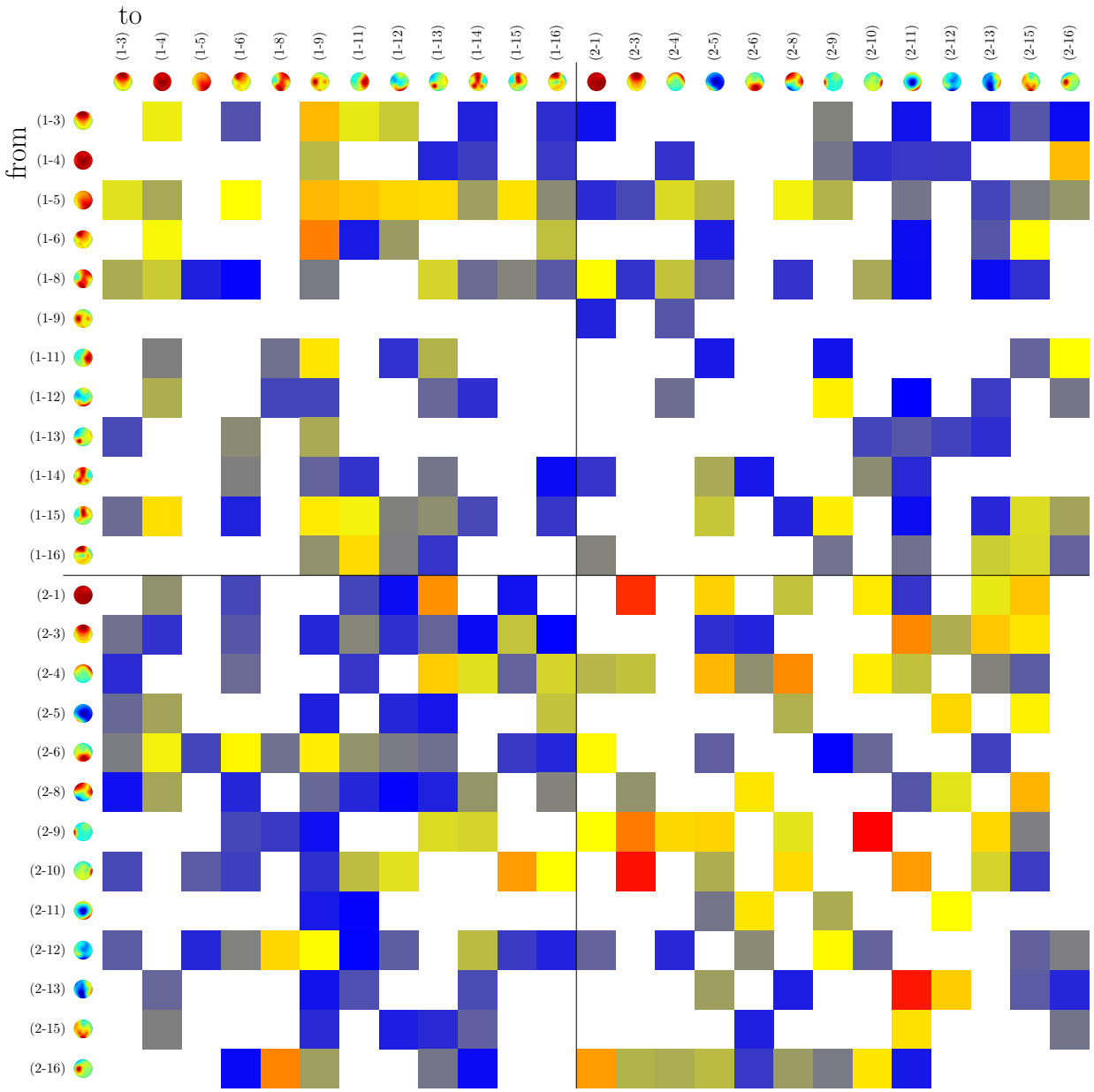


Figure E.32.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment eight and γ-band. No significant changes in connectivity, nor remarkable patterns could be shown.

E.9. Experiment Nine

Eighteen significant changes in within-participant connections found in α-band and two in γ-band.

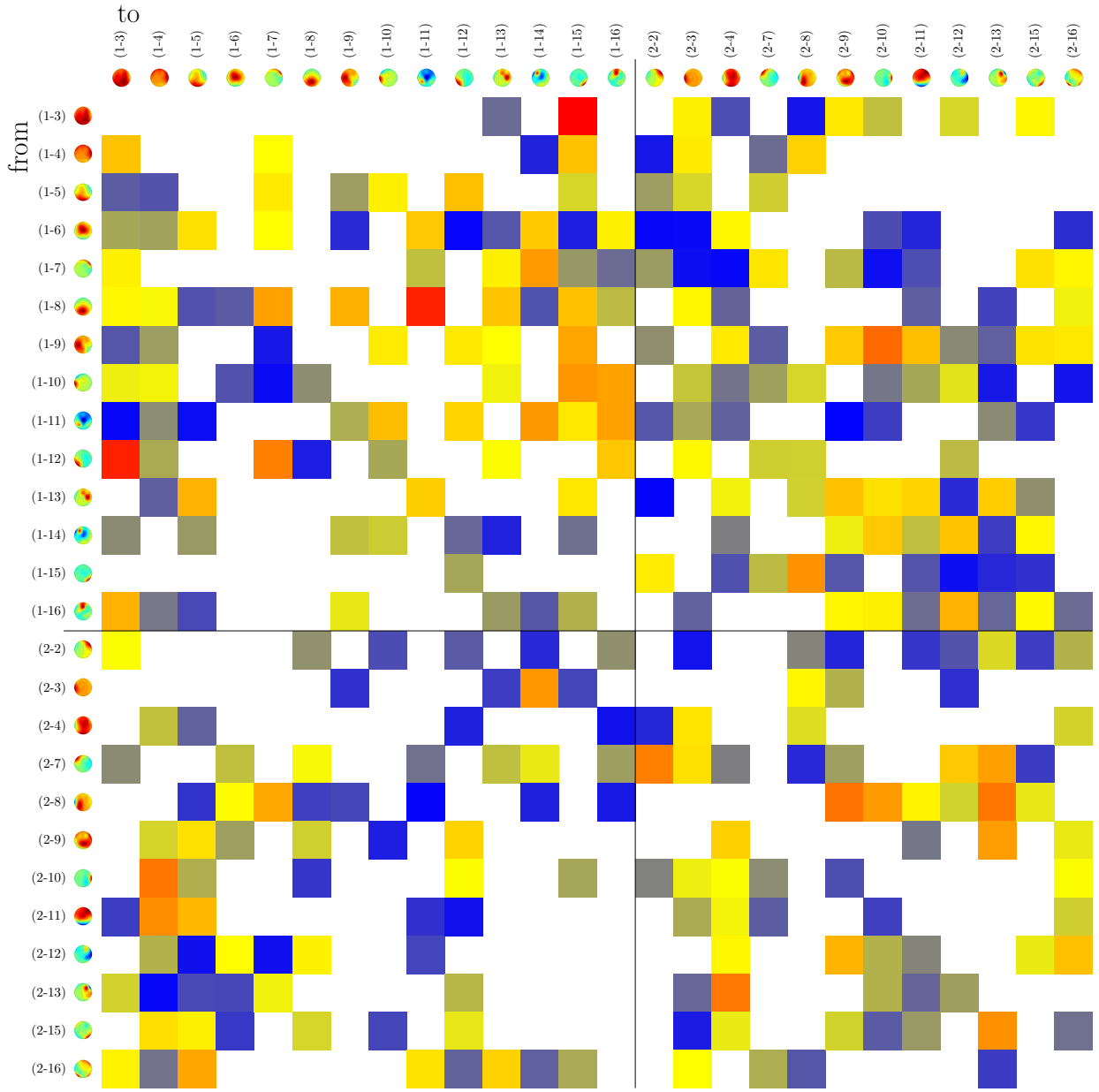


Figure E.33.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment nine and θ-band. No significant changes in connectivity, nor other evidence was found.

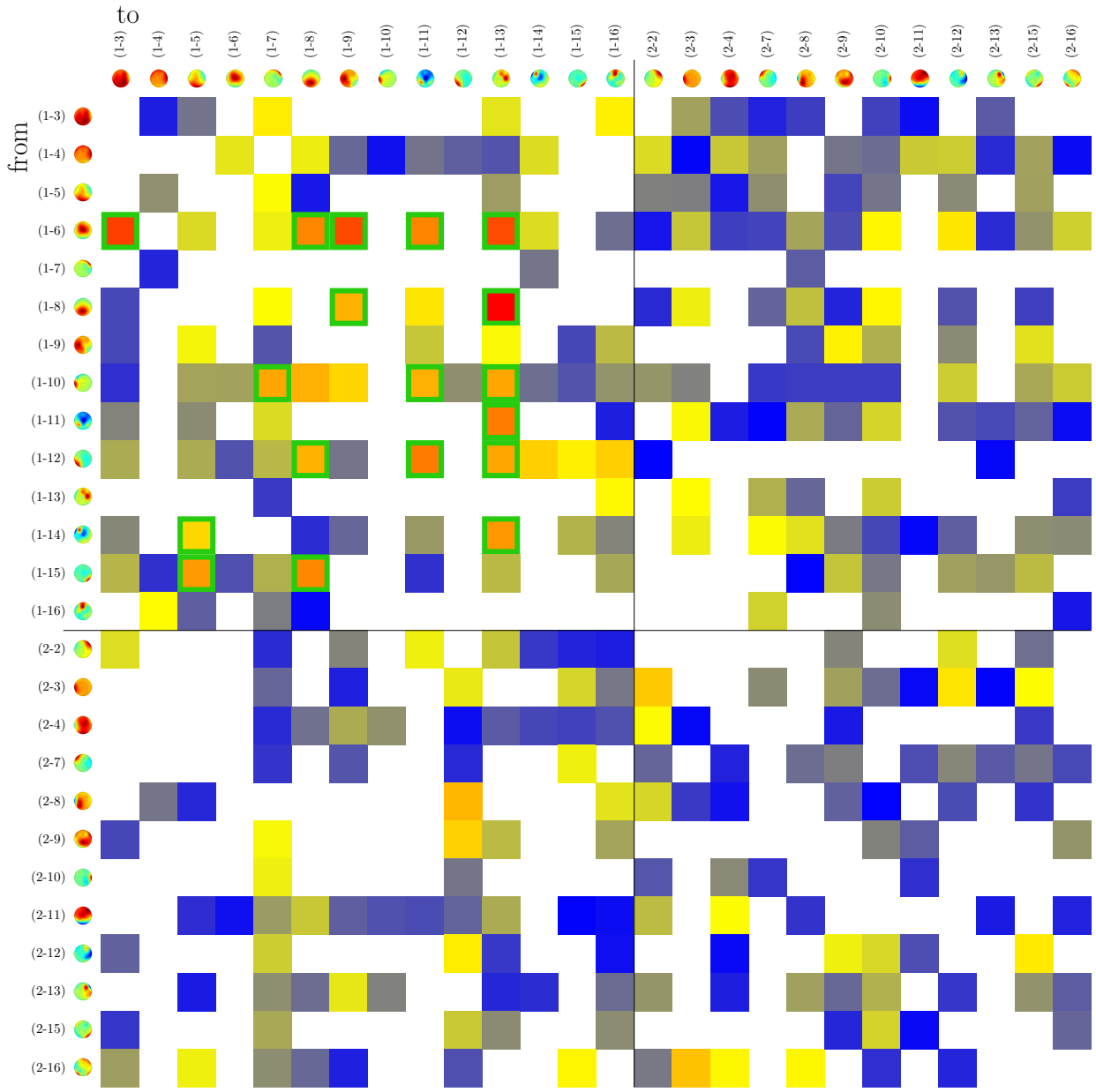


Figure E.34.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment nine and α-band. Eighteen significant changes affecting within-participant connections were found, all for participant one.

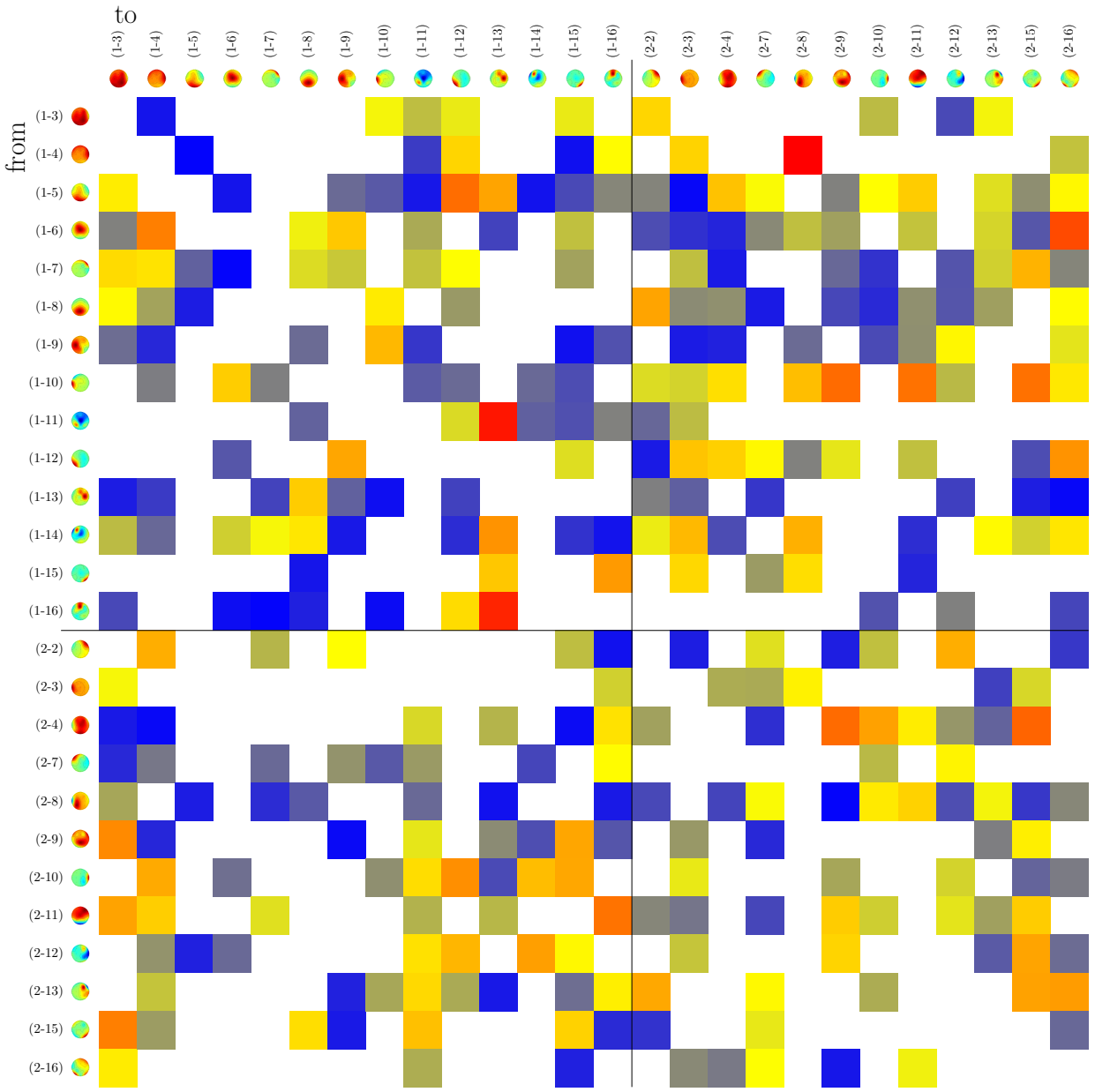


Figure E.35.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment nine and β-band. No significant changes in connectivity were found, nor other interesting observations were made.

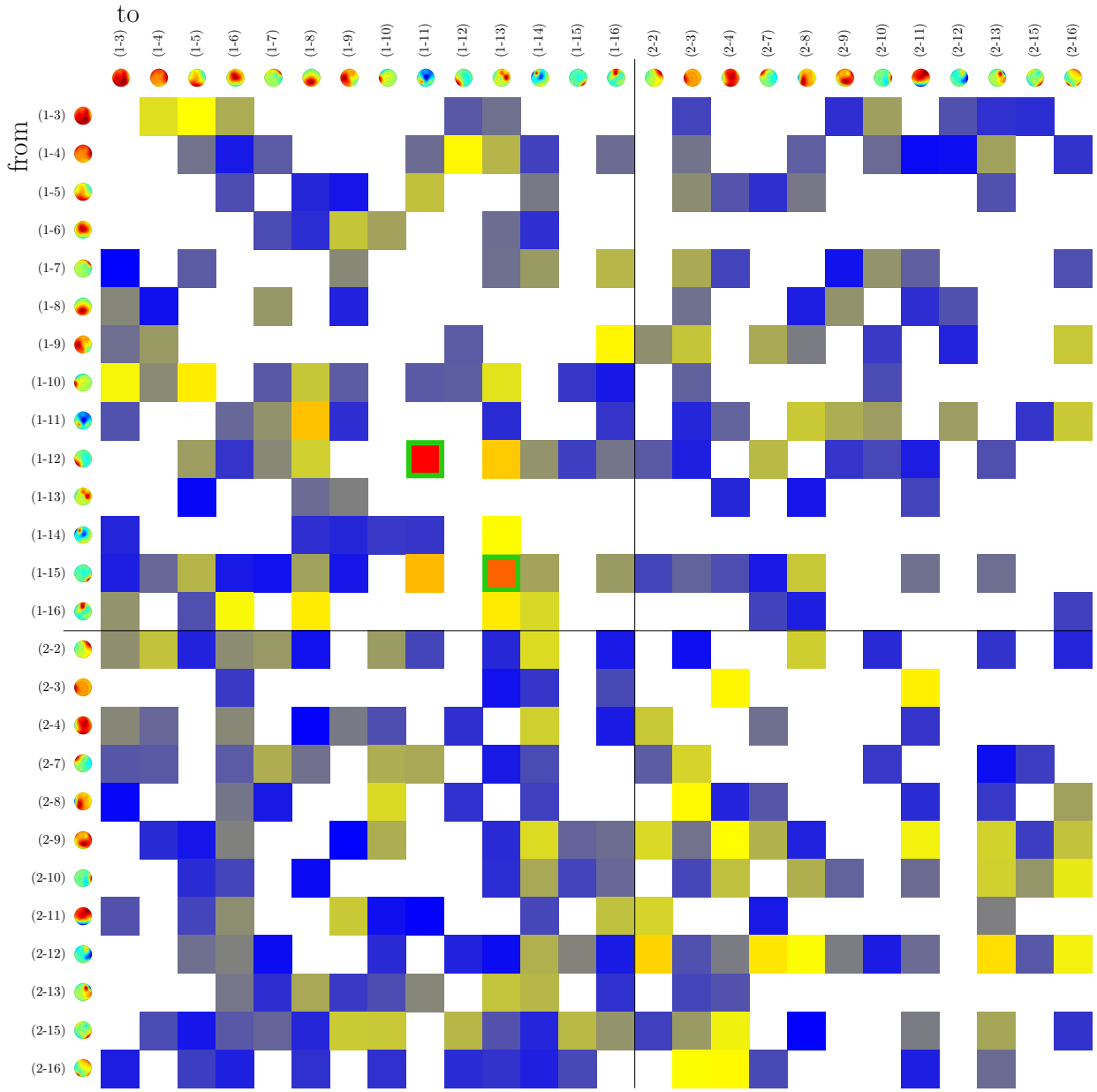


Figure E.36.: Differences in PSI connectivity between epochs recorded during the experiment and the baseline period for experiment nine and γ-band. Two significant changes of within-participant connections were found for participant one. One connecting a strongly parietal-lateral component with even narrower parietal ipsi-lateral component and one connecting a strongly occipital-latera component with a fronto-ipsi-lateral component.

F. Differential PSI: Cooperative vs. Non Cooperative

These tables show the results of the differential PSI analysis contrasting data from cooperative with data from non-cooperative trials. For details on the interpretation of these plots please refer to section 8.6.2 on page 102.

F.1. Experiment One

Some significant changes are found in θ- and α-band which all affect hyper-connections from participant two to participant one.

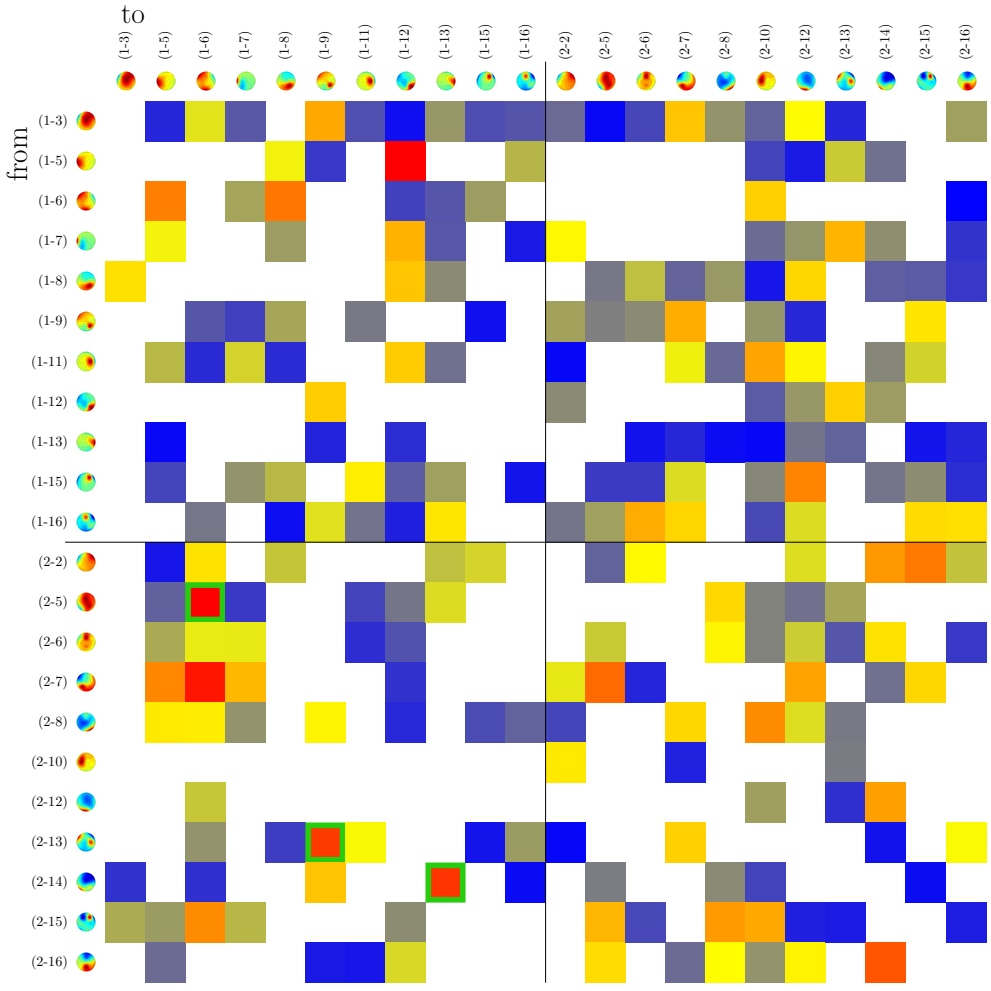


Figure F.1.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment one θ-band. Changes in three connections are significant. All three affected connections are hyper-connections from participant one to participant two. One connects a global component with a very broad lateral component. Another connects a strongly parietal component with one focusing on frontal regions and one parietal spot. The last connects a parietal-lateral component with a strongly lateral component.

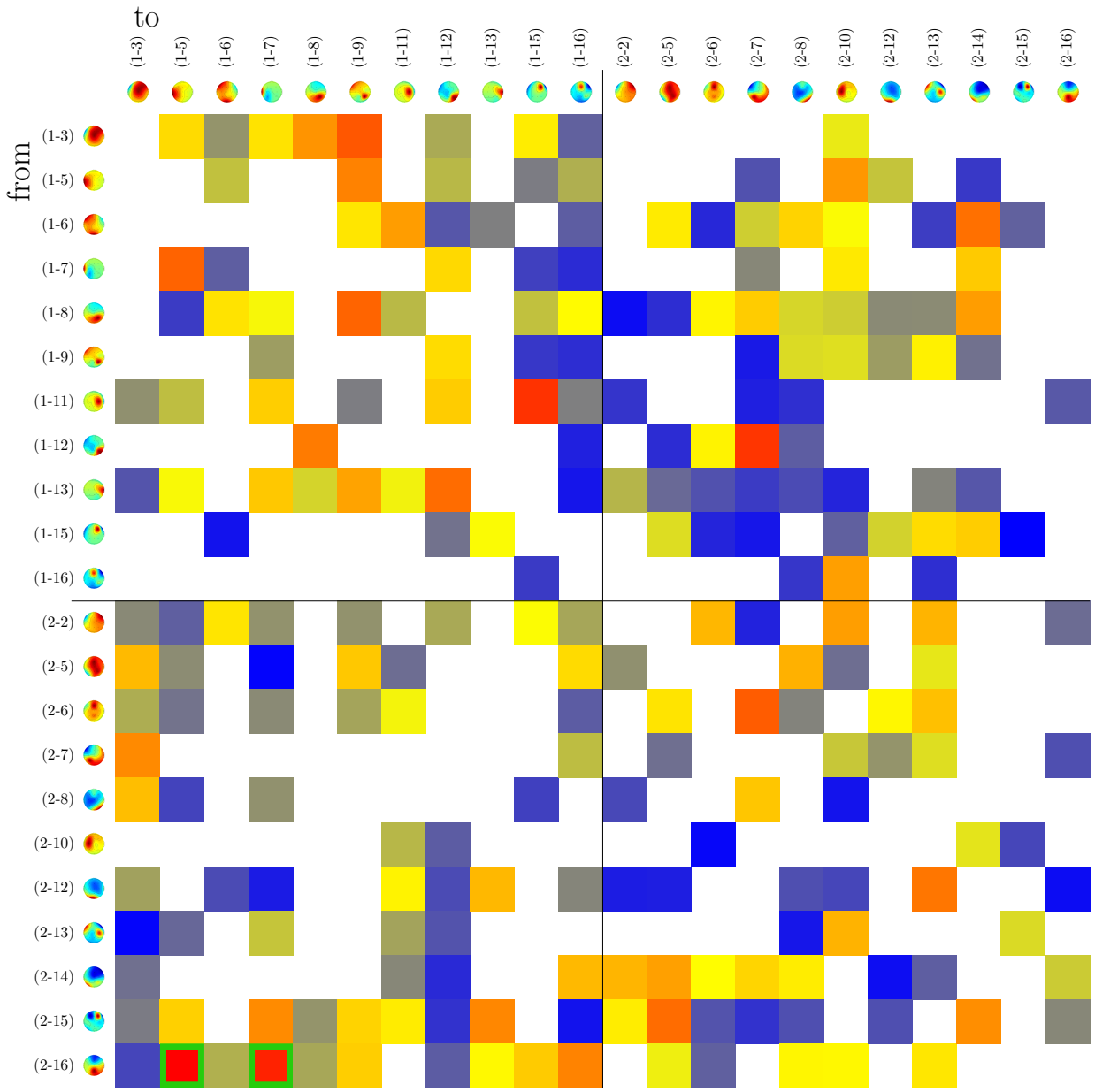


Figure F.2.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment one α-band. Two connections show significant changes. Both hyper-connections from participant two to participant one. For both connections the same fronto-central component of participant one is the sender. The one receiving component of participant one is broader lateral while the other is a strongly lateral component.

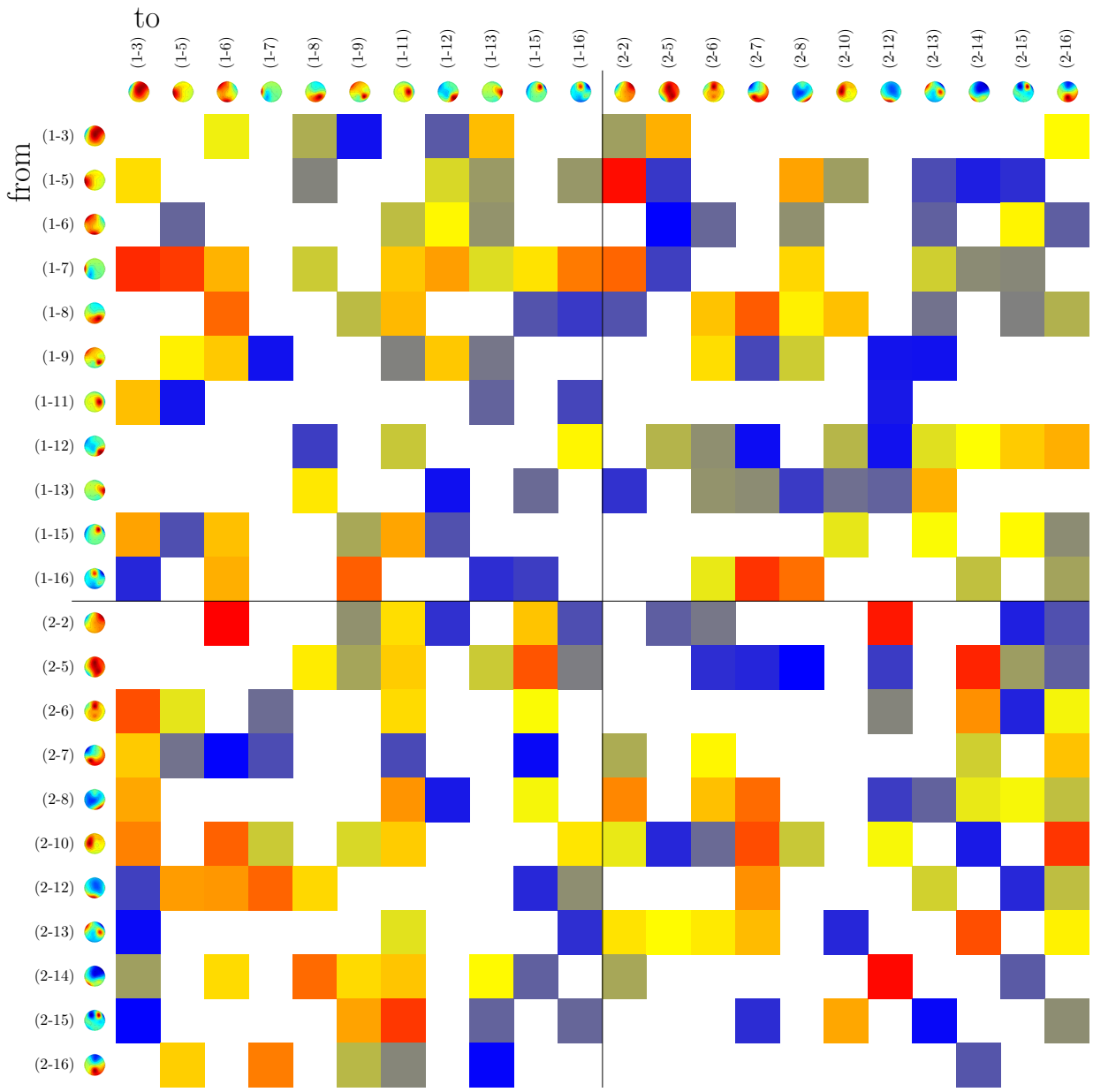


Figure F.3.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment one β-band. Although a number of changes in connectivity have similar high values, none of these reach significance.

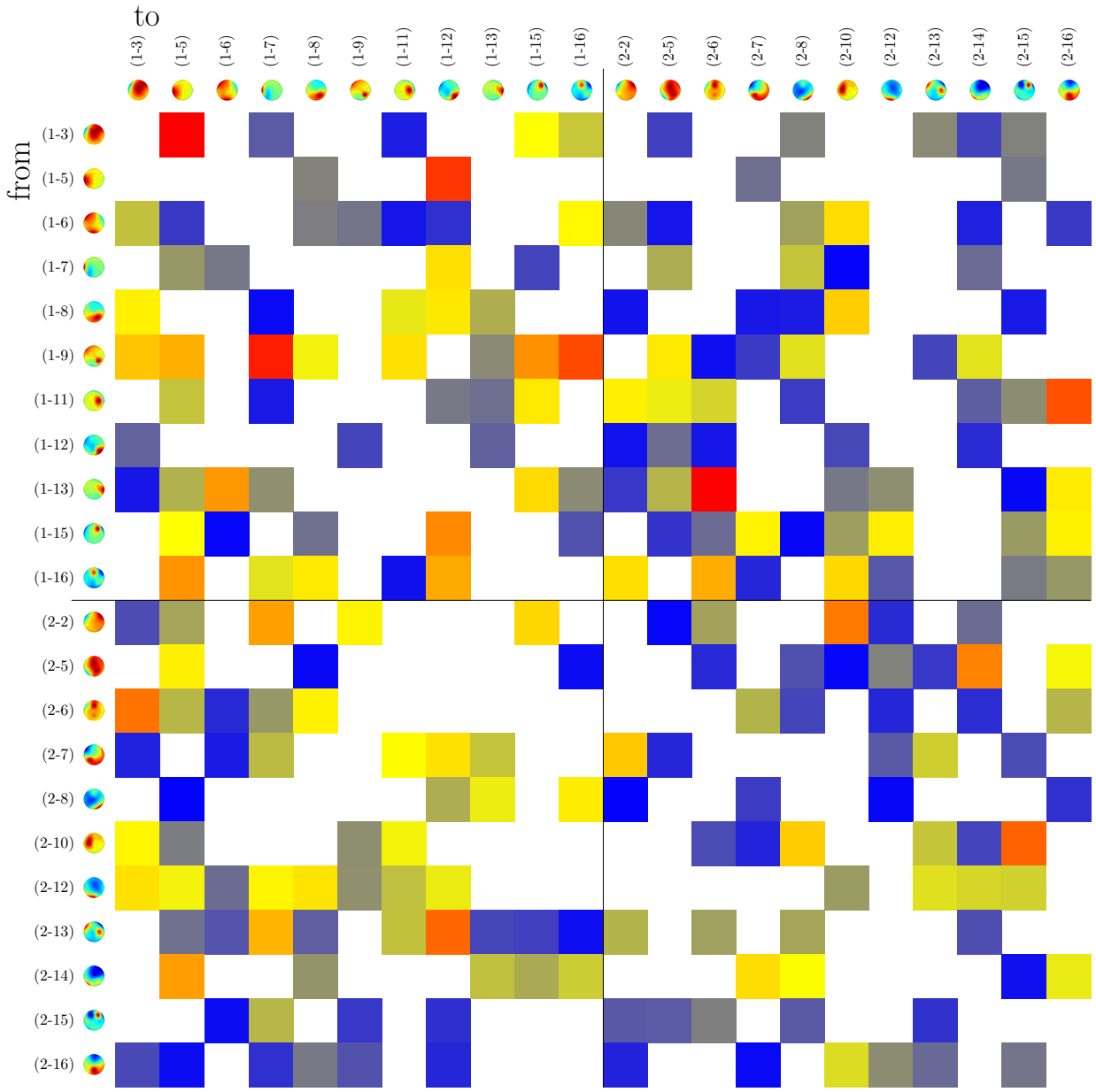


Figure F.4.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment one γ-band. Only few changes in connectivity have high values and none reach significance.

F.2. Experiment Two

No significant changes in connectivity were found for any of the bands.

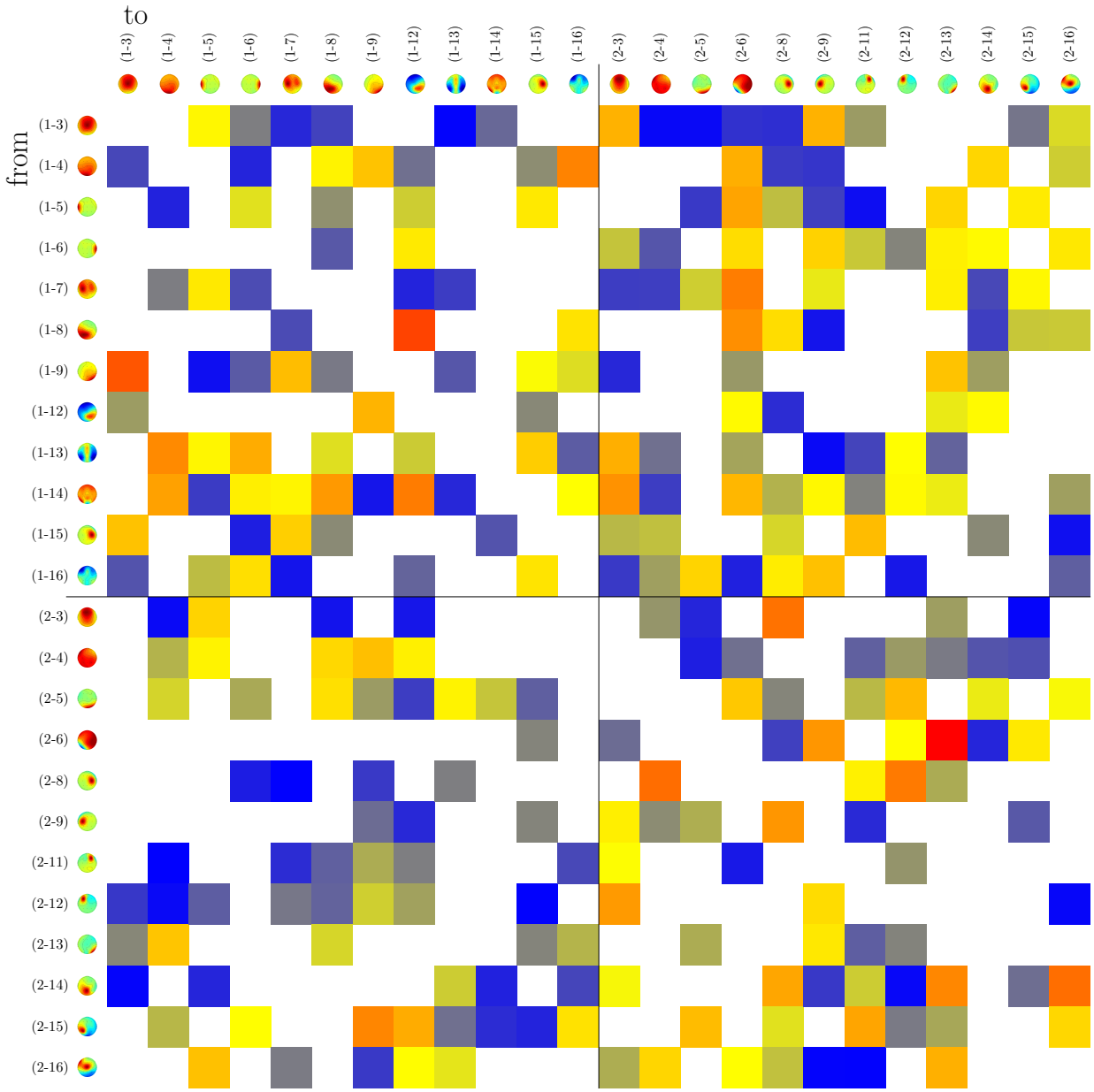


Figure F.5.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment two θ-band. In contrast to experiment one, none of the connectivity changes reaches significance and only very few have high values.

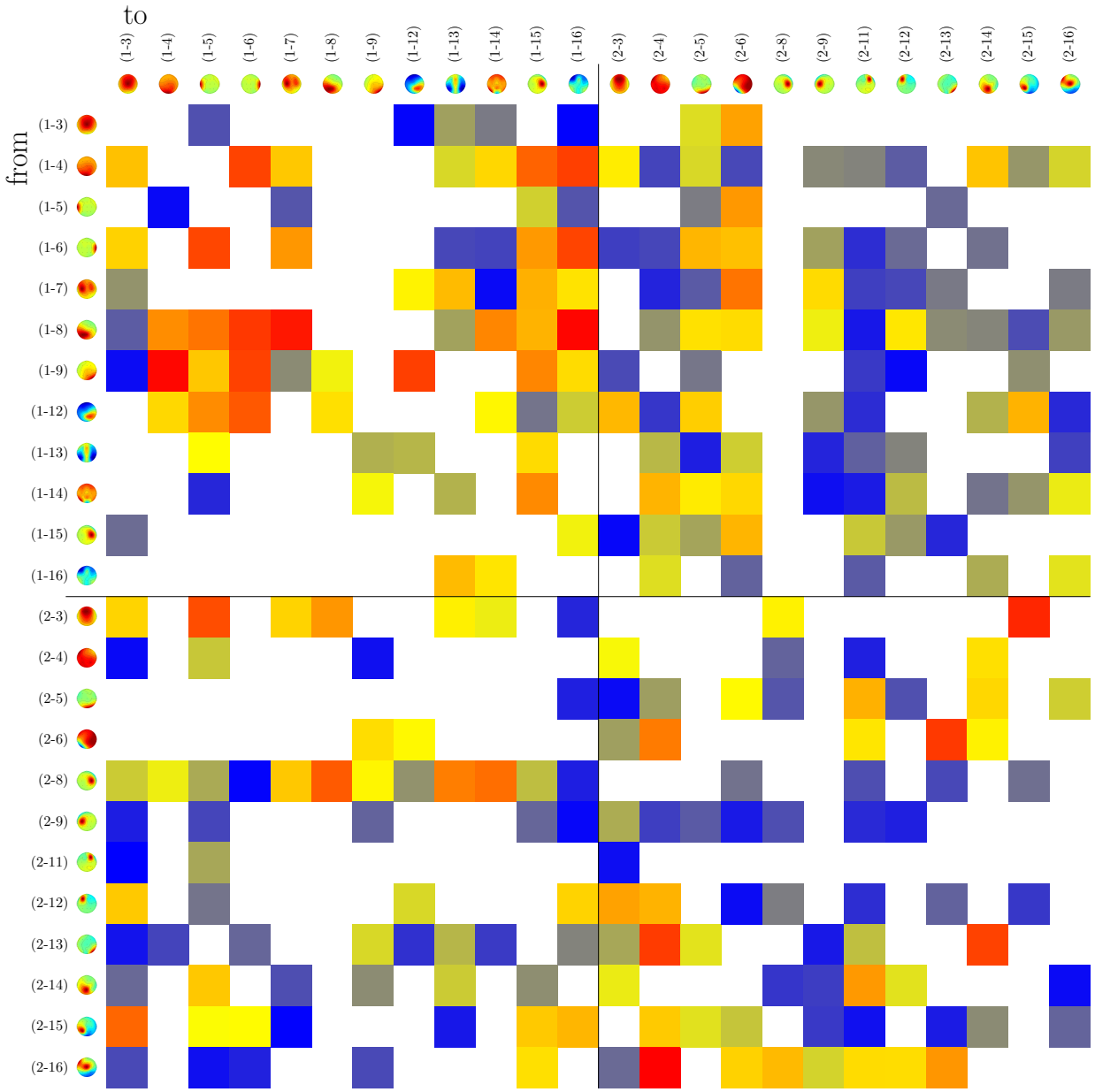


Figure F.6.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment two α-band. No significant changes in any of the connections. It might be remarkable, though, that for one lateral component of participant two (2-8), PSI differences for all hyper-connections it participates in are positive, but only for few of its within-participant connections.

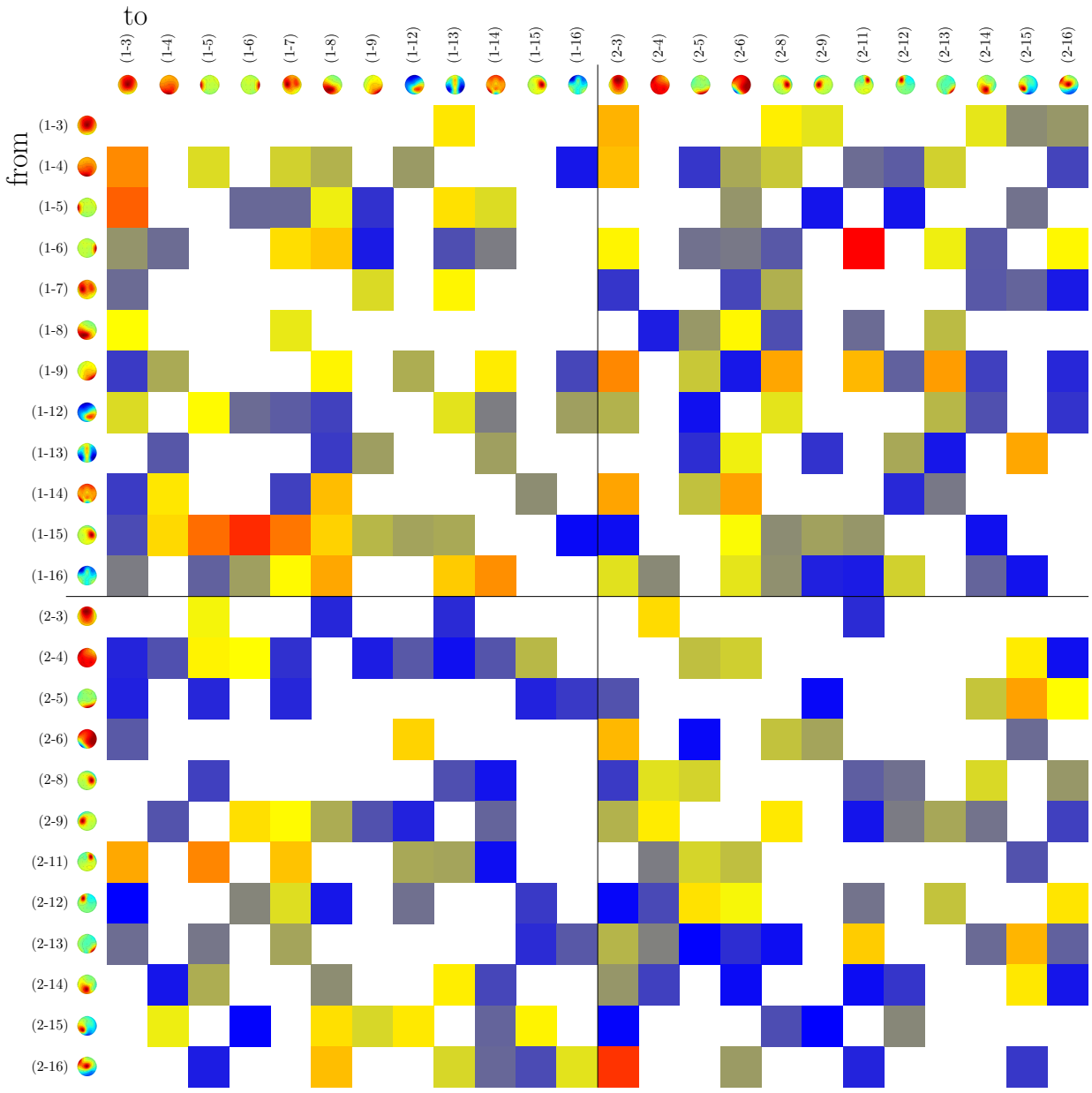


Figure F.7.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment two β-band. Only few connectivity changes reach high values and none reaches significance.

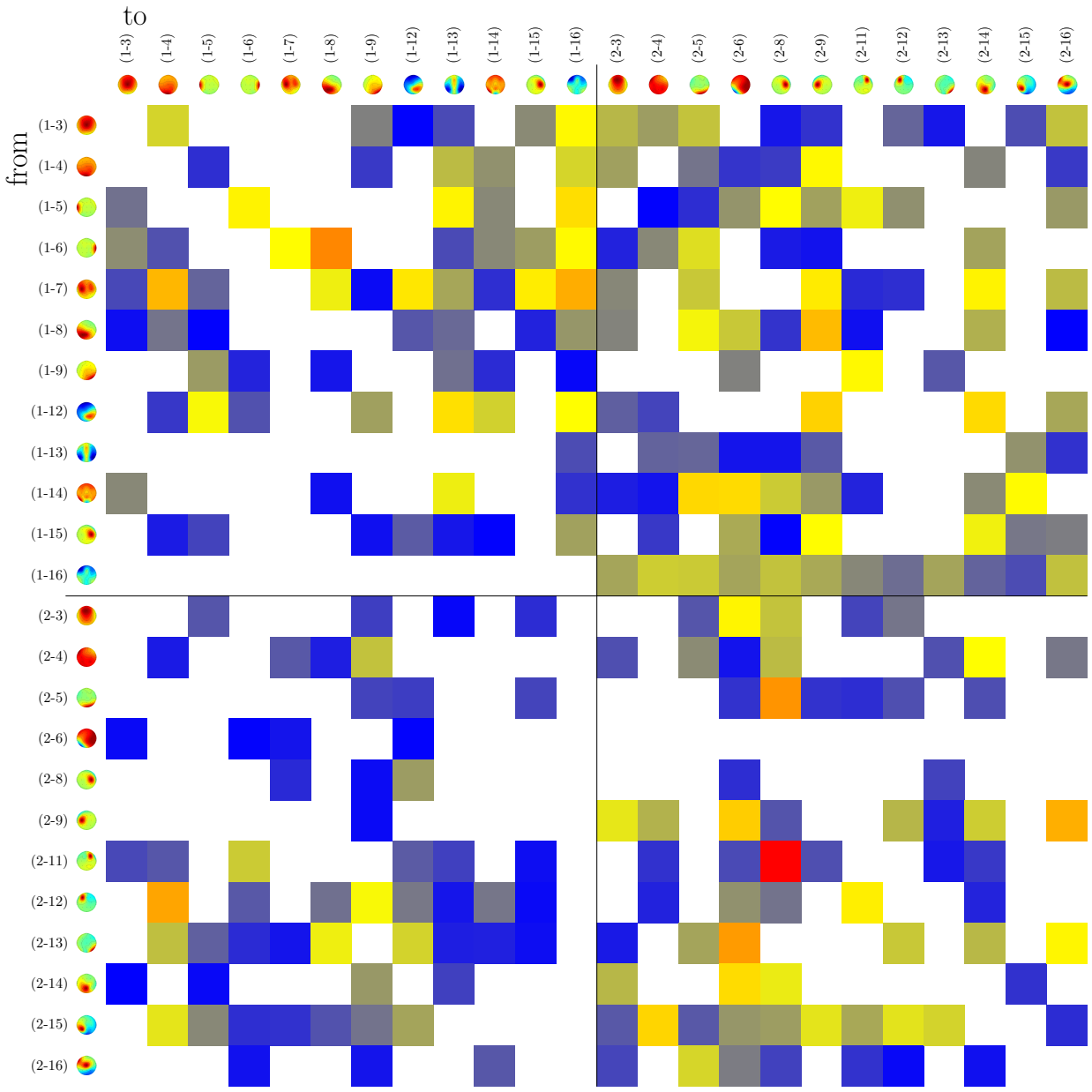


Figure F.8.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment two γ-band. No significant changes in any connection.

F.3. Experiment Three

Significant changes in three connections. All of these are hyper-connections in the γ-band.

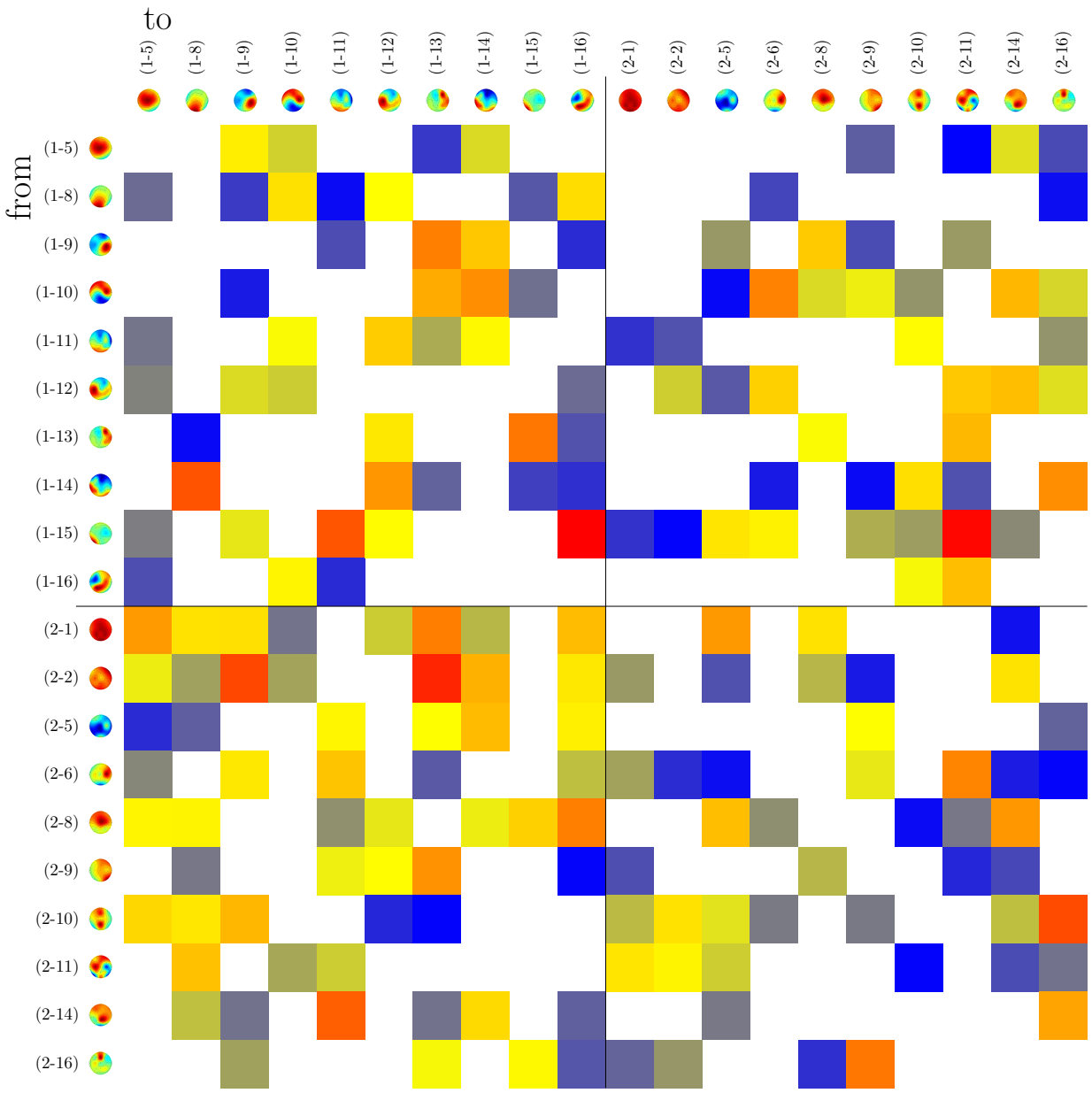


Figure F.9.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment three θ-band. None of the changes reached significance.

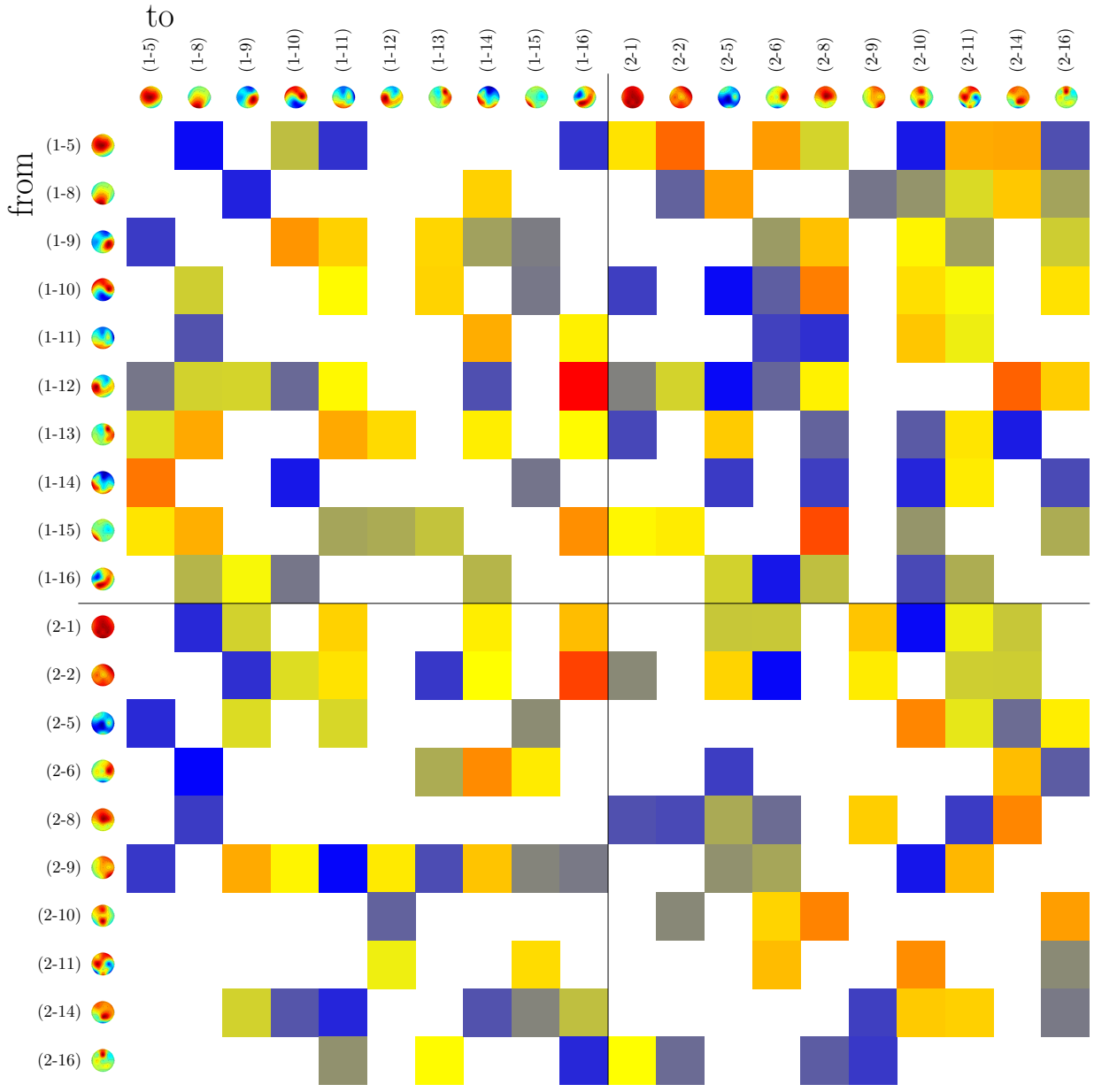


Figure F.10.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment three α-band. None of the changes reached significance.

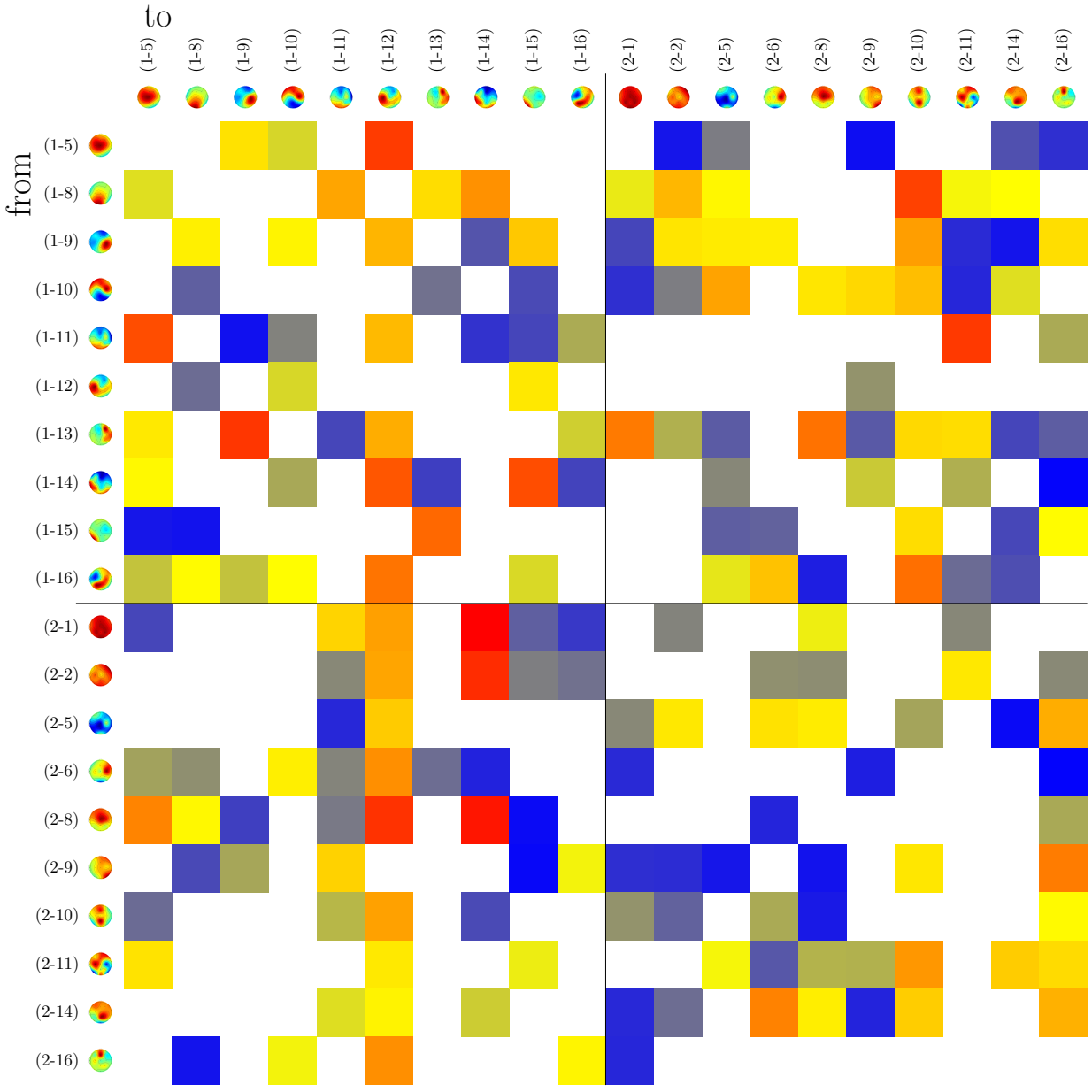


Figure F.11.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment three β-band. None of the changes reached significance.

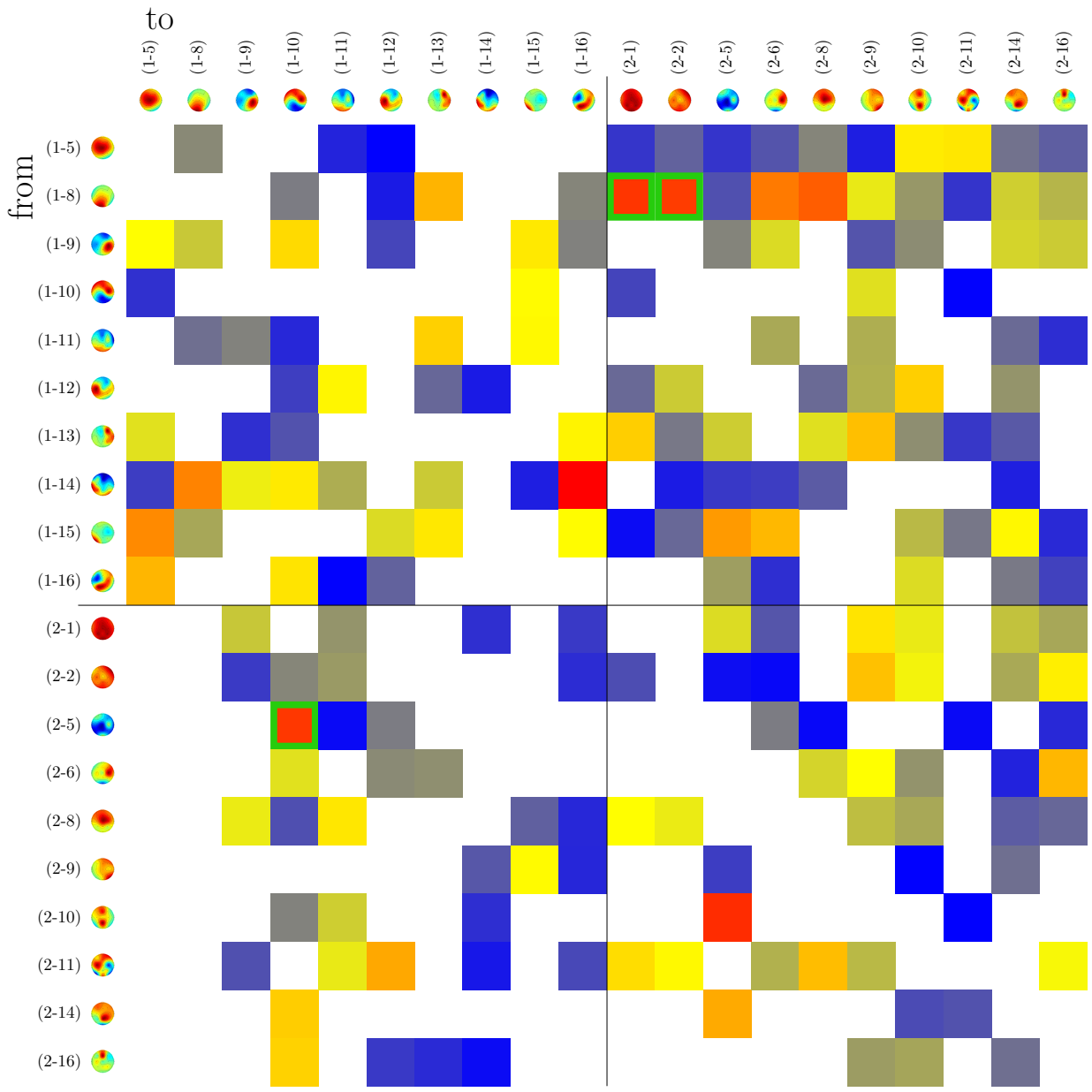


Figure F.12.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment three γ-band. Three significant changes in hyper-connections were found. It is also remarkable that the first two components of participant one all have positive PSI value in relation to participant two's components. Two of these connections, connecting an occipital component of participant one with rather global components of participant two showed significant changes. Furthermore a connection from one frontal component of participant two to a global component of participant one changed significantly.

F.4. Experiment Four

There is little conclusive in the cooperative/non-cooperative PSI analysis results for this data set.

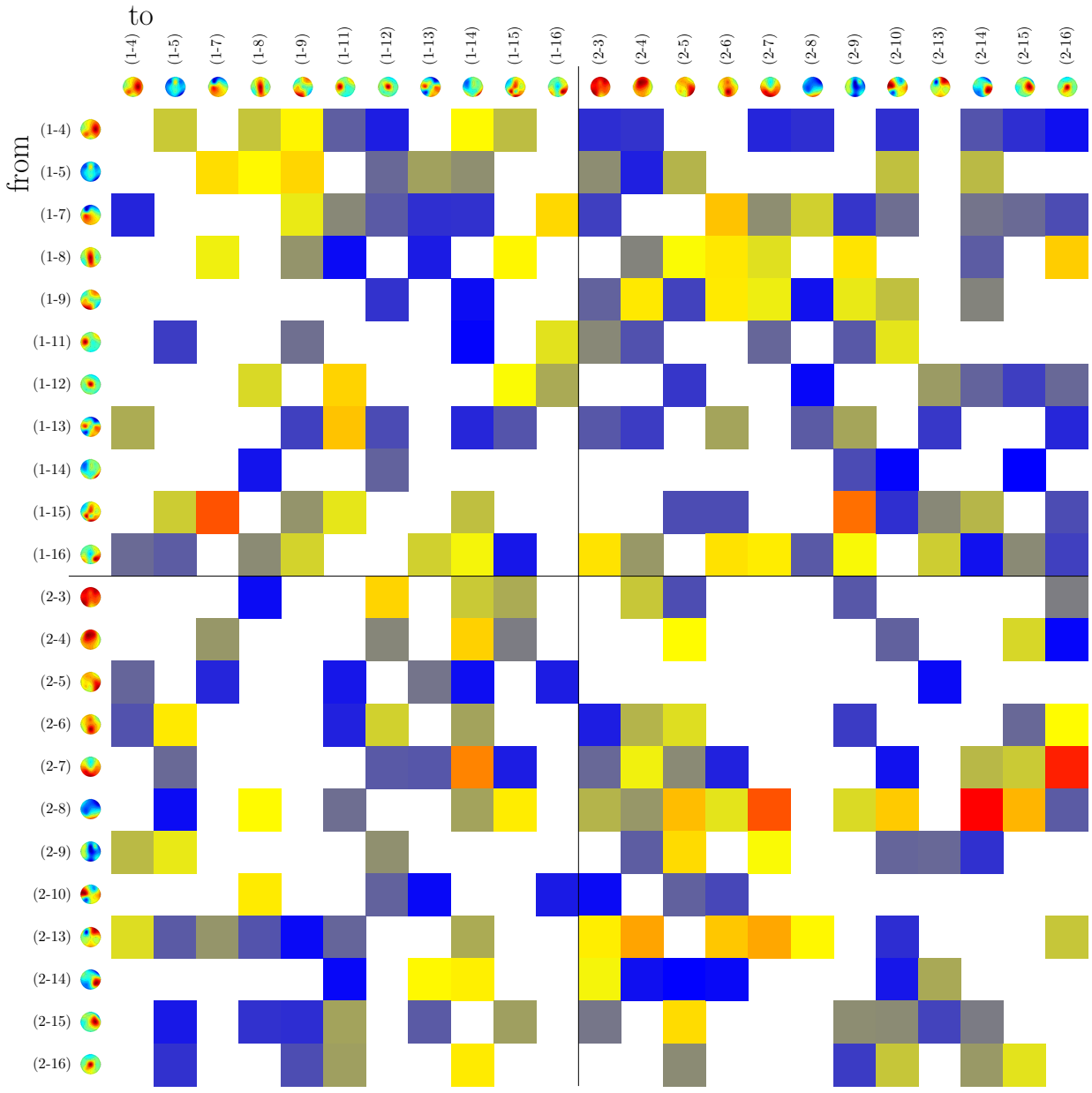


Figure F.13.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment four θ-band. None of the changes reaches significance.

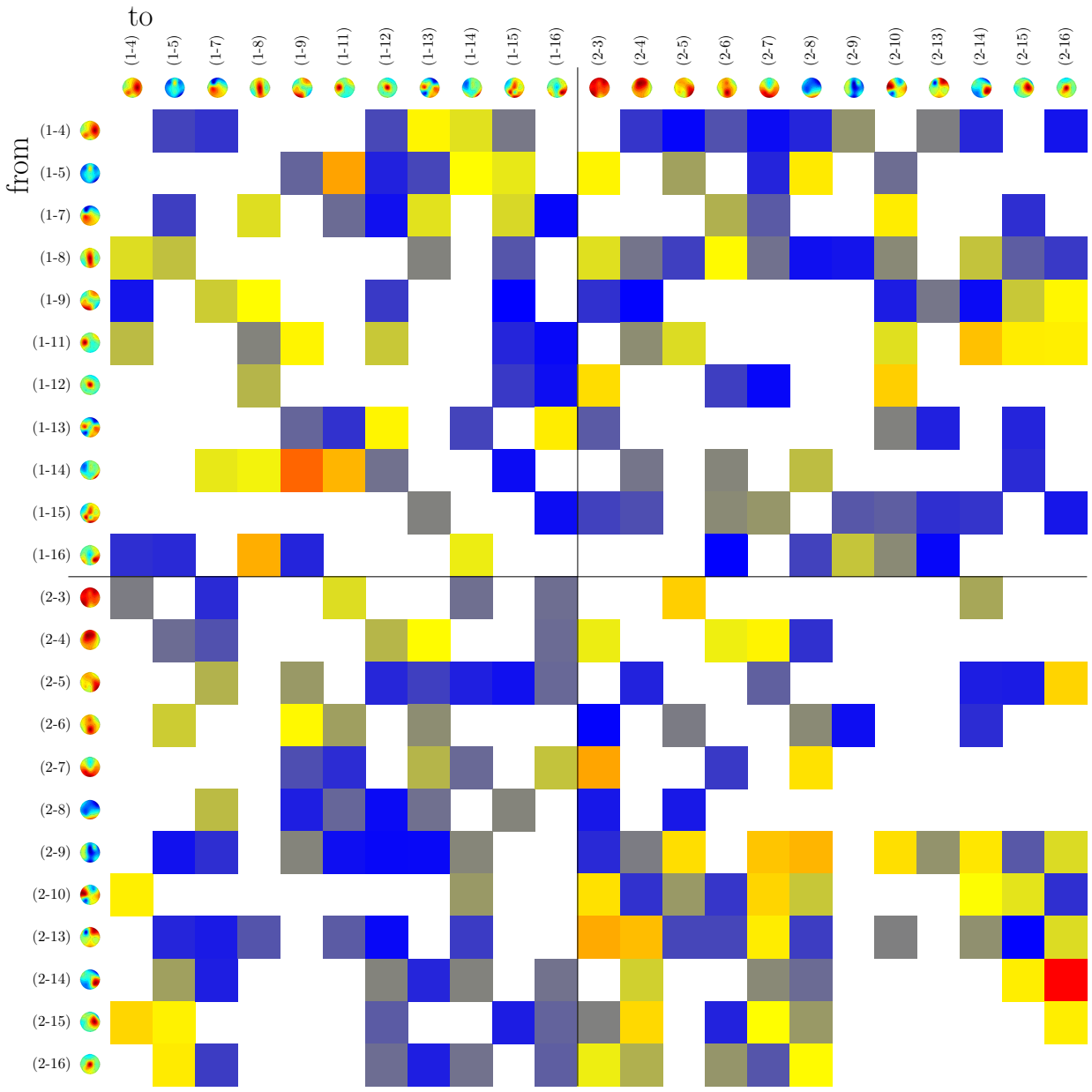


Figure F.14.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment four α-band. One change is noticeably stronger than all others. None are significant.

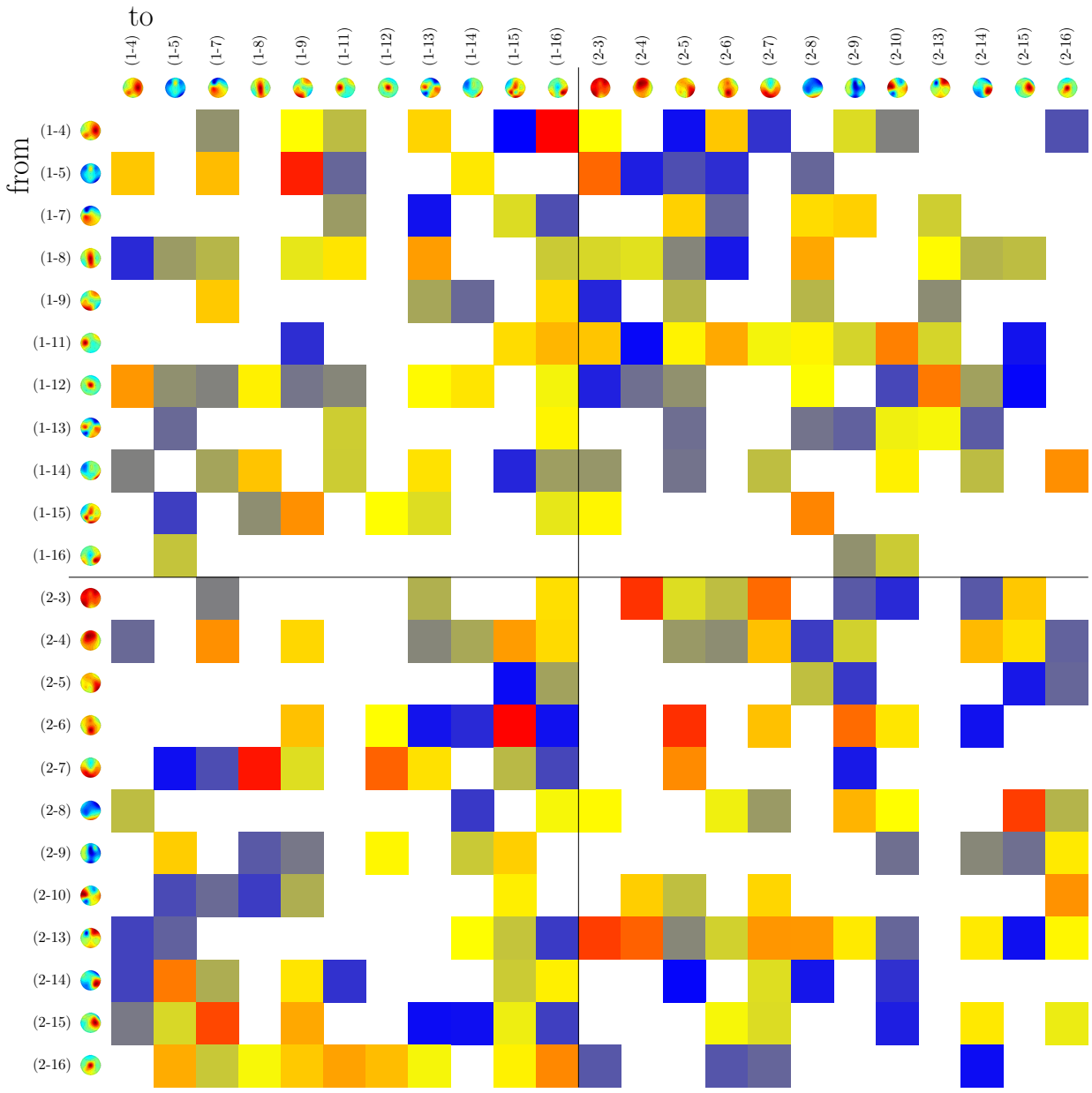


Figure F.15.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment four β-band. No significant changes and no salient patterns in the distribution of connectivity.

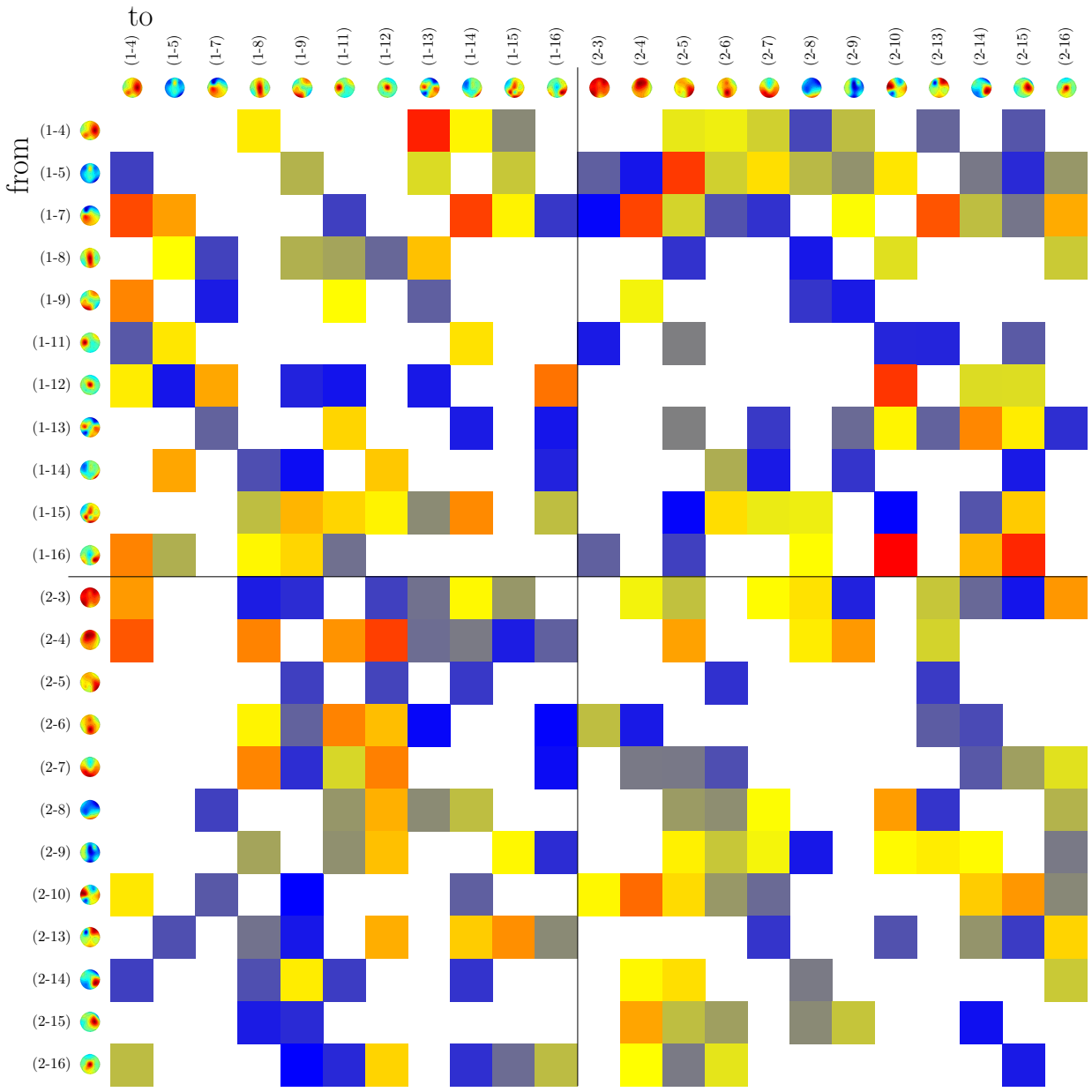


Figure F.16.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment four γ-band. Again no significant changes nor salient patterns.

F.5. Experiment Five

No significant changes could be found in this analysis.

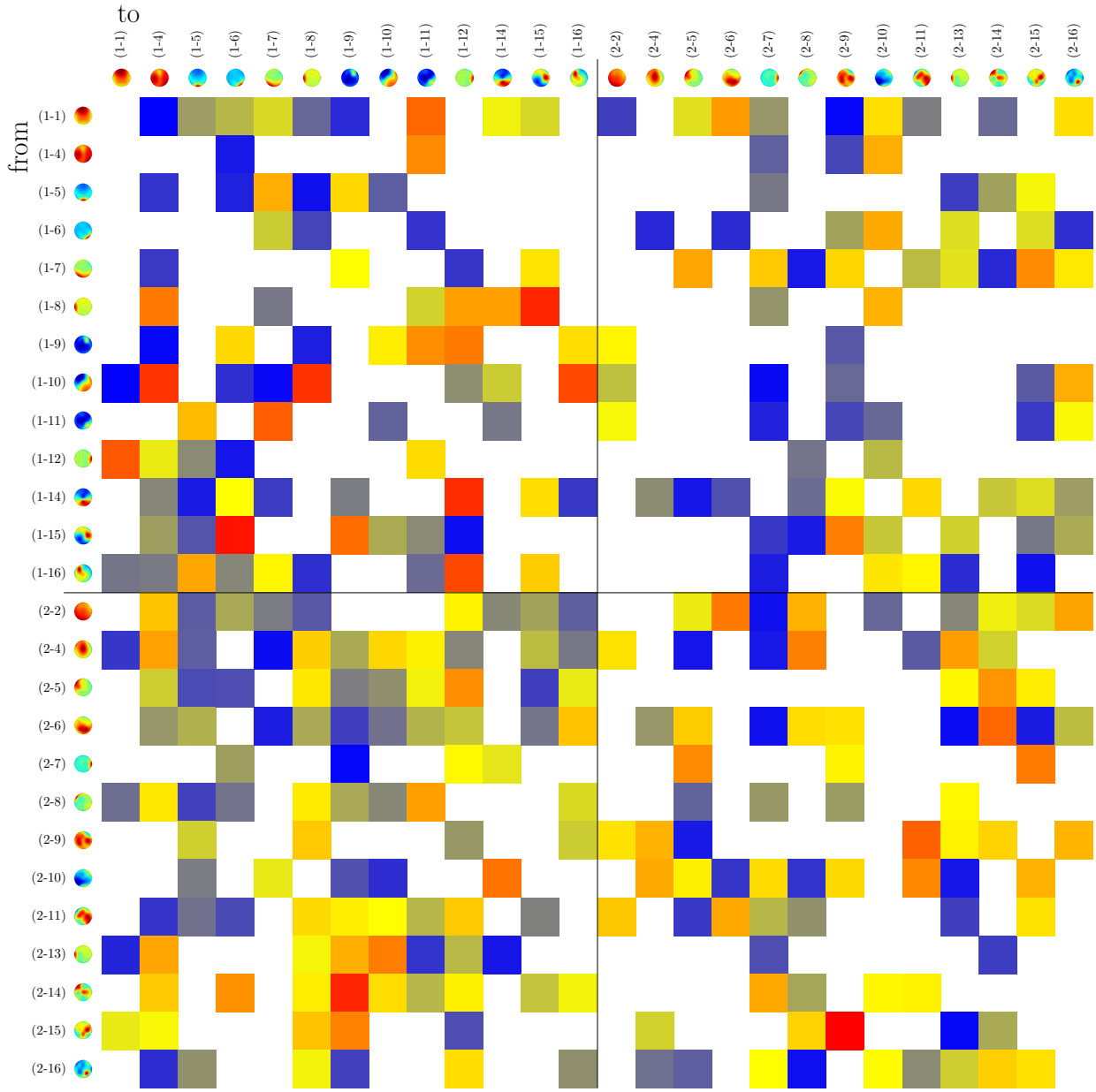


Figure F.17.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment five θ-band. None of the changes reaches significance.

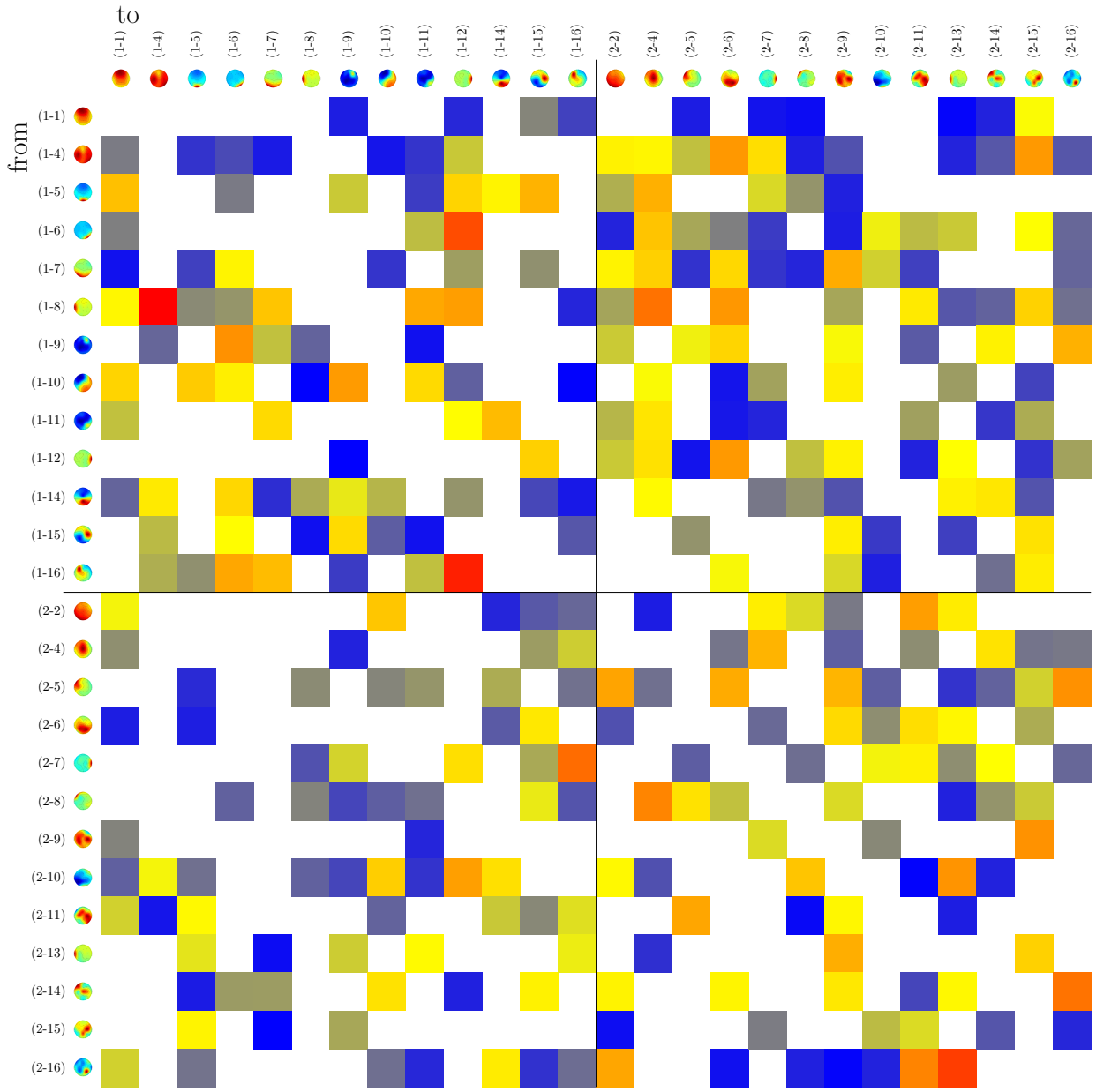


Figure F.18.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment five α-band. No significant changes were found.

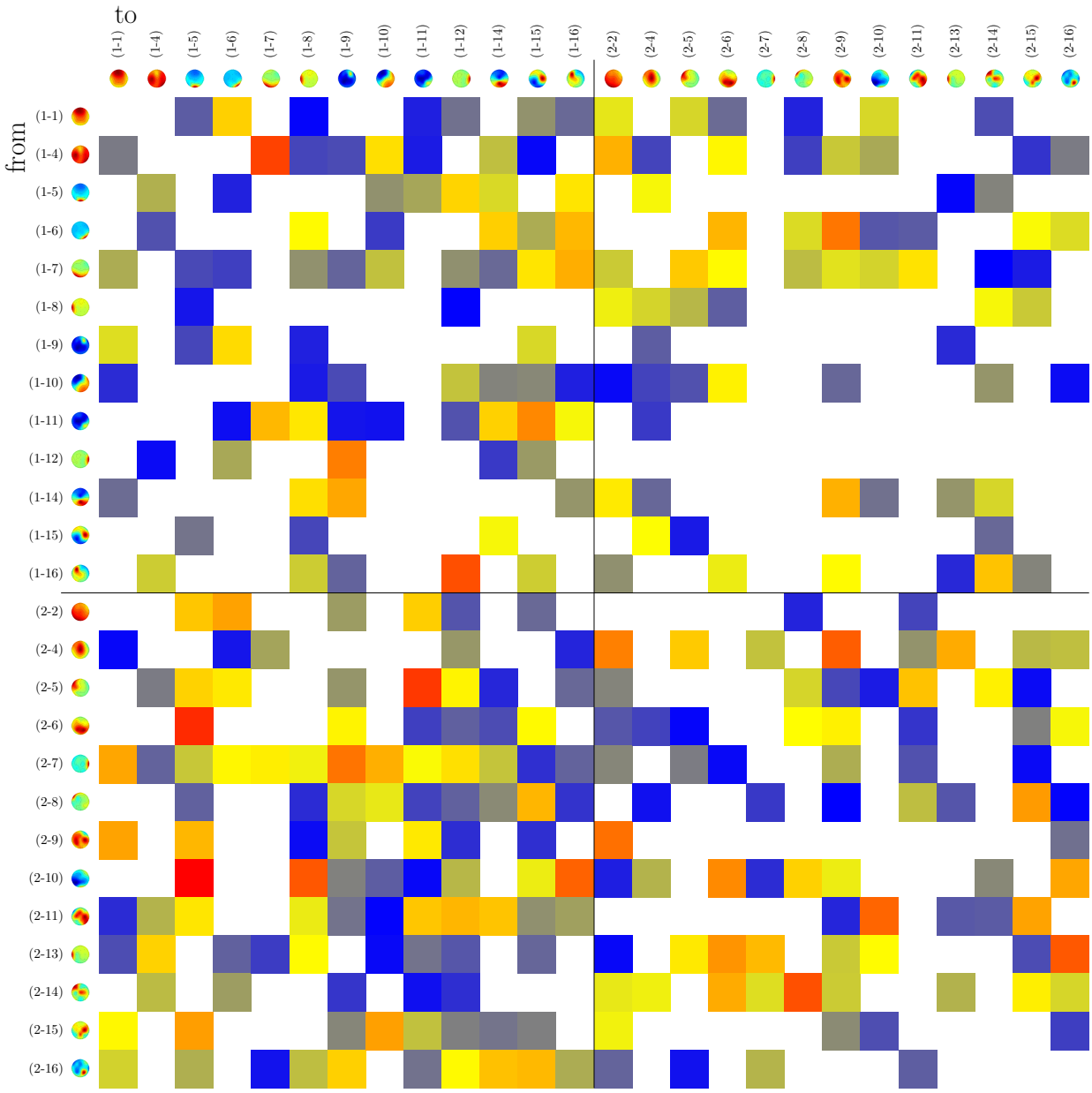


Figure F.19.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment five β-band. No significant changes were found.

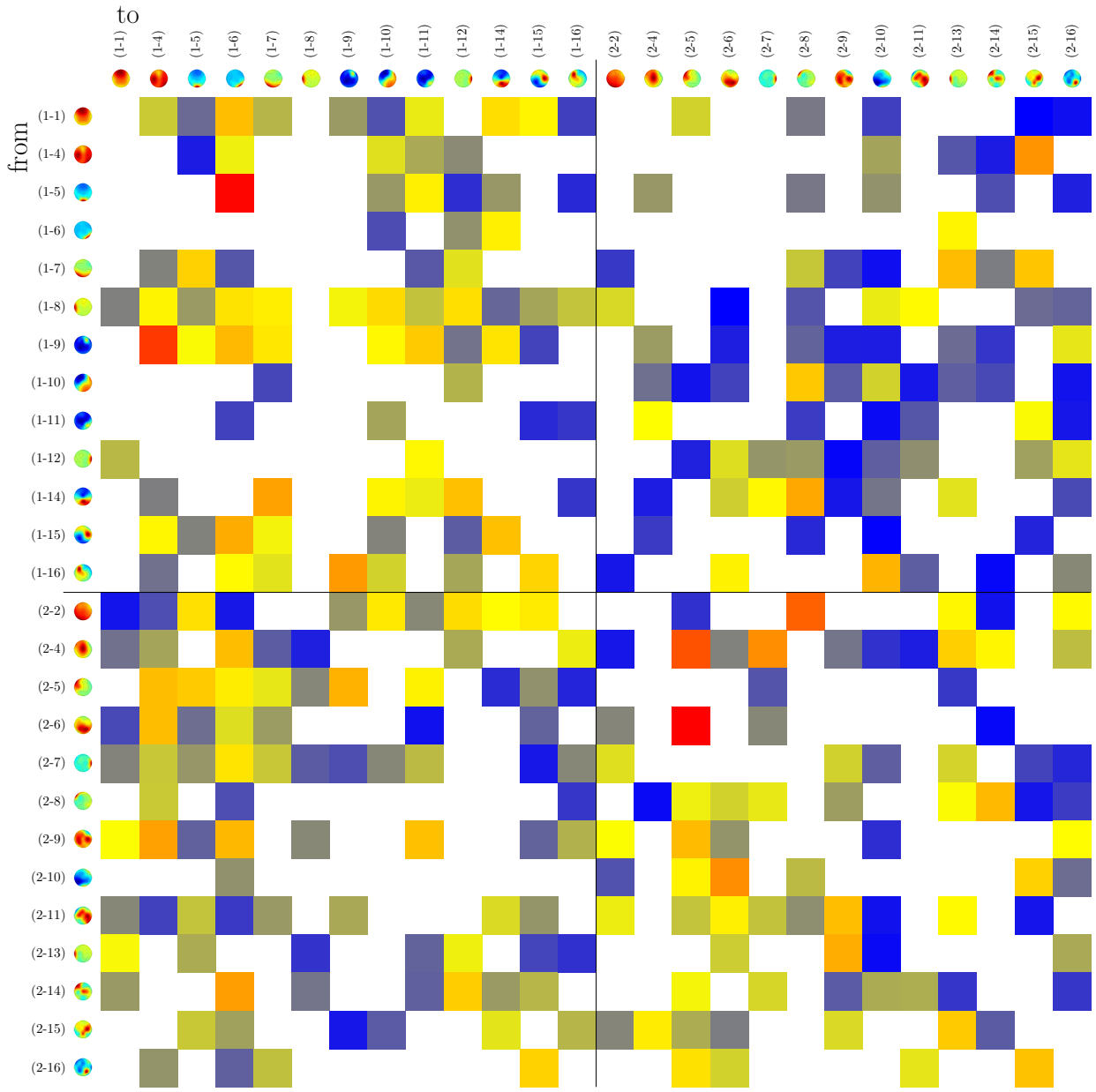


Figure F.20.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment five γ-band. None of the changes is significant.

F.6. Experiment Six

Three connections show significant changes. Two of the affected connections are hyper-connections.

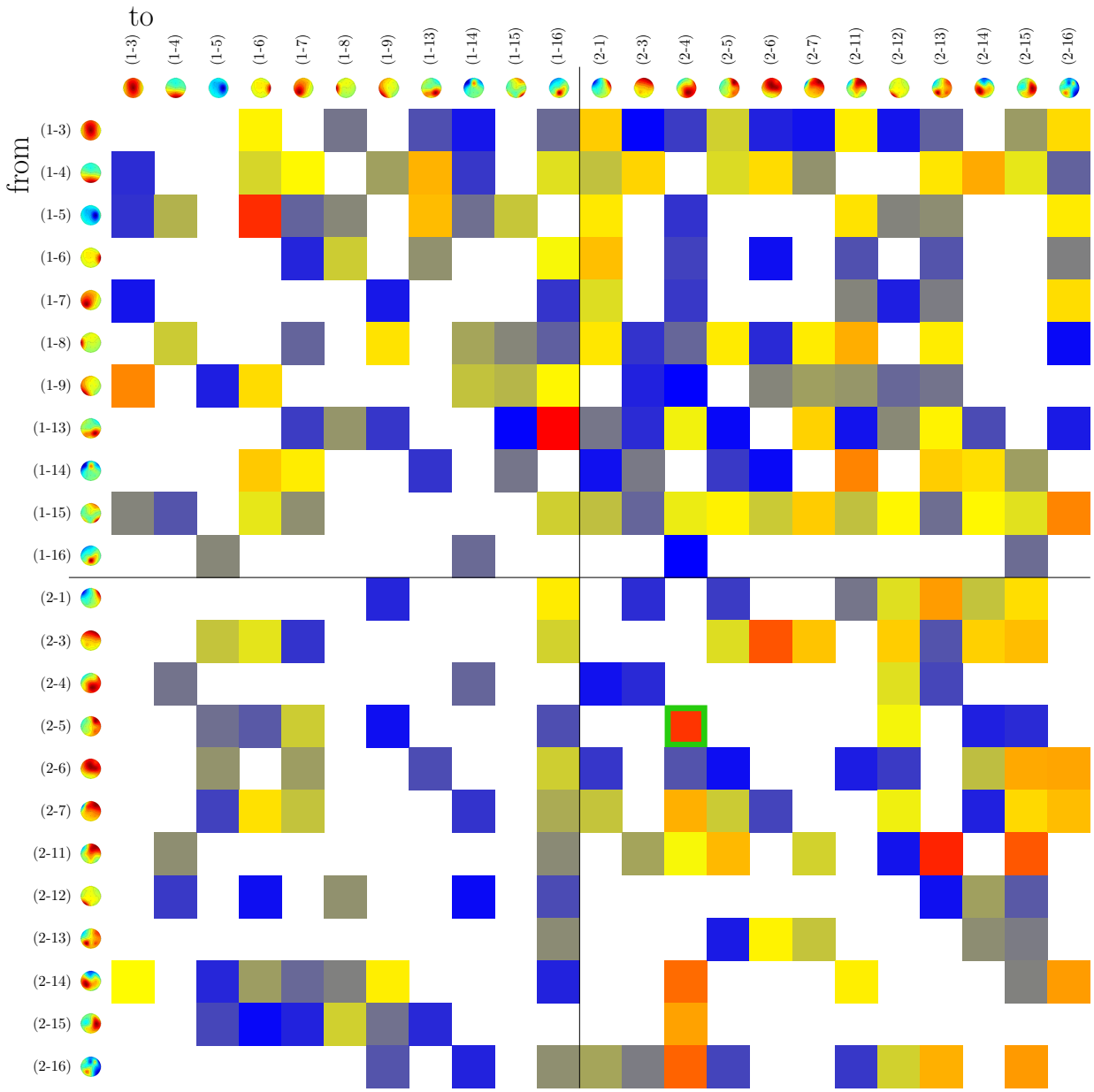


Figure F.21.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment six θ-band. One significant change can be found for a within-participant connection (participant two) from a lateral component towards a broad parietal component with a slight bias to the right.

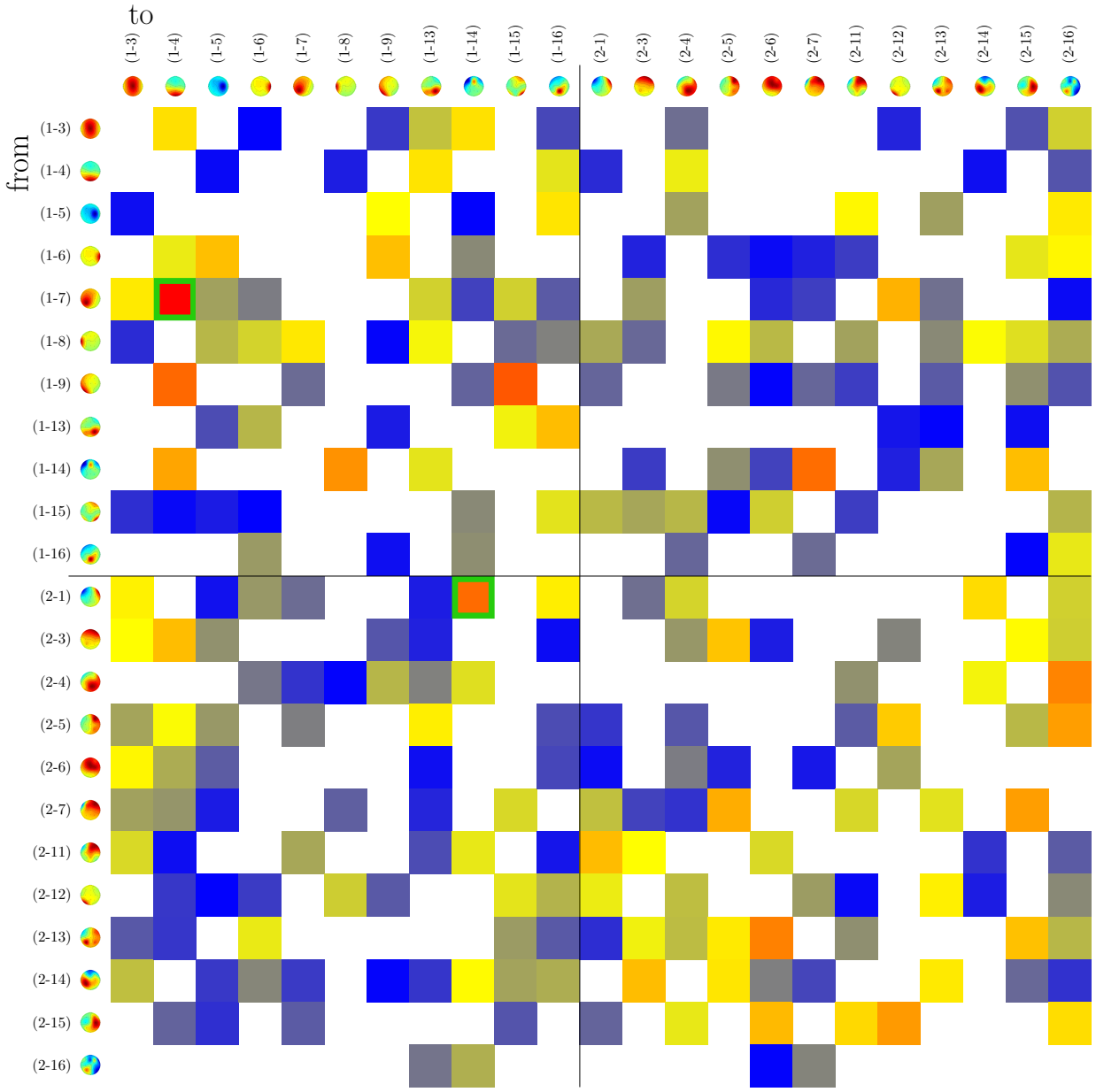


Figure F.22.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment six α-band. Two significant changes: One for a within-participant connection of a broad parietal-lateral component towards a strongly occipital component of participant one. The second for a hyper-connection from a strongly fronto-lateral component of participant two towards a very narrow frontal component of participant one.

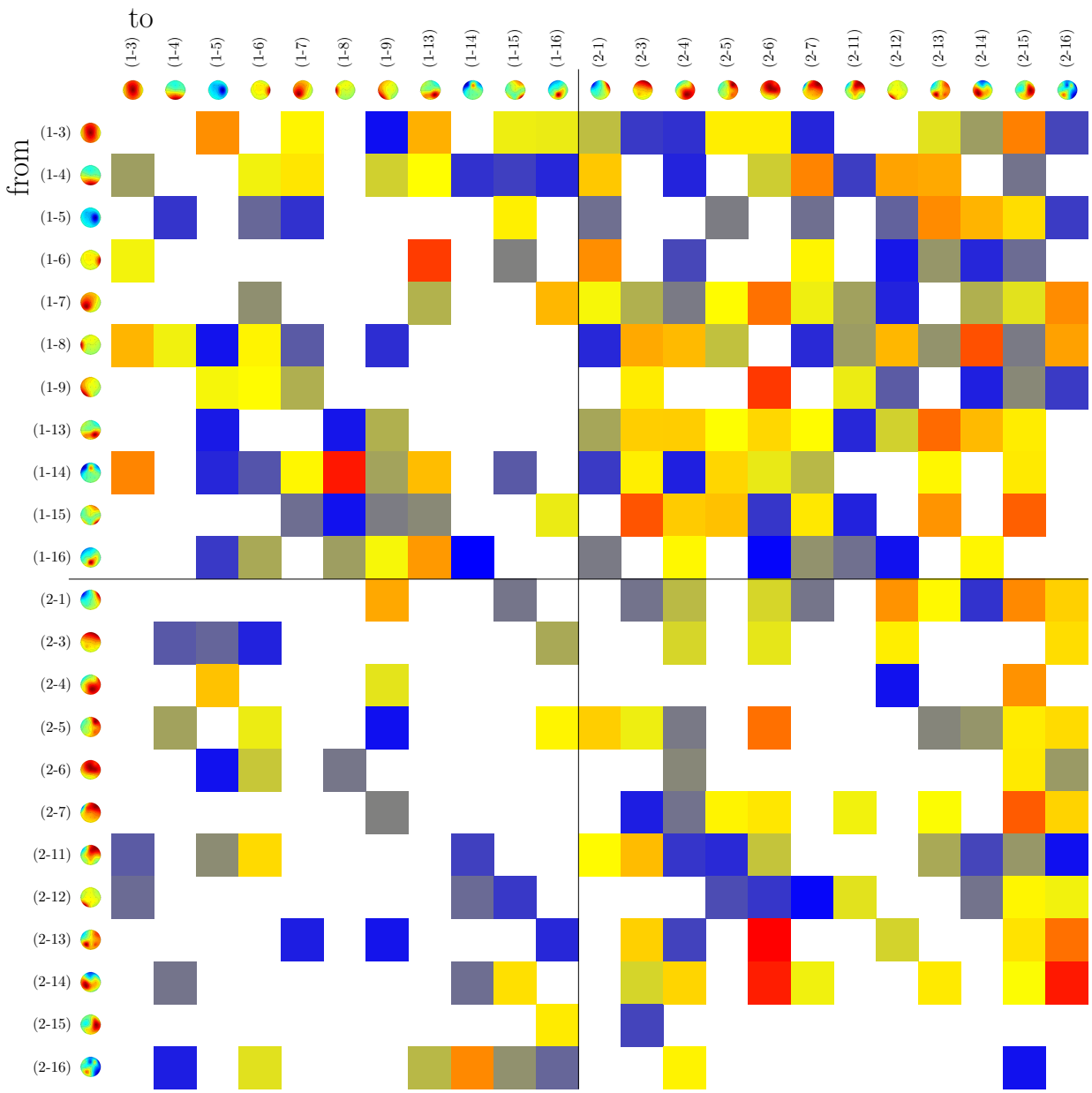


Figure F.23.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment six β-band. Regarding hyper-connections there seems to be a strong bias for connection changes for positive values for connections from participant one to participant two. However, none of the changes in connectivity reaches significance.

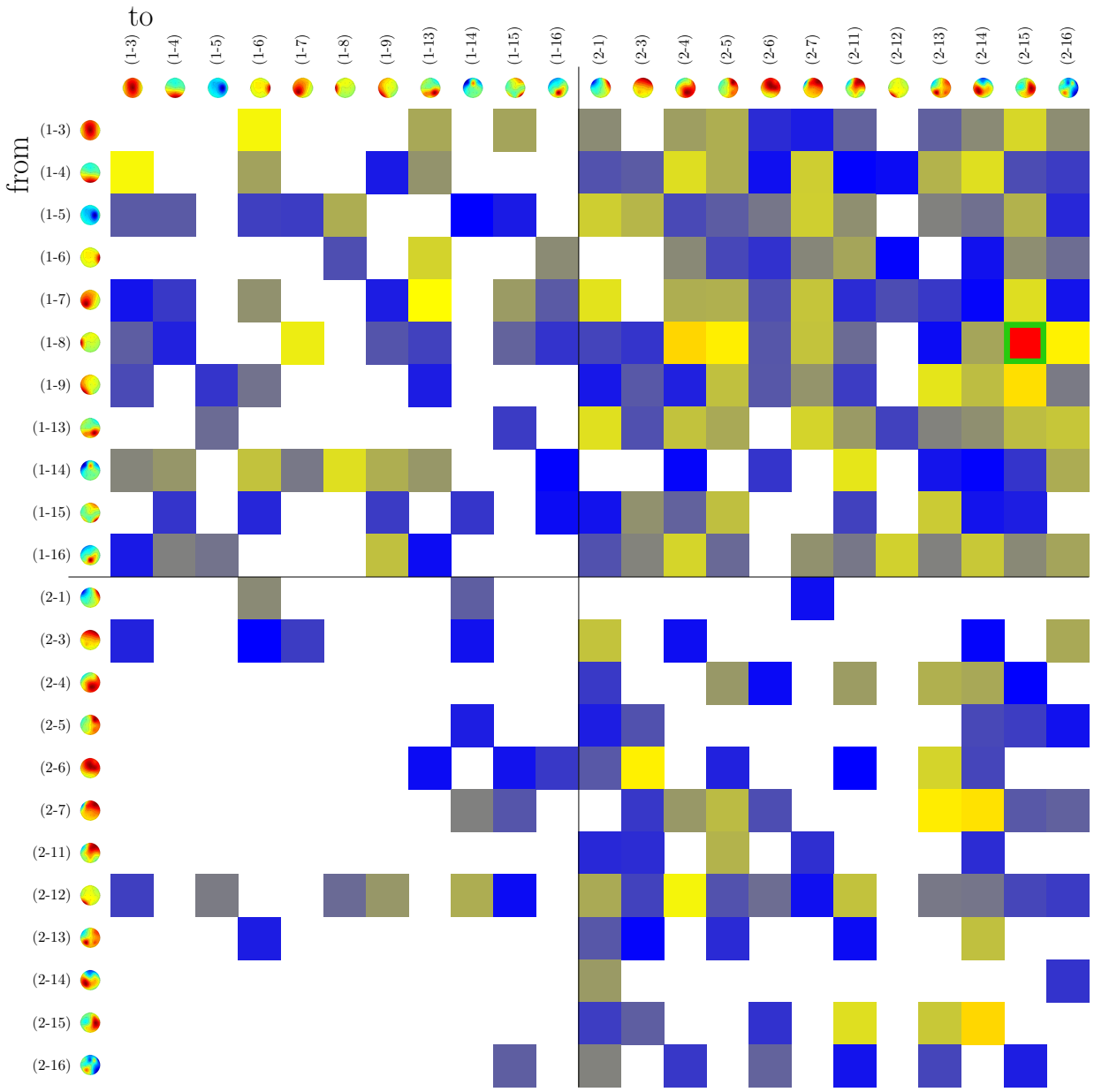


Figure F.24.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment six γ-band. This is pretty remarkable because there is a strong bias for positive value for connections from participant one to participants two among hyper-connections. Furthermore all values are pretty low. At least when compared to the one PSI value for a hyper-connection from a strongly lateral component of participant one to a contra-lateral component of participant two. This is also the only change which reaches significance.

F.7. Experiment Seven

In β-band an interesting bias among hyper-connections can be found, but no change reaches significance.

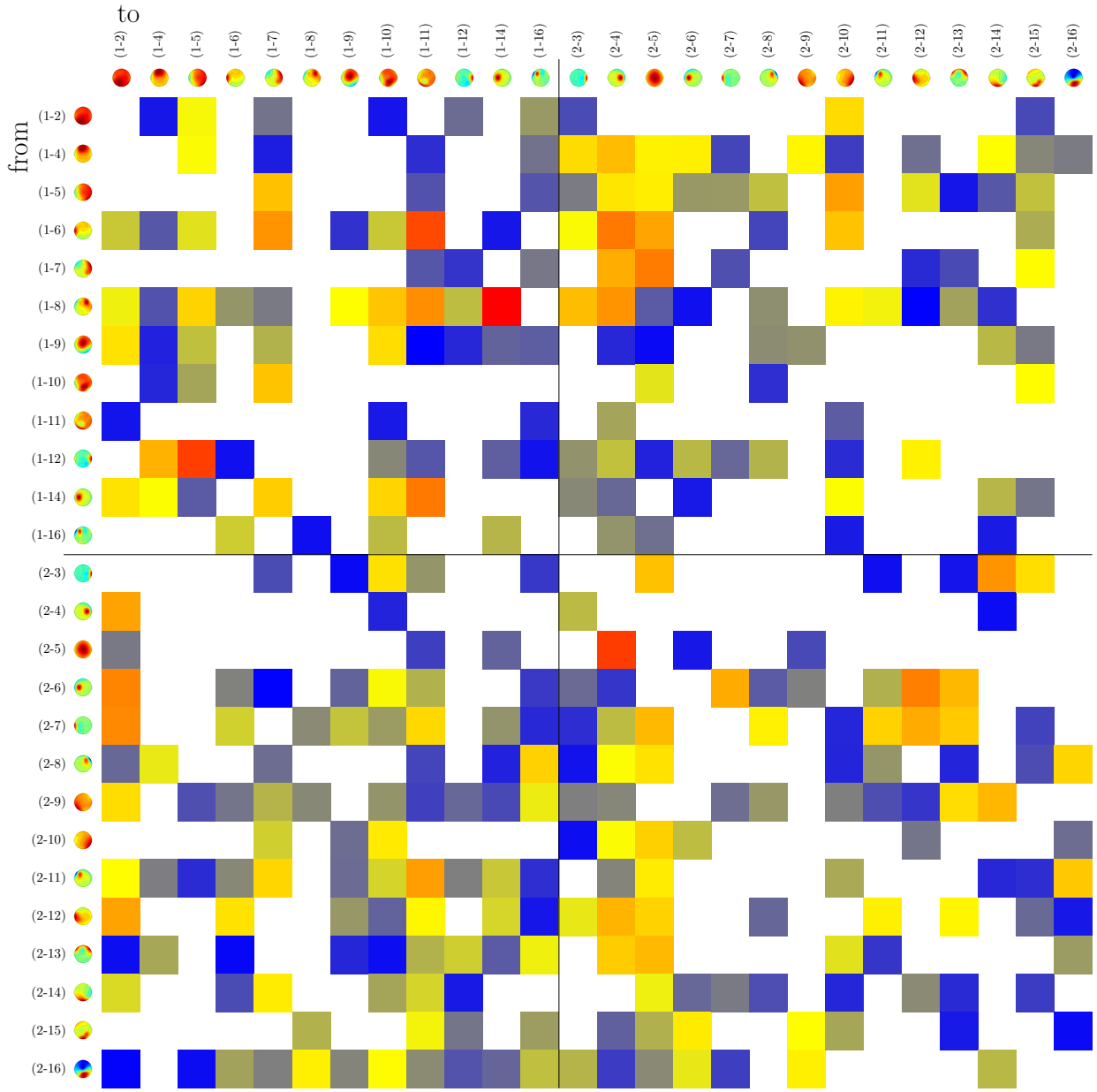


Figure F.25.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment seven θ-band. No significant changes and no interesting pattern in this plot.

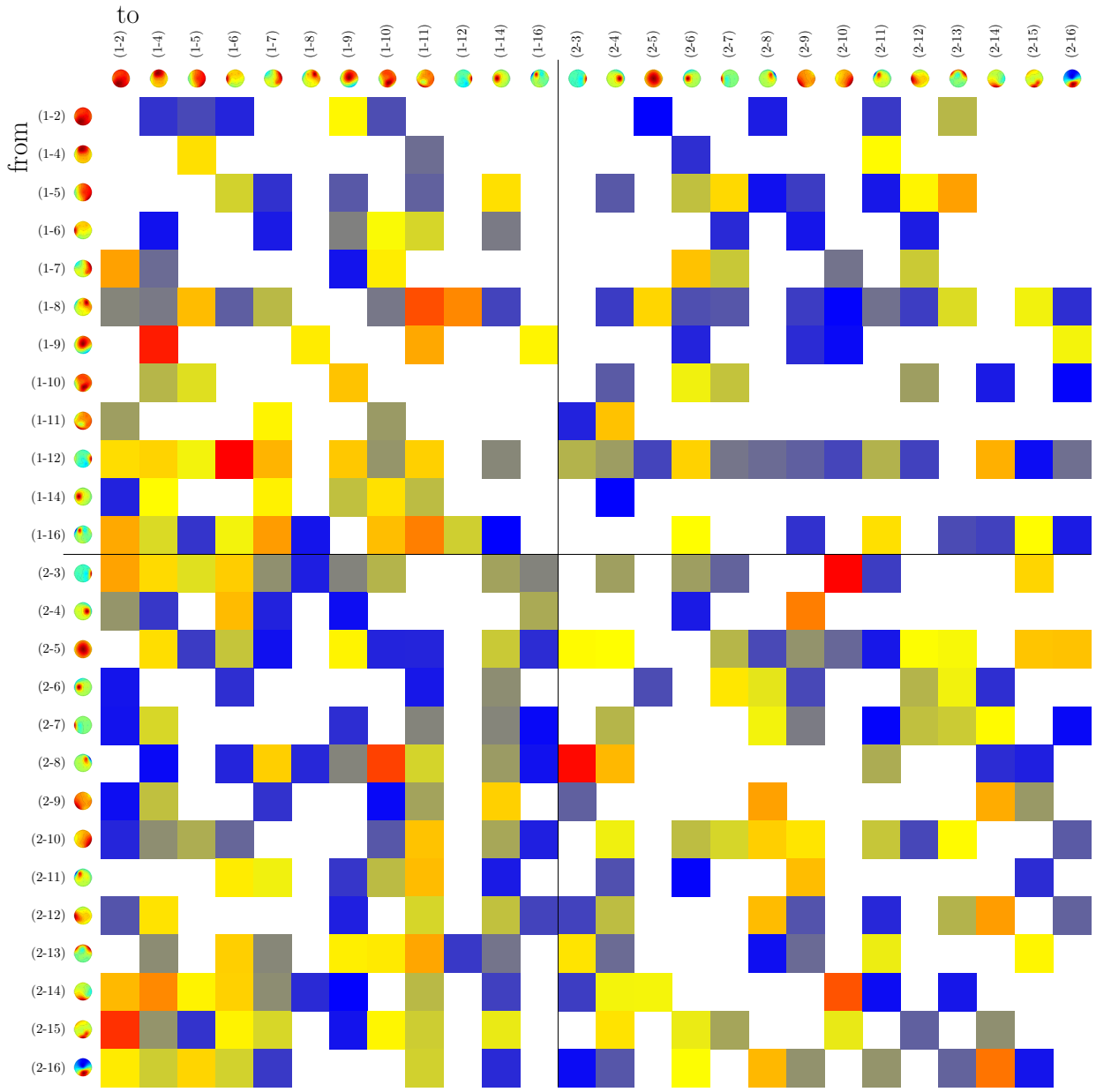


Figure F.26.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment seven α-band. Again, there are no significant changes nor interesting patterns of biases which can be identified.

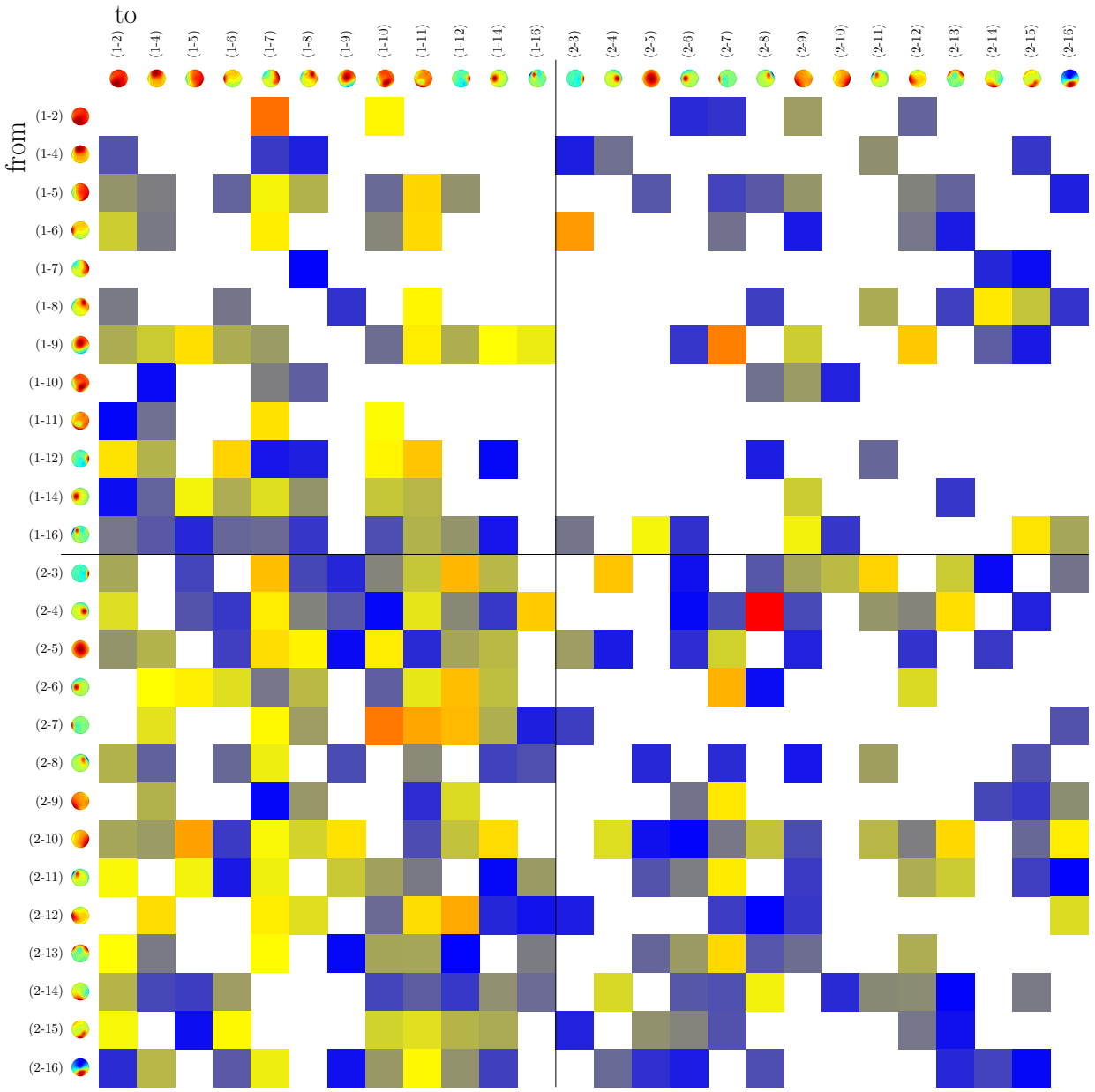


Figure F.27.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment seven β-band. Here a strong bias among hyper-connections can be found, favouring positive value for connections from participant two to participant one. However, none of these changes is significant.

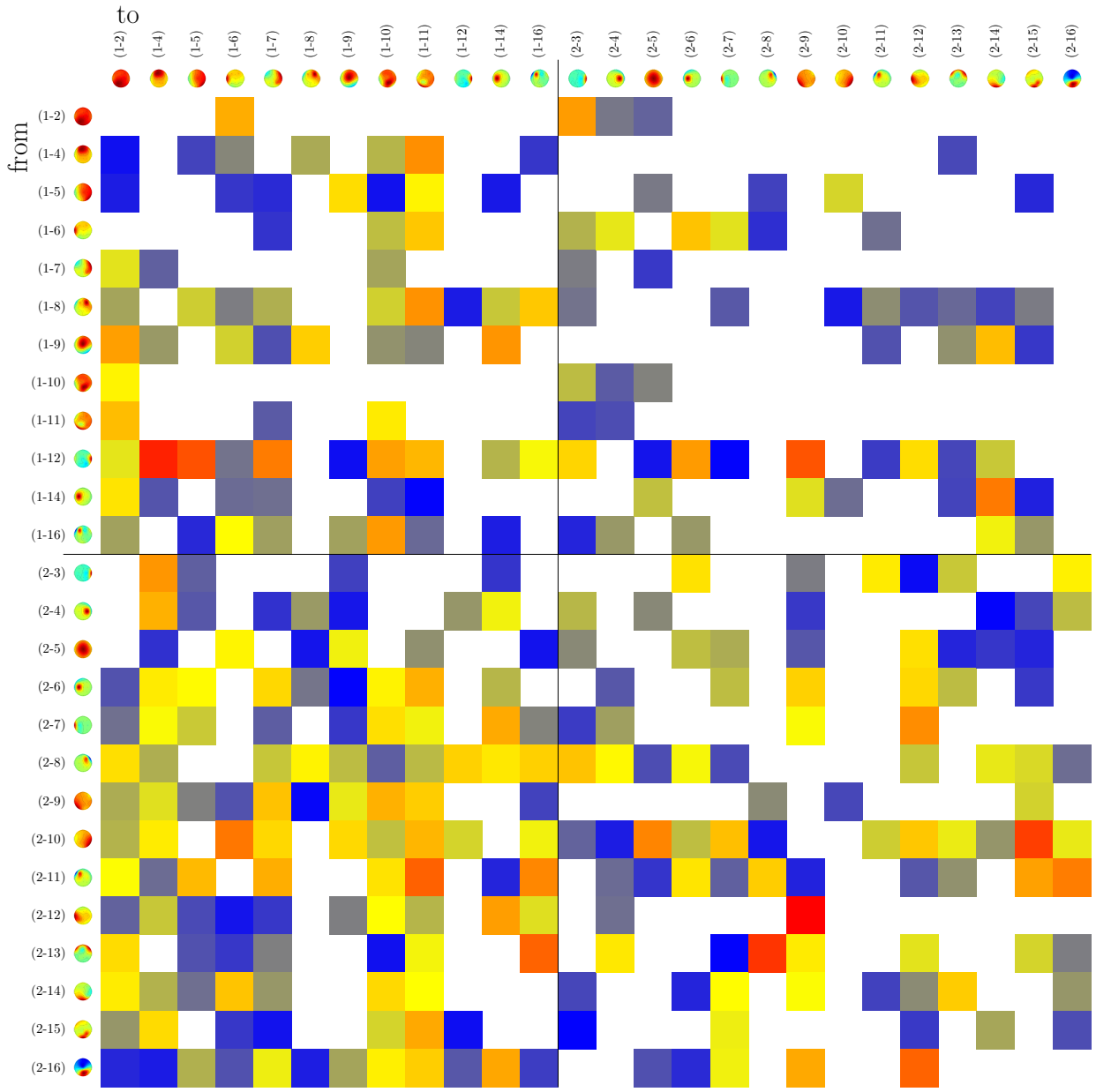


Figure F.28.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment seven γ-band. No significant changes and interesting biases.

F.8. Experiment Eight

One significant change for a hyper-connection could be found in this analysis in the γ-band.

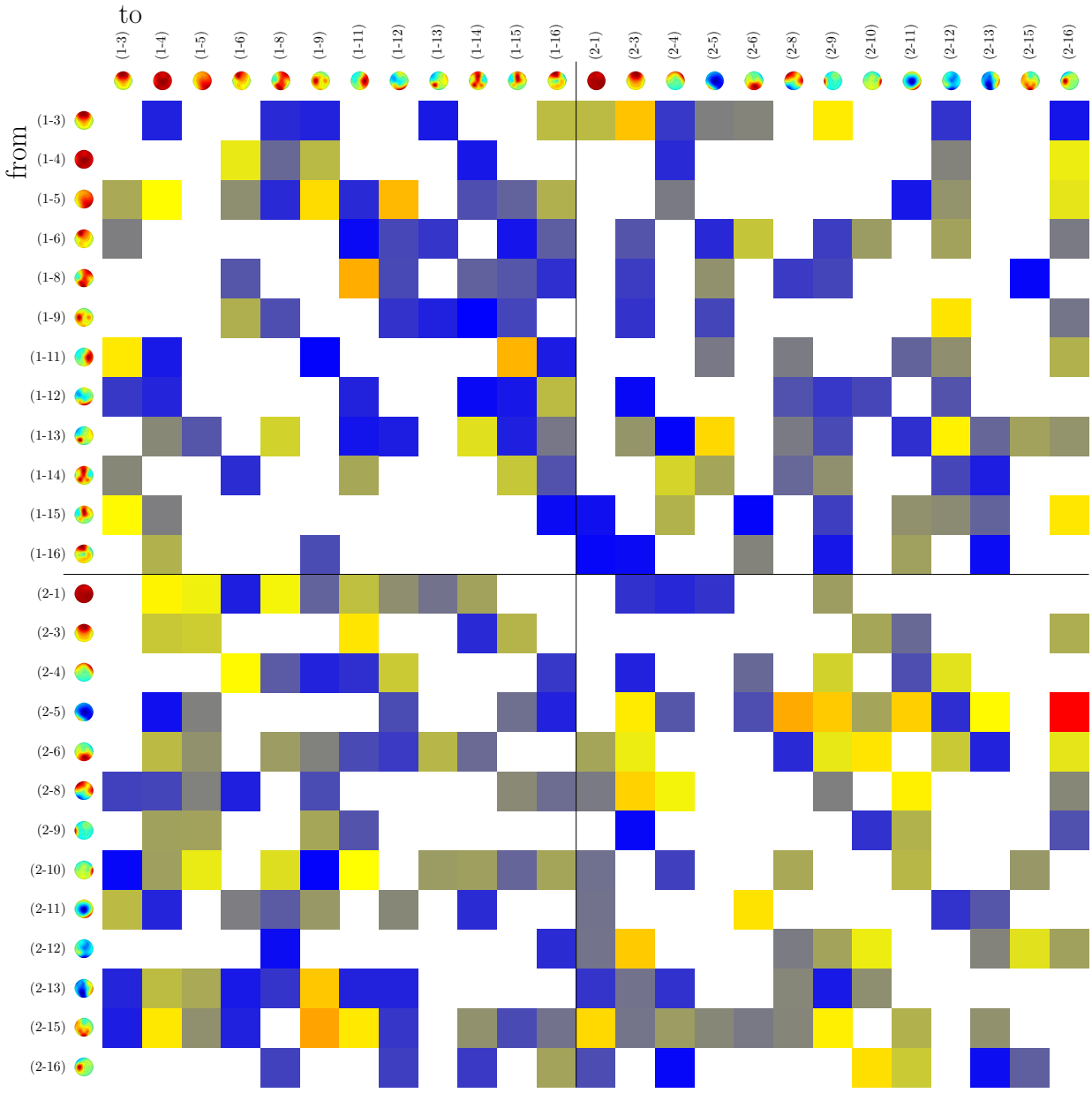


Figure F.29.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment eight θ-band. There is one differential PSI value, which is notable higher than the others, but none reaches significance.

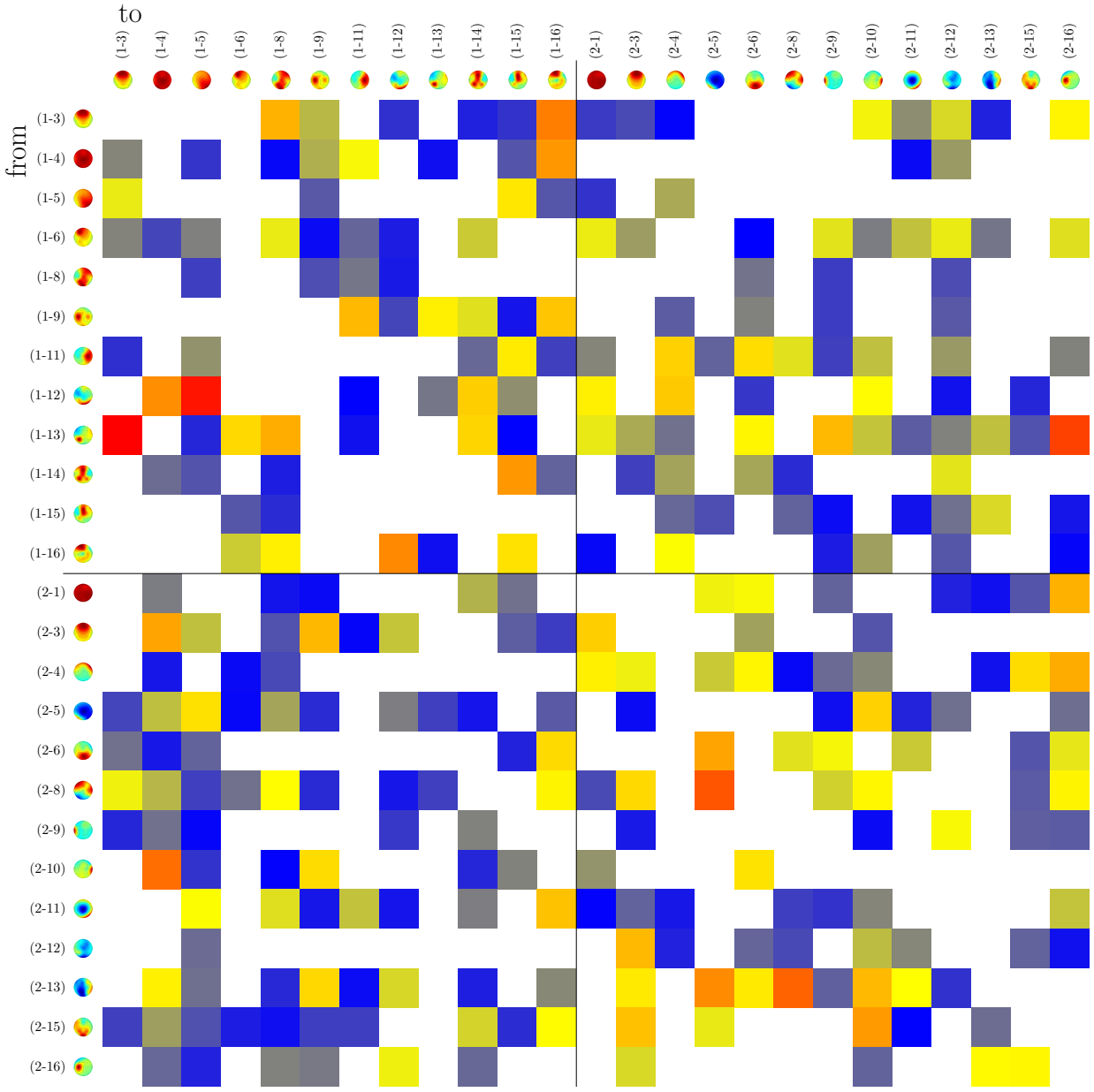


Figure F.30.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment eight α-band. No interesting patterns of significant changes.

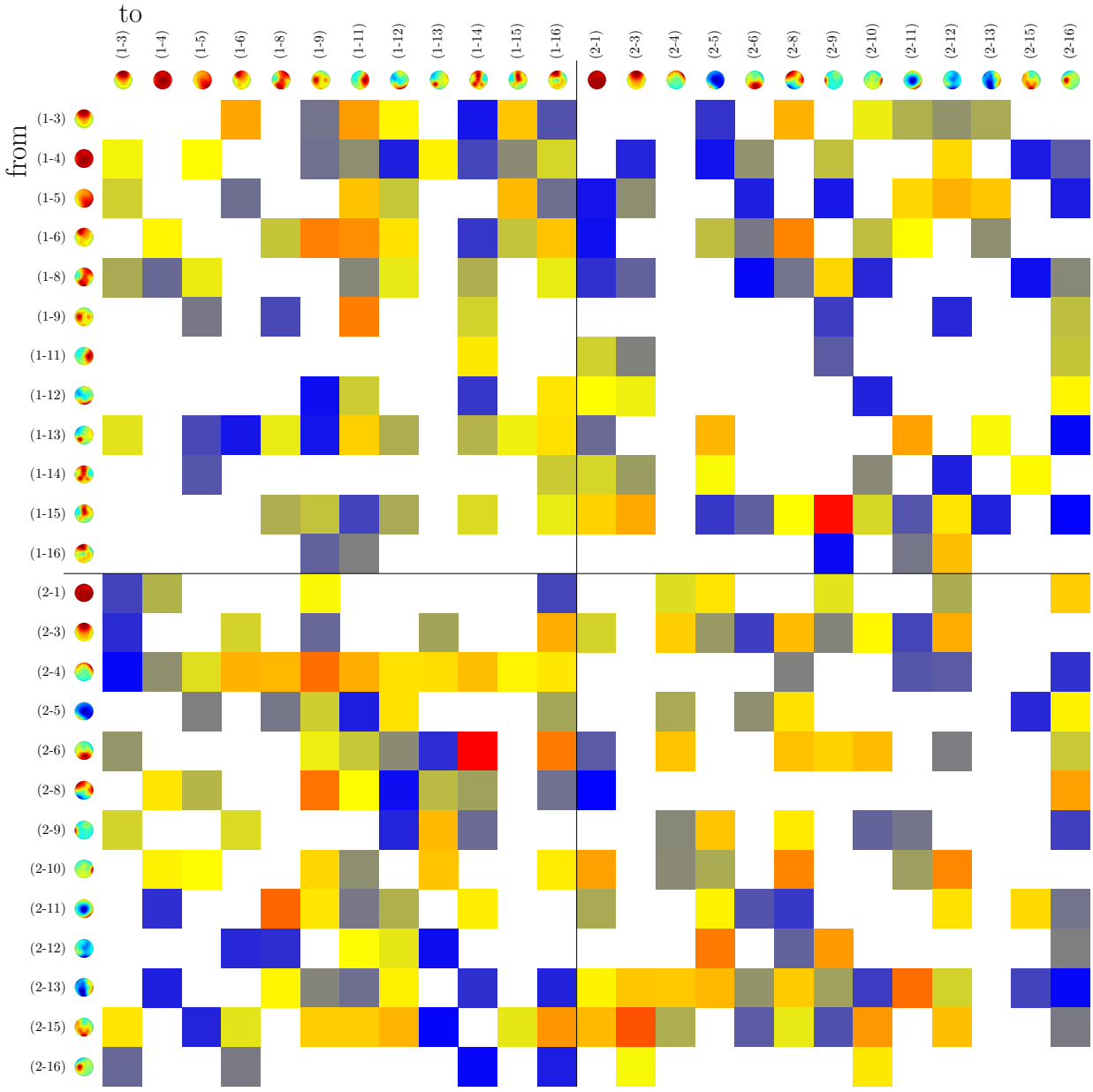


Figure F.31.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment eight β-band. No interesting patterns or significant changes.

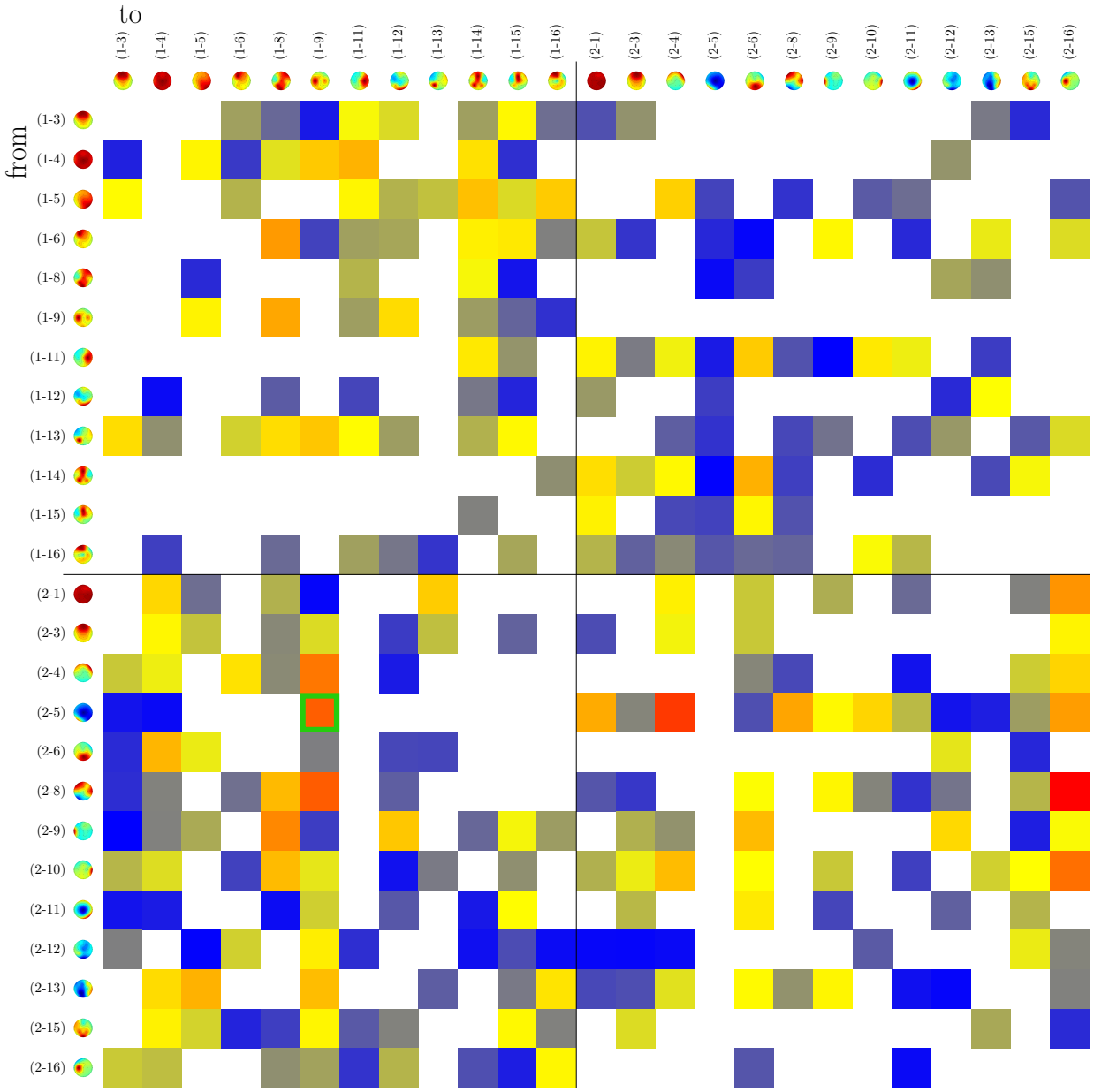


Figure F.32.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment eight γ-band. One change reaches significance, again with a value which is quite a bit smaller than the maximum. And again it affects a hyper-connection. This connects a very narrow occipital-lateral component of participant two with a component (of participant one) having lateral foci on both sides with an emphasis on the left side.

F.9. Experiment Nine

Quite some significant changes for hyper-connections in θ- and α-band.

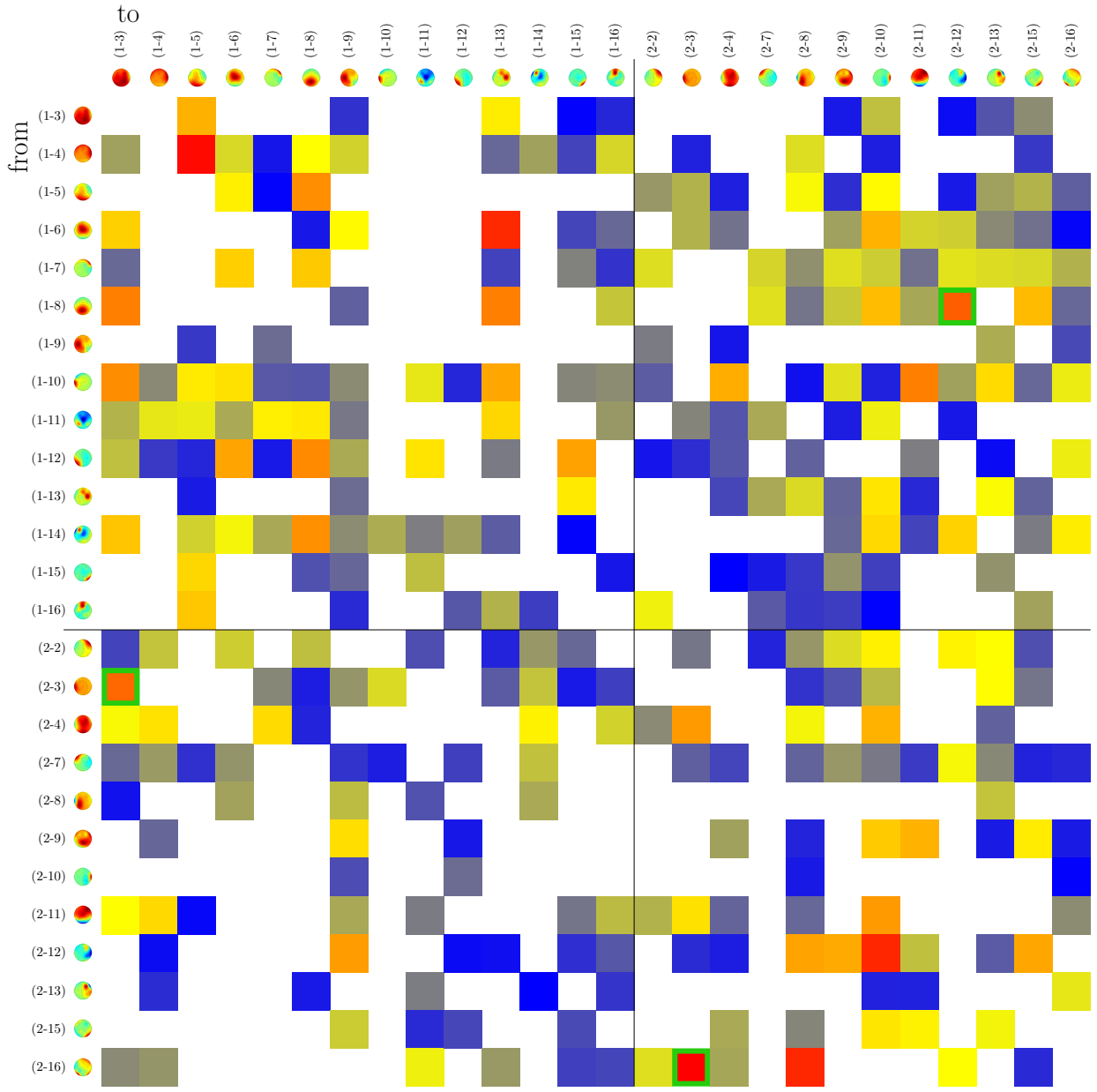


Figure F.33.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment nine θ-band. Three changes reach significance. Two hyper-connections are affected, which are for the first time in opposing directions: One is connecting a very broad lateral component of participant two with a global component of participant one. The other connects a parietal component of participant one with a narrow fronto-lateral component of participant two. Both have more mediocre differential PSI values. Finally there is a significant change for a within-participant connection for participant two connecting a component with three foci (frontal and occipital on both sides) with the same very broad lateral component mentioned before.

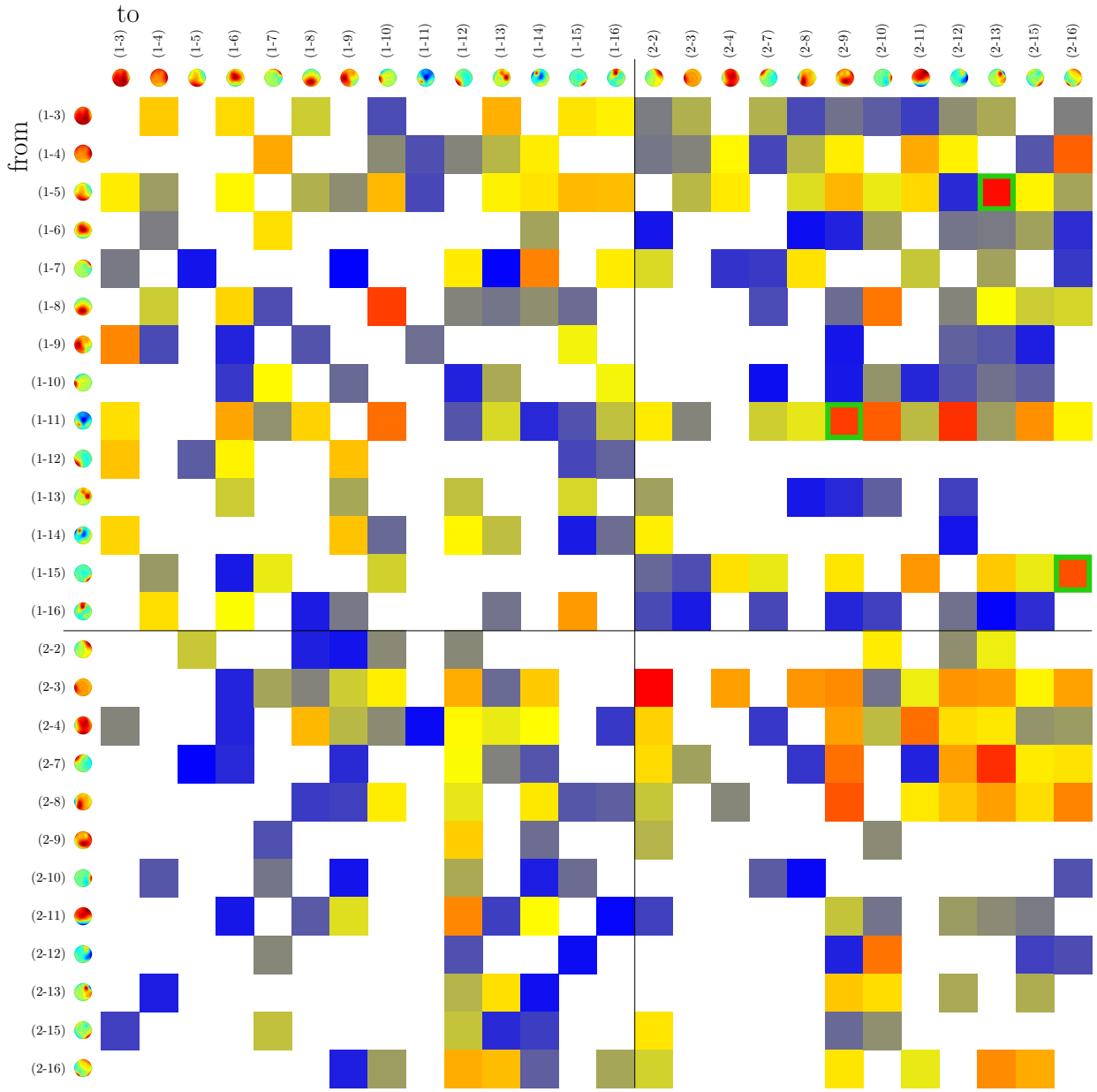


Figure F.34.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment nine α-band. Three hyper-connections show significant changes. All directed from participant one to participant two. One connects a central-lateral component with a very broad parietal component. The second connects an occipital component with central contributions with a strongly fronto-lateral component and the last connects an occipital-lateral component with a component with frontal and occipital contributions.

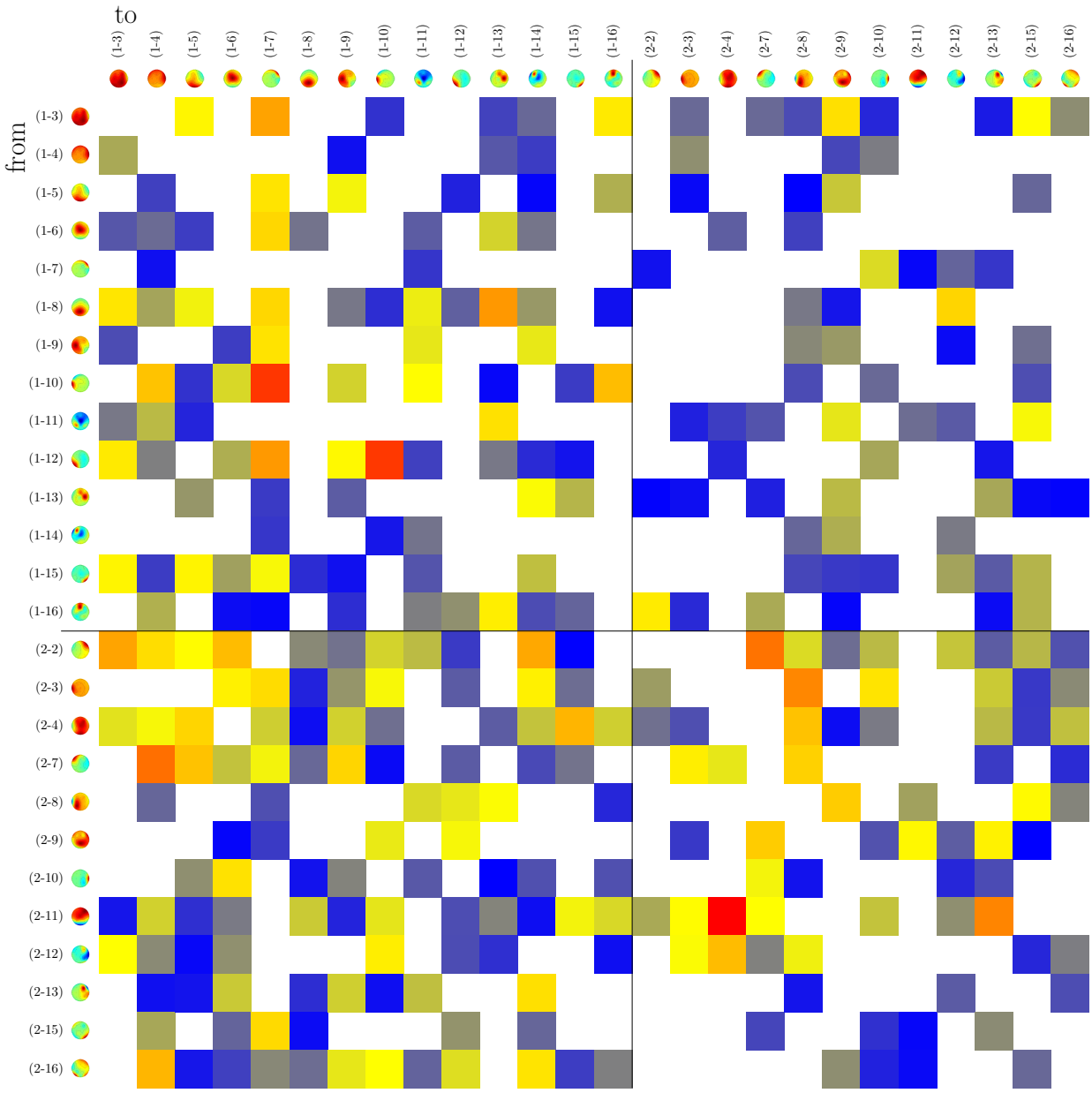


Figure F.35.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment nine β-band. No significant changes or salient patterns in the distribution of connection can be identified.

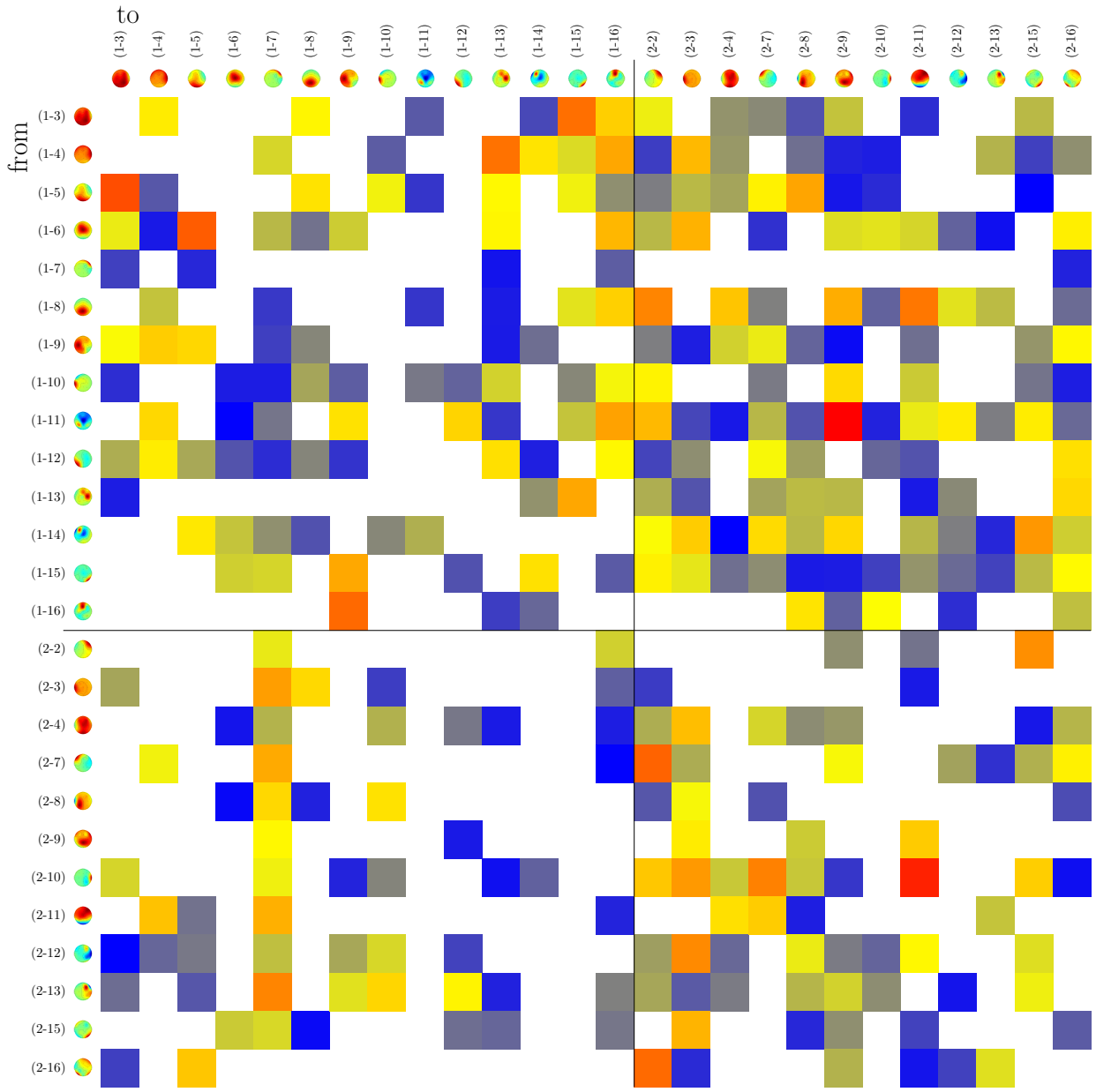


Figure F.36.: Delta in PSI connectivity estimation between cooperative trials and non cooperative trials in experiment nine γ-band. A slight bias among hyper-connections favouring positive values for connections from participant one to participant two can be identified. No changes reach significance.

G. Differential PSI: By Initiator of Robotic Action

These tables show the results of the differential PSI analysis contrasting data from trials which have been initiated by different participants. For details on the interpretation of these plots please refer to 8.6.2 on page 102.

G.1. Experiment One

Significant changes could be found for seven hyper-connection, all in β-band.

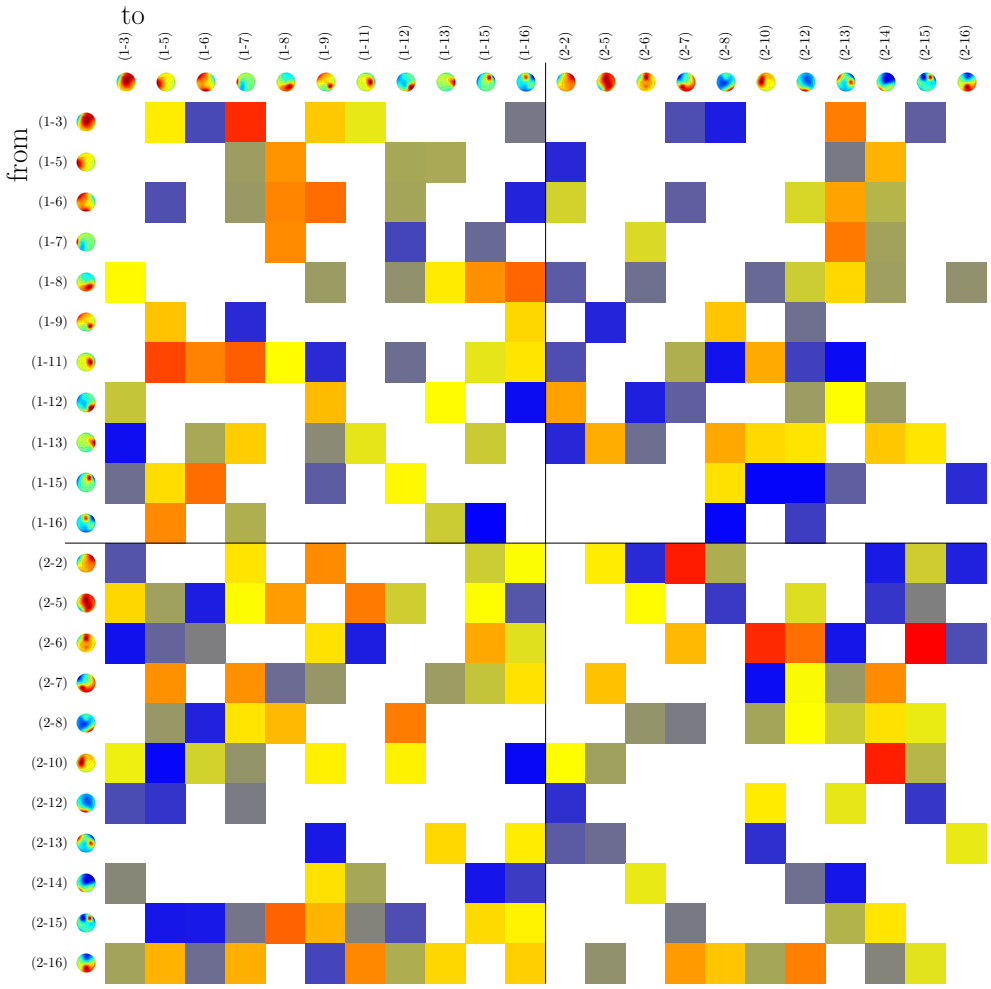


Figure G.1.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment one for θ-band. None of the connections show significant changes and there are no biases nor unusual patterns.

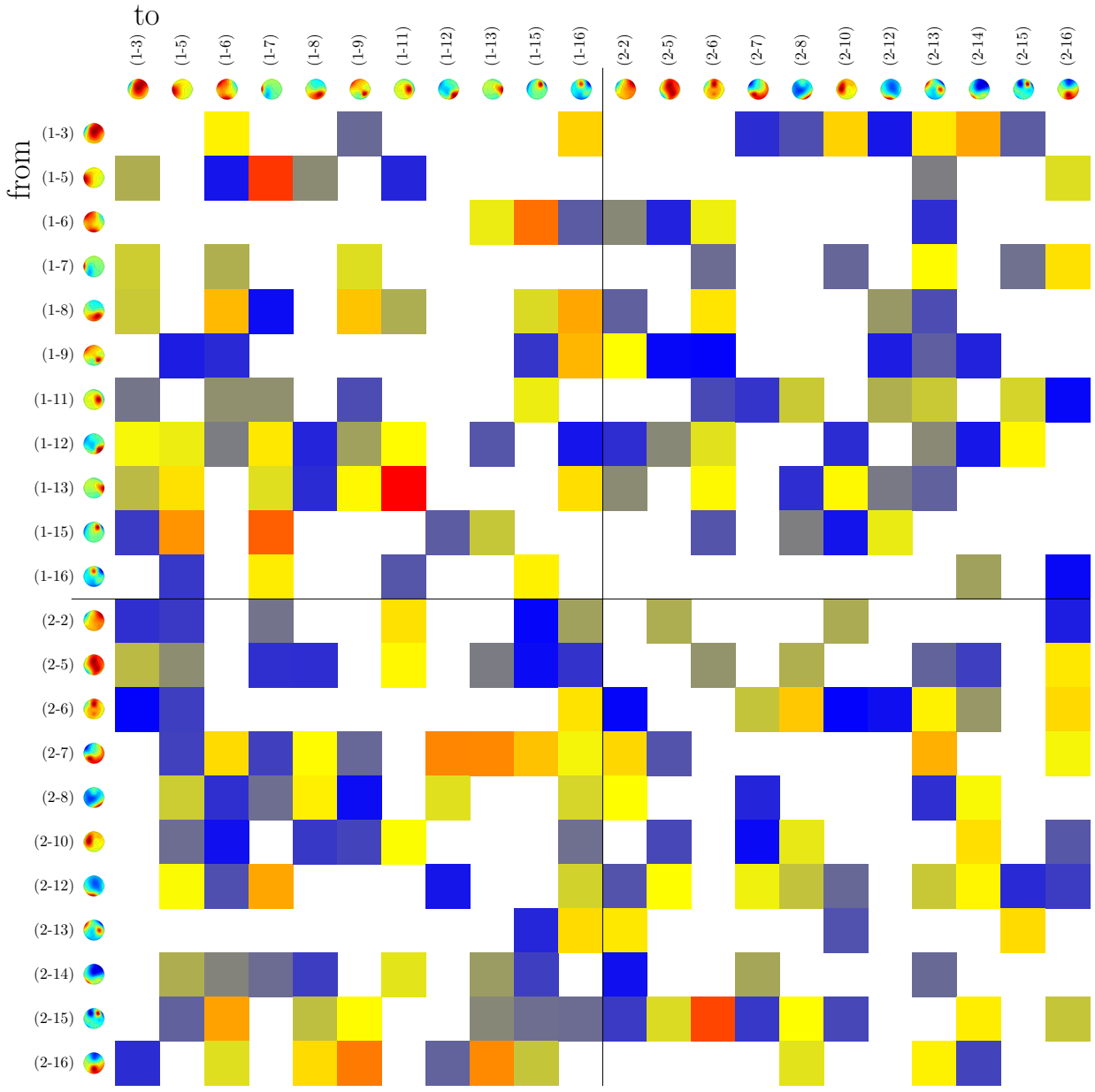


Figure G.2.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment one for α-band. Again no significant changes, biases nor usual patterns in the distribution.

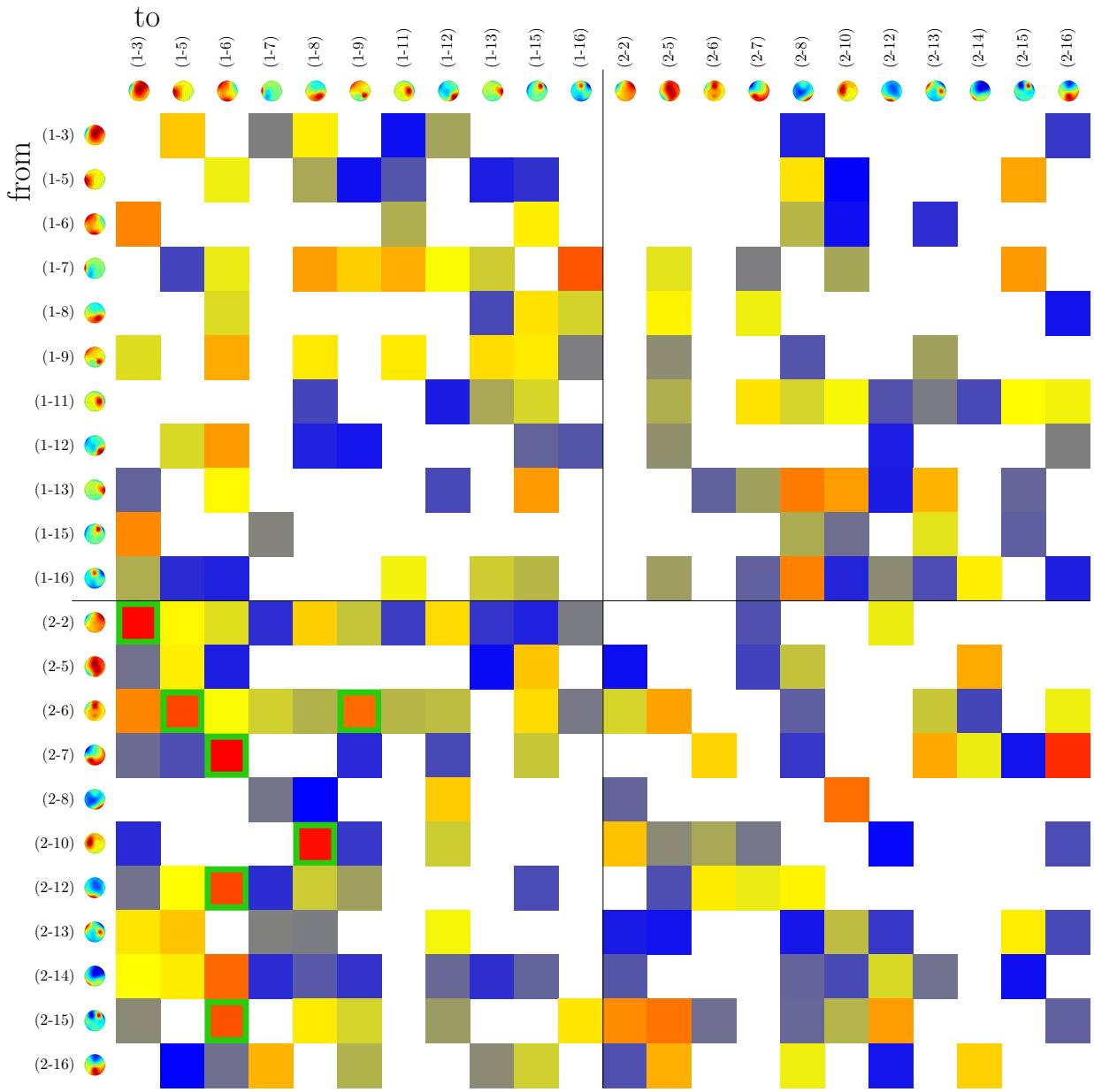


Figure G.3.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment one for β-band. Seven connections, all hyper-connections from participant one to participant two, show significant changes. One component is the recipient for three and thereby for almost half of the affected connection. This is a broad lateral component.

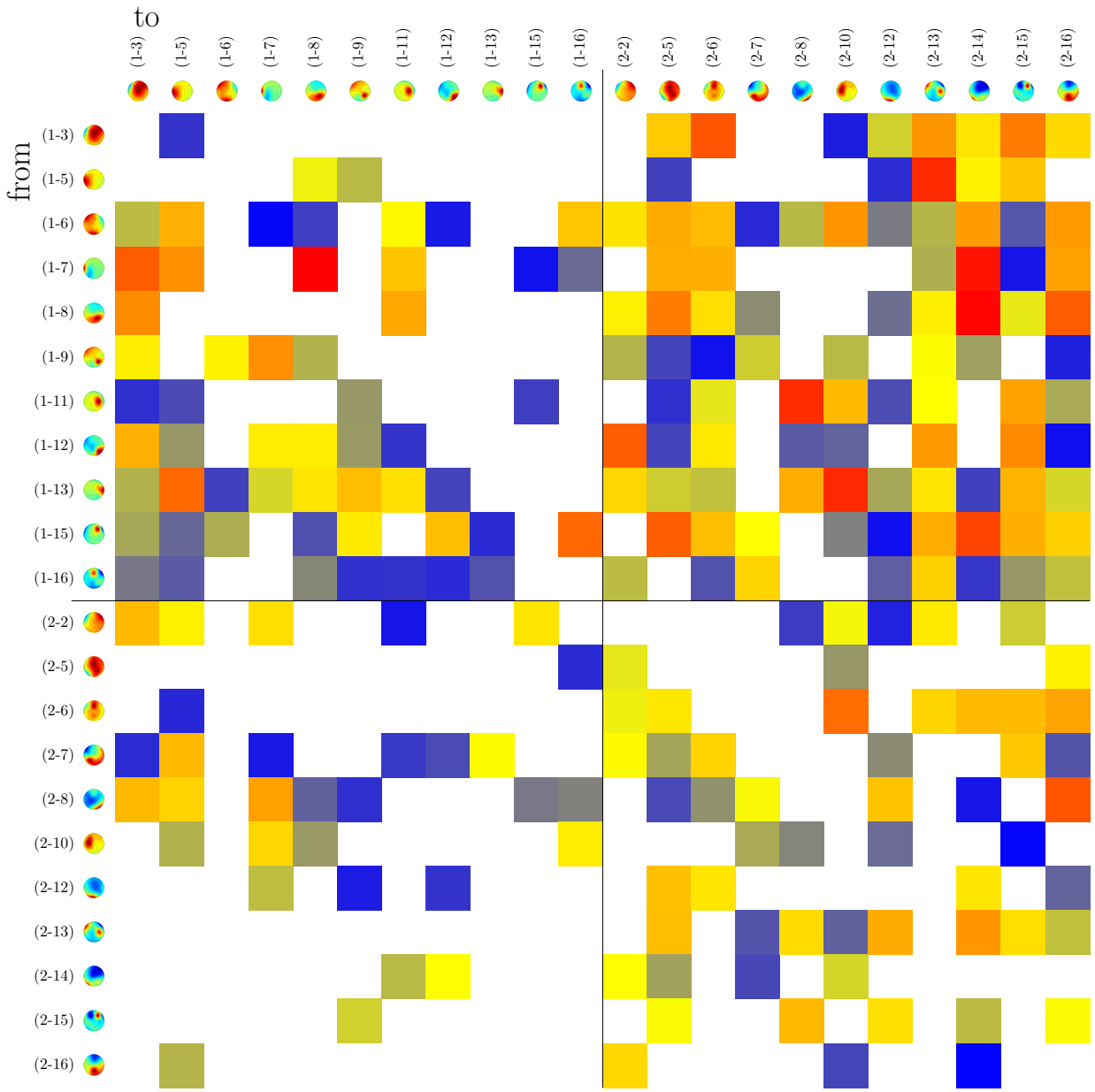


Figure G.4.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively in, experiment one for γ-band. No significant changes, but a pretty strong bias among hyper-connections favouring positive values for connections from participant one to participant two.

G.2. Experiment Two

Nine hyper-connections in γ and θ-band and four within-participants showed significant changes.

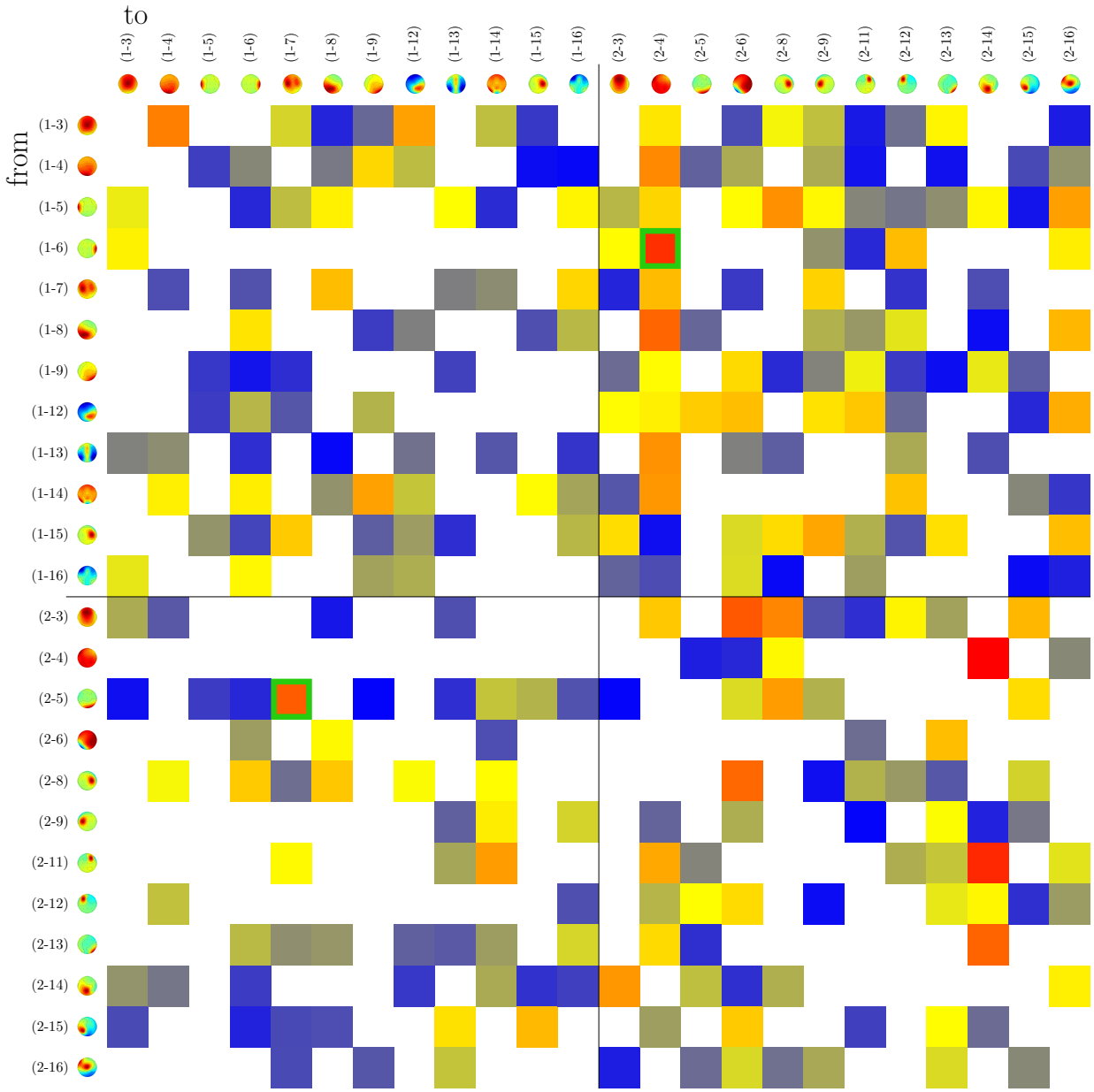


Figure G.5.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment two for θ-band. Two connections showed significant changes. One from a lateral component of participant one to a global component of participant two and the other from an occipital component of participant two to a broad lateral component (of participant two) with foci on both sides.

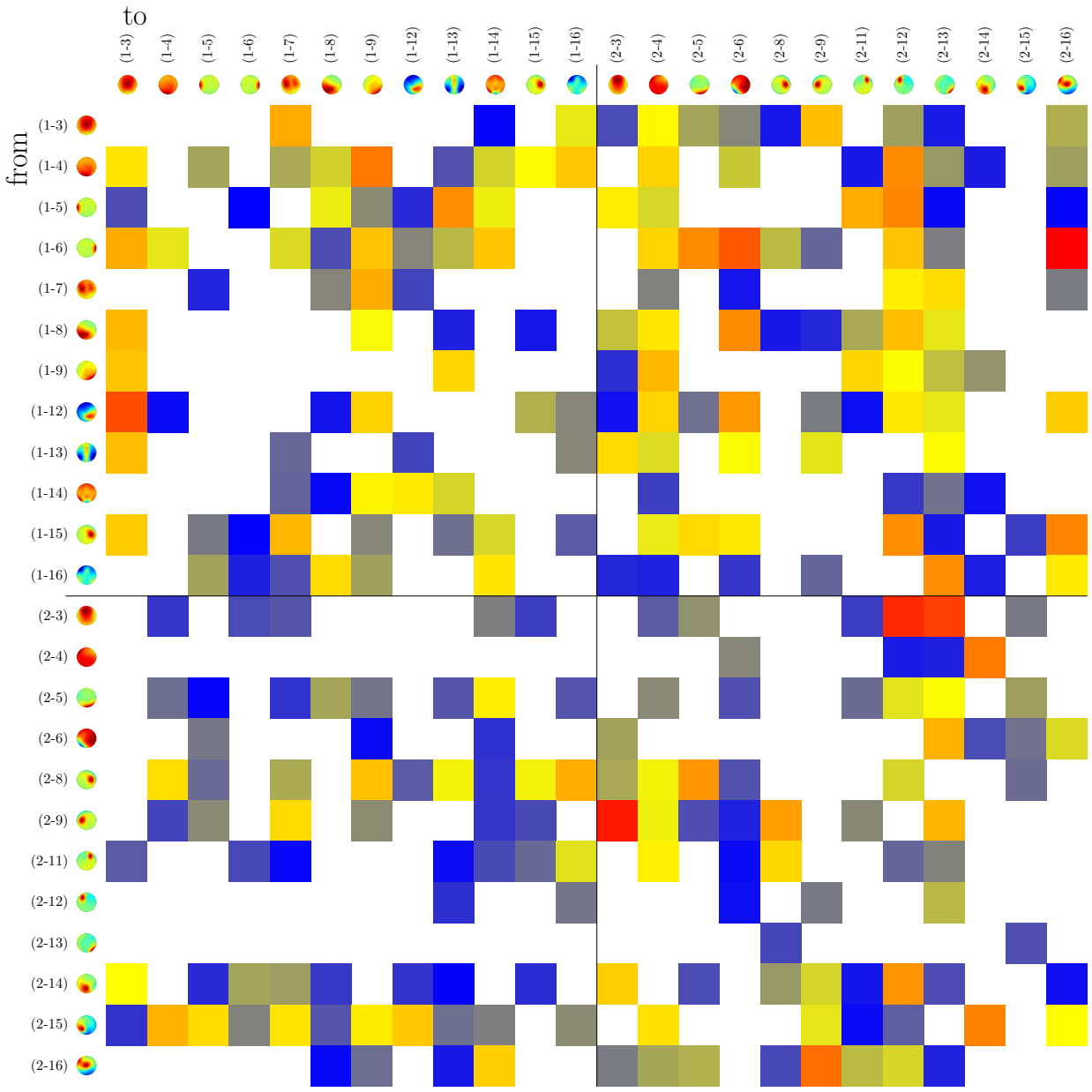


Figure G.6.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment two for α-band. No significant changes where found. The distribution of the connectivity values is unremarkable.

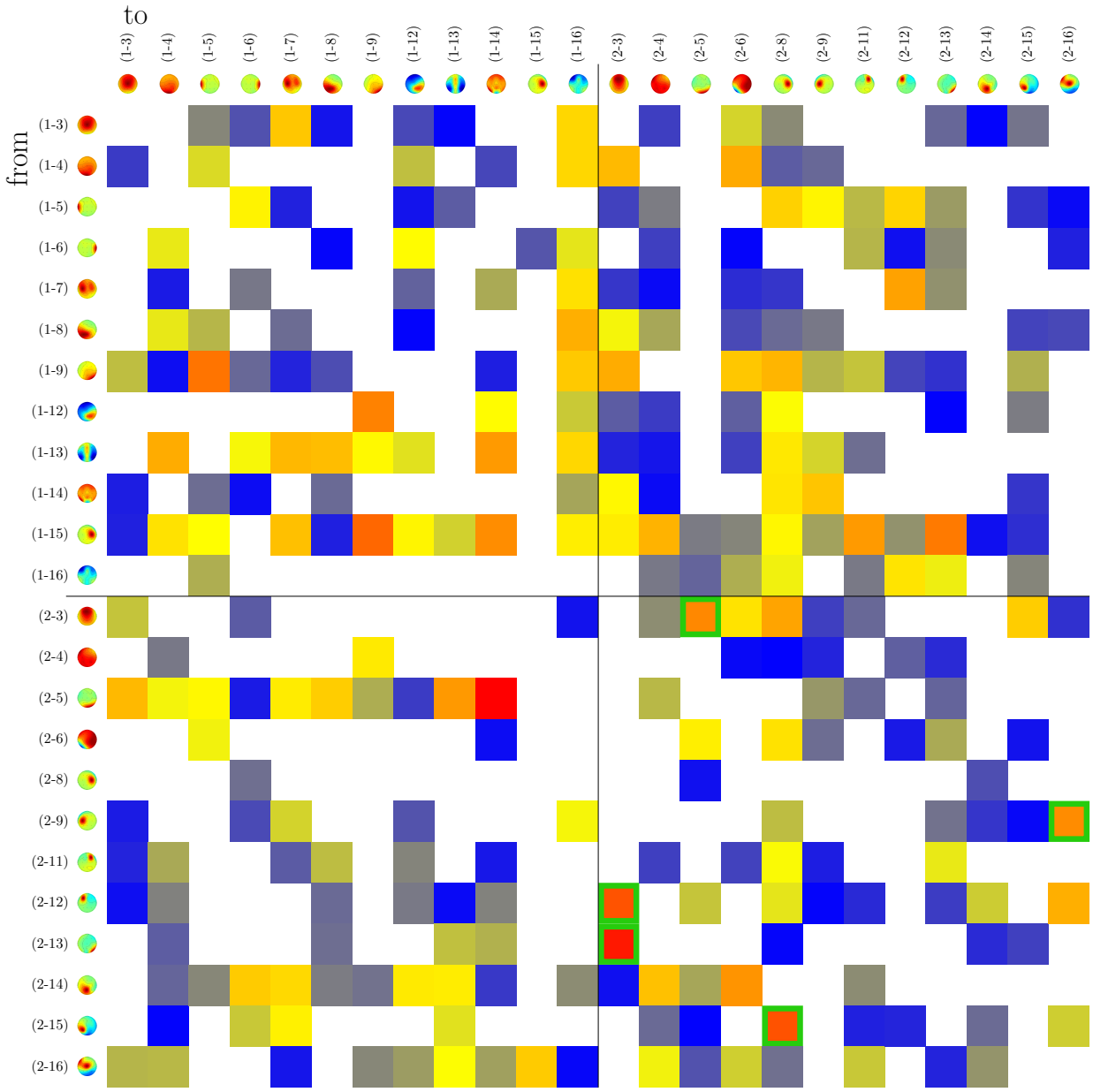


Figure G.7.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment two β-band. For five connections significant changes were identified. All of them were within-participant connections of participant two. A network originating from two components one fronto-lateral, the other occipital-lateral over a broad frontal to a strongly occipital component can be described. A fourth significant change was found between a parietal-lateral to a lateral component. The last significant change is directed from a lateral component towards a central-lateral component.

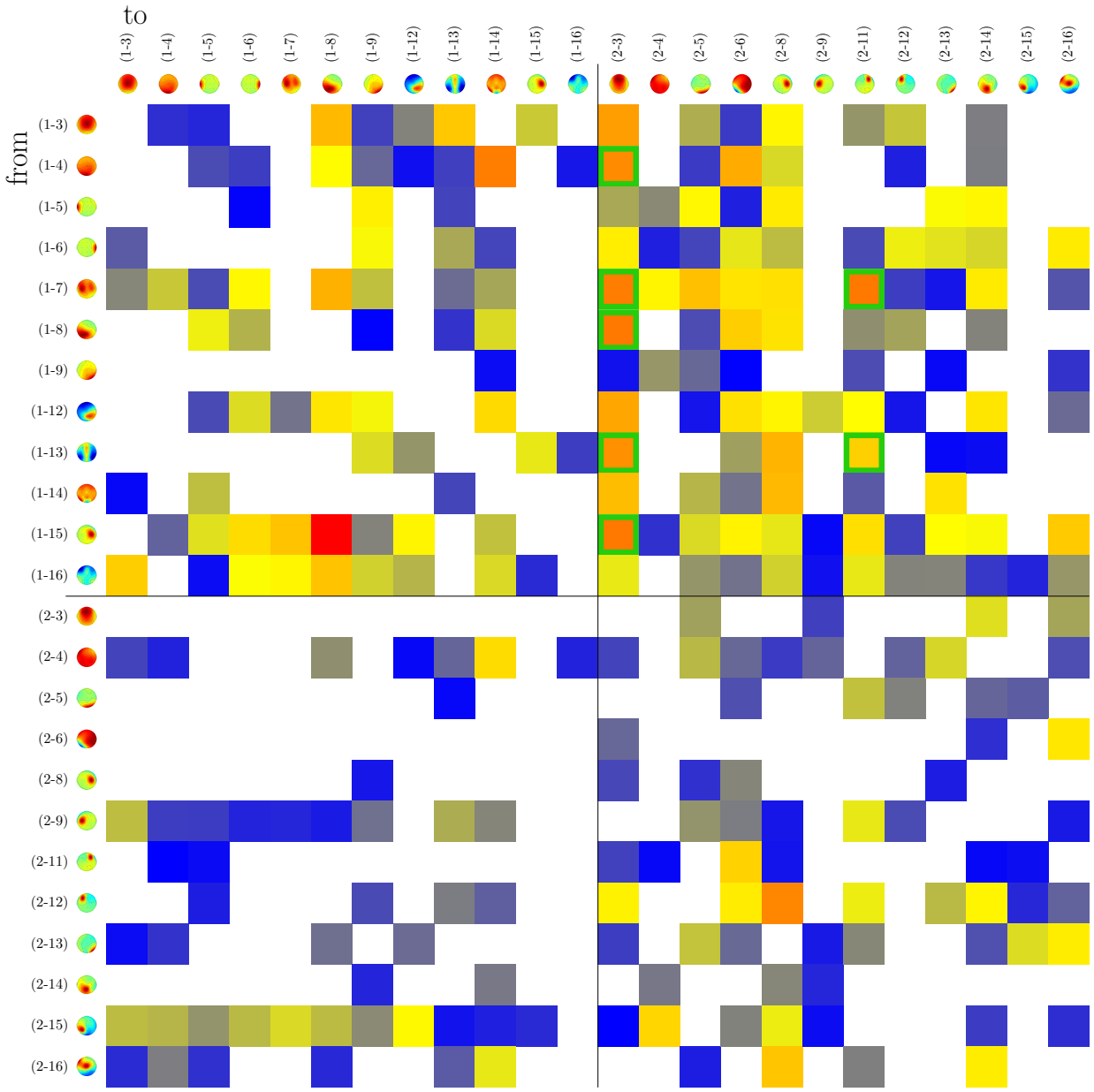


Figure G.8.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment two for γ-band. Seven connection showed significant changes. All of them are hyper-connections from participant one to participant two. Five of them influence a broad frontal component of participant two. Two of these components also have positive values to a fronto-lateral component of participant two.

G.3. Experiment Three

Many hyper-connections and some with-in participant connections were subject to significant changes, all in β- and γ-band.

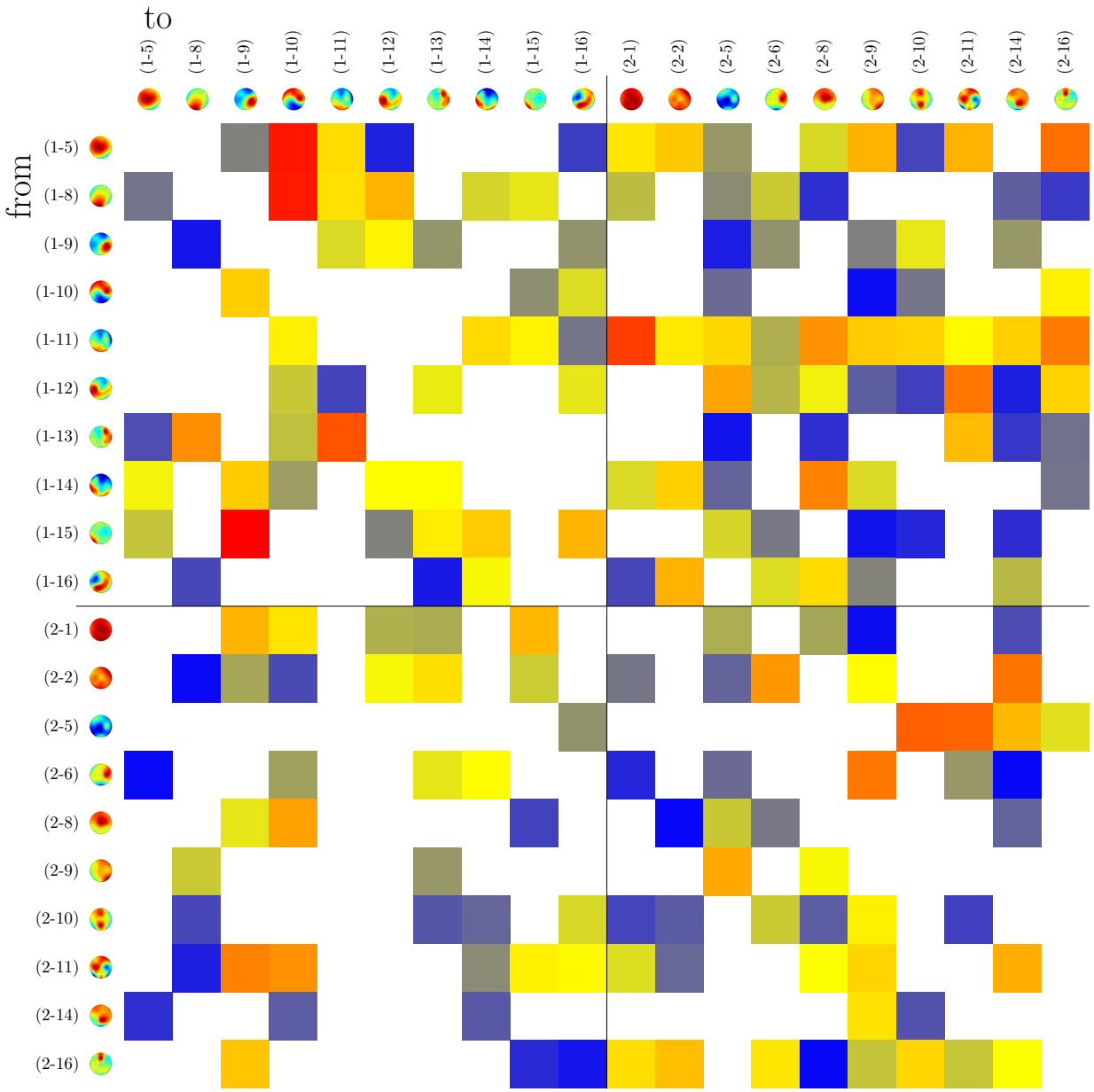


Figure G.9.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment three for θ-band. No significant changes have been found and no distinct patterns in the connectivity values' distribution have been found.

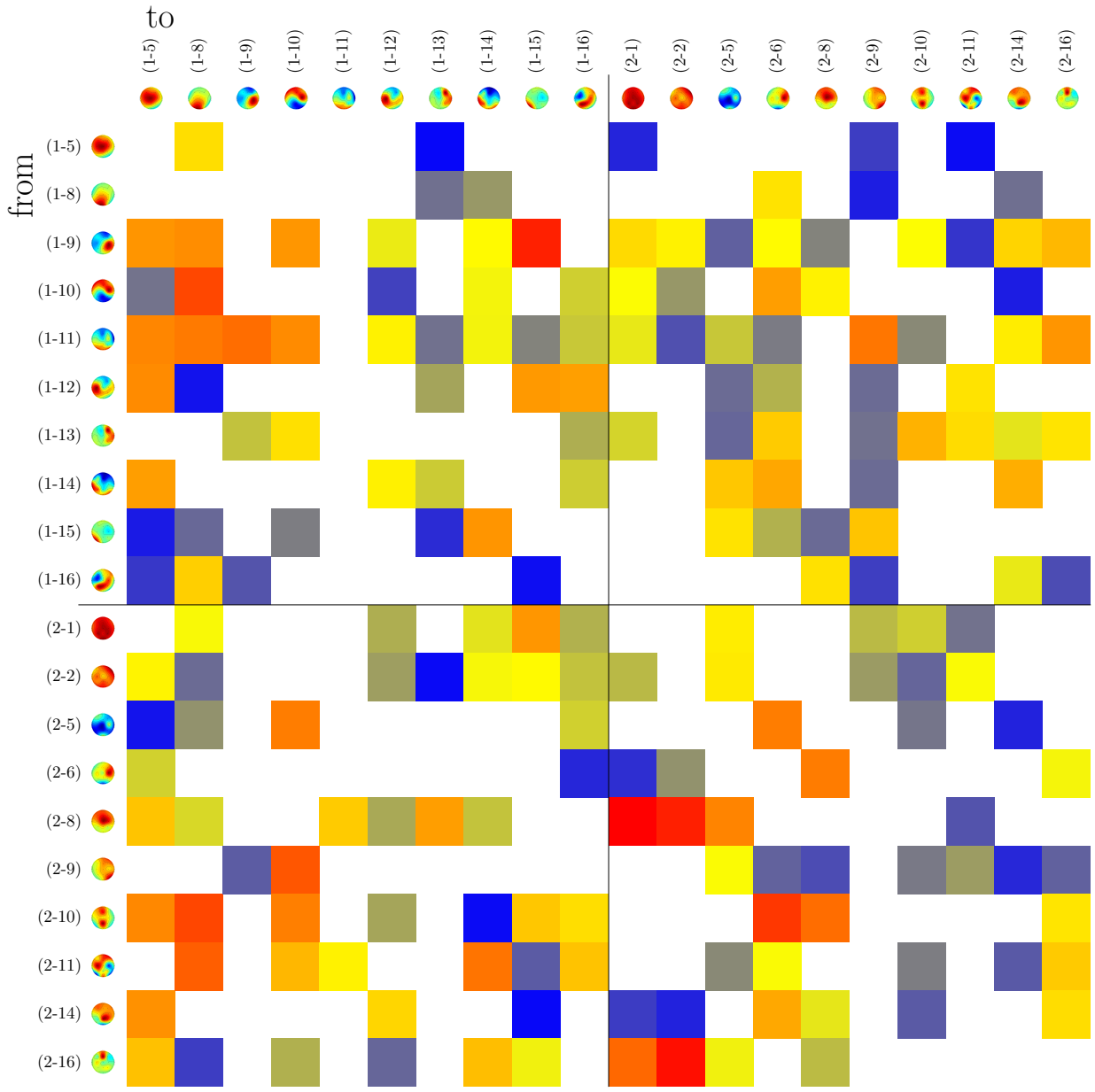


Figure G.10.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment three for α-band. Again no significant changes and remarkable distributions have been found.

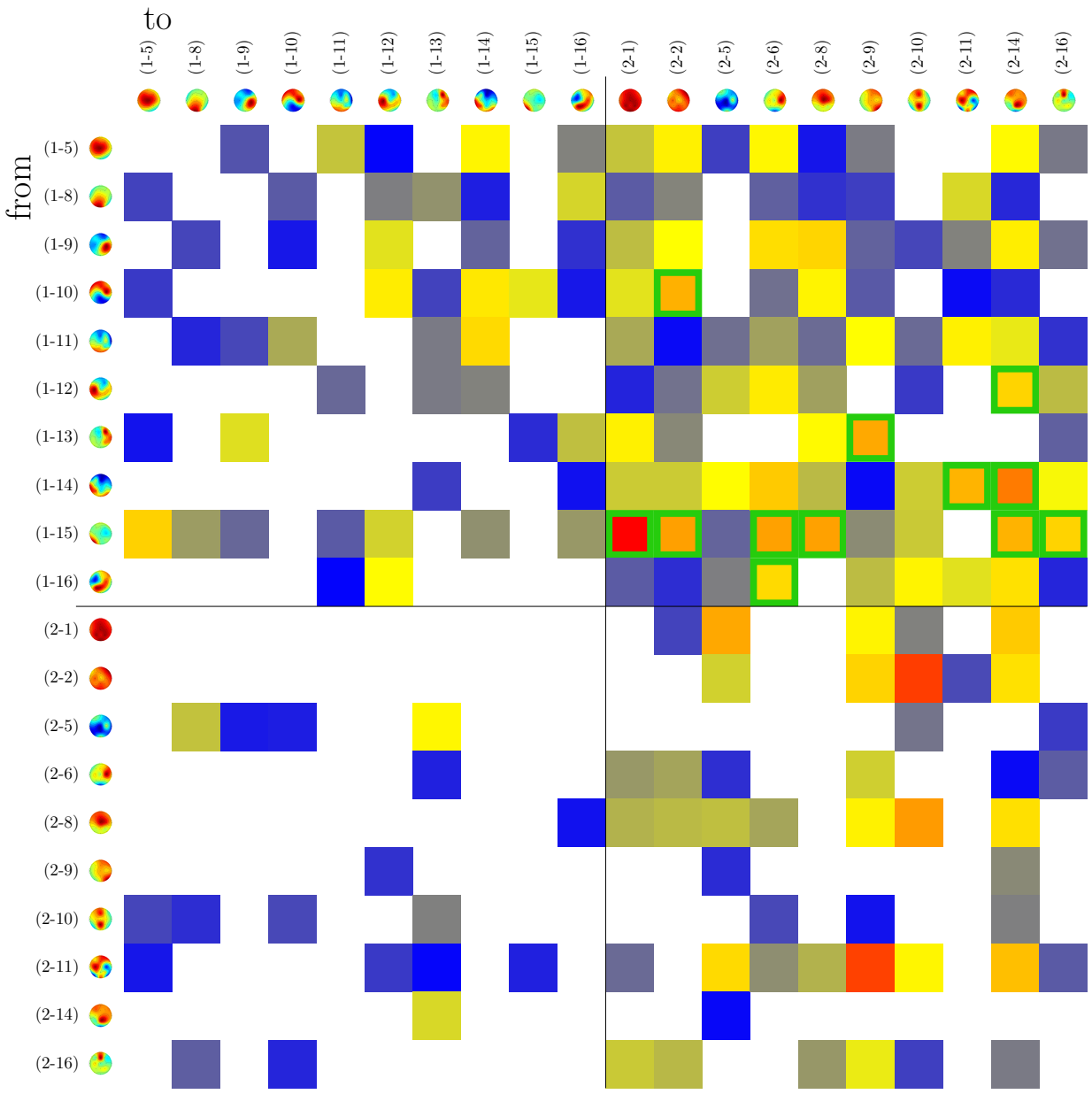


Figure G.11.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment three for β-band. Twelve connections showed significant changes. All of them were hyper-connections from participant one to participant two. There is a general strong bias among hyper-connections favouring positive values for connections from participant one to participant two. One occipital-lateral component of participant one is sender for six of connections which underwent significant changes.

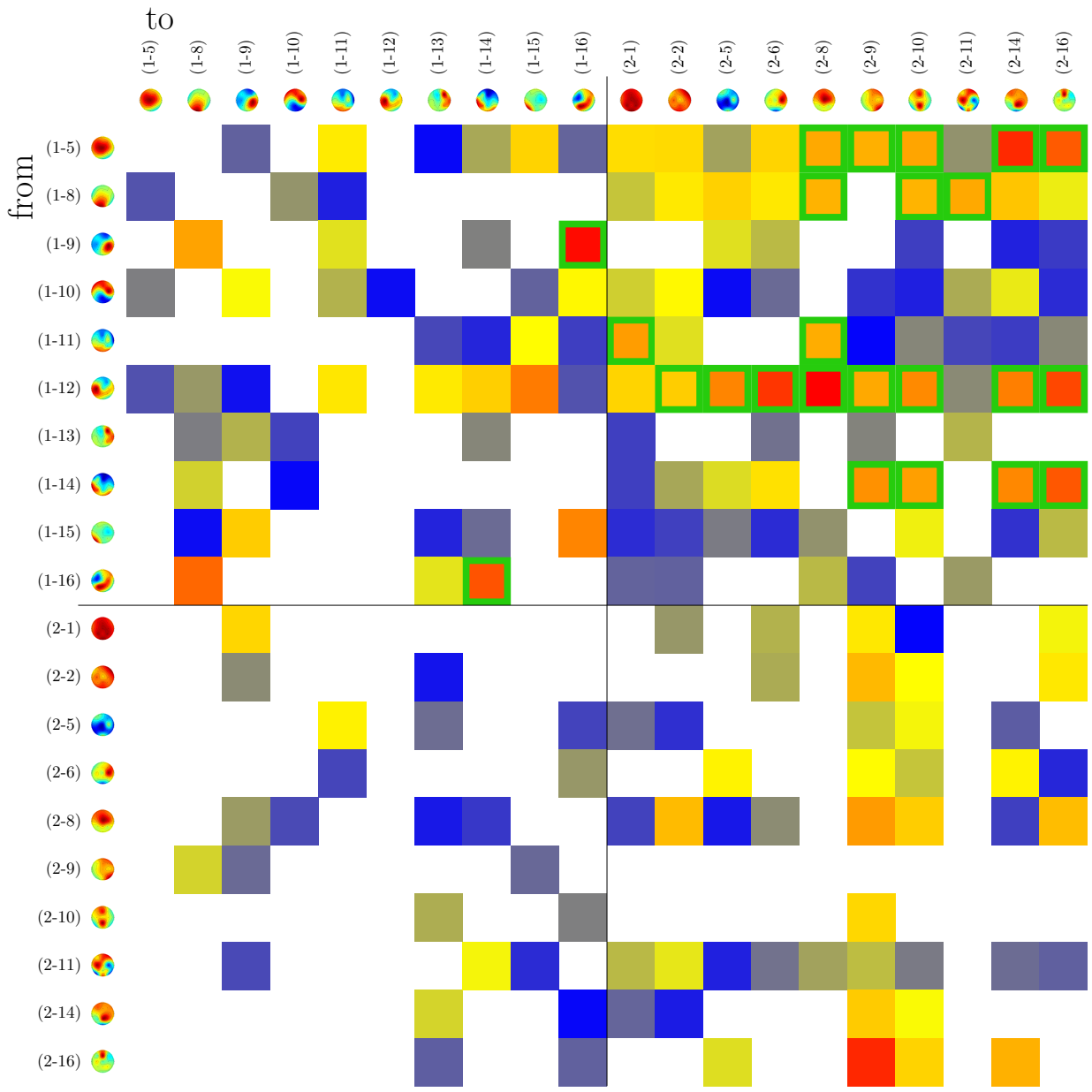


Figure G.12.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment three for γ-band. Twenty-two hyper-connections and two within-participant connections showed significant changes. All of them originate from components of participant one. One lateral component which stretches along the coronal plane is the origin of eight hyper-connections affected by significant changes. In total only five components of participant one are the origin for all hyper-connections subject to significant changes.

G.4. Experiment Four

In contrast to the previous experiment only three connections could be found which show significant differences between data groups and these are in θ- and α-band.

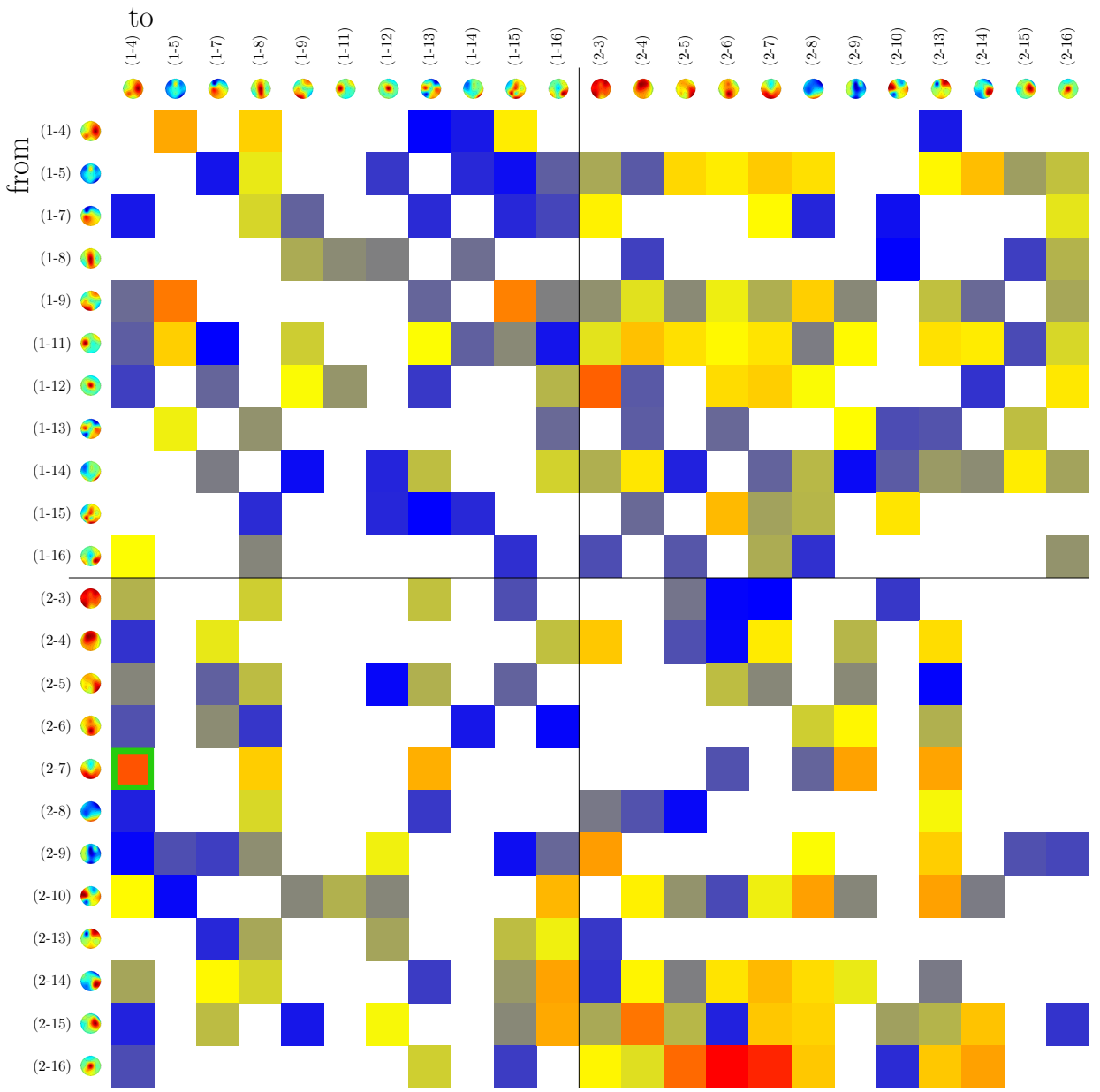


Figure G.13.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment four for θ-band. One hyper-connection from an occipital component of participant two to lateral component extending along the coronal axis of participant one showed significant changes. Apart from that no remarkable patterns in the distribution of differential PSI values were found.

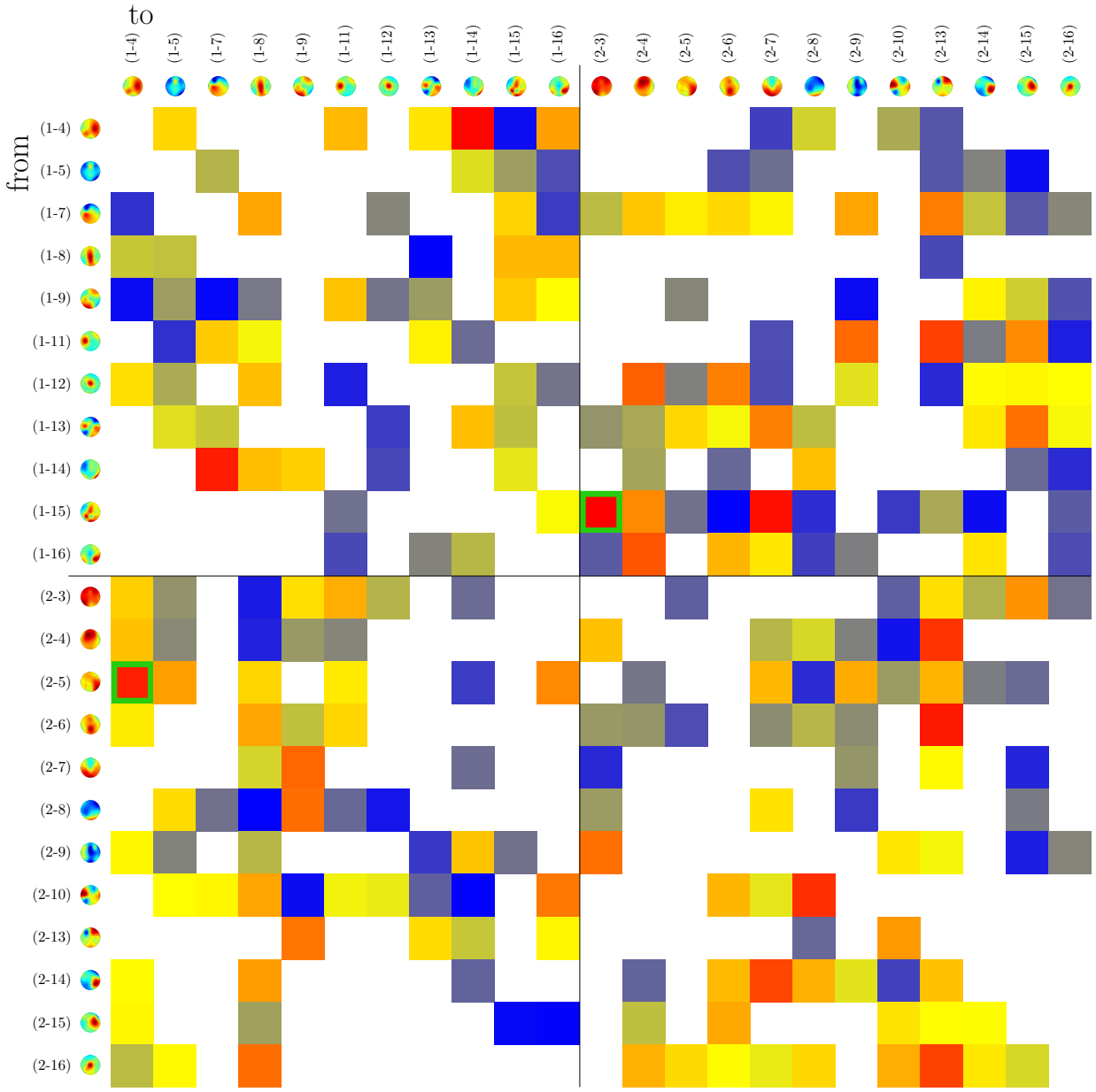


Figure G.14.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment four for α-band. One connection from a parietal-lateral component of participant two to an ispi-lateral component of participant changed significantly. A second connection subject to significant change connects an occipital component extending along the sagittal axis of participant one with a rather global component of participant two.

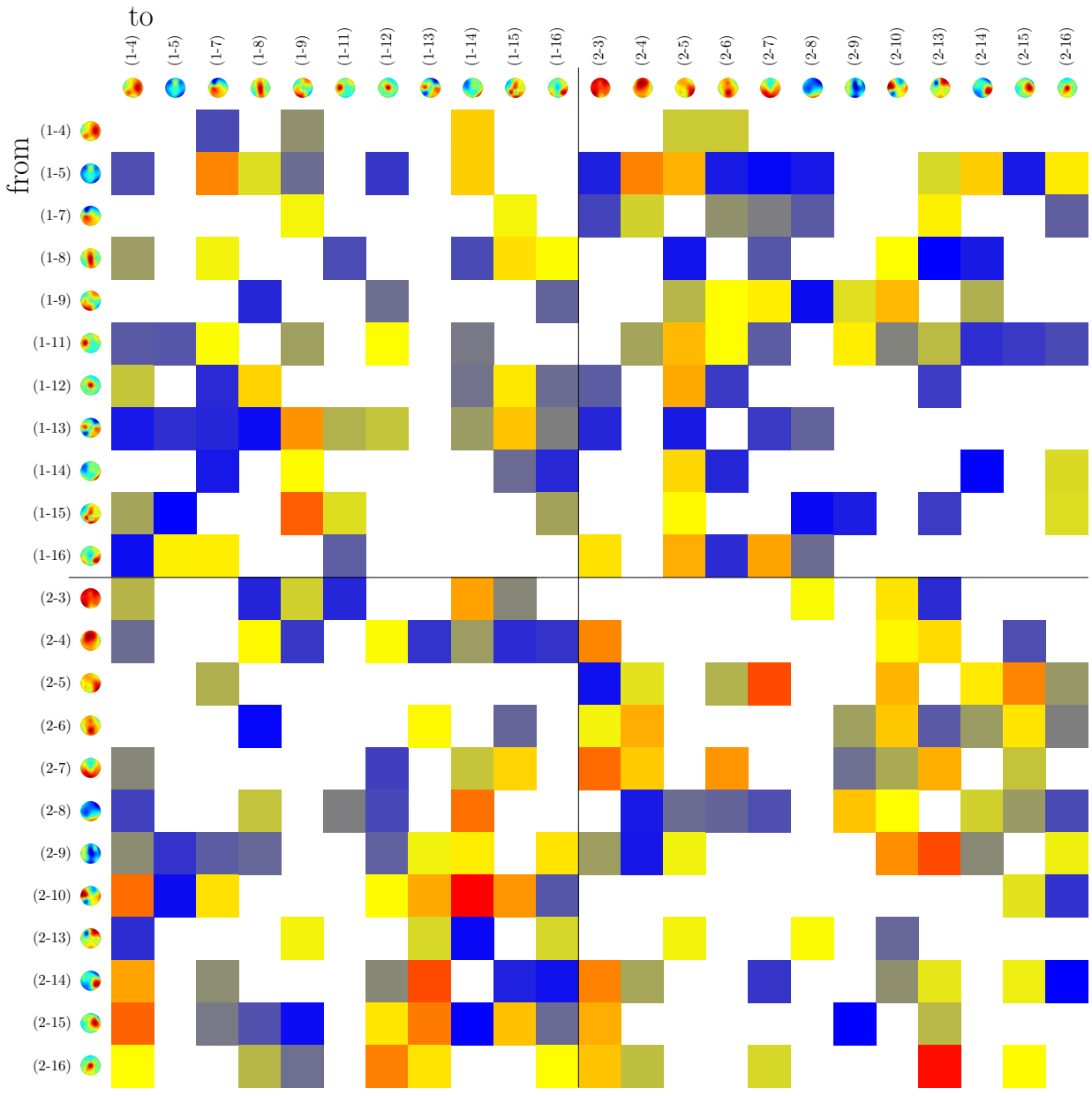


Figure G.15.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment four for β-band. No significant changes and no abnormalities in the PSI-values' distribution was found.

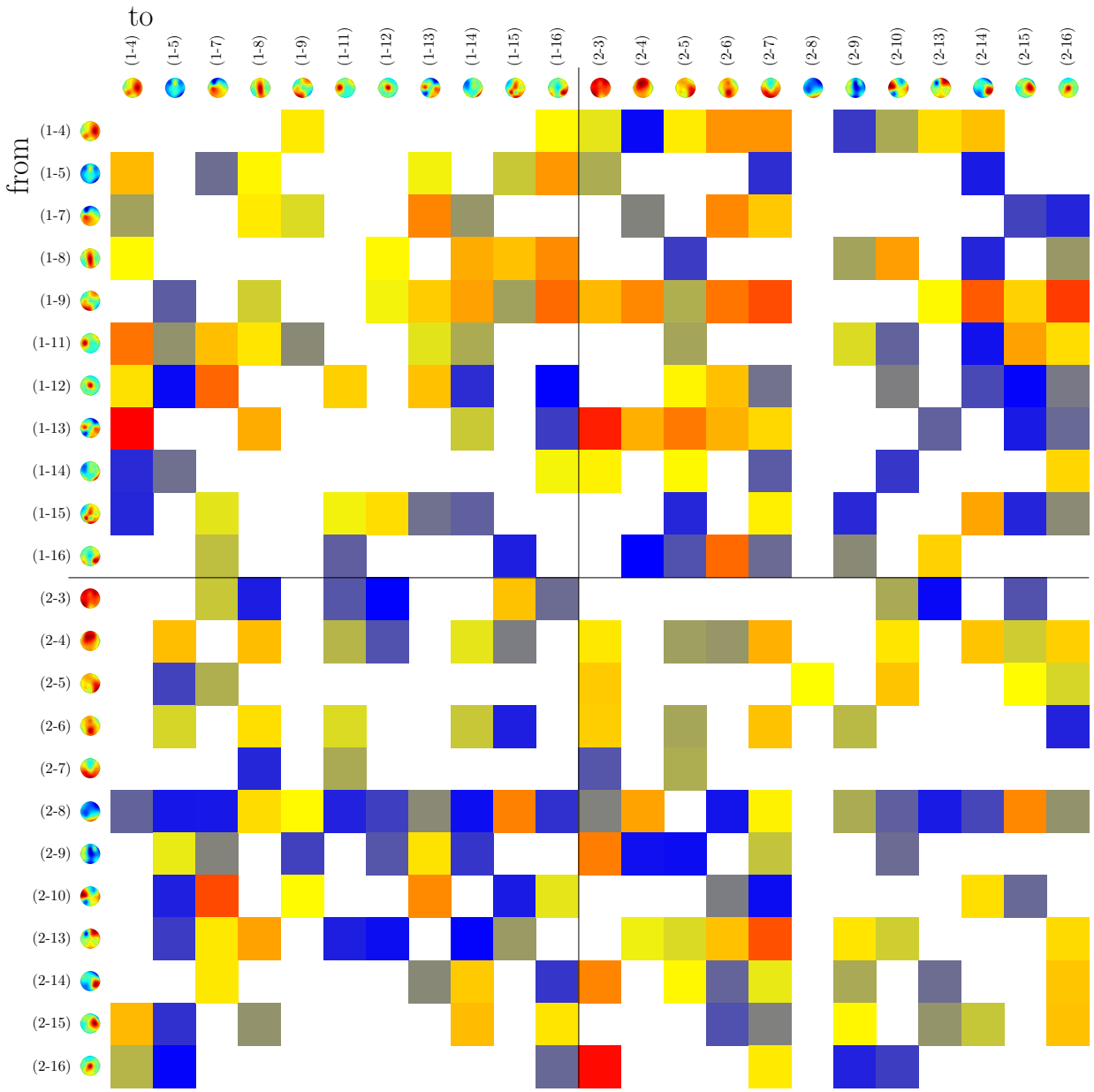


Figure G.16.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment four for γ-band. Again no significant changes and no remarkable biases or patterns.

G.5. Experiment Five

This analysis of the fifth experiment's data set revealed little conclusive evidence.

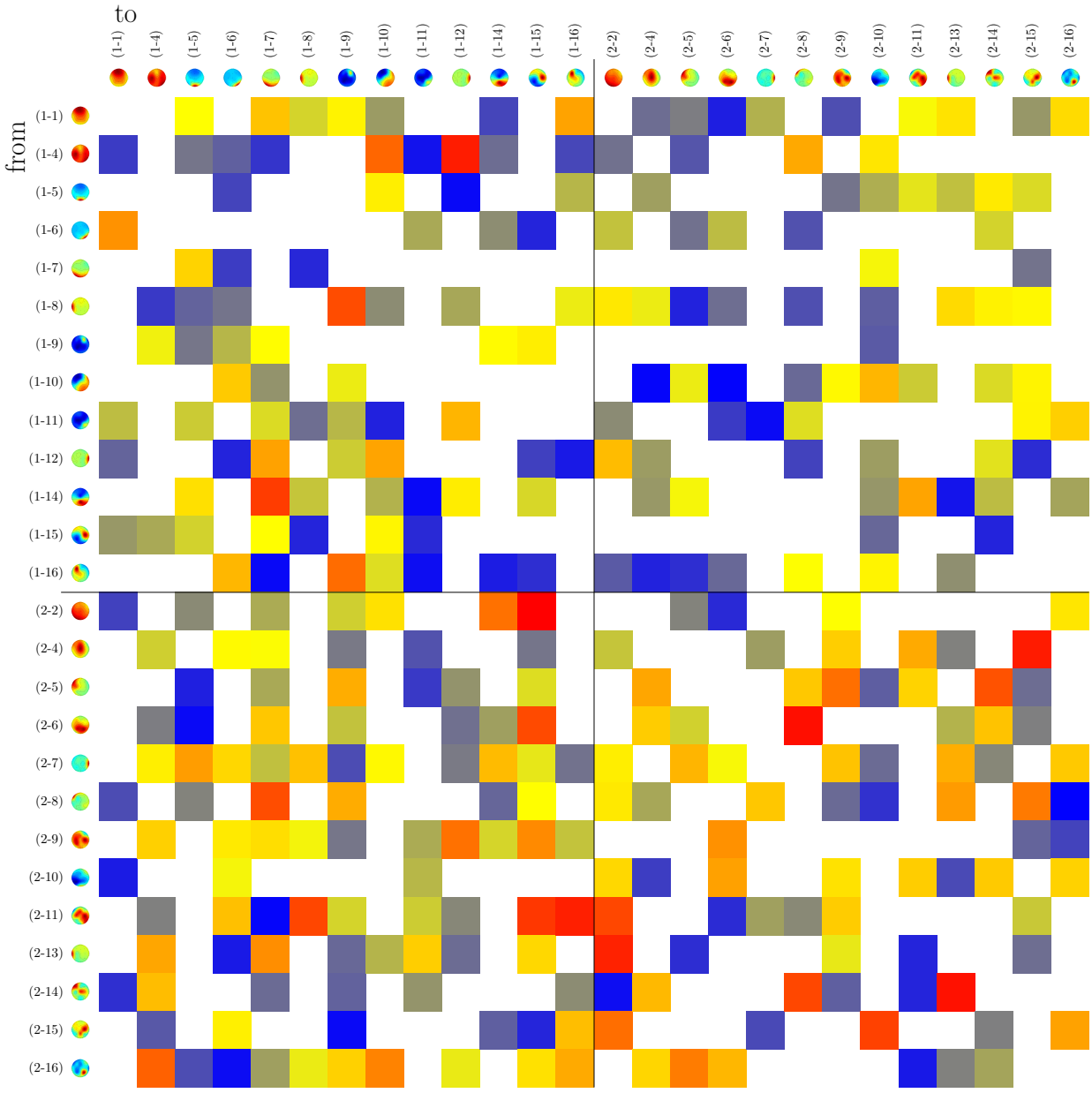


Figure G.17.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment five for θ-band. No significant changes and an unremarkable distribution.

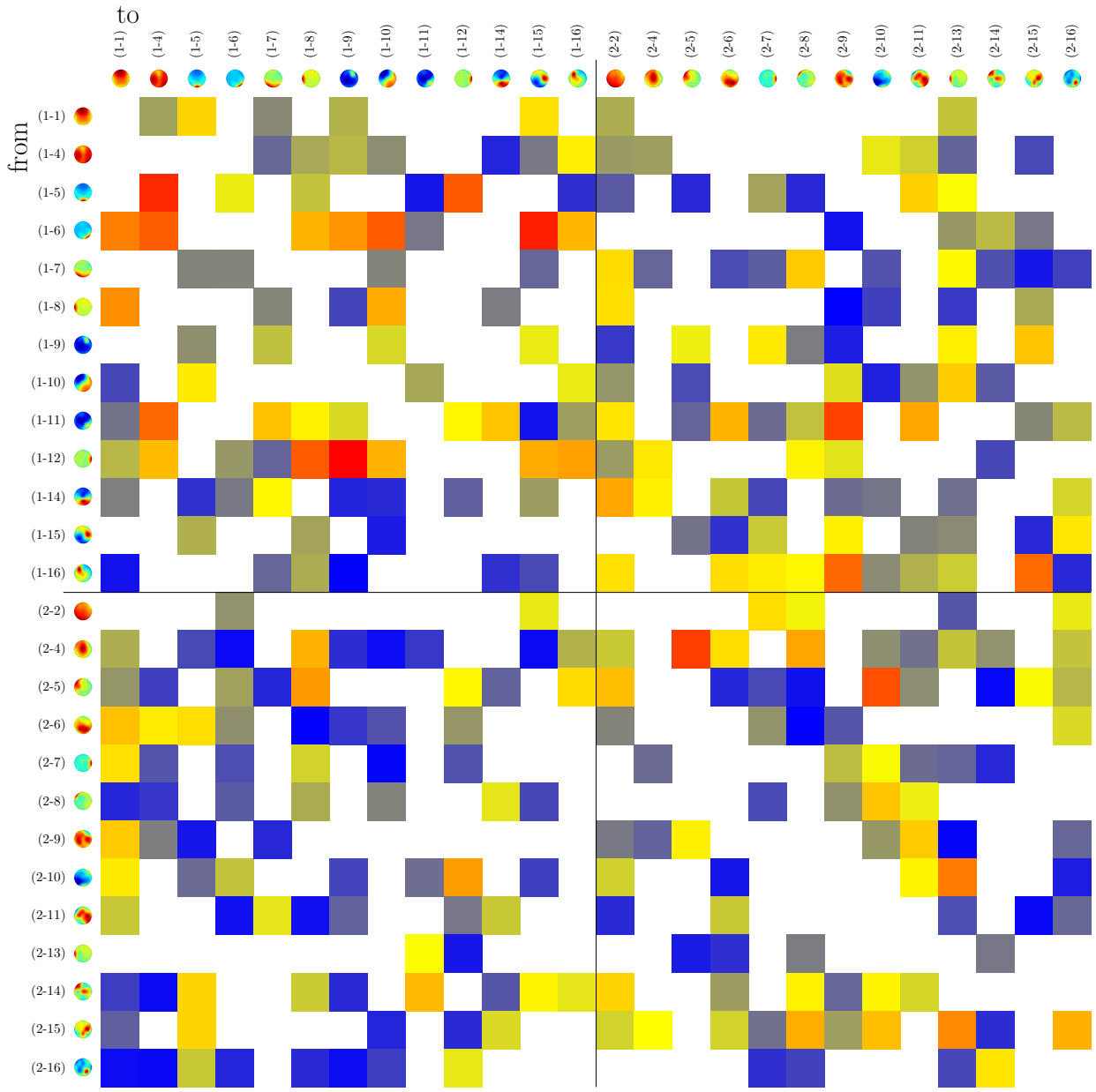


Figure G.18.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment five for α-band. Again no significant changes and no remarkable biases or patterns.

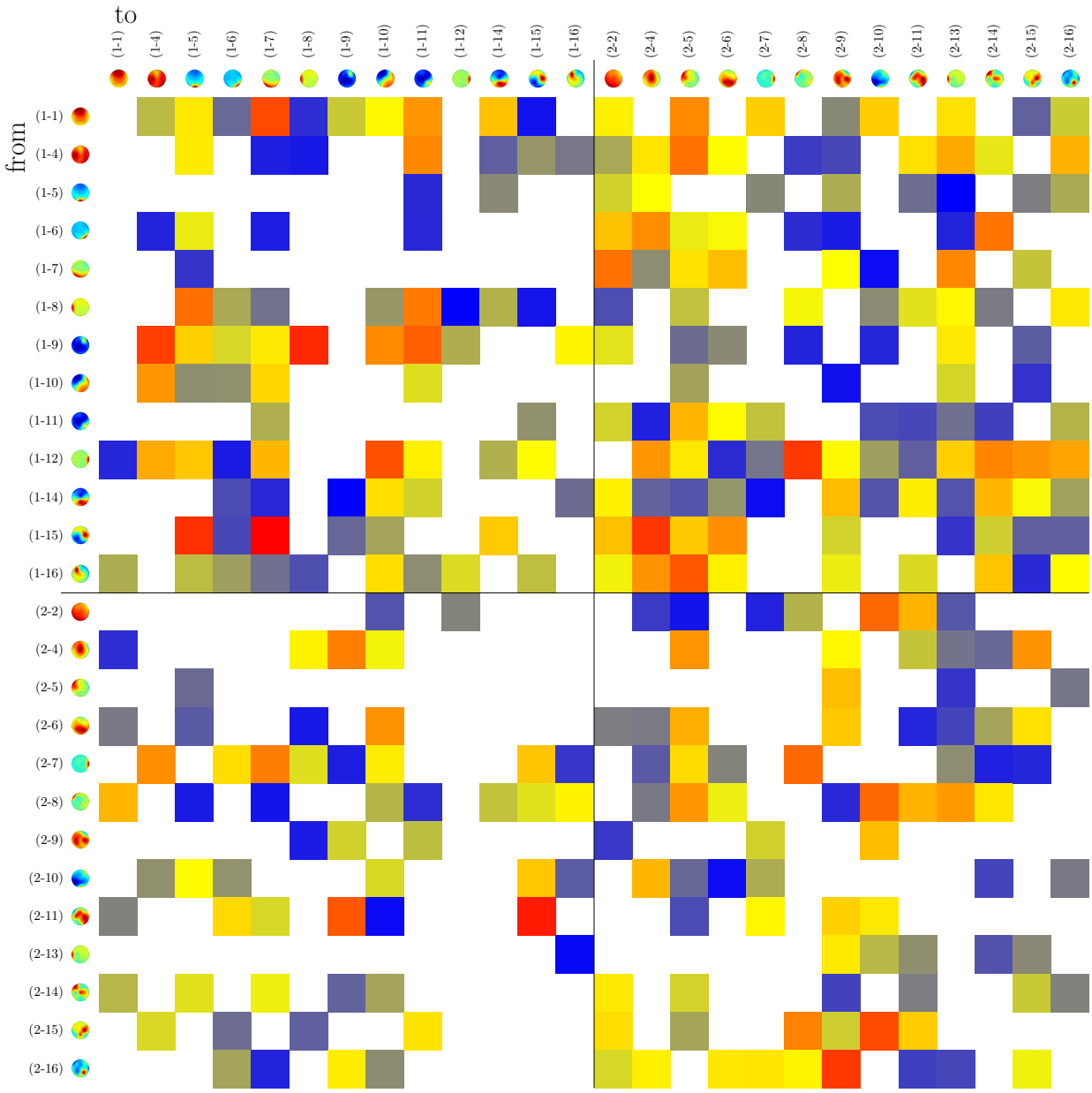


Figure G.19.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment five for β-band. Once more no significant changes or other remarkable features.

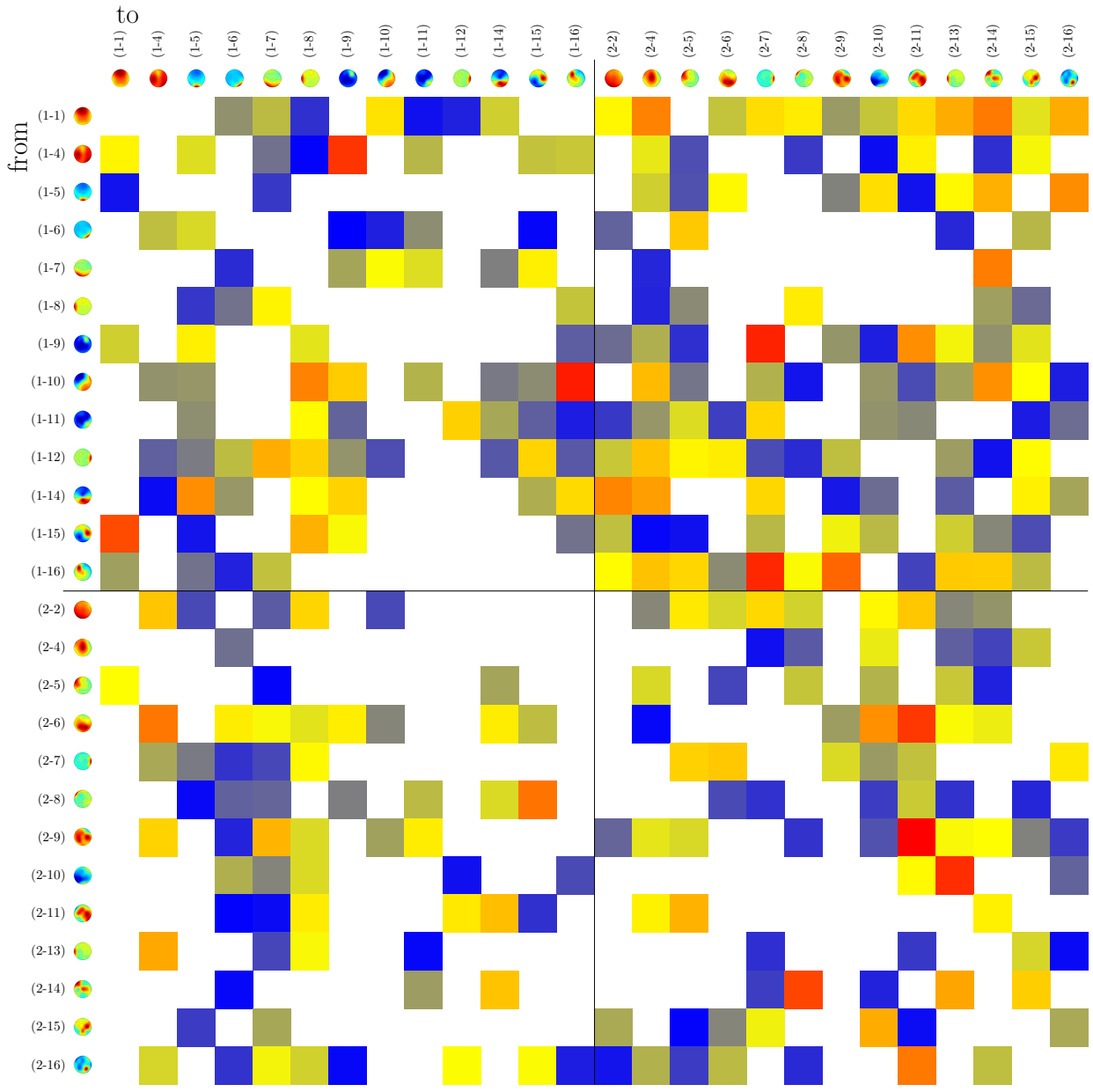


Figure G.20.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment five for γ-band. No significant changes, no biases and no remarkable patterns have been found.

G.6. Experiment Six

Much more significant changes than in experiment five, namely 19. In contrast to the first four experiments, significant connectivity changes appear predominantly in θ- and α-band.

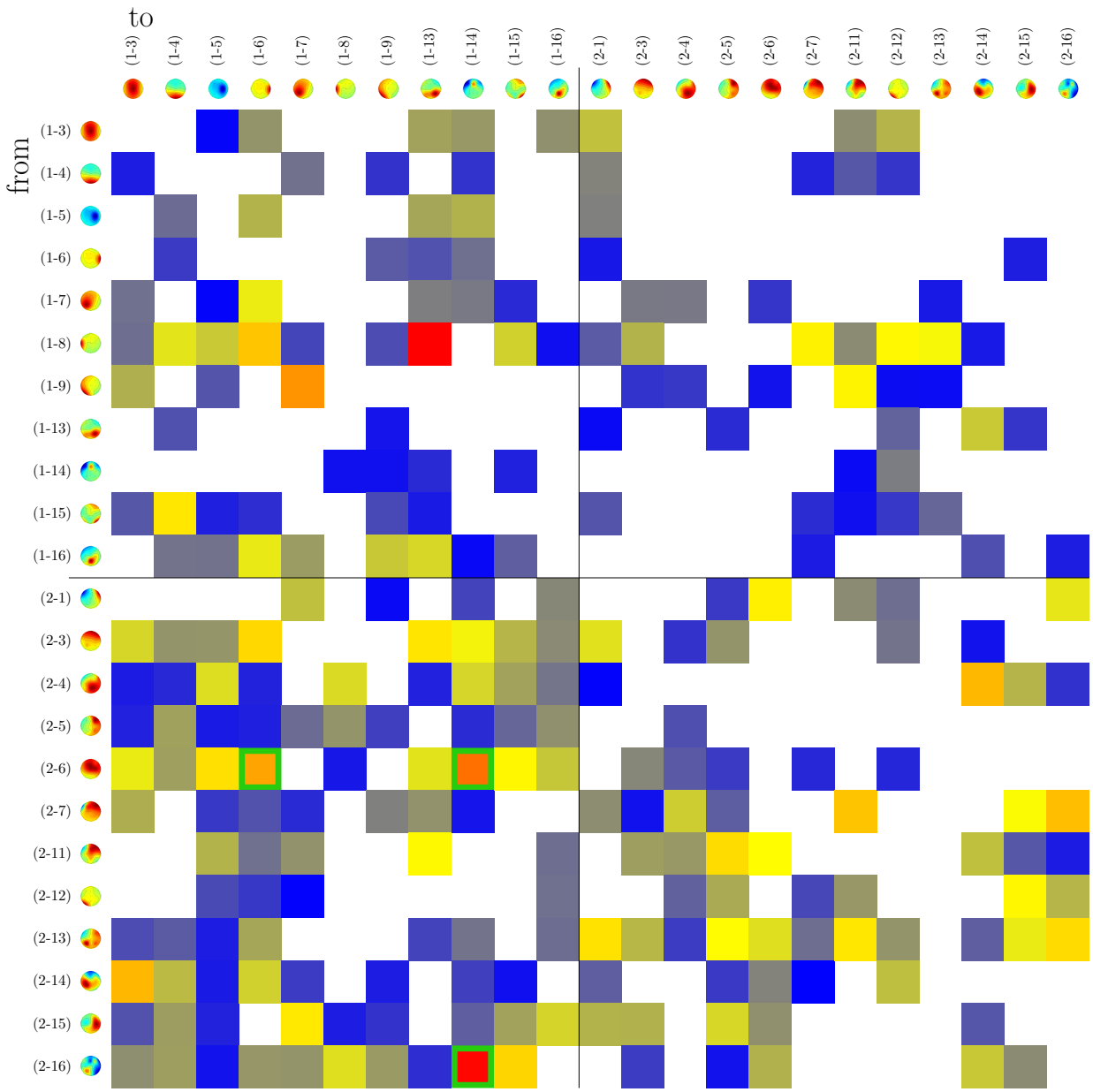


Figure G.21.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment six for θ-band. Three hyper-connections from participant two to participant one showed significant changes. Two of them originate from a broad central component and influence a lateral and a narrow fronto-central component. The latter is also the receiver of the third connection showing significant changes, which originates from a narrow parietal-lateral component.

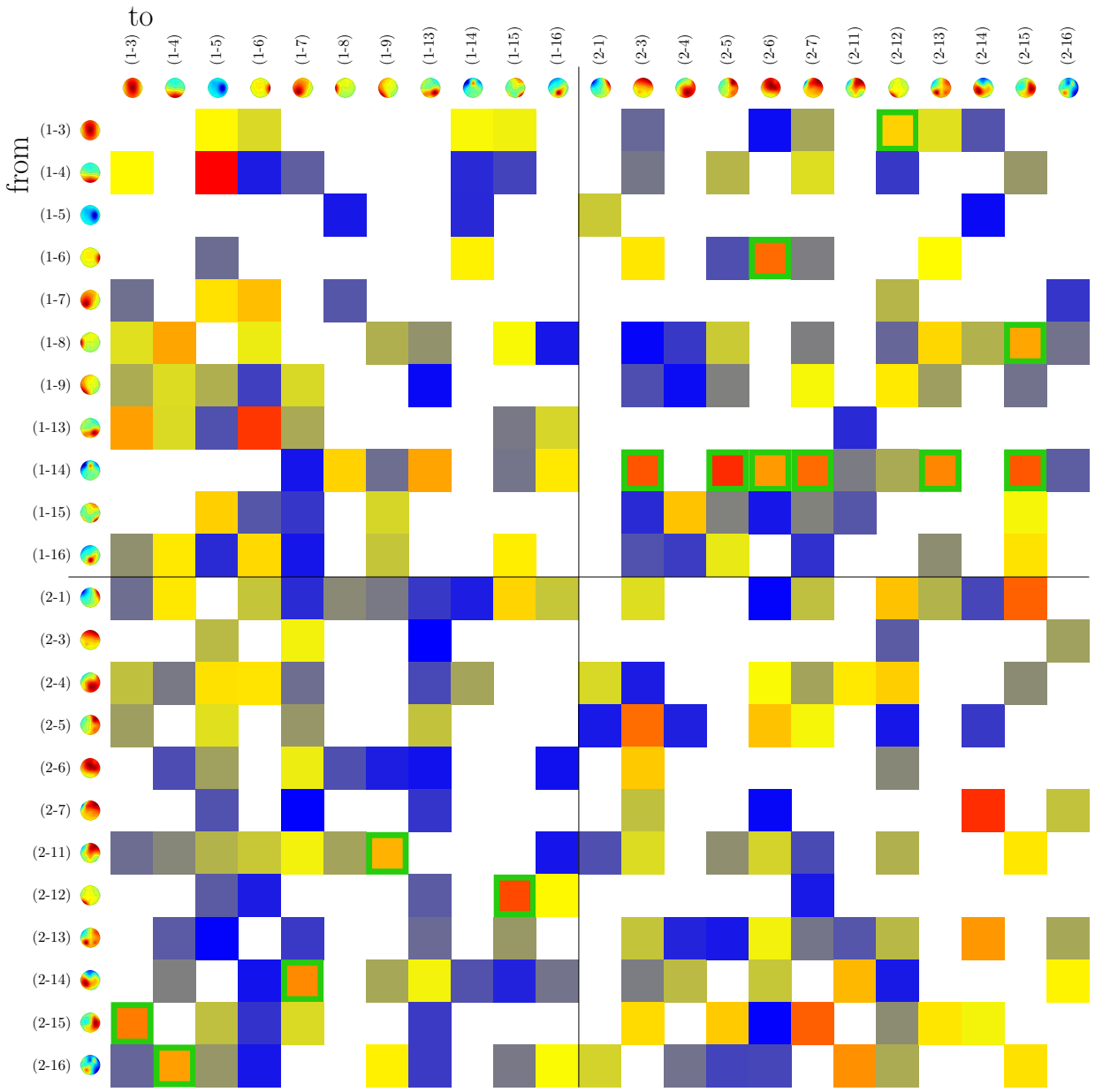


Figure G.22.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment six for α-band. A total of 14 connections were subject to significant changes. Five directed from participant two to participant one and the others from participant one to participant two. Remarkable is that a narrow fronto-central component of participant one is the source for five of these connections. The complex network spanned by these connections depicted in figure 8.7 on page 107.

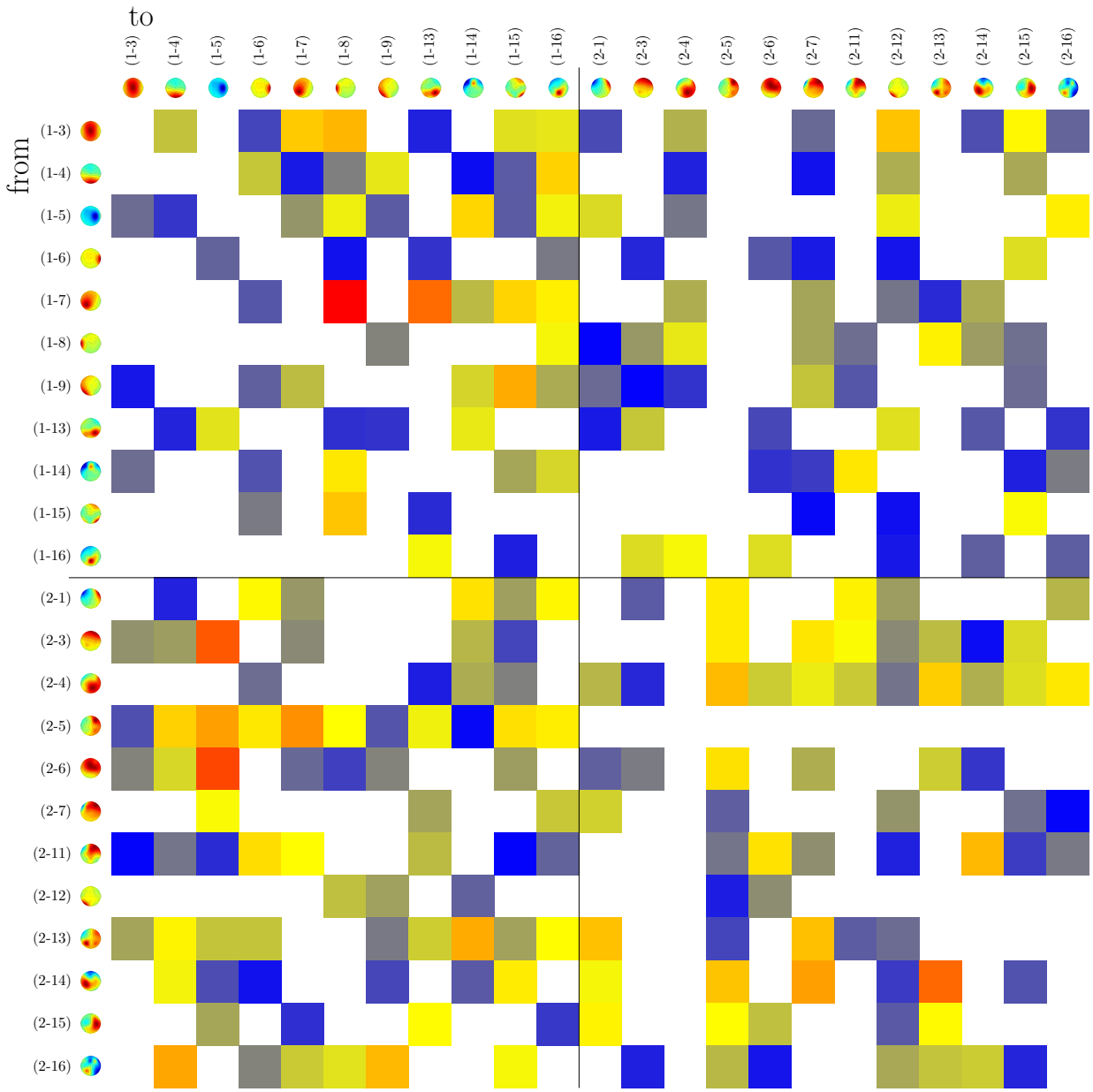


Figure G.23.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment six for β-band. No significant changes have been found and the distribution of connectivity values is unremarkable.

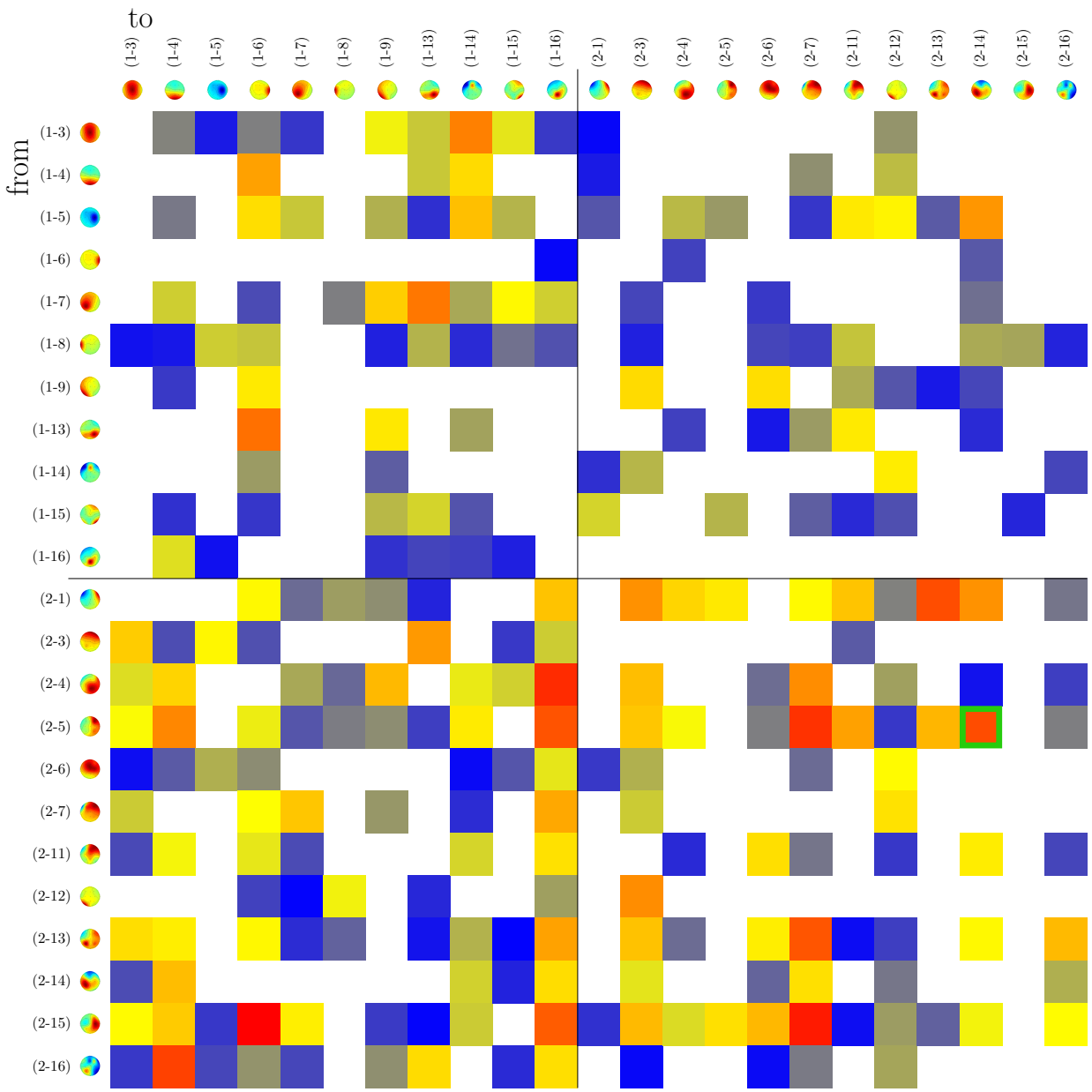


Figure G.24.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment six for γ-band. One within-participant connection from a lateral to parietal-lateral component of participant two underwent significant change. Furthermore there is a slight bias among hyper-connections favouring connections with participant two as the sender.

G.7. Experiment Seven

Four significant changes have been found. Three hyper-connections in the θ-band and one within-participant connection in β-band are affected.

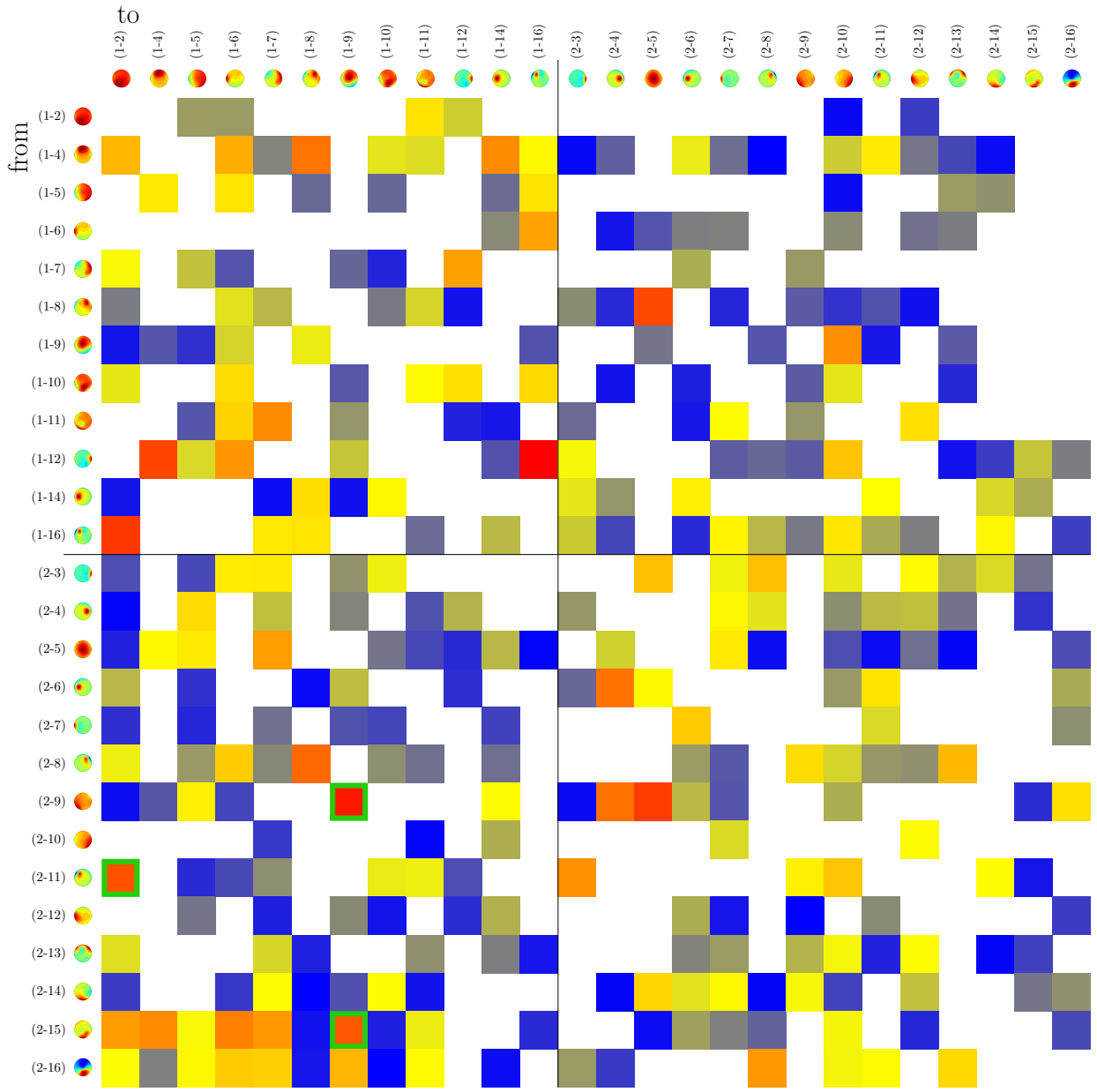


Figure G.25.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment seven for θ-band. There are three significant changes. All of them affect hyper-connections from participant two to participant one. One connects a narrow fronto-lateral component with a global component. One connects an occipital component with a fronto-central component. That receiver-component is also the receiver of the third connection, which originates from a broad parietal-lateral component.

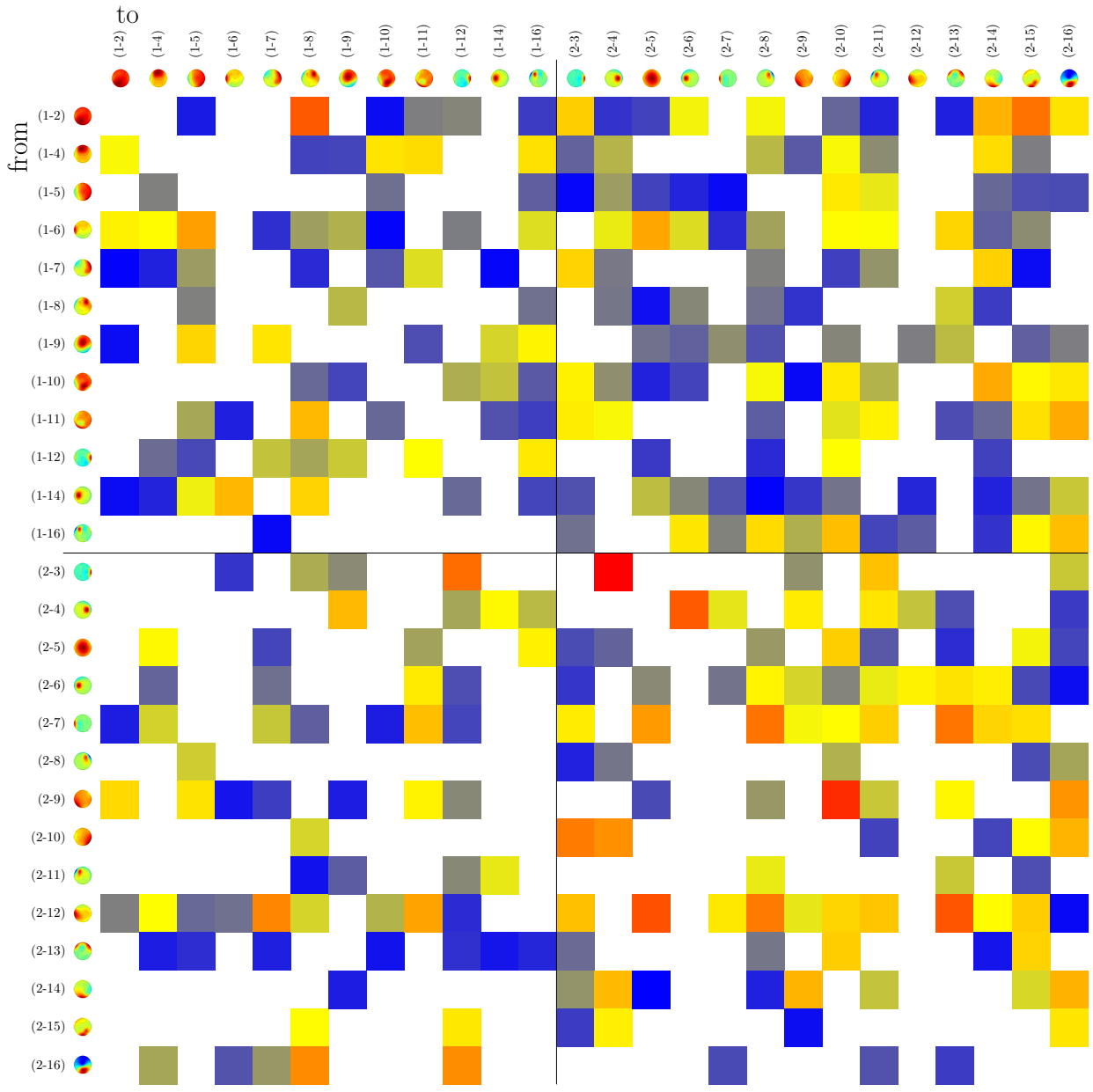


Figure G.26.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment seven for α-band. There are no significant changes. However, there seems to be a bias among hyper-connections favouring participant one as the sender.

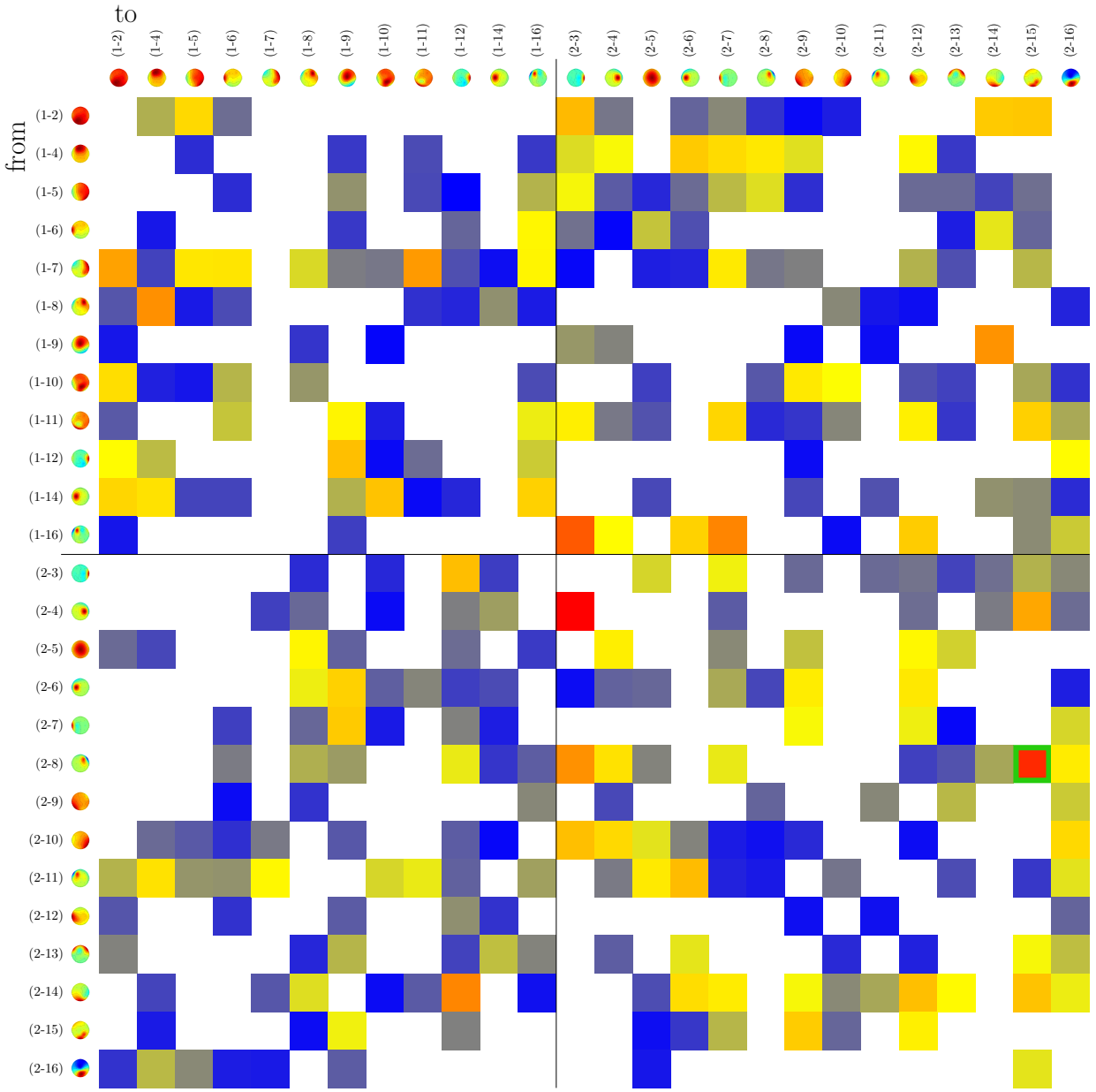


Figure G.27.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment seven for β-band. There is one significant change, which affects a within-participant connection for participant two, connecting a fronto-lateral component an occipital component.

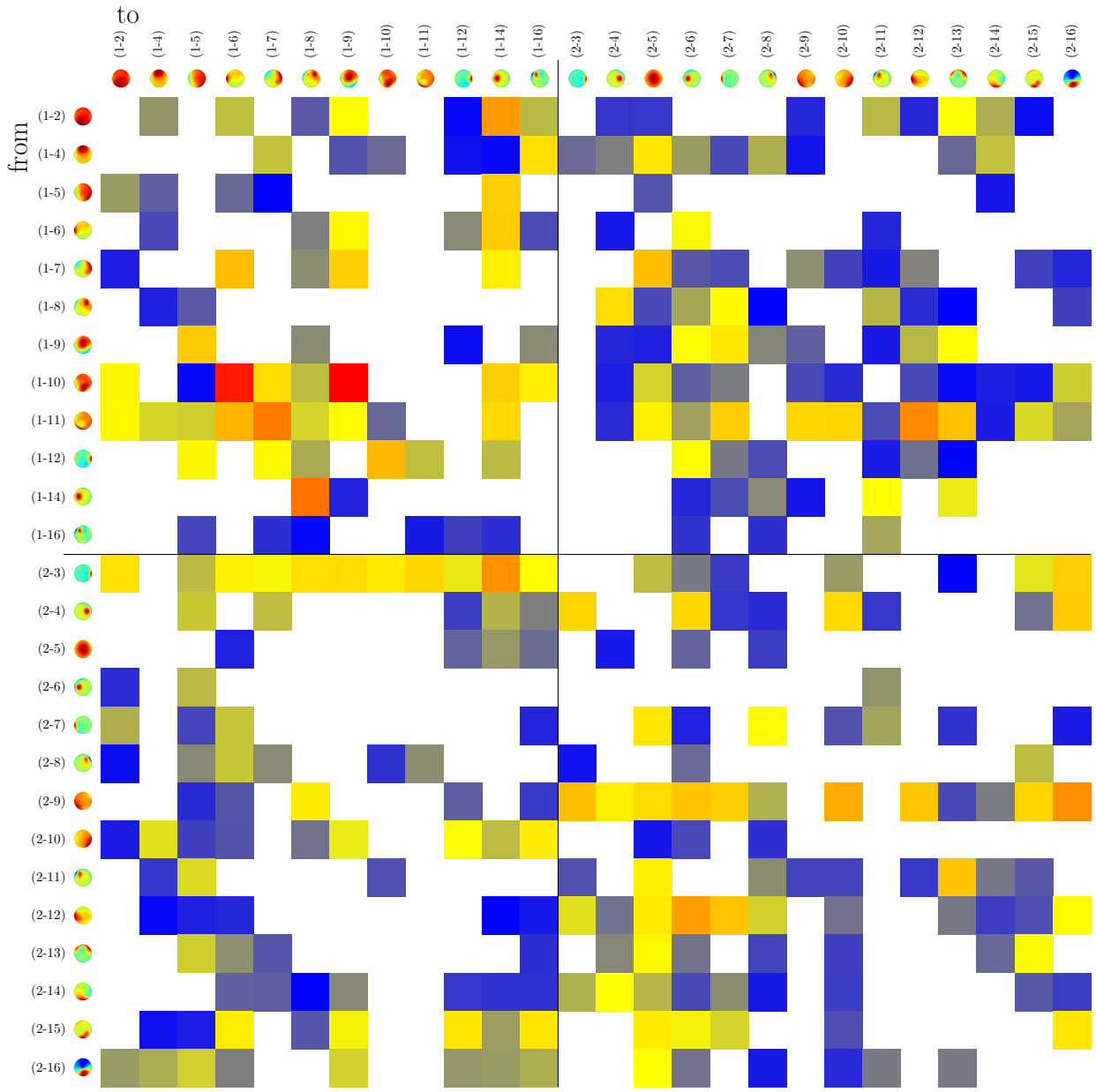


Figure G.28.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment seven for γ-band. There are no significant changes, nor remarkable patterns or biases.

G.8. Experiment Eight

No significant connectivity changes and little other remarkable observation were made in this analysis.

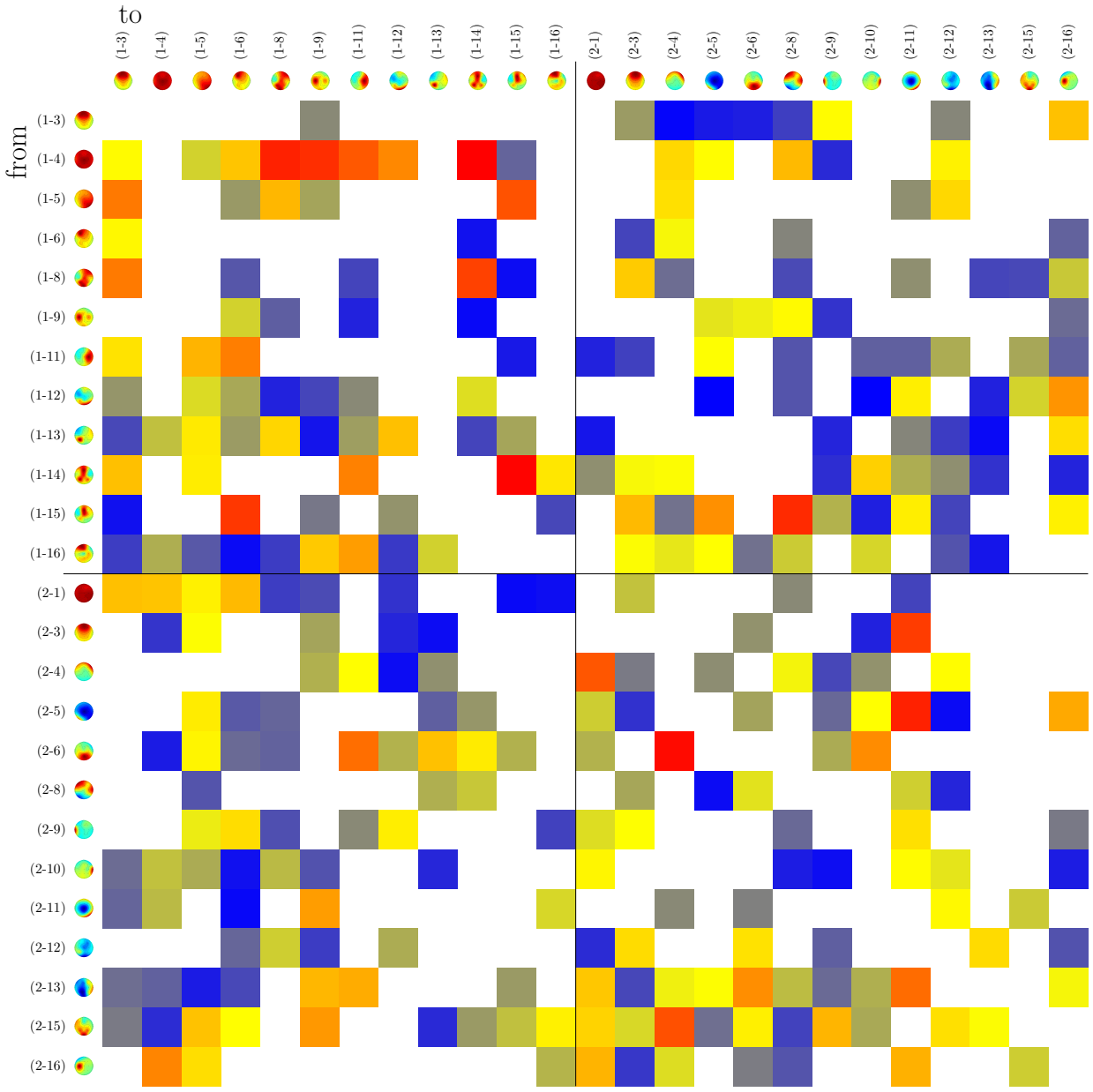


Figure G.29.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment eight for θ-band. There are no significant changes, nor remarkable patterns, nor biases observable in the connectivity estimation.

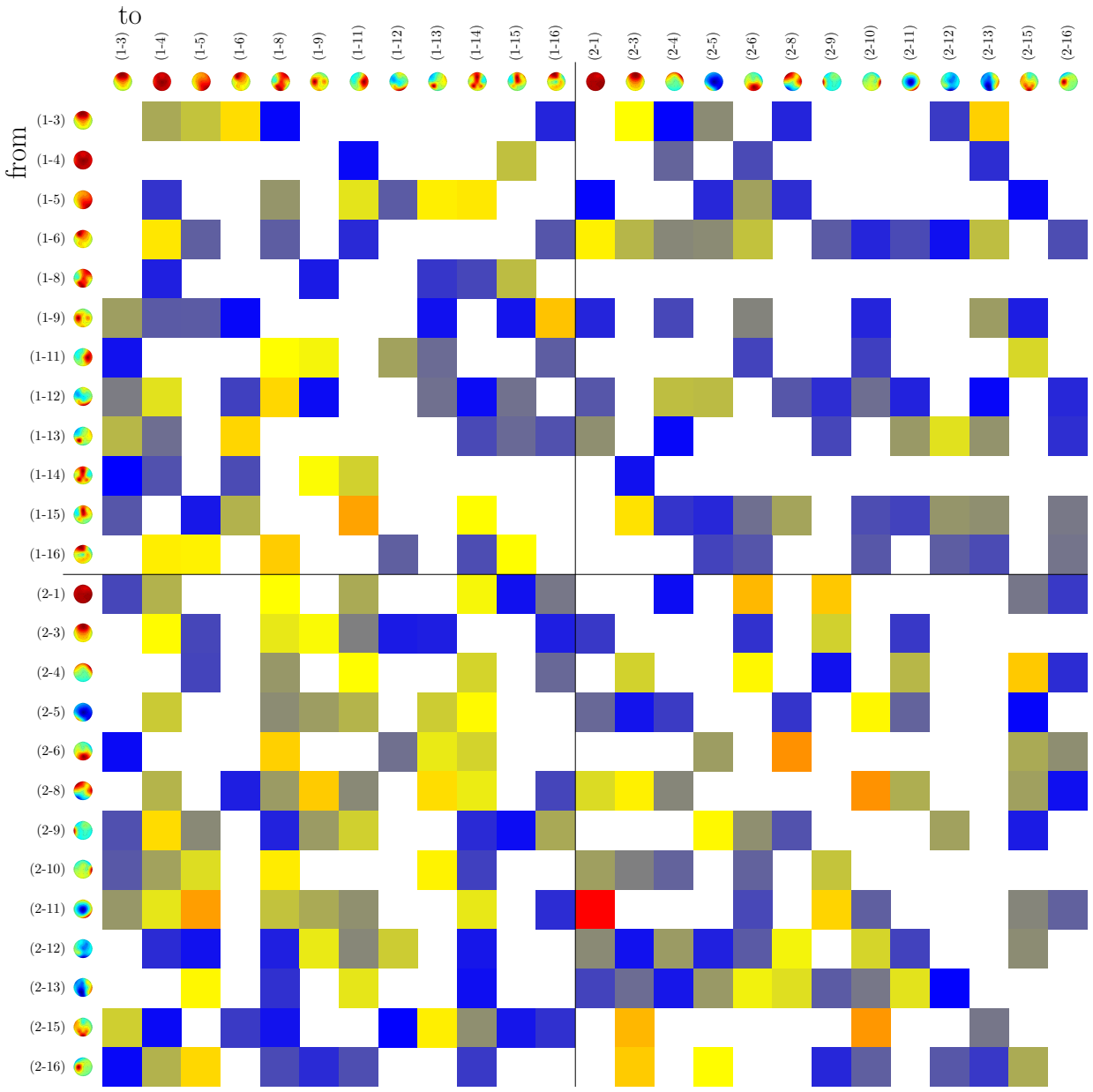


Figure G.30.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment eight for α-band. No significant changes could be found. It is, however, remarkable that one single, differential connectivity estimate surmounts all others.

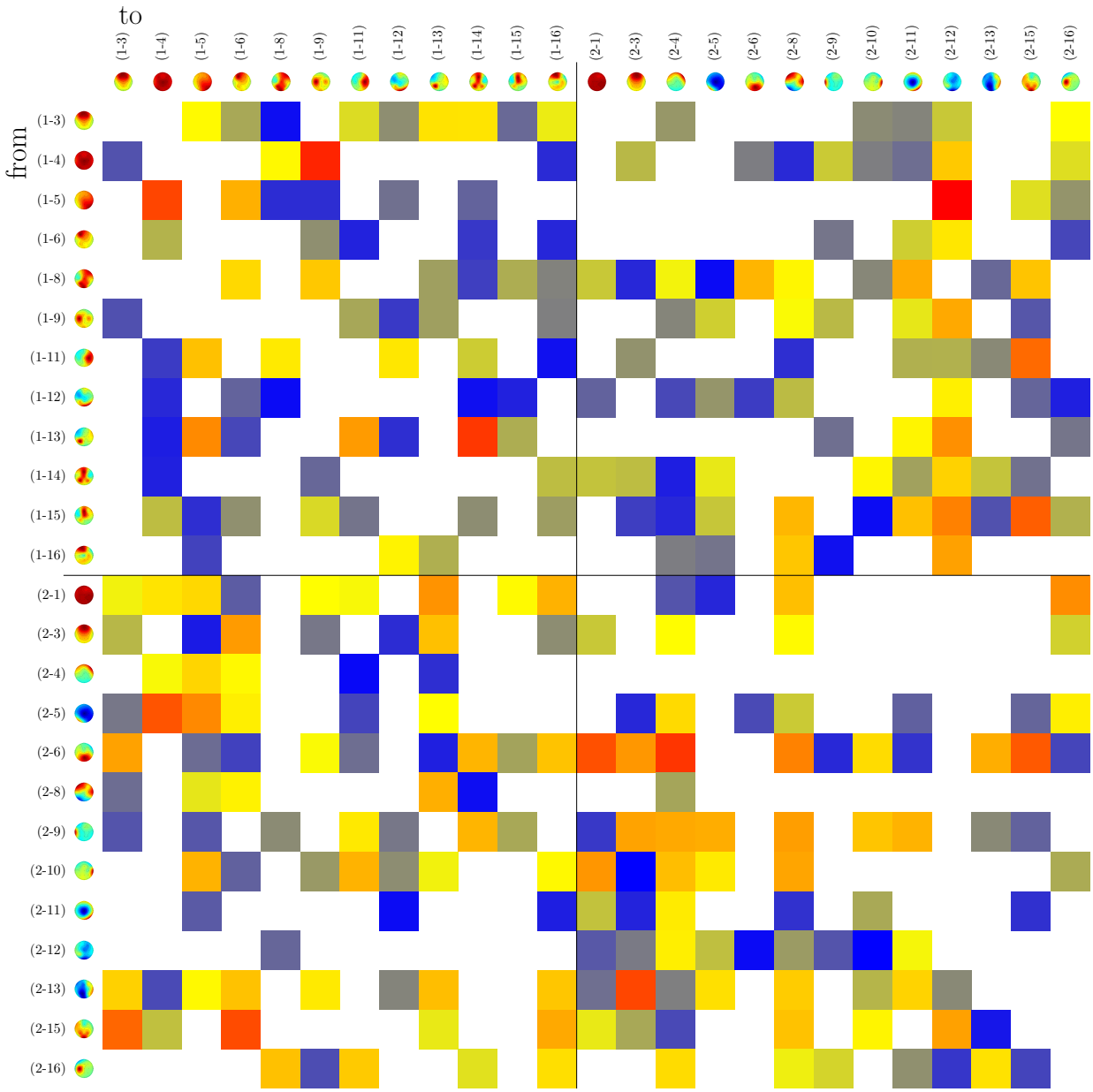


Figure G.31.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment eight for β-band. No significant changes were found. No interesting patterns, nor biases in the differential connectivity estimation.

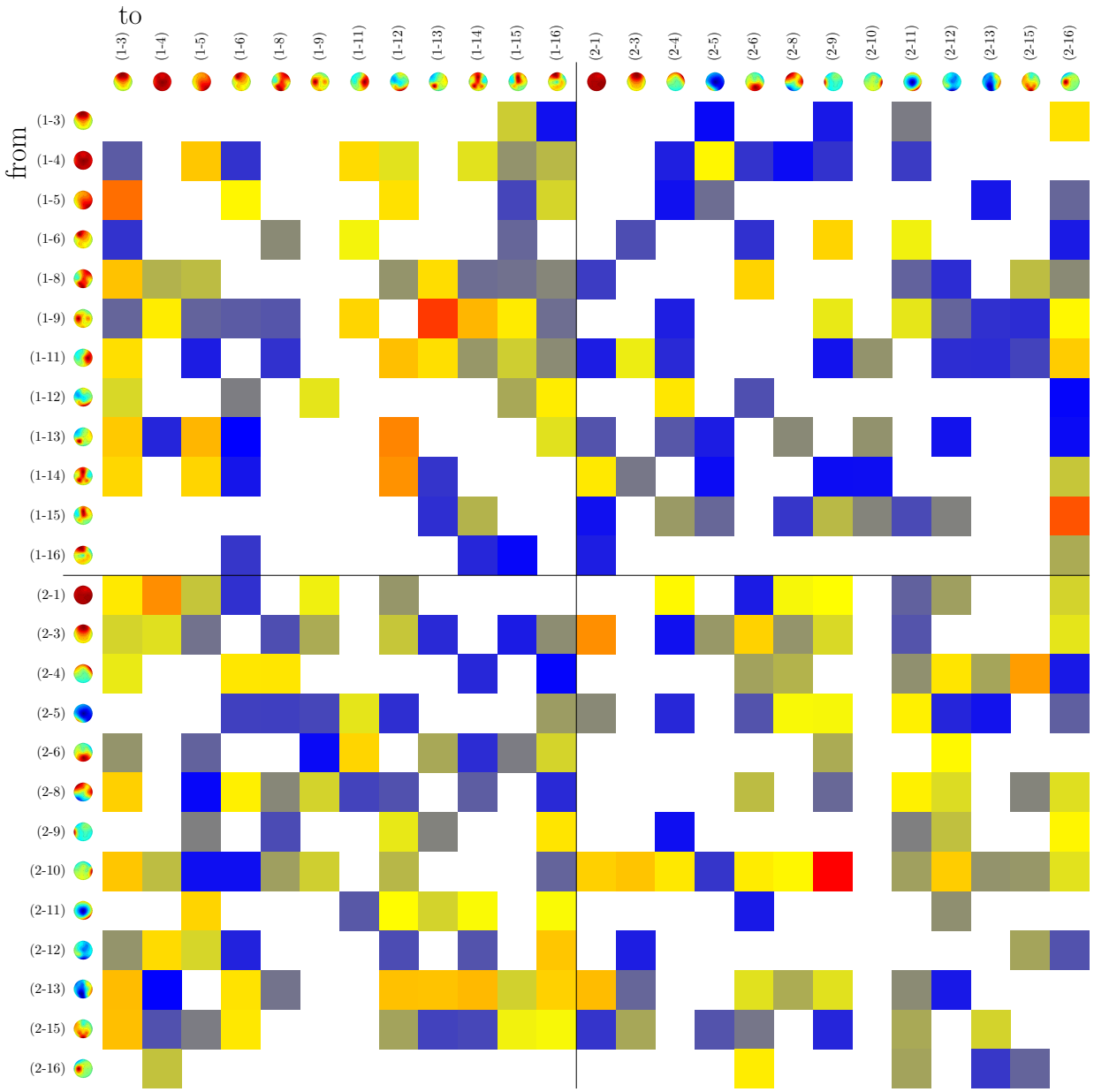


Figure G.32.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment eight γ-band. No significant changes, nor salient patterns, nor biases in the differential connectivity estimation.

G.9. Experiment Nine

No significant changes were found and little interesting observations were made in this analysis.

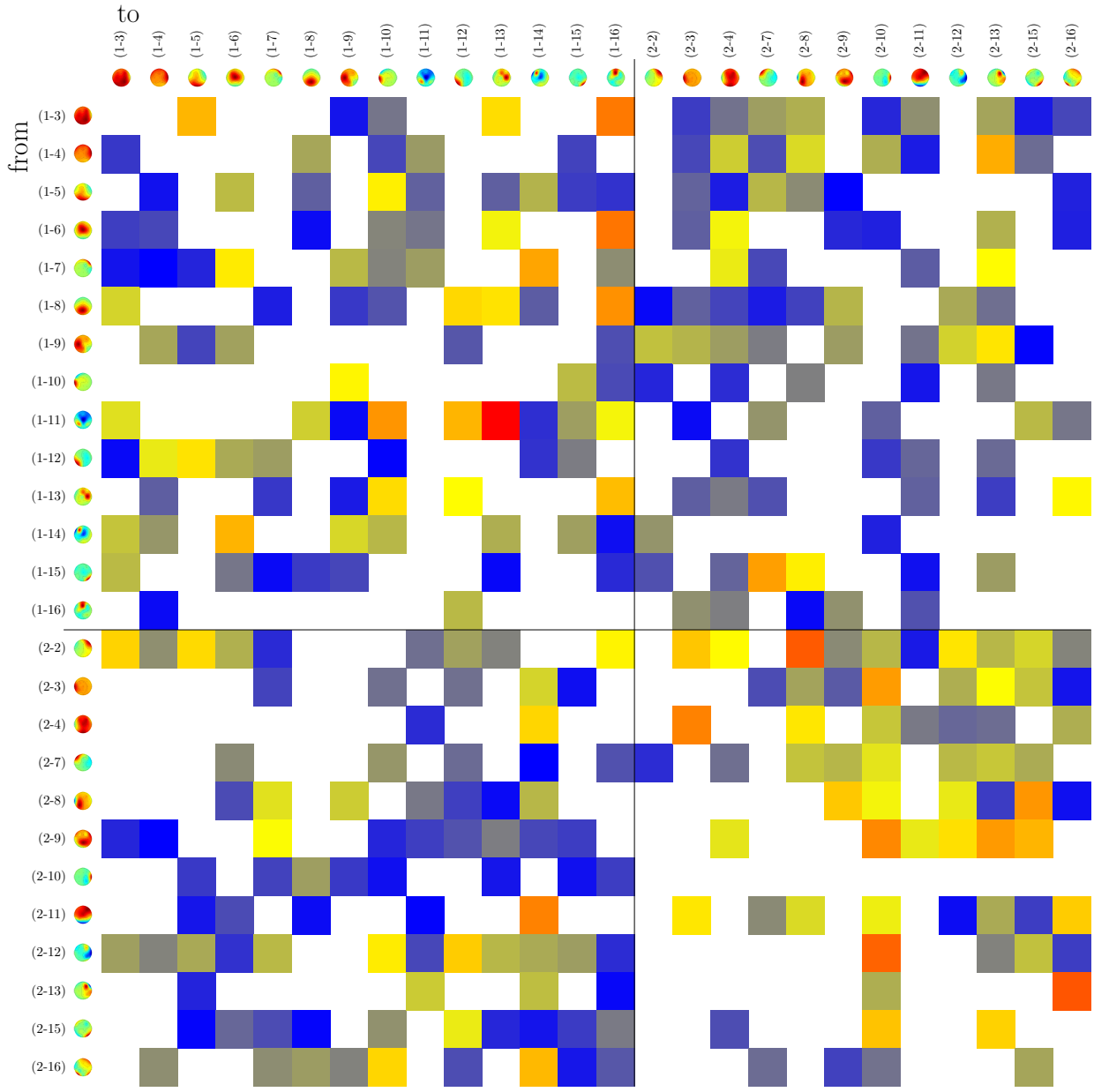


Figure G.33.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment nine for θ-band. No significant changes were found. The distribution of differential connectivity values lacks any interesting patterns or biases.

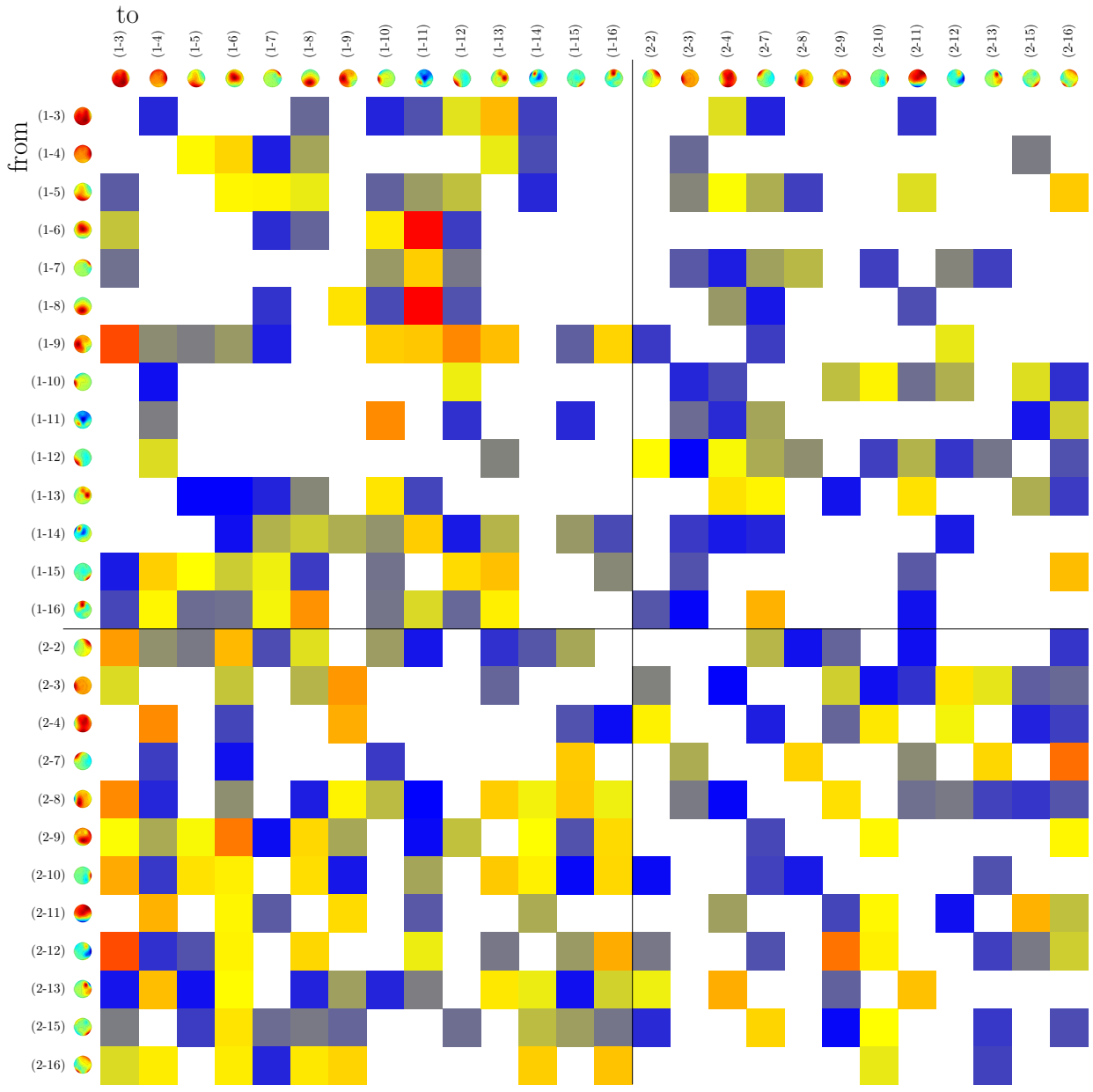


Figure G.34.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment nine for α-band. Again no significant changes, nor interesting observations.

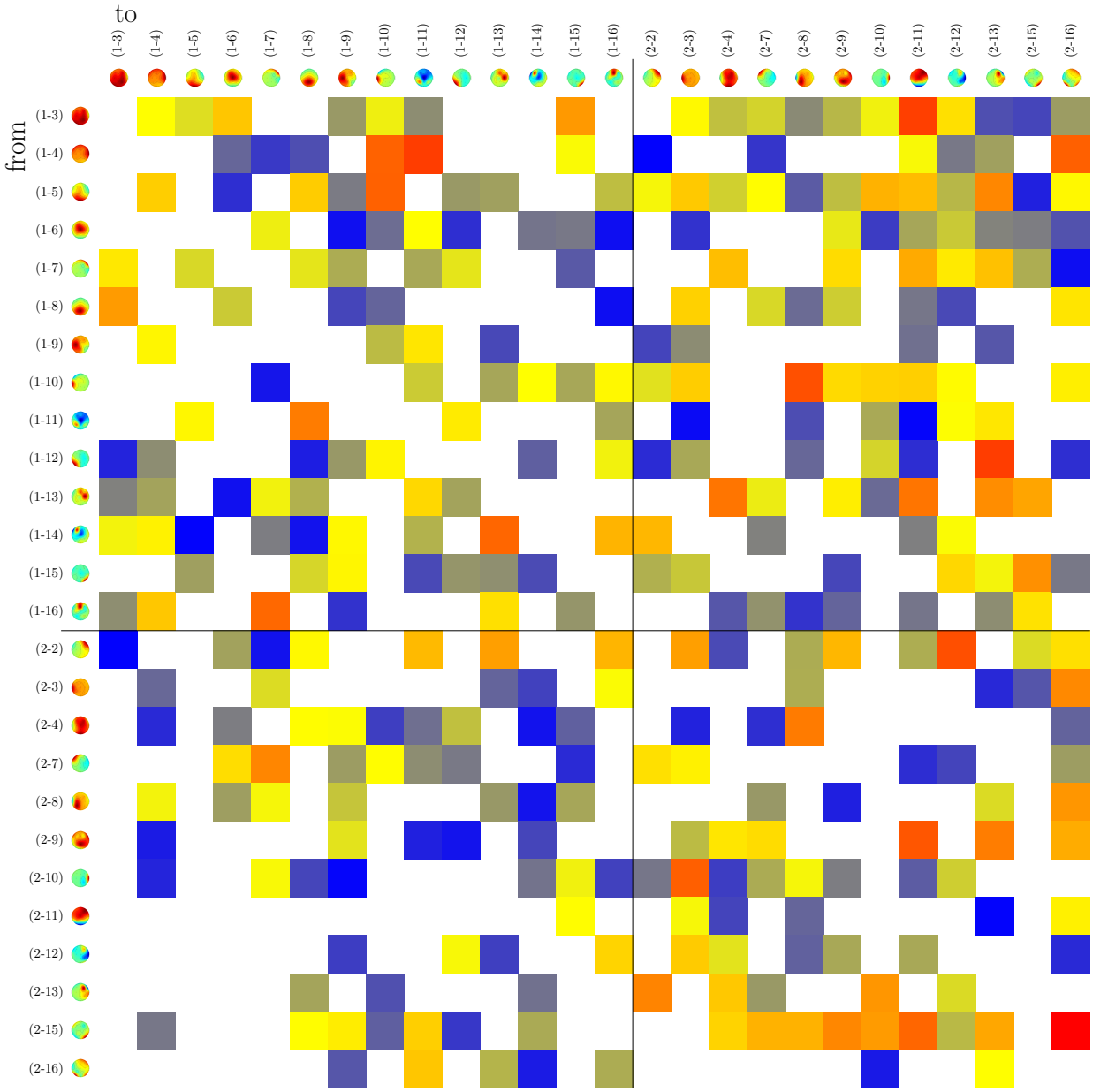


Figure G.35.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment nine for β-band. There seems to be a slight bias among hyper-connection preferring connections for which participant one as a sender. But no change reaches significance.

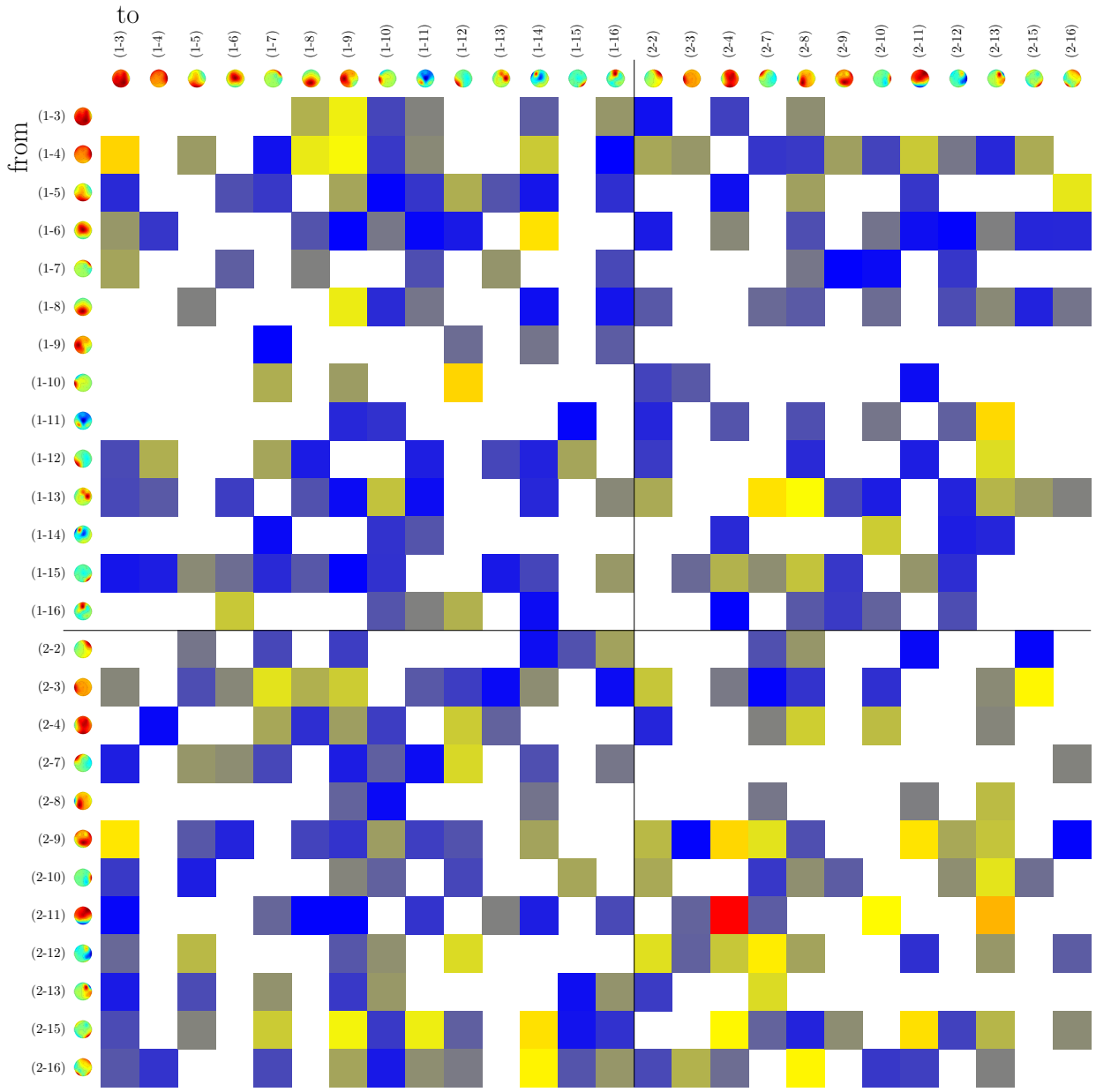


Figure G.36.: Differences in PSI connectivity between epochs for which the robot action had been initiated by participant one/two, respectively, in experiment nine for γ-band. Again no significant changes, but one within-participant connection's differential PSI value surmounts all others.

H. RPDC Analysis

H.1. Experiment One

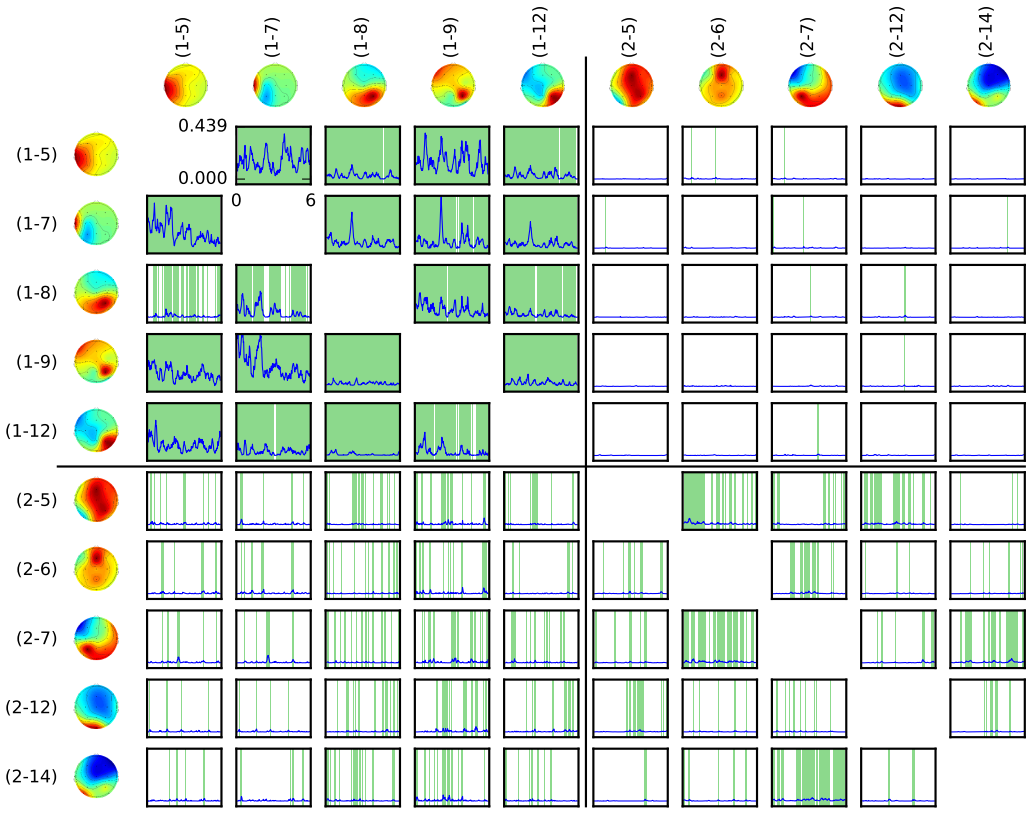


Figure H.1.: RPDC analysis of experiment one in the θ-band. The within-participant connections of participant one reach much higher values and are deemed significant much more often that those of participant two. The PDC values of within-participant one connections are also relatively high compared to those of other experiments. Furthermore, hyper-connections with participant two as a sender are significant more often than those in the reverse direction.

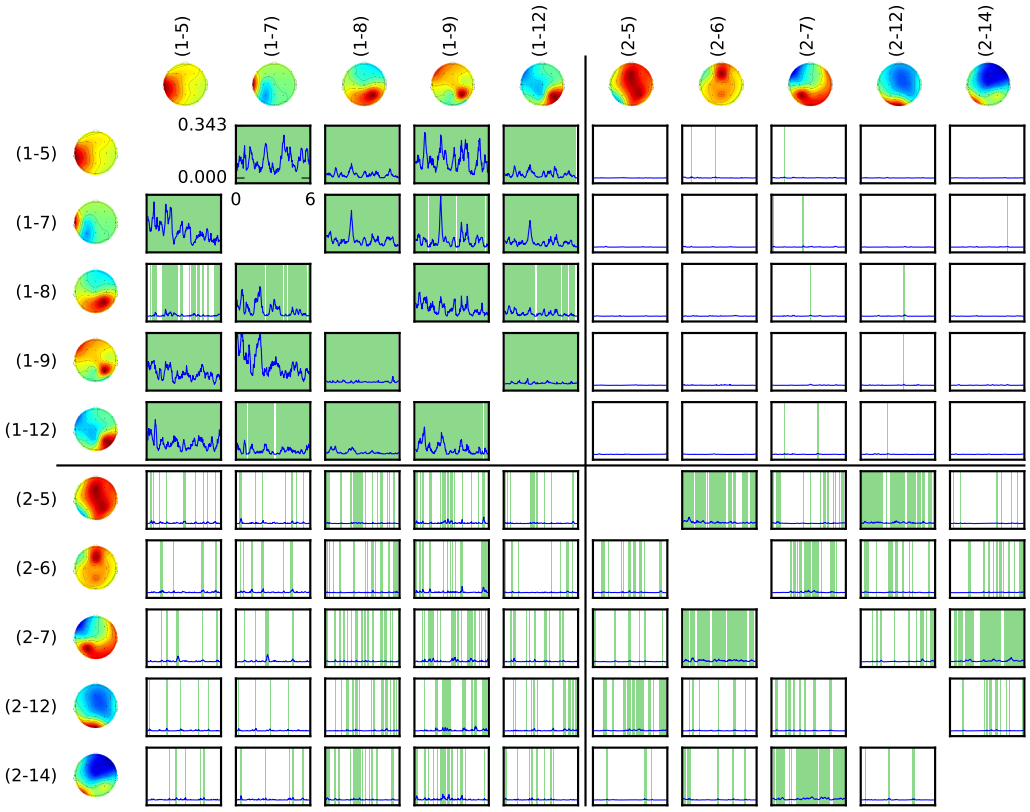


Figure H.2.: RPDC analysis of experiment one in the α-band. The results are similar to those of the θ-band. Much higher and more significant PDC values for participant one and for hyper-connections from participant two to participant one. The PDC values are a bit lower than for θ-band, but still pretty high.

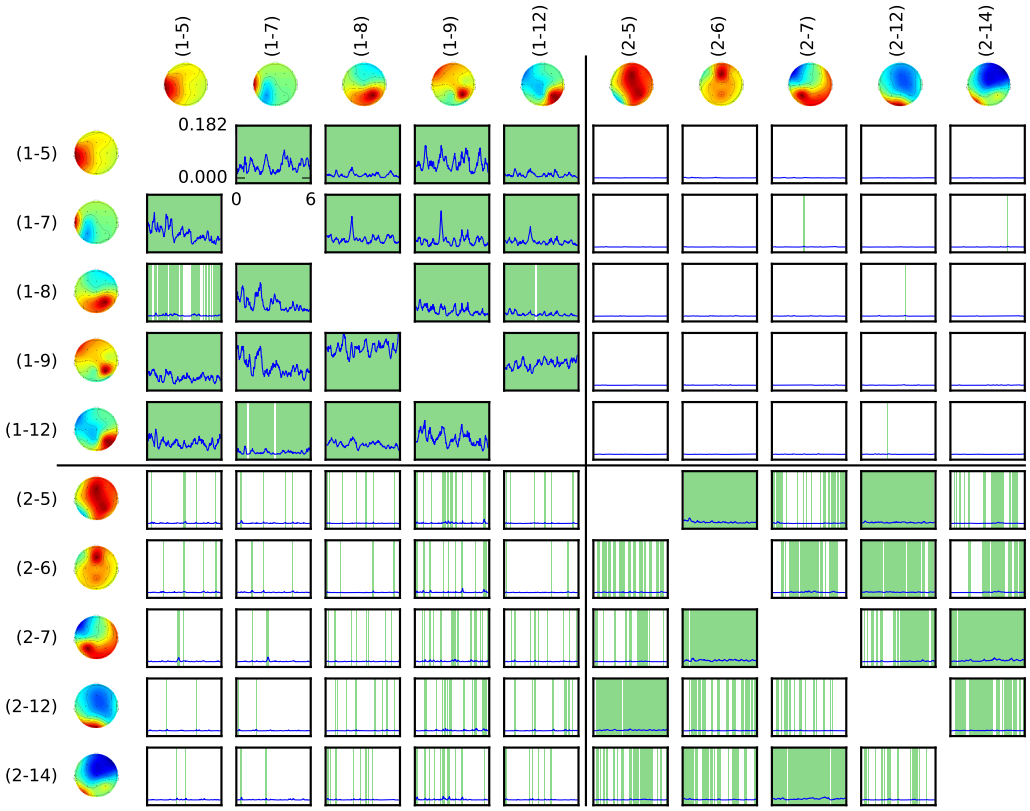


Figure H.3.: RPDC analysis of experiment one in the β-band. Again the came observations can be made. Again the value range is lower than for α.

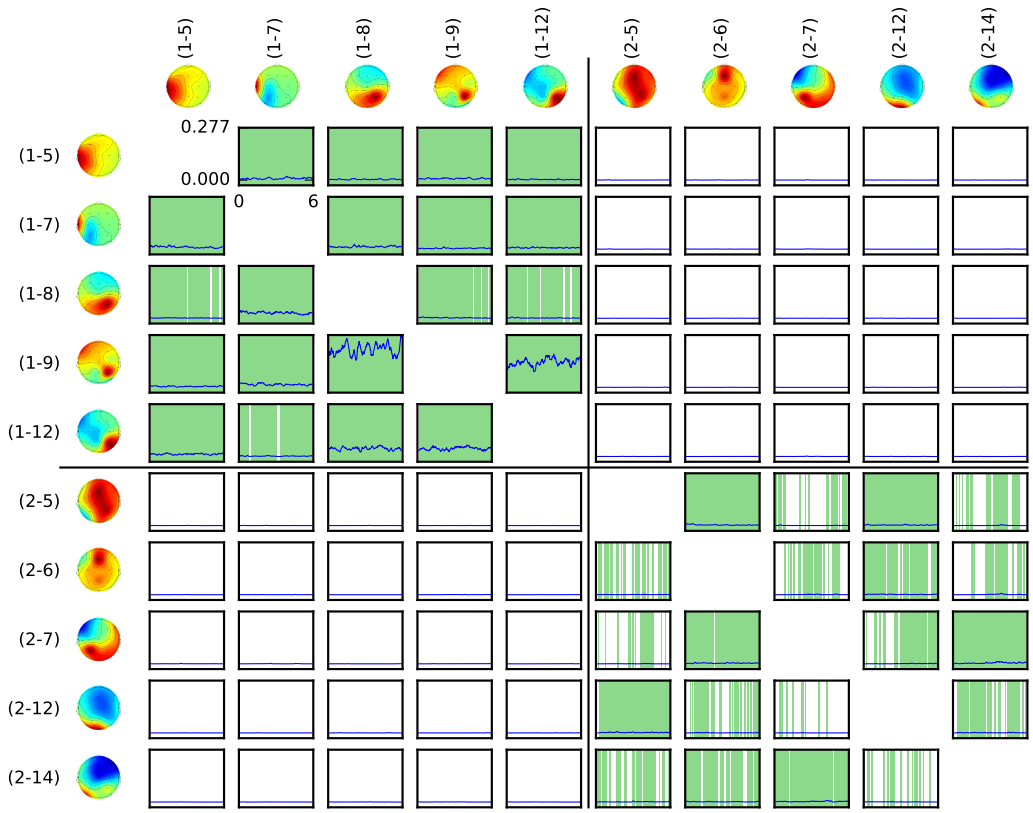


Figure H.4.: RPDC analysis of experiment one in the γ-band. Again the highest PDC values can be found for within-participant one connections. However, only two of these connections have these high values, while the others show pretty low values again. There are no significant results for hyper-connections, but more significant results for within-participant two connections. The value range is a bit higher than for β-band.

H.2. Experiment Two

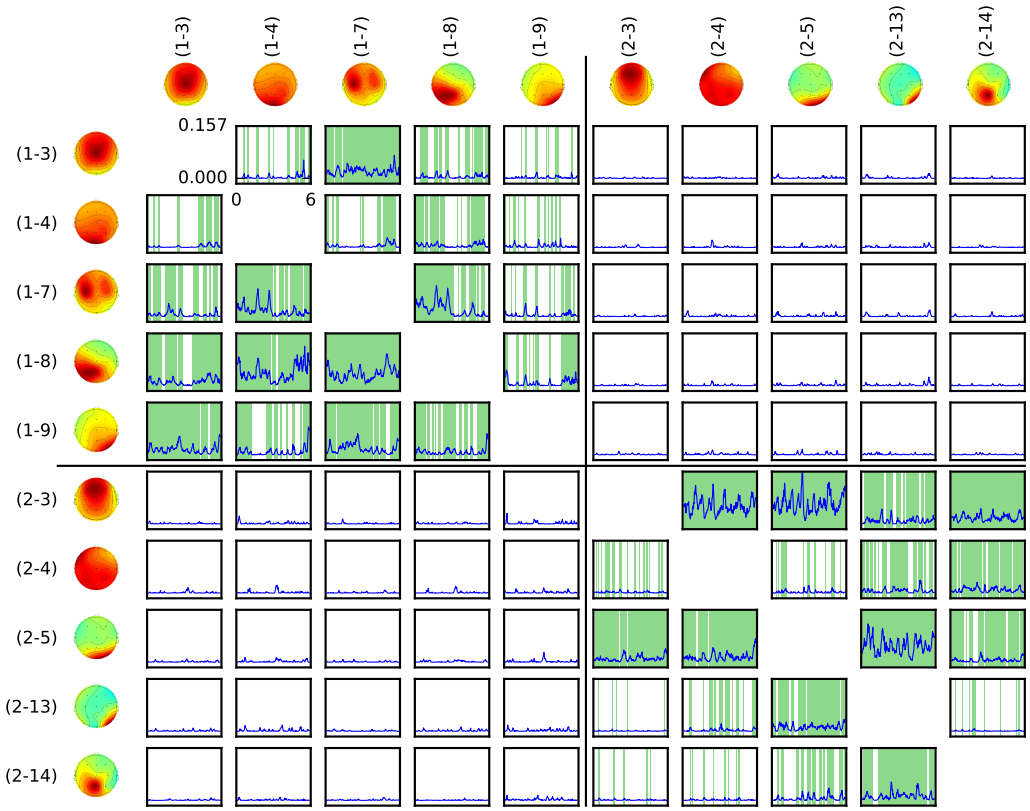


Figure H.5.: RPDC analysis of experiment two in the θ-band. In contrast to experiment one we here only have a comparable value range for within-participant connections of both participants. No significant results for hyper-connections could be found.

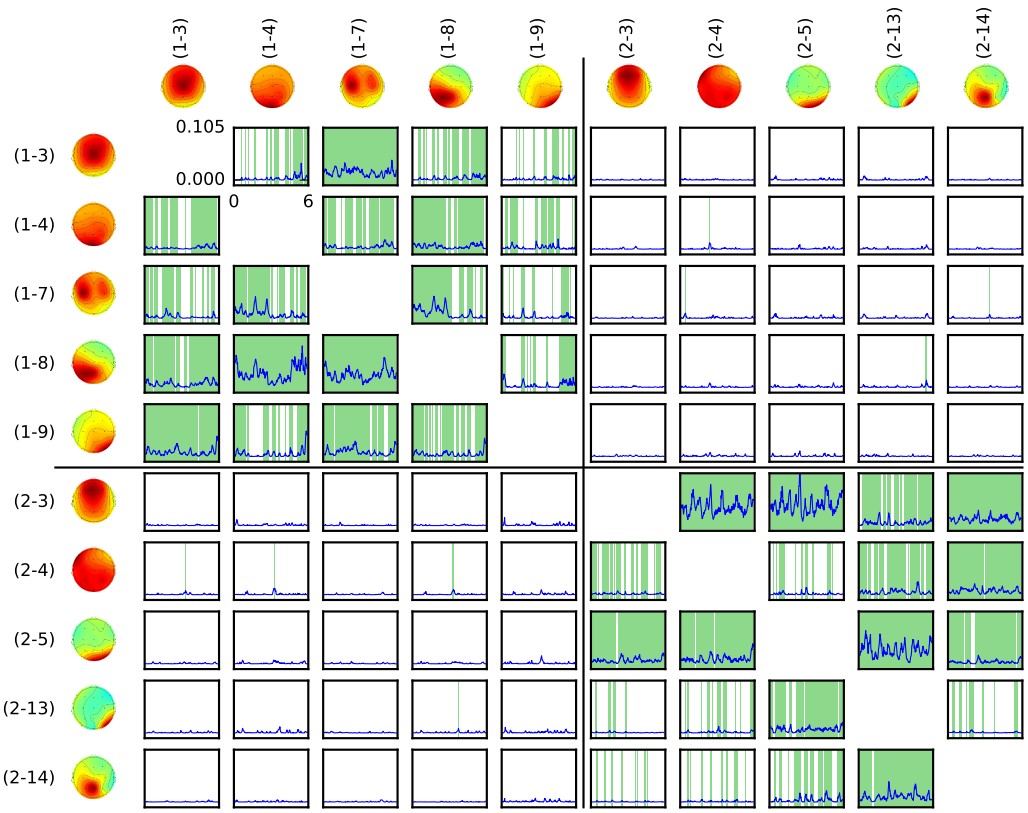


Figure H.6.: RPDC analysis of experiment two in the α-band. Only very few segments of the PDC values for hyper-connections are significant. Those are aligned with peaks in the PDC-curve. The connection from (2-3) to (2-4) was recognised as significant in the corresponding PSI baseline vs. experiment analysis. Here the connections has high, significant values over the entire six seconds.

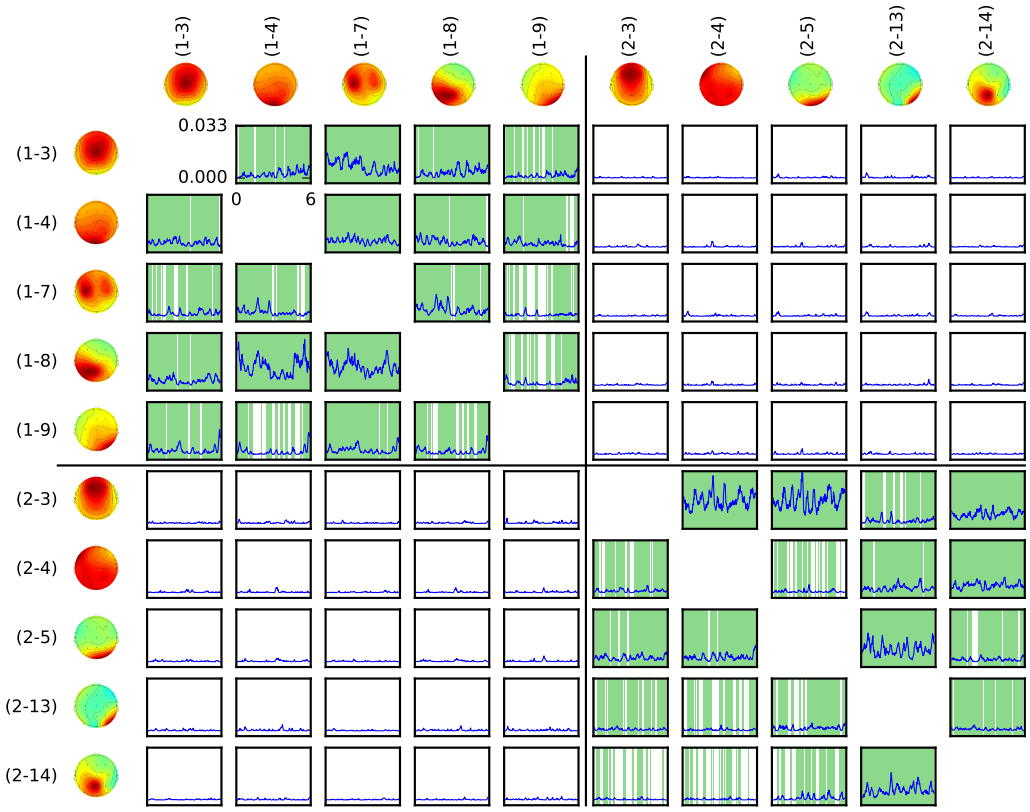


Figure H.7.: RPDC analysis of experiment two in the β-band. Again no significant hyper-connections could be found. The PDC values are pretty low, but still the within-participant connections show a coverage with periods deemed significant, comparable to other analyses.

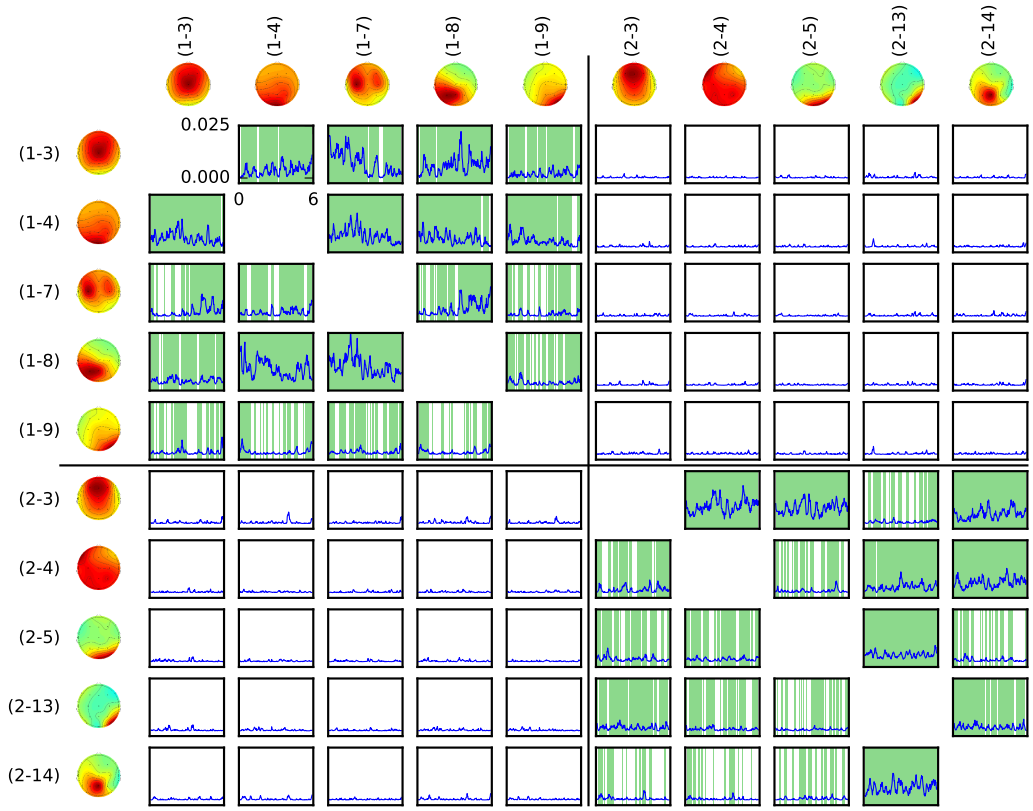


Figure H.8.: RPDC analysis of experiment two in the γ-band. The values are even lower than for β-band, yet the distribution of significant values is comparable.

H.3. Experiment Three

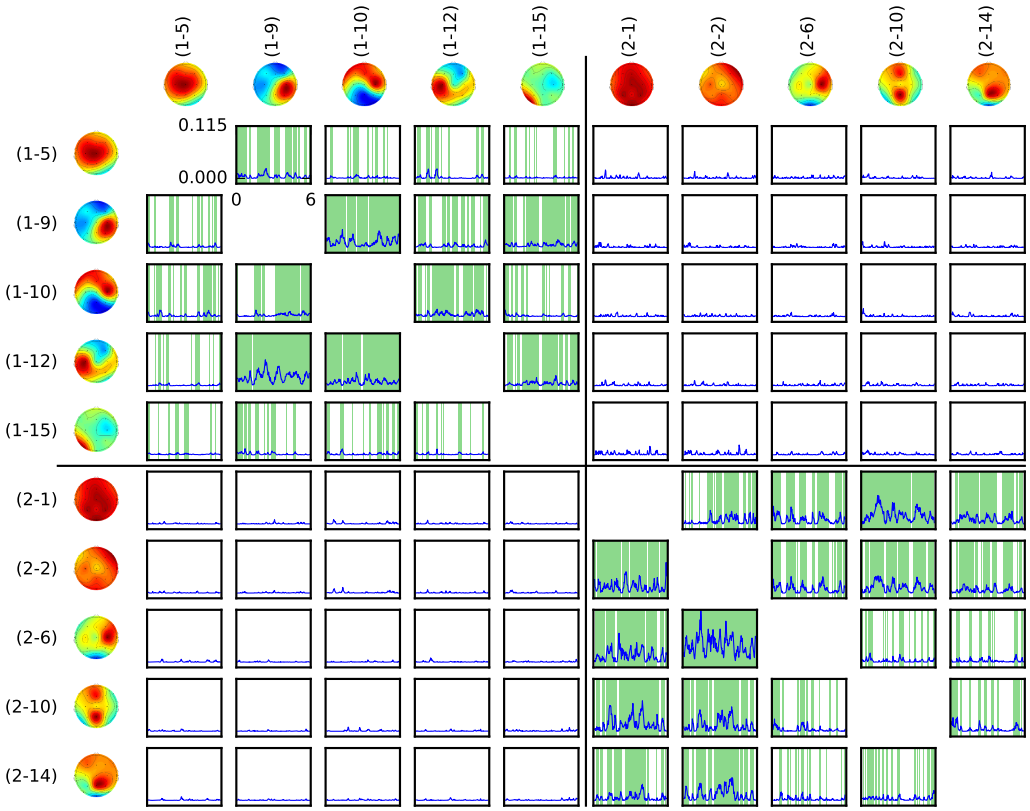


Figure H.9.: RPDC analysis of experiment three in the θ-band. I could not show any significant results regarding hyper-connections. Remarkable are some similarities in the PDC curves of (2-10) $\rightarrow$ (2-1) and (2-10) $\rightarrow$ (2-2) as well as for (2-1) $\rightarrow$ (2-10) and (2-2) $\rightarrow$ (2-10). This could be an indication that (2-1) and (2-2) as well as (2-10) and (2-14) are affected by a common driver possibly not included in the analysis.

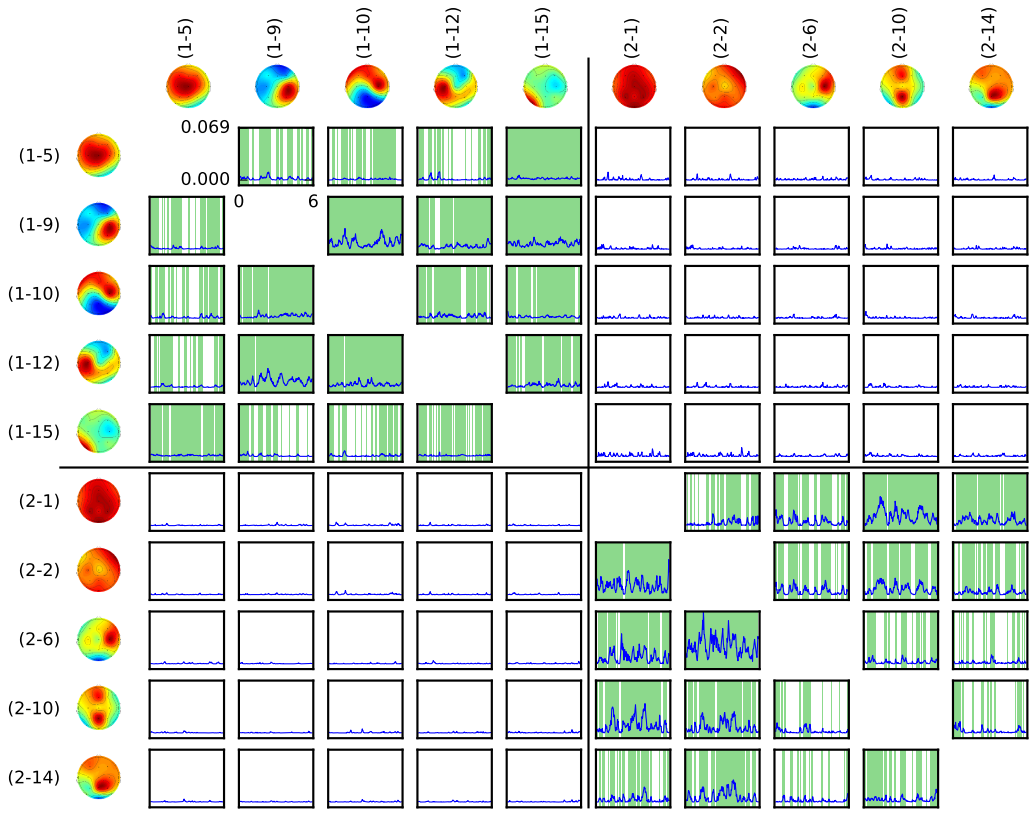


Figure H.10.: RPDC analysis of experiment three in the α-band. The values are generally pretty low. No significant values were found for hyper-connections. (1-15) $\rightarrow$ (1-9) was recognised as significant in the corresponding differential PSI analysis. Here the PDC values are particularly low, but still over large periods significant.

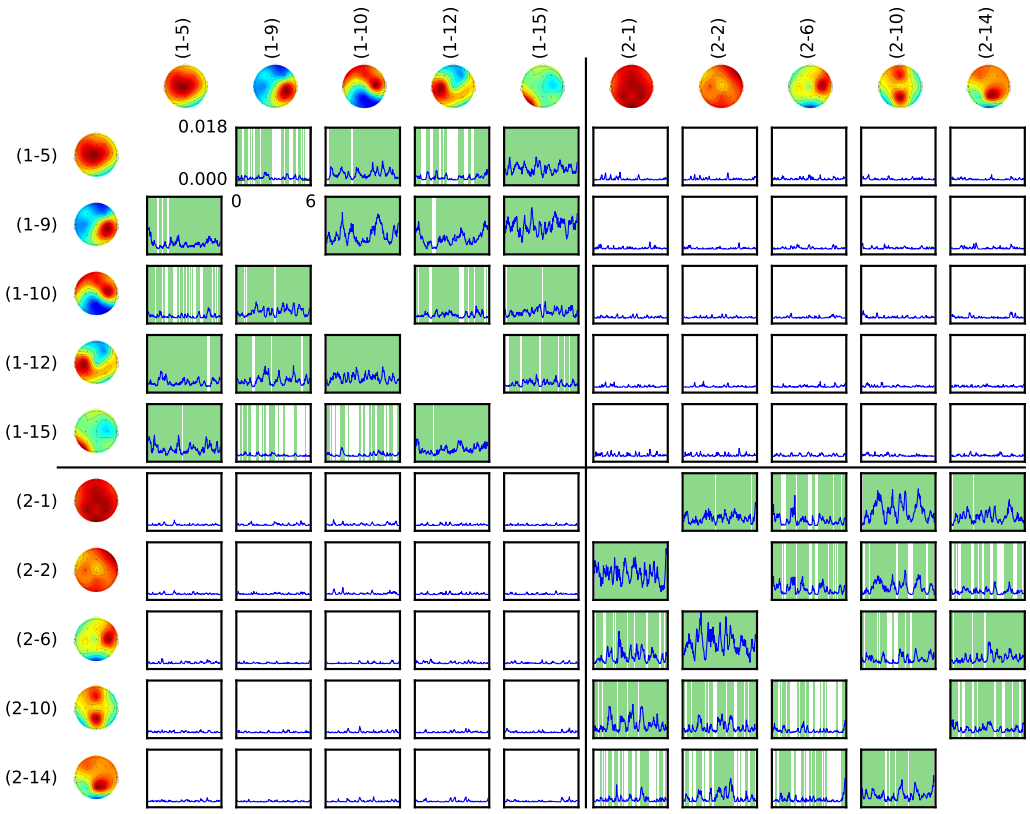


Figure H.11.: RPDC analysis of experiment three in the β-band. The values are remarkably small, but within the same range for both participants. Again we see that small values do not necessarily mean that fewer periods are significant.

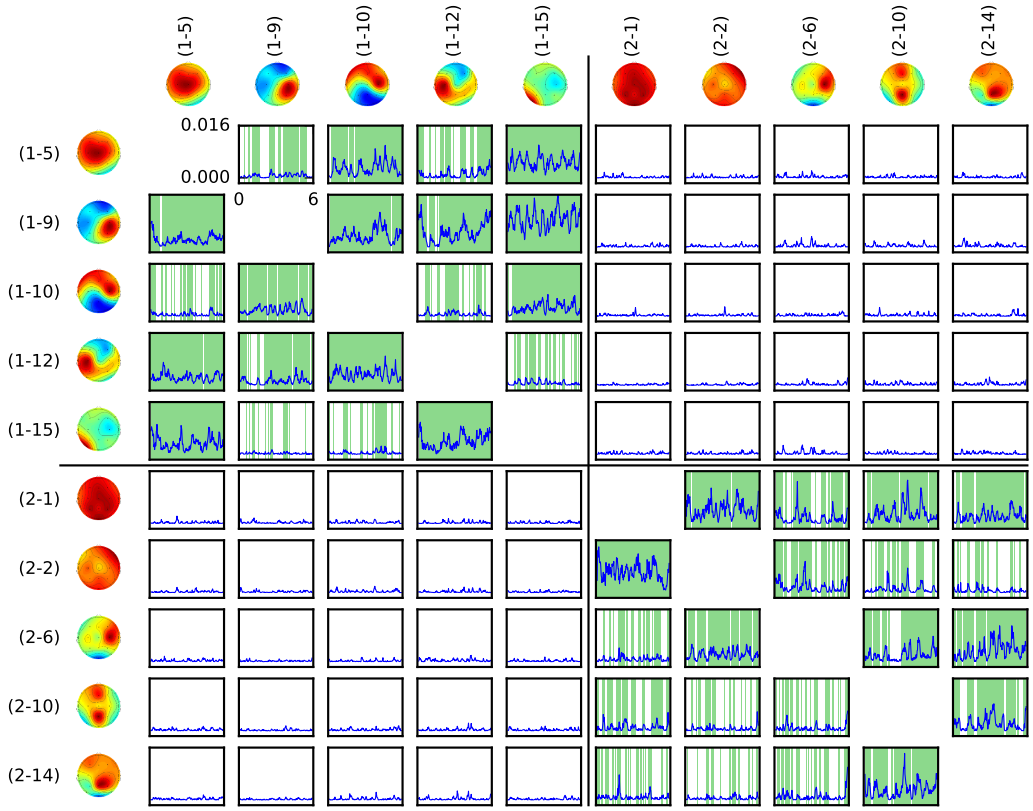


Figure H.12.: RPDC analysis of experiment three in the γ-band. The values are very small again, but in a similar range for both participants.

H.4. Experiment Four

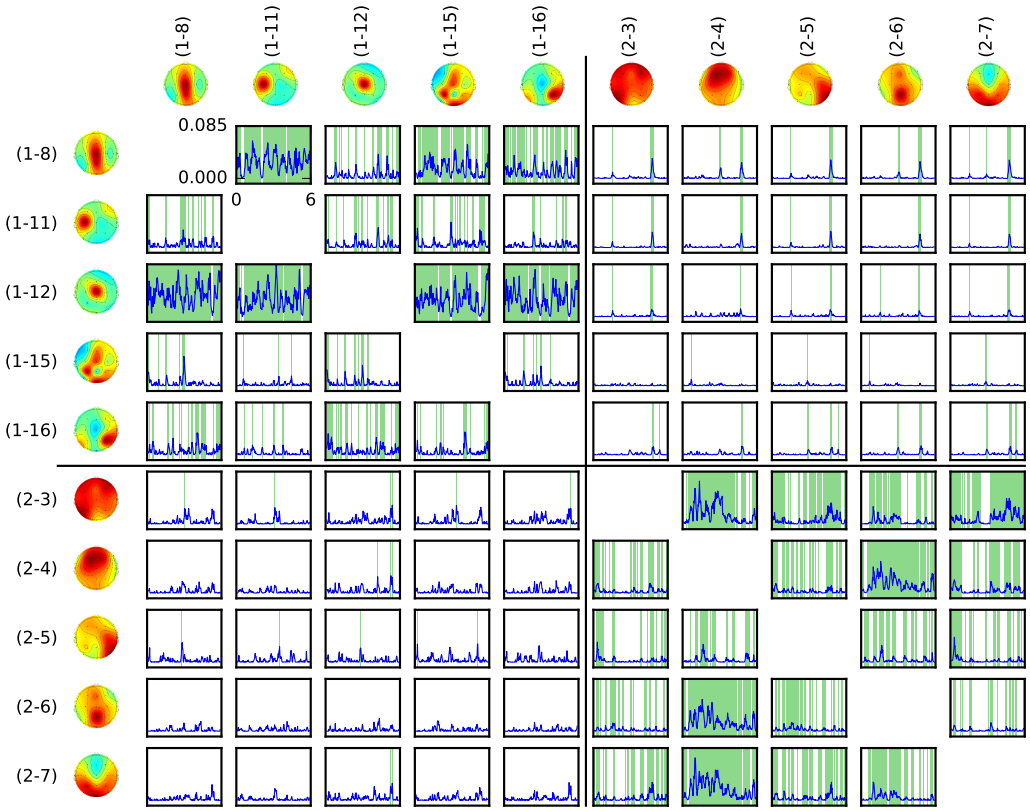


Figure H.13.: RPDC analysis of experiment four in the θ-band. In contrast to other results, PDC values are significant for the entire six seconds for none of the connections. (1-12) is remarkable as it acts as a strong sender for all other components of participant one. For many of the hyper-connections from participant one to participant two a significant peak in PDC values can be spotted towards the end of the of the time frame. If we assume that there is actually another component, not included in the analysis which acts as a common driver for all of these components (either directly or indirectly), including this component would result in this peak vanishing in the plots of the other components, due to the multi-variant nature of PDC.

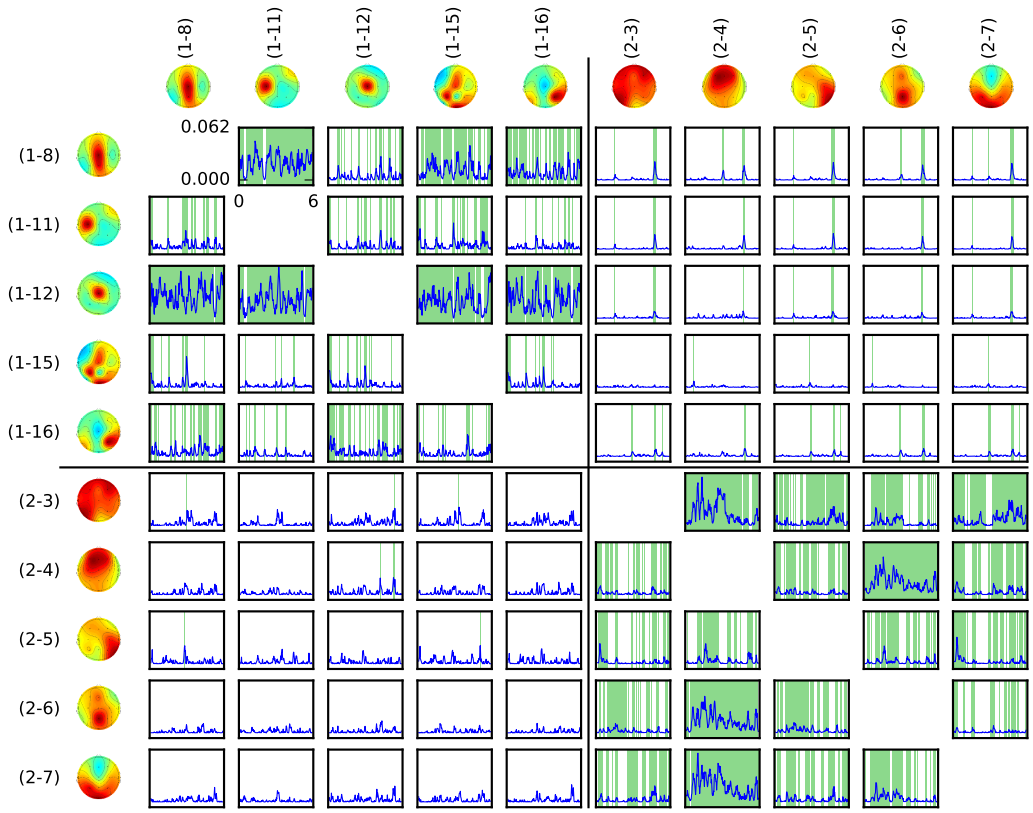


Figure H.14.: RPDC analysis of experiment four in the α-band. The same consistently occurring peak in the hyper-connections from participant one to participant two can be observed.

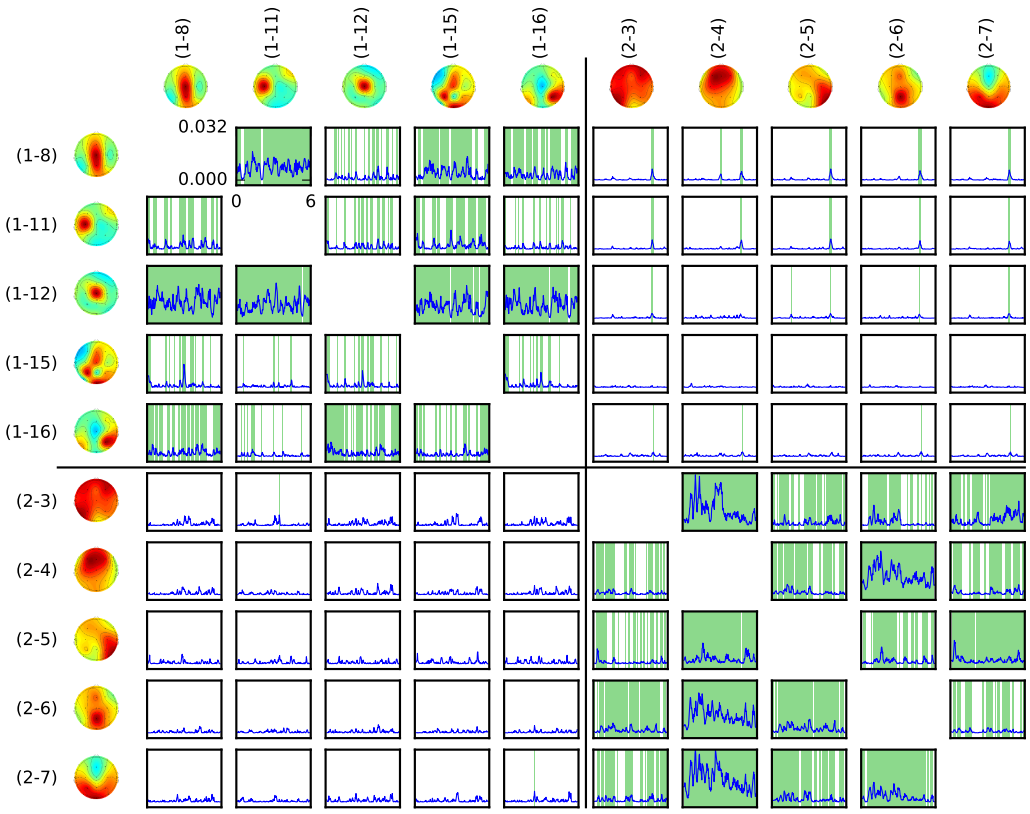


Figure H.15.: RPDC analysis of experiment four in the β-band. Again the same peak can be seen, but for fewer of the connections. In this band it seems to be present mostly in connection with components (1-8) and (1-11) as a sender.

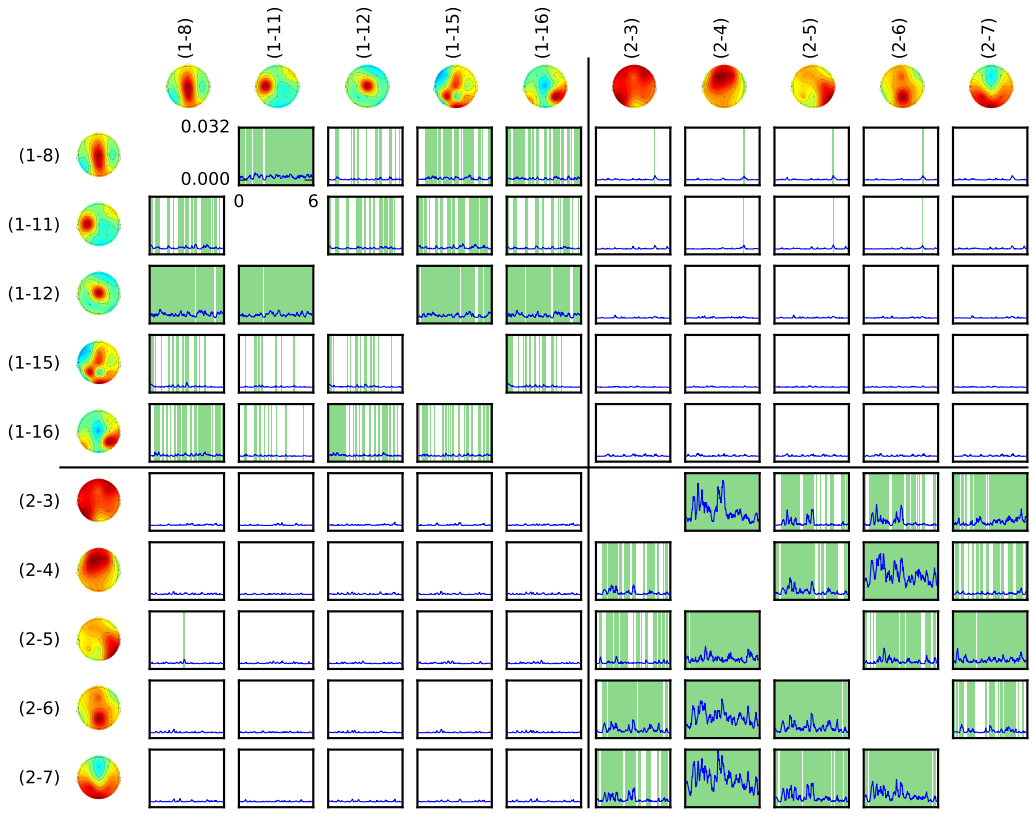


Figure H.16.: RPDC analysis of experiment four in the γ-band. We still have some short significant periods among hyper-connections. However, the consistently occurring peak has vanished.

H.5. Experiment Five

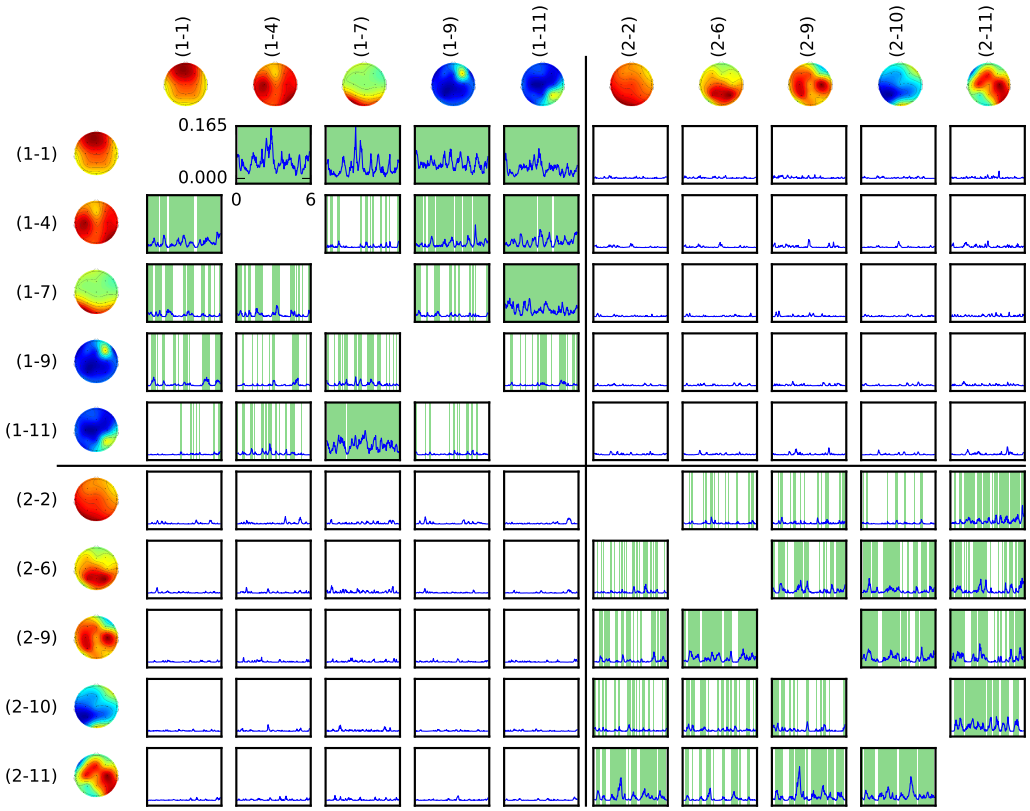


Figure H.17.: RPDC analysis of experiment five in the θ-band. The lack significant connectivity for some within-participant connections is remarkable, compared to the other results.

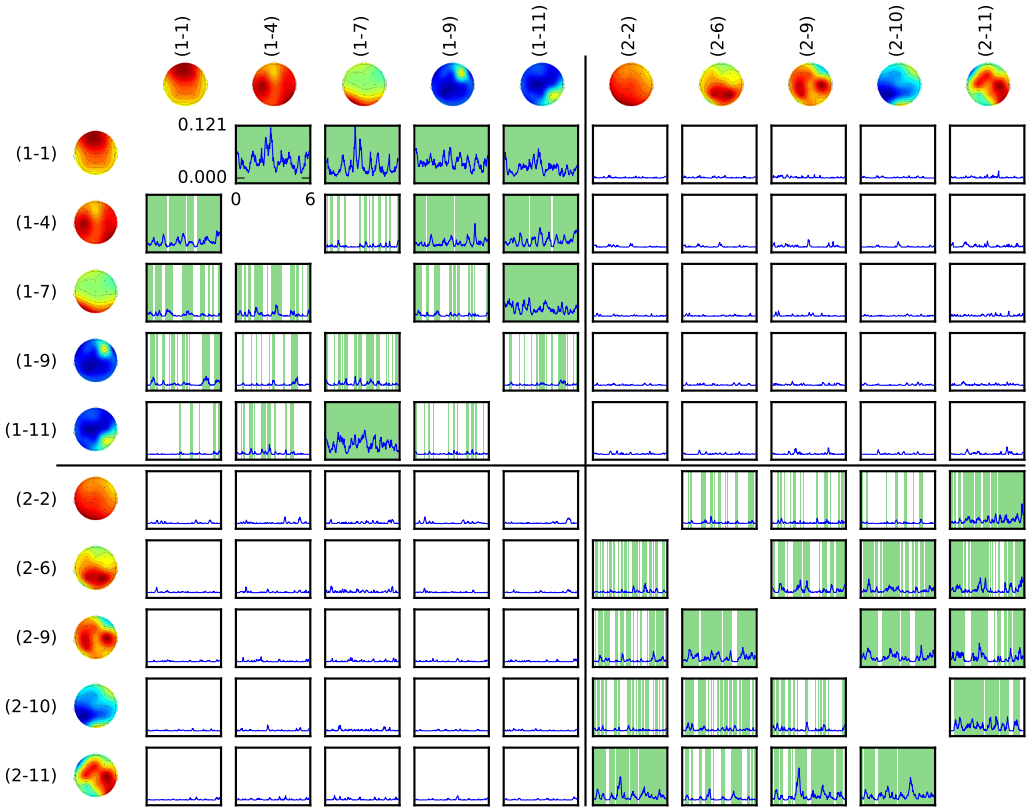


Figure H.18.: RPDC analysis of experiment five in the α-band. Component (1-1) has a strong influence on the other within-participant components. The other way round, (1-11) is being influenced by most other within-participant components.

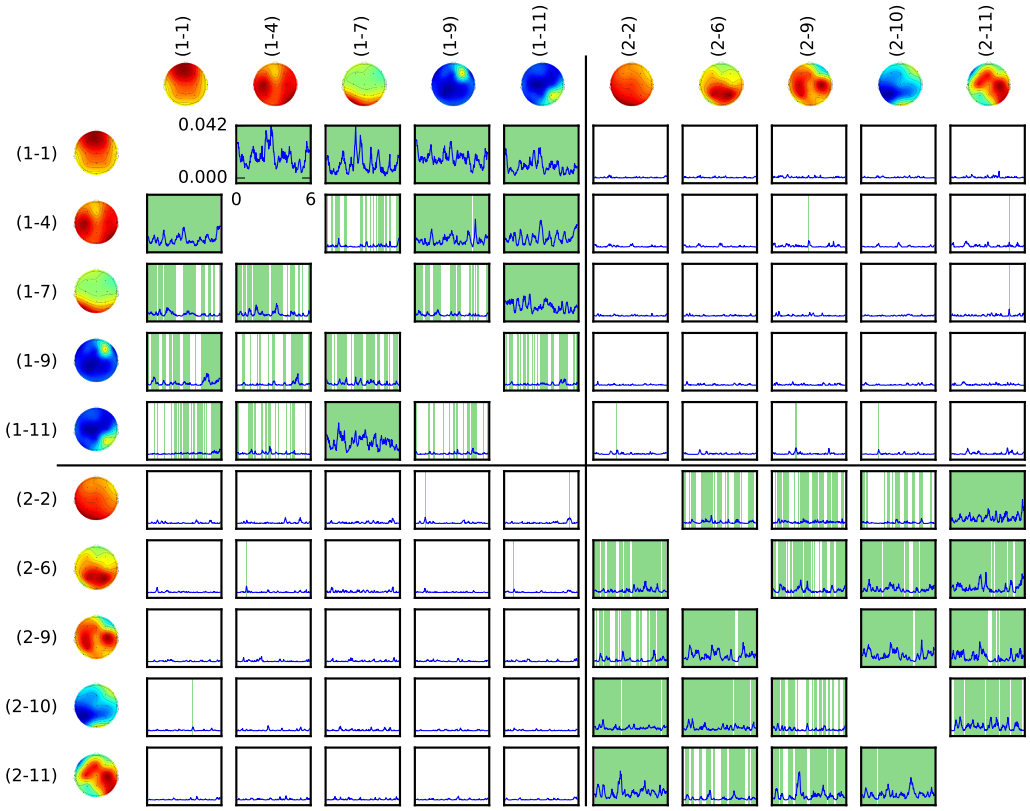


Figure H.19.: RPDC analysis of experiment five in the β-band. The same observations as for α-band can be made here. Additionally we have few, very short significant peaks among hyper-connections.

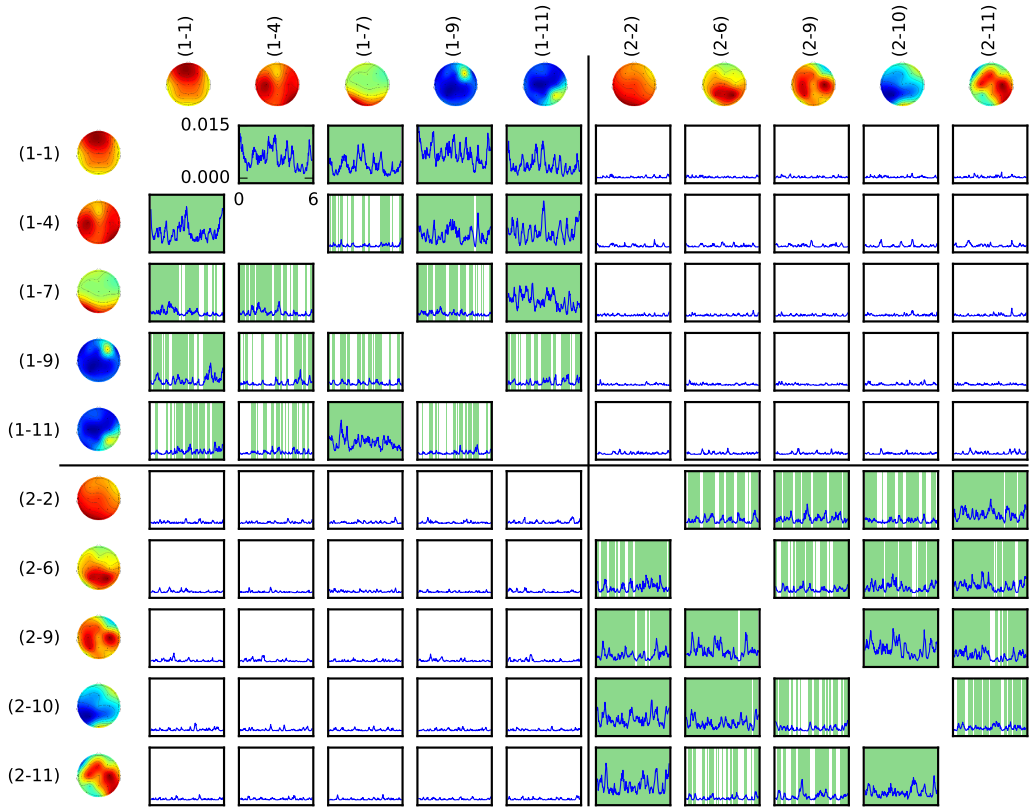


Figure H.20.: RPDC analysis of experiment five in the γ-band. Values are particularly low. Apart from that the same observations about the components (1-1) and (1-11) as in α- and β-band can be made.

H.6. Experiment Six

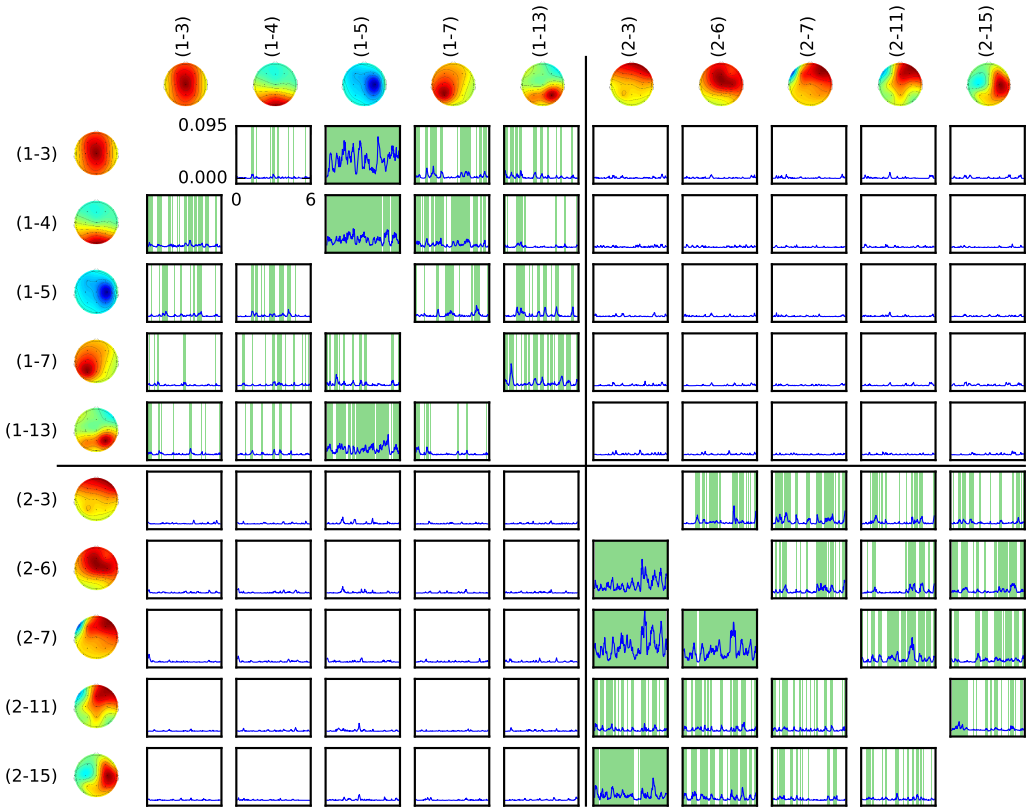


Figure H.21.: RPDC analysis of experiment six in the θ-band. PDC values are remarkably sparsely significant for both participants.

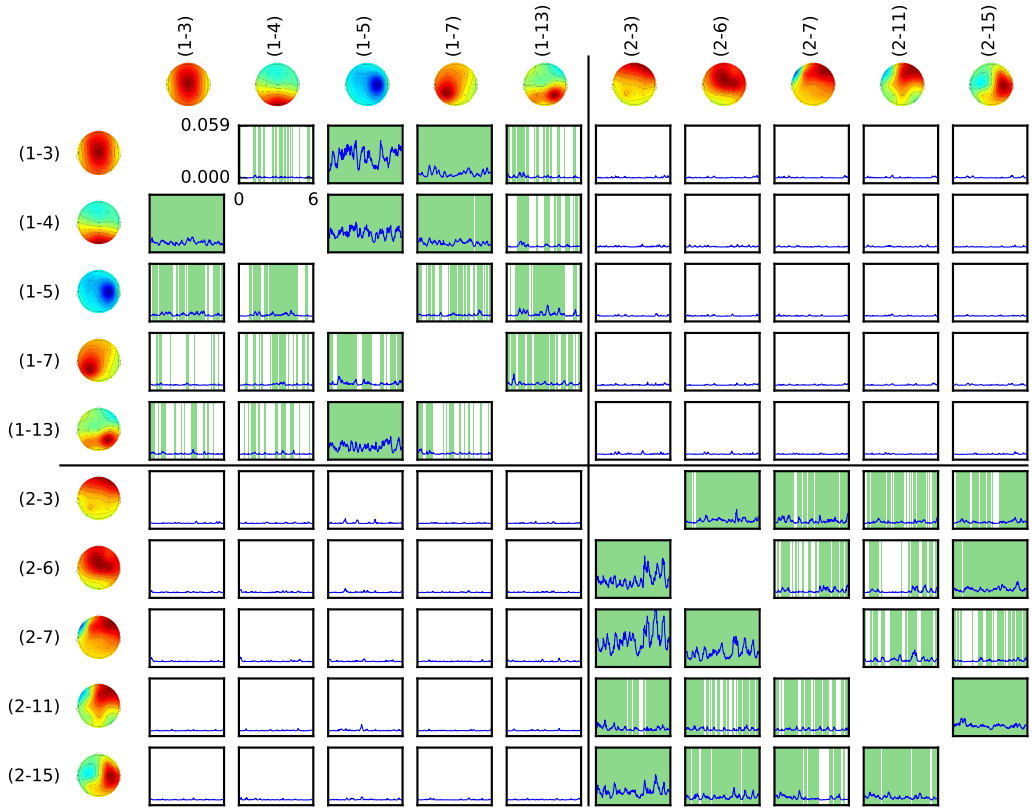


Figure H.22.: RPDC analysis of experiment six in the α-band. Wider periods are significant as compared to θ-band, but still its remarkably sparse.

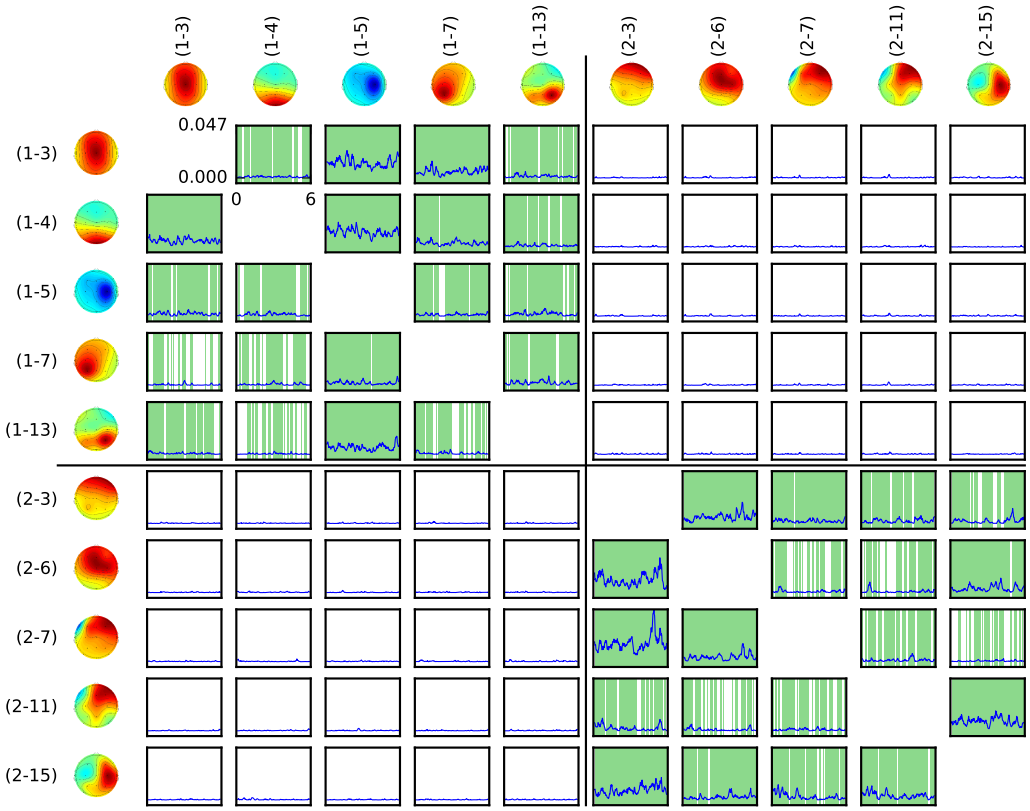


Figure H.23.: RPDC analysis of experiment six in the β-band. Although the value range is half of that of the θ-band, much more significant PDC values were found.

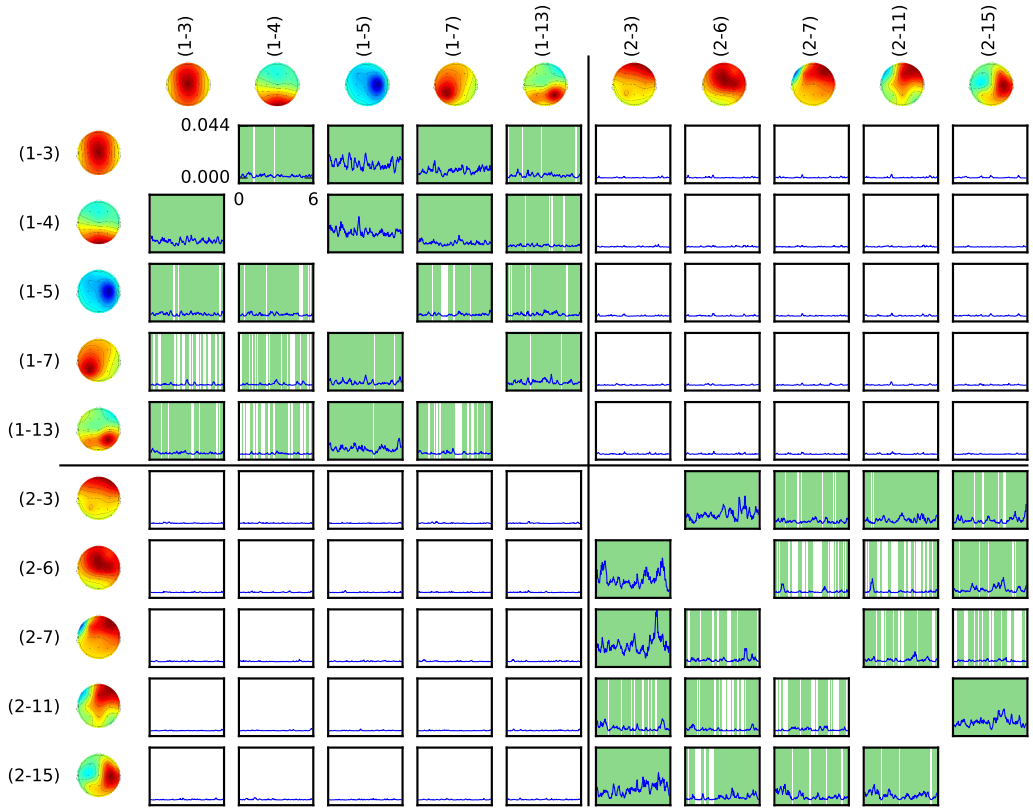


Figure H.24.: RPDC analysis of experiment six in the γ-band. The general impression is pretty similar to the β-band. Also the (low) value range is the same.

H.7. Experiment Seven

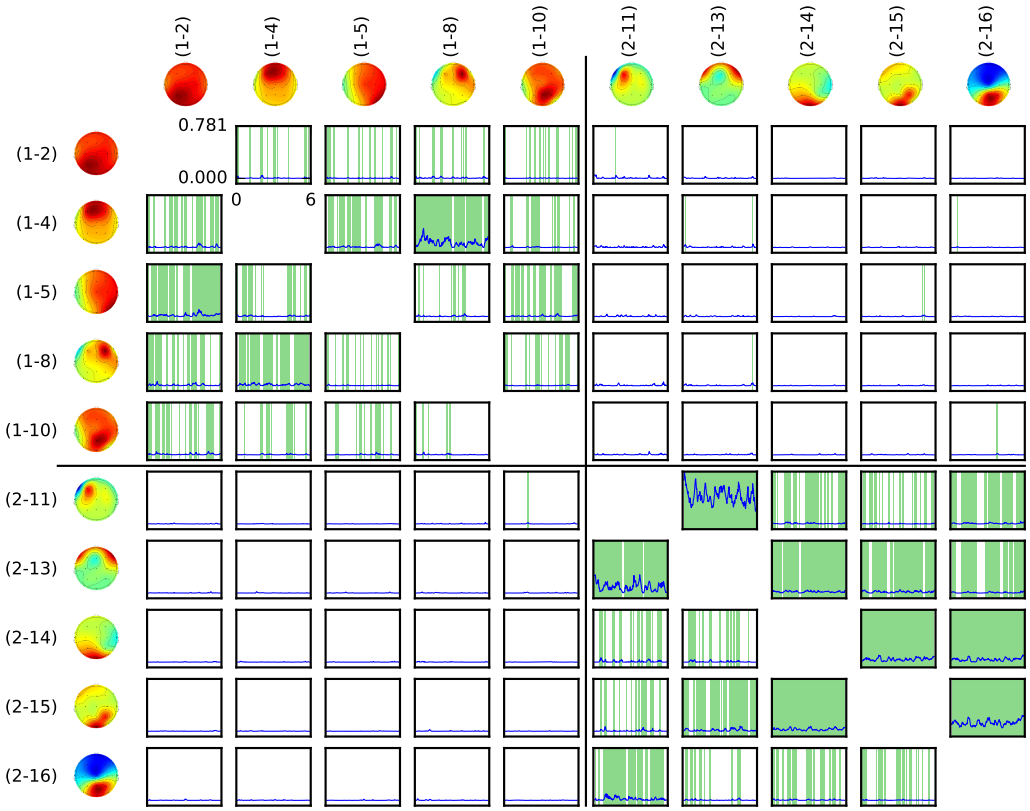


Figure H.25.: RPDC analysis of experiment seven in the θ-band. PDC values are pretty large, but only for few connection. Relatively sparse significant periods for within-participant one connections were found. Some significant hyper-connectivity was found.

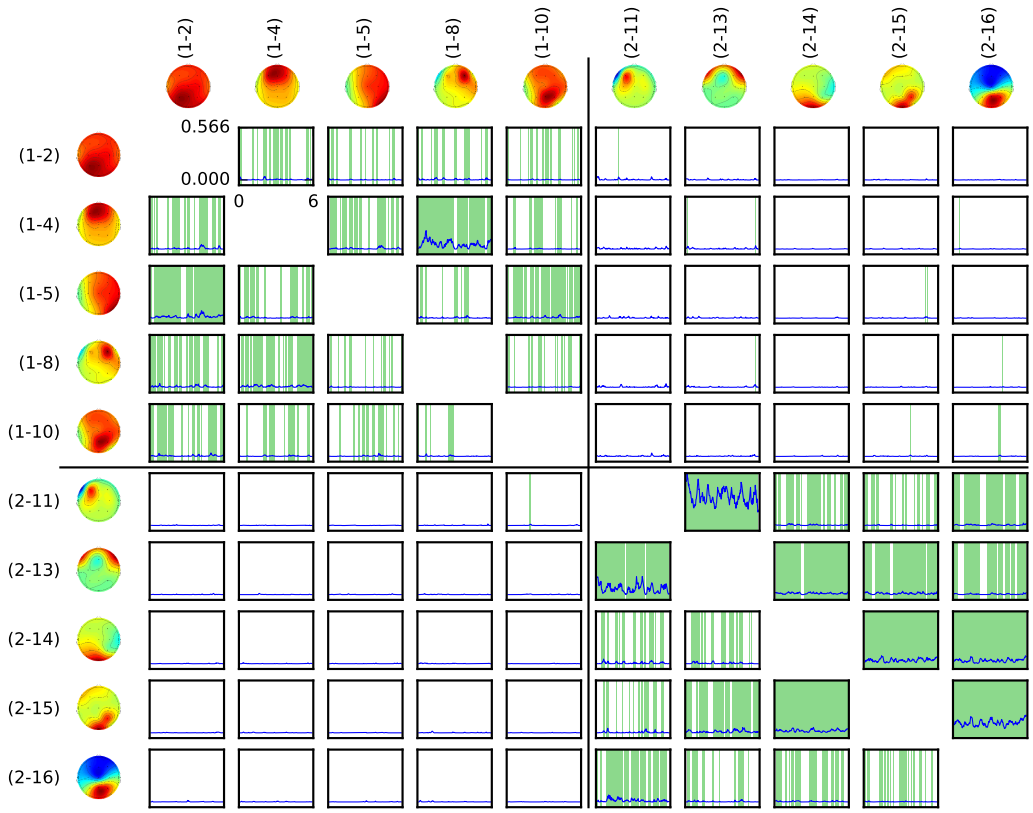


Figure H.26.: RPDC analysis of experiment seven in the α-band. The values are a little smaller compared to the results from the θ-band. The same components share the highest values.

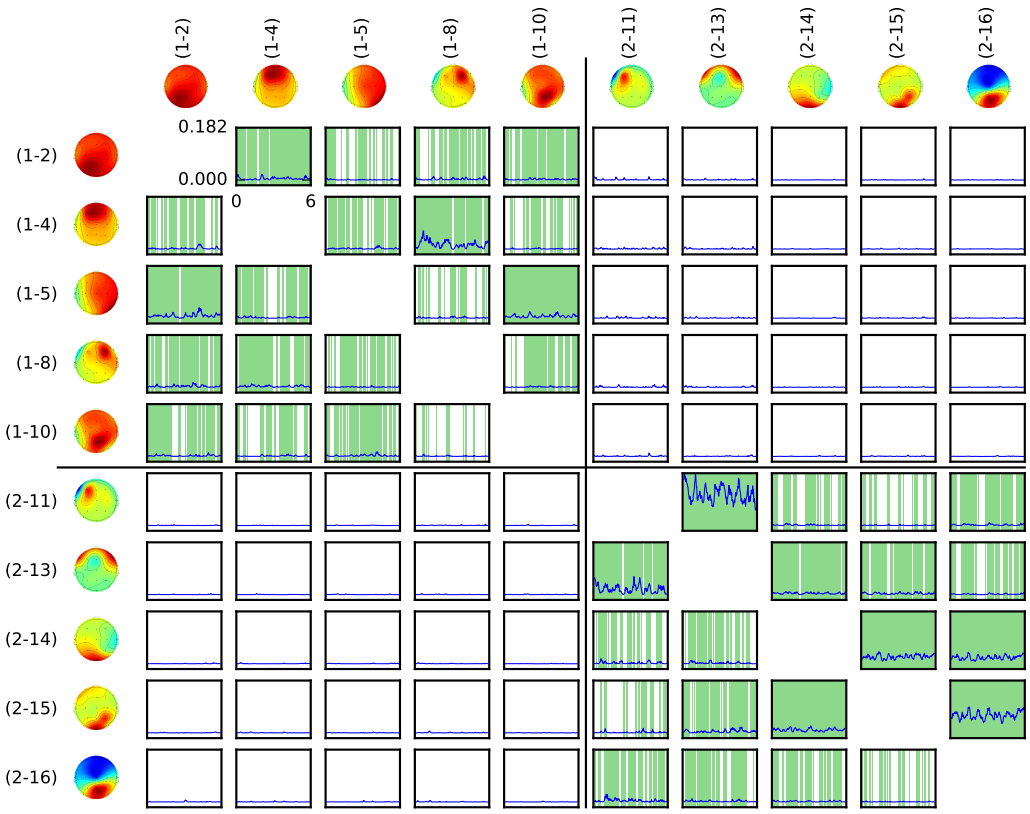


Figure H.27.: RPDC analysis of experiment seven in the β-band. The values range is no longer remarkable (other than for θ- and α-band). However, the same connections, which had particularly high PDC values in these bands, stand out here again.

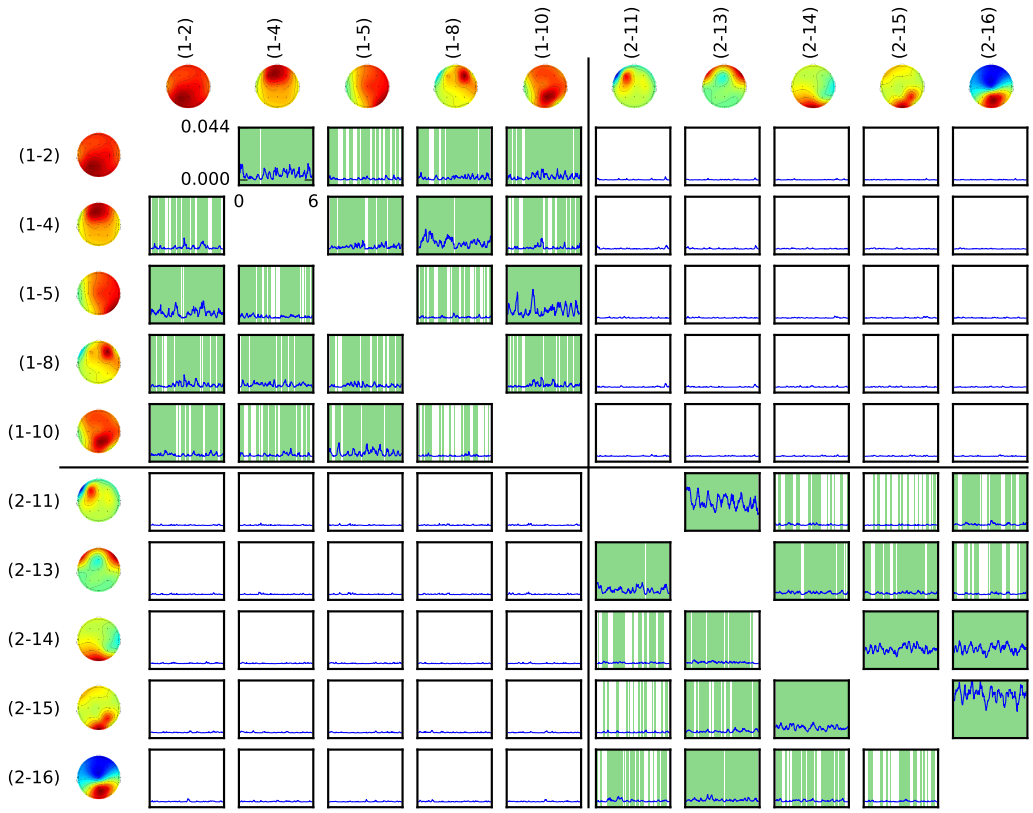


Figure H.28.: RPDC analysis of experiment seven in the γ-band. The values are now relatively low. And again the same components are active.

H.8. Experiment Eight

This is the only experiment for which I was forced to reduce the number of components to eight (four + four) to obtain a stable MVar model.

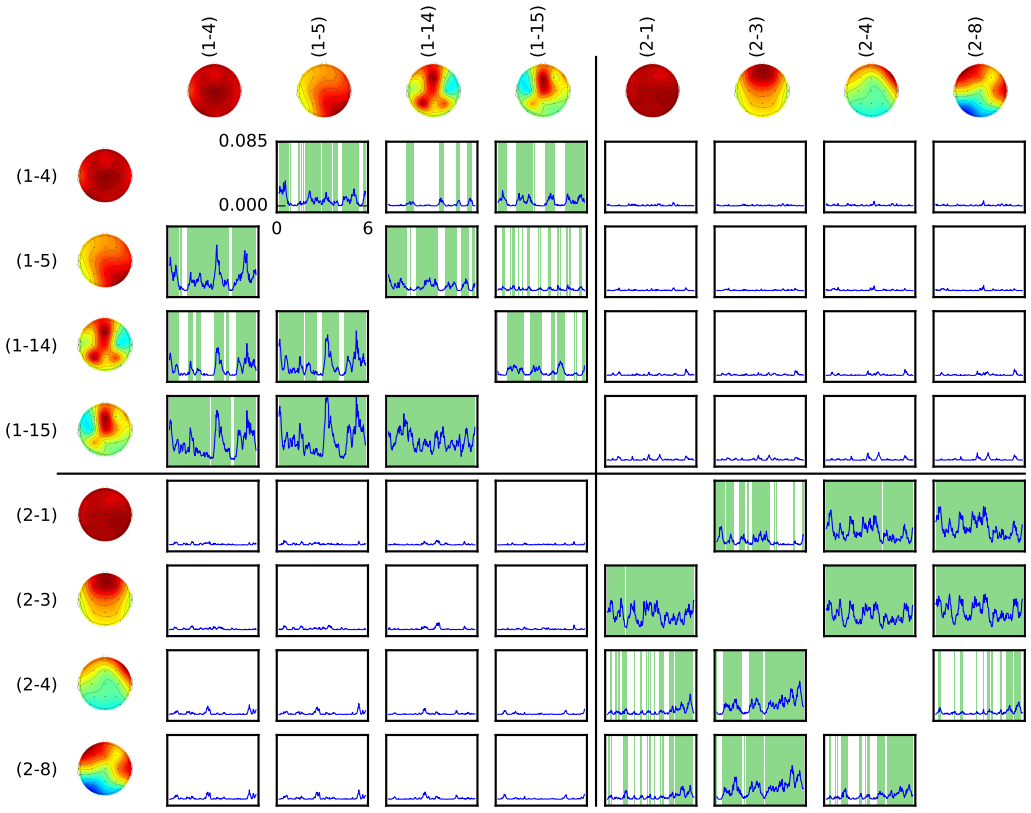


Figure H.29.: RPDC analysis of experiment eight in the θ-band. The connections from components (2-1) and (2-3) to (2-4) and (2-8) show a remarkably similar course over the six seconds. Some similarities can be found in those within participant connections with (1-4) and (1-5) as a receiver.

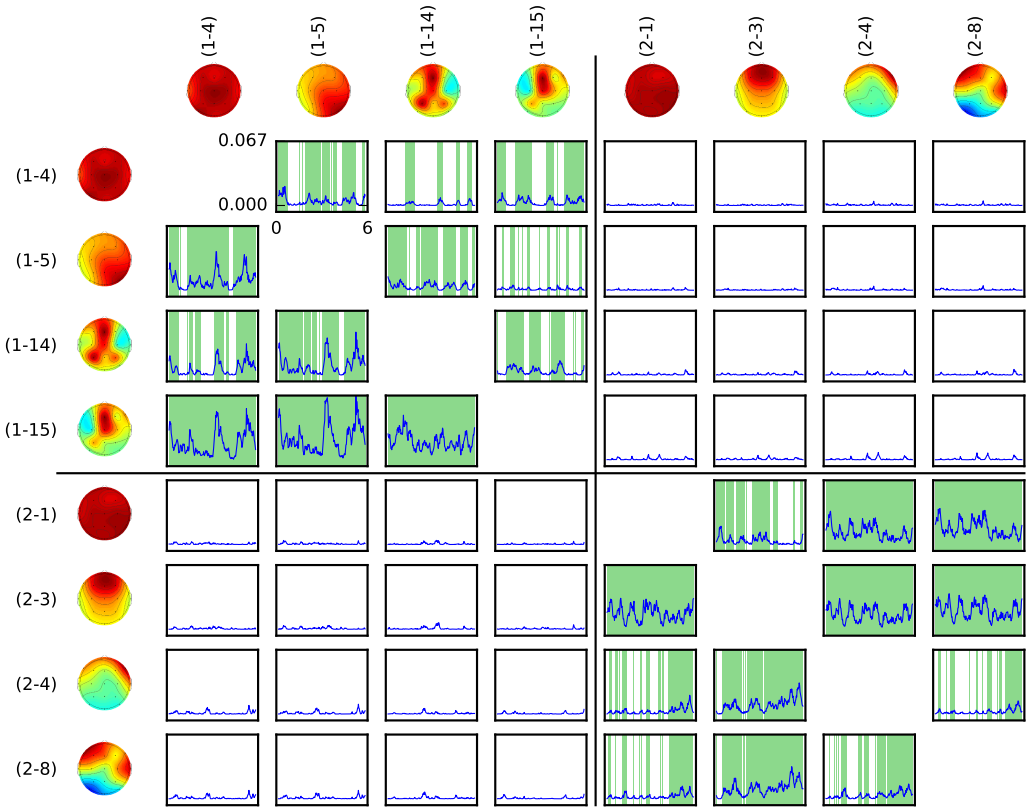


Figure H.30.: RPDC analysis of experiment eight in the α-band. The similarities in the time course of the connections named in θ-band analysis persist.

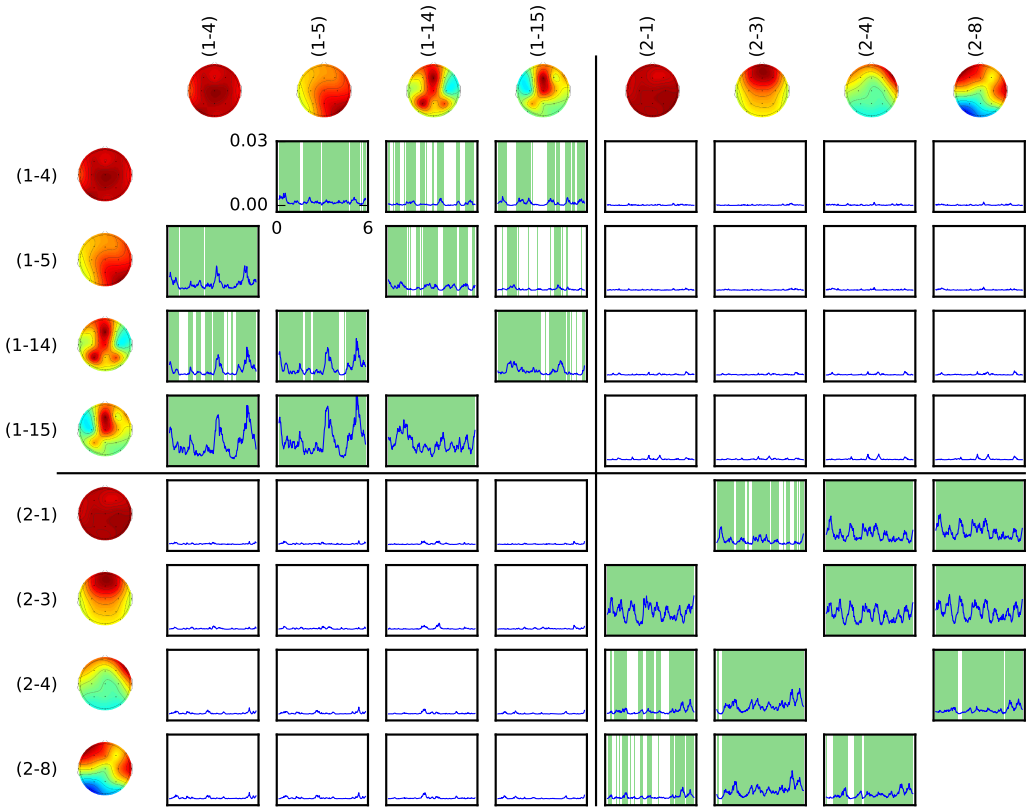


Figure H.31.: RPDC analysis of experiment eight in the β-band. Again PDC time courses are remarkably similar for many connections.

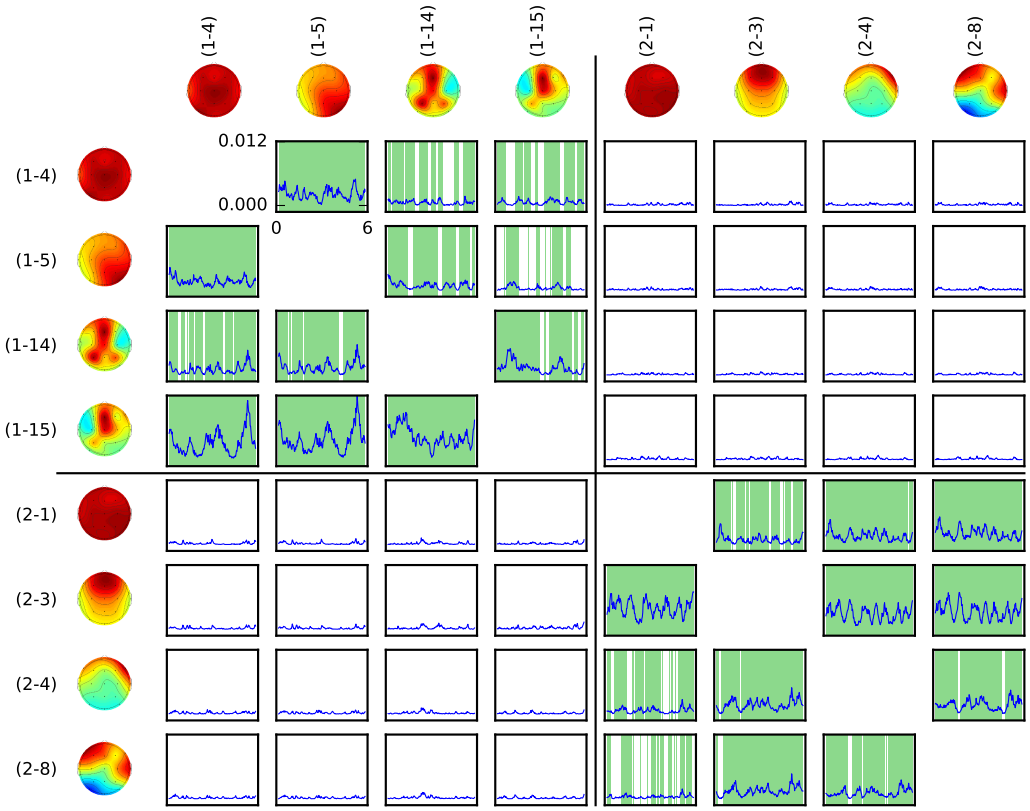


Figure H.32.: RPDC analysis of experiment eight in the γ-band. And again the time courses of different connections are remarkably similar.

H.9. Experiment Nine

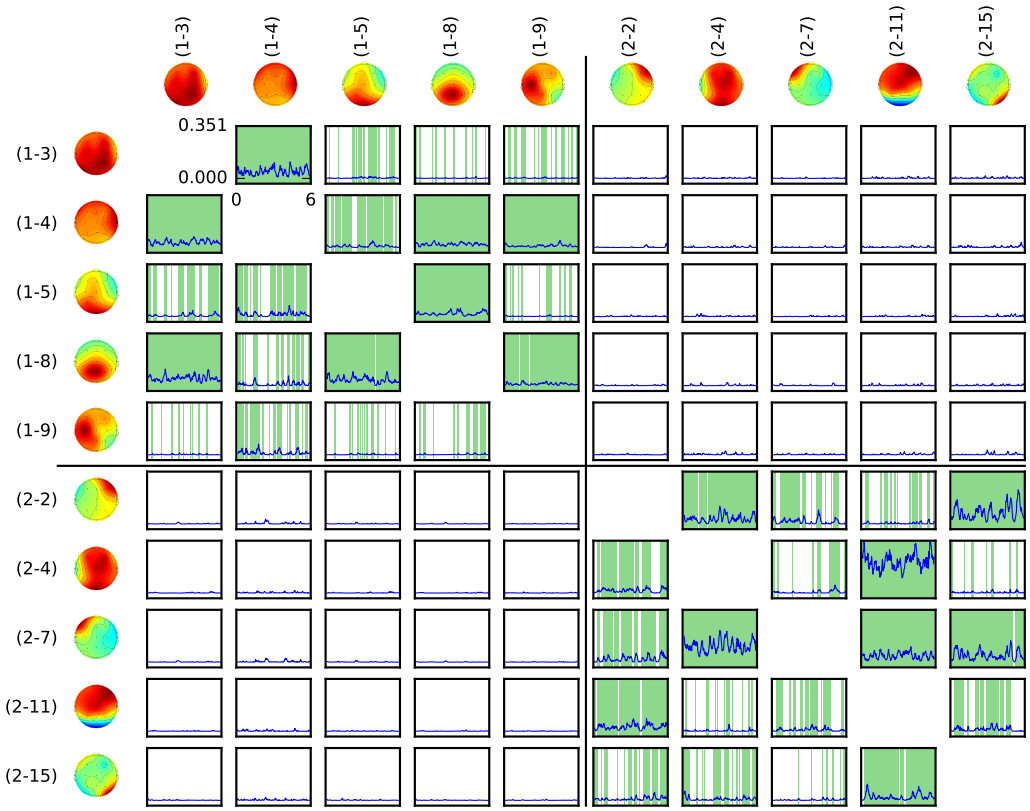


Figure H.33.: RPDC analysis of experiment nine in the θ-band. For some within-participant connection remarkably sparse periods are deemed significant, compared with other analyses.

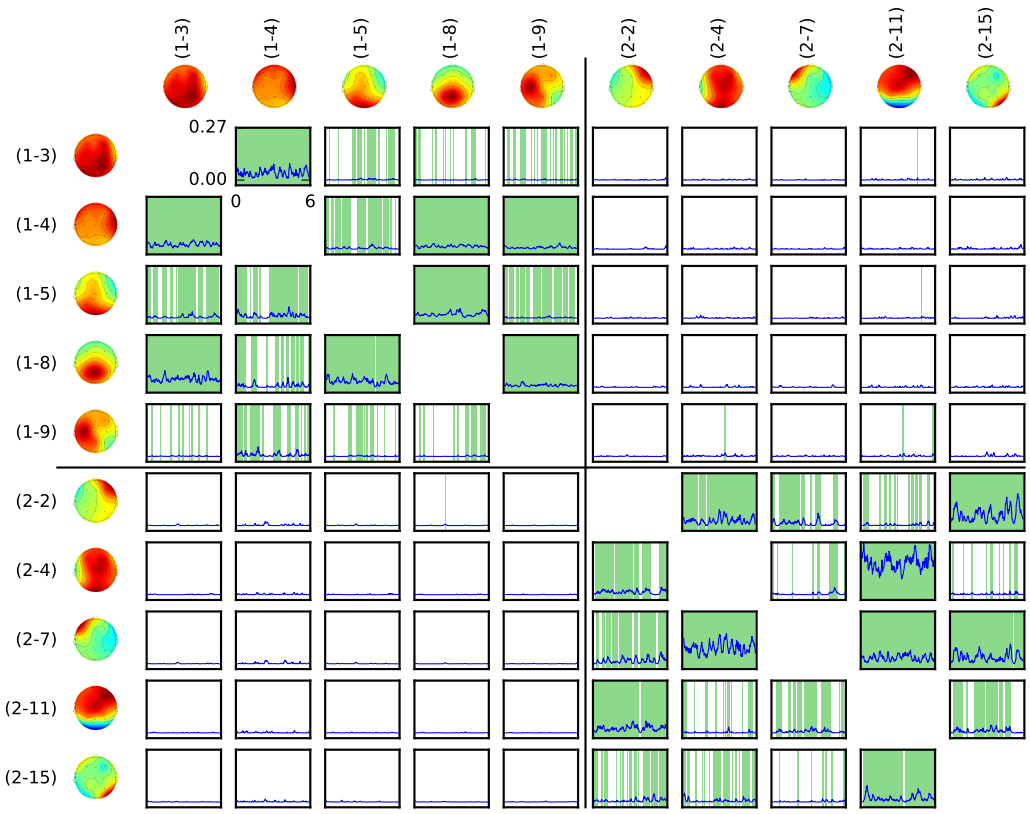


Figure H.34.: RPDC analysis of experiment nine in the α-band. Again, some within-participant connections for both participants are deemed significant only sparsely. Additionally the connection (1-8) $\rightarrow$ (1-9) was deemed significant in the corresponding differential PSI analysis. Here its PDC values are not particularly high, but it is significant for the entire six seconds.

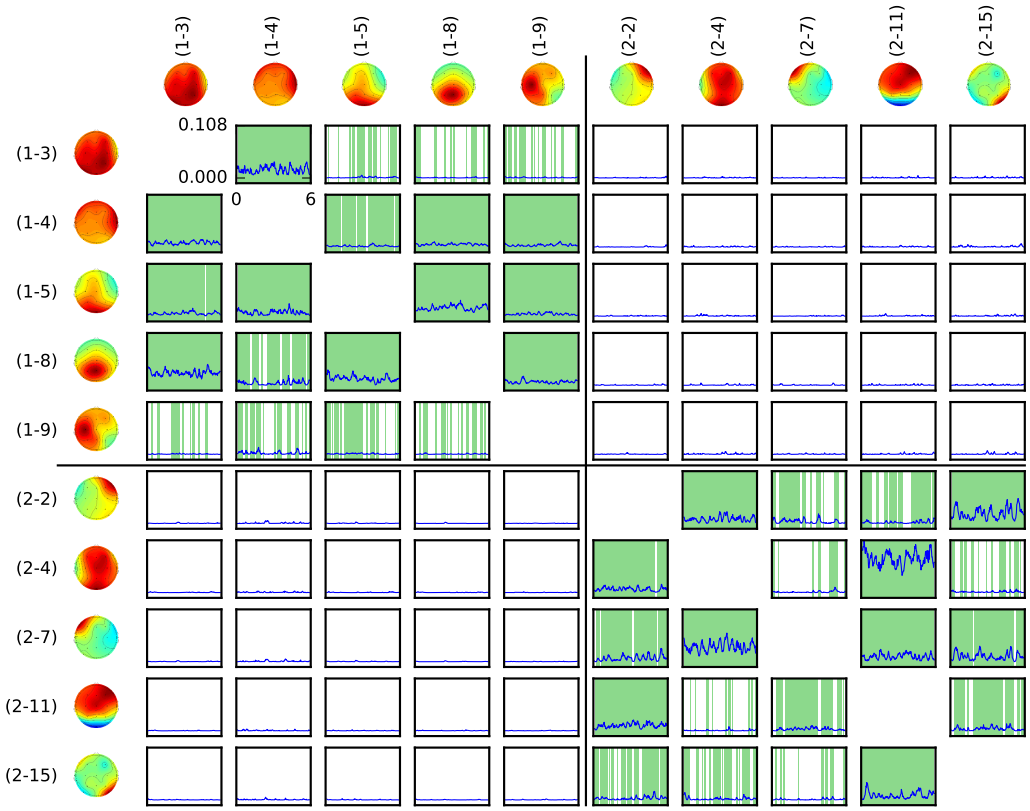


Figure H.35.: RPDC analysis of experiment nine in the β-band. Especially component (1-8) stand out with very sparsely significant connectivity.

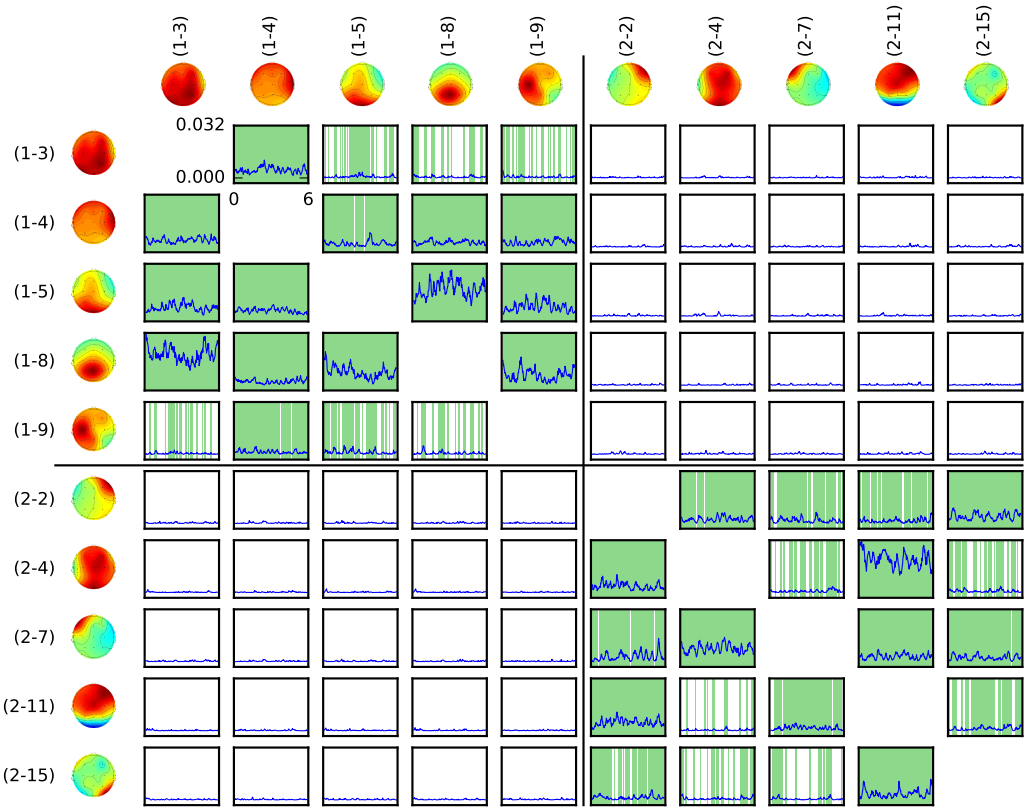


Figure H.36.: RPDC analysis of experiment nine in the γ-band. The impression is pretty similar to β-band.

Bibliography

Astolfi, L., Babiloni, F., Babiloni, C., Carducci, F., Cincotti, F., Basilisco, A., Rossini, P. M., Salinari, S., Ding, L., Ni, Y., and He, B. (2004). Time-varying cortical connectivity by high resolution eeg and directed transfer function: Simulations and application to finger tapping data. In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4405–4408. IEEE.

Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Salinari, S., Marciani, M. G., Wilke, C., Doud, A., Yuan, H., He, B., and Babiloni, F. (2009). Estimation of the cortical activity from simultaneous multi-subject recordings during the prisoner's dilemma. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1937–1939. IEEE.

Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Salinari, S., Vecchiato, G., Toppi, J., Wilke, C., Doud, A., Yuan, H., He, B., and Babiloni, F. (2010a). Simultaneous estimation of cortical activity during social interactions by using EEG hyperscannings. In *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 2814–2817. IEEE.

Astolfi, L., Cincotti, F., Mattia, D., Marciani, M. G., Baccalà, L. A., De Vico Fallani, F., Salinari, S., Ursino, M., Zavaglia, M., and Babiloni, F. (2006). Assessing cortical functional connectivity by partial directed coherence: Simulations and application to real data. *IEEE Transactions on Biomedical Engineering*, 53(9):1802–1812.

Astolfi, L., Toppi, J., Borghini, G., Vecchiato, G., He, E. J., Roy, A., Cincotti, F., Salinari, S., Mattia, D., He, B., and Babiloni, F. (2012). Cortical activity and functional hyperconnectivity by simultaneous EEG recordings from interacting couples of professional pilots. In *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4752–4755. IEEE.

Astolfi, L., Toppi, J., Borghini, G., Vecchiato, G., Isabella, R., De Vico Fallani, F., Cincotti, F., Salinari, S., Mattia, D., He, B., Caltagirone, C., and Babiloni, F. (2011a). Study of the functional hyperconnectivity between couples of pilots during flight simulation: An EEG hyperscanning study. In *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2338–2341. IEEE.

Astolfi, L., Toppi, J., De Vico Fallani, F., Vecchiato, G., Cincotti, F., Wilke, C., Yuan, H., Mattia, D., Salinari, S., He, B., and Babiloni, F. (2011b). Imaging the social brain by simultaneous hyperscanning during subject interaction. *IEEE Intelligent Systems*, 26(5):38–45.

Astolfi, L., Toppi, J., De Vico Fallani, F., Vecchiato, G., Salinari, S., Mattia, D., Cincotti, F., and Babiloni, F. (2010b). Neuroelectrical hyperscanning measures simultaneous brain activity in humans. *Brain topography*, 23(3):243–256.

Astolfi, L., Toppi, J., Vogel, P., Mattia, D., Babiloni, F., Ciaramidaro, A., and Siniatchkin, A. (2014). Investigating the neural basis of cooperative joint action. An EEG hyperscanning study. In *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4896–4899. IEEE.

Babiloni, F., Astolfi, L., Cincotti, F., Mattia, D., Tocci, A., Tarantino, A., Marciani, M., Salinari, S., Gao, S., Colosimo, A., and De Vico Fallani, F. (2007a). Cortical activity and connectivity of human brain during the prisoner's dilemma: An EEG hyperscanning study. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4953–4956. IEEE.

Babiloni, F., Cincotti, F., Mattia, D., De Vico Fallani, F., Tocci, A., Bianchi, L., Salinari, S., Marciani, M., Colosimo, A., and Astolfi, L. (2007b). High resolution EEG hyperscanning during a card game. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4957–4960. IEEE.

Babiloni, F., Cincotti, F., Mattia, D., de Vico Fallani, F., Tocci, A., Bianchi, L., Salinari, S., Marciani, M., Colosimo, A., and Astolfi, L. (2007c). Simultaneous tracking of multiple brains activity with high resolution EEG hyperscannings. In *Proceedings of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging*, pages 196–199. IEEE.

Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159.

Bell, C. J., Shenoy, P., Chalodhorn, R., and Rao, R. P. N. (2008). Control of a humanoid robot by a noninvasive brain-computer interface in humans. *Journal of Neural Engineering*, 5(2):214–220.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B*, 57(1):289–300.

Bennett, C. M., Baird, A. A., Miller, M. B., and Wolford, G. L. (2009). Neural correlates of interspecies perspective taking inthe post-mortem atlantic salmon: An argument forproper multiple comparisons corrections. *Journal of Serendipitous and Unexpected Results*, 1(1):1–5.

Berger, H. (1938). Über das Elektrenkephalogramm des Menschen. *Nova Acta Leopoldina*, 6(38):173–309.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Blankertz, B., Curio, G., and Müller, K.-R. (2001). Classifying single trial EEG: Towards brain computer interfacing. In *Proceedings of the 14th Advances in Neural Information Processing Systems Conference*, pages 157–164. MIT Press.

Bonnet, L., Rennes, I., Lotte, F., Bordeaux, I., and Anatole, L. (2013). Two brains, one game: Design and evaluation of a multi-user BCI video game based on motor imagery. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):185–198.

Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: A cautionary note. *Frontiers in Human Neuroscience*, 7(881):1–17.

Carlson, T. and Del R. Millan, J. (2013). Brain-controlled wheelchairs: A robotic architecture. *IEEE Robotics and Automation Magazine*, 20(1):65–73.

De Vico Fallani, F., Nicosia, V., Sinatra, R., Astolfi, L., Cincotti, F., Mattia, D., Wilke, C., Doud, A., Latora, V., He, B., and Babiloni, F. (2010). Defecting or not defecting: How to "read" human behavior during cooperative games by EEG measurements. *PLOS One*, 5(12):e14187.

Debener, S., Emkes, R., De Vos, M., and Bleichner, M. (2015). Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Scientific Reports*, 5(16743):1–11.

Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., Vankov, A., and Makeig, S. (2011). EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing. *Computational Intelligence and Neuroscience*, 2011(130714):1–12.

Ding, M., Bressler, S. L., Yang, W., and Liang, H. (2000). Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: Data preprocessing, model validation, and variability assessment. *Biological Cybernetics*, 83(1):35–45.

Dodel, S., Cohn, J., Mersmann, J., Luu, P., Forsythe, C., and Jirsa, V. (2011). Brain signatures of team performance. In *In Processdings of the 6th International Conference on Foundations of Augmented Cognition. Directing the Future of Adaptive Systems.*, pages 288–297. Springer.

Dumas, G. (2011). Towards a two-body neuroscience. *Communicative & Integrative Biology*, 4(3):349–352.

Dumas, G., Chavez, M., Nadel, J., and Martinerie, J. (2012a). Anatomical connectivity influences both intra- and inter-brain synchronizations. *PLOS One*, 7(5):e36414.

Dumas, G., Martinerie, J., Soussignan, R., and Nadel, J. (2012b). Does the brain know who is at the origin of what in an imitative interaction? *Frontiers in Human Neuroscience*, 6(128):1–11.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLOS One*, 5(8):e12166.

Egetemeir, J., Stenneken, P., Koehler, S., Fallgatter, A. J., and Herrmann, M. J. (2011). Exploring the neural basis of real-life joint action: Measuring brain activation during joint table setting with functional near-infrared spectroscopy. *Frontiers in Human Neuroscience*, 5(95):1–9.

Elbrechter, C., Haschke, R., and Ritter, H. (2011). Bi-manual robotic paper manipulation based on real-time marker tracking and physical modelling. In *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1427–1432. IEEE.

Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510–523.

Finke, A., Rudgalwis, B., Jakusch, H., and Ritter, H. (2012). Towards multi-user brain-robot interfaces for humanoid robot control. In *12th IEEE-RAS International Conference on Humanoid Robots*, pages 532–537. IEEE.

Friston, K., Frith, C., and Frackowiak, R. (1993a). Time-dependent changes in effective connectivity measured with PET. *Human Brain Mapping*, 1(1):69 – 79.

Friston, K., Worsley, K., Frackowiak, R., Mazziotta, J., and Evans, A. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1(3):210–220.

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. (1993b). Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow & Metabolism*, 13(1):5–14.

Fritsch, J., Kleinehagenbrock, M., Haasch, A., Wrede, S., and Sagerer, G. (2005). A flexible infrastructure for the development of a robot companion with extensible HRI-capabilities. In *Proceedings of the 22nd IEEE International Conference on Robotics and Automation*, pages 3408–3414. IEEE.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Hachmeister, N., Riechmann, H., Ritter, H., and Finke, A. (2011). An approach towards human-robot-human interaction using a hybrid brain-computer interface. In *Proceedings of the 13th International Conference on Multimodal Interfaces*, page 49–52. ACM.

Haufe, S., Nikulin, V. V., Müller, K.-R., and Nolte, G. (2012). A critical assessment of connectivity measures for EEG data: A simulation study. *NeuroImage*, 64(2013):120–133.

Jasper, H. H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology*, 10(2):370–375.

Kaminski, M. J. and Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. *Biological Cybernetics*, 65(3):203–210.

Kaper, M. and Ritter, H. (2004). Generalizing to new subjects in brain-computer interfacing. In *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4363–4366. IEEE.

Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E., and Yamaguchi, Y. (2013). Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Scientific Reports*, 3(1):1–8.

Kuś, R., Kamiński, M., and Blinowska, K. J. (2004). Determination of EEG activity propagation: Pair-wise versus multichannel estimate. *IEEE Transactions on Biomedical Engineering*, 51(9):1501–1510.

Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, 8(4):194–208.

Le Van Quyen, M., Foucher, J., Lachaux, J., Rodriguez, E., Lutz, A., Martinerie, J., and Varela, F. J. (2001). Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *Journal of Neuroscience Methods*, 111(2):83–98.

Lee, L., Harrison, L. M., and Mechelli, A. (2003). The functional brain connectivity workshop: Report and commentary. *Network: Computation in Neural Systems*, 14(2):1–15.

Lenhardt, A., Kaper, M., and Ritter, H. J. (2008). An adaptive P300-based online brain–computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(2):121–130.

Lindenberger, U., Li, S.-C., Gruber, W., and Müller, V. (2009). Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience*, 10(1):1–12.

Lotte, F., van Langhenhove, A., Lamarche, F., Ernest, T., Renard, Y., Arnaldi, B., and Lécuyer, A. (2010). Exploring large virtual environments by thoughts using a brain–computer interface based on motor imagery and high-level commands. *Presence: Teleoperators and Virtual Environments*, 19(1):54–70.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Lyon, L. (2017). Dead salmon and voodoo correlations: Should we be sceptical about functional mri? *Brain*, 140(8):e53–e53.

Maycock, J., Dornbusch, D., Elbrechter, C., Haschke, R., Schack, T., and Ritter, H. (2010). Approaching manual intelligence. *KI-Künstliche Intelligenz*, 24(4):287–294.

Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M. (2011a). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48(2):229–240.

Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M. (2011b). ADJUST tutorial an automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48(2):229–240.

Mullen, T. (2010). *Source information flow toolbox (SIFT) theoretical handbook and user manual.* Swartz Center for Computational Neuroscience, California, San Diego.

Münßinger, J. I., Halder, S., Kleih, S. C., Furdea, A., Raco, V., Hösle, A., and Kübler, A. (2010). Brain painting: First evaluation of a new brain-computer interface application with ALS-patients and healthy volunteers. *Frontiers in Neuroscience*, 4(182):1–11.

Naeem, M., Prasad, G., Watson, D. R., and Kelso, J. A. S. (2012). Electrophysiological signatures of intentional social coordination in the 10-12 Hz range. *NeuroImage*, 59(2):1795–1803.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology*, 115(10):2292–2307.

Nolte, G., Ziehe, A., Nikulin, V. V., Schlögl, A., Krämer, N., Brismar, T., and Müller, K. R. (2008). Robustly estimating the flow direction of information in complex physical systems. *Physical Review Letters*, 100(23):1–4.

Pfurtscheller, G. and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, 110(11):1842–1857.

Plomp, G., Quairiaux, C., Michel, C. M., and Astolfi, L. (2014). The physiological plausibility of time-varying Granger-causal modeling: Normalization and weighting by spectral power. *NeuroImage*, 97:206–216.

Quigley, M., Conley, K., Gerkey, B., FAust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., and Mg, A. (2009). ROS: An open-source robot operating system. In *Proceedings of the ICRA Workshop on Open Source Software*, pages 1–6. IEEE.

Ramoser, H., Müller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8(4):441–446.

Riechmann, H. (2014). *Exploiting code-modulating, visually-evoked potentials for fast and flexible control via brain-computer interfaces.* PhD thesis, Bielefeld University.

Riedenklau, E., Hermann, T., and Ritter, H. (2011). Saving and restoring mechanisms for tangible user interfaces through tangible active objects. In *Proceedings of the International Conference on Human-Computer Interaction*, pages 110–118. Springer.

Riedenklau, E., Hermann, T., and Ritter, H. (2012a). An integrated multi-modal actuated tangible user interface for distributed collaborative planning. In *Proceedings of the 6th International Conference on Tangible, Embedded and Embodied Interaction*, page 169–174. ACM.

Riedenklau, E., Petker, D., Hermann, T., and Ritter, H. (2012b). Embodied social betworking with gesture-enabled tangible active objects. In *Proceedings of 6th International Symposium on Autonomous Minirobots for Research and Edutainment (AMIRE)*, pages 235–248. Springer.

Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069.

Saito, D. N., Tanabe, H. C., Izuma, K., Hayashi, M. J., Morito, Y., Komeda, H., Uchiyama, H., Kosaka, H., Okazawa, H., Fujibayashi, Y., and Sadato, N. (2010). "Stay tuned": Inter-individual neural synchronization during mutual gaze and joint attention. *Frontiers in Integrative Neuroscience*, 4(127):1–12.

Sanei, S. (2013). *Adaptive Processing of Brain Signals*. Wiley.

Sänger, J., Müller, V., and Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Frontiers in Human Neuroscience*, 6(312):1–19.

Sänger, J., Müller, V., and Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Frontiers in Human Neuroscience*, 7(234):1–14.

Saygin, A. P. and Ishiguro, H. (2010). The perception of humans and robots: Uncanny hills in parietal cortex. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, pages 2004–2008. Cognitive Science Society.

Schelter, B., Timmer, J., and Eichler, M. (2009). Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. *Journal of Neuroscience Methods*, 179(1):121–130.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (2013). Toward a second-person neuroscience. *The Behavioral and Brain Sciences*, 36(4):393–414.

Schwarz, H. R. and Köckler, N. (2006). *Numerische mathematik*. Teubner Verlag.

Schwenkreis, P., El Tom, S., Ragert, P., Pleger, B., Tegenthoff, M., and Dinse, H. R. (2007). Assessment of sensorimotor cortical representation asymmetries and motor skills in violin players. *European Journal of Neuroscience*, 26(11):3291–3302.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2):70–76.

Steffen, J., Elbrechter, C., Haschke, R., and Ritter, H. (2010). Bio-inspired motion strategies for a bimanual manipulation task. In *Proceedings of the 10th IEEE-RAS International Conference on Humanoid Robots*, pages 625–630. IEEE.

Steffen, J., Haschke, R., and Ritter, H. (2007). Experience-based and tactile-driven dynamic grasp control. In *Proceedings of the 20th International Conference on Intelligent Robots and Systems*, pages 2938–2943. IEEE.

Supp, G. G., Schlögl, A., Trujillo-Barreto, N., Müller, M. M., and Gruber, T. (2007). Directed Cortical information flow during human object recognition: Analyzing induced EEG gamma-band responses in brain's source space. *PLOS One*, 2(8):e684.

Sutter, E. E. (1992). The brain response interface: Communication through visually-induced electrical brain responses. *Journal of Microcomputer Applications*, 15(1):31–45.

Tognoli, E., Lagarde, J., DeGuzman, G. C., and Kelso, J. A. S. (2007). The phi complex as a neuromarker of human social coordination. *Proceedings of the National Academy of Sciences of the United States of America*, 104(19):8190–8195.

Tonin, L., Carlson, T., Leeb, R., and Del R. Millan, J. (2011). Brain-controlled telepresence robot by motor-disabled people. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pages 4227–4230. IEEE.

Twardon, L. and Ritter, H. (2015). Interaction skills for a coat-check robot: Identifying and handling the boundary components of clothes. In *In Proceedings of the 32nd International Conference on Robotics and Automation*, pages 3682–3688. IEEE.

Ückermann, A., Elbrechter, C., Haschke, R., and Ritter, H. (2014). Real-time hierarchical scene segmentation and classification. In *Proceedings of the 27th IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1–7. IEEE.

Wienke, J. and Wrede, S. (2011). A middleware for collaborative research in experimental robotics. In *2011 IEEE/SICE International Symposium on System Integration*, pages 1183–1190. IEEE.

Wolpaw, J., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791.

Yuan, H., He, B., Babiloni, F., and Design, A. E. (2010). Imaging the social brain: Multi-subjects EEG recordings during the "Chicken's game". In *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 1734–1737. IEEE.

Yun, K., Watanabe, K., and Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific Reports*, 2(959):1–8.