Delta rhythms as a substrate for holographic processing in sleep and wakefulness

Henrik Daniel Kjeldsen

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Newcastle University
Faculty of Medical Sciences
Institute of Neuroscience
Systems Neuroscience: From Networks to Behaviour

September 2013

Abstract

We initially considered the theoretical properties and benefits of so-called holographic processing in a specific type of computational problem implied by the theories of synaptic rescaling processes in the biological wake-sleep cycle.
This raised two fundamental questions that we attempted to answer by an experimental *in vitro* electrophysiological approach. We developed a comprehensive experimental paradigm based on a pharmacological model of the wake-sleep-associated delta rhythm measured with a Utah micro-electrode array at the interface between primary and associational areas in the rodent neocortex.

We first verified that our *in vitro* delta rhythm model possessed two key features found in both *in vivo* rodent and human studies of synaptic rescaling processes in sleep: The first property being that prior local synaptic potentiation in wake leads to increased local delta power in subsequent sleep. The second property is the reactivation in sleep of neural firing patterns observed prior to sleep. By reproducing these findings we confirmed that our model is arguably an adequate medium for further study of the putative sleep-related synaptic rescaling process. In addition we found important differences between neural units that reactivated or deactivated during delta; these were differences in cell types based on unit spike shapes, in prior firing rates and in prior spike-train-to-local-field-potential coherence. Taken together these results suggested a mechanistic chain of explanation of the two observed properties, and set the neurobiological framework for further, more computationally driven analysis.

Using the above experimental and theoretical substrate we developed a new method of analysis of micro-electrode array data. The method is a generalization to the electromagnetic case of a well-known technique for processing acoustic microphone array data. This allowed calculation of: The instantaneous spatial energy flow and dissipation in the neocortical areas under the array; The spatial energy source density in analogy to well-known current source density analysis.

We then refocused our investigation on the two theoretical questions that we hoped to achieve experimental answers for: Whether the state of the neocortex during a delta

rhythm could be described by ergodic statistics, which we determined by analyzing the spectral properties of energy dissipation as a signature of the state of the dynamical system; A more explorative approach prompting an investigation of the spatiotemporal interactions across and along neocortical layers and areas during a delta rhythm, as implied by energy flow patterns.

We found that the *in vitro* rodent neocortex does not conform to ergodic statistics during a pharmacologically driven delta or gamma rhythm.
We also found a delta period locked pattern of energy flow across and along layers and areas, which doubled the processing cycle relative to the fundamental delta rhythm, tentatively suggesting a reciprocal, two-stage information processing hierarchy similar to a stochastic Helmholtz machine with a wake-sleep training algorithm. Further, the complex valued energy flow might suggest an improvement to the Helmholtz machine concept by generalizing the complex valued weights of the stochastic network to higher dimensional multi-vectors of a geometric algebra with a metric particularity suited for holographic processes.

Finally, preliminary attempts were made to implement and characterize the above network dynamics *in silico*. We found that a qubit valued network does not allow fully holographic processes, but tentatively suggest that an ebit valued network may display two key properties of general holographic processing.

Acknowledgements

I would like to thank my supervisors: Miles Whittington, Marcus Kaiser, Jennifer Simonotto and Mark Cunningham. Their patience, open-mindedness and willingness to allow me to pursue my own course, balanced with timely guidance, input and supervision has made this time particularly enjoyable and well-spend.

I would also like to thank all the remaining members of the respective labs, without whose help and support on a daily basis, I would simply not have been able to carry the physical workload.

I thank my family and friends, especially Lars Bredvig who over the years has put in thousands of hours discussing the fundamental issues forming the starting point for this work.

Publications

Kjeldsen, H. D., Traub, R. D., Whittington, M. A. (2012). “A neocortical delta frequency generator modulates layer2/3 sparse spiking via nested theta rhythms.” Society for Neuroscience.

Carracedo, L. M., Kjeldsen, H. D., Cunnington, L.,Jenkins, A.,Schofield, I.,Cunningham, M. O., Davies, C. H., Traub, R. D., Whittington, M. A. (2013). "A Neocortical Delta Rhythm Facilitates Reciprocal Interlaminar Interactions via Nested Theta Rhythms." The Journal of Neuroscience **33**(26): 10750-10761.

Table of Contents

List of Figures

Chapter 1. Introduction

defining the problem and scope of investigation

The setting for the present thesis is the Wellcome Trust funded 4-year MRes/PhD programme, 'Systems Neuroscience: From Networks to Behaviour' at the Institute of Neuroscience within Newcastle University. A major purpose of this programme is to bring in new talent from technical fields into the biomedical field of neuroscience. This goes some way in explaining the rather unusual focus for a biomedical thesis, which is decidedly technological rather than medical. This means that we have sought primarily a specific technological innovation, using quantifiable biomedical phenomena to refine and constrain (to some extent) this innovation. While the resulting novel approach to analysis and interpretation of neurodynamic data has a great deal of scope for application in the biomedical field, insufficient time was available to exploit this potential here. The reason for seeking this technological breakthrough in a neuroscientific setting is specifically that it seems to be a form of technology, which Nature and evolution has figured out and implemented in the brain already. What this technology is, and why it appears to be critically important if we wish to truly understand brain function, is the subject of this chapter. It is hoped it will provide explanation as to why we thus wished to reverse-engineer this unknown solution.

1.1 The issue of intractable problems

Of the many important problems in science that one can choose to study, few carry the practical significance of that of 'intractability'. It is a sort of meta-problem in the sense that it concerns which other types of problems we can hope to solve with the application of computers. The use of computers has undoubtedly been of great benefit in many fields, and it has even been suggested that once a field adopts computers as a fundamental tool, once a field becomes computerized, so to say, then the accelerating increase in computer speed, known as Moore's law, empowers a similar law of accelerated advance in the given field (Kurzweil 2005). However, the problem of intractability drops like a spanner into the machinery of this logic: The fundamental misconception lies in the notion that a problem can be intractable at the moment, because we lack the necessary computer speed, but will become tractable in the future as enough computational speed becomes available according to Moore's law. This is not

necessarily wrong, but it overlooks the real issue and clouds the proper understanding of intractability: There is a very large group of important problems that for any problem size of practical importance are truly intractable (see e.g. Aaronson 2004b). This means that even if we can write down a sequence of steps (an algorithm) necessary to solve the problem, we can never finish running through those steps no matter how fast we try. The number of steps simply grows exponentially as a function of problem size, and any relevant size is big enough that the resulting number of steps is effectively infinite. This is the true meaning of intractability; that the speed of computation, current or future, is irrelevant.

That being said, there is however a further point of great importance that needs to be made clear: Intractable does not necessarily mean impossible; it only means impossible for a computer following an algorithm. Further to that, we should specify that a computer does not just mean a device which carries out an algorithm; it also means that the environment or universe in which this device operates follows certain rules, for example, the above statement "no matter how fast we try", would implicitly assume that nothing in the device could move faster than light, and possible parallel devices could not communicate faster than light.

Hope of achieving tractability runs short when one realizes that these rules, our modern physical theories that describe our environment in extreme detail, are themselves algorithms or equivalently formal mathematical systems. It would seem to imply that nothing can exist that is not equivalent to a formal system or an algorithm on a computer, thus we can never hope to construct a tractable computing device within these laws. This foreboding idea is known as the physical Church-Turing thesis (Piccinini 2007), but it is not universally accepted, and it is clear that neither classical nor quantum theory is really complete (Einstein, Podolsky et al. 1935), i.e. they cannot necessarily describe everything that can exist. Further, the same proofs that originally put limits on the power of formal systems, the proofs that made us aware of the problem of intractability in the first place, did so by contrasting formal systems with something that is clearly beyond them, and clearly exists: The brain and the mind. We will give a brief and informal account of these proofs and their historical context in the next subsection, but it should be noted that we will not be trying to introduce or review the field of complexity theory, whose domain we are trespassing, especially because we

suspect that the point of view normally taken in that field is not particularly helpful to our own endeavor.

1.1.1 *Interpreting Gödel's theorems*

In a golden period of mathematics, between the world wars, there was great optimism that everything in Nature could finally be described and understood under one grand formal mathematical system. This pursuit was the dominant programme of mathematics at the time, known as Hilbert's programme. Nevertheless, it was all stopped dead in its tracks before it really got rolling as Kurt Gödel introduced a surprising sort of logic of paradox that proved that it could not be done (Gödel 1931). Simply put, Gödel said, "There is no proof of this statement". Indeed there is no formal proof; its truth cannot be shown by any formal mathematical method or by any amount of computation, but *we* can do something else, *we* can simply *see* that it is true; by the might of mental insight. It is recommended at this point to take a minute to verify this in your own mind. It is a beautifully simple point, which demonstrates that the brain is not a computer and cognition not simply a collection of applied algorithms. Once this point of view sets in, it becomes self-reinforcing and eventually self-evident. Much like free will is self-evident to those who think they have it: Not because no evidence is needed, but because the evidence is all around. There is nothing mystical about self-evidence, and it is of course at the very heart of science, because self-evidence is the basis of all foundational axioms from which all scientific reasoning derives.

Intractable problems are in fact solved right and left when one learns to recognize them. Most of what we do in our daily lives is utterly intractable; this is the reason why artificial intelligence has had so little general success (Dreyfus 2007). For example, when IBM announced that they could simulate an entire cat brain (IBM 2009), a neuroscientist might have claimed that they had not included enough neurobiological detail, but that debate is mute in a sense, because there is no theoretical foundation telling us which level of detail is required for cat cognition. On the other hand, we can simply ask if that IBM cat can catch a mouse, and realize that the answer is no, because catching a mouse is an intractable problem. The same essentially goes for the billion euro EU project aimed at simulating an entire human brain, the human brain project, which might be hopeless in the same sense, because from our understanding of intractability the brain cannot ever be computationally simulated no matter the detail.

There is of course hope that in the process we might somehow discover how the brain achieves tractability, but the idea that simply simulating every thinkable detail will reproduce a mind is presently false, because that simulation will never be able to run, unless something fundamental changes in our understanding of computation and simulation. There have been estimates of the computational power needed to run the human brain simulation, and of the power required to power and cool the needed computers, in the order of one full nuclear power plant (Kogge 2011; Furber 2012), which clearly shows that something is fundamentally wrong when we are trying to simulate a device, the brain, that does not use enough power to heat a cup of tea.

A few course corrections might be in order at this point: The mention of free will and the comparison with its self-evidence might lead the reader to think that we intend to debate personal philosophy rather than hard science, but that is not the case. In fact, the idea that the question of free will is outside the scope of science is an unfortunate result of one of a few foundational axioms of science, and we already know from quantum Bell's theorem-type experiments (about quantum entanglement) that we must give up at least one of these foundational axioms in order to explain the experimental outcomes (Werbos 2008), and we also know that the current consensus, known as the Copenhagen Interpretation, which gives up the Axiom of Reality (objective reality), is not workable (Hestenes 2009). It is indeed a paradoxical situation that the consensus is that the consensus is wrong, but that is nevertheless the state of affairs in physical science; even a Dane can see that there is something rotten in the Kingdom of Denmark (paraphrasing David Hestenes' comment on the Copenhagen Interpretation).

The question of free will ties in with the validity of the Axiom of Causality, so is not a matter of mere philosophical taste, but rather a perfectly valid scientific question that can and must be both argued and settled scientifically, because it is of essential importance to all science through the foundational axioms that are supposed to be self-evident, but are obviously overdue for re-evaluation due to Bell's theorem-type experiments. In the same way, as we will see in more detail later, the foundational axioms of science have direct relations to the issue of intractability through the definition of the environment in which computation takes places, and which defines the limits of computation. It goes both ways though; the power of computations that we can observe experimentally make requirements on the rules of the environment, and can potentially help decide which axiom(s) must be abandoned, affirmed or reinstated. This

is the position taken here; *that the computational power of the brain to solve intractable problems tells us something essential about the physical environment in which it exists.*

The commonly accepted idea that all interpretations of quantum theory are equivalent and only a matter of personal philosophy, because they lead to the same experimental outcomes, is arguably wrong, because different interpretations give up different axioms, and, for instance, the existence or not of an objective reality is obviously not without consequence, nor is the existence of free will. Of similar importance is the way these axioms of science implicitly define some of our terminology, and potentially limit or bias what can be scientifically stated without paradox or other epistemic problems. Axioms must be unconditionally true or they blind our formal systems of logic, and in extension us, to the truth; the fact that we already know that we have an unstable set of foundational axioms should encourage us to explore all options carefully.

Gödel's theorems are often misunderstood exactly because they unsettle basic axioms and fundamental assumptions, but admittedly there is to some extent room and need for interpretation; the account given here tries to follow the comprehensive interpretation given by Roger Penrose (Penrose 1989; Penrose 1994), but we do not know whether he agrees that it does. For reasons that will become clear later we will however depart from his use of the term non-computation, to describe processes that solve intractable problems, and instead keep the door open that computation might achieve the same thing if the rules of the environment of the computing device are different than assumed by Penrose's conventional definition of computation. This change in terminology might be unwise, but sometimes an over-established terminology can hold back understanding of new perspectives or simply overcomplicate things, and we suspect that this might be the case here.

1.1.2 *The Gödel letter and the modern Millennium Prize problem of P vs. NP*

Although Gödel's theorems changed mathematics and science in fundamental ways, it seems Gödel himself might eventually have had a change of heart, perhaps by the same sort of considerations of the brain and mind as above: In a letter to John von Neumann (after whom the modern computer architecture is rightly named) in 1956 (see Lipton 2013), 25 years after his theorems, he proposed that there might be a sort of loop-hole in his own theorems' limits on computation. We do not know what his detailed reasoning

was, but he likely knew that all physical theories were and are subject to his theorems, and we can reasonably speculate that he considered the brain to be physical, thus leading to the paradox (and paradox was his speciality): How can the physical brain, thus being equivalent to a computer, produce a mind capable of surpassing all computers. The letter is generally considered the birth of the field of complexity theory (Lipton 2013), but our reading departs on the suggestion of a loop-hole, which complexity theory has since rejected. It is also sometimes considered a gesture to von Neumann who was dying, and who would have appreciated staying in the game, so to say, till the very end. However, our interpretation suggests that it was no mere gesture, but rather a genuine request for help and support from an old hero. This is not clear from the letter alone, but as we will see later Gödel had already found a physical basis for the suspected loop-hole, although he never to our knowledge made the connection in public (and we can therefore only speculate that he did in private) we know that he was sorely disappointed in the reception of this basis as first presented without a connection to his theorems (Yourgrau 2006), so before presenting any such connection (which he never did though) he would have needed to gather considerable support from other giants of the field. It is also worth mentioning here that when Gödel was first presenting parts of the results that would soon form his famous theorems, at a conference discussion, von Neumann was the only one to grasp the significance, and he even produced his own proof, which he revealed to Gödel who had just gone to print with his own now famous version. This must have been an enormous relief; that someone at the top of mathematics understood what he had achieved.

Gödel continued to gather evidence for his loop-hole basis for the rest of his life (Yourgrau 2006), but his wider interpretation of what it meant had the consequence that it was not widely accepted or even understood. This failure to assert an important possibility in physics has likely contributed to the precarious situation we have today. From a letter of genuine inquest has grown a completely one-sided discourse where there is no longer any real question, only a matter of finding a final proof of an already settled truth. The consensus for the last half century since Gödel has indeed been that there is no loop-hole, that tractability is impossible, and that the brain therefore does not solve intractable problems. At the same time we know that certain problems that the brain certainly solves are intractable, but instead of challenging our preconceptions we try to explain away the paradox; we will return to the details of this in the next section. Science has tried very hard to prove that the consensus is correct, but all we have

succeeded in proving is that it is much harder to prove than initially thought (Lance 2009). In fact, it is the most sought after proof at the top of the list of Millennium Prize problems each worth a million dollars. The prize is given for a proof that P ≠ NP, which is shorthand for the consensus that a tractable device cannot be build (this is not formally very accurate; there are many subtle details, but it is the spirit of it). The opposite possibility, that NP = P, is not mentioned, and is generally considered ridiculous, because, well, it would be too good to be true, or absurd to use the technical term. How many times has the absurd turned out to be the truth in the history of science (Feyerabend 2010; Kuhn and Hacking 2012)? In this thesis we will promote that neither P ≠ NP nor NP = P is the case, and that something more like NP ≈ P is the real situation. The difficulty of the current situation is then due to a fundamental misstep in raising the P versus NP issue in the first place, in assuming that the question makes sense, and that the distinction must be real, while in reality it could just as well be an artifact of a forceful introduction of a limited notion of computation in an environment defined, as a basic axiom of science, to only allow a limited form of computation.

1.2 Does the brain somehow solve intractable problems?

It is clearly of fundamental importance to the purpose of this thesis whether the brain really solves intractable problems or not. The counter-consensual hypothesis made at the very onset is that it does indeed solve intractable problems; not as part of any conscious problem solving ability, but rather intrinsically in basic subconscious processes at the neuronal level. If it were to turn out that the brain does in fact not solve intractable problems, there might still be some value to this study in the sense that historically we have learned important lessons from considering things that eventually or even immediately turn out to be wrong; but it must be made clear that we can assume it to be true with just as much justification as the consensus assumes it to be wrong. We simply do not know either way. That there even *is* a consensus is out of old habit and not based on real evidence (Aaronson 2005a), and the neglect of consideration to the alternatives is quite irrational and of course against scientific principles. That being said, let us briefly consider the evidence in our favour; the only 'evidence' against, that we know of, being the lack of hard proof and acceptance of our position, and the notion that it would simply be too good to be true, i.e. absurd.

1.2.1 *Gödel's theorems imply that the brain solves intractable problems*

As we have already covered, Gödel's theorems themselves imply and are based on the distinction between mind and machine, with a further distinction between formal proof and informal proof by mental insight. The informal proof that there is no formal proof is indeed by insight as the reader should have experienced firsthand in the previous section. However, before we go on to the more concrete evidence there is another issue that is often mentioned as a way around it, so it might be worth covering it first: Considering that humans are fallible it is easy to think that perhaps brains do not really solve intractable problems, but rather approximate solutions; that is, that there is a difference in difficulty of exact and approximate solutions of intractable problems. This is however in our opinion unlikely to be the case. The central issue is of course the question of when an approximation is good enough; if the reader is already familiar with the coarse graining of phase space when calculating entropy we might adopt a similar view that the approximation and exact solution must look macroscopically the same, but even this criterion is not that well defined anyway (Baranger 2001). Let us instead consider a simpler analogy: The original problem of intractability is due to the (exponential) size of the search space, let us represent this space by the space of our Milky Way galaxy. The exact solution might then be represented by Earth; the relative scales might even be fairly realistic for smaller problems. Now let us consider the size of the space of the approximate solutions; if we make this too big then we must conclude that there was no real problem in the first place since an approximate solution is easy to find in this case. We could make it the size of our solar system, enormously bigger than Earth, but still enormously difficult to find in the entirety of the galaxy. In this picture we can see that an approximate solution, the solar system instead of Earth, although much much bigger in search space is not much easier to find in the scale of the entire search space of the galaxy. So the tentative conclusion is that even if the brain is operating with approximate solutions it does not make the intractable problems so much easier as to be suddenly tractable.

The field of complexity theory usually takes the approximation approach, and hundreds of different schemes are known, none of which provide general approximate solutions to all or most intractable problems in practical settings. One of the more popular approaches, simulated annealing (Kirkpatrick, Gelatt et al. 1983), should be mentioned here because it relates to ideas that we will consider later, and it will be important to understand how what we are proposing is different from this idea. At first sight or going

by the name alone it seems that this approach is based on a known physical phenomenon, annealing, which makes it particularly interesting, because it happens more or less naturally. However, the simulation part makes the issue a lot more complicated, because we do not know how to efficiently simulate physical annealing, so different approximation schemes are again used. These schemes are typically the same as used to approximate Bayesian statistics, which we will elaborate on in the next subsection, as the first example of intractable problems, which the brain appears to solve.

Another way to see that perhaps approximations are not the key to information processing in the brain is to consider problems that have only yes/no answers, but are still algorithmically intractable; in this case we can see that an approximate yes or no does not make much sense. It might seem that we could invoke a heuristic along the lines of, if one answer leads to a logical paradox, choose the other answer, but this is something Gödel showed is algorithmically impossible in the details of his proof.

1.2.2 *The Bayesian brain is intractable*

Helmholtz suggested that the brain is a statistical inference machine estimating the probabilities of alternative explanations of sensory input (Helmholtz 1860). This idea has since achieved a deep foundation in Bayesian probability theory (Westheimer 2008), and the resulting idea of a so-called Bayesian brain, is now generally accepted as a theoretical foundation in neuroscience (Jaynes 1988). Further, it is known that Bayesian methods are optimal in the sense that there is no better way of making a given inference. Unfortunately, it is almost universally forgotten or ignored that this is only true in theory; in practice direct Bayesian methods are simply intractable, and in order to use them we must resort to various approximations like Monte Carlo methods and variations, and it seems quite likely that neuroscientists would be more reluctant to accept the Monte Carlo brain. It might be implicitly assumed that the brain has found a tractable way to be Bayesian, but from the previous sections it can be understood that this is not possible within the current framework of what the brain is, a computer. We will consider this in more detail and look at the alternatives later; first we continue with further examples of how we can see that the brain really deals with intractable problems.

1.2.3 *Associative memory is intractable*

Associative or content-addressable memory is about context. The problem with context is that in general there is no way of knowing what is relevant and what is not until both have been considered. This means that everything must be considered. Searching a semantic network for the relevant context is known to be intractable in the simplest version where concepts are represented by individual nodes (Cook 1971). It is however also possible to represent concepts by patterns across large numbers of nodes, where each node can also participate in multiple concepts; this idea was developed by Hinton, who called it distributed representations (Hinton 1990). The concept of distributed representation obviously fits much better with neural networks, both biological and other, but in itself does not provide a means of addressing the network or memory in a tractable way, which is also reflected in Hinton's further conceptualization of *reduced* distributed representations that maintain the computational properties of the full representations, but at a much lower cost at the expense of some uncertainty. We will return to actual attempts at implementing such reduced representations in chapter 3.

1.2.4 *Vision is intractable*

It has been formally shown that general visual search (finding a general object in a general visual scene) is an intractable problem (Tsotsos 2005). This should not be confused with simpler examples of specialized systems that are able to recognize specific objects in specific situations and environments; this can mostly be achieved with sufficient effort, but it does not imply that the system can do it in general, i.e. for general objects in general situations. On the other hand it should also not be confused with more difficult situations that require the system to actually understand what it is recognizing, like situations of complex occlusion, distortion or deformation.

The proof of intractability was given in the context of biological vision, but the authors go on to conclude that the brain does not actually solve it. Rather than conclude that the brain handles intractable visual search they suggest that the problem is reduced to a tractable problem by the process of attention (Tsotsos 2005; Tsotsos and Bruce 2008). This is clearly not reasonable unless attention itself is then an intractable problem, and we would be back to square one. If we could simply reduce intractable problems to tractable ones by a tractable process there would be no intractable problems in the first place. Again, it could be argued that the brain has simply found a way through millennia

of evolution, which we will not disagree with; in fact, this is also our view, but it must then be accepted that this way is not by any of the known theories of attention (Itti, Rees et al. 2005), to the extent that these are formulated in enough detail to be computationally simulated.

1.2.5 *Natural language is at least intractable*

Recognizing and responding to natural language, or simply translating between natural languages in text form is very often misunderstood as something that is on the horizon of artificial intelligence, because much research and attention is focused on the problem, and many products and services promote or even hype the idea. However, it is clear that this cannot be very successful without the actual understanding of what is being said (Kanerva 2010). Understanding is arguably more difficult than single intractable problems; it could seem that a combination of several intractable problems might achieve understanding, but it might as well be that it requires actual consciousness, and that that is even more difficult than a combination of several intractable problems. It might also be that solving intractable problems requires understanding, which is certainly in line with Gödel's theorems, and certainly a property of the brain and mind, but the point here is not to speculate on these issues, but rather to argue that language is another example of a problem that is at least intractable, though routinely solved by the brain.

1.3 How does the brain solve intractable problems?

Having argued that the brain is able to solve intractable problems we will now consider suggestions of how it might actually do so. It is clear that it is extremely difficult to come up with anything that is not equivalent to a formal system, which of course also complicates its description, since a normal logical exposition is a formal system itself, and therefore might not be able to directly capture the essence of the idea. Presumably a description of such an idea must then be given in such a way as to invoke an insight in a manner similar to the informal proof of Gödel's statement encountered above. Ultimately the only way to be really convincing in lieu of formal proof is by demonstration, i.e. by solving a general intractable problem in a recognizable way. This is really also all that is needed in order for the first suggestion (that we will consider below) to be entirely convincing, but unfortunately it has everything but such a demonstration.

1.3.1 *Friston's ergodic suggestion*

Karl Friston has proposed a grand unified theory of the brain, whose apparent explanatory power spans a vast expanse of bioscience (Friston, Kilner et al. 2006; Friston and Kiebel 2009; Feldman and Friston 2010; Friston 2010; Friston, Daunizeau et al. 2010). It is rather difficult to grasp in its mathematical detail, so the following merely reflects our reading of it, with the implied understanding that it might not be ultimately correct. The theory is first and foremost a Bayesian brain theory, but the necessary approximation and subsequent tractability is achieved in a non-standard way known as variational Bayes. It is also a strange attractor theory, the most salient point of which is that the system is indeterministic (Gleick 1997); however as indeterminism is not enough to solve intractable problems (Penrose 1994), we will not dwell further on this aspect of the theory. In any case it is exceedingly difficult to tell the flavours of indeterministic dynamical systems apart (Mandell and Selz 1997).

The foundational idea of Friston's theory is that the goal of brains of biological agents is to minimize entropy, but rather than evaluate this inaccessible or intractable measure directly they minimize another measure, free energy, which is a bound on entropy under certain conditions. These conditions, as far as we understand it, are the conditions of ergodicity, i.e. that the long-term statistics of the system are the same as the short-term statistics of many representative systems. This is unfortunately rather unlikely to hold in our opinion; in general complex systems are not ergodic (Grigolini, Aquino et al. 2009), and the assumption of ergodicity has already failed spectacularly several times in the history of science: Most famously in the failure of statistical physics to predict the spectrum of blackbody background radiation, known as the ultraviolet catastrophe, which led Planck to suggest that energy is quantized, eventually leading to the theory of quantum mechanics. It also failed in what is known as the Fermi-Pasta-Ulam problem (Porter, Zabusky et al. 2009), which sprung from one of the first computer simulations of complex systems (weights connected by springs), and stood unresolved for a decade (partly because it was classified at Los Alamos). The anecdote is that they accidentally left the simulation running overnight, and in the morning instead of ending in the predicted thermalization the simulated system displayed a recurrence of past states, revealing that it was not ergodic and chaotic, but quasi-periodic. The resolution of this problem is also rather spectacular, but unfortunately outside our present scope (see instead e.g. Hoffman and Wayne 2008).

We should of course equally consider Friston's arguments in favour of ergodicity, but to our knowledge there are none; rather the motivation seems to be that if it holds then the brain can make good enough predictions to enable a general predictive coding scheme, which leads to great explanatory powers in a long list of desirable properties from perception and action, over learning and attention, to neurodevelopment and evolution (Friston and Kiebel 2009). This is indeed an enticing proposition, and it is tempting to accept ergodicity to gain a unified explanation. However, it might be essentially equivalent to postulating an oracle in a black box, then showing that the oracle (by being an oracle) can make good enough predictions to enable a predictive coding scheme and everything that follows, and then concluding that the postulate of an oracle in a box must be true based on its extraordinary explanatory powers alone. The present thesis can possibly be accused of the same problem itself, but we will try to be upfront about the issue.

Interestingly, Friston has proposed free-energy robots based on his theory (Friston 2010), and a demonstration on any of a number of intractable robotic or other artificial intelligence problems would indeed be a very welcome (informal) proof of the theory. However, although variational Bayes is already used in artificial intelligence, it is not known as a general solution, which we assume it would be by now if it was.

To be more constructive we can consider how we could experimentally determine whether the brain is an ergodic system; this can potentially be done by considering cortical oscillations as a signature of the dynamical system properties: When the spectrum of oscillations can be described by a power law, we know that certain values of the coefficient of the power law imply that the system under study is ergodic or not (Grigolini, Aquino et al. 2009). Nevertheless, there has been some confusion about exactly what to look at, e.g. local field potentials, neuron firing rates or perhaps a signal corrected for filtering in the tissue (Bedard, Kroger et al. 2006), so it is problematic to consult specific values in the literature. In theory, though, it is not confusing: We are supposed to look at energy dissipation, as far as we understand it. There has unfortunately been no way of directly measuring energy dissipation, so the confusion is understandable. In chapter 6 we will present a new method of measuring energy dissipation *in vitro* with micro-electrode arrays. An important point that we can take away from the literature without looking at specific values is that the values are very different between sleep and wakefulness (Bedard, Kroger et al. 2006), so it could be the

case that the brain is ergodic in a half-cycle. This would perhaps not help Friston's theory, which we think requires ergodicity all the time, but it might be useful for other ideas as we will see later.

Going back to the original goal of minimizing entropy, which we certainly agree with (and which was first noted by Maxwell and debated ever since (Schneider 1991)), we can ask if it is reasonable that entropy can even be bounded by a smooth function at all. Although we can put formulas to entropy it is a measure that cannot be said to be fully understood; it is intimately tied to the problem of the arrow of time (Coveney and Highfield 1992), but a full exposition is beyond our current scope. We simply propose instead that for instance at a moment of eureka, the entropy (change) cannot be bounded, and that these moments of sudden insight are essential parts of our cognition, which we cannot do away with in any model no matter how alluring.

1.3.2 *Penrose's interpretation (but not his suggested solution)*

Roger Penrose places major importance on just the concept of insight as being beyond anything computers can achieve. His view follows from Gödel's proposition of a truth that cannot be formally proven, but is easily understood by mental insight. Penrose's exposition over two long books (Penrose 1989; Penrose 1994) is comprehensive and convincing, but he might have made a crucial strategic error that is only really obvious in hindsight: He proposed a solution. A clear distinction between the well-founded issues he raises, and the very tentative solution he proposes must be made.

It seems that many (neuro)scientists consider firstly his tentative ideas, and in rejecting them they, perhaps inadvertently, reject the major point of his work, which cannot be discarded so lightly. This point is that the brain is not a computer. The point is just as clearly that the brain is also not a quantum computer. This might be surprising, because Penrose is commonly associated with the idea of a quantum mind, but he specifically states that the brain is not any form of quantum computer that we know of, in fact it does not use any known quantum effect. He does later (in the tentative part) suggest that it uses an unknown quantum phenomenon, and that microtubules might be at the heart of such a phenomenon in the brain (Hameroff and Penrose 1996), but the important distinction between issue and solution should still be respected.

Another important theme in Penrose's view is the connection between physics and mind, which is also somewhat underappreciated. Few neuroscientists would agree that there is any tangible connection between, say, black holes and the brain, but if they can accept Friston's premise that biological agents resist the rising of entropy (perhaps even lower it), and accept that it is a fundamental feature, then the connection is inescapable.

Penrose puts even more emphasis on the need to understand how entropy really works in this regard and in general, and it seems that the real question is not how brains can possibly resist the Second Law of Thermodynamics, but rather where this law comes from in the first place, since there is no *a priori* or theoretical reason for such asymmetry of past and future. This difficulty also led Hawking to cast an ad hoc chronology protection conjecture (Hawking 1992), which is a fancy version of the Axiom of Causality. Penrose's version of this axiom is a little more specific, known as the Weyl curvature hypothesis (Hawking and Israel 1979; Penrose 1982), it states that for some reason the Weyl-part of the space-time curvature tensor starts out (at the Big Bang) at zero, reflecting no distortion of space-time (the other half, the Ricci-part, is infinite and reflects compression). The symmetry in the maths on the other hand suggests that an initial white hole singularity (e.g. the Big Bang), should have high entropy (and infinite Weyl-tensor) just like a final singularity (a Big Crunch black hole), but that is not what we find in reality. This is known as the arrow of time problem (Coveney and Highfield 1992; Ionescu 2008), and highlights the intimate relationship between entropy and time, with rising entropy in a clear sense defining the direction of time. Lowering entropy is then in a more tentative sense the same as going in the reverse direction of time.

In order to even begin to understand how the Big Bang could possibly have been born in a state of extreme low entropy we must consider the question of what was before it; it is often stated that this question does not make any sense since time itself was created with the Big Bang, but then the question really only changes to why it happened at all. This is the old question of first cause, which is an area littered with debris from millennia's worth of philosophic, religious and scientific struggles. On the fringe of science Arthur Young has constructed an argument from first principles (Young 1976), which reverses the order of cause and effect by introducing purpose; e.g. the barn burned down and the pigs were roasted, roasted pigs taste good, so barns were burnt to roast pigs. The purpose of roasting pigs in a sense reverses the causality of the fire

causing the pigs to roast, the fire being caused by the purpose to roast. However, we cannot introduce purpose into science at this level without breaking with science as most scientists see it today, but if we disregard this problem then purpose might indeed be an adequate first cause. The real breaking point is obviously that we cannot be allowed to assign a divine purpose to the Big Bang, but the argument itself really says nothing about divinity of purpose; if we had the technology, the purpose of mankind would do just fine. We bring this up to really understand the implications of entropy lowering brains when taken to the extreme conclusion; bar any major extinction events brains will keep getting smarter and keep lowering entropy at an accelerated pace, while the rest of the Universe will gravitate towards higher entropy (so total entropy is still rising).

If we are still around at the very end we will see some spectacular events as the last supermassive black holes collide, causing gravity waves in the fabric of space-time itself. This last part is surprisingly not just speculation, because the signature of such events has been found in the cosmic microwave background radiation in the form of concentric circles of anomalous low variance (Gurzadyan and Penrose 2010). Obviously this seems impossible, since it has not happened yet, but the theory that predicted these events gives a simple explanation: We are seeing the signature of collisions at the end of a previous universe, part of a chain of universes called aeons; a signature in gravity waves, which is able to pass through the singularity to the next aeon (Penrose 2010). So in this view there was something before the Big Bang, and the first cause of one universe is the last effect of the previous. Neat; except for the question of the first cause of the first universe in the chain. The alternative explanation of these observations is that we are not seeing the end of the previous aeon, but rather the end of our own, but this is of course to be considered unphysical (Penrose 2010) as it breaks with our conventional understanding of causality, although the mathematics are equivalent, and although Einstein's equations allow such loops in time (Gott 2002).

It is the hope that this brief exposition highlights that there are fundamental issues in our understanding of entropy (and in direct extension, of the concepts of information and noise (Von Baeyer 2004)) in physics as well as in the brain, and that the connections sometimes made by physicists, but mostly ignored or rejected in neuro-science, are not just frivolous ideas, but should be seriously considered, not least if Friston's premise is accepted.

It might still be difficult to see how black holes are really relevant to neuroscience, but they are as much (if not more) mathematical objects as real objects, and the mathematical object poses some really good questions when it comes to understanding information processing, entropy and computation. Information is supposed to be indestructible as a consequence of (unitary evolution in) quantum mechanics, but this poses a fundamental problem when information falls into a black hole and most definitely becomes inaccessible to the outside. This is known as the black hole information paradox (Hawking 1976) and has been popularized by a bet by Hawking and Thorne versus Preskill about what ultimately happens to the information; by now Hawking has conceded the bet to Preskill (Hawking 2005), agreeing that information is not destroyed forever, but released in the form of Hawking radiation (correlated with the original information) over millennia, but the question is obviously still wide open. Let us consider in a little more detail what we are taking about; a thought experiment known as Hawking's box is illustrative (Penrose 2001): Consider a box of cosmic scale, the box contains enough mass to eventually form black holes. Now consider the phase space describing the state of this box. The volume of a compartment in phase space is the entropy associated with the state(s) described by that compartment. The phase space of the box will have two large compartments, one describing all the states where there is at least one black hole in the box, the other describing all the states where there are none. The time evolution of the state of the box is given by arrows flowing from compartment to compartment, eventually passing from the no black hole side to the black hole side (compartment). The volume of any phase space must remain constant by Liouville's theorem, it is incompressible, but what if we allow information to be destroyed by a black hole? This causes the arrow flow lines to converge and merge, and reflects an illegal compression of phase space, but we might save the overall balance sheet as long as an equivalent amount of phase space volume or information is created somewhere else. In a classical phase space there is no mechanism to possibly allow such a thing, but in quantum phase space (Hilbert space) the realm of possibilities is still somewhat uncharted. Hilbert space volume is conserved under unitary evolution, i.e. reversible evolution, but decohering measurements collapsing to classical states are irreversible, and so is the creation of Bell-states, maximally entangled states, which are an essential resource in for example quantum teleportation (Bennett, Brassard et al. 1993). Certainly we can imagine that information teleportation could play the role needed; if this sounds too farfetched it is worth considering what the alternative of correlated Hawking radiation entails: Close to the black hole event horizon quantum

fluctuations (which are always occurring everywhere) create virtual particle pairs out of nothing in singlet Bell-states, which immediately sum (annihilate) back to nothing, e.g. an electron-positron pair (a positron is the time-reverse of an electron, its anti-particle (Feynman 1990)); however occasionally the anti-particle falls into the black hole and the particle escapes, the negative mass of the anti-particle then means that the black hole loses mass, i.e. evaporates. This process in itself does not mean that information is re-created although some information is carried by the escaping particle; for the information to be the same as originally fell into the black hole long before, we must consider the concept which made Hawking eventually give up his position on the information paradox: The idea is that when matter falls in, its information content is not destroyed, but rather imprinted on the event horizon surface of the black hole. This is now known as the holographic principle (Susskind 1994), because of its similarity to general holography, the details of which we will return to in chapter 3. Then as the black hole eventually starts to evaporate via Hawking radiation this radiation is imbued with the information stored in the horizon.

The entire process sounds quite farfetched, but the consequence that black hole entropy is proportional to its surface area rather than to its volume is generally accepted. This view of black holes is intimately related to a more general version where the entire information of the Universe is considered to be represented on a dual space surface of one less dimension, known as the AdS/CFT correspondence (Maldacena 1999). A not popularly known fact is that this correspondence in its natural form is between a hologram (the surface) and a universe with an extra dimension of time; only by a (forced) mathematical so-called unwrapping procedure is it then turned into a universe respecting our normal notions of causality (Penrose 2001).

Penrose rules out information creating white holes (other than the Big Bang) on clear observational grounds, but perhaps overlooks the most obvious scenario in which the white holes to look for are not actual Big Bang-like singularities, but rather precursors to such events, in the same general entropy sense that stars are precursors to black holes. We can succinctly say that it appears that brains are to white holes as stars are to black holes (in terms of entropy). So in this simplified view stars and black holes destroy information, while brains and white holes create information. The assumed indestructibility of information aside, it does not seem unreasonable that brains create information, perhaps simply by identifying and removing noise.

1.3.3 *The brain is too cold to be a classical computer*

We have already seen that Gödel's theorems imply that the brain is not a classical computer; here we will consider a more practical perspective from the point of view of the physical temperature of the brain as a measure of information processes within. The argument that the brain is too hot to be a quantum computer is well-known (Tegmark 2000), and it is implicitly understood that it is then a classical computer (if nothing else then for lack of alternatives). We will revisit this argument shortly, but first let us consider a similar but less known argument that simply states that the brain is too cold to be classical computer. This requires a bit more explanation, but it is not a complicated consideration: If we assume that the brain is able to solve intractable problems (as we have argued) we run into a basic problem, which is that all known computation comes at a cost; the cost might seem to vary and also appears to be rapidly decreasing (Moore's law), but that is the economic cost; there is a more fundamental cost, which within known physics has a lower limit (Berut, Arakelyan et al. 2012), known as Landauer's limit (Landauer 1961). From the energy bill it would seem that the cost is energy, but that is still the economic cost; the energy is conserved, the amount the goes into the computer is exactly equal to the amount that goes out (mainly heating the user's office or lap).

What actually changes is the entropy content of the energy; that is, it enters as low entropy energy and leaves as high entropy energy, i.e. heat. So the real cost is (negative) entropy, and as an associated cost is the generation of an amount of heat, which is necessary to balance the books of entropy accountancy. In order to lower the entropy (do the computation) the processor must raise the entropy in its environment. Since it takes an exponential amount of computation to solve an intractable problem we can conclude that a processor must dissipate an exponential amount of heat to do so, which the brain is clearly too cold to be anywhere near doing. The substrate of the computation does not matter here, be it silicon or biology, and the effectiveness of the process is also irrelevant; the argument that nature has found a more effective way does not apply, unless we mean a way that is not a classical computer. Even in that case it would be extremely surprising if that way was a way to reduce the cost below the fundamental limit; it could however be possible to pay in a different way, so to say. If the brain's lowering of entropy did not have to be balanced by an increase in entropy in the local environment, but rather by an increase in entropy in the larger environment, say in the Sun (which is the largest account holder anyway), it would be past the heat dissipation

concern. This kind of suggestion could also include a notion of entropic credit, so that the account could be pre-paid or perhaps settled later, but neither is how classical computation works.

We should add that the above is not true under ergodic conditions where exponential complexity does not require exponential dissipation, which is true for physical annealing, but not for simulated annealing (Hastings and Waner 1985).

1.3.4 *Is the brain really too hot to be a quantum computer?*

We have seen that Gödel's theorems also imply that the brain is not a quantum computer in the conventional sense, but at this point the reader should also have come to appreciate the reasons for considering quantum-like effects in the brain in any case. We have introduced the main arguments why classical physics is not enough to produce cognition, and although the same arguments apply to quantum theory we will soon argue that it is still the most likely place to uncover phenomena, which could eventually defeat these arguments.

The main argument against quantum effects playing a role in cognition is that the brain is too hot. This is because quantum coherence is extremely short-lived at physiological temperatures, where it is very difficult to isolate the quantum system from the classical environment. Physicist Max Tegmark and others have made calculations estimating how long quantum coherence could be maintained in a brain-like environment at physiological temperatures, and concluded that the timescale ($\sim 10^{-13}$-10^{-20} seconds) is orders of magnitude too short to be able to have any influence on neurons (Tegmark 2000; Litt, Eliasmith et al. 2006). This is a reasonable consideration, but it very much reflects the current main problem in quantum computing, which is also specifically to maintain quantum coherence (avoiding decoherence) for long enough to perform computations, and one could therefore suspect that this might have coloured the reasoning somewhat.

There are two potential ways around this temperature roadblock; the first is the simplest and just involves changing the assumptions of the calculations: Tegmark assumes that we are dealing with microscopic quantum objects that are inherently temperature sensitive due to the scale. However, there is a clear trend that systems that are not

microscopic can still behave in a quantum coherent way, both in general (Vedral 2011) and in biology (Ball 2011); this was originally thought to be impossible, and to our knowledge is still theoretically impossible, since there is no complete explanation, but many different experiments document it, most famously with macroscopic C60 carbon buckyballs (Arndt, Nairz et al. 1999), or with Schrödinger's kittens, which are macroscopic packets of light (Bruno, Martin et al. 2013; Lvovsky, Ghobadi et al. 2013). The upshot for Tegmark's temperature argument is that perhaps temperature is not as important as was assumed, and indeed quantum coherent energy transport has since been found in certain proteins at ambient temperatures (Collini, Wong et al. 2010). The natural extension of the trend that macroscopic objects can behave in a quantum mechanical way is to propose that perhaps all classical objects can behave as if they were quantum under the right (but rare) conditions. The problem is to understand what those conditions are; the conditions in the brain are certainly rare, but there is nothing obvious about them that suggest a way to achieve quantum phenomena.

The other potential way around the temperature problem comes from quantum computing. The first somewhat effective way to deal with unwanted decoherence in quantum computing was the introduction of quantum error correcting codes, which take advantage of quantum parallelism to introduce massive redundancy rather than actually solving the decoherence problem itself. A more interesting (but much less developed) approach is based on the curious discovery of decoherence-free subspaces (Lidar, Chuang et al. 1998); subsystems that are passively protected from decoherence otherwise caused by unwanted interactions with the classical environment. It should be noted that even if we could prevent unwanted decoherence it is not thought that quantum computers can solve intractable problems. However, it only takes very minor changes to the quantum computing formalism to gain this ability, but it has not been proven in practice, and it is often taken to imply that a given change violates quantum mechanics or rather violates fundamental axioms in the interpretation of quantum mechanics. In this sense quantum mechanics is considered an island in theory space (Aaronson 2004a), because any slight change causes fundamentally different results; but this is really only to be expected considering that we must give up at least one of our fundamental axioms (as discussed above) to explain the results in the first place, and depending on which we give up the interpretation is very different. We will outline the fundamental axioms and their present relevance in more detail in chapter 3.

1.3.5 *For now define a holographic computer to be just right*

If we rule out that the brain is a classical computer, and also rule out that it is a quantum computer we are indeed left with very little; it is exceedingly difficult to come up with physical concepts that cannot be described by either or both formalisms. We will however for the sake of inquiry postulate a third kind of computer, which we, for reasons that will become clear in the chapter 3, call a holographic computer (holographic computation is considered in more detail in chapter 3). The properties of this hypothetical device are just right, in the sense that it is somehow unaffected by Gödel's theorems, and is thus able to solve intractable problems. It is part of the definition that it is allowed to achieve this by violating any fundamental axioms as long as it provides a coherent interpretation of the consequent physical phenomena. In this way we form an initial assumption, which is that the brain is such a holographic computing device, and that it as such holds the secret as to exactly what this means. All neuroscientists must agree that the brain holds the secret to cognition, but our point of view takes it further in the sense that this is also assumed to be the secret to understanding real physics that is not captured by the classical and quantum formalisms as we understand them now.

1.4 A reverse-engineering perspective on the brain

The concept of reverse-engineering is a well-known approach to a specific kind of problem where a reference device is given in some form, but cannot be taken apart in the traditional sense to simply reveal its mechanistic operation, as is the case with the brain. A good example is the problem of artificial intelligence, which is a valid target of reverse-engineering, because intelligence is known to occur naturally; whereas examples where reverse-engineering would not apply are other science-fiction concepts such as teleporters or tractor-beams, because no such naturally occurring phenomena are known. Reverse-engineering is however not just a recipe of serial steps to carry out; rather it can be considered an art-form in which creative thought is necessary. It is necessary, because in order to initiate any investigation a solution or direction towards a potential solution must first be imagined, and fundamental (but tentative) constraints must be established. It is usually a poor strategy to test for specific end-point solutions unless ones guess is so inspired as to be right; rather a strategy should be laid out that partitions the vast problem space, so that further experiments narrow the space down until a less miraculous insight can be distilled. It is also inherent to the reverse-

engineering approach that the only real success is the (re)implementation of the critical features of the device in question; famously, Feynman's blackboard at his death had scribbled on it, "what I cannot create, I do not understand" (among many other things) (Gleick 2011). Creation does not always require full prior understanding, but it carries an ultimate form of proof, which is much less sensitive to interpretation than many other forms.

The starting point that we will consider here springs from the idea that the enormous advantage of tractability might come at a steep price; the cost which seems most universal and relevant is simply the cost of sleep: We spend about a third of our lives asleep in what seems to be an utterly unproductive and defenceless state that as far as everything but the brain is concerned could just as well be spent in a state of wakeful rest. It could of course be argued that the cost of sleep is greatly tempered by social structures, and that there could even be a benefit of sleep forcing the need for social structures, but at face value it is hard to see why we would sleep if it was not absolutely essential. *We therefore hypothesize that the wake-sleep cycle is necessary to facilitate information processing of a holographic nature,* as postulated above.

1.4.1 *The two concurrent and complementary approaches of this thesis*

The theoretical framework of holographic processes (detailed in chapter 3) suggests several general features that might be present in systems that potentially use holographic phenomena to solve difficult problems. It is however unclear how to implement such a system, so this is where an experimental, electrophysiological approach comes in: We will consequently look for traces of these enabling features *in vitro*, and then for crucial hints as to how they might be re-implemented *in silico*. In return the concurrent *in silico* approach is intended to provide constraints on our hunt for said hints. It is the hope that the two approaches will converge to a point where we have *in vitro* experimental evidence for one or more features and hints, and *in silico* evidence for the scope of computational usefulness of the system under study. If our hypothesis proves even partially correct we would then consider a specific route to a novel technological implementation, time permitting. We will detail these aims at the ends of the next two chapters, respectively introducing each of the two approaches.

Chapter 2. The Wake-Sleep Cycle

neurophysiological background and framework

This chapter provides a short introduction to the neurophysiological background of the electrophysiological approach, and thus forms the framework for this experimental work. We will specifically consider the wake-sleep cycle, but limited to the aspects deemed most relevant to the reverse-engineering perspective specified in the previous chapter: If we are to investigate the hypothesis that the brain is a holographic computer we must consider how information is fed into the device (sensory processing during the wake-state) and how this representation is operated upon to generate desired/useful outcomes (cortical dynamics during sleep and their relation to prior activity during the wake-state).

2.1 Is sleep the price of tractability?

In the introduction we set forth the speculative idea that we might start a reverse-engineering process of tractability in the brain by considering one of the main costs of operating a brain, namely the need to sleep. Now we must consider which aspects of the wake-sleep cycle are most likely to be relevant to solving difficult computational problems. Fortunately, neuroscience has produced a rather comprehensive and successful theory, which deals almost exactly with the kind of problem we are interested in; we will refer to that theory here as the synaptic rescaling hypothesis of the wake-sleep cycle (Turrigiano, Leslie et al. 1998; Tononi and Cirelli 2006).

2.2 The wake-sleep cycle and the synaptic rescaling hypothesis

The basic idea of the synaptic rescaling hypothesis is simple: While we are awake and experiencing sensory events, we build up synaptic potentiation encoding (but not necessarily organizing) new information. Sensory inputs are presented to the cortex in a patterned manner (i.e. they are pre-processed by sense organs and subcortical structures like the thalamus (Ramkumar, Jas et al. 2013)). The inputs are presented in a highly layer-specific manner (see below) and, as far as neuroscientists can tell to date, are represented both spatially and temporally within the functionally-specific primary sensory area concerned, and connected to associational areas higher up the cortical

hierarchy (see Einhauser and Konig 2010). In this way the internal (to the brain) representation of sensory input can be thought of as a distribution of information across the cortical mantle (an 'engram', Lashley 1949). The nature of this information is thought of classically as a population of bits in the form of action potentials in neurons. However, the temporal arrangement of these action potentials (bits) is tightly controlled by oscillations in neuronal populations; shared, periodic fluctuations in neuronal membrane potentials from hyperpolarised to relatively depolarised levels (Whittington, Cunningham et al. 2011). This dependence of cortical representations on oscillations has three main consequences for the hypothesis under test: Firstly, oscillations are not classical 'bits' of information. They can exist in extreme states (peak/trough), but also in *any* state in-between. Secondly, their presence in any given neuron is inescapably dependent on the state of multiple other neurons, i.e. they are 'entangled' together by the network process that generates them in the first place (see the discussion on Bell-states above and qubit structure in chapters 3 and 7). Thirdly, from a neurophysiological perspective, the temporal arrangement of cortical activity afforded by oscillations has a powerful influence on synaptic plasticity (Kopell, Whittington et al. 2011; Ainsworth, Lee et al. 2012). In other words, the very process of representing sensory information in cortex has an inescapable synaptic plastic component.

This process of synaptic plasticity is naturally limited by an upper degree of potentiation, a limit that we feel as increasing sleep pressure as we approach it (Dijk 2009). Then when we fall asleep a different process starts whereby the synaptic potentiation built up during the day is brought back down to a nominal level (Esser, Hill et al. 2007), so that another wake-sleep cycle can start over the next day. This all sounds simple, but it involves some rather profound computational problems: It appears that it is not enough to simply scale down all synaptic connections by the same amount, because synaptic plasticity is by default noisy, and is 'unsupervised' in the sense that it is not necessarily even given what is noise and what is information. There are several ways to look at it; we can think of a rescaling process as data compression, which is formally equivalent to pattern recognition or to inverse problems, like learning from examples (Vito, Rosasco et al. 2005). These are ways to formalize the computational problem faced by the rescaling process, and since these formalizations have been studied in some detail already, we can apply such prior work to evaluate ideas like scaling down all connections. If such a simple normalization process was enough it would probably not reflect a very interesting (computational) problem, and we should

perhaps consider rescaling as merely some housekeeping process, which is not particularly relevant to the computational problem at hand. Certainly, from a purely metabolic perspective sleep appears to have multiple restorative purposes at the cellular level (Vyazovskiy and Harris 2013), but this does not in itself rule out a more computational purpose. In addition, as we shall see from data presented in chapter 5, some neurons show activity patterns during (a model of) deep sleep that cannot at all be seen as restorative in the above sense as they are extremely metabolically demanding.

However, there are several ways in which the brain could potentially deal with noisy potentiation: The simplest well-known approach is to re-scale in some frequency or frequency-like domain; removing the smallest Fourier-like coefficients is a regularization process, which blindly removes the worst noise. The brain might well be doing something just like this, but in spite of extensive study and numerous applications this sort of process has not yet been able to tackle intractable problems, and we must consider that if this is what the brain is doing, it is doing it in a way which we have yet to understand. So we are still left with the question of how the brain selects which connections to scale; we will have much more to say about this question below, and again following the results of chapter 5.

Neuroscience fortunately seems to provide a straightforward answer in the form of a process, which is best called replay (Peyrache, Khamassi et al. 2009). We will consider this process in some detail, but first let us reflect on a possible mechanism for rescaling itself. The question is what characteristics are associated with synaptic depotentiation, and what their dynamics are over a wake-sleep cycle. One characteristic, in the form of delta rhythms, is both associated with depotentiation at a long-term synaptic plasticity level (Malinow and Malenka 2002; Cingolani, Thalhammer et al. 2008), and with prior potentiation by stimulation (Huber, Ghilardi et al. 2004; Vassalli and Dijk 2009). The point is that increased local potentiation by some form of local stimulation during wakefulness leads to stronger local delta rhythms during deep sleep. Let us therefore further characterize delta rhythms in relation to the wake-sleep cycle.

2.2.1 *Delta rhythms and the wake-sleep cycle*

Delta rhythms are primarily associated with deep sleep where they are particularly prominent. The two main sleep phases (with further divisions), non-REM and REM,

display strongest delta in deep non-REM sleep, which characteristically declines until the beginning of a REM phase, during which it is substantially lower. At the end of the REM phase delta power returns to a level slightly stronger than at the beginning of the REM phase, and the decline starts over until the next REM phase. The non-REM delta phases get progressively shorter as the REM phases get progressively longer (McCarley 2007). Except for the minor increases following REM phases, delta power continually declines throughout the night; eventually reaching wake levels. It should be noted that delta is far from the only active oscillation during sleep, as we will also see shortly for wakefulness, delta in sleep is orchestrating a nested hierarchy of smaller, higher frequency oscillations (Valderrama, Crépon et al. 2012; Carracedo, Kjeldsen et al. 2013)

Although delta is most prominent and clearly dominant in sleep, it is still very significant in wakefulness. It might be necessary to make a distinction between two types of delta in wakefulness: One type associated with sleep delta and the other with primary sensory processing. In support of the first type, experiments show that delta increases in wakefulness with sleep pressure (the need to sleep due to prolonged wakefulness and sensory experience) (Franken, Dijk et al. 1991). This has recently been proposed to be due to actual local sleep (Vyazovskiy, Olcese et al. 2011), i.e. parts of the brain sleeping while you are awake, and it might be an attractive explanation of performance decrease after prolonged wakefulness (Van Dongen, Caldwell et al. 2006).

The second kind, related to sensory processing, might be more interesting here since it suggests a purpose that goes beyond synaptic rescaling in sleep or local sleep in wakefulness. In sensory processing the delta rhythm is prominent in primary sensory areas, and it has been suggested that it acts as a carrier wave for higher, nested frequencies in a processing hierarchy where delta entrains theta and theta in turn entrains higher frequency oscillations like gamma (30-100hz) (Lakatos, Shah et al. 2005), which is typically associated with sensory processing and attention through synchrony and coherence. This is to some extent supported by results showing that delta is phase reset by the onset of visual stimulation, and that in-phase stimulation improves reaction time while out-of-phase stimulation is suppressed (Lakatos, Karmos et al. 2008; Schroeder and Lakatos 2009). In development delta power changes with progressing development of layer 5, while theta power follows the development of layer 2/3 (Campbell and Feinberg 2009). Delta power is particularly strong in infants, and

through childhood, but declines abruptly around puberty and continues to decline with maturation, however its ability to phase-lock with stimuli is not necessarily affected by age (Schmiedt-Fehr, Duhl et al. 2011).

One of the most interesting features of the neocortex is the layered structure defined by differences in cell types, cell type distributions and connectivity. Functionally, layer 4 is considered an input layer, and layer 5 is considered an output layer, reflected to some extent in decreasing size of layer 4 towards higher areas that receive no input directly from sense organs. Layer 4 primarily passes information to layers 2/3, from where it is relayed back down past layer 4 to layer 5 (Thomson, West et al. 2002). Layer 4 also has reciprocal connections with layer 6 (Tarczy-Hornoch, Martin et al. 1999), which feed information back to thalamus and also direct connections to output layer 5 activated on receipt of strong stimuli (Ainsworth, Lee et al. 2011). Layer 4 has the highest concentration of inhibitory interneurons, while layers 2/3 and 5 have more pyramidal cells.

The profile of delta power across layers in associational rodent neocortices in a pharmacological delta model is very specific, with highest power in layer 5. In this layer, it is powered by an NMDA receptor-driven network of intrinsic bursting cells activating a $GABA_B$ receptor-mediated source of slow inhibition. Regular spiking cells in layer 5 generate outputs at theta frequency, which correlate temporally with delta-nested theta frequency EPSPs in layer 2/3 principal cells. In contract, outputs from the intrinsic bursting cells correlate with IPSPs in these superficial layers (Kjeldsen, Kaiser et al. 2012; Carracedo, Kjeldsen et al. 2013). The most noticeable aspect of the delta phase relations between layers is an abrupt and almost complete phase reversal between layers 5 and 1 (Carracedo, Kjeldsen et al. 2010).

2.2.2 *Replay and rescaling*

There is almost abundant evidence for replay in sleep of patterns of spiking neural activity seen during prior wakefulness (Destexhe, Hughes et al. 2007), primarily in hippocampus (Nadasdy, Hirase et al. 1999; Foster and Wilson 2006), but also in cortex (Peyrache, Khamassi et al. 2009), and in coordination between hippocampus and cortex (Ji and Wilson 2007; Born 2010). It is then not too far off to propose that replay has a similar effect to rehearsal or repetition, and thus selects out information, which is to be

spared rescaling as defined as depotentiation. However, have we then not only really moved the problem to the question of how patterns are selected for replay in the first place? We could say that it is simply the most salient information (Diekelmann and Born 2010), but we do not know which is most salient in advance, unless it is simply the strongest, in which case we are back to scaling by a simple normalization process. Another way we could resolve the situation is to say that the salient features are determined in advance by some information processing step occurring as the synaptic potentiation is build up, and perhaps then coded in the strength of connections; this would move at least half of the computational problem to the wake-state side, which might be quite reasonable as a way to share the load so to say. In this case we would ideally need to try to identify such a process in the wake-state, and to try to characterize its computational properties. To some extent evidence for this is already available when considering the interaction between temporal structure afforded by sensory-evoked cortical rhythms and the spike timing-dependent plasticity framework (Lee, Sen et al. 2009).

So a closer (theoretical) inspection might suggest that replay is more of an epiphenomenon than an actual answer to the question of how connections are selected. Arguably any reorganization or noise cleaning process must involve some degree of reactivation of the information in question, so it is not surprising to find replay in any case, but we must be prepared that it is not a self-contained process, and that there could be some intimate connection to processes occurring already in the wake-state, most likely through some neural synchrony or coherence scheme as discussed above (Engel, Fries et al. 2001; Kaiser and Lutzenberger 2005; Benchenane, Peyrache et al. 2010), perhaps of nested gamma rhythms (Lakatos, Shah et al. 2005), as associated with sensory processing.

In addition to replay in sleep there is also evidence for replay during the wake-state (Foster and Wilson 2006; Karlsson and Frank 2009), during REM sleep (Louie and Wilson 2001), during sleepwalking (Oudiette, Constantinescu et al. 2011) , and even of what is better called preplay, perhaps reflecting mental planning processes (Buhry 2011). The relative time scales of replay also vary from real-time to very accelerated (Nadasdy, Hirase et al. 1999; Euston, Tatsuno et al. 2007), and there are even reports of reverse replay (Foster and Wilson 2006), which is particularly interesting in relation to the Helmholtz machine concept that we will consider shortly.

This all further complicates the picture, and it would also be hard not to notice a burning need to further integrate REM dream sleep into the rescaling and replay framework. This has of course been considered already, it is for example known that dream sleep displays less reactivation than deep sleep, at least in rodents (Diekelmann and Born 2010), but although we agreed that this reflects different processes, the suggestion that consolidation occurs during dream sleep seems awkward in our view, and is also still controversial (Vertes 2004). Another potential difficulty with replay is the question of which neural code (or codes (Ainsworth, Lee et al. 2012)) is being replayed; this is not really a problem for place cells, where a route through the environment describes a very basic neural code, but in general neural codes are unknown, and a given, observed pattern of activity might not code for anything of interest. Certainly the findings of, for example, accelerated replay put some limits on the types of neural codes that can possibly be reflected by this phenomenon, but it is still too early to tell. In extension there is the fundamental question of what really constitutes replay and how it is quantified (Chi, L. Rauske et al. 2003); ultimately, judgement of the various forms of correlations will have to wait until we understand neural coding more fully.

From our point of view the essential question is whether replay is a concept that can be applied to help improve existing computational approaches to compression, pattern recognition or other inverse problems, or even to suggest entirely new approaches. If replay is selective and only revives the right, relevant information the answer is clearly positive, but the concept of replay itself does not come with an adequate explanation of how such selection would work, unless it is the most salient information, the selection of which would then have to be explained instead.

2.2.3 *Delta orchestrated interaction looks like a Helmholtz machine*

Our lab has previously suggested that reciprocal interaction across layers during delta is suggestive of a Helmholtz machine running a wake-sleep learning algorithm (Carracedo, Kjeldsen et al. 2013). This is interesting from several perspectives; let us first consider the concept of the Helmholtz machine and how the wake-sleep algorithm operates (Dayan, Hinton et al. 1995; Hinton, Dayan et al. 1995): The basic characteristic of the Helmholtz machine is that it is composed of two sets of stochastic weights (Dayan and Hinton 1996), which converge onto the same probability distribution

(although the weight sets will be different in general) during a weight learning process, which alternates the flow of information from one set to the other; during one half-cycle one set of weights is updated, during the other half-cycle the other set is updated. The updates proceed as the difference between the input and the outcome of a back and forth pass through both sets of weights. This algorithm is called the wake-sleep algorithm, because it can be mapped onto a wake-sleep cycle as follows: During wakefulness sensory input activates the first set of (so-called recognition) weights, the outcome of which is fed into the second set of (so-called generative) weights producing another outcome, which is compared to the original input, the second set being updated by the difference. During the following sleep half-cycle a random 'dream' activates the second set, generating an outcome that is fed to the first set producing yet another outcome, which is compared to the dream, and then the first set is updated with the difference.

We can take this analogy to biology further by also mapping it onto the neocortex's anatomical and functional structure; tentatively this mapping identifies one set of weights with primary sensory areas and the other set with associational areas, this also allows the reciprocal flow directions to map onto bottom-up and top-down processing in a natural way. In addition, the above wake-sleep analogy considers only the dream, REM-state of sleep. But what about deep sleep? Here we further develop the idea that prior wake-sleep reciprocal interactions with the Helmholtz scheme also occur at a much more time-compressed level; that of individual delta rhythm periods.

Thus if we take the reciprocal interaction across layers over one full delta period as one full cycle of the wake-sleep algorithm, the above mapping then immediately suggests that the reciprocal interaction found across layers should also be in play along layers, and that the interaction with weights as well as the comparisons (for updates) should happen twice on each cycle, i.e. major computational events are to occur with double the fundamental delta frequency.

We would ideally need to reconsider how the abstract wake-sleep algorithm relates to the actual wake-sleep cycle, since one delta cycle appears to correspond to one wake-sleep cycle in the algorithm; certainly delta rhythms are also prominent in wakefulness, but the characteristic pattern over the entire wake-sleep cycle is not explained by the Helmholtz machine alone. This is not necessarily unsalvageable, since the abstract terms of the Helmholtz machine can be manipulated to some extent, we might for

example hypothesize that the characteristic changes in delta power reflects changes in the thresholds of the stochastic weights, i.e. changes in the so-called temperature (Kirby 2006), adding a potentially interesting dynamic. We will return to these issues a little further down once we have covered a bit more of the important features of the Helmholtz machine concept.

The Helmholtz machine concept is particularly interesting for our search for clues to tractability, because it is exactly an attempt to make the Bayesian brain concept tractable. Although it is not known to be a general solution, it can be shown to minimize a free energy measure similar if not identical to Friston's (Kirby 2006). The question then arises if there is a similar assumption to Friston's ergodicity involved (see also section 1.3.1). This is difficult to glean from the concept itself; however, thinking about such potential equivalences prompted the consideration that two sets of real valued weights can be seen as one set of complex valued weights, which in network terms then makes it a quantum neural network (Peruš 2000; Perus, Bischof et al. 2007). This would then have interesting implications for quantum mechanics as well as for brains (to the extent they can be described by Helmholtz machines). Indeed, as revealed by a search thus directed, someone faster and with the skills to produce the rigorous proofs had recently shown the equivalence of a stochastic Helmholtz machine running a wake-sleep algorithm with a basic quantum evolution by Schrödinger's equations (Chapline 1999; Chapline 2004; Chapline 2008). The mapping is even quite obvious once spotted, since the quantum evolution consists of both forwards and backwards (in time) parts mapping naturally to the reciprocal wake and sleep half-cycles of the Helmholtz machine.

However, the author notes that simulating quantum mechanics via a Helmholtz machine is not enough to make the process tractable, so we do not directly gain any new computational power. We can tentatively identify the quantum evolution, which forms a quantum reciprocal diffusion process (Levy and Krener 1996), with the recently popularized concept of quantum annealing (see e.g. Ohzeki and Nishimori 2011), which is a form of quantum computing that is much closer to classical conditions than the previous quantum computing paradigm (Rose and Macready 2007). In fact, if we combine the hypothesis from above identifying delta power with the temperature concept of the stochastic weights, with the profile of delta power during the sleep half-cycle, we get a very suggestive picture of an annealing process, as the delta power

profile corresponds to the temperature profile of a (classical) annealing schedule, and the physical, structural changes during annealing corresponds exactly to the purpose of synaptic rescaling. It also provides a more natural role for dream sleep as causing a temporary increase in the abstract temperature of the system as is done during an annealing schedule. The actual temperature of the brain is known to follow a similar schedule, lowering during deep sleep and increasing during dream sleep (Muzur 2005). We would then suggest that dream sleep reflects random potentiation rather than actual memory consolidation (Siegel 2001; Diekelmann and Born 2010). However, the situation is complicated by the diverse nature of different memory processes consolidated during the wake-sleep cycle. Declarative memory (trivially put, the learning of basic facts and their interrelations) is consolidated preferentially by deep (delta) sleep, whereas procedural memory (again, trivially put, the learning of sequence-dependent cognitive strategies) is critically dependent on REM sleep and not deep sleep (Smith 1995; Dotto 1996). In an attempt to focus, we will consider only delta rhythms as a substrate for the 'sleep' component of our experimental model and consequent computational approach to understanding how the brain may solve intractable problems. What we can take from this section is that it does appear that, seen from the annealing perspective, consolidation occurs during deep sleep by depotentiating noise.

The difference between quantum and classical annealing is that one relies on thermal fluctuations the other on quantum fluctuations (both describable by the stochastic weight temperature concept of the Helmholtz machine), which means that the interpretations of how energy minima are reached are also different: Energy barriers are jumped over in the classical process, but tunnelled through in the quantum process; also in the quantum case the system is considered to be at multiple energy levels concurrently, and only as the energy barriers are raised (by reducing fluctuations) it becomes increasingly likely to end up in the lowest single energy state (Battaglia, Santoro et al. 2005; Rose and Macready 2007). Classical annealing is an ergodic process, so there is potentially some common ground with Friston's ideas, but from the delta power profile we would suggest that the process would only occur during sleep. Quantum annealing is probably not ergodic in the classical sense of eventually visiting all states, but potentially in a more quantum sense of visiting all states simultaneously.

Following up on the view taken at the end of the first chapter we might propose a hypothesis of what we will then call holographic annealing, which possesses the same

abstract computational properties as quantum annealing, but without a direct requirement of actual quantum microscopic processes, and more importantly at the same time is tractable. We can see already that if the brain can be described by processes similar to a stochastic Helmholtz machine with a wake-sleep algorithm it can simulate basic quantum processes, which in turn enable quantum annealing, without involving actual quantum mechanics. This simulation would of course need to be tractable for the scheme to work in practice (which is part of what we mean by holographic), and this is the real unknown factor, to which we hope to uncover clues, by examining the relevant processes in an electrophysiological *in vitro* model of the wake-sleep cycle.

2.3 Electrophysiological aims

On the background established above we finally set forth the aims of the electrophysiological approach. Implicit in all these aims is the underlying aim of keeping our eyes open to any and all hints of tractable processes to help shape the computational approach, the background of which is outlined in the next chapter.

2.3.1 *Establish a simplified model of the wake-sleep cycle*

We first aim to design and implement a reduced and simplified *in vitro* model of the major aspects of the wake-sleep cycle by pharmacological manipulation at the interface between sensory and associational areas in the rodent neocortex. The validity of this model will only be confirmed by success of our further aims as follow below.

2.3.2 *Investigate the influence of prior activation on delta power*

Using the substrate of the wake-sleep model of the first aim, we will pharmacologically stimulate proximal to sensory areas in layer 4 to simulate sensory input. This will allow a micro-electrode array investigation of the influence of prior activation on subsequent delta power in both sensory and associational areas, following previous studies as outlined in the background above.

2.3.3 *Investigate reactivation in delta of stimulated units*

Further, given the success of the previous aims we will investigate the influence of prior activation on unit activity across the states of the wake-sleep model. Specifically we are interested in replay of unit activity for the many reasons given above, and in

possible coherence during stimulation as potential explanations of increased delta power via increased potentiation, and of selection of information for replay.

2.3.4 *Investigate possible ergodic dynamics in the wake-sleep model*

Given that the above aims are met we will know that the model is a reasonable (although reduced) representation of the wake-sleep cycle. On this background we will then investigate whether any or all of the model states comply with the statistics of ergodic dynamics in order to evaluate the many important consequences mentioned above.

2.3.5 *Expand on Helmholtz machine interaction*

Finally we wish to test the predictions listed above by the tentative mapping of the Helmholtz machine concept onto the sensory and associational areas of the rodent neocortex during a sleep-like delta rhythm. Candidates for this investigation are pairwise causality measures, as well as multi-channels extensions; we will however also consider the possibility of other novel options.

Chapter 3. Holographic Processes

computational background and framework

This chapter builds a conceptual framework around the computational properties of holographic processes; we will thus try to define the key aspects of holography, which might enable tractability, and then specify how we might look for them in our electrophysiological experiments. We will also try to explain in more detail what we mean by holographic processes, but we will not fully review the myriad aspects and interpretations of holography in general (see instead e.g. (Talbot 1996)). A clearer overview of the set of foundational axioms and their respective implications for computation (alluded to above in the introduction) will be given in the context of understanding the close relations of quantum and holographic computing.

3.1 Gabor's discovery of the holographic process

The form of holography most people are familiar with is also the form which was invented or discovered first, by the great Hungarian electrical engineer and physicist Dennis Gabor in Britain in 1947, and which earned him a Nobel Prize in 1971 among many other prizes. This form is optical coherent holography, and is far from all there is to holography, but the anecdote is quite insightful: Gabor had the simple idea that in order to get the best possible microscope image one had to use all the incoming light in the imaging process instead of filtering it in various ways to obtain various limited views. Holography literally means the whole information. The remarkable holographic principle which fulfilled this idea required phase information to be recorded and used in the image reconstruction process. With the advent of the coherent laser in 1964 Gabor's idea became practical reality, and led to improved microscopy as envisioned. Gabor and others soon realized that the natural simplicity and power of holography meant that it was possible that biological evolution could have taken advantage, and suggested that biological memory used holographic principles (Gabor 1969; Westlake 1970; Poggio 1973).

3.2 Holographic processes in the brain?

These ideas were taken further by the neurosurgeon Karl Pribram who noticed the remarkable analogy of the failure of the search for the memory engram, with the key property of holographic recordings that information is stored equally everywhere in the hologram (Pribram 1971). Pribram's so-called holonomic brain theory essentially predicted that the brain processes information in the frequency domain (Pribram 1986; Pribram 1991); an idea which is still considered a valid and even well-supported hypothesis in neuroscience (Palva, Palva et al. 2005), but which is rarely considered in Pribram's holographic perspective. Part of the reason for this neglect is likely due to his collaboration with acclaimed physicist David Bohm whose holographic interpretation of quantum mechanics interfaced naturally with Pribram's ideas (Wilber 1982), and added a generalization of frequency domain processing to a more abstract higher dimensional frequency domain (Bohm and Hiley 1995), which unfortunately also made the theory less accessible. The integration of these ideas from scholars of disparate fields demonstrates the striking generality of a holographic principle, which applies equally across dimensions and scale. We can also begin to see that this principle involves a duality between a real space and a Fourier-like space, in which certain computations might be easier, just like circular convolution is point-wise multiplication in Fourier space, or multiplication is addition in log space. We will return to more details of this duality shortly.

Perhaps the most fascinating example of the generality of holography is the form known as generalized, near-field or incoherent holography, which achieves the same holographic effects, but without the need for a coherent reference; in fact, in some ways it achieves more because additional information can be recovered from the incoherent near-field of the radiation sources, as we will see later in chapter 6, where we introduce a new micro-electrode array data analysis method based on this concept.

3.2.1 *Holographic interpretations allow tractable processes*

Bohm's holographic interpretation of quantum mechanics begs the question of how our hypothesized holographic computation might be related. The answer is surprisingly straightforward: Although Bohm considered his interpretation just that, an interpretation, because he foresaw no technological consequences, no differences in predictions or experimental outcomes, we now know that it allows a special form of

measurement, which defeats intractable problems (Aaronson 2005a), at least in theory. This has then been used to demonstrate the inconsistency of the theory, because it allows something, which other theories (considered to be equivalent) do not (Aaronson 2005a). However, in general it is still unknown how measurement is supposed to work in quantum theory, a problem known among other names as the measurement problem (Schlosshauer 2005).

In any case, Bohm's special measurement and consequent computational ability was achieved by giving up the Axiom of Locality (that interactions are local), which is specifically the connection to holography where information is distributed equally (non-locally) over the entire hologram. It is however not the only way to incorporate holographic properties into an interpretation of quantum mechanics; Gabor's notion of the whole information might be applicable to other configurations of axioms, perhaps with additional insights into the classical-quantum 'measurement' interface as we will see shortly.

3.3 Potential loop-holes in Gödel's theorems

Before we get further into other quantum interpretations' implications for intractable problems let us however step back and consider how Gödel's theorems can possibly be defeated in the first place.

3.3.1 *Gödel's loop-hole in this own theorems*

There is no better place to start than with Gödel's own loop-hole, which is also conveniently located in the complementary part of physics, namely in Einstein's relativity theories. When these theories came out a major philosophical problem arose, because an important implication was that time was not a privileged dimension, and as everything else, was relative. If time was relative then the fundamental and otherwise self-evident assumption of causality was in trouble; and with it, the foundation of science, especially of reductionist science. However, science quickly recovered: It took some jumping through hoops defining a preferential frame of reference by the mean movement of all mass in the universe, but eventually a notion of cosmic time was accepted, and relative time was relegated to special relativistic cases, thus reaffirming normal causality (Yourgrau 2006).

Then Gödel did it again, another ground shaking disruption of a major and generally accepted paradigm of science, and on the occasion of Einstein's seventieth birthday in 1949 no less (Gödel 1949). But this time he was unsuccessful, although his math and logic were impeccable, the radical implications could just not be accepted, especially by an already jaded community. The simplicity of his result should not be confused with the difficulty of obtaining it: He gave an entirely new solution to Einstein's equations, which arranged matter and momentum in a way that made the resulting universe possess an overall rotation. The breaking point was that this rotation allowed a phenomenon that is now known as closed time-like curves, which basically just means loops in time.

It is commonly stated that Gödel's universe is different from our own, so there is nothing to worry about, but Gödel's logic was different, and instead effectively stated that the most fundamental physical laws of the Universe should not depend on the specific arrangement of matter and momentum; these laws are given by the equations, not by specific solutions; a phenomenon of fundamental importance to these laws in one solution, should affect the laws for all solutions. This logic is not difficult to follow, but the implications are difficult to accept, and Gödel's own conclusions might have thrown the baby out with the bathwater, and might have been a major factor in its failure to gain acceptance: Gödel considered that time loops inevitably lead to paradox, and that time therefore had to be an illusion in itself, or rather that time was a Platonic ideal, which then confirmed Platonic dualistic philosophy (Gödel 1988). This was however a bridge too far, too mystical rather than scientific, and could not be accepted by science almost by definition.

Gödel continued to look for evidence of an overall rotation of the Universe for the rest of his life; today studies of more than 15000 galaxies have shown a small net rotation (Longo 2011), but the validity of the result is still open. We have also since realized that the putative existence of closed time-like curves would allow even classical computers to solve intractable problems with ease; a point which is usually used to argue against their existence (Aaronson 2005a). However, it is in any case an important point, and as implied earlier we think that Gödel might have had an inkling in his letter to von Neumann, although he did not say so directly, but he was also notoriously cautious with his ideas, especially after the disappointment with the Gödel universe. Interestingly, von Neumann had already stated the year before that quantum mechanics was not compatible with classical notions of causality.

The first direct study of closed time-like curves in quantum mechanics is however due to David Deutsch (well-known as the father of quantum computing), and revealed that quantum computing close to closed time-like curves would be ridiculously powerful (Deutsch 1991). More recently a new development has however upset this conclusion slightly: Seth Lloyd, who is famous for coming up with the first feasible design for a quantum computer, has analysed the situation in more detail and proposed a different instantiation of closed time-like curves, different from the kind considered by Deutsch (Lloyd, Maccone et al. 2011a). In Deutsch's study closed time-like curves lead back to a different past than the proverbial time-traveller remembers, leading to the Many-Worlds Interpretation of quantum mechanics that Deutsch is also a proponent of. Lloyd's closed time-like curves lead back to the same past, and where Deutsch's version naturally resolve time-travel paradoxes, because the pasts are not directly overlapping (being in different parallel universes or so), Lloyd simply requires that the trip back in time is impossible and cannot occur if and only if it leads to paradox; which is a new and more specific formulation of the Novikov self-consistency principle as less formally formulated in the 1980s (Friedman, Morris et al. 1990). This new formulation has in fact been verified experimentally by probabilistically sending a photon back in time to try to kill its past self with 50% probability (Lloyd, Maccone et al. 2011b); the experimental outcome is then consistent with the idea that the trip always fails if the murder is/would have been successful. Technically this process involves a somewhat dubious step called post-selection (Lloyd, Maccone et al. 2011c), which means that a selection is made post hoc on the desired outcome. From a classical computational modelling perspective it would seem that in order to achieve tractability we would more likely need a process that we would then call pre-selection (although obviously pre- and post-selection might be the same in a time loop). It is too early to conclude on this work, but it certainly evokes the idea that Gödel and Deutsch's conclusions about closed time-like curves are extreme, opposite ends of a spectrum, where there might be a golden path in the middle, which neither allows paradox nor unreasonably powerful computation. Our sought after holographic computation might lie further down such a path.

Lloyd's new consistency condition (that time loops leading to paradox do not occur) is in fact not too far from existing techniques; quantum calculations are done on a reduced representation called a density matrix (because the full representation is hopelessly

costly), which specifically only represents the physically relevant states, so from all quantum possibilities and impossibilities only a much smaller subspace is considered. Effects forward as well as backwards in time are naturally present in all this; it is only by specific choice that we normally consider only the time-forwards part of the solution. Lloyd's approach suggests that we can consider both, because even if we do, only the possible can happen; it is an assumption or even an axiom, but the understanding of why or how this comes about is still missing, but perhaps not too far off.

This might then prompt a re-evaluation of the foundational axioms of science; specifically that we might be able to give up the Axiom of Causality, but recover its effect from the interaction of the remaining axioms and the addition of a no-paradox or self-consistency axiom. This is not an entirely new idea, and as we shall see older versions might help explain the why and how of Lloyd's axiomatic self-consistency condition.

3.3.2 *Holographic computation as the interaction of past and future*

A problem which Richard Feynman called one of the last remaining problems of classical physics is the self-interaction of the electron, that is, the electromagnetic field influences the electron, which influences the field, which influences the electron; an infinite regress, which is profoundly difficult to deal with in calculations or even in concept (Feynman 1965). Wheeler and Feynman came up with a radical approach already in 1945 in the context of classical radiation theory, called the absorber theory of radiation, in which an emitter cannot radiate without coordinating with an absorber (Wheeler and Feynman 1945). The nature of this coordination is the radical point, which entails that both the emitter and absorber emit waves in both directions of time, these waves then interfere and cancel out in the emitter's past and the absorber's future, but reinforce in between emitter and absorber to generate an exclusively future-going time-forward wave from emitter to absorber. A very elegant concept, which also came with the experimental prediction that absorbers would display a form of pre-acceleration in anticipation of the incoming wave, they called this a witness of the interaction of past and future; something which we could potentially look for in our own experiments, but which without an exact calculation of the time scale involved might be difficult to evaluate; rather we can bring this concept to bear on neuroscience in a broader (if not too serious) sense with surprising ease: One of the classic mainstream results of

neuroscience, which have also had heavy influence on popular philosophy, is that motor neurons were found to respond before an actual conscious decision to move had been made (Libet, Gleason et al. 1983); this was later taken to imply that free will is an illusion (Wegner 2002), a point of enormous philosophical importance, but which in Wheeler and Feynman's framework might have a completely different interpretation as the aforementioned witness. In addition, in the framework we are building here it also has enormous practical importance, because we are beginning to see that with the paradox free interaction of past and future comes the power to solve intractable problems even with classical means. These same ideas of past and future interacting have also been generalized to the quantum case; firstly in Cramer's Transactional Interpretation of quantum mechanics (Cramer 1986), later refined into the Possibilist Interpretation (Kastner 2012), which again echoes the idea of the Novikov self-consistency principle that only the possible happens, even under or especially under paradoxical conditions.

Paul Werbos, who was highly instrumental in ending the debate whether neural networks could actually learn complex representations, by inventing the first feasible training algorithm, known as back-propagation (Werbos 1974), later extended to back-propagation through time for recurrent networks (Werbos 1990), has done impressive work towards maturing these ideas conceptually as well as mathematically and experimentally (Werbos 2000). The structured understanding of the impact of fundamental axioms on interpretations of quantum mechanics is due to this work (Werbos 2008), and Werbos also states that what he calls the Backwards-Time Interpretation provides a solution to a fundamental computational problem known as the closure of turbulence (Werbos 2002), which we have already encountered in the electron self-interaction problem, and which we will encounter again later as it is intimately related to other intractable problems. He also notes that the problem was in fact a mathematical illusion, much like our position on the NP versus P problem mentioned in the introduction.

Although not everyone agree that the interaction of past and future really provides a way to achieve tractable computing (Bennett, Leung et al. 2009), we can see that there are really not many other options, especially with this many positive indications. We suspect that the disagreements are essentially down to prior expectations of whether or

not tractability is achievable in any way at all, but this of course says nothing about which position is ultimately correct.

Finally, in order to tie the concept of interaction of past and future up with the principles of holography we must consider a final generalization of holographic processes, the generalization to time, known as dynamic holography. The key ingredient of dynamic holography is exactly the dynamic interaction of two opposing waves traveling in opposite directions of time (Zheng, Liu et al. 2005); implemented in optical setups via a so-called phase-conjugate mirror, which not merely reflects, but time-reverses incoming waves (Feldman, Bigio et al. 1982; Lehar 2010). In Bohm's terminology this dynamic process is described as continuous *enfolding* and *unfolding* of reality between, respectively, a holographic dual space and a real space; a process very similar to the rotation of a qubit as we will see later. This is of course only one possible description of a difficult to describe concept.

3.4 Holographic processes in practice *in silico*

So far we have considered aspects of holography which are mostly theoretical, but since we are taking a practical experimental approach we also need to consider where we are in terms of practical implementations.

3.4.1 *Holographic neural networks via circular convolution*

The concept of reduced representation, which we encountered in the introduction (section 1.2.3), reflects a deep insight into distributed representations: Distributed representations represent concepts (or information in general) as patterns across a large number of abstract neurons; many concepts can be represented in the same set of neurons by different patterns, and all neurons can participate in many patterns or concepts. Representing two concepts at the same time in a fixed set of neurons produces a combined pattern that is similar to each individual pattern, but reduced in a certain sense compared to both full, separate patterns; this is the idea of a reduced representation.

The main strength of the reduced representation is that unlike an arbitrary abstract representation, like a data pointer in computer programming, the reduced representation is similar to the full representation and can therefore be used directly in information

processing without retrieving the full representation. Hinton described this reduced processing as intuitive inference, while continued processing to retrieve the full representations reflects rational inference (Hinton 1990).

However, at the time of conception these ideas were mainly theoretical, and it was not known how to actually implement a reduced representation in a practical way. This problem was solved by Hinton's student, Tony Plate who introduced holographic reduced representation (Plate 1991).

The holographic reduced representation simply binds two patterns together by their circular convolution, producing a pattern of the same length thus reducing the storage requirement by half. Applying either full pattern by circular deconvolution to the reduced representation retrieves the other full signal. The circular convolution and deconvolution process is consequently a mathematical analogy to the holographic process, not only at the superficial level, but also in terms of the relevant equations (Rabal 2001). Since the length of the patterns or rather the number of neurons involved is fixed, it can be seen that the binding process can be repeated over and over with different elements, and in this way complete cognitive structures, concept hierarchies for instance, can be build up (though still ultimately limited by the memory size).

Much of the power of holographic reduced representations can be understood by considering the task of comparing two arbitrary hierarchical structures: Normally we would need to compare each element of the hierarchies pairwise at the level of the elements' full representations, and then sum up an overall similarity score. Holographic reduced representations allow a much faster similarity estimate by simply comparing the reduced representations, the process Hinton referred to as intuitive inference. So a very large number of hierarchies can be compared very quickly by intuitive inference, and then the best candidates can be expanded to less reduced representations for a better similarity estimate if necessary. The ability to process structures, like hierarchies, puts the connectionist holographic reduced representation systems on a level with symbolic systems: It was originally claimed that connectionist systems in general could not achieve a high level of systematicity (Fodor and Pylyshyn 1988; Fodor and McLaughlin 1990; Niklasson and Gelder 1994), i.e. the systematic processing of symbolic structures (in the connectionist case, without actual symbols). It was however predicted by (Eliasmith 1997) and finally shown by Jane Neumann that holographic reduced

representations (and some other distributed representations) are able to achieve systematicity level five (Neumann 2002), which means generalization over unseen structures and even over structures of higher complexity than the training examples.

Plate has created both pure feed-forward networks and more complicated recurrent networks (Plate 1993). These networks are typically trained by back-propagation through time methods (Werbos 1990), but Neumann also showed that her networks could be trained equally well by a method she called one-shot learning, which essentially amounts to adding up and thresholding associations in the memory (Neumann 2001). This method is a huge simplification of training methods in general and its simplicity is perhaps more biologically plausible given that even individual neurons have potent combinatorial abilities to handle their inputs and thresholding to determine their output probabilities.

Holographic associative memories have been criticized for their lower storage capacity than other matrix memories, but this objection overlooks that such is the price of their excellent associative properties. A more appropriate measure is their generative capacity, the ability to generate new patterns within the same parameters as the example patterns, which has also been investigated by Plate and found to be excellent (Plate 2003).

A more serious concern is due to the fact that deconvolution is not well-defined, and is therefore not guaranteed to produce stable results. This led Plate to replace deconvolution with correlation, which is only possible for a certain kind of data, called noise-like data. Real world data is not usually noise-like and must therefore be mapped onto noise-like replacements. Since two noise-like representations cannot be similar in the same way as their real world counterparts a problem arises if we wish to compare noise-like representations themselves (not just structural comparisons). In other words, representations lose their intrinsic similarity when they are replaced by noise-like representations. Plate has suggested adding a conventional associative memory to learn a transformation between noise-like data and real data, but this could possibly counter-act the advantage of using a holographic associative memory in the first place, and it certainly goes against the principle of Occam's razor.

We have previously followed a suggestion to replace the inverse in the deconvolution step by a Moore-Penrose pseudo-inverse operation (Schönemann 1987). The problem with the conventional inverse is that dividing by very small numbers produces very big numbers, and this effectively amplifies noise in the deconvolution process. The Moore-Penrose pseudo-inverse on the other hand is a least-squares approximation to the inverse, and is the best possible compromise. With the pseudo-inverse, deconvolution of real data becomes stable, and there is no need for noise-like replacements. This has been demonstrated on automatic vowel recognition in an earlier project that also adapted Neumann's one-shot learning method to the new approach (Kjeldsen 2008).

This approach was however exceedingly difficult to generalize to higher dimensional data (such as images), as the conceptual complexity of the regularization in the pseudo-inverse grows considerably with each added dimension.

A very elegant solution to the deconvolution problem is however to change the fundamental algebraic conception so that deconvolution becomes well-defined; this is achieved in geometric algebra, as considered next.

3.4.2 *Geometric algebra*

Becoming acquainted with geometric algebra can be quite a shocking experience: The first lesson of geometric algebra is that the way we are used to multipling vectors together is wrong, or perhaps not wrong, but at least there is a much better way: Geometric algebra is distinguished from regular vector algebra by the rules of multiplication and addition, and by the generalization of the vector concept to a multi-vector concept; in geometric algebra the (geometric) product has two parts (added together), one is an inner product and the other is an outer product. These two parts represent different aspects of the geometric product with different, and in a sense opposite geometrical interpretations (Aerts, Czachor et al. 2009). By combining the inner and outer product the geometric product achieves properties that neither hold on their own. The geometric algebra is real-valued, but a concept subsuming the imaginary unit is easily constructed, and with the major advantage that it is formulated in general for spaces of any dimension, meaning that an algebraic expression generalizes from complex numbers to hypercomplex numbers, like quaternions. This is not normally the case; it usually becomes increasingly difficult to work in higher dimensional complex

spaces from the point of view of implementing the basic operations. The geometric product absorbs the complexity of such generalizations, for example concepts like qubits and higher dimensional qudits can be formulated in geometric algebra from basic operations; for this reason it has been called quantum computation without quantum mechanics (Aerts and Czachor 2007), because the algebra remains real. An additional advantage is that a problem with complex numbers, or specifically with the imaginary unit, namely that we do not always know how to interpret it, disappears; in geometric algebra we only have to deal with an interpretation of geometry, which is usually intuitive (Gull, Lasenby et al. 1993b).

The number of basic geometric algebra steps needed to solve intractable problems is not an exponential of the problem size (Schott and Staples 2010); however, this does not directly allow one to solve intractable problems in practice, because the basic operation of the geometric product, at least in the obvious implementations, is itself intractable, since it is an all-to-all operation with each element of a multi-vector, and because the number of possible multi-vectors is set by the dimension of the algebra, which is reflected in the size of individual multi-vector elements. So the intractable problem has merely been moved to a deeper level in geometric algebra in this sense, but there might be some advantages in having the problem focused in the single point of the geometric product.

If we can come up with a tractable version of the geometric product, or at least operate only within a set of conditions under which the product is tractable, we can solve intractable problems in practice with geometrical algebra. One suggestion in this direction is a version of the geometric product referred to as the projected product, which came about in the context of reading out solutions from a memory multi-vector of holographic reduced representations (Aerts, Czachor et al. 2009). The idea is that some questions have answers that can only exist in a limited subspace of the full space of the algebra; this prior knowledge is used to only carry out the geometric product operations that end up in this subspace. The problem then is knowing in advance and in general which product elements end up in the right subspace, without spending more resources finding out than it takes to simply do the full product and afterwards select out the right subspace.

3.4.3 *Holographic neural networks in geometric algebra*

The geometric algebra version of holographic reduced representations replaces convolutions and deconvolutions with the geometric product (Patyk 2010), and therefore the implementation of the geometric product is the main aspect of an implementation of holographic reduced representations in geometric algebra. Within the framework of distributed representations holographic reduced representations have an equivalent binary representation called binary spatter codes (Kanerva 1994); in this formalism binding is achieved by binary XOR. The binary spatter codes allow a convenient implementation of the geometric product by representing the presence or absence of geometric so-called basis-blades as a binary code and implementing the geometric product as binary XOR (with a certain sign-correcting factor) (Aerts, Czachor et al. 2006). This implementation of the geometric product is called a bitmap representation on an additive basis.

Distributed representations have also been based on tensor products (Smolensky 1990), and holographic reduced representations can be considered a lossy compression of tensor product representations; the geometric product version lies in between these two extremes in terms of the operations and memory size required (Aerts, Czachor et al. 2009): The convolution based implementation uses a fixed memory size, while the geometric product causes the memory to grow, although not as rapidly as the tensor product.

Perhaps the easiest way to implement the geometric product is to express it with matrix multiplications: A matrix representation, called Cartan's representation, of the geometric product can be built from a series of so-called Pauli matrices (which are also fundamental building blocks of quantum mechanics) (Dorst, Fontijne et al. 2009), but the representation is very sparse and the size of the required matrices grows as 2^N, where N for holographic reduced representations outside toy-domains is typically in the order of thousands. This means that the matrix representation is very inefficient, although insights might be gained by directly inspecting the matrix representations visually.

It is clear that the practical implementations of holographic neural networks have not yet caught up with the main theoretical aspects outlined above, but it is important to notice that once the step to geometric algebra is taken the lines between holographic

neural networks and the quantum formalism become severely blurred by several mathematical similarities and identities. It is the hope that considering it from both sides will lead to new insights about both, and it is particularly encouraging that geometric algebra already reveals that quantum structures and processes can be achieved without actual quantum mechanics represented by the imaginary unit.

3.5 Computational aims

The computational aims of this thesis in the first instance revolve around implementing a framework based on holographic neural networks in geometric algebra, which can serve as a substrate for investigating the relations between holographic and quantum computing, especially relating to the question of tractability. In the second instance it is the hope that this will prepare us to take up any hints or clues found in the electrophysiological approach. We outline the first instance aims below:

3.5.1 *Holographic neural networks in geometric algebra*

We will implement a holographic neural network in geometric algebra, and investigate basic computational properties; the concept of the projected-product is of particular interest as a candidate for a tractable version of the full geometric product. We will attempt to relate this concept to the quantum formalism where such measurement concepts are perhaps more developed.

3.5.2 *Cartan's matrix embedding of holographic neural networks*

We will visually examine holographic memory structures in Cartan's matrix embedding of geometric algebra, with the purpose of identifying patterns which could suggest paths to tractable geometric products in the optimized (non-embedded) version.

3.5.3 *Quantum neural network memory structure*

We also aim to investigate properties of simulated quantum neural networks, which might reveal differences or similarities with the holographic counterpart. The hope is to understand why quantum neural networks are intractable in simulation, and to identify potentially important differences in the holographic geometric algebra formalism.

3.5.4 *Qubit structure in geometric algebra*

As a secondary approach to the previous aim we will implement qubit structures in geometric algebra to allow easier comparison with traditional quantum computing, and to facilitate discussion of comparisons with holographic methods within the established terminology of quantum computing with concepts like superposition, entanglement, measurement and teleportation.

Chapter 4. Methods

general electrophysiological and computational methods

This methods chapter provides the most general methods applied in the experiments and analysis reported in the following results chapters. For easier reference the more specific methods are however given in the respective results chapters.

4.1 Animal and slice preparation

All animals were young adult male Wistar rats weighing in excess of 150g obtained from Charles River Laboratories Inc. and housed in the Comparative Biology Centre at Newcastle University. Animals were allowed to acclimatise for at least two days prior to experiments, and were on a light-dark cycle with light 8am-8pm, with food and water *ad libitum*, and environmental enrichment in the form of cardboard boxes.
All procedures were conducted in accordance with the UK Animals (Scientific Procedures) Act, of 1986, under personal and project licenses granted by the UK Home Office.

Animals were placed in a bell jar to induce light anaesthesia by exposure to 3 ml of an inhalant anaesthetic, isoflourane (Abbott Laboratories., Kent, UK).
An intramuscular injection of ketamine ($\geq$100mg/kg^{-1}, Pfizer Ltd., Kent, UK) and xylazine ($\geq$10 mg/kg^{-1}, Millpledge Veterinary, Retford, UK), was then administered in the gluteal area of the hind leg. It was confirmed that all tail pinch, pedal withdrawal and corneal reflexes were abolished. The thoracic cavity was cut open and the rib cage partly removed in order to gain access to the heart. A catheter was inserted into the left ventricle of the heart followed by a small incision in the right atrium to allow for manual intracardial perfusion with approximately 50 ml of a solution of cold, oxygenated, sucrose artificial cerebrospinal fluid (saCSF, see below). An incision was made along the midline of the head and neck of the animal, and the head was separated from the neck with scissors. The skull was split along the sagittal suture accessed from the severed spinal column, allowing for the skull and dura to be peeled away. The brain was excised and placed in a petri-dish of cold, oxygenated saCSF. The frontal cortex, brain stem and cerebellum were cut away, and the remaining tissue was glued to a chuck of a Leica VT1000 vibratome (Leica Microsystems, Nussloch GmbH, Germany)

on the caudal surface of the brain. The chuck was placed into the cutting chamber filled with cold oxygenated saCSF.

Coronal sections were cut at a thickness of 450 µm, and transferred to a petri-dish filled with cold oxygenated saCSF. Three slices per hemisphere around the relevant area were usually obtained; this area was located using visual landmarks with reference to the rat brain atlas (Paxinos and Watson 1998). Slices were lightly trimmed down with a scalpel to allow easier handling and better perfusion in the recording chamber.

Coronal sections were transferred to a holding chamber to allow to stabilize at room temperature for thirty minutes in oxygenated normal artificial cerebrospinal fluid (aCSF, see below). Two slices were then placed in a recording chamber on a double layer of lens cleaning tissue and maintained at the interface between warm humidified carbogen gas (95% O_2/5% CO_2) and normal aCSF, and perfused by circulating oxygenated normal aCSF through the chamber using a peristaltic pump at a rate of about 1.2 ml/minute at 34°C using a flow heater. Slices were left to stabilize in the recording chamber for thirty minutes before recording or drug application.
Recipes for aCSF and saCSF are given below:
aCSF (in mM): 126 NaCl, 3 KCl, 1.25 NaH_2PO_4, 24 $NaHCO_3$, 1.2 $MgSO_4$, 1.6 $CaCl_2$, and 10 glucose.
saCSF (in mM): 252 sucrose, 3 KCl, 1.25 NaH_2PO_4, 24 $NaHCO_3$, 2 $MgSO_4$, 2 $CaCl_2$, and 10 glucose.

4.2 Pharmacology

Below we list the recipes for the relevant pharmacological manipulations.

4.2.1 *Gamma wake-state model*

The persistent gamma wake (active attending) model state was induced by bath application of 400 nM of the kainate receptor agonist, kainic acid ((2*S*,3*S*,4-*R*)-carboxy-4-(1methylethenyl)-3-pyrrolidineacetic acid)) (kainate from here on).

4.2.2 *Delta sleep-state model*

The delta sleep model state was induced by bath application of low cholinergic drive with 4 μM of the non-hydrolysable acetylcholine analogue carbachol, and low dopaminergic tone by blocking dopamine D1 receptors with 10 μM SCH23390.

4.2.3 *Sensory stimulation model*

Stimulation was 1mM glutamate in aCSF applied as a micro drop of approximately 50-150 nl with a blunted glass electrode of type GC120TF (Harvard Apparatus). See section 5.2.3 for details of application.

4.3 Electrophysiology

The electrophysiology carried out in this thesis was primarily multi-channel recordings, but also single channel glass electrodes and paired glass electrodes were used as described below.

4.3.1 *Glass electrodes*

Extracellular local field potentials were recorded using glass micro-electrodes pulled from thin-walled borosilicate glass capillaries of type GC120TF (Harvard Apparatus), mounted in electrode holders with a silver wire connector inside the glass electrode, which was filled with normal aCSF. Estimated resistance was less than 5 MOhm. All glass electrode experiments presented were done with paired electrodes in layer 2/3 and layer 5 in associational neocortex, aligned on the same line perpendicular to the curvature of the cortex. All recordings were 60 seconds long and sampled at 5 kHz.

4.3.2 *Utah array*

Micro-electrode array recordings of both local field potentials and unit activity were done with silicon electrode arrays, from here on called Utah arrays. The arrays were square, 10x10 electrodes with a separation of 0.4 mm in each direction, and a shank length of 1.2 mm, and electrode impedances ranging from 230-370 kOhm. Electrode arrays were super-glued onto a machined 4mm diameter, 8cm length plastic shank and mounted on a 3D patch manipulator.

The unfiltered signal sampled at 30 kHz was used for spike detection in the Matlab software Wave_clus (Quiroga, Nadasdy et al. 2004) (with the recorded unit activity

given by the recording software only used as reference) for a more uniform approach between experiments. In Wave_clus spike detection thresholds (negative going) were set at five times the estimated standard deviation of the noise, and units were defined by superparamagnetic clustering; clusters which did not display reasonable spike shapes and inter-spike-intervals (particularly those corresponding to 50Hz noise) were manually rejected.

4.4 Complex continuous cross wavelet semblance analysis

We implemented a wavelet analysis technique for pairwise correlations of local field potentials and in extension for functional connectivity estimates (Lee, Harrison et al. 2003; Fingelkurts, Fingelkurts et al. 2005; Li, Guo et al. 2009). The wavelet method called wavelet semblance (Cooper and Cowan 2008) or the wavelet cross-spectrum (Grinsted, Moore et al. 2004) has several advantages over time-series correlations and Fourier based measures: In general wavelet analysis offers more flexible trade-offs between time and frequency resolution, and time-frequency maps are the natural outcome without special extensions like short-time Fourier transforms or multi-taper methods. Also, wavelet analysis does not assume stationarity, and signal amplitude and phase can easily be considered separately, which gives more flexible measures of correlation. Wavelet analysis can also be applied to spike data (Nedungadi, Rangarajan et al. 2009; Makarov, Pavlov et al. 2010), but the process requires high time-resolution, and is therefore very computationally intensive. We return to this point later when we will consider spike-train-to-local-field-potential coherence specific to chapter 5.

In our efforts to optimise and port the wavelet analysis (used for post hoc analysis of all the data presented in this thesis) to the graphics processing unit (GPU) we realized that the wavelet transform is essentially convolution with wavelet shapes at many different scales. Consequently the process can possibly be done much faster in Fourier space where (circular) convolution is element-wise multiplication. Following this idea we found an existing implementation (Torrence and Compo 1998), and since FFT parallelizes well and is supported on the GPU in Matlab it was simple to achieve very significant speed-ups. The Fourier space version of the wavelet transform has since been included in the latest Matlab version (2012b), but does not support complex wavelets, which are necessary to take full advantage of the transform. It does also not support GPU.

4.5 Computational modelling

The computational approach of this thesis does not operate at the level of equivalent circuit models or other detailed modelling of the neurobiological operations of neurons, to then build up to some functional, observational or behavioural aspect. Instead our approach is in a sense the opposite; we start with the most abstract properties of an algebra as the formal mathematical foundation of anything a detailed model can achieve. It is the algebra itself which possesses the features of an abstract model of cognition, and the details of how the algebra is implemented is considered to be multiple realizable as long as the algebraic properties are fulfilled. Many cognitive abilities can for example be characterized as inverse problems albeit, in general, difficult ill-posed and ill-conditioned inverse problems (Vito, Rosasco et al. 2005). These, are also commonly encountered in physics and engineering, where a very wide range of approximation approaches are known, but without general solutions. Inversion is a basic algebraic operation, and for this and other reasons we changed the basic algebraic language to a little-known formulation, which not only unifies the range of different approaches given in standard algebra (Lasenby, Lasenby et al. 2000), but also solves inverse problem is a very elegant and natural way. A brief introduction to geometric algebra was given in section 3.4.2. We give the detailed methods of our implementations in chapter 7.

Chapter 5. Results

verification and extension of the in vitro wake-sleep model

5.1 Introduction and aims

This first results chapter presents initial efforts to verify that our *in vitro* delta rhythm model was an adequate substrate for further investigation of sleep-associated synaptic rescaling processes. For this purpose we developed an extensive experimental paradigm to mimic and model some of the primary features of the wake-sleep cycle, i.e. during a persistent gamma model of the wake-state (active attending) we introduced a stimulation model of sensory input to induce synaptic potentiation, and subsequently switched to the delta model of the sleep-state (non-REM) to form one whole evolution of the wake-sleep cycle model. The validity of all these models was thus under scrutiny, and was evaluated by their combined ability to reproduce two key aspects of synaptic rescaling processes found in both *in vivo* and human studies as detailed in chapter 2. The experiments presented in this chapter were thus designed with two main aims in mind as specified in the following subsections.

5.1.1 *Delta power and prior stimulation*

The first and foremost aim was to attempt to extent the *in vitro* delta model to the finding that prior stimulation of local cortical areas produces stronger delta power in the same local areas during subsequent sleep, as seen in a number of studies *in vivo* and in human subjects (see section 2.2.1 for details). We further aimed to investigate whether the presence of a persistent gamma rhythm, as a model of wakefulness, enhanced the effect of stimulation.

5.1.2 *Reactivation and replay*

The second aim extended our inquiry into the domain of unit spiking, and set out to investigate whether units that responded to stimulation are reactivated, deactivated or unchanged during subsequent delta; and if so, to examine the characteristics of the respective unit types. We considered reactivation a limited form of replay, since it is a prerequisite of replay, but does not in itself imply any particular neural code, which was

anyway unknown in our case, since we were not dealing with place cells like most previous studies as discussed in chapter 2.

Figure 5.1: Overview of experimental paradigm

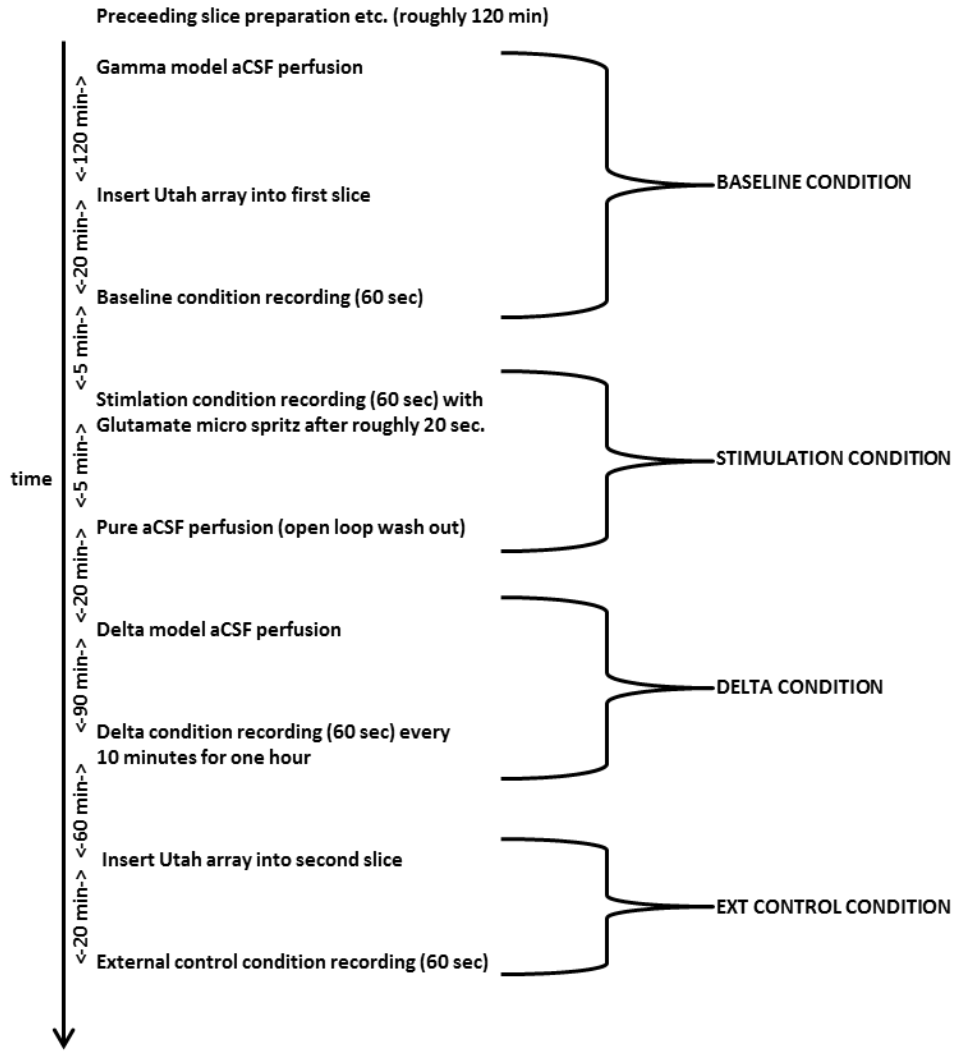


Figure 5.1: Overview of experimental paradigm

The experimental paradigm consisted of four distinct conditions: The baseline condition was defined by gamma model perfusion for 2 hours, then the insertion of the Utah array with 20 minutes recovery time, and culminated in a 60 second recording of baseline activity. Figure 5.2 shows an example of a baseline condition recording. The stimulation condition followed immediately after the baseline condition, as soon as micro spritzing was ready, typically after 5 minutes. The spritzing of one micro drop (50-150 nl) of aCSF containing 1mM Glutamate was done approximately 20 seconds into the stimulation condition recording. Figure 5.3 shows an example of a stimulation condition recording. The gamma model was then washed out until 25ml pure aCSF had run through. The delta condition then started with 90 minutes of perfusion with the delta model aCSF, and continued for another hour with 60 second recordings every 10 minutes. Figure 5.4 shows an example of a delta condition recording. When the Utah array was taken out the temporary marks left in the slice were immediately micro-photographed for post hoc electrode localization. Finally the Utah array was inserted into the secondary unstimulated slice, and after 20 minutes recovery time a final recording was taken as the external control condition. This slice was also micro-photographed.

5.2 Methods

The methods described here are specific to this chapter; for general methods see chapter 4. We start with an overview of the experimental paradigm, and then consider each experimental condition and the associated methods in more detail. We will also give a short introduction to the various statistical tests used throughout.

5.2.1 *Overview of the experimental paradigm*

An overview of the experimental paradigm is shown in Figure 5.1, and is further described in steps as follows:

Following slice preparation as described in chapter 4, two slices were placed into the recording chamber in the presence of gamma model aCSF perfusion. We then waited two hours for the model to take effect. Then we inserted the Utah array into one slice and allowed 20 minutes for the slice to recover. Initially did one 60 second recording as the baseline condition. Figure 5.2 shows an example of a baseline condition recording, and further details are given in the next subsection. Then did one 60 second stimulation condition recording during which (after approximately 20 second) a stimulation model aCSF micro drop spritz was introduced in layer 4 at the edge of the Utah array proximal to primary areas. Figure 5.3 shows an example of a stimulation condition recording, and further details are given two subsections hence. Washed out gamma model by breaking the perfusion loop, so that runoff was discarded and continually replaced by pure aCSF, closed the loop again when 25ml of pure aCSF had run through. Switched to the delta model aCSF, and waited one and a half hour for the model to take effect. Recorded for 60 seconds every ten minutes for one hour as the delta condition. Figure 5.4 shows an example of a delta recording, and further details are given three subsections hence. Took out array, and micro-photographed array marks left in slice for post hoc electrode localization. Inserted array into second slice (that was not stimulated) and allowed 20 minutes for the slice to recover. Did one 60 second recording as an external control condition (the external controls were a precaution in case internal controls in the form of channels not responding to stimulation were not enough; this was not found to be the case). Took out array and micro-photograph array marks left in second slice.

5.2.2 *Baseline condition - gamma wake-state model*

To separately evaluate the effect of the gamma wake-state model we quantified the presence or absence of a gamma rhythm; obviously the gamma model itself was always

present, but as we were trying to model a wake-like gamma rhythm we set this as a reasonable criterion. A gamma rhythm was deemed to be present when there was a clearly dominant peak in the spectrum in the gamma band (25-100 Hz) after non-causal notch filtering at 50 Hz. The gamma model aCSF contained 400nM kainate. Incubation time was two hours with perfusion at a rate of approximately 1.2 ml per minute, with an additional 20 minutes recovery time after insertion of the Utah array.

5.2.3 *Stimulation condition - sensory input model*

To model sensory input we introduced pharmacological stimulation in layer 4 primary areas. Stimulation was 1mM Glutamate in aCSF applied as a micro drop spritz with a blunted glass electrode of type GC120TF (Harvard Apparatus). Using a micro-manipulator to briefly touch the stimulation electrode to the slice as close to the array as possible, one micro drop of approximately 50-150 nl was applied. Figure 5.6.A indicates approximate stimulation location in layer 4 at the top of the array proximal to primary areas.

Stimulation was deemed successful for each sorted unit when an abrupt increase (at least tripling over 1-2 seconds) in firing rate could be identified in the firing histograms with one second bins (as seen in Figure 5.5, top red unit), and was further quantified by its peak firing rate as the maximal histogram firing rate after response onset (also indicated in Figure 5.5.

5.2.4 *Delta condition - delta sleep-state model*

The delta sleep model state was induced by perfusion with aCSF containing low cholinergic drive with 4 μM of the non-hydrolysable acetylcholine analogue carbachol, and low dopaminergic tone by blocking dopamine D1 receptors with 10 μM SCH23390.

Incubation time was 1.5 hours, followed by repeated recordings every 10 minutes for an hour; of these six recordings the recording with strongest delta power was chosen for further analysis. Delta power was determined at a fundamental peak in the signal spectrum in the delta band (1-4 Hz). The analysis of the putative increase in delta power following stimulation was carried out by comparing channels (i.e. locations) that showed spiking response to stimulation (as described above) with those channels that

did not. The potentially confounding factor of location was controlled for as described below in 5.2.7.

An alternative approach would have been to use the external control slices that were not stimulated at all, but due to large variability in delta power between slices in general additional confounding factors would have been more difficult to control for.

Figure 5.2: Example of baseline condition recording

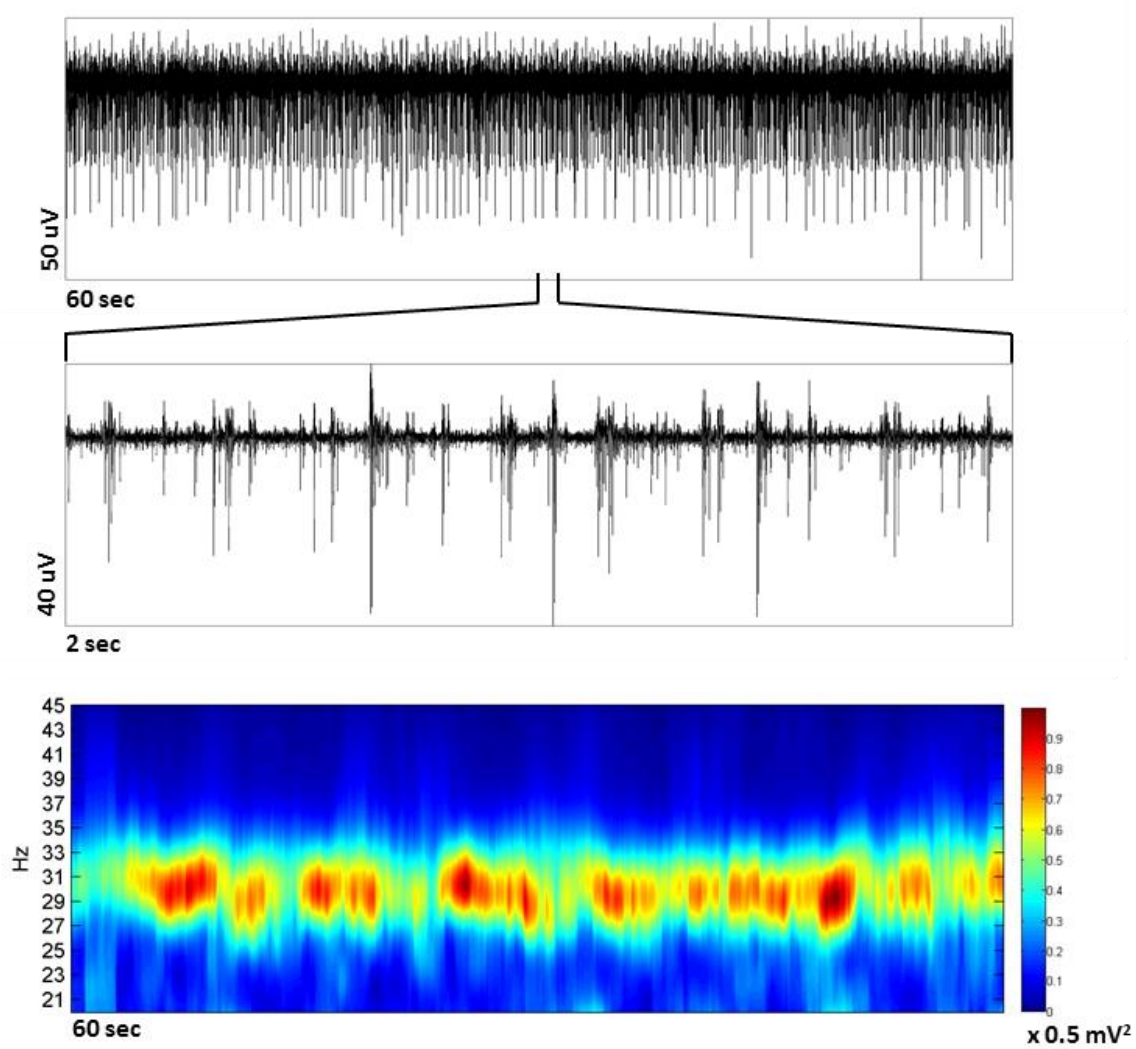


Figure 5.2: Example of baseline condition recording

The top trace shows an example of a baseline condition recording of 60 seconds at 30 kHz sample rate. At least one spike size is clearly visible in the full trace, and zooming in on 2 seconds of data, shown in the middle trace, reveals 2 or 3 additional spike sizes that were later thresholded and sorted with the Wave_clus Matlab software. The bottom complex continuous wavelet scalogram reveals an on-going gamma oscillation at around 30 Hz. The presence of a gamma oscillation was an experimental factor, but not an experimental requirement, so this example is special in this sense, and should not be taken to indicate that all experiments or channels displayed clear gamma oscillations. A gamma oscillation was determined to be present when there was a clear peak in the power spectrum in the gamma band (25-100 Hz) after non-causal notch filtering at 50 Hz.

Figure 5.3: Example of stimulation condition recording

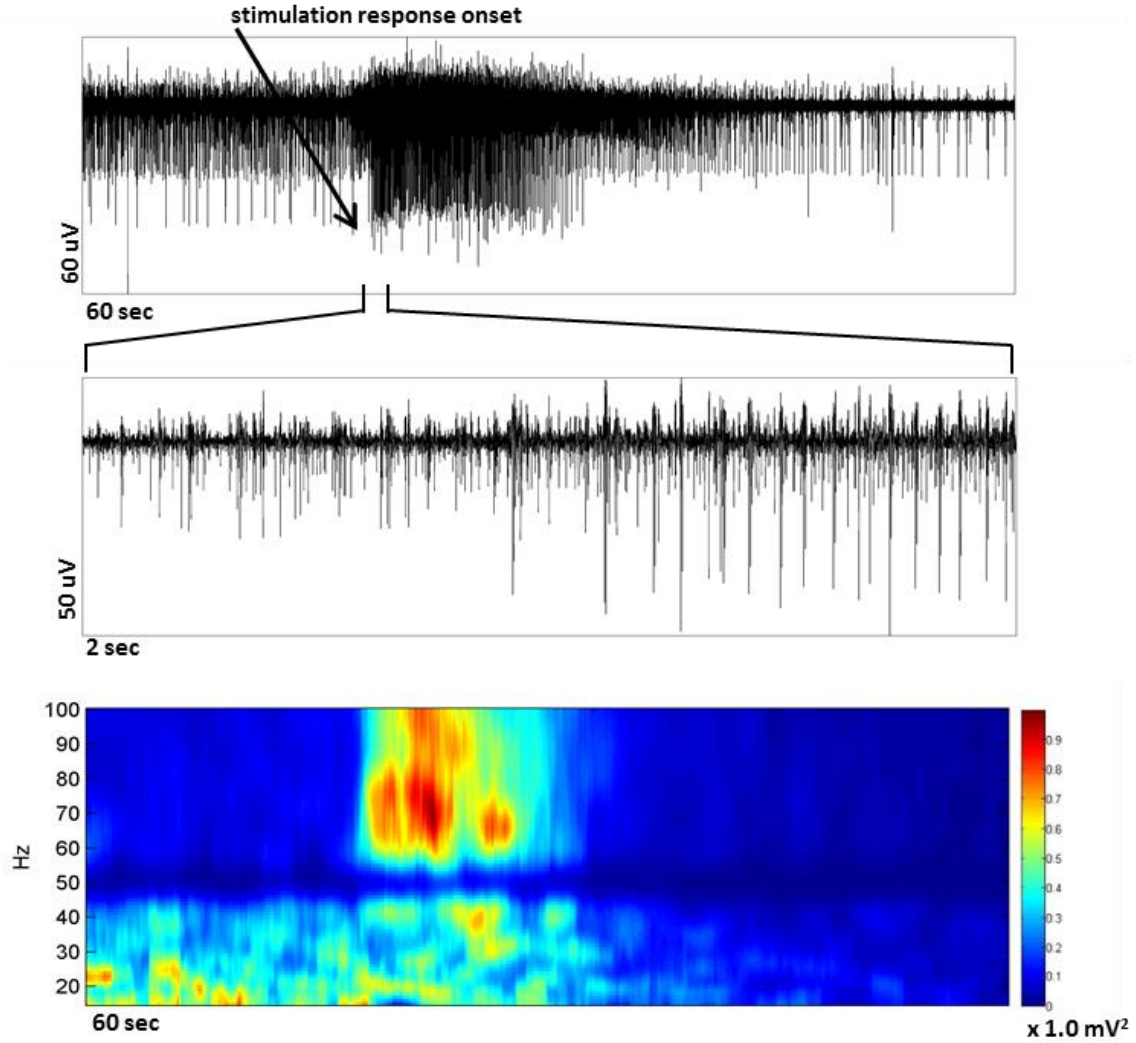


Figure 5.3: Example of stimulation condition recording

The top trace shows an example of a stimulation condition recording of 60 seconds at 30 kHz sample rate. About 20 seconds into the recording the stimulation response onset is clearly visible as condensed firing (indicated by the arrow), zooming in on a 2 second segment at the onset, shown in the middle trace, reveals increased spike sizes as well. The complex continuous wavelet scalogram shown at the bottom indicates a strong response in the gamma band (25-100 Hz) particularly in the high gamma band (60-100 Hz) in this case. 50 Hz noise was filtered out with a non-causal notch filter. In many examples the response power tapered off a bit more slowly starting at the higher frequencies. Response power as indicated by the scalogram was not used to quantify responses, instead the firing rates of individual units after sorting were found to be more reliable, and also allowed quantification on unit basis instead of on channel basis.

Figure 5.4: Example of delta condition recording

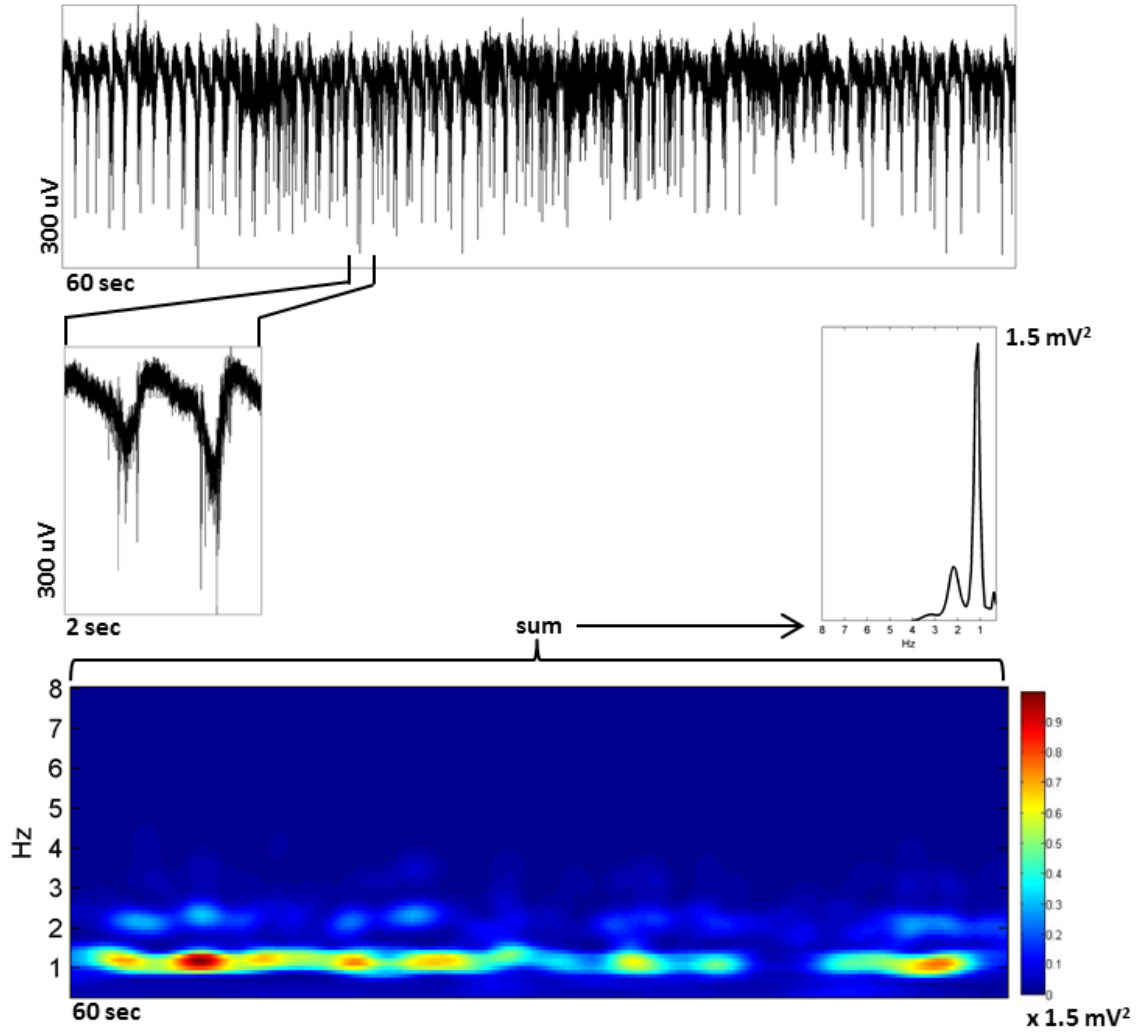


Figure 5.4: Example of delta condition recording

The top trace shows an example of a delta condition recording of 60 seconds at 30 kHz sample rate. The 2 second zoom-in shown in the middle-left trace confirms preferential firing in bursts on the trough of the delta period. The complex continuous wavelet scalogram shown at the bottom indicates a strong fundamental delta rhythm as well as a weaker harmonic, the summation over time into a wavelet spectrum (also on a linear scale) shown in the middle-right plot shows this more clearly. A weak harmonic was typical, but far from discernible in all cases. Delta power was determined as the peak in the power spectrum at the fundamental delta frequency.

5.2.5 *Spike sorting and classification*

The raw recordings of each condition (baseline, stimulation and delta) at 30 kHz sample rate were concatenated, thresholded and sorted manually with the Wave_clus Matlab software, which uses wavelet decomposition and superparamagnetic clustering in wavelet space of unit spike shapes. Units are the commonly used designation of extracellular measures of spikes in membrane potential of single or multiple coherent cells. Each unit was then given a set of binary classifications indicating if it responded to stimulation, and for each condition indicating if it was active during that condition; if a unit showed less than 20 spikes during a 60 second epoch of any condition it was deemed inactive in that condition. Figure 5.5 shows example units for one channel, and the corresponding classifications.

5.2.6 *Spike-train-to-local-field-potential phase semblance*

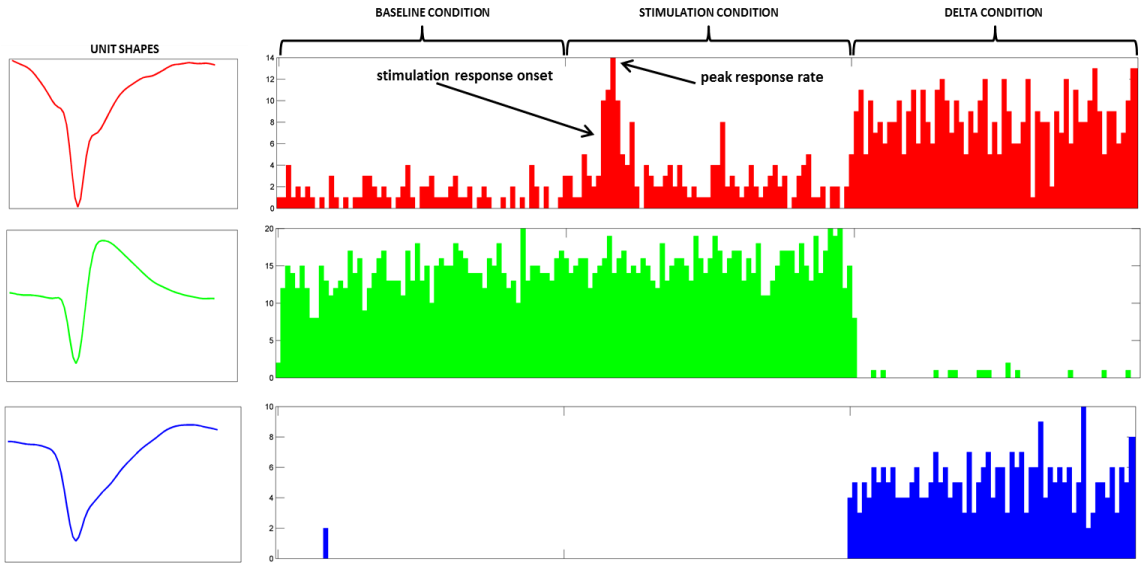
In order to evaluate the temporal alignment of spikes with the phase of local field potential oscillations during stimulation we generalized the complex continuous wavelet semblance approach detailed in chapter 4, to the case of spike-train-to-local-field-potential semblance, where the spike-train is represented as delta functions at each spike sorted spike timestamp. This approach had previously been verified in (Vinck, Battaglia et al. 2012), and is equivalent to considering individual spikes via spike-triggered averages under certain conditions (Vinck, Battaglia et al. 2012). By using a complex wavelet approach the phase and amplitude of the signals can be separated (Cooper and Cowan 2008), so that the phase-only semblance measure is independent of the spike rate, which forms the signal amplitude, and might have confounded our analysis when spike rates of compared groups were different.

5.2.7 *Statistics*

The delta power data under consideration was not normally distributed as determined by normality tests in SPSS; we therefore used the non-parametric Wilcoxon rank sum test and the Kruskal–Wallis test with multiple comparison provided by Matlab. We also used the McNemar difference of proportion test in SPSS to test differences between percentages of units that responded to stimulation and were active or not during delta, and to test percentage differences in unit shape groupings.

We further used generalized linear models in SPSS to evaluate a number of experimental factors. These models try to fit the variation in the outcome variable with a function of the included factors; analysis proceeds by first including all factors and their possible interactions, then reducing the model by removing insignificant factors until the quality of the model degrades. We only present reduced models. For the delta power model the normal distribution was used, which was adequately fulfilled by the log10-transformed data. The binomial distribution was used for the unit reactivation model.

Figure 5.5: Example of spike sorting and classification



Stimulation response	Active during baseline	Active during stimulation	Active during delta
1	1	1	1
0	1	1	0
0	0	0	1

Figure 5.5: Example of spike sorting and classification

Recordings from the three main experimental conditions were concatenated, so that the same units could be identified and followed across all three conditions. Stimulation responses were identified as an at least triple increase in firing rate over 1-2 seconds, and the peak response rate was quantified as the maximal firing rate after response onset. The figure shows an example of three different units from the same channel; the top red unit shows response to stimulation and was also active during the delta condition, the middle green unit was not active during delta (if less than 20 spikes in any condition these were deemed to be spurious events due to the limits of the clustering process), while the bottom blue unit was only active during delta. The response and presence of each unit was thus entered into a binary classification table as indicated by the bottom table, e.g. the top red unit reads: showed stimulation response and was active during all three conditions. Each condition was 60 seconds long, time bins of one second were used for this figure.

5.3 Results

We follow the logical order of first considering the stimulation responses, then their effect on local delta power, and finally whether responding units were reactivated or deactivated during delta, and what the unit characteristics were.

5.3.1 *Stimulation response*

Since it was known that delta power varies with layer (Carracedo, Kjeldsen et al. 2013) we needed not only to identify which channels showed responses, but also where they were located, so that the location could be controlled for. For this purpose we used post hoc localization by identifying the marks left by the array in micro-images of the slices, as indicated in the example in Figure 5.6.A. Figure 5.6.B and Figure 5.6.C show the trend of the spatial response patterns for ten cases, which could be easily aligned, respectively as absolute counts of responding units (i.e. units like the top red unit in Figure 5.5) for each electrode location, and as average peak response rates measured as indicated in Figure 5.5. In our analysis we controlled for layer and area differences by statistical modelling (see section 5.2.7) as well as by splitting the full data set by layers and areas.

5.3.2 *Delta power and prior stimulation response*

We were not only interested in the overall effect of stimulation on delta power, but also in any and all effects of layer and area, as the slices used contained a stimulated primary sensory area and an adjacent associational area. We introduced a range of data splits by layer and by area, and by both layer and area at the same time. Since the statistical power is reduced with every split, and since the best statistical method is not easily given, we compared several approaches. We considered an N of 19, where N is a slice, for a total of 566 channels in neocortex with clear delta rhythms. Figure 5.7 reflects the initial and most straight-forward statistical analysis of the raw delta power values log10 transformed to mitigate inter-slice variation. This analysis revealed an overall significant effect ($p=0.00197$) when disregarding layer 4 of stimulation response to increase delta power, but when spilt into layers and areas only layer 6 of associational cortex (i.e. the non-stimulated region) was significant ($p=0.0087$).

Figure 5.8 reflects the same analysis, but on values normalized by the slice norm to mitigate inter-slice variation, which enhanced the effect across deep associational areas

($p=8.6220e^{-04}$). Figure 5.9 again reflects the same analysis, but on delta power values standardized by subtracting the mean and dividing by the standard deviation within slice to mitigate inter-slice variation, which also agreed well with the previous analysis.
We then employed a generalized linear model to evaluate the effect of several additional factors on delta power; by including slice identity in the model we mitigated the inter-slice variability, and by including layer and area we controlled for layer and area effects. We also included the presence of gamma as a factor to evaluate if gamma enhanced the effect of stimulation. The results of this analysis are shown in Figure 5.10, and revealed agreement with the previous measures, as well as additional effects in some primary areas as shown. There was also an overall effect of gamma, but without any clear spatial pattern in the splits; however there was no effect of the interaction of gamma and stimulation response (and it was subsequently removed from the model), which was the real point of interest. A variation of this approach, where we instead nested each factor within the slice identity, the effect of which was to remove the assumption that factors are the same in each slice, e.g. layer 2 can be considered different in different slices (which might be justified in the sense that they would not necessarily have the same baseline potentiation), is shown in Figure 5.11. This analysis also agreed with the previous measures, and additionally enhanced effects in superficial layers, which was to be expected from the higher degrees of freedom of the model.

Overall all the approaches agreed on the overall effect of stimulation; when broken down into layers and areas there was some variation, but with a clear centre of gravity in deep associational areas.

Figure 5.6: Stimulation response localization and quantification

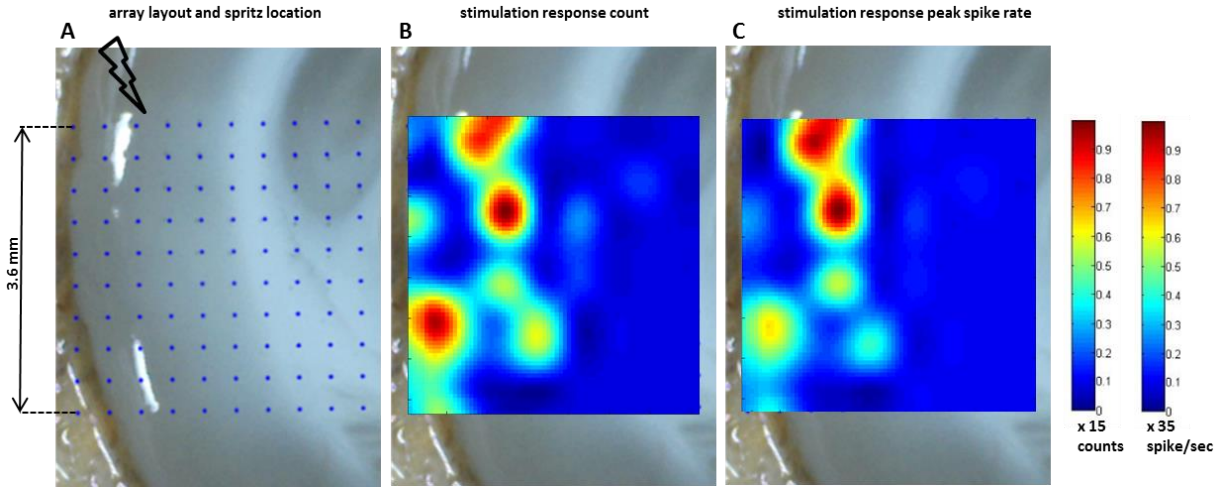


Figure 5.6: Stimulation response localization and quantification

A) Example of post hoc Utah array localization; dots indicate locations of marks left by the array, and the cartoon lightning indicates the approximate location of the stimulation spritz in layer 4 at the upper edge of the array proximal to primary areas. B) In ten cases the locations of the arrays were similar enough to be easily aligned to produce maps of the simulation response trends. The total counts of responding channels for each location are indicated in this map. C) The average peak response rate is indicated in this map. Although the previous map indicates that responses occur about as often in superficial layers distal to stimulation as in deep layers proximal to stimulation, this map reveals that the sizes of these responses are smaller as would be expected from a basic distance effect.

Figure 5.7: Effect of stimulation response on delta power #1

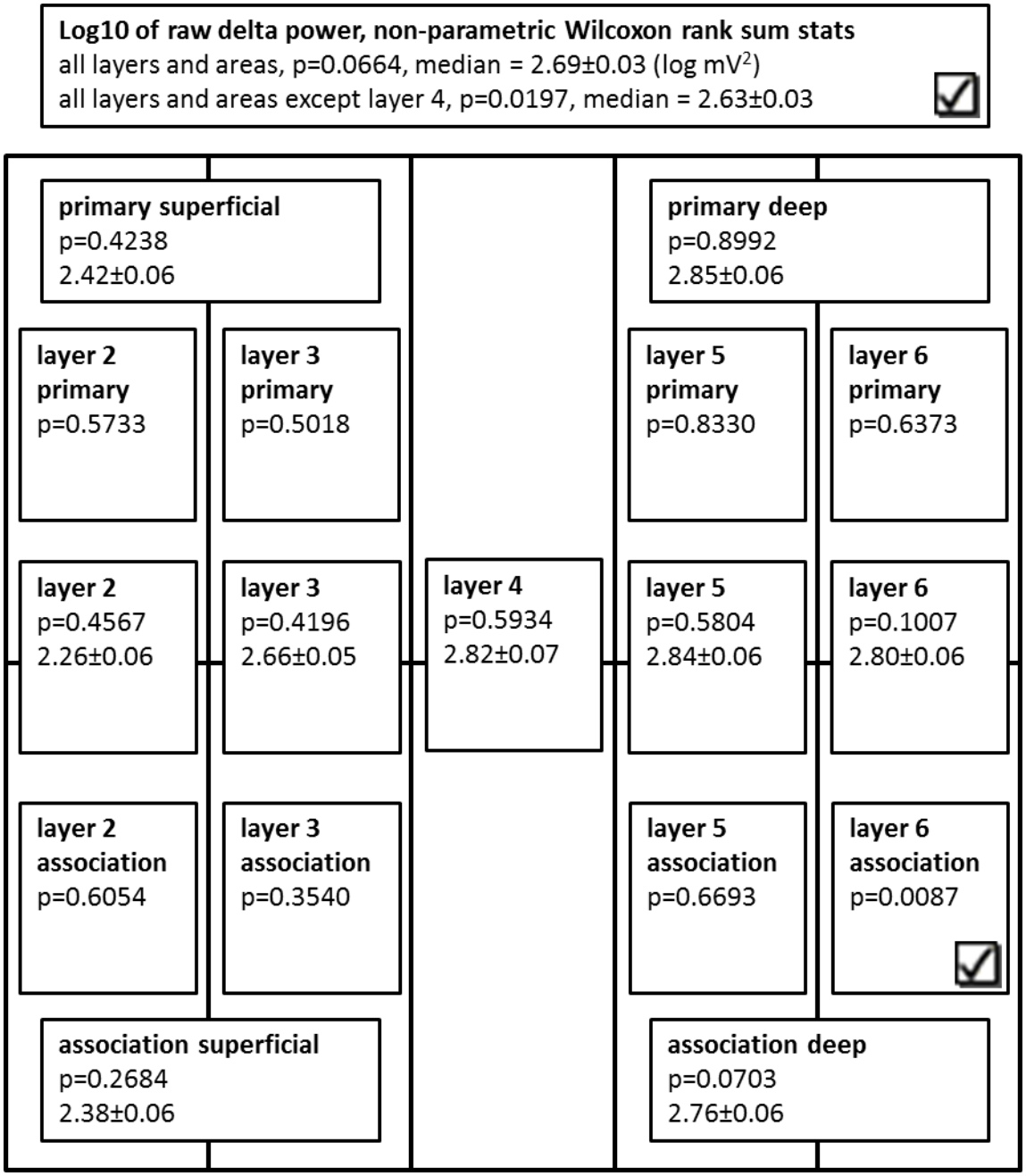


Figure 5.7: Effect of stimulation response on delta power #1

This is a comparison between the delta power on channels that responded to stimulation and channels that did not respond to stimulation. This is therefore a comparison of two groups. Here we compared the logarithm of the raw delta power values to mitigate the large variation, no other normalization was used. Overall there is a significant difference when discounting layer 4, but when broken down into cortical subareas only layer 6 association is significant. Check marks are shorthand for significance of the variable of interest.

We have also indicated median and standard error log power values for the main areas and layers across both groups.

Figure 5.8: Effect of stimulation response on delta power #2

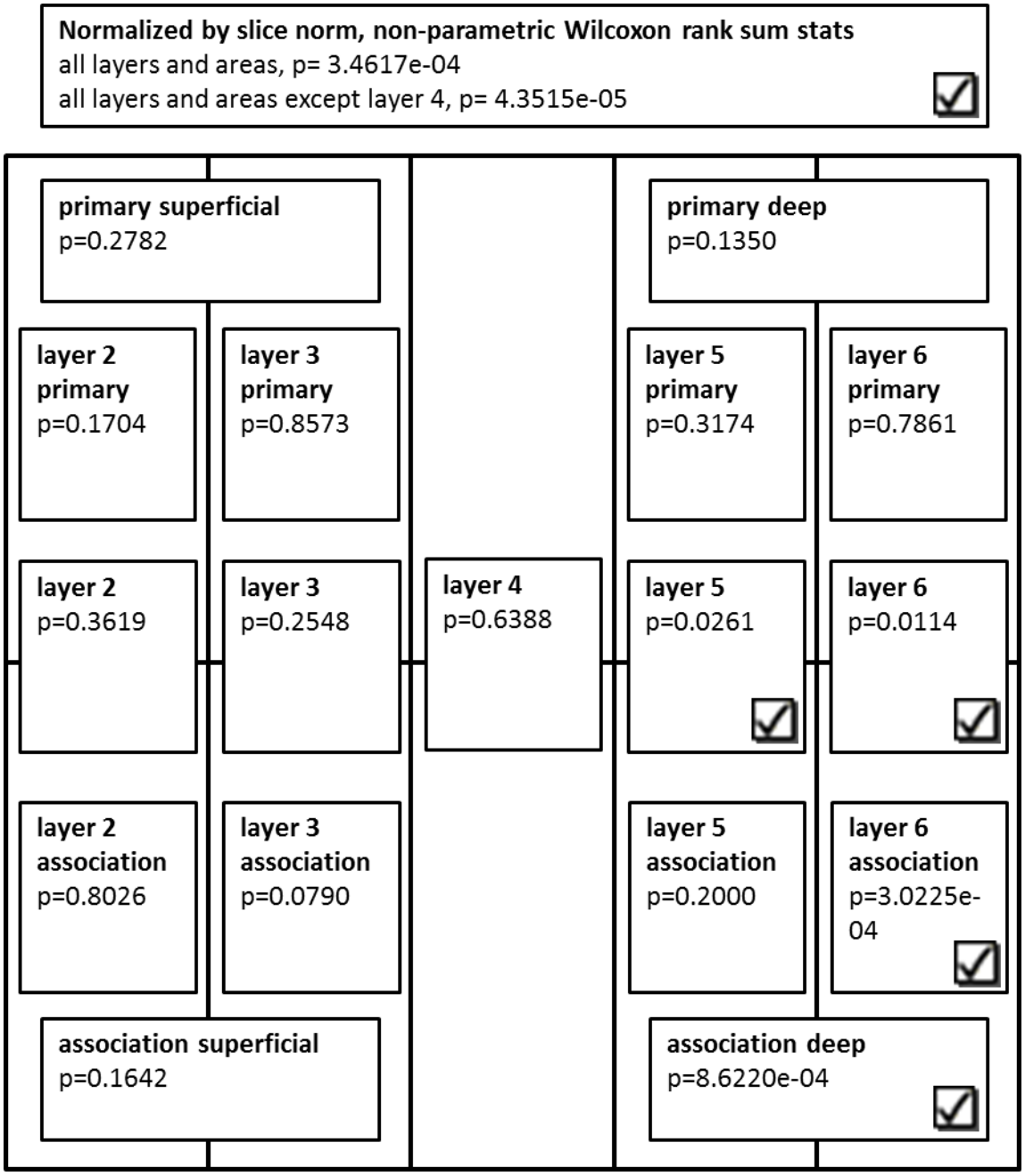


Figure 5.8: Effect of stimulation response on delta power #2

This is a comparison between the delta power on channels that responded to stimulation and channels that did not respond to stimulation. This is therefore a comparison of two groups. Here we compared delta power values that have been divided by the slice norm to mitigate the large variation. Overall there is a significant difference, but when broken down into cortical subareas only deep layers and association areas are significant. Check marks are shorthand for significance of the variable of interest.

Figure 5.9: Effect of stimulation response on delta power #3

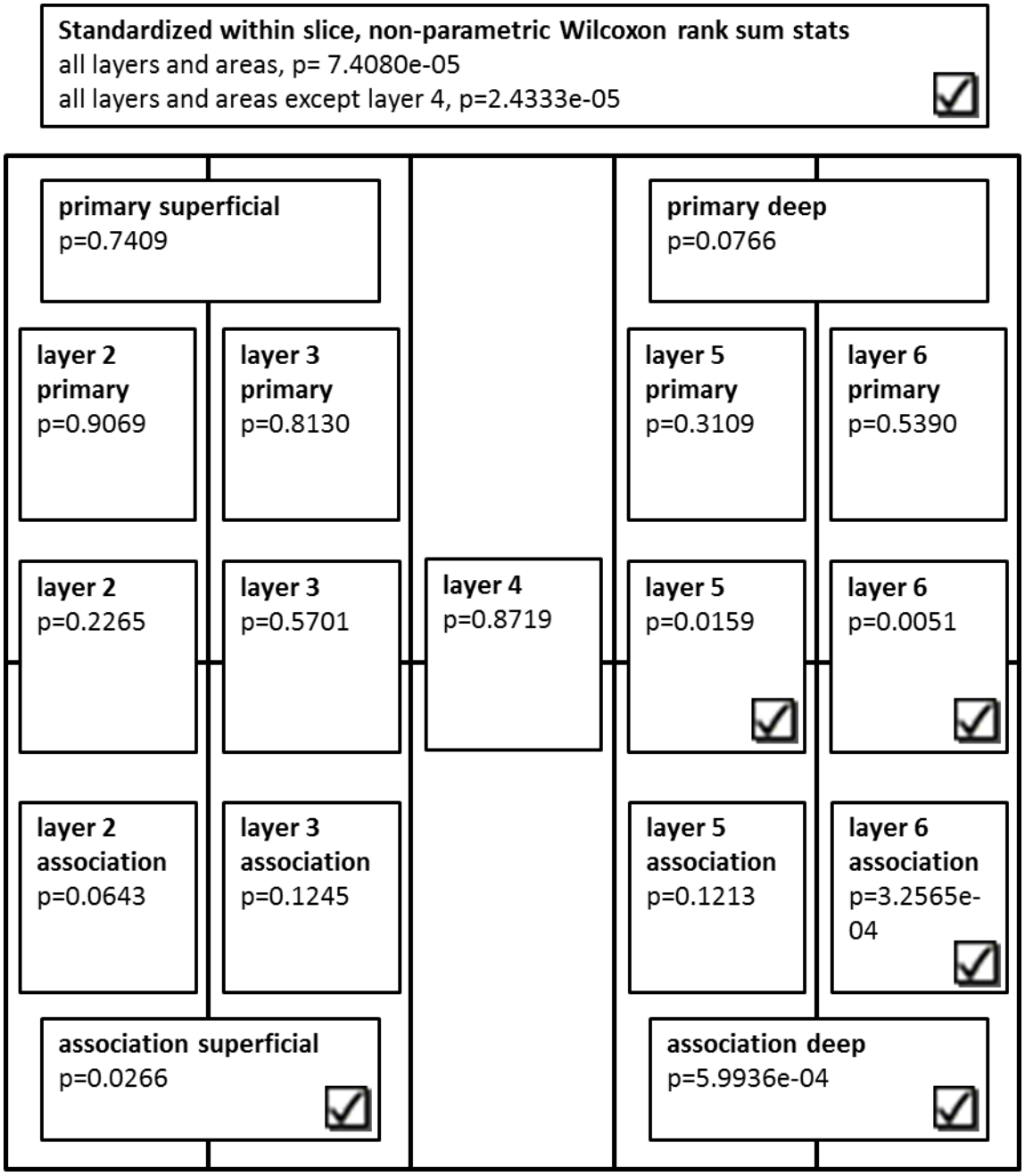


Figure 5.9: Effect of stimulation response on delta power #3

This is a comparison between the delta power on channels that responded to stimulation and channels that did not respond to stimulation. This is a comparison of two groups. Here we compared delta power values that had been standardized (subtract mean and divide by standard deviation) to mitigate the large variation. Overall there is a significant difference, but when broken down into cortical subareas only deep layers and association areas are significant. Superficial association was also significant, but the effect disappeared when broken down further. Check marks are shorthand for significance of the variable of interest.

Figure 5.10: Effect of stimulation response on delta power #4

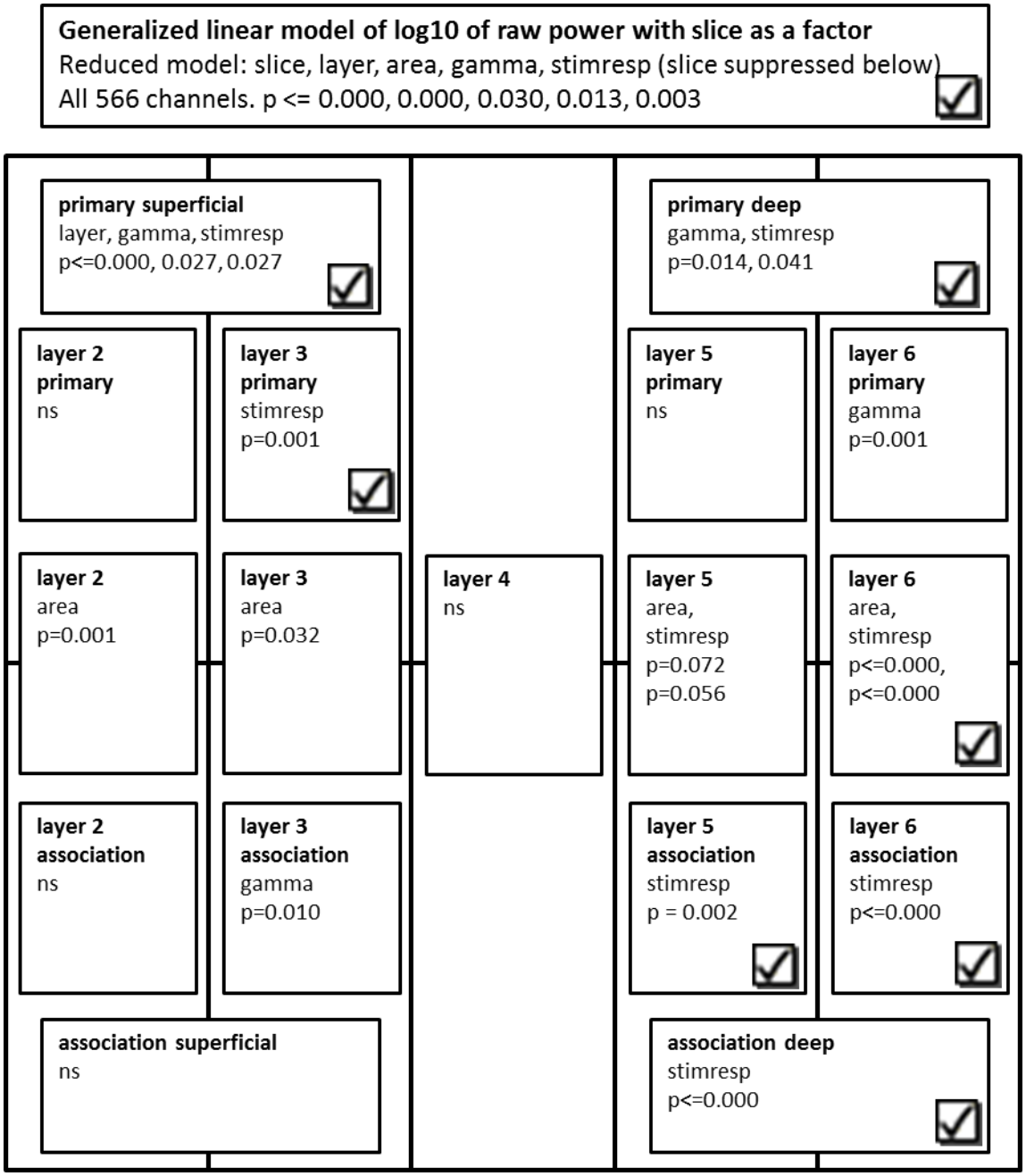


Figure 5.10: Effect of stimulation response on delta power #4

This is a generalized linear model of the logarithm of the raw delta power. This is not a comparison of two groups; instead it is a test whether a number of factors significantly influence the outcome (log delta power). Overall all the listed factors were significant, which also meant that the included factors were controlled for, so that the significance of the factor we were interested in (stimresp) was not confounded by, in this case, layer and area. However, we also wanted to know where the significant differences in layers and areas are coming from, so we broke down the data by these factors and ran new models without them. Check marks are shorthand for significance of the variable of interest, other significant or lose to significant factors are listed in the respective splits.

Figure 5.11: Effect of stimulation response on delta power #5

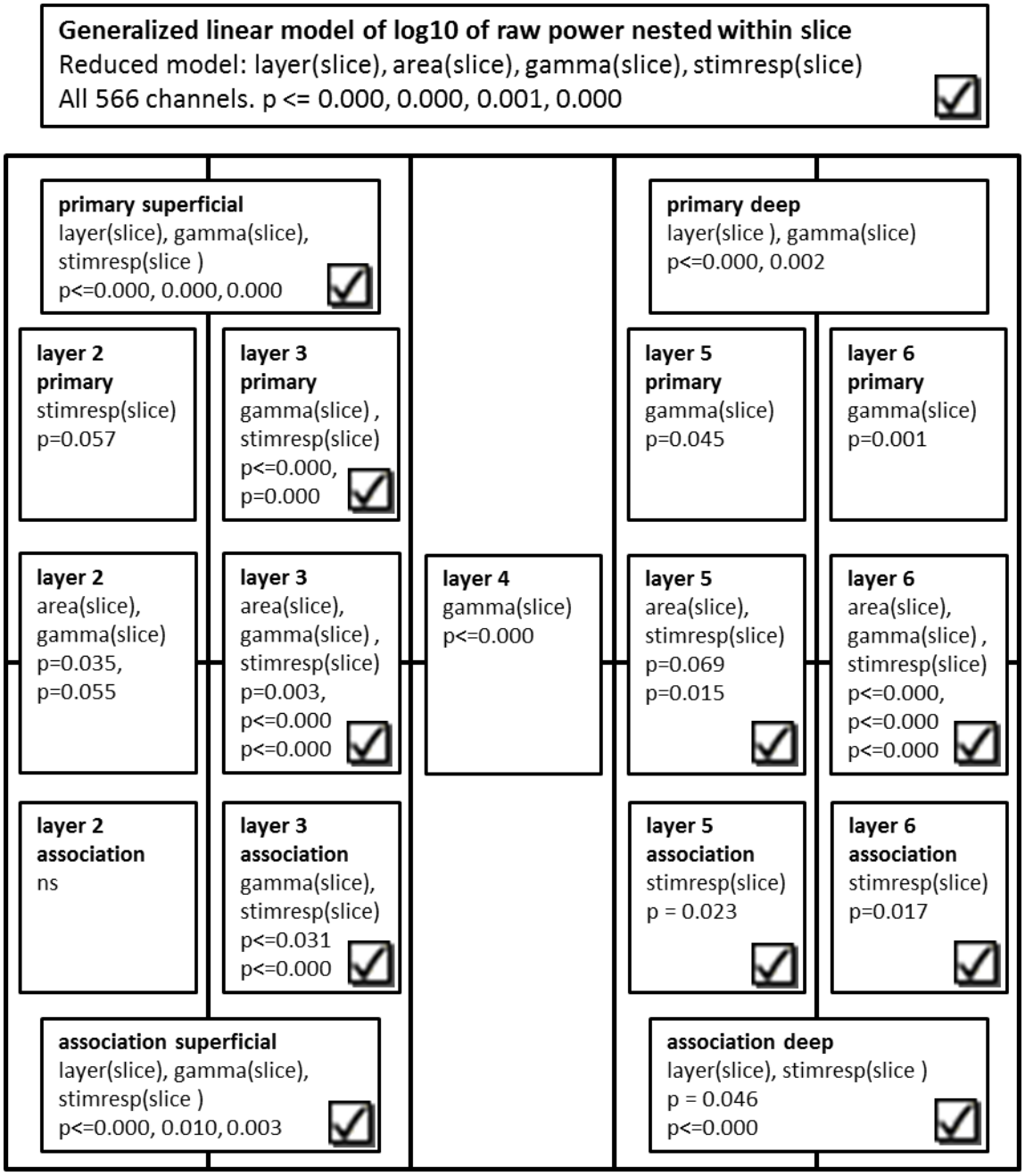


Figure 5.11: Effect of stimulation response on delta power #5

This is a generalized linear model of the logarithm of the raw delta power with all other factors nested within the slice factor. This is not a comparison of two groups; instead it tests whether a number of factors significantly influence the outcome (log delta power). By nesting factors within slice we are making stronger assumptions about the hierarchy of factors, meaning that layers, area etc. are not considered the same in each slice. This might be justified in the sense that layer x in one slice will not necessarily have the same baseline synaptic potentiation as layer x in another slice. The relationship of layers within a slice is the still same for all slices. The model has N times higher degrees of freedom. Overall all the listed factors were significant, which also meant that the included factors were controlled for, so that the significance of the factor we were interested in (stimresp) was not confounded by, in this case, layer and area. However, we also wanted to know where the significant differences in layers and areas was coming from, so we broke down the data by these factors and ran new models without them. Check marks are shorthand for significance of the variable of interest.

5.3.3 *Unit reactivation and deactivation during delta*

We also utilized multiple types of statistical analysis to investigate whether units that responded to stimulation were more likely to be active again during delta than units that did not respond. We present results for N = 16 slices; the remaining three Ns considered in the previous analysis were not excluded, but merely skipped during preparation of the analysis in order to save time.

Figure 5.12 shows the result of the simplest analysis by comparing the percentages of responding and not responding units that were active again during delta. This revealed an overall significant effect ($p<=0.000$), which focused in associational areas when the data was split up by respective layers and areas. The generalized linear model shown in Figure 5.13 also confirmed this result. Out of a total of 1155 units, 257 responded to stimulation. It is worth noting the spatial overlap in terms of significance of this reactivation analysis with the previous section's analysis of increased delta power, focussing in on deep as well as superficial associational areas.

5.3.4 *Differences between reactivating and deactivating units*

We started by investigating unit shapes according to whether they responded to stimulation, and by when they were otherwise active. Out of 257 units responding to stimulation, 129 where active again during delta (reactivated), while the remaining 128 units were not active during delta (deactivated). This revealed a clear size difference between reactivated and deactivated units as shown in Figure 5.14. To formalize this analysis we then mixed the two groups of unit shapes, and blindly sorted and clustered them into two clusters with the Wave_clus Matlab software, to form a comparison of the number of units from each group belonging to each cluster as shown in Figure 5.15. A difference of proportion test revealed a highly significant difference ($p<=0.000$) between the two groups. We then further compared the baseline firing rates of these two groups, shown by layer in Figure 5.16, and also compared the peak stimulation response rates, shown by layer in Figure 5.17. In both cases the same analysis by areas instead of layers produced the same results.

In light of the group differences, reactivated versus deactivated units, and of the spatial overlap of increased delta power and reactivated units, shown in previous sections, we then sought a mechanism to tie these findings closer together. To address this question

we considered a group comparison of the temporal alignment of spike-trains-to-local field-potential phase during stimulation, with the hypothesis that this might underlie differences in potentiation, which in turn underlies differences in delta power, and perhaps also explains the selective reactivation of one group of units. An example of this analysis is shown in Figure 5.18, indicating enhanced phase coherence of reactivated units during stimulation, stemming mainly from the high gamma band where the response is also strongest (see e.g. Figure 5.3). The statistical results of this analysis across all experiments is shown in Figure 5.19, and confirmed that in general reactivated units were much better aligned to the local field potential than units that were later deactivated. The difference was significant ($p= 0.0432$) even when considering the entire gamma band (30-100 Hz), and not just the high gamma band.

Figure 5.12: Probability of reactivation

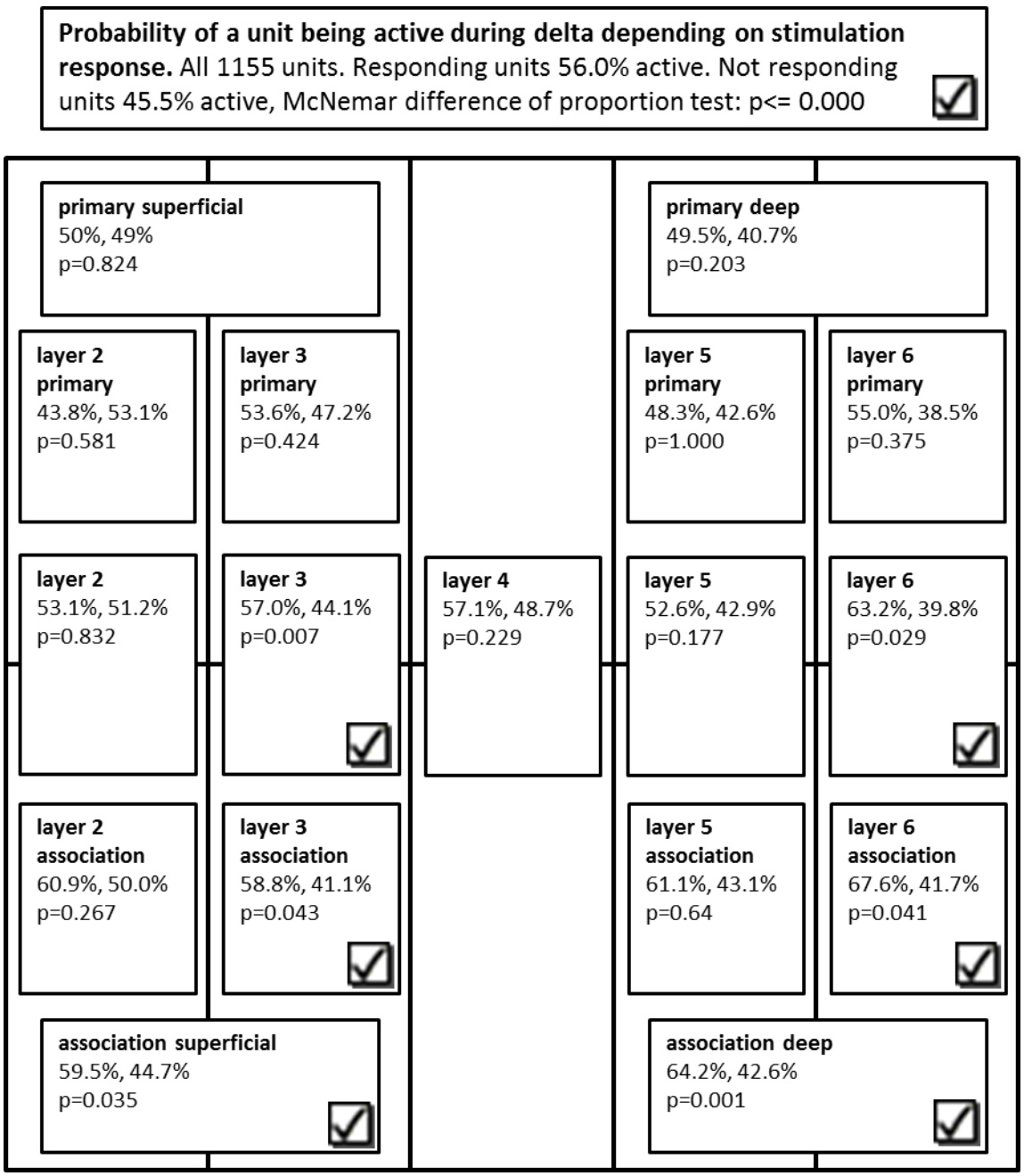


Figure 5.12: Probability of reactivation

This is a comparison of the probability of a unit being active during delta depending on whether it responded to stimulation during the stimulation condition. This is a comparison of proportions of two groups with the McNemar difference of proportion test in SPSS. Overall there was a significant difference, and when broken down by area it can be seen that this difference came from association areas. Check marks are shorthand for significance of the variable of interest.

Figure 5.13: Statistical model of reactivation

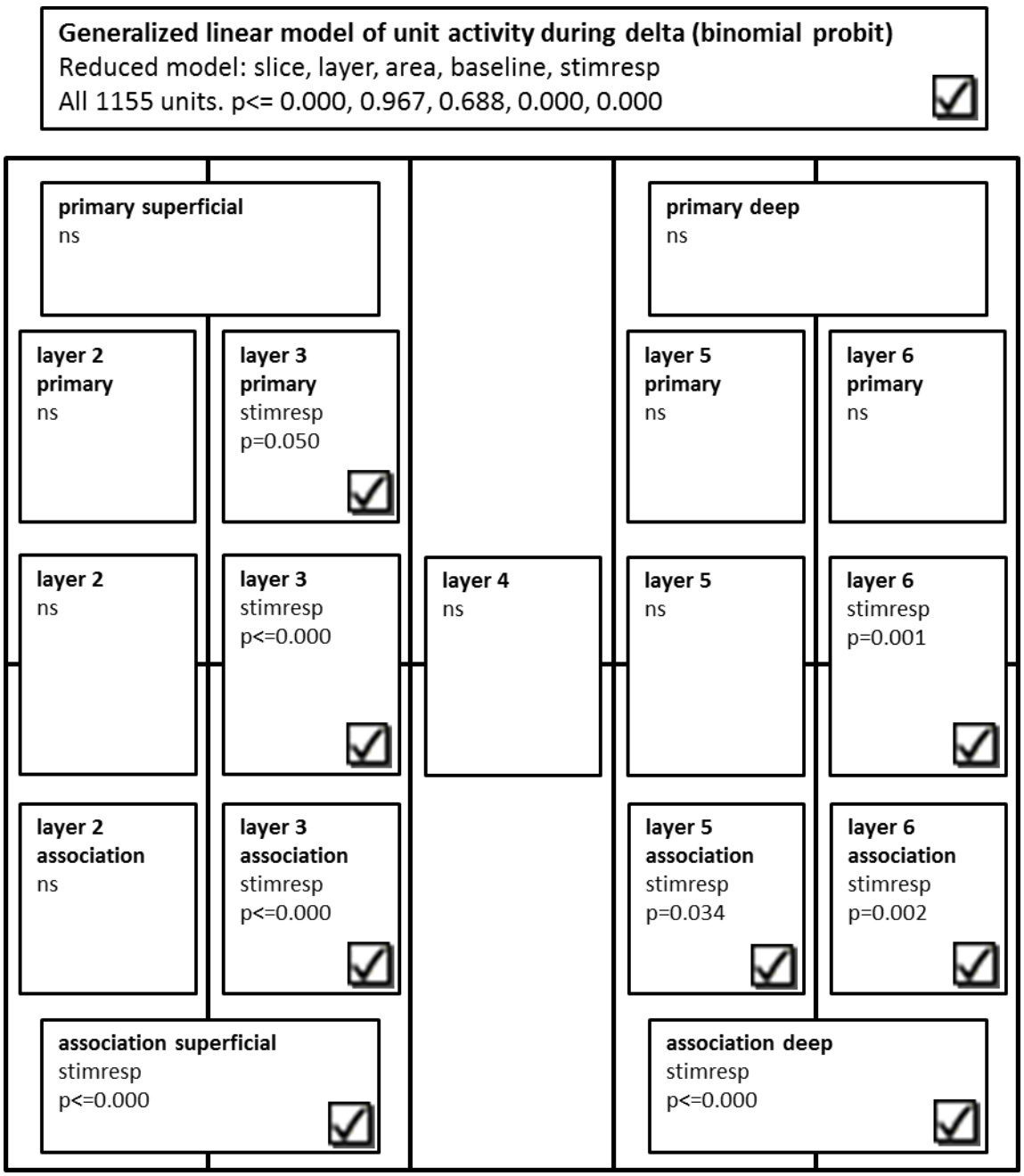


Figure 5.13: Statistical model of reactivation

This is a generalized linear model of the binary outcome of whether a unit is active during delta. This is not a comparison of two groups; instead it is a test whether a number of factors significantly influence the outcome (active during delta). We were interested in the effect of stimulation response. Overall the stimulation response of a unit significantly predicted whether it will be active again during delta. Layer and area were however not significant, so splitting the data on these factors was not predicted to make a difference. However, as the splits showed there were differences in agreement with the previous figure. Check marks are shorthand for significance of the variable of interest.

Figure 5.14: Unit shapes sorted by when they were active

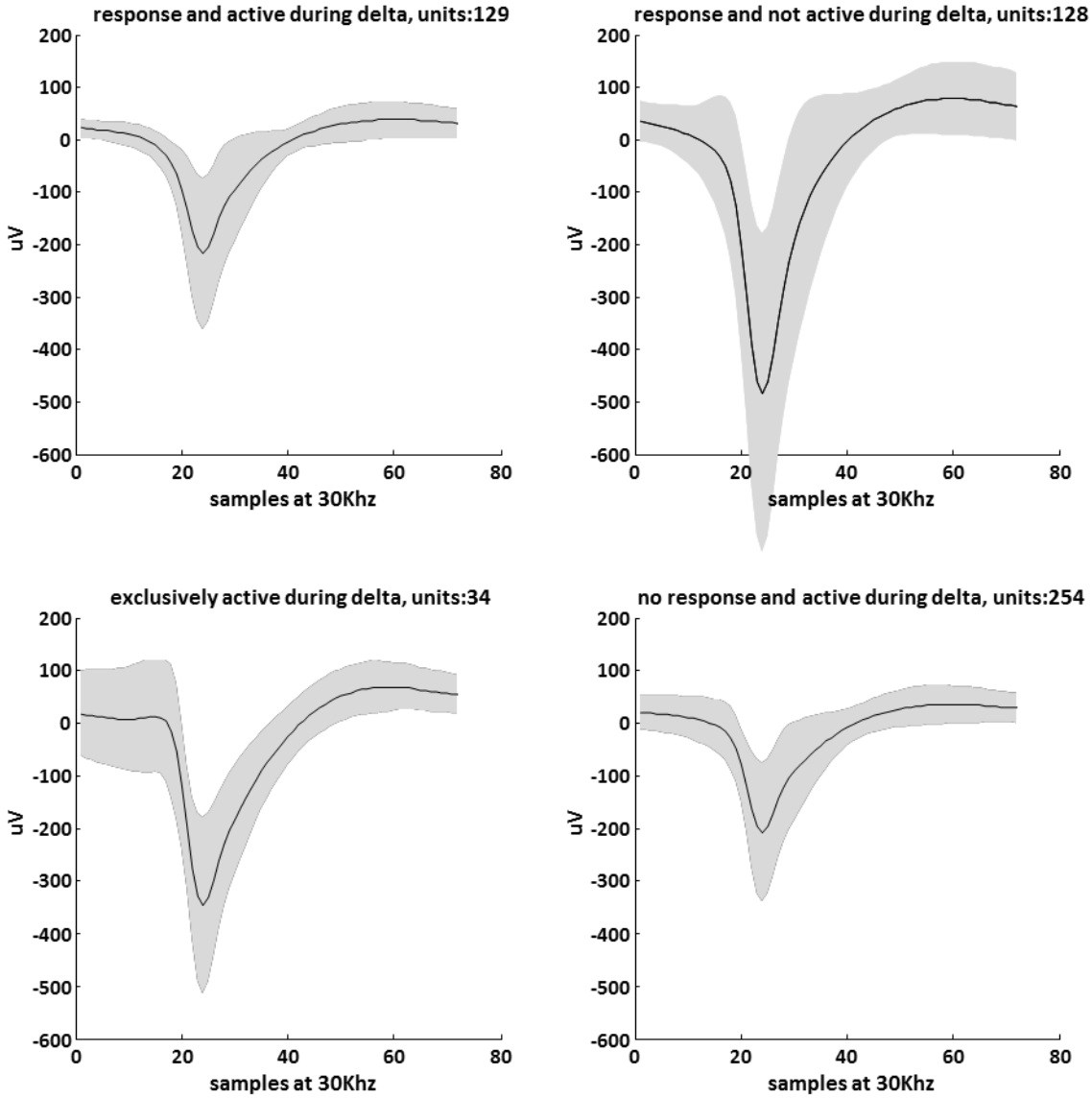


Figure 5.14: Unit shapes sorted by when they were active

Average spike shapes with standard deviation indicated. We used the classifications of Figure 5.5 to produce combined categories of when each unit was active and responded. Top left plot is spike shapes of units that responded to stimulation and were also active during delta, i.e. reactivated units. Top right plot is spike shapes of units that responded to stimulation, but were not active during delta, i.e. deactivated units. Bottom left plot is spike shapes of units that were only active during delta. Bottom right plot is spike shapes of units that did not respond to stimulation, but were active during delta. The difference between reactivated and deactivated units is quite clear, but is further quantified in the following figures.

Figure 5.15: Clustering by Wave_clus of unit shapes from all responding units

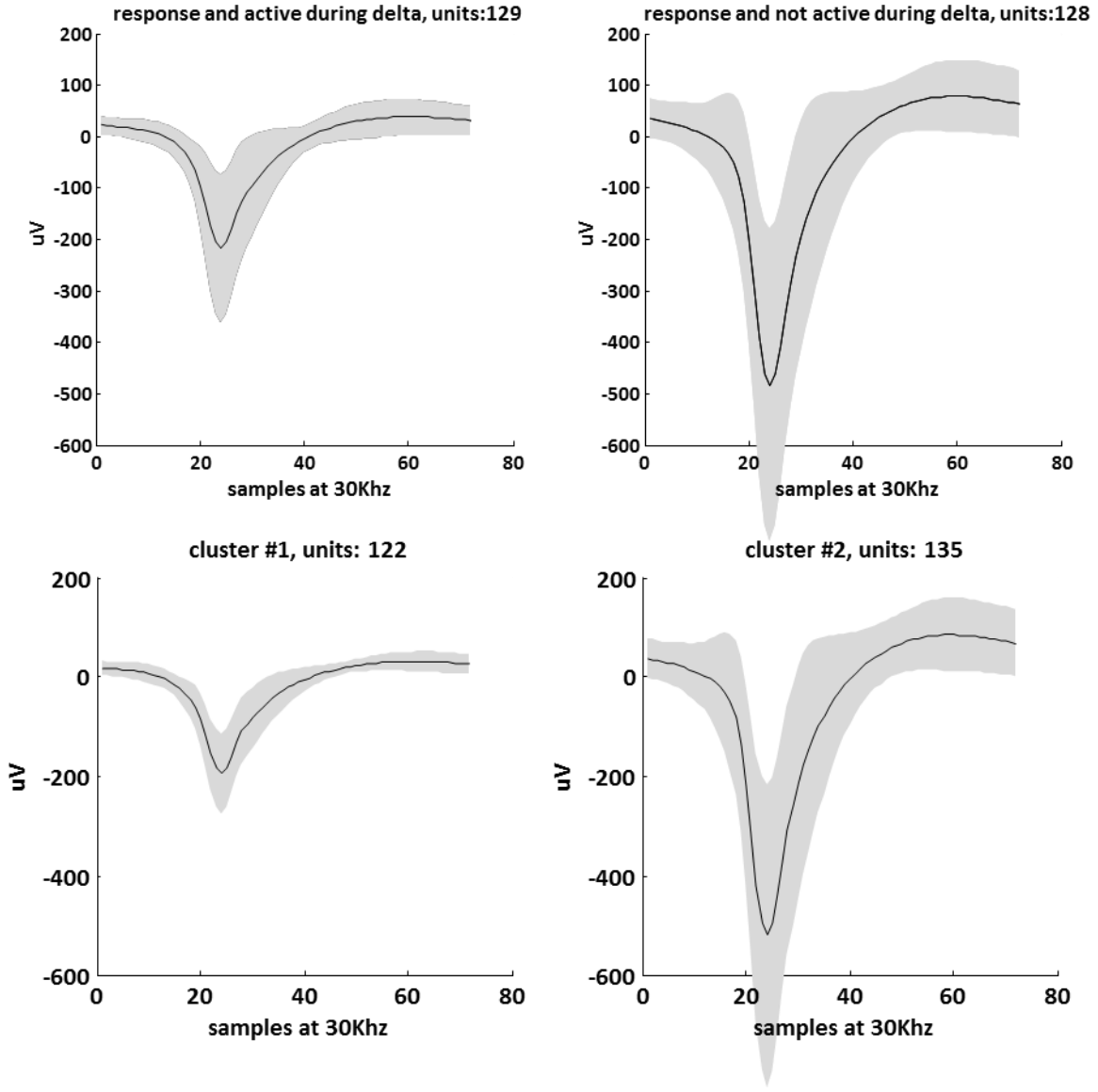


Figure 5.15: Clustering by Wave_clus of unit shapes from all responding units

Top two plots are spike shapes grouped by reactivation or not (from previous figure). Bottom two plots are the same spike shapes, but sorted by wave_clus. A McNemar difference of proportion test revealed that a significantly ($p <= 0.000$) bigger proportion of units that were active again during delta belonged to cluster #1, i.e. based on the spike shapes, reactivated and deactivated units were different cell types.

We also quantified all cells on a bursting criteria during delta; this kind of bursting was specified as the repeated firing of bursts of three or more spikes on the delta period on more than half of the (on average) 60 consecutive periods inspected. McNemar difference of proportion tests revealed no difference in the tendency to burst between units that had previously responded to stimulation and units that had not responded ($p=0.435$); there was also no difference when considering layer 5 (the focal point of bursting units) alone ($p=0.851$).

Figure 5.16: Baseline firing rates of reactivated and deactivated units

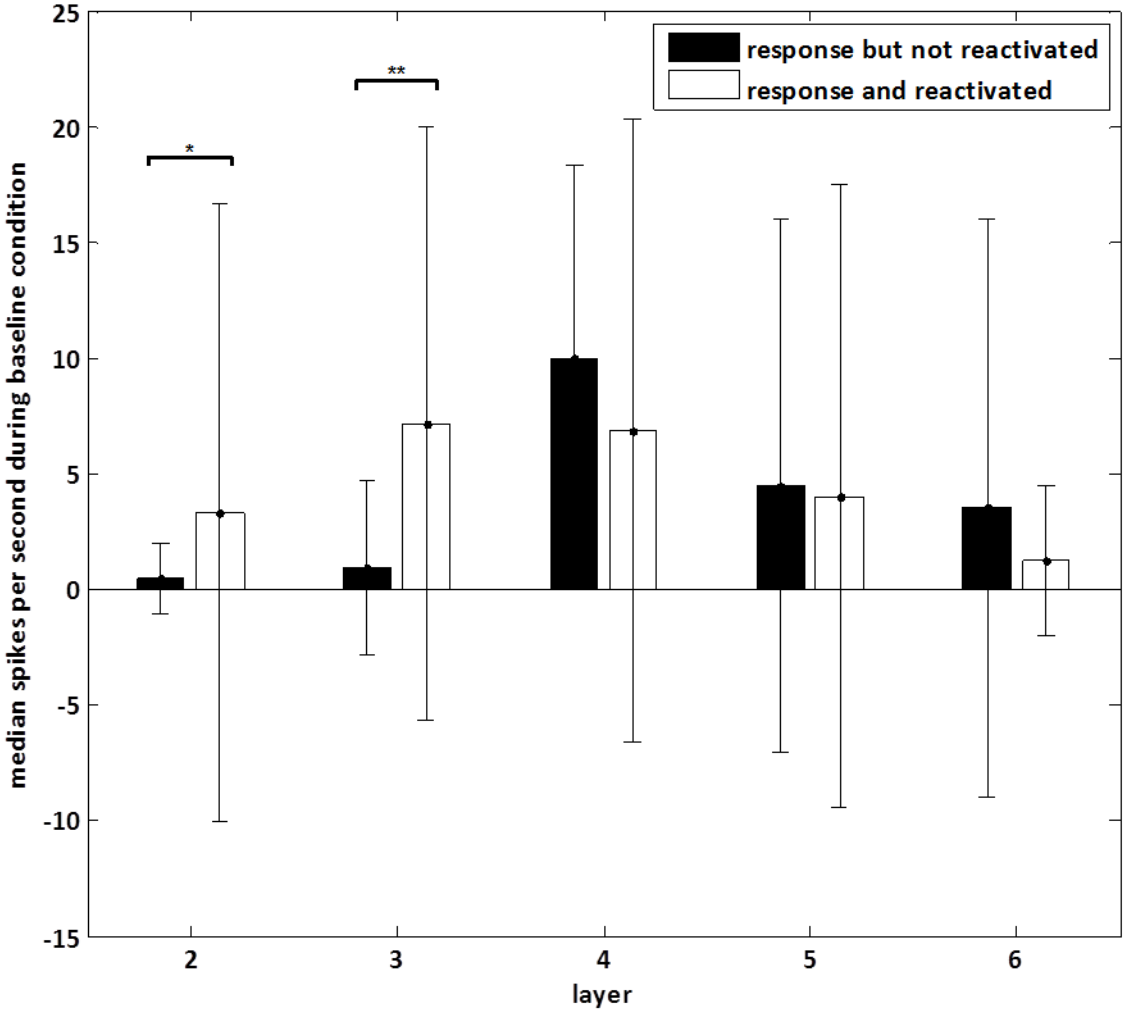


Figure 5.16: Baseline firing rates of reactivated and deactivated units

Comparing the firing rates (in the baseline condition before stimulation) of units that were reactivated with units that were deactivated revealed much sparser firing of units in superficial layers that were not later reactivated ($p=0.0228$ and $p=0.0093$, non-parametric Wilcoxon rank sum statistics). Splitting the data by areas produced the same result. It is interesting that superficial and deep layers did not show the same effect; we have not yet formed a hypothesis to explain the differences in terms of reactivation, but it is already known that deep and superficial layers produce distinct characteristics in this gamma model, with superficial layers dominated by gap-junction-dependent low gamma oscillation, while deep layers take over with pyramidal-interneuron high gamma rhythms, when the system is driven harder, as also seen during stimulation, c.f. Figure 5.3.

Figure 5.17: Peak response firing rates of reactivated and deactivated units

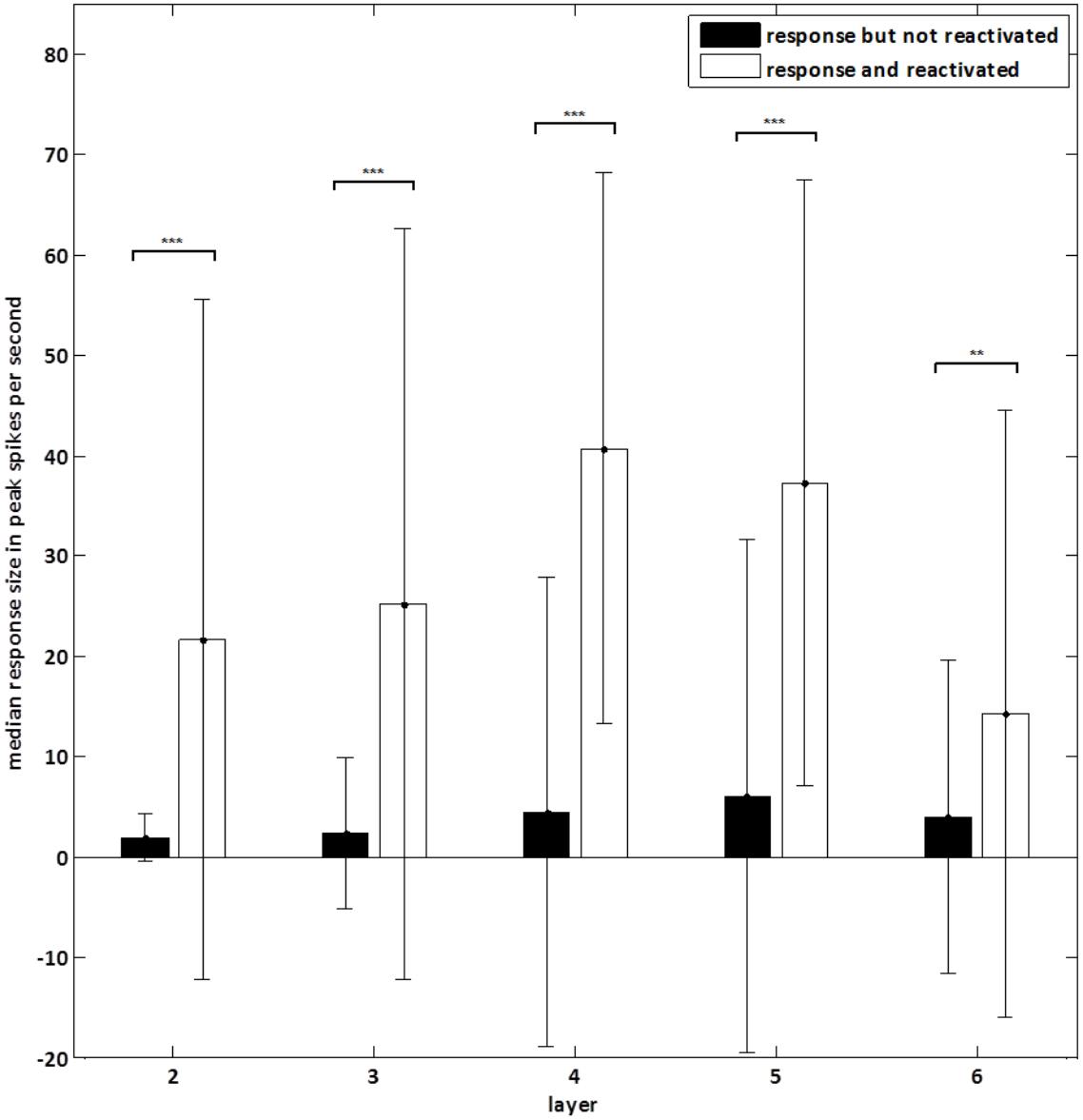


Figure 5.17: Peak response firing rates of reactivated and deactivated units

The peak firing rate of the response of units that were reactivated was much higher than units that were not reactivated. Left to right, p= $2.0177e^{-07}$, 9.1198^{e-10}, $1.8698e^{-05}$, $3.2043e^{-10}$, 0.0093, non-parametric Wilcoxon rank sum statistics. Splitting the data by areas produced the same result. It is interesting that the effect, which was limited to superficial layers during the baseline condition, is seen across all layers during stimulation; this might then be related to stimulation kicking off the deep high gamma rhythms mentioned in the previous figure. This would then mean that reactivated units are involved in this rhythm, and taken together with the spatial overlap of increased delta and reactivation, might further suggest that this rhythm somehow also underlies increased potentiation.

Figure 5.18: Example of spike-train-to-local-field-potential phase semblance

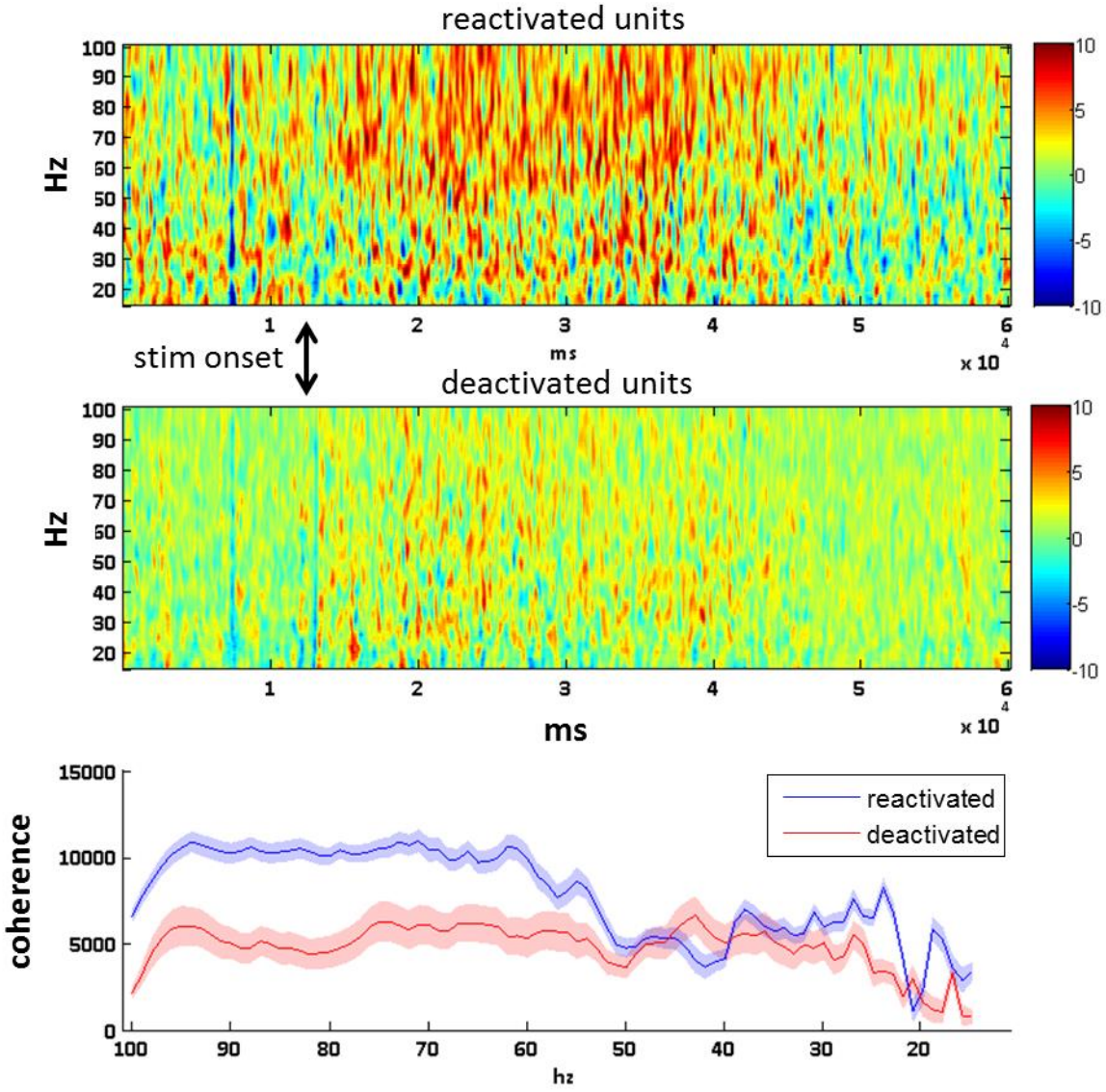


Figure 5.18: Example of spike-train-to-local-field-potential phase semblance

Top two panels: Time-frequency spike-train-to-local-field-potential phase semblance (coherence) plots of the average of 13 reactivated units versus 12 deactivated units from the same slice during stimulation.

Bottom panel: Summation over time of above plots, reactivated in blue, deactivated in red. Notice the main difference in the high gamma band as predicted by the high gamma stimulation responses. The local field potentials were non-causally notch filtered at 50 Hz.

The next figure provides the corresponding statistics across all experiments.

Figure 5.19: Spike-train-to-local-field-potential coherence statistics

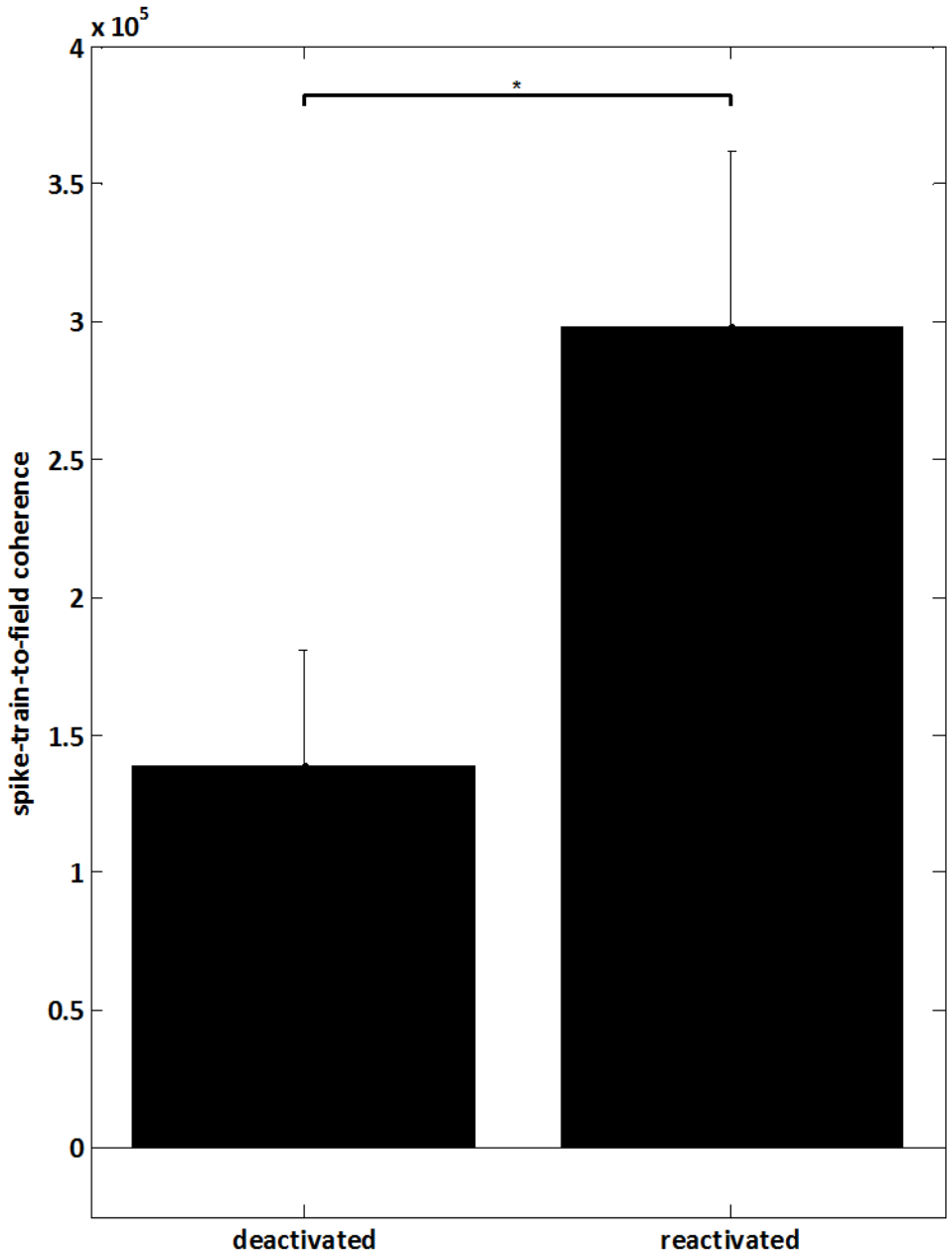


Figure 5.19: Spike-train-to-local-field-potential coherence statistics

Comparing the spike-train-to-local-field-potential phase semblance (coherence) during stimulation of all responding units depending on whether they are reactivated or deactivated during delta, revealed a significant (p= 0.0432, non-parametric Wilcoxon rank sum statistics, median indicated) difference favouring reactivated units. This comparison considered the entire gamma band (30-100 Hz), however as indicated in the previous figure the effect could be enhanced by considered only the high gamma band.

5.4 Discussion

Using a reduced *in vitro* model of two of the main aspects of the wake-sleep cycle the experiments described in this chapter sought to investigate the effects of stimulus response on subsequent delta power, as well as the possibility of reactivation of specific cell types during this delta rhythm. The main findings were that: 1) Layer 4 primary sensory stimulation predominantly activated deep layers proximal to the simulation site. In addition weaker activation of both deep and superficial adjacent association cortex was seen. 2) Response to stimulation was significantly correlated with a subsequent boost in delta power. This boost was seen mainly in deep association cortex, but some evidence for superficial layer involvement and primary sensory area involvement was also seen in a manner dependent on the normalisation of the data and the statistical tests used. 3) Many different unit shapes and behaviours were seen, with sufficient data collected to analyse the type of unit and its anatomical location with respect to reactivation incidence during delta. Reactivated units were predominantly seen in association cortex (where delta power also increased strongest). They were significantly different in their profile compared to non-reactivated units, significantly different in terms of their activation rates during stimulus and the degree of phase locking to stimulus-induced gamma rhythms.

These findings generally support the validity of the model for probing the relationship between deep sleep-associated delta rhythms and sensory stimulation during prior wake states. They also provide a useful substrate for examining the nature of replay responses to sensory stimulation, not just from the superficial point of view of delta carrying fragments of wakefulness (Destexhe, Hughes et al. 2007), but from a computational perspective where replayed aspects of responses to sensory stimuli may be computationally acted upon to selectively extinguish or reinforce certain features for long-term memory. However, while generally useful, we must critically consider the findings in context to published precedents before we can proceed to more detailed analysis of the dynamics of the system.

5.4.1 *Spatial patterns of stimulus response*

We chose a persistent gamma rhythm model as a substrate for wakefulness as activity within the gamma band is involved in multiple aspects of cortical computation. Most of these stem from the close relationship between this rhythm, particularly the faster

gamma components seen in the field potential example illustrated in this chapter (e.g. Figure 5.3), and spike generation in cortical neurons (Rasch, Gretton et al. 2008; Ray, Hsiao et al. 2008). Gamma rhythms have the ability to temporally control and segregate spikes in multiple spatially separate neurons (Engel and Singer 2001; Fries 2005). Thus there is good evidence that the resulting patterns of synchrony and coherence in cortex form the basis of a code for many aspects of sensory representation and short-term memory (Engel, Fries et al. 2001; Kaiser and Lutzenberger 2005). The possibility of replay of such a code is a fundamental topic of this thesis.

There are, however, multiple subtypes of gamma rhythms, e.g. see (Whittington, Cunningham et al. 2011), and the stimulus responses observed here differ in frequency from the background gamma induced by the kainate application to the bathing medium. Interaction between high frequency and lower frequency gamma rhythms has been studied in a similar *in vitro* system (Ainsworth, Lee et al. 2011) and interesting laminar activation profiles result depending on stimulus intensity. A similar situation is seen *in vivo* in studies with changes in sensory stimuli (Ray, Hsiao et al. 2008), suggesting that output from sensory areas (arising from deep layers (layer 5 and 6) may be dramatically influenced by relative stimulus intensity. In this thesis the predominant activation of primary sensory cortex to stimulation was in these deep layers. This suggests that the stimulus used (glutamate application to layer 4) was strong enough to counter the on-going 'contextual' (Petersen and Crochet 2013) influence of the gamma rhythm in superficial layers and may, in part, explain the absence of direct activation and only weak replay in this layer/area (see below).

Activity coding for a sensory input in a primary cortical area is usually passed bottom-up to higher order associational areas. This is seen for visual stimuli *in vivo* (Roberts, Lowet et al. 2013) and primary auditory cortex activation *in vitro* (Roopun, Lebeau et al. 2010). However, the laminar specificity of these bottom-up connections is the subject of much debate (Hilgetag, O'Neill et al. 2000; Hilgetag and Grant 2010). The activation profile seen in this thesis suggests strongly a divergent pathway of connectivity from the deep layers of primary sensory cortex to both deep and superficial layers of association cortex (Figure 5.6). In this way the model used here provides a rich substrate to investigate any changes in sensory responses (modelled as layer 4 excitation) both in primary areas and the associational areas, more closely related to delta rhythm generation (Carracedo, Kjeldsen et al. 2013).

5.4.2 *Delta enhancement in relation to stimulus response*

In humans, local changes in delta power specific to locations stimulated with transcranial magnetic stimulation (TMS) in prior wake-states has been reported (Huber, Esser et al. 2007). Similar reports of enhancement of delta activity following sensory stimulation have also been reported in other experimental systems. Activation of discrete cortical areas during waking shows altered delta power during subsequent sleep episodes (Massimini, Ferrarelli et al. 2007; De Gennaro, Fratello et al. 2008). Delta rhythm generation during deep sleep is associated with enhancement of procedural and declarative memory (Huber, Ghilardi et al. 2004; Aeschbach 2009). Evidence also exists for stimulus-specific, regional potentiation of delta rhythms (e.g. hand vibration (Kattler, Dijk et al. 1994)), or whisker stimulation in rodent models (Vyazovskiy and Tobler 2008), and reactivation of spike timing sequences during sleep is associated with 'learned rules' (Peyrache, Khamassi et al. 2009). Functional clustering algorithms, applied to electrophysiological signals during sleep also show state-dependence consistent with known processes involved in memory consolidation (Feldt, Waddell et al. 2009), and evidence exists for sleep-dependent enhancement of visual cortical circuit remodelling (Aton, Seibt et al. 2009).

The correlation between prior sensory experience and delta power has led to theories relating delta rhythms to modification of synaptic plastic changes occurring during prior wakefulness. Selective enhancement and extinction of plastic changes have been proposed as synaptic rescaling (Tononi and Cirelli 2006), i.e. the restoration of a mean baseline of synaptic weights in cortex. Computational models predict that as delta rhythms decline during sleep so does overall cortical synaptic strength (Riedner, Vyazovskiy et al. 2007), and AMPA receptor phosphorylation levels are potentiated during wakefulness, but depressed after sleep (Vyazovskiy, Cirelli et al. 2008). Where are the loci of such plastic changes, and exactly what operation are they performing on memory of prior sensory events? The predominant occurrence of delta generators in associational areas (Carracedo, Kjeldsen et al. 2013), and the general pattern of dominance of outputs from layer 5 (the layer most strongly activated in the present model) from higher order cortical areas (Rouiller, Simm et al. 1991) suggest subtle synaptic changes favouring the establishment of longer-range functional connections than those induced by sensory input, thus perhaps contextualising prior experiences. However, large-scale alterations on functional cortical circuitry during sleep have been

reported (Esser, Hill et al. 2009) with a strong trend towards uncoupling of longer-range connections.

From the above it is clear the field to date is rather confused as to the nature of the dynamics of delta sleep pertinent to operations on sensory codes held in memory. We therefore began by rigorously analysing the changes in stimulus-induced delta power in our model. These analyses showed that the strongest changes occurred in deep layers of association cortex, and not the primary sensory area stimulated and activating these layers. This observation, taken alone, supports the idea of delta rhythms reinforcing higher cortical representations (contexts) of prior sensory events. This does not fit the highly region-specific delta enhancements seen in the above TMS human studies. However, significant changes in delta power could also be seen in the responding primary sensory loci depending on the constraints of the statistical analysis performed. It is possible that it is the interplay between enhanced delta rhythms in these two areas (sensory and associational) that form the crux of the dynamics pertinent to whatever operation delta is performing on representations of prior sensory experience; this will be dealt with in more detail in the next chapter.

Finally, it can be seen from the present data that superficial layers do not readily display changes in delta power depending on prior stimulation. There are several reasons why this may be the case: Firstly, superficial layers do not, per se, generate delta rhythms at all (Carracedo, Kjeldsen et al. 2013). Instead they receive delta- (and theta-) patterned drive from their corresponding delta-generating deep layers. Secondly, the model used a persistent gamma rhythm as a wake-state baseline. This model selectively generates gamma in superficial layers alone (Cunningham, Halliday et al. 2004) and, as discussed in section 5.4.1, this rhythm imparts a template on neuronal activity such that only template-matched patterns of input (cortical response to stimulation) are processed. As the stimulus used in the model was very crude (just a spatially constrained excitation of layer 4), it is possible that no overt changes in superficial layer activity were induced. This is borne out by the lack of strong unit activity responses seen in primary superficial layers. Thirdly, the nature of any code for sensory information held in superficial layers is undoubtedly complex; involving higher order spike correlations (Ohiorhenuan, Mechler et al. 2010) that may not have influenced the basic spike-rate measures used in this thesis.

5.4.3 *The selectivity of unit reactivation*

Despite not considering higher-order spike correlation patterns briefly discussed above, our basic measure of spike-rate was found to be very useful when investigating the phenomenon of replay during delta rhythms. Most available evidence suggests a critical role for prefrontal cortex in conjunction with sharp wave/ripple outputs from hippocampus (Peyrache, Khamassi et al. 2009; Benchenane, Peyrache et al. 2010). Spike patterns associated with a learned behavioural task were replayed with high fidelity, and the incidence of replay was found to be related to global levels of unit activity in the putative cells involved.

Relating unit activity (the extracellular measure of spiking used in these and many similar studies) to specific neuronal subtypes is however fraught with uncertainty. Spike width has been used to separate interneurons out from principal cells, e.g. see (Peyrache, Dehghani et al. 2012). By this criterion alone it should be safe to say that the units in green in Figure 5.5 represent interneuron activity. The fact that they fire robustly and regularly during gamma and not delta seems to fit this. However, spike width, measured intracellularly from identified neurons reveals a very broad range of widths and activity patterns within interneuron and principal cell subgroups. For example regular spiking and fast rhythmic bursting (FRB) principal cells have wide and narrow spike shapes respectively (Traub, Contreras et al. 2005), and the narrow spiking FRB neurons also fire robustly and regularly during gamma. Similarly, basket-type interneurons and dendrite-targeting interneurons have narrow and wide spike shapes that completely overlap the principal cells just described (compare data in (Armstrong and Soltesz 2012) with (Fuchs, Zivkovic et al. 2007)). This study did therefore not attempt to subclassify units on the basis of cell type.

Despite this difficult point it was clear that different units behaved differently during the stages of the wake-sleep cycle model used here. Some units appeared to be entirely related to the generation of the gamma rhythm and some to the delta rhythm (Figure 5.5), but the most interesting subclass were those that responded with an increase in spike rate to stimulation. Within this subclass two statistically different spike shapes corresponded to units that reactivated or deactivated during delta. The reasons for this stark contrast appeared most obviously to do with the timing of spike-trains relative to the field potential gamma rhythm during stimulation. Those units that aligned precisely with the gamma rhythm replayed whereas those that had more variable timing did not.

Gamma rhythms are an ideal substrate for assembly formation, an assembly being a collection of neurons firing together following stimulation (Ainsworth, Lee et al. 2012). Such simultaneous spiking affords the advantage of supralinear summation of convergent input from the assembly neurons onto down-stream targets, and can thus have a very large effect on synaptic plasticity. Combination of high frequency (stimulus-related) and lower frequency (background) gamma rhythms such as those used in the present model set-up frequency differences that, through standard spike timing-dependent plasticity rules (Feldman 2012), would be expected to boost interlaminar connections in a highly direction-specific manner (Lee, Sen et al. 2009; Ainsworth, Lee et al. 2012).

5.4.4 *Summary*

The data presented in this chapter demonstrate a complex relationship between stimulus presentation and subsequent delta rhythms in terms of both spatial (anatomical) location of plasticity-induced changes and the prior history and (perhaps) neuronal subtype replayed during delta rhythms following 'sensory' stimulation. The clear relation between stimulus response and replay, and the putative relationship of this to synaptic plasticity strongly indicates a role for delta rhythms in further processing of cortical representations of prior sensory input. But what is the nature of this further processing, and how may we resolve some of the more complex spatial patterns of changes seen? We will begin an attempt to unravel these phenomena in the next chapter.

Chapter 6. Results

ergodicity, energy flow and dissipation

6.1 Introduction and aims

This second results chapter builds on the framework established by the results of the previous chapter. Here we took our analysis further to address the two theoretical questions posed by our initial considerations of holographic processes. We outline these two focal points in the subsections below.

6.1.1 *Ergodic statistics*

If the spectrum of energy dissipation of the system in both the wake model state and the sleep model state can be described by power laws with certain specific coefficients it is implied that the system dynamics can be described by ergodic statistics, which would then allow Friston's free-energy minimization principle to be in operation (Friston 2010). If only the sleep-state model is described by ergodic statistics the system could possibly be described by a classical annealing process during this half-cycle (see section 2.2.3).

6.1.2 *Spatiotemporal energy flow*

Previous work in our lab suggested reciprocal interaction of deep and superficial layers in associational areas orchestrated by the delta rhythm, as detailed in section 2.2.3. This was taken to suggest a Helmholtz machine-like process of synaptic weight learning. However, under certain assumptions (as described in section 2.2.3) the Helmholtz machine framework suggests that there would have to be an interaction along layers as well, also orchestrated by the delta rhythm. The flow of energy is not a direct measure of causality, but is likely to reflect causality to some extent, and is far less removed from the actual underlying biological processes.

6.2 Methods

The methods described here are specific to this chapter; for general methods see chapter 4. We introduce the relevant tools in the order that we used them.

6.2.1 *Non-parametric wavelet Granger causality*

A standard correlation or coherence measure gives no information about which signal (if any) drives the other. Estimating correlation as a function of lag gives some idea of this important aspect, but a more general idea is to consider how one signal predicts the other. This is the approach taken by Granger causality, which estimates the variance of the autoregressive prediction error of one signal as the past information of the other signal is included in the model (Ding, Chen et al. 2006). If the variance is reduced the included signal has Granger casual influence on the other. The analysis is based on parametric fitting of linear regression models, and provides Granger causality graphs as a function of frequency. However, we wanted to base the Granger causality method on wavelet analysis, and achieve a time-frequency Granger causality map instead of a parametric time-average, so that we could study the time course in detail where reciprocal interactions might otherwise cancel out.

This line of thought led us to a recent approach that does just that (Dhamala, Rangarajan et al. 2008). The main challenge in implementing this non-parametric Granger causality was a spectral factorization step. Many known algorithms exist (Sayed and Kailath 2001), but no code libraries seem available. The original non-parametric Granger causality uses Wilson's algorithm (Wilson 1972), however we found it simpler to implement the Kolmogorov cepstral method (Kolmogorov 1941), which although intended to be Fourier based seemed to work well with wavelet coefficients as well. A Matlab toolbox implementation of non-parametric Granger causality has since come out (Oostenveld, Fries et al. 2011), and we have verified that our version produce same or similar results, and found it to be orders of magnitude faster.

The question of causality in the presence of multiple signals is immediately relevant to our micro-electrode array study. However, although several techniques exist, like partial directed coherence (Baccalá and Sameshima 2001) and multivariate Granger causality (Barrett, Barnett et al. 2010), these do not provide solutions that are guaranteed to be accurate. This is because of the nature of the problem, which implies an infinite regression that thus has to be cut off at some point; a version of the closure of turbulence problem considered in chapter 3. This problem is amplified with increasing numbers of signals. In the case of two signals there is a risk of spurious causality estimates, because a third (unknown) node might actually mediate the causality, making

it so-called indirect causality, but accounting for all other possible nodes becomes an intractable problem when the number of signals is large. Some approaches compute all pair-wise causality estimates, and then try to eliminate indirect causality heuristically (Zou, Ladroue et al. 2010), but this becomes exceedingly difficult when we consider not only averages, but pair-wise time-frequency maps. For these reasons we sought a different approach to multi electrode array causality estimation, which is presented in section 6.2.3 below.

6.2.2 *Ergodicity measures (power-law fitting and limit torus trajectories)*

Before a power-law coefficient can be determined a power-law curve must be fitted to the dissipation spectrum; this fit must in itself be good enough that we can say with significant confidence that the spectrum is actually best described by a power-law at all. This is a complex process for which we chose to use an existing Matlab toolbox, which directly provides p-value calculations for the best fits (Clauset, Shalizi et al. 2009). The power-law coefficient (n, below) intervals for which ergodicity is known (Grigolini et al. 2009) are: Ergodic: $n<1$, $n=2$ (Brownian). Non-ergodic: $1<n<2$, $n=3$ (black body). The status of other intervals is unknown to our knowledge.

A secondary method that we have only considered to a preliminary degree is described next: A dynamical system with irrationally spaced characteristic frequencies form a limit-N-torus in phase-space, and the trajectories on the torus can reveal ergodicity (Grebogi, Ott et al. 1985): If there are two rationally related characteristic frequencies ($N=2$) the trajectories eventually close up, but if the frequencies are irrationally related the trajectories will eventually cover the entire torus, which indicates ergodicity (Mackenzie 2006). This immediately suggests that the brain is ergodic from previous findings of phi related characteristic frequencies (Roopun, Kramer et al. 2008), however as N increases the picture becomes much more complex and irrationally related frequencies no longer guarantee ergodicity. The trend seems to be that ergodicity becomes increasingly unlikely with increasing N, but because trajectories on N-tori become extremely difficult to visualize above $N=4$, it is in fact unknown (Grebogi, Ott et al. 1985).

6.2.3 *Near-field electromagnetic holography*

It can be difficult to relate causality estimates to the actual underlying activity; for example should causality from A to B be reflected in current source density calculations with A as source and B as sink, or should the causality estimate be calculated on the current source density in the first place, or how should causality estimates for local field potentials be compared to same for spike data, and so on. Causality estimates are inherently far removed from the underlying processes, and it was thus our goal to find an approach what would be more directly based on all simultaneous measurements, and that would also unify local field potential and spike data in the same process.

From the preceding discussions of holography it is obvious that this general technique revolves around just such sought-after features, but the question was how to actually apply this principle to the case of micro-electrode array data. However, this line of thought allowed us to uncover an analogous application in a different field where the technique is known as real-time near-field acoustic holography (Hald 2001; Thomas, Grulier et al. 2010). The idea is to treat the array recordings as samples of a hologram in the recording plane; the hologram can then be backwards propagated in space into the medium back to the sources giving rise to the mixed signals at the recording plane. This is done in a higher dimensional frequency space known as K-space where a complex Green's function propagator reconstructs the signal at progressively deeper planes into the medium. This process is however unstable and must be regularized by appropriate filtering (Paillasseur, Thomas et al. 2011); the reason for instability is that high frequency exponentially decaying, evanescent waves in the near-field are inverted and become exponentially amplified. With carefully balanced filtering this phenomenon nevertheless has the curious property that it allows spatial reconstruction resolutions beyond the resolution of the array (Williams and Maynard 1980), i.e. super resolution. This surpasses mere interpolation, and is reminiscent of recent compressive sensing techniques that also violate Nyquist's well-known sampling theorems (Duarte, Davenport et al. 2008).

Further investigation revealed that the analogy was accurate under the relevant conductive conditions with the equations carrying over directly, only now with complex wave numbers, and electromagnetic instead of acoustic interpretations (Williams and Valdivia 2010). The detailed analogies are: Microphone array corresponds to micro-

electrode array. Acoustic pressure and particle velocity correspond to the electric field and the magnetic field. Acoustic intensity corresponds to the Poynting vector.
There is also an important analogy between the acoustic properties of the acoustic medium, and the complex conductivity of the slice medium (Nicolas, Furstoss et al. 1998). We obtained the relevant conductivity values from direct measurements in rodent neocortex given in the literature by layer, and across and along layers (Goto, Hatanaka et al. 2010).

The relevant equations are given as:
Electrical field in recording plane, **E = gradient(micro-electrode array potential)**
Wave number **K**, complex expression of conductivity **rho**, and frequency.
K-space propagator **G**, complex function of propagation distance.
Magnetic field in reconstruction plane, **$\underline{H}$ = K x ($\underline{E}$ * G)**, underscore is K-space transform**.**
Poynting vector in reconstruction plane, **S = ($\underline{E}$ * G) x H**
Current density approximation, **J = rho * ($\underline{E}$ * G)**
Current source density, **CSD = divergence(J)**
Energy source density, **ESD = divergence(S)**
Energy dissipation, **D = dot(J, ($\underline{E}$ * G))**
With **x** being the cross product, and constants suppressed for clarity.

To replace causality estimates we instead considered the Poynting vectors representing spatiotemporal energy flow. The current source density in this implementation has the advantage that it is based on conductivity measures for each layer, but the measure that we are calling energy source density (by analogy of the equations) turned out to be more clearly defined spatially. The measure of energy dissipation was also of particular interest, considering the intimate relations to computation discussed earlier in chapter 1.

The concept of super resolution, although well-documented in the literature (Peter, George et al. 2002; Fouque, Garnier et al. 2006), can be difficult to accept, so we thought hard about a way to justify that the technique works in the present implementation as well. For this purpose we got hold of an example data set of a hippocampal rodent section recorded with the BioChip 4096 array from 3Brain (3Brain 2013) with 64x64 channels, then down-sampled the array grid at 16x16 and 32x32 points evenly over the full grid area. This then allowed us to visually compare the

original 64x64 grid with super resolved versions based on the extracted down-sampled data.

Figure 6.1 first shows two comparisons of the reduced resolution raw data with the full resolution raw data, as snapshots taken at random time points during major neuronal events. Figure 6.2 then moves on to show the energy source density based on the full grid data versus reduced data sets. Finally, Figure 6.3 shows a comparison of the super resolved outcomes versus cubic spline interpolated outcomes based on the same reduced data. We were thus visually satisfied with the technique, and did not pursue a more detailed quantification of the similarities. It should be noted that agreement was not obvious during quiescent periods of low neural activity, which we estimated was not much above the noise level of the array in any case.

Figure 6.4 shows an example snapshot with three views (energy dissipation, energy source density and energy flow vector field) of the application of our technique at three times the original array resolution, and with the reconstruction plane 100 microns into the slice from the recording plane of the array. For presentation in the results section to follow, and in order to quantify energy flow relevant to our investigation we averaged the spatiotemporal energy flow in each of the four neocortical areas (superficial primary, deep primary, superficial association, deep association, that we also considered in the previous chapter), over time bins one eighth of the delta period, locked to the delta period over 60 seconds. We used the preceding non-parametric Granger causality estimations to decide the alignment of the time bins.

It should be noted that it is quite possible to greatly improve spike localization and sorting with this technique as well, since sorting does not rely on spike shapes, but rather on the spatial 3D localization of the spike sources. The amplifying backwards reconstruction process means that spikes buried in noise (as well as faint contributions to clearer spikes) come back into focus at their originating depth planes, and the increased resolution means that overlapping spikes can be spatially distinguished at their points of origin. However, further work is required to make this process feasible for any large amount of spikes, and we did not have the time to apply it in this project.

Figure 6.1: Full resolution versus reduced resolution array data

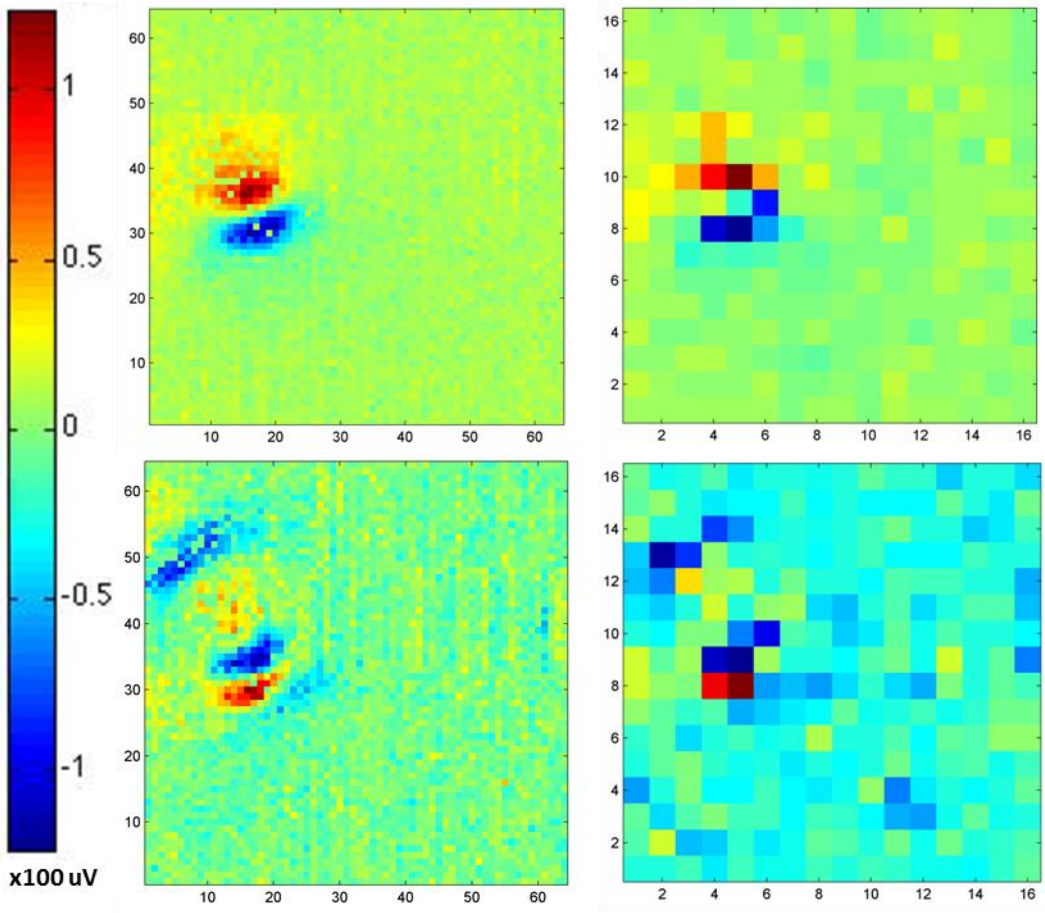


Figure 6.1: Full resolution versus reduced resolution array data

This figure exemplifies the outcomes of down-sampling full array 64x64 resolution to a reduced resolution of 16x16 channels. The data shows strong hippocampal events at otherwise random time points. Top and bottom plots are two different time points. Data taken from example data sets on the 3Brain website. Left is full resolution, right reduced resolution.

Figure 6.2: Energy source density at full and super resolved resolutions

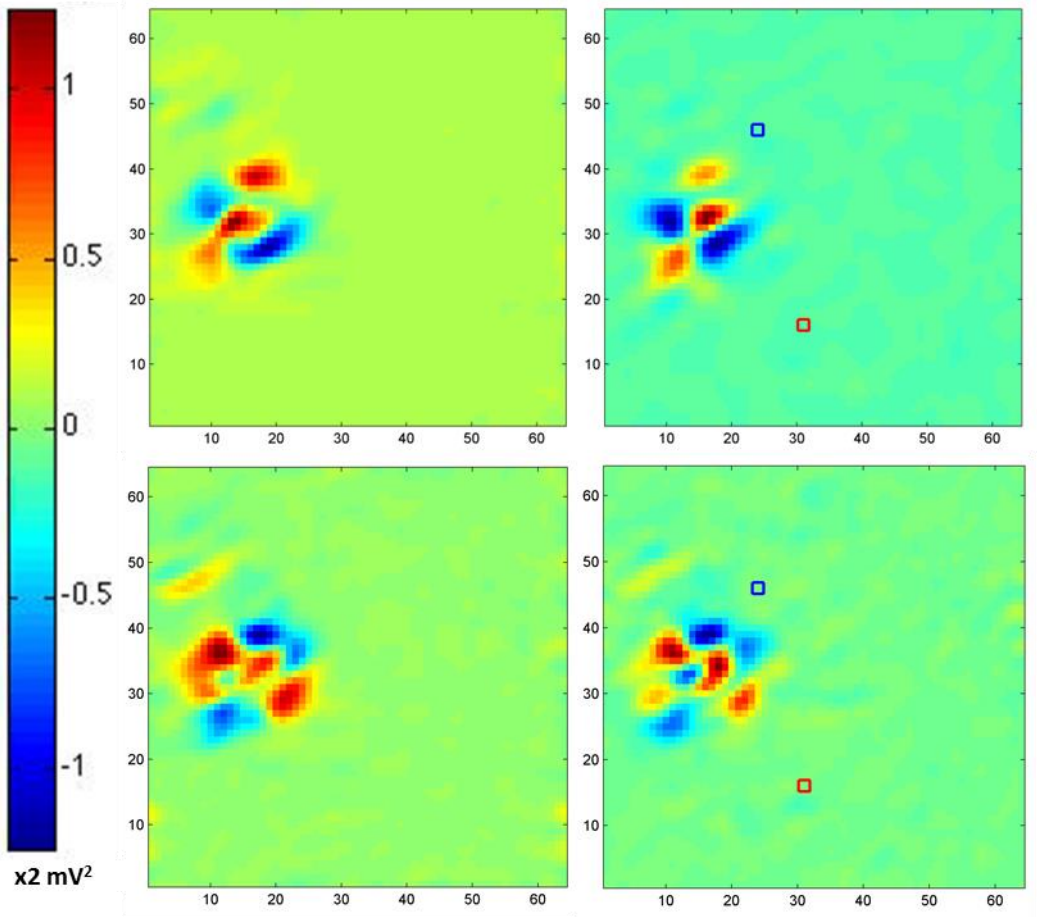


Figure 6.2: Energy source density at full and super resolved resolutions

Comparison of energy source density based on full 64x64 resolution data without super resolution (right) and super resolved energy source density based on reduced 16x16 resolution data (left top) and on 32x32 reduced resolution data (left bottom). Top and bottom plots are two different time points. The square markers in red and blue indicate virtual electrodes for time series comparison (not shown) at random spatial points, which revealed that agreement was only good when neural activity was above noise level.

Figure 6.3: Comparison of super resolution and interpolation

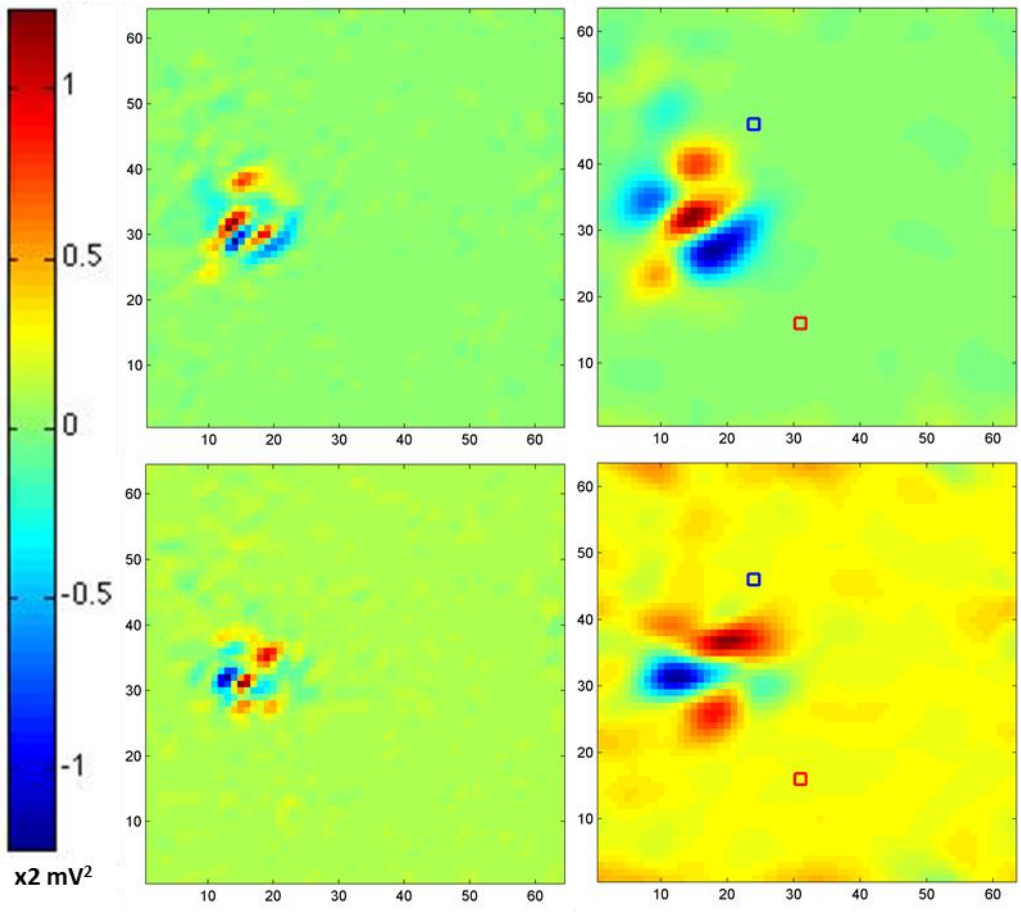


Figure 6.3: Comparison of super resolution and interpolation

Here we compared super resolution (left) against cubic spline interpolation (right) based on reduced 32x32 array data restored to original 64x64 resolution. Top and bottom plots are two different time points. The super resolution view revealed additional detail over the interpolation.

Figure 6.4: Example snapshot of near-field electromagnetic holography

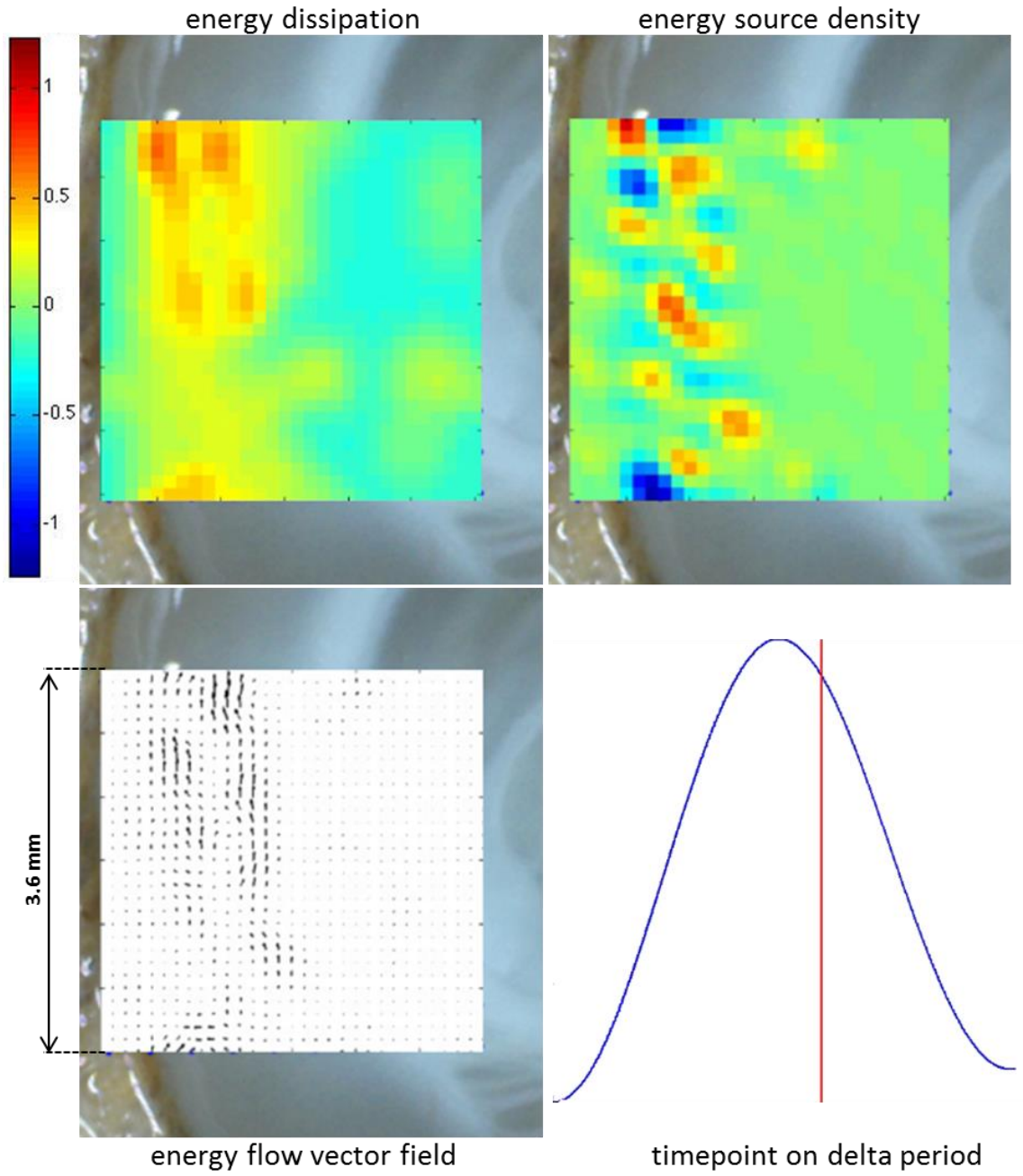


Figure 6.4: Example snapshot of near-field electromagnetic holography

This is a snapshot from a movie covering an entire delta period (about one second) as the delta locked average of 60 seconds of data at 30 kHz sampling rate. The vertical red line indicates the relevant timepoint of the snapshot. This reconstruction was made at three times the original spatial array resolution in both directions. The reconstruction plane is 100 microns into the slice from the recording plane of the array. The energy source density shows alternating sources and sinks along and across neocortical layers, while the energy flow vector field reveals that the actual energy flow is along layers towards primary areas at this time point. The energy dissipation shows a characteristic pattern of dissipation both deep and superficial, but clearly separated. The time course of the movie shows that this dissipation pattern spreads from primary towards association areas from about the time of the delta period peak to the timepoint shown.

6.3 Results

We first addressed the question of ergodicity as determined by the spectrum of energy dissipation. We then moved on to the second question investigating spatiotemporal interactions in neocortex; initially we applied our non-parametric Granger causality measure to a dataset of paired glass electrodes across layers in layer 2/3 and layer 5 perpendicular to the neocortical curvature in the delta model, then we applied our near-field electromagnetic holography technique to the micro-electrode array dataset from the previous chapter. Finally we studied the ratios of energy dissipation to energy flow in the two model states of our main paradigm, respectively the gamma wake-state model and the delta sleep-state model as described in the previous chapter.

6.3.1 *Ergodicity measures for wake and sleep state models*

The results of our investigation of ergodicity can be summarized briefly:
In neither the wake-state model nor in the sleep-state model was it possible to fit power-laws to the dissipation spectrums with significant confidence ($p = 0.22 \pm 0.03$, mean±se). Forcing the fits appeared to produce random coefficient values in the non-ergodic range from 1 to 2 (1.44±0.13). For these reasons we will not report the results in detail, but can conclude that we could not confirm ergodic statistics with this measure; however, whether this really rules out ergodicity remains unknown. We did not carry out a full analysis of the secondary measure, as inspecting a large part of the spectral peaks in the sleep state model revealed (close to) harmonic rather than irrationally spaced frequencies, which also suggested that the system was not ergodic.

6.3.2 *Temporal delta-locked interaction of layer 2/3 and layer 5*

We first set out to elaborate on previous findings in our lab of layer 2/3 and layer 5 interactions with a more detailed causality estimation on local field potentials in the delta and theta bands (Carracedo, Kjeldsen et al. 2013).

Out of 42 paired electrode recordings 13 showed a non-trivial highly similar pattern of non-parametric wavelet Granger causality averaged to the delta period over 60 seconds. No other distinctive patterns were discernible in the remaining cases; some did not allow a proper delta-locked average due to poor delta, others appeared to be insufficiently aligned to the perpendicular axis (potentially of some interest to the concept of cortical columns, but this was not pursued further).

The analysis summarized in Figure 6.5 shows layer 5 being causal to layer 2/3 at delta frequency throughout the delta period, while the causality of higher frequency content (considered up to 15 Hz) alternates between deep and superficial, with layer 5 causality locked to the trough of the delta period where intrinsic bursting cells of the same layer are known to fire in characteristic delta locked bursts. Causality from superficial layers focused in the theta band, but appeared somewhat weaker than the reverse causality on the other half-cycle of the delta period.

We now take this robust observation further and consider whether the nesting and reversing causality at theta frequencies within delta has a correlate in energy flow within cortex.

Figure 6.5: Delta-locked interaction of layer 2/3 and layer 5

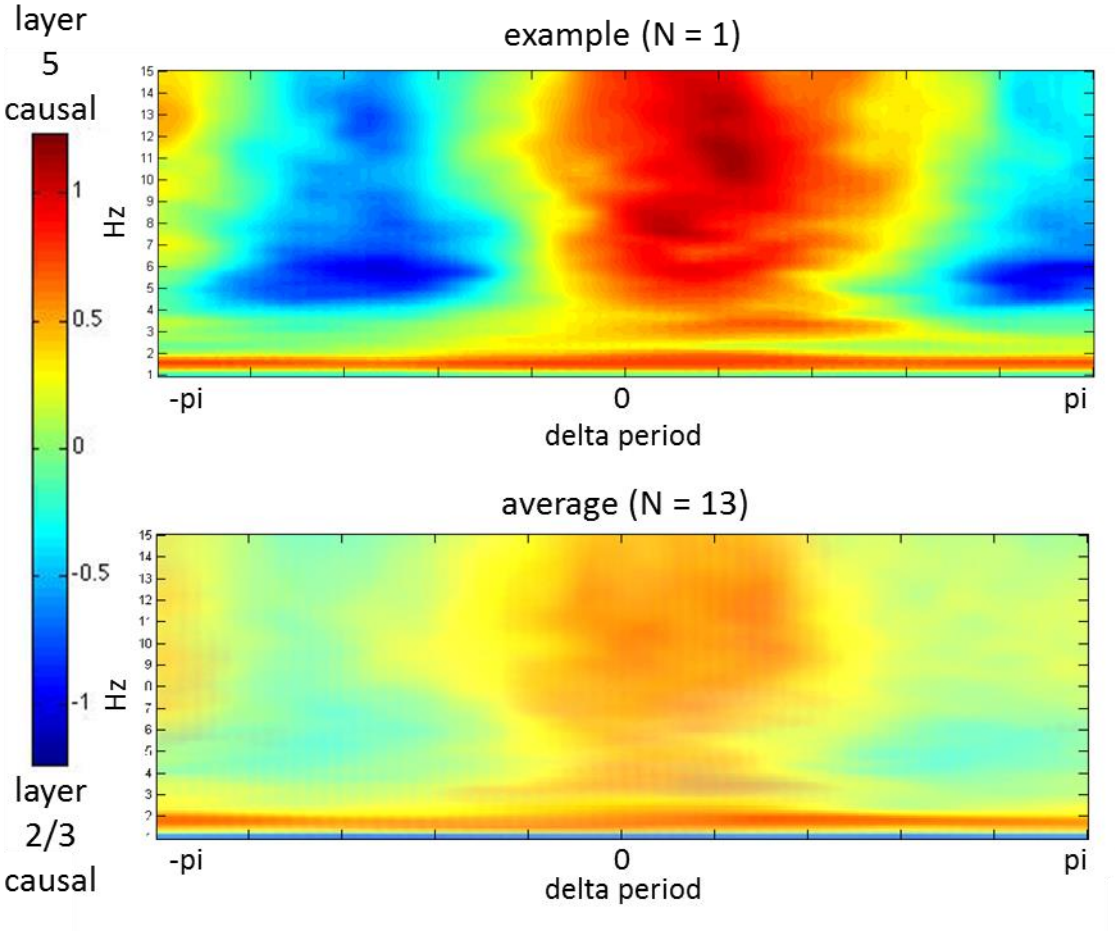


Figure 6.5: Delta-locked interaction of layer 2/3 and layer 5

The temporal interaction of layer 2/3 and layer 5 over a delta period average from 60 second recordings. Signals were decimated to 1 kHz and the resolution of the wavelet analysis was 0.1 Hz in the range 0.5-15 Hz. Top panel shows an example of layer 5 being causal throughout the delta period at delta frequency, while higher theta frequencies alternate between layers. The zero phase point corresponds to the trough of the delta period where bursting activity is focused. The bottom panel shows the average interaction pattern for N = 13. This result confirms earlier findings in our lab of layer 2/3 and layer 5 interaction during delta as discussed in chapter 2.

6.3.3 *Spatiotemporal energy flow in the neocortex*

The tentative prediction of the Helmholtz machine concept considered in section 2.2.3 was that a periodic interaction along layers in addition to the interaction across layers would be necessary to facilitate the wake-sleep learning approach. In extension, if the delta period sets the fundamental rhythm of this process it should occur twice in each period, one for each wake-sleep associated direction.

The examples of total energy flow in Figure 6.6, and total dissipation in Figure 6.7 give hope to this idea by revealing that while the local field potential has only a weak first harmonic, the energy flow and dissipation have much stronger relative first harmonics, and both fundamental and harmonic are locked to the field potential as indicated by the wavelet semblance measures. However, for dissipation, as revealed later in Figure 6.19, it turned out that this phase lock was in anti-phase in the case of the harmonic. Considering deep and superficial layers separately further revealed nested theta components specific to superficial layers; Figure 6.8 exemplifies this quite clearly for energy dissipation.

In Figure 6.9 to Figure 6.16 we investigated the average energy flow for an N of nine (slices) of each of the four areas, superficial primary, deep primary, superficial association, deep association, in eight time bins over the average delta period from 60 second recordings. In all cases the energy flow was strongest at the trough of the delta period, but the directions both across and along layers reversed on each side of the trough as seen in the first and last figures of the set. A smaller secondary reversal appeared twice on each side of the delta period peak, and in several cases (in order of confidence; red, blue, turquoise, purple, green) around smaller increases in energy flow; naturally associated with the secondary peak in the energy flow spectrum exemplified by Figure 6.6.

These patterns were clearly orchestrated by the delta rhythm, but broke down when filtering out higher frequency bands, revealing that nested rhythms were in fact contributing most of the actual energy flow. Progressively filtering out higher bands starting with the gamma band weakened the secondary pattern, ending with completely abolishing it below the theta band (4-9 Hz). The primary pattern was not visibly affected until below theta band, and in this case the pattern remained although considerably weakened. We illustrate two cases of this behaviour in Figure 6.17.

Figure 6.6: Example of total energy flow over one minute

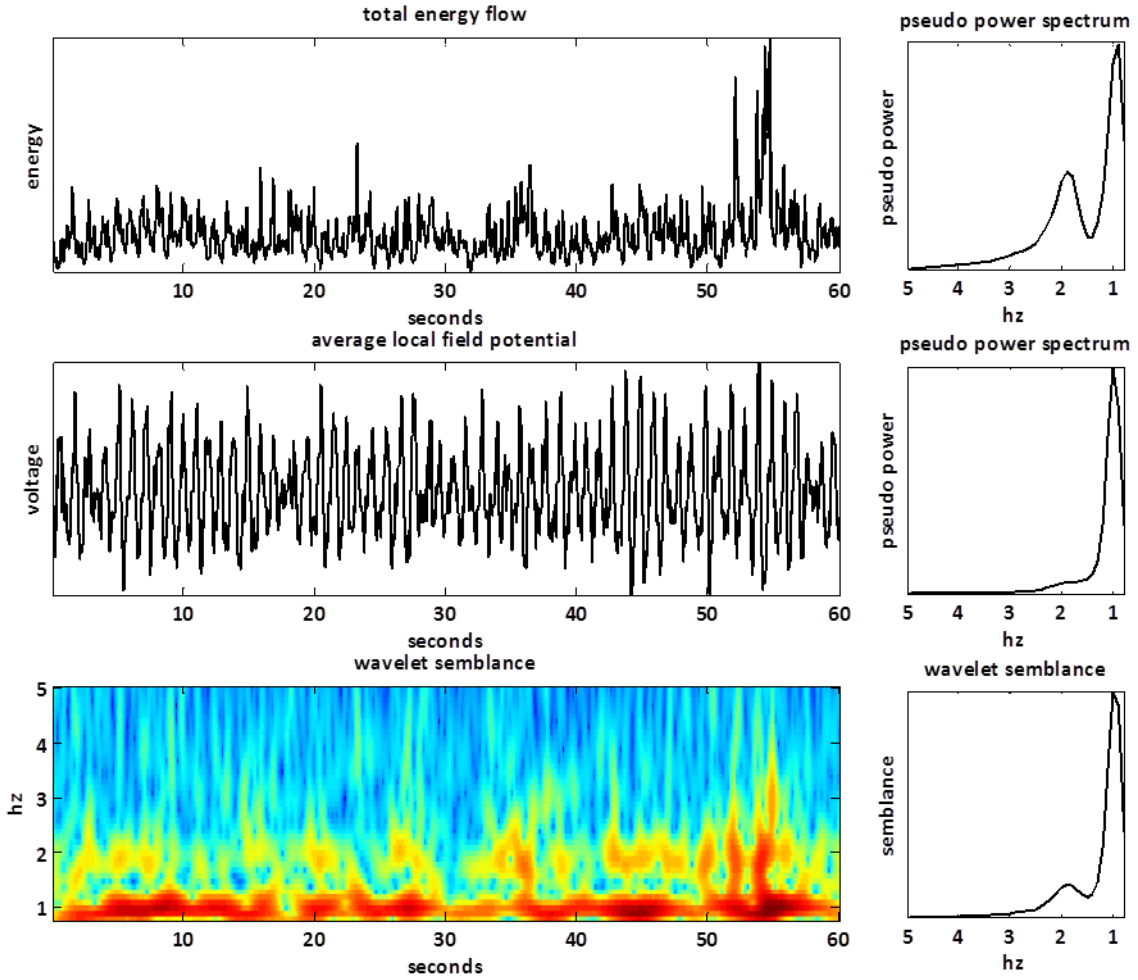


Figure 6.6: Example of total energy flow over one minute

The total energy flow and average local field potential with pseudo (wavelet) spectrums on the right. The spectrum of energy flow reveals a strong secondary harmonic in addition to the fundamental delta rhythm. The bottom panel shows the phase semblance with the local field potential indicating that energy flow is locked to the field delta rhythm.

Figure 6.7: Example of total energy dissipation over one minute

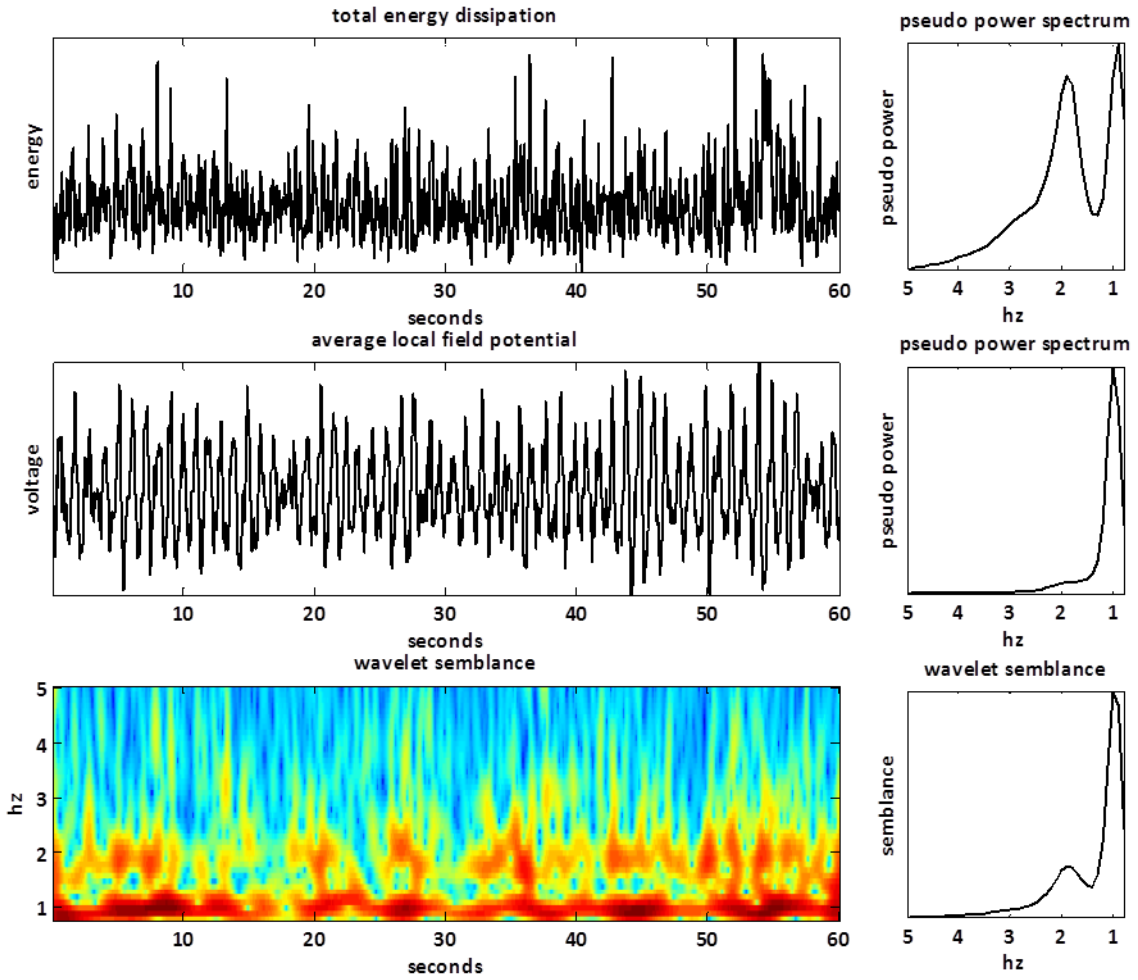


Figure 6.7: Example of total energy dissipation over one minute

The total energy dissipation and average local field potential with pseudo (wavelet) spectrums on the right. The spectrum of energy dissipation reveals an even more pronounced secondary harmonic in addition to the fundamental delta rhythm. The bottom panel shows the phase semblance with the local field potential indicating that dissipation is locked to the field delta rhythm.

Figure 6.8: Example of dissipation in deep versus superficial layers

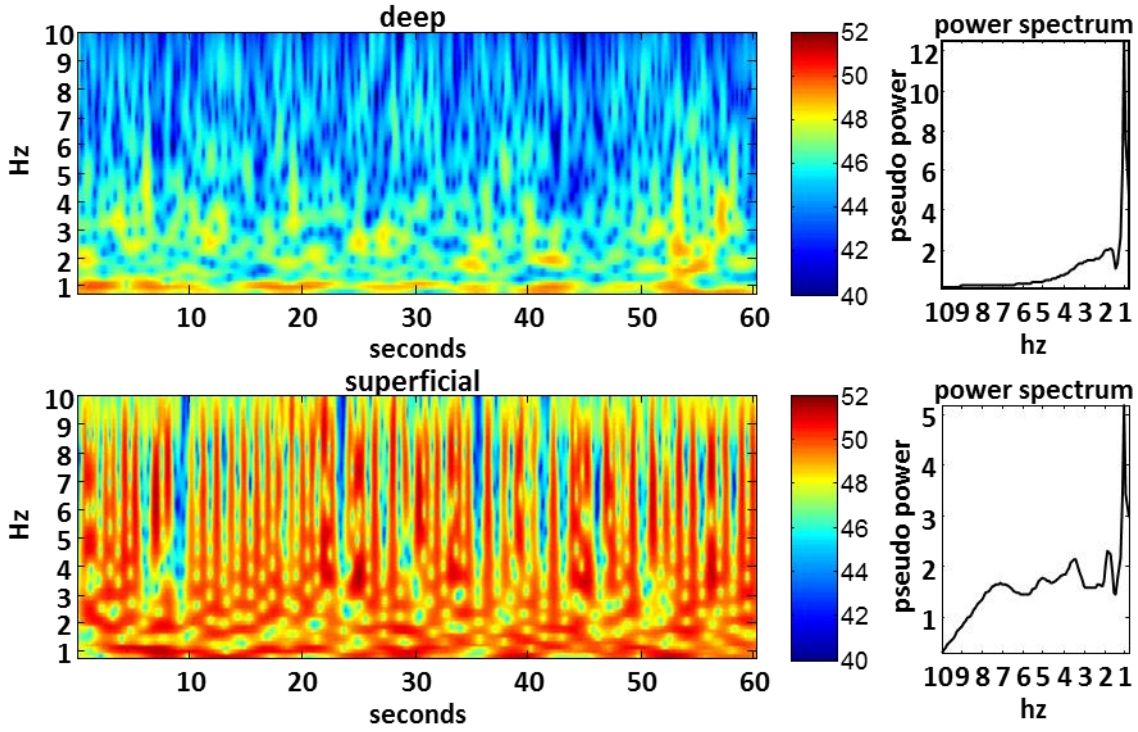


Figure 6.8: Example of dissipation in deep versus superficial layers

Considering field, flow or dissipation by deep and superficial layers separately revealed nested theta components in superficial layers. Here we show example wavelet spectrograms of energy dissipation over 60 seconds during the delta sleep-state model. Spectrums on the right sum over time to produce a condensed view of the frequency content.

Figure 6.9: Energy flow vectors for nine slices. First eighth.

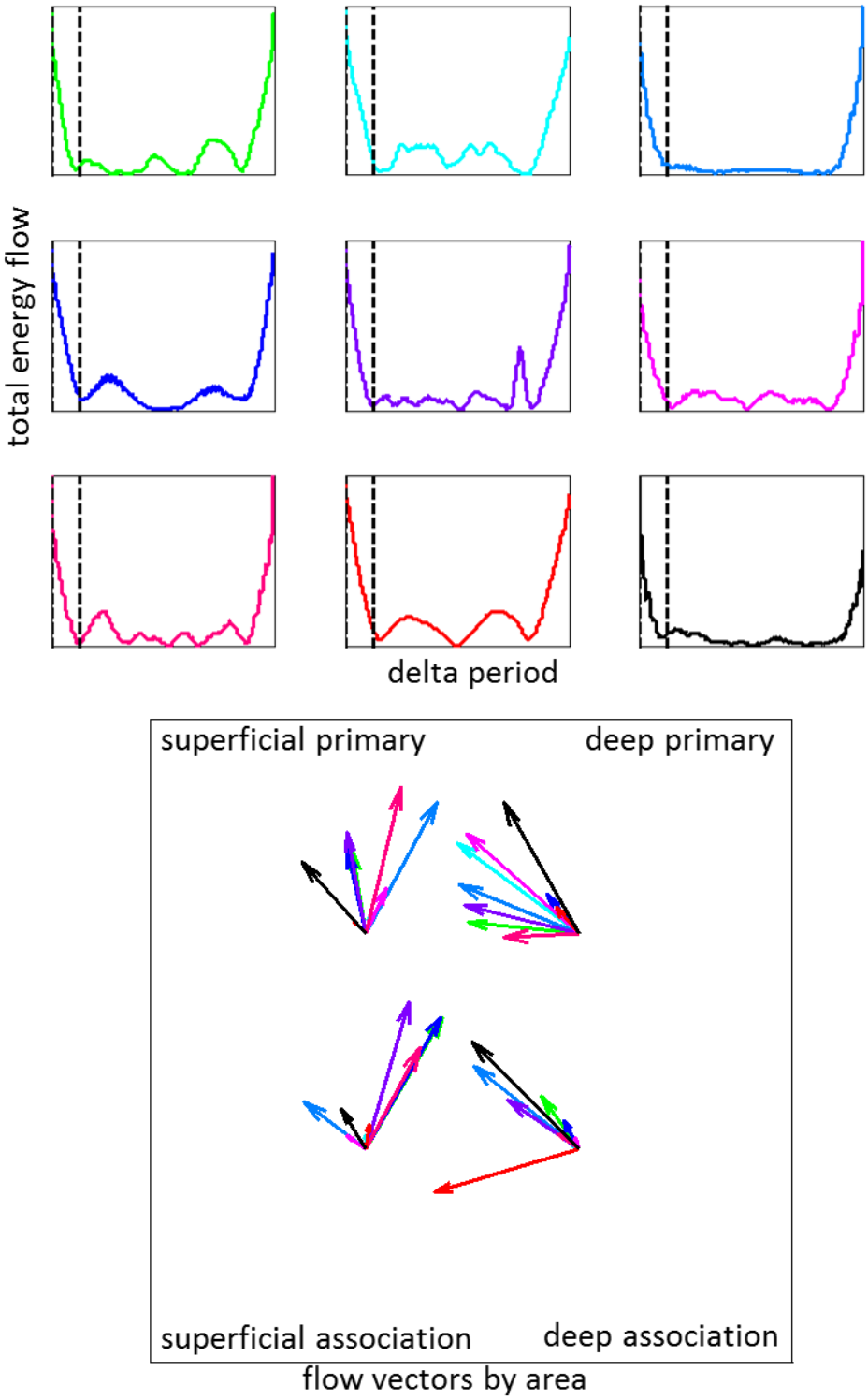


Figure 6.9: Energy flow vectors for nine slices. First eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. The energy flows are in reasonable agreement with deep layers projecting upwards, and superficial layers projecting forwards towards primary areas.

Figure 6.10: Energy flow vectors for nine slices. Second eighth.

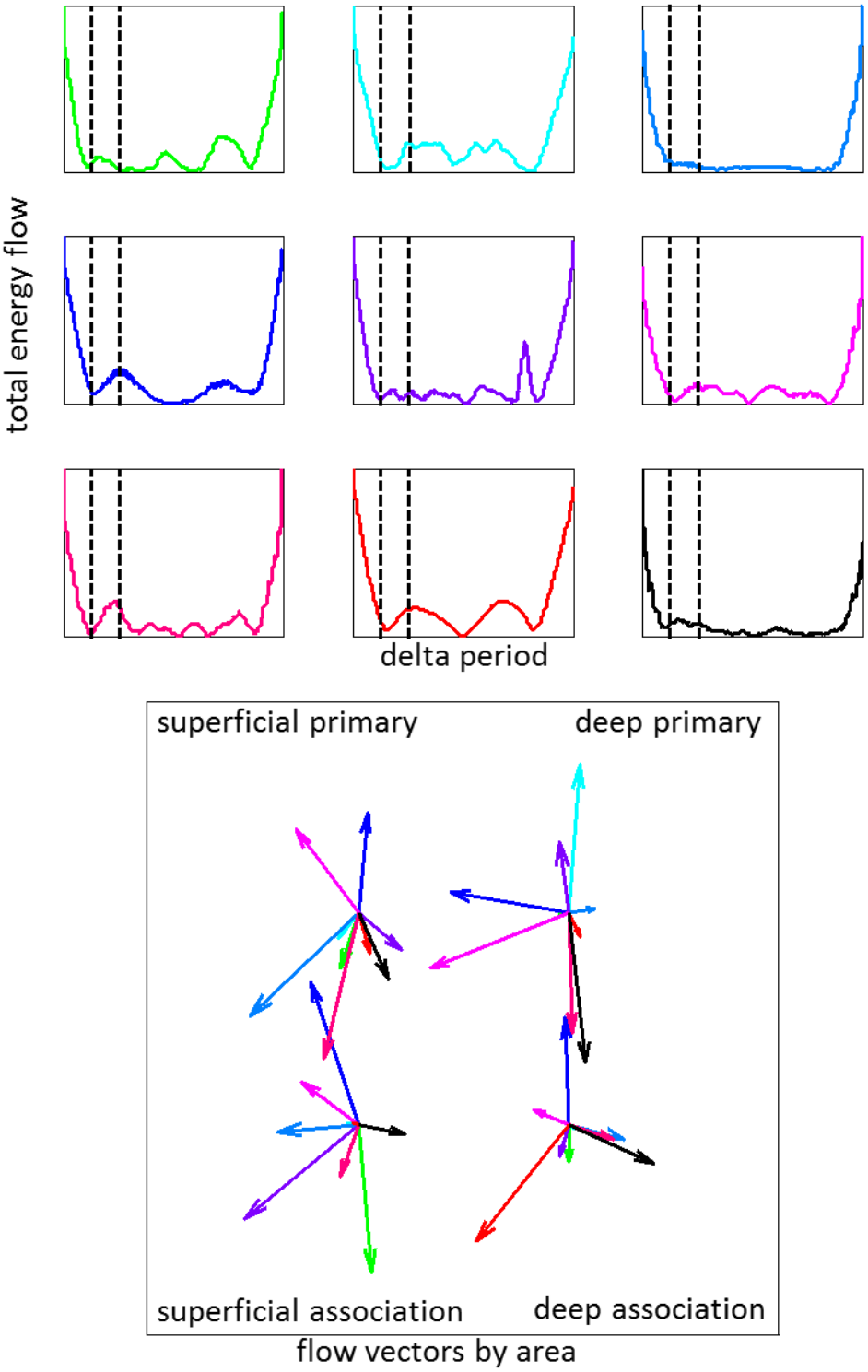


Figure 6.10: Energy flow vectors for nine slices. Second eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. This time bin does not reveal a clear pattern of energy flow and might be considered an intermediate state.

Figure 6.11: Energy flow vectors for nine slices. Third eighth.

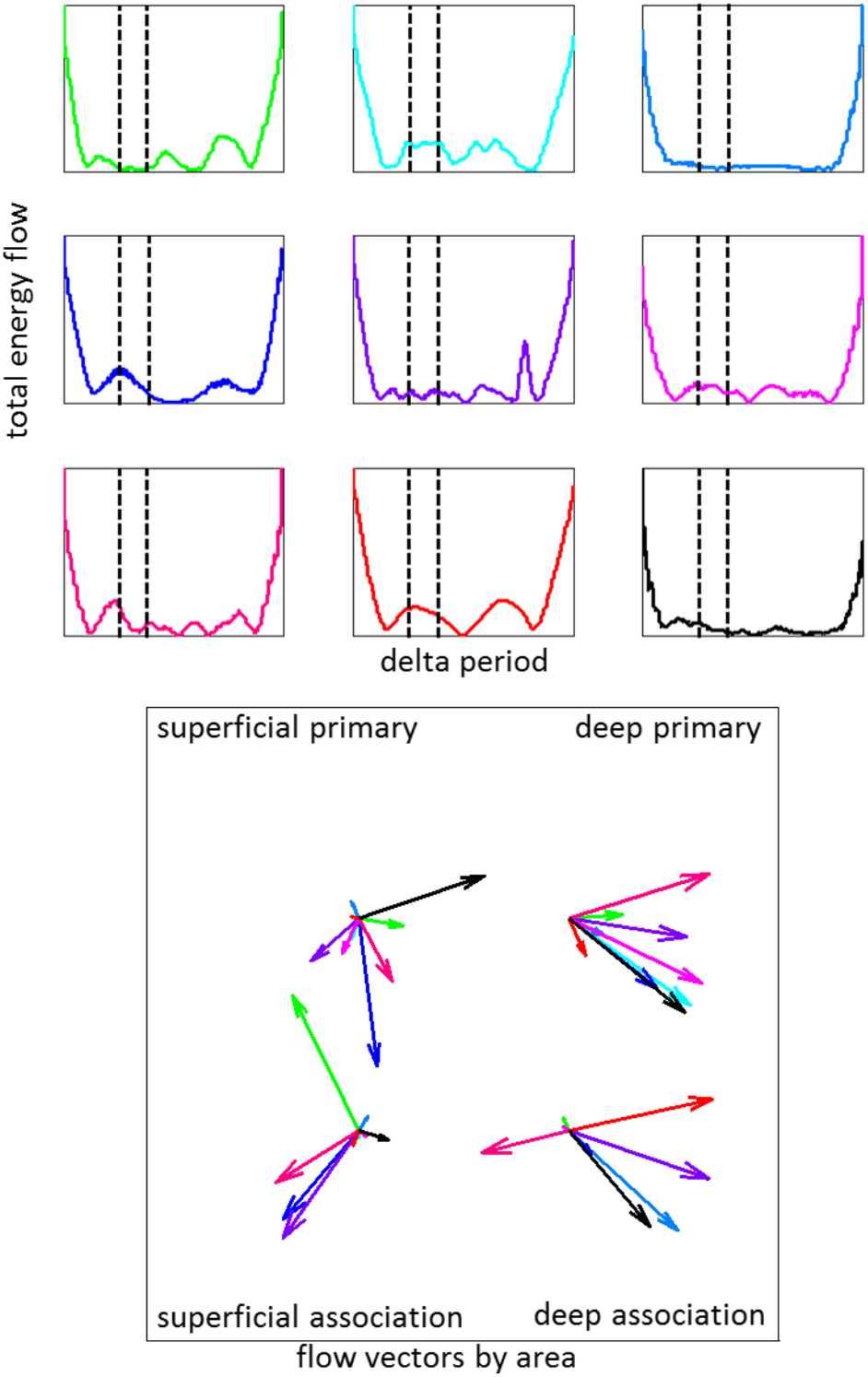


Figure 6.11: Energy flow vectors for nine slices. Third eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. This flow pattern reveals the first secondary reversal around the first secondary increase in energy flow in several cases.

Figure 6.12: Energy flow vectors for nine slices. Fourth eighth.

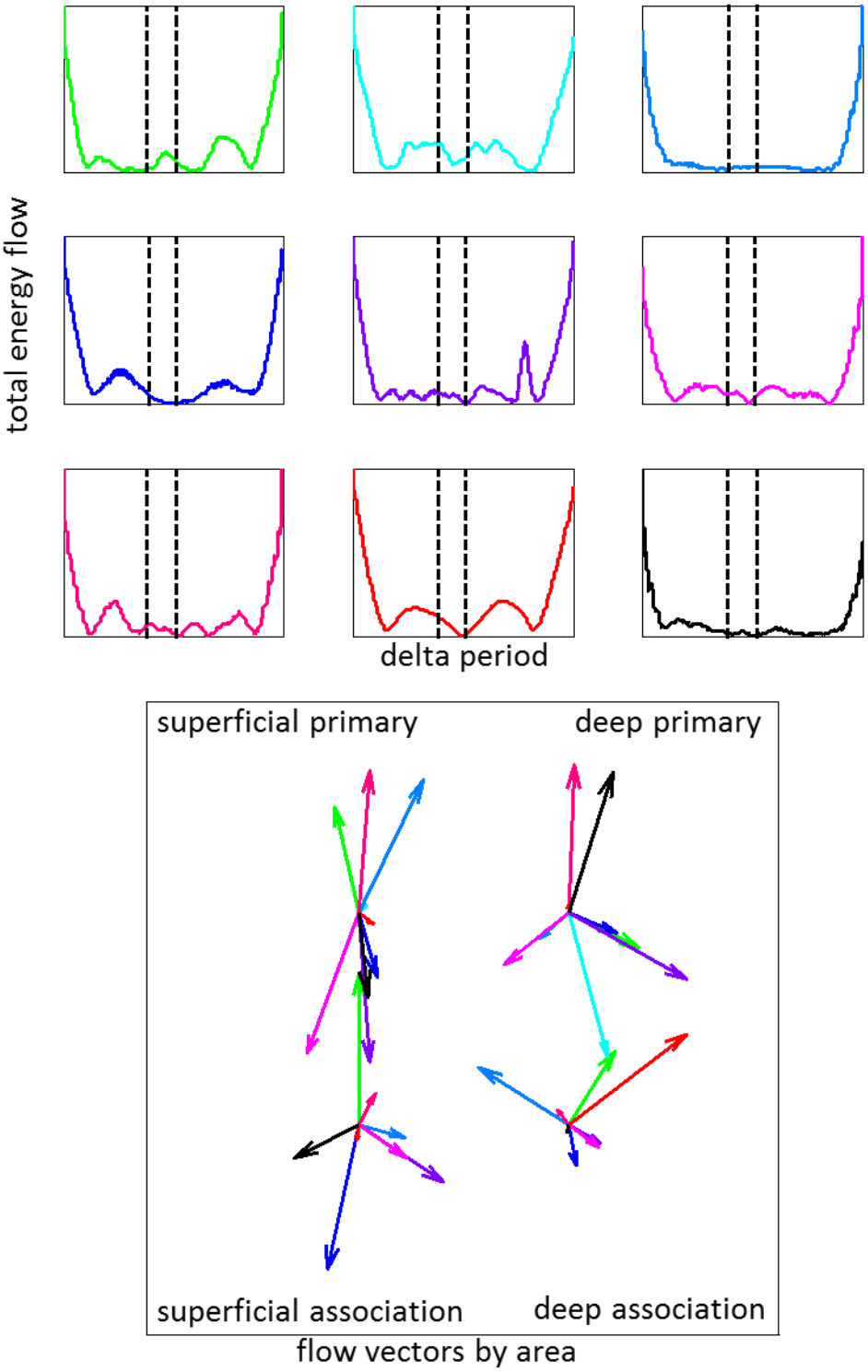


Figure 6.12: Energy flow vectors for nine slices. Fourth eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. This time bin does not reveal a clear pattern of energy flow and might be considered an intermediate state.

Figure 6.13: Energy flow vectors for nine slices. Fifth eighth.

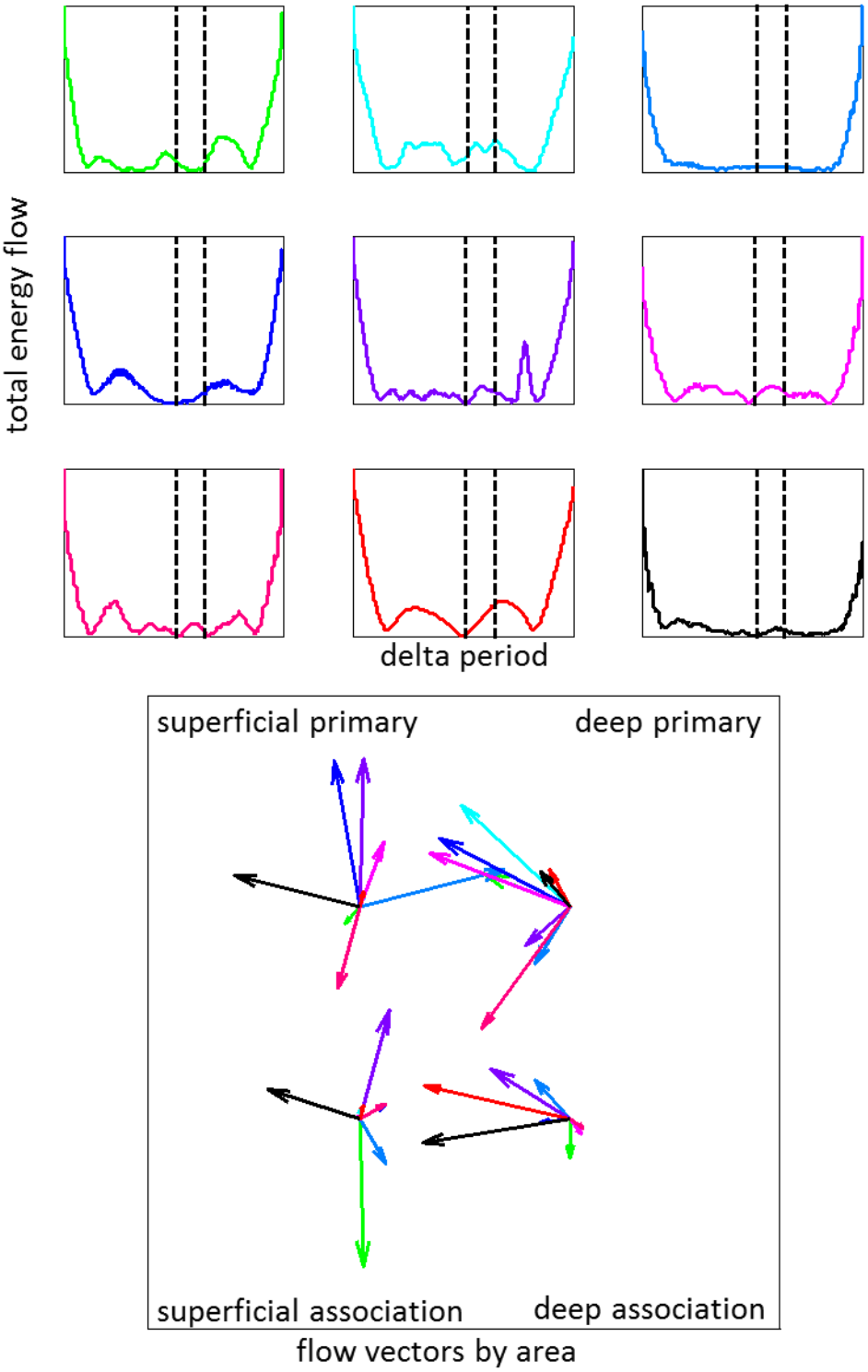


Figure 6.13: Energy flow vectors for nine slices. Fifth eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. This flow pattern reveals the second secondary reversal around the second secondary increase in energy flow in several cases.

Figure 6.14: Energy flow vectors for nine slices. Sixth eighth.

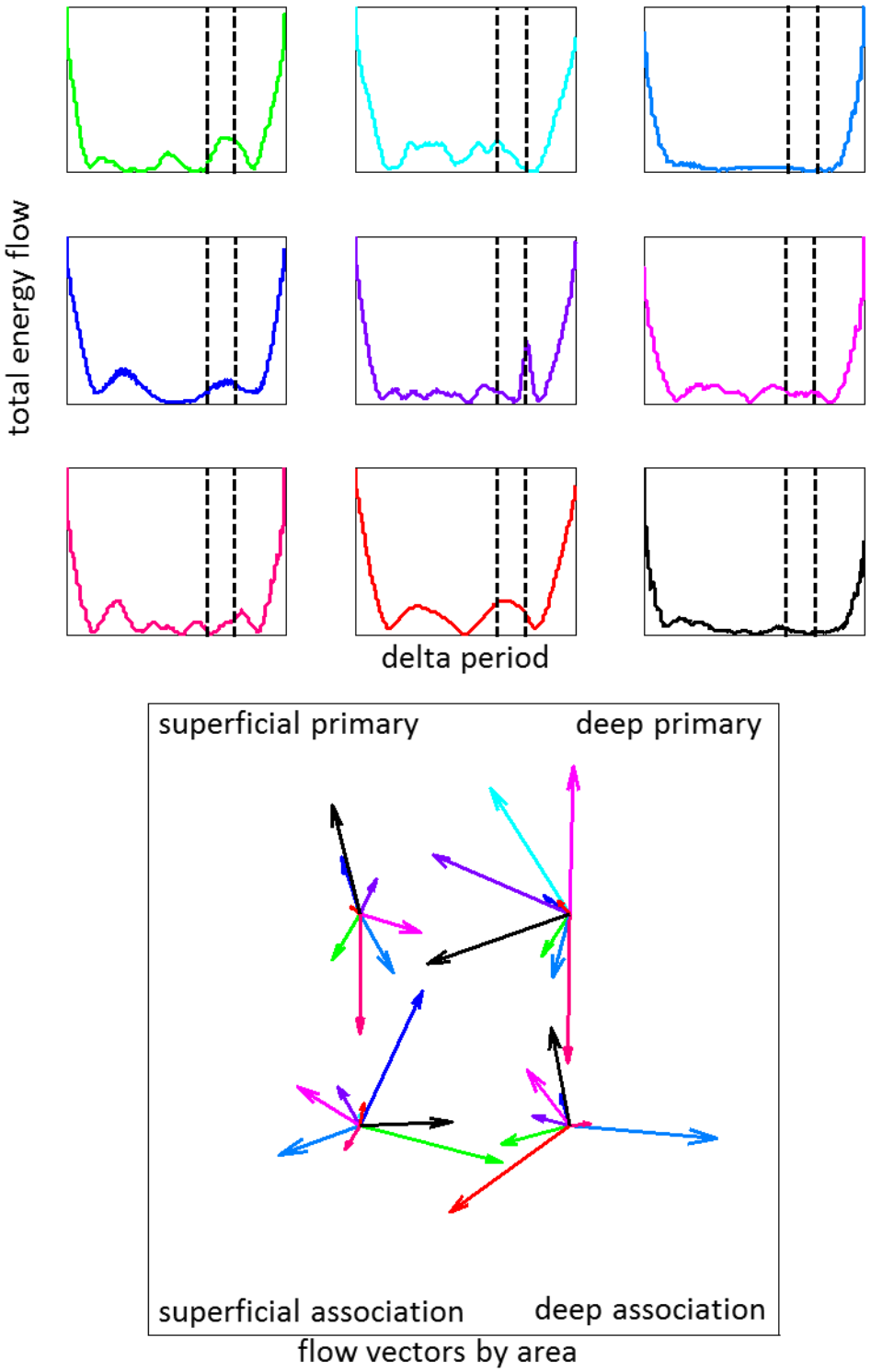


Figure 6.14: Energy flow vectors for nine slices. Sixth eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. This flow pattern also reveals the second secondary reversal around the second secondary increase in energy flow in several cases.

Figure 6.15: Energy flow vectors for nine slices. Seventh eighth.

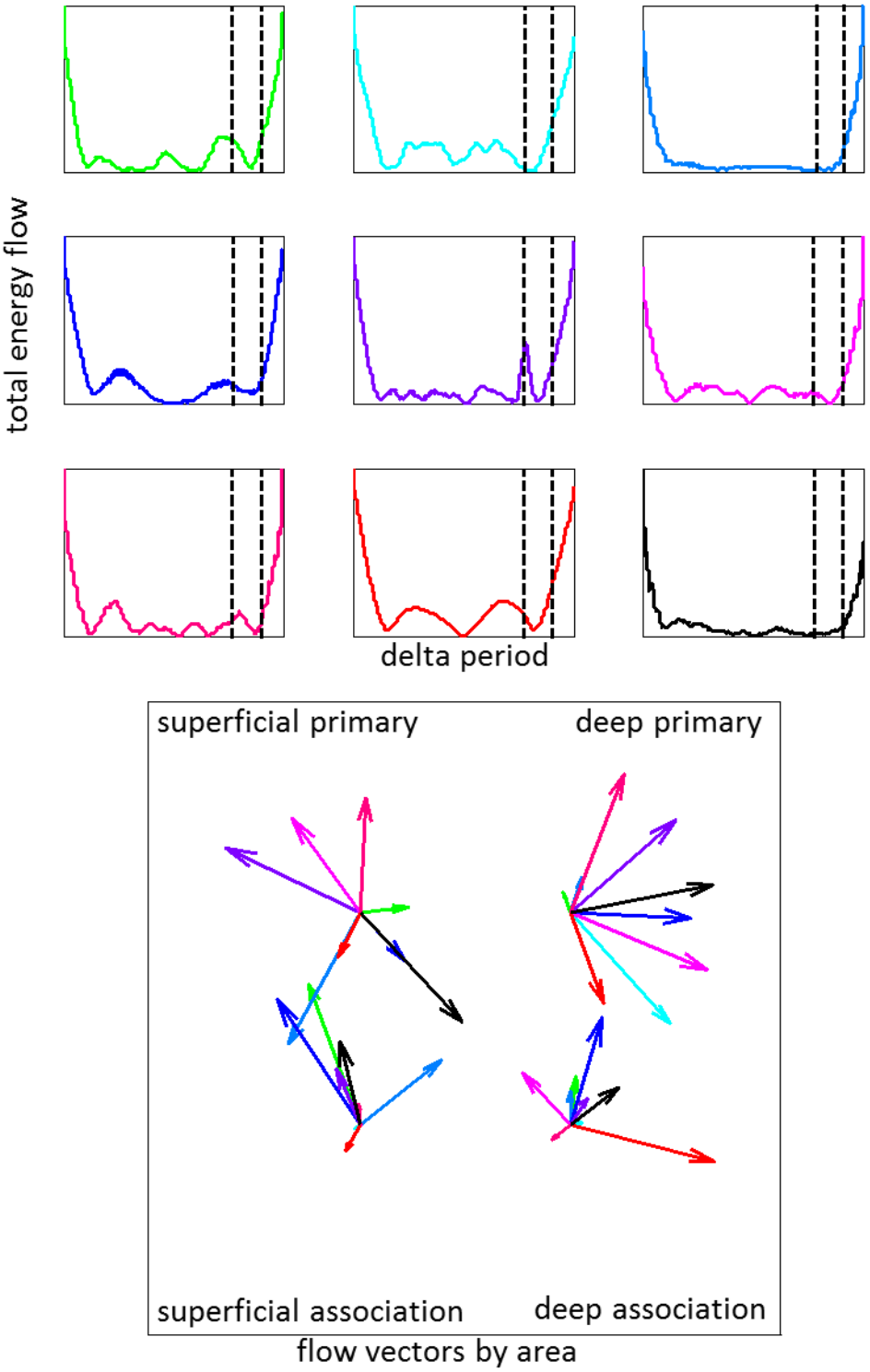


Figure 6.15: Energy flow vectors for nine slices. Seventh eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. This flow pattern reveals the other side of the second secondary reversal around the second secondary increase in energy flow in several cases.

Figure 6.16: Energy flow vectors for nine slices. Eighth eighth.

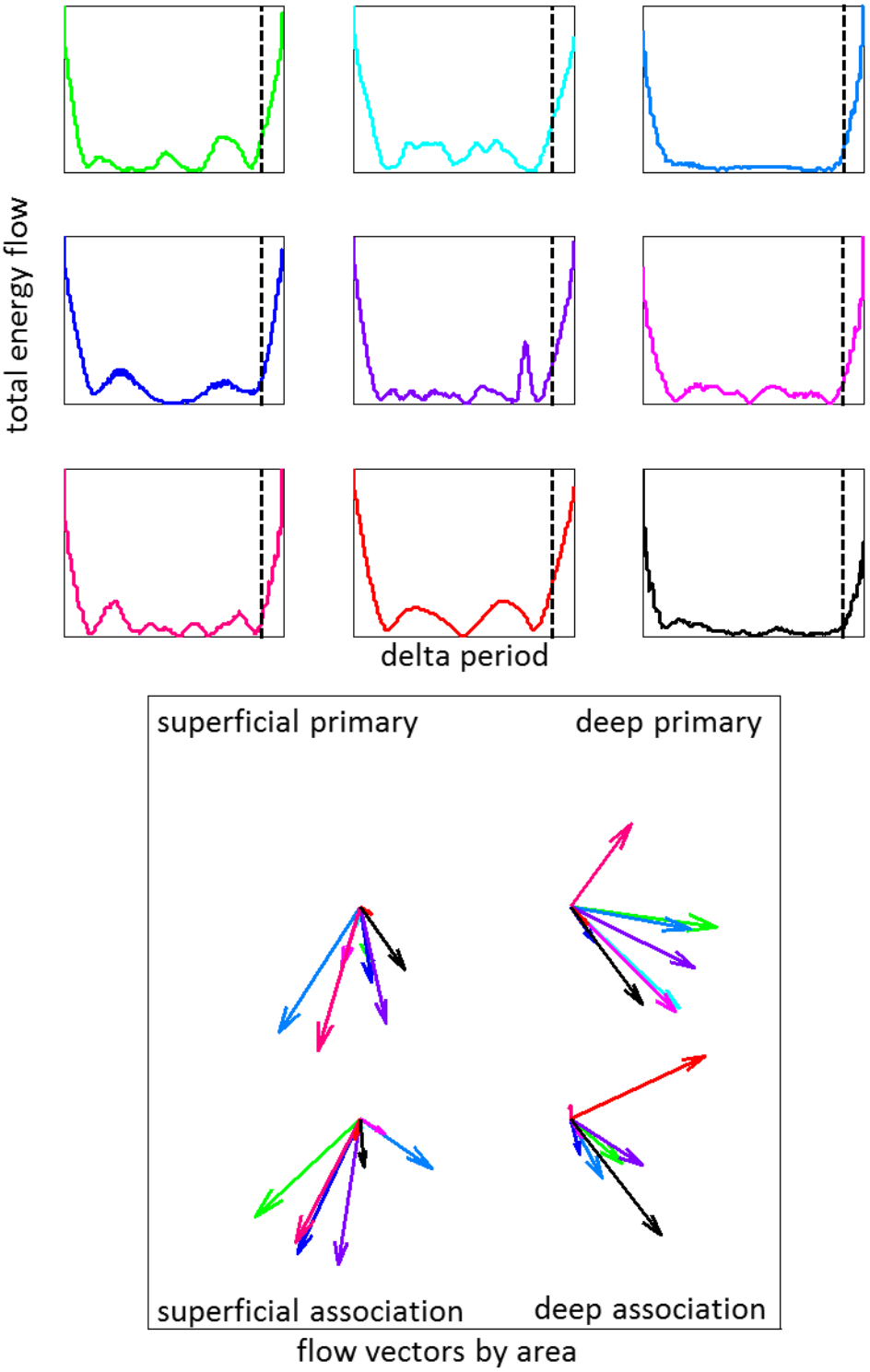


Figure 6.16: Energy flow vectors for nine slices. Eighth eighth.

Top nine plots show total energy flow over averaged delta periods for individual slices. The timescales, i.e. the length of the delta period, and the magnitudes are all scaled to unity, for uniform presentation, so are not on the same scales, except for the location of the eighth part of the delta period being examined, indicated by the vertical dashed lines. Bottom plot shows total energy flow vectors for each area for each slice corresponding by color to the above nine plots. Each set of four vectors with the same color are on the same scale. This scale is not shared across slices. The energy flows are in reasonable agreement with deep layers projecting downwards, and superficial layers projecting backwards towards association areas. This is the other side of the primary reversal at the other side of the strongest energy flow.

Figure 6.17: Higher frequencies contribute to slow flow patterns

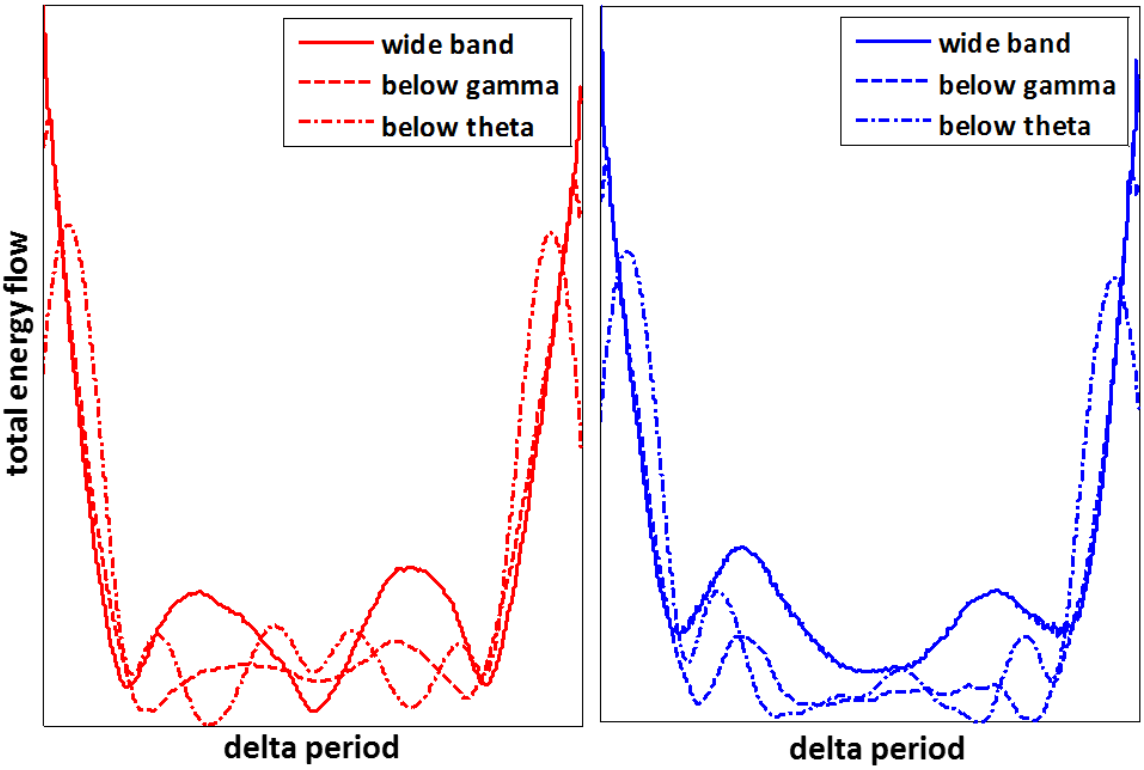


Figure 6.17: Higher frequencies contribute to slow flow patterns

The energy flow patterns of the previous figures appear mainly at delta and first harmonic frequencies, but the energy content of these patterns was not limited to these low frequencies, as illustrated in this figure. Progressively filtering out higher frequency bands appeared to diminish the flow patterns, especially the secondary first harmonic pattern. We picked the two clearest examples, corresponding by colour to the previous set of figures. The wide band went up to 500 Hz, below gamma means below 30 Hz, and below theta means below 5 Hz. This also illustrated that the overall patterns of energy flow were not trivially linked to the frequency-content of the same energy.

6.3.4 *Poynting vector flow/dissipation-ratio in wake and sleep models*

We were further interested in the possible computational differences between the wake- and sleep-state models in terms of dissipation of energy related to computation as discussed earlier in chapter 1. To investigate this aspect we calculated the ratio of total energy dissipation to total energy flow in each state, asking how much of the energy flow was used for dissipation as an indication of irreversible computation.

However, we first formed an overview of energy flow and dissipation in both model states. Figure 6.18 compares the spatial distributions of energy flow and dissipation between the wake-state and sleep-state models. The clear differences in terms of spatial distributions suggested that further quantification was warranted, and would likely reveal further details as to the nature of these more superficial (average) differences.

We then considered delta period averages of both energy flow and energy dissipation in the nine cases from the flow vector study above; a single pattern emerged in at least three cases as shown in Figure 6.19: While flow and dissipation were correlated at delta frequency, they appeared to be anti-correlated at the harmonic.

Instead of the absolute energy flow considered up till now, we finally decided to consider the real and imaginary parts separately in the hope that any differences might provide further insight, and perhaps help interpret the physical or computational meaning of imaginary dissipation and flow.

Figure 6.20 compares the real dissipation ratios between the wake-state model and the sleep-state model; only a small part of the energy flow is dissipated in both states, but the relative dissipation is stronger in the sleep state, suggesting that more information (potentially noise) is deleted (the irreversible process), consistent with the synaptic rescaling idea. In terms of entropy this result suggests that local internal entropy is lowered in both states, i.e. structure is created, by raising entropy in the environment by radiating heat (energy in an unstructured high entropy state).
Figure 6.21 shows the same comparison, but for the imaginary energy flow and dissipation. Again the two states are different, but this view reveals a negative imaginary dissipation ratio in the wake state model, which has no known biophysical interpretation to our knowledge; we will consider possible explanations in the discussion below.

Figure 6.18: Overview of energy dissipation and flow

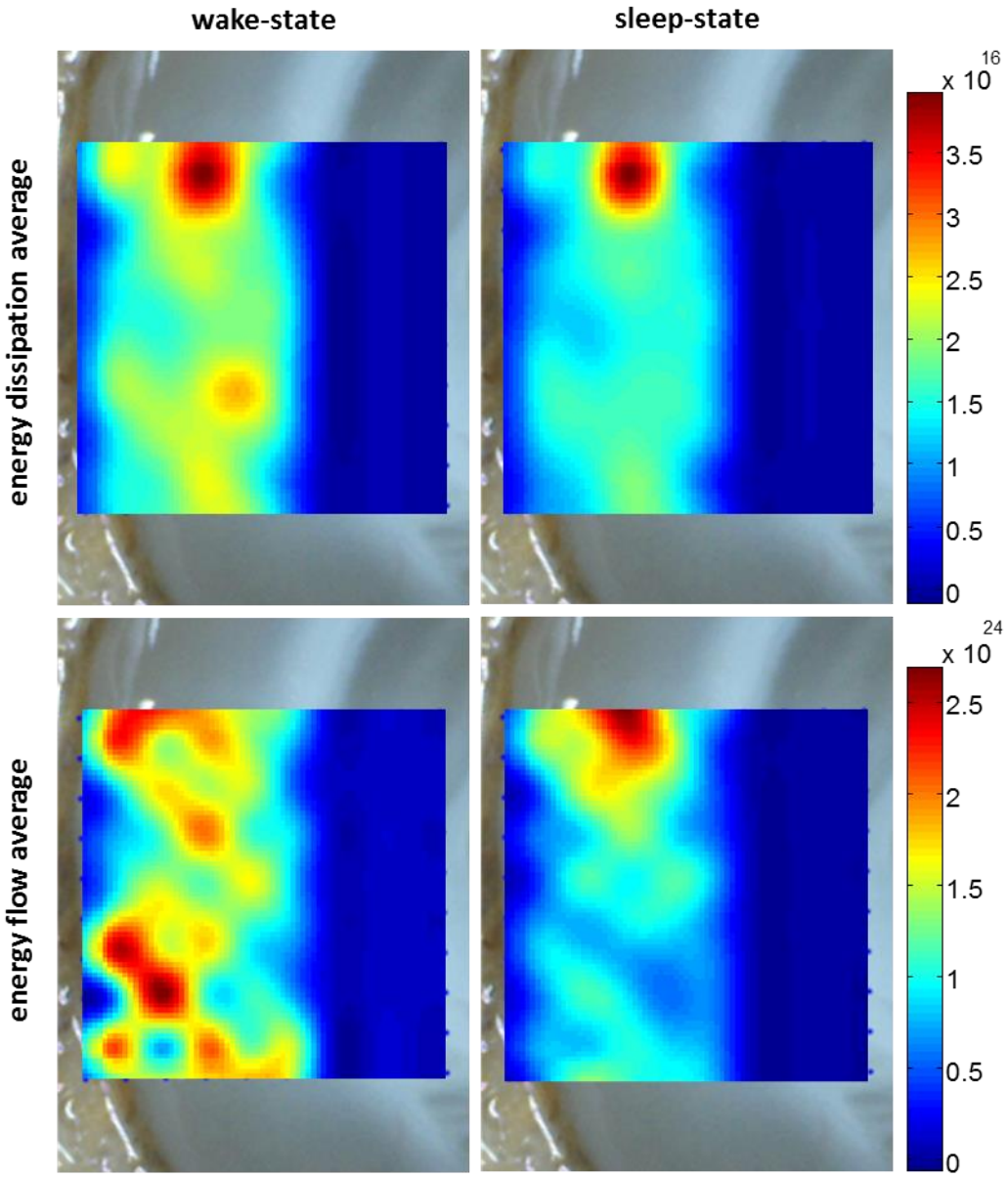


Figure 6.18: Overview of energy dissipation and flow

Using the same easily aligned data subset as in Figure 5.6 we formed average spatial maps of energy dissipation and flow in both model wake- and sleep-states. Both in terms of dissipation and flow the wake-state appeared to be more active; the spatial pattern was also more distributed, whereas the sleep-state seemed to focus towards deep primary areas, where delta is also generally stronger (compare Figure 5.7). These visually identifiable spatial differences set the stage for further quantification. The maps were created based on the original array resolution with subsequent cubic spline interpolation, i.e. not super-resolution, due to the computational demands.

Figure 6.19: Delta period average energy flow and dissipation

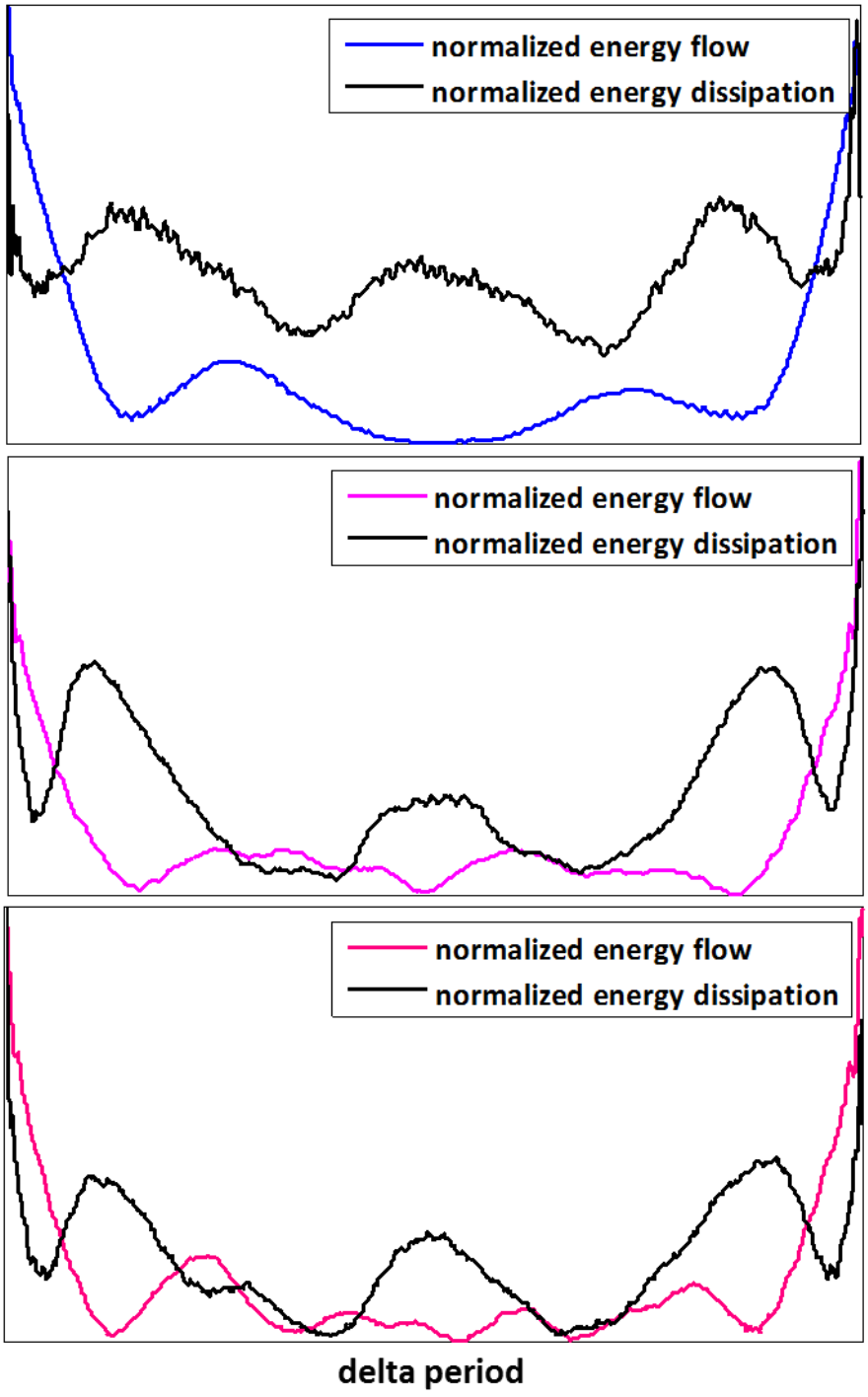


Figure 6.19: Delta period average energy flow and dissipation

Three examples of energy dissipation correlating with energy flow at delta frequency, but anti-correlating at the harmonic. Flow and dissipation were normalized to be presentable on the same scale, as energy flow was orders of magnitude stronger as shown for example in Figure 6.18.

Figure 6.20: Comparison of real energy dissipation/flow ratios

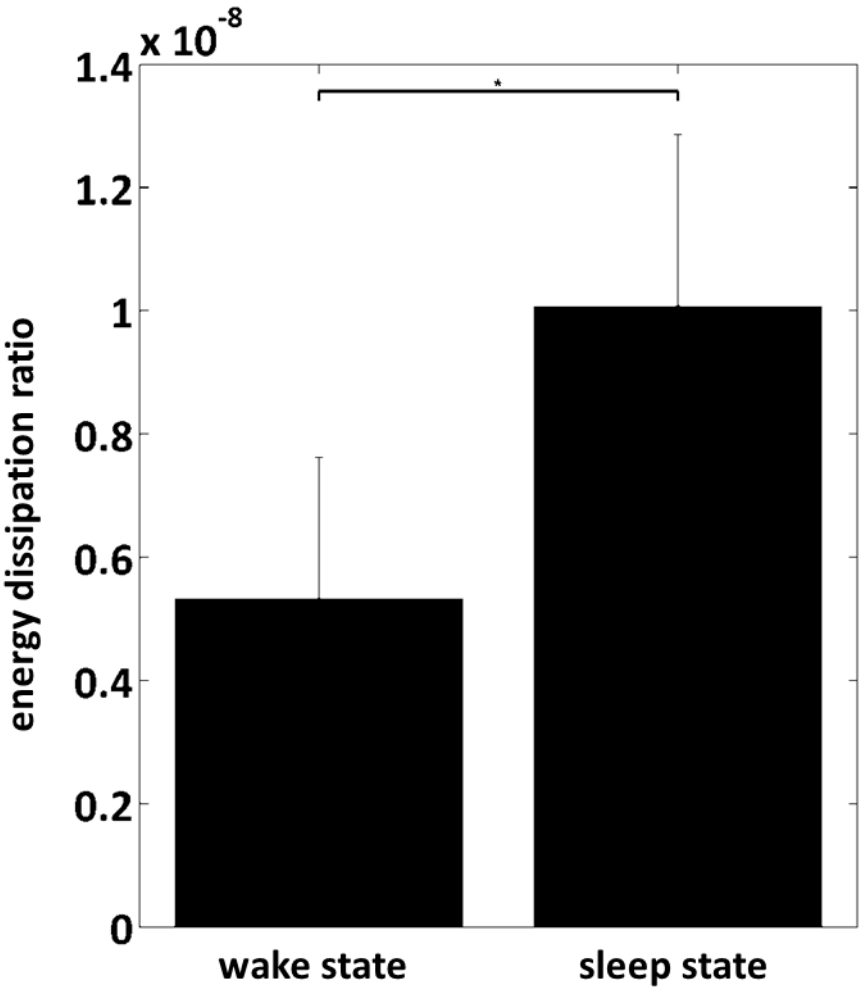


Figure 6.20: Comparison of real energy dissipation/flow ratios

Here we considered the ratios of the real part of total dissipated energy to total energy flow in the wake-state model and in the sleep-state model. These are totals of all channels in neocortex over one minute. Only a small part of the energy flow was dissipated in both cases, but the sleep state model dissipated a relatively higher amount.

Figure 6.21: Comparison of imaginary energy dissipation/flow ratios

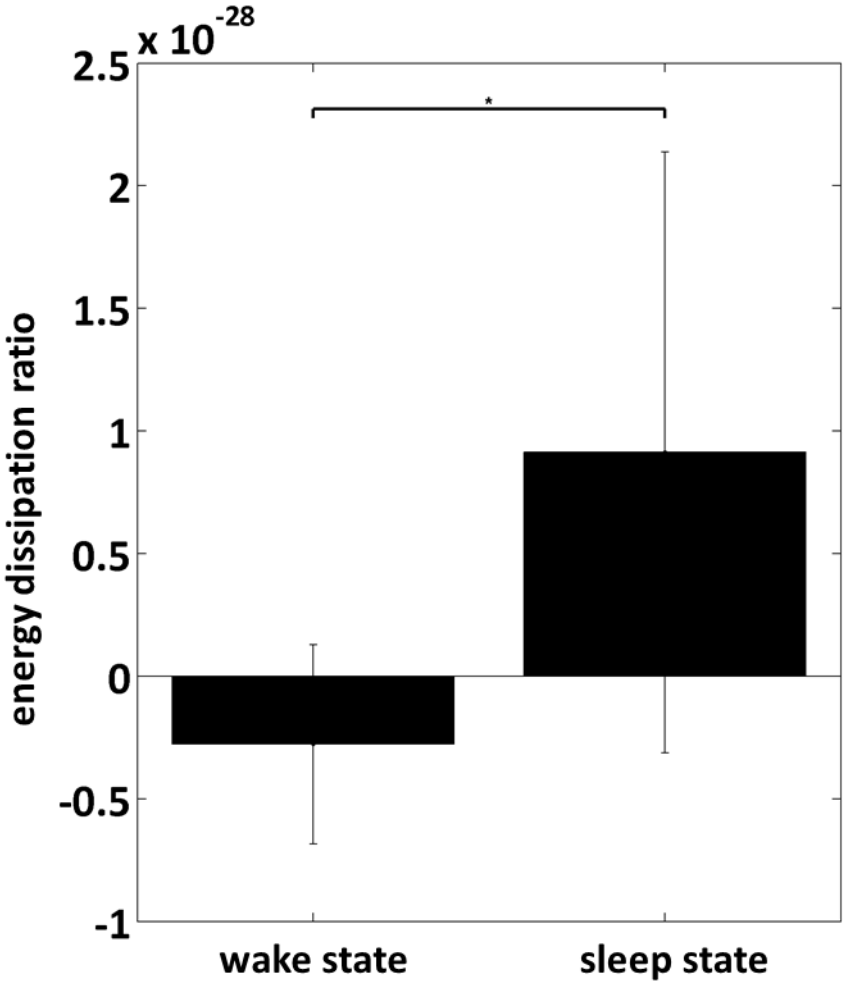


Figure 6.21: Comparison of imaginary energy dissipation/flow ratios

Here we considered the ratios of the imaginary part of total dissipated energy to total energy flow in the wake-state model and in the sleep-state model. These are totals of all channels in neocortex over one minute. Only a very small part of the imaginary energy flow was dissipated in both cases, but the sleep state model dissipated a relatively higher amount, and curiously the wake state imaginary dissipation ratio was negative, which has no known biophysical explanation.

6.4 Discussion

Building on the framework established by the results of the previous chapter, and using a new method of micro-electrode array analysis providing a physically based dissipation measure, as well as causality estimates in the form of energy flow vector fields, we investigated two aspects of the dynamics of a simplified wake-sleep model in the rodent neocortex at the interface between primary sensory areas and adjoining associational areas.

The question of ergodicity was relevant both for Friston's larger proposal of a free energy minimization scheme in the brain (Friston 2010), and (in particular for our current purpose) for the possibility that the free energy minimization of the Helmholtz machine concept might require similar statistical constraints, but also for the possibility that delta power during sleep reflects a classical annealing schedule, also implying ergodic dynamics.

The Helmholtz machine concept mapped onto the neocortex as described in section 2.2.3 provided two tentative predictions: 1) That reciprocal delta locked interaction across layers should be accompanied by interaction along layers as well. 2) That comparison of the outcomes of the two sets of weights occurs at twice the speed of the fundamental cycle of the wake-sleep algorithm.

6.4.1 *Ergodicity in the wake-sleep model states*

Firstly, we were unable to show that the dynamical system of the neocortex conforms to ergodic statistics in either of the two wake-sleep model states. This suggests that the system is not ergodic, but we have not shown this to be the case in a rigorous way, since the methods of determination did not apply to the considered spectrums of dissipation. If the states were ergodic the methods should have applied, but it is unknown if this implies non-ergodicity, since we could also not show the system to be specifically non-ergodic.

Nevertheless, power-laws have been found in many different settings, in human electroencephalograms (Grigolini, Aquino et al. 2009) and fMRI (He 2011), as well as in animals (Bedard, Kroger et al. 2006), and both based on field potentials and on extracellular multi-unit spike events (Beggs and Plenz 2003; Plenz and Thiagarajan

2007). These studies have mainly focused on concepts like self-organized criticality and scale-free networks, which are computationally attractive constructs (Pu, Gong et al. 2013), however not in as clear a sense as we consider ergodicity here, given the relations to tractability outlined in chapter 1.

It has already been noted that fitting a power law is not in itself a guarantee that it is the best suited distribution, and it can indeed be exceedingly difficult to tell the difference between distributions (Chu-Shore, Westover et al. 2010), and since we have used a more rigorous criterion (Clauset, Shalizi et al. 2009) than most studies, it is difficult to assess the disagreement between previous fits and the present failure to achieve significant fits at all.

In addition, we have of course considered a different measure than previous work; as far as we can tell there is agreement that the right measure is indeed dissipation of energy, and that neural activity (as spikes or field potential) reflects this dissipation to some extent. From a metabolic point-of-view this appears to be correct, since a spike reasonably has an associated metabolic energy cost, but from a computational point-of-view it is only true if the process is irreversible in the sense that the entropy of the energy is raised. By more directly measuring dissipation we hope to have bypassed such uncertainties. It is also interesting to note that peak energy flow occurs during the most active part of the delta duty cycle, peak negativity in deep layers corresponds to peak depolarisation of layer 5 intrinsically bursting cells (Carracedo, Kjeldsen et al. 2013). Of further interest is the abrupt reversal of flow seen either side of this peak. Given the temporal arrangement of nested theta rhythms in the delta rhythm this strongly suggests that the two theta periods identified previously may represent a biological substrate for horizontal as well as interlaminar interactions. The previous study cited above, and the Granger measures used here revealed a superficial-deep/deep-superficial layer repeated pattern of causal interaction. The energy flow data shown in this chapter show that, at a time in the delta period when the first theta period would be expected in superficial and deep RS cells (see figs 2 & 3 of (Carracedo, Kjeldsen et al. 2013)), energy is being transferred from associational to primary sensory areas (‘top down’). In contrast, during the second nested theta period energy flow is reversed, primary sensory to association cortex (‘bottom up’). The relationship between these findings and both traditional ideas about bottom-up and top-down processing, and the Helmholtz machine are considered in more detail below.

6.4.2 *Helmholtz-like interaction across and along layers*

Our basic paired-electrode investigation of causality across layers agrees with previous work (Carracedo, Kjeldsen et al. 2013), and provides a more detailed demonstration through the complex continuous wavelet based non-parametric Granger causality time-frequency maps. We did not specifically investigate the relation between causality estimates and energy flow, but the primary pattern of flow agrees with the pattern of causality reversal (compare Figure 6.5 with Figure 6.9 and Figure 6.16, and see previous section). However, we also found more detailed information in the micro-electrode array approach in the form of a secondary peak in the energy flow spectrum (which was much weaker in the local field potential), and in the secondary reversals occurring with this harmonic. Although the harmonic was also in the delta band we found that the flow patterns of first, and in particular second reversals at harmonic frequency were powered by higher nested theta and gamma frequencies, again highlighting the idea that although delta is the main orchestrator it is not the only functional aspect in this process, the whole being far more spectrally distributed. This was also apparent in the fact that dissipation appeared to be anti-phased with the flow at the harmonic (Figure 6.19), suggesting different computational roles.

Overall these patterns agree with the predictions made based on the Helmholtz machine framework, that there should be a flow of information along layers as well as across, and that there should be major computational events occurring with twice the frequency of the fundamental rhythm of the wake-sleep algorithm (not the biological wake-sleep cycle). Currently the biological origin of the weaker, second flow reversal is unclear. The timing within one delta cycle suggests coincidence with a predominantly hyperpolarised state in the main delta generating neurons in layer 5. However, the nesting of theta frequency activity demonstrated in (Carracedo, Kjeldsen et al. 2013) was seen to be extremely labile; occurring over a very narrow range of membrane potentials in the cells concerned. Excitation above this narrow band resulted in near-persistent theta activity. It is tempting to suggest, therefore, that the more labile (in terms of occurrence in each slice) second flow reversal may reflect continued, or at least prolonged theta activity in deep and superficial RS cells.

These findings might also fit well with ideas of interacting bottom-up and top-down processes, most commonly studied in the visual system (McMains and Kastner 2011), but to our knowledge the specific delta orchestrated alternating pattern along layers

found here has no neurobiological precedents. Alternating patterns have been found in slightly different settings though; for example, in wake-state monkey temporal cortex between sensory processing and putative memory processing (Takeuchi, Hirabayashi et al. 2011), and in rat auditory cortex between sensory-evoked and spontaneous activity (Sakata and Harris 2009).

6.4.3 *Energy flow and dissipation*

The present finding that the percentage of energy flow dissipated in the sleep-state model is larger than in the wake-state model appears superficially to conflict with the fact that metabolic demands are reduced in sleep (Vyazovskiy and Harris 2013), but these are not the same measures; in fact, the energy flow itself was also found to be higher in the wake-state model. However the percentage dissipated further indicates how much of the energy is used in computationally irreversible processes, at least in the framework established in the first three chapters. Although the metabolic cost of potentiation cannot be recovered (it is not reversible), and although there must also be a metabolic cost of depotentiation, from a computational entropy perspective the establishment of connections is reversible, but the deletion or rescaling of connections cannot be undone (unless the original stimulation is reintroduced or perhaps replayed). So this suggests a strong computational role for sleep, but as we saw in the previous chapter this is only in connection with equally important computations in the preceding wake-state in the form of spike-train-to-field coherence, presumably to predetermine which connections to keep alive (by replay) during sleep depotentiation. As we saw in the previous chapter, this predetermination appeared to be, in part at least, related to the degree of spike-train-to-field correlation in the high gamma band during stimulation.

Surprisingly there is no accepted interpretation of the imaginary part of the instantaneous Poynting vector (Czarnecki 2006; Emanuel 2007; Balci, Hocaoglu et al. 2010; W. Duana 2012), although it is a fundamental concept in electromagnetism, constructed merely as the cross product of electric and magnetic fields. Confirming this peculiar situation is the fact that it is mirrored in its acoustic counterpart where an agreed interpretation also appears to be missing (Domenico, Nicola et al. 1996; W. Duana 2012). The imaginary part of the time-averaged (as opposed to instantaneous) Poynting vector is normally considered to reflect energy flow associated with non-radiating electromagnetic waves (the real part being the radiating waves), but the

interpretation is somewhat fussy in the sense that the direction of flow is not explained by such standing waves; it is normally considered to represent back and forth flows, but these flows do not equal out, and the net flow and its direction is not really explained. Again the situation is mirrored in acoustics where the imaginary part of the time-averaged intensity is also associated with standing waves with some difficulty (Stanzial, Prodi et al. 1996).

We can perhaps venture an explanation as to why it has been so difficult to come to grasps with this issue in both fields; there appears to be an intrinsic interpretational problem with the use of complex numbers, although immensely useful in general, they introduce the imaginary unit, which is almost by definition uninterpretable. This only really becomes apparent when one considers the alternative of the geometric algebra introduced in section 3.4.2.

It is however not difficult to come up with a possible interpretation of the imaginary part of the instantaneous Poynting vector, since it is already known that energy transport happens in at least two distinct ways: The first is via traveling waves corresponding to the radiative real part as mentioned above. The second is via resonant energy transfer, a phenomenon which occurs in the exponentially decaying near-field of resonators whose standing waves oscillate at the same (or harmonic) frequency (Kurs, Karalis et al. 2007; Karalis, Joannopoulos et al. 2008). Importantly, resonant energy transfer is however not limited to only near-field distances, since any series of resonators, each only within near-field distance of the previous and the next resonator, will form a bridge or tunnel between any points on the series like a waveguide. This is important because of the interpretation of on one hand the real radiative part as a locally progressive, propagating or traveling effect, and on the other the imaginary resonant part as a non-local and virtually instantaneous phenomenon. In biological terms these parts would correspond, respectively, to propagating waves of activity (i.e. poorly synchronised, percolative neuronal activations), and to synchronous activity where multiple cortical subregions, and the neurons therein, are tightly, temporally coactive. It is tempting to suggest that the negative imaginary part of the dissipation/flow ratio seen in the wake-state model may represent the use of such synchronous connections via the gamma rhythms generated by the stimuli (Traub, Whittington et al. 1996).

Let us further attempt to apply this interpretation to the dissipation ratio findings from above: Again let us view the situation from the entropy perspective related to dissipation as previously described; then in the wake-state local entropy is lowered presumably by the creation of local order (synaptic potentiation), but this enhanced local order at the same time breaks the global coherence signified by a negative dissipation ratio or equivalently an increase in entropy in a global perspective. As discussed earlier there is always the question of scope when considering entropy, there is no contradiction in lowering entropy in a local perspective and at the same time raising it in a larger perspective, presumably it is the enormous increase in the Sun at the solar system level, that allows biological life on Earth to lower entropy in a local perspective without per se going against the Second Law of Thermodynamics (Penrose 2001).

In the sleep-state on the other hand both local and global perspectives display a decrease in entropy (via positive dissipation). The suggestion is then that the sleep state facilitates a global or at least long-range structuring of neocortex, which is lacking in the wake state. The suggestion is also that the physical mechanism for this process is a non-local information processing perspective carried by resonant energy transfer under the straight-forward interpretation of the Poynting vector given above. Potentially this view can help unify the idea of contextualization during sleep with the conflicting finding (Esser, Hill et al. 2009) of uncoupling of propagating longer-range connections, as discussed in the previous chapter, since the suggested resonant connections are specifically not propagating.

The suggestion that global entropy increases in the wake-state fits nicely with the increase in sleep pressure, but might be at variance with ideas of communication through coherence (Fries 2005); this might depend on whether there is a cost associated with such communication in the sense that coherence is a resource that is spend on communication. This point of view would provide a semi-mechanistic explanation for sleep pressure, but we are not aware of direct evidence that sensory-related coherence and sleep pressure anti-correlate during wakefulness. We do know that neural activity increases with sleep pressure (Mignot and Huguenard 2009), which could indicate a compensatory mechanism; however, this increase was also associated with increased synchrony (Vyazovskiy, Olcese et al. 2009), unfortunately it was unclear to us whether this increase was normalized by the level of neural activity. From an abstract (quantum)

computational point of view sensing the environment is associated with a loss of coherence in exchange for local physical manifestations, while building a coherent resource in the first place requires de-coupling from the environment, consistent with the characteristics of the wake-sleep cycle. We will reconsider such processes in the next chapter.

It should be noted that the concept of resonant energy transfer goes against the view of the brain as an electrical circuit (Kandel, Schwartz et al. 2012), but for an electrical engineer this view is not necessarily acceptable in the first place, since in practice electrical circuits of this scale and nature are extremely difficult to keep isolated, and it is as much an antenna design problem as a circuit design problem. Some neuroscientists might like to counter with the fact the axons are insulated by myelin, but that is not really how insulation works; even if you are in a Faraday cage, but leave the door open the slightest you will have a cell phone signal, meaning that any gap in insulation defies the purpose of preventing antenna characteristics, indeed not all axons are myelinated in the brain, and even those that are can have many hundreds of microns of unmyelinated initial segment before insulation starts. This does obviously not deny that there is another important purpose of myelination, maintaining the axonal chemical micro-environment for example. Further, it has previously been noted that equivalent circuit models contain resistors, which imply a heat-loss associated with action potential propagation that does not match actual measured values (Heimburg and Jackson 2007); the proposed resolution of this conundrum was to consider it a reversible heat-loss, i.e. negative dissipation as the process is reversed, so the present finding of negative dissipation might not be without theoretical precedents.

These final ideas are obviously very tentative and hinge on several equally tentative interpretations, which cannot be said to be proven in any way; further, the interpretations might be coloured by our desire to find a physical substrate in the brain for processes, which could somehow go beyond algorithmic limitations discussed earlier. So let us finally discuss whether this could be the case in the hypothetical framework built so far: From the exposition given in the background chapters it is clear that we are looking for an effect that might banish the Axiom of Causality, but would allow a non-axiomatic causality to re-emerge naturally from the remaining axioms. The obvious questions then revolve around the nature of resonant energy transfer. We might consider it in a semi-classical framework like Wheeler and Feynman's absorber theory

(see section 3.3.2), or even in an acoustic framework where the evanescent wave-mechanical coupling equations are still the same, but the most general perspective is given by quantum theory, where the same phenomenon is also the foundation of quantum annealing, which we favoured in section 2.2.3.

We would expect to be able to calculate how long, say, an electron spends in the tunnel during a quantum tunnelling process, however quantum mechanics provides three possible answers, all of which are unsatisfactory: The first is simply that we are not allowed to ask, the second is exactly no time at all, and the third is some amount of imaginary time. This again highlights the issues with interpreting the imaginary unit, which this example has been used specifically to illustrate in the geometric algebra context (Gull, Lasenby et al. 1993a). Unfortunately the geometric algebra version does not so far provide an answer as well, so we are left with experimental measures, which also have great difficulties in finding agreement amongst themselves. The emerging consensus however appears to be that it does travel faster than light, while no information can be communicated faster than light by the process (Nimtz 2011; Bancal, Pironio et al. 2012); the same argument as for quantum teleportation, perhaps suggesting a relation. However, detecting the arrival of an electron (to measure the arrival time) logically entails the communication of information, so the argument appears weak; on the other hand it is clear that electrons do not arrive before they leave, even with faster than light speed they start in the past and arrive in the future, so causality is maintained (Nimtz 2011). This is exactly the sort of process we are looking for, where causality is not enforced by the fundamental speed limit of Special Relativity theory, but appears to be in force anyway, even when the speed limit is broken. Again it should be noted that we are not trying to look at things purely objectively at this point; here we are specifically trying to apply our theoretical bias to form an interpretation, which fits with the same bias, so that we can conceptualize an approach to realizing similar processes *in silico*. This is the focus of the next chapter.

6.4.4 *Validity of near-field electromagnetic holography*

A final point that we should not gloss over is of course that the method of near-field electromagnetic holography, at least in this setting and implementation, is not a tried and tested method (as the analogous method is in acoustics), and it is quite possible that we have made both conceptual and implementational mistakes along the way. We have

tried to avoid such pitfalls, but time and resources did not allow a completely rigorous verification; instead we have relied on the reasonable agreement of the outcomes of our method with prior knowledge of the neocortical system. The main points of agreement are: 1) The energy flow across layers agrees with the causality estimates. In general the broad anatomical structure of the slices can be made out in the energy flow, clearly separating cortical and subcortical structure, as well as deep and superficial layers. Primary and associational areas also appear as distinct units when considering the time course of energy flow. 2) The dissipation of energy during delta is focused in layer 5 and layer 2/3 (see Figure 6.4), which coincides with the spatial distribution of delta power and unit activity (Carracedo, Kjeldsen et al. 2013). In addition, the stronger dissipation events appear to spread along layers in a coherent and realistic fashion. 3) The general patterns of energy source density agree with current source density patterns (Carracedo, Kjeldsen et al. 2010), but are more refined and of greater interpretational scope.

6.4.5 *Summary*

We did not find any evidence of ergodic dynamics in either the wake-state model or the sleep-state model. However, a conclusion of non-ergodic dynamics could also not be fully drawn.

We confirmed the predictions made by a more detailed mapping of the Helmholtz machine concept onto neocortical areas. This was shown by energy flow vectors fields as a potential measure of causality, but closer to the underlying neurobiology than conventional causality estimates.

In addition, the ratios of dissipation to flow were found to be different between wake and sleep model states, suggesting different computational loads and perspectives. The difficult finding of negative imaginary dissipation in the wake-state was interpreted in our framework in a way which allows a path forward for the abstract computational modelling approach. This is the topic of the following chapter.

Chapter 7. Results

search for holographic tractability in silico

7.1 Introduction

This chapter presents the results of our computational modelling efforts. We focussed on geometric algebra implementations of holographic and quantum neural network concepts, as a mathematical foundation free of the difficult to interpret imaginary unit, which we identified in chapter 3 as a source of both conceptual and practical issues.

Although qubits and quantum-like computation have been mapped onto biological neurons and oscillations before (Orlov 2002; Burger 2011), our aim here is more abstract in the sense that we wish to identify (holographic) processes that can potentially go beyond conventional quantum computing. The specific mapping is thus a secondary concern, and is probably only likely to be correct once the fundamental holographic computational substrate has been identified, since the potential mappings of qubits are many, but the elusive mapping between a holographic computational construct and enabling physical/biological processes is likely singular.

We first aimed at reproducing holographic neural networks in geometric algebra on a basic test case as previously formulated in (Aerts, Czachor et al. 2009), with specific focus on the projected product as a potentially tractable version of the full geometric product, and as a conceptual link to quantum measurements that are also conceptually projections on limited subspaces. We then moved on to create an optimized version in order to test more advanced cases of structural and hierarchical data structures as first demonstrated in (Neumann 2001) without geometric algebra. In order to connect with quantum computation terminology we implemented quantum neural networks both with conventional complex numbers and in geometric algebra; for the same purpose we also considered qubit structures in geometric algebra.

Finally, inspired by the results of the previous chapter we undertook that construction and initial characterization (in geometric algebra) of a computational communication resource called an ebit.

7.2 Methods

The computational modelling used in this thesis was not about neurobiological detail, but rather about computational properties of principal network concepts at the most abstract algebraic level; at this level the model neurons are perhaps best considered as coupled oscillators, i.e. complex- or qubit-valued networks, but the nature of such coupling remains abstract in the sense that we wish to keep the door open to more exotic connections as suggested by chapter 3, and by the interpretation of imaginary dissipation in chapter 6.

7.2.1 *Holographic neural networks in geometric algebra*

We implemented basic geometric algebra operations in Matlab, and verified it against a reference implementation (Dorst, Fontijne et al. 2009). The implementation allowed an optimized path for large-scale processing, as well as a matrix embedded version (Cartan's representation of geometric algebra) for visual inspection of smaller (abstract) memory structures. We followed (Aerts, Czachor et al. 2009) to then implement holographic neural networks in geometric algebra.

7.2.2 *Quantum neural networks*

We implemented quantum neural networks with conventional complex numbers as described in (Peruš 2000; Perus, Bischof et al. 2007), then replaced the imaginary unit, with its geometric algebra equivalent, called the pseudoscalar. For a more original approach we then followed (Matzke 2002) to implement qubit representations in geometric algebra, as building blocks for another type of quantum neural network in geometric algebra, with easier consideration of quantum computation terminology.

7.3 Results

We present the results in the order they were obtained; throughout we focussed on conceptual computational properties rather than rigorous demonstrations and performance tests on specific computation problems.

7.3.1 *Holographic neural networks in geometric algebra in practice*

Traditionally, the basic test case of holographic neural networks is concept-to-value binding in memory of personal data of a fictive character, called 'Pat Smith'. Each value

is retrieved from memory by addressing it with the corresponding concept. Concepts and values are represented as so-called blades in a geometric algebra of limited dimension depending on the number of concepts required.

In Figure 7.1 we demonstrate this basic test case in our geometric algebra implementation. We first addressed the memory via the full geometric product with the reverse (see figure text) of the ‘name’ concept to retrieve the relevant answer, ‘pat’, as well as two noise terms in higher subspaces, specifically the subspace of 3-blades (indicated in the figure by the presence of three ‘eX’ factors, X being a number). Traditionally these noise terms are cleaned up by an additional processing step comparing the noisy answer with all possible answers. The projected product, on the other hand, uses the prior knowledge that answers only appear in the 1-blade subspace where they were encoded, to limit the retrieval process to this subspace. We implemented the projected product both by post hoc removal of noise subspaces from the full geometric product, and by pre-computing which parts of the geometric product end up in the right subspace, and then only computing those. This process was however almost as costly as the full product; in both cases we however avoided the need of the additional clean-up comparison with all possible answers.

The specific implementation of the projected product was however not general in the sense that in more complicated cases answers end up in more complicated subspaces, and the projected product must be designed especially to reflects this; to our knowledge it is unknown whether this design process is always feasible. To test this we implemented a more elaborate test case of learning logical transformations, shown in Figure 7.2. In this case we were also able to design a suitable projected product, but again it was specific to the special case being considered. Figure 7.3 visually compares the outcome of the full geometric product versus the projected product; the full product creates an answer with more than 2500 blades, most of which belong to the noise part. The projected product only produces about 200 blades, only about half of which belong to the noise part. Compared to the simple test case this case also reveals that under more complicated conditions even the correct subspace is also contaminated by some amount of noise, the strength of which is however negligible compared to the strength of the correct answer (the length of the vertical line indicated in the figure).

It is worth noting that addressing the memory with the geometric product is reversible, since by applying the same (or reverse) operator again, the three-part noisy answer recreates the full memory. The projected product, on the other hand, is irreversible, because no more information can be recreated from the noise-free answer. However, since this is not a quantum system we are allowed to simply retain the original copy of the memory, so there is no practical collapse of the memory superposition.

Figure 7.1: Basic test case of holographic reduced representations

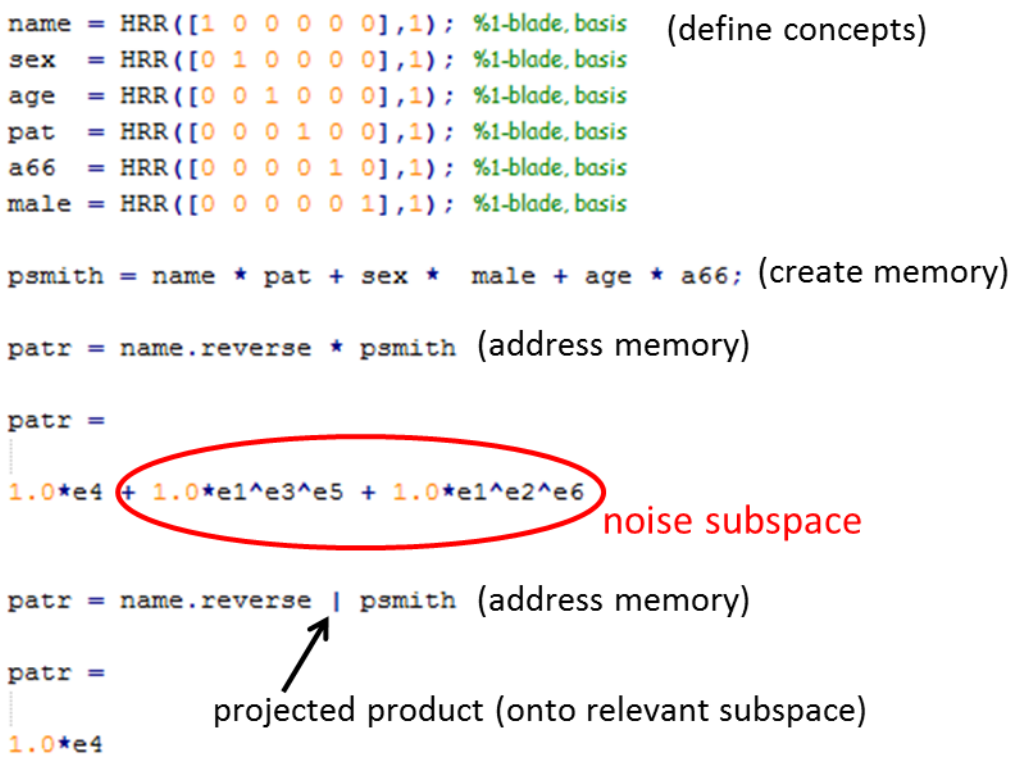


Figure 7.1: Basic test case of holographic reduced representations

Six personal data concepts and values of the fictive character, Pat Smith, are coded as basis blades of a geometric algebra. Concepts and corresponding values are bound by the geometric product, and added to a superposition in a single memory element (psmith above). The memory is addressed via the geometric product with the reverse of the concept whose value is to be retrieved. The reverse is literally formed by reversing the order of the string of zeros and ones in the blade.

The geometric product produces a three-part answer, the first part belonging to the 1-blade subspace, the other two to the 3-blade subspace. The correct answer is only the first part; the other two parts represent noise.

By applying the prior knowledge that only answers in the 1-blade subspace make sense (because all information was encoded in this subspace), we can throw away the last two parts, and get a clean answer. However, the resources spent on calculating these two noise parts are wasted; it would be preferable to simply not calculate them in the first place.

The projected product attempts this feat; in this case by manual design of the reduced calculation, which, although successful, does not generalize to more complex situations.

Figure 7.2: Structural test case of holographic reduced representations

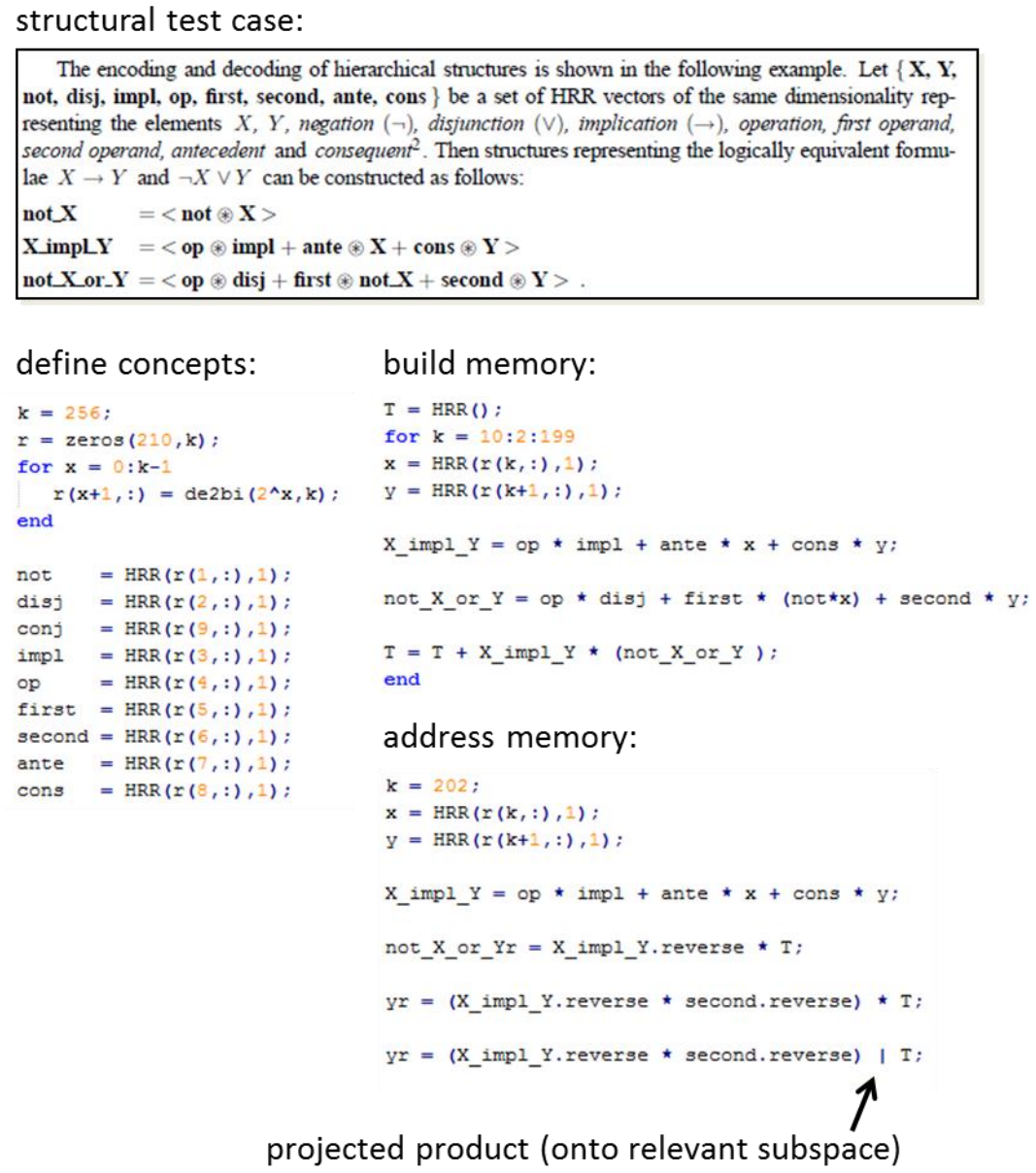


Figure 7.2: Structural test case of holographic reduced representations

Hierarchical structure, in this case formal logical formulae, can be represented and transformed with holographic reduced representations in geometric algebra. Here we learned the transformation between two logically equivalent expressions, and tested the memory on an unseen expression, comparing the true answer y, with the reconstructed answer yr, for both the geometric product and the projected product (redesigned for this specific case). See next figure for a visual representation and comparison of results. The structural test case was taken from (Neumann 2002).

Figure 7.3: Structural test case of holographic reduced representations, outcome

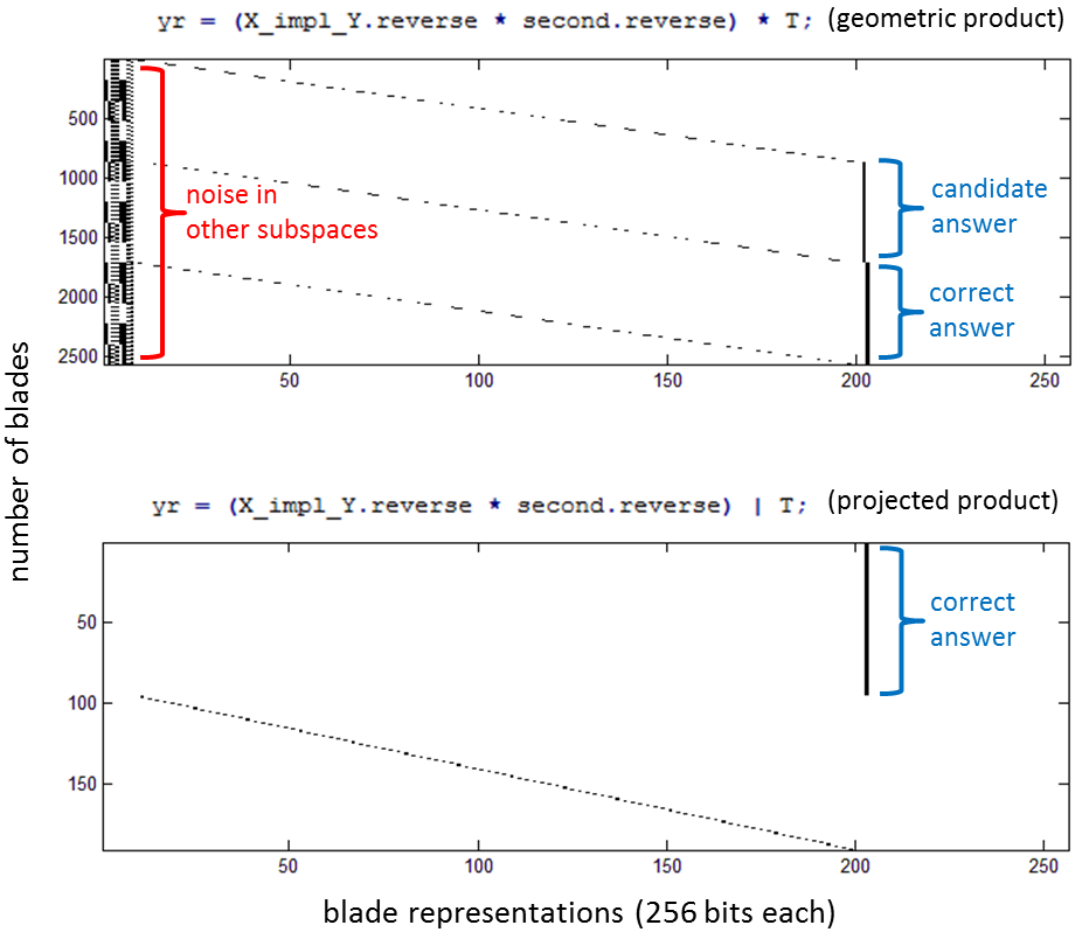


Figure 7.3: Structural test case of holographic reduced representations, outcome

Graphical representations of the outcomes of addressing the transformation memory defined in the previous figure, with the geometric product (top), and the projected product (bottom). The x-axis shows the number of bits set in each blade; one bit corresponds to 1-blade subspace, two bits to the 2-blade subspace etc. The y-axis shows the number of blades, each row is one blade.

In the top panel there are more than 2500 blades, the left columnar pattern of set bits reflects noise in higher subspaces; the strongest single component (blade) is the correct answer, but there is also a strong candidate answer.

The bottom panel answer from the projected product, on the other hand, only produces slightly more than 200 blades, and the correct answer is very clearly the strongest component, although the same subspace also contains some noise (the slanted, dotted line) in this more elaborate example.

The basic test case was implemented in Cartan's matrix embedded version of geometric algebra (described in more detail in chapter 3), which is catastrophically inefficient, because of close-to exponential scaling of matrix sizes. For this reason the more complicated test case was implemented in our optimized version. However, the matrix version allows one to inspect the matrix content, which led (Patyk 2010) to notice that the sparse matrices contain repeating patterns. This discovery resulted in a much reduced version of Cartan's representation, which was implemented simply by dividing by a factor of two in the right places in an otherwise complicated expression (involving repeated application of so-called Pauli matrices), used to generate Cartan's embedding; we refer to this innovation as Patyk's trick.

7.3.2 *Beyond Patyk's trick*

In Figure 7.4 we demonstrate the application of Patyk's trick on a holographic neural network representation embedded in Cartan's matrix form as a 256 trillion element matrix (implementable only due to extreme sparsity exploitable through sparse classes in Matlab), reduced to about 67 million elements by Patyk's trick. However, by exploiting sparseness we were able to consider embedded matrix representations of much larger size than originally considered by Patyk, which led us to the further discovery that Patyk's trick was not the end of the possible exploitations of self-similarity in the data structure; in fact, the fundamental pattern in this case repeated past Patyk's representation down to around 150000 element matrices, in some cases further (while retaining its computational properties).

Unfortunately, exploiting this discovery directly turned out to be difficult; in Patyk's case dividing by a factor of two was enough, but we were unable to find any usable effect of dividing by larger integers, and we were unable to conceptualize another approach within Cartan's representation. Nonetheless, it was encouraging that the holographic neural network possessed such natural self-similarity, perhaps suggesting that in an optimized implementation this redundant data would simply cancel out to reduce the computational load as well as the storage requirements.

Figure 7.4: Patyk's trick hints at further tricks

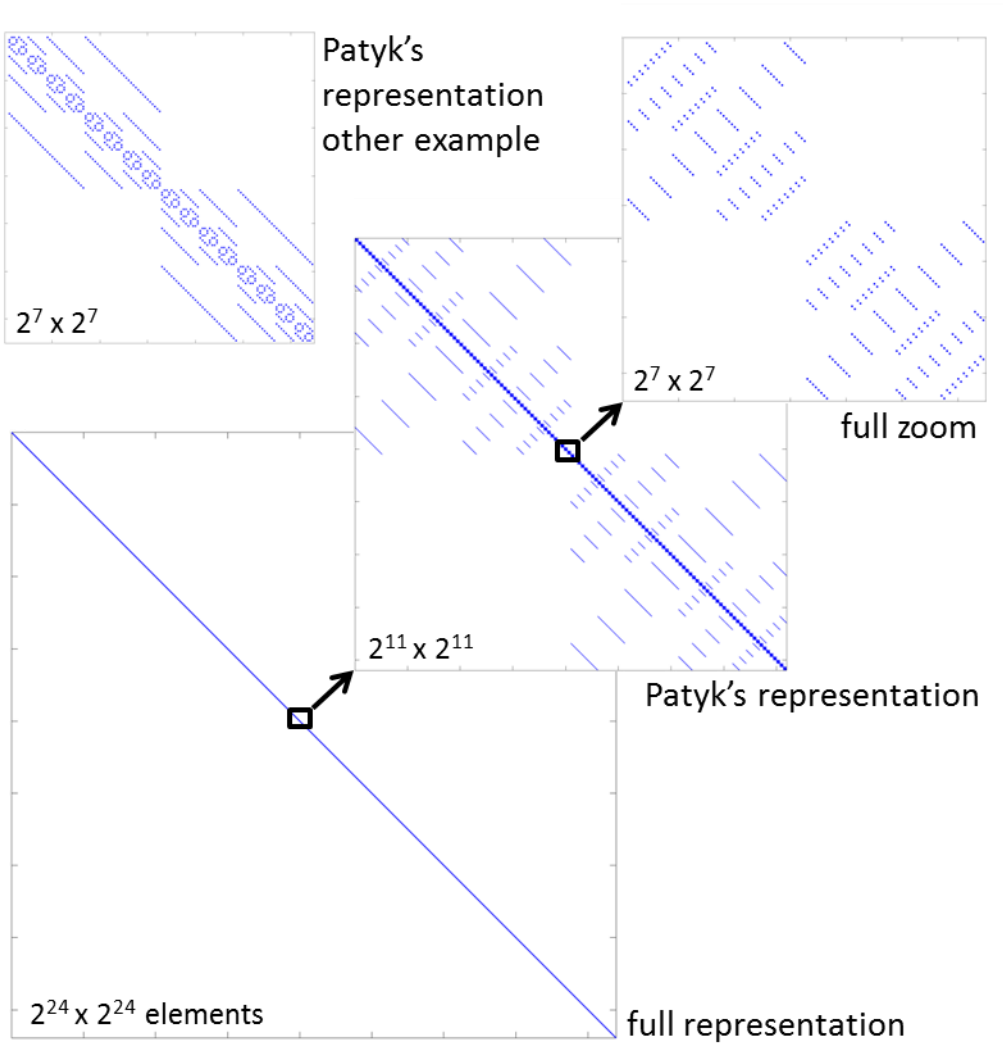


Figure 7.4: Patyk's trick hints at further tricks

The full representation (big square with diagonal) in Cartan's representation takes a 256 trillion element matrix, while Patyk's representation only contains about 67 million elements taken from the centre of the diagonal, indicated by the black square (not to scale). A further zoom-in reveals that the self-similarity however continues below Patyk's representation to a matrix of only about 150000 elements. An additional example of Patyk's representation shows another kind of self-similarity along the diagonal. The computational properties of these matrices are the same, i.e. concepts and values could be retrieved with the same fidelity. The full representation was only implementable by exploiting the extreme sparsity with sparse classes in Matlab.

7.3.3 *Quantum neural networks in geometric algebra in practice*

We started an investigation of quantum neural networks by implementing a standard quantum neural network as a complex valued matrix memory; to verify the implementation we selected an image pair dataset for learning natural human pose estimations from (Ferrari, Marin-Jimenez et al. 2009). Pixel values were coded in the phase angles of complex numbers with magnitudes set to unity. Image sizes were 627x353 by 3 colour values, producing a matrix size requirement of about 440 billion complex values. This made the problem immediately intractable in space (problems can be both intractable in time (computation) and space (storage)), however it was not quite intractable in time, so we exploited a common trick of trading time for space (Peruš 2000; Perus, Bischof et al. 2007): This was achieved by introducing the addressing of the memory (to retrieve an answer) before the full matrix was created, meaning that each matrix elements' contribution to the final answer was generated one-by-one, corresponding to the addressing object, and then discarded to make room for the next calculation of the next part. The trade-off being that the full set of calculations had to be re-done for each memory element. In this way the full memory matrix did not exist at any one time.

Interestingly, this process was somewhat analogous to the quantum collapse that occurs at the time of measurement in real quantum systems; in the same way the superposition never physically exists, and we are unable to copy and reuse the actual quantum memory for further measurements. If the memory could be copied or retained it is known to lead to tractable quantum computation (Aaronson 2005a). There is some indication that this might eventually be possible, since it is known that fast continuous measurements can prevent a quantum system from collapsing, a phenomenon known as the quantum Zeno effect (Misra and Sudarshan 1977).

Figure 7.5 shows the outcome of addressing the quantum neural network to retrieve a pose estimation. The network was able to recall seen poses and to some extent poses that were very close to seen poses; however, it could not generalize to completely unseen poses, which was also expected to be too difficult.

Figure 7.5: Quantum neural network example

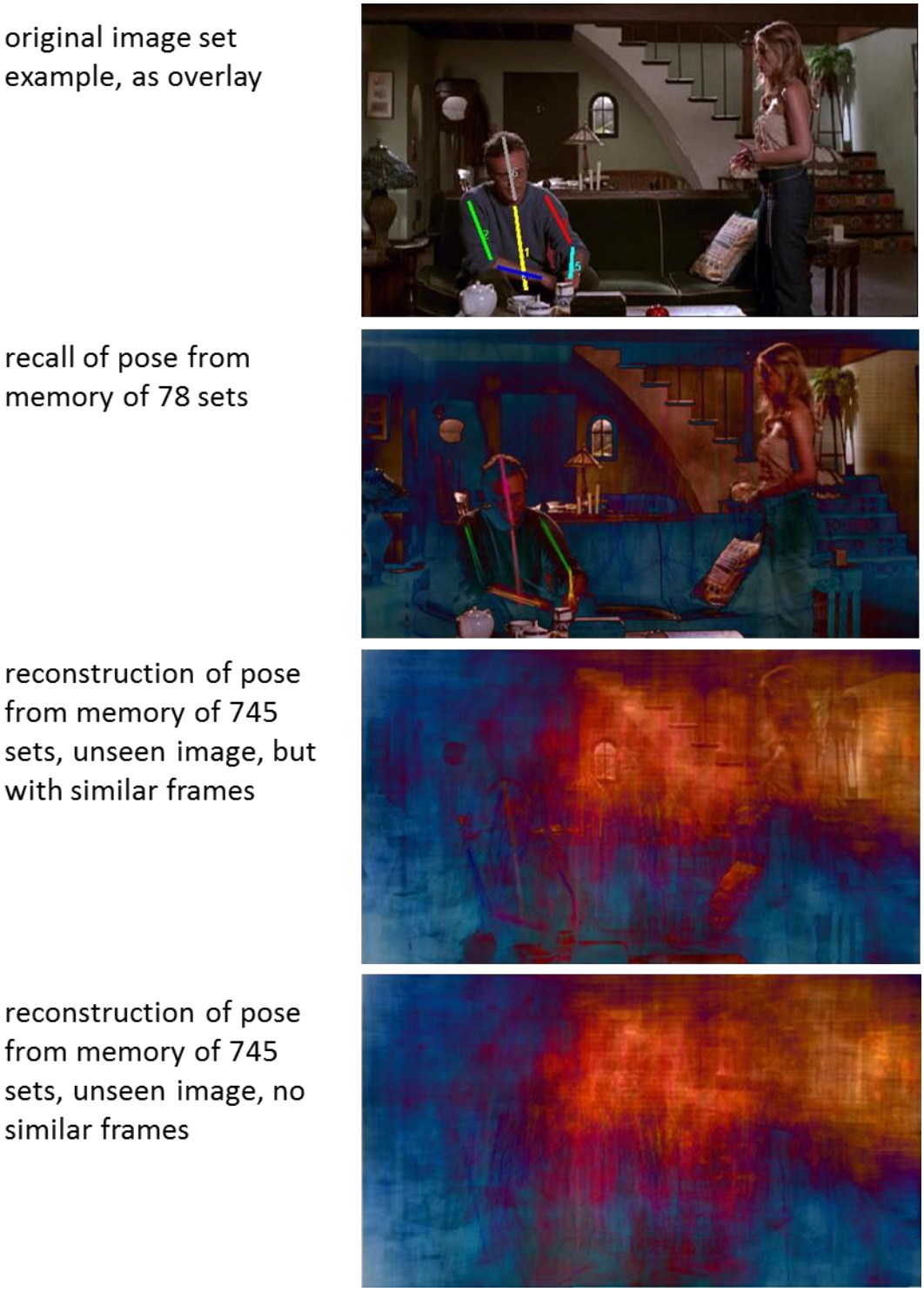


Figure 7.5: Quantum neural network example

Top, example of original image and pose-estimation set, shown with pose-estimation overlaid. Second, recall of seen pose from 78 set memory. Third, weak recall/generalization of unseen pose from 745 element memory with very similar poses. Fourth, failed generalization to unseen pose. Recall/generalization was judged by eye depending on whether pose estimations (coloured line overlays) could be visually discriminated.

Re-implementing the quantum neural network in geometric algebra revealed several interesting aspects. The main conceptual tasks were to replace the imaginary unit with its geometric algebra equivalent, and replace the matrix structure with a multi-vector structure. This led to interesting considerations: The imaginary unit corresponds to a so-called pseudoscalar, the highest grade object (for complex numbers, a 2-blade), of the space of the geometric algebra in question. If all complex numbers are given the same pseudoscalar they essentially end up as one summed-up complex value in the multi-vector memory (which is useless), revealing that it is the matrix structure that prevents this in the conventional version. Giving each complex number its own pseudoscalar means reserving a specific subspace for it, while the size of the full space of the geometric algebra must increase correspondingly to accommodate all necessary subspaces (depending on the number of complex values). No significant time or space appeared to be saved in this setup. However, small or redundant values in the matrix formulation can possibly be set to zero; if enough zeros exist to make the matrix sparse a significant optimization would exist through sparse library routines. On the other hand, in the multi-vector representation all zeros can be immediately discarded (sparseness not required); we did however fail to find a feasible implementation and application of these ideas.

Considering complex numbers in their own subspaces, however, immediately led to the concept of qubits, which we considered next in the hope to achieve further understanding through the terminology of quantum computing.

7.3.4 *Qubits in geometric algebra*

Qubits map into geometric algebra in much the same way as complex numbers, with additional interpretations of specific values as zero or one, and as zero and one at the same time. A key concept is the so-called Hadamard gate, which is a fundamental operator in quantum computing, since it rotates the qubit state between classical states (zero or ones) and quantum superposition states of zero and one (and the opposite, one and zeros) simultaneously. Figure 7.6 shows the geometry of the qubit, and its construction in geometric algebra, as well as the construction and effect of its associated Hadamard gate, which turns out to be simply the pseudoscalar of the space of the qubit, as with complex values above.

Figure 7.6: Qubit structure and Hadamard gate in geometric algebra

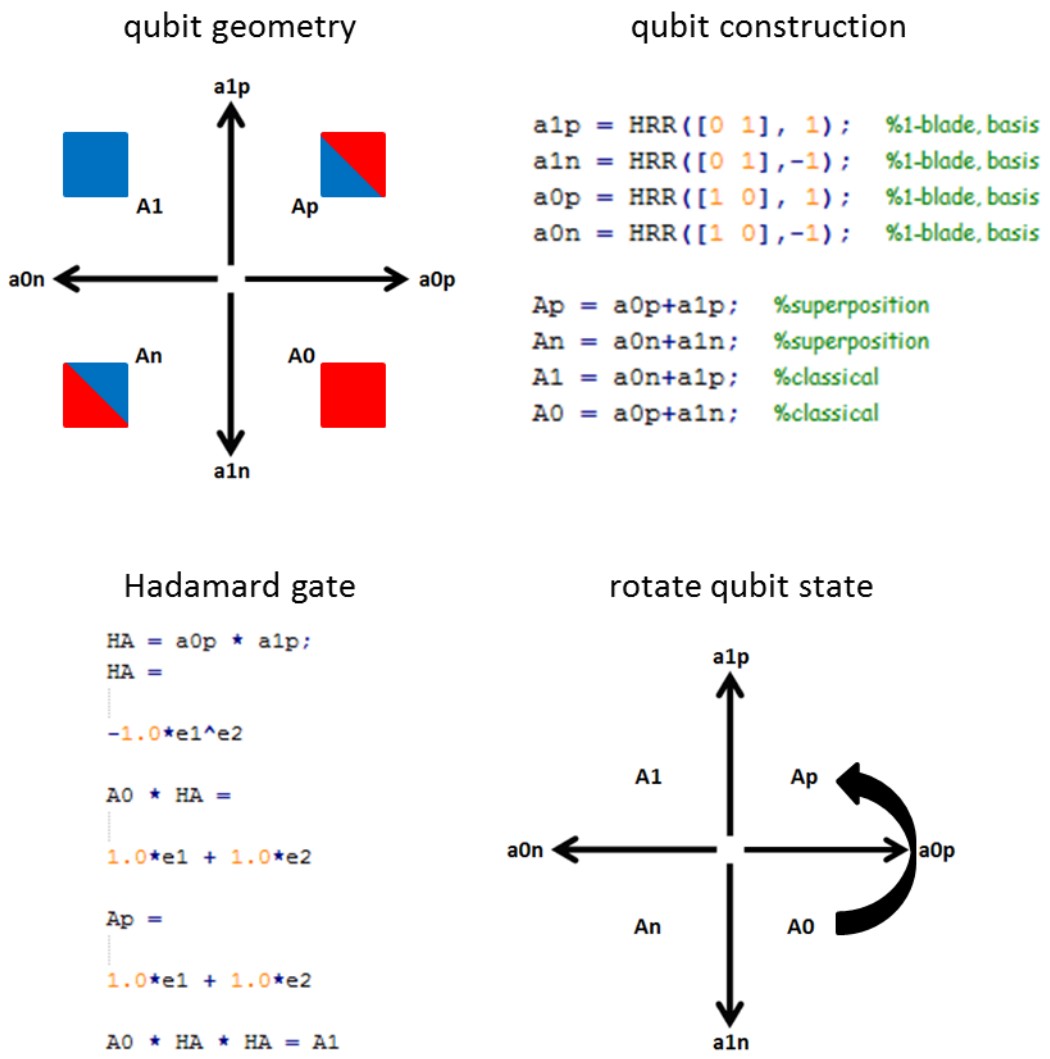


Figure 7.6: Qubit structure and Hadamard gate in geometric algebra

A qubit can be represented and manipulated in geometric algebra without the use of complex numbers. Here we define qubit A, and its associated Hadamard gate HA, which rotates the qubit state 45 degrees with every application, back and forth between classical and superposition states.

Every additional qubit requires its own subspace and its own Hadamard gate; a qubit register is formed via the geometric product, A * B * C etc., and can be put into a separable entangled superposition by sequential application of respective Hadamard gates, A * B * C * HA * HB * HC.

A qubit corresponds (can be directly encoded into) to a ½-spin particle, e.g. an electron, which can be visualized as a spinning top or better a gyroscope; the axis of rotation pointing up is the logical 1, pointing down is 0, the superposition state is represented by the vertical position of the gyroscope spin axis, in which case the gyroscopic forces (precession) create a second axis of rotation causing the gyro to spin around both axes at the same time, i.e. state superposition.

Figure 7.7 further provides a view of the four qubit states in Cartan's embedded matrix representation, as a reference for later, more complex representations. In this example we used a small space, which could only hold a few different qubits, to make the patterns easily visible.

The Hadamard concept from qubits and quantum computing inspired several new considerations: Firstly, we considered the possibility of applying a similar concept to the quantum neural network developed earlier, since this appeared to be a new possibility afforded by the re-formulation in geometric algebra; we could however not conceptualize an immediate benefit to the memory size problem faced by this approach, although the dual-space transformation by the pseudoscalar is similar to an eigenspace transformation, and several optimizations might lie down this path. Instead, we returned to the holographic neural network concepts to consider the possible benefits of a dual space representation in this case.

It turned out that the projected product, being a projection onto a subspace, can be formulated in the dual space in a simpler way. Figure 7.8 revisits the basic test case, this time transforming it to the dual space by application of the pseudoscalar, and addressing it via the outer product instead of the full geometric product, then transforming it back by another application of the pseudoscalar, to reveal a dual space version of the projected product. The cost of the dual space transformations is negligible compared to the cost of the geometric product, although the transformations are applied via the geometric product, because the pseudoscalar itself is a single element multi-vector. Unfortunately, the savings afforded by the outer product instead of the geometric product were insignificant, at least in the implementations that we could come up with, and in addition it did not generalize to the more complicated structure case, for which we instead found that various combinations of the inner products of geometric algebra produced a projected product in dual space. This was encouraging since these are also potentially cost saving, but we were ultimately unable to identify a general pattern to exploit this in general settings.

Figure 7.7: Qubit states in Cartan's matrix embedding

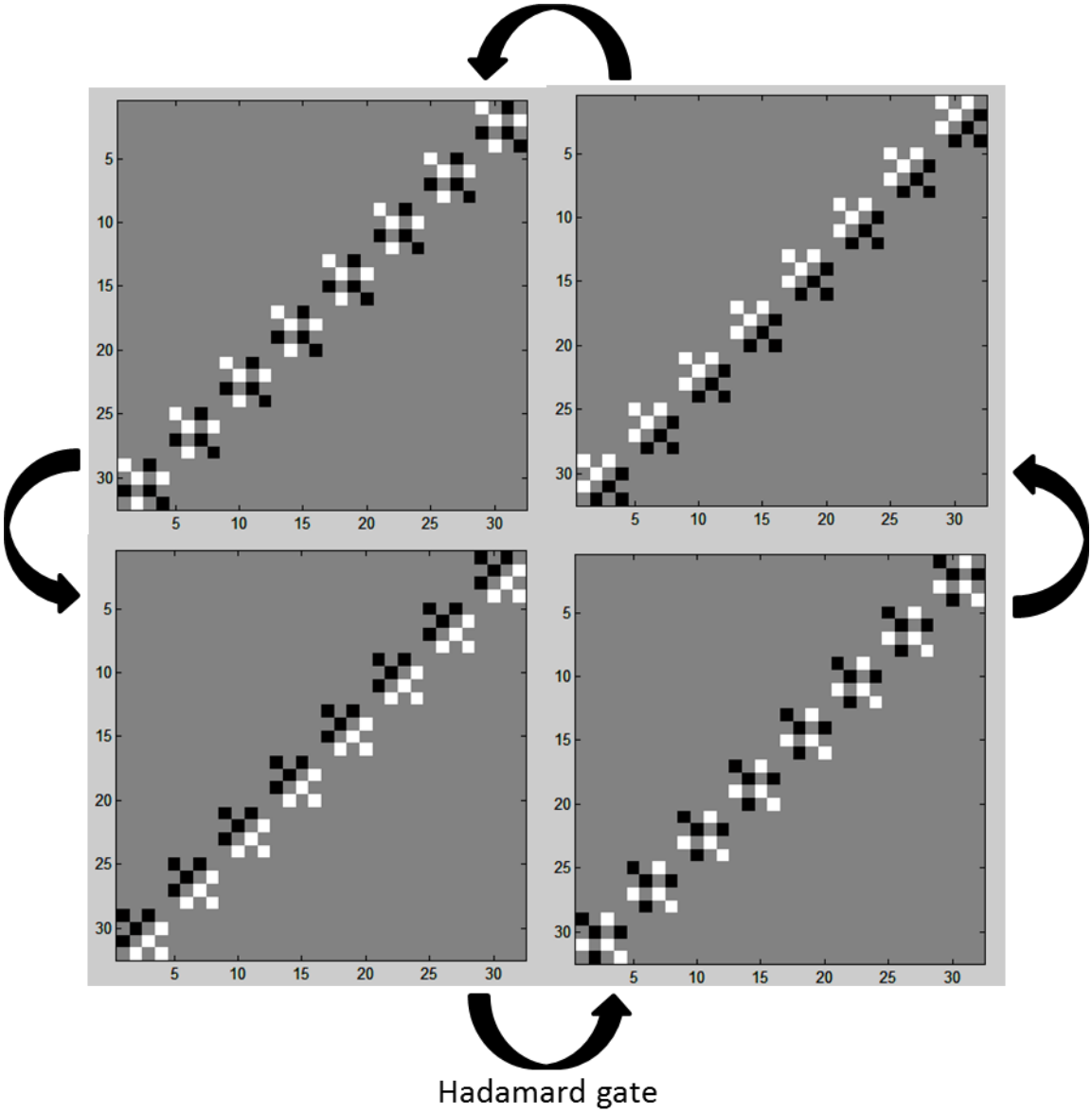


Figure 7.7: Qubit states in Cartan's matrix embedding

Here we provide a view of the four qubit states in Cartan's matrix embedding via Patyk's trick, for later reference. This is one example; each qubit will in general appear different, and the size of the matrix (the overall space of the algebra) will need to grow to accommodate further qubits, this example can hold a few different qubits.

Figure 7.8: Projected product as outer product in dual space

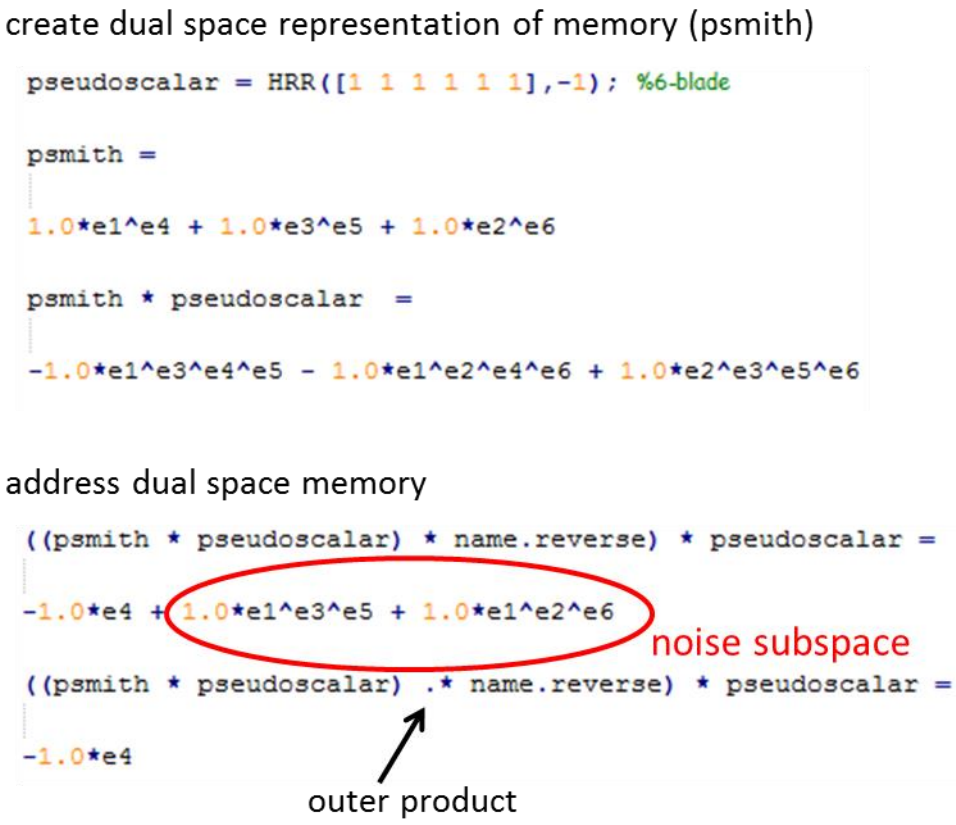


Figure 7.8: Projected product as outer product in dual space

Transforming the holographic neural network memory, psmith, into its dual space representation by application of the pseudoscalar allowed a different instantiation of the projected product directly through the outer product. This approach did however not generalize to the more complicated structure case, where the more restrictive (but potentially computationally cheaper) inner products were needed. We did however not find the general structure to the inner product versions of the projected product.

7.3.5 *Potential resources for tractable holographic computation, ebits*

Through the conceptualization of tractable holographic computation developed over the first three chapters we arrived at the idea of paradox-free interaction of past and future, as a way to circumvent Gödel's theorems, giving up the Axiom of Causality, but without losing causality itself. The question of which physical phenomenon could possibly carry such interaction in the brain (and elsewhere for that matter) was left open until the end of chapter 6, where we, on the basis of imaginary dissipation measurements, suggested that resonant energy transfer was a likely candidate.

The idea of quantum-like computation without actual quantum mechanics as suggested by formulations in geometric algebra (chapter 3 and this chapter), was supported by the notion that paradox-free interaction of past and future gives equal computational power to quantum and classical systems (both beyond previous limits); put another way, the interaction of past and future makes classical systems look quantum, and the concept of a distinct quantum level is then an illusion created by the failure to consider past-future interactions in both the classical and quantum pictures.

These ideas, although tentative, suggested that there is no fundamental difference between classical wave-mechanical resonance phenomena, and their mathematically equivalent quantum counterparts of quantum tunnelling and teleportation. For these reasons we decided to consider the geometric algebra equivalent of a quantum communication resource, known as ebits, used primarily in quantum teleportation protocols.

A single ebit is generally formed from two qubits that are maximally entangled into their shared Bell-states; this transformation in geometric algebra is achieved by the concurrent application of the qubits respective pseudoscalars (Hadamard gates) (as opposed to the sequential application used to sequentially rotate both qubits into their superposition states) (Matzke, Manthey et al. 2003). Figure 7.9 illustrates this process seen from the perspective of their Cartan matrix representations. The Bell-state matrix representations revealed that information is deleted in the process, because the number of matrix elements containing non-zero values is halved compared to the superposition states; this also revealed the irreversible nature of the concurrent Hadamard gates, because the deletion cannot be undone (Matzke 2002). Further applications of the concurrent Hadamard gates rotate the ebit between the four Bell-states, in analogy to the

single qubit case. This suggested that ebits can be used as computational elements in the same way as qubits, which is also the idea of decoherence-free subspaces considered in section 1.3.4, as a way to protect quantum states from unwanted interaction with the environment, especially the warm environment of the brain.

Following this concept we then encoded the basic test case from above into both qubits and ebits; in both cases we were able to retrieve correct answers with noise via the geometric product, and correct answers without noise via the projected product as the outer product in dual space, in the same way as above. We were however unable to identify any overt advantages of the ebit case in terms of computational costs or computational properties. We suspected that this was due to the simple nature of the basic test case, but the reality of the time-schedule for this thesis did not permit further investigations into more complex cases. To some extent this would also first have had to have been conceptualized as suitable test protocols did not appear to be readily available in this network perspective, although ebits have long been used in quantum teleportation protocols.

Figure 7.9: Ebits via irreversible concurrent Hadamard gates

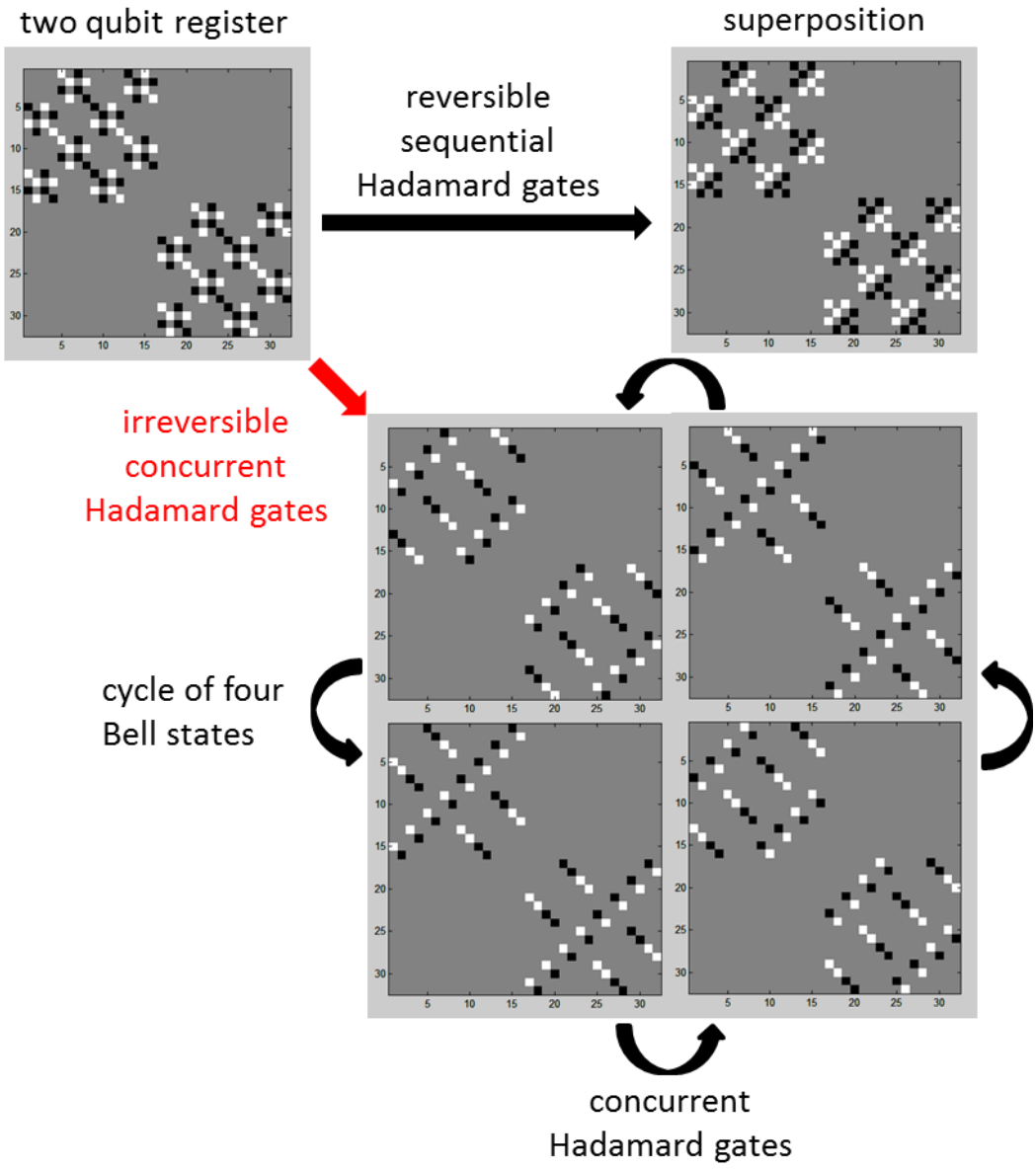


Figure 7.9: Ebits via irreversible concurrent Hadamard gates

A single ebit is formed from two qubits (a single qubit is formed from two bits) by the concurrent application of both qubits' Hadamard gates. As opposed to the sequential application, A * B * HA * HB, which separably, reversibly entangles qubits, the concurrent application, A * B * (HA + HB), inseparably, irreversibly entangles qubits into their maximally entangled Bell-state cycle, which is similar to the cycle of the qubits, and can be used in the same way to encode a single classical bit. As can be seen by considering Cartan's representations above, the reversible, sequential application creates an entangled superposition, but the irreversible, concurrent application deletes information (fewer squares are used in the Bell-state representations, half to be exact). This creates a quantum communication resource, the ebit, which in a sense holds an entropic credit since it has already pre-dissipated an amount of energy without computing specifically anything.

7.4 Discussion

We did not achieve the ultimate goal of the computational efforts presented in this chapter, which was to implement a tractable holographic computing concept able to solve intractable problems. We did however, within the time available, uncover several potentially important hints in the process:

7.4.1 *The projected product as a tractable geometric product*

The idea of the projected product as a potentially tractable version of the full geometric product resonates with the measurement problem of quantum mechanics (Schlosshauer 2005), where it is known that certain forms of measurement (projections) not allowed in the standard interpretation, but potentially allowed in holographic interpretations (Bohm and Hiley 1995), afford tractable quantum computation able to solve intractable problems (Aaronson 2005a). These kinds of measurements can also be seen as projections onto Bell-states, which when post-selected to only allow consistent results also provide tractable computation of otherwise intractable problems (Lloyd, Maccone et al. 2011b), and not only at the quantum level, but also classically (Aaronson 2005b). This selective consideration of only the states, which are physically relevant and paradox-free, was perhaps also reflected in the representational redundancy, which became apparent in Cartan's matrix representations of holographic memories. This can be seen in Figure 7.4, going beyond Patyk's trick; but we were unable to exploit it beyond said trick at this time. There is of course also a chance that the self-similarity of Cartan's representations is due to the construction of the representations themselves, but since this did not appear to be the case for Patyk's trick (Patyk 2010), it would also not be the case for further tricks.

7.4.2 *The reason quantum neural networks are intractable?*

We also found a natural interface between intractable quantum neural networks, and the measurement problem of real quantum systems: In the same way as the network memory never existed, and only provided an answer at the time of measurement, real quantum systems never physically display their coherent superposition states, and collapse to a single answer at the time of measurement. In both cases this prevents memory reuse, which at least in the quantum case is known to solve intractable problems, if it were possible (Aaronson 2005a).

7.4.3 *Ebits as entropic credit and communication resource*

In chapter 6 we suggested that the patterns of dissipation in the wake and sleep model states potentially reflected the notion of entropic credit outlined in the introduction, and that this credit was built during sleep and spend on communication through coherence during the wake-state. This fits with the concept of ebits in at least two ways; the creation of ebits (blindly) deletes redundant information in an irreversible way, in a sense pre-dissipating energy before any actual computation, the ebits then become a communication resource, which (like a qubit) is spend (destroyed) when it is used (measured). The specific type of communication is characterized by long-range correlations, and at least potentially occurs at super-luminal speeds in the paradox-free sense considered throughout.

Returning to the question of mapping these computational concepts onto biophysical processes it is worth noting that ebits based on Bell-states are not quite unique; in fact, if we apply a slightly different version of the concurrent Hadamard gate, by changing (HA + HB +...) to (HA – HB –…), we get another set of states equivalent to the Bell-states, called the Magic-states (Matzke 2002). This four-cycle of states are ninety degrees out of phase with the four-cycle of the Bell-states, but interchangeable at every second state. It is then hard not to notice the similarity with the electric and magnetic field concepts studied for near-field electromagnetic holography in section 6.2.3. While entry into Bell- or Magic-states is irreversible (i.e. exclusively time-forward), the rotation between the four states within is reversible (by the opposite rotation). The analogy with electric and magnetic fields forming the Poynting energy flow by their cross-product might then suggest that a similar combination of Bell- and Magic-states could produce a causal time-forwards energy flow (without assuming it in advance). An important hint to understanding such processes might be found in the fact that in Fourier dual space multiplying and dividing by the pseudoscalar, i.e. the imaginary unit, produces, respectively, derivatives and integrations, in a sense opposite directions of time. Fundamentally, there are only four time derivatives forming a cycle of position, velocity, acceleration, jerk (control), and back to position (Young 1976). We hope that future work will elucidate the true nature of these potential analogies.

Relatively little attention has been paid to ideas of tractable computing concepts in general, making it difficult to discuss; however, we particularly note (Pawłowski 2006) providing abstract, tractable geometric algebra algorithms for several intractable

problems, but also noting that these are considered unrealizable in physical systems, including quantum systems. We also note the concept of a duality quantum computer (Gui-Lu 2006; Gudder 2008), which goes beyond quantum computing parallelism by considering a dual space representation of the entire quantum computer. Within this framework, quantum interference provides computational properties similar to maximally entangled ebits, which were concluded to defeat intractable problems. This concept, however, also remains unproven.

Chapter 8. Conclusion

Having already discussed the main findings of this thesis in the respective results chapters above, we will finally give a more general evaluation of the two concurrent approaches and their putative convergence. We will thus highlight the main novel elements and their interactions along the way.

8.1 Evaluation of the electrophysiological approach

The electrophysiological approach adequately fulfilled the initial aims set out in section 2.3, and in addition yielded multiple somewhat unexpected opportunities for further investigation, as well as potentially important hints to help guide the computational approach.

We established a simplified *in vitro* model of the biological wake-sleep cycle, by firstly demonstrating that prior sensory-related activation in the wake-state positively affected delta orchestrated processes in the sleep-state (section 5.3.2 and 5.4.2); in line with established theories and prior findings in the broader field. We further considered the notion of reactivation of wake-state, sensory-related neural activity in the subsequent sleep-state (section 5.3.3 and 5.4.3), and in the process uncovered a potential mechanistic explanation of reactivation and increased delta power in the form of enhanced spike-train-to-local-field-potential phase coherence of reactivated neural units during prior activation in the wake-state (section 5.3.4 and 5.4.3). These data raise the interesting possibility that synaptic rescaling during deep sleep is prior use-dependent. That is to say, connections to neurons that were recruited into cortical representations of prior sensory stimuli within the 'ensemble' conceptual framework (Nicolelis, Baccala et al. 1995) were reactivated. In contrast, neurons that responded with relatively poor temporal precision to the 'sensory' input were deactivated. This can be seen as a form of 'pre-processing' of cortical activity in which only the strongest (in terms of temporal-precision) neuronal activity patterns are selected for further processing within the Helmholtz scheme (see below). The biological mechanism underlying this replay-selection remains unclear. Prior work showed that neuronal spiking in regular spiking cells was extremely variable, with very small changes in membrane potential (ca. 5 mV) taking these cells from a state of no spiking to continuous spiking at theta frequencies

(Carracedo, Kjeldsen et al. 2013). It is tempting to suggest that replay selection may therefore simply be a 'hard-thresholding' process, with only the strongest excitatory connections surviving.

As a bio-computational framework we built upon the concept of the Helmholtz machine, and formed a novel hypothesis about the spatiotemporal interactions along and across layers in the rodent neocortex during the sleep-state model (section 2.2.3). To test these predictions we introduced a novel analysis technique, near-field electromagnetic holography, which allowed us several new views on the data at hand, particularly energy flow vector fields and an energy dissipation measure, both instantaneous and spatial quantifications of putatively important aspects of the underlying neurobiological processes (section 6.2.3). The quantification of energy flow vector fields appeared to confirm our Helmholtz machine-based hypothesis, and strengthened and extended our previous work with new detail (section 6.3.2, 6.3.3 and 6.4.2).

Additional findings of energy dissipation and flow relations in the neocortex (section 6.3.4) allowed us to propose a novel interpretation of the general electromagnetic concepts involved (section 6.4.3), which in turn provided a potentially important hint for the final direction of the computational approach. These data also revealed previously unknown subtleties in the spatiotemporal dynamics of cortex during deep sleep, and as we argue in chapter 6, suggests different computational processes occurring at different times on each delta cycle. However, what those processes may be, and how they are biologically generated, will have to be the subject of further work.

8.2 Evaluation of the computational approach

The computational approach fulfilled the initial computational aims (section 3.5), and also managed to take up at least one interesting hint from the electrophysiological analysis.

We focussed on geometric algebra formulations (section 3.4.2 and 3.4.3), which turned out to be wise in the sense that the main computational constructions of holographic, qubit and ebit networks could be relatively easily formulated within the same framework, allowing novel points of comparison, with further understanding of potential equivalences (section 7.4). In addition, the test case of hierarchical structure

transformation (section 7.3.1) had not been considered in a geometric algebra formulation before, and its successful implementation in turn strengthened the geometric algebra approach.

We did reach some roadblocks along the way; the investigation of the projected product appeared very promising (section 7.3.1), but did not yield a final, general solution. The same is true of the hints that it could be possible to take Patyk's trick further (section 7.3.2); although very encouraging, we failed, in the time available, to find a general way to exploit it, and also did not manage to translate this hint into the equivalent form in the more optimized geometric algebra representations.

The formulation of qubits and ebits in geometric algebra was not in itself novel, but the abstract point-of-view in terms of the search for tractability provided a new focus in considering these computational construct. On the other hand, and more concretely, the visual representations afforded by the Cartan's matrix embeddings were novel and useful ways to visualize otherwise difficult concepts (section 7.3.4 and 7.3.5).

Although we purposely focused on the basic computational constructs, it would have been desirable to also have had time for more rigorous testing on various computational problems. Based on prior experience we nevertheless feel confident that we were able to recognize important computational constructs and properties without building our approach from the very bottom in the smallest possible steps.

8.3 Evaluation of convergence

Admittedly, the work load of the electrophysiological approach came to slightly dominate the efforts of this thesis, it was however never intended that the two approached should necessarily be equally weighted.

By the nature of the reciprocal interaction, both sides continuously influenced each other to the extent that we did not see them as two separate approached once we got deeper into it. Perhaps this is the best measure of convergence. In addition, and of a more concrete nature, we consider the interpretation of energy flow and dissipation relations on the electrophysiological side, leading to the introduction of the ebit concept on the computational side. This specific point of convergence was not forced by one

side, but rather followed naturally from both sides, with both leading up to the same point (in abstract computational terms). This meant that once the electrophysiological approach had provided the initial hint, the computational approach was immediately ready to take up a potentially relevant computational construct, which would otherwise have been very far off in terms of implementation.

In terms of the other direction of influence it is clear that it was the considerations of holographic processes in general, and the practical experience with the same, that eventually led to the conceptualization and implementation of the near-field electromagnetic holography analysis technique, which came to great use in the electrophysiological approach, in turn feeding back to the computational side.

Ultimately, we did not manage to reverse-engineer tractability, but we did gain enough headway to be satisfied for now. We did arguable not reach a dead-end, and the road ahead appears open; should we choose to go down it, the work involved picks up from the end of chapter 7 with the construction and characterization of further ebit based networks, particularly ones combining both Bell- and Magic-state ebits.

Glossary

Annealing

Annealing is a technique known primarily from metallurgy where it is used to improve the structural properties of steel. The process involves repeated heating and slow cooling of the object to be annealed. This controlled temperature treatment releases internal stress and rearranges the internal molecular structure into a (closer to) minimal energy state. As a result the steel becomes stronger and less brittle. The process is also commonly applied to glass objects.

Arrow of time

The arrow of time refers to the phenomenological experience that time has a single direction; the term is most often used in connection with the problem of its origin: While other dimensions are bidirectional, the dimension of time appears to be unidirectional; we do not know why. In entropy terms it is equivalent to the Second Law of Thermodynamics, whose origin is also unknown.

Bell-state

A Bell-state is two or more qubits that exhibit perfect correlation no matter their separation in space-time; this effect is the basis of quantum teleportation. A Bell-state of two qubits can encode a single bit of information (just like a single qubit can) and is referred to as an ebit.

Black hole

A black hole is a mathematical singularity predicted by general relativity theory where the singularity refers to gravity or equivalently to space-time curvature. At such extreme gravity not even light can reach escape velocity, and the singularity is therefore referred to as a black hole. A black hole is an object of maximal entropy; if all matter eventually falls into a single black hole it will be the point of maximal entropy of the Universe. It is believed that there is a black hole at the centre of our galaxy and most other galaxies.

Cosmic microwave background radiation

The cosmic microwave background radiation is almost uniformly distributed radiation coming from all directions across the sky, strongest in the microwave band. It is a few degrees above absolute zero, and is believed to be the first light in the Universe following the Big Bang.

Dissipation

Dissipation is a term for the entropy of energy increasing irreversibly. Dissipation can also refer to non-energy related increases in entropy, such as diffusion or mixing, but the use in this work refers to energy.

Ebit

An ebit is a two-qubit Bell-state used to encode one bit of information.

Ergodicity

Ergodicity is a term used to describe certain statistical properties of dynamical systems. In its simplest description it means that the long-term statistical properties of a single system are the same as the short-term statistical properties of many systems that are representative of the single system. This affords a computational and experimental short-cut, because the observation of a system for a long time can be replaced with observation of many systems for a short time; thereby saving considerable time. The general properties of ergodicity are however not conceptually clear; there are degrees such as strong and weak ergodicity, and sometimes ergodicity is taken to imply a further property known as mixing, and vice versa, i.e. sometimes the concepts are taken to be identical.

Event horizon

An event horizon is a boundary in general relativity around a singularity such as a black hole. The event horizon is the point of no return for matter and information falling into the black hole, even light cannot return once it has crossed the event horizon, so the blackness of a black hole starts at the horizon. No communication is believed to be possible from inside the event horizon to the outside universe, thus the inside of the event horizon is unknowable for the outside. It is related to, but different form, the idea of a causal horizon, which applies to points separated by a distance greater than the light year age of the Universe, meaning the points can never be in communication even if the

information travels at the speed of light. The term event horizon also applies to a hypothetical future technological singularity, and is sometimes taken to mean that the nature of such an event is unknowable before the event, or rather before the event horizon has been crossed. However, the event horizon of a technological singularity is that of a white hole singularity, i.e. the reverse of that of a black hole, so that information is not prevented from escaping, but rather from entering from outside the horizon.

General relativity

General relativity is the current theory of gravity stating that gravity is caused by the curvature of space-time.

Graphics processing unit

A graphics processing unit is a kind of computer processor characterized by having a small instruction set, but the ability to process many threads of instruction simultaneously, called parallel processing, by the use of multiple processor cores. For certain problems, like the fast Fourier transform, which are said to be parallel, this speeds up overall processing. In terms of functionality it is believed that parallel processing can be perfectly replaced by a single sufficiently faster processor.

Millennium Prize

The Millennium Prize is an award promised by the Clay Mathematics Institute in 2000. It awards a million dollars for the solution of each of seven problems of mathematics. As of 2013 one problem solution has been accepted. The proposed solution to the closure of turbulence problem given by Werbos, as mentioned earlier, might be a candidate for a second solution.

NP-problem

An NP-problem is a problem that belongs in the complexity class NP as in Non-Polynomial time. This means that to solve the problem a classical computer (in a classically causal universe) takes more time than a polynomial function of the problem size, typically an exponential.

Singlet state

A singlet state is a quantum state where the quantum numbers, e.g. the spin values, of the particles involved sum to zero. This is for instance the case when quantum fluctuations create and annihilate pairs of particles and antiparticles.

Singularity

The term singularity refers to a value going to infinity. This can for example be the value of gravity, space-time compression or distortion, or in a more abstract sense technological capability.

Space-time

Space-time is the fabric of the Universe itself. Generally refers to three dimensions of space and one dimension of time, although any combination of spatial and time dimensions is a space-time. According to general relativity the curvature of space-time induces gravity.

Superparamagnetic clustering

Superparamagnetic clustering is a data clustering algorithm named after the physical phenomenon of superparamagnetism, which the algorithm is modelled after. Clustering occurs when the abstract temperature of the material (the data) is lowered until the state passes from paramagnetic to superparamagnetic. It is a temperature driven energy minimization scheme, and as such is conceptually similar to annealing.

Technological singularity

A technological singularity is a hypothesised future event where technological progress goes to infinity. It follows from the observation that technological progress accelerates further technological progress leading to an exponential increase in technological capability. A technological singularity is logically a minimal entropy point like a white hole singularity, but where the white hole Big Bang singularity is in the past the technological singularity is in the future, it is therefore not generally believed that the two can be identified with each other, which would require the technological capabilities of the technological singularity to include some form of backwards time travel (without paradox). It is unknown whether we have crossed the event horizon of a technological singularity, but if the past and future white hole singularities are identified with each other the point might be mute altogether.

Theorem

A theorem is a generally accepted proof; not a theory or conjecture.

Theory of everything

A theory of everything is the ultimate goal of physics. It is generally considered to mean a unification of general relativity with quantum theory.

Quantum theory

Quantum theory is a mathematical framework dealing with the fundamental level of Nature starting from the Planck scale defining the smallest piece, a quantum, of energy, time, space etc. The framework is separate from the interpretation of what it means.

Qubit

A qubit is the fundamental unit of quantum computation in the same way a bit is the fundamental unit of classical computation. It can be in the same two states of zero or one like a conventional bit, but also in both states at once. An easy way to physically model this is to consider a spinning gyroscope with a single axis of rotation; turning it upside down and vice versa gives right or left rotation corresponding to zero or one, but if it is turned on its side gyroscopic precession forces a second axis of rotation, so that both right and left rotation happens at the same time, one on each axis. In the same way a qubit can be encoded in the spin of an electron.

White hole

A white hole is a minimal entropy singularity in general relativity. It is the time reverse of a black hole, or alternatively an anti-matter black hole (anti-matter being the time reverse of matter). The Big Bang is a white hole. It is unknown why the Big Bang was a low entropy singularity as it does not follow from time-reversal of a black hole alone. A hypothetical technological singularity would presumable be a white hole; its low entropy then follows from its technological capability.

References

3Brain. (2013). "Demo files." 2013.

Aaronson, S. (2004a). "Is quantum mechanics an island in theoryspace?" arXiv preprint quant-ph/0401062.

Aaronson, S. (2004b). Limits on Efficient Computation in the Physical World, UNIVERSITY of CALIFORNIA, BERKELEY. PhD Thesis.

Aaronson, S. (2005a). "Guest column: NP-complete problems and physical reality." ACM Sigact News 36(1): 30-52.

Aaronson, S. (2005b). "Quantum computing, postselection, and probabilistic polynomial-time." Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 461(2063): 3473-3482.

Aerts, D. and M. Czachor (2007). "Cartoon computation: quantum-like computing without quantum mechanics." Journal of Physics A: Mathematical and Theoretical 40(13): F259.

Aerts, D., M. Czachor, et al. (2006). "On Geometric Algebra representation of Binary Spatter Codes." arXiv:cs/0610075v2 [cs.AI].

Aerts, D., M. Czachor, et al. (2009). "Geometric analogue of holographic reduced representation." Journal of Mathematical Psychology 53(5): 389-398.

Aeschbach, D. (2009). "Slow waves and learning: beyond correlations." Sleep. 32(10): 1253-1254.

Ainsworth, M., S. Lee, et al. (2011). "Dual Gamma Rhythm Generators Control Interlaminar Synchrony in Auditory Cortex." The Journal of Neuroscience 31(47): 17040-17051.

Ainsworth, M., S. Lee, et al. (2012). "Rates and rhythms: a synergistic view of frequency and temporal coding in neuronal networks." Neuron. 75(4): 572-583. doi: 510.1016/j.neuron.2012.1008.1004.

Armstrong, C. and I. Soltesz (2012). "Basket cell dichotomy in microcircuit function." The Journal of Physiology 590(4): 683-694.

Arndt, M., O. Nairz, et al. (1999). "Wave–particle duality of C60 molecules." Nature 401(6754): 680-682.

Aton, S. J., J. Seibt, et al. (2009). "Mechanisms of sleep-dependent consolidation of cortical plasticity." Neuron. 61(3): 454-466. doi: 410.1016/j.neuron.2009.1001.1007.

Baccalá, L. A. and K. Sameshima (2001). "Partial directed coherence: a new concept in neural structure determination." Biological Cybernetics 84(6): 463-474.

Balci, M. E., M. H. Hocaoglu, et al. (2010). "Comments on “the role of Poynting's vector in polyphase power calculations”." European Transactions on Electrical Power 20(8): 1255-1255.

Ball, P. (2011). "Physics of life: The dawn of quantum biology." Nature 474.

Bancal, J.-D., S. Pironio, et al. (2012). "Quantum non-locality based on finite-speed causal influences leads to superluminal signalling." Nature Physics.

Baranger, M. (2001). "Chaos, complexity, and entropy." Wesleyan University Physics Dept. Colloquium.

Barrett, A. B., L. Barnett, et al. (2010). "Multivariate Granger causality and generalized variance." Physical Review E 81(4): 041907.

Battaglia, D. A., G. E. Santoro, et al. (2005). "Optimization by quantum annealing: Lessons from hard satisfiability problems." Physical Review E 71(6): 066707.

Bedard, C., H. Kroger, et al. (2006). "Does the 1/f Frequency Scaling of Brain Signals Reflect Self-Organized Critical States?" Physical Review Letters 97(11): 118102.

Beggs, J. M. and D. Plenz (2003). "Neuronal avalanches in neocortical circuits." J Neurosci. 23(35): 11167-11177.

Benchenane, K., A. Peyrache, et al. (2010). "Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal- Prefrontal Network upon Learning."

Bennett, C. H., G. Brassard, et al. (1993). "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels." Physical Review Letters 70(13): 1895.

Bennett, C. H., D. Leung, et al. (2009). "Can Closed Timelike Curves or Nonlinear Quantum Mechanics Improve Quantum State Discrimination or Help Solve Hard Problems?" Physical Review Letters 103(17): 170502.

Berut, A., A. Arakelyan, et al. (2012). "Experimental verification of Landauer/'s principle linking information and thermodynamics." Nature 483(7388): 187-189.

Bohm, D. and B. J. Hiley (1995). The undivided universe: an ontological interpretation of quantum theory, Routledge.

Born, J. (2010). "Slow-wave sleep and the consolidation of long-term memory." World J Biol Psychiatry. 11(Suppl 1): 16-21.

Bruno, N., A. Martin, et al. (2013). "Displacement of entanglement back and forth between the micro and macro domains." Nat Phys 9(9): 545-548.

Buhry, L. A., Amir H; Cheng, Sen (2011). "Reactivation, Replay, and Preplay: How It Might All Fit Together." Neural Plasticity 2011.

Burger, J. R. (2011). "Qubit-wannabe Neural Networks." arXiv preprint arXiv:1108.2819.

Campbell, I. G. and I. Feinberg (2009). "Longitudinal trajectories of non-rapid eye movement delta and theta EEG as indicators of adolescent brain maturation." Proc Natl Acad Sci U S A. 106(13): 5177-5180. Epub 2009 Mar 5123.

Carracedo, L. M., H. Kjeldsen, et al. (2010). "Neocortical delta rhythms are wave-packet generators." SfN 2010.

Carracedo, L. M., H. Kjeldsen, et al. (2013). "A Neocortical Delta Rhythm Facilitates Reciprocal Interlaminar Interactions via Nested Theta Rhythms." The Journal of Neuroscience 33(26): 10750-10761.

Chapline, G. (1999). "Quantum mechanics as information fusion." arXiv preprint quant-ph/9912019.

Chapline, G. (2004). "QUANTUM MECHANICS AND PATTERN RECOGNITION." International Journal of Quantum Information 02(03): 295-303.

Chapline, G. (2008). "Quantum Mechanics, Pattern Recognition, and the Mammalian Brain." AIP Conference Proceedings 1051(1): 15-21.

Chi, Z., P. L. Rauske, et al. (2003). "Detection of spike patterns using pattern filtering, with applications to sleep replay in birdsong." Neurocomputing 52-54.

Chu-Shore, J., M. B. Westover, et al. (2010). "Power Law versus Exponential State Transition Dynamics: Application to Sleep-Wake Architecture." PLoS ONE 5(12): e14204.

Cingolani, L. A., A. Thalhammer, et al. (2008). "Activity-dependent regulation of synaptic AMPA receptor composition and abundance by beta3 integrins." Neuron. 58(5): 749-762. doi: 710.1016/j.neuron.2008.1004.1011.

Clauset, A., C. Shalizi, et al. (2009). "Power-Law Distributions in Empirical Data." SIAM Review 51(4): 661-703.

Collini, E., C. Y. Wong, et al. (2010). "Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature." Nature 463(7281): 644-647.

Cook, S. A. (1971). The complexity of theorem-proving procedures. Proceedings of the third annual ACM symposium on Theory of computing. Shaker Heights, Ohio, USA, ACM.

Cooper, G. R. J. and D. R. Cowan (2008). "Comparing time series using wavelet-based semblance analysis." Computers & Geosciences 34(2): 95-102.

Coveney, P. V. and R. Highfield (1992). The arrow of time: a voyage through science to solve time's greatest mystery, Fawcett Columbine.

Cramer, J. G. (1986). "The transactional interpretation of quantum mechanics." Reviews of Modern Physics 58(3): 647.

Cunningham, M. O., D. M. Halliday, et al. (2004). "Coexistence of gamma and high-frequency oscillations in rat medial entorhinal cortex in vitro." J Physiol. 559(Pt 2): 347-353. Epub 2004 Jul 2014.

Czarnecki, L. S. (2006). "Could Power Properties of Three-Phase Systems be Described in Terms of the Poynting Vector?" IEEE Transactions on Power Delivery 21(1).

Dayan, P., G. Hinton, et al. (1995). "The Helmholtz Machine." Neural Computation 7: 889-904.

Dayan, P. and G. E. Hinton (1996). "Varieties of Helmholtz Machine." Neural Networks 9(8).

De Gennaro, L., F. Fratello, et al. (2008). "Cortical plasticity induced by transcranial magnetic stimulation during wakefulness affects electroencephalogram activity during sleep." PLoS One. 3(6): e2483. doi: 2410.1371/journal.pone.0002483.

Destexhe, A., S. W. Hughes, et al. (2007). "Are corticothalamic 'up' states fragments of wakefulness?" Trends Neurosci. 30(7): 334-342. Epub 2007 May 2003.

Deutsch, D. (1991). "Quantum mechanics near closed timelike lines." Physical Review D 44(10): 3197.

Dhamala, M., G. Rangarajan, et al. (2008). "Analyzing information flow in brain networks with nonparametric Granger causality." NeuroImage 41(2): 354-362.

Diekelmann, S. and J. Born (2010). "The memory function of sleep." Nat Rev Neurosci. 11(2): 114-126. doi: 110.1038/nrn2762. Epub 2010 Jan 1034.

Dijk, D. J. (2009). "Regulation and functional correlates of slow wave sleep." J Clin Sleep Med. 5(2 Suppl): S6-15.

Ding, M., Y. Chen, et al. (2006). "Granger Causality: Basic Theory and Application to Neuroscience." from arxiv.org/abs/q-bio/0608035v1.

Domenico, S., P. Nicola, et al. (1996). "Reactive acoustic intensity for general fields and energy polarization." The Journal of the Acoustical Society of America 99(4): 1868-1876.

Dorst, L., D. Fontijne, et al. (2009). Geometric Algebra for Computer Science (Revised Edition): An Object-Oriented Approach to Geometry, Morgan Kaufmann.

Dotto, L. (1996). "Sleep stages, memory and learning." CMAJ: Canadian Medical Association Journal 154(8): 1193.

Dreyfus, H. L. (2007). "Why Heideggerian AI failed and how fixing it would require making it more Heideggerian." Artif. Intell. 171(18): 1137-1160.

Duarte, M. F., M. A. Davenport, et al. (2008). "Single-pixel imaging via compressive sampling." Signal Processing Magazine, IEEE 25(2): 83-91.

Einhauser, W. and P. Konig (2010). "Getting real-sensory processing of natural stimuli." Curr Opin Neurobiol. 20(3): 389-395. doi: 310.1016/j.conb.2010.1003.1010. Epub 2010 Apr 1029.

Einstein, A., B. Podolsky, et al. (1935). "Can quantum-mechanical description of physical reality be considered complete?" Physical review 47(10): 777.

Eliasmith, C. (1997). "Structure Without Symbols: Providing A Distributed Account of Low-level and High-level Cognition." Southern Society for Philosophy and Psychology.

Emanuel, A. E. (2007). "About the Rejection of Poynting Vector in Power Systems Analysis." Electrical Power Quality and Utilisation XIII(1).

Engel, A. K., P. Fries, et al. (2001). "Dynamic predictions: oscillations and synchrony in top-down processing." Nat Rev Neurosci. 2(10): 704-716.

Engel, A. K. and W. Singer (2001). "Temporal binding and the neural correlates of sensory awareness." Trends Cogn Sci. 5(1): 16-25.

Esser, S. K., S. Hill, et al. (2009). "Breakdown of effective connectivity during slow wave sleep: investigating the mechanism underlying a cortical gate using large-scale modeling." J Neurophysiol. 102(4): 2096-2111. doi: 2010.1152/jn.00059.02009. Epub 02009 Aug 00055.

Esser, S. K., S. L. Hill, et al. (2007). "Sleep homeostasis and cortical synchronization: I. Modeling the effects of synaptic strength on sleep slow waves." Sleep. 30(12): 1617-1630.

Euston, D. R., M. Tatsuno, et al. (2007). "Fast-Forward Playback of Recent Memory Sequences in Prefrontal Cortex During Sleep." Science 318(5853): 1147-1150.

Feldman, B. J., I. J. Bigio, et al. (1982). "Through the Looking Glass with Phase Conjugation " Los Alamos Science 3.

Feldman, D. E. (2012). "The spike-timing dependence of plasticity." Neuron. 75(4): 556-571. doi: 510.1016/j.neuron.2012.1008.1001.

Feldman, H. and K. Friston (2010). "Attention, uncertainty and free-energy." Frontiers in Human Neuroscience 4.

Feldt, S., J. Waddell, et al. (2009). "Functional clustering algorithm for the analysis of dynamic network data." Phys Rev E Stat Nonlin Soft Matter Phys. 79(5 Pt 2): 056104. Epub 052009 May 056107.

Ferrari, V., M. J. Marin-Jimenez, et al. (2009). "Pose search: retrieving people using their pose " Proceedings of IEEE Conference in Computer Vision and Pattern Recognition, .

Feyerabend, P. (2010). Against Method, Verso Books.

Feynman, R. P. (1965). "The Development of the Space-Time View of Quantum Electrodynamics." Nobel Lectures, Physics 1963-1970, Elsevier Publishing Company, Amsterdam, 1972

Feynman, R. P. (1990). QED: The Strange Theory of Light and Matter, Penguin Press Science.

Fingelkurts, A. A., A. A. Fingelkurts, et al. (2005). "Functional connectivity in the brain--is it an elusive concept?" Neuroscience & Biobehavioral Reviews 28(8): 827-836.

Fodor, J. and B. P. McLaughlin (1990). "Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work." Cognition 35(2): 183-204.

Fodor, J. A. and Z. W. Pylyshyn (1988). "Connectionism and cognitive architecture: A critical analysis." Cognition 28(1-2): 3-71.

Foster, D. J. and M. A. Wilson (2006). "Reverse replay of behavioural sequences in hippocampal place cells during the awake state."

Fouque, J.-P., J. Garnier, et al. (2006). "Time reversal super resolution in randomly layered media." Wave Motion 43(8): 646-666.

Franken, P., D. J. Dijk, et al. (1991). "Sleep deprivation in rats: effects on EEG power spectra, vigilance states, and cortical temperature." Am J Physiol. 261(1 Pt 2): R198-208.

Friedman, J., M. S. Morris, et al. (1990). "Cauchy problem in spacetimes with closed timelike curves." Physical Review D 42(6): 1915-1930.

Fries, P. (2005). "A mechanism for cognitive dynamics: neuronal communication through neuronal coherence." Trends Cogn Sci. 9(10): 474-480.

Friston, K. (2010). "The free-energy principle: a unified brain theory?" Nat Rev Neurosci 11(2): 127-138.

Friston, K. and S. Kiebel (2009). "Predictive coding under the free-energy principle." Philosophical Transactions of the Royal Society B: Biological Sciences 364(1521): 1211-1221.

Friston, K., J. Kilner, et al. (2006). "A free energy principle for the brain." J Physiol Paris. 100(1-3): 70-87. Epub 2006 Nov 2013.

Friston, K. J., J. Daunizeau, et al. (2010). "Action and behavior: a free-energy formulation." Biol Cybern. 102(3): 227-260. Epub 2010 Feb 2011.

Fuchs, E. C., A. R. Zivkovic, et al. (2007). "Recruitment of parvalbumin-positive interneurons determines hippocampal function and associated behavior." Neuron. 53(4): 591-604.

Furber, S. (2012). "Low-Power Chips to Model a Billion Neurons." IEEE Spectrum.

Gabor, D. (1969). "Associative Holographic Memories " Optical Science and Technology 13(2).

Gleick, J. (1997). Chaos - The Amazing Science of the Unpredictable, Minerva.

Gleick, J. (2011). Genius: The Life and Science of Richard Feynman, Open Road Media.

Gödel, K. (1931). "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I." Monatshefte für Mathematik und Physik 38(1): 173-198.

Gödel, K. (1949). "An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation." Rev. Mod. Phys. 21.

Gödel, K. (1988). Eine Bemerkung über die Beziehungen zwischen der Relativitätstheorie und der idealistischen Philosophie. Zeit im Wandel der Zeit. P. Aichelburg, Vieweg+Teubner Verlag. 23: 161-167.

Goto, T., R. Hatanaka, et al. (2010). "An evaluation of the conductivity profile in the somatosensory barrel cortex of Wistar rats." Journal of Neurophysiology 104(6): 3388-3412.

Gott, J. R. (2002). Time Travel in Einstein's Universe: The Physical Possibilities of Travel Through Time, Houghton Mifflin.

Grebogi, C., E. Ott, et al. (1985). "Attractors on an N-torus: Quasiperiodicity versus chaos." Physica D: Nonlinear Phenomena 15(3): 354-373.

Grigolini, P., G. Aquino, et al. (2009). "A theory of 1/f noise in human cognition." Physica A: Statistical Mechanics and its Applications 388(19): 4192-4204.

Grinsted, A., J. C. Moore, et al. (2004). "Application of the cross wavelet transform and wavelet coherence to geophysical time series." Nonlin. Processes Geophys. 11(5/6): 561-566.

Gudder, S. (2008). "Duality quantum computers and quantum operations." International Journal of Theoretical Physics 47(1): 268-279.

Gui-Lu, L. (2006). "General quantum interference principle and duality computer." Communications in Theoretical Physics 45(5): 825.

Gull, S., A. Lasenby, et al. (1993a). "Electron paths, tunnelling, and diffraction in the spacetime algebra." Foundations of Physics 23(10): 1329-1356.

Gull, S., A. Lasenby, et al. (1993b). "Imaginary numbers are not real—The geometric algebra of spacetime." Foundations of Physics 23(9): 1175-1201.

Gurzadyan, V. G. and R. Penrose (2010). "Concentric circles in WMAP data may provide evidence of violent pre-Big-Bang activity." arXiv preprint arXiv:1011.3706.

Hald, J. (2001). "Time Domain Acoustical Holography and Its Applications." SOUND AND VIBRATION.

Hameroff, S. and R. Penrose (1996). "Orchestrated reduction of quantum coherence in brain microtubules: a model for consciousness." Math. Comput. Simul. 40(3-4): 453-480.

Hastings, H. M. and S. Waner (1985). "Low dissipation computing in biological systems." Biosystems 17(3): 241-244.

Hawking, S. W. (1976). "Breakdown of predictability in gravitational collapse." Physical Review D 14(10): 2460-2473.

Hawking, S. W. (1992). "Chronology protection conjecture." Physical Review D 46(2): 603-611.

Hawking, S. W. (2005). "Information loss in black holes." Physical Review D 72(8): 084013.

Hawking, S. W. and W. Israel (1979). General Relativity: An Einstein Centenary Survey, Cambridge University Press.

He, B. J. (2011). "Scale-Free Properties of the Functional Magnetic Resonance Imaging Signal during Rest and Task." The Journal of Neuroscience 31(39): 13786-13795.

Heimburg, T. and A. D. Jackson (2007). "On the action potential as a propagating density pulse and the role of anesthetics." Biophysical Reviews and Letters 2(01): 57-78.

Helmholtz, H. (1860). "Handbuch der physiologischen optik " Southall, J. P. C. (Ed.), Vol. 3. New York: Dover.

Hestenes, D. (2009). "Hunting for Snarks in Quantum Mechanics." AIP Conf. Proc. 1193,.

Hilgetag, C. C. and S. Grant (2010). "Cytoarchitectural differences are a key determinant of laminar projection origins in the visual cortex." Neuroimage. 51(3): 1006-1017. doi: 1010.1016/j.neuroimage.2010.1003.1006. Epub 2010 Mar 1006.

Hilgetag, C. C., M. A. O'Neill, et al. (2000). "Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor." Philos Trans R Soc Lond B Biol Sci. 355(1393): 71-89.

Hinton, G., E. (1990). "Mapping part-whole hierarchies into connectionist networks." Artif. Intell. 46(1-2): 47-75.

Hinton, G. E., P. Dayan, et al. (1995). "The "wake-sleep" algorithm for unsupervised neural networks." Science 268(5214): 1158-1161.

Hoffman, A. and C. E. Wayne (2008). "Counter-propagating two-soliton solutions in the Fermi–Pasta–Ulam lattice." Nonlinearity 21(12): 2911.

Huber, R., S. K. Esser, et al. (2007). "TMS-induced cortical potentiation during wakefulness locally increases slow wave activity during sleep." PLoS One. 2(3): e276.

Huber, R., M. F. Ghilardi, et al. (2004). "Local sleep and learning." Nature. 430(6995): 78-81. Epub 2004 Jun 2006.

IBM. (2009). "www-03.ibm.com/press/us/en/pressrelease/28842.wss#release." 2013.

Ionescu, L. M. (2008). "On The Arrow of Time." arXiv:0708.4180 [physics.gen-ph].

Itti, L., G. Rees, et al. (2005). Neurobiology of Attention, Elsevier Science.

Jaynes, E. T. (1988). How Does the Brain Do Plausible Reasoning? Maximum-Entropy and Bayesian Methods in Science and Engineering. G. Erickson and C. R. Smith, Springer Netherlands. 31-32: 1-24.

Ji, D. and M. A. Wilson (2007). "Coordinated memory replay in the visual cortex and hippocampus during sleep." Nat Neurosci. 10(1): 100-107. Epub 2006 Dec 2017.

Kaiser, J. and W. Lutzenberger (2005). "Cortical oscillatory activity and the dynamics of auditory memory processing." Rev Neurosci. 16(3): 239-254.

Kandel, E., J. Schwartz, et al. (2012). Principles of Neural Science, Fifth Edition, McGraw-Hill Education.

Kanerva, P. (1994). "The Spatter Code for encoding concepts at many levels." In M. Marinaro and P.G. Morasso (eds.), ICANN '94, Proceedings of International Conference on Artificial Neural Networks (Sorrento, Italy), vol. 1, pp. 226-229. London: Springer-Verlag.

Kanerva, P. (2010). What we mean when we say “What’s the dollar of mexico?”: Prototypes and mapping in concept space.

Karalis, A., J. D. Joannopoulos, et al. (2008). "Efficient wireless non-radiative mid-range energy transfer." Annals of Physics 323(1): 34-48.

Karlsson, M. P. and L. M. Frank (2009). "Awake replay of remote experiences in the hippocampus." Nat Neurosci. 12(7): 913-918. doi: 910.1038/nn.2344. Epub 2009 Jun 1014.

Kastner, R. E. (2012). "The possibilist transactional interpretation and relativity." Foundations of Physics 42(8): 1094-1113.

Kattler, H., D. J. Dijk, et al. (1994). "Effect of unilateral somatosensory stimulation prior to sleep on the sleep EEG in humans." J Sleep Res. 3(3): 159-164.

Kirby, K. G. (2006). Helmholtz Machines: Memory, Fluctuation, Dreams. Academy of Media Arts, Cologne.

Kirkpatrick, S., C. D. Gelatt, et al. (1983). "Optimization by Simulated Annealing." Science 220(4598): 671-680.

Kjeldsen, H. (2008). Automatic Vowel Recognition with GPU based Holographic Neural Network. Master's thesis, Groningen University.

Kjeldsen, H., M. Kaiser, et al. (2012). "A neocortical delta frequency generator modulates layer2/3 sparse spiking via nested theta rhythms." Society for Neuroscience.

Kogge, P. (2011). "Next-Generation Supercomputers." IEEE Spectrum.

Kolmogorov, A. N. (1941). "Interpolation and extrapolation of stationary random sequences." Izv. Akad. Nauk SSSR Ser. Mat. 5, 3-14.

Kopell, N., M. A. Whittington, et al. (2011). "Neuronal assembly dynamics in the beta1 frequency range permits short-term memory." Proceedings of the National Academy of Sciences 108(9): 3779-3784.

Kuhn, T. S. and I. Hacking (2012). The Structure of Scientific Revolutions, University of Chicago Press.

Kurs, A., A. Karalis, et al. (2007). "Wireless Power Transfer via Strongly Coupled Magnetic Resonances." Science 317(5834): 83-86.

Kurzweil, R. (2005). The Singularity is Near, Viking Books.

Lakatos, P., G. Karmos, et al. (2008). "Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection." Science 320(5872): 110-113.

Lakatos, P., A. S. Shah, et al. (2005). "An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex." J Neurophysiol. 94(3): 1904-1911. Epub 2005 May 1918.

Lance, F. (2009). "The status of the P versus NP problem." Commun. ACM 52(9): 78-86.

Landauer, R. (1961). "Irreversibility and Heat Generation in the Computing Process." IBM Journal of Research and Development 5(3): 183-191.

Lasenby, J., A. N. Lasenby, et al. (2000). "A unified mathematical language for physics and engineering in the 21st century." Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 358(1765): 21-39.

Lashley, K. S. (1949). "Persistent problems in the evolution of mind." The Quarterly review of biology 24(1): 28-42.

Lee, L., L. M. Harrison, et al. (2003). "A report of the functional connectivity workshop, Dusseldorf 2002." NeuroImage 19(2): 457-465.

Lee, S., K. Sen, et al. (2009). "Cortical gamma rhythms modulate NMDAR-mediated spike timing dependent plasticity in a biophysical model." PLoS Comput Biol. 5(12): e1000602. doi: 1000610.1001371/journal.pcbi.1000602. Epub 1002009 Dec 1000611.

Lehar, S. (2010). "The Constructive Aspect of Visual Perception: A Gestalt Field Theory Principle of Visual Reification Suggests a Phase Conjugate Mirror Principle of Perceptual Computation." Self-published manuscript.

Levy, B. C. and A. J. Krener (1996). "Stochastic mechanics of reciprocal diffusions." Journal of Mathematical Physics 37: 769-802.

Li, K., L. Guo, et al. (2009). "Review of methods for functional brain connectivity detection using fMRI." Computerized Medical Imaging and Graphics 33(2): 131-139.

Libet, B., C. A. Gleason, et al. (1983). "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the unconscious initiation of a freely voluntary act." Brain 106(3): 623-642.

Lidar, D. A., I. L. Chuang, et al. (1998). "Decoherence-free subspaces for quantum computation." Physical Review Letters 81(12): 2594.

Lipton, R. (2013). "Gödel's Lost Letter." 2013.

Litt, A., C. Eliasmith, et al. (2006). "Is the brain a quantum computer?" Cognitive Science 30(3): 593-603.

Lloyd, S., L. Maccone, et al. (2011a). "Quantum mechanics of time travel through post-selected teleportation." Physical Review D 84(2): 025007.

Lloyd, S., L. Maccone, et al. (2011b). "Closed Timelike Curves via Postselection: Theory and Experimental Test of Consistency." Physical Review Letters 106(4): 040403.

Lloyd, S., L. Maccone, et al. (2011c). "A reply to "Problems with modelling closed timelike curves as post-selected teleportation"." arXiv:1108.0153 [quant-ph].

Longo, M. J. (2011). "Detection of a Dipole in the Handedness of Spiral Galaxies with Redshifts z~ 0.04." Physics Letters B 699(4): 224-229.

Louie, K. and M. A. Wilson (2001). "Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep." Neuron. 29(1): 145-156.

Lvovsky, A. I., R. Ghobadi, et al. (2013). "Observation of micro-macro entanglement of light." Nat Phys 9(9): 541-544.

Mackenzie, D. (2006). "Classifying Hyperbolic Manifolds -All's Well that Ends Well." WHATS HAPPENING IN THE MATHEMATICAL SCIENCES 6.

Makarov, V. A., A. N. Pavlov, et al. (2010). "Stability of neural firing in the trigeminal nuclei under mechanical whisker stimulation." Comput Intell Neurosci 340541: 6.

Maldacena, J. (1999). "The Large-N Limit of Superconformal Field Theories and Supergravity." International Journal of Theoretical Physics 38(4): 1113-1133.

Malinow, R. and R. C. Malenka (2002). "AMPA receptor trafficking and synaptic plasticity." Annu Rev Neurosci. 25: 103-126. Epub 2002 Mar 2004.

Mandell, A. J. and K. A. Selz (1997). "Is the EEG a strange attractor? Brain stem neuronal discharge patterns and electroencephalographic rhythms." The impact of chaos on science and society: 64-96.

Massimini, M., F. Ferrarelli, et al. (2007). "Triggering sleep slow waves by transcranial magnetic stimulation." Proc Natl Acad Sci U S A. 104(20): 8496-8501. Epub 2007 May 8494.

Matzke, D., M. Manthey, et al. (2003). "Quantum Geometric Algebra."

Matzke, D. J. (2002). Quantum computation using geometric algebra.

McCarley, R. W. (2007). "Neurobiology of REM and NREM sleep." Sleep Med. 8(4): 302-330. Epub 2007 Apr 2030.

McMains, S. and S. Kastner (2011). "Interactions of Top-Down and Bottom-Up Mechanisms in Human Visual Cortex." The Journal of Neuroscience 31(2): 587-597.

Mignot, E. and J. R. Huguenard (2009). "Resting Our Cortices by Going DOWN to Sleep." Neuron 63(6): 719-721.

Misra, B. and E. C. G. Sudarshan (1977). "The Zeno's paradox in quantum theory." Journal of Mathematical Physics 18(4): 756-763.

Muzur, A. (2005). "Toward an integrative theory of sleep and dreaming." Journal of Theoretical Biology 233(1): 103-118.

Nadasdy, Z., H. Hirase, et al. (1999). "Replay and time compression of recurring spike sequences in the hippocampus." J Neurosci. 19(21): 9497-9507.

Nedungadi, A., G. Rangarajan, et al. (2009). "Analyzing multiple spike trains with nonparametric granger causality." Journal of Computational Neuroscience 27(1): 55-64.

Neumann, J. (2001). Holistic Processing of Hierarchical Structures in Connectionist Networks, University of Edinburgh. PhD Thesis.

Neumann, J. (2002). "Learning the systematic transformation of holographic reduced representations." Cognitive Systems Research 3(2): 227 - 235.

Nicolas, L., M. Furstoss, et al. (1998). "Analogy electromagnetism-acoustics: Validation and application to local impedance active control for sound absorption." The European Physical Journal Applied Physics 4(01): 95-100.

Nicolelis, M. A., L. A. Baccala, et al. (1995). "Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system." Science. 268(5215): 1353-1358.

Niklasson, L. F. and T. I. M. Gelder (1994). "On being systematically connectionist." Mind & Language 9(3): 288-302.

Nimtz, G. (2011). "Tunneling Confronts Special Relativity." Foundations of Physics 41(7): 1193-1199.

Ohiorhenuan, I. E., F. Mechler, et al. (2010). "Sparse coding and high-order correlations in fine-scale cortical networks." Nature. 466(7306): 617-621. doi: 610.1038/nature09178. Epub 02010 Jul 09174.

Ohzeki, M. and H. Nishimori (2011). "Quantum annealing: An introduction and new developments." Journal of Computational and Theoretical Nanoscience 8(6): 963-971.

Oostenveld, R., P. Fries, et al. (2011). "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data." Computational intelligence and neuroscience 2011: 1.

Orlov, Y. F. (2002). "Quantumlike bits and logic gates based on classical oscillators." Physical review. A 66(5).

Oudiette, D., I. Constantinescu, et al. (2011). "Evidence for the Re-Enactment of a Recently Learned Behavior during Sleepwalking." PLoS ONE 6(3): e18056.

Paillasseur, S., J. H. Thomas, et al. (2011). "Regularization for improving the deconvolution in real-time near-field acoustic holography." J Acoust Soc Am. 129(6): 3777-3787. doi: 3710.1121/3771.3586790.

Palva, J. M., S. Palva, et al. (2005). "Phase synchrony among neuronal oscillations in the human cortex." J Neurosci. 25(15): 3962-3972.

Patyk, A. (2010). Geometric algebra model of distributed representations. Geometric Algebra Computing, Springer: 401-430.

Pawłowski, M. (2006). "Superfast algorithms and the halting problem in geometric algebra." arXiv preprint quant-ph/0611051.

Paxinos, G. and C. Watson (1998). The rat brain in stereotaxic coordinates, Academic Press.

Penrose, R. (1982). Time-Asymmetry, Cosmological Uniformity and Space-Time Singularities. Progress in Cosmology. A. W. Wolfendale, Springer Netherlands. 99: 87-88.

Penrose, R. (1989). The Emperor's New Mind, Oxford University Press.

Penrose, R. (1994). Shadows of the Mind: A Search for the Missing Science of Consciousness, {Oxford University Press, USA}.

Penrose, R. (2001). The Road to Reality: A Complete Guide to the Laws of the Universe.

Penrose, R. (2010). Cycles of Time: An Extraordinary New View of the Universe

Peruš, M. (2000). "Neural networks as a basis for quantum associative networks." Neural Netw. World 10(6): 1001-1013.

Perus, M., H. Bischof, et al. (2007). "A Natural Quantum Neural-Like Network." NeuroQuantology 3(3).

Peter, B., P. George, et al. (2002). "Super-resolution in time-reversal acoustics." The Journal of the Acoustical Society of America 111(1): 230-248.

Petersen, C. C. and S. Crochet (2013). "Synaptic computation and sensory processing in neocortical layer 2/3." Neuron. 78(1): 28-48. doi: 10.1016/j.neuron.2013.1003.1020.

Peyrache, A., N. Dehghani, et al. (2012). "Spatiotemporal dynamics of neocortical excitation and inhibition during human sleep." Proceedings of the National Academy of Sciences 109(5): 1731-1736.

Peyrache, A., M. Khamassi, et al. (2009). "Replay of rule-learning related neural patterns in the prefrontal cortex during sleep." Nat Neurosci. 12(7): 919-926. doi: 910.1038/nn.2337. Epub 2009 May 1031.

Piccinini, G. (2007). "Computationalism, the Church–Turing thesis, and the Church–Turing fallacy." Synthese 154(1): 97-120.

Plate, T. (1991). Holographic reduced representations: convolution algebra for compositional distributed representations. Proceedings of the 12th international joint conference on Artificial intelligence - Volume 1. Sydney, New South Wales, Australia, Morgan Kaufmann Publishers Inc.: 30-35.

Plate, T. (1993). "Holographic Recurrent Networks. ." Advances in Neural Information Processing Systems 5.

Plate, T. (2003). Holographic Reduced Representation - Distributed Representation for Cognitive Structures

The University of Chicago Press.

Plenz, D. and T. C. Thiagarajan (2007). "The organizing principles of neuronal avalanches: cell assemblies in the cortex?" Trends Neurosci. 30(3): 101-110. Epub 2007 Feb 2001.

Poggio, T. (1973). "On holographic models of memory." Kybernetik 12(4): 237-238.

Porter, M. A., N. J. Zabusky, et al. (2009). "Fermi, Pasta, Ulam and the birth of experimental mathematics." American Scientist 97.6.

Pribram, K. H. (1971). Languages of the brain: experimental paradoxes and principles in neuropsychology, Prentice-Hall.

Pribram, K. H. (1986). "Convolution and Matrix Systems as Content Addressible Distributed Brain Processes in Perception and Memory." Journal of Neuroliquistics 2(2).

Pribram, K. H. (1991). Brain and perception : holonomy and structure in figural processing. Hillsdale, N.J. :, Lawrence Erlbaum Associates.

Pu, J., H. Gong, et al. (2013). "Developing neuronal networks: Self-organized criticality predicts the future." Scientific Reports 3, Article number:1081doi:10.1038/srep01081.

Quiroga, R. Q., Z. Nadasdy, et al. (2004). "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering." Neural Computation 16(8): 1661-1687.

Rabal, H. (2001). "Holographic analogues." Optik-International Journal for Light and Electron Optics 112(4): 153-156.

Ramkumar, P., M. Jas, et al. (2013). "Feature-Specific Information Processing Precedes Concerted Activation in Human Visual Cortex." The Journal of Neuroscience 33(18): 7691-7699.

Rasch, M. J., A. Gretton, et al. (2008). "Inferring spike trains from local field potentials." J Neurophysiol. 99(3): 1461-1476. Epub 2007 Dec 1426.

Ray, S., S. S. Hsiao, et al. (2008). "Effect of stimulus intensity on the spike-local field potential relationship in the secondary somatosensory cortex." J Neurosci. 28(29): 7334-7343. doi: 7310.1523/JNEUROSCI.1588-7308.2008.

Riedner, B. A., V. V. Vyazovskiy, et al. (2007). "Sleep homeostasis and cortical synchronization: III. A high-density EEG study of sleep slow waves in humans." Sleep. 30(12): 1643-1657.

Roberts, M. J., E. Lowet, et al. (2013). "Robust gamma coherence between macaque V1 and V2 by dynamic frequency matching." Neuron. 78(3): 523-536. doi: 510.1016/j.neuron.2013.1003.1003.

Roopun, A. K., M. A. Kramer, et al. (2008). "Period concatenation underlies interactions between gamma and beta rhythms in neocortex." Front Cell Neurosci. 2: 1. Epub 2008 Apr 2008.

Roopun, A. K., F. E. Lebeau, et al. (2010). "Cholinergic neuromodulation controls directed temporal communication in neocortex in vitro." Front Neural Circuits. 4:8.(doi): 10.3389/fncir.2010.00008.

Rose, G. and W. G. Macready (2007). "An introduction to quantum annealing." DWave Technical Document 712.

Rouiller, E. M., G. M. Simm, et al. (1991). "Auditory corticocortical interconnections in the cat: evidence for parallel and hierarchical arrangement of the auditory cortical areas." Exp Brain Res. 86(3): 483-505.

Sakata, S. and K. D. Harris (2009). "Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex." Neuron. 64(3): 404-418. doi: 410.1016/j.neuron.2009.1009.1020.

Sayed, A. H. and T. Kailath (2001). "A survey of spectral factorization methods." Numerical Linear Algebra with Applications 8(6-7): 467-496.

Schlosshauer, M. (2005). "Decoherence, the measurement problem, and interpretations of quantum mechanics." Reviews of Modern Physics 76(4): 1267.

Schmiedt-Fehr, C., S. Duhl, et al. (2011). "Age-related increases in within-person variability: Delta and theta oscillations indicate that the elderly are not always old." Neurosci Lett. 495(2): 159-163. Epub 2011 Mar 2029.

Schneider, T. D. (1991). "Theory of molecular machines. II. Energy dissipation from molecular machines." Journal of Theoretical Biology 148(1): 125-137.

Schönemann, P. H. (1987). "Some algebraic relations between involutions, convolutions, and correlations, with applications to holographic memories." Biological Cybernetics 56(5-6): 367-374.

Schott, R. and G. S. Staples (2010). "Reductions in computational complexity using Clifford algebras." Advances in applied Clifford algebras 20(1): 121-140.

Schroeder, C. E. and P. Lakatos (2009). "The gamma oscillation: master or slave?" Brain Topogr. 22(1): 24-26. doi: 10.1007/s10548-10009-10080-y. Epub 12009 Feb 10510.

Siegel, J. M. (2001). "The REM sleep-memory consolidation hypothesis." Science. 294(5544): 1058-1063.

Smith, C. (1995). "Sleep states and memory processes." Behav Brain Res. 69(1-2): 137-145.

Smolensky, P. (1990). "Tensor product variable binding and the representation of symbolic structures in connectionist systems." Artificial Intelligence 46(1-2): 159-216.

Stanzial, D., N. Prodi, et al. (1996). "Reactive acoustic intensity for general fields and energy polarization." The Journal of the Acoustical Society of America 99: 1868.

Susskind, L. (1994). "The world as a hologram." arXiv preprint hep-th/9409089.

Takeuchi, D., T. Hirabayashi, et al. (2011). "Reversal of Interlaminar Signal Between Sensory and Memory Processing in Monkey Temporal Cortex." Science 331(6023): 1443-1447.

Talbot, M. (1996). The Holographic Universe, HarperCollins.

Tarczy-Hornoch, K., K. A. C. Martin, et al. (1999). "Intracortical Excitation of Spiny Neurons in Layer 4 of Cat Striate Cortex In Vitro." Cerebral Cortex 9(8): 833-843.

Tegmark, M. (2000). "Importance of quantum decoherence in brain processes." Physical Review E 61(4): 4194.

Thomas, J. H., V. Grulier, et al. (2010). "Real-time near-field acoustic holography for continuously visualizing nonstationary acoustic fields." J Acoust Soc Am. 128(6): 3554-3567. doi: 3510.1121/3551.3504656.

Thomson, A. M., D. C. West, et al. (2002). "Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2-5 of adult rat and cat neocortex: triple intracellular recordings and biocytin labelling in vitro." Cereb Cortex. 12(9): 936-953.

Tononi, G. and C. Cirelli (2006). "Sleep function and synaptic homeostasis." Sleep Med Rev. 10(1): 49-62. Epub 2005 Dec 2022.

Torrence, C. and G. P. Compo (1998). "A practical guide to wavelet analysis." Bull. Am. Meteorol. Soc., 79, 61-78, .

Traub, R. D., D. Contreras, et al. (2005). "Single-column thalamocortical network model exhibiting gamma oscillations, sleep spindles, and epileptogenic bursts." J Neurophysiol. 93(4): 2194-2232. Epub 2004 Nov 2193.

Traub, R. D., M. A. Whittington, et al. (1996). "A mechanism for generation of long-range synchronous fast oscillations in the cortex." Nature. 383(6601): 621-624.

Tsotsos, J. K. (2005). Computational Foundations for Attentive Processes. Neurobiology of Attention. L. Itti, G. Rees and J. Tsotsos, Academic Press, Elsevier.

Tsotsos, J. K. and N. D. B. Bruce (2008). "Computational foundations for attentive processes." Scholarpedia 3: 6545.

Turrigiano, G. G., K. R. Leslie, et al. (1998). "Activity-dependent scaling of quantal amplitude in neocortical neurons." Nature. 391(6670): 892-896.

Valderrama, M., B. Crépon, et al. (2012). "Human Gamma Oscillations during Slow Wave Sleep." PLoS ONE 7(4): e33477.

Van Dongen, H. P., J. A. Caldwell, Jr., et al. (2006). "Investigating systematic individual differences in sleep-deprived performance on a high-fidelity flight simulator." Behav Res Methods. 38(2): 333-343.

Vassalli, A. and D. J. Dijk (2009). "Sleep function: current questions and new approaches." Eur J Neurosci. 29(9): 1830-1841. Epub 2009 Apr 1828.

Vedral, V. (2011). "Living in a quantum world." Scientific American 304(6): 38-43.

Vertes, R. P. (2004). "Memory consolidation in sleep; dream or reality." Neuron. 44(1): 135-148.

Vinck, M., F. P. Battaglia, et al. (2012). "Improved measures of phase-coupling between spikes and the Local Field Potential." J Comput Neurosci. 33(1): 53-75. doi: 10.1007/s10827-10011-10374-10824. Epub 12011 Dec 10821.

Vito, E. D., L. Rosasco, et al. (2005). Learning from examples as an inverse problem. Journal of Machine Learning Research.

Von Baeyer, H. C. (2004). Information: The New Language of Science, Harvard University Press.

Vyazovskiy, V. V., C. Cirelli, et al. (2008). "Molecular and electrophysiological evidence for net synaptic potentiation in wake and depression in sleep." Nat Neurosci. 11(2): 200-208. doi: 210.1038/nn2035. Epub 2008 Jan 1020.

Vyazovskiy, V. V. and K. D. Harris (2013). "Sleep and the single neuron: the role of global slow oscillations in individual cell rest." Nat Rev Neurosci. 14(6): 443-451. doi: 410.1038/nrn3494. Epub 2013 May 1032.

Vyazovskiy, V. V., U. Olcese, et al. (2011). "Local sleep in awake rats." Nature 472(7344): 443-447.

Vyazovskiy, V. V., U. Olcese, et al. (2009). "Cortical firing and sleep homeostasis." Neuron. 63(6): 865-878. doi: 810.1016/j.neuron.2009.1008.1024.

Vyazovskiy, V. V. and I. Tobler (2008). "Handedness Leads to Interhemispheric EEG Asymmetry During Sleep in the Rat." Journal of Neurophysiology 99(2): 969-975.

W. Duana, J. P., K. V. Horoshenkovb, R. Kirbya (2012). "A comparison between measured and predicted complex intensity in a flanged cylindrical pipe." Proceedings of the Acoustics 2012 Nantes Conference.

Wegner, D. M. (2002). The Illusion of Conscious Will, Cambridge University Press.

Werbos, P. J. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. Harvard University. Ph.D dissertation.

Werbos, P. J. (1990). "Backpropagation through time: what it does and how to do it." Proceedings of the IEEE 78.

Werbos, P. J. (2002). "Classical ODE and PDE Which Obey Quantum Dynamics." INTERNATIONAL JOURNAL OF BIFURCATION AND CHAOS IN APPLIED SCIENCES AND ENGINEERING 12.

Werbos, P. J. (2008). "Bell's Theorem, Many Worlds and Backwards-Time Physics: Not Just a Matter of Interpretation." International Journal of Theoretical Physics 47(11): 2862-2874.

Werbos, P. J. D., Ludmila (2000). "The Backwards-Time Interpretation of Quantum Mechanics - Revisited With Experiment." arXiv:quant-ph/0008036.

Westheimer, G. (2008). "Was Helmholtz a Bayesian?" Perception. 37(5): 642-650.

Westlake, P. (1970). "The possibilities of neural holographic processes within the brain." Kybernetik 7(4): 129-153.

Wheeler, J. A. and R. P. Feynman (1945). "Interaction with the Absorber as the Mechanism of Radiation." Reviews of Modern Physics 17(2-3): 157-181.

Whittington, M. A., M. O. Cunningham, et al. (2011). "Multiple origins of the cortical gamma rhythm." Dev Neurobiol. 71(1): 92-106. doi: 110.1002/dneu.20814.

Wilber, K. (1982). The Holographic Paradigm and Other Paradoxes: Exploring the Leading Edge of Science, Shambhala.

Williams, E. G. and J. D. Maynard (1980). "Holographic imaging without the wavelength resolution limit." Physical Review Letters 45(7): 554.

Williams, E. G. and N. P. Valdivia (2010). "Near-Field Electromagnetic Holography in Conductive Media." Antennas and Propagation, IEEE Transactions on 58(4): 1181-1192.

Wilson, G. T. (1972). "The Factorization of Matricial Spectral Densities." SIAM Journal on Applied Mathematics 23(4): 420-426.

Young, A. M. (1976). The geometry of meaning, Delacorte Press/S. Lawrence.

Yourgrau, P. (2006). A World Without Time: The Forgotten Legacy of Gödel and Einstein.

Zheng, W., B. Liu, et al. (2005). "Holographic associative memory by phase conjugate of four-wave-mixing in Sc:Fe:LiNbO3 crystal." Optics Communications 246(4-6): 297-301.

Zou, C., C. Ladroue, et al. (2010). "Identifying interactions in the time and frequency domains in local and global networks-A Granger Causality Approach." BMC Bioinformatics 11(1): 1-17.

Appendix A: Near-field electromagnetic holography

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Near-field electromagnetic holography framework for the Utah array
% Henrik D. Kjeldsen, 2013, Matlab code
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

sr_ori = 30000; %original sample rate
sr = 15000; %sample rate to use for NEH
time = 0.1; %length of time segment in seconds

gridn = 10; %recording array size
gridx = 10;
gridy = 10;
gridmul = 3; %super-resolution factor
pad = 1; %padding factor, 1 = no padding, 2 = double size (not used)

%conductivity, rho, values along and across layers
%mS/cm cond(l,d) l = layer; 1 = 2/3; 2 = 4; 3 = 5; 4 = 6; d = 1
%is along layers, d = 2 is across; values from Goto 2010.
%5 = whitematter from other source

cond(1,1) = 3.19;
cond(1,2) = 2.31;

cond(2,1) = 3.25;
cond(2,2) = 2.40;

cond(3,1) = 3.53;
cond(3,2) = 2.28;

cond(4,1) = 2.94;
cond(4,2) = 2.68;

cond(5,1) = 2.25;
cond(5,2) = 2.25;

%array to hold conductivity values
gcond = zeros(3,gridn*gridmul,gridn*gridmul);

if gridmul == 3
%expand layer mask for super-resolution
layermask2 = zeros(30);
X = 1:3:30;
Y = 1:3:30;
for x = 1:10
    for y = 1:10
        layermask2(X(x):X(x)+3,Y(y):Y(y)+3) = layermask(x,y);

    end
end
layermask2 = layermask2(1:30,1:30);
%populate conductivity array
gcond(4,find(layermask2 == 6)) = sqrt(cond(4,1).*cond(4,2));
gcond(4,find(layermask2 == 5)) = sqrt(cond(3,1).*cond(3,2));
gcond(4,find(layermask2 == 4)) = sqrt(cond(2,1).*cond(2,2));
gcond(4,find(layermask2 == 3)) = sqrt(cond(1,1).*cond(1,2));
```

```
gcond(4,find(layermask2 == 2)) = sqrt(cond(1,1).*cond(1,2));
gcond(4,find(layermask2 == 0)) = sqrt(cond(5,1).*cond(5,2));
else
%populate conductivity array
gcond(4,find(layermask == 6)) = sqrt(cond(4,1).*cond(4,2));
gcond(4,find(layermask == 5)) = sqrt(cond(3,1).*cond(3,2));
gcond(4,find(layermask == 4)) = sqrt(cond(2,1).*cond(2,2));
gcond(4,find(layermask == 3)) = sqrt(cond(1,1).*cond(1,2));
gcond(4,find(layermask == 2)) = sqrt(cond(1,1).*cond(1,2));
gcond(4,find(layermask == 0)) = sqrt(cond(5,1).*cond(5,2));
end

er = 3400;  %relative permittivity
%Joy 1999 cortex, white matter is 2720
e0 = 8.854e-12; %vacuum
e = e0 * er; %specific

ur = 0.99999103; %relative
u0 = 4e-7*pi; %H/m free space permeability
u = u0 * ur; %specific

%setup frequency range
f = -sr/2:1/time:(sr)/2;
f = f([1:(length(f)-1)/2 (length(f)-1)/2+2:end]);
f = f([1:(length(f))/2+1 (length(f))/2+2:end]); %odd
find(f == 0) %must be empty (no zeros)
%in radians
w = 2*pi*f;

%calculate characteristic wave-number
ehat = e*(1 + i*repmat(squeeze(gcond(4,:,:)), [1 1
ceil(sr*time)]) ./shiftdim(repmat((w*e), [gridn*gridmul 1
gridn*gridmul ]),2));
k = sqrt(ehat.*u.*shiftdim(repmat((w), [gridn*gridmul 1
gridn*gridmul ]),2));

%spatial wave-numbers (Utah array spacing is 0.0004m) (K-wave toolbox)
kx = getWavenumbers(gridn*gridmul*pad+1, 0.0004/gridmul, 1);
ky = getWavenumbers(gridn*gridmul*pad+1, 0.0004/gridmul, 1);
kx = kx(2:end);
ky = ky(2:end);
kx = fftshift(kx);
ky = fftshift(ky);

%calculate spatially dependent kz depending on near- or far-field
kz =
complex(zeros(gridn*gridmul*pad,gridn*gridmul*pad,round(sr*time*pad)));
for x = 1:gridn*gridmul*pad
    for y = 1:gridn*gridmul*pad
         for t = 1:length(k)*pad
              if kx(x)^2+ky(y)^2 <= k(x,y,t)^2
                kz(x,y,t) = sqrt(k(x,y,t)^2-kx(x)^2-ky(y)^2);
%far-field
              else
                kz(x,y,t) = -1i*sqrt(kx(x)^2+ky(y)^2-k(x,y,t)^2);
%near-field
              end
         end
    end
end
```

```
%collect wave-vectors in one structure
K = zeros(3,gridn*gridmul*pad,gridn*gridmul*pad,ceil(sr*time)*pad);
K(1,:,:,:) = repmat(kx,[1 gridn*gridmul*pad ceil(sr*time)*pad]);
K(2,:,:,:) = repmat(ky,[1 gridn*gridmul*pad ceil(sr*time)*pad]);
K(3,:,:,:) = kz;

%create k-space filter
clear K_window
alpha = 0.4;
kc = 12000;
for Ym=1:gridx*gridmul*pad
for Xm=1:gridx*gridmul*pad
kr=abs(sqrt(kx(Xm)^2+ky(Ym)^2));
if kr<=kc
    K_window(Ym,Xm)=1-(0.5*exp(((kr/kc)-1)/alpha));
else
    K_window(Ym,Xm)=0.5*exp((1-(kr/kc))/alpha);
end
end
end
K_window = repmat(K_window, [1 1 ceil(sr*time)]);

%load NS6 data and layout file (Blackrock library)
openNSx('read');
myArrayMap = KTUEAMapFile('SN 1017-000170.cmp')

%re-arrange and filter data
for ch = 1:96
[x y] = myArrayMap.getChannelColumnRow(ch);

tmp = double(NS6.Data(ch,30*sr_ori+1:30*sr_ori+time*sr_ori)); %take
segment of data of length time starting 30 sec in

%filter out 50hz
tim = 0:(1/sr_ori):(time);
temp = idealfilter(timeseries(tmp(1:ceil(time*sr_ori)),tim(1:end-
1)),[48 52],'notch');
df = temp.Data;

%decimate for target sample rate
dec = sr_ori/sr;
df = decimate(df,dec);

%create data structure
map(11-(y+1),(x+1),:) = df;
end

%apply mask to remove bad channels
map = map.*repmat(deltamask, [1 1 ceil(time*sr)]);

%pre-populate "missing" super-resolution data-points with
%interpolated super-resolution values similar to k-wave
%toolbox; actual super-resolution values are then filled in
%during propagation
del = [1 0 0];
mapinterp = papoulisgerchberg(map,del,gridmul);
%function by Karim Krichane

%calculate E field in recording plane from potentials
[Ex Ey Ez] = gradient((mapinterp),0.0004/gridmul,0.0004/gridmul,1/sr);
```

```
%setup k-space propagator
z = 0.0001; %propagation distance in meters
G = (exp(-1i*kz*z));

%propagate and filter E field
E0(1,:,:,:) = fftshift(fftn(-Ex)).*G.*K_window;
E0(2,:,:,:) = fftshift(fftn(-Ey)).*G.*K_window;
E0(3,:,:,:) = fftshift(fftn(-Ez)).*G.*K_window;

%calculate H field
H0 = cross((K),E0);

Hx = ifftn(ifftshift(squeeze(H0(1,:,:,:))));
Hy = ifftn(ifftshift(squeeze(H0(2,:,:,:))));
Hz = ifftn(ifftshift(squeeze(H0(3,:,:,:))));

%create structures to hold all field directions at
%reconstruction plane
H(1,:,:,:) = Hx;
H(2,:,:,:) = Hy;
H(3,:,:,:) = Hz;

E(1,:,:,:)  = ifftn(ifftshift(squeeze(E0(1,:,:,:))));
E(2,:,:,:)  = ifftn(ifftshift(squeeze(E0(2,:,:,:))));
E(3,:,:,:)  = ifftn(ifftshift(squeeze(E0(3,:,:,:))));

%Poynting vector
S = cross(E,H);

%current density
J(1,:,:,:) = repmat(squeeze(gcond(4,:,:)),[1 1
ceil(sr*time)]).*squeeze(E(1,:,:,:));
J(2,:,:,:) = repmat(squeeze(gcond(4,:,:)),[1 1
ceil(sr*time)]).*squeeze(E(2,:,:,:));
J(3,:,:,:) = repmat(squeeze(gcond(4,:,:)),[1 1
ceil(sr*time)]).*squeeze(E(3,:,:,:));

[pomx,pomy,pomz]=meshgrid((1:gridx*gridmul),(1:gridy*gridmul),1:ceil(s
r*time));

%current source density
CSD =
divergence((pomx),(pomy),(pomz),squeeze(J(1,:,:,:)),squeeze(J(2,:,:,:)
),squeeze(J(3,:,:,:)));

%energy source density
ESD =
divergence((pomx),(pomy),(pomz),squeeze(S(1,:,:,:)),squeeze(S(2,:,:,:)
),squeeze(S(3,:,:,:)));

%dissipation
D = dot(J,E);

%it is helpful to filter these measures before visualization
```

Appendix B: Granger causality

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Non-parametric Granger Causality based on complex wavelets
% Henrik D. Kjeldsen, 2012, Matlab code
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%get some example data contained in spike structure
LII = (spike.lfps{find(cell2mat(spike.location) == (86))});
LV = (spike.lfps{find(cell2mat(spike.location) == (46))});

%or for csd
LII = spike.csd(:,ind2sub([10 10],82));
LV = spike.csd(:,ind2sub([10 10],42));

%or straight to csd so skip till granger
c2 = spike.cw03to15hz{ind2sub([10 10],82)};
c1 = spike.cw03to15hz{ind2sub([10 10],42)};

%or spikes, does not work with granger if too singular
ts = (spike.timestamp{1,find(cell2mat(spike.location) == (26))});
d = round(ts(find(ts > 0))*1000);
d = d(find(0 < d & d < 60000));
LII = zeros(60000,1);
LII(d) = 1;

ts = (spike.timestamp{1,find(cell2mat(spike.location) == (35))});
d = round(ts(find(ts > 0))*1000);
d = d(find(0 < d & d < 60000));
LV = zeros(60000,1);
LV(d) = 1;

%setup scales/freqs of interest of wavelet analysis
%1/1000 is 1000 samples per sec. adjust as required
scale_array=frq2scal([20:1:45], 'cmor1-1', 1/1000, 10); %low gamma
scale_array=frq2scal([0.3:0.1:15], 'cmor1-1', 1/1000, 10);
%delta and some

nscales=wrev(scale_array);

%get the actual freqs used
f = scal2frq(nscales,'cmor1-1',1/1000);
ff = length(f);

%pre-filter
tmp = idealfilter(timeseries(LV(1:60000),tim),[0.3 15],'pass');
LV = tmp.data;
tmp = idealfilter(timeseries(LII(1:60000),tim),[0.3 15],'pass');
LII = tmp.data;

%do wavelet analysis of both signals, change the number of samples
below
c1=(cwt(LV(1:60000),nscales,'cmor1-1'));
c2=(cwt((LII(1:60000)),nscales,'cmor1-1'));

%inspect scalogram
sc1 = wscalogram('',c1,'power',0);
```

```
imagesc(log(sc1))
%set axis labels, you have to match to number of scales
set(gca,'YTick',[1:10:ff])
set(gca,'YTickLabel',f(1:10:ff))
xlabel('ms')
ylabel('hz')
%do a spectrum by summing over time
plot(log(sum(sc1,2)))
set(gca,'XTick',[1:10:ff])
set(gca,'XTickLabel',f(1:10:ff))
xlabel('hz')
%do semblance
ctc=c1.*conj((c2));
%get phase relations
tmpp = cos(angle(ctc));
%show phase semblance
imagesc(smooth2a((tmpp'),1,1000))
%set axis labels, you have to match to number of scales
set(gca,'YTick',[1:10:ff])
set(gca,'YTickLabel',f(1:10:ff))
xlabel('ms')
ylabel('hz')
%sum over time
plot((sum(tmpp',2)))
set(gca,'XTick',[1:10:ff])
set(gca,'XTickLabel',f(1:10:ff))
xlabel('hz')

%only if from fft wavelet analysis
c1 =c1';
c2 =c2';

%setup variables for granger analysis
k11 = c1.*conj((c1));
  k12  =c1.*conj((c2));
   k21  =c2.*conj((c1));
    k22  =c2.*conj((c2));

    %set to match number of scales/freqs
        s = 1:ff;
    %set to number of time steps (samples)
        tt = 1:60000;

    %zero granger variables, one for each direction
    I12 = zeros(length(s),length(tt));
    I21 = zeros(length(s),length(tt));

    %do granger
    for t = tt

        %%%%11
        f1k = k11(s,t);

        y = real(ifft(log(abs([wrev(f1k) ;(f1k)]))));
        n = length(y);

        w = [1;2*ones(n/2-1,1);ones(1-rem(n,2),1);zeros(n/2-1,1)];
        ym = real(ifft(exp(fft(w.*y))));

        ymm = fft(ym);
        E11 = ymm(1)*transpose(ymm(1));
```

```
        H11 = ymm*inv(ymm(1));

        %%%%%12
          f1k = k12(s,t);

        y = real(ifft(log(abs([wrev(f1k) ;(f1k)]))));
        n = length(y);

        w = [1;2*ones(n/2-1,1);ones(1-rem(n,2),1);zeros(n/2-1,1)];
        ym = real(ifft(exp(fft(w.*y))));

        ymm = fft(ym);
        E12 = ymm(1)*transpose(ymm(1));

        H12 = ymm*inv(ymm(1));

        %%%%%22
          f1k = k22(s,t);

        y = real(ifft(log(abs([wrev(f1k) ;(f1k)]))));
        n = length(y);

        w = [1;2*ones(n/2-1,1);ones(1-rem(n,2),1);zeros(n/2-1,1)];
        ym = real(ifft(exp(fft(w.*y))));

        ymm = fft(ym);
        E22 = ymm(1)*transpose(ymm(1));

        H22 = ymm*inv(ymm(1));

        %%%%%12
        f1k = k21(s,t);

        y = real(ifft(log(abs([wrev(f1k) ;(f1k)]))));
        n = length(y);

        w = [1;2*ones(n/2-1,1);ones(1-rem(n,2),1);zeros(n/2-1,1)];
        ym = real(ifft(exp(fft(w.*y))));

        ymm = fft(ym);
        E21 = ymm(1)*transpose(ymm(1));

        H21 = ymm*inv(ymm(1));

dbl2 = ((E22 ).* abs(H22+(E21/E11)*H21).^2) ;
%from first to second signal
I12(s,t) = log((k22(s,t)./sqrt(abs(dbl2(s+max(s))))));

dbl2 = ((E11 ).* abs(H11+(E12/E22)*H12).^2) ;
%from second to first signal
I21(s,t) = log((k11(s,t)./sqrt(abs(dbl2(s+max(s))))));

    end
```

Appendix C: Geometric algebra

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Cartan's matrix representation of Geometric Algebra
% Henrik D. Kjeldsen, 2012, Matlab code
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%unit/identity matrix
I = single([1 0;0 1]);

%Pauli's matrices
a1 = single([0 1;1 0]);
a2 = single([0 -i;i 0]);
a3 = single([1 0;0 -1]);

%SPARSE VERSION
%unit/identity matrix
I = sparse(([1 0;0 1]));

%Pauli's matrices
a1 = sparse([0 1;1 0]);
a2 = sparse([0 -1i;1i 0]);
a3 = sparse([1 0;0 -1]);

%setup basis blades, optimized version below
N = 12;
B = cell(2);
for i = 1:N
    B{i} = a1;
end
for i = 1:N
   if mod(i,2)
        K = (i + 1)/2;
    for j = 1:N-K-1

        B{i} = kron(B{i} ,a1);
    end
    B{i}  =  kron( B{i} ,a3);
    for j = 1:K-1

        B{i}  = kron( B{i} ,I);
    end
   else
        K = i/2;
    for j = 1:N-K-1

        B{i} = kron(B{i} ,a1);
    end
    B{i}  =  kron( B{i} ,a2);
    for j = 1:K-1

        B{i}  = kron( B{i} ,I);
    end
   end
end
```

```
%setup basis blades, optimized (Patyk's signature version)
N = 24;
B = cell(2);
for i = 1:N
    B{i} = a1;
end
for i = 1:N
   if mod(i,2)
        K = (i + 1)/2;
    for j = 1:N/2-K

        B{i} = kron(B{i} ,a1);
    end
    B{i}  =  kron( B{i} ,a3);
    for j = 1:K-1

        B{i}  = kron( B{i} ,I);
    end
   else
        K = i/2;
    for j = 1:N/2-K

        B{i} = kron(B{i} ,a1);
    end
    B{i}  =  kron( B{i} ,a2);
    for j = 1:K-1

        B{i}  = kron( B{i} ,I);
    end
   end
end


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%"Pat Smith" toy example

%define concepts and values
Pat = B{1};
male = B{2};
a66 = B{3};
name = B{4};
sex = B{5};
age = B{6};

%create memory
PSmith = name*Pat+sex*male+age*a66;

%%bigger example
%%jane
Jane = B{13};
female = B{15};
a25 = B{17};
hername = B{19};
hersex = B{21};
herage = B{23};

%create memory
PSmithJDoe =
name*Pat+sex*male+age*a66+hername*Jane+hersex*female+herage*a25;
```

```
%address memory to recover value of name
PatR = inv(name)*PSmith;
PatR = inv(name)*PSmithJDoe;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%qubits and ebits in cartan's matrix representation

%basis blades for qubit 1
a0 = B{1};
a1 = B{3};

%basis blades for qubit 2
b0 = B{5};
b1 = B{7};

%basis blades for qubit 3
c0 = B{9};
c1 = B{11};

%basis blades for qubit 4
d0 = B{13};
d1 = B{15};

%define qubits
A0 = +a0-a1;
A1 = -a0+a1;
Ap = +a0+a1;
An = -a0-a1;

B0 = +b0-b1;
B1 = -b0+b1;
Bp = +b0+b1;
Bn = -b0-b1;

C0 = +c0-c1;
C1 = -c0+c1;
Cp = +c0+c1;
Cn = -c0-c1;

D0 = +d0-d1;
D1 = -d0+d1;
Dp = +d0+d1;
Dn = -d0-d1;

%define Hadamard gates
Sa = a0*a1;
Sb = b0*b1;
Sc = c0*c1;
Sd = d0*d1;

%define the value one
ma = (Sa*Sa);
mb = (Sb*Sb);
mc = (Sc*Sc);
md = (Sd*Sd);

%other example for qutrits
T0 = a0*a1*b0;
T1 = a0*a1*b1;
Tp = a0*b1*b0;
Tn = a1*b1*b0;
```

```
%Pauli gates
Pa = ma + Sa;
Pb = mb + Sb;
Pc = mc + Sc;
Pd = md + Sd;

(A0*B0) %two-qubit register
(A0*B0*Sa*Sb) %separably entangled qubits

((A0*B0*(Sa+Sb))) %inseparably entangled qubits (Bell state)

%all four Bell states
BELL0 = ((A0*B0*Sa*Sb*(Sa+Sb)));
BELL1 = ((A0*B0*Sa*Sb*(Sa+Sb)*(Sa+Sb)));
BELL2 = ((A0*B0*Sa*Sb*(Sa+Sb)*(Sa+Sb)*(Sa+Sb)));
BELL3 = ((A0*B0*Sa*Sb*(Sa+Sb)*(Sa+Sb)*(Sa+Sb)*(Sa+Sb)));
BELL4 = ((A0*B0*Sa*Sb*(Sa-Sb)*(Sa+Sb)*(Sa+Sb)*(Sa+Sb)*(Sa+Sb)));
%same as first state

%all four magic states
MAGIC0 = ((A0*B0*Sa*Sb*(Sa-Sb)));
MAGIC1 = ((A0*B0*Sa*Sb*(Sa-Sb)*(Sa-Sb)));
MAGIC2 = ((A0*B0*Sa*Sb*(Sa-Sb)*(Sa-Sb)*(Sa-Sb)));
MAGIC3 = ((A0*B0*Sa*Sb*(Sa-Sb)*(Sa-Sb)*(Sa-Sb)*(Sa-Sb)));
MAGIC4 = ((A0*B0*Sa*Sb*(Sa-Sb)*(Sa-Sb)*(Sa-Sb)*(Sa-Sb)*(Sa-Sb)));
%same as first state
```