University of Sussex

On the relation between complex brain activity and consciousness

Michael Manfred Schartner

Submitted for the degree of Doctor of Philosophy
Department of Informatics
University of Sussex, UK
September 2016

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Preface

Parts of this thesis have been previously published. Chapter 3 was presented as a poster at ASSC15 and consequently published as a research article in the journal Plos One [191]. Chapter 4 is currently under review for publication as a research article in the journal Neuroscience of Consciousness [192]. Parts of Chapter 6 have been presented as a poster at ASSC16.

Signature:

Michael M. Schartner

Acknowledgements

This work was crucially fostered by my supervisors Adam Barrett and Anil Seth, who supported and challenged me throughout my time at Sussex University. Both took great efforts to form me with constructive criticism into a better scientist and I greatly appreciate the innumerable lessons learnt. A shared fascination for the nature of consciousness motivated this highly collaborative work.

Key insights of this work were shaped further in discussions with Lionel Barnett, Chris Buckley, Pedro Mediano and Enzo Tagliazucchi. I'm also most thankful for the proactive and effective collaboration with the teams that provided data sets for re-analysis: Quentin Noirhomme, Melanie Boly, Marie-Aurelie Bruno and Steven Laureys for EEG data obtained at the Université de Liège. Andrea Pigorini, Steve A. Gibbs, Gabriele Arnulfo, Simone Sarasso, Lino Nobili and Marcello Massimini for depth electrode data obtained with facilities of the University of Milan and Niguarda Hospital Milan. Robin Carhart-Harris and Suresh Muthukumaraswamy for MEG data obtained at Cardiff Univesity.

Many thanks also to Murray Shanahan, Joon-Young Moon and again Pedro Mediano for highly useful support with computer simulations of coupled oscillators. I'm further grateful for the many stimulating conversations within the seminars of the Artificial life reading group at the University of Sussex and with my colleagues at the Sackler Center of Consciousness Science, in particular Acer Chang, Keisuke Suzuki, David Schwartzmann, Georgina Heron and Warrick Roseboom.

Consciousness is everything we have,
and everything we are.

Giulio Tononi [233]

Abstract

Why does it feel like something to be awake? I.e. how is consciousness generated by the body, the brain in particular? Seeking to map phenomenological properties of any first person experience to neural activity patterns, theories of consciousness suggest a correlation between a specific type of neural dynamical complexity and the level of consciousness: When awake and aware, all brain regions are to a certain extent connected and there is diversity in the interactions. In support of this, Casali et al. (2013) have used EEG and transcranial magnetic stimulation to show extensively that brain response activity to direct perturbation is the more diverse across regions and time, the higher the level of consciousness. The spatio-temporal diversity of the response signal is quantified by a single index, the perturbational complexity index (PCI), using a Lempel-Ziv compression algorithm. Motivated by this result, and given that spontaneous neural signals are easier to obtain than response signals to perturbation, this thesis proposes measures - based on Lempel-Ziv compression and entropy - to quantify spontaneous neural signal diversity across channels and observations. Our measures' sensitivity and specificity to conscious level is demonstrated by re-analysing resting state scalp EEG during propofol-induced anaesthesia and depth electrode recordings during sleep stages, resulting in consistently higher scores for subjects that are awake than being in propofol-induced anaesthesia or non-rapid eye movement sleep. In addition we demonstrate that our measures score higher for states induced by psychedelic substances by re-analysing resting state magnetoencephalography (MEG) data. We further explore in computer simulation how our measures and PCI behave as a function of connectivity of coupled oscillators, informing models of brain mechanisms associated with the loss of consciousness. While our measures may be weaker than PCI in terms of specificity and sensitivity to conscious level, they are quick and easy to compute and applicable to readily available resting state data. This thesis provides strong evidence that cortical signal diversity is a hallmark of consciousness, as predicted by integrated information and complexity theories of consciousness.

Contents

Chapter 1

Introduction

1.1 Motivation

The question how mind arises from matter has occupied thinkers of all times and is often called the mind-body problem [69]. Investigating the problem with modern scientific methods has only recently become an established research domain [155, 131, 33]. This science of consciousness aims to explain in terms of neural activity why we experience at all and what we experience. For a system like a human to have an experience differs starkly from other system properties, such as temperature or weight. Unlike temperature, which is externally measurable, experience is only directly accessible by being the system. However, an external observer can search for neural mechanisms that generate experience by testing proposed mappings between phenomenological properties (i.e. aspects of what it feels like to be the system) and neural activity patterns. The current attempts to solve this non-trivial mind-body problem are scientific models of the physical basis of consciousness. Even though these theories of mind are still in a phase of conception, they can now be tested using modern technology. For the first time in human history real progress on the mind-body problem can be made.

This progress is of great importance for a more complete description of reality, given that the conscious observer is the starting point for all scenarios in physics and a world without a conscious observer resembles a show without audience, being arguably without meaning for anyone. A better understanding of why and how we experience, i.e. consciousness[1], promises also to be of great practical use, including medical improvements, such as more reliable techniques to monitor anaesthetic depth, treating comatose patients effectively or investigating psychedelic substances as a cure for depression. For the more distant future, implications for ethics are conceivable. E.g. a clear understanding of the extent that a being can experience at all will help to know when it can feel pain or dis-

[1]The term consciousness has nuanced meanings, from having any experience at all, also known as phenomenal consciousness [31, 32], to requiring to memorise the experience at least for an instant to act on it e.g. by report, known as access consciousness [32]. Another distinction is between basic experience and higher order experience, as for example proposed by Edelman who distinguishes primary consciousness that other mammals may have (most basic experiences but requiring memory to act on them) from higher order consciousness that only comes together with human level language [87]. Throughout this thesis the word consciousness refers to phenomenal consciousness, i.e. any experience that an organism may have, in the sense that it feels like something to be this organism at that moment [162].

tress, in particular when the being is non-communicative such as animals, young children or maybe man-made systems.

Science progresses by a mutually informing interplay of experiment and theory. The experiments in consciousness science aim to establish external signs of consciousness, indicating if a human has any experience at all and if so what is the content of the experience. I.e. on the one side one can investigate signs for average properties of experiences during some extended period of time and call this a state with a particular level of consciousness [36]. For example wakeful rest can be seen as a state with a higher level of consciousness than the state of slow-wave sleep. Such levels of consciousness are closely related to global states of consciousness, characterising an organism's overall conscious condition and being distinguished from each other not only due to a different range of average properties of experiences, but also on behavioural and physiological grounds [26]. On the other side, one can look more closely at aspects of each single experience, known as conscious content [117], in the sense of what a particular experience is like. For example seeing a red colour is a different conscious content than feeling pain.

For each of these two foci, conscious level and conscious content, various behavioural and neural signatures were proposed [131]. These are motivated by taking our own consciousness as a starting point - the only fact we can be absolutely certain of, as already stated by Descartes - and arguing that other people that have a similar anatomy and behave similarly should also experience similarly. Behavioural markers of consciousness are for example the extent that simple commands can be obeyed to identify different levels of consciousness, in particular in the clinic, and verbal reports to identify conscious content, for example describing what one is seeing. An example of a neural marker of consciousness is oxygen consumption of the brain as measured by functional magnetic resonance imaging (fMRI), measured globally, across the whole brain, to indicate conscious level and locally to identify differences for different types of conscious content.

Theories of consciousness differ in ambition, from the global neuronal workspace model that states that a percept becomes conscious when its representation in neuronal activity is broadcast to many cortical regions - as introduced by Baars [11, 12] and developed by Dehaene [78, 79] - to Giulio Tononi's integrated information theory that bravely proposes rules how level and content of any experience can be exactly mapped to state transitions in networks of causally connected nodes. The latter is built on proposed mappings between phenomenal properties that every experience has and neural mechanisms, aiming to to map the first to the third person perspective on experience. I.e. seeking how features of internal subjective experience are caused by externally observable changes in the system. This idea started with Giulio Tononi and Gerald Edelman proposing a balance of functional segregation and integration of brain activity as a hallmark of conscious cognition [228]. They defined functional segregation within a neural system in terms of the relative statistical independence of small subsets of the system and functional integration in terms of significant deviations from independence of large subsets. Starting from this proposed hallmark of consciousness, Tononi and Edelman went further and suggested that every experience we have is firstly integrated, in the sense that all the features of the experience

such as colour and sound are experienced together as a whole, and secondly differentiated, meaning that each experience is different and a vast repertoire of experiences exists [229]. They then suggest neural activity patterns to reflect these phenomenological properties of integration and differentiation, inspiring the creation of various measures of brain activity to capture neural analogues of conjoined integration and differentiation, in the sense of interconnected, yet inhomogeneous, and containing a diverse set of states.

A recent empirical breakthrough provided striking evidence for conjoined differentiation and integration to be connected to consciousness. It was shown that the EEG response to transcranial magnetic stimulation (TMS) is more spatially widespread ("integration") and spatio-temporally diverse ("differentiation") when awake than when unconscious. A single index, called the perturbational complexity index (PCI), was devised to quantify the extent of both, the diversity and spatial range of the EEG response (PCI will be introduced in Section 1.4.9, defined and illustrated in Section 2.2). This measure's specificity and sensitivity across a wide range of clinically defined states of consciousness is unprecedented, including anaesthesia induced by various drugs, rapid-eye-movement sleep (REM), deep sleep and disorders of consciousness such as the minimally conscious state in which patients show only the faintest signs of consciousness [131]. This clear empirical finding confirms the theoretical prediction that conjoined differentiation and integration are hallmarks of consciousness, encouraging to continue the pursuit of finding maps between phenomenology and neural activity patterns to advance the understanding of consciousness.

1.2 Thesis overview

Inspired by the impact of the perturbational complexity[2] index for the science of consciousness, this thesis investigates the diversity of spontaneous electrophysiological thalamocortical signals during different global states of consciousness and their potential to reflect phenomenological differentiation. We do so by introducing and applying three new measures, designed to capture different flavours of signal diversity across channels and observations.

Briefly, these three measures are defined for multidimensional continuous time series, consisting of several channels and many more observations. Amplitude coalition entropy (ACE) quantifies the diversity over time of binary activity patterns, where such an activity pattern is defined by channels with amplitude beyond a threshold for the same observation. Synchrony coalition entropy (SCE) measures the diversity over time of binary synchrony patterns and Lempel-Ziv complexity (LZc) estimates the diversity of activity patterns by a Lempel-Ziv compression algorithm.

We define these measures in detail in Chapter 2 and illustrate their sensitivity to signal

[2]The word complexity here should be understood in the sense of Kolmogorov's algorithmic complexity [132], defined for a string of symbols to be the length of the shortest computer program that produces the string as output. Algorithmic complexity is also known as algorithmic randomness, since the closer the string to being a random permutation of all available symbols, the larger the algorithmic complexity, as approximations show [257]. Algorithmic complexity is not to be confused with dynamical complexity, the latter indicating some balance between randomness and order. Entropy and Lempel-Ziv complexity are both ways to capture signal diversity and are asymptotically identical for certain processes as will be shown in Section 1.4.7.

diversity across channels and observations using surrogate data, a simple autoregressive model and a Kuramoto model. We also present the perturbational complexity index (PCI) here in detail and illustrate the key result while its empirical results are summarised in Section 1.4.9.

Chapter 3 describes our analysis of resting state EEG of healthy subjects that are in wakeful rest, mild propofol sedation and propofol-induced general anaesthesia. All three spontaneous signal diversity measures score for all 7 subjects substantially lower for general anaesthesia as opposed to wakeful rest and for most subjects scores for mild sedation lie in between.

Chapter 4 covers the behaviour of ACE, SCE and LZc across sleep stages such as rapid-eye-movement sleep (REM), non-rapid-eye movement sleep (NREM) and wakeful rest, when computed for intracranial depth electrode signals from epilepsy patients. The robust decrease of all three measures for all 10 subjects for NREM sleep, in comparison to wakeful rest or REM sleep, further supports the measures' sensitivity and specificity to conscious states in line with clinical classification. We further show that changes in signal diversity with conscious level as captured by all three measures are a global effect without strong local variation across brain regions. We find evidence however that independent of conscious state, signal diversity is largest in the frontal lobe when compared to other brain regions, suggesting a connection of structural diversity and signal diversity.

In Chapter 5 the increase of spontaneous signal diversity for psychedelic states is reported, inferred from the application of our spontaneous signal diversity measures to resting state magnetoecephalography (MEG) data. Conscious states induced by the psychoactive substances LSD, ketamine or psilocybin are compared to placebo, reanalysing datasets from three independent experiments, one for each drug. Our results provide first direct evidence that measures of spontaneous cortical signal diversity that are sensitive to other levels of consciousness, such as sleep stages or propofol anaesthesia, score higher for the psychedelic state than wakeful rest. Implications for a multidimensional characterisation of conscious states are discussed, in particular as we directly compared brain signal diversity scores to subjective ratings of the psychedelic experience.

Finally, Chapter 6 is about the behaviour of the perturbational complexity index (PCI) and our spontaneous signal diversity measures (ACE, SCE and LZc) for computer simulations of coupled Stuart-Landau oscillators. After investigating the signal diversity measures' dependence on mean coupling and topology of the oscillators, brain-mechanism-inspired adjacencies are presented for which all 3 spontaneous signal diversity measures decrease, as empirically observed for unconscious states, with increasing coupling of the oscillators. Limitations and strengths of the model are discussed with particular focus on activity time series similarity between oscillators and local field potential as well as a separation of signal diversity across time alone and signal diversity across channels and time.

Conclusions about all empirical findings and simulation are presented in Chapter 7. Supplementary results for the analysis of EEG during propofol anaesthesia, depth electrode recordings during sleep and MEG during psychedelic states are presented in Appendices A

to C, respectively.

1.3 Thesis contribution

New measures of multidimensional signal diversity were introduced and demonstrated to robustly index the level of consciousness during REM sleep, wake, propofol-induced anaesthesia and psychedelic states. We provided first evidence for the psychedelic state to lie above conscious states such as wakeful rest and REM sleep on a one-dimensional scale defined by neural signal diversity. By describing and interpreting spontaneous signal diversity measures in comparison to the perturbational complexity index (PCI), a clearer distinction of spontaneous and perturbational signal diversity signatures of consciousness is provided. The application of these spontaneous signal diversity measures and PCI to systems of coupled oscillators is further novel, advancing the understanding of signal diversity signatures of consciousness and the coupling mechanisms of macroscopic brain regions that lead to such signal features. A key further contribution is the discussion of our results in the context of theories of consciousness, arguing that signal diversity reflects phenomenological differentiation. Overall this thesis encourages the pursuit of the search for representations of general phenomenological features of brain activity patters in order to advance complexity theories of consciousness.

Michael Schartner did all analyses and simulations in collaboration with Adam Barrett and Anil Seth. All empirical data sets were obtained from external research groups, and pre-processing was discussed with each group. The EEG data for the propofol-anaesthesia analysis were provided by Quentin Noirhomme, Melanie Boly, Marie-Aurelie Bruno and Steven Laureys at the Université de Liège. The intra-cranial depth electrode data for the analyses of sleep stages were provided by Andrea Pigorini, Steve A. Gibbs, Gabriele Arnulfo, Simone Sarasso, Lino Nobili and Marcello Massimini at the University of Milan and Niguarda Hospital Milan. MEG data for the analyses of altered states of consciousness induced by psychoactive substances were provided by Robin Carhart-Harris and Suresh Muthukumaraswamy, obtained at Cardiff Univesity.

1.4 The wider context: Literature review

> "You", your joys and your sorrows,
> your memories and your ambitions,
> your sense of personal identity and
> free will, are in fact no more than the
> behaviour of a vast assembly of nerve
> cells and their associated molecules.
>
> *Francis Crick [71]*

1.4.1 Studying consciousness: From philosophy to neuroscience

Falling each night into dreamless sleep changes the quality of experience abruptly for most of us. A rich inner mental life during waking hours may include the sound of a bird or the image of the sky while during phases of dreamless sleep at night, all experience seems absent. We can share our impressions with others by describing what we experienced during waking hours, be it by language or forms of art, while this seems impossible after certain epochs of sleep, suggesting that we all noticed this change of awareness between dreamless sleep and wakefulness each day. The absence of experience during phases of sleep arguably made us notice that we experience at all [80].

Humans have ever since wondered what this internal perspective is and how it relates to the body, the so called mind-body problem. Ancient Western philosophers such as Democritus and Epicurus proposed that mind arises from matter, stating "there is only atoms and the void" [147], a philosophy broadly called physicalism, while others such as Descartes in the 17th century argued that mental substance is distinct from matter, an idea known as substance dualism, with a wide opinion spectrum about the extent that mind depends on matter [69]. Also in Eastern philosophy there are ancient accounts for both positions, e.g. the Yogi school proposing substance dualism [138] while the ancient Hindu Charvaka school advocating physicalism [60].

Eastern spiritual traditions brought about a rich culture of introspection and meditation as tools to study the mind [172]. In the West experimental psychology started to develop in the 19th century, notably influenced by William James in 1890 [123], and provided further tools to study consciousness. This direct study of consciousness by psychologists stopped after James due to a conviction called behaviourism, stating that human and animal behaviour can be explained and should be studied in terms of conditioning alone, without appeal to thoughts or feelings, as argued for by Watson and Skinner [246, 215]. I.e. a key tool to study consciousness, subjective reports as raw data, became only accepted again by psychologists towards the end of the 20th century [28, 80].

Interest in the neuro-scientific study of the mind-body problem substantially increased from around 1990 onward, fuelled by technological advances such as magnetic resonance imaging and faster data processing due to better computers. Publication of studies on neural correlates of consciousness appeared, in particular by Francis Crick and Christof Koch [70]. They proposed that looking for neural correlates alone allows the science to progress independently from philosophical disputes [72]. Neural correlates of consciousness have been defined as "the minimal neural mechanisms that are jointly sufficient for any one conscious percept, thought or memory, under constant background conditions" [184, 128, 131]. A distinction is made between neural correlates of states with a particular level of consciousness, this level being higher for example in wake than dreamless sleep, and content specific neural correlates, with content in the sense of what each experience is like, for example seeing a red apple is a different conscious content than hearing an ambulance siren and each triggers different neural correlates when perceived.

Consciousness science has since become a multidisciplinary research focus with its own journals such as "Neuroscience of Consciousness" and conferences such as "The Associa-

tion for the Scientific Study of Consciousness" (ASSC). The academic background of active participants ranges from philosophy over neuroscience and biology to physics and mathematics. In addition to the ones sketched above, many more philosophical and scientific opinions on consciousness influence the current debate. A compact presentation of the work of influential and currently active philosophers and scientists of consciousness can be found in recent books by Cavanna and Nani [55] as well as Miller [155]. Comprehensive introductions to the science of consciousness aimed at the general reader were for example published by Koch [130], Bor [38] and Dehaene [80].

1.4.2 Measuring consciousness: The clinic as a starting point

In contrast to the philosophical debate on what is consciousness and whether can we study it, clinicians need to define and measure levels of consciousness in order to treat patients effectively. E.g. assessing whether a patient experiences pain is crucial in order to know if more of an anaesthetic substance should be administered before a painless surgery can commence. Consciousness as measured and defined by clinicians is a practical starting point in the discussion about measures and theories of consciousness.

A spectrum of conscious states was proposed by clinicians to be situated in a two-dimensional space, where on one axis the extent of awake behaviour is measures (arousability), meaning signs of vigilance such as opening eyes or command following, while the other axis is labelled level of consciousness, i.e. the intensity or diversity of experiences averaged over an extended period of time, as sketched in Fig. 1.1 (see [139, 36] for a more detailed version). States range from lowest conscious level and lowest arousability for coma and general anaesthesia to highest level of consciousness and arousability for conscious wakefulness. Certain states such as the vegetative state (also known as unresponsive wakefulness syndrome [140]) have high arousability, as e.g. wake-sleep cycles are observed, yet lowest awareness, since patients in such a state do not experience anything by definition.

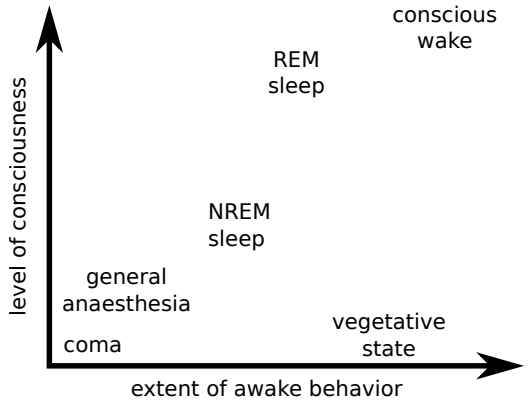

Figure 1.1: **Examples of global states of consciousness** Global states of consciousness can be classified using two dimensions, one being the level of consciousness, i.e. intensity or diversity of experiences, and the other being the extent of awake behaviour, such as opening of eyes or reaction to stimuli. More refined such classifications are used in the clinic. Adapted from [139].

The tools for clinicians to classify the state of consciousness of patients can be mainly divided into analysing either behaviour or brain signals, with a combination of both being

most promising [201]. behavioural measures include refined versions of the Glasgow coma scale [224], where the patients' motor, verbal and eye responses to stimuli are classified (see [180] for a comprehensive descriptions of the diagnosis of stupor and coma). behavioural measures cannot be used to detect consciousness for states such as locked-in syndrome, where - in the most severe cases - patients cannot move any muscle voluntarily to give signs but are still conscious.

The direct analysis of patients' brain activity as a way to classify their level of consciousness has been realised using various technologies. For example the BOLD signal of magnetic resonance imaging machines (fMRI) was used to show that a richer repertoire of functional configurations is present in wakeful rest as opposed to anaesthesia [25] or non-rapid-eye movement sleep [221]. As another example, the activity patterns seen via fMRI while subjects watch a movie clip differ consistently across populations of conscious and unconscious subjects, which allowed for instance the identification of a patient as being conscious who has been entirely behaviourally non-responsive for 16 years by showing him the movie clip and classifying the fMRI activity pattern as similar to that of healthy conscious subjects [161]. Also differences in brain metabolism could be used to assess patients' level of consciousness [216, 139], in particular for distinguishing minimally conscious patients (some faint level of consciousness is sometimes present) from patients with unresponsive wakefulness syndrome (vegetative state, without any inner mental life), deemed a particularly challenging and clinically urgent distinction between two global states of consciousness [34].

Electrophysiological means provide another popular modality to obtain brain signals for the assessment of the level of consciousness. In particular electroencephalography (EEG), where electrodes capture local field potential at the scalp, is often analysed for signatures of consciousness [214]. Automatic feature detection in EEG signals, such as performed in a method known as bispectral index [212], is a tool for anaesthetic monitoring found in many clinics, which does indicate levels of consciousness up to a point, yet the signal features remain unknown and thus do not inform theories of consciousness. The analyses of spectral changes provides well defined EEG signatures of conscious states, for example showing that slow wave power (e.g. delta brain waves, that comprise frequencies from 0.5 to $4Hz$) is elevated in unconscious states such as anaesthesia or deep sleep when compared to wakeful rest [163, 76, 245, 48, 152, 83]. These correlates have proven to be of practical use for anaesthetic monitoring and sleep-staging, yet fail to indicate conscious level in other cases, such as distinguishing minimally conscious patients (that have brief periods with conscious experiences) from vegetative state patients (a state without any experiences) [175].

Informed by theory, a new approach is to analyse the response EEG activity to transcranial magnetic stimulation and quantifying how diverse this EEG response is across EEG channels and observations. The response signal diversity is quantified by an index called the perturbational complexity index (PCI). It did not only outperform all other behaviourally independent measures of conscious level in terms of telling minimally conscious state patients apart from vegetative state patients [53], it further captures interpretable signatures that are important to further advance theories of consciousness. We'll present

PCI in full detail in Sections 1.4.9 and 2.2.

There is further the search for anatomical structures that facilitate and mediate the level of consciousness (see [131] for a recent review, and [98] for a broader evolutionary perspective). For example certain brain regions, such as the reticular activating system in the brainstem, are indispensable for healthy consciousness in humans while activity in other brain parts such as the cerebellum has little effect on consciousness and its contents [131]. In fact, two patients were described that each lacked the cerebellum altogether yet had a fairly normal level of consciousness [144, 252]. As another example, the claustrum may play a crucial role for consciousness [73], supported by a case study of a patient whose consciousness disappeared when electric current was sent into the claustrum via a depth electrode [134].

Each neural correlate of consciousness helps to extend the description of consciousness and restricts proposed theories thereof [201], while conversely, most robust correlates of consciousness used in the clinic are increasingly the outcome of applied science of consciousness, as the success of the theory-based PCI demonstrates. After this introductory illustration of practical notions and measures of consciousness used in the clinic, the remainder of this chapter provides an overview of trends towards a theory of consciousness.

1.4.3 Conscious level and conscious content

Clinically defined levels of consciousness delineate the overall "richness" of awareness, from complete absence of any experience as in coma to a rich inner mental life as in wakeful rest [227]. Such levels of consciousness form one dimension of global states of consciousness, the latter characterising an organism's overall conscious condition and being distinguished from each other not only due to a different range of average properties of experiences (conscious level), but also on behavioural and physiological grounds [26]. When considering a single experience one can more finely analyse its phenomenological properties, such as being aware of a particular scene, a particular feeling or a combination thereof. Such properties are collectively called conscious content [117] and vary from instance to instance during global states of consciousness, such as REM sleep or wakeful rest.

Subjective report is seen by many as the gold standard of what is consciously perceived, and often used to find conscious-content specific neural correlates of consciousness. Experimental designs range from studies about binocular rivalry, where perception alternates between two different images and the corresponding neural activity for each is sought [30], to various visual masking techniques, e.g. forward and backward masking, helping to separate a hierarchy of mechanisms that lead to a conscious percept [41] (see [131] for more examples of conscious-content specific neural correlates of consciousness). There are cases though where subjective reports are not reliably indicating if something was experienced or not, as most of us know from reporting dreams, where experiences are easily forgotten. A wide range of other experiments further illustrate how subjective reports can fail in indicating experience, from change blindness (the change of two radically different percepts is not notice) to inattentional blindness (the perception of objects cannot be reported if they were unexpectedly presented outside of the focus of attention) and many more, as

summarised here [137]. The great variety of experimental designs to study conscious content reflects the vast diversity of possible experiences. Cognitive studies of consciousness are seamlessly mingled with the wider literature on psychophysics and the psychology of perception.

A theoretical challenge lies in the connection of conscious level and conscious content. Overgaard suggested for example tying levels of consciousness to the average clearness or vividness of contents [170]. Bayne et al. developed this connection further by suggesting the characterisation of global states of consciousness along several dimensions, where arousability may be one, brain signal diversity may be another and yet another may correspond to properties of average conscious content. An example of such average content was suggested to be the average level of abstraction about perceived objects, in the sense that perceiving only the colour of a cup is less abstract than consciously perceiving its purpose as a container [26, 97]. Giulio Tononi's integrated information theory of consciousness tries to explain both, conscious content and conscious level, directly and fundamentally as independent properties of state transitions as will be discussed in more detail in Section 1.4.5.

1.4.4 Complexity theories of consciousness

Forms of physicalism are widely accepted amidst consciousness scientists, e.g. in the sense that "we are food rearranged" and consciousness is a particularly interesting property of matter, as phrased by Tegmark [225]. I.e. it feels like something to be matter (paraphrasing Thomas Nagel [162]), to the extent that the building blocks are arranged in a particular dynamical structure that enables rich inner life. Barrett [23] and Pockett [177, 178] refined this formulation of physicalism by suggesting that consciousness is a property of particular electromagnetic field dynamics, brought about for example by the movement of charged particles in the brains of conscious humans. This suggestion followed by arguing that if consciousness is a fundamental attribute of matter and given that physics posits force fields or quantum particles as most fundamental building blocks of matter, consciousness must be described in terms of quantum particles or force fields. Since we know that signalling between neurons is mainly electrical and that quantum effects such as quantum coherence are most unlikely in systems with "high" temperature such as the brain [129], force fields seem more plausible than quantum particles. In comparison to the other force fields, the electromagnetic force field appears as the best candidate for the fundamental substrate of consciousness in human brains, as is argued by comparison of the four forces of nature: strong, weak, electromagnetic and gravitational. The strong and weak forces don't propagate over distances much larger than the width of the nucleus of an atom, and gravity alone cannot generate complex structures by virtue of being solely attractive; in contrast, the electromagnetic field can propagate over macroscopic scales, is both repulsive and attractive, and is fundamentally what enables non-trivial chemistry and biology [23].

The theoretical challenge lies in the description of dynamical properties related to consciousness, following from the conviction that consciousness is a process, as William James famously proposed already in 1890 [123]. Tononi, Sporns and Edelman suggested a fine balance of functional segregation and integration of brain activity as a characterisation

of overall brain dynamical complexity, where full integration means that all brain regions are active in the same way at the same time while full segregation means that each brain region's activity is independent from that of other regions [228].

Driven by the assertion that qualities of experience must somehow be mapped to neural activity, these thoughts on neural complexity were directly connected to consciousness by Tononi and Edelman [229, 88]. In particular they stated that each experience is integrated in the sense that it is composed of a complete scene instead of independent aspects, and each experience is differentiated in the sense that whatever is perceived rules out an enormous number of other possible experiences one could have in that moment. Hence "integration" and "differentiation" in neural activity patterns were said to reflect corresponding phenomenological properties.

1.4.5 Integrated information theory of consciousness

Tononi and collaborators extended and specified in detail the proposed map from phenomenological properties of any conscious experience on one side to neural processes on the other in what is called integrated information theory of consciousness. The theory has been evolving since 2004 with the latest summary published in 2016 [231, 232, 15, 166, 234, 235]. In addition to differentiation and integration, the list of phenomenological properties to be mapped to neural activity was extended, including "internal existence" (meaning that each experience exists independent of external observers), "composition" (each experience consists of phenomenological structure) and "exclusion" (each experience only exists at its particular spatio-temporal scale). Analogies of these particular properties, that any experience proposedly has and which are thus called axioms in the theory, are sought in state transitions of idealised networks, i.e. a graph of binary nodes where each node can either be on or off and edges are causal functions. Knowing all possible state transitions of the network, the theory allows a particular transition to be mapped to a particular conscious content and in addition to its intensity, i.e. interpreted as conscious level, using various mathematical expressions, many of which are based on mutual information. The most recent formulation of the theory is centred around the postulate that a system is conscious to the extent that it influences itself (has cause-effect power upon itself). I.e. the theory is boldly stating that an experience is identical to a particular network state transition [235]. The theory further postulates that the spatio-temporal grain with maximum intrinsic cause-effect power is the only one relevant for the system's experience and the potential consciousness creation of all other intrinsic cause-effect mechanisms are suppressed, implying in a sense that there cannot be independent minds in a connected system [235].

Integrated information theory (IIT) does not require the network to be biological, the causal architecture alone determines if the system experiences. Consequently the theory asserts experience - and be it nearly zero for all practical purposes - to any system that has state changes and thus IIT can be seen as a version of panpsychism (the idea that everything experiences [198]). Yet this is not to say that IIT would propose consciousness to be equally distributed across matter, quite the opposite. IIT describes consciousness as a strongly

graded property across systems, depending on the system's intrinsic cause-effect power, i.e. how much the system has an effect on itself [235]. This implies that complicatedness alone does not correlate with the level of consciousness of the system, as for instance complicated feed-forward implementations of computer simulations are proposed as systems with very low integrated information [234].

Overall, integrated information theory of consciousness is arguably the most discussed and advanced theory in the context of the body mind problem in recent years, with numerous advocating publications [235, 53, 226, 167], heralding it for example as the only theory that starts from experience itself instead of the brain and that tries to attack the hard-problem of consciousness[3] head-on [235]. It is criticised by others, to overstate the usefulness of existing measures of integrated information given that they require inelegant searches for the maximum intrinsic cause-effect power across all temporal and spatial scales of the system [24]. Further, Cerullo belittles IIT's strength in explaining consciousness by arguing that his own, absurd, ad-hoc theory has the same explanatory power [57], and Aaronson challenges the theory's measures of integrated information by applying one to a simple network for which the measure's score grows with network size, while the network seems too simple to be conscious [1, 2].

1.4.6 Global workspace theory, recurrent processing and predictive coding

Besides integrated information theory, other theories of consciousness are popular in the community of consciousness scientists. Global workspace theory, introduced by Baars [11, 12] and further developed by others [78, 206, 79], postulates that neural activity related to conscious percepts is broadcast to many brain areas (called the global workspace), possibly inspiring the "integration" notion found later in complexity theories of consciousness. I.e. unconscious processing, termed subliminal, is represented by a number of distinct distributed local processes that compete for conscious access. Experience takes shape when widespread activity in these nodes is broadcast into a global workspace [155]. The theory inspired computer simulations [77, 207] and informed recent experiments focusing on subliminal processing [81, 86]. Global workspace theory was also discussed in the context of conscious level, interpreting findings of more long range information flow across the cortex in conscious as opposed to unconscious states [127].

Other theories stress the importance of feed-back information projection in the brain for conscious perception. For example the recurrent processing theory, developed mainly by Lamme [136], initially motivated from experiments about the separation of experiencing and having access to the experience (being able to report or act on it). It has developed into a general theory of conscious visual perception [155] with recent experimental support

[3]The hard problem of consciousness science, a term coined by Chalmers [58], is to explain why and how any physical processes - such as those occurring within brains - give rise to conscious experiences at all. Chalmers contrasts this with the 'easy problem' of explaining how the complex biophysical operations of the brain enable all brain functions that don't require consciousness, for example image recognition. In a strict sense, IIT does only attempt to answer the question "how" is consciousness created by physical processes, not "why", i.e. leaving out half of the answer to the hard-problem. Claiming that consciousness is a property of a particular electromagnetic field configuration may offer more towards the "why" [24].

[85]. Unlike global workspace theory, recurrent processing theory states that consciousness does not necessarily need the involvement of the complete fronto-parietal network, but posterior cortical areas are sufficient as long as sufficient recurrent interactions occur [136].

Feedback information projection is also a central idea when applying predictive coding to explain the content of experience. First proposed by Helmholtz, predictive coding is the view that the human brain constantly generates hypothesis about the world which are corrected using sensory information. Hohwy suggested to apply this viewpoint to explain conscious content [118], for example for binocular rivalry experiments as follows: alternation between the conscious perception of two competing stimuli occurs when firstly, there is no single hypothesis to explain a stimulus with high probability and, secondly, when one stimulus dominates, the bottom-up signal for that stimulus is explained away while the bottom-up signal for the suppressed stimulus is not [116]. Others used the framework to explain particular conscious content such as interoceptive awareness [203].

All here described theories of consciousness converge on certain points, such that consciousness requires neural activity in distributed thalamocortical networks and that feedback projections among them are important. While direct experimental support for IIT seems weaker than for the other theories mentioned above, IIT differs strongly in ambition and scope. However, complexity theories of consciousness [229] which subscribe more weakly to the concepts of IIT - a particular overlap is that both "integration" and "differentiation" are hallmarks of consciousness - have strong experimental support by the results of the perturbational complexity index (PCI). Global workspace, recurrent processing and predictive coding all provide important correlates of consciousness yet only IIT offers a fundamental explanation for such correlates by its bold claim that experience of a system equals the maximal amount of cause-effect power that the system has on itself. This identity can be used to explain for example that feedback connections are essential for consciousness since no purely feedforward system has intrinsic cause-effect power. IIT is further distinct from the other theories as it takes phenomenological properties as a starting point and not the brain [235].

1.4.7 Differentiation measures

One approach to advance the science of consciousness is to more or less blindly "fish" for correlates of consciousness and then use the results to restrict theories of consciousness, as may have been the case for many currently known correlates [131]. Each such serendipitously found correlate is by definition related to consciousness but not necessarily informative for theories of consciousness, which would be the case for example if the correlate can be clearly mapped to proposed phenomenological properties of experience. As an alternative approach towards better theories of consciousness, one can start from existing theories to inform the design of potential measures of consciousness and apply these to brain signals for different states of consciousness.

Complexity theories of consciousness and IIT stressed that phenomenological "differentiation" must be reflected in the dynamics of the system [229, 88]. I.e. that each experience is nearly unique, ruling out a vast number of possible other experiences one could have,

suggests that a great variety of different brain states are traversed as experiences are constantly changing. This differentiation property can be linked to various signal diversity measures, designed to analyse time series of neural activity.

Measures using various forms of entropy to assess temporal signal diversity are widespread in medical signal analysis, either based on the time domain or frequency domain of the signal. The entropy of a system that can be in N different states, the i^{th} state appearing with probability p_i, is defined as

$$H(\text{system}) = -\sum_{i=1}^{N} p_i \log(p_i). \tag{1.1}$$

There are many ways to define states of the brain from neural time series. Liang et al. [149] recently compared 12 popular entropy measures for EEG during different anaesthetics (sevoflurane and isoflurane), all evaluating the entropy of single EEG channels, concluding that all of them could track conscious state changes when applied to EEG, scoring lower during anaesthesia than wake, with forms of permutation entropy (Tsallis permutation entropy [256] and Renyi permutation entropy) tracking changes best [149]. Permutation entropy defines states in single EEG time series by first dividing the time series into non-overlapping segments of fixed length m and then ordering the observations of each segment by amplitude. Each order of a subsequence is thus a particular permutation of the first m integers. The frequency of that permutation within all occurring permutations is then p_i in Equ. (1.1) (see [149] for the mathematical definition). Even though the authors aim for high sensitivity and specificity for the sake of practical applications in the clinic, their results can be interpreted in light of reflecting phenomenological differentiation.

Other attempts to quantify temporal signal diversity in neural signals, compared across conscious states, include versions of Lempel-Ziv complexity, using an algorithm developed for file compression [4]. The more diverse the signal, the less compressible it is (see detailed definition in Section 2.1). In order to compute Lempel-Ziv complexity for continuous time series, the signal must be discretised, since the Lempel-Ziv algorithm searches for repeated subsequences and won't find any in a continuous signal. There are many ways to discretise continuous time series, for example by binarising them using their mean amplitude over time or using any discretisations proposed in the context of entropy measures [149]. Entropy and Lempel-Ziv complexity for a binned neuronal spike train are asymptotically identical, if the neuronal source is ergodic[4].

[4] As described in [6], a binned spike train can be represented as a binary string $s_n = x_1 x_2 ... x_n, \quad x_i \in \{0, 1\}$, where $x_i x_{i+1}$ indicates concatenation. The entropy (rather entropy rate) of this string can be defined as a function of substring length l ($1 \leq l \leq n$), as $H(l) = -1/l \sum p_i \log_2(p_i)$, where p_i is the probability that the i^{th} substring of length l appears. The true entropy of the neuronal process that generates the spike train is defined as $H(\text{process}) = \lim_{n,l \mapsto \infty} H(l)$ which is in practice impossible to compute, as it would require infinitely extensive sampling. Let $C(s_n)$ be the Lempel-Ziv complexity of the binary sequence s_n obtained by applying LZ-76, i.e. the Lempel-Ziv algorithm from 1976 [146] (see also Chapter 2), then

$$\limsup_{n \mapsto \infty} \frac{C(s_n) \log_2 n}{n} = \lim_{n,l \mapsto \infty} -1/l \sum p_i \log_2(p_i) = H(\text{process}) \tag{1.2}$$

provided that the source of the spike train is ergodic. Ergodic means stationary (i.e. statistical properties of substrings of the same length coincide almost surely) and in addition ensemble averages and time averages coincide almost surely, that is, one can calculate expected values over the substring ensemble

A Lempel-Ziv based measure applied to single EEG time series, binarised by mean activity for $30sec$ segments (6000 observations at $200Hz$), scored robustly higher for wakeful rest than general anaesthesia (for any of sevoflurane, isoflurane, propofol or desflurane) [253]. Computing Lempel-Ziv complexity for time series discretised using the method of permutation entropy (see previous paragraph), was shown to increase discriminative power between wake and anaesthesia [14]. A decrease in Lempel-Ziv complexity during aneasthesia was further found for pentobarbital-anaesthetised rats [210], and rats in deep sleep [3], showing that Lempel-Ziv complexity indexes these states also in rats besides humans.

Single channel entropy and Lempel-Ziv complexity measures did clearly index several conscious states at the group level, yet there were considerable variations across subjects [149]. These signal diversity measures on spontaneous brain recordings indicate the extent of randomness alone, irrespective of whether the diversity ensues from a semi-dependent interplay of connected or independent processes (see Section 2.6.1 for examples).

A different indicator of differentiation in spontaneous brain activity is the statistical properties of neuronal avalanches. A neuronal avalanche can be understood as neural activity that originates at one cortical position and subsequently spreads across other cortical regions. Statistical properties of such avalanches, such as size or length distribution, were found to vary consistently for different states of consciousness. For example avalanche analysis performed on signals from intracranial depth electrodes, where needle-like electrodes are inserted into cortical and subcortical brain regions of epilepsy patients for pre-surgical evaluation, resulted in more large avalanches in NREM sleep as opposed to wakeful rest [182]. This study and others [182, 197, 223] aim to infer dynamical "distance" of states of consciousness from criticality, in the sense that activity in sub-critical states dies out while in supercritical states all brain regions fire in synchrony, as observed during epileptic seizures [63].

The hypothesis of the brain operating close to criticality is still debated [27, 165] with particular criticism of using a power-law distribution of avalanche sizes to infer scale-invariance of the dynamics and that this scale-invariance implies criticality [236, 181]. However, the study by Scott et al. [197] is not using the contentious proximity to a power law of the avalanche size distribution to infer the distance to criticality. Instead it is shown that the distribution of avalanche sizes varies with the pixel size of the activity-videos (displaying voltage changes in a portion of the rat cortex) only during anaesthesia. This adds new evidence that neural activity during wake is more complex than during anaesthesia in the sense that it is scale invariant - given that the distribution of avalanche sizes was the same at different scales of the dynamics (different pixel size for activity videos) for wake but not for anaesthesia.

1.4.8 Integration measures

Experiencing something feels integrated [229, 88], in the sense that each conscious percept is an integrated whole, made up of information from many different sources such as vision,

using the relative frequencies of subsubstrings of one sufficiently long substring. Interestingly, $\frac{1}{n}C(s_n)\log_2 n$ sometimes converges for reasonably small n and can thus be used as an approximation of $H(\text{process})$, see [6] for simulations.

the auditory system, tactile, olfactory, vestibular system, interoception and taste or abstract thoughts. Finding neural correlates of integration that in addition are sensitive to consciousness is thus of theoretical interest.

Integration may be measured by synchrony of activity, especially temporal correlation, as has been already proposed by Crick and Koch in 1990 [70] as a correlate of consciousness, in particular in the high frequency components. Engel and Singer [91] review studies in humans and animals that support the connection of synchrony of neuronal signals of fast oscillations (gamma activity, $40-70Hz$) and consciousness. The reviewed studies describe how gamma synchrony between regions at the cellular level correlates with aspects of the binding-problem, i.e. how brain circuits that each process a different feature of a percept - e.g. sound and vision - contribute to a percept that is perceived as a whole, e.g. an audiovisual experience. In particular experiments on binocular rivalry make it very likely that only strongly synchronised neuronal signals contribute to awareness. However, gamma activity can be equally present during sleep phases with and without dreams [45] and the amplitude of gamma activity during anaesthesia is variable but is often greater than it is in the normal waking state [243]. This renders the relationship of gamma synchrony and consciousness ambivalent [201].

Overall mutual information between signals from different brain regions was further interpreted to indicate integration. The mutual information $I(X,Y)$ between two multidimensional time series of brain activity, X and Y, is the average information, or reduction in uncertainty (entropy), about X, knowing the outcome of Y [21]:

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{1.3}$$

where $H(X)$ is defined as in Equ. 1.1 with X being the system and a particular observation of X, e.g. discretised voltage at all EEG channels at a particular time, being a state of the system. $I(X,Y) = 0$ if X and Y are statistically independent, else $I(X,Y) > 0$. The average mutual information across pairs of fMRI resting state recordings from different brain regions was shown to be lower for propofol-induced anaesthesia than wakeful rest [196]. Similarly, overall mutual information for pairs of brain regions was found to be lower for vegetative state patients as opposed to control subjects, when calculated from resting state EEG recordings and using a particular "symbolic" discretisation of the time series before computing mutual information (denoted as symbolic mutual information) [127].

Although these studies affirm that the mutual influence across many brain regions is necessary for consciousness and may reflect the phenomenological property of each experience being an integrated whole, the mentioned studies show clear changes at the group level only and score less consistent for single subjects [53].

1.4.9 Measures of conjoined differentiation and integration

Several theory-derived consciousness measures aim to capture a form of conjoined differentiation and integration in brain activity. This started with Tononi and Edelman's suggestion that a fine balance of integration and segregation of the neural network dy-

namics is a hallmark of conscious processing and a measure called neural complexity was designed to peak at this balance. Functional segregation within a neural system was defined in terms of the relative statistical independence of small subsets of the system and functional integration in terms of significant deviations from independence of large subsets. Their measure, neural complexity, was defined for n-dimensional time series X, each series assumed to be a stationary stochastic process. Neural complexity of the system X is defined as the average mutual information I (defined as in Equ. 1.3) between bipartitions of X summed over all bipartition sizes k [228]:

$$C_N(X) = \sum_{k=1}^{n/2} \langle I(X_j^k, X - X_j^k) \rangle_{\text{all j}} \tag{1.4}$$

where X_j^k is the j^{th} possible subset of X when choosing k from n system components and $X - X_j^k$ its compliment. I.e. X_j^k represents some k-dimensional time series. $I(X_j^k, X - X_j^k)$ is a measure of statistical dependence of the two parts making up the whole system, being zero if they are statistically independent, else positive. Given that $C_N(X)$ is the sum over all bipartitions, its score is low for systems whose components are characterised either by total independence or total dependence and high for systems whose components show simultaneous evidence of independence in small subsets and increasing dependence in subsets of increasing size [228].

The neural complexity measure was applied to various model systems and behaved as intended [228, 18]. Yet, its application to EEG signals showed rather the opposite as theory predicted, namely an increase for states of reduced consciousness such as propofol anaesthesia [240] and patients with Alzheimer's disease [242]. This may be as stationarity is assumed for the definition of $C_N(X)$ and neural time series are not stationary enough. However a recent application to fMRI data showed higher neural complexity values when a movie as opposed to a scrambled movie is seen [37], which is interpreted to reflect that different meaningful stimuli are informative in different ways for different brain regions, while the scrambled movie is only one stimulus.

Tononi and Edelman's neural complexity measure inspired the creation of many alternative measures of conjoineddifferentiation and integration in one form or another, most based on notions of mutual information between time series from different system (brain) parts. These measures aimed to be "explanatory correlates" of consciousness [202], by defining mappings from proposed universal properties of any experience to neural activity, in particular using Tononi and Edelman's suggestion of conjoinedintegration and differentiation [228].

Tononi et al. proposed several measures of integrated information, most denoted by the Greek letter phi, [231, 232, 15] that aim to capture the extent of information that is generated by the system as a whole over and above that generated by its parts independently. The exact definitions of these measures have changed substantially over the years (see [226] for a compact presentation) with the most recent version being reviewed here [235]. Several of these measures were shown to behave in desired ways for certain model systems, for example they increased with fitness in the evolution of animats when

computed for a 12-node network that constitutes each animat's "brain" [89]. These "pure" phi-measures, that each claim not only to capture correlates of experience but to directly describe content and level of each experience, are only applicable to networks of a few nodes with exact knowledge about the system's states and transitions. I.e. their enormous computational cost and the requirement to know all states and transitions of the system exactly render a direct application of these phi-measures to neural time series impossible.

More practical measures of integrated information were designed to be applicable to time series, under Gaussian assumptions [142, 21, 204, 226]. Several of these practical phi-measures were explored for different networks of autoregressive processes and shown to score high for networks in between all-to-all connectivity or total independence [21, 204] and are still theoretically refined today [226].

Applications of such practical phi-like measures to empirical data are sparse [167], yet examples exist such as reduced information integration found in EEG signals for propofol anaesthesia [142] and a measure called phi* [167], found to be higher for a meaningful stimulus (Charlie Chaplin movie) as opposed to a scrambled movie when computed for fMRI data, concluding higher neural signal diversity to reflect higher diversity across experiences [37]. The extent of support that these empirical results provide for the integrated information theory is not clear, given the absence of rigorous controls for trivial signatures that may influence these measures' scores such as the overall level of correlation or changes in the spectral profile.

Another measure introduced as potentially capturing a form of integration and differentiation in system activity that may be related to consciousness, is causal density [199, 200, 204]. It is defined as the average of all pairwise Granger-causalities between system elements, conditioning on all remaining elements.

Granger causality [109, 106, 107] quantifies information transfer from a (possibly multivariate) stationary stochastic process (time series) $x_1(t)$ to a jointly stationary process $x_2(t)$, optionally conditioned on a third jointly stationary process $x_3(t)$. It is conventionally operationalised in terms of linear autoregressive (predictive) modelling: to calculate the Granger causality $F_{x_1 \mapsto x_2|x_3}$ from x_1 to x_2 conditional on x_3, x_2 is regressed firstly on its own past together with the pasts of x_1 and x_3 (the "full model"), and then on its own past and the past of x_3 only (the "reduced model"). If the residual prediction errors for the full and reduced models are $\xi(t)$, $\xi'(t)$ respectively, then $F_{x_1 \mapsto x_2|x_3}$ is defined as the logarithm of the ratio of the "generalised variances" (determinant of the covariance matrix; see [20]) of the residual errors:

$$F_{x_1 \mapsto x_2|x_3} = \ln \frac{|cov(\xi'(t))|}{|cov(\xi(t))|}. \tag{1.5}$$

Intuitively, $F_{x_1 \mapsto x_2|x_3}$ quantifies the gain in predictive power attained by inclusion of the variable x_1 in the predictor set for x_2. [17, 19] show that Granger causality has a clear information-theoretic interpretation, and for Gaussian processes in particular is equivalent to *transfer entropy* [195, 171]. In practice, Granger causality estimation and statistical inference is limited to neural time series which may be reasonably well-modelled

as stationary, linear autoregressive systems[5].

Causal density for a system described by the n-dimensional time series X, is defined as

$$CD(X) = \frac{1}{n(n-1)} \sum_{i \neq j} F_{x_i \mapsto x_j | X_{[i,j]}} \tag{1.6}$$

where x_i is the i^{th} time series of the system and $X_{[i,j]}$ is the $(n-2)$-dimensional time series formed by omitting the i^{th} and j^{th} series.

Causal density has been shown to peak for a balance of integration and segregation in the connectivity of autoregressive model systems, being low for all-to-all connectivity and also low for independent clusters. While it has been applied to several scenarios in neuroscience, no conclusive connection between causal density and conscious level was reported [205].

The strongest supporting evidence for the connection of consciousness to a notion of conjoinedintegration and differentiation of brain activity - and thus potentially reflecting these two properties shared by all experiences - was provided by the perturbational complexity index, which we'll define in detail in Section 2.2. This measure quantifies the signal diversity in the response activity to a pulse of transcranial magnetic stimulation (TMS) to the thalamocortical system as measured by electroencephalography (EEG). Its scores are higher, the more brain regions show a temporally diverse response activity to the TMS pulse. I.e. the temporal signal diversity is captured across channels and across time (being the potential neural analogue of the differentiation across experiences) and increases the more channels show a complicated response to magnetic perturbation (the potential neural analogue of integration of the experience, implying that many brain regions need to interact to form the experience). The measure's specificity and sensitivity across a wide range of clinically defined states of consciousness is unprecedented [131]. Comparing perturbational complexity for different conscious states was first introduced in 2005 and over the last 10 years actively applied to different conscious states [153, 101, 187, 189, 53, 190], with perturbational complexity being consistently higher for conscious states such as wakeful rest, REM, ketamine-induced REM [190], wakeful rest in locked-in syndrome patients, as opposed to unconscious states such as NREM sleep, propofol-, xenon-, midazolam-induced anaesthesia and unresponsive wakefulness syndrome (vegetative state) [53]. Its behaviour further matched clinical classification of conscious states for intermediate levels of consciousness, such as mild propofol sedation, light sleep (phase N1), minimally conscious state patients and emerging minimally conscious state patients [53], states for which PCI scores lay in between those for wakeful rest and general propofol anaesthesia. I.e. the measure's scores reliably indicate clinically defined states of consciousness at the single subject level while providing at the same time strong support for the theoretical claim that conjoinedintegration and differentiation are hallmarks of consciousness.

Motivated by the landmark study of PCI, yet seeing its need to perturb the brain repeatedly as a limitation, this thesis investigates spontaneous signal diversity measures and contrasts them with PCI in terms of phenomenological mapping and correlation with con-

[5]Note that Gaussian assumptions are not required; however, for non-Gaussian systems, the information-theoretic interpretation may be more tenuous.

scious level. In the next chapter we'll define our spontaneous signal diversity measures and PCI, illustrate the behaviour of spontaneous signal diversity measures via model systems and present PCI's proposed sensitivity to "integration" and "differentiation" in detail.

Chapter 2

Definition of signal diversity measures

In this chapter the signal diversity measures are introduced that form the basis of this thesis. We define versions of Lempel-Ziv complexity and measures of coalition entropy to assess spontaneous signal diversity across channels and observations. We also describe the perturbational complexity index (PCI) here in detail and contrast it with our Lempel-Ziv complexity measure for spontaneous signals. Subsequently the behaviour of our spontaneous signal diversity measures is illustrated using computer-generated time series, employing surrogate data created by increasingly ordered random sequences, an autoregressive model and a Kuramoto model. Importantly, these signal diversity measures are all model-free and only minimal assumptions about the data are needed.

Python implementation of all three spontaneous signal diversity measures can be found online in the Supplementary Information of [191].

2.1 Lempel-Ziv complexity (LZc, LZsum, LZs)

Lempel-Ziv complexity quantifies the number of distinct substrings of a binary[1] string. The greater the degree of randomness, the greater the number of different sub-strings that will be present, and thus the higher the Lempel-Ziv complexity.

There are several different versions of the Lempel-Ziv complexity c of a binary string, in particular the algorithms introduced by Lempel and Ziv in 1976 [146], often abbreviated LZ76 and 1978 [255], LZ78, which is the algorithm we used for computing the measure LZc. LZ78 is defined in the following way. Let S be a binary string $S = s_1 s_2 \dots s_n$. Then c is the number of distinct 'words' in the string, obtained as follows. The dictionary is taken to contain the words '0' and '1' already at the outset. The next word to go in the dictionary is the first two digits of S. Then the algorithm proceeds iteratively: suppose we have already obtained the first r words $w_1, \dots, w_r$ from $s_1 s_2 \dots s_k$. Then w_{r+1} is the

[1]The procedure can be readily generalised for a discrete string composed of any number of different symbols. In the context of consciousness science, see for example Sitt et al. [214] that use a discretisation of the continuous time series into 32 equal bins before computing a classical open source compression algorithm (gzip) to implement Lempel-Zic complexity. Their results were not affected by the number of symbols, testing a range from 8 to 128 for the number of bins (symbols).

shortest string $\sigma_k^p =: s_k s_{k+1} \dots s_{k+p-1}$ such that $\sigma_k^p \notin \{w_1, \dots, w_r\}$; that is, the next word is the shortest sub-string of S that starts from the final digit of the previous word and is not already in the dictionary. An example string is $S = 01011010001101110010$, which gives the list of words $\{0, 1, 01, 10, 011, 101, 100, 00, 0110, 0111, 1001\}$. For a string consisting only of zeros, a relatively small dictionary is obtained: if $S = 00000000000000000000$ then the dictionary is $\{0, 1, 00, 000, 0000, 00000, 000000\}$. For LZ76, the new word is searched throughout $s_1 s_2 \dots s_k$ instead of just within the list of already found words as for LZ78.

Computing the Lempel-Ziv complexity of continuous time series, for example EEG channels, requires a binarisation. Casali et al. computed this measure for event-related, as opposed to resting EEG data, and so used a threshold relative to pre-stimulus activity (baseline) to define the binarisation, as we'll introduce in detail in Section 2.2. For the computation of Lempel-Ziv complexity for spontaneous signals we based the binarisation threshold on the absolute value of the analytic signal of a channel's time series. The threshold T_i for the i^{th} channel was chosen as the mean of the absolute value of the analytic signal of the i^{th} channel. Applying this binarisation to several EEG channels results in a binary matrix, with rows corresponding to channels (time series) and columns corresponding to time (observations).

When analysing the Lempel-Ziv complexity of multidimensional time series, there are several ways to transform the binarised data matrix into one binary string that is then divided into blocks of words by the LZ78 algorithm (we adapted open source code for the implementation [258]) and the number of words is the measure's score. The following two variants are used in this thesis. Firstly, concatenating the data observation by observation, resulting in the measure that we denote by LZc as shown in Fig. 2.1, which thus quantifies signal diversity across channels and across observations. Secondly, computing Lempel-Ziv complexity for each time series separately, resulting in single channel Lempel-Ziv complexity - measuring temporal signal diversity only - denoted by LZs. The average LZs across channels is denoted by LZsum.

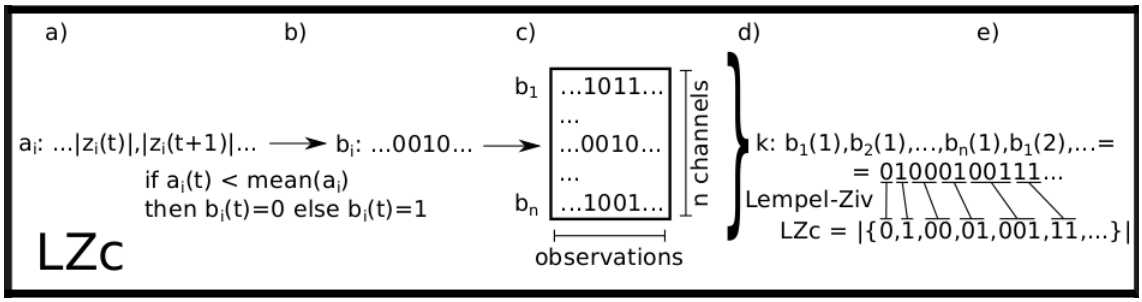

Figure 2.1: **Schematic of the computation of LZc.** a) x_i is the activity of the i^{th} channel and a_i is the (Hilbert) amplitude of x_i. b) b_i is binarised a_i, using the mean activity of a_i as binarisation threshold. c) After binarisation of all n signals, d) the multidimensional time series is concatenated observation-by-observation to one binary string k and then e) repeated patterns are searched and listed into a dictionary of binary words via a Lempel-Ziv algorithm. Lempel-Ziv complexity LZc is proportional to the size of this dictionary.

We normalise LZc (LZs) by dividing the raw value by the value obtained for the same binary input string randomly shuffled[2]. LZc's raw score for a binary string of fixed length is maximal if the string is entirely random. Thus the normalised LZc values indicate the level of complexity on a scale of 0 to 1.

2.2 Perturbational complexity index PCI

Lempel-Ziv complexity is also key for the perturbational complexity index (PCI), a measure that quantifies the diversity of the brain response activity to magnetic perturbation. PCI is colloquially known as "zap and zip", since first the brain is magnetically "zapped" then the response activity's diversity assessed by Lempel-Ziv compressibility, i.e. "zipped". PCI's most robust indication of levels of consciousness (Section 1.4.9 listed global states of consciousness successfully indexed by PCI) while at the same time being interpreted to capture brain activity patterns that reflect phenomenological features of "differentiation" and "integration" made us investigate multidimensional spontaneous signal diversity. The comparison of spontaneous and perturbational signal diversity as "explanatory"[3] correlates of consciousness is a central question of this thesis. The following technical details about PCI further apply for our computation of PCI for a Stuart-Landau oscillator model in Chapter 6.

2.2.1 Definition and methods

In order to quantify the diversity of the response activity to perturbation, first a perturbation must be applied and the response activity in the EEG across cortical regions defined. The transcanial magnetic perturbation was a short magnetic pulse, applied to a cortical region, say the visual cortex[4], and the EEG activity across the whole scalp (60 sensors) up to $300ms$ after the pulse was considered to reflect the response to this perturbation sufficiently. The signals were transformed from sensor space into source space, using the 3-spheres BERG method and the inverse problem was solved by the Weighted Minimum Norm constraint applied to an "empirical" Bayesian approach. This procedure turned the 60 EEG sensor channels into 3000 source channels (see supplement of [53] for more details).

Perturbing the brain about 100 times per subject (see [53] for varying trial numbers across different experiments) allowed an average response waveform across these trials to be found for each source. This average waveform was then binarised by a threshold T obtained from the pre-perturbation activity (from -500 to $-1ms$) across all trials. T was chosen as the 99th percentile of the distribution of the maximum absolute values of pre-perturbation activity sampled from all trials (perturbations), using the bootstrap-

[2]The variation in LZc for 50 different shufflings of the same input string of 25000 binary digits was under 0.002% of the mean result across those 50, showing that the data matrices we analysed - typically 10 channels times 2500 observations - were sufficiently large such that there was negligible variance arising from basing the normalisation on just a single random shuffling.

[3]Explanatory correlates of consciousness are brain activity patterns that can be directly mapped to phenomenological properties, characterizing all conscious experiences [202].

[4]The results did not vary qualitatively when other cortical target locations for the TMS stimulation were used [53].

based statistical procedure described here [151]. With this threshold, a binary response matrix $SS(x,t)$ ("significant sources") with 3000 rows (one for each source x) and 109 columns (one for each observation t, being $300ms$ sampled at $362.5Hz$) was filled, setting $SS(x,t) = 1$ if the activity of source x at time t was above the threshold T and $SS(x,t) = 0$ otherwise. Thus for one subject in one state and a chosen target region for the magnetic perturbation, one binary response matrix SS was obtained from approximately 100 trials (TMS stimulations), representing the brain response to the magnetic perturbation.

A Lempel-Ziv algorithm is now used to determine the number of new binary words in each column of the SS matrix, while keeping track of binary words found already in previous columns as described in Fig. 2.2. The resulting Lempel-Ziv complexity score, i.e. the length of the dictionary of binary words, denoted as $c(SS)$, depends on the size L of the SS matrix, L being the product of SS's number of rows and number of columns. The values of $c(SS)$ are large in practice and it strongly depends on the source entropy $H(L)$ of SS, defined as

$$H(L) = -p\log_2(p) - (1-p)\log_2(1-p) \tag{2.1}$$

where p the fraction of ones across all entries of the SS matrix. Thus $c(SS)$ is normalised by its asymptotic value for a random binary string of length L [126], i.e. as $L \mapsto \infty$, $c(SS) \mapsto LH(L)/\log_2(L)$. Thus the perturbational complexity index is defined as

$$PCI = \frac{c(SS)\log_2(L)}{LH(L)}. \tag{2.2}$$

2.2.2 Differentiation and integration

Fig. 2.3 shows Casali et al.'s illustration of the difference of the SS between wakeful rest and deep sleep (NREM). The large black frames contain the SS matrices for the same subject, one for each state of consciousness, rows being source channels and columns observations (dimensions of the SS matrices are 3000 sources times approx. 100 observations, representing $300ms$). The corresponding PCI values are 0.51 for wake and 0.23 for deep sleep, indicating higher algorithmic complexity for spatio-temporal response to TMS during wake than deep sleep, for this representative subject. These PCI differences are clearly in line with visual inspection of continuous response waveforms between the two states. For wake, the continuous waveforms show greater diversity across channels but also, for each channel, across observations. This diversity in the response to perturbation is a clear marker of differentiation, one of the theoretically proposed neural hallmarks of consciousness.

In order to argue for integration, note that the strength of the TMS stimulation in both states, WR and NREM, was the same, causing an electric field strength at the target of $90V/m$, but for deep sleep, far fewer sources showed an above-threshold response (Fig. 2.3A,B). This is seen as a sign that the effective connectivity between macroscopic brain regions is lower in deep sleep than wake. Effective connectivity in this sense can be seen as a measure of integration, the other of the theoretically proposed neural hallmarks of consciousness, proposedly reflecting the phenomenological property of "wholeness" of

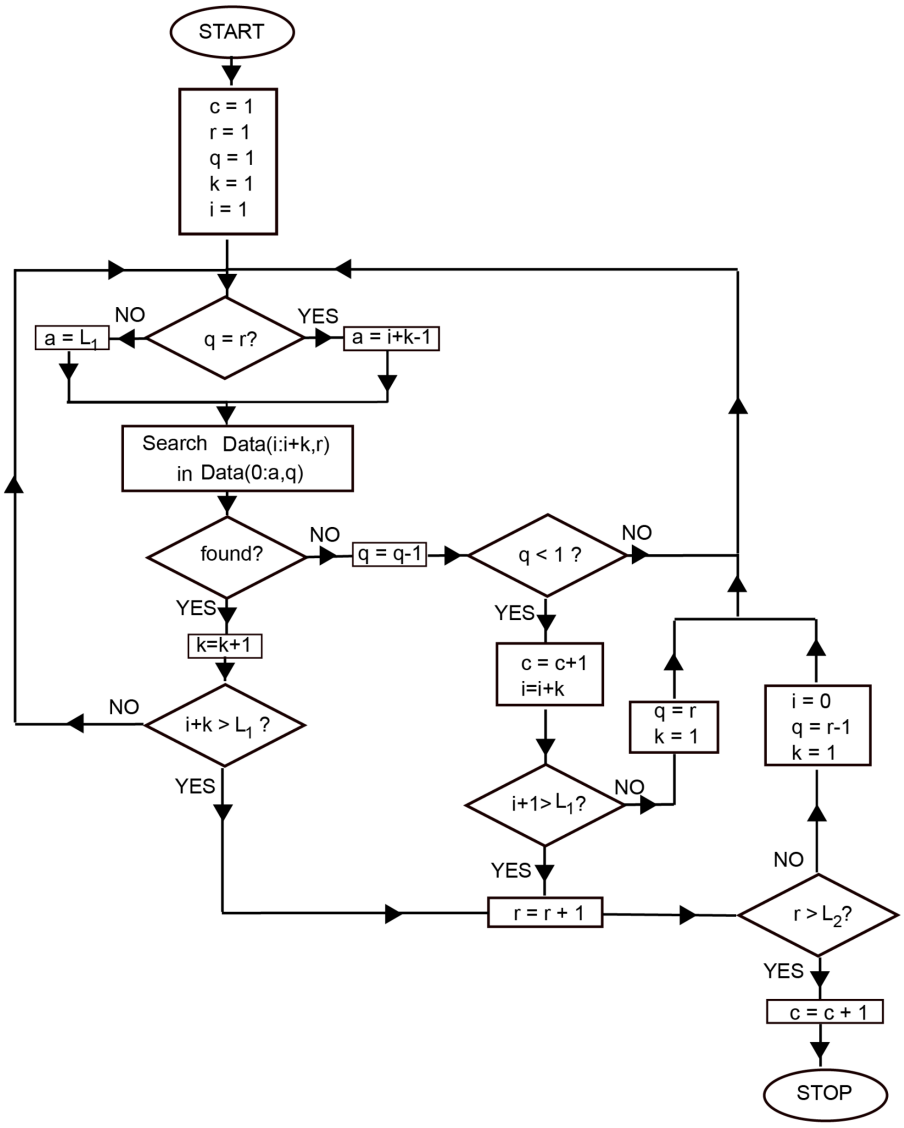

Figure 2.2: **Lempel-Ziv algorithm applied to SS matrix** Flow-diagram of the Lempel-Ziv algorithm (based on LZ-76) applied by Casali et al. to estimate the algorithmic complexity of the binary response matrix SS to transcanial magnetic perturbation. x and t are discrete indices of the spatial and temporal dimensions ($x = 1, ..., L_1$ and $t = 1, ..., L_2$), $l(t)$ is the number of spatiotemporal samples of $SS(x, t)$ up to time t. r, q, a are indices to keep track of location and length of found binary patterns. The input ("Data") is the binary SS matrix, having $L1$ rows (row i for the i^{th} source, $i = 0, .., L1$) and $L2$ columns (column j for the j^{th} observation, $j = 0, .., L2$). Patterns of length k are strings of k bits, $Data(i : k, j) = SS(i + 1, j)SS(i + 2, j)...SS(i + k, j)$, with $0 < i + k \leq L1$. The algorithm outputs the non-normalised Lempel-Ziv complexity c of the SS matrix, say $c(SS)$. Copied from Fig.S3 of [53], reprinted with permission from AAAS.

every experience, binding together features of each experience into a consciously perceived integrated whole.

Yet a high fraction of channels with significant response is not sufficient to indicate elevated effective connectivity and thus integration. In addition, diversity in the response

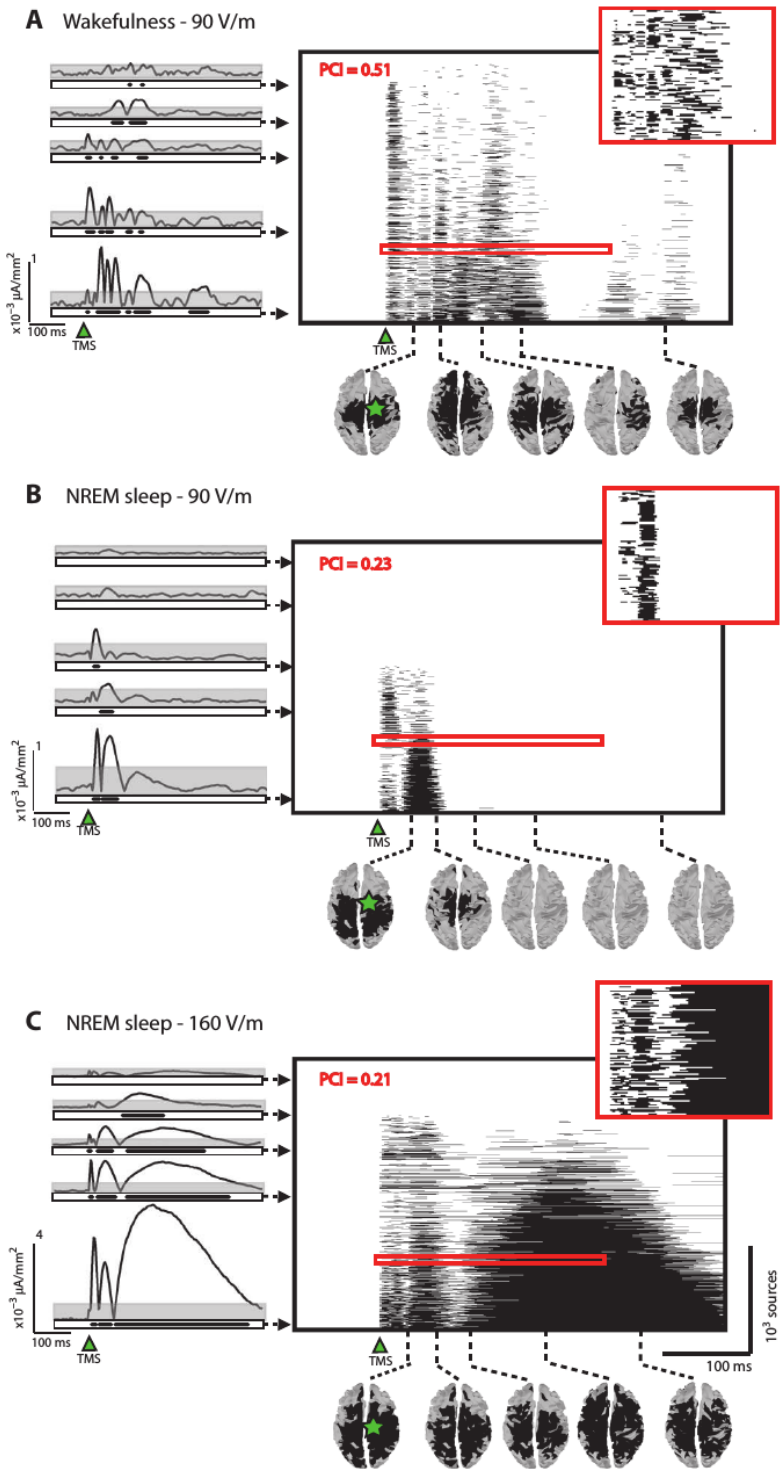

Figure 2.3: **PCI simultaneously measures integration and differentiation.** The binary matrix of significant sources $SS(x,t)$ for a subject stimulated with TMS during (A) alert wakefulness and (B) NREM sleep (stimulus at an intensity of 90 V/m) and (C) NREM sleep (stimulus at an intensity of 160 V/m. In each matrix, sources are sorted from bottom to top according to their total amount of significant activation during the post-stimulus period. The time series on the left of each SS matrix show the TMS-response activity for some representative sources, averaged over approximately 100 trials. The grey area indicates the binarisation threshold for each source activity. The cortical topographical maps below each SS matrix show the spatial extent of significant cortical activations at selected time points. Green star, site of TMS stimulation. Adapted from Fig.2 in [53], reprinted with permission from AAAS.

activity is needed. This can be seen in Fig. 2.3C, showing the binary response matrix (SS) for NREM sleep again but this time with the amplitude of the magnetic perturbation nearly doubled (inducing an electric field of $160V/m$ at target) and thus forcing significant response activity upon most channels. Yet PCI scores are still comparably low as for NREM with stimulation strength at $90V/m$, in line with its interpretation to measure integration

(effective connectivity) as the fraction of channels that show significant and diverse response activity. That PCI captures exactly this was further confirmed with computer simulations where PCI increases monotonically with the fraction of non-zero random channels in SS.

In summary, PCI captures integration by the fraction of channels with strong enough and diverse response (i.e. how "far" the response signal to TMS perturbation spreads across sources) and differentiation in terms of response signal diversity across sources and observations.

2.3 Differences between LZc and PCI

The LZc measure used here differs from Casali et al.'s PCI in several ways. PCI captures compressibility of the EEG response to TMS whereas our LZc is applied to spontaneous EEG. Reflecting this, there is a difference in obtaining the binarisation threshold. We used for LZc the mean of the absolute value of the analytic signal whereas Casali et al. used the 99th percentile of the amplitudes during prestimulus activity. PCI is computed across a $0.3sec$ data segment that starts directly after magnetic stimulation; our results for the EEG/propofol analysis were most robust for segment lengths of between $0.2sec$ and $20sec$, although segment lengths shorter than $1sec$ substantially affected the results for one subject only (Appendix A.2). Before computing PCI, the channels of the binary response matrix SS are ordered by activity, for LZc the activity of each channel was the same by construction. For LZc the binarised multidimensional activity is concatenated observation-by-observation and the LZ78 algorithm applied, whereas the LZ76 algorithm used for PCI Fig. 2.2 integrates the concatenation in the algorithm, not counting words that are split across two multidimensional observations (columns of SS). The LZ78 score is normalised in the case of LZc by the LZ78 score for the same input shuffled in time. Casali et al. normalise by the asymptotic upper bound: $LH(L)/log_2(L)$, where L is the total number of binary matrix entries and $H(L)$ the source entropy of this matrix (supp. mat. [53]). We also applied PCI's LZ76 algorithm and its normalisation to binary multidimensional spontaneous activity for our EEG/propofol data set and found similar results as for the computationally faster LZc.

2.4 Coalition entropy measures (ACE, SCE)

Besides LZc we explored alternative measures that capture spontaneous signal diversity across channels and across observations. Several measures of entropy for single channels have been previously applied to quantify temporal signal diversity and index global states of consciousness (Section 1.4.7), discretising continuous signals in various ways, for example by binning the activity into several equally sized bins. Here we introduce coalition entropy measures based on the entropy over time of activity or synchrony patters across multiple channels.

The idea of coalition entropy was introduced by Shanahan [209, 249]. In its original form it measures the entropy (over time) of the constitution of the set (coalition) of channels that are active, given a binarisation scheme for classifying channels as either active or

inactive. It was shown to peak for intermediate overall system synchronisation of coupled Kuramoto oscillators - i.e. where signal diversity over "channels" and observations was maximal - when applied to coalitions of synchronies of oscillator communities[5]. We call our version of this measure amplitude coalition entropy (ACE). Here we compute it using the same binarisation as described for LZc, i.e. taking for each channel the mean of the absolute value of the analytic signal as the threshold. As for LZc we normalise ACE by dividing the raw value by the value obtained for the same binary input shuffled (the upper bound), where shuffling means that the position of each digit was randomly changed. Note Shanahan's original version differs slightly from the version used here by utilising a fixed absolute binarisation threshold (which is not applicable to real EEG data, as activity varies widely across electrodes), and taking the asymptotic analytical upper bound as the normalisation (which is again not reached for the shape of data matrix analysed here, since the time-series are not long enough for all possible coalitions to be sampled). The entropy (over time) of the constitution of the set of channels that are active (assuming n channels in total) was implemented by mapping each of the possible $n \times 1$ dimensional binary observations to a distinct integer and then computing the entropy of the resulting sequence of integers.

We also introduce a new variant of coalition entropy, synchrony coalition entropy, denoted SCE. This measures the uncertainty, and hence diversity, over time of the constitution of the set of channels that are in synchrony - rather than active - schematically described in Fig. 2.4 and formally defined as follows:

For data $\boldsymbol{X}_t$, consisting of channels $X_{i,t}$, $i = 1, \ldots, n$, we consider two channels to be in synchrony at time t if the absolute value of the difference between their instantaneous Hilbert phases is less than 0.8 radians (approximately 45 degrees). Then we define coalition time-series $\boldsymbol{\Psi}_t^{(i)}$ by $\Psi_{j,t}^{(i)}$ taking the value 1 if channels i and j are synchronised at time t and taking the value 0 otherwise. The coalition entropy of $\boldsymbol{X}_t$ with respect to channel i is the entropy of $\boldsymbol{\Psi}_t^{(i)}$ (over time), normalised as a proportion of its maximum possible value N:

$$\mathrm{SCE}^{(i)} = -\frac{1}{N} \sum_{\boldsymbol{\psi}} p(\boldsymbol{\Psi}_t^{(i)} = \boldsymbol{\psi}) \log p(\boldsymbol{\Psi}_t^{(i)} = \boldsymbol{\psi}) \,. \tag{2.4}$$

The overall SCE is then the mean value of the $\mathrm{SCE}^{(i)}$ across channels. The upper bound SCE would arise from completely random coalition time-series in which each entry is 1

[5]Kuramoto oscillators are represented as unit length complex numbers whose phase changes with time (see Section 2.6.3 for details), e.g. the state of the k^{th} oscillator at time t is equal to $e^{i\theta_k(t)}$. The synchrony ϕ of a community (group) c of n oscillators at time t was defined as the modulus of the sum of all complex numbers divided by n:

$$\phi_c(t) = |\frac{1}{n} \sum_{k=1}^{n} e^{i\theta_k(t)}|. \tag{2.3}$$

For low global synchrony, defined as the the mean of $\phi_c(t)$ across all times and communities, there is also low synchrony within each community, thus their respective synchronies are mostly zero across communities and observations t, given the fixed binarisation threshold of 0.8. When global synchrony is high, most communities are synchronised thus $\phi_c(t) = 1$ for most communities and observations. In these both extreme cases, very low and very high global synchrony, the diversity of synchronies across communities and time is low, thus coalition entropy is low. Coalition entropy peaks for intermediate global synchrony, as here diversity across communities and observations is highest. See [209] for full details.

with probability 0.5. Such time-series are generated (with the same dimensions as those arising from the data) to obtain the normalisation factor N. Note that SCE does not score exactly 1 for shuffled input data - unlike ACE and LZc - as the probability at a give time-point of two shuffled channels being in synchrony is less than 0.5.

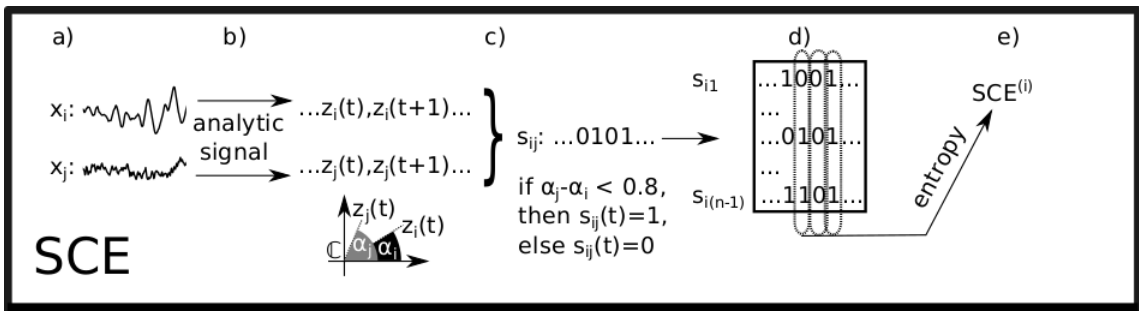

Figure 2.4: **Schematic of the computation of SCE** a) Two time series. b) The analytic signals of these two, which are complex signals with the real part being the original signal and the imaginary part being the Hilbert transform of the original signal. c) A binary synchrony time series is created for this pair of signals; a 1 indicates that the phases of the complex values of the analytic signals are similar (difference of less than 0.8 modulo 2π). d) Such time series are obtained to represent each channel's synchrony with seed channel i. e) $\text{SCE}^{(i)}$ is the entropy over observations in the resulting data matrix Ψ_i. The overall SCE is then the mean value of $\text{SCE}^{(i)}$ across choices of seed channel i.

2.5 'Phase shuffling' normalisation

As an alternative binarisation for the spontaneous signal diversity measures, a spectral-profile-preserving phase shuffling was applied to control for signal diversity changes due to spectral profile changes only. I.e. instead of normalising each signal diversity measure (LZc, LZs, ACE, SCE) for a single segment, using a score obtained from the same data segment but shuffled in time, an alternative normalisation was applied, using the measure's average score across "phase-shuffled" data segments, randomly picked from all available segments, each data segment individually randomised in a spectral-power-profile preserving way.

I.e. this 'phase shuffling' normalisation is obtained from phase-randomised surrogate data as follows. From the complete data from a given subject in a given state, a segment is randomly chosen. Each time series of that segment is expressed as a superposition of sinusoids using fast-Fourier transform. Then the phase of each sinusoid is independently randomly changed, before applying inverse Fourier transform. The signal diversity measure is computed for 100 such phase-randomised data segments. The mean of these 100 scores is then used to normalise the measure's score for the original data segments. Averaging over 100 such surrogate data segments suffices to obtain negligible variance in the normalisation factor.

We indicate that the 'phase shuffling' normalisation was used by attaching an N to the measure's name, i.e. LZc_N, LZs_N, ACE_N and SCE_N.

2.6 Basic models to illustrate ACE, SCE and LZc's behaviour

As a demonstration of how ACE, SCE and LZc indicate signal diversity across space (channels) and time (observations), and how their behaviour can differ, we apply them to computer generated time series before computing them on real data. First, we increasingly order random strings and show the measures decrease with increasing order. Secondly, we investigate the three measures' scores as a function of coupling strength and level of noise correlation in an autoregressive model, as an illustrative example dynamical process. Finally, we use a particular setting of a system of coupled Kuramoto oscillators to demonstrate that SCE can behave completely different to ACE and LZc. As another model, we analyse signal diversity extensively with coupled Stuart-Landau oscillators in Chapter 6.

2.6.1 Ordering random signals

Here we consider binary matrices, which represent the coalitions of either active (for LZc and ACE) or synchronous (for SCE) channels at each observation. We plotted the dependence of the measures against increasing order in this input matrix as follows. First a random binary matrix of 25 channels $\times$ 2500 observations was created (an entry being 1 with probability 0.5) and one by one each channel was replaced by a copy of the first channel. Thus at first the input matrix was fully random, then after one iteration there were two identical channels, at second iteration three identical channels (with random channel indices), until at the final iteration a matrix was obtained consisting of 25 identical channels. Coalition entropy and LZc all score 1 for the initial random matrix and decrease as the number of duplicate channels increases, approaching zero when all channels of the matrix are identical, see Fig. 2.5 (repeated 100 times with identical results). Thus, as expected, LZc and ACE increase monotonically as the level of diversity of the matrix of activations increases, and so does SCE with the matrix of synchronies[6].

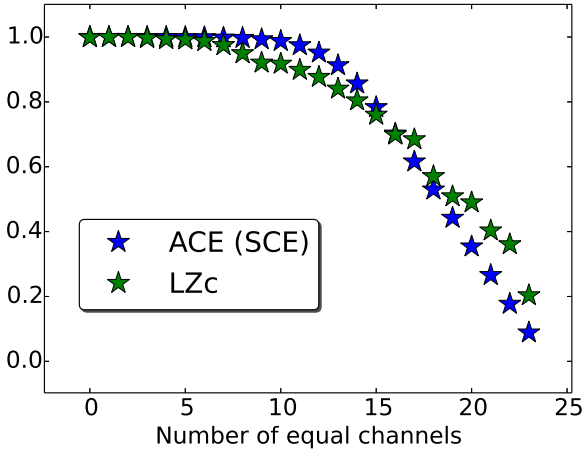

Figure 2.5: **signal diversity measures for increasingly regular activation/synchrony matrices.** Given a random binary matrix of activations (for ACE and LZc) or synchronies (for SCE), with increasing number of duplicated channels, the signal diversity measures monotonically decrease with the number of equal channels.

In the method just described, each channel remains fully random while the cross-channel variability is reduced. An alternative manipulation of diversity is to alter the regularity of

[6]See [126] for further basic bahavior of Lempel-Ziv complexity for coupled logistic maps and cellular automata.

each channel independently. We sorted increasingly long subsequences of each binary time series. That is, for each time series a subsequence of fixed length was chosen, starting at a different randomly chosen time in each time series. Then this subsequence of a given series is replaced by the same subsequence but ordered, placing all zeros to the left and all ones to the right within this subsequence. Now the length of the subsequence is increased stepwise, resulting in increasingly ordered binary matrices for which we computed the measures. We found the same trend as seen in Figure 2.5, with both LZc and coalition entropy (ACE and SCE) increasing as the level of diversity increases.

In sum, ACE, LZc and SCE are computed from binary matrices of activations (LZc, ACE) or synchronies (SCE). Each of the measures quantifies the level of diversity in such a matrix. We verified explicitly that if the matrix is a random matrix with some time-series identical, then all the measures decrease monotonically with the number of identical time-series (see Figure 2.5). LZc counts the number of distinct patterns in the concatenated observations, or equivalently computes how incompressible the matrix is, while ACE and SCE are based on the average surprise (entropy) of a single $n \times 1$ observation drawn at random from the matrix (n being the number of time series). Thus, the measures quantify the degree of diversity of a matrix in slightly different ways, diverging slightly for intermediate levels of diversity, but under normalisation, all tend to zero in the limit of an infinite matrix with every time-series the same, and all score one in the limit of an infinite matrix in which each entry is independent and has equal chance of being one and zero.

2.6.2 Autoregressive model

The previous analysis of increasingly equal channels allowed the conclusion that ACE, SCE and LZc increase with increasing diversity across channels, i.e. system differentiation. It was not possible though to interpret an independent property of the system as integration, which is why we investigate next how the spontaneous signal diversity measures LZc, ACE and SCE behave for an autoregressive (AR) system as a function of independently varied levels of coupling ("integration") and noise correlation ("differentiation").

We used the following autoregressive model:

$$\boldsymbol{X}_t = aA\boldsymbol{X}_{t-1} + \boldsymbol{\epsilon}(c)_t \tag{2.5}$$

The resulting multidimensional time series consists of N channels and M observations, and can be seen to describe the activity of a system of N nodes that influence each other according to a connectivity matrix A, whose entries are all multiplied by the coupling factor a. The multidimensional observation of the system at time t is an N-dimensional vector $\boldsymbol{X}_t$, initiated for $t = 0$ as standard normal random values. Any consecutive multidimensional observation of the system, $\boldsymbol{X}_t$ is equal to $A\boldsymbol{X}_{t-1} + \boldsymbol{\epsilon}(c)_t$. I.e. each node's activity at time t is a linear combination of the activity of all nodes at the previous time-step, as specified by the matrix multiplication $aA\boldsymbol{X}_{t-1}$. In addition, noise is added to each node at each time step, chosen from an N-dimensional standard normal distribution that has a uniform covariance matrix, all entries equal to c, except main diagonal entries that are always 1. This set-up allows the coupling strength between nodes to be regulated uniformly by

coupling factor a and, independently, the covariance of the added noise to be regulated by the noise correlation factor c.

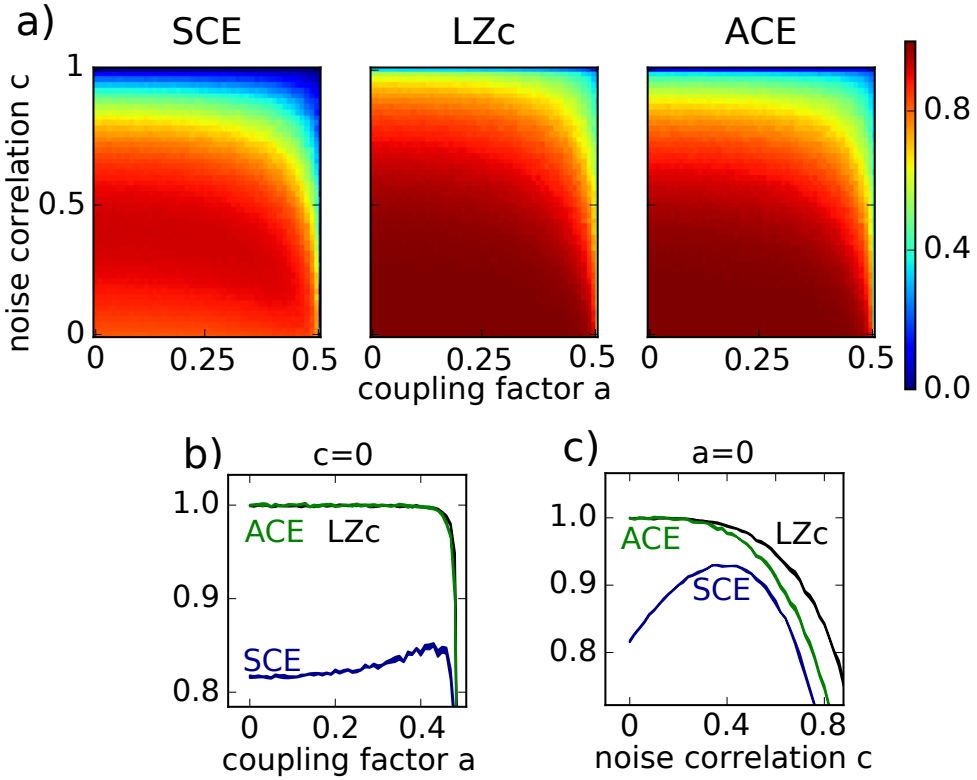

Figure 2.6: **SCE behaves differently to ACE and LZc for an AR system.** Equ. 2.5 was used to generate 10×2500 time series with varying noise correlation c and coupling strength a. a) The scores of ACE, SCE and LZc are displayed for combinations of c and a, showing that ACE and LZc decrease monotonically with increasing coupling and correlation while SCE has a weak local maximum in c and a. b) SCE peaks for $a = 0.44$ while ACE and LZc stay constant and then decrease, showing that SCE indicates an intermediate level of "integration". c) SCE peaks for $a = 0.44$, an intermediate value of noise correlation while LZc and ACE decrease monotonically with increasing noise. Line-width indicates standard error across trials.

We created system activity for $N = 10$ channels and 2500 observations (3000 were created of which the first 500 were discarded), with adjacency matrix $A = (a_{ij})$ chosen such that $a_{ij} = 1$ if $j = i - 1$ or $j = i + 1$, else $a_{ij} = 0$, with circular boundary conditions, i.e. $1 - 1 = N$ and $N + 1 = 1$, $i, j = 1, ..., N$. A topological interpretation for A is that the nodes sit homogeneously distributed on a circle and only nearest neighbours are bidirectionally connected, i.e. a ring lattice. We varied the noise correlation c in 50 steps from 0 to 1. When $c = 0$ the noise is completely uncorrelated, whereas $c = 1$ results in fully correlated noise across channels. The coupling factor a was varied in 50 steps from 0 to 0.5. For $a = 0$ the system's activity at time t equals $\boldsymbol{\epsilon}(c)_t$, i.e. temporally un-correlated but spatially (across nodes) correlated noise. For $a > 0.5$ the system becomes unstable and activity of each channel grows to infinity[7].

[7]For $a = 0.5$ the time series becomes a unit root process, i.e. a random walk which has still finite mean

The measures ACE, LZc and SCE were computed repeatedly (10 trials) for system activity of 10 channels and 2500 observations and each combination of a and c. Their scores as a function of a and c are shown colour-coded in Fig. 2.6a. While LZc and ACE quite similarly and monotonically decrease for increasing noise correlation c and increasing coupling a, SCE shows a weak local maximum for both, a and c. This behaviour is further illustrated in Fig. 2.6b, where the noise is completely uncorrelated ($c = 0$) and SCE increases with increasing coupling a before it decreases rapidly, while ACE and LZc nearly identically stay maximal before decreasing rapidly for very high coupling. Fig. 2.6c further shows, setting $a = 0$, that SCE peaks for an intermediate value of noise correlation c while ACE and LZc monotonically decrease with increasing noise correlation. This difference in the behaviour of SCE and ACE/LZc was also seen for other fixed values of coupling and correlation, respectively, as can be seen in Fig. 2.6a. Similar yet even more pronounced behaviour was found for a system of only $N = 4$ channels (not shown).

In conclusion, LZc and ACE both similarly and monotonically decreased with increasing noise correlation and coupling for this autoregressive model, yet not completely identically. ACE and LZc thus clearly measure signal diversity only, which may be called differentiation of the dynamical system. SCE by contrast showed local maxima for both, noise correlation and coupling, indicating that it does capture a balance of integration and differentiation for this model system, however weakly.

2.6.3 Kuramoto model

We computed the three measures for simulated continuous data created with a Kuramoto model, a model known to display rich dynamics [39, 209]. We used the same parameters as Shanahan [209] (with a small exception, see below). That is, we simulated 8 communities of 32 oscillators, with the activity of each oscillator being a unit complex number with phase θ. For the i^{th} oscillator the evolution of its phase in time is given by:

$$\frac{d\theta_i}{dt} = 1 + \frac{1}{N+1}\sum_{j=1}^{N} K_{i,j}\sin\left(\theta_j - \theta_i - \alpha_{i,j}\right). \tag{2.6}$$

The matrix K of coupling strengths was such that each oscillator is fully connected to its own community and has 32 random connections to oscillators in other communities. N denotes the number of connections per oscillator. Inter-community coupling strengths were set to 0.4, and intra-community coupling strengths to 0.6. As shown in the equation, all natural frequencies are 1. An Euler method step size of 0.05 was used and 1500 time steps were generated, starting from random initial phases. The first 500 time steps were discarded. The phase lag parameter $\beta_{i,j} = \frac{\pi}{2} - \alpha_{i,j}$ was varied uniformly for inter-community connections, and held constant at 0.15 for intra-community connections. (Keeping the intra phase lag fixed is the only difference from the parameters in [209]). A small value of β corresponds to a large phase lag and leads to little synchrony whereas larger values of β correspond to smaller phase lags and leads to greater inter-community synchrony. A single time-series was generated for each community by taking the real part of the aver-

but possibly infinite variance.

age instantaneous activity over all of its oscillators. Thus the activity $X_c(t)$ of the c^{th} community C_c is given by

$$X_c(t) = \mathrm{Re}\left(\frac{1}{32}\sum_{j\in C_c} e^{i\theta_j(t)}\right). \tag{2.7}$$

Figure 2.7 shows the results for the Kuramoto simulations, plotting LZc, ACE and SCE against the phase-lag parameter β_{ext}. In addition a measure of phase synchrony is plotted (PhaseSync), being the mean of $\Psi_t^{(i)}$ over all observations t and channels i, as described in the definition of SCE. This plot illustrates (i) the similarity of LZc and ACE, which are both based on the diversity of the dynamics of amplitude fluctuations, and (ii) that SCE is a distinct measure of signal diversity to LZc and ACE, by virtue of being based on diversity in synchrony patterns. For all values of β_{ext}, fluctuations in the amplitude of each community propagate to different communities, and this leads to diversity in relative amplitude reflected in the high values of LZc and ACE. In contrast, SCE is (i) low for small values of β_{ext} (large phase lag) since there is little synchrony between communities, (ii) low for large values of β_{ext} (small phase lag) since there is almost total synchrony between communities, (iii) high for intermediate values of β_{ext} in which synchrony between communities is able to fluctuate. Also note that LZc and ACE have a minimum at $\beta_{\text{ext}} = 0.15$ whereas SCE has a maximum there.

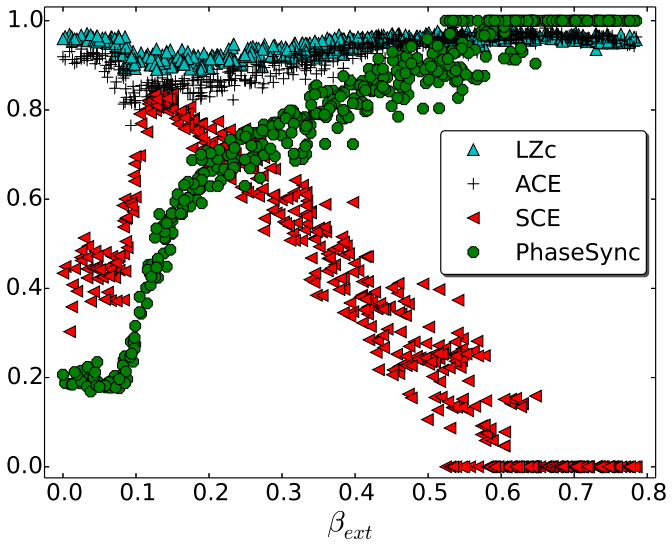

Figure 2.7: **LZc, ACE and SCE for Kuramoto model.** The data were obtained by varying phase-lags for inter-community interactions. A small value of β_{ext} corresponds to a large phase lag and leads to little inter-community synchrony (as measured by phase synchrony, PhaseSync - indicated as green discs, see text for its computation) whereas larger values of β_{ext} correspond to smaller phase lags and lead to greater inter-community synchrony. LZc and ACE show similar dependence on β_{ext} whereas SCE peaks where the former two have a minimum.

In order to assess perturbational signal diversity in the Kuramoto model, one would have to perturb the system. Each oscillator in the Kuramoto model has fixed unit ampli-

tude and is entirely characterised by phase alone. It is thus not possible to perturb the system stronger than setting all phases equal, which did not result in interesting response activity when averaging over many perturbations. To overcome this we explored a system of coupled Stuart-Landau oscillators, a generalisation of the Kuramoto model, charaterising each oscillator by time-varying phase and in addition time-varying amplitude. This allows a much stronger perturbation of the system by increasing the amplitude of each oscillator substantially. It further has the advantage that a complex (in the sense of having real and imaginary part) time series of an oscillator can be directly interpreted as the analytic signal of an EEG channel. We therefore extended our explorations of signal diversity in coupled oscillator systems to the Stuart-Landau system as will be presented in Chapter 6. Before that, Chapters 3 to 5 cover the behaviour of spontaneous signal diversity measures for empirical data of different global states of consciousness.

Chapter 3

Signal diversity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia

3.1 Abstract

Emerging neural theories of consciousness suggest a correlation between a specific type of neural dynamical complexity and the level of consciousness: When awake and aware, causal interactions between brain regions are both integrated (all regions are to a certain extent connected) and differentiated (there is inhomogeneity and variety in the interactions). In support of this, recent work by Casali et al. (2013) has shown that Lempel-Ziv complexity correlates strongly with conscious level, when computed on the EEG response to transcranial magnetic stimulation. Here we investigated complexity of spontaneous high-density EEG data during propofol-induced general anaesthesia. We consider three distinct measures: (i) Lempel-Ziv complexity, which is derived from how compressible the data are; (ii) amplitude coalition entropy, which measures the variability in the constitution of the set of active channels; and (iii) the novel synchrony coalition entropy (SCE), which measures the variability in the constitution of the set of synchronous channels. We show that these three distinct 'flavours' of signal diversity robustly decrease in spontaneous EEG during general anaesthesia.

3.2 Introduction

The first set of global states of consciousness we explore with the spontaneous signal diversity measures in this thesis are propofol-induced anaesthesia, propofol-induced mild sedation and wakeful rest. Further global states analysed in later chapters are sleep stages (Chapter 4) and psychedelic states (Chapter 5).

As introduced in detail in Chapter 1, the idea that the level and range of consciousness relates in some way to dynamical complexity of brain activity is becoming increasingly prominent [11, 228, 229, 231, 88, 139, 102]. A common way to conceptualise dynamical

complexity in this setting is as simultaneous *differentiation* (subsets of the system being dynamically distinct) and *integration* (the system as a whole exhibiting coherence), and this idea draws from what is taken to be a fundamental property of conscious experience, namely that each conscious scene is composed of many different parts and is different from every other conscious scene (differentiation), yet each conscious scene is experienced as a coherent whole (integration) [228, 229, 231, 88]. A number of different measures of neural dynamical complexity have been proposed based on information sharing and transfer [228, 231, 230, 199, 200, 15, 20, 204, 166]. Properties of these measures have been explored on simple models, for instance neural network activity of artificial agents [199, 89]. However, these measures, based on information sharing and transfer, rely on restrictive assumptions, such as stationarity and linearity, that limit the conclusions that can be drawn when applied to real brain data [22, 59].

A series of recent studies, investigating perturbational signal diversity, take a more pragmatic approach to investigating the relationship between consciousness and complexity, as we presented in Sections 1.4.9 and 2.2. PCI values obtained for conscious subjects were consistently higher than for unconscious subjects, to the extent that a single classifier threshold could be applied: when the PCI value was above the threshold the subject was always conscious and when the PCI value was below the threshold the subject was always unconscious [53].

On spontaneous steady-state EEG data, several measures of signal diversity have been computed on single time-series, reflecting local signal diversity over time rather than differentiation and/or integration across a network, see Section 1.4.7. Briefly, these measures include various forms of spectral entropy [100, 244, 125] and again Lempel-Ziv complexity [42, 43, 188, 253, 16, 46, 108, 124, 250] and all of them have a tendency to decrease during general anaesthesia . Recently these measures have also been applied to auditory evoked potentials in disorders of consciousness patients [214], and values were found to correlate with behaviourally-diagnosed level of consciousness, although there was no single cross-subject threshold for classifying subjects as conscious or unconscious.

Given that previous spontaneous signal diversity measures captured temporal signal diversity only, we contributed for the first time multi-dimensional spontaneous signal diversity measures (LZc, SCE and ACE), capturing spontaneous signal diversity across channels and across observations. Here we investigated LZc, SCE and ACE as defined in Chapter 2 on multi-dimensional spontaneous EEG data from subjects undergoing propofol-induced general anaesthesia. For spontaneous EEG signals, LZc strictly only reflects differentiation (and not integration); it computes diversity in patterns of activity in both space and time. ACE is similar to Lempel-Ziv complexity, in the sense that it quantifies variability in space and time of the activity. By contrast, SCE is conceptually different because it quantifies variability in the relationships between pairs of channels, as was illustrated with SCE's non-monotonic behaviour for increasing diversity in time series of an autoregressive process, see Section 2.6.2.

We computed Lempel-Ziv complexity and the coalition entropy measures for sets of equally spaced channels across the whole scalp and also for sets of channels restricted

respectively to the frontal, parietal, temporal and occipital lobes, on full broadband signals and on frequency-restricted signals, in each case comparing results for data from wakeful rest, mild sedation and general anaesthesia. We contrasted these measures' ability to indicate conscious level on these data, with that of control measures not based on signal signal diversity, including normalised delta power [158, 217, 214].

3.3 Methods

3.3.1 Ethics statement

The data analysed in this study were obtained from a previous study [158] with procedures approved by the Ethics Committee of the Faculty of Medicine of the University of Liège.

3.3.2 EEG data acquisition and preprocessing

Spontaneous high-density EEG recordings (256 electrodes, EGI [90]) were re-analysed from 7 healthy subjects, sampled at $1000Hz$, before, during and after propofol-induced general anaesthesia. Propofol is an intravenous anaesthetic that is widely used in surgical settings and which reversibly induces a state of diminished responsiveness behaviourally similar to non-rapid eye movement sleep [148]. States of consciousness were defined behaviourally using the Ramsay scale [183]. The 4 different states labelled here are: wakeful rest (WR), mild sedation (MS; slower response to command, Ramsay scale score 3), loss of consciousness (LOC) with clinical unconsciousness (no response to command, Ramsay scale score 5), and recovery of consciousness (wakeful rest after propofol, WRa) [158]. Propofol was administered as described in [158]: 'A simple constant rate infusion of propofol was used together with the computerised Marsh model to predict when to manually adjust infusion rate to maintain predicted steady-state propofol levels, although the main goal was to achieve the range of clinical states (Ramsey scores).' Average arterial blood concentrations of propofol were $1.91 \pm 0.52mcg/mL$ for MS and $3.87 \pm 1.39mcg/mL$ for LOC [158]. 20 minutes of recording during each of these 4 states were obtained for each subject.

The data were pre-processed as follows. First epochs of recording and complete channels that showed obvious artefacts were rejected by visual inspection. Artefacted channels were identified by their extreme amplitudes and irregular behaviour throughout the complete recording and entirely removed. Artefacted epochs, displaying abnormally high amplitudes for several seconds across all channels (typical for muscle movement), were manually excised and the remaining data concatenated. Then $50Hz$ and $100Hz$ frequency noise (artefacts from electricity mains) were removed by Butterworth notch filtering. Next the data were down-sampled from $1000Hz$ to $250Hz$ and spatially filtered by computing the surface Laplacian.

Surface Laplacian is a method for spatial filtering of EEG data, performed in order to increase topographical specificity (i.e. to make each channel's activity more directly indicative of brain activity under the channel's electrode). It reduces the effect of volume-conduction (electrical fields tangentially conducted along the skull) so the filtered signal reflects more closely the local brain activity - radial dipoles in gyral crowns [65]. For

each electrode, the surface Laplacian can be implemented as a subtraction of a weighted sum of its nearest neighbours' electrical activity. This filtering method was applied using MATLAB with BCIlab and EEGlab [82].

Finally linear de-trending and baseline subtraction was performed for each channel of each segment. After preprocessing the length of the time series varied per subject and per condition between 9-14min, i.e. approximately half the length of the raw data. Figure 3.1 illustrates example 10sec segments of EEG data for WR and for LOC. Analyses were performed using such non-overlapping $10sec$ segments for a total number of on average 60 segments of EEG recording per subject and per condition.

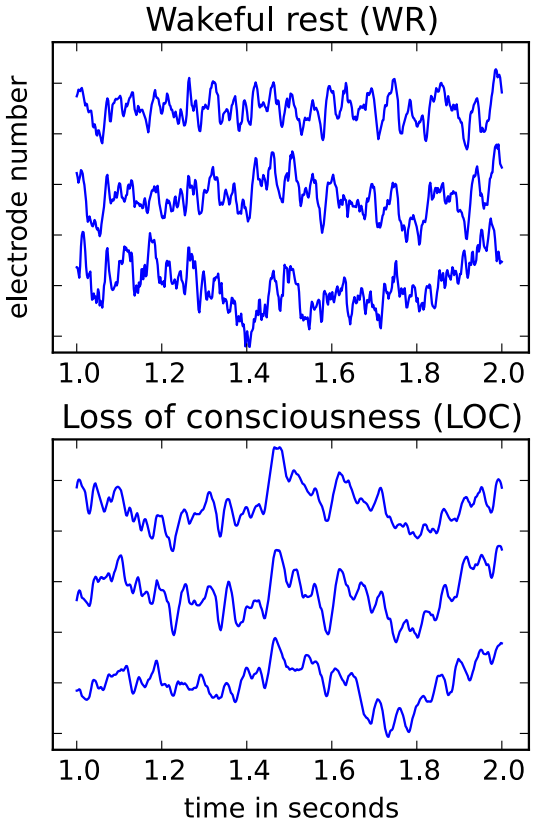

Figure 3.1: **Two 2sec EEG segments from 3 example channels.** The segment in the top panel is during wakeful rest (WR) and the segment in the bottom panel is during propofol-induced loss of consciousness (LOC); both segments are shown after pre-processing and in addition normalised by standard deviation to highlight spectral changes (for more data details, see [22, 158]). The recordings for LOC display visibly stronger slow waves (low-frequency components) as compared to those for WR.

3.3.3 EEG channel selection

Analyses were performed on automatically selected electrodes, first 25 taken from across the whole cortex, see Figure 3.2a, and secondly 25 taken exclusively from either frontal, parietal, temporal or occipital cortex, Figure 3.2b. A k-medoids clustering algorithm [120] was implemented in Python to automatically select electrodes spatially uniformly distributed over a particular region (either the whole cortex or a specific lobe).

The k-medoids cluster algorithm takes two arguments as input, first a set of 3D coordinates of points in space (i.e., the EEG electrode coordinates) and second the number of desired clusters. It outputs labels for each point, indicating to which cluster it belongs and also if it is this cluster's representative. The cluster representative is the point in a cluster which has the least mean Euclidean distance to all other points of that cluster. The algorithm finds these labels by initialising cluster representatives as randomly chosen points of

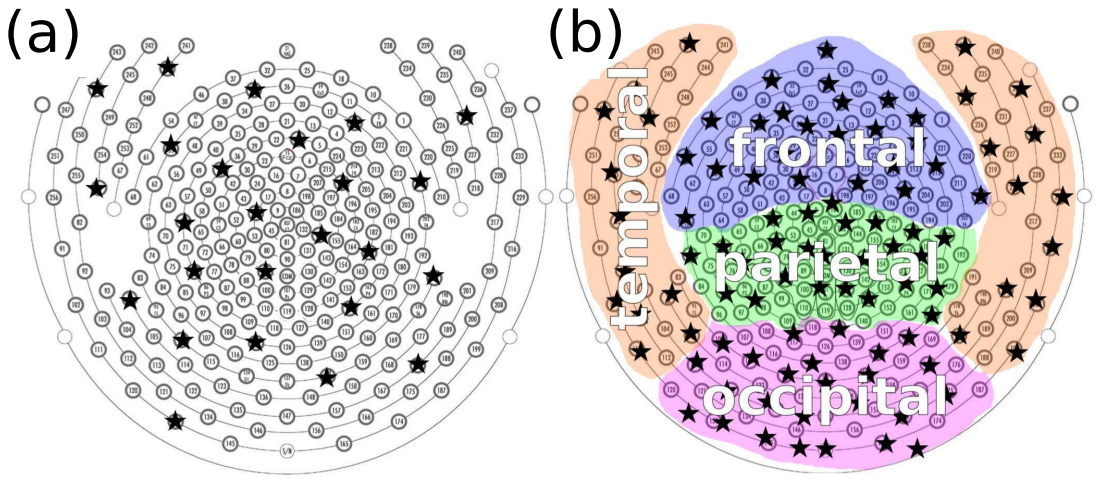

Figure 3.2: **Channel selection from the high density (256 electrodes) EEG.** The electrode layout is from manufacturer EGI [90]. The homogeneously distributed channel selection is shown for: a) analyses across the whole scalp; b) analyses restricted to a certain lobe. Chosen channels are indicated as black stars. See text for details.

the set, as many as there are desired clusters. Then it iteratively adjusts labels until all points are divided into approximately equally large regional clusters. The representatives of the clusters are then approximately homogeneously distributed across the cloud of all points in space, irrespective of the cloud's shape.

3.3.4 Statistics

For within subject comparison of states, we considered differences in the scores of a measure to be substantial if the effect size as measured by Cohen's d was greater than 0.8 (a threshold indicating large effect size [220]). For a given subject, measure and state pair, we computed Cohen's d as the difference of the mean scores across segments for each state respectively, divided by the pooled standard deviation.

In addition we compared for each measure and state pair the mean scores across subjects. Given the independence of different subjects, we applied a Wilcoxon rank sum test for each measure and state pair, corrected for false discovery rate (FDR), using the Benjamini-Hochberg procedure [29].

3.4 Results

3.4.1 Broadband signal for whole cortex

Figure 3.3 shows the mean values across $10sec$ segments of ACE, SCE and LZc during wakeful rest before sedation (WR), mild sedation (MS), LOC and shuffled data (in time domain shuffled WR) for each subject, computed for 25 EEG channels automatically selected to be spread evenly via k-medoids clustering across the whole cortex. For all subjects, the three measures ACE, SCE and LZc score higher for WR than for LOC, nearly all with high effect

size (Cohen's $d > 0.8$ for all measures and subjects except subject 1 for all measures and subject 3 for measure SCE, compare with Table 3.1.) Values for MS typically lie between those for WR or LOC, but the differences between MS vs. WR or LOC are less consistent than those between WR vs. LOC. For all three measures a single threshold can be drawn that separates WR from LOC across all subjects, as indicated by the cyan lines.

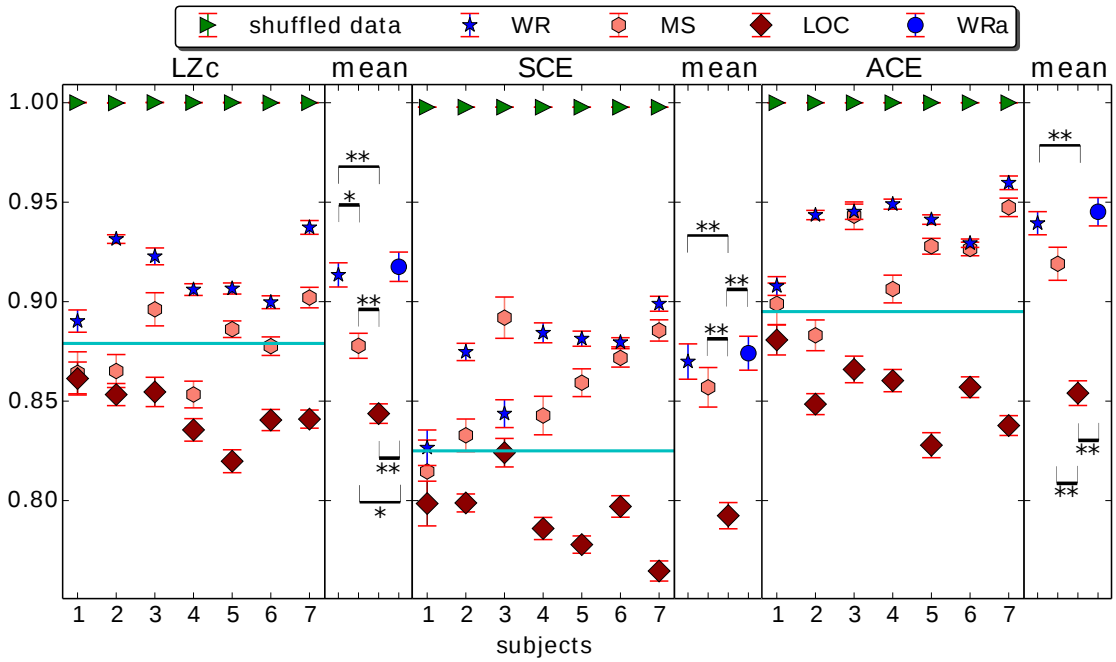

Figure 3.3: **LZc, SCE and ACE computed as averages over multiple 10sec segments of EEG of the 7 subjects before and during anaesthesia.** States shown are wakeful rest (WR) before propofol, mild sedation (MS), LOC and wakeful rest emerging from propofol sedation (WRa, not shown for single subject results). Measures are computed across 25 channels spread evenly across the whole cortex. The measures score highest for shuffled WR data, and consistently across subjects higher for WR as opposed to LOC. Error bars indicate standard error across segments, cyan horizontal lines are example thresholds for each of the measures, separating WR from LOC for all 7 subjects. For each single subject plot, the mean and standard error across its 7 values per state is displayed in the narrow plot to its right, with the title 'mean'. For these mean values across subjects, significant differences between state pairs are shown by a double asterisk if $p < 0.01$ and a single asterisk if $p < 0.05$ (Wilcoxon rank sum test, FDR corrected for multiple comparison). See Table 3.1 for effect size comparison.

Figure 3.3 further displays the mean scores of each measure - LZc, SCE and ACE - across the 7 subjects. By these mean scores, all three measures score higher for waking and mild sedation states (WR, WRa, Ms), than for LOC ($p < 0.01$, Wilcoxon rank sum test corrected for false-discovery rate using the Benjamini-Hochberg method). Also all three measures score higher for WR and WRa than for MS, though only significantly so for LZc ($p < 0.05$, see Figure 3.3).

The scores for all three states of consciousness and all three measures lie within 25% of the scores obtained for random input, see Figure 3.3. This proximity of the measures' scores to 1 for WR (as well as LOC) contrasts with the pattern of the overall lower PCI scores

	γ	β	α	θ	δ	sumCov	LZc	SCE	ACE
WR/WRa	1 4 2	0 1 6	4 3 0	6 1 0	2 4 1	0 7 0	0 6 1	1 6 0	0 6 1
WR/MS	4 2 1	0 0 7	6 1 0	7 0 0	2 2 3	0 3 4	4 3 0	1 6 0	2 5 0
WR/LOC	6 0 1	2 1 4	3 0 4	5 1 1	0 1 6	0 1 6	6 1 0	5 2 0	6 1 0
WRa/MS	5 2 0	0 2 5	3 3 1	5 2 0	1 3 3	0 5 2	2 5 0	1 6 0	2 5 0
WRa/LOC	6 0 1	4 1 2	1 1 5	4 1 2	0 1 6	0 2 5	6 1 0	6 1 0	6 1 0
MS/LOC	6 0 1	6 0 1	1 1 5	4 0 3	0 1 6	0 7 0	3 4 0	5 2 0	5 2 0

Table 3.1: **Effect size comparison per measure and state pair.**For each measure and state pair, the three numbers correspond to how many subjects out of 7 had higher score for the left state with Cohen's $d > 0.8$ (left digit), no substantial difference, $d < 0.8$, (middle) and higher score for the right state with Cohen's $d > 0.8$ (right). The results were obtained from applying the measures to the broadband signal from 25 k-medoids chosen electrodes from the whole cortex. Here WRa is wakeful rest emerging from propofol sedation.

presented by Casali et al. for EEG input from propofol anaesthetised subjects (Figure 4a in [53]). There are however several differences between PCI and our measures, see Section 2.3. Notably, PCI is computed after ordering the channels according to their overall level of activation following the TMS stimulation. This ordering reduces signal diversity. There is no analogous step in the computation of LZc, ACE and SCE. By design, all channels are equally 'active' according to the thresholding procedure we use on the spontaneous data.

We compared the ability of the signal diversity measures to discriminate the different states of consciousness with that of normalised spectral power bands and a simple correlation measure, sumCov, which equals the mean of the absolute values of correlation coefficients between all channels[1]. Normalised spectral power bands were obtained by fast Fourier transform of each channel's $10sec$ time series, then averaged over the 25 channels and grouped into frequency intervals, normalised such that the summed power across all bands equals one. Frequency bands are defined, following convention, as δ =1-4Hz, θ=4-8Hz, α =8-13Hz, β =13-30Hz, γ =30-70Hz. All results are summarised in Table 3.1. The compared states were WR, MS, WRa (wakeful rest after propofol) and LOC.

We also assessed, via ROC curve analysis, the extent to which classification thresholds can be applied to each measure across subjects to discriminate between state pairs (e.g. WR versus LOC). Given the set of mean scores for each subject for a pair of states, one computes the ROC curve for discriminating between the states by plotting for each possible classification threshold the hit rate (y-axis) versus the false alarm rate (x-axis) when classifying a mean score as indicating the more conscious state when it exceeds the threshold. For example, in a comparison of LZc between WR and LOC, for a given threshold C, the hit rate is the proportion of subjects for which the mean LZc is greater than C during WR, and the false alarm rate is the proportion of subjects for which the

[1]We further compare LZsum and LZc, which showed similar behaviour, however, mild sedation and wakeful rest was more clearly discriminated by LZc as opposed to LZsum (compare Fig. A.4 with Fig. 3.3). This showed that temporal signal diversity alone was similarly indicative of conscious state a signal diversity across channels and observations, which is a novel contribution.

mean LZc is greater than C during LOC. An area under the ROC curve (AROC) of 1 indicates that there exists a classification threshold for which for all 7 subjects the mean score exceeds the threshold for the more conscious state and is sub-threshold for the less conscious state. An AROC of 0 also indicates perfect discriminability between the states, but with the measure scoring higher for the less conscious state, i.e. there exists a threshold for which for all 7 subjects the mean score exceeds the threshold for the less conscious state and is sub-threshold for the more conscious state. An AROC of 0.5 indicates that the measure has no ability to discriminate between the states; for each threshold, the hit rate equals the false alarm rate. For an interactive applet that illustrates ROC analysis see [119].

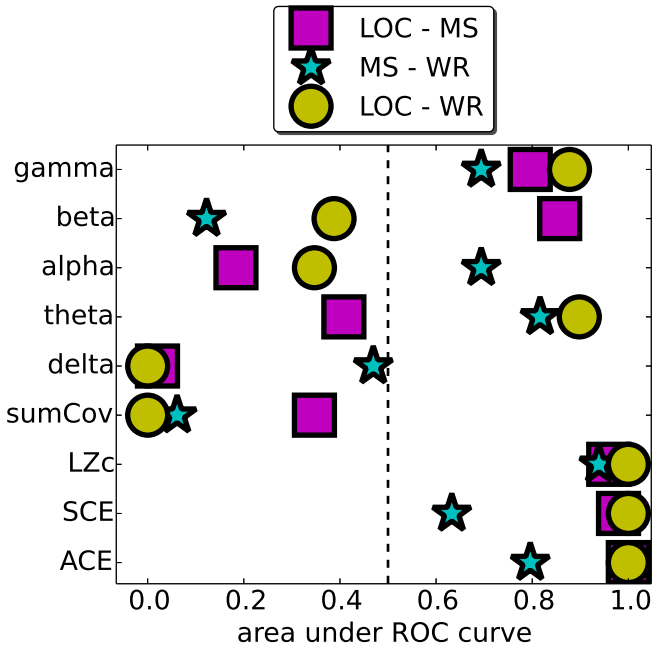

Figure 3.4: **Discriminative power across subjects of nine measures, as measured by area under the ROC curve (AROC).** Each symbol represents a state pair as indicated. The AROC is computed from the mean scores for the 7 subjects, obtained from the broadband signal from 25 electrodes from the whole cortex. The closer the AROC is to 0 or 1, the better the measure is at discriminating the given state-pair, close to 1 signifying that the measure tends to be greater for the more conscious state and close to 0 signifying that the measure tends to be greater for the less conscious state. When the AROC is 0.5 there is no discriminative power; hit rate equals false alarm rate for all classification thresholds. LZc, ACE and SCE have nearly maximal discriminative power for state pairs LOC/MS and LOC/WR. The measure sumCov fails to discriminate LOC/MS yet has strong (inverse) discriminative power for LOC/WR and MS/WR. Normalised delta band power discriminates LOC/WR and LOC/MS strongly yet MS/WR poorly.

As shown in Figure 3.4, LZc, SCE and ACE give nearly perfect cross-subject discrimination between LOC/WR and between LOC/MS, scoring lower for LOC than any other state. Discrimination was weaker for MS/WR. In addition, normalised delta power and sumCov give near perfect cross-subject discrimination between LOC/WR (delta also for LOC/MS, and sumCov also for MS/WR), scoring low for WR and highest for LOC. Delta

band power strongly discriminating LOC/WR (see also Appendix A.4 and Table 3.1) is in line with previous studies [158, 217, 214], reflecting the presence of slow-waves during LOC [217].

3.4.2 Controlling for changes in power spectrum, effects of number of electrodes and segment length

Importantly, to check whether changes in ACE, LZc and SCE reflected more than mere changes to the power spectrum, we computed the measures normalised by their values for phase-randomised surrogate data. Phase shuffling was performed by Fourier transforming, and then applying inverse Fourier transform with the addition of an independent random phase to each channel. When repeating the analysis displayed in Figure 3.3 no major changes in the measures' behaviour were found. (See Appendix A.3.)

Next, we tested dependence of the results on the number of electrodes across which the measures were computed. Different numbers of electrodes were chosen equally distributed across the whole cortex via k-medoids. We tested 5, 10, 50, and 100 electrodes in addition to the 25 considered above, for the measures LZc, SCE and ACE, all computed from $10sec$ segments. The results are all broadly the same as for 25 electrodes (see Appendix A.1).

As a final control, the role of segment length was explored. Analysed across segments of length other than $10sec$, we found for LZc, SCE and ACE broadly identical results in almost all cases for all tested segment lengths (0.2, 0.4, 0.8, 1.2, 2, 4, 6, 10 and 20 seconds). There was just one subject, subject 1, for whom LZc, SCE and ACE were only significantly greater in WR than LOC for segments of length $2sec$ or more. We conclude that the discriminative power of the measures is robust across a range of segment lengths (see Appendix A.2).

3.4.3 Correlation between changes in different measures

We computed correlations between all pairs of measures. The ratio of the score for WR and LOC was obtained for each subject and measure and then used to compute the Pearson correlation coefficient across subjects, as well as a 2-tailed p-value for that correlation. In other words, for each pair of measures we considered the correlation of measure1(LOC)/measure1(WR) and measure2(LOC)/measure2(WR) across subjects. For the state pair WR/LOC all correlations with $p > 0.05$ were ignored and the remainder are plotted in Figure 3.5, ordered by magnitude of the correlation. All pairs chosen from the measures ACE, LZc, SCE correlated significantly, with Pearson correlation coefficient (r) greater than 0.75 and $p < 0.05$. Beta power changes correlated strongly with changes to ACE, even though beta sometimes increases and sometimes decreases during LOC (see Table 3.1). All significant correlations were positive except between alpha and gamma power, which is flipped for convenience in Figure 3.5 but indicated in red. Interestingly, change in delta power does not correlate significantly with changes in any of the other 8 measures, confirming again that changes in signal diversity do not simply reflect changes to the overall power spectrum.

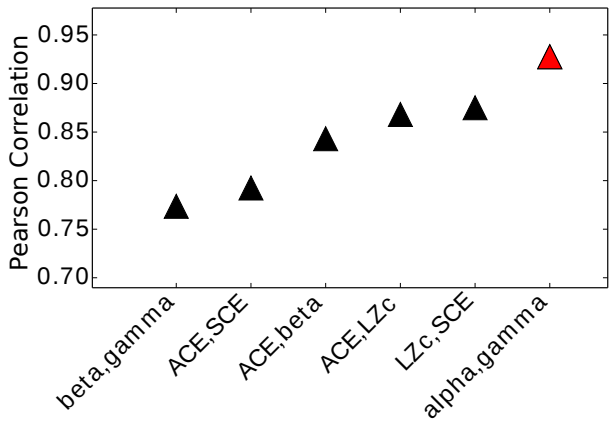

Figure 3.5: **Pearson correlations between measures.** Each black triangle indicates the positive correlation for the given measure pair listed in increasing order along the x-axis. The correlation was computed across subjects for the ratio of WR/LOC. Each correlation of all possible pairs of 9 measures was tested for significance and only those with 2-tailed p-value smaller than 0.05 are plotted. The red triangle indicates negative correlation. All 3 pairs out of ACE, LZc, SCE correlate significantly. There is no significant correlation of delta power with any other of the 9 measures for the state pair WR/LOC.

3.4.4 Grouping EEG into lobes

To assess whether changes in signal diversity during LOC occurred across the whole cortex, or whether more local changes could be identified, we analysed the measures on each of the four lobes separately. The electrodes were divided into four groups, each corresponding to one of the four main anatomical lobes: temporal, occipital, frontal and parietal, see Figure 3.2. Within each lobe, 25 channels were automatically chosen via k-medoids and from them the measures LZc, SCE and ACE computed in the same way as for the whole cortex.

LZc and ACE scored consistently higher for WR than for LOC for all four lobes for all subjects. This was with Cohen's $d > 0.8$ for all lobes for all subjects, except subject 1, see Figure 3.6. SCE scored higher for WR than for LOC for 6 out of the 7 subjects for each lobe, with Cohen's $d > 0.8$ for all subjects and lobes, except subject 1 for all lobes and except subject 3 for all lobes other than the occipital lobe (see red subject labels in Figure 3.6). Thus the discriminative power of the measures was similar when computed across a single lobe to when computed across the whole cortex.

3.4.5 Frequency filtered whole cortex recordings

We also analysed the LZc, SCE and ACE measures on the whole cortex data restricted to different frequency bands, using the same bands as before and also a high-pass filter, excluding all frequencies below $1Hz$. Butterworth filters were applied to restrict the signal to the respective frequency bands prior to computing the signal diversity measures. The results for the 25 electrodes - automatically chosen via k-medoids across the whole cortex - are summarized in Table 3.2.

The findings suggest that the elevated signal diversity in the WR state arises more from activity in low and high frequency bands (delta, beta, gamma) than from intermediate frequency bands (theta, alpha), and that signal diversity is not just a property of the low

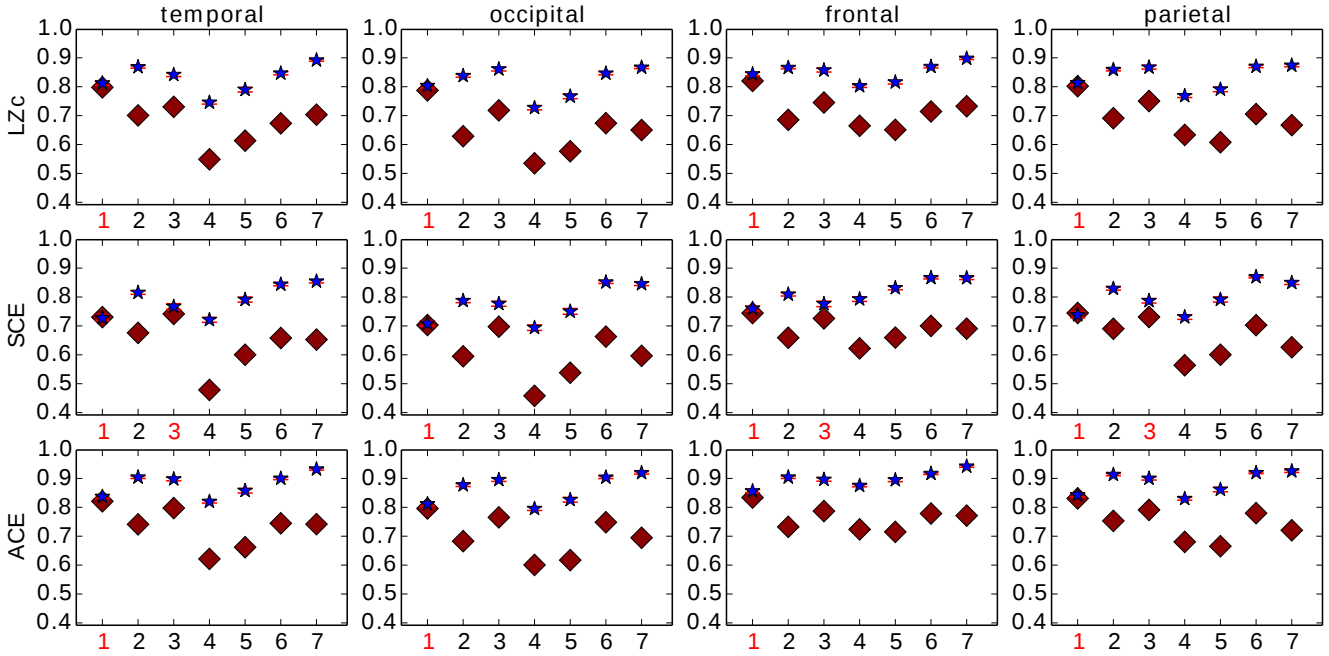

Figure 3.6: **Analysis of 25 k-medoids chosen electrodes for each of four lobes.** Blue star is WR, red diamond is LOC. The red error bars indicate standard error. LZc and ACE score for all subjects higher for WR as opposed to LOC, SCE does so for 6/7 subjects. The measures behave similarly for different brain regions. If the effect size Cohen's $d < 0.8$ for the scores of a given subject, the subject's label is printed in red.

	delta	theta	alpha	beta	gamma	$> 1Hz$
LZc	5 2 0	1 4 2	5 2 0	5 2 0	6 1 0	5 2 0
SCE	6 1 0	3 3 1	4 2 1	**7 0 0**	6 0 1	6 1 0
ACE	**7 0 0**	1 4 2	4 3 0	6 1 0	**7 0 0**	5 2 0

Table 3.2: **Effect size comparison per measure and frequency band for WR/LOC.**For each measure and frequency-band-filtered input, the three numbers display how many subjects out of 7 had higher score for WR than LOC with Cohen's $d > 0.8$ (left digit), no substantial difference, $d < 0.8$, between WR and LOC (middle), and lower values in WR when compared to LOC with Cohen's $d > 0.8$ (right). High-pass-filtered input data are labelled by $> 1Hz$.

frequency components of the spectrum.

3.5 Discussion

We have analysed three different signal diversity measures on spontaneous EEG data from subjects undergoing propofol-induced general anaesthesia: Lempel-Ziv complexity (LZc), synchrony coalition entropy (SCE) and amplitude coalition entropy (ACE). All three of these measures robustly distinguished loss-of-consciousness (LOC) from wakeful resting (WR) on the broadband signal, giving higher mean values for WR as compared to LOC across subjects, a range of segment lengths, and number and location of electrodes (in-

dividual lobes versus whole cortex). On analyses restricted to specific frequency bands, the discriminative power of the signal diversity measures was highest for low and high frequency bands (delta, beta, gamma), and lower for intermediate bands (theta, alpha). The measures ACE and LZc also had some ability to discriminate mild sedation (MS) from WR.

We combined several approaches to verify that LZc, SCE and ACE capture more than just spectral changes between the states. First, we found that re-normalising the measures by values obtained after phase randomisation did not affect the results (see section Appendix A.3). In addition, when comparing the discriminative power (as measured by AROC) of the three signal diversity measures with that of normalised spectral power (on the non-randomised data) we found that normalised delta power discriminated between the states WR and LOC as well as the three signal diversity measures did, but it could not discriminate between MS and WR (see Figures 3.4, 3.3, S3 and Table 3.1), unlike LZc. Further, all three signal diversity measures tend to behave monotonically with respect to depth of sedation (see Figure 3.3), unlike delta or gamma power (see Figure S3). Normalised spectral power in other bands all had weaker discriminative power for WR/LOC. Finally, we did not see significant correlations between the change in delta power and changes to any of the signal diversity measures, amongst the 7 subjects (Figure 3.5).

The consistency of our findings across the four cortical lobes, over different numbers of channels, segment lengths, or normalisation (spectral-profile preserving or not) shows that the decrease in spontaneous EEG signal diversity during general anaesthesia is very robustly measurable. In ongoing work we are exploring the behaviour of these measures on a finer spatial scale to see at what level, if any, regional differences can be detected.

Our simulation with the Kuramoto model (Section 2.6.3) demonstrated that SCE, with its analysis of phase (as opposed to amplitude) synchrony, can exhibit different behaviour from LZc and ACE (see Figure 2.7). We varied a parameter that controls the overall level of synchrony, and found that SCE peaked strongly at an intermediate level of overall synchrony (as one would expect), while LZc and ACE actually exhibited a small dip where SCE peaked. This simulation, together with the slightly different results for the three measures when applied to EEG (see Figure 3.3), demonstrate that at least two distinct 'flavours' of signal diversity were decreasing in the spontaneous EEG during anaesthesia. Chapter 6 presents additional simulations with Stuart-Landau oscillators to examine more closely situations in which these three measures converge and diverge.

Our results complement those of Casali et al. [53] who measured Lempel-Ziv complexity of the EEG response to transcranial magnetic stimulation (TMS). In their paradigm, high levels of Lempel-Ziv complexity corresponded to conjoined differentiation and integration by virtue that a high score could only be obtained when the neuronal activity in response to the magnetic perturbation spreads far (integration) and evolves in a non-stereotypical way (differentiation). Their variant measure, called the Perturbational Complexity Index (PCI), also scored consistently higher for WR than for propofol anaesthetised subjects (LOC) [53], in addition to also scoring relatively low values for other unconscious states, namely, non-rapid eye movement sleep and vegetative state.

In summary, we have demonstrated a correlation between level of sedation and three distinct signal diversity measures during propofol-induced general anaesthesia. There are now a number of candidate 'neural correlates of consciousness' [70]. We have added to the list multi-dimensional signal diversity of spontaneous EEG. Our measures derive from theory and thus are 'explanatory correlates' [202], are quick and easy to compute, and do not involve the need to stimulate or perturb the brain.

Testing the measures on data from other manipulations of conscious level adds to a more complete indication of the specificity and sensitivity of these measures to the level of consciousness associated with diverse states. Analyses of signal diversity during sleep states are presented next in Chapter 4 and during psychedelic states in Chapter 5.

Chapter 4

Spontaneous depth electrode signal diversity decreases globally and locally during NREM sleep

4.1 Abstract

Key to obtaining a scientific understanding of how the brain generates consciousness is the characterisation of the neural signatures of changes in level of consciousness during sleep. Here we analysed three measures of signal diversity on spontaneous depth electrode recordings from 10 epilepsy patients during wakeful rest and different stages of sleep: (i) Lempel-Ziv complexity, which is derived from how compressible the data are; (ii) amplitude coalition entropy, which measures the variability in the constitution of the set of channels active above a threshold; (iii) synchrony coalition entropy, which measures the variability in the constitution of the set of synchronous channels. When computed across sets of channels that are broadly distributed across multiple brain regions, all 3 measures decreased substantially in all subjects during early-night non-rapid eye movement (NREM) sleep. This decrease was partially reversed during late-night NREM sleep, while the measures scored similar to wakeful rest during rapid eye movement (REM) sleep. This global pattern was in almost all cases mirrored at the local level by groups of channels located in a single region. In testing for differences between regions, we found elevated signal diversity in the frontal lobe. Our results provide further evidence that the level of consciousness correlates with neural dynamical complexity.

4.2 Introduction

As shown in Chapter 3, we found evidence for a robust and spatially uniform decrease in three distinct flavours of spontaneous multi-dimensional EEG signal diversity during propofol induced general anaesthesia. These three distinct flavours of signal diversity were captured by LZc, ACE and SCE as introduced and defined in Chapter 2.

Here we apply the these spontaneous signal diversity measures to depth electrode

recordings taken during non-rapid eye movement (NREM) sleep, REM sleep, and wakeful rest (WR) from ten epilepsy patients undergoing pre-surgical evaluation. Complementing our analysis of spontaneous signal diversity of EEG under propofol, our main focus with this data set from depth electrodes is to firstly investigate signal diversity changes during sleep phases and secondly use the high spatial resolution of depth electrodes to compare regional with global signal diversity changes. Electrode locations varied across subjects and spanned cortical and sub-cortical regions. We tested the extent to which changes in the signal diversity measures were consistent and detectable irrespective of which regions were covered. This regional analysis of signal diversity is relevant for comparison with other signatures of the different sleep stages, some of which have been reported to differ strongly across cortical regions (e.g. slow wave and sleep spindle propagation [7, 163]) while other signatures are exhibited more evenly across the cortex (e.g. average power spectra [56]).

A recent study of avalanche events during wake and sleep states found more large avalanches in NREM than REM or WR, indicating increased correlations between cortical areas in NREM sleep and more fragmented cortical dynamics in WR and REM [182]. Given the similarity of our dataset with the one studied in [182], we additionally replicated results from that study, and discuss the implications in conjunction with the observed behaviour of the signal diversity measures.

We found that all three signal diversity measures scored substantially lower in all subjects during NREM than during WR, when computed across sets of channels broadly distributed across multiple cortical and sub-cortical regions. This global pattern was in almost all cases mirrored at the local level when analysing the measures across groups of channels located in just a single region. We also found evidence for higher signal diversity in the frontal lobe than the other cortical lobes.

4.3 Methods

4.3.1 Ethics statement and data protection

In agreement with the HORIZON 2020 requirement, the protocol used to collect the data analysed here has been drawn up in accordance with the EU standards of good clinical practice and with the Declaration of Helsinki (current revision) and is approved by the Ethics Committee of the Niguarda Hospital of Milan (protocol number: ID 939, Niguarda Hospital, Milan, Italy). All data related to the study participation are treated confidentially in compliance with good clinical practice as well as in compliance with Italian specific national laws on the protection of individuals. Patients are informed that personal data are collected and stored electronically, that can be used for purposes of scientific research and that dissemination of the results can take place only in an anonymous and / or aggregate form. Patients are informed that they have the right to access the stored data, and to update or modify erroneous data.

4.3.2 Patients and data acquisition

The data were derived from a dataset collected during the pre-surgical evaluation of ten neurosurgical patients with a history of drug-resistant, focal epilepsy. All subjects were candidates for surgical removal of the epileptogenic zone. The recordings were obtained from stereotactically implanted depth multi-lead electrodes (Stereo-EEG, SEEG), inserted for the precise localisation of the epileptogenic zone and connected areas [67]. The investigated hemisphere, the duration of implantation, the location and number of recording sites were determined based on non-invasive clinical assessment.

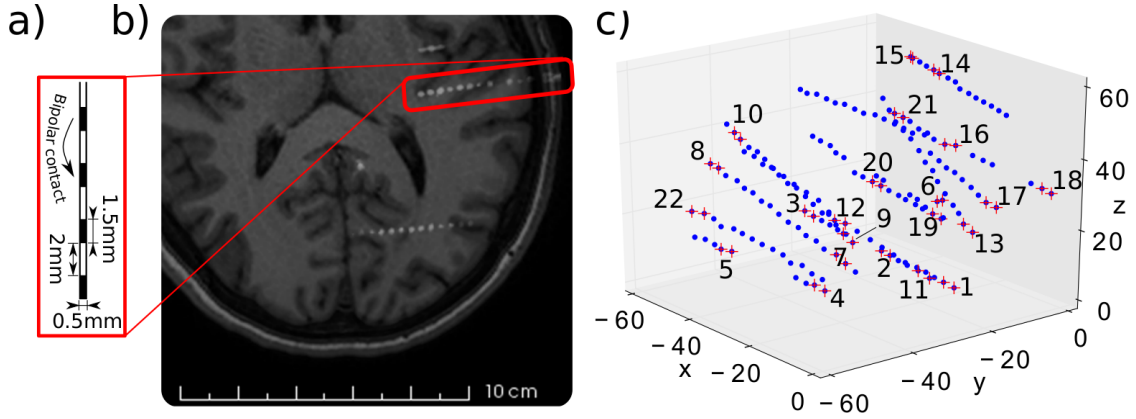

Figure 4.1: **Depth electrode dimensions, location and channel definition.** a) Outline of a multi-lead intracerebral electrode. b) Horizontal section of CT scan of subject 1, showing a multi-lead intracerebral electrode (red rectangle) [9]. c) In MNI space in mm, the location of initial contacts (blue dots) are partially overlaid by red crosses, marking the contacts chosen for bipolar-montage. Numbers index the resulting channels as used for analysis. Channels were obtained by applying bipolar-referencing to neighbouring contact pairs that were not discarded (see text for details).

SEEG activity was recorded from platinum-iridium semiflexible multi-contact intracerebral electrodes, with a diameter of 0.8mm, a contact length of 1.5mm, an inter-contact distance of 2mm and a maximum of 18 contacts per electrode (Dixi Medical, Besancon France), see Fig. 4.1. The individual placement of the electrodes was ascertained by post-implantation tomographic imaging (CT) scans (post-CT), and Montreal Neurological Institute (MNI) coordinates obtained for each contact. Full details on the contact localisation procedure can be found here [9]. Briefly, post-CT was coregistered to pre-implant MRI by an algorithm based on rigid affine transformation and mutual information. Next the post-CT scan was thresholded and skull-stripped in order to find and remove radiological artefacts. Given the planned entry points on the skull and electrode pin dimensions, the axis direction was estimated and each recording position iteratively computed. A region-of-interest around each contact was defined and a most probable anatomical label assigned, using the Destrieux atlas. Single subject channel positions were projected to MNI space by internal transformation files defined in the matlab toolbox Freesurfer.

In addition scalp EEG activity was recorded from two platinum needle electrodes placed

during surgery on the scalp at standard 10-20 positions Fz and Cz. Electro-ocular activity was recorded from the outer canthi of both eyes, and submental electromyographic activity was also recorded. Both EEG and SEEG signals were recorded using a 192-channel recording system (NIHON-KOHDEN NEUROFAX-110) with a sampling rate of 1000 Hz. Data were recorded and exported in EEG Nihon-Kohden format. Recordings were referenced to a contact located entirely in the white matter.

4.3.3 Selection of recording contacts and data preprocessing

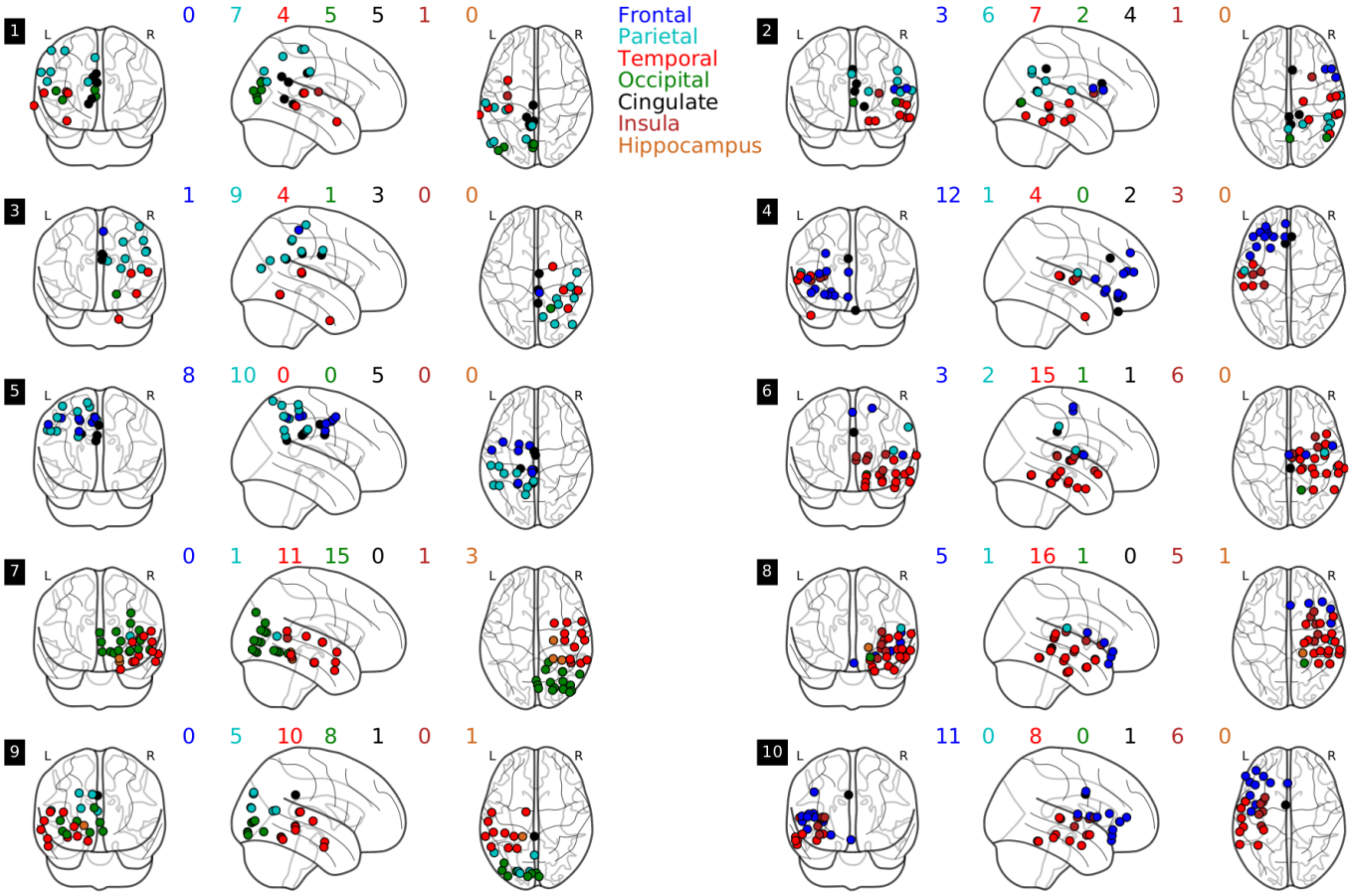

Figure 4.2: **Channel locations for the 10 subjects.** For each subject, coloured dots indicate the positions of channels that were used for analysis. The locations are plotted using MNI coordinates on a standard glass brain (python nilearn [5]); see text for details. Numbers in black blocks indicate subject number, coloured digits count channels per subject and region.

In each subject recordings were made from up to 194 contacts (blue dots in Fig. 4.1c). For the present analysis, selection of recording sites was based on the following criteria: we excluded from the analysis those contacts that (i) were located in the epileptogenic zone (as confirmed by post-surgical assessment), (ii) were located over regions of documented alterations of the cortical tissue (e.g. Taylor Dysplasia) as measured by the radiographic assessment, or (iii) exhibited spontaneous or evoked [239] epileptiform SEEG activity dur-

ing wakefulness or NREM. Contacts located in white matter, assessed by MRI, were also excluded from analysis. Data samples were taken from each subject from four different states: wakeful rest (WR), non-rapid eye movement sleep early at night (NREMe), non-rapid eye movement sleep late at night (NREMl) and rapid eye movement sleep (REM). Sleep scoring was obtained according to [213] using one scalp EEG derivation, together with one bipolar electrooculographic (EOG) and one electromyographic (EMG) derivation. All NREM epochs were collected during stage N3 according to [213]. NREMe corresponds to the first stable NREM (stage N3) episode and NREMl to the last stable NREM (stage N3) episode of the night [174]. By using only NREM in stage N3, possible fluctuation of the level of consciousness due to subliminal processing in NREM stage N1 and N2 [8] are avoided. The data samples were imported from EEG Nihon Kohden format into Matlab and converted using a customised Matlab script.

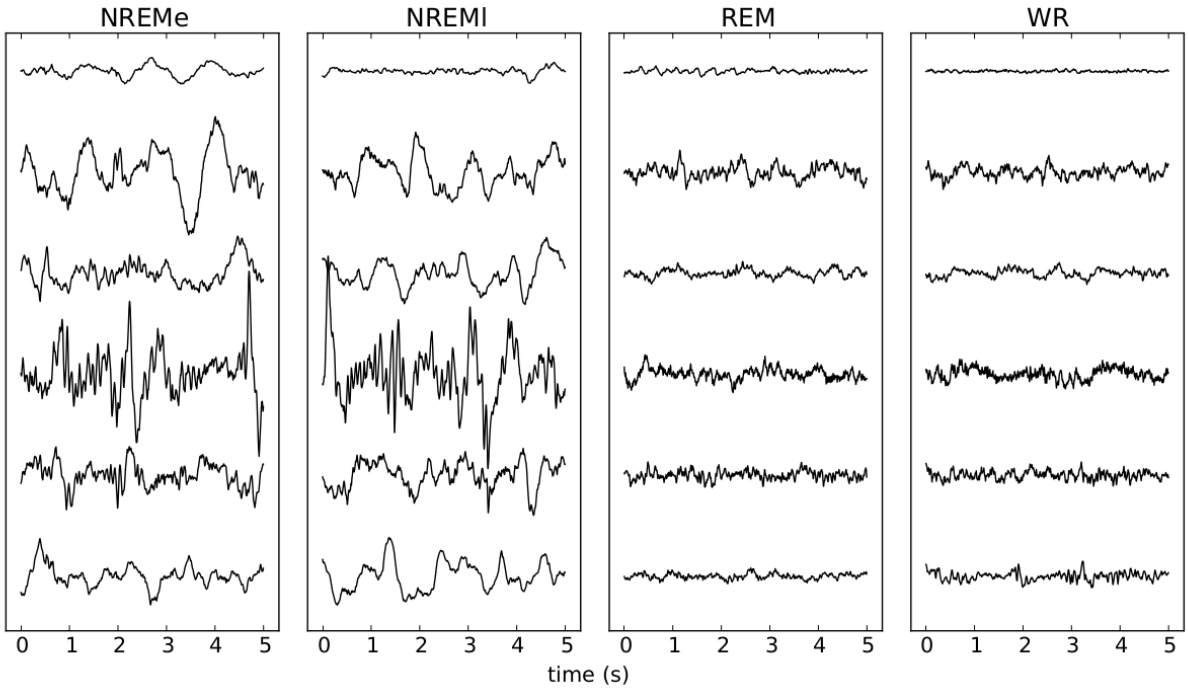

Figure 4.3: **Sample segments from 5 channels for each state.** Segments are shown after pre-processing prior to normalisation by standard deviation. The recordings for non-rapid eye movement sleep early at night (NREMe) display visibly stronger slow waves (low-frequency components) as compared to those for rapid-eye movement sleep (REM) or wakeful rest (WR).

Bipolar montages were calculated by subtracting the signals from adjacent contacts of the same depth-electrode (see Fig. 4.1a) to minimise common electrical noise and to maximise spatial resolution [54, 104]. To further minimise volume conduction artefacts, at most every third (bipolar) channel from each electrode was retained for analysis. The number of retained channels per patient varied between 18 and 31 (red crosses in Fig. 4.1c show an example choice of contact pairs; this is subject 1 in Fig. 4.2). For most analyses we used 18 channels per subject, selected as follows. A first electrode was chosen at

random and the m channels on that electrode were all selected, ordered from the innermost outward. The process was repeated until 18 channels had been selected. (See Fig. 4.1c for an illustration of this ordering). Fig. 4.2 depicts the channels' anatomical locations.

The data from the selected channels was further preprocessed as follows. No epochs were removed, as visual inspection did not result in the detection of severely artefacted epochs. The data samples were downsampled to 250Hz and divided into 10s segments. Linear de-trending, baseline subtraction and normalisation by standard deviation was performed for each channel of each segment. After preprocessing the length of the retained data sample for each subject and state varied between 7-16min. Fig. 4.3 illustrates representative channel activity for the different states.

4.3.4 Statistics

Analyses were performed using non-overlapping segments of length 10s for a total length between 7min and 16min of SEEG recording per subject and per state. The mean and standard error of the signal diversity measures' scores were computed over these segments. At the single subject level, the effect size of differences between states was measured using Cohen's d [64]. We call an effect size high if $d > 0.8$ [220]. For group level comparisons, a t-test was applied, with correction for false discovery rate (FDR) via the Benjamini-Hochberg procedure.

4.4 Results

4.4.1 Global analyses

We first computed the signal diversity measures across broadly-distributed sets of channels. Specifically, for each subject we used the first 18 channels (out of the 18 to 31 available; see Methods for channel ordering), whose precise sets of locations varied across subjects, but which spanned at least two cortical regions (see Fig. 4.2). Fig. 4.4a shows the mean values across 10s segments of LZc, ACE and SCE for each subject during wakeful rest (WR), non-rapid eye movement sleep early at night (NREMe), non-rapid eye movement sleep late at night (NREMl) and rapid eye movement sleep (REM). Further, Fig. 4.4b shows mean results across subjects. Classification of effect sizes for differences between states at the single subject level are shown in Tab. 4.1. For all subjects, the three measures ACE, SCE and LZc score higher for WR than NREMe with high effect size (Cohen's $d > 0.8$). For all but one measure and subject, scores are also higher for REM than NREMe with high effect size. Values for NREMl typically lie in between those for WR or NREMe, with high effect sizes for most subjects (see state pair NREMl/NREMe in Tab. 4.1). The exception is LZc for NREMl versus NREMe, for which the effect size is small for most subjects. The difference in values between WR and REM is small for most cases ($d < 0.8$). Overall, signal diversity across all available brain regions is higher for REM and WR - states associated with consciousness - as opposed to NREMe or NREMl.

In Appendix B, the behaviour of the signal diversity measures is compared to that of normalised spectral power in the various frequency bands (δ, θ, α, β, γ), that of a simple

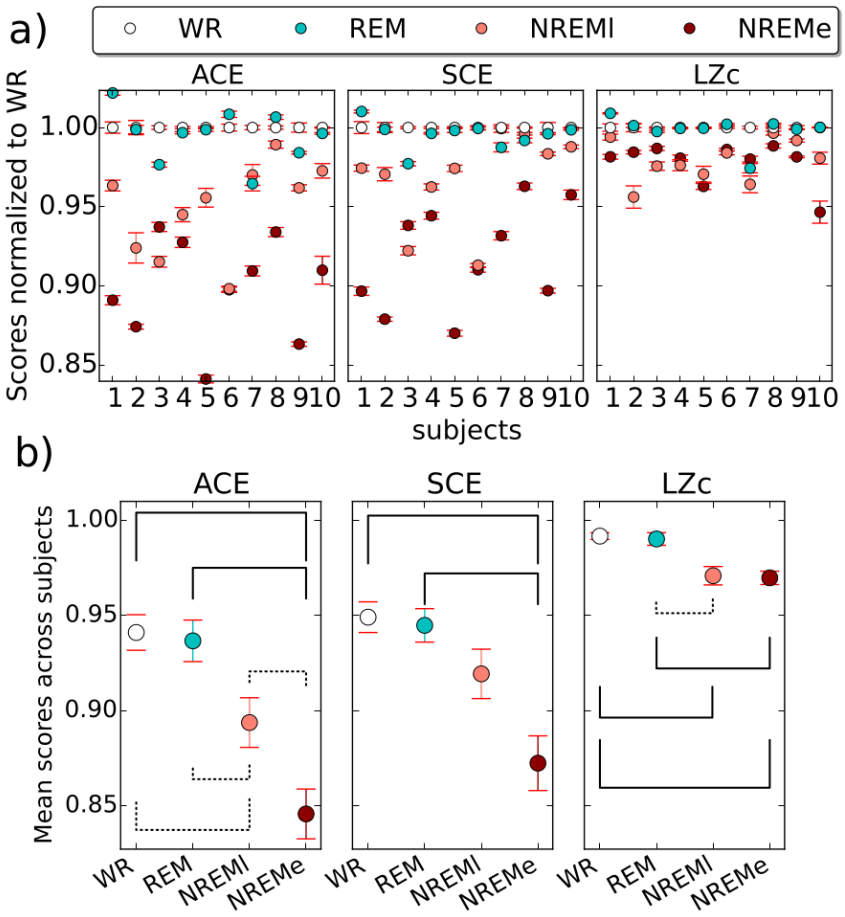

Figure 4.4: **ACE, SCE and LZc computed across broadly-distributed sets of 18 channels.** (a) Individual subject plots. Plotted points show mean over 10s segments. Scores are normalised by the score for WR (in addition to each measures' normalisation as specified in their definition). The measures score higher for WR as opposed to NREMe or NREMl, consistently across subjects. The measures score similar values for REM and WR. Error bars indicate standard error across segments. See Tab. 4.1 for effect sizes at the single subject level. (b) Mean scores across the 10 subjects. For each measure, the mean and standard error across subjects are displayed, here not normalised relative to scores for WR. Significant differences between state pairs are shown by a solid line if $p < 0.01$ and a dotted line if $p < 0.05$ (t-test, FDR corrected for multiple comparison).

correlation measure, sumCov, which equals the mean of the absolute values of correlation coefficients between all channels, and that of LZsum, the mean Lempel-Ziv complexity of single channels (see Tab. B.1). As expected, normalised delta power is consistently substantially higher in NREMe than WR, while beta and gamma power are in almost all cases substantially lower. The high delta band power during NREM sleep reflects the presence of slow-waves [163, 76, 245, 48, 152, 83].

For each state we investigated the extent to which the signal diversity measures correlated with each other, as well as with normalised spectral power in the various frequency bands. Fig. 4.5 indicates the pairs of measures and states that showed the strongest correla-

	ACE	SCE	LZc
WR/NREMe	10, 0, 0	10, 0, 0	10, 0, 0
WR/NREMl	8, 2, 0	8, 2, 0	8, 2, 0
WR/REM	4, 5, 1	2, 8, 0	1, 9, 0
REM/NREMe	10, 0, 0	10, 0, 0	9, 1, 0
REM/NREMl	9, 1, 0	8, 2, 0	8, 2, 0
NREMl/NREMe	7, 2, 1	8, 2, 0	2, 8, 0

Table 4.1: **Effect size comparison per measure and state pair.**For each measure and state pair, the three numbers correspond to how many subjects out of 10 had higher score for the left state with Cohen's $d > 0.8$ (left digit), no substantial difference, $d < 0.8$, (middle) and higher score for the right state with Cohen's $d > 0.8$ (right), computed across trials. The results were obtained from applying the measures to $10s$ segments from 18 channels as in Fig. 4.4.

tions (Pearson coefficient r of absolute value greater than 0.7). Strong positive correlations ($r > 0.7$) were found for ACE/SCE for all states, and ACE/LZc for all states except REM. Weaker, yet still significant correlations were observed between SCE and LZc (see Tab. B.2 and for scatter plots of scores for each segment Fig. B.2). The signal diversity measures SCE and ACE also showed strong negative correlations ($r < -0.7$) with delta power during NREM sleep states but only weaker correlations during WR and REM, and $r > 0.7$ was observed for SCE versus gamma power during NREMl. Weaker, yet still significant correlations were observed for LZc versus delta power, ACE versus gamma power and LZc versus gamma power (see Tab. B.2). The imperfect correlation between all the signal diversity measures indicates that they are capturing not entirely equivalent properties of the dynamics.

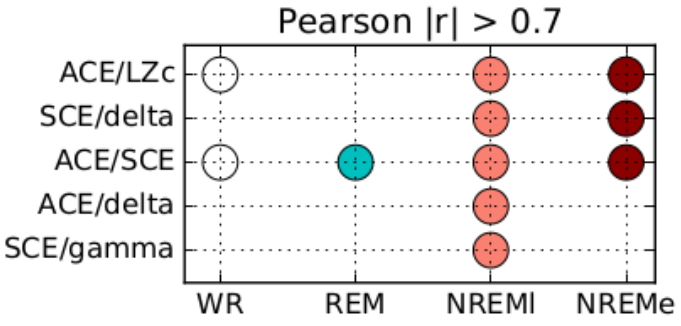

Figure 4.5: **Thresholded correlations between signal diversity measures and frequency bands for each state.** For a given state and subject the Pearson correlation between pairs of measures was computed across 10s segments. For each state and measure pair (ACE, SCE, LZc and spectral power in the 5 canonical frequency bands, correlation of two spectral power bands not shown) a disc coloured by state is displayed if the absolute value of the Pearson correlation r averaged across subjects is greater than 0.7. (These correlations are all significant at $p < 0.05$, corrected for false discovery rate.)

We further showed that the observed differences in ACE, LZc and SCE scores between states remain intact in most cases when renormalising the measures based on phase-

randomised surrogate data, thus demonstrating that the measures are reflecting more than just changes in spectral power, in spite of the strong (negative) correlation of two of the signal diversity measures with delta power for NREM states (see Appendix B.5).

4.4.2 Local analyses

4.4.2.1 Differences between states

Fig. 4.4 shows that, when computed across widely distributed sets of channels, ACE, SCE and LZc take lower values during NREM than WR. In order to test if or to what extent there are regional differences in the decrease of signal diversity with NREM sleep, we used the following three approaches.

Firstly, we applied ACE, SCE and LZc to individual regions, following the classification in Fig. 4.2. For each subject, each region with four or more channels was analysed; where there were more than four channels, a quartet was picked at random. Depending on the subject-specific distribution of electrodes, there were two, three or four regions analysed per subject. The results (just for WR and NREMe) are shown in Fig. 4.6, and in general mirror the global result. The scores for LZc and ACE were in almost all cases greater for WR than for NREMe with large effect size. The consistency of SCE was weaker than for the other two measures (see Tab. B.3 for score counts). Importantly however, despite several instances of SCE showing increased values during NREMe, there was no region for which this was observed for more than one subject.

Secondly, in order to perform an analysis on the maximum possible number of regions for each subject, we computed the Lempel-Ziv complexity of each single channel (LZs, same computation as LZc with trivial concatenation, i.e. a measure of temporal signal diversity only). Averaged across all channels per region and subject, LZs scored for 48 out of 50 region/subject pairs lower during NREMe than WR (see Tab. B.3 and Fig. B.4). Considering individual channels further, we noted also that normalised delta power is higher for NREMe than WR for 49 out of 50 region/subject pairs (compare Fig. B.4 with Fig. B.5).

Thirdly, we investigated the signal diversity of the interaction of a channel in one region with a group of channels in another region. To this end we utilised the local synchrony coalition entropy $SCE^{(i)}$ of a group of target channels in one region with respect to a seed channel i in another region (see Chapter 2 and Appendix B.4). When taking 3 target channels we observed lower values during NREMe than WR for 107 out of 129 choices of seed and target regions for all subjects. Despite this inconsistent decrease of $SCE^{(3)}$ with NREM sleep, there were no exceptional choices of seed and target region that went against the trend for more than one subject (Fig. B.6). Similar results were obtained when taking 2 target channels (Fig. B.7) with 109 out of 143 choices of seed and target regions for all subjects showing lower values during NREMe than WR, also without discernible pattern for the 34 cases with LZs higher in WR than NREMe. We conclude that there is no consistent local deviation from the global behaviour of SCE. Finally, we tested complexity of synchrony (CS) between pairs of channels, defined as the Lempel-Ziv complexity of their synchrony time series (see Appendix B and Fig. B.8). CS scored higher for WR than

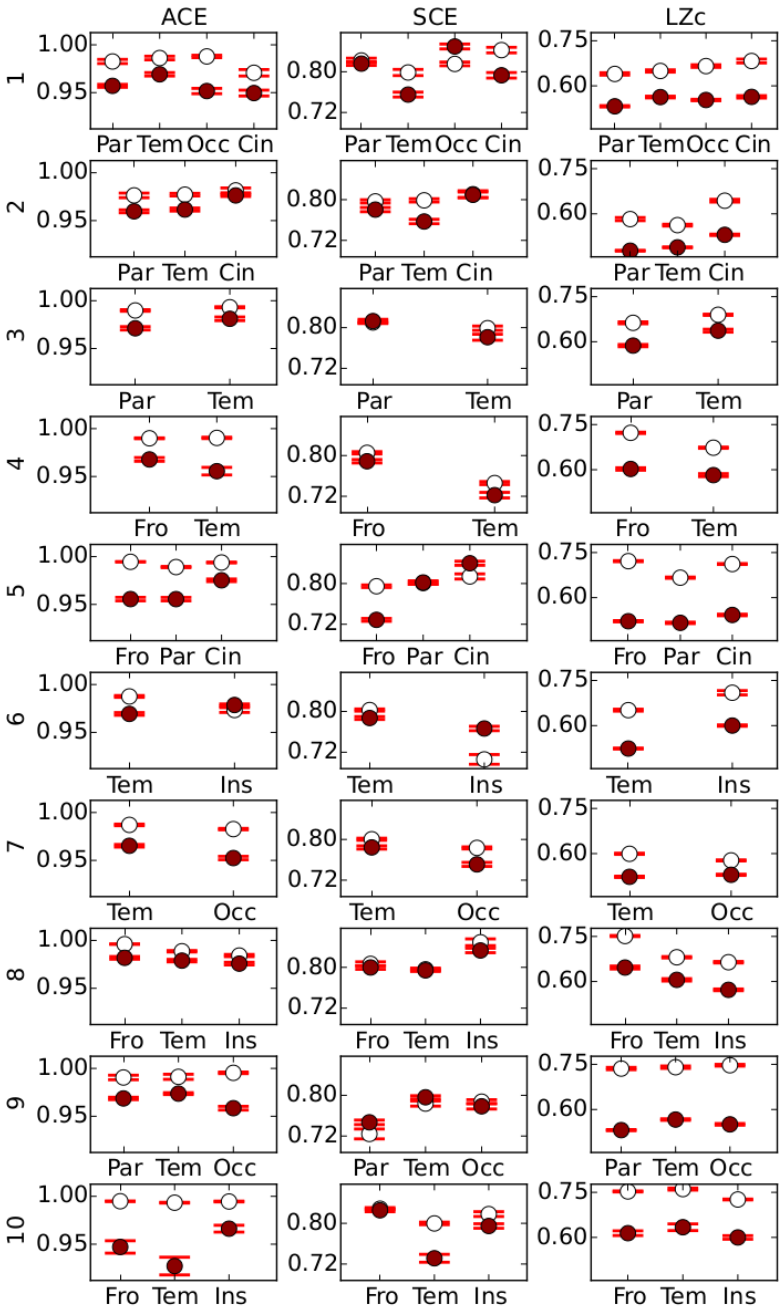

Figure 4.6: **ACE, SCE and LZc scores computed for channel quartets restricted to single regions.** The results for WR (white) and NREMe (red) were obtained from applying the measures repeatedly to quartets of channels within regions as specified by first 3 letters of region names (as listed in Fig. 4.2). LZc and ACE score for all regions and subjects higher for WR than NREMe (with just a single exception for ACE); SCE scores less consistently here.

NREMe for 122/139 pairs of regions.

In summary, all locally applied signal diversity measures behave predominantly according to the main trend with just a few anomalies. There was no consistent pattern to anomalies across regions or subjects.

4.4.2.2 Differences between regions for a given state

Differences in signal diversity scores between regions were overall less pronounced and less consistent than differences in signal diversity between states (see Fig. 4.6). Here we investigate whether there are trends of regional signal diversity differences at the group level.

Mean scores for LZc, ACE and SCE were computed from channel quartets restricted to each of the four cortical lobes. All measures scored higher in the frontal lobe compared to other lobes for REM and WR, although an ANOVA test for a significant variation across the lobes gave a significant p value (< 0.05; p-values uncorrected in this section) only for LZc when pooling across states ($F = 4.1$; $p = 0.01$). Given the available data, this analysis had $n = 4, 5, 9, 3$ respectively for frontal, parietal, temporal and occipital cortex.

We repeated the analysis with LZs (single-channel Lempel-Ziv complexity), computable whenever the subject has just one or more channels located in the given lobe. This led to the slightly larger sample sizes of $n = 7, 9, 9, 7$ respectively for frontal, parietal, temporal and occipital cortex. For each state, LZs scored highest for the frontal lobe, see Fig. 4.7. A 1-way ANOVA test for significant variation between lobes yielded $(F, p) = (2.3, 0.10), (1.9, 0.15), (0.8, 0.49), (1.6, 0.20)$ for state WR, REM, NREMl, NREMe respectively (F is the statistic, p the associated p-value from the F-distribution). A t-test as a post-hoc analysis indicated a significant difference at $p < 0.05$ (uncorrected) between the frontal and temporal lobe in states WR and REM. When pooling values from all states for a given lobe, the ANOVA test did indicate a significant variation of LZs across lobes, with $(F, p) = (4.5, 0.005)$. By contrast, an ANOVA across lobes for the difference LZs(WR)-LZs(NREMe) yielded $(F, p) = (1.6, 0.2)$, confirming that the drop in signal diversity with NREMe does not differ substantially across lobes. For comparison, an ANOVA between lobes for delta power yielded $(F, p) = (1.8, 0.16), (2.2, 0.11), (0.4, 0.74), (0.85, 0.47)$, with a significant (t-test, $p < 0.05$, uncorrected) difference between the frontal and temporal lobe (delta higher in temporal) in states WR and REM; pooled across states gave $(F, p) = (2.7, 0.05)$.

In summary, we found evidence for greater signal diversity in the frontal lobe compared to the other cortical lobes and confirmed that signal diversity changes across states are more pronounced than across regions.

4.4.3 Signal diversity in different frequency bands

To test whether the relationship between signal diversity and conscious level is restricted to specific frequency bands, we re-analysed the ACE, SCE and LZc signal diversity measures on the data after frequency filtering (using Butterworth filters). The results for 18 channels per subject (chosen as above) are shown as average scores across subjects in Fig. 4.8 for

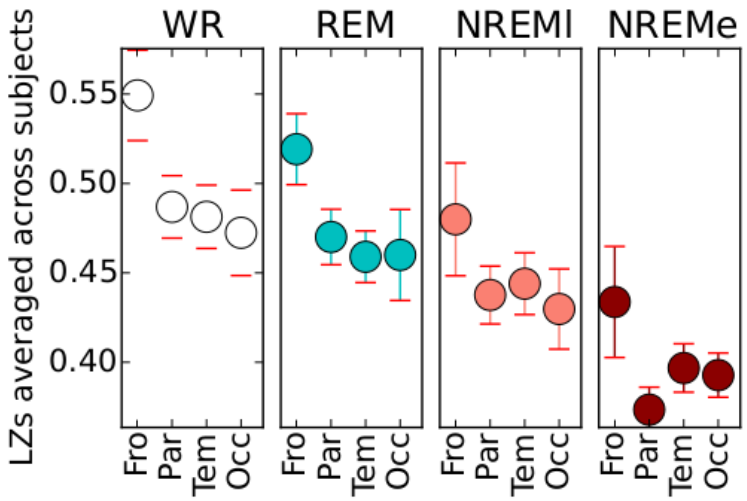

Figure 4.7: **Mean LZs scores for the four cortical lobes in each state.** For each subject, LZs scores were averaged across all channels from the given lobe. The mean of these scores across subjects is displayed with standard error across subjects for each state. The labels on the x-axis give the first 3 letters of the lobe name (frontal, parietal, temporal, occipital) from 7, 9, 9 and 7 subjects respectively. LZs scores higher for the frontal lobe than for the other lobes (ANOVA pooled across states for difference between regions gives $F = 4.5$ and $p = 0.005$; t-tests for differences between pairs of lobes came out non-significant ($p > 0.05$) except for between Fro and Tem for states WR and REM (uncorrected for multiple comparison).

6 different frequency bands and for 6 different high-pass cutoffs. Fig. 4.8a shows that the elevated signal diversity of ACE, SCE and LZc in the WR state is most pronounced in the delta (1-4Hz) and alpha (8-13Hz) ranges, and the general trend is at least weakly present in all bands. Also in Fig. 4.8b higher signal diversity in WR than NREMe is visible for all high-pass cutoffs, yet this difference becomes smaller with increasing high-pass cutoff.

In summary, higher signal diversity in WR than NREMe is present for almost all tested frequency bands and high-pass cutoffs, and it is strongest when frequencies smaller than 4Hz are left in the signal. This shows that the decrease of signal diversity with NREMe is amplified by an increase of low-frequency waves, yet still exists after most of the applied frequency filters.

4.4.4 Avalanche statistics

Recent studies of avalanche events have shown that avalanche size distributions follow a power law to good approximation over a wide range of scales, during both wake and sleep states [182, 185]. However, large avalanches were found to be more frequent in NREM sleep than WR or REM, indicating increased correlations between cortical areas in NREM sleep and more fragmented cortical dynamics in WR and REM, in line with our findings with the signal diversity measures. Since the present dataset is similar in structure to the one analysed in [182], we replicated from that study some analyses of the distribution of avalanche events during wake and sleep states.

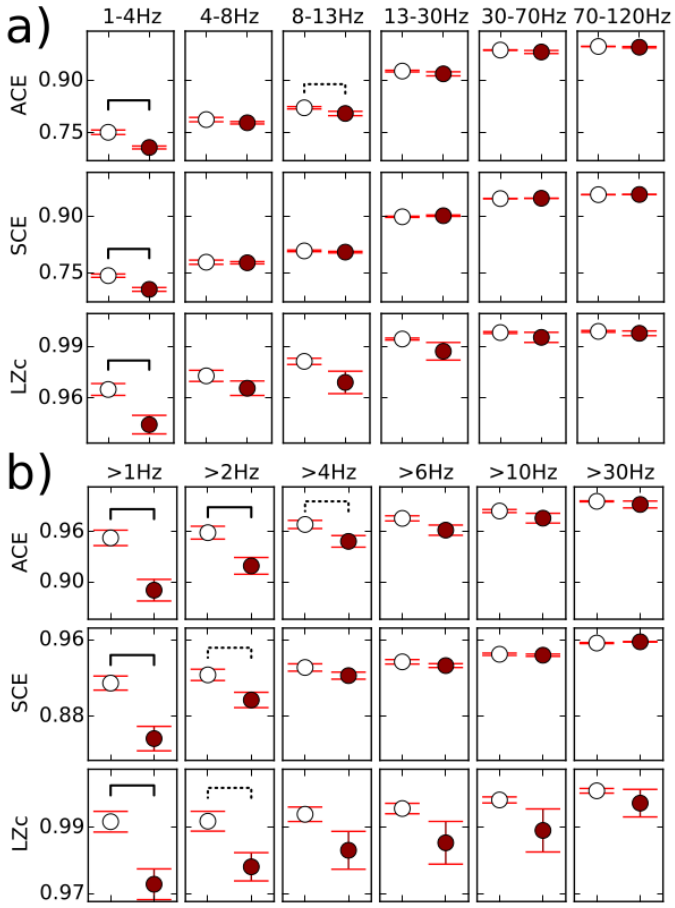

Figure 4.8: **ACE, SCE and LZc for frequency filtered data.** (a) Bandpass filtering (frequency ranges indicated in each column). (b) High-pass filtering (frequency cutoffs indicated in each column). Scores were obtained from applying the measures to 10s segments of frequency filtered data from 18 channels and then averaging across all segments from all subjects. WR is in white and NREMe in red. Significant differences between state pairs are shown by a solid line if $p < 0.01$ and a dotted line if $p < 0.05$ (uncorrected t-test across subjects). Error bars show standard error computed from the mean scores for each of the 10 subjects.

We defined events and avalanches of events in the same way as in [182], in brief as follows. For each positive deflection between two zero crossings of a channel time-series, the area under the deflection was calculated. An event was said to have occurred whenever this area exceeded a threshold. The threshold was set for each channel such that there was the same event rate of 13 events per second for all channels. The time-series is then binned (we took the bin size to be half the mean inter-event interval as in [182]), assigning a 1 to a bin if an event occurs in it, and a 0 otherwise. An avalanche is defined as a cluster of events: in each time bin during an avalanche, there is an event occurring in at least one channel. Avalanches are preceded and followed by time bins in which there are no events. Three quantities that characterise avalanches are analysed: (i) avalanche size s, which is the total number of binary events that occur during the avalanche; (ii) avalanche

duration d, which is the number of time bins spanned by the avalanche; (iii) the branching parameter σ which is the average number of events in one time bin divided by the number of events in the preceding time bin, given that there was at least one event in the preceding time bin.

For the avalanche analyses we used 18-31 channels per subject, selected and preprocessed as described in the Methods section. We successfully replicated the findings of Priesemann et al.'s study [182], that mean size $\bar{s}$ and duration $\bar{d}$ of avalanches, and the mean branching parameter $\bar{\sigma}$ for avalanche events were all greater during NREM sleep (we used NREMe only) than during WR or REM, see Fig. 4.9a. Further, for all states the avalanche distribution was approximately following a power law, however large avalanches were more frequent in NREMe than WR or REM (Fig. 4.9b). We discuss these findings in conjunction with our findings for the signal diversity measures in the Discussion.

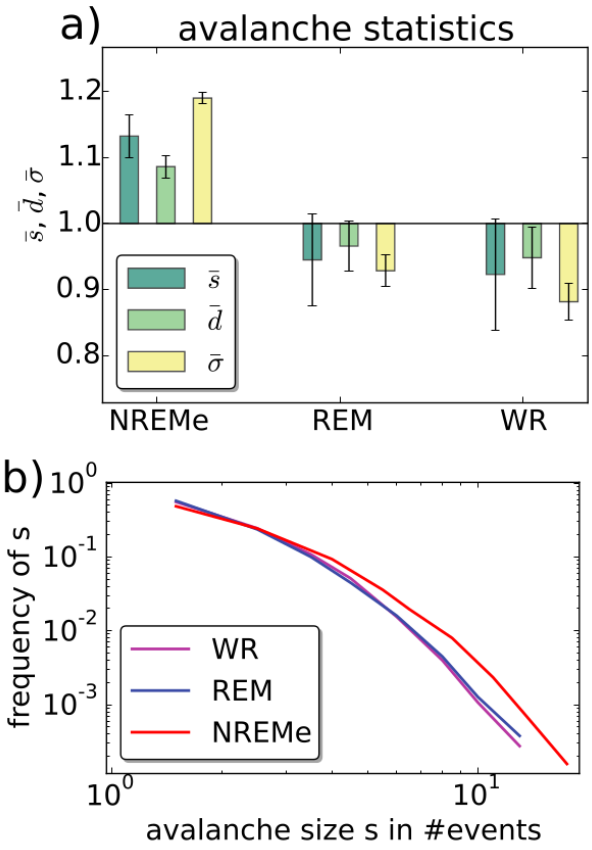

Figure 4.9: **Avalanche analysis as introduced by Priesemann et al., Figs. 7 and 4 in [182], for the present dataset.** (a) States shown are NREMe, REM and WR. Scores of each measure (mean avalanche size $\bar{s}$, mean avalanche duration $\bar{d}$, mean branching parameter $\bar{\sigma}$ across 10 subjects) are shown for each state, normalised by mean score across all three states. Error bars indicate standard error across 10 subjects. All three measures have substantially higher scores for NREMe than WR or REM ($p < 0.01$, FDR corrected t-test across subjects). (b) The avalanche size distribution for each state (not normalised) approximates a power law. Large avalanches are more frequent in NREMe than in WR or REM.

4.4.5 Controls

It is important to verify that changes in signal diversity with conscious level, as measured by ACE, LZc and SCE, are not merely reflecting well known changes in spectral properties of the data (e.g. increase in delta power during NREM sleep). To address this we performed a surrogate data analysis, following our previous study [191]. Specifically, we

computed the measures normalised by their values for phase-randomised surrogate data. Phase shuffling was performed by Fourier transforming each channel, and then applying inverse Fourier transform with the addition of an independent random phase. When repeating the analysis displayed in Fig. 4.4 with the phase surrogate normalisation, the consistency of the measures' behaviour across subjects was slightly reduced (see Fig. B.9) yet the same general behaviour was observed. Thus we have confirmed that the observed changes in signal diversity measures are indeed not merely reflecting spectral changes.

As an alternative analysis of the extent that the measures' sensitivity to conscious level depends on spectral changes, surrogate data were created with the spectral profile of wakeful rest (WR) and non-rapid eye movement sleep (NREMe), then the high-pass frequency filters of Fig. 4.8b were applied prior to computing the measures ACE, LZc and SCE. All three measures score higher for the spectrum corresponding to wakeful rest, also after applying a 4Hz high-pass cutoff - showing that the measures are sensitive to low-frequency component of the signal - however for cutoffs of 10Hz, all three measures score higher for NREMe than WR (Fig. B.1). Crucially, the measures did still score higher for high-pass filtered depth-electrode recordings at 10Hz (Fig. 4.8b), confirming that spectral changes alone are not enough to explain these measures' sensitivity to conscious level.

When omitting bi-polar referencing as a pre-processing step, results are barely affected (Fig. B.10). This shows that bi-polar referencing as a pre-processing step of the data is not crucial for the detection of signal diversity changes in WR/NREMe.

We varied the time-series' sampling rate, computing ACE, SCE and LZc for data sampled at 10, 50, 150, 350, 500, 750 and 1000Hz. The segment length and channel number was fixed at 2500 observations and 18 channels, respectively. (Note that we analysed the measures' behaviour for different segment lengths as well as different numbers of channels in our previous study [191], finding similar results for a range of different segment lengths (0.2-20s at 250Hz) and channel numbers (5-100). (See next paragraph for even shorter segment lengths.) All three measures scored for all subjects higher for WR than NREMe, for a sampling rate at least 150Hz. Consistency was considerably weaker for the low sampling rate of 50Hz and the results were totally inconsistent across subjects for 10Hz, see Fig. B.11.

When searching for the lower bound of segment length, for which the measure's behaviour was still consistent (for 18 channels at 250Hz), we found that for 500 observations (2s) per analysed segment still all measures scored for all subjects higher in WR than NREMe, as was found for 2500 observations (10s) for the main analysis. For 100, 50 and 30 observations per segment (0.4s, 0.2s and 0.12s respectively) there were at most two subjects showing higher scores for NREMe than WR for each measure. At segments that were 15 observations long (0.06s), inter-subject consistency was lost for all three measures, see Fig. B.12. Note that for fewer observations the influence of spatial signal diversity - as opposed to temporal signal diversity - on the score increases.

In summary, our findings linking low signal diversity to loss of consciousness depend on sampling data at biophysically plausible timescales and ensuring segment lengths are long enough for statistical analysis.

4.5 Discussion

We have analysed the behaviour of three measures of signal diversity on spontaneous depth electrode data recorded during wakeful resting (WR) and diverse sleep states: Lempel-Ziv complexity (LZc), amplitude coalition entropy (ACE) and synchrony coalition entropy (SCE). All three of these measures scored substantially lower (high effect size, $d > 0.8$) during early night NREM (NREMe) sleep than during WR in all 10 subjects. This is in spite of the recordings being taken from different sets of regions in each subject (as prescribed by clinical requirements for epileptic focus detection). For the majority of subjects, LZc, ACE and SCE scores during late-night NREM (NREMl) sleep were in between those for NREMe and WR. By contrast, there was no overall significant difference in the scores of any of the measures between WR and REM sleep.

Scores for the three signal diversity measures tended to be imperfectly positively correlated across subjects, indicating that they are capturing similar yet not entirely equivalent signal changes. This can be seen in model simulations in [191], which show SCE behaving in some cases differently from ACE and LZc. Further, both ACE and SCE strongly correlated inversely with delta power. The signal diversity measures do however capture more than just spectral changes between states. This was explored in detail in [191], and we confirmed this again on the present dataset, via re-computation of the measures on surrogate phase-randomised data (Fig. B.9). When analysing the measures on frequency restricted data, we found that the changes in signal diversity were most pronounced in the delta and alpha frequency bands.

The pattern of results obtained was in almost all cases preserved when the measures were computed across groups of channels restricted to a single cortical lobe (or specified sub-cortical region, see Fig. 4.6). Scores for NREMe were almost always lower than for WR, and exceptions to this followed no discernible anatomical pattern.

We also tested whether there were observable differences in local signal diversity between different cortical lobes. We found that on average the measures score higher for channels located in frontal cortex compared to parietal, temporal or occipital cortex, irrespective of the state. Given that there are substantial differences in anatomical structure between brain regions - e.g. notably structural connectivity tends to be much more dense within more anterior cortical regions [156, 114] - our results are suggestive of the denser connectivity of frontal cortex supporting increased signal diversity.

Significant differences in NREM sleep electrophysiology between cortical lobes have been observed previously [7, 163], namely inhomogeneities in the direction of propagation and frequencies of synchronous sleep spindles and slow waves. However, when comparing different regions in terms of the magnitude of the difference in signal diversity between WR and NREMe, we found no evidence for differences between regions. Thus any inhomogeneities in sleep spindle and slow wave events are unlikely to have significantly affected signal diversity scores in our data. This is likely due to the signal diversity measures' scores being averages across channels and segments, and thus being sensitive to steady state properties and not to transient events like spindles. Other studies on slow waves did also find regional differences in slow wave propagation, for example waves originating

more frequently in frontal regions, [211], but brain areas consistently recruited by the slow oscillation have not been identified [75]. This is in line with our observations on the homogeneity of delta power for single channels (Appendix B and Fig. B.5) and a study by Cavelli et al. [56], showing that sleep-stage dependent spectral coherence varied equally across the cortex.

Recent studies of avalanche events have shown that avalanche size distributions follow a power law to good approximation over a wide range of scales, during both wake and sleep states [182, 185] while large avalanches are more frequent in NREM than WR. We replicated the summary statistics from the study in [182] which utilised a dataset similar to the present one. Specifically, we also found that for all states the avalanche size distribution followed approximately a power law. Similar proximity to a power law does not imply that avalanche distributions during each sleep stage are similarly scale-free, since power-law scaling is not necessarily indicating scale-free relations as was shown in general for thresholded stochastic processes by Touboul and Destexhe [236] and discussed by [181, 182]. Unequivocally, we found mean avalanche size, duration and branching parameter all to be greater during NREMe sleep than WR or REM, indicating increased correlations between cortical areas in NREMe sleep and more fragmented cortical dynamics in WR and REM. This is in line with our finding of higher signal diversity as captured by ACE, LZc and SCE for WR and REM than NREMe. More large avalanches in NREM than WR contrast with the finding that the EEG response to transcranial magnetic stimulation (TMS) is less widespread during NREM sleep than WR [53]. However, this holds for low intensity perturbations only, as for high enough perturbational intensity the response signal spreads as far as during WR but less diversely. A possible explanation is that widespread sub-cortical drive is responsible for the increase in large avalanches, while a simultaneous decrease in cortico-cortical effective connectivity is responsible for the drop in responsiveness to TMS stimulation, in line with relay models of thalamo-cortical interaction [154, 68]. Further, if the subcortical drive results in more stereotypical cortical activity then that could explain the decrease in signal diversity as measured by ACE, SCE and LZc. We explored such potential mechanisms with the Stuart-Landau model of coupled oscillators as will be presented in Chapter 6.

In summary, we have found that three measures of signal diversity, capturing distinct aspects of signal diversity in space and time, all robustly decrease during NREM sleep, across local and global brain networks. Our application of these signal diversity measures represent new contributions to the statistical characterisation of cortical signals for different sleep stages, providing well defined signatures. Importantly, complementing decreased signal diversity of cortical response activity to TMS stimulation [53], our measures provide direct evidence for a breakdown of differentiation between regions and diversity of brain states explored, during states of unconsciousness, as predicted by integrated information and complexity theories of consciousness [229, 200].

Chapter 5

Diversity of spontaneous MEG signals increases for psychoactive doses of ketamine, LSD and psilocybin

5.1 Abstract

What is the level of consciousness of the psychedelic state? It has been proposed that certain measurements of brain dynamics, such as the repertoire of connectivity motifs that form and fragment across time, separate different states of consciousness, including the psychedelic state. Empirically, measures of signal diversity such as entropy and Lempel-Ziv complexity score higher for REM sleep and wakeful rest than for states with lower conscious level like non-rapid eye movement sleep or propofol-induced anaesthesia, when applied to spontaneous neural signals. Here we use these measures to examine whether the diversity of spontaneous magnetoencephalographic (MEG) signals changes during altered states of consciousness induced by three psychedelic substances: psilocybin, ketamine and lysergic acid diethylamide (LSD). We find reliably higher spontaneous signal diversity across all three manipulations, even when controlling for accompanying spectral changes. These changes are brain-wide though with a focus in occipital-parietal regions. We also uncover selective correlations between changes in signal diversity and phenomenological reports of the intensity of psychedelic experience. This is the first time that these measures have been applied to the psychedelic state and, crucially, the first time they have yielded scores exceeding those for normal waking consciousness. Our findings have implications for relating quantitative measures of neural diversity to phenomenological properties of consciousness, and they refine our understanding of how changes in conscious level relate to changes in conscious content.

5.2 Introduction

Recently, following early suggestions that increased conscious level may be related to an increased range of conscious contents [201, 36], there has been growing interest in characterising how conscious level and conscious content may relate [26, 170]. One empirical approach to this question is to apply emerging measures of conscious level to experimental manipulations that primarily affect conscious content. Here, we capitalise on the profound effects on conscious phenomenology elicited by psychedelic compounds, specifically LSD, psilocybin, and ketamine. These drugs normally have profound and widespread effects on conscious experiences of self and world. More specifically, they appear to "broaden" the scope of conscious contents, vivifying imagination [50] and positively modulating the flexibility of cognition [51]. At the same time, they are not accompanied by global losses of consciousness or the marked changes in physiological arousal as seen in sleep or anaesthesia. These observations raise the question of whether theoretically-grounded measures of conscious level would be changed in the psychedelic state.

One disadvantage of the PCI approach is that it requires brain stimulation, which limits its applicability. A complementary approach is therefore to measure signal diversity of spontaneous neural activity recorded under various manipulations of conscious level. Following early studies of anaesthetics [253, 99, 100] and natural sleep states [150, 46], we have found reliable reductions in neural signal diversity with diminished conscious level across a range of measures and experimental manipulations, focusing on spontaneous electrophysiological recordings. These measures include versions of Lempel-Ziv complexity (LZc, LZs), which derive from the (lack of) compressibility of binarised matrices of observations; amplitude coalition entropy (ACE), which reflects the entropy over time of most active channels, and synchrony coalition entropy (SCE), which reflects the entropy over time of synchronous channels. These measures of signal diversity robustly index levels of propofol sedation [191] and sleep stages [8, 192] when applied to spontaneous electrophysiological recordings. As with the PCI studies, these measures were reliably higher for conscious than for unconscious conditions.

Measures of entropy and Lempel-Ziv complexity both capture the diversity of a signal. The concepts are closely related, as can be illustrated with a binned neuronal spike train, for which entropy and Lempel-Ziv complexity are asymptotically identical [6]. A binned spike train is a binary string and the entropy of the process generating this string is defined using the probability distribution of the appearance of each possible substring of each possible length (which is in practice impossible to compute as it would require infinitely extensive sampling). In the limit of an infinitely long binary string, the string's Lempel-Ziv complexity [146] is almost surely identical to the entropy of the process generating the string, provided the process is ergodic. Interestingly, Lempel-Ziv complexity for certain processes converges for reasonably small sample lengths and can thus be used as an approximation of the entropy of the process that generated the string, as was shown in simulations [6]. This suggests that Lempel-Ziv complexity may more accurately estimate the true entropy of a process than direct approximate entropy measures. The latter approximate the true probability distribution of any substring of any length by only esti-

mating the frequency of some particular substrings, often called symbols [150]. Such an approximate entropy measure depends only on the histogram of the different symbols and is thus invariant under reordering of the symbols. By contrast, ordering the string of symbols such that they are grouped by type would substantially reduce the string's Lempel-Ziv complexity, further suggesting that Lempel-Ziv complexity estimates the true entropy of a signal more accurately than direct approximate entropy measures.

Functional MRI-based measures of entropy have previously been found to be greater in the psychedelic state than in normal waking consciousness [49, 222, 141] and this effect has been related both theoretically [49] and empirically [141], to the phenomenal qualities of the psychedelic state. Given that Lempel-Ziv complexity can quantify the true entropy of certain stochastic processes more accurately than direct approximate entropy measures [6], it is arguably more sensitive for signal diversity than entropy measures that have been applied previously with psychedelics. However, and crucially, no such measures have ever been applied to data derived from EEG or MEG recordings of the psychedelic state. EEG/MEG data have far higher temporal resolution than fMRI and therefore are much better suited for signal diversity analyses.

Here, we sought to test the hypothesis that three different psychedelic drugs (psilocybin, LSD and sub-anaesthetic ketamine), known to produce unusual altered states of consciousness, characterised by rich phenomenal content, would yield scores of signal diversity exceeding those for normal waking consciousness.

We did this by re-analysing multidimensional spontaneous MEG recordings using our measures of spontaneous signal diversity. We compared signal diversity for two conditions: post-placebo and post-psychedelic drug. The data come from three different experiments, in each a different drug was administered to different participants: lysergic acid diethylamide (LSD) [52], ketamine (KET) [160] and psilocybin (PSIL) [159]. We further examined whether changes in measured signal diversity could be related to subjective phenomenological descriptions obtained following drug administration, to shed additional light on the complex relations linking conscious level and conscious content.

5.3 Methods

5.3.1 Data and preprocessing

5.3.1.1 Overview

We re-analysed MEG recordings from healthy subjects with open eyes, after taking a placebo and after taking a psychedelic drug. The data come from three different experiments, in each a different drug was administered intravenously to different participants: lysergic acid diethylamide (LSD) [52], ketamine (KET) [160] and psilocybin (PSIL) [159]. After artefact removal and source modelling (see following sections for details), we analysed 2s segments of 90 source channels at 600Hz: 5-7min data for 15 participants for LSD, 6-10min data for 19 participants for KET and 2-5min data for 14 participants for PSIL, each time comparing resting state MEG for the drug condition with a placebo condition. Prior to computing signal diversity measures, the data was further low-pass filtered with

lowcut at 30Hz.

5.3.1.2 Ethics statement

All studies were approved by a UK National Health Service research ethics committee and participants gave informed consent to participate.

5.3.1.3 Participants and Drug dose

For all three studies (PSIL [159], KET [160], LSD [52]) participant exclusion criteria have been described in detail. Briefly: younger than 21 years, pregnancy, personal or immediate family history of psychiatric disorder, substance dependence, cardiovascular disease, claustrophobia, blood or needle phobia, a significant adverse response to a hallucinogenic drug or a medically significant condition rendering the volunteer unsuitable for the study. All participants had previous experience with a hallucinogenic drug but not within 6 weeks of the study (for LSD and PSIL only). For KET the additional exclusion criteria were to be smoker, being female or having a body mass index outside the range of 18-30 kg/m^2.

LSD and PSIL were each administered intra-venously at a fixed single dose of $75\mu g$ and $2mg$, respectively, within one minute. By contrast KET was administered with an initial bolus of $0.25mg/kg$ delivered over one minute followed by maintenance infusion at a rate of $0.375mg/kg/h$ for forty minutes. PSIL and KET data were obtained immediately after drug administration whereas for LSD the data were obtained four hours after drug administration due to LSD's slow pharmacodynamics.

5.3.1.4 Data acquisition and preprocessing

Participants lay in supine position for KET and LSD but were seated for PSIL. Participants' pulse rate and blood oxygenation level were continually monitored throughout the experiment via a probe over their left hand index finger. Whole-head MEG recordings were made using a CTF 275-channel radial gradiometer system sampled at 1200 Hz ($0-300Hz$ band-pass). An additional 29 reference channels were recorded for noise cancellation purposes and the primary sensors were analysed as synthetic third-order gradiometers. For LSD and KET, in addition to the MEG channels, we recorded participants' ECG: horizontal and vertical electro-oculograms as well as electromyograms from bilateral frontalis and temporalis muscles and participant compliance was in addition monitored via an eye-tracking camera.

All MEG recordings were band-pass filtered ($1-150Hz$), downsampled to $600Hz$ and segmented into epochs of 2s length. Each epoch was then visually inspected, and those with gross artifacts (e.g. head movements, jaw clenches) were removed from the analysis. An automated algorithm was used to remove further epochs contaminated with muscle artefacts. In this algorithm, a set of 30 gradiometer sensors were predefined at the edge of the MEG dewar (vacuum flask), as these are most likely to be contaminated by muscle artefacts. Using Hanning windowed fourier transformations, the mean spectral power for these sensors in the $105-145Hz$ frequency band for each epoch was calculated. If the resulting

power averaged across these sensors exceeded $10fT^2$ then that epoch was eliminated from subsequent analysis. On the remaining epochs independent component analysis (ICA) was performed, as implemented in Fieldtrip/EEGLAB, to identify and remove ocular, muscle and cardiac artifacts from the data. For LSD and KET, any components that showed a correlation ($r > 0.1$) in the time domain with the EOG/EMG electrodes were automatically removed, whereas these were identified manually for the PSIL data. Likewise, any components that showed correlations ($r > 0.1$) with similarly filtered EOG/EMG channels after being bandpass filtered in the range $105 - 145Hz$ were removed. Visual inspection was also used to remove artifact components.

Source modelling of the data was performed using the fieldtrip toolbox [168]. For each participant individual forward models were generated from their individual structural MRI scan [164]. In order to reduce the data, an atlas-based beamformer approach was used [115]. Broadband virtual sensor time-series were constructed using a linearly constrained minimum variance beamformer [241] at 90 cortical and subcortical seed locations as specified in the automated anatomical labelling atlas [238].

5.3.2 Spontaneous signal diversity measures

We compute the spontaneous signal diversity measures LZc, LZs, ACE and SCE and their "phase-shuffle" normalised versions LZc_N, LZs_N, ACE_N and SCE_N as defined in Chapter 2. Single channel Lempel-Ziv complexity, LZs, averaged across all channels is denoted as LZsum or LZs when clear from context.

5.3.3 Normalised spectral power and phase coherence (PC)

The behaviour of the signal diversity measures is compared to that of normalised spectral power and phase coherence. We defined power bands as: $\delta = 1 - 4Hz$, $\theta = 4 - 8Hz$, $\alpha = 8-15Hz$, $\beta = 15-30Hz$, $\gamma = 30-70Hz$. The power of a spectral band was computed using Welch's method [247] for each 2s segment of each of the 90 sources, normalised by the sum of the power of all 5 bands, then averaged across sources and trials per participant.

As a measure of synchrony we use the mean phase coherence (PC) across all pairs of channels. Let $z_k(t) = r_k(t)e^{i\theta_k(t)}$ and $z_j(t) = r_j(t)e^{i\theta_j(t)}$ describe the analytic signals of two channels at time t then

$$PC_{kj} = \left|\frac{1}{T}\sum_{t=1}^{T} e^{i(\theta_k(t)-\theta_j(t))}\right|, \qquad PC(\text{system}) = \frac{1}{\binom{N}{2}}\sum_{0<k<j\leq N} PC_{kj} \quad (5.1)$$

with T being the length of the segment (1200 observations= 2s) and $N = 90$ the number of sources. Reported PC scores are averages across trials. Scores of PC lie between 0 and 1, with 1 indicating perfect synchrony and 0 indicating no synchrony at all.

5.3.4 Questionnaire scores

Across the three experiments different questionnaires were employed for participants to retrospectively evaluate their psychedelic experience. Here, we analyse a subset of the questionnaire items that were common across all three experiments, i.e. for PSIL, KET and LSD:

strange Things looked strange.
geom I saw geometric patterns.
vivid My imagination was extremely vivid.
time My perception of time was distorted.
space My sense of size and space was distorted.
ego I experienced a disintegration of my 'self' or 'ego'.
muddle My thinking was muddled.
merge I experienced a sense or merging with my surroundings.
control I feared losing control of my mind.
spirit The experience had a spiritual or mystical quality.
peace I felt a profound inner peace.
float I felt like I was floating.
past I saw events from my past.
sounds Sounds influenced things I saw.

For all three experiments subjective questionnaires were completed retrospectively on the day of the experiments after most drug effects had subsided. For KET and PSIL this was typically an hour after drug delivery had ceased but was approximately ten hours after LSD due to its relatively prolonged pharmacodynamic profile. Participants answered each question using a visual analogue scale format with a bottom anchor of "no, not more than usually" and a top anchor of "yes, much more than usually". In addition, we consider the mean score over all these questions as an index of overall intensity of the psychedelic state. We call this index "total".

5.3.5 Statistics

Analyses were performed using non-overlapping segments of length 2s for a total length between $2min$ and $10min$ of MEG recording per participant and state. For each segment, the signal diversity measures ACE, SCE and LZc were computed for 30 random picks of 10 channels, and the mean across these 30 scores was considered the score for the segment. For LZs the mean across all 90 channels was set as score for a segment. The mean and standard error of the diversity measures' scores were computed over these segments. At the single participant level, the effect size of differences between states was measured using Cohen's d [64]. We call an effect size high if $d > 0.7$. For group level comparisons, a two-sided t-test was applied, with Bonferroni correction (by the number of measures) where indicated.

5.4 Results

5.4.1 Increased spontaneous signal diversity for all three drugs

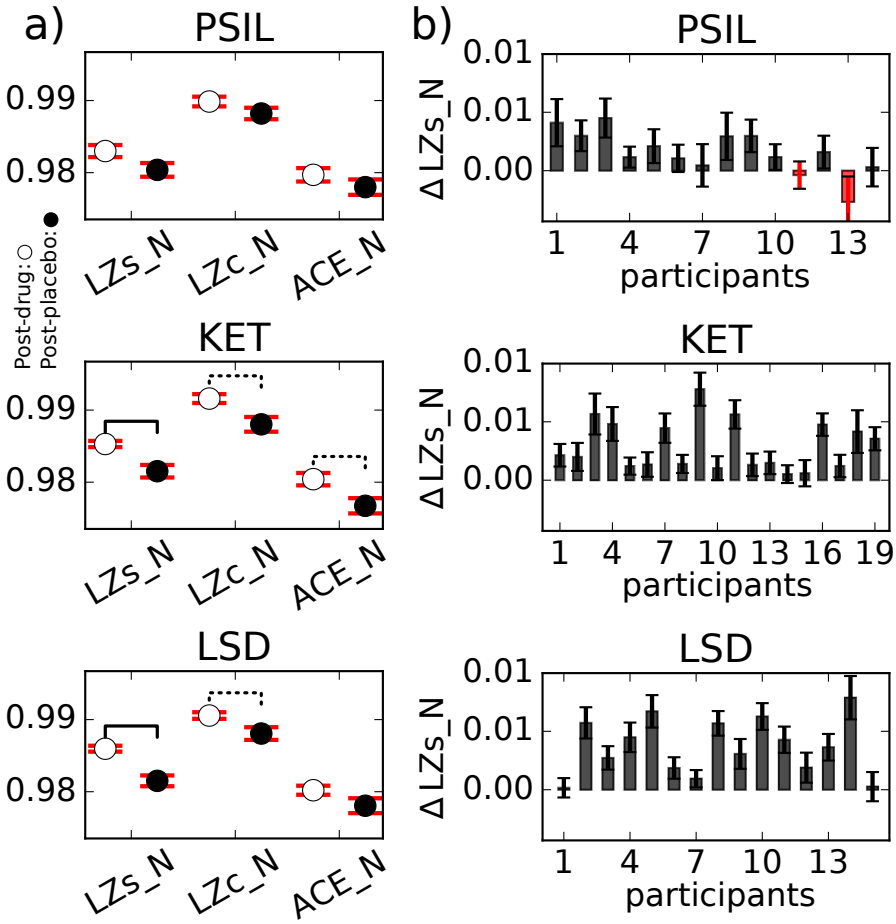

Figure 5.1: **Increased spontaneous signal diversity for PSIL, KET and LSD.** a) Mean scores across participants for the signal diversity measures LZs_N, LZc_N and ACE_N are higher for each of the three drug conditions (white discs) than for the corresponding placebo conditions (black discs). A solid line across conditions indicates $p < 0.001$ and a dotted line $0.001 < p < 0.05$, uncorrected, obtained from a two-sided t-test. b) Each panel displays the increase in single channel Lempel- Ziv complexity with PSIL, KET and LSD, respectively, compared to placebo (ΔLZs_N). The error bars indicate standard error across trials. Except from two participants for PSIL (indicated in red), the average LZs_N score across trials was higher for the drug than for placebo, for all participants and drugs. A t-test across participants gave (uncorrected) $p = 0.1891,\quad 0.00085,\quad 0.00004$ for PSIL, KET, LSD respectively (compare with Tab. 5.1).

Our main question was whether spatio-temporal signal diversity of MEG recordings increases in the psychedelic state, so we computed LZs, LZc, ACE and SCE for all three drug and placebo conditions and compared their scores on the participant and group level. To exclude the possibility that any observed changes could be attributed to changes in overall spectral profile, we also computed these diversity measures normalised by their scores for phase-shuffled data, denoted by "_N" in the measure's name: ACE_N, LZc_N,

	PSIL (14)	KET (19)	LSD (15)
ACE	12, [1,7,6]	18, [0,6,13]	14, [0,2,13]
ACE_N	10, [0,14,0]	17, [0,19,0]	10, [0,15,0]
LZc	12, [0,12,2]	19, [0,11,8]	14, [0,9,6]
LZc_N	11, [0,14,0]	19, [0,19,0]	12, [0,15,0]
LZs	12, [1,5,8]	17, [0,5,14]	15, [0,1,14]
LZs_N	12, [0,14,0]	19, [0,19,0]	15, [0,15,0]
SCE	8, [2,10,2]	17, [2,9,8]	8, [3,7,5]
SCE_N	4, [0,14,0]	10, [1,18,0]	4, [0,15,0]
PC	9, [0,10,4]	5, [9,9,1]	10, [3,5,7]
δ	10, [1,11,2]	5, [9,10,0]	4, [6,8,1]
θ	9, [2,8,4]	13, [2,7,10]	3, [11,4,0]
α	2, [9,4,1]	1, [15,3,1]	1, [12,3,0]
β	2, [5,8,1]	4, [8,10,1]	5, [8,5,2]

Table 5.1: **Effect size comparison per measure and state pair.** For each measure and drug, the first digit equals the number of participants for which the measure was higher in the drug condition than placebo. The number triplet lists for how many participants the measure scored lower with high effect size (left), higher with high effect size (right) or with low effect size (middle). Effect size is said to be high if Cohen's d was greater than 0.7. The cell-colour of a table entry is green, if the uncorrected p-value of a t-test across participant, for conditions placebo and drug, was below 0.05. If the Bonferroni corrected p-value was below 0.05, the cell-colour is yellow.

LZs_N and SCE_N.

Across all drugs and all measures, we found increased signal diversity in the psychedelic state as compared to the placebo condition, with most comparisons reaching statistical significance. The largest increase for all three drugs when compared to placebo was found for LZs_N, with higher average scores for 12/14, 19/19 and 15/15 participants for PSIL, KET and LSD respectively (Fig. 5.1b), resulting in higher LZs_N at the group level with p values at 0.19, 0.0009 and 0.00004, respectively, (Fig. 5.1a). Notably, the two participants who did not have higher scores (in PSIL) also had the lowest average score across all subjective ratings of the intensity of the psychedelic experience ("total") among all participants (see also Sections 5.3.4 and 5.4.3). Considering other measures, LZc_N and ACE_N also scored higher for the majority of participants for all three drugs with $p<0.05$ for KET and LSD (LZc_N only, see (Fig. 5.1a). The weaker consistency in the behaviour of LZc_N and ACE_N suggests that changes in spatial signal diversity across channels is weaker than changes in temporal signal diversity as captured by LZs_N, when comparing the psychedelic state against placebo.

For comparison with the signal diversity measures, we computed phase coherence and normalized spectral power for various frequency bands, for all three drugs and their placebo condition. All 90 sources were used to compute these measures (Section 5.3.3). For each participant, drug, and measure, Cohen's d was used to estimate effect sizes. Tab. 5.1 lists for each measure and drug the number of participants with higher score of the measure in the drug condition and further how many of those had different scores with high effect size, i.e. $d > 0.7$. Tab. 5.1 further indicates the number of participants for which each mea-

sures' score was higher for the drug condition than placebo and indicates p values for the measure's change at the group level. For all measures, effect sizes at the single participant level are lower for the normalisation using phase-shuffled as opposed to time-shuffled data, showing that the diversity measures depend to some extent on average spectral changes. Strongest changes in the average power spectrum were seen for the alpha band ($8-15Hz$), which was decreasing for all drugs with high effect sizes for the majority of participants. As an indicator of the global level of signal synchrony we computed phase coherence (PC) and found inconsistent behaviour. Neither SCE nor SCE_N showed a consistent change across drugs and scores of SCE and SCE_N also differed substantially.

To examine whether changes in signal diversity showed any anatomical localisation, we compared local signal diversity changes across drugs, using spectral-profile controlled single-channel Lempel-Ziv complexity LZ_N. For a given drug, participant and source channel, ΔLZ_N=LZ_N(drug)-LZ_N(placebo) was obtained as an average across all 2s data segments. The t-statistic obtained for the ΔLZ_N scores across participants was corrected for false discovery rate (FDR) and then mapped in colour onto a standard MNI brain, shown in Fig. 5.2. For all three drugs, substantial increases in LZ_N can be seen in occipital-parietal areas, despite differences in pharmacological target region and psychological effects for each drug. The regions with significant changes for PSIL also had significant changes for KET and LSD, and are thus the regions with significant changes across all drugs. These spatial distributions, with maximal locations in occipital and parietal areas are consistent with the localisation of alpha-band changes that were previously reported [52, 160, 159], although here we used significantly reduced spatial resolution in order to make the data more amenable to complexity analysis.

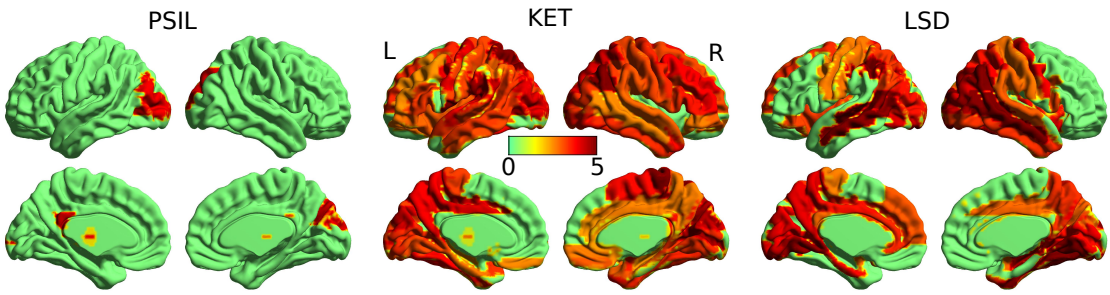

Figure 5.2: **t-statistics across participants of Δ LZs_N per source channel.** Group level changes in single channel Lempel-Ziv complexity, ΔLZs_N, are indicated as t-statistic for each of the 90 source channels and 3 drugs. I.e. for a given drug and source channel, the FDR-corrected t-statistic of ΔLZs_N across participants is setting the colour of this channel's location on the standard MNI brain, using BrainNet viewer [251]. For all three drugs, an increase in LZs_N can be seen in occipital-parietal areas.

In summary, we found an robust increase in signal diversity for all three drugs at the group level, independent of average spectral changes, as measured by LZs_N, LZc_N and ACE_N, with effects being strongest for KET. We further confirm a consistent decrease in normalised alpha power for all three drugs, in line with previous analyses of this data

[52, 160, 159].

5.4.2 Correlations between neurophysiological measures

To test whether the various measures (both signal diversity measures and others) were reflecting distinct features in the data, we computed correlations between the changes in these measures, across participants, for each comparison between drug and placebo. Specifically, we computed the Pearson correlation across participants for the score differences of 14 measures (measure(drug)-measure(placebo)): the diversity measures ACE, LZs, LZc and SCE, their "_N" versions (normalised by their score for phase-randomised data), phase coherence and normalised spectral power in the delta, theta, alpha, beta and gamma band. For a given participant, trial and measure, we subtracted this measure's score for the placebo condition from that of the drug condition to obtain a score difference for one trial. The average across trials gave then this measure's difference in score for this participant. The Pearson correlation r is then computed for such scores for two measures across all participants for a given drug and indicated in colour in Fig. 5.3 if $r > 0.5$ in order to highlight moderate and stronger correlations only. For clarity, the upper triangular portion of the symmetric correlation matrices is hidden in Fig. 5.3a .

We found strong correlations across most diversity measures for all three drugs. Slight inconsistencies of these correlations across drugs show that the measures capture not identical signal features, in line with their varying behaviour as listed in Tab. 3.1. Most diversity measures show stronger correlations with normalised spectral power bands when compared to their versions normalised by their scores for phase-randomised data (indicated by "_N"), verifying that the phase-shuffling normalisation indeed reduced the measures' sensitivity to spectral changes, as intended. For all drugs there is a clear negative correlation of phase coherence and SCE, resulting from PC measuring synchrony while SCE measures asynchrony across channels. For KET and PSIL, ACE anti-correlates moderately with PC. (See Fig. C.2 for full correlation matrix.)

5.4.3 Neurophenomenological correlations

Having established that signal diversity measures increase in the psychedelic state, we next asked whether these increases were related to subjective phenomenological descriptions about the state. We therefore computed the Pearson correlation for all combinations of the measures LZs_N, LZc_N and ACE_N with the scores of 14 subjective ratings of the participants' psychedelic experience, as defined in Section 5.3.4. The results are shown in Fig. 5.3. Supporting a relation between phenomenology and signal diversity, the measure LZs_N - which showed the strongest overall response to the psychedelic state across drugs - correlates significantly with the total score of all questions ("total") for PSIL; also recall that the two participants with the lowest total score across all questions for PSIL were the two outlying participants for ΔLZs_N as shown in Fig. 5.1b (i.e. these participants did not show an increase in LZs_N under PSIL). These observations support the notion that LZs_N correlates with the intensity of the psychedelic experience. In addition, strong correlations were found between all 3 signal diversity measures and the total score across

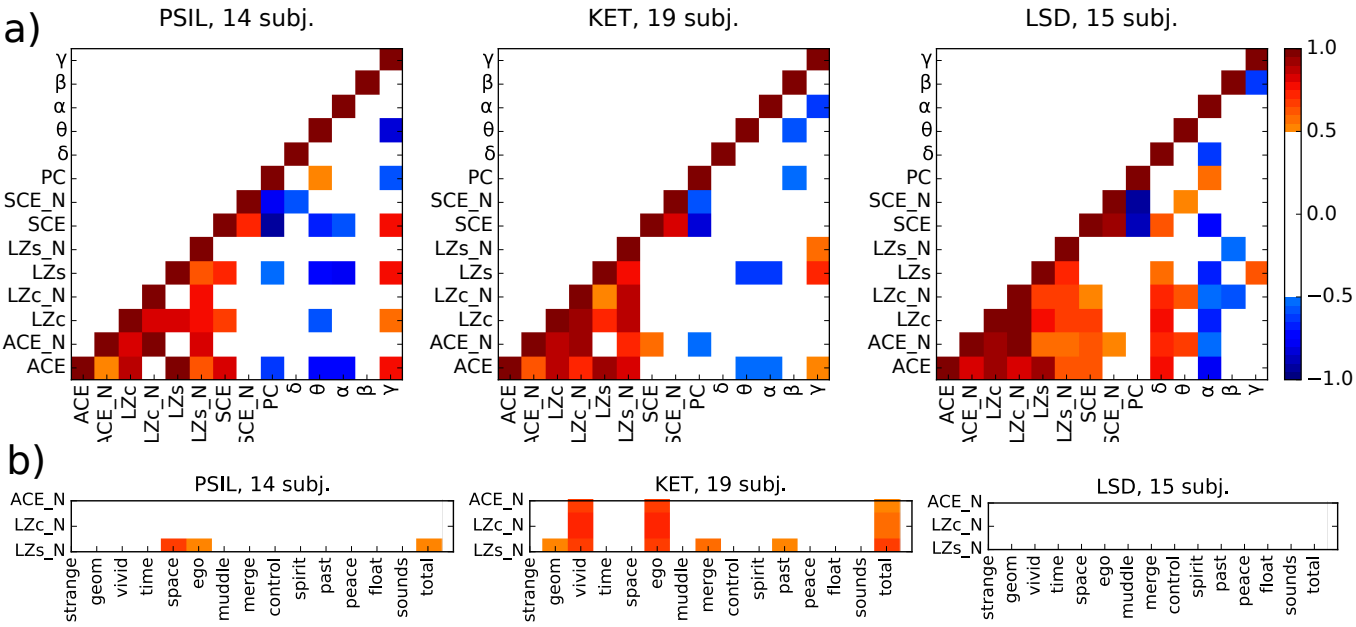

Figure 5.3: **Correlations across measures and questionnaire answers.** a) For each drug, a matrix indicates in colour the Pearson correlation, r, of the score difference between drug and placebo condition (averaged across trials) of any measure pair across participants. The upper triangular entries and entries with $|r| < 0.5$ are set white to highlight strong correlations only. These correlations indicate broad patterns only as they are not controlled for multiple comparisons as each experiment had a limited sample size. Across drugs, signal diversity measures show high correlation with each other, yet inconsistently, as they capture different flavours of signal diversity. Most signal diversity measures show more correlations with normalised spectral power bands when compared to their "_N" versions, supporting that normalisation by phase-shuffling reduced the measures' sensitivity to spectral changes. b) The measures ACE_N, LZc_N and LZs_N correlate strongly with certain questionnaire scores most clearly for KET. No consistent correlation across drugs was found for any combination of the measure's scores and scores for a particular question. See Section 5.3.4 for definition of the questions. See Fig. C.2 for full correlation matrix.

questions for KET. For this drug all three diversity measures also showed further strong correlations with specific phenomenological dimensions, in particular ego dissolution and vivid imagination. However, substantially fewer strong correlations were found between measures and specific phenomenological dimensions for PSIL and especially LSD. Indeed, for LSD no strong correlation was seen, while for PSIL only LZs_N correlated strongly with three questions. As we discuss later, the surprising absence of specific neurophenomenological correlations for LSD (given its strong effect on signal diversity measures) could be accounted for by the timing of acquisition of the subjective reports - which was much later for LSD. For certain phenomenological dimensions, such as "space", and "ego", the subjective ratings were weakest for LSD in comparison to the other drugs, whereas this was not the case for all dimensions and also, importantly, for the scores of "total" difference in phenomenology, LSD lay above KET and below PSIL, suggesting that the overall intensity of the psychedelic experience induced by LSD was comparable to the other drugs,

see Fig. C.1.

In sum, subjective ratings of the psychedelic experience correlate most strongly with spontaneous signal diversity for KET, and while some strong correlations exist also for PSIL, no neuroneurophenomenological correlations were observed for LSD. The total score across all questions, being an index of the overall intensity of the psychedelic experience, correlates strongly with LZs_N for PSIL and KET. Given the constraints on the acquisition of subjective reports these findings tentatively support a relationship between psychedelic phenomenology and neural signal diversity as discussed below.

5.5 Discussion

We have demonstrated, for the first time, that measures of neural signal diversity that are known to be sensitive to conscious level, are also robust indicators of the psychedelic state. We found that the psychedelic state induces increased brain-wide signal diversity as compared to a baseline placebo, across a range of measures and three different psychedelic compounds. The measures LZc_N, LZs_N and ACE_N all scored higher at the group level for the drugs PSIL, KET and LSD, with strongest increases seen for single channel Lempel-Ziv complexity LZs_N for KET and LSD. Importantly, by utilising phase-shuffled surrogate data we excluded that the observed increases in signal diversity could be explained by changes in the spectral profile induced by the drugs. Together, these findings constitute a new neural correlate for the psychedelic state.

Despite the differing pharmacological mechanism of action of KET, LSD and PSIL, we observed a clear similarity in the cortical localisation of changes in signal diversity measures - with relatively overlapping distributions centered over occipital and parietal cortices. These areas are strikingly similar to the locations of alpha power decreases that we have previously reported [52, 160, 159], yet spectral changes alone do not account for changes in signal diversity as our "phase-randomisation" control showed. The primary psychedelic effects of LSD and PSIL are thought to be mediated via 5HT2A receptors [113], which although distributed throughout the neocortex, have somewhat higher levels in occipital-parietal areas [94]. Conversely, KET's primary mechanism of action is as an NMDA antagonist whose receptors are located quite ubiquitously across the cerebral cortex [66, 122]. One could speculate that some of the shared phenomenological and electrophysiological effects of these drugs may be mediated by the known interactions between 5HT2A receptors and NMDA receptors [96, 10]. This could account for the localised and overlapping areas of both signal diversity and decreased alpha power that we have observed with all three drugs. However, the non-NMDA receptor effects of KET cannot be discounted, in particular its interactions with opioid receptors and HCN channels [61, 254]. Alternatively, it may simply be that MEG/EEG is well tuned to measure changes in these cortical areas due to local synchronisation properties of the underlying circuits in these areas.

Neural correlates of consciousness are particularly valuable when they account for phenomenological properties [229, 202]. For KET, strong correlations across participants were found between scores of LZc_N, LZs_N and ACE_N and specific subjective ratings of

the psychedelic experience, in particular for ego-dissolution, vividness of the experience as well as the total score across all questions - reflecting overall intensity of the psychedelic experience. For PSIL, LZs_N showed strong correlation with three questions, including the total "intensity" score, while for LSD no strong correlations between any signal diversity measures and subjective ratings were found. The stronger correlations found for KET but not for PSIL or LSD may be connected to the different time delay between the peak effect of the drug and obtaining the subjective rating. For PSIL and LSD the subjective ratings were given many hours after the experiment, when the psychedelic effects were either much weaker or no longer present, while for KET the ratings were given while the effects of the drug were still weakly present. However, for the scores of "total" difference in phenomenology, LSD lay above KET and below PSIL, suggesting that the overall intensity of the psychedelic experience induced by LSD was not lower than for the other drugs.

While we [159, 160, 52] and others [133] have used the methodological approach of correlating across participants' subjective states with electrophysiological measures it should be noted that this is an imperfect approach. In particular, the ratings of any individual are heavily influenced by their individual biases and histories - each participant has their own yardstick for evaluating the strength of an experience. Further the retrospective nature relies on recall of the experience as a single entity and does not capture the dynamics of the psychedelic experience. Future experiments which seek to capture temporal variation in experience through use of multiple probe items or perhaps even by spontaneous self-report and retrospective coding may help to more tightly tie neurophysiological measures such as those we have used here to subjective experiences.

Correlations of perturbational and spontaneous signal diversity with conscious states support integrated information and complexity theories of consciousness that emphasise diversity of phenomenology as a key property of consciousness that must be reflected in its neural correlates [228, 229, 200, 232, 234]. Perturbational [53, 190] and - with weaker specificity and sensitivity - spontaneous signal diversity measures capture types of neural signal diversity - across broadly distributed brain regions - that correlate with changes in conscious level across a broad range of states involving diminution of overall level of consciousness from a baseline of conscious wakeful rest [191, 192]. Pragmatically, these results suggest an operationally useful one-dimensional scale for level of consciousness, with wakeful rest and REM sleep at the top and coma and propofol-induced general aneasthesia at the bottom. Our findings of *increased* spontaneous signal diversity for KET, PSIL and LSD presented here, represent the first observations of an increase in theoretically-motivated measures of conscious level with respect to the baseline of wakeful rest. These results broaden the scope of application of signal diversity measures relevant to conscious level, showing that the one-dimensional scale extends in both directions from the baseline state. While it may be tempting to describe the psychedelic state as a "higher" state or level of consciousness on the basis of our findings, any such description needs to be cautiously interpreted. The measures applied in this paper focus on signal diversity, rather than the simultaneous existence of integration and differentiation (or integrated information) that are emphasised within complexity and integrated information theories of conscious-

ness. Further research should examine how measures that more directly reflect co-existing integration and differentiation, such as the PCI index, behave in the psychedelic state.

Our results may also be interpreted as supporting the notion that neural signal diversity indicates the variation of inner mental life over time. On this view the average duration of a conscious percept is decreased and /or the qualitative differences across conscious percepts is increased in the psychedelic state, in line with the entropic brain hypothesis [50].

More generally, our results provide an example of how quantitative measures of neural dynamics can bridge the gap between studies of conscious content and studies of conscious level. In terms of phenomenology, the primary features of the psychedelic state have to do with changes in conscious content, rather than global alterations of level of consciousness as seen in sleep and anaesthesia. Our findings are in line with intuitive suggestions that increases in conscious level correspond to increases in the range of possible conscious contents [26, 36]. Recently, efforts to finesse the relationship between level and content have been made in the context of integrated information theory [105] and in multidimensional descriptions of conscious level [26]. Interestingly, Bayne et al. suggest a multidimensional classification of conscious levels, with one dimension being for example the depth of context an observer grasps on average, for example whether only the colour of an object is perceived, or whether the purpose of the object is also perceived. Distinctions like this may be useful in characterising the phenomenology of the psychedelic state. Further research into the relation between phenomenology and the detailed expression of measures of complexity and diversity at local and global levels will help refine and constrain these emerging ideas.

In sum, we found increased global signal diversity for the psychedelic state induced by KET, PSIL or LSD, suggesting the psychedelic state to lie above conscious states such as wakeful rest and REM sleep on a one-dimensional scale defined by neural signal diversity. Our findings are relevant for the definition of conscious level as an average property of instantaneous experience, given that participants in the psychedelic state are fully conscious yet their experience differs substantially from normal wakefulness.

Chapter 6

Signal diversity measures of consciousness compared across networks of coupled Stuart-Landau oscillators

6.1 Abstract

Measures of spontaneous and perturbational signal diversity were found to score higher for conscious than unconscious states, when applied to distributed electrophysiological signals from the human cortex. Here we apply our spontaneous signal diversity measures ACE, SCE and LZc and Casali et al.'s perturbational complexity index PCI to time series of coupled Stuart-Landau oscillators. The measures' behaviour was investigated as a function of mean coupling and topological changes in the connectivity of the oscillators. We found networks for which an increase in the coupling strength of some oscillators only induced a decrease of all three spontaneous signal diversity measures, which was not observed for homogeneously increased coupling strength or topological changes at fixed mean coupling. Using a direct comparison between oscillator dynamics and EEG for wakeful rest and deep sleep we describe strengths and limitations of the Stuart-Landau model to produce signal diversity changes as observed empirically for different levels of consciousness.

6.2 Introduction

As introduced in detail in Chapter 1, various measures that quantify signal diversity across space and time have been shown to index clinical levels of consciousness when computed for distributed electrophysiological signals from the cortex. These measures include, most prominently, the perturbational complexity index (PCI) that was shown to index a wide range of clinically defined states of consciousness with unprecedented precision (see Section 1.4.9 and [131]), including intermediate levels of sedation and subtle cases of disorders of consciousness. Also our measures of spontaneous signal diversity, ACE, SCE and

LZc, robustly index several conscious states: propofol sedation (Chapter 3), non-rapid eye movement sleep, REM sleep, wakeful rest (Chapter 4) and states induced by psychoactive substances (Chapter 5). Here we consider our spontaneous signal diversity measures and PCI together for a simulation of electrophysiological signals from the cortex. Our aim is twofold: Firstly we want to deepen the understanding of differences and commonalities of the signal diversity that each of these measures captures. Secondly, we want to explore mechanisms to alter these signal features all at the same time and draw analogies to macroscopic coupling changes in the brain that may accompany the loss of consciousness.

Here we reach towards these goals by exploring a computer simulation of coupled Stuart-Landau oscillators and assessing the signal diversity of the oscillators' activity time series. The Stuart-Landau model can be seen as a generalisation of the famous Kuramoto model, that we used already in Section 2.6.3. The latter was introduced more than 40 years ago by Yoshiki Kuramoto [135] and has since been widely studied as a non-linear system to investigate synchronisation [219] and used as the canonical oscillator model for brain activity [193, 40, 74]. Oscillators in the Kuramoto model are complex numbers that rotate around the origin at a fixed distance but with varying speed, influenced by other oscillators. Interactions depend sinusoidally on the phase difference between each pair of oscillators.

The Kuramoto model is of limited use as we want to perturb the system (in order to compute PCI) and model each oscillator's complex time series as the analytic signal of an electrophysiological signal, requiring that each oscillator's amplitude varies with time. These requirements are met by the Stuart-Landau model, a generalisation of the Kuramoto model, with each oscillator being characterised by time-varying amplitude in addition to phase. The Stuart-Landau model is a canonical model for interacting components that individually display limit-cycle oscillatory behaviour [157] and like the Kuramoto model, it has been used to investigate synchronisation in general [194, 143] and to model neural dynamics more specifically. For example Popovych et al. [179] showed how time delayed feedback controls synchronisation and may be a treatment for epilepsy and by Moon et al. [157] who demonstrated that the number of connections of an oscillator to other oscillators (degree) determines if an oscillator leads or lags in activity. They further suggest a degree-dependent coupling reduction to model the loss of consciousness as will be disussed further in Section 6.4.2.

We want to assess PCI (Section 2.2), we perturb the oscillator model - in analogy to trans-cranial stimulation (TMS) of the brain - and assess the signal diversity of the average response signal. We use the pre-perturbation network activity to asses spontaneous signal diversity - via ACE, SCE and LZc - and post-perturbation activity to assess perturbational signal diversity. We model for the first time global effects of TMS across the cortex with coupled oscillators. However, local effects of TMS on cortical activity have been modelled before with a detailed 33000 spiking neuron model [95].

Our exploration of signal diversity in the Stuart-Landau model focuses on variations of the network structure, i.e. how oscillators are coupled together. We firstly explore in general the behaviour of spontaneous and perturbational complexity as a function of homogeneously increased coupling strength for several network structures with topological

properties ranging from dense and regular to sparse and random. Next we analyse signal diversity changes under fixed mean coupling strength for various topological changes of the network, investigating our measures for systematic variations of the density of connections in random networks, the small-world index in community-structured networks and a degree-dependent weight reduction as used by [157], finding that topology influences only some of our measures whereas mean coupling clearly affects all of them.

Given that the homogeneous changes in coupling did not result in a decrease in all signal diversity measures together, which would be necessary to model a change in conscious state, we increased the complicatedness of the network changes by fixing the coupling between some oscillators while changing that of others. We designed two networks inspired by proposed coupling mechanisms at play between macro brain regions when consciousness is lost. The first network models the loss of consciousness by increasing coupling of one region only, modelling "thalamic drive", an increase of which has been suggested to correlate with the loss of consciousness [154, 103]. For the second network, a community-structured network, we investigated if a decrease in long range coupling could model the loss of consciousness, as suggested by [237, 186]. We found a decrease of the three spontaneous signal diversity measures ACE, LZc and SCE, for an increase of "thalamic drive" in the first network - supporting the "thalamic drive" hypothesis - and a drop of long range coupling in the second network - but only while strongly increasing short range coupling at the same time. PCI was not found to decrease together with spontaneous signal diversity measures here, but rather increase with mean coupling strength.

We discuss our model in direct comparison to depth electrode EEG for wakeful rest and anaesthesia, comparing activity time series and power spectra in addition to the behaviour of spontaneous and perturbational signal diversity measures. We use this comparison further to exemplify the difference of spatio-temporal signal diversity and temporal signal diversity only.

Collectively, our results demonstrate the extent to which coupled Stuart Landau oscillators can be used to model global neural signatures of consciousness, deepen the understanding of signal diversity measures of conscious level and contribute to the discussion about possible coupling-mechanisms at play in the brain when losing consciousness.

6.3 Methods

6.3.1 Set-up of Stuart-Landau oscillators

As a generalisation of the Kuramoto model that we introduced in Section 2.6.3, the Stuart-Landau model consists of N coupled oscillators, each a complex number, $z = |z|e^{i\theta}$ with amplitude $|z(t)|$ and phase $\theta(t)$ changing with time (Fig. 6.1). They are oscillators as they oscillate around the origin of the complex plane with natural frequency ω_j, influenced by themselves, the other oscillators of the system and noise η.

The dynamics of the oscillator model that we use here can be expressed via coupled stochastic differential equations. Either by two coupled equations of real numbers, one for the evolution of the j^{th} oscillator's amplitude, r_j, and one for its phase, θ_j:

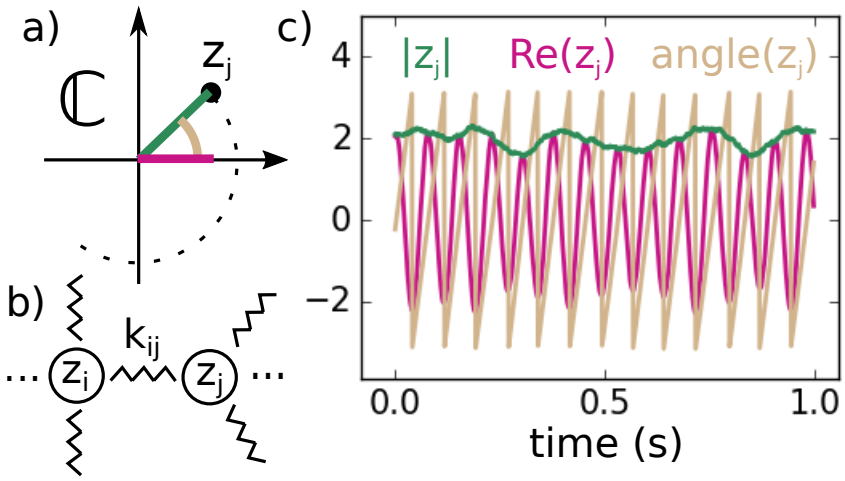

Figure 6.1: **Coupled Stuart Landau oscillators**. a) Each oscillator is a complex number, circling around the origin with natural frequency ω_j. As a generalisation of the Kuramoto model, each oscillator is characterised by phase, $angle(z_j)$, and amplitude, $|z_j|$. b) The j^{th} oscillator influences the i^{th} oscillator with coupling strength k_{ij} and fixed phase lag and is subject to noise. c) Example activity of a coupled oscillator. We use the oscillator time series of complex numbers to model the analytic signal of an EEG signal.

$$\frac{dr_j(t)}{dt} = (r_0^2 - r_j(t)^2)r_j(t) + \left(\frac{1}{N}\sum_{k=1}^{N} K_{jk} r_k(t)\cos(\theta_k - \theta_j - \alpha)\right) + f\sqrt{\eta_1(t)^2 + \eta_2(t)^2} \quad (6.1)$$

$$\frac{d\theta_j(t)}{dt} = \omega_j + \left(\frac{1}{N}\sum_{k=1}^{N} K_{jk}\frac{r_k(t)}{r_j(k)}\sin(\theta_k - \theta_j - \alpha)\right) + \arctan(\frac{\eta_1(t)}{\eta_2(t)}) \quad (6.2)$$

or, alternatively, more compactly with complex numbers in one equation for the j^{th} oscillator:

$$\frac{dz_j(t)}{dt} = (r_0^2 + i\omega_j - |z_j(t)|^2)z_j(t) + \frac{e^{-\alpha i}}{N}\sum_{k=1}^{N} K_{jk} z_k(t) + f\left(\eta_1(t) + i\eta_2(t)\right) \quad (6.3)$$

where $N = 64$ is the total number of oscillators in the system and $(r_0^2 + i\omega_j - |z_j(t)|^2)z_j(t)$ governs the dynamics of an uncoupled single oscillator with natural frequency ω_j and baseline amplitude r_0 (baseline follows from factor r_0 when considering the uncoupled steady state solution). r_0 needs to be greater than zero for the oscillator to reach a stable limit cycle, we chose $r_0 = \sqrt{2}$. This part of the equation with general parameters is also known as the normal form of the Hopf bifurcation, being the simplest model to capture essential behaviour near the Hopf bifurcation point [157][1]. $\frac{1}{N}e^{-\alpha i}\sum_{k=1}^{N} K_{jk} z_k(t)$

[1]In the mathematical theory of bifurcations, a Hopf bifurcation is a critical point where a dynamical

constitutes the influence of other oscillators on z_j's evolution, where i is the imaginary unit and K an adjacency matrix. We set $\alpha = 2\pi/10$ as a phase lag for the interaction between any oscillator pair to approximate time delay.

The Stuart-Landau model reduces to the simpler Kuramoto model by setting the amplitudes of all oscillators constant to 1 in Equ. 6.2. The Stuart-Landau model itself can be derived from the more complicated Wilson-Cowan neural mass model, in which two coupled differential equations with sigmoid activation functions model the behaviour of excitatory and inhibitory neuronal populations (see derivation in Supplement of [157]).

Complex noise is added to the evolution equation of each oscillator via η_1 and η_2, both i.i.d. $\sim \mathcal{N}(0, 1)$. If not otherwise stated, the noise level was chosen via factor $f = 1/2$, multiplying the complex noise, such that the standard deviation of an uncoupled oscillator's amplitude $|z|$ was equal to 10% of its baseline $r_0 = \sqrt{2}$. The noise-induced fluctuation of an uncoupled oscillator's amplitude increases with the square root of the noise factor f, see Fig. 6.2.

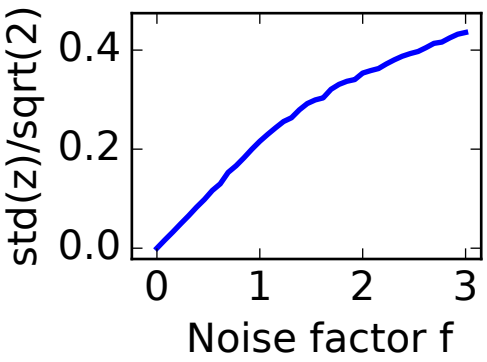

Figure 6.2: **Amplitude fluctuations as a function of noise factor** f The standard deviation of the steady-state amplitude $|z|$ of an uncoupled oscillator grows with $\sqrt{f}$, f being the noise factor multiplying the complex noise in Equ. 6.3. Values shown are averages over 10 trials and all 64 uncoupled oscillators for noise factors between 0 and 3 in 40 equal steps.

The natural frequencies of the oscillators ω are fixed at one particular sample of independently drawn $\mathcal{N}(20\pi, 15)$ variables. The natural frequency of $10 * 2\pi$ sets an uncoupled oscillator to have 10 full rotations in a time step of 1, which we call one second. This choice allows to interpret the natural frequency of an oscillator to be on average 10Hz, with standard deviation of approximately $2Hz$ simulating the alpha bandwidth (8-13Hz) of human EEG [39].

We solve the stochastic system described with Equ. 6.3 with 64 oscillators for 4.5s, initiating the oscillators' phases randomly and setting all amplitudes equal to $\sqrt{2}$. This results in 64 complex time series of 4500 observations at $1000Hz$ each. The first 1500 observations are discarded to avoid initial transients. All non-perturbational measures were computed from segments of 2500 observations, i.e. of length 2.5s. The last 500 observations were used to asses perturbational signal diversity, applying a perturbation to the system at time point 4s, in analogy to a transcranial magnetic stimulation (TMS) of the brain. At the moment of perturbation, amplitudes of 16 oscillators (with index 32 to 48) were set to 35 times uncoupled baseline, and then the system continued to evolve

system's stability switches and a periodic solution arises. Under reasonably generic assumptions about the dynamical system, a small-amplitude limit cycle in phase space branches from the fixed point.

according to Eqn. 6.3. These central oscillators were chosen for perturbation as they did not cause any bias for the used adjacency matrices. The strength of the perturbation was chosen such that approximately half of the oscillators showed a response to perturbation for intermediate coupling of a homogeneous coupling matrix. The 500 observations after perturbation contained most of the diversity in the response activity to perturbation.

Computations were performed in parallel on a high-performance cluster with basic scripts in the python programming language. The complex version of the system of coupled stochastic differential equations, Equ. (6.3), was solved with the Euler-Maruyama method, using integration step size 0.1 and sampling rate 1000Hz (defining a time step of 1 as being one second). The perturbation was implemented by adding 50 to the real parts of perturbed oscillators for one numerical time step only.

6.3.2 Measures and statistics

6.3.3 Phase coherence

As in [157], the level of synchrony of the system is measured by the mean phase coherence (PC) across all pairs of oscillators. Let $z_k(t) = r_k(t)e^{i\theta_k(t)}$ and $z_j(t) = r_j(t)e^{i\theta_j(t)}$ describe two oscillators at time t then

$$PC_{kj} = \left|\frac{1}{T}\sum_{t=1}^{T} e^{i(\theta_k(t)-\theta_j(t))}\right|, \qquad PC(\text{system}) = \frac{1}{\binom{N}{2}}\sum_{0<k<j\leq N} PC_{kj} \quad (6.4)$$

with T being the length of their time series and N the number of oscillators. There are 2016 channel pairs when choosing from 64. Reported PC scores are averages and standard error across 15 trials, with initial phases and noise varying across trials. PC is close to zero for uncoupled oscillators and close to one for highly coupled oscillators.

6.3.4 Spontaneous signal diversity

ACE, LZc, SCE and LZsum were computed as described in Chapter 2, using the complex oscillator time series instead of the complex analytic signal of an EEG signal. For all numerical experiments the measures were computed for 10 oscillators and 2500 observations (2.5s) and averaged across trials as follows. If the adjacency matrix did not contain random connections, 15 trials were computed. For each such trial, 10 oscillators were randomly picked from 64 oscillators, 10 times. The reported score of each of the 4 spontaneous signal diversity measures is thus the average and standard error of 150 values (15 times new initial phases and noise and for each such trial 10 different channel samples). If the adjacency matrix under consideration did contain random connections, this procedure was repeated 10 times and grand average scores reported.

6.3.5 Perturbational signal diversity

We assessed perturbational signal diversity from the response activity $|z|$ of oscillators for segments of 500 observations (0.5s) post perturbation. In analogy to Casali et al.'s empirical procedure (see Section 2.2), the superposition across many trials is used to separate spontaneous activity from response activity. This can be seen in Fig. 6.3a where the time series are nearly flat prior to the perturbation, showing that spontaneous activity averages to a nearly constant time series. We found that averaging across 150 trials for the same adjacency - but different initial phases and noise - sufficiently averaged out spontaneous fluctuations.

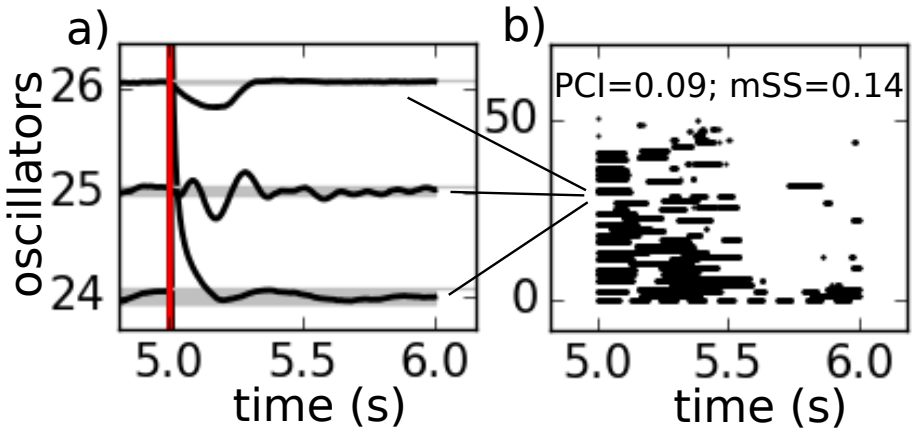

Figure 6.3: **Illustration of perturbational response.** a) Averages of three example oscillator amplitude time series $|z(t)|$, the red line indicating the time of perturbation. The bottom oscillator is directly perturbed - visible by it's high amplitude at the time of perturbation - unlike the other two. The upper edge of a grey horizontal bar indicates the binarisation threshold. b) Binarised response activity of the whole system of 64 oscillators, ordered by response activity, with PCI and mSS scores for this example. Adjacency matrix parameters are as denoted by point p3 in Fig. 6.7e.

The continuous wave forms visible as response to perturbation, illustrated in Fig. 6.3a, were then binarised using pre-perturbation activity. I.e. the binarised response activity of the i^{th} oscillator was obtained by computing $mean + 2 \times std$ of the i^{th} oscillator's averaged pre-perturbation activity of length 2500 observations and using this number as a threshold for averaged post-perturbation activity. That is, for each observation the binary response activity is 1 if $|z_i|$ was greater than the threshold, else zero. After performing this binarisation for all 64 oscillators, a binary response matrix was obtained, consisting of 64 oscillators and 500 binary observations, as exemplified in Fig. 6.3b. This is our analogy to the matrix of significant sources SS, described by Casali et al. [53] as the average response EEG activity to TMS perturbation.

From this binary response matrix SS we computed two measures to quantify the system's response to perturbation. Firstly, we denote the mean of all entries of this matrix as mSS and use it to assess the intensity of the response to perturbation. Secondly, we compute the perturbational complexity index, PCI, as defined in Section 2.2.

Each binary response matrix SS was obtained from the average system activity of 150 trials. Reported PCI scores per adjacency are average scores across 15 different binary response matrices. For adjacencies with random entries, this total procedure was repeated 10 times and grand average scores of PCI reported.

6.4 Topology versus mean coupling

We first explored the signal diversity measures' behaviour for simple parameter variations in the oscillator system, testing if simple system changes can cause spontaneous and perturbational signal diversity measures to change in the same direction, which would demonstrate signal diversity changes similar to those empirically observed during wakeful rest and NREM sleep. To this end we explored the dependence of the signal diversity measures on coupling strength and topology for several canonical networks. The topology of the network is defined by a binary adjacency matrix, defining directed connections between oscillators. In this section, the coupling strength between oscillators is homogeneously changed by multiplying the adjacency matrix by a factor. The sum of all entries of the adjacency matrix (i.e. all existing connections between oscillators) divided by the total number of oscillators squared (i.e. the number of connections in an all-to-all adjacency) is what we call "mean coupling". We will focus on the behaviour of the signal diversity measures for detailed variations of mean coupling for several fixed topologies before investigating topological changes at fixed mean coupling.

6.4.1 Noise and spontaneous signal diversity

In order to illustrate basic system dynamics as a function of noise and coupling strength, we first consider homogeneous all-to-all coupling of 64 oscillators. We justify our choice of noise level and show that if noise is present, we don't find all three spontaneous measures to change together with increasing coupling strength.

Oscillator time series, $|z(t)|$, were compared for three different values of mean coupling (0, 10, 20) and three different noise levels (0, 0.1 and 0.3, denoting the ratio of the standard deviation of an uncoupled oscillator's amplitude to its baseline). If noise is absent and the mean coupling is set to zero, all oscillators' amplitudes settle at $\sqrt{2}$, as seen in Fig. 6.4a. Increasing mean coupling diversifies the dynamics, caused by interactions with other oscillators, until the system becomes increasingly synchronised (e.g. mean coupling at 20) and for even higher mean coupling, all oscillators' amplitudes become constant (not shown). By contrast, adding noise causes diverse system dynamics for the uncoupled as well as the strongly coupled case, since noise influences each oscillator's amplitude directly in addition to other oscillators, see Fig. 6.4a. The more noise is added, the less interactions between oscillators contribute to the diversity of the system's dynamics.

Let us next consider how phase coherence PC and spontaneous signal diversity (ACE, LZc, SCE) are influenced by mean coupling and noise level. For each noise level (0, 0.1 and 0.3), mean coupling was increased from 0 to 20 in 20 steps. PC monotonically increases with increasing mean coupling for all three noise levels, while the behaviour of

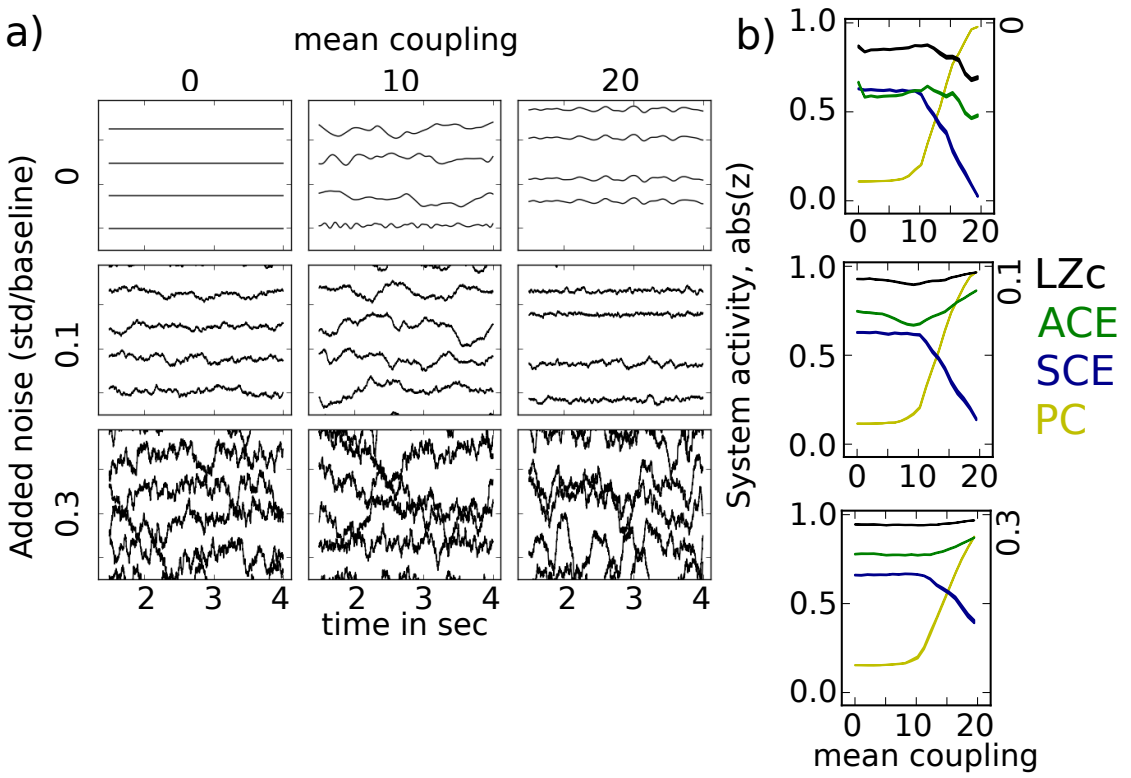

Figure 6.4: **Exemplifying system dynamics under different noise levels.** 64 oscillators were coupled all-to-all for different levels of mean coupling and noise. a) Each panel exemplifies spontaneous activity $|z(t)|$ for 4 oscillators. b) For noise at level 0, 0.1 and 0.3 and three levels of mean coupling, the measures PC, SCE, ACE and LZc were computed. Independent of noise level, PC increases monotonically with mean coupling, as the system becomes increasingly synchronised. The behaviour of LZc and ACE depends strongly on noise level. Both increase with greater coupling in the presence of noise, as noise is the main source of diversity in activity in the strongly synchronised system.

LZc and ACE differs across noise levels, as shown in Fig. 6.4b. For the case without noise, LZc and ACE show a local maximum for intermediate coupling and then decrease with increasing coupling, as the time series become increasingly synchronised. If noise is present at the 0.1 level, LZc and ACE show a local minimum for intermediate coupling and then increase with greater coupling. The increase with high coupling is also seen for noise at 0.3. This behaviour follows from noise being the main source of diversity in activity in the strongly synchronised system, resulting in more random binary sequences under ACE and LZc's binarisation by mean activity. By contrast, SCE's behaviour is barely affected by the amount of noise, which follows from its binarisation being based on the difference of phases of two oscillators being below a threshold. In contrast to the constant influence of noise on amplitude variation, noise's influence on phase variation decreases the more synchronised the system becomes as all oscillators stabilise each other in oscillating with the mean natural frequency. Importantly, in the presence of noise, the scores of the three measures do not all change in the same direction as coupling is increased, SCE behaving

differently to ACE and LZc.

In our simulations shown in Section 2.6, SCE increased with ACE and LZc for increasing signal diversity for the increasingly identical random channels and also for a broad range of parameter space for the autoregressive model. We did see though for our Kuramoto experiment that SCE can behave completely different to LZc and ACE. Hence the insight from the previous models, that SCE can behave differently to LZc and ACE, while ACE and LZc behave similarly, is again confirmed here. However, we did see all three measures behave similarly for most empirical analysis, which is thus an important criterion for modelling empirically observed spontaneous signal diversity changes for different levels of consciousness, justifying the application of all three measures.

6.4.2 Increase of mean coupling for fixed topologies

With the noise level fixed at 0.1, we investigated the dependence of the signal diversity measures on mean coupling for three more networks, with a focus on a detailed change of mean coupling. The choice of the three additional adjacencies presented in Fig. 6.5 is motivated by choosing examples that differ clearly in small-worldness (indicating a balance between network segregation and integration, ranging from maximal order, as in all-to-all connectivity, to maximal disorder, as in a graph with random connections [44]) as well as the density of connections (from maximal density in all-to-all connectivity to sparse connectivity at the limit of the graph being disconnected[2]).

The three adjacencies shown in Fig. 6.5 are defined as follows. As an example for a sparse and disordered network we chose a symmetric Erdös-Rényi random network (ER), with connection probability $p = 0.07$, chosen as small as possible such that the network is still connected. Symmetric Erdös-Rényi random graphs are constructed by setting each of the $n(n-1)$ possible edges with probability p [93, 111].

As network samples for intermediate randomness and sparse connections we chose a community-structured adjacency with maximal small-world index before becoming disconnected (see also Section 6.4.4). As an example of a dense, intermediately random network, we chose a spatially embedded network, suggested by [92, 121] to reflect mammalian cortical connectivity of axonal pathways as obtained via retrograde tract tracing. Using the procedure and parameters suggested by [92], this network is constructed by starting from placing all nodes randomly in a cube of $50mm$ side-length. Then nodes are randomly connected all-to-all with probability given by an exponentially decaying function of Euclidean distance between the nodes in the cube (exponent chosen as $-0.188mm^{-1}$). Of all these edges (connections) a fraction of 0.66 is randomly selected. Connection weights are set proportional to the connection probabilities, i.e. the greater the Euclidean distance between two nodes in the cube, the smaller the weight of that connection. This network is called spatially embedded as there is this clear map from the adjacency matrix to the 3D graph.

Fig. 6.5 displays the behaviour of PC and signal diversity measures as a function of

[2] An adjacency matrix represents a graph. A Graph is said to be disconnected if there exists at least one node from which there is no path through the graph to at least one other node.

mean coupling for the sparse ER network, the community structured network as well as the spatially embedded network. PC and spontaneous signal diversity measures were computed for 20 equally spaced values of mean coupling, ranging from 0 to 20. For all three networks PC increases and SCE decreases with increasing mean coupling, LZc and ACE show a local minimum - least pronounced for the ER network - similarly to the behaviour seen for all-to-all coupling, displayed in Fig. 6.4b, illustrating how topology only weakly influences signal diversity. Important for our later discussion on empirical modelling, for none of these four networks - all-to-all, ER, community-structured or spatially embedded - SCE was decreasing together with LZc and ACE for increasing mean coupling. Measures of perturbational complexity, mSS and PCI, only computed for 4 values of mean coupling, show both a slight increase with mean coupling for all three networks.

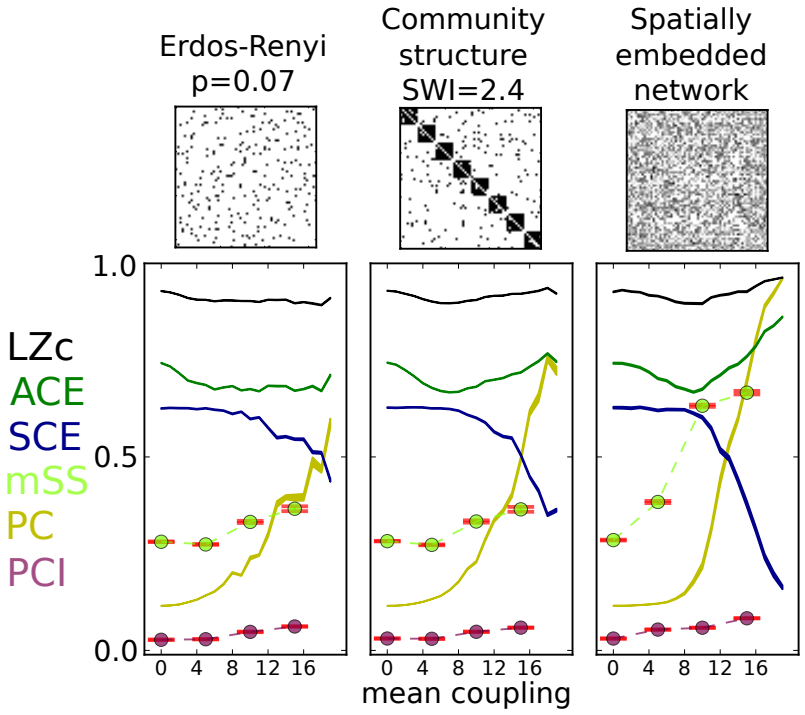

Figure 6.5: **System dynamics as a function of mean coupling.** With noise fixed at 0.1, the measures' dependence on mean coupling is shown for three networks. Their behaviour is similar to what was found for all-to-all coupling, illustrating how topology only weakly influences signal diversity. Importantly, SCE does not decrease with LZc and ACE for any of the shown networks. mSS and PCI increase with increasing mean coupling. Linewidth (or error bars) indicates standard error across trials.

The results summarised in Fig. 6.5 are focussed on the change in signal diversity with coupling increasing in small steps. Next we'll focus on further changes in topology at only three different values of fixed mean coupling. All adjacencies in this chapter were tested to correspond to connected graphs.

6.4.3 Random graph density

Aiming to investigate the influence of network density on system dynamics, we computed ACE, LZc, SCE, mSS, PCI and PC for symmetric Erdös-Rényi random graphs with connection probabilities 0.07, 0.3, 0.5, 0.75 and 1, each for three mean coupling strengths, 5, 10 and 15, chosen to correspond to low, intermediate and high phase coherence - as informed by Fig. 6.4. Erdös-Rényi random graphs are constructed by setting each of the $n(n-1)$ possible edges with probability p [93, 111].

Fig. 6.6a shows the measures's scores and adjacencies. Considering changes across mean coupling, all measures are affected (in line with Fig. 6.4b for noise being 0.1), in particular PC clearly and PCI weakly increase with increasing coupling. For fixed mean coupling at 5 or 10, no clear dependence on network density is visible for any of the measures, except an increase of mSS from sparsest to higher density, indicating a limited spread of the perturbation in the sparsest network ($p = 0.07$). However for mean coupling at 15, Fig. 6.6a, PC and ACE are increasing with network density while SCE decreases. The behaviour of LZc and PCI is barely affected by density also for fixed mean coupling at 15.

In sum we see that network density influences the dynamics more strongly at high than at low mean coupling. The spontaneous signal diversity measures ACE, SCE and LZc do not decrease together with the density of random graphs.

6.4.4 Small-world index

Dynamical complexity as measured by causal density has been found to depend on small-world index [208] in networks of spiking neurons, causing us to investigate the influence of small world index on our signal diversity measures for coupled oscillators. The small-world index for an adjacency K was computed as described in [208]. Briefly, the ratio of clustering coefficient and average shortest path length of K was divided by the ratio of clustering coefficient and average shortest path length of a random graph with the same degree distribution than K (the average of 100 trials was used; implemented with the networkx library in python [111]). Small-worldness is further a widely discussed property of brain connectivity networks. E.g. based on diffusion magnetic resonance imaging, Hagmann suggested a method to obtain an adjacency matrix to reflect average network properties of neuronal connections throughout the human cortex. The resulting adjacency matrix was sparse with high small world index [112].

We defined an adjacency where every 8 consecutive oscillators belong to the same community. Each oscillator is connected to an oscillator of the same community with probability p and to one from a different community with $1 - p$ (adapted from [209]). The small world index for this adjacency was found to increase with increasing p, until at $p = 0.96$ the graph became disconnected. At this limit, the maximal small world index for our network of 64 oscillators was found to be 2.4. $p = 0.84$ resulted in a small-world index of 1.0, which we used for comparison.

Fig. 6.6b displays the scores of ACE, LZc, SCE, mSS, PCI and PC for these two community-structured adjacencies with small-world index 1.0 and 2.4, respectively, for three different levels of mean coupling: 5, 10, 15. The adjacency matrix with higher

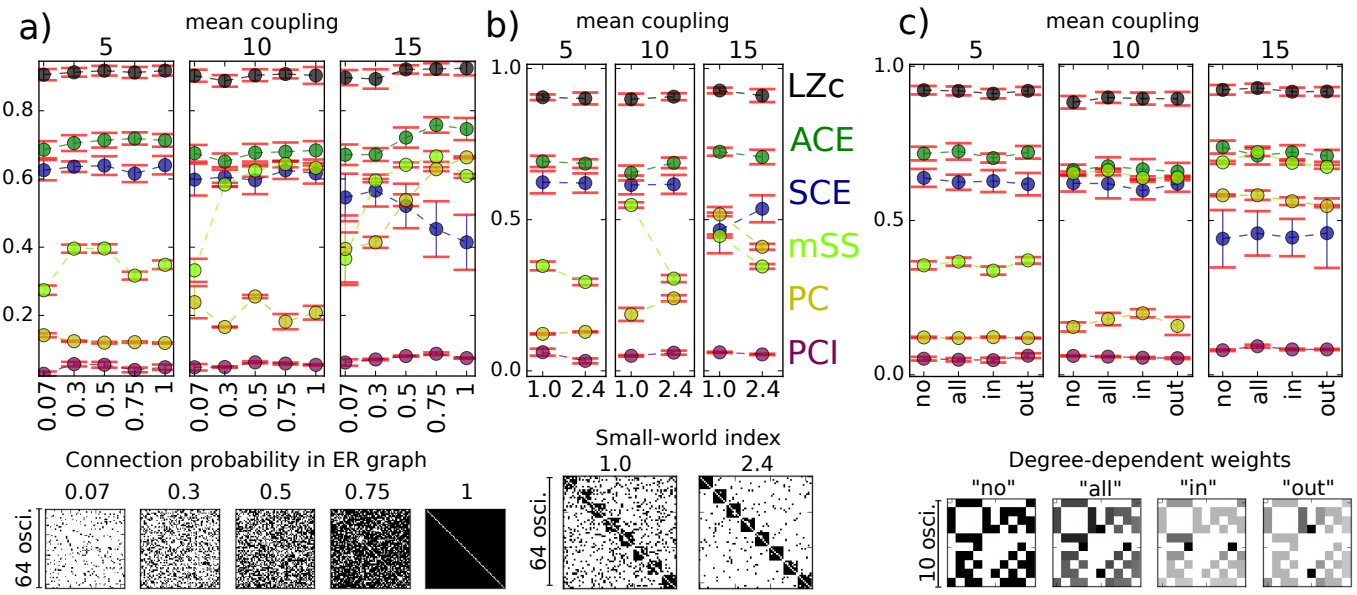

Figure 6.6: **Changing topology at fixed mean coupling** The behaviour of LZc, ACE, SCE, mSS, PC and PCI is shown for systematically changed adjacencies for 3 levels of mean coupling: 5, 10, 15. a) Symmetric Erdös-Rényi random networks with connection probability being 0.07, 0.3, 0.5, 0.75 and 1 differ in network density. PC, SCE and ACE depend on network density for mean coupling at 15, as the only consistent changes. b) Two community-structured adjacencies are compared, having small-world index 1.0 and 2.4 respectively. For all three levels of mean coupling, only mSS is consistently affected, scoring lower for the higher small-world index, indicating wider spread of the perturbation for more inter-community connections. c) Degree-dependent weight reduction (method from Moon et al. [157]) has no influence on any of the measures tested here. For clarity, adjacencies are shown for first 10 oscillators only. Linewidth indicates standard error across trials.

small-world index has visibly more intra- than inter-community connections. For all three coupling levels, mSS is lower for the higher small-world index, indicating a wider spread of the perturbation in the more densely connected network (as also seen for Fig. 6.6a). The three spontaneous measures, PC and PCI are not consistently affected by the change in small-world index, yet vary with the level of mean coupling, again in line with Fig. 6.4b.

In sum, only mSS's behaviour was consistently affected, scoring lower for the higher small-world index for all tested levels of mean coupling. This indicates that the response to the perturbation spreads further if there are more inter-community connections.

6.4.5 Degree-dependent coupling weight

Moon et al [157] claim for the Stuart-Landau model that the node degree (the number of connections of a single oscillator to other oscillators) indicates if a node (oscillator) leads (low node degree) or lags (high node degree) in phase, as measured by directed phase-lag-index. Directed phase-lag between two oscillators, say $z_k(t) = r_k(t)e^{i\theta_k(t)}$ and $z_j(t) = r_j(t)e^{i\theta_j(t)}$, is defined as the mean signum of the phase difference across observations,

$$dPLI_{kj} = \langle sgn(\theta_k - \theta_j) \rangle,$$

and interpreted to indicate which oscillator leads the other one. I.e. if $dPLI_{kj} > 0$ then oscillator k leads oscillator j in activity.

Moon et al. further modelled propofol anaesthesia with Stuart-Landau oscillators by decreasing the coupling weights in a degree-dependent way and argue that the observed homogenisation in directed phase-lag-index across oscillators is similar to what one observes across EEG channels for propofol-sedated subjects [157]. Note that the degree-dependent coupling decrease that Moon et al. propose as a model for loss of consciousness clearly reduces mean coupling, the latter strongly influencing our signal diversity measures as we showed earlier in Fig. 6.4. Given that this structural change to the topology is different to the networks we considered so far and has been connected to modelling the loss of consciousness, we investigated if the degree-dependent weight reduction introduced by Moon et al. has effects on our signal diversity measures' behaviour when fixing the level of mean coupling.

The weights of symmetric Erdös-Rényi random adjacencies (connection probability 0.5) were decreased in the following three degree-depending ways (compare with Fig. 6.6c). Firstly, dividing each entry of the adjacency matrix by the number of non-zero entries of its row reduces weights by in-degree, given that the i^{th} row specifies weights with which the i^{th} oscillator is influenced by each of the N oscillators. We denote the adjacency matrix after this process as "in". Secondly, dividing each entry by the number of non-zero entries of its column reduces weights by out-degree, given that the i^{th} column specifies weights with which the i^{th} oscillator influences each of the N oscillators. We denote the adjacency matrix after this process as "out". Thirdly, dividing each entry by the number of non-zero entries of its column and row summed together reduces weights by in and out degree. We denote the adjacency matrix after this process as "all".

Fig. 6.4c shows that none of the measures was affected by any of the degree-dependent weight reductions. The adjacency changes associated with loss of consciousness by Moon et al. [157] did not influence spontaneous nor perturbational complexity as measured here, with fixed mean coupling.

Similarly as for the adjacency changes described in Fig. 6.6a,b the measures did change with increasing mean coupling level in line with their behaviour for increasing coupling of all-to-all connectivity, Fig. 6.4b. In particular, an increase in PCI and mSS with increasing coupling can be seen for all three experiments.

In summary for all three topology experiments shown in Fig. 6.6, the overall level of coupling has a visible effect on all tested measures (in line with the measures' behaviour seen for all-to-all coupling) whereas changes in topology alone influence only some measures under certain levels of coupling. SCE decreases while ACE and LZc increase for increasing density of connections for high mean coupling, as was seen for random adjacencies (compare $p = 0.3$ with $p = 1$ at mean coupling 15 in Fig. 6.6a) and also for the community-structured adjacency (Fig. 6.6b at mean coupling 15). This suggests that increasing the density of

connections while keeping the mean coupling fixed has a similar effect on the measures' behaviour to increasing coupling for fixed all-to-all connectivity. The difference of SCE and ACE/LZc's behaviour is again due to SCE being nearly insensitive to noise, unlike ACE and LZc, which are both maximal for noise dominated dynamical regimes, as we explained in Section 6.4.1. None of the explored changes in topology caused the scores of ACE, LZc and SCE to change together in the same direction, as would be required to model changes in spontaneous signal diversity as observed across state of consciousness. We found such changes for more complicated adjacency changes, as will be explained in the next section.

6.5 Modelling change of conscious level

Here we aim directly to model empirically observed signal diversity changes across wakeful rest and NREM sleep, as captured by spontaneous and perturbational signal diversity measures. We first introduce non-homogeneous coupling changes that cause the scores of ACE, SCE and LZc to decrease together - albeit not together with PCI - and then compare the signals of this model more closely with depth electrode signals obtained during WR and NREM sleep to evaluate the model's merit.

6.5.1 Modelling thalamic drive and long range coupling

The experiments with varying topology and mean coupling of Section 6.4 did not result in all spontaneous signal diversity measures' scores to change together, when noise was present. The presence of noise is important to model EEG, so we no longer focus on the case without noise for our model. Coupling changes in Section 6.4 were applied homogeneously to all oscillators for various topologies. We next explored ways to change the coupling heterogeneously, i.e. changing the coupling strength only between some oscillators while keeping the remaining connection strengths fixed.

We found a drop of all three spontaneous signal diversity measures for two adjacency matrices that both were inspired by different hypotheses about coupling mechanisms between macro brain regions causing consciousness to fade. Firstly, an increase in thalamic influence on cortical activity was suggested as a cause of the loss of consciousness [154, 103]. Secondly, several studies argue for a decrease of long range coupling when consciousness is lost [237, 186]. In addition to a drop of all three spontaneous signal diversity measures, we searched for conditions such that the oscillator time series still visually resembled healthy EEG time series. I.e. the diversity in the time series should not be entirely due to noise - which would resemble iso-electric activity of a brain-dead patient - nor too periodic as seen in epilepsy. Restricting phase coherence to lie between 0.5 and 0.7 excluded such either fully decoherent or fully coherent system dynamics for these adjacencies.

Fig. 6.7a shows our "thalamic" adjacency, chosen such that the first 12 oscillators were said to represent thalamic activity whereas the remaining 52 oscillators represent cortical activity. The adjacency matrix was divided into 3 blocks, each having homogeneous coupling strength, representing connections between macro-brain regions: s1 (thalamus driving cortex), s2 (inter-cortical coupling) and s3 (cortex driving thalamus) respectively.

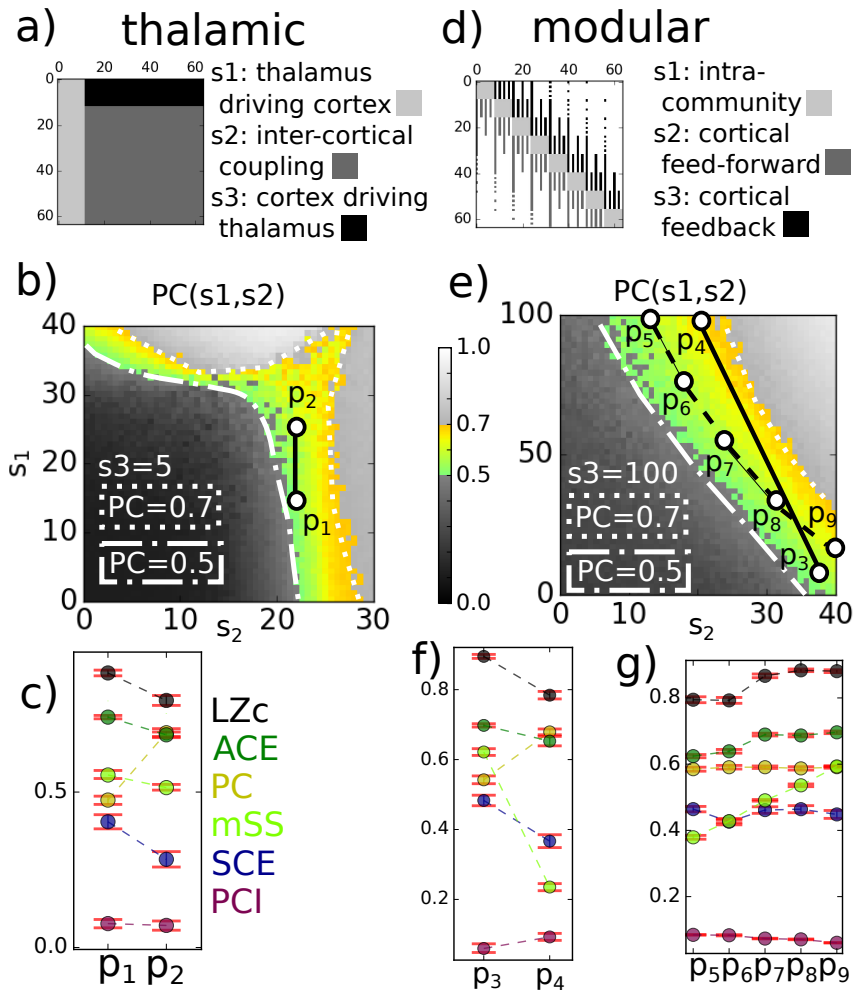

Figure 6.7: **Modelling thalamic drive and long range coupling.** Two adjacency matrices for which point pairs in parameter space were found such that all three spontanenous signal diversity measures decrease together, as observed empirically for EEG (Chapters 3 and 4). a) A block adjacency with the first 12 oscillators representing the thalamus and the remaining 52 the cortex, allowing to increase "thalamic drive" only. b) Phase coherence (PC) as a function of "thalamic drive" and "inter-cortical" coupling for a range of coupling parameters with regions greyed out where it is possible that dynamics are either fully coherent or fully decoherent, including a point pair c) for which all three measures drop. d) Modelling long range coupling with a modular network. e) Phase coherence (PC) as a function of intra-community coupling $s1$ and "cortical feed-forward" coupling $s2$ with an example point pair, for which LZc, ACE and SCE drop together f) by decreasing long range coupling and increasing intra-community coupling. The mean of the binarised perturbation response matrix (mSS) shows a decrease with ACE, SCE and ACE for the two example point pairs while PCI does not change consistently. g) A path of points where PC is constant while mSS, PCI, ACE and LZc change.

An example section of the explored parameter space is shown in Fig. 6.7b, displaying the dependence of PC on s1 and s2 while $s3 = 100$ (points are greyed out if $PC < 0.5$ or $PC > 0.7$, excluding regions where fully coherent or fully decoherent dynamics occur). $s3 = 100$ was chosen after several lower or higher values for s3 were explored and found

to induce a weaker effect of all three spontaneous measures to decrease together. An example point pair is highlighted, for which all three spontaneous measures decreased, Fig. 6.7c, while only the coupling parameter "thalamic drive" s1 increases, in line with the thalamic drive hypothesis. This hypothesis states that consciousness fades when effective connectivity between cortical regions is distorted by strong, synchronising influence from slow-wave activity in the thalamus [154, 103].

Fig. 6.7d shows a modular adjacency that divides the 64 oscillators into communities of 8, each oscillator being fully connected to all other oscillators of the same community, then there are exponentially (powers of 2) fewer connections with increasing distance between oscillator indices. The matrix was constructed to resemble a sparse small-world network with 8 communities[3]. The adjacency matrix is divided into 3 parts, each part being multiplied by a different coupling parameter: The 8 communities on the diagonal (multiplied by s1, representing intra-community coupling), all lower diagonal entries except the communities (multiplied by s2, representing cortical feed-forward coupling) and all upper diagonal entries except the communities (multiplied by s3, representing cortical feedback coupling). I.e. inter-community coupling represents long-range coupling.

The example section of parameter space shown in Fig. 6.7e illustrates a point pair (p3, p4, connected by straight line) for which LZc, ACE and SCE decrease with decreasing feed-back long-range coupling and strongly increasing intra-community coupling, as shown in Fig. 6.7f. Although the decrease of all three spontaneous signal diversity measures with decreasing feed-back long range coupling is in line with some studies [237, 186], a simultaneous increase of intra-community coupling is not mentioned in the literature, hence this is an intriguing possibility raised by the model.

The example parameter space further shows a path (dotted line) along which PC is constant but LZc, ACE, mSS increase, showing that these measures can vary independently to phase coherence and thus capture different signal features. SCE by contrast is fairly constant along this path of constant PC, in line with the observation from Fig. 6.4 that SCE tends to behave inversely like PC, Fig. 6.7g. This is as expected, given that SCE is designed to capture diversity of phases across channels while PC captures the opposite, coherence of phases.

ACE, LZc and SCE drop together for both models of loss of consciousness. In contrast to empirical observation, PCI's behaviour stayed equal for the first and even increased for the second model of losing consciousness. The thalamic drive model is attractive as it is possibly the simplest realisation of non-homogeneous change of coupling that causes ACE, SCE and LZc to decrease while PC increases, further only a 40% increase in coupling strength resulted in the desired signal diversity changes and mSS was increasing with

[3]The construction of this matrix was performed as follows. The adjacency matrix contains an entry for every ordered pair {i,j} of oscillator indices for $i, j = 0, 1, ..., 63$. When dividing the matrix into a regular grid of 64 tiles, each having size 8 entries times 8 entries, the 8 diagonal tiles represent the fully connected communities. The distance between the k^{th} and the l^{th} tile is measured as $\Delta_{k,l} = |k - l|$ with $k, l = 0, 1, ..., 7$, only considering horizontal or vertical paths. Now starting from a tile on the main diagonal (representing a community), say with index k, tiles with distance Δ_{kl} being $0, 1, ..., 7$ will be filled with $64, 32, 16, 8, 4, 2, 1, 0$ entries, respectively. The filling of a tile works such that its 64 entries are sequentially indexed column by column and then only every $(64/x)^{th}$ entry is filled, for $x = 64, 32, 16, 8, 4, 2, 1, 0$, respectively, as a function of distance.

increasing spontaneous signal diversity, in line with empirical observation (Section 2.2). It further is clearly in line with the thalamic drive hypothesis. The community-structured model has the shortcoming that strong increase in short-range coupling contrasts with empirical observations and possibly simply displays the same effect as for the "thalamic" setting where a strong increase in some of the connections causes all three measures to decrease. We further explored a hybrid-model, imposing the community-structure of the modular network onto the region in the thalamic adjacency that represents the cortex, without gaining further insights.

6.5.2 Spatial versus temporal diversity: Comparing models with depth electrode recordings

Here we compare power spectra and time series between depth electrode recordings and oscillators, using sleep and wakeful rest as empirical conditions and the point pair p3, p4 from the previous section as two points for which ACE, SCE and LZc changed in the same direction. We further use this comparison to highlight the difference between temporal and spatial signal diversity by the additional computation of single channel Lempel-Ziv complexity (LZs), capturing temporal signal diversity only whereas ACE, SCE and LZc capture signal diversity across channels and observations.

The spontaneous signal diversity measures ACE, SCE and LZc applied to depth electrode recordings were found to be lower in NREM sleep than WR, see Fig. 6.8c for a recapitulation (full details in Chapter 4, NREMe indicates non-rapid eye movement sleep early at night). By definition, all three measures capture both, temporal and spatial signal diversity. Computing the Lempel-Ziv complexity of single channels and subsequently averaging scores across all channels results in a measure that only captures temporal signal diversity. As defined in Chapter 2, we call this measure LZsum. Applied to depth, the score of LZsum is substantially higher for WR than NREMe for all tested subjects (average across subjects shown in Fig. 6.8c), indicating that changes in temporal signal diversity alone differ substantially for depth electrode recordings across the conditions WR and NREM sleep.

Visual inspection of example depth electrode recordings (Fig. 6.8b) indicates less low-frequency components of the signal in WR than NREMe, both for the local-field-potential (LFP) as well as the absolute value of the analytic signal (amplitude). Comparing LFP with its amplitude, we see that signal sections with negative LFP are flipped smoothly to the positive side. The average spectral profile (Fig. 6.8a) further confirms that there are less slow-frequency components for WR than NREMe. PC increased for NREM sleep yet is close to the low score of 0.2 for both conditions (Fig. 6.8c), indicating that signal diversity across channels is high.

By contrast, consider the point pair p3 and p4 as described in Fig. 6.7e for which ACE, SCE and LZc decreased together with increasing mean coupling (short range coupling strongly increased while long range coupling weakly decreased). Importantly, LZsum does not differ across conditions p3 and p4 (Fig. 6.8f), indicating that temporal signal diversity does not differ on average across channels. This is further evident in the nearly identical

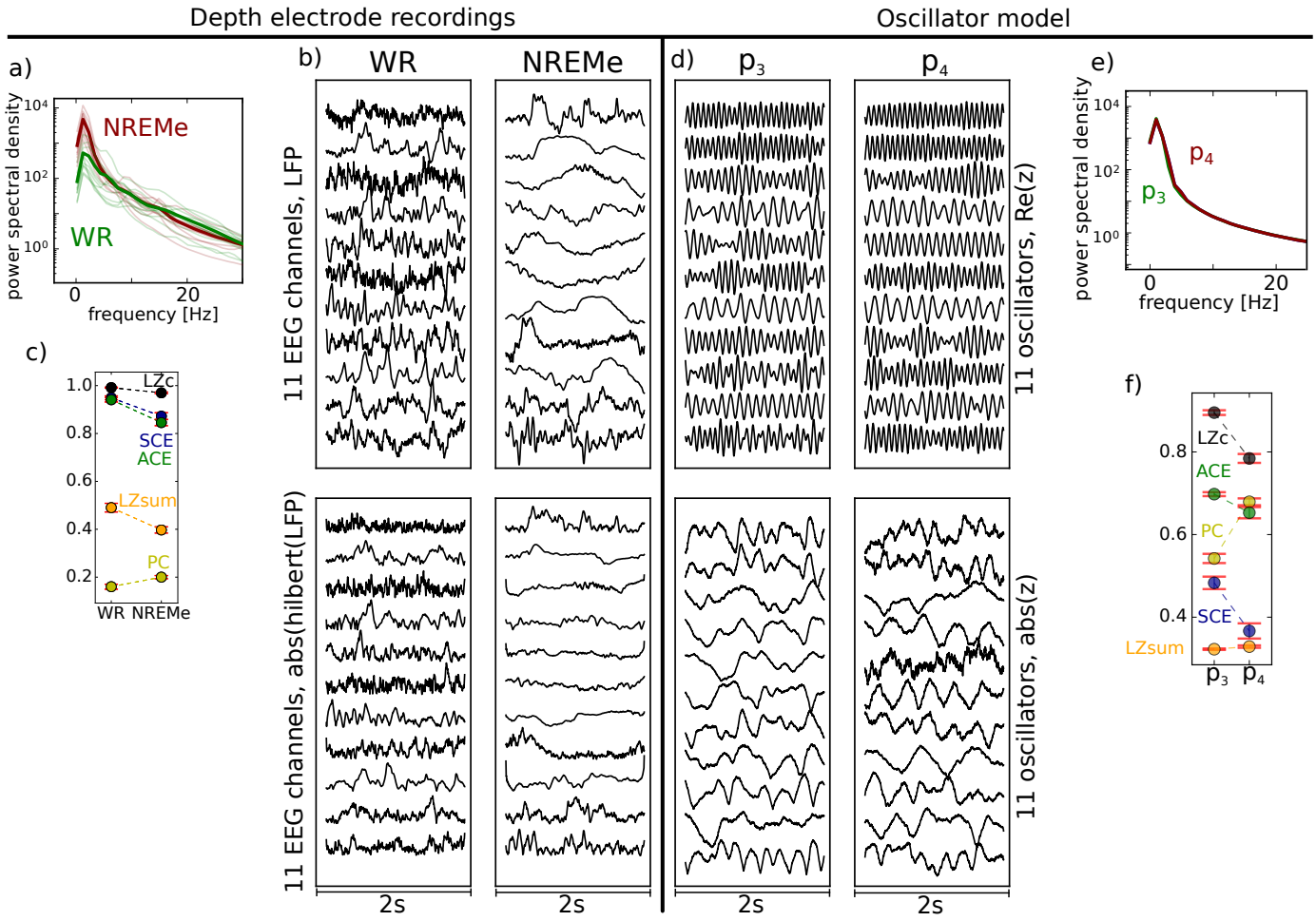

Figure 6.8: **Depth electrode recordings compared to oscillator signals.** a) The spectral power of depth electrode recordings for wakeful rest (WR) and non-rapid eye movement sleep (NREMe) is shown, averaged across channels and trials per subject (thin lines) and across subjects (thick lines). The signal for NREM displays more low-frequency components, also visible in b) when comparing example time-series (normalised by standard deviation). c) Shows a decrease in ACE, LZc and ACE in NREM, as well as a decrease in LZsum which measures single channel (temporal) signal diversity only. Signal diversity across channels is higher for WR than NREM as indicated by a lower PC value. e) Spectral power of oscillator activity, averaged across oscillators and trials, is nearly identical for conditions p3 and p4 of the modular adjacency setting described in Fig. 6.7e, also visible in d), showing a similar distribution of temporal diversity in the activity time series of oscillators for both conditions. f) The scores of LZc, ACE and SCE decrease with increasing coupling and PC, the latter being considerably higher than for WR and NREM. Importantly, LZsum does not differ across conditions p3 and p4, indicating that only spatial but not temporal signal diversity changes.

power spectra for conditions p3 and p4 (Fig. 6.8d) and visual inspection of example time series of oscillator activity (Fig. 6.8e). This constitutes thus a clear demonstration that LZc does capture signal diversity across channels and observations.

Unlike LFP and the absolute value of the analytic signal of LFP, the time series of the real part of oscillators' activity differs strongly to time series of their amplitudes. Visual inspection of the oscillator time series further suggests that temporal changes in oscillator

amplitude are more similar to LFP than temporal changes in the real part of oscillators, the latter being highly sinusoidal. Oscillator amplitude fluctuations look reasonably similar to depth electrode recordings, while appearing smoother, which may be adjusted by increasing the noise level.

In summary, our example point pair showed changes in signal diversity across oscillators but not time whereas empirically, signal diversity decreases across channels and time when consciousness is lost. A more extensive exploration of heterogeneous coupling changes may induce temporal signal diversity in the amplitudes oscillators in addition. This oscillator network example demonstrates that changes in ACE, LZc and SCE may result from changes in spatial diversity only, and clearly shows that LZc captures signal diversity across channels in addition to signal diversity across observations. Amplitude time series drawn from the model more closely resembled amplitude time series from the empirical data, than the highly sinusoidal real part of the oscillators.

6.6 Discussion

Our aim was to manipulate the dynamics of coupled Stuart-Landau oscillators such that the system's signal diversity changes as observed for electrophysiological brain signals across different levels of consciousness. I.e. a decrease of all spontaneous signal diversity measures as well as the perturbational complexity index as observed for distributed electrophysiological signals from the human cortex when the level of consciousness decreases.

Empirically, response activity to TMS "travels" further across the cortex in WR than NREM sleep, indicating reduced effective connectivity between cortical regions in NREM (Section 2.2). This suggested as a first hypothesis that a decrease in coupling in the oscillator model should reduce signal diversity. By contrast, phase coherence for spontaneous depth electrode signals was lower in WR than NREM and, in the model, found to increase monotonically with increasing mean coupling of the oscillators. Irrespective of the hypothesis, neither decrease nor increase of mean coupling, equally applied to all connections between oscillators, did result in a change of all signal diversity measures in the same direction, in particular as SCE behaved differently to LZc and ACE. This was also the case when exploring the measures' behaviour for topological changes of the oscillator network at fixed levels of mean coupling.

We thus explored more complicated changes in the oscillator network. When we increased the coupling strength for some connections only, we did find a drop of all three spontaneous signal diversity measures (ACE, SCE, LZc). This was achieved in particular by the "thalamic" drive adjacency where coupling strength was increased only for a section of connections - being arguably the simplest heterogeneous coupling change. This finding is in line with the hypothesis that consciousness is lost when thalamic activity synchronises activity throughout the cortex via strong coupling [154, 103] and therefore prevents diverse cortico-cortical interaction.

We further showed that for our "modular" adjacency, a weak decrease in certain inter-community connections and a simultaneous strong increase in intra-community coupling also lead to the decrease of ACE, SCE and LZc. While the three spontaneous signal di-

versity measures decreased together for this "modular" adjacency change, signal diversity across time remained the same, as measured by single channel Lempel-Ziv complexity (LZ-sum). This is a further restriction of our model, since temporal signal diversity (quantified by LZsum), changes in depth electrode recordings across conscious states.

Another limitation of our model is that we did not find a setting where PCI changed in the same direction as all three spontaneous signal diversity measures. By contrast, PCI weakly increased with increasing mean coupling for our different oscillator models. Empirically observed response wave-forms are clearly more complicated than response wave-forms of our model and PCI scores substantially higher: PCI scores for the brain range from 0.15 in unconscious states up to 0.7 in wakeful rest [53] while for our oscillator model the highest observed PCI score was 0.13. More complex response activity to perturbation may be obtained by considering time-delayed interactions of Stuart-Landau oscillators, as applied by Moon [157] and/or increasing the number of oscillators and grouping their activity.

Despite the model's shortcomings, an important observation is that our spontaneous signal diversity measures did only decrease together for heterogeneous and not for homogeneous coupling changes, when noise is present. This suggests that the coupling changes between brain regions involved in the loss of consciousness require unbalanced interaction of dynamically independent brain regions, such as the thalamus and the whole cortex, in line with a proposed relay model of unbalanced cortico-thalamic interaction suggested to explain the transition from wake into deep sleep [154]. The role of the thalamic influence in the mechanism of losing consciousness is not clear, however, as e.g. the delta rhythm survives complete thalamectomy, suggesting a cortical origin of strong dynamical changes [218].

The direct comparison of depth electrode signal changes across wakeful rest and deep sleep and those observed in oscillator dynamics for our "modular" and "thalamic" model showed further for both, depth electrode recordings and oscillator models, an increase in phase coherence. This suggests that signal diversity across depth electrode channels is lower in deep sleep as opposed to wake, in line with our oscillator models where phase coherence also increases while the three spontaneous signal diversity measures decrease. We further saw that oscillator amplitude fluctuations ($|z|$) have good visual resemblance to depth electrode recordings, in sharp contrast to their real parts, which are much more regular oscillations, yet have previously been used to model EEG [179].

It is a drawback of ACE, SCE and LZc not to peak for intermediate values of phase coherence in the presence of noise, given that the dynamical regime in between full independence and full synchrony of the oscillators is most similar to healthy brain dynamics. As expected, SCE (measuring diversity of phases) increases fairly monotonically with decreasing PC (measuring coherence of phases). ACE and LZc mainly detect how much influence the added noise has to the signal diversity (high for fully independent and fully synchronised oscillators, low in between as fluctuations due to other oscillators dominate) instead of capturing the richer repertoire of coalitions between oscillators for intermediate phase coherence. This was the case in the initial formulation of coalition entropy as introduced by Shanahan [209], which peaked for intermediate phase coherence in a Kuramoto

model, however without noise and defined for synchronies of oscillator communities unlike for single oscillator amplitudes as in our case.

In conclusion, in this first exploratory study on signal diversity measures sensitive to conscious state applied to time-series of coupled Stuart-Landau oscillators, we showed that heterogeneous coupling changes between oscillators are required for the scores of all three spontaneous signal diversity measures to change in the same direction. This was realised by increasing the coupling of only a fraction of all connections, supporting the hypothesis that strong coupling between the cortex and the thalamus prevents effective connectivity between cortical regions and thus stops the sustenance of consciousness. A shortcoming of our model is that PCI scores do not change together with scores of ACE, SCE and LZc and the response activity to perturbation is very simple in comparison to that of EEG/TMS for any global state of consciousness. More complicated responses may be obtained by exploring time-delayed coupling between oscillators, which may further be used to induce changes in temporal signal diversity (LZsum) in addition to changes in signal diversity across channels (ACE, SCE, LZc), as observed empirically but not found here. It should be further investigated how the distribution of natural frequencies determines the average spectral profile of the oscillator amplitude fluctuations, aiming to model spectral changes in addition to spontaneous and perturbational signal diversity changes in line with empirical observations (e.g. stronger slow wave component in the depth electrode recordings during NREM than WR). The presented results already illustrate the model's potential to clarify differences of signal diversity that correlate with consciousness and mechanisms that generate such signal features.

Chapter 7

Conclusions

This thesis argues that brain signal diversity is a hallmark of consciousness, reflecting phenomenological differentiation. Measures that capture spontaneous cortical signal diversity across regions and time were introduced and shown to robustly index several global states of consciousness. In comparison to wakeful rest, signal diversity was shown here to be lower for propofol-induced anaesthesia (Chapter 3), comparable for REM sleep, lower for NREM sleep (Chapter 4) and higher for psychedelic states (Chapter 5). The relation of signal diversity and other notions of signal complexity was clarified using these empirical results in conjunction with computer simulations and, most importantly, connected to complexity theories of consciousness.

The basic behaviour of our novel measures of multidimensional spontaneous signal diversity, ACE, SCE and LZc was illustrated with simple simulations of time series, such as increasingly ordering initially random strings or manipulating coupling and noise correlation in an auto-regressive model. Both models showed that all three measures clearly capture signal diversity (Section 2.6) in the sense of algorithmic complexity, i.e. randomness and irregularity[1]. There are subtle differences, though. While LZc and ACE monotonically increased with increasing signal diversity, SCE showed a peak for intermediate diversity in the auto-regressive model, induced by either coupling strength or noise correlation, or both. This shows that SCE captures not only signal diversity as LZc and ACE, but also dynamical complexity in the sense of peaking somewhere in between complete order and complete randomness of the signal for this system. SCE was further shown to behave completely different than LZc and ACE for a particular setting of the Kuramoto model and several settings in our Stuart-Landau model, where SCE was nearly unaffected by noise unlike ACE and LZc. Thus SCE's behaviour only matches that of ACE and LZc for certain dynamical regimes.

Also for the empirical analyses we found differences between the behaviour of SCE and LZc/ACE. SCE's scores changed less consistently across sleep states than those of LZc and

[1]There is rich mathematical literature that aims to formalise notions of randomness, with the concept of Kolmogorov complexity frequently occurring. The latter is defined for a string of symbols to be the number of lines of computer code needed to generate the string. The more lines of code are needed, the higher the Kolmogorov (or algorithmic) complexity of the string. This idea of quantifying randomness (or diversity), although proven to be impossible to compute exactly, has inspired many alternative definitions of randomness, aiming to define notions of disorder precisely [84].

ACE for the local depth electrode analysis conducted on only 4 channels and the analysis of MEG during psychedelic states, using 10 channels. However, when comparing wake with propofol anaesthesia and wake with different sleep stages, SCE's behaviour was very similar to that of ACE and LZc when analysing more then 10 channels broadly distributed across the cortex. This is a strong indication that all three "flavours" of signal diversity index these states.

This fact further guided us in the search for Stuart-Landau oscillator system dynamics that are such that all three measures change in the same direction, hence the subtle differences of the measures allowed a more precise characterisation of the average signal complexity changes between these global states of consciousness. It allowed us to show that heterogeneous coupling changes between oscillators are required for the scores of all three spontaneous signal diversity measures to change in the same direction. This was realised by increasing the coupling only of a fraction of all connections between oscillators, supporting the hypothesis that strong coupling between the cortex and the thalamus prevents effective connectivity between cortical regions and thus makes it impossible for consciousness to occur. A shortcoming of our model is that PCI scores do not change together with scores of ACE, SCE and LZc and the response activity to perturbation is very simple in comparison to that of EEG/TMS for any global state of consciousness. Yet the results presented in Chapter 6 illustrate the promising potential of the Stuart-Landau model to clarify signal features that correlate with consciousness and mechanisms that generate these. Future work should aim to seek parameter changes in the model that induce changes in additional neural signatures of consciousness faithful to empirical observation, such as PCI, delta power and avalanche size distribution, to propose mechanisms that affect these different measures ideally all at once as observed empirically.

Besides the subtle differences in the three multidimensional signal signal diversity measures (SCE, LZc, ACE) for empirical data, importantly, also single channel Lempel-Ziv complexity (LZs) - capturing temporal diversity of spontaneous EEG/MEG signals only[2] - indicated conscious states for all experiments presented here with equal or even higher consistency than the multidimensional signal diversity measures. LZs may thus be the simplest spontaneous signal diversity measure of consciousness for the global states we investigated here, nearly independent of channel location in the cortex (as can be seen in Fig. B.4) and fairly unaffected across a wide range of sampling rates or data segment lengths. It is an important contribution of this thesis to introduce different measures of spontaneous signal diversity across channels and observations and compare their behaviour with measures of temporal signal diversity only, providing measures of different types of signal diversity that all correlate with conscious level.

Clearly, the fact that LZs is maximal for a random signal - and nearly maximal for any of the empirical signals analysed here - cannot imply that neural processes are highly random themselves. At the spatial scale of EEG or MEG signals - both being superpositions of the

[2]ACE, SCE and LZc are capturing signal diversity across channels and observations: SCE and ACE would score very low and LZc intermediate for, say, 10 maximally random but identical signals whereas LZsum would have maximal score. Given the large values of SCE, ACE and LZc for all here-tested global states of consciousness, temporal signal diversity (across observations for a single channel) comes together with spatial signal diversity (across channels), else SCE and ACE would score substantially lower.

electromagnetic field of countless neural circuits[3] - much of the information about the fine structure of underlying dynamics is averaged away, yet the large scale signal diversity can only be induced by a richer repertoire of states at a smaller scale .

Edelman and Tononi suggested that a large repertoire of different neural activity patterns directly reflects the large repertoire of different experiences, i.e. mapping neuronal to phenomenological diversity, called "differentiation" [229]. The rapid succession of different experiences in wakeful rest must accordingly be reflected by higher signal diversity than during the absence of experiences, e.g. in NREM sleep. Irrespective of the link to phenomenology, the observed changes in signal diversity trivially imply that neural processes during wakeful rest are on average more diverse than processes during NREM sleep[4].

We found lower spontaneous signal diversity for NREM/propofol anaesthesia than WR at all tested cortical regions, as supported in particular by the LZs analyses of all depth electrode channels for WR and NREMe (Chapter 4), showing higher signal diversity for WR than NREM for 48 out of 50 channels. This implies that processes in all these distributed cortical areas are affected by the state of consciousness, clearly in line with the broader empirical consensus that activity of many distinct regions, particularly across the thalamocortical system, is essential for generating the rich contents of any experience [131]. It is obvious that for example an audio-visual experience should correlate with activity in the auditory and visual cortex, so many different brain regions must be active to account for many different properties of the given experience.

This "differentiation" is also key for the entropic brain hypothesis of Carhart-Harris et al. [49], which states that there is correlation between the degree of overall randomness (entropy, differentiation) in brain dynamics and "vividness of cognition", locating brain states such as coma, anaesthesia and deep sleep at relatively low entropy, wakeful rest at intermediate entropy and REM sleep and psychedelic states at relatively high entropy [49]. In support of this, other signal diversity measures applicable to a single time series, such as approximate entropy [176] and permutation entropy [16], have been found to decrease during anaesthesia and sleep [42, 43, 46, 124]. Our results are consistent with these findings and, importantly, we provided first evidence that psychedelic states lie above wakeful rest and REM on a one-dimensional scale defined by spontaneous neural signal diversity (Chapter 5).

Tononi and Edelman [229] suggested further that each conscious experience is perceived as an integrated whole and this should be reflected in neural activity by brain regions being effectively connected to bring together all aspects of the given experience, for example sound and colour. This property of "wholeness" of each experience on the phenomeno-

[3]Neuronal circuits are regarded as the main unit for information processing in the human brain and their activity gives rise to electrical currents. Such currents from all active cellular processes within a volume of brain tissue superimpose at a given location in the extra-cellular medium and generate an electric potential that can be compared to another one found at another location. The difference in potential constitutes an electric field that can be measured with electrodes, e.g. directly with scalp EEG or intracranial depth electrodes with very high temporal resolution [47]. Currents further induce magnetic fields that can be detected with MEG, also at high temporal resolution and allowing even better spatial resolution than scalp EEG.

[4]Also highly complex brain processes occur during slow wave sleep - such as memory consolidation - yet unconsciously.

logical side and the effective connectivity on the neuronal side was called "integration". Effective connectivity between brain regions is realised via the ongoing, recursive, highly parallel signaling within and among brain areas, also known as reentry [229].

To what extent can we say that ACE, SCE and LZc reflect *complexity,* in the sense of co-existing differentiation and integration amongst the EEG signals [229]? Regarding integration, variations over time to the set of active or synchronous channels could arise in the absence of causal interactions; for instance, each channel could be exhibiting its own individual chaotic dynamics, and be evolving in isolation. Thus the measures ACE, SCE and LZc would clearly not in general for any system capture integration. However, structural constraints mitigate against fully random activity in the context of the thalamocortical system. Given that most cortical regions tend in general to receive strong driving input from sub-cortical regions [110], it would seem that diverse activity across EEG channels is likely to require a certain amount of functional integration [53, 153, 214]. Thus ACE, SCE and LZc, when applied to multi-dimensional EEG recordings of brain activity, may cautiously be considered to also correlate with integration in EEG dynamics, and therefore to track complexity in the sense of simultaneous differentiation and integration. Only SCE has been shown to peak at intermediate levels of signal diversity for an autoregressive model, showing that for certain systems it captures complexity in the sense of co-existing differentiation and integration directly, however this was not the case for the empirical data where SCE behaved predominantly like ACE and LZc. Thus, for the dynamical regimes of the empirical data analysed here, our measures of spontaneous signal diversity do not directly quantify "integration" in the sense of effective connectivity.

By contrast it was argued that PCI indicates effective connectivity directly, by the fraction of source channels with significant response signals to the magnetic perturbation. E.g. substantially more source channels show significant response activity during WR in comparison to NREM sleep is interpreted to indicate higher effective connectivity between brain regions in WR than in NREM sleep[5].

The behaviour of spontaneous and perturbational signal diversity measures for different global states of consciousness both clearly support the proposed map from phenomenological differentiation to neural signal differentiation. The map from phenomenological wholeness, i.e. integration, to effective connectivity between brain regions as indicated by how far the response signal to magnetic perturbation spreads can also convincingly be argued for. However, in addition to "integration" and "differentiation", Tononi suggested further phenomenological features of any experience, e.g. the exclusion principle states that each experience happens at its own spatio-temporal grain, requiring that only one experience can be generated by the brain at any given time [235], which seem impossible to be connected to the coarse average brain signal features as measured by spontaneous or perturbational signal diversity measures for global states of consciousness.

However, the arguments for neural signal diversity to reflect phenomenological diversity are convincing and should be investigated for finer levels of consciousness and ideally more directly connected to conscious content. Our finding of increased spontaneous sig-

[5]This difference is only visible as long as the strength of the magnetic perturbation is not too high as to activate all sources in both states of consciousness (discussed in Section 2.2).

nal diversity for psychedelic states of consciousness is an important step in this direction (Chapter 5). We provided a concrete suggestion for a multidimensional characterisation of global states of consciousness, as asked for by Bayne et al. [26], for the psychedelic state with spontaneous MEG signal diversity as one dimension and subjectively rated difference of overall experience as another.

Perturbational signal diversity has not yet been analysed for the psychedelic state, yet it would be interesting to see if PCI is consistently higher here, in particular since PCI's behaviour was in line with the behaviour of spontaneous signal diversity measures for the comparison of WR to NREM and propofol anaesthesia. It would be surprising and against theoretical considerations to find two global states of consciousness for which spontaneous and perturbational signal diversity anticorrelated consistently.

Perturbational and spontaneous signal diversity are complimentary aspects to characterise the diversity of neural activity, both with practical strengths and weaknesses. Perturbational signal diversity is deduced from the average response signal across many repeated perturbations, resulting in a response signal that is largely free of noise. This is most likely the main reason for PCI's higher specificity and sensitivity to conscious states in comparison to spontaneous signal diversity measures, as can be seen for example when comparing our results and those presented in [53] for intermediate states of propofol sedation or, as another example, comparing single channel Lempel-Ziv complexity with PCI for subtle differences of conscious level in patients with disorders of consciousness (see [214], where single channel Lempel-Ziv complexity is called K-complexity).

Despite the weaker sensitivity and specificity to conscious states, our measures of spontaneous signal diversity do not require a large number of trials as is needed for PCI, and are thus an attractive additional analysis choice for studies involving different global states of consciousness, e.g for sleep stages, as has been already demonstrated by Andrillon el al that used our LZc definition [8] to index light and deep sleep. Spontaneous signal diversity can further be estimated from short 2sec segments of activity, thus allowing continuous tracking of signal diversity.

We expect ACE, SCE and LZc to act synergistically with other measures besides PCI, for example a coarse analysis of the power spectrum. This would be in line with recent work by Sitt et al. [214] analysing auditory evoked potentials in EEG from traumatic brain injury patients. They combined spectral measures with measures of the signal diversity of a single time-series to achieve greater discriminative power (as measured by area under the ROC curve in classifying the states of traumatic brain injury patients).

Other theory-based measures that could be explored in conjunction with the signal diversity measures described here include 'causal density' [204] and various versions of integrated information (as measured by the quantity 'phi' [166]), integration measures based on long-range functional connectivity such as weighted symbolic mutual information [127] and global graph-theoretic measures such as 'efficiency' of connectivity graphs [62]. Moreover, the computation of the LZc and ACE measures based on a simplification of the continuous EEG to a binary signal necessarily discards some information content from the signals analysed. It is possible that utilising a transformation of the continuous signal into

a repertoire of 'symbols', as in [127], could increase discriminative power.

Collectively this thesis argues that brain signal diversity, spontaneous and perturbational, is a robust explanatory correlate of consciousness, constituting a general and well defined feature of neural activity that can be clearly interpreted to reflect the phenomenological diversity across different conscious experiences. Our findings encourage to continue the theoretical efforts to connect general phenomenological properties to general neural activity patterns, that may one day clarify what complex dynamics are required such that it feels like something to be a system.

Bibliography

[1] Aaronson, S. (2014). Giulio Tononi and Me: A Phi-nal Exchange. Accessed August 1 2016. http://www.scottaaronson.com/blog/?p=1823

[2] Aaronson, S. (2014). Why I Am Not An Integrated Information Theorist (or, the Unconscious Expander). Accessed August 1 2016. http://www.scottaaronson.com/blog/?p=1799

[3] Abasolo, Daniel, et al. "Lempel-Ziv complexity of cortical activity during sleep and waking in rats." Journal of neurophysiology 113.7 (2015): 2742-2752.

[4] Aboy, Mateo, et al. "Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis." IEEE Transactions on Biomedical Engineering 53.11 (2006): 2282-2288.

[5] Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, Gramfort A, Thirion B and Varoquaux G (2014) Machine learning for neuroimaging with scikit-learn. Front. Neuroinform. 8:14. doi: 10.3389/fninf.2014.00014

[6] Amigo, Jose M., et al. "Estimating the entropy rate of spike trains via Lempel-Ziv complexity." Neural Computation 16.4 (2004): 717-736.

[7] Andrillon T, Nir Y, Staba RJ, Ferrarelli F, Cirelli C, Tononi G, Fried I (2011) Sleep spindles in humans: insights from intracranial EEG and unit recordings. J Neuroscience

[8] Andrillon, Thomas, et al. "Neural Markers of Responsiveness to the Environment in Human Sleep." The Journal of Neuroscience 36.24 (2016): 6583-6596.

[9] Arnulfo, Gabriele, et al. "Phase and amplitude correlations in resting-state activity in human stereotactical EEG recordings." Neuroimage 112 (2015): 114-127.

[10] Arvanov, Viktor L., et al. "LSD and DOB: interaction with 5-HT2A receptors to inhibit NMDA receptor-mediated transmission in the rat prefrontal cortex." European Journal of Neuroscience 11.9 (1999): 3064-3072.

[11] Baars B (1988) A cognitive theory of consciousness. Cambridge University Press, New York.

[12] Baars, Bernard J. "In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness." Journal of Consciousness Studies 4.4 (1997): 292-309.

[13] Baars, Bernard J. "The global brainweb: An update on global workspace theory." Science and Consciousness Review 2 (2003).

[14] Yang Bai, Zhenhu Liang, Xiaoli Li (2015) A permutation Lempel-Ziv complexity measure for EEG analysis, Biomedical Signal Processing and Control, Volume 19

[15] Balduzzi D, Tononi G (2008) Integrated information in discrete dynamical systems: Motivation and theoretical framework. PLoS Comput Biol 4(6): e1000091.

[16] Bandt C, Pompe B, (2002) Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett. 88, 174102-1.

[17] Barnett, Lionel, Adam B. Barrett, and Anil K. Seth. "Granger causality and transfer entropy are equivalent for Gaussian variables." Physical review letters 103.23 (2009): 238701.

[18] Barnett, Lionel, Christopher L. Buckley, and Seth Bullock. "Neural complexity and structural connectivity." Physical Review E 79.5 (2009): 051914.

[19] Barnett, Lionel, and Terry Bossomaier. "Transfer entropy as a log-likelihood ratio." Physical review letters 109.13 (2012): 138105.

[20] Barrett A, Barnett L, Seth A (2010) Multivariate Granger causality and generalized variance. Phys Rev E 81: 041907.

[21] Barrett, Adam B., and Anil K. Seth. "Practical measures of integrated information for time-series data." PLoS Comput Biol 7.1 (2011): e1001052.

[22] Barrett A, Murphy M, Bruno M, et al. (2012) Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. PLoS ONE, 7(1): e29072.

[23] Barrett, Adam B. "An integration of integrated information theory with fundamental physics." Frontiers in psychology 5 (2014).

[24] Barrett, Adam B. "A comment on Tononi and Koch (2015)'Consciousness: here, there and everywhere?'." Phil. Trans. R. Soc. B 371.1687 (2016): 20140198.

[25] Pablo Barttfeld, Lynn Uhrig, Jacobo D. Sitt, Mariano Sigman, Bechir Jarraya, and Stanislas DehaeneSignature of consciousness in the dynamics of resting-state brain activityPNAS 2015 112 (3) 887-892; published ahead of print January 5, 2015, doi:10.1073/pnas.1418031112

[26] Bayne, T., Hohwy, J., Owen, A. M. (2016). Are There Levels of Consciousness?. Trends in cognitive sciences, 20(6), 405-413.

[27] Beggs, John M., and Nicholas Timme. "Being critical of criticality in the brain." Frontiers in physiology 3 (2012): 163.

[28] Benjamin Jr, Ludy T. A brief history of modern psychology. Blackwell Publishing, 2007.

[29] Benjamini, Yoav; Hochberg, Yosef (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing" (PDF). Journal of the Royal Statistical Society, Series B. 57 (1): 289âĂŞ300. MR 1325392.

[30] Blake, Randolph, and Nikos K. Logothetis. "Visual competition." Nature Reviews Neuroscience 3.1 (2002): 13-21.

[31] Block, Ned. "How many concepts of consciousness?." behavioural and brain sciences 18.02 (1995): 272-287.

[32] Block, Ned. "How can we find the neural correlate of consciousness?." Trends in neurosciences 19.11 (1996): 456-459.

[33] Block, Ned, et al. "Consciousness science: real progress and lingering misconceptions." Trends in cognitive sciences. 18.11 (2014): 556-557.

[34] Blume, Christine, et al. "Across the consciousness continuum-from unresponsive wakefulness to sleep." Frontiers in human neuroscience 9 (2015).

[35] Boly M (2011) Measuring the fading consciousness in the human brain. Curr. Opin. Neurol. 24, 394-400.

[36] Boly, Melanie, et al. "Consciousness in humans and non-human animals: recent advances and future directions." Frontiers in psychology 4 (2013): 625.

[37] Boly, Melanie, et al. "Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study." PloS one 10.5 (2015): e0125337.

[38] Bor, Daniel. The ravenous brain: How the new science of consciousness explains our insatiable search for meaning. Basic Books, 2012.

[39] Botcharova, Maria, Simon F. Farmer, and Luc Berthouze. "Markers of criticality in phase synchronization." Frontiers in systems neuroscience 8 (2015).

[40] Breakspear M, Heitmann S, Daffertshofer A. Generative models of cortical oscillations: neurobiological implications of the kuramoto model. Front Hum Neurosci. 2010;4(November):190.

[41] Breitmeyer, Bruno G., and Haluk Ogmen. "Recent models and findings in visual backward masking: A comparison, review, and update." Perception and psychophysics 62.8 (2000): 1572-1595.

[42] Bruhn J, Ropcke H, Hoeft A (2000) Approximate entropy as an electroencephalographic measure of anesthetic drug effect during desfluraneanesthesia. Anesthesiology 92: 715-26.

[43] Bruhn J, Ropcke H, Rehberg B, Bouillon T, Hoeft A (2000) Electroencephalogram approximate entropy correctly classifies the occurrence of burst suppression pattern as increasing anesthetic drug effect. Anesthesiology 93:981-5.

[44] Bullmore, Edward T., and Danielle S. Bassett. "Brain graphs: graphical models of the human brain connectome." Annual review of clinical psychology 7 (2011): 113-140.

[45] Bullock, T. H., et al. "Temporal fluctuations in coherence of brain waves." Proceedings of the National Academy of Sciences 92.25 (1995): 11568-11572.

[46] Burioka N, Miyata M, Cornelissen G, et al. (2005) Approximate entropy in the electroencephalogram during wake and sleep. Clin EEG Neurosci 36:21-4.

[47] Buzsaki, Gyorgy, Costas A. Anastassiou, and Christof Koch. "The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes." Nature reviews neuroscience 13.6 (2012): 407-420.

[48] Buzsaki G, Logothetis N, Singer W. (2013) Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. Neuron

[49] Carhart-Harris R, Leech R, Hellyer P, et al. (2014) The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. Front. Hum. Neurosci. 8:20 10.3389/fnhum.2014.00020.

[50] Carhart-Harris, Robin L., et al. "The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs." Psychoanalytical neuroscience: Exploring psychoanalytic concepts with neuroscientific methods (2015): 140.

[51] Carhart-Harris, R. L., et al. "The paradoxical psychological effects of lysergic acid diethylamide (LSD)." Psychological medicine 46.07 (2016): 1379-1390.

[52] Carhart-Harris, Robin L., et al. "Neural correlates of the LSD experience revealed by multimodal neuroimaging." Proceedings of the National Academy of Sciences 113.17 (2016): 4853-4858.

[53] Casali A, Gosseries O, Rosanova M, et al. (2013) A Theoretically Based Index of Consciousness Independent of Sensory Processing and behaviour. Sci. Transl. Med. 5, 198ra105.

[54] Cash, S.S., Halgren, E., Dehghani, et al (2009) The human K-complex represents an isolated cortical down-state. Science 324, 1084-1087.

[55] Cavanna, Andrea Eugenio, and Andrea Nani. Consciousness: theories in neuroscience and philosophy of mind. Springer, 2014.

[56] Matias Cavelli, Santiago Castro, Natalia Schwarzkopf, Michael H. Chase, Atilio Falconi, Pablo Torterolo, Coherent neocortical gamma oscillations decrease during REM sleep in the rat, Behavioural Brain Research, Volume 281, 15 March 2015, Pages 318-325, ISSN 0166-4328

[57] Cerullo, Michael A. "The Problem with Phi: A Critique of Integrated Information Theory." PLoS Comput Biol 11.9 (2015): e1004286.

[58] Chalmers, D.J. Facing up to the problem of consciousness. Journal of consciousness studies 2, 200-219 (1995)

[59] Chang J, Pigorini A, Massimini M, et al. (2012) Multivariate autoregressive models with exogenous inputs for intracerebral responses to direct electrical stimulation of the human brain. Front. Hum. Neuroscience

[60] Chattopadhyaya, Debiprasad. "Lokayata a Study in Ancient Indian Materialism." (1959).

[61] Chen, Xiangdong, Shaofang Shu, and Douglas A. Bayliss. "HCN1 channel subunits are a molecular substrate for hypnotic actions of ketamine." The Journal of Neuroscience 29.3 (2009): 600-609.

[62] Chennu S, Finoia P, Kamau E, et al. (2014) Spectral Signatures of Reorganised Brain Networks in Disorders of Consciousness. PLoS Comput Biol 10(10): e1003887. doi: 10.1371/journal.pcbi.1003887

[63] Chialvo DR. 2010 Emergent complex neural dynamics. Nat. Phys. 6, 744-750.

[64] Cohen J. (1992) A power primer.Psychological Bulletin, Vol 112(1)

[65] Cohen M (2014) Analyzing Neural Time Series Data. ISBN: 9780262019873.

[66] Conti, Fiorenzo, et al. "Cellular localization and laminar distribution of NMDAR1 mRNA in the rat cerebral cortex." Journal of Comparative Neurology 343.4 (1994): 554-565.

[67] Cossu M, Cardinale F, Colombo N, Mai R, Nobili L, Sartori I, Lo Russo G. (2005) Stereoelectroencephalography in the presurgical evaluation of children with drug-resistant focal epilepsy. J Neurosurg (Pediatrics 4)

[68] Coulon P, Budde T, Pape H-C (2012) The sleep relay: the role of the thalamus in central and decentral sleep regulation. Pflugers Arch 463(1):53-71.

[69] Crane, Tim, and Sarah Patterson, eds. History of the mind-body problem. Routledge, 2012.

[70] Crick F, Koch C (1990) Towards a neurobiological theory of consciousness. Seminars in Neuroscience Vol2.

[71] Crick, Francis. Astonishing Hypothesis: The Scientific Search for the Soul. Scribner reprint edition. ISBN 0-684-80158-2, 1995.

[72] Crick, Francis. "Visual-Perception-Rivalry and Consciousness." Nature 379.6565 (1996): 485-486.

[73] Crick, F. C. and Koch, C. What is the function of the claustrum? Phil. Trans. R. Soc. B 360, 1271-1279 (2005).

[74] Cumin, D., and C. P. Unsworth. "Generalising the Kuramoto model for the study of neuronal synchronisation in the brain." Physica D: Nonlinear Phenomena 226.2 (2007): 181-196.

[75] Thien Thanh Dang-Vu, Manuel Schabus, et al (2008)Spontaneous neural activity during human slow wave sleepPNAS 2008 105 (39) 15160-15165;

[76] Deboer T. (2013) behavioural and electrophysiological correlates of sleep and sleep homeostasis. Curr Top Behav Neurosci Oct 19.

[77] Dehaene, Stanislas, Michel Kerszberg, and Jean-Pierre Changeux. "A neuronal model of a global workspace in effortful cognitive tasks." Proceedings of the National Academy of Sciences 95.24 (1998): 14529-14534.

[78] Dehaene, Stanislas, Claire Sergent, and Jean-Pierre Changeux. "A neuronal network model linking subjective reports and objective physiological data during conscious perception." Proceedings of the National Academy of Sciences 100.14 (2003): 8520-8525.

[79] Dehaene, S., et al. Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy. Trends in Cognitive Sciences 100. (2006): 204-11.

[80] Dehaene, Stanislas. Consciousness and the brain: Deciphering how the brain codes our thoughts. Penguin, 2014.

[81] Del Cul, Antoine, Sylvain Baillet, and Stanislas Dehaene. "Brain dynamics underlying the nonlinear threshold for access to consciousness." PLoS Biol 5.10 (2007): e260.

[82] Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 134(1):9-21.

[83] Destexhe A, Contreras D, Steriade M. (1999) Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. J Neurosci 19(11):4595-608.

[84] Downey, Rodney G., and Denis R. Hirschfeldt. Algorithmic randomness and complexity. Springer Science and Business Media, 2010.

[85] Drewes, Jan, et al. "Recurrent Processing in the Formation of Shape Percepts." The Journal of Neuroscience 36.1 (2016): 185-192.

[86] Ebner, Christian, et al. "Open and closed cortico-subcortical loops: A neuro-computational account of access to consciousness in the distractor-induced blindness paradigm." Consciousness and cognition 35 (2015): 295-307.

[87] Edelman, Gerald. "Consciousness: the remembered present." Annals of the New York Academy of Sciences 929.1 (2001): 111-122.

[88] Edelman G (2003) Naturalizing consciousness: A theoretical framework. Proc Natl Acad Sci USA.

[89] Edlund J, Chaumont N, Hintze A, et al. (2011) Integrated Information Increases with Fitness in the Evolution of Animats. PLoS Comput Biol 7(10): e1002236. doi:10.1371/journal.pcbi.1002236.

[90] Electrical Geodesics, Inc.500 East 4th Ave. Suite 200, Eugene, OR 97401, USA, `www.egi.com`.

[91] Engel, Andreas K., and Wolf Singer. "Temporal binding and the neural correlates of sensory awareness." Trends in cognitive sciences 5.1 (2001): 16-25.

[92] Ercsey-Ravasz, Maria, et al. "A predictive network model of cerebral cortical connectivity based on a distance rule." Neuron 80.1 (2013): 184-197.

[93] Erdos P, Renyi A. 1959. On random graphs: I. Publ. Math. 6:290-97

[94] Erritzoe, David, et al. "Brain serotonin 2A receptor binding: relations to body mass index, tobacco and alcohol use." Neuroimage 46.1 (2009): 23-30.

[95] Modeling the Effects of Transcranial Magnetic Stimulation on Cortical CircuitsSteve K. Esser, Sean L. Hill, Giulio TononiJournal of Neurophysiology Jul 2005, 94 (1) 622-639; DOI: 10.1152/jn.01230.2004

[96] Farber, Nuri B., et al. "Serotonergic agents that activate 5HT2A receptors prevent NMDA antagonist neurotoxicity." Neuropsychopharmacology 18.1 (1998): 57-62.

[97] Fazekas, Peter, and Morten Overgaard. "Multidimensional models of degrees and levels of consciousness." Trends in Cognitive Sciences (2016).

[98] Feinberg, Todd E., and Jon M. Mallatt. The Ancient Origins of Consciousness: How the Brain Created Experience. MIT Press, 2016.

[99] Ferenets, R.; Lipping, T.; Anier, A.; Jantti, V.; Melto, S.; Hovilehto, S., "Comparison of entropy and complexity measures for the assessment of depth of sedation," in Biomedical Engineering, IEEE Transactions on , vol.53, no.6, pp.1067-1077, June 2006

[100] Ferenets R, Vonluchene A, Lippig T, et al. (2007) behaviour of Entropy/Complexity Measures of the Electroencephalogram during Propofol-induced Sedation: Dose-dependent Effects of Remifentanil. `http://lib.ugent.be/catalog/pug01:391118`

[101] Ferrarelli F, Massimini M, Sarasso S, et al. (2010) Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. PNAS 2010 107 (6) 2681-2686.

[102] Fingelkurts A, Fingelkurts A, Bagnato S, et al. (2013) Dissociation of vegetative and minimally conscious patients based on brainoperational architectonics: factor of etiology. Clin EEG Neurosci. 44:209-220.

[103] Franks NP. General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. Nat Rev Neurosci. 2008;9: 370-386. doi: 10.1038/nrn2372. pmid:18425091

[104] Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Cohen, L.,Naccache, L., 2009. Converging intracranial markers of conscious access. PLoS Biol.7, e61.

[105] Gallimore, A. R. (2015). Restructuring consciousness -the psychedelic state in light of integrated information theory. Frontiers in Human Neuroscience, 9, 346. http://doi.org/10.3389/fnhum.2015.00346

[106] Geweke, John. "Measurement of linear dependence and feedback between multiple time series." Journal of the American statistical association 77.378 (1982): 304-313.

[107] Geweke, John F. "Measures of conditional linear dependence and feedback between time series." Journal of the American Statistical Association 79.388 (1984): 907-915.APA

[108] Gosseries O, Schnakers C, Ledoux D, et al. (2011) Automated EEG entropy measurements in coma, vegetative state/unresponsive wakefulness syndrome and minimally conscious state. Functional Neurology 26(1): 25 - 30.

[109] Granger, C. W. J. (1969). "Investigating Causal Relations by Econometric Models and Cross-spectral Methods". Econometrica 37 (3): 424-438

[110] Guillery R, Sherman S (2002) Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. Neuron. 33(2):163-75.

[111] Hagberg, Aric A., Schult, Daniel A., and P. Swart. "Exploring network structure, dynamics, and function using NetworkX." Proceedings of the 7th Python in Science Conferences (SciPy 2008). Vol. 2008. 2008.

[112] Hagmann, Patric. From diffusion MRI to brain connectomics. Diss. Institut de traitement des signaux, Universite de Lausanne, 2005.

[113] Halberstadt, Adam L. "Recent advances in the neuropsychopharmacology of serotonergic hallucinogens." Behavioural brain research 277 (2015): 99-120.

[114] van den Heuvel MP, Sporns O. Rich-club organization of the human connectome. J Neurosci. 2011;31:11.

[115] Hillebrand, Arjan, et al. "Frequency-dependent functional connectivity within resting-state networks: an atlas-based MEG beamformer solution." Neuroimage 59.4 (2012): 3909-3921.

[116] Hohwy, Jakob, Andreas Roepstorff, and Karl Friston. "Predictive coding explains binocular rivalry: An epistemological review." Cognition 108.3 (2008): 687-701.

[117] Hohwy, J. (2009). The neural correlates of consciousness: new experimental approaches needed? Conscious. Cogn. 18, 428-438

[118] Hohwy, Jakob. The predictive mind. Oxford University Press, 2013.

[119] Hopley L, van Schalkwyk J, `http://www.anaesthetist.com/mnm/stats/roc/Findex.htm`. Accessed 11 December 2014.

[120] de Hoon M, Imoto S, Nolan J, Miyano S (2004) Open Source Clustering Software. Bioinformatics, 20 (9): 1453–1454.

[121] Horvat, Szabolcs, et al. "Spatial Embedding and Wiring Cost Constrain the Functional Layout of the Cortical Network of Rodents and Primates." PLoS Biol 14.7 (2016): e1002512.

[122] Huntley, G. Wt, et al. "Distribution and synaptic localization of immunocytochemically identified NMDA receptor subunit proteins in sensory-motor and visual cortices of monkey and human." The Journal of neuroscience 14.6 (1994): 3603-3619.

[123] W. James, The Principles of Psychology (Holt, New York, 1890).

[124] Jordan D, Stockmanns G, Kochs E et al (2008) Electroencephalographic order pattern analysis for the separation of consciousness and unconsciousness: An analysis of approximate entropy, permutation entropy, recurrence rate, and phase coupling of order recurrence plots. Anesthesiology 109:1014-22.

[125] Kaskinoro K, Maksimow A, Langsjo J, et al. (2011) Wide inter-individual variability of bispectral index and spectral entropy at loss of consciousness during increasing concentrations of dexmedetomidine, propofol, and sevoflurane.Br. J. Anaesth., doi:10.1093/bja/aer196.

[126] F. Kaspar, H. G. Schuster, Easily calculable measure for the complexity of spatiotemporalpatterns. Phys. Rev. A 36, 842-848 (1987).

[127] King JR, Sitt JD, Faugeras F, et al. Information sharing in the brain indexes consciousness in noncommunicative patients. Curr Biol. 2013;23:1914-1919.

[128] Koch, Christof. "The Quest for Consciousness: A Neurobiological Approach (Englewood, CO: Roberts and Company)." (2004).

[129] Koch, Christof, and Klaus Hepp. "Quantum mechanics in the brain." Nature 440.7084 (2006): 611-611.

[130] Koch, Christof. Consciousness: Confessions of a romantic reductionist. MIT press, 2012.

[131] Koch, Christof, et al. "Neural correlates of consciousness: progress and problems." Nature Reviews Neuroscience 17.5 (2016): 307-321.

[132] Kolmogorov, Andrei N. "On tables of random numbers." Sankhya: The Indian Journal of Statistics, Series A (1963): 369-376.

[133] Kometer, Michael, et al. "Psilocybin-induced spiritual experiences and insightfulness are associated with synchronization of neuronal oscillations." Psychopharmacology 232.19 (2015): 3663-3676.

[134] Koubeissi, Mohamad Z., et al. "Electrical stimulation of a small brain area reversibly disrupts consciousness." Epilepsy and behaviour 37 (2014): 32-35.

[135] Kuramoto, Yoshiki (1975). H. Araki, ed. Lecture Notes in Physics, International Symposium on Mathematical Problems in Theoretical Physics 39. Springer-Verlag, New York. p. 420

[136] Lamme, Victor AF, and Pieter R. Roelfsema. "The distinct modes of vision offered by feedforward and recurrent processing." Trends in neurosciences 23.11 (2000): 571-579.

[137] Lamme, Victor AF. "Towards a true neural stance on consciousness." Trends in cognitive sciences 10.11 (2006): 494-501.

[138] Larson, Gerald James (2008). The Encyclopedia of Indian Philosophies: Yoga: India's philosophy of meditation. Motilal Banarsidass. ISBN 978-81-208-3349-4.

[139] Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. Trends Cogn. Sci. 9, 556-559. doi: 10.1016/j.tics.2005.10.010

[140] Laureys, Steven, et al. "Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome." BMC medicine 8.1 (2010): 1.

[141] Lebedev, A. V., et al. "LSDâĂŘinduced entropic brain activity predicts subsequent personality change." Human brain mapping (2016).

[142] Lee, UnCheol, et al. "Propofol induction reduces the capacity for neural information integration: implications for the mechanism of consciousness and general anesthesia." Consciousness and cognition 18.1 (2009): 56-64.

[143] Lehnert, Judith. Controlling synchronization patterns in complex networks. Springer, 2015.

[144] Lemon, R. N., and S. A. Edgley. "Life without a cerebellum." Brain 133.3 (2010): 652-654.

[145] Lempel A, Ziv J (1976) On the complexity of finite sequences. IEEE Trans. Inform. Theory 22, 75-81.

[146] Lempel, Abraham, and Jacob Ziv. "On the complexity of finite sequences." IEEE Transactions on information theory 22.1 (1976): 75-81.

[147] Leszl, Walter. "Democritus' works: from their titles to their contents." Democritus: Science, The Arts, and the Care of the Soul. Brill, 2006. 11-76.

[148] Lewis L, Weiner V, Mukamel E, et al. (2012) Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. Proc. Natl. Acad. of Sci. Plus, 109(49):E3377-86

[149] Liang, Z., Wang, Y. (2015). EEG entropy measures in anesthesia. Frontiers in Computational Neuroscience, 9, 16. http://doi.org/10.3389/fncom.2015.00016

[150] Liu, Zhiyong, and Jinwei Sun. "Sleep Staging from the EEG Signal Using Multifractal Detrended Fluctuation Analysis." 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC). IEEE, 2015.

[151] Lv, Jing, David M. Simpson, and Steven L. Bell. "Objective detection of evoked potentials using a bootstrap technique." Medical engineering & physics 29.2 (2007): 191-198.

[152] Massimini M, Huber R, Ferrarelli F, Hill S, Tononi G. (2004) The sleep slow oscillation as a traveling wave. J Neurosci 24(31):6862-70.

[153] Massimini M, Ferrarelli F, Huber R, et al (2005) Breakdown of cortical effective connectivity during sleep, Science, 309, pp. 22.

[154] McCormick, David A., Matthew J. McGinley, and David B. Salkoff. "Brain state dependent activity in the cortex and thalamus." Current opinion in neurobiology 31 (2015): 133-140.

[155] Miller, Steven M., ed. The Constitution of Phenomenal Consciousness: Toward a science and theory. Vol. 92. John Benjamins Publishing Company, 2015.

[156] Modha D, Raghavendra S (2010) Network architecture of the long-distance pathways in the macaque brain; PNAS

[157] Moon J-Y, Lee U, Blain-Moraes S, Mashour GA (2015) General Relationship of Global Topology, Local Dynamics, and Directionality in Large-Scale Brain Networks. PLoS Comput Biol 11(4): e1004225. doi: 10.1371/journal.pcbi.1004225

[158] Murphy M, Bruno M, Riedner B, et al. (2011) Propofol anesthesia and sleep: a high-density EEG study. Sleep 34:283-291A.

[159] Muthukumaraswamy, Suresh D., et al. "Broadband cortical desynchronization underlies the human psychedelic state." The Journal of Neuroscience 33.38 (2013): 15171-15183.

[160] Muthukumaraswamy, S. D., Shaw, A. D., Jackson, L. E., Hall, J., Moran, R., Saxena, N. (2015). Evidence that Subanesthetic Doses of Ketamine Cause Sustained Disruptions of NMDA and AMPA-Mediated Frontoparietal Connectivity in Humans. The Journal of neuroscience : the official journal of the Society for Neuroscience, 35 (33), 11694-11706.

[161] Naci, Lorina, et al. "A common neural code for similar experiences in different individuals." Proceedings of the National Academy of Sciences 111.39 (2014): 14277-14282.

[162] Nagel, Thomas. "What is it like to be a bat?." The philosophical review 83.4 (1974): 435-450.

[163] Nir Y, Staba RJ, Andrillon T, Vyazovskiy VV, Cirelli C, Fried I, Tononi G (2011) Regional slow waves and spindles in human sleep. Neuron

[164] Nolte, Guido. "The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors." Physics in medicine and biology 48.22 (2003): 3637.

[165] Nonnenmacher, Marcel, et al. "Signatures of criticality arise in simple neural population models with correlations." arXiv preprint arXiv:1603.00097 (2016).

[166] Oizumi M, Albantakis L, Tononi G (2014) From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. PLoS Comput Biol 10(5): e1003588. doi:10.1371/journal.pcbi.1003588.

[167] Oizumi, Masafumi, et al. "Measuring integrated information from the decoding perspective." PLoS Comput Biol 12.1 (2016): e1004654.

[168] Oostenveld, Robert, et al. "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data." Computational intelligence and neuroscience 2011 (2010).

[169] Oostenveld, Robert, et al. "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data." Computational intelligence and neuroscience 2011 (2010).

[170] Overgaard, Morten, and Rikke Overgaard. "Neural correlates of contents and levels of consciousness." Frontiers in psychology 1 (2010): 164.

[171] Palus, Milan, et al. "Synchronization as adjustment of information rates: detection from bivariate time series." Physical Review E 63.4 (2001): 046211.

[172] Patel, Kartikeya C. "Eastern Traditions, Consciousness, and Spirituality." The Oxford Handbook of Psychology and Spirituality (2012): 343.

[173] Perez F, Granger B (2007) IPython: A System for Interactive Scientific Computing. Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, `http://ipython.org`

[174] Pigorini, Andrea, et al. "Bistability breaks-off deterministic responses to intracortical stimulation during non-REM sleep." Neuroimage 112 (2015): 105-113.

[175] Pilge, Stefanie, et al. "Differences between state entropy and bispectral index during analysis of identical electroencephalogram signals: A comparison with two randomised anaesthetic techniques." European Journal of Anaesthesiology (EJA) 32.5 (2015): 354-365.

[176] Pincus S (1991) Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America 88(6):2297-2301.

[177] Pockett S. (2000). The Nature of Consciousness: A Hypothesis. Lincoln; NE: iUniverse.com

[178] Pockett S. (2012). The electromagnetic field theory of consciousness: a testable hypothesis about the characteristics of conscious as opposed to non-conscious fields. J. Conscious. Stud. 19, 191-223

[179] Popovych, Oleksandr V., Christian Hauptmann, and Peter A. Tass. "Control of neuronal synchrony by nonlinear delayed feedback." Biological cybernetics 95.1 (2006): 69-85.

[180] Posner, J. B., Saper, C. B., Schiff, N. D. and Plum, F. Plum and Posner's Diagnosis of Stupor and Coma (Oxford University Press, 2007).

[181] Priesemann, Viola, Matthias HJ Munk, and Michael Wibral. "Subsampling effects in neuronal avalanche distributions recorded in vivo." BMC neuroscience 10.1 (2009): 1.

[182] Priesemann V, Valderrama M, Wibral M, Le Van Quyen M. Neuronal avalanches differ from wakefulness to deep sleep-evidence from intracranial depth recordings in humans. PLoS Comput Biol. 2013;9:e1002985. doi: 10.1371/journal.pcbi.1002985.

[183] Ramsay M, Savege T, Simpson B, Goodwin R (1974) Controlled sedation with alphaxalonealphadolone. Br Med J 2: 656-9.

[184] Rees, G., Kreiman, G. and Koch, C. Neural correlates of consciousness in humans. Nat. Rev. Neurosci. 3, 261-270 (2002).

[185] Ribeiro TL, Copelli M, Caixeta F, Belchior H, Chialvo DR, et al. (2010) Spike Avalanches Exhibit Universal Dynamics across the Sleep-Wake Cycle. PLoS ONE

[186] Riedner, B. A., Vyazovskiy, V. V., Huber, R., Massimini, M., Esser, S., Murphy, M., Tononi, G. (2007). Sleep Homeostasis and Cortical Synchronization: III. A High-Density EEG Study of Sleep Slow Waves in Humans. Sleep, 30(12), 1643-1657.

[187] Rosanova M, Gosseries O, Casarotto S, et al. (2012) Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. Brain 135, 1308-1320.

[188] Sara M, Pistoia F (2010) Complexity loss in physiological time series of patients in a vegetative state. Nonlinear Dynamics Psychol Life Sci;14:1-13.

[189] Sarasso S, Rosanova M, Casali A, et al. (2014) Quantifying Cortical EEG Responses to TMS in (Un)consciousness. Clinical EEG and Neuroscience, January 2014 vol. 45 no. 1 40-49.

[190] Sarasso, Simone, et al. "Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine." Current Biology 25.23 (2015): 3099-3105.

[191] Schartner M, Seth A, Noirhomme Q, Boly M, Bruno M-A, et al. (2015) Complexity of Multi-Dimensional Spontaneous EEG Decreases during Propofol Induced General Anaesthesia. PLoS ONE 10(8): e0133532. doi: 10.1371/journal.pone.0133532

[192] Schartner, Michael M., et al. "Global and local complexity of intracranial EEG decreases during NREM sleep." Neuroscience of Consciousness (2016).

[193] Schmidt R, LaFleur KJR, de Reus MA, van den Berg LH, van den Heuvel MP. Kuramoto model simulation of neural hubs and dynamic synchrony in the human cerebral connectome. BMC Neuroscience. 2015;16:54. doi:10.1186/s12868-015-0193-z.

[194] Schöll, Eckehard, Sabine HL Klapp, and Philipp Hövel. Control of self-organizing nonlinear systems. Springer, Berlin, 2016.

[195] Schreiber, Thomas. "Measuring information transfer." Physical review letters 85.2 (2000): 461.

[196] Schrouff, Jessica, et al. "Brain functional integration decreases during propofol-induced loss of consciousness." Neuroimage 57.1 (2011): 198-205.

[197] Scott G, Fagerholm E et al (2014) Voltage Imaging of Waking Mouse Cortex Reveals Emergence of Critical Neuronal Dynamics;

[198] Seager, William. "Panpsychism." Encyclopedia of Cognitive Science (2005).

[199] Seth A (2005) Causal connectivity of evolved neural networks during behaviour. Network Comput Neural Sys 16:35-54.

[200] Seth A, Izhikevich E, Reeke G, Edelman G (2006) Theories and measures of consciousness: An extended framework. Proc Natl Acad Sci USA 103.

[201] Seth, Anil K., et al. "Measuring consciousness: relating behavioural and neurophysiological approaches." Trends in cognitive sciences 12.8 (2008): 314-321.

[202] Seth, Anil. "Explanatory correlates of consciousness: theoretical and computational challenges." Cognitive Computation 1.1 (2009): 50-63.

[203] Seth, Anil K., Keisuke Suzuki, and Hugo D. Critchley. "An interoceptive predictive coding model of conscious presence." Frontiers in psychology 2 (2011).

[204] Seth A, Barrett A, Barnett L (2011) Causal density and integrated information as measures of conscious level. Philos. Trans. A Math. Phys. Eng. Sci. 369, 3748-3767.

[205] Seth, Anil K., Adam B. Barrett, and Lionel Barnett. "Granger causality analysis in neuroscience and neuroimaging." The Journal of Neuroscience 35.8 (2015): 3293-3297.

[206] Shanahan, Murray, and Bernard Baars. "Applying global workspace theory to the frame problem." Cognition 98.2 (2005): 157-176.

[207] Shanahan, Murray. "A cognitive architecture that combines internal simulation with a global workspace." Consciousness and cognition 15.2 (2006): 433-449.

[208] Shanahan, Murray. "Dynamical complexity in small-world networks of spiking neurons." Physical Review E 78.4 (2008): 041924.

[209] Shanahan M (2010) Metastable chimera states in community-structured oscillator networks. Chaos, Vol:20, ISSN:1054-1500.

[210] Shaw, Fu-Zen, et al. "Algorithmic complexity as an index of cortical function in awake and pentobarbital-anesthetized rats." Journal of neuroscience methods 93.2 (1999): 101-110.

[211] Sheroziya, Maxim, and Igor Timofeev. "Global intracellular slow-wave dynamics of the thalamocortical system." The Journal of Neuroscience 34.26 (2014): 8875-8893.

[212] Sigl JC, Chamoun NG (Nov 1994). "An introduction to bispectral analysis for the electroencephalogram". Journal of Clinical Monitoring 10 (6): 392-404.

[213] Silber, M., S. Ancoli-Israel, M. Bonnet, et al . 2007. The visual scoring of sleep in adults. J. Clin. Sleep Med. 15: 121-131.

[214] Sitt, Jacobo Diego, et al. "Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state." Brain 137.8 (2014): 2258-2270.

[215] Skinner, Burrhus Frederic. "behaviourism at fifty." Science 140.3570 (1963): 951-958.

[216] Stender, Johan, et al. "The Minimal Energetic Requirement of Sustained Awareness after Brain Injury." Current Biology 26.11 (2016): 1494-1499.

[217] Steriade M, McCormick D, Sejnowski T (1993) Thalamocortical oscillations in the sleeping and aroused brain. Science. 262:679-85.

[218] Steriade, Mircea, and Robert W. McCarley. "Intrinsic electrophysiological properties of brainstem and forebrain neurons." Brain Control of Wakefulness and Sleep (2005): 155-209.

[219] Strogatz, S. H. (2000) From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. Physica D 143, 1-20

[220] Sullivan G M, Feinn R (2012) Using Effect Size - or Why the P Value Is Not Enough. Journal of Graduate Medical Education, 4(3), 279-282.

[221] Tagliazucchi E, et al. (2013) Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep. Proc Natl Acad Sci USA 110(38):15419-15424.

[222] Tagliazucchi, Enzo, et al. "Enhanced repertoire of brain dynamical states during the psychedelic experience." Human brain mapping 35.11 (2014): 5442-5456.

[223] Tagliazucchi, E., Chialvo, D. R., Siniatchkin, M., Brichant, J. F., Bonhomme, V., Noirhomme, Q., et al. (2015). Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. Journal of the Royal Society Interface

[224] Teasdale, G. M., and L. Murray. "Revisiting the Glasgow coma scale and coma score." Intensive care medicine 26.2 (2000): 153-154.

[225] Tegmark, Max. "Consciousness as a state of matter." Chaos, Solitons and Fractals 76 (2015): 238-270.

[226] Tegmark M (2016) Improved Measures of Integrated Information. PLOS Computational Biology 12(11)

[227] Tindall, Suzie C. "Level of consciousness." (1990), Butterworths, Boston.

[228] Tononi G, Sporns O, Edelman G (1994) A measure for brain complexity: Relating functional segregation and integration in the nervous system. Proc. Natl. Acad. Sci. USA. 91, 5033-5037.

[229] Tononi G, Edelman G (1998) Consciousness and complexity. Science 282, 1846-1851.

[230] Tononi G, Sporns O (2003) Measuring information integration. BMC Neurosci 4: 31.

[231] Tononi G (2004) An information integration theory of consciousness. BMC Neurosci. 5, 42

[232] Tononi G. Consciousness as integrated information: a provisional manifesto. Biol Bull. 2008; 215: 216-242. doi: 10.2307/25470707. pmid:19098144

[233] Tononi, Giulio. Phi: A Voyage from the Brain to the Soul. Pantheon Books, 2012.

[234] Tononi, Giulio, and Christof Koch. "Consciousness: here, there and everywhere?." Phil. Trans. R. Soc. B 370.1668 (2015): 20140167.

[235] Tononi, Giulio, et al. "Integrated information theory: from consciousness to its physical substrate." Nature Reviews Neuroscience (2016).

[236] Touboul, Jonathan, and Alain Destexhe. "Can power-law scaling and neuronal avalanches arise from stochastic dynamics?." PloS one 5.2 (2010): e8982.

[237] Tsai, Pei-Jung, et al. "Local awakening: regional reorganizations of brain oscillations after sleep." Neuroimage 102 (2014): 894-903.

[238] Tzourio-Mazoyer, Nathalie, et al. "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain." Neuroimage 15.1 (2002): 273-289.

[239] Valentin, A., Anderson, M., Alarcon, G., Seoane, J.J.G., Selway, R., Binnie, C.D., Polkey, C.E., (2002). Responses to single pulse electrical stimulation identify epileptogenesis in the human brain in vivo. Brain 125, 1709-1718.

[240] Van Putten, M. J. A. M., and C. J. Stam. "Application of a neural complexity measure to multichannel EEG." Physics Letters A 281.2 (2001): 131-141.

[241] Van Veen, Barry D., et al. "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering." IEEE Transactions on biomedical engineering 44.9 (1997): 867-880.

[242] van Walsum, A-M. van Cappellen, et al. "A neural complexity measure applied to MEG data in Alzheimer's disease." Clinical neurophysiology 114.6 (2003): 1034-1040

[243] Vanderwolf, C. H. "Are neocortical gamma waves related to consciousness?." Brain research 855.2 (2000): 217-224.

[244] Viertio-Oja H, Maja V, Sarkela M, et al. (2004) Description of the entropy(TM) algorithm as applied in the datex-ohmeda S/5TM entropy module. Acta Anaesthesiol Scand 48:154-61.

[245] Vyazovskiy VV, Harris KD. (2013) Sleep and the single neuron: the role of global slow oscillations in individual cell rest. Nat Rev Neurosci 14(6):443-51.

[246] Watson, John B. behaviourism. Read Books Ltd, 2013.

[247] Welch, Peter D. "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms." IEEE Transactions on audio and electroacoustics 15.2 (1967): 70-73.

[248] Welch, Terry A. "A technique for high-performance data compression." Computer 17.6 (1984): 8-19.

[249] Wildie M, Shanahan M (2012) Metastability and Chimera States in Modular Delay and Pulse-coupled Oscillator Networks. Chaos, vol. 22, 043131.

[250] Wu D, Cai G, Yuan Y, et al. (2011) Application of nonlinear dynamics analysis in assessing unconsciousness: A preliminary study, Clinical Neurophysiology, Volume 122, Issue 3, Pages 490-498, ISSN 1388-2457, `http://dx.doi.org/10.1016/j.clinph.2010.05.036`.

[251] Xia, Mingrui, Jinhui Wang, and Yong He. "BrainNet Viewer: a network visualization tool for human brain connectomics." PloS one 8.7 (2013): e68910.

[252] Yu, Feng, et al. "A new case of complete primary cerebellar agenesis: clinical and imaging findings in a living patient." Brain (2014): awu239.

[253] Zhang, X-S., Rob J. Roy, and Erik W. Jensen. "EEG complexity as a measure of depth of anesthesia for patients." IEEE Transactions on Biomedical Engineering 48.12 (2001): 1424-1433.

[254] Zhou, Cheng, et al. "Forebrain HCN1 channels contribute to hypnotic actions of ketamine." The Journal of the American Society of Anesthesiologists 118.4 (2013): 785-795.

[255] Ziv, Jacob, and Abraham Lempel. "Compression of individual sequences via variable-rate coding." IEEE transactions on Information Theory 24.5 (1978): 530-536.

[256] Zunino, L., Perez, D., Kowalski, A., Martin, M., Garavaglia, M., Plastino, A., et al. (2008). Fractional Brownian motion, fractional Gaussian noise, and Tsallis permutation entropy. Phys. A Statist. Mech. Appl. 387, 6057-6068. doi: 10.1016/j.physa.2008.07.004

[257] Zurek, Wojciech H. "Thermodynamic cost of computation, algorithmic complexity and the information metric." Nature 341 (1989): 119-124.

[258] `http://rosettacode.org/wiki/LZW_compression#Python`

Appendix A

Supplementary material for propofol/EEG study

A.1 Dependence on channel number

In order to test whether the number of channels of the input data affected the behaviour of the measures for the state pair WR/LOC, we tried 5, 10, 50, and 100 k-medoids selected electrodes in addition to the 25 (see Fig. A.1). We found all three measures' scores to be higher for WR than LOC for all tested channel numbers. For all subjects and measures, the differences between WR and LOC had high effect sizes (Cohen's $d > 0.8$) except for one subject for all three measures and channel numbers and two more subjects for measure SCE and certain channels (see subject label colour in Fig. A.1). Despite varying effect sizes, this control suggests that the measures' scores are across subjects higher for WR than LOC for a broad range of channel numbers.

A.2 Dependence on segment length

As a second control we tested whether segment length influences the behaviour of our measures. Analysed across segments of length other than $10sec$, we found for ACE, SCE and LZc very similar results in almost all cases for all tested segment lengths (0.2, 0.4, 0.8, 1.2, 2, 4, 6, 10 and 20 seconds); except for one subject the behaviour of the measures did not vary with segment length (Fig. A.2). That is, the measures scored higher for WR than LOC with high effect size (Cohen's $d > 0.8$) across all tried segment lengths for 6 of the 7 subjects. Subject 1 was an exception, with LZc, SCE and ACE showing the same behaviour as for the other subjects only for segment lengths greater than 4 seconds. In sum, this control shows that the behaviour of the measures is robust across a large range of segment lengths.

A.3 Control for changes in power spectrum

To test directly whether changes in the signal diversity measures arose from more than just changes in the spectral profile of the input signal, we compared two different types

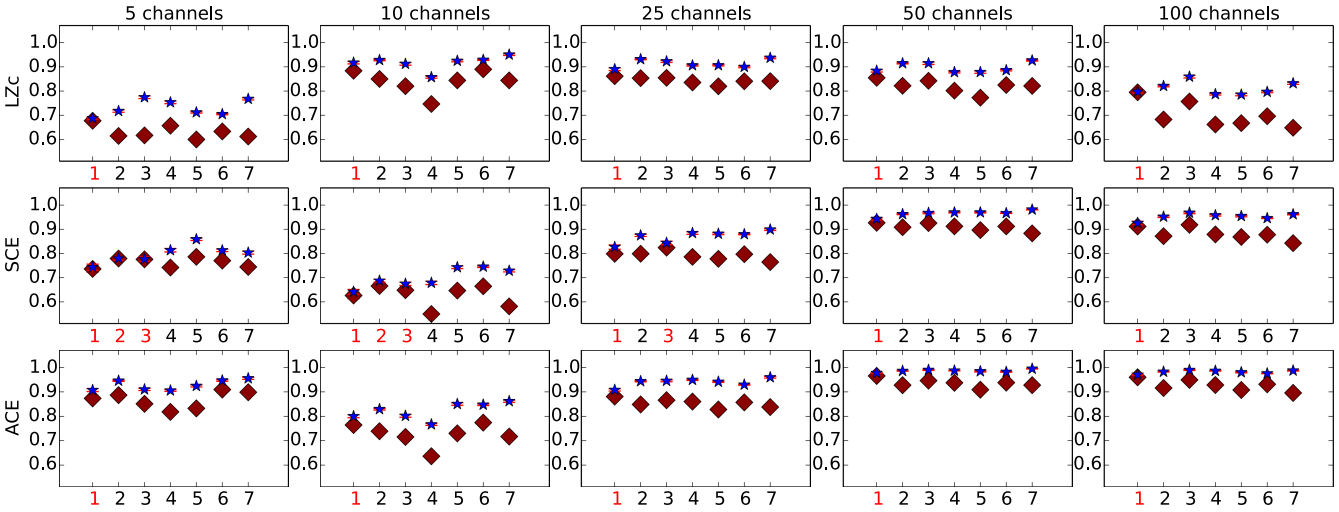

Figure A.1: **WR (asterisks) and LOC (diamonds) scores of LZc, SCE and ACE for different numbers of electrodes.** $\{5, 10, 25, 50, 100\}$ electrodes were chosen via k-medoids from the whole cortex. For all subjects and channel numbers, the score of any of the three measures is higher for WR than for LOC. Red subject labels indicate small effect size, i.e. Cohen's $d < 0.8$. For LZc and ACE only subject 1 shows small effect size for any channel number. For SCE up to three subjects show small effect size for certain channels.

of normalisation of the measures. First, as used for the results in the main text, 'time shuffling', obtained from surrogate data for which the measure reaches its upper bound, but that has a flattened spectral profile (see Chapter 2). Second, 'phase shuffling', obtained from surrogate data with a conserved spectral profile. We find qualitatively similar results for both types of normalisation, and thus confirm that changes in the signal diversity measures reflect more than just spectral changes.

The 'phase shuffling' normalisation is obtained from phase-randomised surrogate data as follows. From the complete data from a given subject in a given state, a segment is randomly chosen. Each time series of that segment is expressed as a superposition of sinusoids using fast-Fourier transform. Then the phase of each sinusoid is independently randomly changed, before applying inverse Fourier transform. The signal diversity measure is computed for 100 such phase-randomised data segments. The mean of these 100 scores is then used to normalise the measure's score for the original data segments. Averaging over 100 such surrogate data segments suffices to obtain negligible randomness in the normalisation factor.

Fig. A.3 is a comparison of the three signal diversity measures' scores per subject for the two normalisation methods. As can be seen when comparing the measures' results subject by subject, the qualitative behaviour of all three measures is very similar under both normalisations, implying that the signal diversity we measure is not trivially connected to spectral properties.

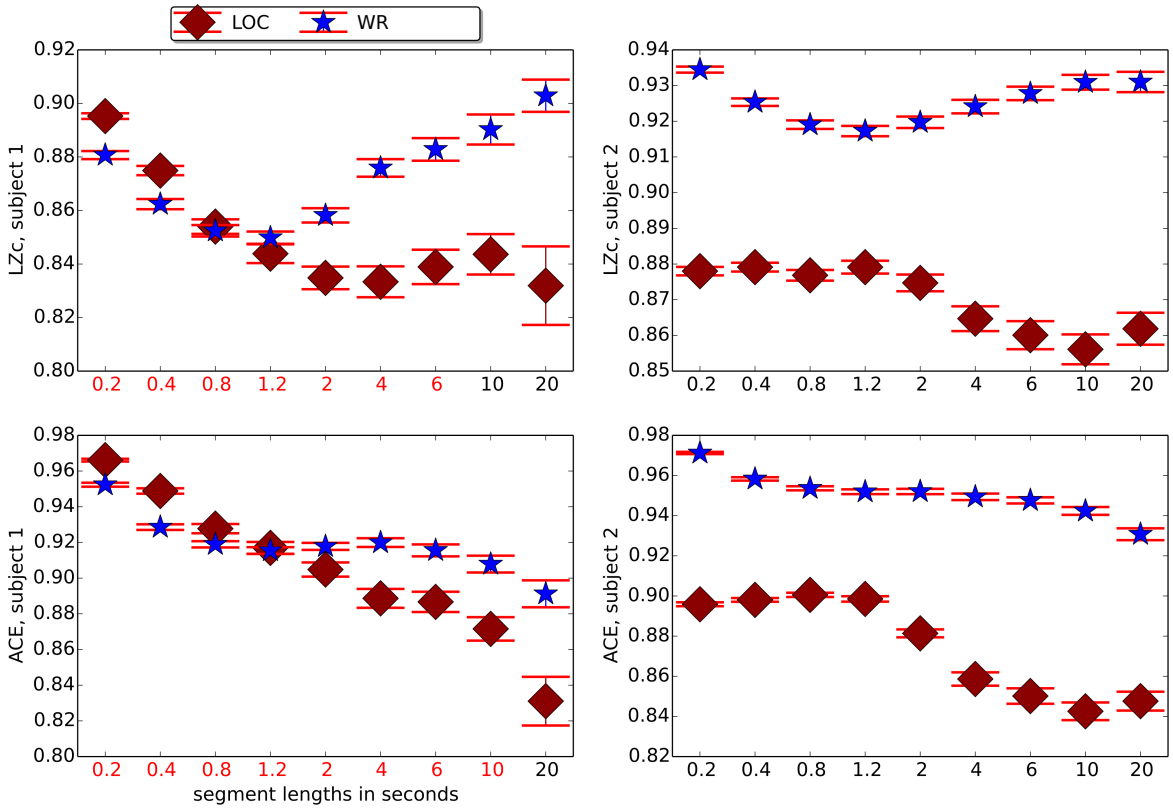

Figure A.2: **Sensitivity of the measures to segment length.** Segment lengths are indicated in seconds, coloured red if the difference between WR and LOC has effect size $d < 0.8$. There is one out of 7 subjects, subject 1, for which the segment length changed the measures LZc, SCE and ACE from being greater for WR than for LOC to the opposite, i.e. smaller for WR than for LOC. This change happened at a segment length close to 1 second. This can be seen in the left two panels, which indicate respectively the score of LZc and ACE, each for the two states WR and LOC for subject 1. For the remaining 6 of the 7 subjects, the segment length of the input had no influence on the order of the LZc (SCE, ACE) scores for WR and LOC, as is exemplified in the right two panels with subject 2. (Similar results hold for SCE, not shown).

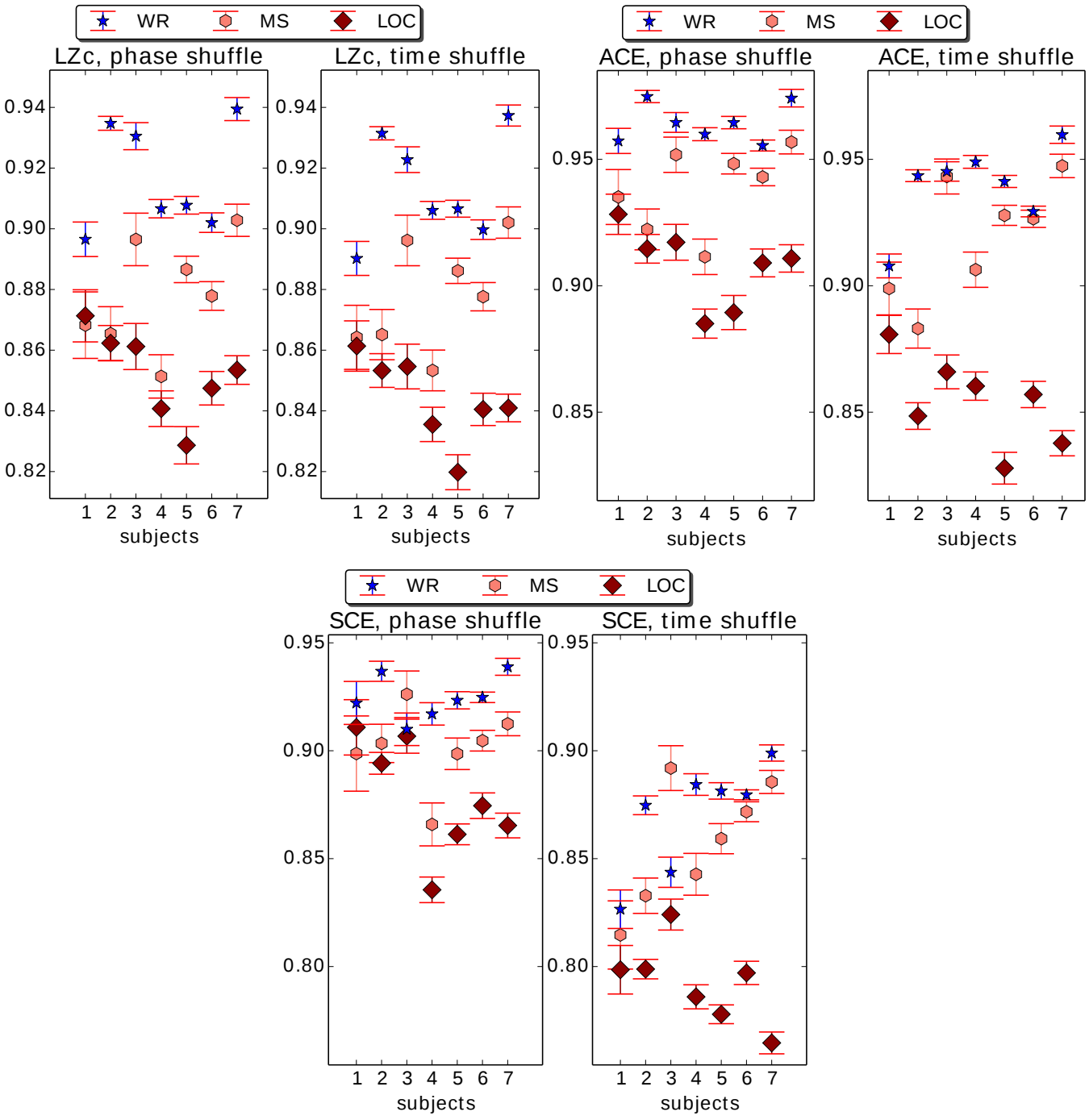

Figure A.3: **Invariance of the measures' qualitative behaviour under spectral-profile-preserving normalisation.** LZc, SCE and ACE computed as averages over multiple 10sec segments of EEG of the 7 subjects before and during anaesthesia, as described in caption for Fig. 3.3, yet here for two different normalisations. The titles indicate which measure and which normalisation is displayed (see main text for details). The measures' qualitative behaviour is very similar for the two normalisations.

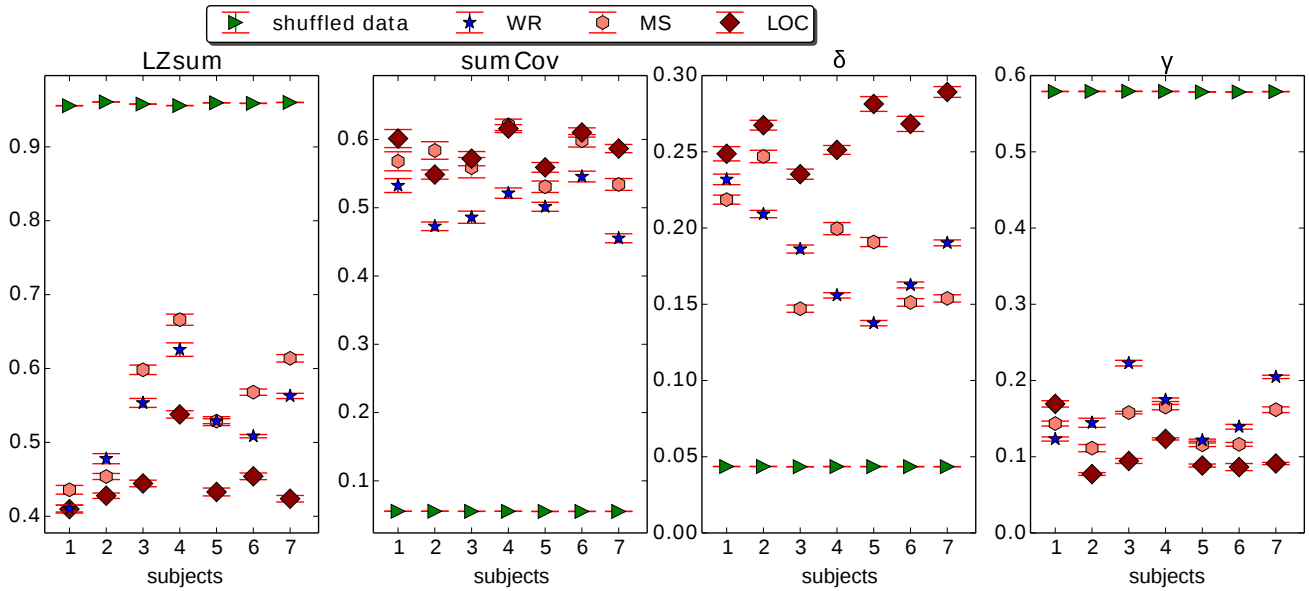

Figure A.4: **Mean LZsum, sumCov, normalised delta power and normalised gamma power**, computed for broadband signal of 25 channels from the whole cortex, 10*sec* segments, i.e. the same input as for Fig. 3.3 and Tab. 3.1. States shown are WR, MS, LOC and shuffled data. sumCov and delta power discriminate WR/LOC consistently across subjects, whereas neither scores for MS in between WR and LOC consistently across subjects. LZsum separates WR/LOC for 6/7 subjects and has poor discrimination for MS/LOC, unlike LZc. Error bars indicate standard error across segments.

A.4 Results for alternative measures

Besides LZc, SCE and ACE, the following alternative measures scored for certain state pairs consistently different across subjects (compare with Fig. 3.3 and Tab. 3.1). Fig. A.4 displays that normalised delta power and sumCov (average correlation of the signals) are both consistently higher for LOC than for WR across subjects. Note that neither of them scores for MS consistently in between WR and LOC, whereas LZc, SCE and ACE do so for all except subject 2 for SCE. The Lempel-Ziv variant LZsum and gamma power also discriminate WR/LOC well, except for the first subject, see Fig. A.4. Interestingly, LZsum performs less well in discriminating MS/WR, showing that the way in which the binary matrix was concatenated influenced the outcome. I.e. the observation-by-observation concatenation used for LZc results in capturing randomness across channels as opposed to capturing average randomness per channel, as used in the computation of LZsum.

Appendix B

Supplementary material for sleep/depth electrode study

B.1 Surrogate data with empirical power spectra

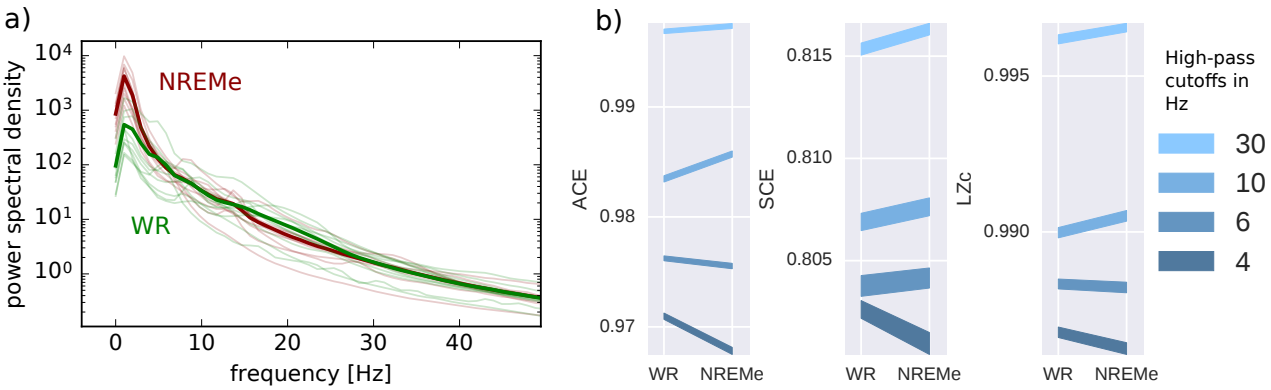

Figure B.1: **Surrogate data with average spectral profile for wake or NREM sleep.** a) The depth electrode data set of sleeping subjects (Chapter 4) was used to obtain the average spectral power density of wakeful rest (WR) and non-rapid eye movement sleep (NREMe), respectively. The thin lines show power spectra per subject (averaged across channels per subject) and the thick line is the average across subjects. Delta power density is clearly higher for NREMe than WR. b) Surrogate data with each spectrum was created (see text for details) and ACE, SCE and LZc computed for both spectra after high-pass filtering at the indicated cutoffs (linewidth indicating standard error across 100 trials). All three measures score higher for the spectrum corresponding to wakeful rest for a 4Hz high-pass cutoff, however for cutoffs of 10Hz, all three score higher for NREMe than WR. Importantly, the measures did still score higher for a high-pass filtered depth-electrode recording at 10Hz, see Fig. 4.8, showing that spectral changes alone are not enough to explain that they index conscious level.

In order to test to what extent the sensitivity of the measures to conscious level depends on spectral changes, surrogate data were created with the spectral profile of wakeful rest (WR) and non-rapid eye movement sleep (NREMe), then different high-pass frequency filters applied prior to computing the measures ACE, LZc and SCE. Fig. B.1a show spectral power density averaged across channels and trials (thin lines) and across subjects (thick

lines) of the depth-electrode data set of subjects in different sleep stages (Chapter 4). The spectra differ visibly, with NREM-sleep having more spectral power in the delta range (1-4Hz) compared to wakeful rest. Surrogate time series with these power spectral densities were obtained as follows. The discrete fast Fourier transform of a uniform random signal was computed, the amplitudes normalised to one and then multiplied by the desired spectrum before computing the inverse Fourier transform. ACE, SCE and LZc were each computed for 100 data segments of 10 channels and 2500 observations (corresponding to 10s at 250Hz), for each state (WR, NREMe) and high-pass filter cutoff (4Hz, 6Hz, 10Hz, 30Hz).

All three measures score higher for the spectrum corresponding to wakeful rest for a 4Hz high-pass cutoff, however for cutoffs of 10Hz, all three score higher for NREMe than WR. Importantly, the measures did still score higher for a high-pass filtered depth-electrode recording at 10Hz, see Fig. 4.8, showing that spectral changes alone are not enough to explain their power of indexing conscious level.

B.2 Supplementary measures for global analysis

In addition to ACE, SCE and LZc, we computed other measures for the global analysis of 18 channels per subject. These were normalised spectral power bands, LZsum (the mean Lempel-Ziv complexity of single channels) and sumCov (the mean of the absolute values of correlation coefficients between all channels). Tab. B.1 summarises the results. Normalised spectral power bands were obtained by fast Fourier transform of the segment, averaged over the channels and grouped into frequency bands, normalised such that all bands' scores together sum to one. Frequency bands are defined - following convention - as δ =1-4Hz, θ=4-8Hz, α =8-13Hz, β =13-30Hz, γ =30-70Hz. As expected, normalized delta power is consistently substantially higher in NREMe than WR, while beta and gamma power are in almost all cases substantially lower. LZsum behaved very similarly to LZc for state pairs WR/NREMe and REM/NREMe, yet was slightly less consistent at discriminating between WR or REM and NREMl. sumCov's behaviour slightly resembled that of delta power, yet was overall less consistent and with weaker effect sizes across subjects.

B.3 Correlation between measures

Correlations between ACE, SCE, LZc and spectral power in the 5 canonical frequency bands were analysed for each state separately and the complete results summarised in Tab. B.2. Correlations between all three signal diversity measures exist, showing that ACE, SCE and LZc are similarly sensitive to certain signal features, yet variations across states are also apparent. Scatter plots for each subject and selected measure pairs are shown in Fig. B.2, illustrating variation across subjects, with subject 7 as a visible outlier. These imperfections in correlations indicate that the three measures capture not entirely identical types of spatio-temporal signal diversity. Strong correlations were also found between signal diversity measures and spectral power in certain bands - especially delta - showing that the measures are affected by the spectral profile of the signal, yet the

	ACE	SCE	LZc	LZsum	sumCov	δ	θ	α	β	γ
WR/NREMe	**10 0 0**	**10 0 0**	**10 0 0**	**10 0 0**	0 4 6	**0 0 10**	5 1 4	7 2 1	9 1 0	9 0 1
WR/NREMl	8 2 0	8 2 0	8 2 0	6 4 0	0 5 5	1 0 9	4 4 2	7 3 0	7 2 1	5 4 1
WR/REM	4 5 1	2 8 0	1 9 0	4 4 2	0 7 3	0 2 8	4 3 3	7 3 0	3 5 2	2 3 5
REM/NREMe	**10 0 0**	**10 0 0**	9 1 0	9 1 0	1 2 7	0 1 9	4 2 4	5 2 3	**10 0 0**	8 2 0
REM/NREMl	9 1 0	8 2 0	8 2 0	7 2 1	1 1 8	1 3 6	2 7 1	3 3 4	7 2 1	5 3 2
NREMl/NREMe	7 2 1	8 2 0	2 8 0	8 2 0	1 8 1	0 2 8	2 5 3	6 1 3	8 2 0	7 3 0

Table B.1: **Comparisons of broader set of measures per state pair, using 18 channels per subject.** For each measure and state pair, the three numbers correspond to how many subjects out of 10 had higher score for the left state with Cohen's $d > 0.8$ (left digit), no substantial difference, $d < 0.8$, (middle) and higher score for the right state with Cohen's $d > 0.8$ (right). The results were obtained from applying the measures to the broadband signal of 18 channels.

decrease in measured signal diversity for WR/NREMe is not only due to spectral changes (see Fig. B.9). Tab. B.2 further shows an overall tendency for correlations to be strongest in state NREMl.

B.4 Local analysis

As described in Chapter 2, Lempel-Ziv complexity of single channels (LZs, same computation as LZc with trivial concatenation) scored predominantly lower during NREMe than WR, with some outliers. For each subject, a panel in Fig. B.4 displays LZs' scores for all 4 states (WR, REM, NREMl, NREMe), the region label of each channel as x-tick labels and gives in the panel title the number of channels for which LZs was higher for WR than NREMe with Cohen's $d > 0.8$ (see also figure caption). Across all subjects, 247 channels were analysed here. For 206 of those LZs scored higher for WR than NREMe with Cohen's $d > 0.8$, for 35 $d < 0.8$ and for 6 LZs scored lower for WR than NREMe with Cohen's $d > 0.8$. I.e. there were 6 channels for which LZs did not display the trend found globally. These 5 outliers are in subject 8 (Hi, In)[1], subject 9 (Hi, Oc) and subject 10 (Fr, In), i.e. different regions and different subjects, thus they do not indicate a local difference from the global trend. When considering LZs scores per subject and region, averaged across channels in the same region, LZs scored lower during NREMe than WR for 48 out of 50 subject region pairs (see Tab. B.3).

Fig. B.5 shows the scores for normalised delta power for the same data as Fig. B.4 (also applied to $10s$ segments), displaying delta(WR)<delta(NREMe) for most subjects. As can be seen by addition of the scores in the panel titles, there are in total 247 channels. For 210 of them, delta(WR)<delta(NREMe) with Cohen's $d > 0.8$, for 26 $d < 0.8$ and for 11 delta(WR)>delta(NREMe) with Cohen's $d > 0.8$. All these 11 channels - with delta behaving against the main trend - were in the same subject, in Hi, Oc and Te. Unlike

[1]Macro-region label abbreviations: Fr (frontal), Pa (parietal), Te (temporal), Oc (occipital), Ci (cingulate), In (insula), Hi (Hippocampus).

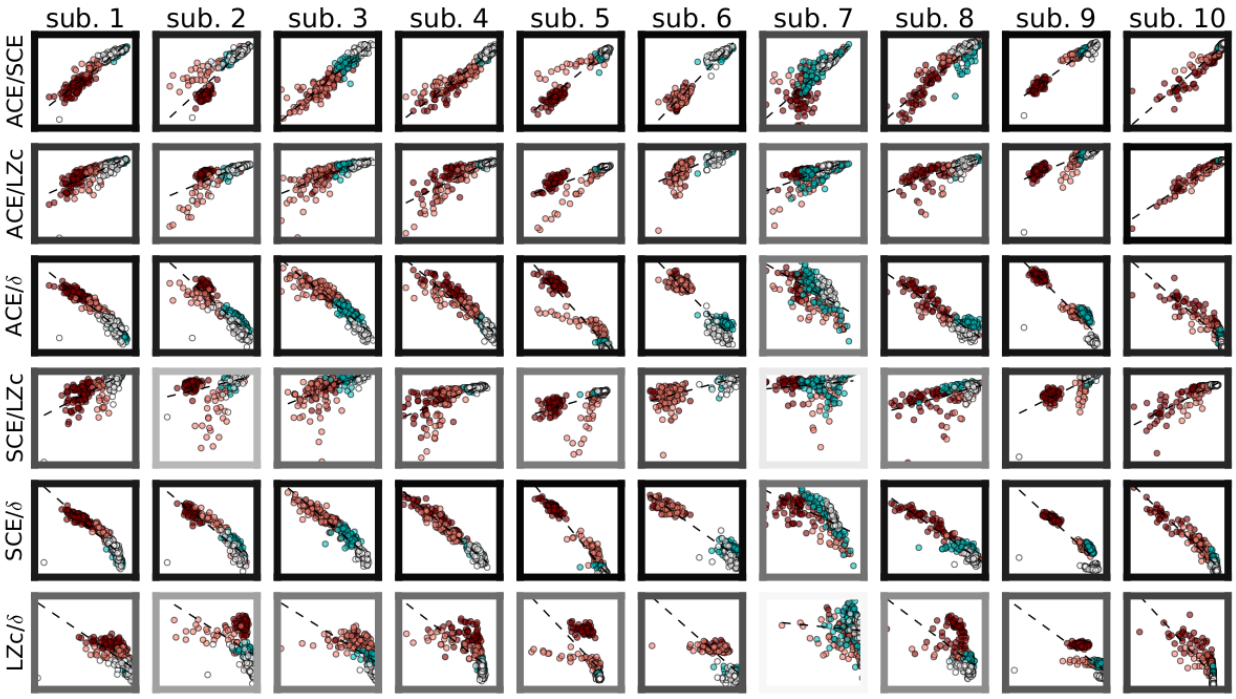

Figure B.2: **Scatter plots illustrating correlation of pairs of measures**. Each panel shows a scatter plot for a pair of measures (chosen from ACE, LZc, SCE and normalised delta power) and one of the ten subjects. Each dot - coloured by state as in main text figures - has as x value the score for a 10s segment of the first measure of the pair indicated for each row and as y value the score of the second measure. The best linear fit is indicated with a dotted line. The frame of each panel is coloured according to the absolute value of the Pearson correlation coefficient; white signifying 0, black signifying 1 and linearly scaled grey tones showing values in between. Measure pairs ACE/SCE, ACE/LZc and SCE/delta show high correlation across subjects. Subject 7 shows low correlation for all displayed measure pairs, demonstrating that the signal diversity measures are not all sensitive to the same dynamics. (compare with Tab. B.2)

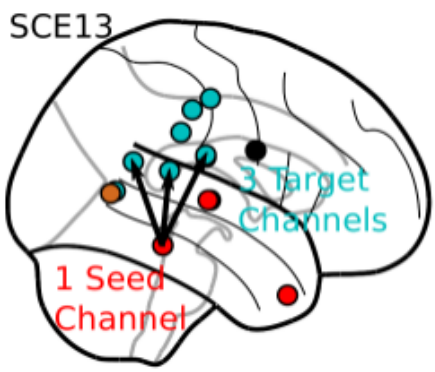

Figure B.3: **Schematic of the computation of SCE13**. This measure quantifies the diversity of the interaction between a seed channel in one region with three target channels in another region using the entropy $SCE^{(i)}$ of the constitution of the set of target channels in synchrony with the seed channel.

delta power, LZs did score according to the main trend also for these channels in subject 7 (compare Fig. B.4 with Fig. B.5), further adding evidence for signal diversity varying

Measure pair	WR	REM	NREMl	NREMe
ACE/LZc	0.71	0.6	0.86	0.72
ACE/SCE	0.72	0.73	0.84	0.71
ACE/delta	-0.56	-0.6	-0.76	-0.66
ACE/theta	-0.27	n.s.	n.s.	n.s.
ACE/alpha	n.s.	n.s.	n.s.	0.29
ACE/beta	0.5	0.43	0.62	0.5
ACE/gamma	0.57	0.66	0.65	0.54
LZc/SCE	0.51	0.39	0.62	0.37
LZc/delta	-0.32	-0.29	-0.56	-0.31
LZc/theta	n.s.	n.s.	n.s.	n.s.
LZc/alpha	n.s.	n.s.	n.s.	n.s.
LZc/beta	0.23	0.18	0.4	n.s.
LZc/gamma	0.28	0.26	0.45	0.26
SCE/delta	-0.46	-0.62	-0.82	-0.75
SCE/theta	n.s.	n.s.	n.s.	n.s.
SCE/alpha	n.s.	n.s.	n.s.	0.33
SCE/beta	0.44	0.48	0.69	0.65
SCE/gamma	0.36	0.5	0.7	0.62

Table B.2: **Pearson correlation for all measure pairs, pooled across subjects.**For a given state and subject the Pearson correlation of two different measures' scores across 10s segments was computed. For any state and pair of measures (chosen from ACE, SCE, LZc and spectral power of the 5 canonical frequency bands, excluding pairs of spectral power alone) the value of the Pearson correlation averaged across subjects, r, is displayed if $p < 0.05$ (obtained by a FDR corrected t-test across the values for the 10 subjects, setting 0 correlation as null hypothesis). If the p-value was above 0.05, the table entry was set to not significant (n.s.).

globally and locally robustly across subjects when changing state from WR to NREMe. Considering delta power per subject averaged across all channels of the same region, delta power is higher for NREMe than WR for 49 out of 50 region/subject pairs. The outlying region/subject pair is Oc in subject 7.

As described in Section 4.4.2.1 and illustrated in Appendix B.4, we explored the signal diversity, as measured by $SCE^{(i)}$, of the interaction of a seed channel in one region with a group of target channels in another. We call this measure SCE1N for the case of N target channels and computed it for $N = 2$ and $N = 3$ for all possible region pairs (seed channel in one region, N target channels in another region) for each subject. If there were multiple channel choices for a given region pair (e.g. the first subject had 7 parietal and 5 occipital channels, so when computing SCE13 for a $10s$ segment there were multiple choices for one channel in the parietal lobe and 3 channels in the occipital lobe) up to 10

different random combinations were chosen per $10s$ segment and results averaged across them. We also computed complexity of synchrony (CS) as being Lempel Ziv compression of the synchrony time series of two channels, thus capturing temporal signal diversity thereof. The scores of SCE13, SCE12 and CS can be compared between pairs of regions for single subjects with results shown in Figs. B.6 to B.8, respectively.

For each of these three measures of the diversity of the interaction of a channel in one region with channels in another region, we observed lower values during NREMe than WR for almost all choices of seed and target regions for all subjects. When using Cohen's $d > 0.8$ as a criterion for strong effect size, SCE13, computed for 129 pairs of regions across all subjects, scored higher for WR with $d > 0.8$ for 24, for 103 $d < 0.8$ and for 2 lower for NREMe than WR with $d > 0.8$. This overall result across subjects for SCE13 can be expressed as a triplet: $\{24, 103, 2\}$. The analogous triplet for SCE12 is $\{11, 130, 2\}$ and that for CS is $\{98, 33, 8\}$ (for CS we included also reflexive region pairs). The high fraction of cases with small effect size for SCE12 (130/142, as opposed to 24/129 for SCE13, and 33/139 for CS) indicates that the small number of possible synchrony coalitions for the case of 2 target regions leads to SCE12 being less effective as a measure of signal diversity.

There were no exceptional choices of seed and target region that went against the trend for more than one subject. Hence this SCE1N analysis did not yield evidence for a change in score against the main trend involving some region consistently in more than one subject, further supporting that signal signal diversity changes with sleep globally. Tab. B.3 summarises the consistency across regions of all the measures considered.

State pair	ACE	SCE	LZc	LZs	SCE13	SCE12	CS	delta	gamma
WR/NREMe	26 1	20 7	27 0	48 2	107 22	109 34	122 17	1 49	42 8
WR/NREMl	26 1	16 11	27 0	41 9	93 36	96 47	103 36	6 44	33 17
WR/REM	14 13	19 8	17 10	36 14	70 59	77 66	101 38	14 36	18 32
REM/NREMe	25 2	17 10	27 0	46 4	102 27	114 29	105 34	2 48	42 8
REM/NREMl	26 1	16 11	24 3	36 14	90 39	98 45	86 53	11 39	35 15
NREMl/NREMe	18 9	19 8	25 2	42 8	94 35	90 53	122 17	6 44	47 3

Table B.3: **Summary of local results (without applying statistics).**Considering all subject and region combinations, the left-hand number in each cell gives the total occurrences of the state on the left scoring higher, and the right-hand number gives the total occurrences of the state on the right scoring higher. ACE, LZc and SCE were computed across 4 channels (thus for these measures there were 27 subject and region combinations in total, see Fig. 7). For SCE12 and SCE13 there were respectively 129 and 143 combinations of subject, target region and seed region, and for CS there were 139 combinations. LZs, delta power and gamma power were computed for each single channel in each subject (247 channels in total, resulting in 50 region-subject combinations when averaging per subject across channels in the same region).

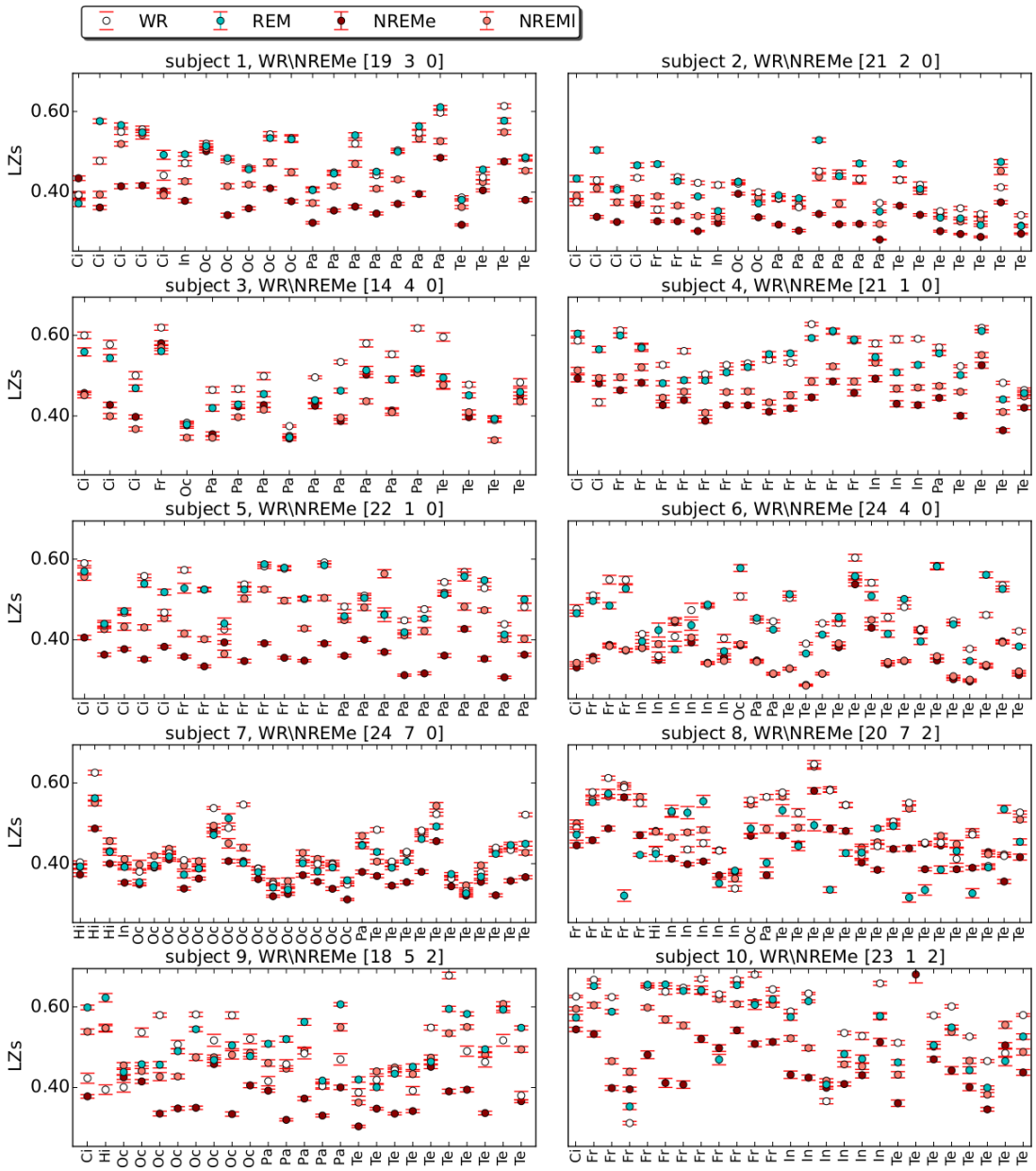

Figure B.4: **Lempel Ziv signal diversity of single channels (LZs) for each subject and state.** States shown are WR, REM, NREMl and NREMe. LZs was computed across all 18-31 channels per subject, the channel's region is indicated as an x-tick label (first 2 letters of the anatomical regions as listed in Fig. 4.2). Error bars indicate standard error across 10s segments. All 10 panels share the same y-axis. The numbers in the triplet in the title of each panel indicate respectively how many channels scored higher (with Cohen's $d > 0.8$) for WR than NREMe (left digit), non-substantially different (Cohen's $d < 0.8$, middle digit) or lower (with Cohen's $d > 0.8$, right digit). For the clear majority of channels, LZs scored substantially higher for WR as compared to NREMe. The few outliers are not consistently in one region, showing that there is no local effect against the global trend (unlike for delta power, compare with Fig. B.5).

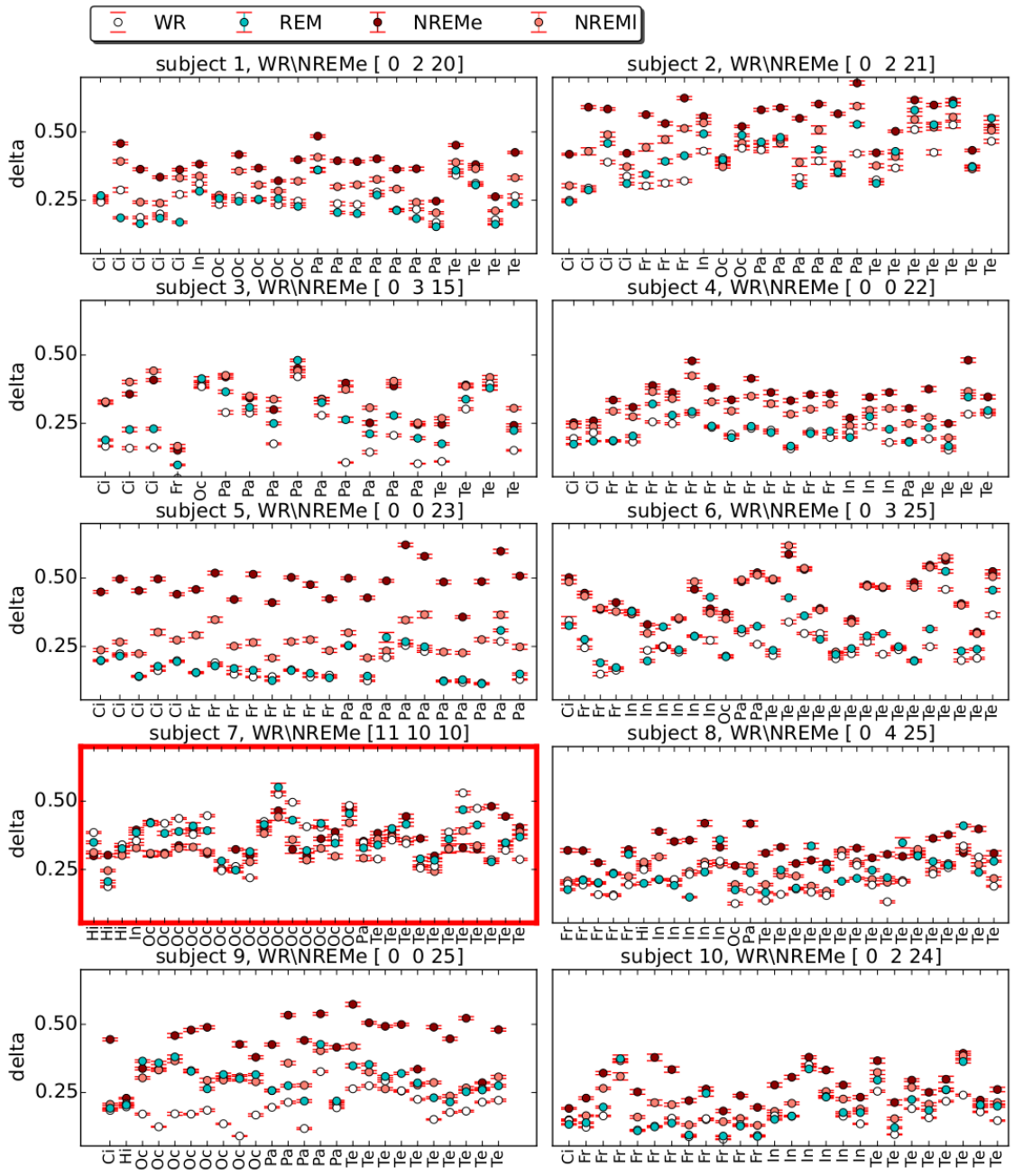

Figure B.5: **Normalised delta power of single channels for each subject and state.** States shown are WR, REM, NREMl and NREMe. Delta power was computed across all 18-31 channels per subject, the channel's region is indicated as an x-tick label (first 2 letters of the anatomical regions as listed in Fig. 4.2). Error bars indicate standard error across 10s segments. All 10 panels share the same y-axis. The numbers in the triplet in the title of each panel indicate respectively how many channels scored higher (with Cohen's $d > 0.8$) for WR than NREMe (left digit), non-substantially different (Cohen's $d < 0.8$, middle digit) or lower (with Cohen's $d > 0.8$, right digit). For all regions in all subjects (except occipital lobe in subject 7 - 10 channels, panel in red frame), delta power is lower for WR as compared to NREMe.

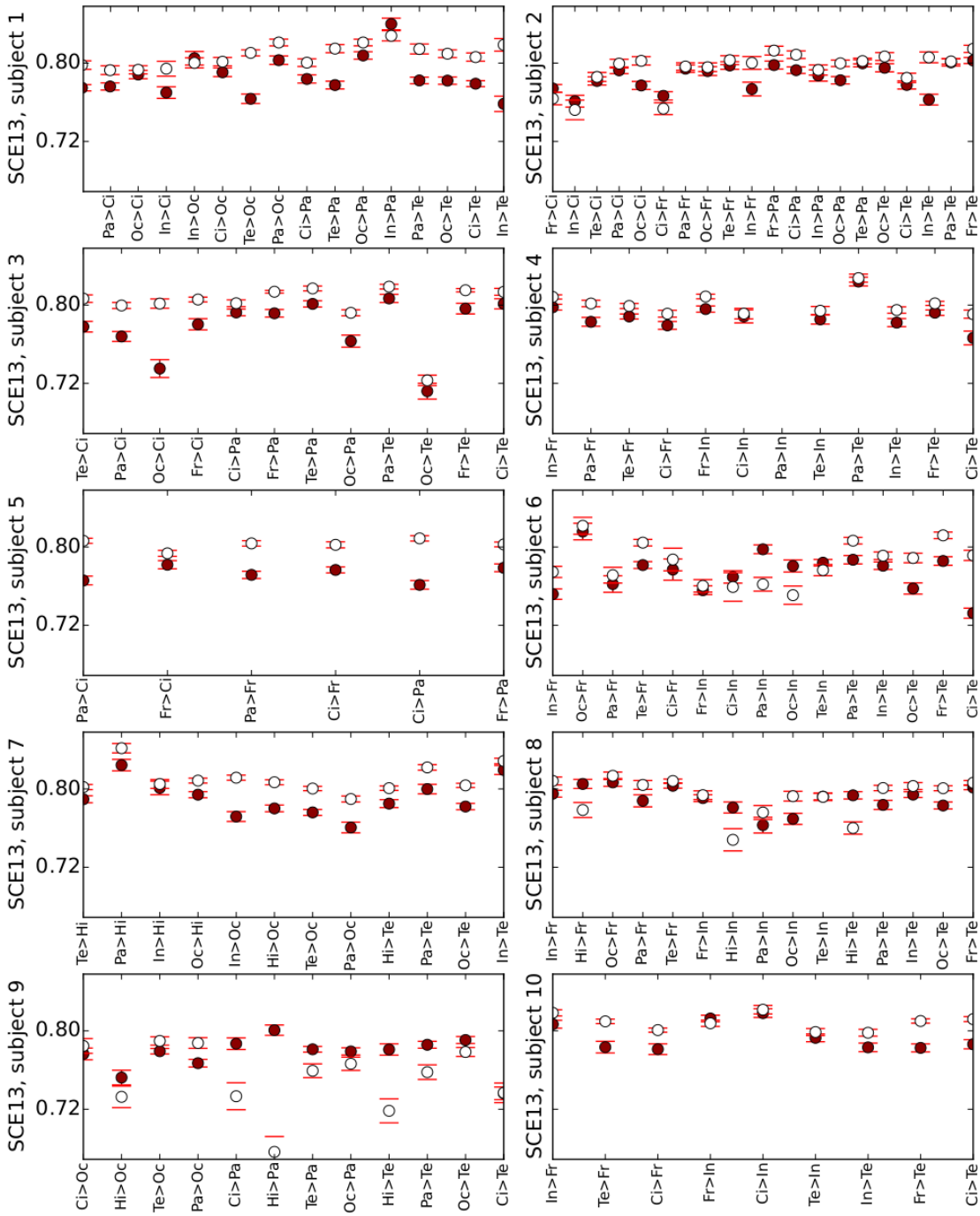

Figure B.6: **SCE13 computed for each possible combination of seed and target regions for each subject.** The labels along the x-axis give the pairs of anatomical regions, abbreviated as in the previous figures, with e.g. Oc>Fr denoting that the seed is located in the occipital lobe and the targets are located in frontal lobe.

B.5 Spectral profile preserving shuffling

To confirm that changes in ACE, SCE and LZc scores between states were due to signal changes beyond changes in power spectrum, we applied a spectral-profile preserving

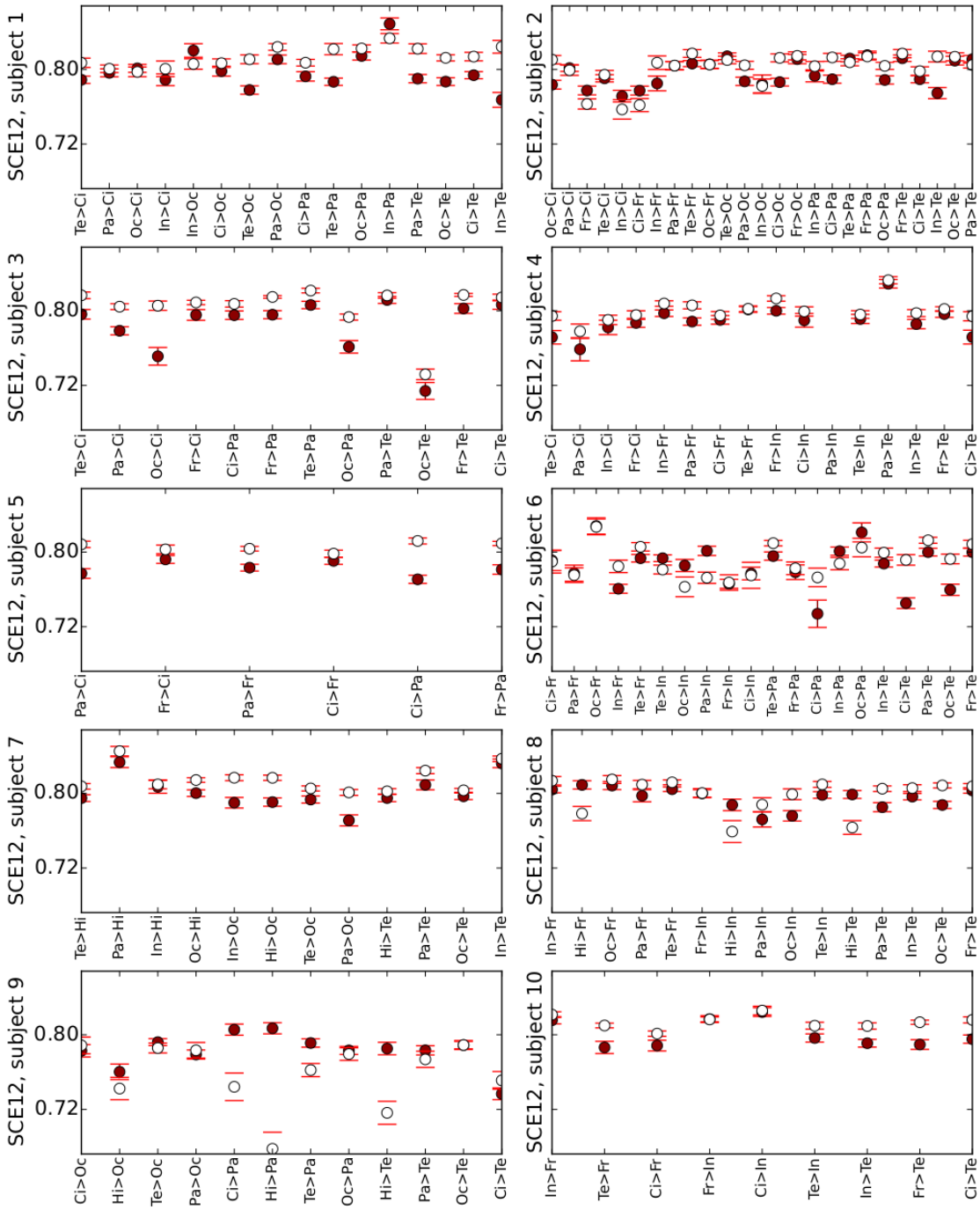

Figure B.7: **SCE12 computed for each possible combination of seed and target regions for each subject.** The labels along the x-axis give the pairs of anatomical regions, abbreviated as in Fig. B.6.

shuffling to the data and used the measures' scores for them as a normalisation for the non-normalised scores of the measures obtained from the un-shuffled data. Our study of the same signal diversity measures ACE, SCE and LZc applied to EEG recordings of propofol-sedated subjects contains a detailed description of this phase-shuffling control [191]. Fig. B.9 compares the behaviour of ACE, SCE and LZc for normalisation by phase-

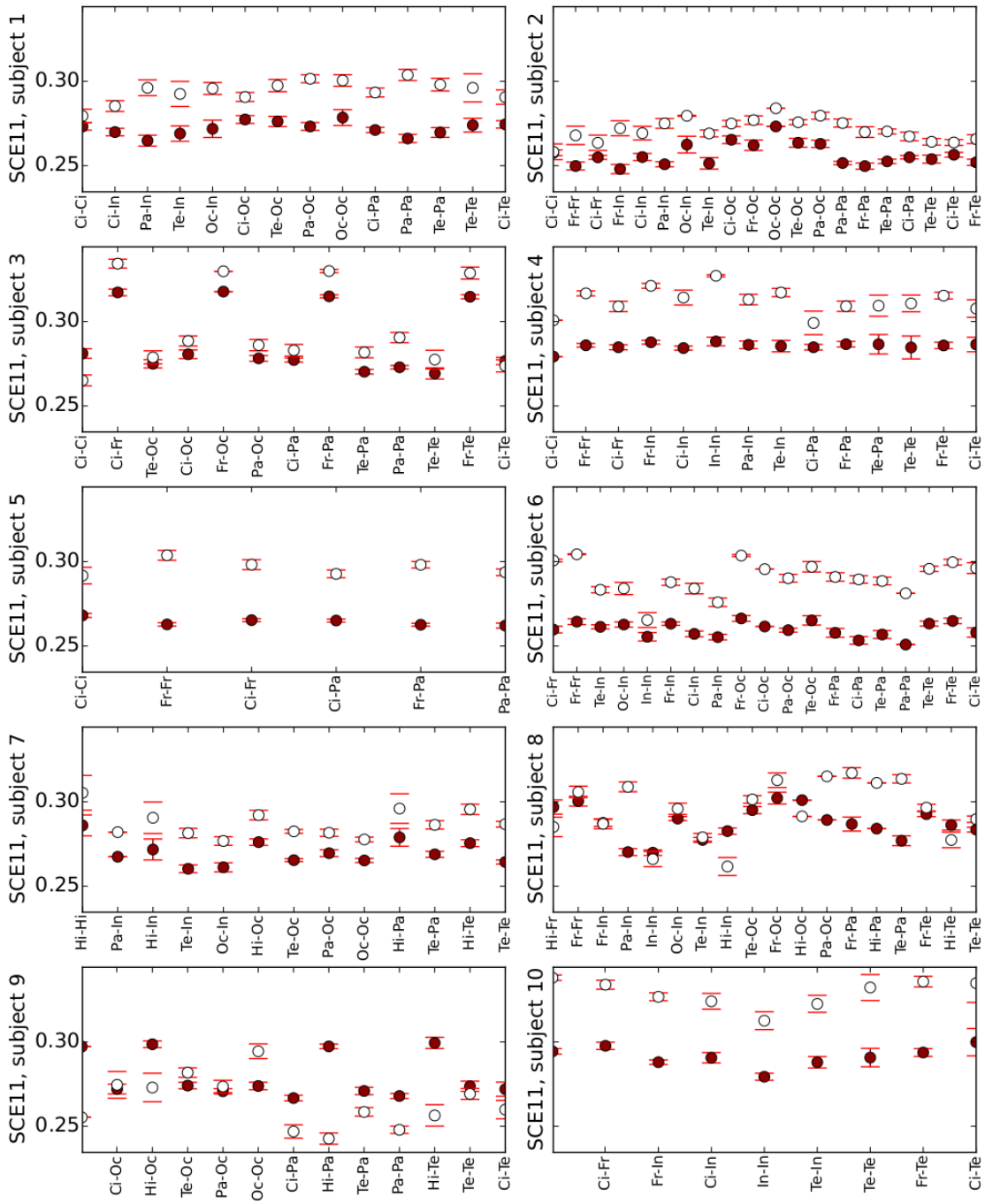

Figure B.8: **CS computed for each non-ordered pair of regions (including reflexive pairs) for each subject.** The labels along the x-axis give the pairs of anatomical regions, abbreviated as in previous figures.

shuffled data (indicated by underscore N in the measures' names) and for our standard normalisation by time-shuffled data. It can be seen that the decrease in the measures during NREM sleep are only partially attenuated by applying the phase-shuffling normalisation, and thus that this can be attributed to changes in signal properties that go beyond changes to the power spectrum.

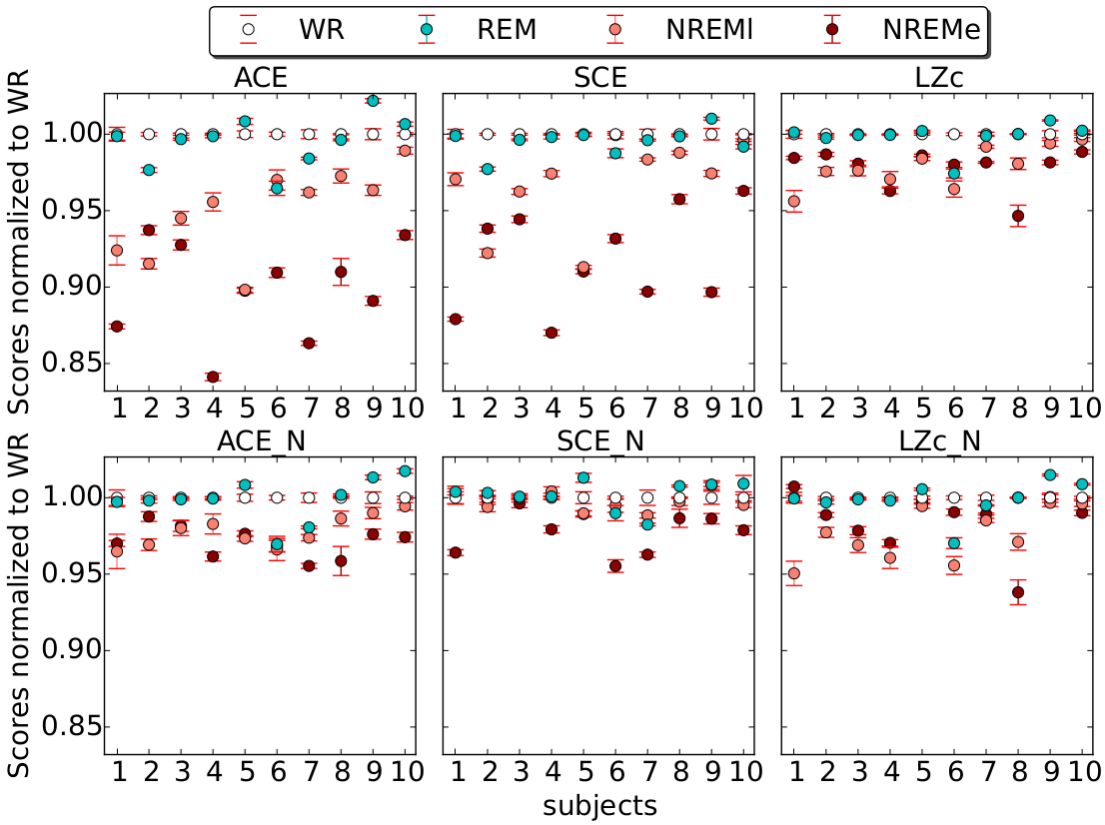

Figure B.9: **Control for changes in spectral power profile.** (Top) ACE, SCE and LZc computed across 18 channels, normalized using their scores for time-shuffled data, as in Fig. 4.4. (Bottom) ACE_N, SCE_N and LZc_N, which are the same measures but normalized using their scores for phase-shuffled data. The decrease in the measures during NREM sleep are only partially attenuated by applying the phase-shuffling normalisation, and thus can be attributed to changes in signal properties that go beyond changes to the power spectrum.

B.6 Influence of bi-polar referencing

When omitting bi-polar referencing as a pre-processing step (i.e. maintaining a common reference, thus obtaining mono-polar data, see Methods) and applying the three measures otherwise identically, the measures behave very similar. Fig. B.10 compares the measures' scores computed on data where bipolar-referencing was applied (identical to Fig. 4.4) with scores obtained when omitting the bi-polar re-referencing step. This shows that bi-polar referencing as a pre-processing step of the data is not crucial for the detection of signal diversity changes between WR and NREM sleep.

B.7 Different sampling rates under fixed input dimensions

Prior to computing ACE, SCE and LZc for the same data set as in the global analysis (Fig. 4.4), we varied the time-series' sampling rate while segment length and channel

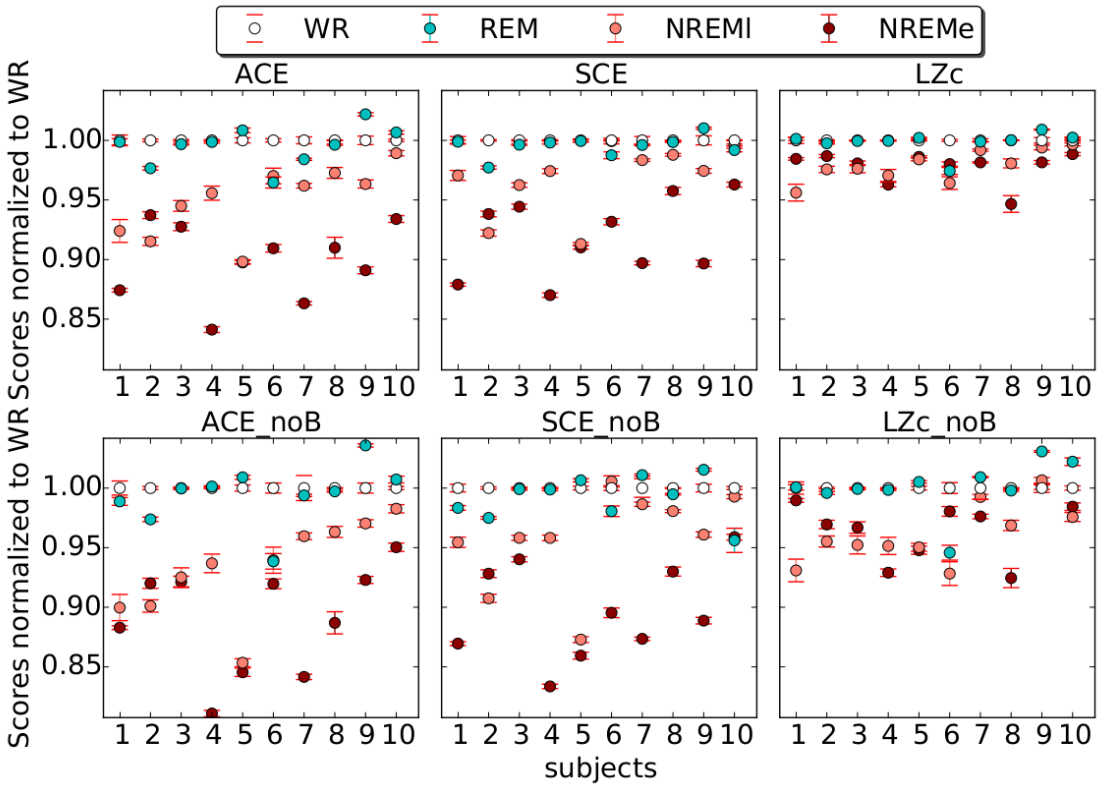

Figure B.10: **Comparison of bipolar and monopolar montage.** ACE, SCE and LZc computed across 18 channels using (top) the bipolar montage used for the main analyses; (bottom) monopolar referencing (same data omitting the bipolar re-referencing). Plotted points show mean across 10s segments, and error bars show standard error.

number were fixed at 2500 observations and 18 channels, respectively. (Note that we analysed the measures' behaviour for different segment lengths as well as different numbers of channels in our previous study [191], finding similar results for a range of different segment lengths (0.2-20s at 250Hz) and channel numbers (5-100). See next subsection for even shorter segment lengths.) We computed ACE, SCE and LZc for data sampled at 10, 50, 150, 350, 500, 750 and 1000Hz, as indicated in the title of each panel column in Fig. B.11. All three measures scored for all subjects higher for WR than NREMe, for a sampling rate at least 150Hz. Consistency was clearly weaker for 50Hz and the results seemed completely random at 10Hz. Average scores across subjects for either state and each measure decreased consistently with increasing sampling rate (not visible in Fig. B.11 due to normalisation by WR score), indicating that in total across the segment length there is less temporally diverse activity in a 2.5s segment (2500 observations at 1000Hz) as opposed to a 17s segment (2500 observations at 150Hz).

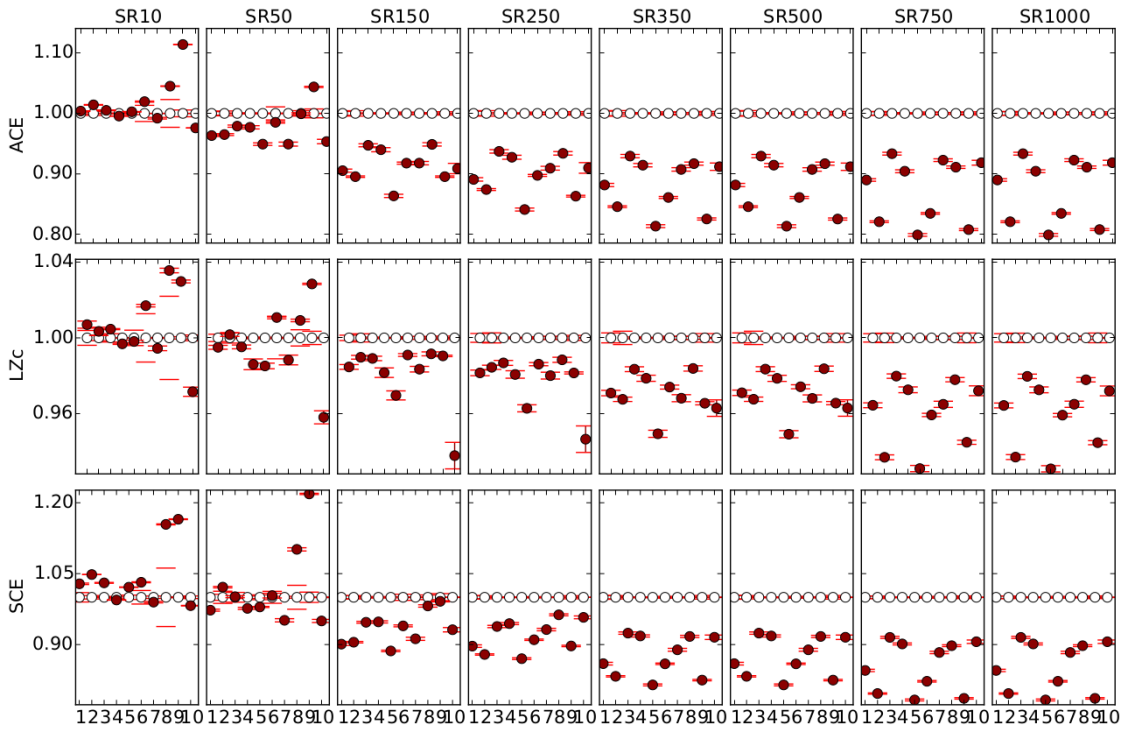

Figure B.11: **ACE, SCE and LZc for 18 channels and 2500 observations for different sampling rates**. Each column of panels displays the three measures' scores for all subjects and the sampling rate indicated in the column's title ($SR = 10Hz, 50Hz, 150Hz, 250Hz, 350Hz, 500Hz, 750Hz, 1000Hz$). Each measure's score is higher for WR than NREMe for every subject and sampling rate of at least $150Hz$. For sampling rate smaller or equal to $50Hz$ consistency of the scores decreases strongly for all 3 measures. Scores are normalized by the result for WR.

B.8 Different number of observations under fixed sampling rate

When searching for the lower bound of segment length, for which the measure's behaviour was still consistent (for 18 channels at $250Hz$), we found that for 500 observations (corresponding to $2s$ at $250Hz$) per analysed segment still all measures scored for all subjects higher in WR than NREMe, as was found for 2500 observations ($10s$) for the main analysis. For 100, 50 and 30 observations per segment there were at most two subjects showing higher scores for NREMe than WR for each measure. At segments that were 15 observations long, results were completely inconsistent for all three measures.

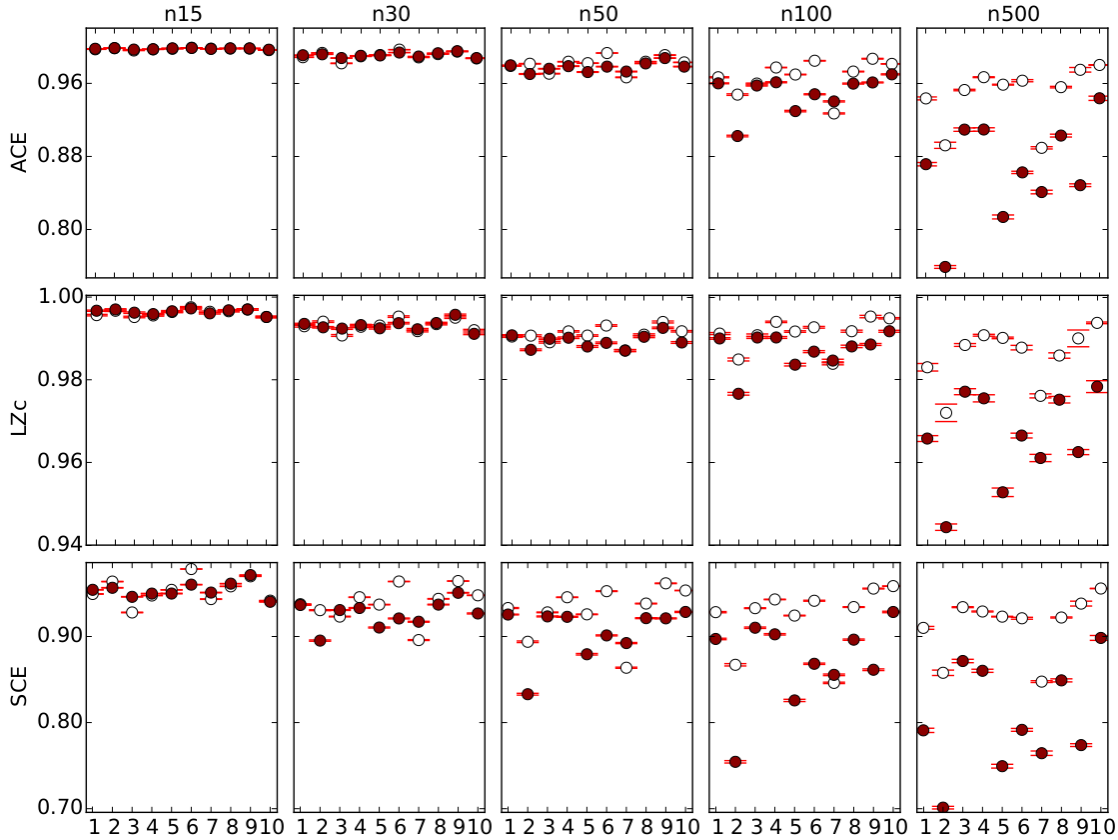

Figure B.12: **ACE, SCE and LZc for 18 channels for different numbers of observations at** $250Hz$**.** Each column of panels displays the three measures' scores for all subjects and the number of observations indicated in the column's title. For segment lengths below $2s$ (corresponding to 500 observations at sampling rate $250Hz$), the measures' scores are no longer consistently higher for WR (white) than NREMe (red). Scores are normalized by the result for WR.

Appendix C

Supplementary figures for MEG/psychedelics analysis

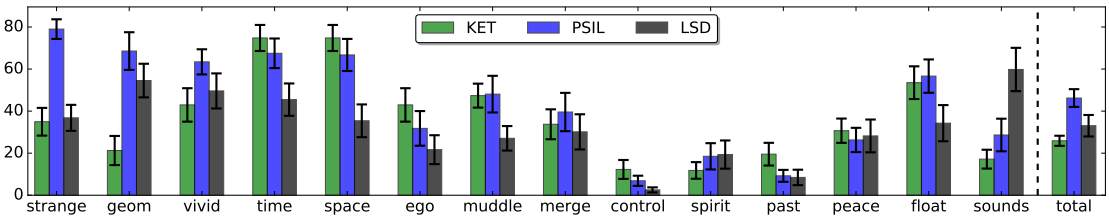

Figure C.1: **Subjective ratings** The changes in subjective ratings under each drug condition are shown as averages with standard error bars across subjects. The "total" of all changes across these 14 questions is shown normalised by factor 20 in order to fit the scale. No consistent differences are apparent across drugs.

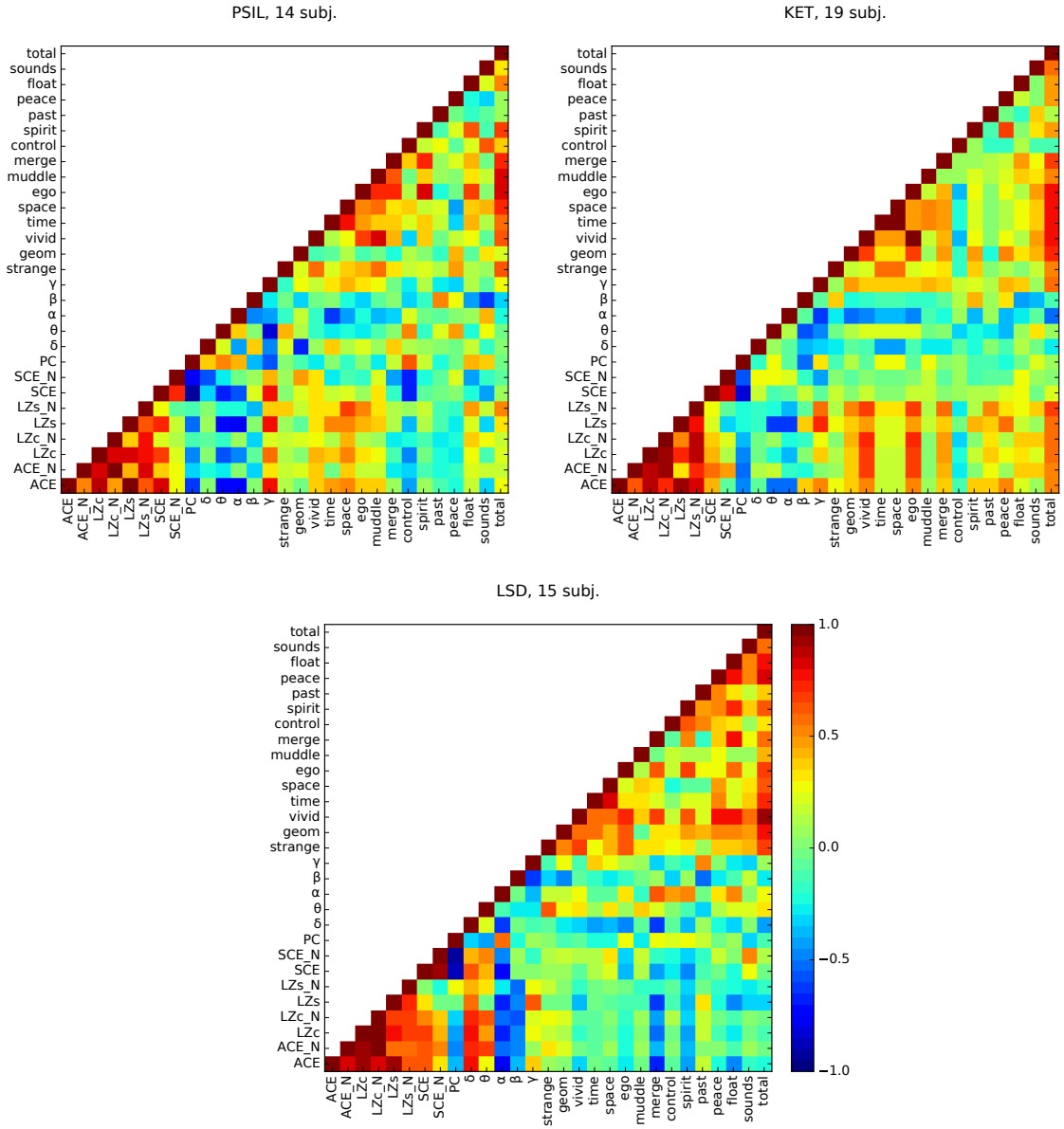

Figure C.2: **Full correlations across measures and questionnaire answers** Here are the correlation results of Fig. 5.3 in full. For each drug, a matrix indicates in colour the Pearson correlation, r, of the score difference between drug and placebo condition (averaged across trials) of any measure (subjective rating) pair across subjects. See main text for details.