Mathematical modelling of magnetoencephalographic data

VRIJE UNIVERSITEIT

MATHEMATICAL MODELLING OF MAGNETOENCEPHALOGRAPHIC DATA

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. T. Sminia,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op vrijdag 1 juli 2005 om 15.45 uur
in de aula van de universiteit,
de Boelelaan 1105

door

Fetsje Bijma

geboren te Leeuwarderadeel

Voorwoord

Het feit dat, en de hoedanigheid waarin dit proefschrift nu af is, is te danken aan vele mensen om mij heen. Zonder hen had ik er waarschijnlijk al lang de brui aan gegeven of was het resultaat een stuk minder fraai geweest.

Allereerst bedank ik mijn promotor Rob Heethaar voor het mogelijk maken en begeleiden van mijn promotietraject. Alle hulp bij de formele gang van zaken alsmede het verlengen van mijn aanstelling heb ik zeer op prijs gesteld.

En dan Jan de Munck. Ondanks dat je aanvankelijk dacht dat ik Chinees was - de Friese taal zorgt soms voor wat misverstanden - bleek je een erg goede begeleider. En dat is niet alleen door de inhoudelijke discussies die we hadden waarin je je enthousiasme (vakidiotie?) op mij over wist te brengen en je prompte hulp bij problemen en vragen, maar vooral ook door al je opmerkingen over de dingen die er helemaal niet toe deden en die ook onvermijdelijk deel uitmaakten van onze conversaties. Ik vraag me af of al die *waanzin* onze communicatie nu heeft *extended* of juist *constrained*. Ik blijf vasthouden aan het eerste! En wat dat Fries betreft: je bent toch maar mooi aan de Friese rijwielelfstedentocht verslaafd geraakt.

De collega's in het AG project, Hilde, Koen, Lourens en Raoul, wil ik bedanken voor alle opbouwende discussies en feedback in onze verschillende bijeenkomsten. De hulp van Koen bij het analyseren van datasets en het controleren van manuscripten heb ik erg gewaardeerd.

A special word of thanks is to Arye Nehorai, whom I visited in Chicago for a research summer. Thank you for giving me the possibility to visit your lab in such a nice city, which I have enjoyed very much.

Fernando Lopes da Silva wil ik bedanken voor zijn onderwijs in de neurofysiologie. Deze opgedane kennis heeft mij in staat gesteld de eerste bladzijden van mijn proefschrift te schrijven.

Geert, hoe moet ik jou bedanken? Je met grote regelmaat terugkerende hulp bij computerproblemen van allerlei aard is een van de noodzakelijke ingrediënten van dit proefschrift geweest. Van simpele dingen als het installeren van een nieuw programma tot ernstiger zaken als een gecrashte harde schijf; jij was altijd bereid om mij te helpen. Naast al die computeraangelegenheden konden we bovendien onze passie voor muziek delen, waar ik ook erg blij mee ben geweest.

Mijn kamergenoten Jeroen, Ilonka en Geert (alweer!) bedank ik voor de leuke en collegiale sfeer op onze kamer. Mede door jullie ging ik met plezier naar mijn werk. Ilonka, jou bedank ik in het bijzonder voor het detecteren van de lelijke typefout op de eerste pagina van mijn proefschrift. Stel je voor, zeg . . .

Alle andere collega's van de MEG-groep, Dennis, Arent, Hanneke, Sandra L. en Bob,

bedank ik voor de leuke sfeer in de groep. I thank the Portuguese clan, Inês, Sónia, Isabella (officially Polish, though highly contaminated by Portuguese roommates), Sandra T., Teresa and Mafalda for contributing to a very enjoyable atmosphere and good women gossip. I also have appreciated the good contact with the non-Portuguese foreign colleagues, Patrik, Fabrice, Keith and others.

I am grateful to Gordon Haid from VSM MedTech Ltd. for kindly offering financial support for printing this thesis.

En dan alle andere mensen, die niets van MEG af weten en dat graag zo willen houden.

Om te begjinnen Berber. As ik wer ris yn de safolste aio-dip delsakke wie, hiest altyd wol in opkleurenje opmerking. Dêr bin ik ekstraprotfolle bliid mei west. En dêrom stiest aanst neist my as paranimf.
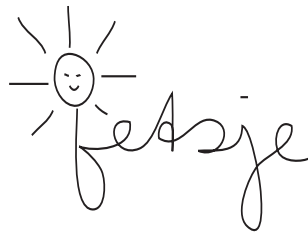
Ook mijn andere paranimf, Waldemar, bedank ik. Samen muziek maken is heerlijk. Het heeft me de afgelopen tijd heel goed gedaan en wat mij betreft, blijven we lekker zo doorgaan.

Ik bin Heit en Mem tankber foar alle stipe dy't ik krigen ha yn myn stúdzje en myn wurk. Ek Piter en de fierdere famylje ha der de ôfrûne jierren oan bydroegen dat ik sa fier kommen bin.

De hulp van Dirk-Jan en Berber bij de Fryske gearfetting en lay-out aangelegenheden heb ik erg gewaardeerd. Dirk-Jan, je bleek een zeer behulpzame helpdesk wat betreft LATEX en taalkundige problemen. Zelfs het Fries lijkt een eitje voor je, terwijl je een jaar geleden van *bûter, brea en griene tsiis* nog geen (groene) kaas gegeten had.

Al mijn vriendjes en vriendinnetjes die mijn aio-perikelen de afgelopen jaren hebben aangehoord ben ik dankbaar voor hun luisterende oren. In het bijzonder wil ik Henk noemen, aan wie ik met name in het begin veel heb gehad (behalve dan met die klapband natuurlijk . . .).

Wie bewaart, die heeft wat. En daarom noem ik Hans als laatste. Aan jou had en heb ik veel steun. Enorm bedankt!

Contents

Chapter 1

Introduction

1.1 Magnetoencephalography and its application

In an active human brain, information is passed on from one neuron to an other by small currents (see section 1.1.2). When a sufficient number of neurons is active synchronically, the interaction of these neurons produces both a small electrical potential distribution on the skin and a small magnetic field outside the head. This small magnetic field can be measured using the sensitive Magnetoencephalography (MEG) technique, while the potential differences are measurable by Electroencephalography (EEG). Whereas EEG was developed already in the 1920s, the first MEG was registered in 1972 [12]. Nowadays, MEG is used on a wider scale in clinical applications and for research purposes.

One of the main applications of MEG is *functional mapping*. Different functional areas in the brain, e.g. the auditory or visual area, can be depicted using MEG by presenting a stimulus to the subject. For example, if one is interested in the visual area of the brain, different pictures are presented to the patient in order to activate the visual area of the brain and the MEG is measured meanwhile. From these measured MEG signals one can determine where this active area of the brain is located. Varying this method over different functionalities of the brain, one can map each functionality to a location in the brain. The resulting map is interesting on itself in terms of functional anatomy; yet, for brain tumour patients this map is of clinical importance. When the location of a brain tumour coincides with the regular location of one of the functionalities, it is important to know which effect the tumour has on that functional area. The tumour may have pushed the functional area away, or may have grown together with the functional area. In the former case, the neurosurgeon can (hopefully) resect the tumour, whereas in the latter case resection would yield functional deficits.

A second major, clinical, application of MEG is the investigation of epilepsy. Epilepsy patients are routinely treated with medical drugs. However, for some patients the drugs do not suppress the epileptic seizures (attacks) sufficiently. For these drug resistant epilepsy patients, neurosurgery can be an option to become seizure free. The possibility of such a surgery depends on the location and extent of the epileptic brain tissue. In case of focal epilepsy, the epileptic seizures consist of so-called *epileptic spikes*. A focal epileptic spike is a short burst of activity at a localised place in the brain. The

MEG technique can be used to record many of these localised spikes and determine the location of their underlying source. This information is important to the neurosurgeon in order to decide which part of the brain should be resected in order to make the patient seizure free.

The third utilisation of MEG is research on the MEG technique it self. Because MEG is a rather young technique, both the apparatus and the signal processing methods engaged are being improved gradually. These improvements include increasing the number of MEG sensors, improving the accuracy of functional source localisation and improving the noise attenuation.

1.1.1 Instrumentation

Neuromagnetic signals are very small compared to other magnetic fields surrounding us. For example, the earth's magnetic field is larger by a factor of 10^8 to 10^9. The magnetic field measured by MEG typically ranges from 50 fT to 500 fT (1 fT equals 10^{-15} Tesla). These weak signals can be measured using the superconductive quantum interference device, the SQUID. This sensitive detector of magnetic flux is the basis of the MEG technique. MEG sensors are embedded in liquid helium at a temperature of 4 K to keep them superconductive. Furthermore, the MEG scanner (Figure 1.1(a)) is positioned in a magnetically shielded room to keep external disturbances away. Sources of significant external disturbance include moving vehicles, power line fields, monitors and cell phone networks.
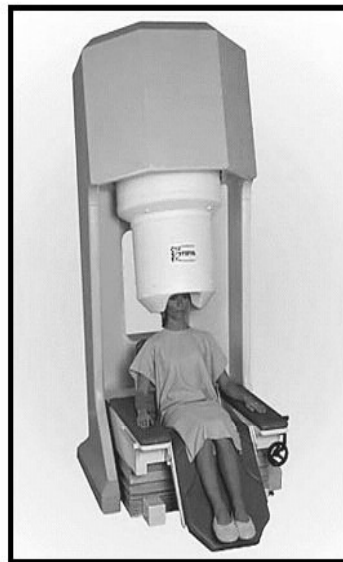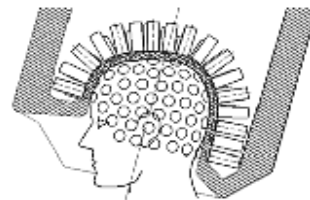
(a) MEG scanner   (b) MEG helmet

Figure 1.1: *The MEG scanner (a) and the MEG helmet of the whole head system (b).*

Using a whole head system, the magnetic field is measured on several sensors covering the whole brain. These sensors are inside the MEG helmet, that is the lower

part of the dewar, the white vessel, in Figure 1.1(a). A cross section of this helmet is drawn in Figure 1.1(b). From this figure, one can see that MEG is a fully noninvasive technique; the patient only has to put the head in the helmet. The patient is not exposed to radioactive injections, strong magnetic fields or x-rays, as is the case in Positron Emission Tomography (PET), Single-Photon Emission Computerised Tomography (SPECT), Magnetic Resonance Imaging (MRI) and Computer Tomography (CT).

1.1.2   Physiological background

Electrophysiological activity in the brain consists of information that is processed by neurons (nerve cells) [4, 37, 60, 86]. Nerve cells pass on their messages through synapses along their dendrites, see Figure 1.2 and 1.3. A synapse is a place where two neurons touch each other and messages are passed on from the *presynaptic cell* to the *postsynaptic cell*. In the resting state, neurotransmitter is present in small vesicles in the axon terminal of the presynaptic cell, see Figure 1.3. The outer shell of a cell is called the cell membrane, and is surrounded by ions on both sides in different concentrations. In Figure 1.3 one of the dendrites of the postsynaptic cell is drawn together with its membrane, which is surrounded by ions, both inside and outside the membrane. These ions cause a potential difference between the inner side (more negative) and the outer side (more positive) of the dendrite, the so-called membrane potential. This is a balancing state between the inner and outer concentrations of the different ions (e.g. Na^+ and K^+). The membrane is selectively permeable to these ions, that is, ions can travel through the membrane. Messages in the presynaptic cell travel by rapid action potentials towards the axon terminal. An action potential is a fast moving disturbance of the balance of the potential membrane. Upon arrival at the axon terminal, the signal is conveyed through neurotransmitter into the synapse cleft. The molecules of this substance bind to receptor molecules at the outer side of the dendrite of the postsynaptic cell (Figure 1.3). This, in turn, causes ion channels to open. The type of ion channel that is opened, depends on the neurotransmitter and the receptor. In the case of opened Na^+-ion channels, positively charged Na^+-ions flow inwards, while in the case of opened K^+-ion channels, positively charged K^+-ions flow outwards. This ion flow disturbs the balance of the membrane potential at the place of the synapse: a postsynaptic potential (PSP). The neighbouring ions in the dendrite of the postsynaptic cell will react by a positive current away from the synapse in case of opened Na^+-ion channels or a positive current towards the synapse in case of opened K^+-ion channels. Thus, a current, or ion flow, starts travelling through the dendrite (that is, one of the legs of the postsynaptic cell, see Figure 1.2). This current flow is called the primary current. Because of the permeability of the membrane, in the case of Na^+-ions, part of these travelling positive ions will flow outwards somewhere along the dendrite, while the others will flow until the end, the axon hillock (Figure 1.2). A PSP that causes a positive current flow away from the synapse is called an Excitatory PSP (EPSP). In the case of K^+-ions, the current direction is opposite and positive ions in the dendrite are pulled towards the synapse. Some of these ions will come from outside the dendrite and will be pulled inwards while the others will come from inside. This kind of PSP is called an Inhibitory PSP (IPSP). The PSP mechanism has two effects: firstly, the
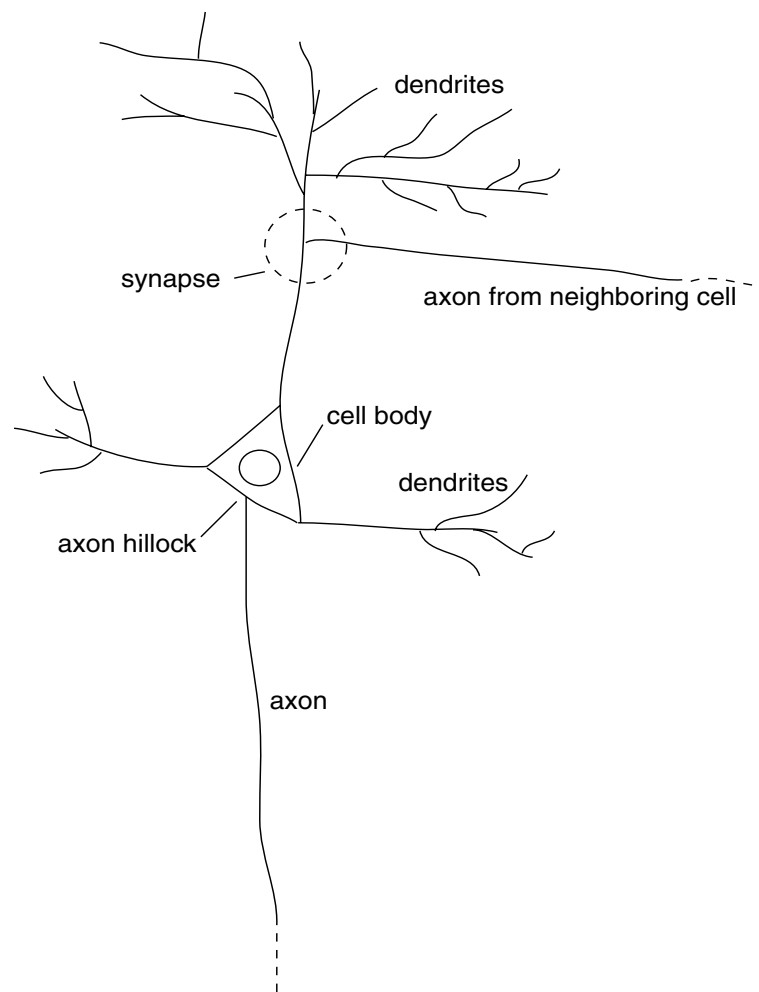
Figure 1.2: *Illustration of a nerve cell with a synapse along one of the dendrites. A magnified view of the synapse is presented in Figure 1.3.*

potential at the axon hillock is changed (increased in case of EPSP and decreased in case of IPSP) and secondly, an extracellular *return* or *secondary* current is generated. When the potential at the axon hillock has reached a certain threshold by many (mainly EPSP) synapses as a result of the first effect, a new action potential is generated at the axon hillock and the postsynaptic cell will pass on the message to other neurons. The second effect is caused by those ions that flow through the membrane and cause an extracellular current. The effects of this current can be measured at a certain distance. EEG measures the potential fluctuations due to these return currents, whereas both the primary and secondary currents contribute to the MEG signal. In Figure 1.4 the primary and secondary currents are illustrated.

The primary and secondary currents due to one excited neuron are too small to be measured by MEG/EEG. For a measurable signal typically thousands of parallel arranged *pyramidal cells* are excited simultaneously, and the superposition of their neural currents produces a measurable magnetic field and electrical potential outside
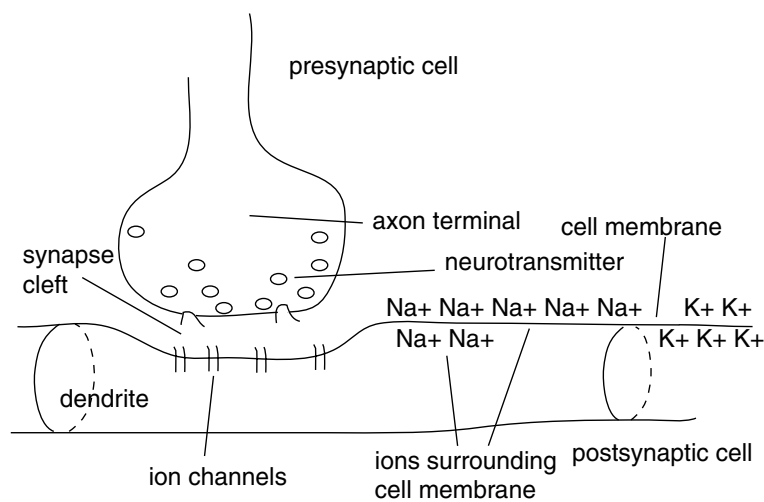
Figure 1.3: *Synapse*
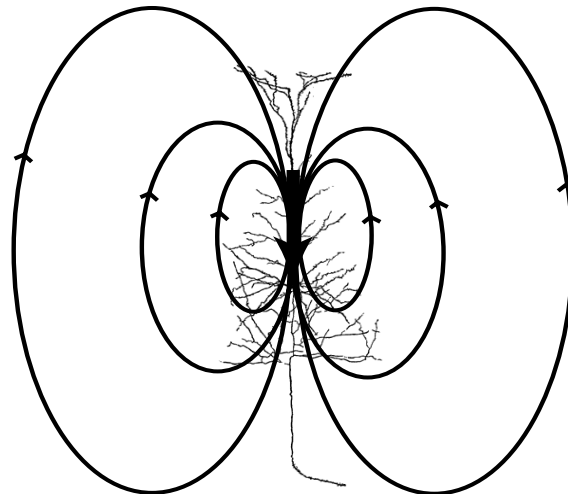
Figure 1.4: *Illustration of the currents around an excited nerve cell. The thick arrow pointing downwards in the middle indicates the primary current along the cell's dendrite. The thinner loops indicate the return currents through the volume conductor. This source configuration is equal to the source configuration of a current dipole.*

the head. The cortex, the grey matter at the outer side of the cerebrum (see Figure 1.5), contains many of these pyramidal cells that are arranged perpendicular to the surface of the cortex. The white matter underneath the grey matter mainly consists of long neuron fibers, the connecting 'highways' between different parts of the cortex. The signals that are measured by the MEG/EEG techniques are essentially generated in the grey matter, i.e. along the outer surface of the brain. The action potentials are in general too small and unsynchronised to be caught by the MEG/EEG technique.

The *inverse problem* of MEG/EEG involves the determination of the underlying sources of measured MEG/EEG signals. In order to determine the origin of the signals,
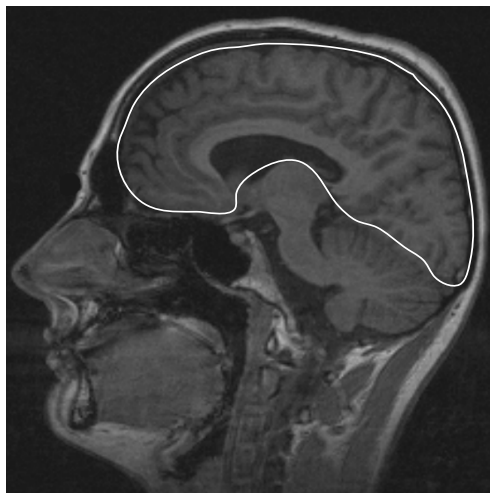
Figure 1.5: *Sagittal MRI image of the human brain. The area within the white line is called the cerebrum. The difference between grey and white matter in the cerebrum is clearly visible: the darker edges of the brain indicate grey matter, while the lighter inner part is white matter.*

models for electromagnetic brain sources are needed. The underlying electrophysiology of neural brain activity, as presented above, is important to understand when designing models for the sources of brain activity. Usually, these sources are modelled as point sources, as an equivalent of thousands of nearby sources. The *forward problem* deals with the computation of the magnetic field and the electric potential due to such a point source, throughout and outside the head. In order to solve the inverse problem, the source parameters are tuned such that the difference between the measured signals and the predicted forward model becomes minimum in some optimal sense. The next two sections deal with the forward problem and the inverse problem, respectively. An overview of different aspects of these two topics can be found in [3].

1.2   Forward Problem

1.2.1   General formulation

The electric and magnetic fields of a bioelectric source in conducting tissue are computed using the quasi-static approximation of the Maxwell's equations [78, 92]. Denoting the primary source current density by $\mathbf{J}^p$, and assuming that the conductivity of the tissue is σ and the magnetic susceptibility is μ_0 throughout the tissue, these equations state

$$\mathbf{E} = -\nabla V \tag{1.1}$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \tag{1.2}$$
$$\nabla \cdot \mathbf{B} = 0 \tag{1.3}$$
$$\mathbf{J} = \mathbf{J}^p + \sigma \mathbf{E} \tag{1.4}$$

where $\mathbf{E} = \mathbf{E}(\mathbf{r}) \in \mathbb{R}^3$ denotes the electric field at location $\mathbf{r} \in \mathbb{R}^3$ in the conductor, $\mathbf{B}(\mathbf{r}) \in \mathbb{R}^3$ the magnetic induction and $V(\mathbf{r}) \in \mathbb{R}$ the electric potential. $\sigma\mathbf{E}(\mathbf{r}) \in \mathbb{R}^3$ denotes the volume (secondary) current and $\mathbf{J}(\mathbf{r}) \in \mathbb{R}^3$ the total current density. The magnetic field $\mathbf{B}(\mathbf{r})$ for any $\mathbf{r}$ now follows from the Biot-Savart law:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_G \mathbf{J}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} dV \tag{1.5}$$

where the integral is taken over the conductor G and $\|.\|$ denotes Euclidean vector length. Combining equations (1.1) and (1.4) the total current density $\mathbf{J}(\mathbf{r})$ in the integrand of equation (1.5) can be written as

$$\mathbf{J}(\mathbf{r}) = \mathbf{J}^\mathbf{P}(\mathbf{r}) + \mathbf{J}^\mathbf{s}(\mathbf{r}) = \mathbf{J}^\mathbf{P}(\mathbf{r}) - \sigma\nabla V(\mathbf{r}), \tag{1.6}$$

where $\mathbf{J}^s$ denotes the secondary current.

The head comprises different tissues, each having its own characteristic conductivity. Assume that the conductivity in the volume conductor G is piecewise constant and define N subvolumes G_n, $n = 1, \ldots, N$, such that $\sigma = \sigma_n$ within G_n. Furthermore, the bounding surface of G_n is denoted by S_n. Using equation (1.6) and Stokes' Theorem with the boundary condition that no current flows inwards or outwards, the integral expression in equation (1.5), becomes Geselowitz' formula [30]:

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{n=1}^{N} (\sigma_n^- - \sigma_n^+) \int_{S_n} V(\mathbf{r}')\mathbf{n}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} dS \tag{1.7}$$

where

$$\mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_G \mathbf{J}^\mathbf{P}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} dV, \tag{1.8}$$

σ_n^- and σ_n^+ are the conductivities at the inner and outer sides of S_n and $\mathbf{n}(\mathbf{r})$ denotes the outward normal. $\mathbf{B}_0(\mathbf{r})$ is the magnetic field due to $\mathbf{J}^\mathbf{P}$ in an unbounded homogeneous space. In order to calculate the expression in equation (1.7), one needs know $V(\mathbf{r})$ on the surfaces S_n. By a similar calculation the electric potential $V(\mathbf{r})$ can be expressed by the following integral equation

$$\frac{\sigma_n^- + \sigma_n^+}{2} V(\mathbf{r}) = \sigma_N V_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{n=1}^{N} (\sigma_n^- - \sigma_n^+) \int_{S_n} V(\mathbf{r}')\mathbf{n}(\mathbf{r}') \cdot \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} dS \tag{1.9}$$

for $\mathbf{r}$ on one of the surfaces S_n and

$$V_0(\mathbf{r}) = \frac{1}{4\pi\sigma_N} \int_G \mathbf{J}^\mathbf{P}(\mathbf{r}') \cdot \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} dV. \tag{1.10}$$

After solving this integral equation for $V(\mathbf{r})$, the magnetic field can be calculated from equation (1.7). However, solving equation (1.9) is complicated in general for two reasons. First of all, for a general source configuration and an arbitrary head shape there is no analytical solution to this equation. And secondly, the conductivities are poorly known. The first problem is eliminated by assuming the head shape to be spherical and the source to be a dipole source. This special case is addressed in the next section.

In general, one uses the Boundary Element Method to solve equation (1.9) [17]. As subvolumes G_n, three regions are usually distinguished in the head: the brain, the skull and the scalp. The realistic shapes of the surfaces S_n are obtained using the anatomical information from other data modalities as MRI and CT [36, 105]. This procedure is rather time consuming compared to finding the best fitting spherical head model. The conductivity values of the different layers are more complicated to estimate noninvasively. These values can be estimated either *in vivo* or *in vitro*. In vivo measurement is measurement by leaving the tissue in its position in the living human body. In vitro measurement is measurement by taking the tissue outside the living human body. Whereas in vivo measurement is best in terms of keeping the experimental settings as natural as possible, the in vitro measurement usually allows for a better measurement setup. Different conductivity estimation techniques have been performed, and have, unfortunately, resulted in greatly varying values, e.g. [32, 35, 80]. In this connection it should be mentioned that the EEG forward problem is much more sensitive to errors in estimated conductivity values and head shape than the MEG forward problem. An illustration of this remark is found in the next section.

1.2.2 Spherical conductor and current dipole source

When the head shape is assumed to be spherical, the magnetic field $\mathbf{B}(\mathbf{r})$ outside the head is independent of the conductivities [92]. For the radial component of $\mathbf{B}(\mathbf{r})$ this can be seen easily from equation (1.7). The radial component is $\mathbf{B}(\mathbf{r}) \cdot \mathbf{e}_r$, where $\mathbf{e}_r$ is the unit radial vector. Using equation (1.7) we find for $\mathbf{r}$ outside the head

$$\mathbf{B}(\mathbf{r}) \cdot \mathbf{e}_r = \mathbf{B}_0(\mathbf{r}) \cdot \mathbf{e}_r - \frac{\mu_0}{4\pi} \sum_{n=1}^{N} (\sigma_n^- - \sigma_n^+) \int_{S_n} V(\mathbf{r}')[\mathbf{n}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3}] \cdot \mathbf{e}_r dS = \mathbf{B}_0(\mathbf{r}) \cdot \mathbf{e}_r,$$

$$(1.11)$$

because $\mathbf{n}(\mathbf{r}')$ is aligned with $\mathbf{r}'$ and $\mathbf{e}_r$ with $\mathbf{r}$. Hence the triple product vanishes. Because $\mathbf{B}_0(\mathbf{r})$ is independent of the conductivities σ_n, the radial component of $\mathbf{B}(\mathbf{r})$ is independent of the σ_n. For the remaining components of $\mathbf{B}(\mathbf{r})$ this also holds true [92].

When the source is assumed to be a current dipole point source [15], there exists an analytical solution for $\mathbf{B}(\mathbf{r})$. Modelling the source of brain activity as a dipole source is based on the idea that at a certain distance, an active patch of cortex can be regarded as an active point source. In other words, the superposition of primary currents of neighbouring pyramidal neurons can at sensor level be regarded as dipolar activity. A dipolar point source is determined by a location $\mathbf{r_0} \in \mathbb{R}^3$ and an orientation (or moment) $\mathbf{q} \in \mathbb{R}^3$. The electric potential and magnetic field of the dipole in a homogeneous medium are

$$V_0(\mathbf{r}) = \frac{1}{4\pi\sigma}\mathbf{q} \cdot \frac{\mathbf{r} - \mathbf{r_0}}{\|\mathbf{r} - \mathbf{r_0}\|^3} \tag{1.12}$$

$$\mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi}\mathbf{q} \times \frac{\mathbf{r} - \mathbf{r_0}}{\|\mathbf{r} - \mathbf{r_0}\|^3}. \tag{1.13}$$

Because the current $\mathbf{J}(\mathbf{r})$ is zero outside the conductor, $\mathbf{B}(\mathbf{r})$ is both divergence and curl free there. Therefore, $\mathbf{B}(\mathbf{r})$ can be expressed as the gradient of a magnetic potential

$U(\mathbf{r})$:

$$\mathbf{B}(\mathbf{r}) = -\mu_0 \nabla U(\mathbf{r}). \tag{1.14}$$

Using equations (1.11) and (1.13) this magnetic potential can be solved analytically, yielding for $\mathbf{B}(\mathbf{r})$ outside the spherical conductor [92]

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi F^2}(F(\mathbf{q} \times \mathbf{r_0}) - (\mathbf{q} \times \mathbf{r_0} \cdot \mathbf{r})\nabla F) \tag{1.15}$$

with F only dependent on $\mathbf{r}$ and $\mathbf{r}_0$. This expression for $\mathbf{B}(\mathbf{r})$ does not depend on the conductivities, and hence, by superposition, for a general source configuration $\mathbf{B}(\mathbf{r})$ does not depend on the conductivities σ_n, as was shown in equation (1.11) for the radial component only. Moreover, from equation (1.15), it follows that $\mathbf{B}(\mathbf{r})$ vanishes outside the conductor if $\mathbf{q}$ and $\mathbf{r}_0$ are aligned. Consequently, $\mathbf{B}(\mathbf{r})$ is zero for radial sources in a spherical conductor, i.e. the field of a radial dipole, or equivalently, the radial component of any dipole, cannot by measured by MEG.

In sum, the magnetic field outside a spherical conductor does not depend on the conductivities and a radial dipole source does not produce a magnetic field outside a spherical conductor. These two findings are fundamental for MEG.

1.2.3   Variations over time

The forward field on each MEG/EEG sensor generated by a dipole at location $\mathbf{r}_0$ can algebraically be expressed by a vector $\mathbf{f} \in \mathbb{R}^{I \times 1}$. I denotes the number of MEG/EEG sensors. Because the electrical potential, and therefore also the magnetic field, is linear in $\mathbf{q}$, the field on each MEG sensor is a linear combination of the field in each of the three orthogonal directions of the moment vector. So is the electric potential on each EEG sensor. For the ith MEG sensor this can be written as

$$\mathbf{f}_i = q_x \Phi_{i,1} + q_y \Phi_{i,2} + q_z \Phi_{i,3} \tag{1.16}$$

where $\Phi \in \mathbb{R}^{I \times 3}$ is the *unit dipole field matrix* or, *unit lead field matrix* dependent of $\mathbf{r}_0$. The forward model for all sensors can be combined in one matrix expression:

$$\mathbf{f} = \Phi \mathbf{q}. \tag{1.17}$$

This is the forward model for one time instant.

When considering a time window, consisting of J time instants (samples), the forward model becomes a matrix $\mathbf{F} \in \mathbb{R}^{I \times J}$. This *spatiotemporal* forward model [93] depends on how the dipole parameters, i.e. location and moment, change over time. Three possibilities can be distinguished: the location and orientation of a *stationary dipole* are fixed over time, though the amplitude can change; the orientation of a *rotating dipole* changes over time, while its location is fixed; a *moving dipole* has both a varying location and a varying orientation. The forward field of a stationary dipole is

$$\mathbf{F} = \Phi \mathbf{q} S, \tag{1.18}$$

with $S \in \mathbb{R}^{1 \times J}$ the row vector containing the amplitude of the dipole at each time instant. The forward field of a rotating dipole is

$$\mathbf{F} = \Phi S, \tag{1.19}$$

with $S \in \mathbb{R}^{3 \times J}$ the matrix containing the amplitude in the three orthogonal directions at all time instants. The forward field of a moving dipole cannot be parameterised any further than in equation (1.17). In that case, the vector $\mathbf{f}$ changes from each time instant to time instant.

In the case of p active dipoles, the forward field becomes the superposition of the fields of each dipole separately. For stationary dipoles this is

$$\mathbf{F} = \Phi \Omega S, \tag{1.20}$$

with $\Phi \in \mathbb{R}^{I \times 3p}$, $\Omega \in \mathbb{R}^{3p \times p}$ and $S \in \mathbb{R}^{p \times J}$. The matrix Φ now contains in each set of 3 columns the unit forward dipole field of one dipole. The matrix Ω is a block diagonal matrix with p 3×1 blocks along its diagonal, each containing a normalised moment vector $\mathbf{q} = (q_x, q_y, q_z)^t$:

$$\Omega = \begin{pmatrix} \begin{bmatrix} q_x^1 \\ q_y^1 \\ q_z^1 \end{bmatrix} & \mathbf{0} & \cdots & \\ \mathbf{0} & \begin{bmatrix} q_x^2 \\ q_y^2 \\ q_z^2 \end{bmatrix} & \mathbf{0} & \vdots \\ \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ & \cdots & \mathbf{0} & \begin{bmatrix} q_x^p \\ q_y^p \\ q_z^p \end{bmatrix} \end{pmatrix} \tag{1.21}$$

The matrix S contains the amplitudes of all p dipole sources. For multiple rotating sources, the spatiotemporal model becomes

$$\mathbf{F} = \Phi S, \tag{1.22}$$

with $\Phi \in \mathbb{R}^{I \times 3p}$ and $S \in \mathbb{R}^{3p \times J}$. Φ is the same as in the case of multiple stationary dipoles, while S now contains the amplitudes in all three orthogonal directions for all the dipoles.

In the case of stationary dipoles, the expression preceding S, that is, $\Phi \Omega$ in equation (1.20) is called the *dipole field matrix*, or *lead field matrix* or *forward field matrix* and is denoted by Ψ. In the literature, the unit lead matrix is often abbreviated as lead field matrix, though in this thesis the terms are kept separately for stationary dipoles. For rotating dipoles, the two terms, lead field and unit lead field matrix, automatically indicate the same matrix, matrix Φ in equation (1.22). To keep the notation consistent, the forward field matrix is in this thesis denoted by Ψ, also for rotating dipoles. The forward field matrix Ψ is dependent on all source locations and orientations, i.e. on all *spatial* parameters. The matrix S is called the *source time function matrix*, because the amplitude row vectors can be regarded as discrete source time functions. In this way, all *temporal* parameters are contained in S. In all, this yields the general spatiotemporal forward model

$$\mathbf{F} = \Psi S. \tag{1.23}$$

1.3   Inverse Problem

The inverse problem deals with finding the dipole sources that have generated the measured EEG/MEG signals. When solving the inverse problem, the difference between the predicted forward field based on source parameters and the measured field is minimised. In general, there is no unique solution to this problem, due to possible magnetically silent sources. As mentioned above, a radial source in a spherical conductor produces no magnetic field outside the conductor. Moreover, because of the finite number of MEG/EEG sensors, the number of sources that can be localised is limited. Hence, restricting constraints are necessary to find a solution. A second problem with the inverse problem lies in the fact that measured signals are always corrupted with *noise*. This noise consists of both internal and external noise. The former indicates brain activity that is not of particular interest in the experiment, the so-called background activity. The latter denotes noise that is caused by external sources, e.g. moving vehicles, moving iron objects (buttons, buckles, necklaces etc.), as mentioned in section 1.1.1. The magnetically shielded room (MSR) wards off most of the external noise; however, the brain noise is not barred by the MSR. Therefore, one needs to separate two parts in the measured signal: the signal of interest, and the remainder. The next section is about the separation between signal and noise.

Two different ways of source localisation exist: *dipole localisation* and *spatial scanning*. The former method fits a number of dipoles to the measured data, while the latter technique scans the brain for activity, that is, computes the activity at all locations. Section 1.3.2 deals with dipole localisation and the spatial scanning technique is presented in 1.3.3.

1.3.1   Signal Plus Noise model

To localise certain areas in the brain, the area of interest is usually activated by the presentation of an external stimulus to the subject. This way, a brain response is evoked in the area that is to be localised. Examples of this kind of experiments are Visual Evoked Field (VEF), Auditory Evoked Field (AEF) and Somatosensory Evoked Field (SEF) experiments. Each time the stimulus is presented, the brain responds to the stimulus. An idealised example of such an *evoked response* is given in Figure 1.6. In practise, however, the measured MEG signal consists of both the evoked response
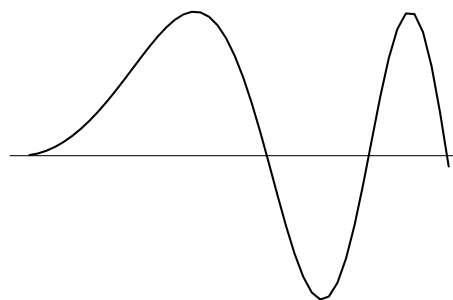
Figure 1.6: *An example of an evoked response by an external stimulus*

and additional noise. Therefore, the measured signal is blurred, as illustrated in Figure 1.7.

(a)

(b)

(c)

(d)
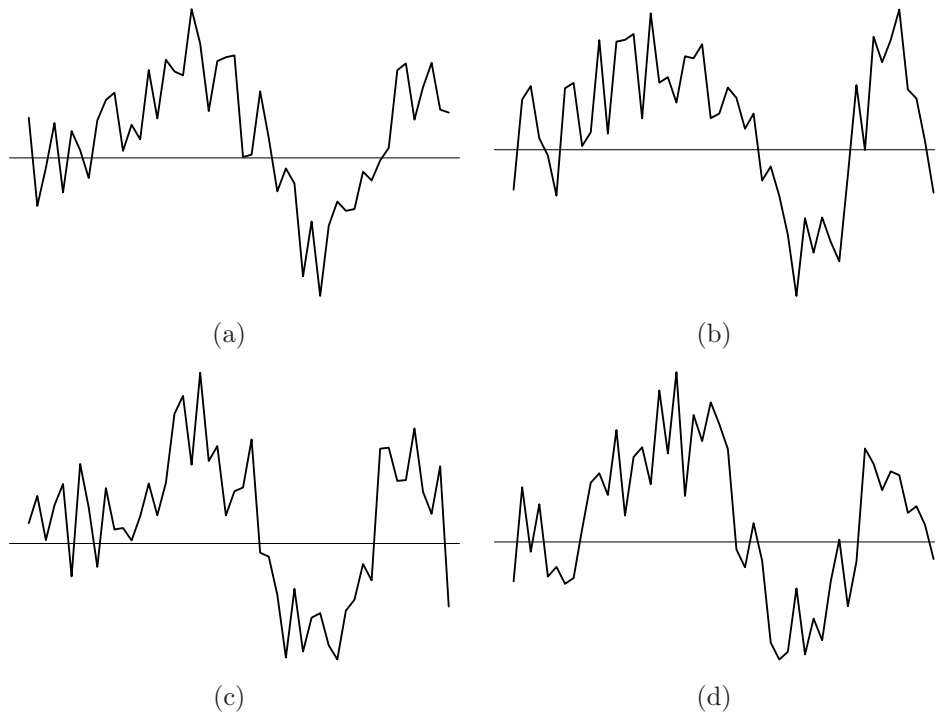
Figure 1.7: *Different measurements of the evoked response and additional noise*

The standard way of 'cleaning' the signal is to present the stimulus several times and to average the measured signals following each stimulus presentation. The idea behind this averaging is that the measured signals consist of a fixed evoked response which is the same each time the stimulus is presented, and a varying noise part. Synonyms for the noise part are *remainder*, *residual* and *error*. This is called the Signal Plus Noise (SPN) model. The SPN model can be formulated as

$$r_{ij}^k = r_{ij} + \varepsilon_{ij}^k \tag{1.24}$$

where r_{ij}^k is the measured signal on sensor i at time sample j in trial (stimulus presentation) k, r_{ij} is the evoked response which is fixed over trials, and ε_{ij}^k is the varying additional noise. The name *residual* is used somewhat ambiguously. When dealing with the model, as in equation (1.24), the residual indicates the difference between measurement and model, e.g. the noise. In the case of real data, the residual of a certain trial k often indicates the difference between the measurement in trial k and the average measurement. Throughout this thesis, *residual* will indicate the latter, i.e. the difference between measurement and average, whereas *noise* will be used for the difference between measurement and model, i.e. the ε-component in equation (1.24). Regarding the statistical properties of the noise and residuals, these are usually assumed to be the same (see section 1.4.3).

In the noise component, both the internal and external noise are assembled. The internal noise mainly consists of ongoing background activity. Background activity is activity of the brain that is not specifically related to a task, and also appears in the resting state. When a subject is at rest, the brain produces at different parts of the cortex spontaneous activity. The most prominent example of this background activity is oscillations [13, 59] (see Figure 1.8). Depending on the aim of the MEG study, these oscillations are or are not of interest. In studies on spontaneous brain activity, it is exactly these oscillations that form the interesting brain signal, while in evoked response experiments this kind of brain activity is considered as unwanted noise or even as 'error'.

Figure 1.8: *An example of oscillating background activity in spontaneous activity with varying amplitude over a time window of 5.89 seconds.*

The averaging technique is rather standard in evoked field MEG data processing, although the basic assumption is debatable. When neurons in the brain are excited by an external stimulus they will *fire*, that is, generate action potentials. However, these action potentials are only propagated when the potential at the axon hillock exceeds a certain threshold, as was explained in Section 1.1.2. The precise reaction to the external stimulus, therefore, depends on the situation at cell level at the moment of stimulus presentation. This, on turn, is related to the ongoing background activity, which constantly varies. Therefore, one may assume that the exact response, on cell level, is different from trial to trial. However, on macroscopic level, like MEG and EEG, these differences may not be visible and the SPN model may well explain the recorded signals. In cognitive experiments, however, where a reaction is asked from the subject, e.g. by pressing a button, the reaction time to the stimulus will depend on the level of attention of the subject and will certainly vary to some extent over trials. One of the goals of this thesis is to validate the SPN model in different ways for different kinds of MEG data, see chapters 2, 3, and 4.

Within the SPN paradigm, the best estimator for the evoked response r_{ij} is the average over trials of the measured signals r_{ij}^k. The more stimuli presented to the subject, the more responses that can be averaged and the more reliable the estimator for the evoked response becomes. In Figure 1.9 this is illustrated; in Figure 1.9(a) the average of the four measured signals of Figure 1.7 is plotted and Figure 1.9(b) shows the average of 100 such signals. Clearly, by taking more measured signals into account in the averaging, the signal becomes smoother and better approaches the true evoked response. In practise, one usually averages between 100 and 600 measured responses to a stimulus.

(a) Average of 4 measurements   (b) Average of 100 measurements
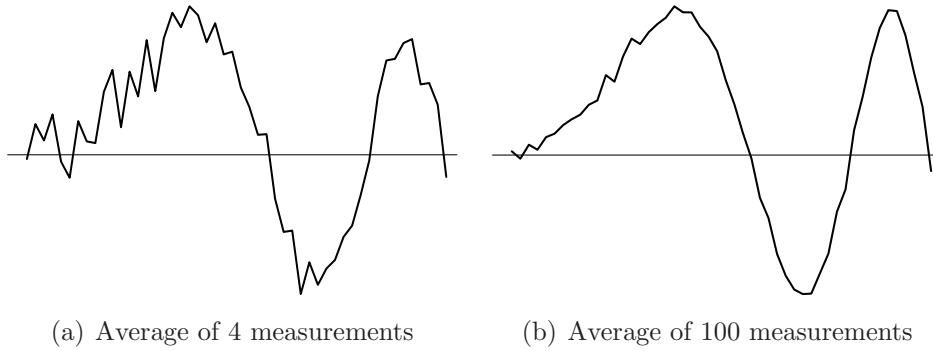
Figure 1.9: *Averages of 4 (figure a) and 100 (figure b) measured signals as in Figure 1.7. The average in figure b is smoother than the average in figure a.*

1.3.2   Dipole localisation

The measured signals of trial k of the SPN model in equation (1.24) are gathered in a matrix $R^k \in \mathbb{R}^{I \times J}$

$$(R^k)_{ij} = r_{ij}^k \tag{1.25}$$

and likewise the noise of trial k in the matrix $E^k \in \mathbb{R}^{I \times J}$

$$(E^k)_{ij} = \varepsilon_{ij}^k. \tag{1.26}$$

Then the model for dipole localisation is

$$R^k = \Psi S + E^k. \tag{1.27}$$

Now the forward field, ΨS, based on a certain number of dipole sources is fitted to the average measured signal. The matrix containing the average signal is

$$\bar{R} := \frac{1}{K} \sum_{k=1}^{K} R^k. \tag{1.28}$$

Fitting the model in equation (1.27) is performed by minimising the sum of the squared differences between the forward field matrix and the averaged measured data matrix, $\bar{R}$, from equation (1.28). This is called the least squares (LS) paradigm. The cost function for this minimisation is

$$\min_{\xi, \eta, S} \|\bar{R} - \Psi(\xi, \eta) S\|_F^2. \tag{1.29}$$

Here ξ is the vector containing all dipole location parameters and η the vector containing all orientation parameters. The norm $\|.\|_F$ is the Frobenius norm [31]. For any matrix $A \in \mathbb{R}^{n \times m}$, $A = (a_{i,j})_{i,j}$, this norm is defined as:

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{i,j}|^2}. \tag{1.30}$$

Using the trace function [31] of square matrices $B \in \mathbb{R}^{n \times n}$, $B = \{b_{i,j}\}$,

$$tr(B) = \sum_{i=1}^{n} b_{i,i} \tag{1.31}$$

the squared Frobenius norm can be alternatively expressed as

$$\|A\|_F^2 = tr(A^t A) = tr(AA^t). \tag{1.32}$$

The standard approach for finding the values for ξ, η and S that minimise the norm in (1.29) is differentiation of the cost function. Then the first derivatives with respect to each parameter are equated to zero and the remaining equations are solved for the parameter. Differentiation with respect to matrices and vectors is explained in [63]. The matrix S, that consists of linear parameters only, is solved in the following way:

$$\partial_S \|\bar{R} - \Psi(\xi, \eta)S\|_F^2 = \partial_S tr[(\bar{R} - \Psi S)(\bar{R}^t - S^t \Psi^t)] = 2tr([-\bar{R}^t \Psi + S^t \Psi^t \Psi]dS). \tag{1.33}$$

Equating this derivative to zero implies

$$-\bar{R}^t \Psi + S^t \Psi^t \Psi = 0 \implies \hat{S} = (\Psi^t \Psi)^{-1} \Psi^t \bar{R}. \tag{1.34}$$

When this estimator for S is substituted in equation (1.29) we get the updated cost function

$$\min_{\xi, \eta} \|\bar{R} - \Psi(\Psi^t \Psi)^{-1} \Psi^t \bar{R}\|_F^2 = \min_{\xi, \eta} \|(\mathbf{I}_I - \Psi(\Psi^t \Psi)^{-1} \Psi^t)\bar{R}\|_F^2 = \min_{\xi, \eta} \|P_\Psi^\perp \bar{R}\|_F^2. \tag{1.35}$$

Here $P_\Psi = \Psi(\Psi^t \Psi)^{-1} \Psi^t \in \mathbb{R}^{I \times I}$ is the projection matrix onto the column space of Ψ and $P_\Psi^\perp = \mathbf{I}_I - P_\Psi \in \mathbb{R}^{I \times I}$ is the projection onto the space orthogonal to the column space of Ψ. The parameters in ξ and η are more complicated to solve because the resulting cost function in (1.35) depends in a nonlinear way of these parameters. Therefore, equating the derivative to zero yields equations that cannot be solved straightforwardly. The common alternative way to find the optimal spatial parameters is a nonlinear search algorithm preceded by a global search.

Using the dipole localisation method, one has to choose the number of dipole sources that is to be fitted to the average data. A priori, this number is unknown and has to be estimated from the data. Increasing this number will lower the cost function in equation (1.35). However, increasing the number too far, results in estimated dipoles that describe the noise in the data, rather than the evoked responses. Furthermore, the inverse solution becomes unstable when too many sources are fitted. This shows that estimation of the number of dipole sources is an essential part of the dipole localisation problem. Because there is no standard way of determining this number, this is the main weakness of the dipole localisation method. Different methods have been designed to tackle this problem, e.g. [53, 67]. An other possibility lies in model selection procedures. One can try different numbers of sources and choose the right forward model (i.e. the right number of dipoles) by applying model selection procedures [110, 111].

Extensions of the point dipole source can be found in distributed (or extended) sources, e.g. [50, 114]. Distributed sources can be 1D (line) shaped sources, 2D (surfaces)

sources or 3D (volume) sources. Using a distributed model one can model a piece of cortex more realistically. Another extension is using multi-pole sources in stead of dipolar sources [45, 77].

Due to the nonlinear parameters, the dipole localisation method becomes cumbersome when many dipoles have to be fitted. To circumvent the estimation of these nonlinear parameters, alternative methods have been sought. The rather intuitive spatial filtering methods, presented in the next section, scan the entire volume for activity, accordingly avoiding the estimation of the nonlinear sources location parameters.

1.3.3   Spatial scanning

A different method to map the underlying sources of measured MEG/EEG signals is found in the spatial scanning technique. Multiple Signal Classification (MUSIC) [74, 94] and beamformer approaches like the Linearly Constrained Minimum Variance (LCMV) method [90, 107] are the main scanning methods. The basic idea is to divide the brain into many voxels and to scan the head voxel by voxel. For each voxel the source power is computed based on the measured data. This procedure yields a 3D image of the source power throughout the brain, which can for example be fused with MRI images.

In the MUSIC paradigm, the sources are found in two steps. First, the cost function (1.35) is minimised with respect to the projection $P_{\Psi}^{\perp}$ and in the second step a scan over source locations is performed to find the optimal source locations. In case of independent dipole sources, this split-up approach is an approximation of the ML estimation method presented in section 1.3.2. The optimal projection is found by using the Singular Value Decomposition (SVD) of the average data matrix $\bar{R}$. The SVD of an arbitrary matrix $A \in \mathbb{R}^{n \times m}$ is

$$A = U \Delta V^t \tag{1.36}$$

where $U \in \mathbb{R}^{n \times n}$ is an *orthogonal matrix*, that is $U^t U = U U^t = \mathbf{I}_n$, and $V \in \mathbb{R}^{m \times m}$ is also orthogonal. If $n > m$ the matrix $\Delta \in \mathbb{R}^{n \times m}$ is $\Delta = \begin{pmatrix} \Delta_0 \\ \mathbf{0} \end{pmatrix}$ where $\Delta_0 \in \mathbb{R}^{n \times n}$ is a diagonal matrix and $\mathbf{0}$ is the null matrix of the appropriate dimension. In case $n \leq m$, the matrix $\Delta = \begin{pmatrix} \Delta_0 & \mathbf{0} \end{pmatrix}$, where $\Delta_0 \in \mathbb{R}^{m \times m}$ is a diagonal matrix and $\mathbf{0}$ is the null matrix of the appropriate dimension. The diagonal matrix Δ_0 contains the singular values of A. For the average data matrix this decomposition is

$$\bar{R} = U \Delta V^t. \tag{1.37}$$

In order to determine the number of sources, usually a scree plot of the singular values in Δ is made. This decreasing series of singular values theoretically shows a bend after, say, p values. This number indicates the rank of the signal subspace, while the remaining values correspond to the noise subspace. In case of independent single sources, this rank p equals the number of dipoles. In case of sources with linearly dependent amplitudes, this method does not yield the right number of dipoles in general. After having determined the number of sources, p, the dimension of the forward field matrix Ψ is set to $I \times p$ and the optimal projection $\hat{P}$ of rank $I - p$ that minimises $\|P\bar{R}\|_F^2$ is given by

$$\hat{P} = \begin{pmatrix} U_{p+1} & \dots & U_I \end{pmatrix} \begin{pmatrix} U_{p+1} & \dots & U_I \end{pmatrix}^t \tag{1.38}$$

where U_q is the qth column of the matrix U. The value of the cost function then becomes

$$\|\hat{P}\bar{R}\|_F^2 = tr[\left(\begin{array}{ccc} U_{p+1} & \ldots & U_I \end{array}\right)^t U\Delta\Delta^t U^t \left(\begin{array}{ccc} U_{p+1} & \ldots & U_I \end{array}\right)] = \sum_{s=p+1}^{\min(I,J)} \sigma_s^2. \qquad (1.39)$$

The values σ_s are the singular values of $\bar{R}$, i.e. $\Delta_0 = diag(\sigma_1, \ldots, \sigma_{\min(I,J)})$.

   The second step in the MUSIC algorithm is to find the source (location and orientation) parameters $\zeta = (\zeta_1, \ldots, \zeta_p) \in \mathbb{R}^{1 \times 6p}$ such that the space orthogonal to the column space of the dipole field matrix $\Psi(\zeta) = (\Psi(\zeta_1), \ldots, \Psi(\zeta_p)) = (\Psi_1, \ldots, \Psi_p) \in \mathbb{R}^{I \times p}$ is as close as possible to $\hat{P}$. This minimisation can equivalently be expressed as

$$\min_{\zeta} \|\hat{P}\Psi(\zeta)\|_F^2 = \min_{\zeta} \|\left(\begin{array}{ccc} \hat{P}\Psi_1 & \ldots & \hat{P}\Psi_p \end{array}\right)\|_F^2 = \min_{\zeta} \sum_{n=1}^{p} \|\hat{P}\Psi(\zeta_n)\|_F^2$$

$$= \sum_{n=1}^{p} \min_{\zeta_n} \|\hat{P}\Psi(\xi_n)\|_F^2. \qquad (1.40)$$

In other words, the joint cost function for all p sources is the sum of the cost function for each source separately. This minimisation is subject to the constraint that no two sources ζ_{p_1} and ζ_{p_2} are equal to each other in order to assure a correct rank of the lead field matrix $\Psi(\zeta) = (\Psi(\zeta_1), \ldots, \Psi(\zeta_p))$. Hence, the minimisation can equally well be performed source after source. In order to avoid minima in the cost function that are caused by small entries in the lead field matrices (i.e. sources with small gain) the cost function is normalised:

$$Cost(\zeta) = \frac{\|\hat{P}\Psi(\zeta)\|_F^2}{\|\Psi(\zeta)\|_F^2} = \|\left(\begin{array}{ccc} U_{p+1} & \ldots & U_{\min(I,J)} \end{array}\right)\Psi(\zeta)\|_F^2. \qquad (1.41)$$

For each location ξ in the brain, the optimal orientation η that minimises this cost function is found by solving a generalised eigenvalue problem [74]. The optimal source parameters are found by 3D scanning: for each location ξ in the brain the cost function $Cost(\zeta) = Cost(\xi, \eta)$ is evaluated. The final step in the MUSIC algorithm is to invert these cost values and select the p highest peaks in the $\frac{1}{Cost(\zeta)}$-graph as sources, which is equivalent to finding the p locations with smallest cost values. In practise, it appears that this intuitively created method, using the two-step approach, yields (under some conditions) approximately the same solutions as the dipole localisation approach presented in section 1.3.2.

   In the LCMV beamformer method (or spatial filtering) the measured signal is filtered for each possible source location such that only activity from that specific location passes the filter. If we denote the unit dipole field matrix of a dipole at location ξ by $\Phi(\xi) \in \mathbb{R}^{I \times 3}$, as in equation (1.17), the optimal spatial filter for location ξ_0 can be expressed by a matrix $W(\xi_0) \in \mathbb{R}^{I \times 3}$ such that

$$W(\xi_0)^t \Phi(\xi) = \begin{cases} \mathbf{I}_3 & \text{for} \quad \xi = \xi_0 \\ \mathbf{0} & \text{for} \quad \xi \neq \xi_0 \end{cases} \qquad (1.42)$$

where $\mathbf{I}_n$ indicates the n-dimensional identity matrix. The meaning of equation (1.42) is that the activity from location ξ_0 passes the filter, while the activity of all other

locations is barred. In the absence of noise, the output of the filter applied on the forward field in equation (1.17) is

$$W(\xi_0)^t \mathbf{f}(\xi_0) = W(\xi_0)\Phi(\xi_0)\mathbf{q} = \mathbf{q}. \tag{1.43}$$

The condition in equation (1.42) causes the norm of the filter $W(\xi_0)$ to become rather big if $\Phi(\xi)$ and $\Phi(\xi_0)$ are different though nearly linearly dependent. In such a case $\Phi(\xi_0) = \Phi(\xi) + \delta$ for some small δ and the condition becomes $1 = W(\xi_0)\Phi(\xi_0) = W(\xi_0)\Phi(\xi) + W(\xi_0)\delta = W(\xi_0)\delta$. Because δ is small, $W(\xi_0)$ is forced to be large. A large norm of $W(\xi_0)$, in turn, causes a large gain to noise and other sources at locations different from ξ and ξ_0. Denoting the measured signal at the I sensors by a vector $\mathbf{r} \in \mathbb{R}^{I \times 1}$ and the output of the filter by $\mathbf{s} \in \mathbb{R}^{3 \times 1}$ the filter output is

$$\mathbf{s} = W(\xi_0)^t \mathbf{r}. \tag{1.44}$$

In LCMV spatial filtering, the variance of $\mathbf{s}$ is minimised, which can mathematically be expressed as

$$\min_{W(\xi_0)} \mathcal{C}_\mathbf{s} = \min_{W(\xi_0)} W(\xi_0)^t \mathcal{C}_\mathbf{r} W(\xi_0) \quad \text{subject to} \quad W(\xi_0)\Phi(\xi_0) = \mathbf{I}_3 \tag{1.45}$$

where $\mathcal{C}_h$ denotes the covariance matrix of variable h. The optimal matrix $W(\xi_0)$ is then given by [107]

$$\hat{W}(\xi_0)^t = [\Phi(\xi_0)^t \mathcal{C}_\mathbf{r}^{-1} \Phi(\xi_0)]^{-1} \Phi(\xi_0)^t \mathcal{C}_\mathbf{r}^{-1}. \tag{1.46}$$

This filter is then computed and applied for all possible locations ξ_0 in the brain from which activity may be expected, yielding the estimated moment vector $\mathbf{q}_0$ for each possible location ξ_0. In that way, a map of the source power is obtained, from which the peaks are detected as being the underlying sources.

In case of uncorrelated (white) noise (see next section) and uncorrelated sources these spatial scanning methods perform acceptably. However, whereas dipole localisation is statistically well defined, the scanning methods lack a clear statistical foundation. In particular, this applies for the estimation of the covariance matrix $\mathcal{C}_\mathbf{r}$ in the LCMV method. Furthermore, the variations over time and temporal correlations are not incorporated in the method. Instead, correlations over time are usually assumed to be zero. In case of correlated sources, the performance of these methods degrades and the computed source power map seriously differs from the true source configuration. This problem becomes more severe when the sources are further apart [107].

The dipole localisation method allows for a transparent incorporation of both the spatial and the temporal correlations. Furthermore, all parameters, both noise and model parameters, can be estimated within one statistical paradigm, the maximum likelihood paradigm (see section 1.4.3). Therefore, dipole source localisation based upon the ML method is used in this thesis.

1.4 Correlations in the background activity

1.4.1 Correlation

Correlations between signals indicate dependencies. For example, if two signals are positive correlated (that is, have a positive covariance) they will tend to show the same

pattern. This means that if one is positive, the other is likely to be also positive. Signals can also have a negative correlation, then they tend to show opposite sign. Correlations are scaled between -1 and +1. Perfectly positive correlated signals have correlation 1 and are proportional to each other having the same sign. Perfectly negative correlated signals have correlation -1, and are proportional to each other with opposite sign. A zero correlation, indicates that both signals behave unconnectedly. The mathematical formulas for the correlation and covariance of two stochasts X and Y are

$$Cov(X,Y) = E[(X - EX)(Y - EY)] \tag{1.47}$$

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{E[(X - EX)^2]E[(Y - EY)^2]}} \tag{1.48}$$

where EZ denotes the expected value of stochast Z.

Background activity in MEG measurements is correlated both in space and over time. In space, this means that nearby MEG sensors tend to measure the same brain activity. The further the sensors are away from each other, the smaller their covariance gets. This is because the nearer the channels the more likely they will measure the same activity, while further apart from each other, they are likely to measure different, local activity. The meaning of the correlations over time, i.e. the temporal correlations, is that two measurements that are close in time, will tend to find the same value. The farther the two measurements are lagged the smaller the connection between the two measurements becomes. An important temporal feature in background activity is oscillations. Alpha, beta and mu rhythms are examples of characteristic oscillations in
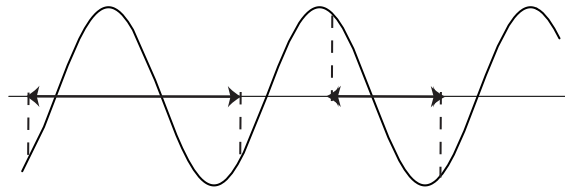
Figure 1.10: *An oscillating signal. Time instants that are one period apart (arrow at the left hand side) show the same signal and therefore have perfect correlation, while instants that are half a period apart (arrow at right hand side) show opposite signs and have a negative correlation.*

spontaneous or resting state activity. The frequency of the alpha and mu rhythms is 8-13 Hz, while frequencies of beta waves range from 14 Hz to 60 Hz [86]. Oscillations cause certain patterns in the temporal correlation. Namely, two time instants that are one period apart, will show the same signal (see Figure 1.10) and therefore their correlation will equal 1, while instants that are half a period apart will show opposite signals and a correlation of -1. Increasing the time lag by a multiple of periods does not alter the correlation in case of perfect oscillations. For nonoscillatory brain activity the temporal covariance decreases with time lag between the two time instants and will vanish for big time lags. The idea behind this assumption is that nonoscillatory brain activity consists of transient bursts at different places in the cortex that are independent

of each other. In Figure 1.11 the covariance shape as function of time lag is presented for oscillatory (Figure 1.11(a)) and nonoscillatory signals (Figure 1.11(b)).

(a) Covariance as function of time lag of purely oscillatory brain activity

(b) Covariance as function of time lag of random nonoscillatory brain activity
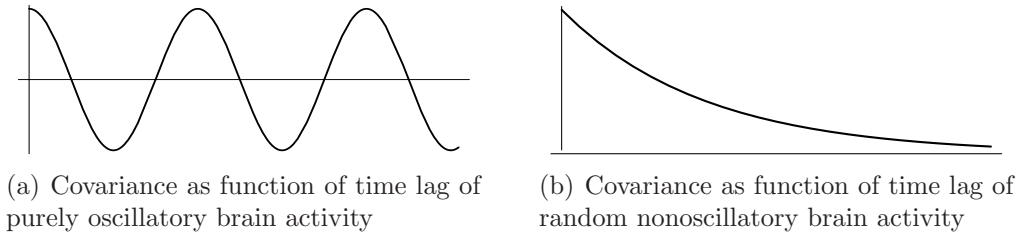
Figure 1.11: *Covariance shapes of different kinds of brain activity*

1.4.2   Temporal nonstationarity

The temporal covariance of MEG signals is called nonstationary if it does not only depend on relative time (time lag between the signals) but also on absolute time (time instant). In fact, this defines *weak stationarity*, though because no other kind of temporal stationarity is discussed in this thesis, this the word *weak* is omitted throughout this thesis. For the special case of a zero time lag, i.e. the variance of the signal,

$$Var(X) = E[(X - EX)^2] \tag{1.49}$$

nonstationarity means that the expected amplitude deviation from the average varies with time. In other words, at certain instants there is more variation around the average than at others. In Figure 1.12 an example of nonstationary signals is presented. The signals in this example are sinusoids with varying amplitude. The variance is also a sinusoid, though with the double frequency. This is a common phenomenon: the variance of nonstationary oscillatory brain activity has the double frequency (Figure 1.12), whereas the covariance has the same frequency as the oscillatory signals (Figures 1.10 and 1.11(a)).

As an example, consider one MEG sensor in an evoked response measurement (e.g. SEF, VEF). The measured signals are denoted by r_j^k, for J time instants, $j = 1, \ldots, J$, and K trials, $k = 1, \ldots, K$. The temporal covariance matrix of these signals is a $J \times J$ matrix with entries $Cov(r_j, r_{j'})$. Based on the measurements the usual, unbiased, estimator for this matrix is the sample covariance matrix

$$\widehat{Cov}(r_j, r_{j'}) = \frac{1}{K-1} \sum_{k=1}^{K} (r_j^k - \bar{r}_j)(r_{j'}^k - \bar{r}_{j'}), \tag{1.50}$$

where $\bar{r}_j$ denotes the average measured signal at time instant j. Following the SPN model, the recorded signal is the sum of a fixed and a random part. The random noise part consists of background activity and some external noise. When this noise is indeed not influenced by the presented stimulus, the temporal covariance will be stationary, i.e. the noise is not varying in average amplitude over time. In case the SPN
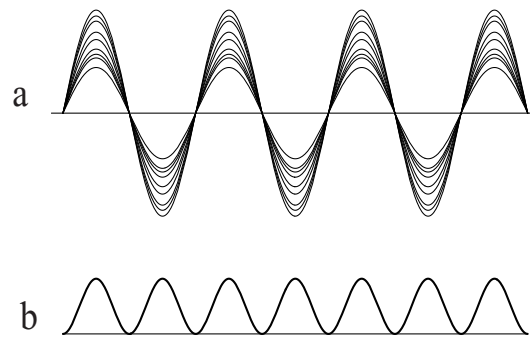
Figure 1.12: *An example of different signals (realisations) of a nonstationary process. The variation of the signals in figure a around their average varies with time. At instants where all signals cross zero there is no variation, while at others the variation is maximum. In figure b the variance is plotted on the same time scale. The time instants where all signals show the same value in figure a correspond to a zero variance in figure b.*

assumption $(r_{ij}^k = r_{ij} + \varepsilon_{ij}^k)$ is incorrect, the estimator in equation (1.50) can produce a nonstationary structure. Two possibilities can be distinguished: the fixed signal part, r_{ij}^k, is not fixed though varies from trial to trial or the noise part, ε_{ij}^k, is influenced by the presented stimulus. Examples of the former are amplitude and latency jitter in the response. An example of the latter is phase locking.

(a) Example of three responses with different amplitudes

(b) Residuals of the three examples in figure a

(c) Variance of the residuals in figure b
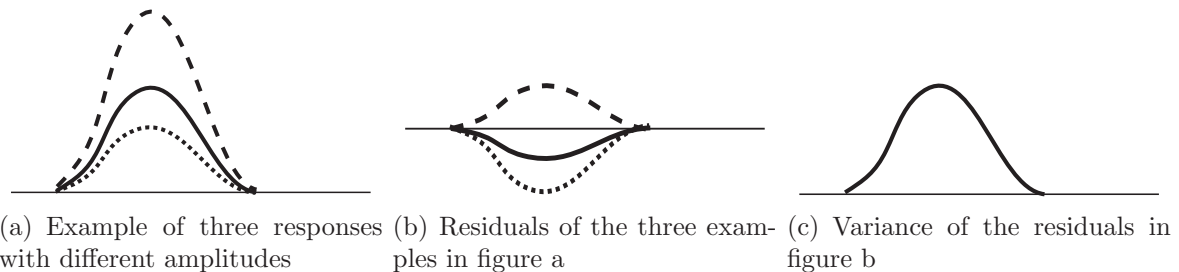
Figure 1.13: *An example of three responses with varying amplitude. The three responses in figure a have different amplitudes, resulting in nonzero residuals in figure b. Figure c shows the estimated temporal variance of these three signals. This variance shows a similar shape as the response wave form, and is, therefore, not constant over time.*

The two different types of response jitter are illustrated in Figures 1.13 and 1.14. In these figures the variance caused by the response part of the signal is plotted. There will be additional (stationary) variance due to the noise part. The second example of invalidity of the SPN assumption is stimulus induced phase locking of oscillatory background activity. This means that the phase of an ongoing wave in the background activity is reset by the stimulus. This also generates a nonstationary temporal covariance. Furthermore, in this case the separation between noise and signal becomes delicate. Namely, when the ongoing activity, normally belonging to the background

(a) Example of three responses with different latencies

(b) Residuals of the three examples in figure a

(c) Variance of the residuals in figure b
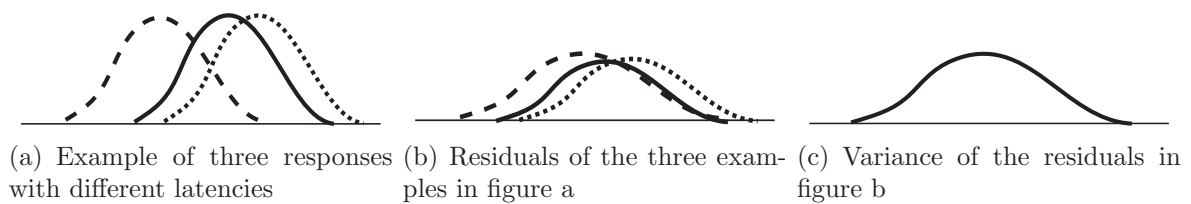
Figure 1.14: *An example of three responses with varying latency. The three responses in figure a have different latencies, resulting in nonzero residuals in figure b, that have a wider shape than the initial wave form. Figure c shows the estimated temporal variance of these three signals. This variance shows a shape that is wider than the response wave form.*

noise, is phase reset, it will build up a nonzero average, and hence it will be regarded as part of the average response when all signals are averaged. Moreover, after averaging these oscillations will not show up in the residuals after stimulus presentation. Hence, regarding the variance of the residuals, there will be a decrease in the variance after stimulus presentation.

In all, the validity of the SPN is linked to the temporal stationarity of the background noise. When the SPN assumption is correct, the temporal structure is stationary, while incorrectness may yield temporally nonstationary estimates.

1.4.3   Correcting for spatiotemporal covariance in dipole localisation

In the cost function of the dipole localisation method presented in equation (1.29) no compensation is made for the spatiotemporal correlations, and therefore this method is called the *ordinary least squares* (OLS) method. This method is rather intuitive; however, it is not optimal. The OLS estimators for the parameters are inefficient. An inefficient parameter is a parameter of which the standard deviation does not approach the Cramèr Rao lower bound when an infinite number of data is available [95]. Consequently, inefficient parameters have higher standard deviations than theoretically necessary. In order to have efficient estimators, one needs to take into account the spatiotemporal correlations. This is called the generalised least squares (GLS) method.

The spatiotemporal covariance of two measured MEG signals, r_{ij}^k and $r_{i'j'}^{k'}$, as in equation (1.24) is given by

$$Cov(r_{ij}^k, r_{i'j'}^{k'}) = E(r_{ij}^k - \bar{R}_{ij})(r_{i'j'}^{k'} - \bar{r}_{i'j'}) = E(\varepsilon_{ij}^k \cdot \varepsilon_{i'j'}^{k'}) = Cov(\varepsilon_{ij}^k, \varepsilon_{i'j'}^{k'}),   \quad (1.51)$$

where we assumed that the average signal $\bar{r}_{ij}$ is a good estimator for the expected value of r_{ij}^k and, hence, noise and residuals coincide. A priori, these correlations are unknown and need to be estimated from the measured data. Theoretically, the dimension of the spatiotemporal covariance matrix is $IJK \times IJK$. However, in order to estimate the matrix, multiple independent measurements are needed. It is common practise to assume that the different trials, corresponding to index k, are independent. Hence, the dimension of the matrix is reduced to $IJ \times IJ$. This, in turn, is still huge: I may equal

150 and J, say 300, which yields a 45000×45000 matrix to estimate. Using the trials as independent measurements, one needs at least IJ trials in order to obtain a nonsingular sample covariance matrix (see equation 1.50). Hence, this approach is not feasible in practise. To reduce the dimensionality of the parameter space, the spatiotemporal covariance matrix is parameterised: in [16, 20, 42] the Kronecker product (KP) model of a temporal matrix T and a spatial matrix X is presented. The expression in equation (1.51) is then parameterised as

$$Cov(\varepsilon_{ij}^k, \varepsilon_{i'j'}^{k'}) = \delta_{kk'} T_{jj'} X_{ii'}, \tag{1.52}$$

where δ denotes the Kronecker delta:

$$\delta_{kk'} := \begin{cases} 1 & \text{for} \quad k = k' \\ 0 & \text{for} \quad k \neq k'. \end{cases} \tag{1.53}$$

The spatiotemporal index, that is, the index of the spatiotemporal covariance matrix, is the index of $vec(E^k)$ in equation (1.26). Denoting the spatiotemporal covariance matrix of the MEG residuals by $\Sigma \in \mathbb{R}^{IJ \times IJ}$, the KP model is expressed as

$$\Sigma = T \otimes X. \tag{1.54}$$

The Kronecker product possesses elegant properties [31, 104], which are used in this thesis. The most important formulas for the MEG/EEG inverse problem are

$$(T \otimes X)^{-1} = T^{-1} \otimes X^{-1} \tag{1.55}$$
$$|T \otimes X| = |T|^I \, |X|^J, \tag{1.56}$$

where $|A|$ denotes the determinant of a matrix A.

The distribution of the MEG residuals in (1.27) is usually assumed to be Gaussian:

$$vec(E^k) \sim \mathcal{N}(0, T \otimes X), \tag{1.57}$$

for all k. For one trial, the corresponding marginal probability density function (pdf) becomes

$$f_{\Psi, S, \Sigma}(E^k) = \frac{1}{(2\pi)^{\frac{IJ}{2}}} \frac{1}{|T|^{\frac{I}{2}}|X|^{\frac{J}{2}}} \exp\left[-\frac{1}{2}[vec(E^k)]^t (T^{-1} \otimes X^{-1}) vec(E^k)\right]. \tag{1.58}$$

Substituting $E^k = R^k - \Psi S$, the pdf of the joint distribution of all trials becomes

$$f_{\Psi, S, \Sigma}(E^1, \ldots, E^K) =$$

$$\frac{1}{(2\pi)^{\frac{IJK}{2}}} \frac{1}{|T|^{\frac{IK}{2}}|X|^{\frac{JK}{2}}} \exp\left[-\frac{1}{2}\sum_{k=1}^{K}[vec(R^k - \Psi S)]^t (T^{-1} \otimes X^{-1}) vec(R^k - \Psi S)\right]$$

$$= \frac{1}{(2\pi)^{\frac{IJK}{2}}} \frac{1}{|T|^{\frac{IK}{2}}|X|^{\frac{JK}{2}}} \exp\left[-\frac{1}{2}\sum_{k=1}^{K} tr[(R^k - \Psi S)^t X^{-1}(R^k - \Psi S)T^{-1}]\right], \tag{1.59}$$

which is just the product of the K marginal pdfs. For the last equality in equation (1.59) the following property was used [63]

$$tr(ABCD) = (vec(A^t))^t(D^t \otimes B)vec(C) \tag{1.60}$$

for any $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{p \times q}$ and $D \in \mathbb{R}^{q \times m}$.

In the GLS paradigm the sum in the exponent in (1.59) is minimised with respect to all model parameters. In the maximum likelihood (ML) method the pdf is regarded as a function of the model and covariance parameters, rather than a function of the measured data. This function is called the likelihood function. Using the ML method, the likelihood function is maximised with respect to all model and covariance parameters. When the spatiotemporal covariance is known, that is X and T are known, these two methods, ML and GLS, coincide and the cost function becomes:

$$\min_{\xi,\eta,S} \sum_{k=1}^{K} tr[(R^k - \Psi(\xi,\eta)S)^t X^{-1}(R^k - \Psi(\xi,\eta)S)T^{-1}$$

$$= \min_{\xi,\eta,S} tr[(\bar{R} - \Psi(\xi,\eta)S)^t X^{-1}(\bar{R} - \Psi(\xi,\eta)S)T^{-1} + \gamma \tag{1.61}$$

where

$$\gamma = \sum_{k=1}^{K} tr[(R^k)^t X^{-1} R^k T^{-1}] - tr[\bar{R}^t X^{-1}\bar{R}T^{-1}] \tag{1.62}$$

does not depend on the parameters in ξ, η and S. Estimation of the parameters is similar to the OLS estimation explained in section 1.3.2, that is, by differentiation of the cost function.

When the covariance matrices X and T are unknown, which is generally the case, these matrices can be estimated within the ML paradigm. In such a case, the likelihood in (1.59) is maximised with respect to both the covariance and the model parameters, i.e. with respect to X, T, ξ, η and S. Differentiating and successively equating to zero for all these parameters, yields a complicated system, of which the estimators for all parameters are expressed in terms of the other parameters. This system yields the exact solution, though has to be solved iteratively, which requires much more time than reasonable. Therefore, the ML estimation method is usually split into two parts: first, the noise parameters are estimated and secondly, the model parameters are estimated as in the GLS method. In the first step the likelihood function in equation (1.59) is differentiated with respect to X and T. In this step, the parameters ξ, η and S are still unknown, and the expression ΨS in the cost function needs to be replaced by a known expression. For this substitution the ML estimator for the entire term (ΨS) is used:

$$\widehat{\Psi S}_{ML} = \frac{1}{K} \sum_{k=1}^{K} R^k = \bar{R}. \tag{1.63}$$

This estimator can be derived by differentiation of the likelihood function in equation (1.59) with respect to the entire term (ΨS). In other words, in the first step, the noise is replaced by the residuals, assuming that the statistical properties of noise and

residuals are comparable. After this substitution, the likelihood function is maximised with respect to X and T, yielding the following system of estimators:

$$\hat{T}_{ML} = \frac{1}{IK} \sum_{k=1}^{K} (R^k - \bar{R})^t X_{ML}^{-1} (R^k - \bar{R}) \tag{1.64}$$

$$\hat{X}_{ML} = \frac{1}{JK} \sum_{k=1}^{K} (R^k - \bar{R}) T_{ML}^{-1} (R^k - \bar{R})^t. \tag{1.65}$$

This system has to be solved iteratively, starting with $T = I_J$ for example. Then, in the second step of the ML method, these estimated covariance matrices are used as the true matrices in the GLS estimation of the model parameters with cost function (1.61).

Estimating and understanding these covariance matrices and more advanced models for the spatiotemporal covariance is one of the main goals of this thesis. As stated in section 1.4.2, nonstationarities in these covariance matrices are indicators for the invalidity of the SPN model. This and related topics are discussed in detail in chapters 2, 3 and 4.

1.5 This thesis

1.5.1 Temporal nonstationarities

Incorrectly assuming the Signal Plus Noise model (1.24) leads to a nonstationary temporal covariance structure (section 1.4.2 and [101]). To investigate the validity of the SPN, trial-to-trial variations are investigated in Chapters 2, 3 and 4.

In Chapter 2 the fixed response model of the SPN model is extended to a response model that allows varying amplitude over trials (as in Figure 1.13). This model is not only applicable in evoked response experiments but also is very interesting in application to epileptic spike data. It is shown that allowing the amplitude to vary removes the temporal nonstationarities in the statistical structure of the residuals of the spike data.

In chapter 3 the ML-estimated temporal covariance matrix is parameterised by only 6 parameters in order to understand the temporal properties of the MEG residuals. In the presented Poisson modulated alpha model, the background activity is assumed to consist of frequency specific alpha activity and additional random noise. It appears in that chapter that the ML-estimated temporal structure can be explained by an, in principle, stationary model up to an error of less than 1 % relative matrix power.

Temporal nonstationarities are also addressed in the discussion in chapter 7. In this final chapter some special phenomena that can cause artifactual temporal nonstationarities are examined.

1.5.2 Improvement of the spatiotemporal covariance model

The KP model for the spatiotemporal covariance is based on the assumption that time and space are not correlated. In other words, the temporal covariance is fixed over MEG sensors, while the spatial covariance is fixed in time. This, though, is a simplification of the reality. For example, the alpha rhythm is not equally distributed over the sensors

with respect to the other background activity, and, therefore, contradicts the KP assumption. For this reason, an extension of the KP model is sought in chapter 4: a sum of Kronecker products. This extension has the advantage that multiple spatiotemporal patterns can be accounted for by different terms. Furthermore, higher order terms may show nonstationarities that do not appear in the single KP model. Hence, the sum of KP model contains valuable information about the statistical characteristics of the background noise. Regarding the practical application in dipole localisation, this model is hampered by computational considerations. A nonconventional amount of computer power would be needed in order to use this model for dipole fitting. Nonetheless, it appears from the analysis in chapter 4, that the improvement of the accuracy of the source parameters is expected to be only marginal, compared to the use of the single KP. Apparently, for dipole localisation one KP is adequate, while a sum of KP is interesting for its physiological information. For further improvement of the source localisation method, the emphasis should be on improving the forward model rather than on the spatiotemporal covariance model.

1.5.3   Improving of the forward model

Following the suggestion of chapter 4, improvement of the forward modelling is explored in chapters 5 and 6 by combining MEG data of multiple related experiments in a single model. The Coupled Dipole Model and the Extended Coupled Dipole Model are presented in these two chapters, which can be used to localise the sources in multiple related MEG/EEG data sets simultaneously. This integrated method is applicable to data sets in which the same sources are active or proportional source time functions occur. This approach stabilises the solution to the inverse problem and relatively increases the signal-to-noise ratio (SNR). Both models use a basic set of common sources and a basic set of common source time functions (wave shapes). The set of common sources contains all sources that are active in at least one of the data sets in the analysis. The number of both the spatial and the temporal common components is chosen by the user. The model for each data set under consideration is a linear combination of these common components. In the Coupled Dipole Model the user is allowed to specify in more detail this linear combination based upon prior information or assumptions, whereas the Extended Coupled Dipole Model estimates more objectively this linear combination. Furthermore, the statistical transparency of the Extended Coupled Dipole Model allows for the computation of the statistical significance of the estimated source activity, which can be illustrated by the confidence regions around the estimated source time functions.

Chapter 2

A maximum likelihood estimator for trial-to-trial amplitude variations in noisy MEG/EEG data sets

2.1 Introduction

Spontaneous brain activity in MEG and EEG signals is caused by dendritic and return currents of neurons receiving signals from connected neurons (see section 1.1.2). Part of this activity, like the alpha and mu rhythms are enhanced in the absence of visual and motor input, respectively. When a stimulus is presented to a subject several parts of the brain that are involved in processing the stimulus will show increased neuronal activity and will therefore act as the generators of the brain response, which can be recorded by MEG or EEG. When the same stimulus is applied repeatedly, the brain response is usually extracted from the recorded data by a simple averaging technique. The question is how the averaged signal should be interpreted.

The (often implicit) assumption underlying response averaging is the Signal Plus Noise model (section 1.3.1):

$$r_{ij}^k = r_{ij} + \varepsilon_{ij}^k \tag{2.1}$$

for $k = 1, \ldots, K$. Here r_{ij}^k is the recorded signal at channel i, time sample j and trial k. r_{ij} is the spatiotemporal pattern of the brain response, which is assumed to be constant over trials. And ε_{ij}^k is the background noise, which is assumed to have a Gaussian distribution and which is in general correlated over sensors and time samples. It is straightforward to show that the maximum likelihood (ML) estimator for the constant brain response r_{ij} is the simple average over trials of the recorded data.

Nevertheless, it is known from experiments that the constant response assumption is false in general. Examples of trial-to-trial variations in human EEG are habituation effects [79, 113], P300 effects and event related synchronisation and de-synchronisation

effects [51, 84]. However, because these effects are generally small, it is not straightforward to prove their presence in the data and to distinguish them from the background noise. This problem becomes even more severe if one realises that the spatial and temporal properties of the background noise are very similar to those of the brain response under study, since the background noise mainly consists of spontaneous brain activity.

In a recent paper [101] is was demonstrated that neglecting trial-to-trial variations results in an estimate of the background noise of which the variance is nonstationary over the time interval of interest. However, that paper does not show how trial-to-trial variations can be estimated from 'first principles'. In [44] and [85] a ML model is formulated yielding estimators of amplitude and latency jitters of single trials in a multi-trial evoked potential experiment. A restriction of these papers is that they are based on the assumption of uncorrelated background noise, whereas it is known that for instance the alpha rhythm is both correlated in time and over sensors (e.g. [16, 42]). Furthermore, when estimating the response, the trial-to-trial variations and the noise characteristics one has to realise that these components are mutually dependent, because the measured signal consists of their sum.

In this chapter, a model is presented to describe amplitude jitter in spatiotemporal MEG (e.g. habituation effects) and ML estimators for all model parameters are derived without the restricting assumption of uncorrelated background noise. This model is applied to data of standard SEF experiments in which the subject is electrically stimulated at the median nerve. Since the model is equally well applicable to the simultaneous analysis of epileptic spikes of varying polarity, an example of that kind of data is discussed also.

2.2   Methods

2.2.1   Model

In the data model it is assumed that the recorded brain signal r_{ij}^k consists of a constant spatiotemporal pattern r_{ij} multiplied by a trial dependent amplitude factor α^k, and a Gaussian noise part ε_{ij}^k

$$r_{ij}^k = \alpha^k r_{ij} + \varepsilon_{ij}^k \tag{2.2}$$

for $k = 1, \ldots, K$. The noise part ε_{ij}^k is assumed to be statistically independent from trial to trial and its spatiotemporal covariance is modelled as a Kronecker product $T \otimes X$ (e.g. [63]) of a spatial covariance matrix X and a temporal covariance matrix T. This can be expressed as (see equation 1.52):

$$\mathcal{E}[\varepsilon_{ij}^k \varepsilon_{i'j'}^{k'}] = \delta_{k,k'} X_{ii'} T_{jj'}. \tag{2.3}$$

In this model, the unknown parameters are contained in α^k, r_{ij}, X and T and need to be estimated from the measured data r_{ij}^k.

The underlying idea of the ML paradigm is to express the probability density function of the noise as a function of the a priori unknown parameters, as the likelihood function, and to find the parameters for which this likelihood is maximum. Following the definition in equation 1.25, the array containing the measured data in trial k is

denoted by R^k, and the fixed responses r_{ij} are gathered in matrix R. With the above assumptions the likelihood function is given by

$$f_\varepsilon(\alpha, R, X, T) = \frac{\exp\left[-\frac{1}{2}\sum_k tr[(R^k - \alpha^k R)T^{-1}(R^k - \alpha^k R)^t X^{-1}]\right]}{(2\pi)^{\frac{IJK}{2}}|T|^{\frac{IK}{2}}|X|^{\frac{JK}{2}}}$$ (2.4)

Here I is the number of channels, J is the number of time samples and K is the number of trials.
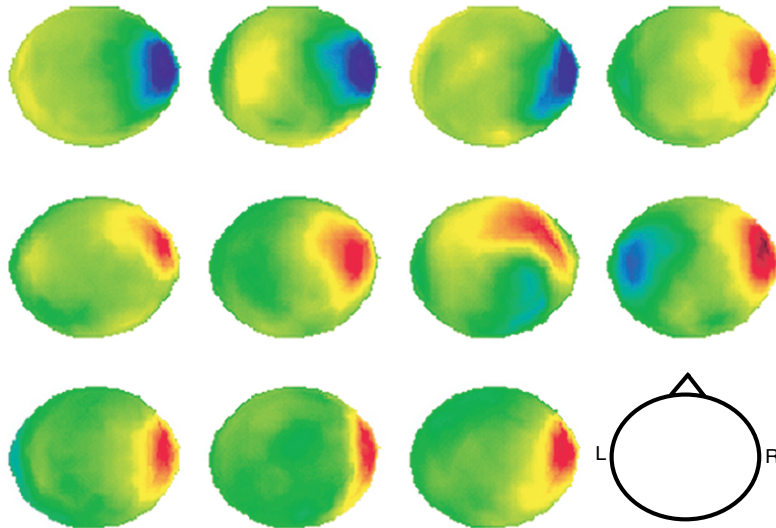
Figure 2.1: *A selection of maps corresponding to the markers set by a technician on different spikes. The spatial patterns of these spikes are similar, though the polarity changes from spike to spike.*

The ML-estimators of α, R, X and T are derived by equating the corresponding derivatives of equation (2.4) to zero and solving the estimators from the resulting equations. This method yields as ML estimator for R the weighted average of the single trial data:

$$R = \frac{1}{K}\sum_k \alpha^k R^k.$$ (2.5)

From this estimator one can see that, in case it is assumed that no trial-to-trial variations are present, i.e. that $\alpha = (1, \ldots, 1)^t$, then this estimate of the brain response reduces to the simple average over trials. The estimators for the weights α^k appear to be the elements of the eigenvector with the largest eigenvalue of the following system:

$$\sum_{k_2} tr[R^{k_1}T^{-1}(R^{k_2})^t X^{-1}]\alpha^{k_2} = \lambda_{max}\alpha^{k_1}.$$ (2.6)

For the spatial and temporal covariances a system of equations can be derived in which T is expressed in terms of X and vice verse:

$$X = \frac{1}{J}\left(\frac{1}{K}\sum_k R^k T^{-1}(R^k)^t - RT^{-1}R^t\right)$$ (2.7)

$$T \;=\; \frac{1}{J}\Big(\frac{1}{K}\sum_{k}(R^k)^t X^{-1} R^k - R^t X^{-1} R\Big). \tag{2.8}$$

Note that in equations (2.5 - 2.8), the symbols α, R, X and T have been used to indicate the estimators of the true values. Equations (2.5 - 2.8) have to be solved iteratively. First, $\alpha = (1, \ldots, 1)^t$ and $T = I_J$ are taken, and then equations (2.7 - 2.8) are solved iteratively [20]. These first estimates of the covariances are substituted into (2.6) to obtain an update of α, which is substituted back into (2.5), etc. In practise, it appears that a few iterations between equations (2.5 - 2.6) and equations (2.7 - 2.8) are sufficient to obtain a stable solution.

2.2.2   Data

Somatosensory Evoked Field (SEF) data were measured from nine healthy subjects by applying an electrical median nerve stimulation to both the left and right wrist (N=8), or only to the left wrist (N=1). The stimulus intensity was individually adjusted such that a twitch of the thumb appeared. A regular inter-stimulus interval was used, which was set at 1 s. MEG data were acquired on a whole head 151 channel system of CTF, using a sample rate of 2083 Hz. Off-line, bad epochs and bad channels were marked and removed from the data. The number of remaining good responses varied between 450 and 570. Data were offset corrected using a pre-stimulus interval of 100 ms, which is optimal to reduce the nonstationary influence of the offset correction (see chapter 3). No band pass filtering was applied. Two time intervals were selected to investigate trial-to-trial variations. The early time window extended from 23 to 37 ms post-stimulus, the late window from 30 to 120 ms post-stimulus. These windows were selected based on the distribution of the averaged MEG data power as a function of time.

As stated above, another interesting application of the presented model is the simultaneous analysis of multiple epileptic spikes of varying polarity. For this kind of data, averaging according to the SPN would cause the signal to vanish because of the varying polarities of the spikes. For this study, MEG data of an epileptic patient were collected during a period of one hour, and stored in six data files of 10 minutes each. First, the data were motion-corrected [19] and spikes were marked at the maximum signal amplitude by an experienced MEG technician. Then for each spike marker a symmetric data window was cut, 110 ms left and 110 ms right of the marker. These spikes were subjected to cluster analysis [106], after which 254 spikes remained having a similar topography, but a varying amplitude and polarity. The field maps of a selection of these spikes are presented in Figure 2.1. One observes that most graphs only show one polarity, while the other one falls outside the MEG helmet. This is a common picture for MEG spikes, in particular in case of temporal lobe epilepsy.

2.3   Results

2.3.1   SEF data

The algorithm presented in the previous section usually converged in 8 to 12 iterations. The iteration stopped (converged) when the relative difference in matrix power between

an estimate and the next estimate was below 10^{-10}. For the late responses (187 samples) this resulted in a computation time of about 2 h (Pentium IV, 1 GHz) or more. For the early responses (29 samples) this time amounted typically 15 min. Figure 2.2 shows
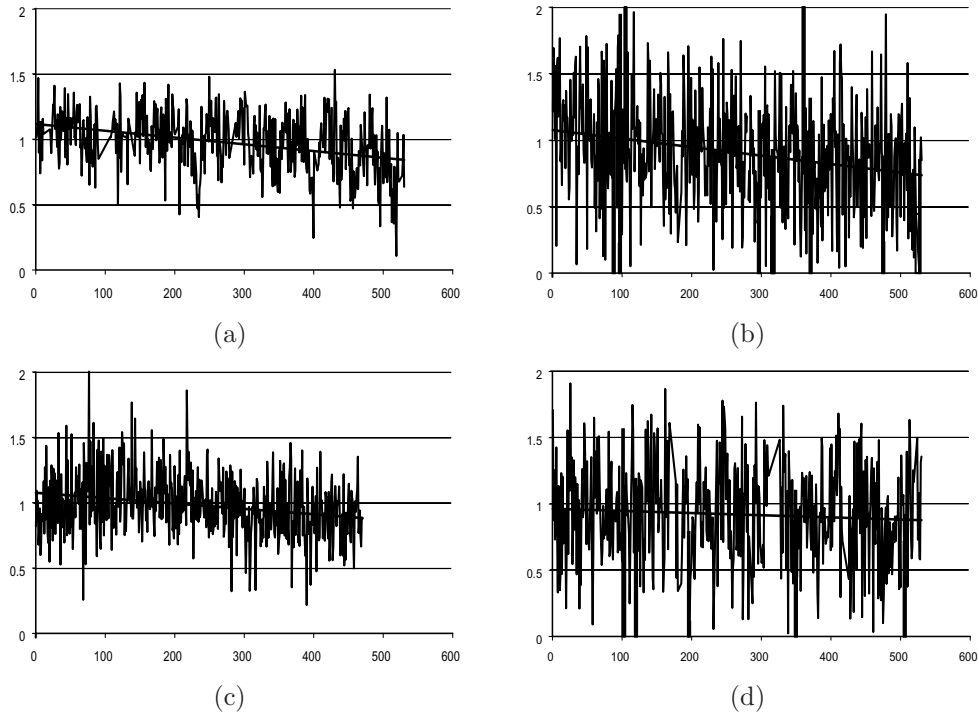
Figure 2.2: *The estimated amplitudes for the late time window (30 - 120 ms post-stimulus) of the SEF data sets for four subjects. The horizontal scale runs from 0 to 600 seconds (trials). The vertical scale is in arbitrary units, where a constant value of 1 indicates the solution for the simple SPN model with no amplitude variations. In all cases, there are relative fast variations from one trial to the next and also slower variations on scales between 50 and 200 s. Furthermore, all these cases show a negative trend indicating that the response gets weaker over time.*

four typical examples of the estimated trial-to-trial amplitude variations in the (late) SEF response. In this figure the estimated trial multiplication factor α^k is plotted as function of trial or equivalently, as function of stimulus time in seconds. In all cases, there are relative fast variations from one trial to the next and also slower variations on scales between 50 and 200 s. To extract systematic behaviour from all data sets, a straight line was fitted to the estimated amplitudes, using the stimulus times of each trial as x-coordinate and the estimated α^k as y-coordinate. The quality of these line fits, was in all cases comparable to the examples presented in figure 2.2. In all cases presented in Figure 2.2, the slope of the fitted trend line was negative. The data of the early components and the data of the other subjects were treated similarly, yielding an estimated linear trend for all data sets. Figure 2.3 summaries all these estimated trends for both the early and late responses for all subjects. From these graph it appears that for the early responses the trend varies strongly over subjects, whereas for the late
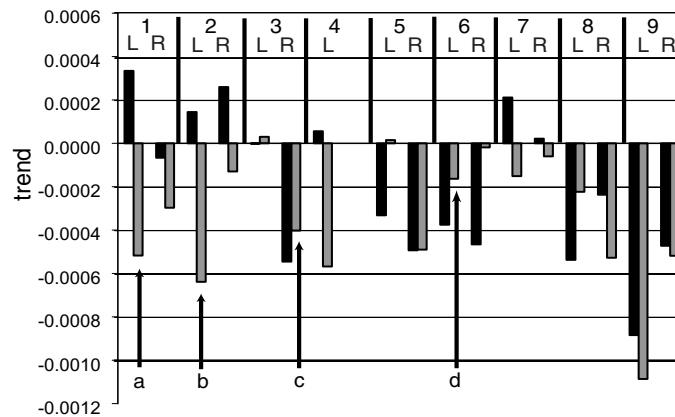
Figure 2.3: *An overview of the trends in the SEF responses for different data sets. The numbers on top indicate subject. For each subject (except subject 4) both left (L) and right (R) SEF stimulation were performed. The black bars present the trends of the early response windows (20-35 ms), and the grey bars the late responses (30-120 ms). The arrows correspond to the cases shown in figure 2.2.*

responses all trends are negative or slightly positive at most.

2.3.2   Spike data

Figure 2.4 shows the results of the amplitude analysis for the spike data. Since the spikes do not regularly occur over time, the estimated weight factors α^k are now plotted as function of spike occurrence in stead of time. It appears that 60 of the 254 spikes have a negative polarity with respect to the others. Furthermore, the 'negative' spikes seem to have slightly smaller amplitude than the 'positive' ones.
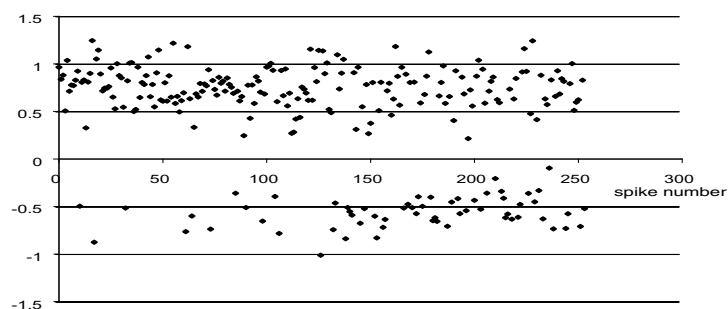
Figure 2.4: *The amplitude factors as function of spike number for different epileptic MEG spikes. It appears that 25 % of the spikes have a different polarity than the remaining 75 %.*

When the spikes are simply averaged, which is done in the first iteration of the estimation of amplitude modulation factors, the temporal variance should change in time with a pattern corresponding to the shape of the averaged event, as predicted in

[101]. This effect is demonstrated in Figure 2.5. After 9 iterations, when the algorithm converges, the variance in Figure 2.5(a) has decreased and becomes much more constant. Figures 2.5(b) and 2.5(c) show that not only the variance, but the entire covariance matrix T becomes much more stationary. The remaining nonstationarities may be explained by a small misalignment in the spike markers, which introduce latency jitters.

<center>(a)   (b)   (c)</center>
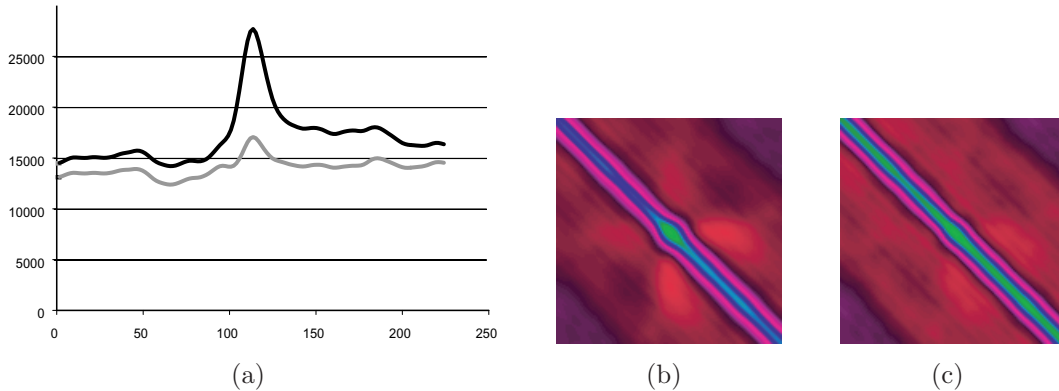
Figure 2.5: *The estimated temporal variances (a) and covariance structures plotted in grey scale (b, c) in the first and last iteration of the algorithm. In figure a, a nonstationary structure is seen for the first iteration (black line), while the estimated variance in the last iteration (grey line) is much more constant. These variances are the diagonals of the temporal covariance matrices, that are plotted entirely in figures b (first iteration) and c (last iteration). These figures show that the erroneous assumption of no amplitude variation, i.e. the starting assumption in the first iteration, yields a nonstationary structure in the entire covariance matrix that has disappeared in the last iteration.*

2.4   Discussion

This chapter presents an ML estimation model to extract trial-to-trial amplitude variations in brain response and to distinguish these variations from spatially and temporally correlated background noise. The application of this model to 17 SEF data sets shows that single trial SEF data show a systematic weakening of the SEF amplitudes. This negative trend in the response can be explained as habituation or nerve fatigue effects. Similar effects are also present during visual stimulation, both in EEG and NIRS (Near Infra Red Spectroscopy) data, as presented in [79]. In that study, the responses were averaged over 15 trials (of .33 s) and trends were computed over 12 of these subsequent averages. In all subjects (N=12) negative trends were found in the component from 100 to 135 ms, whereas for the earlier component (from 75 to 110 ms) no systematic negative trend was found. This finding is similar to our results for the SEF data, although the analysis methods are quite different. Strong habituation effects in group-averaged auditory fMRI data have been reported in [83]. That study as well as other studies on

habituation contain several implicit assumptions that are avoided with the ML method presented here. The price to be paid is the relative long computation time.

The application of the model to spike data can be considered as a validation of the method, because in these data the amplitude variations are so strong that they are already visible in the raw data (Figure 2.1). However, the model presented in this paper can be of practical use in the analysis of multi-spike data, when the number of spikes is limited, when they have a low signal-to-noise ratio or when they appear in different polarities. Such a change in polarity without a change of pattern could indicate a small shift of the underlying current dipole from one gyrus to the next. This is not an uncommon situation with MEG spikes, as was shown in a recent paper [106]. In that paper different spike types were automatically grouped by computing their mutual Euclidean signal distances and applying a clustering algorithm. When in that algorithm in stead of Euclidean distances, a negative correlation measure would be used, spikes with the same spatial pattern but different polarity would show up in the same cluster, thereby increasing the number of events per cluster. Applying this alternative distance measure, in combination with the current model, would increase the signal-to-noise ratio, compared to the case of subgroup averaged spikes.

In [5] the same SEF data sets are used to study the stationarity of the background noise (see chapter 3). It is concluded in that chapter that, when accounting for non-stationarities caused by the baseline correction, the temporal covariance matrix of the background noise can be explained for 99% or more by a stationary noise model. Nevertheless, the present study shows that there are weak trial-to-trial variations in the data sets. It is apparently so that the amount of trial-to-trial variations should exceed a certain threshold in order to become visible in the temporal covariance of the background noise, as is the case with the spike data set.

The central idea behind the presented method is to consider the determination of brain responses as a parameter estimation problem, using a mathematical model to describe single trial data. Therefore it is, at least in principle, straightforward to extend our model to physiologically more advanced models, such as suggested in e.g. [65, 82]. However, one should also realise that including, for instance, simple latency jitters in the model, would increase the number of nonlinear parameters by 300 to 1000 (one for each trial), which is probable become problematic due to multiple local maxima in the likelihood function. In the presented model, which is also a nonlinear model, the nonlinear amplitude parameters can be solved simultaneously, by solving an eigenvalue problem. Another aspect of more advanced models is that an increase of the number of parameters is accompanied with, on the one hand, an improvement of the goodness of fit, and on the other hand, a larger variance in the estimated parameters. It can be argued that the ML framework is very well suited to address this dilemma objectively. However, to perform these ideas in practise, a nonconventional amount of computer power is required. For that purpose a network of parallel computers could be used.

Chapter 3

A mathematical approach to the temporal stationarity of background noise in MEG/EEG measurements

3.1   Introduction

Background noise in MEG/EEG measurements is correlated both in space and in time. When estimating dipole source parameters one has to take into account the covariance of this noise. Moreover, the study of the background noise is also important for its own sake regarding the ongoing debate about the meaning of averaged brain responses in relation to the background noise [44, 65, 85, 89, 101]. When it is assumed that the recorded signal is a simple superposition of the brain response and the background noise, as in the Signal Plus Noise (SPN) model (see equation 1.24), the measured signal r_{ij}^k at channel i and time sample j in trial k is modelled as

$$r_{ij}^k = r_{ij} + \varepsilon_{ij}^k, \tag{3.1}$$

where r_{ij} is the brain response caused by the stimulus and ε_{ij}^k the measured noise. The SPN model is based on the assumption that the brain response r_{ij} does not vary over trials, whereas ε_{ij}^k does vary from trial to trial.

The spatial covariance of the background noise in equation (3.1) has already been studied in detail [39, 61, 62, 96, 109]; it has been shown that, by taking into account these spatial correlations, the accuracy of the estimated parameters in the source localisation is improved. Recently, in addition to the spatial covariance, also the temporal correlations were incorporated in the source localisation. It was demonstrated in [20, 42] that this addition generally improves the dipole estimation further. Both these studies were based on the SPN model.

In [20] and [42] the spatiotemporal noise covariance matrix Σ is modelled as the Kronecker product (KP) of a spatial covariance matrix X and a temporal covariance

matrix T in order to reduce the dimensionality of the parameter space to a feasible size (see [16] and equation (1.52)):

$$\Sigma = T \otimes X. \tag{3.2}$$

Note that X and T are not unique due to a common factor; therefore, one is normalised. Furthermore, different trials k and k' are assumed to be uncorrelated. The basic idea of the KP parametrisation is the assumption that the noise covariance between two residuals at channels i and i' at time samples j and j' respectively can be factored as the product of a spatial term $X_{ii'}$ and a temporal term $T_{jj'}$

$$E(\varepsilon_{ij}^k \ \varepsilon_{i'j'}^{k'}) = \Sigma_{ij,i'j'} \ \delta_{k,k'} = X_{ii'} T_{jj'} \delta_{k,k'}. \tag{3.3}$$

Here $\delta_{k,k'}$ denotes the Kronecker delta function, defined in (1.53).

In [42] a parametric model is used for both X and for T, where the matrix elements are assumed to depend only on sensor distance and time difference, respectively. In [20], maximum likelihood (ML) estimators are derived for R, X and T without further assumptions. The resulting estimators are for the brain response

$$\hat{R} = \frac{1}{K} \sum_{k=1}^{K} R^k, \tag{3.4}$$

which is the usual averaged signal, and for the covariance matrices:

$$\hat{X} = \frac{1}{JK} \Big[\sum_{k=1}^{K} R^k \hat{T}^{-1} (R^k)^t - \hat{R} \hat{T}^{-1} \hat{R}^t \Big] \tag{3.5}$$

$$\hat{T} = \frac{1}{IK} \Big[\sum_{k=1}^{K} (R^k)^t \hat{X}^{-1} R^k - \hat{R}^t \hat{X}^{-1} \hat{R} \Big]. \tag{3.6}$$

Equations (3.5 - 3.6) need to be solved iteratively, starting with $T = I_J$ for example. In the expressions for the estimators, I, J and K are the numbers of channels, time samples and trials respectively and R^k is the $(I \times J)$ data matrix of the kth trial. A^t denotes the transpose and A^{-1} the inverse of matrix A.

The statistics of ε_{ij}^k express properties of the ongoing background activity. If the background noise is modelled as the electromagnetic field of randomly distributed stationary dipole sources, assuming these dipoles are statistically independent and their positions are independent of their amplitudes, then the spatiotemporal covariance presents itself as a Kronecker product as in equation (3.2) [16].

Regarding the spatial part of the KP, different models have been investigated [16, 18, 42]. The model for the spatial covariance derived in [16] can be interpreted as a function of sensor distance. A similar stationary model was presented [42]. Although the KP structure was already suggested in [16], in that study the temporal structure was explicitly left unspecified. Now the question arises whether the temporal covariance has a stationarity property similar to that of the spatial covariance. Stationarity in the temporal domain (see section 1.4.2) indexcovariance!temporal is the property that the temporal covariance only depends on time difference and is independent of time, so that T is constant along subdiagonals (i.e. Toeplitz), cf. [42].

To investigate the possibility of a physiologically adequate parametric model for T, equations ((3.4) - (3.6)) were applied to several data sets. Figure 3.1 presents an example of $\hat{T}$. In Figure 3.1(a) the matrix is visualised by plotting the average along subdiagonals as function of time difference. For the j_0^{th} subdiagonal this average is

$$\frac{1}{J - j_0} \sum_{j=0}^{J-j_0-1} T_{j,j+j_0}. \tag{3.7}$$

In Figure 3.1(a) this average covariance (black line) and its standard deviation (grey band) are plotted as function of time lag together with the diagonal of $\hat{T}$, the variance, (red line) as function of time. In Figure 3.1(b) the same matrix is plotted in colour.

(a)   (b)
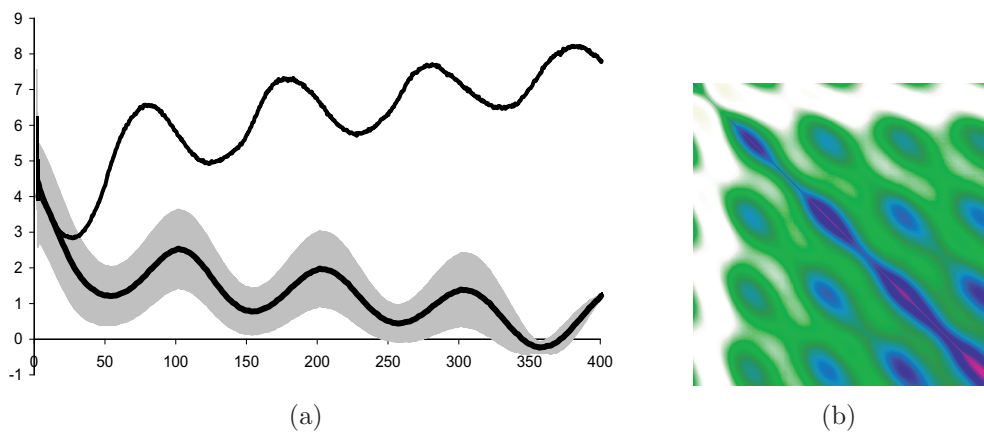
Figure 3.1: *Two different visualisations of the same ML-estimated temporal covariance matrix. (a): lower line: average temporal covariance as in formula (3.7) ($\pm$ SD in the grey band) as function of time difference (ms); upper line: temporal variance as function of time (ms). The vertical scale in figure (a) is in arbitrary units because of the common factor of X and T (3.2). The horizontal line denotes the time in ms. (b): entire temporal covariance matrix (400 ms by 400 ms) plotted as bitmap. The colour scale is as in Figure 3.2.*

Figure 3.2: *The colour scale used in Figure 3.1(b). White indicates zero, purple is positive, red is negative.*

In order to plot the matrix, its entries are scaled such that the entry that is largest in absolute value equals 1 or -1. The colour scale used for this kind of covariance plots is presented in Figure 3.2. As with usual printing of matrices, all entries are arranged in a square, and in stead of values, corresponding colours are plotted. Nonstationarities and oscillations in the temporal domain can now easily be detected: a stationary temporal

covariance matrix has a constant value (colour) along its (sub)diagonals and oscillations in the covariance are reflected by a line pattern parallel to the diagonal.

The matrix in Figure 3.1 shows a nonstationary covariance structure. First of all, the variance in Figure 3.1(a)(upper black line), i.e. the diagonal of the matrix, is not constant over time but oscillates. The same holds true for the subdiagonals (Figure 3.1(b)). The average covariance in Figure 3.1(a) (lower black line) also oscillates, though this does not indicate nonstationarity. The underlying reason is that the covariance is a function of time lag, rather than a function of time. The oscillation visible in the covariance indicates a 10 Hz oscillation in the background activity (see section 1.4.1). In Figure 3.1(b) this oscillation is visible in the direction perpendicular to the diagonal. However, the SD-band around this average covariance witnesses again nonstationarity of ε_{ij}^k: the rather high and fluctuating SD shows that on some subdiagonals the variation around the average covariance is higher than on others. In Figure 3.1(b) the oscillations in both the diagonal and the anti-diagonal direction result in blue spots in the figure, whereas in case of temporal stationarity one would expect a line pattern parallel to the diagonal. In all, Figure 3.1 shows an example of a nonstationary temporal covariance matrix.

The oscillations in Figure 3.1 suggest that the background noise is partly generated by alpha activity. Therefore in the present chapter a parametrisation of T is proposed that is based on a noise model consisting of two components: alpha activity and additional random noise. The covariance of this additional noise is modelled as an exponentially decreasing function of time lag (i.e. as low frequency noise). The presented ongoing alpha model is based on the assumption of ongoing alpha activity, i.e. one everlasting alpha wave. The more advanced Poisson modulated alpha model assumes that the alpha activity consists of separated waves which occur every now and then and have finite duration. This assumption is more realistic because raw data clearly show separated waves.

In the next section the OAM and the PoMAM are presented and preprocessing of the data by an offset correction is considered in detail. Technical details of this section are moved to the appendix, in order to keep the text compact. The results section shows the application of the PoMAM to SEF data sets. And finally, the results are discussed in the last section.

3.2   Methods

3.2.1   Ongoing alpha model

In the ongoing alpha model (OAM) the temporal component of the background noise is modelled as the sum of an ongoing alpha wave and additional random noise. For convenience the model uses continuous time, and is converted to discrete time in its application to our experiment. The model for the noise in the kth trial, $\varepsilon_k(t)$, is formulated as

$$\varepsilon_k(t) = \Omega \sin\left(\omega t + \tau_k\right) + \eta_k(t) \tag{3.8}$$

where Ω is the amplitude, ω the frequency of the alpha activity and t indicates time instant. Because of the random inter stimulus interval the phase of the wave, τ_k, is

random for each trial, i.e. the stochast τ_k has the uniform distribution in $[-\pi, \pi]$. For the additional noise, $\eta_k(t)$, it is assumed that

$$E(\eta_k(t_1)\eta_k(t_2)) = \sigma^2 e^{-\kappa|t_2-t_1|} \tag{3.9}$$

where $\kappa > 0$ and σ^2 is the variance. This means that the additional noise is temporally stationary and its covariance decreases with time lag. Furthermore, the alpha wave is assumed to be independent of the additional noise. This yields the temporal covariance for an arbitrary trial k

$$\begin{aligned}
Cov(t_1, t_2) &= E(\varepsilon_k(t_1) \cdot \varepsilon_k(t_2)) \\
&= E[(\Omega \sin(\omega t_1 + \tau_k) + \eta_k(t_1)) \cdot (\Omega \sin(\omega t_2 + \tau_k) + \eta_k(t_2))] \\
&= \frac{\Omega^2}{2\pi} \int_{-\pi}^{\pi} \sin(\omega t_1 + \tau_k) \cdot \sin(\omega t_2 + \tau_k) d\tau_k + \sigma^2 e^{-\kappa|t_2-t_1|} \\
&= \frac{\Omega^2}{2} \cos(\omega(t_2 - t_1)) + \sigma^2 e^{-\kappa|t_2-t_1|}.
\end{aligned} \tag{3.10}$$

From (3.10) it is clear that the OAM is a stationary model: $Cov(t_1, t_2)$ only depends on time lag $t_2 - t_1$. In Figure 3.3 an example of the OAM covariance is presented. This
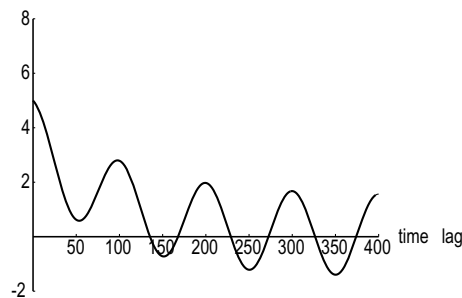
Figure 3.3: *The covariance in $(fT)^2$ of the OAM as function of time lag (ms) for $\omega = 20\pi$ rad s^{-1}, $\Omega^2 = 3$ (fT)^2, $\sigma^2 = 3.5$ (fT)^2 and $\kappa = 10s^{-1}$.*

figure shows that the OAM covariance does not vanish for large time difference as was the case in the experimental data presented in Figure 3.1(a), though remains oscillatory. For this reason the more realistic Poisson modulated alpha model is introduced.

3.2.2   Poisson modulated alpha model

In the Poisson modulated alpha model (PoMAM) the alpha activity is modulated by an interrupted Poisson process. A Poisson process with intensity parameter λ is a statistical process generating events at random with mean intermediate time $\frac{1}{\lambda}$ [11, 99]. The time between two consecutive events has the Exponential(λ) probability density function:

$$f_\lambda(t) = \lambda e^{-\lambda t} 1_{[0,\infty)}(t) \tag{3.11}$$

where $1_{[a,b]}(t) = 1$ for $t \in [a, b]$ and zero for $t \notin [a, b]$.
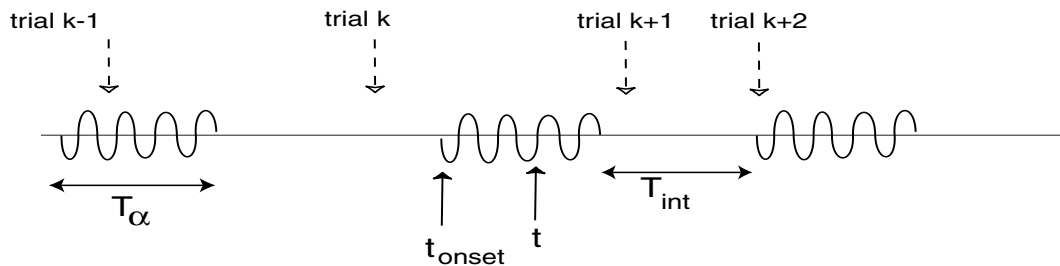
Figure 3.4: *Occurrence of alpha waves according to the PoMAM. Because of the random inter stimulus interval trial onsets can occur in any phase of the wave. T_{int} denotes the time between waves, T_α the duration of a wave. t_{onset} indicates the onset of a wave in trial k.*

In the covariance model the events stand for the onsets of α-waves (Figure 3.4). After each event the process is disrupted for the duration of the wave, which is fixed to T_α for all wave, after which it resumes to generate the next event (onset). Different waves are assumed to be uncorrelated.

In the PoMAM two stochastic processes are operating simultaneously, namely the Poisson process (generating wave onsets) and the random phase process. These two processes are assumed to be statistically independent.

Assuming the amplitude of a wave to be constant and equal to Ω, the amplitude time function (envelope) of a wave which started at $t = 0$ is $\Omega \, 1_{[0,T_\alpha]}(t)$. Let $\varepsilon_k^\alpha(t) = \varepsilon_k(t) - \eta_k(t)$ denote the alpha part of the noise. In Figure 3.4 the alpha activity at time t in trial k due to the wave started at t_{onset} is

$$
\begin{aligned}
\varepsilon_k^\alpha(t) &= \Omega 1_{[0,T_\alpha]}(t - t_{onset}) \sin\left(\omega(t - t_{onset}) + \tau_k\right) \\
&= \Omega 1_{[0,T_\alpha]}(t - t_{onset}) \sin\left(\omega t + \tau_k'\right)
\end{aligned}
\tag{3.12}
$$

where $\tau_k' = -\omega t_{onset} + \tau_k$. $\tau_{k'}$ has the same probability density function as τ_k, namely the uniform distribution in $[-\pi, \pi]$. Therefore, in the sequel τ_k' will be denoted as τ_k again. From the model (3.12) it is clear that the Poisson process only comprises the amplitudes of the waves, so that the expected value of the error is still zero because of the random phase τ_k.

The covariance in the PoMAM is

$$
Cov(t_1, t_2) = E(\varepsilon_k(t_1) \cdot \varepsilon_k(t_2)) = Cov^\alpha(t_1, t_2) + Cov^\eta(t_1, t_2)
\tag{3.13}
$$

because the alpha activity and the additional noise are assumed to be independent. In this expression the second term is the same as in the OAM:

$$
Cov^\eta(t_1, t_2) = \sigma^2 e^{-\kappa|t_2 - t_1|}.
\tag{3.14}
$$

The computation of the alpha part of the covariance is complicated because there are different possibilities for t_1 and t_2. First of all, if t_1 and t_2 are in the same wave, then $Cov^\alpha(t_1, t_2) \neq 0$ in general. Secondly, if one of the two instants, say t_1, is in a wave, and t_2 is not in the same wave (not in any wave or perhaps in another wave) then

$Cov^{\alpha}(t_1, t_2) = 0$, because either $\varepsilon_k^{\alpha}(t_2) = 0$ or t_2 is in another wave, and different waves are assumed to be independent. And finally, if both instants are not in any wave, $Cov^{\alpha}(t_1, t_2) = 0$ obviously. Therefore only the case of t_1 and t_2 being in the same wave has to be considered.

In order to compute the covariance of the PoMAM the Total Probability Theorem [11] applied to a function $g(X)$ of a stochast X is used:

$$E(g(X)) = P(A) \times E(g(X)|A) + P(A^c) \times E(g(X)|A^c) \tag{3.15}$$

where A^c stands for the complement of event A. Defining A to be the event 't_1 *and* t_2 *are in the same wave*' and $g(X)$ to be $\varepsilon_k^{\alpha}(t_1)\varepsilon_k^{\alpha}(t_2)$, which is a function of both stochastic processes, (3.15) yields

$$Cov^{\alpha}(t_1, t_2) = E(\varepsilon_k^{\alpha}(t_1) \cdot \varepsilon_k^{\alpha}(t_2)) = P(A) \times E(\varepsilon_k^{\alpha}(t_1)\varepsilon_k^{\alpha}(t_2)|A). \tag{3.16}$$

In this formula the results from Appendix 3.5.1 ($P(A)$) and Appendix 3.5.2 ($E(\varepsilon_k^{\alpha}(t_1)\varepsilon_k^{\alpha}(t_2)|A)$) are substituted ((3.31) and (3.34)):

$$Cov^{\alpha}(t_1, t_2) = \frac{T_{\alpha} - |t_2 - t_1|}{T_{\alpha} + \frac{1}{\lambda}} \frac{\Omega^2}{2} \cos\left(\omega(t_2 - t_1)\right) 1_{[-T_{\alpha}, T_{\alpha}]}(t_2 - t_1). \tag{3.17}$$

This can be written as

$$Cov^{\alpha}(t_1, t_2) = \gamma(\lambda, T_{\alpha})\frac{\Omega^2(T_{\alpha} - |t_2 - t_1|)}{T_{\alpha}} 1_{[-T_{\alpha}, T_{\alpha}]}(t_2 - t_1)\frac{1}{2}\cos\left(\omega(t_2 - t_1)\right), \tag{3.18}$$

where

$$\gamma(\lambda, T_{\alpha}) = \frac{T_{\alpha}}{T_{\alpha} + \frac{1}{\lambda}}. \tag{3.19}$$

One can regard (3.18) as the product of three factors: the relative amount of alpha waves $\gamma(\lambda, T_{\alpha})$, the convolution of the amplitudes and the covariance of the OAM with unit amplitude.

In sum, the total PoMAM covariance is

$$Cov(t_1, t_2) = \gamma(\lambda, T_{\alpha})\frac{\Omega^2(T_{\alpha} - |t_2 - t_1|)}{T_{\alpha}} 1_{[-T_{\alpha}, T_{\alpha}]}(t_2 - t_1)\frac{1}{2}\cos\left(\omega(t_2 - t_1)\right) + \sigma^2 e^{-\kappa|t_2 - t_1|}. \tag{3.20}$$

Note that this expression is entirely parametric. Furthermore, the covariance model in equation (3.20) is stationary, i.e. only depends on lag $t_2 - t_1$.

For a more general envelope function of the waves $\Phi(t)$ the covariance can be calculated in a similar way. The convolution of the amplitudes then becomes

$$\frac{1}{T_{\alpha}}\int_0^{T_{\alpha}} \Phi(s)\Phi(s + |t_2 - t_1|)ds \tag{3.21}$$

and the equivalent of (3.20) becomes

$$Cov(t_1, t_2) = \gamma(\lambda, T_{\alpha})\frac{1}{T_{\alpha}}\int_0^{T_{\alpha}} \Phi(s)\Phi(s + |t_2 - t_1|)ds\frac{1}{2}\cos\left(\omega(t_2 - t_1)\right) + \sigma^2 e^{-\kappa|t_2 - t_1|}. \tag{3.22}$$

An example of the PoMAM with fixed amplitude, similar to the OAM example presented in Figure 3.3, is plotted in Figure 3.5. This figure shows that the PoMAM covariance indeed vanishes for big time lags as is the case for the experimental data in Figure 3.1(a).
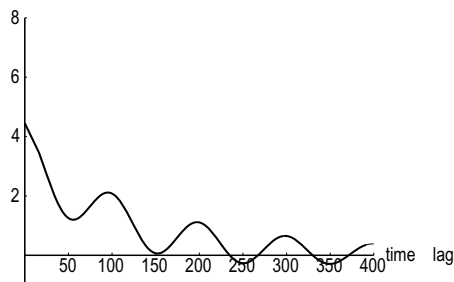
Figure 3.5: *The covariance in $(fT)^2$ of the PoMAM with fixed amplitude as function of time lag (ms) for $T_\alpha = 600$ ms, $\lambda = \frac{1}{0.6}$ s^{-1}, $\omega = 20\pi$ rad s^{-1}, $\Omega^2 = 3$ $(fT)^2$, $\sigma^2 = 3.5$ $(fT)^2$ and $\kappa = 10$ s^{-1}.*

3.2.3 Baseline Correction

Due to external influences in MEG/EEG measurements the baselines of the single channel signals are usually shifted over an unknown offset which can be quite large. To correct for these shifts one has to carry out an offset removal. One standard way of performing this baseline correction (BC) is subtracting per channel the average over a pre-stimulus interval. In this section the influence of this correction on the temporal covariance matrix is studied.

Let $[t_0 - T_c, t_0]$ be the interval over which the correction is calculated. Then the formula for the corrected noise of the kth trial, $\varepsilon_k^c(t)$, is

$$\varepsilon_k^c(t) = \varepsilon_k(t) - \frac{1}{T_c} \int_{t_0 - T_c}^{t_0} \varepsilon_k(t')dt'. \tag{3.23}$$

To compute the (co)variance of the corrected signal, the *corrected (co)variance*, one has to calculate

$$\begin{aligned}
Cov^c(t_1, t_2) &= E(\varepsilon_k^c(t_1) \cdot \varepsilon_k^c(t_2)) \\
&= Cov(t_1, t_2) - \frac{1}{T_c} \int_{t_0 - T_c}^{t_0} [Cov(t_1, t') + Cov(t_2, t')]dt' \\
&\quad + \frac{1}{T_c^2} \int_{t_0 - T_c}^{t_0} \int_{t_0 - T_c}^{t_0} Cov(t', t'')dt'dt''. \tag{3.24}
\end{aligned}$$

The first term in (3.24) is the stationary uncorrected covariance and the last term is a constant dependent on T_c and t_0. The second term, though, is in general not stationary.

In Figure 3.6 the effect of the BC is illustrated. If a signal contains alpha activity having nonzero average over the BC-window, a vertical shift is introduced by the correction. Moreover this introduced variance varies with time: the variation in signals is periodically bigger or smaller. If the BC-window equals one alpha period, the average alpha activity will be zero and this oscillating additional variance will not occur.
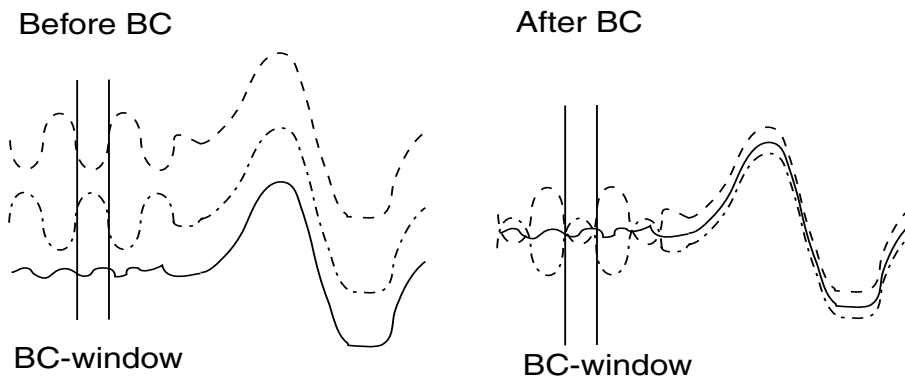
Figure 3.6: *The effect of the baseline correction on three signals (solid line: without alpha activity in BC-window; dashed and dot-dashed lines: with alpha activity in BC-window). When the average alpha activity during the BC-window is not zero the BC introduces a vertical shift in the signal (dashed and dot-dashed lines).*

For the OAM formula (3.24) can be computed (partly) analytically using (3.10) (see Appendix 3.5.3):

$$Cov^c(t_1, t_2) =$$

$$\Omega^2 \Big[\frac{1}{\omega^2 T_c^2}(1 - \cos(\omega T_c)) + \frac{1}{2}\cos(\omega(t_2 - t_1))$$

$$- \frac{2}{\omega T_c}\sin\frac{\omega T_c}{2}\cos(\frac{\omega(t_2 - t_1)}{2})\cos(\omega(\frac{t_1 + t_2}{2} - t_0 + \frac{T_c}{2}))\Big]$$

$$+ \sigma^2 e^{-\kappa|t_2 - t_1|} - \frac{\sigma^2}{T_c}\int_{t_0 - T_c}^{t_0}(e^{-\kappa|t_1 - t'|} + e^{-\kappa|t_2 - t'|})dt' + \frac{\sigma^2}{T_c^2}\int_{t_0 - T_c}^{t_0}\int_{t_0 - T_c}^{t_0}e^{-\kappa|t' - t''|}dt'dt''.$$

$$(3.25)$$

As an example the corrected variance ($t_1 = t_2 = t$) is plotted in Figure 3.7 for three different values of T_c for both the OAM (Figure 3.7(a)) and for empirical data (Figure 3.7(b)). It is clear from these figures that the corrected temporal covariance structure is not stationary and that this nonstationarity highly depends on the choice of baseline correction. Although there are differences between the OAM figure and the empirical figure the similarity in effect of baseline correction is striking.

There is an α-oscillation in the variance for $T_c = 25$ ms (dashed line) and $T_c = 50$ ms (dotted line), but this oscillation vanishes for $T_c = 100$ ms (solid line). This is clarified by the term within square brackets of formula (3.25): the amplitude of this sinusoid in $\frac{t_1 + t_2}{2}$ contains the factor $\sin\frac{\omega T_c}{2}$, which is in the case of Figure 3.7(a) equal to $\sin\frac{T_c \pi}{100}$ (taking time in ms) and is zero for $T_c = 100$ ms. Furthermore, the phase of this oscillation is $-t_0 + \frac{T_c}{2}$, so the phase shift between the dashed and the dotted lines is 12.5 ms. The average variance is minimum for the solid line in this figure ($T_c = 100$ ms) since then $\frac{1}{\omega^2 T_c^2}(1 - \cos(\omega T_c)) = \frac{1}{4\pi^2}(1 - \cos\frac{\pi 100}{50}) = 0$, see (3.25).

In all cases, the variance is minimum at $t = -\frac{T_c}{2}$, halfway the correction window. This drop is caused by the second part of formula (3.25). Taking $t_1 = t_2 = t$, this corrected variance due to $\eta(t)$, consists of a constant, a nonstationary term and another
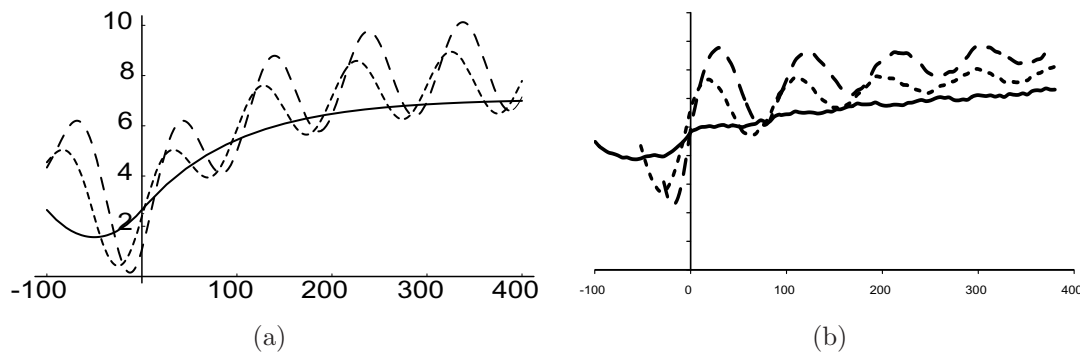
Figure 3.7: *Corrected OAM variance and corrected empirical variance for three different values of T_c as function of time. The horizontal axis indicates time in ms. $t_0 = 0$ ms. (a): corrected OAM variance as function of t with $\omega = 20\pi$ rad s^{-1}, $\Omega^2 = 2$ $(fT)^2$, $\sigma^2 = 3.5$ $(fT)^2$, $\kappa = 10$ s^{-1} for different values of T_c. (b): corrected empirical variances. The dashed line corresponds to $T_c = 25$ ms, the dotted line to $T_c = 50$ ms and the solid line to $T_c = 100$ ms in both figures.*

constant. The integrand in the middle term is larger for t closer to the correction window, because then $|t - t'|$ is smaller. Therefore the closer t to this window, the larger the integral. Together with the minus sign in front of the term, this causes the drop. Moreover a decrease in T_c increases the magnitude of the drop (see Appendix 3.5.4).

To obtain stationarity one should choose T_c in such a way that the oscillation vanishes and the drop is minimum. This is achieved when $\omega T_c = 2l\pi$, $l \in \mathbb{N} = \{1, 2, 3, \ldots\}$ i.e. when the correction interval is taken to be l α-periods with $l \in \mathbb{N}$. The bigger l, the smaller the magnitude of the drop, thus the more stationary the matrix. To keep the preprocessing feasible taking one or two alpha periods as BC window is adequate.

The covariance of the corrected error $\varepsilon_k^c(t)$ in the PoMAM is calculated using (3.20) and (3.24). For the PoMAM with fixed amplitude formula (3.24) becomes a parametric representation for the entries of the temporal covariance matrix T. The main difference between the simple OAM model and this parametric PoMAM is the decreasing amplitude of the oscillations in the (co)variance. The effect of the baseline correction on the stationarity is the same for both models. Therefore it is sufficient to examine formula (3.25) for the simpler OAM in stead of the more complicated formula (3.24) to investigate stationarity of the PoMAM.

3.3   Results

The PoMAM was fitted to ML-estimated temporal covariance matrices of five subjects to see how well this parametric model describes the abstract and nonphysiological ML-estimate. The ML-estimates were obtained from MEG data of Somatosensory Evoked Field experiments, where the left median nerve was stimulated. Data were acquired on a 151 channel CTF Omega system at a sample rate of 2083 Hz. The number of

trials was approximately 500 for all subjects. No filtering was applied, expect for the baseline correction. Subjects 2 and 3 were stimulated at a constant stimulus rate of 1 Hz, while the inter stimulus interval in subjects 1, 4 and 5 varied uniformly between 800 ms and 1200 ms. Regarding this difference, best fitting parameter values appeared to reproduce very well within subjects between data sets with random and regular stimulation. Therefore the comparison of parameter values for subjects 2 and 3 with those for the other subjects is justified.

The parameters λ (the intensity of the Poisson process) and Ω (the amplitude of the waves) were fitted simultaneously in $\tilde{\Omega}^2 \equiv T_\alpha \Omega^2 (T_\alpha + \frac{1}{\lambda})^{-1}$ since they cannot be distinguished in the covariance formula (3.18). Furthermore, an additional term σ_{hf}^2 was added to the main diagonal and $\frac{1}{2}\sigma_{hf}^2$ to the first subdiagonal to model high frequency noise due to the omitted filtering. The parameters to be fitted are ω (alpha frequency), T_α (duration of α-wave), κ (covariance length of additional noise), σ^2 (variance of additional (low frequency) noise), σ_{hf}^2 (variance due to high frequency noise) and $\tilde{\Omega}^2$ (representing amplitude and intensity of alpha activity).

The cost function used is

$$C(\omega, T_\alpha, \kappa, \tilde{\Omega}^2, \sigma^2, \sigma_{hf}^2) = \frac{\|T_{(\omega, T_\alpha, \kappa, \tilde{\Omega}^2, \sigma^2, \sigma_{hf}^2 \mid t_0, T_c)} - \hat{T}\|_F^2}{\|\hat{T}\|_F^2} \times 100\% \qquad (3.26)$$

i.e. the relative squared Frobenius norm of the difference between the model matrix T and the ML-estimated matrix $\hat{T}$. To minimise the cost function (3.26) the Simplex method is used. Using this This iterative method was only used to estimate the three nonlinear parameters (ω, T_α, κ), while the three linear parameters ($\tilde{\Omega}^2$, σ^2, σ_{hf}^2) were fitted in a least squares sense in each iteration. For each subject, the values for t_0 and T_c were taken equal to the corresponding values used in the ML-estimation of T. In this study, the ML matrices were explicitly not taken as stationary as possible, though instead, it was attempted to approximate the ML-estimated matrices as accurately as possible, irrespective of the correction window used. Moreover, in this way, it can be verified whether the effect of the baseline correction is taken into account correctly.

Table 3.1 presents the best fitting parameter values as well as the cost function values for the five subjects. The nonlinear parameters T_α and κ appeared to be rather insensitive in this study, while the parameter ω was most sensitive. The table shows that the PoMAM describes the ML-estimated temporal covariance structure accurately: the ML-estimates can be approximated up to an error of less than one percent in relative squared Frobenius norm by the PoMAM.

3.4   Discussion

The temporal covariance of the background noise in MEG/EEG measurements can be described accurately by the parametric Poisson modulated alpha model (PoMAM). In the PoMAM the temporal noise is modelled in a physiological way as the sum of randomly occurring alpha waves and additional noise. In principle, this model is stationary, though in practise the temporal stationarity is destroyed by the baseline correction, which is apparent from Figure 3.6 and formulas (3.24) and (3.25). Taking the offset

subject	stim	T_c	$\frac{\omega}{2\pi}$	T_α	$\frac{1}{\kappa}$	$\tilde{\Omega}^2$	σ^2	σ_{hf}^2	C
1	random	48	10.61	287	22.9	19980	13610	21099	0.4 %
2	regular	25	9.85	359	73.8	23527	15382	22952	0.5 %
3	regular	50	11.03	289	16.3	22651	12274	16804	0.7 %
4	random	50	8.85	536	14.7	8497	10402	11787	0.6 %
5	random	49	11.92	61	85.3	24746	11746	17293	0.4 %
average			10.45	306	42.6	19880	12683	17987	0.58 %
unit		ms	Hz	ms	ms	$(\text{fT})^2$	$(\text{fT})^2$	$(\text{fT})^2$	

Table 3.1: *Best fitting values (between double bars) for the six parameters of the Poisson modulated alpha model together with the residual error for five different subjects. The number of time samples is 500. The values for t_0 and T_c were taken equal to the corresponding values in the ML-estimated matrices. Subjects 2 and 3 had eyes open, while the others had eyes closed.*

correction into account correctly, the ML-estimated temporal covariance matrix can be described by the PoMAM up to an error of less than one percent (Table 3.1) using only six parameters. It is, however, not straightforward how to estimate the PoMAM parameters based on the raw data.

Some parameters (especially T_α and κ) appeared to be rather insensitive. This redundancy suggests that a model with even less parameters is possible. Bearing this insensitivity in mind, it may well be possible to compose a *standard* PoMAM matrix T. Such a standard matrix can be based on standard values for T_α, κ, $\tilde{\Omega}$, σ and σ_{hf}, while it should still depend on the sensitive parameter ω). It may be expected that this matrix will perform considerably better in source localisation than the commonly used $T = \sigma^2 I$ does.

The temporal stationarity of the background noise highly depends on the window used for the baseline correction (see equation (3.25)). Taking the length of this window to be equal to a multiple of alpha periods one obtains the most stationary temporal covariance. However, if one is interested in a particular sample after the stimulus (e.g. the N20 response in a SEF experiment) advantage can be taken of the nonstationarity due to the BC. In Figure 3.8 the empirical variances at $t = 20$ ms and $t = 60$ ms after the stimulus in a SEF data set are plotted for several values of T_c. This figure shows that for the N20 response the optimal BC-length is 100 ms, whereas for $t = 60$ ms a BC-window of 70 ms yields the minimum variance. This can be explained by formula (3.25): if the time sample t and the PoMAM parameters are substituted then this formula becomes a parametric expression in t_0 and T_c, which can be minimised with respect to T_c, taking $t_0 = 0$.

In source localisation analyses where no temporal correlations are estimated or are taken into account, in other words, where it is assumed that $T = \sigma^2 I$ or equivalently, that $Cov(t_1, t_2) = \sigma^2 \delta(t_1 - t_2)$, the baseline correction also alters the temporal covariance matrix. Due to the correction a positive constant is added to all entries in T (see (3.24)), and, therefore, it is inconsistent to fix the baseline corrected temporal covariance matrix to $\sigma^2 I$ in such a case.
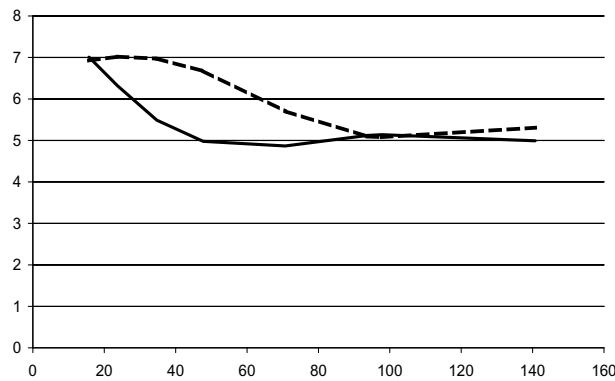
Figure 3.8: *The empirical variance in a SEF data set at $t = 20$ ms (dashed line) and $t = 60$ ms (solid line) post-stimulus for different values of T_c (horizontal axis in ms), the length of BC-window; $t_0 = 0$.*

In section 1.4.2, chapter 2 and [101], it was shown that trial-to-trial variations in the response cause nonstationarities in the background noise when trials are averaged according to the Signal Plus Noise model. In the current SEF experiment the assumption of the SPN model leads to temporally stationary background noise; the observed nonstationarity in the matrix is caused by the preprocessing and does not originate from the data (cf. [65]). Therefore, taking into account that in this study, contrary to others, both spatial and temporal correlations are incorporated, the present study not disapprove the SPN model.

In the PoMAM the alpha activity is modelled as a random phase process with constant amplitude. In case of phase locking of the alpha activity (see section 1.4.2), the phase of α-waves would be modulated by the stimulus. In that case, a different probability density function for the phase τ_k is more likely, as well as a different amplitude time function for the waves as in (3.22) (cf. [65]). Nonetheless, the small fit error of the PoMAM shows that the current data can be well explained without the assumption of such a stimulus modulated alpha model.

In sum, the Poisson modulated alpha model is a mathematical model, describing the temporal noise covariance in a physiological and accurate way. For practical application of this model further study of the parameter estimation and the effect on source localisation is needed.

3.5   Appendix

3.5.1   Probability t_1 and t_2 in the same wave

In this appendix, the probability that the alpha process with stochastic intermediate time intervals is such that t_1 and t_2 occur in one and the same alpha wave is computed.

For $|t_2 - t_1| > T_\alpha$, this probability is obviously zero, since no wave can contain both instants in that case. For $|t_2 - t_1| < T_\alpha$ this probability is derived below, assuming $0 < t_1 < t_2$. For $0 < t_2 < t_1$, the result is stated at the end of this appendix.
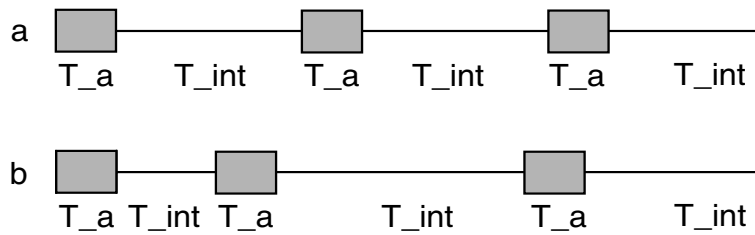
Figure 3.9: *Examples of the alpha process with fixed intermediate time (a) and varying intermediate time (b).*

Assume t_1 and t_2 are given, such that $0 < t_1 < t_2$ and $t_2 - t_1 < T_\alpha$. If the intermediate time between waves were fixed to T_{int}, the durations of the consecutive alpha waves and their intermediate times would be fixed over time (see Figure 3.9(a)). In that case, the probability that this disrupted alpha process is such that both instants t_1 and t_2 occur in one wave would be equal to the ratio

$$\frac{T_\alpha - (t_2 - t_1)}{T_\alpha + T_{int}}. \tag{3.27}$$

One can regard (3.27) as the ratio between favourable phases of the process and all possible phases.

Assuming the intermediate time intervals T_{int} have the Exponential(λ) distribution in stead of being fixed, this probability changes. The probability that needs to be calculated is now an integral over all possible values of T_{int}. Given $T_{int} = \zeta$ (and t_1 and t_2 are in the wave and intermediate time combination corresponding to ζ), the probability that t_1 and t_2 both are in the wave, is

$$\frac{T_\alpha - (t_2 - t_1)}{T_\alpha + \zeta}. \tag{3.28}$$

This probability needs to be integrated over all ζ, multiplied by the probability that $T_{int} = \zeta$. Moreover, the integrand needs to be weighted with the length $T_\alpha + \zeta$ relative to all possible lengths. This weighting is necessary because random selection of t_1 and t_2 is less likely occur in a combination of $T_\alpha + 100$ms than $T_\alpha + 1000$ms, just because the latter combination spans a longer time interval (see Figure 3.9(b)). Hence the integrand becomes

$$\frac{T_\alpha - (t_2 - t_1)}{T_\alpha + \zeta} \times p_{T_{int}}(\zeta) \times \frac{T_\alpha + \zeta}{\int_0^\infty (T_\alpha + \theta)\lambda e^{-\lambda\theta}d\theta} = \frac{T_\alpha - (t_2 - t_1)}{T_\alpha + \frac{1}{\lambda}} \times p_{T_{int}}(\zeta). \tag{3.29}$$

Finally, the probability of 't_1 and t_2 occur in the same wave' is

$$P = \int_0^\infty \frac{T_\alpha - (t_2 - t_1)}{T_\alpha + \frac{1}{\lambda}} \times \lambda e^{\lambda\zeta}d\zeta = \frac{T_\alpha - (t_2 - t_1)}{T_\alpha + \frac{1}{\lambda}}. \tag{3.30}$$

For arbitrary values of t_1 and t_2 the probability is

$$P(t_1 \text{ and } t_2 \text{ in same wave}) = \frac{T_\alpha - |t_2 - t_1|}{T_\alpha + \frac{1}{\lambda}} 1_{[-T_\alpha, T_\alpha]}(t_2 - t_1). \tag{3.31}$$

3.5.2   Computation of equation (3.16)

In this appendix the last term of (3.16):

$$E(\varepsilon_k^\alpha(t_1)\varepsilon_k^\alpha(t_2)|t_1 \text{ and } t_2 \text{ in same wave})$$

is calculated. From (3.12) it is clear that the onset of the wave in which t_1 and t_2 occur determines the amplitude of the activity at both instants. Therefore all possible onsets have to be taken into account.

First assume $t_1 < t_2$. Given the event 't_1 and t_2 are in the same wave' the stochast $S = t_1 - t_{onset}$ has a uniform distribution in $[0, T_\alpha - (t_2 - t_1)]$. Using the formula for the expected value of a function $g(S)$ of a stochast S with probability density function $f_S(s)$ [11]

$$E(g(S)) = \int g(s)f_S(s)ds \tag{3.32}$$

with $g(S) = \varepsilon_k^\alpha(t_1)\varepsilon_k^\alpha(t_2)$ yields

$$E(\varepsilon_k^\alpha(t_1)\varepsilon_k^\alpha(t_2)|t_1 \text{ and } t_2 \text{ in same wave}, t_1 < t_2)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\Bigg[\frac{1}{T_\alpha - (t_2 - t_1)}$$
$$\int_0^{T_\alpha-(t_2-t_1)} 1_{[0,T_\alpha]}(s)\Omega\sin(\omega t_1 + \tau_k)1_{[0,T_\alpha]}(s + t_2 - t_1)\Omega\sin(\omega t_2 + \tau_k)ds\Bigg]d\tau_k$$
$$= \frac{\Omega^2}{2}\cos(\omega(t_2 - t_1)). \tag{3.33}$$

Taking t_1 and t_2 arbitrary in the same wave this remains

$$E(\varepsilon_k^\alpha(t_1)\varepsilon_k^\alpha(t_2)|t_1 \text{ and } t_2 \text{ in same wave }) = \frac{\Omega^2}{2}\cos(\omega(t_2 - t_1)). \tag{3.34}$$

3.5.3   Corrected covariance for OAM

For the simple OAM the corrected covariance (3.24) is calculated using the covariance for the uncorrected case (3.10):

$$Cov^c(t_1, t_2) =$$
$$= Cov(t_1, t_2) - \frac{1}{T_c}\int_{t_0-T_c}^{t_0}[Cov(t_1, t') + Cov(t_2, t')]dt' + \frac{1}{T_c^2}\int_{t_0-T_c}^{t_0}\int_{t_0-T_c}^{t_0}Cov(t', t'')dt'dt''$$
$$= \frac{\Omega^2}{2}\Bigg[\cos(\omega(t_2 - t_1)) - \frac{1}{T_c}\int_{t_0-T_c}^{t_0}[\cos(\omega(t_1 - t')) + \cos(\omega(t_2 - t'))]dt'$$
$$+ \frac{1}{T_c^2}\int_{t_0-T_c}^{t_0}\int_{t_0-T_c}^{t_0}\cos(\omega(t' - t''))dt'dt''\Bigg]$$
$$+ \sigma^2\Bigg[e^{-\kappa|t_2-t_1|} - \frac{1}{T_c}\int_{t_0-T_c}^{t_0}e^{-\kappa|t_1-t'|} + \sigma^2 e^{-\kappa|t_2-t'|}dt' + \frac{1}{T_c^2}\int_{t_0-T_c}^{t_0}\int_{t_0-T_c}^{t_0}e^{-\kappa|t'-t''|}dt'dt''\Bigg] =$$
$$= \frac{\Omega^2}{2}\Bigg[\cos(\omega(t_2 - t_1)) + \frac{1}{\omega T_c}\big(\sin(\omega(t_1 - t_0)) - \sin(\omega(t_1 - t_0 + T_c))\big)$$

$$+\frac{1}{\omega T_c}\left(\sin\left(\omega(t_2-t_0)\right)-\sin\left(\omega(t_2-t_0+T_c)\right)\right)+\frac{1}{\omega^2 T_c^2}\left(2-2\cos\left(\omega T_c\right)\right)\Big]$$

$$+\sigma^2\left[e^{-\kappa|t_2-t_1|}-\frac{1}{T_c}\int_{t_0-T_c}^{t_0}e^{-\kappa|t_1-t'|}+e^{-\kappa|t_2-t'|}dt'+\frac{1}{T_c^2}\int_{t_0-T_c}^{t_0}\int_{t_0-T_c}^{t_0}e^{-\kappa|t'-t''|}dt'dt''\right].$$

$$(3.35)$$

Applying successively the rules

$$\sin x-\sin\left(x+a\right)\quad=\quad-2\sin\frac{a}{2}\cos\left(x+\frac{a}{2}\right) \tag{3.36}$$

$$\cos x+\cos\left(x+a\right)\quad=\quad 2\cos\frac{a}{2}\cos\left(x+\frac{a}{2}\right) \tag{3.37}$$

to the first part of this formula yields (3.25).

3.5.4   Influence baseline correction on variance

The drop in the variance over the correction window (Figure 3.7(a)) is caused by the next to last term in (3.25) (substituting $t_1=t_2=t$):

$$-\frac{2\sigma^2}{T_c}\int_{t_0-T_c}^{t_0}e^{-\kappa|t-t'|}dt'. \tag{3.38}$$

In this appendix it will be proved that a smaller value of T_c (i.e. a shorter correction window) yields a deeper drop. Consider the magnitude of the drop, M, at the minimum point (halfway the correction interval) as function of T_c by setting $t=t_0-\frac{T_c}{2}$ in (3.38):

$$M(T_c)=\frac{2\sigma^2}{T_c}\int_{t_0-T_c}^{t_0}e^{-\kappa|t_0-\frac{T_c}{2}-t'|}dt'. \tag{3.39}$$

This function is positive and it has to be proven that it is decreasing in T_c (i.e. having negative slope), which means that a shorter correction interval yields a deeper drop. To prove $M'(T_c)<0$, first the integral in (3.39) is expressed as

$$M(T_c)=\frac{4\sigma^2}{\kappa T_c}(1-e^{-\kappa\frac{T_c}{2}}) \tag{3.40}$$

and then the derivative is taken with respect to T_c:

$$M'(T_c)=\frac{-4\sigma^2}{\kappa T_c^2}+\frac{4\sigma^2}{\kappa T_c^2}e^{-\kappa\frac{T_c}{2}}+\frac{2\sigma^2}{T_c}e^{-\kappa\frac{T_c}{2}}. \tag{3.41}$$

To prove that the expression in (3.41) is negative it is multiplied by the positive factor $\frac{\kappa T_c^2}{4\sigma^2}$ and $s=\kappa\frac{T_c}{2}$ is substituted

$$\frac{\kappa T_c^2}{4\sigma^2}M'(T_c)=-1+(1+s)e^{-s}. \tag{3.42}$$

Note that s only takes strictly positive values. Now the last expression in (3.42) is always negative because

$$-1+(1+s)e^{-s}<0\Longleftrightarrow 1+s<e^s \tag{3.43}$$

which is true for all $s>0$. This completes the proof: a shorter correction interval (i.e. a smaller value of T_c) yields a deeper drop in the variance. Therefore the larger the baseline correction window the more stationary the covariance matrix.

Chapter 4

The spatiotemporal MEG covariance matrix modelled as a sum of Kronecker products

Adapted from: F. Bijma, J.C. de Munck, R.M. Heethaar: *The spatiotemporal MEG covariance matrix modelled as a sum of Kronecker products*, NeuroImage 2005, *in press*

4.1 Introduction

In MEG measurements, background noise is correlated both in space and in time. Although these correlations are unknown a priori, they are of interest for two reasons: they contain physiological information and they can be used to improve source localisation [16, 61, 96, 109]. These spatiotemporal correlations can be estimated from the measured data. The general spatiotemporal covariance matrix, however, has a large dimension, yielding two main problems, the first of which being its estimation and the second its storage. Estimation would require an unrealistically high number of measurements to achieve nonsingularity and storage would require far more memory than commonly available.

A way of resolving these two problems has been found in the parametrisation of the spatiotemporal covariance matrix through a Kronecker product (KP) [56, 104] of a spatial and a temporal covariance matrix, reducing its dimensionality considerably [16, 20, 42], see equation (1.54). The KP parametrisation assumes that an arbitrary spatiotemporal correlation can be modelled as a product of a spatial and a temporal factor. These two factors are independent of each other; hence, the spatial and temporal correlations are separated from each other in the KP model. Physiologically, this model can be interpreted by assuming background noise to be generated by randomly distributed dipolar sources having amplitude functions independent of the source locations [16].

Applications of the KP model in source localisation methods have revealed that the accuracy improves when the spatiotemporal correlations are taken into account in stead of no or only spatial correlations [20, 42]. Nevertheless, there are two important shortcomings of the KP model. The first deficiency is the rigidness of the KP: the

shape of the temporal cross spectrum is forced to be fixed over all channels. This is a simplification, as is illustrated by the alpha rhythm: the amount of alpha activity relative to other spontaneous activity is not equally distributed over the head. The second point of debate is trial-to-trial variations, which have been discussed in literature [14, 26, 29, 44, 57, 66, 73, 85]. The Signal Plus Noise (SPN) model in evoked field experiments assumes that no such variations occur in the data (section 1.3.1). This model can be formulated as

$$r_k^{i,j} = r^{i,j} + \varepsilon_k^{i,j}, \tag{4.1}$$

where $r_{i,j}^k$ is the measurement in trial k at the ith sensor and j^{th} time instant, $r_{i,j}$ is the trial independent response to the stimulus at the ith sensor and j^{th} time instant and $\varepsilon_{i,j}^k$ is the residual ('noise'). Hence, the SPN model assumes that all trial-to-trial variations are accounted for by the $\varepsilon_{i,j}^k$. The response $r_{i,j}$ is estimated by the average over trials of the measurements $r_{i,j}^k$. When incorrectly assumed, i.e. when $r_{i,j}$ does depend on k, the SPN model leads to nonstationarity in the temporal covariance [101] (section 1.4.2). Nonetheless, it has been shown for Somatosensory Evoked Field data sets that both the temporal (chapter 3) and the spatial covariance [16] estimated under the SPN assumption can be explained by a stationary model. The reason may be that the nonstationarities are suppressed by the rigidness of the KP model: if the majority of the channels show temporally stationary signals, this stationarity will dominate the temporal matrix. To overcome its shortcomings, an extension of the KP model is investigated in this chapter: a sum of Kronecker products.

In the sum of KP model, each term presents a combination of a spatial and a temporal pattern. Unlike the single KP model, the sum model allows for multiple temporal structures with specific spatial patterns, and can, thus, account for temporal nonstationarities in separate terms. The interpretation of the sum of KP model is analogous to that of the single KP model. Assuming that spontaneous background activity is the composite of a number of independent ongoing processes, cf. [58], each of which can be described by a random dipole model as explained above, the spatiotemporal covariance matrix becomes a sum of KP.

Although extending the single KP model to a sum of KP may seem rather straightforward, in the practical application in dipole localisation it becomes quite delicate. The main problem is the inversion of a sum of Kronecker products, which, contrary to a single KP, cannot be performed by inversion of only the smaller dimensional matrices, but requires inversion in the large spatiotemporal dimension. Therefore, the emphasis in this study is primarily on estimating rather than applying the sum of KP. The aim is twofold: firstly, the estimated sum of KP contains information about the validity of the single KP model for dipole localisation; secondly, from the estimated sum physiological information is assembled about the spatial and temporal features in the background activity. Hence, the goal of this chapter is not to present an improved method for source localisation; the accent is on investigating the spatiotemporal MEG residuals.

In the next section, first the findings and formulas of the single KP model are summarised shortly and then the estimators for the sum of KP model are derived and discussed. In the third section the sum model is estimated for data sets of three types (VEF, SEF, AEF) and results are shown. In the final section the results are discussed and conclusions are drawn. The technical details of the model are put together in the

appendices in order to keep the text compact.

4.2 Model

4.2.1 Single KP model

In the single Kronecker product model, the covariance between two MEG residuals, $e_k^{i,j}$ and $e_{k'}^{i',j'}$, is modelled as the product of a temporal and a spatial term (equation 1.52):

$$\mathcal{E}(\varepsilon_{i,j}^k, \varepsilon_{i',j'}^{k'}) = X_{i,i'} T_{j,j'} \delta_{k,k'} \tag{4.2}$$

where $\varepsilon_k^{i,j}$ is the MEG residual measured at sensor i, time sample j in the kth trial and $\delta_{k,k'}$ denotes the Kronecker delta function (see definition 1.53). Thus, different trials are assumed to be independent. The meaning of equation (4.2) is that the temporal covariance matrix T is fixed in space and the spatial covariance matrix X does not vary over time. In other words, space and time are not correlated.

The matrix formula for the Kronecker product model is

$$\Sigma = T \otimes X, \tag{4.3}$$

where $T \in \mathbb{R}^{J \times J}$ is the temporal, $X \in \mathbb{R}^{I \times I}$ the spatial covariance matrix and $\Sigma \in \mathbb{R}^{IJ \times IJ}$ is the spatiotemporal covariance matrix. I denotes the number of sensors and J the number of time samples. The dimensions of these two covariance matrices are much smaller than the dimension of Σ, and by the structure of the Kronecker product the computations are much less demanding [63]:

$$\Sigma^{-1} = T^{-1} \otimes X^{-1} \tag{4.4}$$
$$\det(\Sigma) = \det(T)^I \det(X)^J. \tag{4.5}$$

X and T can be estimated using either the maximum likelihood (ML) paradigm or the least squares (LS) method. In the ML case, the MEG residuals are assumed to have a Gaussian distribution with the KP as the covariance matrix. The likelihood function is maximised with respect to the matrices X and T. In the LS case, the KP is fitted to the spatiotemporal sample covariance matrix and the difference in Frobenius norm is minimised with respect to X and T. The sample covariance matrix $\Sigma_s \in \mathbb{R}^{IJ \times IJ}$ is defined as

$$\Sigma_s \equiv \frac{1}{K-1} \sum_{k=1}^{K} vec(E^k)[vec(E^k)]^t \tag{4.6}$$

where $E_k \in \mathbb{R}^{I \times J}$ is the matrix containing the residuals of trial k,

$$(E^k)_{i,j} = e_{i,j}^k, \tag{4.7}$$

and K is the number of trials (repetitive measurements). In both the ML and the LS case, the estimators for X and T are given by an iterative system. For the ML model

this system is [20] (see section 1.4.3):

$$\hat{X}_{ML} \quad = \quad \frac{1}{JK} \sum_{k=1}^{K} E^k k \hat{T}_{ML}^{-1} (E^k)^t \tag{4.8}$$

$$\hat{T}_{ML} \quad = \quad \frac{1}{IK} \sum_{k=1}^{K} (E^k)^t \hat{X}_{ML}^{-1} E^k \tag{4.9}$$

and for the LS paradigm, see appendix 4.5.1,

$$\hat{X}_{LS} \quad = \quad \frac{1}{K-1} \frac{1}{\|\hat{T}_{LS}\|^2} \sum_{k=1}^{K} E^k \hat{T}_{LS} (E^k)^t \tag{4.10}$$

$$\hat{T}_{LS} \quad = \quad \frac{1}{K-1} \frac{1}{\|\hat{X}_{LS}\|^2} \sum_{k=1}^{K} (E^k)^t \hat{X}_{LS} E^k. \tag{4.11}$$

For the extension to the sum of KP model, the ML formulas become prohibitively complicated because the sum does not maintain the elegant structure for the inverse and the determinant as the single KP does (see equations (4.4) and (4.5)). For the LS paradigm it appears that extension is possible.

4.2.2   Sum KP model

As stated in the introduction, extending the single KP to a sum of KP allows for a more general spatiotemporal covariance structure. The sum model is expressed as

$$\Sigma = \sum_{n=1}^{N} T_n \otimes X_n \tag{4.12}$$

and the corresponding LS cost function is

$$\mathcal{C}_{LS} = \|\Sigma_s - \sum_{n=1}^{N} T_n \otimes X_n\|^2 \tag{4.13}$$

where $\|.\|^2$ denotes the Frobenius norm. In order to minimise the cost function, the algorithm presented by Van Loan to find the best sum of KP approximation to a given matrix is used [104]. The matrix elements in equations (4.12) and (4.13) are rearranged according to Van Loan's shuffle operator $\mathcal{S} : \mathbb{R}^{IJ \times IJ} \longrightarrow \mathbb{R}^{I^2 \times J^2}$ such that the model equation (4.12) is transformed into

$$\mathcal{S}(\Sigma) = \sum_{n=1}^{N} vec(X_n) vec(T_n)^t, \tag{4.14}$$

and the cost function in equations (4.13) becomes

$$\mathcal{C}_{LS} = \|\mathcal{S}(\Sigma_s) - \sum_{n=1}^{N} vec(X_n) vec(T_n)^t\|^2. \tag{4.15}$$

From appendix 4.5.2, containing more details about $\mathcal{S}$, the formula for the shuffled sample covariance matrix follows:

$$\mathcal{S}(\Sigma_s) = \frac{1}{K-1} \sum_{k=1}^{K} E^k \otimes E^k. \tag{4.16}$$

Note that the dimension of this shuffled matrix is $I^2 \times J^2$, which is in general not square.

LS-estimators for X_n and T_n, $n = 1, \ldots, N$, are obtained by minimising $\mathcal{C}_{LS}$ in equation (4.15). This minimisation is equivalent to finding the best rank N approximation of $\mathcal{S}(\Sigma_s)$, which can be obtained from the Singular Value Decomposition (SVD) of $\mathcal{S}(\Sigma_s)$. Write the SVD of $\mathcal{S}(\Sigma_s)$ [31, 63]

$$\mathcal{S}(\Sigma_s) = U\Delta V^t, \tag{4.17}$$

where $U \in \mathbb{R}^{I^2 \times I^2}$ and $V \in \mathbb{R}^{J^2 \times J^2}$ are orthogonal matrices and $\Delta = (\Delta_0, \mathbf{0}) \in \mathbb{R}^{I^2 \times J^2}$ and $\Delta_0 \in \mathbb{R}^{I^2 \times I^2}$ is diagonal. Here it is assumed that $I \leq J$. In case $I > J$ the expression for Δ becomes $\begin{pmatrix} \Delta_0 \\ \mathbf{0} \end{pmatrix}$, with $\Delta_0 \in \mathbb{R}^{J^2 \times J^2}$. The best rank N approximation of $\mathcal{S}(\Sigma_s)$, for $N \leq \min(I^2, J^2)$, is now given by

$$\sum_{n=1}^{N} U_n \sigma_n V_n^t \tag{4.18}$$

where U_n (V_n) denotes the nth column of U (V) and $\sigma_n = \Delta_{n,n}$, the $(n,n)^{th}$ entry of Δ_0. Hence, the estimators for $vec(X_n)$ and $vec(T_n)$ are given by

$$vec(\hat{X}_n) = U_n \tag{4.19}$$
$$vec(\hat{T}_n) = \sigma_n V_n \tag{4.20}$$

for $1 \leq n \leq N$. Note that these estimators are not unique: e.g. multiplying equation (4.19) and dividing equation (4.20) by the same constant yields an equivalent solution. Throughout this section, the normalisation as in equations (4.19) and (4.20) is used. It follows from equation (4.18) that the entire sample covariance matrix can be described by a sum of KP, when N is taken equal to $\min(I^2, J^2)$. Furthermore, from the σ_n the distribution of explained matrix power over the KP terms is obtained:

$$\text{rel pow } n^{th} \text{ term} = \frac{\sigma_n^2}{\|\mathcal{S}(\Sigma_s)\|^2} \times 100\% = \frac{\sigma_n^2}{\|\Sigma_s\|^2} \times 100\%. \tag{4.21}$$

Despite the straightforward application of Van Loan's method, the estimators in equations (4.19) and (4.20) are not convenient in practise, due to the dimensionality of the desired SVD. Therefore, alternative estimators are deduced below.

The alternative way of estimating the terms (T_n, X_n) in the KP sum uses Lagrange multipliers [68]. To find the best rank N approximation it suffices to first find the best rank 1 approximation and successively find all the subsequent terms one after another. The initial term, $n = 1$, corresponds to the best rank one approximation of the sample covariance matrix and is estimated by the system of equations (4.10) and (4.11). As

the higher order terms are estimated one by one, for any p satisfying $1 < p \leq N$, the first $p-1$ terms, $(\hat{T}_n, \hat{X}_n)$ for $n = 1, \ldots, p-1$, will have been estimated at the instant of estimation of the pth term. In other words, the best rank $p-1$ approximation of $\mathcal{S}(\Sigma_s)$ is known and the best rank p approximation has to be estimated. This step is explained in appendix 4.5.3 from which expressions for $\hat{X}_p$ and $\hat{T}_p$ follow:

$$\hat{X}_p = \frac{1}{(K-1)\|\hat{T}_p\|^2} \sum_{k=1}^{K} \left[E^k \hat{T}_p (E^k)^t - \sum_{n=1}^{p-1} tr[\hat{X}_n E^k \hat{T}_p (E^k)^t] \hat{X}_n \right] \quad \text{for } 1 < p \leq N \quad (4.22)$$

$$\hat{T}_p = \frac{1}{K-1} \sum_{k=1}^{K} \left[(E^k)^t \hat{X}_p E^k - \sum_{n=1}^{p-1} \frac{1}{\|\hat{T}_n\|_F^2} tr[\hat{T}_n (E^k)^t \hat{X}_p E^k] \hat{T}_n \right] \quad \text{for } 1 < p \leq N. \quad (4.23)$$

The starting value for this iterative system and for the system in equations (4.10) and (4.11) is to set $T_p = \mathbf{I}_J$ and start with updating X_p. The iteration stops when the relative difference in matrix power between an estimate and the next estimate is less than 10^{-12}, i.e. when $\frac{\|A^s - A^{s+1}\|_F^2}{\|A^s\|_F^2} < 10^{-12}$ for both $A = T_p$ and $A = X_p$, where A^s indicates the sth estimate in the iteration. To verify the iterative estimators in equations (4.22) and (4.23), the SVD estimation method in equations (4.19) and (4.20) was applied to small data sets ($I = 30$, $J = 30$). For this small data sets the first four iteratively estimated KP terms were compared to the estimated terms from the SVD. This comparison showed that both methods yielded identical matrices.

Note that, although the cost function $\mathcal{C}_{LS}$ in equation (4.15) is expressed in terms of the spatiotemporal sample covariance matrix Σ_s, the solution in the iterative system of equations (4.22) and (4.23) does not require the storage of this huge matrix Σ_s in memory. In order to find the solution, apart from the recorded data, one only needs to store the terms (X_n, T_n) in memory, which is of order $(I^2 + J^2)$.

4.2.3   Rewriting the sum of 2 Kronecker products

The terms in the sum of KP are estimated under the constraint of orthogonal 'vecced' matrices, see equations (4.54) and (4.55) of appendix 4.5.3. In order to interpret the matrices (X_n, T_n) in each term as covariance matrices of the underlying physiological processes, these matrices should be positive definite. However, the orthogonality constraint forces the higher order terms to be *indefinite* matrices. This can be explained by the following reasoning. The first term consists of two positive definite matrices, T_1 and X_1, representing the best rank 1 approximation of $\mathcal{S}(\Sigma_s)$. Therefore, there exists a nonsingular matrix $W_1 \in \mathbb{R}^{J \times J}$ such that $T_1 = W_1 W_1^t > 0$. To show that any higher order temporal matrix T_n for $n > 1$ must be indefinite, it is demonstrated below that the assumption of T_n being positive (or negative) definite leads to a contradiction. Assuming T_n to be positive definite implies that there exists a matrix $W_n \in \mathbb{R}^{J \times J}$ such that $T_n = W_n W_n^t > 0$. Substituting W_1 and W_n, the orthogonality constraint can be written as

$$0 = vec(T_1)vec(T_n)^t = tr(T_1 T_n) =$$

$$tr(W_1 W_1^t W_n W_n^t) = tr(W_1^t W_n W_n^t W_1) = \|W_1^t W_n\|^2 > 0 \quad (4.24)$$

which is a contradiction. For T_n negative definite, a similar contradiction can be derived. Hence, T_n must be indefinite. The same holds true for all the higher order spatial matrices X_n. The Kronecker product of two indefinite matrices, A and B, is again indefinite, because the eigenvalues of $A \otimes B$ are given by $\lambda_i \mu_j$, all possible combinations of λ_i an eigenvalue of A and μ_j an eigenvalue of B [104]. Consequently, the higher order terms in the sum, $T_n \otimes X_n$ for $n > 1$, are indefinite. In sum, this implies that the higher order terms cannot be interpreted as physiologically meaningful covariance matrices. Therefore, the estimated sum is converted to an interpretable sum of KP. For this conversion, the freedom of a best rank N approximation is exploited.

In general, the freedom in the shuffled sum of N Kronecker products in equation (4.14) can be exhibited by a nonsingular matrix $H \in \mathbb{R}^{N \times N}$:

$$
\mathcal{S}(\Sigma) = \sum_{n=1}^{N} vec(X_n) vec(T_n)^t = \begin{pmatrix} vec(X_1) & \cdots & vec(X_N) \end{pmatrix} \begin{pmatrix} vec(T_1)^t \\ \vdots \\ vec(T_N)^t \end{pmatrix}
$$

$$
= \left[\begin{pmatrix} vec(X_1) & \cdots & vec(X_N) \end{pmatrix} H \right] \left[H^{-1} \begin{pmatrix} vec(T_1)^t \\ \vdots \\ vec(T_N)^t \end{pmatrix} \right]
$$

$$
= \begin{pmatrix} vec(\tilde{X}_1) & \cdots & vec(\tilde{X}_N) \end{pmatrix} \begin{pmatrix} vec(\tilde{T}_1)^t \\ \vdots \\ vec(\tilde{T}_N)^t \end{pmatrix}. \tag{4.25}
$$

In this rewritten expression, the $\tilde{X}_n$ ($\tilde{T}_n$) matrices are linear combinations of the X_n (T_n) matrices and are thus symmetric. Furthermore, note that this rewriting does not damage the KP structure: $\Sigma_s = \sum_n T_n \otimes X_n = \sum_n \tilde{T}_n \otimes \tilde{X}_n$. To convert the estimated sum to an interpretable sum of KP, one should try to find a matrix H such that the $(\tilde{T}_n, \tilde{X}_n)$ are positive (semi-)definite for $n = 1, \ldots, N$. The remainder of this section concentrates on the special case $N = 2$.

For $N = 2$ the matrix H becomes a nonsingular (2×2) matrix and can be written as

$$
H = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{4.26}
$$

with $ad - bc = 1$ such that $H^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. In practise, it appears that the best orthogonally estimated sum of two KP, $\hat{T}_1 \otimes \hat{X}_1 + \hat{T}_2 \otimes \hat{X}_2$, is not always positive definite, but usually contains some small negative eigenvalues. Although these values are very small, this indicates that the orthogonally estimated sum of two KP is indefinite. Consequently, it is not possible to rewrite this sum of two KP as a sum of two positive definite KP, because a positive definite sum cannot equal an indefinite expression. Therefore, we seek a matrix H that is optimal in a slightly different way. Given the estimated sum, $\hat{T}_1 \otimes \hat{X}_1 + \hat{T}_2 \otimes \hat{X}_2$, the shuffled version of the rewritten sum in terms of a, b, c and d is

$$\left(\begin{array}{cc} vec(\hat{X}_1) & vec(\hat{X}_2) \end{array}\right) \left(\begin{array}{cc} a & b \\ c & d \end{array}\right) \left(\begin{array}{cc} d & -b \\ -c & a \end{array}\right) \left(\begin{array}{c} vec(\hat{T}_1)^t \\ vec(\hat{T}_2)^t \end{array}\right)$$

$$= \left(\begin{array}{cc} a\ vec(\hat{X}_1) + c\ vec(\hat{X}_2) & b\ vec(\hat{X}_1) + d\ vec(\hat{X}_2) \end{array}\right) \left(\begin{array}{c} d\ vec(\hat{T}_1)^t - b\ vec(\hat{T}_2)^t \\ -c\ vec(\hat{T}_1)^t + a\ vec(\hat{T}_2)^t \end{array}\right).$$

Hence,

$$\tilde{X}_1 = a\ \hat{X}_1 + c\ \hat{X}_2 \tag{4.27}$$
$$\tilde{X}_2 = b\ \hat{X}_1 + d\ \hat{X}_2 \tag{4.28}$$
$$\tilde{T}_1 = d\ \hat{T}_1 - b\ \hat{T}_2 \tag{4.29}$$
$$\tilde{T}_2 = -c\ \hat{T}_1 + a\ \hat{T}_2. \tag{4.30}$$

As noticed above, these matrices cannot be all four positive definite. These matrices are uniquely decomposed into a symmetric positive part and a symmetric negative part, the latter of which might be zero for some but not all of the four matrices. This can be expressed as:

$$\tilde{X}_1 = U_{\tilde{X}_1}\Delta^+_{\tilde{X}_1}U^t_{\tilde{X}_1} - V_{\tilde{X}_1}\Delta^-_{\tilde{X}_1}V^t_{\tilde{X}_1} \tag{4.31}$$
$$\tilde{X}_2 = U_{\tilde{X}_2}\Delta^+_{\tilde{X}_2}U^t_{\tilde{X}_2} - V_{\tilde{X}_2}\Delta^-_{\tilde{X}_2}V^t_{\tilde{X}_2} \tag{4.32}$$
$$\tilde{T}_1 = U_{\tilde{T}_1}\Delta^+_{\tilde{T}_1}U^t_{\tilde{T}_1} - V_{\tilde{T}_1}\Delta^-_{\tilde{T}_1}V^t_{\tilde{T}_1} \tag{4.33}$$
$$\tilde{T}_2 = U_{\tilde{T}_2}\Delta^+_{\tilde{T}_2}U^t_{\tilde{T}_2} - V_{\tilde{T}_2}\Delta^-_{\tilde{T}_2}V^t_{\tilde{T}_2} \tag{4.34}$$

where all $U_\#$ and $V_\#$ are matrices with orthogonal columns, and the $\Delta^+_\#$ and $\Delta^-_\#$ are positive diagonal matrices with descending entries along the diagonal. Now a, b, c and d are estimated such that the matrix power corresponding to the negative eigenvalues of the four matrices, i.e. the power of the $\Delta^-_\#$ matrices is minimum. Then the $\Delta^-_\#$ matrices are set to zero, such that the final rewritten sum, denoted by $\breve{T}_1 \otimes \breve{X}_1 + \breve{T}_2 \otimes \breve{X}_2$, only contains positive (semi-)definite matrices:

$$\breve{X}_1 = U_{\tilde{X}_1}\Delta^+_{\tilde{X}_1}U^t_{\tilde{X}_1} \tag{4.35}$$
$$\breve{X}_2 = U_{\tilde{X}_2}\Delta^+_{\tilde{X}_2}U^t_{\tilde{X}_2} \tag{4.36}$$
$$\breve{T}_1 = U_{\tilde{T}_1}\Delta^+_{\tilde{T}_1}U^t_{\tilde{T}_1} \tag{4.37}$$
$$\breve{T}_2 = U_{\tilde{T}_2}\Delta^+_{\tilde{T}_2}U^t_{\tilde{T}_2}. \tag{4.38}$$

Summarising, the cost function used to find a, b, c and d is

$$\mathcal{C}(a,b,c,d) = \frac{\|\Sigma_s - \breve{T}_1 \otimes \breve{X}_1 - \breve{T}_2 \otimes \breve{X}_2\|^2}{\|\Sigma_s\|^2} \tag{4.39}$$

and the relative matrix power of the sample covariance matrix explained by the rewritten sum, $\breve{T}_1 \otimes \breve{X}_1 + \breve{T}_2 \otimes \breve{X}_2$, is defined as

$$\text{rel pow rewr sum} := (1 - \frac{\|\Sigma_s - \breve{T}_1 \otimes \breve{X}_1 - \breve{T}_2 \otimes \breve{X}_2\|^2}{\|\Sigma_s\|^2}) \times 100\%. \tag{4.40}$$
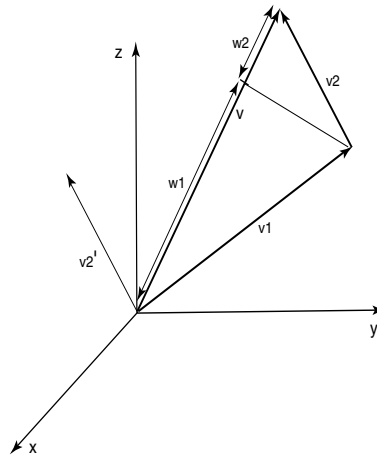
Figure 4.1: *Illustration of the computation of the contribution of the rewritten terms. Vector* v1 (v2) *presents the first (second) rewritten term and the vector* v *corresponds to the rewritten sum. The contribution of* v1 (v2) *to the length of* v *is the length* w1 (w2).

Compared to the contribution of the orthogonally estimated terms in equation (4.21), the relative contribution of the two rewritten terms, $\breve{T}_1 \otimes \breve{X}_1$ and $\breve{T}_2 \otimes \breve{X}_2$, is less well-defined because the terms are not orthogonal anymore. To compute the contribution of the two terms, the vecced KP terms are considered as elements in $\mathbb{R}^{I^2 J^2}$. In Figure 4.1 this embedding is illustrated; the first term is represented by vector v_1 and the second by v_2. The sum is drawn as vector v. The relative contribution of v_1 (v_2) to the length of v is the length w_1 (w_2) divided by $\|v\|^2$:

$$\frac{w_1}{\|v\|^2} = \frac{v_1 \cdot v}{\|v\|^2} \tag{4.41}$$

$$\frac{w_2}{\|v\|^2} = \frac{v_2 \cdot v}{\|v\|^2} \tag{4.42}$$

Here the $a \cdot b$ denotes the inner product of vectors a and b. Note that the sum of the contributions of the vectors equals the contribution of the sum of the vectors, that is 1. This principle is applied to $v_n = vec(\breve{T}_n \otimes \breve{X}_n)$, for $n = 1, 2$. It is important that the angle between v_1 and v_2 is at least 90°, because otherwise, v_2 would point backwards. The angle between the vectors v_1 and v_2' in the graph can be computed from the inner product between $vec(T \otimes X_1)$ and $vec(T_2 \otimes X_2)$. This inner product is $(vec(T_1 \otimes X_1))^t vec(T_1 \otimes X_1) = (vec(T_1)^t vec(T_2)(vec(X_1))^t vec(X_1) = tr(T_1 T_2) tr(X_1 X_2) \geq 0$. Therefore, the angle between v_1 and v_2' is between 0° and 90°. Hence, the angle between the v_1 and v_2 will be at least 90°. The relative explained power of the sum, equation (4.40), is split into two parts, proportional to the relative contributions of the two terms, yielding

$$\text{rel pow } n^{th} \text{ rewr term } :=$$

$$\frac{\|\breve{T}_n\|^2 \|\breve{X}_n\|^2 + tr(\breve{T}_1 \breve{T}_2) tr(\breve{X}_1 \breve{X}_2)}{\|\breve{T}_1 \otimes \breve{X}_1 + \breve{T}_2 \otimes \breve{X}_2\|^2} \frac{\|\Sigma_s\|^2 - \|\Sigma_s - \breve{T}_1 \otimes \breve{X}_1 - \breve{T}_2 \otimes \breve{X}_2\|^2}{\|\Sigma_s\|^2} \tag{4.43}$$

for $n = 1, 2$. The computation of this power distribution requires the computation of $\|\Sigma_s\|^2$. It appeared that, compared to the estimation of the orthogonal KP terms in equations (4.22) and (4.23), the calculation of $\|\Sigma_s\|^2$ requires considerably more computation time.

4.3   Results

The sum of KP model was applied to evoked response MEG data sets of three different kinds: Somatosensory Evoked Field (SEF) (4 subjects), Visual Evoked Field (VEF) (3 subjects) and Auditory Evoked Field (AEF) data (3 subjects). First, for each data set considered, the average signal over trials was subtracted to obtain the MEG residuals and an offset correction over one alpha period was applied. The alpha period was obtained from the frequency spectrum of the raw data. The offset correction over one alpha period is optimal to reduce the introduction of nonstationarities in the temporal covariance due to alpha background activity, as explained in chapter 3 (single KP model) and appendix 4.5.4 (sum KP model). After the first KP was estimated using equations (4.10) and (4.11), the second KP was found from the iterative system in equations (4.22) and (4.23). Then the relative matrix power explained by the first two orthogonally estimated KP terms was calculated according to equation (4.21). To find the optimal values for a, b, c and d a global search was performed for a spatiotemporally downsampled data set. This reduced data set was obtained by downsampling both in space and in time to approximately 30 time samples and 30 sensors. The so obtained optimal values for a, b, c and d were used to rewrite the original data set. Finally, the relative matrix power explained by the rewritten terms was computed according to equation (4.43).

Table 4.1 presents the resulting power distributions of both the orthogonally estimated and the rewritten sum of two KP as well as the values for a, b, c and d for all data sets in this study. This table shows that the first KP describes between 62% and 91% power of the sample covariance matrix, and the second term between 1% and 12%. Rewriting the first two terms into two positive semi-definite terms only reduces the total amount of explained matrix power by a negligible amount (less than 1.4%). As can be seen from Table 4.1, the power distribution over the rewritten terms varies over subjects and data types. These values are directly related to the varying values for a, b, c and d. In the global search over a, b, c and d for the downsampled data sets, it appeared that the cost function contains several local minima that are very close to the global minimum in cost value. Apparently, rewriting the sum is not very sensitive to (small) changes in a, b, c and d.

Table 4.2 contains the positivity percentages of the temporal and spatial matrices for all data sets. By positivity percentages the relative matrix power explained by the positive eigenvalues of matrix is meant. The first orthogonally estimated term is always positive definite, and therefore the entries of the first two columns, XO1 and TO1, all equal 100%. The third and fourth columns contain the positivity percentages of the matrices in the second KP of the orthogonally estimated sum. Clearly, these matrices are far from positive definite, and can be even mainly negative (e.g. subject 5VEF). After rewriting according to equations (4.31 - 4.34), all four matrices (XR1,

S	TW	BPF	O1	O2	O	a	b	c	d	ADD	AL	RS	DIF
1SEF	574	none	84.79	3.10	87.89	1.30	-0.20	1.43	0.55	59.46	27.47	86.93	0.96
2SEF	574	none	71.82	11.98	83.80	1.00	-0.70	0.79	0.45	38.63	44.79	83.42	0.38
3SEF	574	none	77.66	5.80	83.46	0.90	-0.55	1.00	0.50	37.18	45.05	82.23	1.23
4SEF	574	none	86.63	1.16	87.79	1.30	-0.15	1.03	0.65	72.88	14.76	87.64	0.15
5VEF	480	0-50	86.36	2.71	89.07	0.85	-0.20	0.96	0.95	68.13	19.59	87.72	1.34
6VEF	480	0-50	90.77	1.58	92.35	1.15	-0.10	0.80	0.80	83.41	8.92	92.33	0.02
7VEF	480	0-50	90.67	2.25	92.92	0.95	-0.10	0.98	0.95	81.96	10.99	92.90	0.02
8AEF	480	none	62.12	5.16	67.28	0.60	-0.20	2.15	0.95	37.51	29.71	67.22	0.06
9AEF	480	none	77.10	10.35	87.45	1.10	-0.05	0.20	0.90	76.00	11.25	87.26	0.19
10AEF	480	none	78.25	7.02	85.28	1.00	-0.15	0.67	0.90	70.87	14.36	85.23	0.05

Table 4.1: *The power distribution in the orthogonally estimated sum and in the rewritten sum of 2 KP for all data sets in the study. The first column (S) denotes the subject and the kind of data (SEF/VEF/AEF), TW indicates the length of the time window analysed in ms. Used band pass filtering is stated in the BPF column. O1 (O2) denotes the relative matrix power of the sample covariance matrix in the first (second) orthogonally estimated term, and O is the sum of O1 and O2. The columns ADD and AL denote the relative matrix power explained by the rewritten terms, ADD indicates the widespread background activity, and AL the alpha activity. RS is the sum of ADD and AL. DIF is the difference between O and RS. The given values for a, b, c and d are the values used for rewriting the sum.*

S	XO1	TO1	XO2	TO2	XR1	TR1	XR2	TR2
1SEF	100.00	100.00	51.97	74.50	100.00	98.38	100.00	99.75
2SEF	100.00	100.00	52.31	71.40	99.77	99.07	100.00	100.00
3SEF	100.00	100.00	64.49	84.89	98.43	97.12	100.00	100.00
4SEF	100.00	100.00	39.91	66.61	100.00	100.00	97.70	99.77
5VEF	100.00	100.00	38.29	20.98	100.00	100.00	94.36	91.38
6VEF	100.00	100.00	67.93	87.02	100.00	100.00	99.68	99.80
7VEF	100.00	100.00	78.28	91.02	100.00	99.97	99.98	100.00
8AEF	100.00	100.00	54.61	63.48	100.00	99.82	99.99	99.99
9AEF	100.00	100.00	91.51	96.39	100.00	100.00	99.00	99.14
10AEF	100.00	100.00	70.68	86.82	100.00	100.00	99.67	99.85

Table 4.2: *The positivity percentages of the matrices in both terms in the orthogonally estimated sum and in the rewritten sum in equations (4.31 - 4.34). The positivity percentage of a matrix equals the relative matrix power that is accounted for by the positive eigenvalues of that matrix. The first column indicates the subject. Columns XO1 (XO2) and TO1 (TO2) show the percentages of the spatial and temporal matrices in the first (second) term of the orthogonally estimated sum of 2 KP. All entries in the XO1 and TO1 column equal 100% because the first orthogonal term is positive definite. The positivity percentages after rewriting are given in the columns XR1 (XR2) and TR1 (TR2) for the spatial and temporal matrices in both rewritten terms.*

TR1, XR2 and TR2) are well-nigh or completely positive definite. Subject 5VEF is the only case showing positivity percentages below 95%. This table reveals the effect of rewriting in terms of interpretability: before rewriting, the second term does not possess interpretability as a covariance matrix, while after rewriting both terms can be interpreted as covariance matrices. Namely, the rewritten matrices are slightly singular - their small negative parts were set to zero - and this singularity can be interpreted as a light linear dependency amongst the signals.

For one data set of each kind (SEF, VEF, AEF), the results are illustrated here. In Figures 4.2, 4.4 and 4.6 the temporal matrices of the three data sets are shown. The visualisation of the temporal covariance matrices is through plotting the entries of the matrix in colour. In order to plot a temporal covariance matrix, the entries are scaled such that the entry that is largest in absolute value equals 1 or -1. The colour scale used for these covariance plots is equal to the colour scale used in chapter 3 and is presented in Figure 3.2. As with usual printing of matrices, all entries are arranged in a square, and in stead of values, corresponding colours are plotted. Nonstationarities and oscillations in the temporal domain can now easily be detected: a stationary temporal covariance matrix has a constant value (colour) along its (sub)diagonals and oscillations in the covariance are reflected by a line pattern parallel to the diagonal. Figures 4.2 and 4.6 show in subfigures (a) and (b) (the orthogonal terms) a clear oscillation in the covariance. The frequency of this oscillation is approximately 10 Hz; hence this oscillation shows the alpha activity in the background noise. As expected, these matrices are not purely oscillatory, that is, more noise features are present besides alpha activity.

In rewriting the terms, the different noise characteristics of the orthogonally estimated terms are rearranged such that the cost function in equation (4.39) is minimum. It appears that after rewriting, the alpha activity is contained mainly in the second term, while the most of the remaining activity is gathered in the first rewritten term. Note that the cost function in equation (4.39) is not frequency specific. For the VEF data set, presented in Figure 4.4, alpha oscillations in the orthogonal terms are much smaller and are mainly visible in the second term, subfigure (b). Nonetheless, after rewriting, no alpha activity is visible in the first term, whereas the second term mainly consists of alpha activity.

Regarding the nonstationarities, Figures 4.2, 4.4 and 4.6 show predominantly stationary temporal matrices. However, some nonstationary patterns can be detected; that is, the colour along the (sub)diagonal varies somewhat. The main nonstationarity that can be seen from the colour plots is the increase along the diagonal and the subdiagonals. This is the common consequence of the offset correction over the pre-stimulus interval, which artificially pulls the (co)variance over that time window towards zero (chapter 3). Further nonstationarities generally occur in the second orthogonal component (subfigures (b)), and in one or both rewritten terms. These nonstationarities may be caused by beta activity for which the offset correction window is not optimal. The relative matrix power corresponding to these nonstationarities is very small in comparison to that of the alpha activity. Namely, this feature only occurs in the second orthogonal term, and the power in this second orthogonal term is much smaller than the power of the first. Moreover, the second orthogonal term is still mainly stationary.

The rewritten spatial variances of these three data sets are presented in Figures 4.3, 4.5 and 4.7. The spatial covariance matrices are visualised by projecting the variance (the diagonal of the matrix) in colour scale (as defined in Figure 3.2) on the MEG helmet. For all data sets the second term, corresponding to the alpha pattern in the temporal matrix, shows a focal highlighted area in the parieto-occipital area. The spatial distribution of the first term is more widespread, though tends to be more in the temporal region.

In sum, minimising the cost function in equation (4.39) yields two rewritten KP terms, each of which describes a distinct process in the background activity. The first rewritten term describes a rather widespread, not frequency-specific process, while the second term describes the focal alpha activity with its characteristic 10 Hz frequency.

4.4   Discussion

The sum of Kronecker products provides a general model for the spatiotemporal covariance matrix of MEG residuals. Different terms in the sum can describe different, independent phenomena in the ongoing background activity, each of which has its own temporal and spatial characteristics. These processes can be interpreted as generated by randomly distributed dipoles with a certain spatial and temporal distribution. This way, the sum model solves the rigidness drawback of the single KP model. Theoretically, when enough terms are taken into account, the sum describes the sample covariance matrix *exactly*.

(a) First term in orthogonal sum (71.8%)

(b) Second term in orthogonal sum (12.0%)

(c) First term in rewritten sum (38.6%)

(d) Second term in rewritten sum (44.8%)
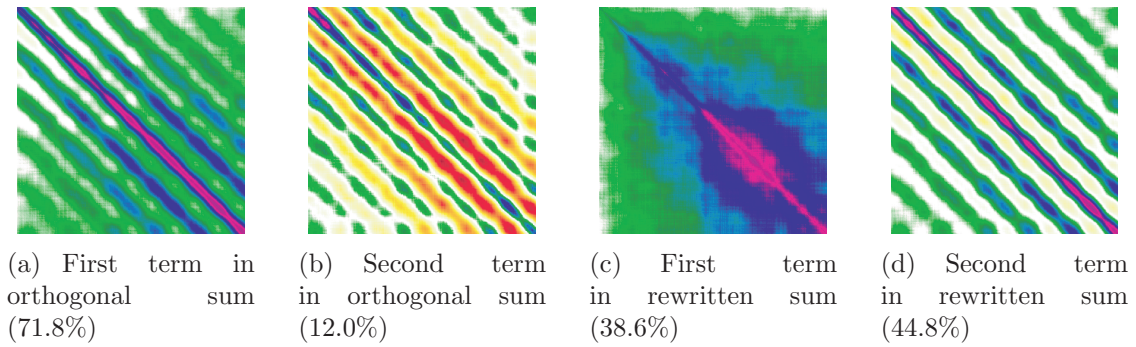
Figure 4.2: *Estimated temporal matrices in the 2SEF data set. Frames a and b show the first two temporal matrices of the orthogonally estimated terms, frames c and d show the first two temporal matrices of the rewritten sum of two matrices. The time scale is 574 ms by 574 ms. The entries of the matrices are plotted in colour. The percentages show the relative matrix power of the sample covariance matrix explained by the KP term.*

(a) First term in rewritten sum

(b) Second term in rewritten sum
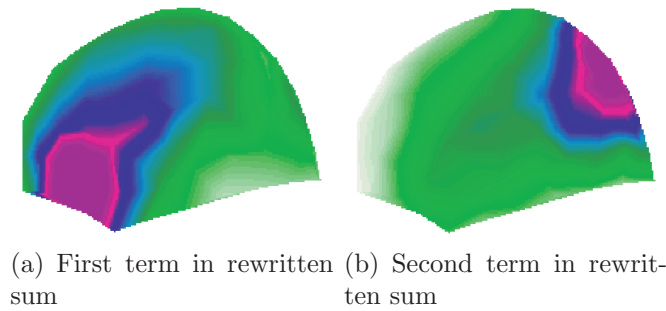
Figure 4.3: *Rewritten spatial variances in the 2SEF data set. The diagonal entries of the spatial covariance matrix, the variances, are plotted in colour on the MEG helmet. Only the left side of the helmet is shown, the right side is similar.*

(a) First term in orthogonal sum (90.7%)

(b) Second term in orthogonal sum (2.2%)

(c) First term in rewritten sum (81.9%)

(d) Second term in rewritten sum (11.0%)

Figure 4.4: *Estimated temporal matrices in the 7VEF data set. The time scale is 478 ms by 478 ms. See description of Figure 4.2.*

(a) First term in rewritten sum

(b) Second term in rewritten sum
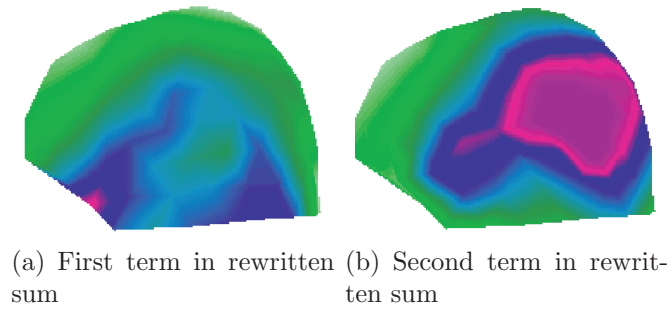
Figure 4.5: *Rewritten spatial variances in the 7VEF data set. The diagonal entries of the spatial covariance matrix, the variances, are plotted in colour on the MEG helmet. Only the left side of the helmet is shown, the right side is similar.*

(a) First term in orthogonal sum (78.3%)

(b) Second term in orthogonal sum (7.0%)

(c) First term in rewritten sum (70.9%)

(d) Second term in rewritten sum (14.4%)

Figure 4.6: *Estimated temporal matrices in the 10AEF data set. The time scale is 318 ms by 318 ms. See description of Figure 4.2.*

(a) First term in rewritten sum

(b) Second term in rewritten sum
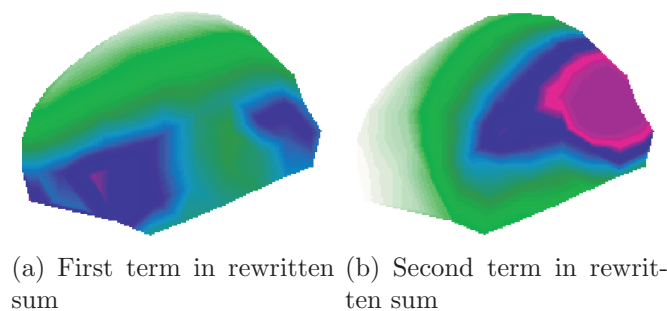
Figure 4.7: *Rewritten spatial variances in the 10AEF data set. The diagonal entries of the spatial covariance matrix, the variances, are plotted in colour on the MEG helmet. Only the right side of the helmet is shown, the left side is similar.*

The first aim of this study is the validation of the single KP model for dipole localisation in terms of accuracy. In practise, it occurred that the first KP term describes roughly between 62% and 91% of the sample covariance matrix and the second between 1% and 12%, whereas the sum of 2 KP explains between 67% and 93%. The higher the order of the term, the smaller the amount of explained power. Taking into account more than 1 KP term in the localisation is not expected to yield a major improvement. Namely, the common practise to neglect all the correlations, i.e. both in space and in time, yields an acceptable accuracy at high signal-to-noise ratio. This accuracy is enhanced by taking into account only the spatial correlations, and a further improvement is achieved when the spatiotemporal covariance, the first KP term, is taken into account [20]. Considering the matrix power explained already by the first KP term, the second and higher order terms are not expected to enhance the localisation accuracy considerably (see Table 4.1). To improve the source localisation further, the emphasis should be on improving the source model, which is likely to be more beneficial. For this reason,in the next two chapters a new forward model is presented and discussed which can be used to analyse multiple related MEG data sets simultaneously. The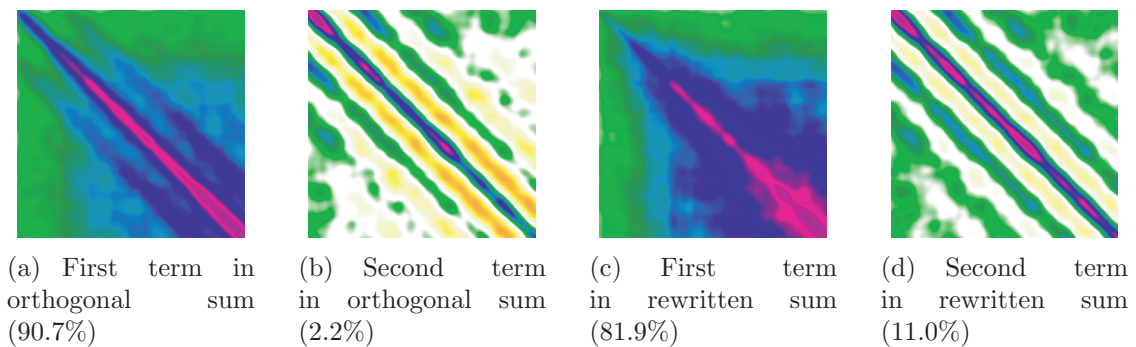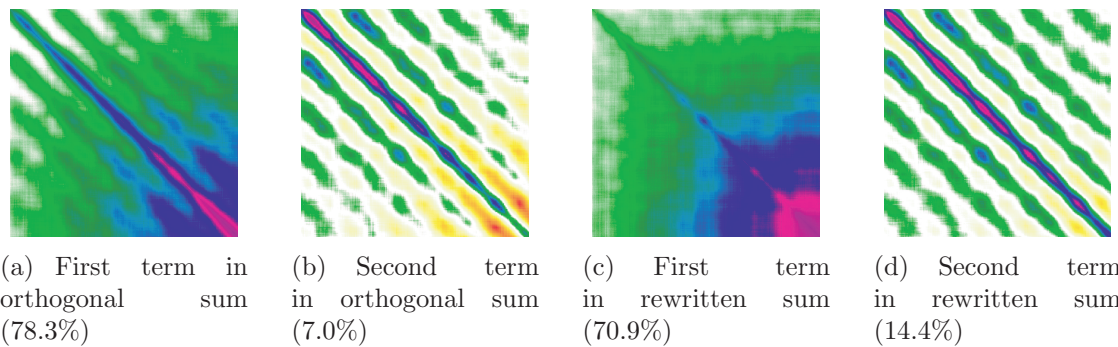 study in the present chapter indicates that the existing covariance model for source localisation, the single KP, is adequate.

Nevertheless, the estimated sum of KP contains interesting physiological information, which is the second goal of the present study. There are two aspects regarding this aim: the separation between alpha activity and the remainder, and nonstationarities. For all subjects, the two terms of the rewritten sum show one term corresponding to alpha activity and the other to additional noise. It is emphasised that this separation automatically comes forth by minimising the cost function in equation (4.39) and is not caused by an a priori constraint. The alpha term is characterised by frequency specific (10 Hz) temporal features and a focal parieto-occipital pattern in space. The additional term shows more widespread characteristics, both in space and in time: there is no frequency specific character in time and the spatial distribution is widespread, though seems to be enhanced in the temporal region. The power distribution over the rewritten terms in Table 4.1 suggests that the VEF data sets contain less alpha background activity than the SEF and AEF data. This can be interpreted in line with the discussion in the literature about whether the visual stimulus resets the phase of the spontaneous alpha rhythm, e.g. [9, 52, 54, 65]. If that applies, subtracting the average includes subtracting a major amount of alpha activity and less alpha activity will remain in the MEG residuals. Nevertheless, for all subjects, including the VEF subjects, separation between the alpha activity and the additional activity is striking, although the entireness of the separation varies slightly over subjects. This separation can be interpreted in line with the Poisson modulated alpha model, which models the background activity in the temporal domain as a sum of alpha activity and additional random activity.

The second aspect of physiological information is about the temporal nonstationarities. As stated in the section 1.4.2, the temporal covariance matrices will contain nonstationarities when the SPN model is assumed incorrectly. However, in the presented results the temporal matrices are mainly stationary. The most apparent nonstationarity is the increase along the (sub)diagonals, caused by the offset correction. Furthermore,

along the diagonal one can see small oscillations, reflected in the coloured bands having varying width. After rewriting the sum, these small oscillations are mainly visible in the alpha term, Figures 4.2(d), 4.4(d) and 4.6(d). A possible explanation for these oscillations again lies in the offset correction: for oscillatory background activity with the beta frequency the offset correction window of one alpha period is not optimal and will introduce a small oscillation along the diagonal and subdiagonals [5]. Further nonstationarities that indicate real trial-to-trial variabilities are very limited in the colour plots. Despite the minor role of these real nonstationarities in terms of matrix power, this study does not disapprove their existence.

In sum, it appears from the distribution of explained matrix power that it is unlikely that dipole localisation will improve considerably by adding more terms to the single KP noise model. Secondly, rewriting the sum of two KP suggests that the noise consists of a focal parieto-occipital alpha part and a more widespread noise part. And finally, nonstationarities due to real trial-to-trial variations in this study appeared to be very limited in terms of matrix power.

4.5   Appendix

4.5.1   Estimating a single Kronecker product

In this appendix expressions for the LS-estimators of X and T in the single KP model are derived. The LS cost function can be written as:

$$\|\Sigma_s - T \otimes X\|^2 = tr(\Sigma_s \Sigma_s^t) + tr(T^2)tr(X^2) - 2\frac{1}{K-1}\sum_{k=1}^{K} tr(E_k T E_k^t X). \qquad (4.44)$$

The optimal X and T are found by differentiating equation (4.44) and subsequently equating the first derivative to zero. Differentiation with respect to matrices is performed according to the rules derived in [63]. For X this yields

$$d_X\|\Sigma_s - T \otimes X\|^2 = 2tr(T^2)tr(XdX) - 2\frac{1}{K-1}\sum_{k=1}^{K} tr(E_k T E_k^t dX)$$

$$= 2tr([tr(T^2)X - \frac{1}{K-1}\sum_{k=1}^{K} E_k T E_k^t]dX) = 0$$

$$\implies tr(T^2)X - \frac{1}{K-1}\sum_{k=1}^{K} E_k T E_k^t = 0 \qquad (4.45)$$

Rewriting this equation yields equation (4.10) as LS-estimator for X and a similar derivation yields equation (4.11) for $\hat{T}_{LS}$.

4.5.2   Van Loan's shuffle operator

The shuffle operator $\mathcal{S} : \mathbb{R}^{IJ \times IJ} \longrightarrow \mathbb{R}^{I^2 \times J^2}$ is defined by

$$\mathcal{S} = vec^{-1} \circ (I_J \otimes K_{J,I} \otimes I_I) \circ vec, \qquad (4.46)$$

where the composition operator $\circ$ is defined as $(L_1 \circ L_2)A = L_1(L_2(A))$ and $vec^{-1} : \mathbb{R}^{I^2 J^2 \times 1} \longrightarrow \mathbb{R}^{I^2 \times J^2}$ is the inverse the usual vec operator $vec : \mathbb{R}^{I^2 \times J^2} \longrightarrow \mathbb{R}^{I^2 J^2 \times 1}$. $K_{p,q} : \mathbb{R}^{1 \times pq} \longrightarrow \mathbb{R}^{1 \times qp}$ is the general commutation matrix [63]:

$$K_{p,q}(vec(A)) = vec(A^t) \tag{4.47}$$

for any matrix $A \in \mathbb{R}^{p \times q}$. To see equation (4.46) the following equality from [63] is used

$$(I_n \otimes K_{m,q} \otimes I_p)vec(A \otimes B) = vec(A) \otimes vec(B). \tag{4.48}$$

for any $A \in \mathbb{R}^{m \times n}$ and any $B \in \mathbb{R}^{p \times q}$. Applying this equality with $A = T$ and $B = X$ and using

$$vec(vec(X_n)vec(T_n)^t) = vec(T_n) \otimes vec(X_n) \tag{4.49}$$

one arrives at equation (4.46). Applying equation (4.46) to equation (4.6) and successively applying equations (4.49) and (4.48) yields the formula for the shuffled sample covariance, $\mathcal{S}(\Sigma_s)$, in equation (4.16).

4.5.3 The LS estimators for the higher order terms in the sum of KP

In this appendix the LS estimators for the higher order terms in the sum of KP model are derived. The pth order term is estimated after the terms $1, \ldots, p-1$ have been estimated. Estimation of the pth term is by differentiation of the cost function in equation (4.15) and applying Lagrange multiplication [68]. Abbreviate

$$\mathbf{x}_n = vec(X_n) \tag{4.50}$$
$$\mathbf{t}_n = vec(T_n). \tag{4.51}$$

Suppose the first $p-1$ terms have been estimated, where the normalisation is chosen such that $\|\mathbf{x}_n\|^2 = 1$ for $n = 1, \ldots, p-1$. The cost function for the pth term is then

$$\|[\mathcal{S}(\Sigma_s) - \sum_{n=1}^{p-1} \mathbf{x}_n \mathbf{t}_n^t] - \mathbf{x}_p \mathbf{t}_p^t\|^2 = \|\mathcal{S}_{p-1} - \mathbf{x}_p \mathbf{t}_p^t\|^2$$

$$= tr(\mathcal{S}_{p-1}^t \mathcal{S}_{p-1} - 2\mathcal{S}_{p-1}^t \mathbf{x}_p \mathbf{t}_p^t + \mathbf{t}_p \mathbf{x}_p^t \mathbf{x}_p \mathbf{t}_p^t), \tag{4.52}$$

where

$$\mathcal{S}_{p-1} = \mathcal{S}(\Sigma_s) - \sum_{n=1}^{p-1} \mathbf{x}_n \mathbf{t}_n^t. \tag{4.53}$$

In other words, the pth term is the best rank-one approximation to $\mathcal{S}_{p-1}$. Considering equations (4.19) and (4.20) this minimisation is subject to

$$\mathbf{x}_n^t \mathbf{x}_p = 0 \quad \text{for } 1 \le n \le p-1 \tag{4.54}$$
$$\mathbf{t}_n^t \mathbf{t}_p = 0 \quad \text{for } 1 \le n \le p-1. \tag{4.55}$$

The derivative of equation (4.52) with respect to $\mathbf{x}_p$ is given by

$$tr\left(-2\mathcal{S}_{p-1}^t d\mathbf{x}_p \mathbf{t}_p^t + \mathbf{t}_p d\mathbf{x}_p^t \mathbf{x}_p \mathbf{t}_p^t + \mathbf{t}_p \mathbf{x}_p^t d\mathbf{x}_p \mathbf{t}_p^t\right) = 2tr((-\mathbf{t}_p^t \mathcal{S}_{p-1}^t + \mathbf{t}_p^t \mathbf{t}_p \mathbf{x}_p^t)d\mathbf{x}_p) \tag{4.56}$$

and, similarly, with respect to $\mathbf{t}_p$

$$2tr((-\mathbf{x}_p^t \mathcal{S}_{p-1} + \mathbf{x}_p^t \mathbf{x}_p \mathbf{t}_p^t)d\mathbf{t}_p). \tag{4.57}$$

The derivatives of the constraints in equations (4.54) and (4.55) are

$$\begin{align}
\mathbf{x}_n^t d\mathbf{x}_p &= 0 \quad \text{for } 1 \le n \le p-1 \tag{4.58} \\
\mathbf{t}_n^t d\mathbf{t}_p &= 0 \quad \text{for } 1 \le n \le p-1 \tag{4.59}
\end{align}$$

The method of Lagrange multiplication now yields the following system of equations

$$\begin{align}
\mathbf{x}_n^t \mathbf{x}_p &= 0 \in \mathbb{R} \ \text{ for } 1 \le n \le p-1 \tag{4.60} \\
\mathbf{t}_n^t \mathbf{t}_p &= 0 \in \mathbb{R} \ \text{ for } 1 \le n \le p-1 \tag{4.61}
\end{align}$$

$$-\mathcal{S}_{p-1}\mathbf{t}_p + \mathbf{t}_p^t \mathbf{t}_p \mathbf{x}_p + \sum_{n=1}^{p-1} \lambda_n \mathbf{x}_n = \mathbf{0} \in \mathbb{R}^{I^2 \times 1} \tag{4.62}$$

$$-\mathcal{S}_{p-1}^t \mathbf{x}_p + \mathbf{x}_p^t \mathbf{x}_p \mathbf{t}_p + \sum_{n=1}^{p-1} \mu_n \mathbf{t}_n = \mathbf{0} \in \mathbb{R}^{J^2 \times 1} \tag{4.63}$$

which has to be solved for $\mathbf{x}_p$, $\mathbf{t}_p$, $\lambda_1, \ldots, \lambda_{p-1}$ and $\mu_1, \ldots, \mu_{p-1}$. The solution for λ_m, $m = 1, \ldots, p-1$, follows from the inner product of $\mathbf{x}_m$ and the vectors in equation (4.62):

$$-\mathbf{x}_m^t \mathcal{S}_{p-1}\mathbf{t}_p + \mathbf{x}_m^t \mathbf{t}_p^t \mathbf{t}_p \mathbf{x}_p + \sum_{n=1}^{p-1} \lambda_n \mathbf{x}_m^t \mathbf{x}_n = 0 \implies \lambda_m = \frac{1}{\mathbf{x}_m^t \mathbf{x}_m} \mathbf{x}_m^t \mathcal{S}_{p-1}\mathbf{t}_p = \mathbf{x}_m^t \mathcal{S}_{p-1}\mathbf{t}_p. \tag{4.64}$$

Likewise, solving the inner product of $\mathbf{t}_m$ and the vectors in equation (4.63) for μ_m yields

$$\mu_m = \frac{\mathbf{t}_m^t \mathcal{S}_{p-1}^t \mathbf{x}_p}{\mathbf{t}_m^t \mathbf{t}_m}. \tag{4.65}$$

Substituting equation (4.64) into equation (4.62) and using equations (4.53) and (4.61) yields

$$0 = -\mathcal{S}_{p-1}\mathbf{t}_p + \mathbf{t}_p^t \mathbf{t}_p \mathbf{x}_p + \sum_{n=1}^{p-1} \mathbf{x}_n^t \mathcal{S}_{p-1}\mathbf{t}_p \mathbf{x}_n = -\mathcal{S}_{p-1}\mathbf{t}_p + \mathbf{t}_p^t \mathbf{t}_p \mathbf{x}_p + \left(\sum_{n=1}^{p-1} \mathbf{x}_n \mathbf{x}_n^t \right) \mathcal{S}_{p-1}\mathbf{t}_p$$

$$\implies \hat{\mathbf{x}}_p = \frac{1}{\hat{\mathbf{t}}_p^t \hat{\mathbf{t}}_p}[I_{I^2} - \sum_{n=1}^{p-1} \mathbf{x}_n \mathbf{x}_n^t]\mathcal{S}_{p-1}\hat{\mathbf{t}}_p = \frac{1}{\hat{\mathbf{t}}_p^t \hat{\mathbf{t}}_p}[I_{I^2} - \sum_{n=1}^{p-1} \mathbf{x}_n \mathbf{x}_n^t]\mathcal{S}(\Sigma_s)\hat{\mathbf{t}}_p \tag{4.66}$$

where the last simplification follows from the constraints in equation (4.55). Similarly, equation (4.65) substituted into equation (4.63) together with equations (4.53), (4.60) and (4.54) yield

$$0 = -\mathcal{S}_{p-1}^t \mathbf{x}_p + \mathbf{x}_p^t \mathbf{x}_p \mathbf{t}_p + \sum_{n=1}^{p-1} \frac{\mathbf{t}_n^t \mathcal{S}_{p-1}^t \mathbf{x}_p}{\mathbf{t}_n^t \mathbf{t}_n}\mathbf{t}_n = -\mathcal{S}_{p-1}^t \mathbf{x}_p + \mathbf{x}_p^t \mathbf{x}_p \mathbf{t}_p + \left(\sum_{n=1}^{p-1} \frac{\mathbf{t}_n \mathbf{t}_n^t}{\mathbf{t}_n^t \mathbf{t}_n} \right) \mathcal{S}_{p-1}^t \mathbf{x}_p$$

$$\implies \hat{\mathbf{t}}_p = \frac{1}{\hat{\mathbf{x}}_p^t \hat{\mathbf{x}}_p}[I_{J^2} - \sum_{n=1}^{p-1}\frac{\mathbf{t}_n\mathbf{t}_n^t}{\mathbf{t}_n^t\mathbf{t}_n}]\mathcal{S}_{p-1}^t\hat{\mathbf{x}}_p = \frac{1}{\hat{\mathbf{x}}_p^t \hat{\mathbf{x}}_p}[I_{J^2} - \sum_{n=1}^{p-1}\frac{\mathbf{t}_n\mathbf{t}_n^t}{\mathbf{t}_n^t\mathbf{t}_n}](\mathcal{S}(\Sigma_s))^t\hat{\mathbf{x}}_p. \quad (4.67)$$

Equations (4.66) and (4.67) are solved iteratively and the normalisation is chosen such that $\|\mathbf{x}_p\|^2 = 1$.

Note that a closed form expression for the estimators can be obtained by substituting equation (4.67) in equation (4.66):

$$\hat{\mathbf{x}}_p = \frac{1}{\hat{\mathbf{t}}_p^t \hat{\mathbf{t}}_p}[I_{I^2} - \sum_{n=1}^{p-1}\mathbf{x}_n\mathbf{x}_n^t]\mathcal{S}(\Sigma_s)[I_{J^2} - \sum_{n=1}^{p-1}\frac{\mathbf{t}_n\mathbf{t}_n^t}{\mathbf{t}_n^t\mathbf{t}_n}](\mathcal{S}(\Sigma_s))^t\hat{\mathbf{x}}_p, \quad (4.68)$$

showing that $\mathbf{x}_p$ is an eigenvector of the $I^2 \times I^2$ matrix

$$[I_{I^2} - \sum_{n=1}^{p-1}\mathbf{x}_n\mathbf{x}_n^t]\mathcal{S}(\Sigma_s)[I_{J^2} - \sum_{n=1}^{p-1}\frac{\mathbf{t}_n\mathbf{t}_n^t}{\mathbf{t}_n^t\mathbf{t}_n}](\mathcal{S}(\Sigma_s))^t. \quad (4.69)$$

corresponding to the largest eigenvalue in order to minimise equation (4.52). Computing eigenvalues and eigenvectors of this $I^2 \times I^2$ matrix is, like computing the SVD of $\mathcal{S}(\Sigma_s)$, in practise not convenient.

Finally, expressions for the iterative estimators for the matrices X_p and T_p are derived from equations (4.16), (4.66) and (4.67).

$$vec(\hat{X}_p) = \frac{1}{(K-1)\|\hat{T}_p\|_F^2}\sum_{k=1}^K[I_{I^2} - \sum_{n=1}^{p-1}vec(X_n)(vec(X_n))^t](E_k \otimes E_k)vec(\hat{T}_p)$$

$$= \frac{1}{(K-1)\|\hat{T}_p\|_F^2}\sum_{k=1}^K[I_{I^2} - \sum_{n=1}^{p-1}vec(X_n)(vec(X_n))^t]vec[E_k\hat{T}_pE_k^t]$$

$$= \frac{1}{(K-1)\|\hat{T}_p\|_F^2}\sum_{k=1}^K\Big[vec[E_k\hat{T}_pE_k^t] - \sum_{n=1}^{p-1}vec(X_n)(vec(X_n))^tvec[E_k\hat{T}_pE_k^t]\Big]$$

$$= \frac{1}{(K-1)\|\hat{T}_p\|_F^2}\sum_{k=1}^K\Big[vec[E_k\hat{T}_pE_k^t] - \sum_{n=1}^{p-1}tr[X_nE_k\hat{T}_pE_k^t]vec(X_n)\Big]. \quad (4.70)$$

This yields equation (4.22) and a similar calculation yields equation (4.23).

4.5.4   Offset correction in case of multiple KP

The expression for the offset corrected residual in discrete time is (cf. equation 3.23)

$$\breve{e}_k^{i,j} = e_k^{i,j} - \frac{1}{L}\sum_{l=1}^L e_k^{i,j_l} \quad (4.71)$$

where $(j_1,\ldots,j_L)$ denotes the time window over which the correction is performed. In chapter 3 an expression is derived for the temporal covariance matrix of the corrected

residuals in equation (3.24), denoted here by T^c, for the single KP model ($\Sigma = T \otimes X$) in terms of the covariance matrix of the uncorrected residuals, T:

$$T^c_{j,j'} = T_{j,j'} - \frac{1}{L} \sum_{l=1}^{L} T_{j_l,j'} - \frac{1}{L} \sum_{m=1}^{L} T_{j,j_m} + \frac{1}{L^2} \sum_{l=1}^{L} \sum_{m=1}^{L} T_{j_l,j_m}. \tag{4.72}$$

For the sum of KP model, $\Sigma = \sum T_n \otimes X_n$, the covariance between two uncorrected residuals, $e_k^{i,j}$ and $e_{k'}^{i',j'}$, is

$$\mathcal{E}(e_k^{i,j}, e_{k'}^{i',j'}) = \delta_{k,k'} \sum_{n=1}^{N} (X_n)_{i,i'} (T_n)_{j,j'}. \tag{4.73}$$

Applying the offset correction of equation (4.71) to the residuals in equation (4.73) one can derive the covariance matrix of the corrected residuals in the case of the sum of KP. Straightforward calculating shows that each temporal matrix in this sum is altered according to equation (4.72).

Chapter 5

The coupled dipole model: an integrated model for multiple MEG/EEG data sets

5.1 Introduction

To investigate the functionality of the human brain, MEG/EEG is often measured in a few different but similar conditions. This way, the influence of a certain experimental parameter on the activity of the brain can be examined. For example, a common experimental paradigm for investigation of the visual cortex is the presentation of checkerboard patterns in different visual fields using varying check sizes. In this kind of experiments, the measured MEG/EEG data of the different conditions will be different, though there are similarities too.

In the source localisation of these data, the Inverse Problem (IP) of MEG/EEG, these similarities can be exploited. The IP is in general ill-posed; assumptions (e.g. head model, source model, number of sources) are necessary to solve the problem. Moreover, often extra constraints (e.g. mirror symmetry) are needed to stabilise the IP. Secondly, a low signal-to-noise ratio (SNR) results in instable solutions to the IP. The SNR of single trials in MEG/EEG measurements is usually too poor to allow source localisation on single trial basis. Therefore, a first approach to take more data into account is to average repeated measurements to increase the SNR (see section 1.3.1). Finally, a third problem with solving the IP using the common equivalent current dipole source model, is instability due to close sources: two closely localised dipole sources having nearly opposite orientations and unrealistically high magnitudes.

In the case of similar conditions the IP can be stabilised using a component model. A component model uses a set of basic spatial and temporal components and the data are described as linear combinations of (some of) these basic components. The similarity

between conditions is reflected in the usage of the same basic components in different conditions.

Component models have been designed before. The Topographic Component Model (TCM) [73] decomposes EEG data of multiple subjects into a sum of topographic components, each component consisting of a scalp distribution and a (nonparametric) time series. In [27] the TCM is applied to real data. Extensions of the TCM can be found in [1, 102, 112].

In [102] the TCM is extended by using dipolar sources [93] as spatial components and parametric, predefined time courses to describe data of different subjects. In [1] the TCM is extended by allowing different durations and different latencies of the temporal components for each condition. In [112] the spatial and temporal components are decoupled and a trilinear model is introduced with so-called loading matrices. In this trilinear model the number of spatial components can be different from the number of temporal components and these loading matrices are placed in between the spatial and the temporal matrices. The spatial components in [112] are again described by scalp distributions in stead of the more elementary dipole sources. Although in some of these studies the correlations of the background activity are mentioned [1, 102], these correlations are neglected in the estimation of the components or the parameters. Therefore these estimation methods seem somewhat ad hoc; a clearly defined statistical framework is not given.

In this chapter a new component model is presented, the coupled dipole model (CDM). The CDM resembles in a way the trilinear model in [112], though is still fundamentally different. The main difference is in the basic idea of the model: the CDM is a parameter estimation method based on the well-defined maximum likelihood (ML) framework. This way, the correlations of the background activity are taken into account in the source localisation, which has been shown to improve the accuracy of source parameters compared to the ordinary least squares approach [20, 42].

The basic spatial components of the CDM are dipole sources, and the basic temporal components are nonparametric time series. The data of each condition are modelled as a linear combination of the basic components. This linear combination is specified by a coupling matrix for each condition, comparable to loading matrices in [112]. However, contrary to the model in that study, all parameters (spatial, temporal and coupling parameters) are estimated simultaneously using the maximum likelihood paradigm, in stead of successively using different Singular Value Decompositions of the rearranged data.

The CDM is applicable when the different conditions contain common information: either common sources or common source time series. In this integrated approach a combination of more data and more constraints is used to solve and stabilise the IP: sources and/or source time series are assumed to be fixed over (part of the) conditions.

In the next section the coupled dipole model is explained. The technical details of the model are put in the appendix. Then the results of the application of the CDM in two MEG simulation studies and in an experimental visual evoked response MEG-study are shown. In the final section the results are discussed and conclusions are drawn.

5.2   Methods

5.2.1   Model

In the coupled dipole model each data set is modelled as a linear combination of basic sources (equivalent current dipoles) and basic source time functions (STF). This linear combination is specified in the coupling matrix C_q for each data set q. The coupling matrices contain the amplitudes of the components, which may vary over data sets.

Assuming that the number of sensors is I and time samples is J, the measured signal of trial k in data set q is stored in the matrix $R_q^k \in \mathbb{R}^{I \times J}$, for $q = 1, \ldots, Q$, $k = 1, \ldots, K_q$. Furthermore, if the number of basic spatial components is P and the number of basic temporal components is Z, the coupling matrices $C_q \in \mathbb{R}^{P \times Z}$. The basic field matrix $A \in \mathbb{R}^{I \times P}$ is the matrix containing the forward fields of the basic dipole sources (DF) and the basic source time function matrix $B \in \mathbb{R}^{Z \times J}$ is the matrix containing all basic STFs as rows. Using these matrices, the model for data set q is formulated as

$$R_q = AC_qB, \qquad\qquad (5.1)$$

where $A = A(\xi, \eta)$ is dependent on the source locations ξ and the source orientations η. The reader is referred to appendix 5.5.1 for a full list of dimensions and variables. In Figure 5.1 the CDM formula (5.1) is illustrated.
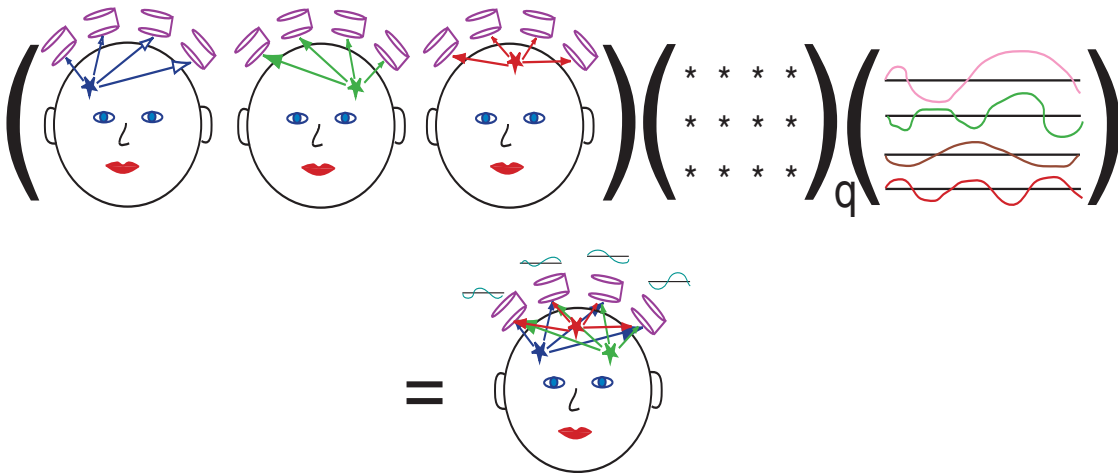
Figure 5.1: *Illustration of the coupled dipole model for the qth data set. The first matrix contains the forward fields of the basic sources (i.e. the contribution of the basic sources to each of the sensors), the second matrix is the coupling matrix for the qth data set and the last matrix is the basic time series matrix. The model for the qth data set, that is, the product of these three matrices, is a linear combination of the depicted basic components, as indicated in the lower picture.*

The CDM provides a general framework that can describe different situations in a flexible way by specifying which entries of C_q are zero and which entries have to be estimated from the data. A simple illustration would consist of two data sets ($Q = 2$), in which one source is active ($P = 1$). The source time function is different in both

data sets ($Z = 2$). In this example $A \in \mathbb{R}^{I \times 1}$ and $B \in \mathbb{R}^{2 \times J}$. Moreover the coupling matrices $C_1, C_2 \in \mathbb{R}^{1 \times 2}$.

$$A = (\ \mathrm{DF}_1 \ ) \ , B = \begin{pmatrix} \mathrm{STF}_1 \\ \mathrm{STF}_2 \end{pmatrix} \ , C_1 = (\ \alpha_1 \quad 0 \ ) \ , C_2 = (\ 0 \quad \alpha_2 \ ) \qquad (5.2)$$

The models for both data sets are

$$R_1 = AC_1B = (\ \mathrm{DF}_1 \ ) \times (\ \alpha_1 \mathrm{STF}_1 \ ) \qquad (5.3)$$

and

$$R_2 = AC_2B = (\ \mathrm{DF}_1 \ ) \times (\ \alpha_2 \mathrm{STF}_2 \ ). \qquad (5.4)$$

In the CDM the basic source parameters, the amplitudes in the coupling matrices and the nonparametric basic source time functions are estimated. For each data set, the linear combination of basic components, characterised by the coupling matrix, has to be specified by the user. Moreover the dimension of the coupling matrices (i.e. the number of basic components) has to be set. In other words, the dimensions and the zero elements of the coupling matrices are set by the user, while the nonzero elements (amplitude parameters) are estimated from the data. Fewer (more) assumptions regarding the similarity of the data sets are reflected by bigger (smaller) dimensions of the coupling matrices and/or fewer (more) zeroes in the coupling matrices. In the extreme case of no assumptions on common components, the dimension of the coupling matrices would be maximum (for each data set separately, spatial and temporal components would be estimated) and each coupling matrix would contain only a few, diagonal nonzero entries, coupling the corresponding spatial and temporal components. In the example (5.2) above, making no assumptions would lead to estimation of two sources ($P = 2$) and two STFs ($Z = 2$) and using two (2×2) coupling matrices:

$$A = (\ \mathrm{DF}_1 \quad \mathrm{DF}_2 \ ) \ , B = \begin{pmatrix} \mathrm{STF}_1 \\ \mathrm{STF}_2 \end{pmatrix} \ , C_1 = \begin{pmatrix} \alpha_1 & 0 \\ 0 & 0 \end{pmatrix} \ , C_2 = \begin{pmatrix} 0 & 0 \\ 0 & \alpha_2 \end{pmatrix}. \qquad (5.5)$$

The models for both data sets would then become

$$R_1 = AC_1B = (\ \mathrm{DF}_1 \ ) \times (\ \alpha_1 \mathrm{STF}_1 \ ) \qquad (5.6)$$

and

$$R_2 = AC_2B = (\ \mathrm{DF}_2 \ ) \times (\ \alpha_2 \mathrm{STF}_2 \ ). \qquad (5.7)$$

The nonzero elements in the coupling matrices are the amplitudes, which are estimated from the data. For all data sets q, the coupling matrix $C_q = C_q(\alpha)$ depends on $\alpha \in \mathbb{R}^{Y \times 1}$, the vector containing all coupling parameters (amplitudes). These coupling parameters determine the magnitude of the temporal components, while the basic STFs in B are normalised, i.e. each row in B has norm 1.

5.2.2 Probability Density Function

When brain activity is evoked by a stimulus, measured data consist, according to the Signal Plus Noise model, of a constant brain response and additional (internal and

external) noise [70] (cf. chapters 2 and 3). If R_q^k denotes the measured data matrix in trial k of data set q this can be formulated as

$$R_q^k = R_q + E_q^k \tag{5.8}$$

where we write R_q for the constant response matrix in data set q and E_q^k for the measured noise matrix. Furthermore, $vec(E_q^k)$ is assumed to have a Gaussian distribution, to be independent over trials $k = 1, \ldots, K_q$ and to have the Kronecker product of a spatial covariance matrix X and a temporal covariance matrix T as spatiotemporal covariance [16, 20, 42] where X and T are constant over k:

$$vec((E_q^k)^t) \sim \mathcal{N}(0, T \otimes X) \quad \text{for all } k. \tag{5.9}$$

In chapter 4 one of the conclusions was that this single Kronecker product model is adequate for dipole localisation. This yields the likelihood function for data set q (see [63] for handling formulas with Kronecker products):

$$L_q(X, T, \xi, \eta, \alpha, B) = \frac{\exp\left(-\frac{1}{2}tr[\sum_{k=1}^{K_q}(R_q^k - AC_qB)^t X^{-1}(R_q^k - AC_qB)T^{-1}]\right)}{(2\pi)^{\frac{IJK_q}{2}}|X|^{\frac{JK_q}{2}}|T|^{\frac{IK_q}{2}}}. \tag{5.10}$$

In practise, different data sets are measured in different trials, and therefore, the assumption of independency over trials implies independency over data sets too. Furthermore, X and T are assumed to be fixed over data sets. This yields the joint likelihood function for all data sets, $L(X, T, \xi, \eta, \alpha, B)$, which is the product over q of the likelihood functions L_q in (5.10):

$$L(X, T, \xi, \eta, \alpha, B) = \prod_{q=1}^{Q} L_q(X, T, \xi, \eta, \alpha, B) =$$

$$\frac{\exp\left(-\frac{1}{2}tr[\sum_{q=1}^{Q}\sum_{k=1}^{K_q}(R_q^k - AC_qB)^t X^{-1}(R_q^k - AC_qB)T^{-1}]\right)}{(2\pi)^{\frac{IJK}{2}}|X|^{\frac{JK}{2}}|T|^{\frac{IK}{2}}} \tag{5.11}$$

where $K = \sum_{q=1}^{Q} K_q$. By maximising (5.11) the maximum likelihood estimators for the noise parameters, X and T, and the signal parameters ξ, η, α and B are derived.

5.2.3   ML-Estimation procedure

The ML-estimators $\hat{X}$, $\hat{T}$, $\hat{\xi}$, $\hat{\eta}$, $\hat{\alpha}$ and $\hat{B}$ are derived from (5.11) by setting the corresponding derivative equal to zero and solving for the estimated parameters. Differentiation of (5.11) is performed using the rules derived in [63]. This yields a complicated system of equations for the estimators: all estimators are expressed in terms of each other, and have to be solved iteratively. The estimator for ξ is even more complex because ξ is a nonlinear parameter. To simplify this estimation procedure, the iterative system is split into two parts as explained in section 1.4.3. In the first, preparative, step the noise parameters are estimated and in the second the model parameters. In the case of known X and T (e.g. based on previous data sets) the first step is left out.

In the first step the expression AC_qB has to be replaced because the model parameters are not yet determined in that step. The substituting term is the ML-estimator for AC_qB as a whole:

$$(\widehat{AC_qB})_{ML} = \frac{1}{K_q} \sum R_q^k =: \bar{R}_q. \tag{5.12}$$

Substituting (5.12) in (5.11) and taking the derivative with respect to the noise parameters, X and T, yields (see appendix 5.5.2):

$$\hat{X}_{ML} = \frac{1}{JK} \sum_{q=1}^{Q} \sum_{k=1}^{K_q} (R_q^k - \bar{R}_q) \hat{T}_{ML}^{-1} (R_q^k - \bar{R}_q)^t \tag{5.13}$$

$$\hat{T}_{ML} = \frac{1}{IK} \sum_{q=1}^{Q} \sum_{k=1}^{K_q} (R_q^k - \bar{R}_q)^t \hat{X}_{ML}^{-1} (R_q^k - \bar{R}_q) \tag{5.14}$$

This system (5.13) and (5.14) is solved iteratively, starting with $T = \mathbf{I}_J$ in (5.13) until convergence of X and T. $\mathbf{I}_J$ denotes the identity matrix of dimension J. In the second step of the parameter estimation either the true X and T or the estimators $\hat{X}_{ML}$ and $\hat{T}_{ML}$, that are assumed to be the true covariances, are substituted in the likelihood function (5.11). For notational simplicity, the subscript ML will be omitted in the sequel. The likelihood has to be maximised with respect to ξ, η, B and α, which is equivalent to the minimisation of the (GLS) cost function $H(\xi, \eta, B, \alpha)$:

$$H(\xi, \eta, B, \alpha) = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} (R_q^k - AC_qB)^t \hat{X}^{-1} (R_q^k - AC_qB) \hat{T}^{-1}]$$

$$= tr[\sum_{q=1}^{Q} K_q (\bar{R}_q - AC_qB)^t \hat{X}^{-1} (\bar{R}_q - AC_qB) \hat{T}^{-1}] + \gamma, \tag{5.15}$$

where

$$\gamma = \sum_{q=1}^{Q} (\sum_{k=1}^{K_q} tr[(R_q^k)^t \hat{X}^{-1} R_q^k \hat{T}^{-1}] - K_q tr[\bar{R}_q^t \hat{X}^{-1} \bar{R}_q \hat{T}^{-1}]). \tag{5.16}$$

Clearly, γ does not depend on the model parameters, and the cost function $H(\xi, \eta, B, \alpha)$ may equally well be replaced by $\tilde{H}(\xi, \eta, B, \alpha)$:

$$\tilde{H}(\xi, \eta, B, \alpha) = tr[\sum_{q=1}^{Q} K_q (\bar{R}_q - AC_qB)^t \hat{X}^{-1} (\bar{R}_q - AC_qB) \hat{T}^{-1}]. \tag{5.17}$$

Setting the derivatives of $\tilde{H}$ with respect to η, B and α to zero, yields the following ML-estimators (the reader is referred to appendices (5.5.3 - 5.5.5) for the mathematical derivations):

$$\hat{B}_{ML} = (\sum_{q=1}^{Q} K_q C_q^t A^t \hat{X}^{-1} AC_q)^{-1} \sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q^t A^t \hat{X}^{-1} (R_q^k)^t \tag{5.18}$$

$$\Phi \, \hat{\eta}_{ML} = \phi \tag{5.19}$$

$$\Psi \, \hat{\alpha}_{ML} = \psi \tag{5.20}$$

where

$$\Phi_{p_1,p_2} = tr[\sum_{q=1}^{Q} K_q C_q B \hat{T}^{-1} B^t C_q^t \frac{\partial A^t}{\partial \eta_{p_2}} \hat{X}^{-1} \frac{\partial A}{\partial \eta_{p_1}}] \tag{5.21}$$

$$\phi_p = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B \hat{T}^{-1} (R_q^k)^t \hat{X}^{-1} \frac{\partial A}{\partial \eta_p}] \tag{5.22}$$

$$\Psi_{y_1,y_2} = tr[\sum_{q=1}^{Q} K_q B \hat{T}^{-1} B^t \frac{\partial C_q^t}{\partial \alpha_{y_2}} A^t \hat{X}^{-1} A \frac{\partial C_q}{\partial \alpha_{y_1}}] \tag{5.23}$$

$$\psi_y = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B \hat{T}^{-1} (R_q^k)^t \hat{X}^{-1} A \frac{\partial C_q}{\partial \alpha_y}]. \tag{5.24}$$

The source positions ξ are determined in a nonlinear search algorithm.

The dimensionality of the problem can be reduced by using the Singular Value Decomposition (SVD) of the data (cf. [20]). In order to take advantage of the vanishing, small eigenvalues of the data, the data are rearranged. For that purpose the following decompositions of the covariance matrices are used

$$\hat{X}^{-1} = W_X W_X^t \ , \ \hat{T}^{-1} = W_T W_T^t. \tag{5.25}$$

Furthermore, we define

$$\mathbf{R} := \begin{pmatrix} \sqrt{K_1} W_X^t \overline{R_1} W_T \\ \vdots \\ \sqrt{K_q} W_X^t \overline{R_q} W_T \\ \vdots \\ \sqrt{K_Q} W_X^t \overline{R_Q} W_T \end{pmatrix} \ , \ \mathbf{C} := \begin{pmatrix} \sqrt{K_1} C_1 \\ \vdots \\ \sqrt{K_q} C_q \\ \vdots \\ \sqrt{K_Q} C_Q \end{pmatrix} \ , \ \mathbf{A} := W_X^t A \ , \ \mathbf{B} := B W_T. \tag{5.26}$$

Then (5.17) can be rewritten as

$$\tilde{H}(\xi,\eta,B,\alpha) = tr[(\mathbf{R} - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB})^t (\mathbf{R} - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB})]. \tag{5.27}$$

Now the SVD of $\mathbf{R}$, containing the stacked prewhitened data of all data sets, is calculated

$$\mathbf{R} = U\Delta V^t \text{ with } \begin{cases} U \in \mathbb{R}^{IQ \times J} & U^t U = \mathbf{I}_J \\ V \in \mathbb{R}^{J \times J} & VV^t = V^t V = \mathbf{I}_J \\ \Delta \in \mathbb{R}^{J \times J} & \Delta = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_J) \end{cases} \tag{5.28}$$

and (5.27) is rewritten

$$\tilde{H}(\xi,\eta,B,\alpha) = tr[(U\Delta - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB}V)^t (U\Delta - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB}V)]. \tag{5.29}$$

The trace in (5.29) is split into terms corresponding to the first J_0 largest eigenvalues of $\mathbf{R}$ and the remaining $J - J_0$ terms:

$$\tilde{H}(\xi,\eta,B,\alpha) = \sum_{l=1}^{IQ} \sum_{j=1}^{J_0} (U\Delta - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB}V)_{l,j}^2 + \sum_{l=1}^{IQ} \sum_{j=J_0+1}^{J} (U\Delta - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB}V)_{l,j}^2.$$

$$\tag{5.30}$$

The dimensionality of the estimation problem is reduced by setting the $J - J_0$ small eigenvalues of $\mathbf{R}$ to zero and choosing $\mathbf{B}$ such that $[(\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB}V]_{l,j} = 0$ for $j > J_0$ for all l. Then the second term in (5.30) will vanish and the remaining cost function is

$$\tilde{H}(\xi, \eta, B, \alpha) = \sum_{l=1}^{IQ} \sum_{j=1}^{J_0} (U\Delta - (\mathbf{I}_Q \otimes \mathbf{A})\mathbf{CB}V)_{l,j}^2. \tag{5.31}$$

If the $J \times J$ truncated diagonal matrix is denoted by

$$\Delta_0 = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{J_0}, 0, \ldots, 0) \tag{5.32}$$

the estimators in (5.18) and (5.21) to (5.24) change accordingly into

$$\hat{\mathbf{B}}_{ML} = (\mathbf{C}^t(\mathbf{I}_Q \otimes \mathbf{A}^t\mathbf{A})\mathbf{C})^{-1}\mathbf{C}^t(\mathbf{I}_Q \otimes \mathbf{A^t})U\Delta_0 V^t \tag{5.33}$$

$$\Phi_{p_1,p_2} = tr[\mathbf{B}^t\mathbf{C}^t(\mathbf{I}_Q \otimes \frac{\partial \mathbf{A}^t}{\partial \eta_{p_2}})(\mathbf{I}_Q \otimes \frac{\partial \mathbf{A}}{\partial \eta_{p_1}})\mathbf{CB}] \tag{5.34}$$

$$\phi_p = tr[V^t\mathbf{B}^t\mathbf{C}^t(\mathbf{I}_Q \otimes \frac{\partial \mathbf{A}^t}{\partial \eta_p})U\Delta_0] \tag{5.35}$$

$$\Psi_{y_1,y_2} = tr[\mathbf{B}^t\frac{\partial \mathbf{C}^t}{\partial \alpha_{y_2}}(\mathbf{I}_Q \otimes \mathbf{A}^t)(\mathbf{I}_Q \otimes \mathbf{A})\frac{\partial \mathbf{C}}{\partial \alpha_{y_1}}\mathbf{B}] \tag{5.36}$$

$$\psi_y = tr[V^t\mathbf{B}^t\frac{\partial \mathbf{C}^t}{\partial \alpha_y}(\mathbf{I}_Q \otimes \mathbf{A}^t)U\Delta_0]. \tag{5.37}$$

Summarising, the estimation procedure looks like:

1. Compute $\bar{R}_q$ for all q and $\hat{X}$ and $\hat{T}$ using (5.13) and (5.14)

2. Perform a global search over source locations to obtain a starting point for the nonlinear (Marquardt) algorithm

3. Iterate until convergence of the cost function $\tilde{H}(\xi, \eta, B, \alpha)$:

   a) Obtain an update for the positions in ξ in the Marquardt algorithm (the first time, the starting point from the global search is taken)

   b) Iterate until convergence of $\tilde{H}(\xi, \eta, B, \alpha)$ for fixed ξ:

      i. Update B using (5.33)

      ii. Iterate until convergence of $\tilde{H}(\xi, \eta, B, \alpha)$ for fixed ξ and B:

         A. Update α using (5.20), (5.36) and (5.37)

         B. Update η using (5.19), (5.34) and (5.35).

In the global search in step (2) a regular grid with locations is computed. For each location, step (3b) is executed and the converging value of the cost function $\tilde{H}$ for that location is computed. The location with the minimum value of the cost function is taken as starting point in step (3). In step (2), alternative initialisation procedures can be used, as outlined in [103].

	position source left (cm)			orientation source left		
	x	y	z	x	y	z
somatosensory	1.63	3.80	3.18	-0.83	0.47	-0.29
visual	-6.37	1.80	-2.82	-0.44	0.22	0.87
auditive	0.63	4.30	-0.82	0.00	0.00	1.00

Table 5.1: *Locations of simulated sources in simulation 1. The positions of the sources are relative to the centre of the spherical volume conductor. Positions and orientations of the sources were taken symmetric in this simulation study, i.e. with opposite y-coordinates. The direction of the x-axis is forward, the y-axis is to the left and the z-axis points upwards.*

5.3   Results

The coupled dipole model was applied in two simulation studies and to one experimental data set. In the first simulation study a symmetric dipole pair was simulated representing three different functional areas: the somatosensory cortex, the auditory cortex and the visual cortex. In the second simulation study data from two dipoles in the visual cortex in one hemisphere were generated in different ratios of activity. The experimental data consisted of Visual Evoked Field (VEF) MEG data. The visual stimulus in this experiment consisted of a checkerboard pattern, presented either in one hemi-field or full field to the subject.

5.3.1   Simulation Study 1

In the first simulation study activity from two single dipole sources was generated. Three surrogate data sets were produced: in the first data set the left source was simulated, in the second the right source and the third data set contained simulated data from both sources. The locations of the sources were taken symmetric and varied over the visual, auditory and somatosensory cortices. True locations of these cortices were based on experimentally located positions, see Table 5.1. Two basic source time functions were used to generate the data, see (5.40) and (5.41). For each location 100 sweeps were generated. The sample rate used was 625 Hz, and a time window of 32 ms was analysed. All data sets consisted of simulated dipole activity and additional white noise with varying signal-to-noise ratio, SNR equal to $\frac{1}{3}$, 1, 3 and 9. The SNR is defined as the ratio between the matrix powers of the surrogate signal and the surrogate white noise:

$$SNR = \frac{tr(R_{sur}^t R_{sur})}{tr(E_{sur}^t E_{sur})}. \tag{5.38}$$

The first data set contained activity from the left source with STF1, the second activity from the right source with STF1, and the third data set contained activity from both sources, both having STF2.

The matrix $A(\xi, \eta)$ contained the forward fields of both dipoles, and B contained

the two normalised basic STFs. The coupling matrices for the three data sets were

$$C_1 = \begin{pmatrix} \alpha_1 & 0 \\ 0 & 0 \end{pmatrix} , C_2 = \begin{pmatrix} 0 & 0 \\ \alpha_2 & 0 \end{pmatrix} , C_3 = \begin{pmatrix} 0 & \alpha_3 \\ 0 & \alpha_4 \end{pmatrix} \quad (5.39)$$

Both input STFs were 10 Hz sinusoids:

$$\text{STF}_1(t) = \frac{1}{n_1} \sin(20\pi t) \quad (5.40)$$

$$\text{STF}_2(t) = \frac{1}{n_2}(\sin(20\pi t - \frac{\pi}{4}) + 0.35) \quad (5.41)$$

where n_1 and n_2 are the normalisation constants, $n_1 = \sqrt{11.5003}$ and $n_2 = \sqrt{10.1852}$. The following coupling parameter values were used

$$\alpha_1 = 1000n_1 \quad (5.42)$$
$$\alpha_2 = 1000n_1 \quad (5.43)$$
$$\alpha_3 = 1000n_2 \quad (5.44)$$
$$\alpha_4 = 1000n_2 \quad (5.45)$$

The absolute values of the magnitudes do not influence the performance of the different localisation methods. Only the relative amount of noise, indicated by the SNR, counts. Therefore, the α-values were chosen such that the absolute magnitudes were comfortable in the analysis. In Figure 5.2 the input STFs are plotted as function of time.
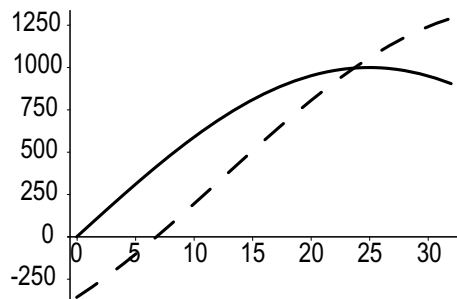
Figure 5.2: *The two basic input STFs for simulation 1 multiplied by 1000, which is the amplitude in all data sets. The solid line indicates STF1, and the dashed line STF2.*

Four different source localisation analyses were performed: the three data sets in three separate analyses as presented in [20] and all data sets simultaneously in the CDM. In the separate analyses the estimation of the source parameters was performed based on the simulated number of sources: one source was localised in data sets 1 and 2, and two sources were estimated in data set 3. The grid for the global search consisted of 100 locations, resulting in 100 possible starting points for data sets 1 and 2 and 100*99/2 = 4950 possible starting combinations for data set 3. In the CDM analysis, two sources, two STFs and the nonzero elements of the coupling matrices in (5.39) were estimated. The same grid was used for the global search, yielding 100*99 =

9900 possible starting combinations of the two sources, because the order of the sources is important in the CDM. Average errors in position and orientation of the estimated dipoles as well as in magnitude of the estimated source time functions were computed for all four analyses. The averages were taken over the sweeps and over both sources (position and orientation) or both STFs (magnitude).

For all three cortex locations the errors show a similar pattern: the integrated analysis yields a lower error than the separate analyses. Depending on the distance between the sources, this difference in performance varies. The errors in position, orientation and magnitude show a similar pattern, therefore only the position error is shown in Figure 5.3. In Figure 5.3(a) the average position error for the simulated dipoles in the somatosensory cortex is shown, in Figure 5.3(b) for the auditory cortex and in Figure 5.3(c) for the visual cortex. Clearly, the difference in performance between the separate approach and the CDM is largest for two active dipoles in the visual cortex (Figure 5.3(c)). The reason for this, is that the (lateral) distance between the simulated visual dipoles is smallest. Therefore, localising the two sources in the classical way will be hampered by the reasons mentioned in the introduction. For all analyses Figure 5.3 shows an improvement (slightly) in position error for the integrated model, indicating that adding more data into the parameter estimation is advantageous.

5.3.2   Simulation Study 2

In the second simulation study two sources in the visual cortex, one in the striate and one in the extrastriate cortex, were simulated in five different ratios of activity. Positions and orientations were taken from experimentally located sources, resulting in the sources being 3 cm apart from each other (Table 5.2). As in simulation 1 data consisted of simulated dipole activity and white noise with varying SNR, SNR $= \frac{1}{3}$, 1, 3 and 9. Each data set contained 100 sweeps, the sample frequency was 625 Hz and a time window of 32 ms was analysed. The input source time functions were two sinusoids, 15 Hz for the striate source and 20 Hz for the extrastriate source

$$\text{STF}_1(t) \quad = \quad \frac{1}{n_3} \sin(30\pi t) \tag{5.46}$$

$$\text{STF}_2(t) \quad = \quad \frac{1}{n_4} \sin(40\pi t). \tag{5.47}$$

where $n_3 = \sqrt{10.4013}$ and $n_4 = \sqrt{8.49828}$ are the normalisation constants. Figure 5.4 shows the two input STFs. The five data sets were generated using the coupling matrices

$$C_1 = \begin{pmatrix} 1000n_3 & 0 \\ 0 & 0 \end{pmatrix}, C_2 = \begin{pmatrix} 750n_3 & 0 \\ 0 & 250n_4 \end{pmatrix}, C_3 = \begin{pmatrix} 500n_3 & 0 \\ 0 & 500n_4 \end{pmatrix},$$

$$C_4 = \begin{pmatrix} 250n_3 & 0 \\ 0 & 750n_4 \end{pmatrix}, C_5 = \begin{pmatrix} 0 & 0 \\ 0 & 1000n_4 \end{pmatrix}. \tag{5.48}$$

Six source localisation analyses were performed: the five data sets separately and all five data sets in the integrated model. Moreover, in the CDM analysis, diagonal coupling matrices with two unknowns were used for all data sets. In other words, in the CDM

(a) Somatosensory cortex

(b) Auditory cortex

(c) Visual cortex
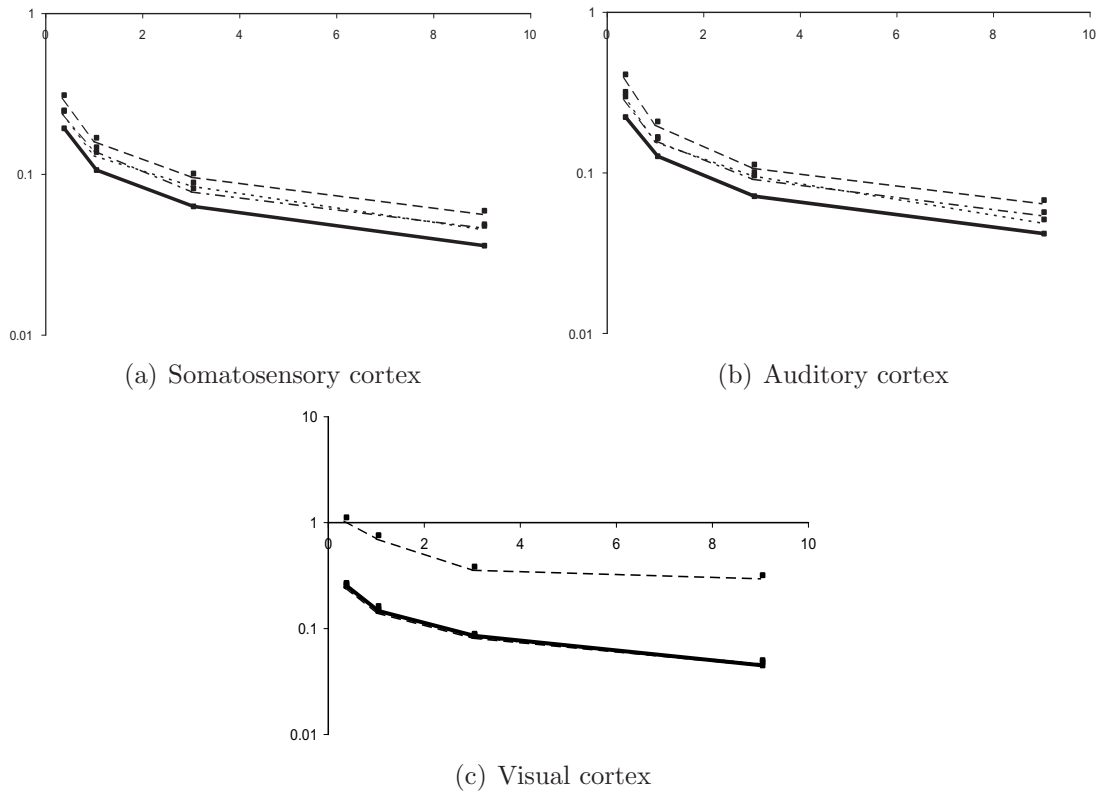
Figure 5.3: *Average position errors of the localised sources in the different cortices in simulation 1 as function on SNR (horizontal axis). The errors are given in cm on a logarithmic scale. The solid line represents the errors of the CDM results. The other lines correspond to the three separate analyses: both sources (dashed line), left source (dotted line), right source (dot-dashed line). The latter two lines approximately coincide in all figures.*
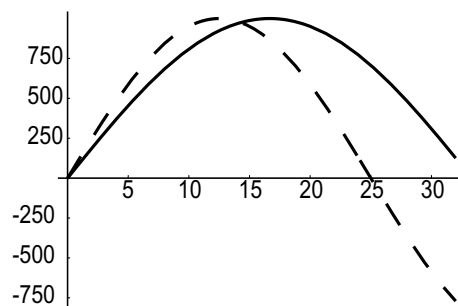
Figure 5.4: *The two basic input STFs for simulation 2 multiplied by 1000, the maximum amplitude (data sets 1 and 5). The solid line indicates STF1 and the dashed line STF2.*

analysis no advantage was taken of the knowledge that data sets 1 and 5 contain activity of only one dipole, while this information was exploited in the separate analyses. For the global search the same grid as in the first simulation study was used. This resulted

	position (cm)			orientation		
	x	y	z	x	y	z
striate	-6.37	1.30	-1.82	-0.15	0.59	0.79
extrastriate	-5.37	3.30	-3.82	0.70	0.60	-0.40

Table 5.2: *Locations of simulated sources in simulation 2. The positions of the sources are relative to the centre of the spherical volume conductor. The direction of the x-axis is forward, the y-axis is to the left and the z-axis points upwards.*

SNR	$\frac{1}{3}$	1	3	9
data set 1	0	0	0	0
data set 2	6	1	1	1
data set 3	0	2	1	4
data set 4	16	16	19	18
data set 5	0	0	0	0
CDM	0	0	0	0

Table 5.3: *The number of outliers (sweeps that fulfill one of the conditions (5.49 - 5.51)) in the six source localisations in simulation study 2.*

in 100 possible starting locations for the separate analyses of data sets 1 and 5, 4950 possible combinations for data sets 2, 3 and 4 and 9900 possible combinations for the CDM (as the order of the sources is of importance in the simultaneous model).

Outliers in the estimated sources were discarded. The outlier criteria used are

$$\text{position error} \quad > \quad 7 \text{ cm} \tag{5.49}$$

$$\text{distance between sources} \quad < \quad 0.5 \text{ cm} \tag{5.50}$$

$$\text{magnitude error (relative power)} \quad > \quad 10 \tag{5.51}$$

The first criterium designates localised sources in implausible positions, which would be discarded in experimental analysis. The latter two criteria often concur and indicate two coinciding sources, usually having opposite orientations and unreasonably high magnitudes, as discussed in the introduction. In Table 5.3 the numbers of outliers in the different localisations are given: only the three data sets containing two active sources yielded outliers in the separate analyses. The simultaneous CDM did not yield any outliers.

Average position errors, orientation errors and magnitude errors were computed for all six analyses. The averages were taken over the 100 sweeps (minus the outlying sweeps) and over both sources and both STFs.

As in the first simulation, the graphs of the three types of error resemble each other. Figure 5.5 demonstrates the position errors for all analyses. The errors in the separate analyses can be divided into two groups; data set 1 and 5, which contain only one active source, show a lower error than the remaining three data sets. This reflects the common feature that one source is easier localised than two sources are. Nevertheless, the CDM
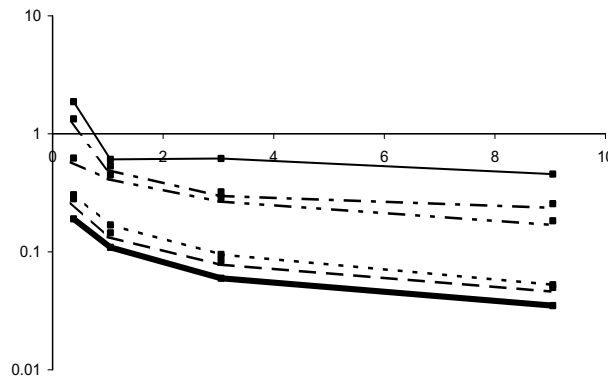
Figure 5.5: *The average position error of the localised sources in simulation 2 as function of SNR. The dashed line corresponds to data set 1, the dot-dashed line to data set 2, the dot-dot-dashed line to data set 3, the thin solid line to data set 4 and the dotted line to data set 5. The thick solid line represents the CDM errors. The errors are given in cm on a logarithmic scale.*

yields the lowest position error. Compared to the separate analyses, using the CDM the position error drops approximately by a factor of 10 for data sets 2, 3 and 4. For data sets 1 and 5 there is still an improvement of roughly a factor of 1.5 in the position error.

5.3.3   Experimental data

The CDM was applied to experimental MEG-data of a Visual Evoked Field experiment. The visual stimuli consisted of checkerboard pattern onsets, presented either in the left or the right hemi-field or full field, thus yielding three data sets per subject. The separate analyses of this kind of conditions show nearby localised sources with similar orientations for the different conditions, indicating that the assumption of common sources is plausible. Moreover, the estimated orientations are opposite [22, 48]; therefore, the solutions to the separate IPs are susceptible for the cancelling dipole problem. For this reason, the behaviour of the separate analyses and the CDM is compared using these visual data.

Data of 5 subjects were considered in this study. The time window of analysis was set to 80-112 ms post-stimulus. The sample rate was 625 Hz, the number of trials was 400 in each data set. Similar to simulation 1, the three data sets were analysed separately, yielding three sets of estimated parameters, and the CDM was applied to all three data sets simultaneously, yielding one solution with estimated common spatial and common temporal parameters.

The check size was 6'. Checks of this size mainly activate the striate cortex in the chosen time interval [22, 81]. As expected, hemi-field stimuli first activate the contralateral hemisphere, yielding a peak at around 90-95 ms. However, between 10-15 ms [91] or 20 ms later [98] the ipsilateral hemi-sphere is activated as well by interhemispheric transfer. Hence, to describe the hemi-field data two sources are needed, having in the

CDM two different STFs: a contralateral STF and an ipsilateral STF. The contralateral STF was also used to describe the time courses of both sources in the bilateral data set, because these stimuli will activate both hemispheres directly [48], although maybe even faster [98]. This yielded the following basic and coupling matrices:

$$A = \left(\begin{array}{cc} \mathrm{DF}_{left} & \mathrm{DF}_{right} \end{array} \right) \;, B = \left(\begin{array}{c} \mathrm{STF}_{contra} \\ \mathrm{STF}_{ipsi} \end{array} \right) \tag{5.52}$$

$$C_{lh} = \left(\begin{array}{cc} 0 & \alpha_2 \\ \alpha_1 & 0 \end{array} \right) \;, C_{rh} = \left(\begin{array}{cc} \alpha_3 & 0 \\ 0 & \alpha_4 \end{array} \right) \;, C_{ff} = \left(\begin{array}{cc} \alpha_5 & 0 \\ \alpha_6 & 0 \end{array} \right), \tag{5.53}$$

where 'lh' corresponds to the left hemi-field stimulation, 'rh' to the right hemi-field stimulation, and 'ff' to the full field stimulation.

In all the analyses, a semi-symmetric source pair was fitted. Semi-symmetric sources have symmetric locations (as in [48] and [22]), but their orientations are free. Nevertheless, the estimated orientations are nearly symmetric (Table 5.4).

The number of locations in the regular grid for the global search for the starting location of the semi-symmetric dipole source was varied over 23, 102 and 466 points in one hemisphere, corresponding to mesh sizes of 4.2 cm, 2.6 cm, 1.5 cm respectively. The solutions to the separate IPs changed when this number is increased from 23 to 102 grid points, but remained the same for the increase from 102 to 466 points. The simultaneous model, though, yielded the same solutions for all three grids. This shows that the CDM is less sensitive to local minima in the cost function.

The results of the separate analyses varied considerably over subjects. For subjects 1 and 2 the separate analyses yielded plausible solutions for all three conditions for both 23 and 102 grid points. For subject 3 all three separate solutions to the IP consisted of a pair of cancelling, coinciding sources (intermediate distance < 0.01 cm). Even using a global search with 466 grid points yielded these implausible solutions for all three separate models. Subject 4 yielded for the right hemi-field data set a physiologically plausible solution for all grids. For the coarse grid, the sources in the full field data set were localised in the cerebellum. For 102 grid points this problem was resolved and the solutions were localised in the visual cortex. The locations of the sources in the left hemi-field data set showed a plausible location, though the ipsi-lateral orientation differs from the usual lateral direction for all grids. The corresponding amplitude is very small; therefore, the total influence of this orientation on the cost function is very small. Subject 5 (see Table 5.4) yielded a solution consisting of a cancelling dipole pair for the full field data set when the coarse grid was used. For the denser grids, this data set yielded an interpretable solution. The right hemi-field data set yielded a location in the visual cortex for 23 grid points, though the ipsilateral orientation was unusual. For 102 and 466 grid points these sources were localised nearer to the midsagittal plane (1.2 cm in between). For all grids left hemi-field data set yielded plausible solutions. The CDM analysis yielded for all five subjects plausible solutions for all conditions for all grids. Summarising, for the coarse grid, 5 of the 15 separate analyses resulted in an implausible solution and for the denser grids, 3 of the 15 conditions did not yield an interpretable solution, whereas the simultaneous model only yielded plausible solutions.

The results of subject 5 are representative for the type of errors that can occur in the separate solutions, therefore the analyses of this subject for the coarse (23 points)

	position left (cm)			orientation left			orientation right		
	x	y	z	x	y	z	x	y	z
LH 23	-5.21	1.29	0.94	0.26	0.96	0.11	0.26	-0.96	0.11
RH 23	-6.14	1.21	0.17	0.17	0.76	0.62	0.00	0.12	0.99
FF 23	-4.42	0.00	-1.15	0.25	-0.05	-0.97	-0.25	0.05	0.97
LH 102	-5.21	1.29	0.94	0.26	0.96	0.11	0.26	-0.96	0.11
RH 102	-5.72	0.32	0.11	0.06	0.88	0.47	0.05	-0.96	-0.27
FF 102	-5.14	1.43	0.89	0.31	0.77	0.56	0.28	-0.96	0.08
CDM	-5.25	1.44	0.85	0.30	0.72	0.62	0.28	-0.95	0.12

Table 5.4: *Estimated sources for subject 5 in the experimental study. The positions of the sources are relative to the centre of the spherical volume conductor. The direction of the x-axis is forward, the y-axis is to the left and the z-axis points upwards. The positions of the two sources have opposite y-coordinates. LH (RH, FF) indicates the left hemi- (right hemi-, full) field stimulation data set.*

and the dense (102 points) grid and the CDM solution are presented. The estimated locations and orientations of the sources for the seven different localisations are reported in Table 5.4. The estimated STFs are shown in Figures 5.6, 5.7 and 5.8. Figure 5.6 shows the STFs estimated in the separate analyses using the coarse grid, Figure 5.7 the estimated STFs for the dense grid and Figure 5.8 displays the estimated STFs for the simultaneous model, that did not change with the number of grid points.

In Figure 5.6(c) an example of the cancelling dipole problem is shown: the magnitudes are unrealistically high, the corresponding locations in Table 5.4 coincide and the estimated orientations are opposite. For the coarse grid the RH data set does not yield the usual lateral orientation for the ipsilateral source. This orientation does not have a substantial influence on the cost function, because the corresponding amplitude (Figure 5.6(b)) is rather small. The amplitudes in the different analyses for the RH data set vary. In Figure 5.7(b) (dense grid) the estimated amplitudes are larger than the amplitudes shown in Figure 5.6(b) (coarse grid) and 5.8(b) (CDM). The corresponding distance to the midsagittal plane for the dense grid (0.32 cm) is smaller than for the coarse grid and the CDM (1.21 cm and 1.44 cm respectively). This may indicate a slight cancellation of dipole activity: the sources are closer and the amplitudes are higher for the dense grid.

5.4   Discussion

The coupled dipole model provides a method to solve the Inverse Problem by analysing multiple MEG data sets simultaneously when these data sets contain common sources or common source time functions. This way, more data are used and more constraints (assumptions) are applied in order to solve and stabilise the IP.

The results of the first simulation show that the position error decreases when the integrated model is used in stead of the separate models. The gain in accuracy depends

(a) Left hemi-field stimulation

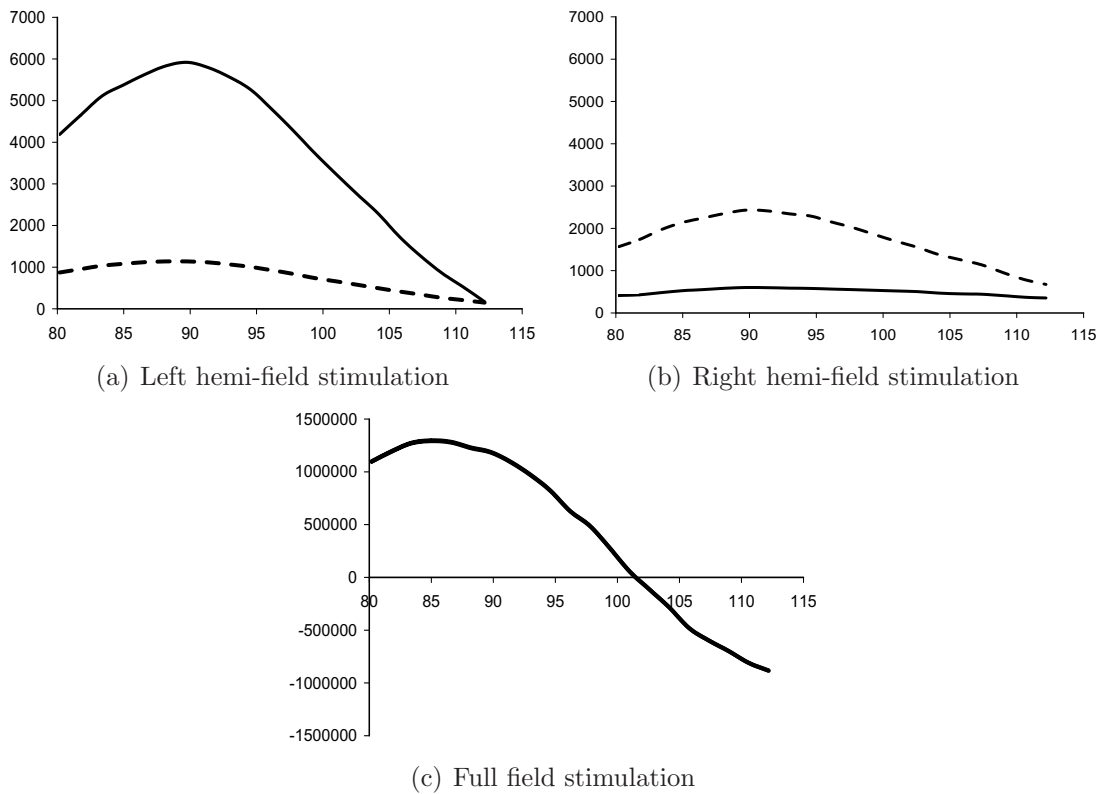(b) Right hemi-field stimulation

(c) Full field stimulation

Figure 5.6: *Estimated STFs of the semi-symmetric sources resulting from the separate analyses using a global search with 23 grid points in the experimental study (subject 5). The vertical axis represents amplitude in nAcm. The horizontal axis denotes time in ms after stimulus. The solid (dashed) lines represents the amplitude of the left (right) source. Note the difference in scaling, in particular in figure c.*

on the locations (intermediate distance) of the sources to be localised and was largest for the simulated sources in the visual cortex.

Simulation 2 displays a considerable improvement in source localisation for the CDM, studying two active dipoles in the visual cortex in five different ratios of activity. For one of the five data sets considered, a quarter striate and three quarters extrastriate activity, the position error drops by a factor of 10 when the CDM is used in stead of a single model. Moreover, outliers only occurred in the separate analyses of data sets 2, 3 and 4, while the CDM did not yield any outliers (see Table 5.3). In Figure 5.5 it can be seen that the position error of data set 4 is higher for SNR=3 than it is for SNR=1. A similar feature was found in the orientation and magnitude error. This is contradicting the fact that for higher SNR the error should decrease. The reason for this slight increase in error lies in the outlier problem. For all SNR values in data set 4, some sweeps were marked as outlier (defined by criteria (5.49 - 5.51)) and discarded. However, some other sweeps, not indicated as outlier, also showed a rather large position error (> 2 cm), but within the outlier criterium of 7 cm. For SNR=1 the number of these semi-outlying sweeps was lower than for SNR=1/3 and SNR=3. Because these sweeps raise the average error considerably, this difference between SNR=1

(a) Left hemi-field stimulation

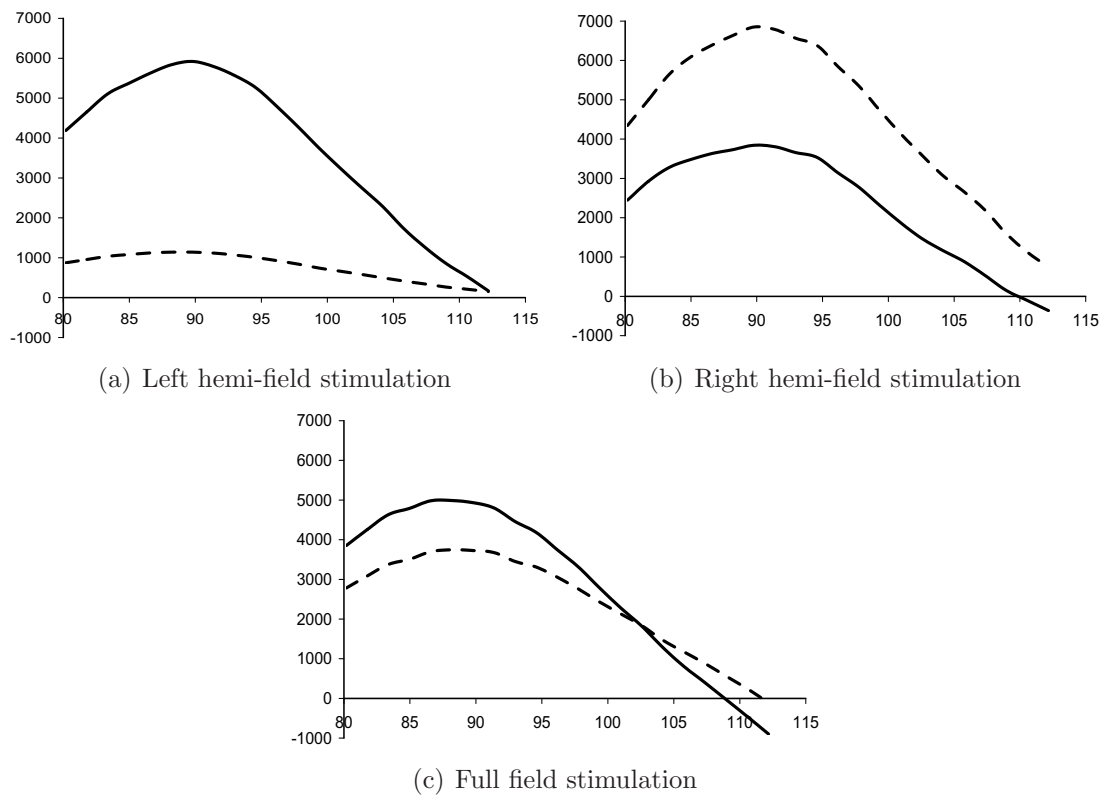(b) Right hemi-field stimulation

(c) Full field stimulation

Figure 5.7: *Estimated STFs of the semi-symmetric sources resulting from the separate analyses using a global search with 102 or 466 grid points in the experimental study (subject 5). The vertical axis represents amplitude in nAcm. The horizontal axis denotes time in ms after stimulus. The solid (dashed) lines represents the amplitude of the left (right) source.*

and the other SNR-values explains the error shape in Figure 5.5. Increasing the number of sweeps substantially or redefining the outlier criteria should eliminate this problem. This shows that the choice of outlier criteria is delicate. Nonetheless, this outlier problem is not present in the CDM (Table 5.3), which is an important advantage of the integrated approach regarding the stability of the solution.

The application of the CDM to experimental VEF data showed that in 5 (coarse grid) or 3 (dense grid) of the 15 conditions considered, no interpretable solution was obtained without using the CDM . The classic (separate) analysis of all but one of these conditions yielded a cancelling dipole source pair and one condition yielded a localised visual source pair in the cerebellum (coarse grid). The CDM yielded in all cases plausible solutions. This demonstrates that the CDM is not sensitive for minima in the cost function that correspond to a cancelling dipole pair. The underlying reason may be that this kind of minima of the separate cost functions occur at different locations for different conditions and will therefore not produce a minimum in the joint cost function.

Another advantage of the simultaneous model is the direct comparability of the estimated magnitudes. Because all parameters are estimated in one analysis, the mag-

(a) Left hemi-field stimulation

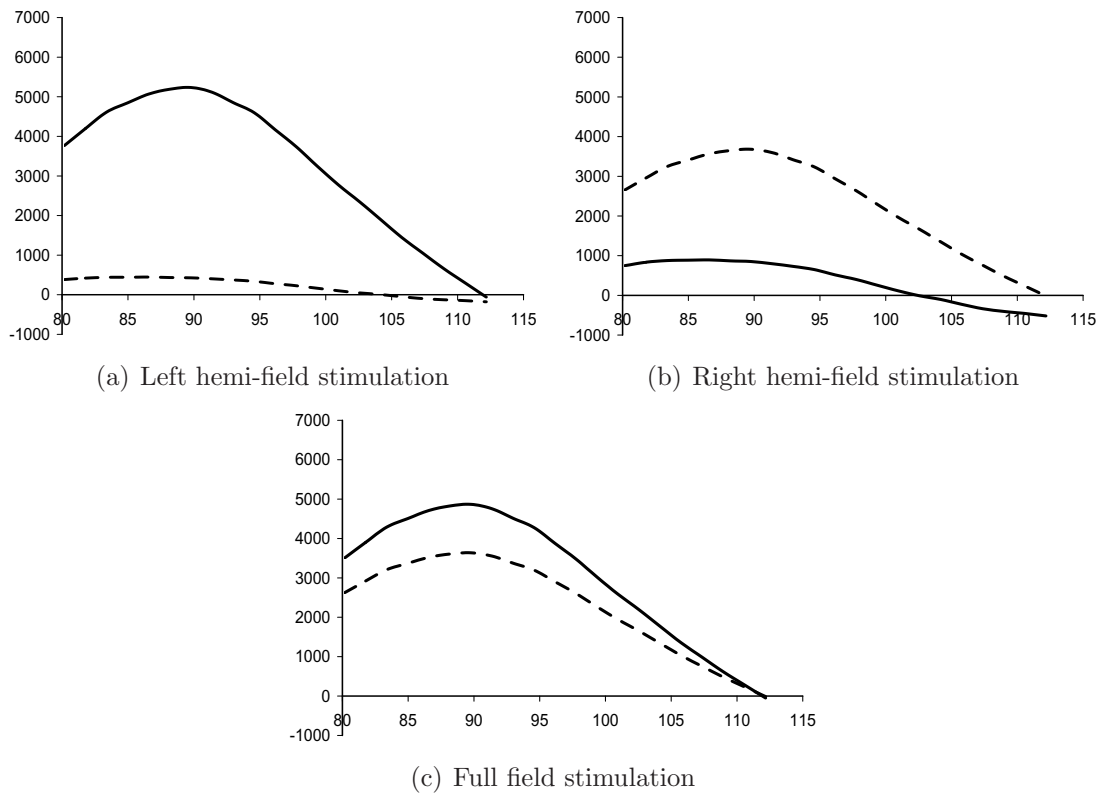(b) Right hemi-field stimulation

(c) Full field stimulation

Figure 5.8: *Estimated STFs of the semi-symmetric sources resulting from the CDM analysis in the experimental study (subject 5). The vertical axis represents amplitude in nAcm. The horizontal axis denotes time in ms after stimulus. The solid (dashed) lines represents the amplitude of the left (right) source.*

nitudes between conditions can directly be compared (Figure 5.8) while for the separate analyses (Figures 5.6 and 5.7) the different conditions use a different scaling and it is difficult to compare the absolute magnitudes between conditions. This difference in scaling is caused by the depth bias (the mutual dependence of depth and amplitude) and possible slight cancellation of the sources (cf. RH for the coarse grid) and the differences in STF shapes between conditions.

Summarising, the experimental application showed that simultaneously estimating the amplitudes and locations in the three data sets yields more stable and comparable estimated parameters, while separate models are more vulnerable for implausible minima and yield in general estimated magnitudes and positions which vary over data sets.

The estimation procedure for the signal parameters presented in this paper is computationally intensive, due to the nested iterations in the estimation procedure. Modifying the setup of the coupling matrices may lead to a less intensive estimation procedure. In the next chapter an extension of this model is presented, in which the user does not select the zero and nonzero entries of the coupling matrix.

Designing the coupling matrices is subject to the choice (or a priori knowledge) of the user. Different users may want to design the coupling matrices in a different way.

This is both a flexibility property of the model and a subjectivity of the CDM. The user can make more or less assumptions by adapting the dimension of the coupling matrices, as was shown in the example in equations (5.2) and (5.5). Such assumptions, in fact, are also made when similar data sets are analysed using separate analyses: results are compared afterwards and conclusions are drawn about the similarity of sources and source time functions in the different data sets. The CDM now provides a way to put this knowledge a priori in the source localisation method, which relatively increases the SNR. If the user is not sure about certain similarities, the coupling matrices should be extended to less assumptions. In the second simulation study the robustness of the CDM was tested by applying less assumptions than possible (fitting two sources when only one was active in data set 1 and 5), and it was shown that the CDM still estimates the sources and STFs accurately.

A possible extension of the CDM would be the simultaneous analysis of data sets of different subjects and/or different conditions. If it is assumed that the spatial and temporal covariances differ only slightly over subjects and conditions (cf. chapter 3), and can therefore be fixated, this extension is rather straightforward. This kind of extension will be helpful if one is interested in grand averages of e.g. amplitude functions or source positions as in cognitive neuroscience [22, 48].

In sum, the CDM combines multiple data sets and extra constraints into one simultaneous analysis. This integrated approach is powerful in order to solve and stabilise the ill-posed Inverse Problem in MEG/EEG, yielding interpretable solutions in cases where separate models not always yield plausible solutions.

5.5   Appendix

5.5.1   Dimensions and variables

The dimensions are defined as

- I - # sensors

- J - # time samples

- K_q - # trials in data set q

- K - # trials in all data sets

- P - # basic sources

- Q - # data sets

- Y - # amplitude parameters

- Z - # basic STFs

and the variables as

- α - $(Y \times 1)$ vector containing amplitude parameters

- η - $(3P \times 1)$ vector containing source orientations

- ξ - $(3P \times 1)$ vector containing source locations

- $A(\xi, \eta)$ - $(I \times P)$ forward field matrix of basic dipoles

- **A** - $(I \times P)$ prewhitened forward field matrix

- B - $(Z \times J)$ basic STF matrix, with basic STFs as rows

- **B** - $(Z \times J)$ prewhitened STF matrix

- C_q - $(P \times Z)$ coupling matrix for data set q

- **C** - $(PQ \times Z)$ stacked coupling matrices

- E_q^k - $(I \times J)$ noise matrix of trial k in data set q

- $\mathbf{I}_n$ - $(n \times n)$ identity matrix

- R_q^k - $(I \times J)$ data matrix of trial k in data set q

- R_q - $(I \times J)$ model matrix for data set q

- $\bar{R}_q$ - $(I \times J)$ average data matrix for data set q

- **R** - $(IQ \times J)$ stacked average prewhitened data matrices

5.5.2   ML-estimators for X and T

This derivation uses the following two matrix derivatives ([63], Chapter 9)

$$d_X(|X|) = |X|tr(X^{-1}dX)$$

and

$$d_X(X^{-1}) = -X^{-1}dXX^{-1}.$$

The derivative of $L(X, T, \xi, \eta, \alpha, B)$ in (5.11) with respect to X is

$$d_X L = L \times \left[\frac{-\frac{JK}{2}d|X|}{|X|} + d_X\left(-\frac{1}{2}tr\left[\sum_{q=1}^{Q}\sum_{k=1}^{K_q}(R_q^k - \bar{R}_q)^t X^{-1}(R_q^k - \bar{R}_q)T^{-1}\right]\right)\right]$$

$$= \frac{1}{2}L \times \left[\frac{-JK|X|tr(X^{-1}dX)}{|X|} + tr\left[\sum_{q=1}^{Q}\sum_{k=1}^{K_q}(R_q^k - \bar{R}_q)^t X^{-1}dXX^{-1}(R_q^k - \bar{R}_q)T^{-1}\right]\right]$$

$$= \frac{1}{2}L \times \left[tr\left([-JK + \sum_{q=1}^{Q}\sum_{k=1}^{K_q}X^{-1}(R_q^k - \bar{R}_q)T^{-1}(R_q^k - \bar{R}_q)^t]X^{-1}dX\right)\right]. \qquad (5.54)$$

The optimal spatial covariance matrix, $\hat{X}_{ML}$, is obtained when $d_X L = 0$, i.e. when

$$[-JK + \sum_{q=1}^{Q}\sum_{k=1}^{K_q} \hat{X}_{ML}^{-1}(R_q^k - \bar{R}_q)T^{-1}(R_q^k - \bar{R}_q)^t]\hat{X}_{ML}^{-1} = 0$$

$$\Longleftrightarrow \hat{X}_{ML} = \frac{1}{JK}\sum_{q=1}^{Q}\sum_{k=1}^{K_q}(R_q^k - \bar{R}_q)T^{-1}(R_q^k - \bar{R}_q)^t, \tag{5.55}$$

provided that $\hat{X}_{ML}$ is nonsingular. Similarly for T the estimator becomes

$$\hat{T}_{ML} = \frac{1}{IK}\sum_{q=1}^{Q}\sum_{k=1}^{K_q}(R_q^k - \bar{R}_q)^t X^{-1}(R_q^k - \bar{R}_q). \tag{5.56}$$

5.5.3   ML-estimator for B

Although the rows in B are normalised, the ML-estimator for B is derived as if B were unconstrained. The rows are normalised afterwards.

$$d_B L =$$

$$= \frac{1}{2}L \times tr[\sum_{q=1}^{Q}\sum_{k=1}^{K_q}(AC_q dB)^t X^{-1}(R_q^k - AC_q B)T^{-1} + (R_q^k - AC_q B)^t X^{-1} AC_q dB T^{-1}]$$

$$= L \times tr[\sum_{q=1}^{Q}\sum_{k=1}^{K_q} T^{-1}(R_q^k - AC_q B)^t X^{-1} AC_q dB]. \tag{5.57}$$

Setting this derivative to zero yields the estimator for B:

$$\sum_{q=1}^{Q}\sum_{k=1}^{K_q} T^{-1}(R_q^k - AC_q \hat{B}_{ML})^t X^{-1} AC_q = 0$$

$$\Longleftrightarrow \sum_{q=1}^{Q}\sum_{k=1}^{K_q}(R_q^k)^t X^{-1} AC_q = \sum_{q=1}^{Q}\sum_{k=1}^{K_q}(AC_q \hat{B}_{ML})^t X^{-1} AC_q$$

$$\Longleftrightarrow \hat{B}_{ML} = (\sum_{q=1}^{Q} K_q C_q^t A^t X^{-1} AC_q)^{-1}\sum_{q=1}^{Q}\sum_{k=1}^{K_q} C_q^t A^t X^{-1}(R_q^k)^t, \tag{5.58}$$

provided that $\sum_{q=1}^{Q} K_q C_q^t A^t X^{-1} AC_q$ is invertible. After this estimator has been calculated, the STFs in B are normalised.

5.5.4   ML-estimator for orientation parameters η in A

The source parameters in A can be split into the linear orientation parameters, η, and the nonlinear position parameters, ξ. If all P sources are modelled as single dipoles or symmetric source pairs, η and ξ will both be $3P \times 1$ vectors. If P semi-symmetric source pairs (symmetric location, free orientation) are used, η will be $6P \times 1$ and ξ will be $3P \times 1$. For notational clarity η and ξ will both be assumed to be $3P \times 1$. For the case of semi-symmetric sources, the dimensions have to be adjusted correspondingly. The linear parameters η are estimated in the usual ML sense; that is, equating the derivative of L equal to zero. The nonlinear parameters ξ are determined using the Marquardt algorithm. The derivative of L with respect to η is calculated by first computing $d_A L$ and subsequently applying the chain rule.

$$d_A L =$$

$$= \frac{1}{2} L \times tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} (dAC_q B)^t X^{-1} (R_q^k - AC_q B) T^{-1} + (R_q^k - AC_q B)^t X^{-1} dAC_q B T^{-1}]$$

$$= L \times tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B T^{-1} (R_q^k - AC_q B)^t X^{-1} dA]. \tag{5.59}$$

Writing $\eta = (\eta_1, \ldots, \eta_{3P})$ and applying the chain rule for the source parameter η_p yields:

$$d_{\eta_p} L = L \times tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B T^{-1} (R_q^k - AC_q B)^t X^{-1} \frac{\partial A}{\partial \eta_p}] d\eta_p \tag{5.60}$$

To obtain the ML-estimate for η_p this derivative is set to zero.

$$d_{\eta_p} L = 0$$

$$\Longleftrightarrow tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B T^{-1} (R_q^k - AC_q B)^t X^{-1} \frac{\partial A}{\partial \eta_p}] = 0$$

$$\Longleftrightarrow tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B^{-1} (R_q^k)^t X^{-1} \frac{\partial A}{\partial \eta_p}] = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B T^{-1} B^t C_q^t A^t X^{-1} \frac{\partial A}{\partial \eta_p}] \tag{5.61}$$

Because the basic field matrix A is linear in η it can be written as

$$A = \sum_{p=1}^{3P} \frac{\partial A}{\partial \eta_p} \eta_p. \tag{5.62}$$

where $\frac{\partial A}{\partial \eta_p}$ is independent of η. Equation (5.61) holds for $p = 1, \ldots, 3P$ and substituting (5.62) into (5.61) yields

$$tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B T^{-1} (R_q^k)^t X^{-1} \frac{\partial A}{\partial \eta_p}] = tr[\sum_{q=1}^{Q} \sum_{p'=1}^{3P} K_q C_q B T^{-1} B^t C_q^t \frac{\partial A^t}{\partial \eta_{p'}} X^{-1} \frac{\partial A}{\partial \eta_p}] \eta_{p'} \quad \forall \ p$$

$$\Longleftrightarrow \Phi \ \eta = \phi \tag{5.63}$$

where

$$\Phi_{p_1,p_2} = tr[\sum_{q=1}^{Q} K_q C_q B T^{-1} B^t C_q^t \frac{\partial A^t}{\partial \eta_{p_2}} X^{-1} \frac{\partial A}{\partial \eta_{p_1}}] \tag{5.64}$$

and

$$\phi_p = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} C_q B T^{-1} (R_q^k)^t X^{-1} \frac{\partial A}{\partial \eta_p}]. \tag{5.65}$$

The linear parameters η_p are solved from the linear system in (5.63).

5.5.5   ML-estimator for amplitude parameters in C

Analogous to the moment parameters, the amplitude parameters α in the coupling matrices C_q are estimated using the chain rule. Writing $\alpha = (\alpha_1, \ldots, \alpha_Y)$ the derivative with respect to α_y is

$$d_{\alpha_y} L =$$

$$= \frac{1}{2} L \times tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} (A d_{\alpha_y} C_q B)^t X^{-1} (R_q^k - A C_q B) T^{-1} + (R_q^k - A C_q B)^t X^{-1} A d_{\alpha_y} C_q B T^{-1}]$$

$$= L \times tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B T^{-1} (R_q^k - A C_q B)^t X^{-1} A d_{\alpha_y} C_q]$$

$$= L \times tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B T^{-1} (R_q^k - A C_q B)^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}] d\alpha_y. \tag{5.66}$$

The ML-estimators for α_y satisfy

$$d_{\alpha_y} L = 0$$

$$\Longleftrightarrow tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B T^{-1} (R_q^k - A C_q B)^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}] = 0$$

$$\Longleftrightarrow tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B T^{-1} (R_q^k)^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}] = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B T^{-1} B^t C_q^t A^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}]. \tag{5.67}$$

Similar to the decomposition of A in (5.62), we can decompose C_q linearly in α:

$$C_q = \sum_{y=1}^{Y} \frac{\partial C_q}{\partial \alpha_y} \alpha_y \tag{5.68}$$

for $q = 1, \ldots, Q$, with $\frac{\partial C_q}{\partial \alpha_y}$ independent of α. Equation (5.67) holds for $y = 1, \ldots, Y$. Formula (5.68) is substituted in (5.67) and we obtain

$$tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} B T^{-1} (R_q^k)^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}] = tr[\sum_{q=1}^{Q} \sum_{y'=1}^{Y} K_q B T^{-1} B^t \frac{\partial C_q^t}{\partial \alpha_{y'}} A^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}] \alpha_{y'} \quad \forall \ y$$

$$\Longleftrightarrow \Psi \ \alpha = \psi \tag{5.69}$$

where

$$\Psi_{y_1,y_2} = tr[\sum_{q=1}^{Q} K_q BT^{-1}B^t \frac{\partial C_q^t}{\partial \alpha_{y_2}} A^t X^{-1} A \frac{\partial C_q}{\partial \alpha_{y_1}}]$$ (5.70)

and

$$\psi_y = tr[\sum_{q=1}^{Q} \sum_{k=1}^{K_q} BT^{-1}(R_q^k)^t X^{-1} A \frac{\partial C_q}{\partial \alpha_y}].$$ (5.71)

This linear system solves the amplitude parameters α in C_q.

Chapter 6

Simultaneous estimation and testing of sources in multiple MEG data sets

6.1   Introduction

The idea of describing measured MEG/EEG data as a linear combination of common (basic) components has been elaborated in many ways. Main examples are the Principal Component Analysis [64], Parallel Factor Analysis (PARAFAC) [71] a special case of which being the Topographic Component Model (TCM) [1, 73, 102] and other trilinear models as in [6, 24, 112]. These models make use of common spatial and temporal components, which all have a certain level of generality (parametrisation). In the TCM each data set is modelled as a weighted sum of parametric spatiotemporal components of which the weights depend on subject and condition. The other models cited decouple the spatial and temporal components by placing a coupling (loading) matrix between the common spatial and common temporal component matrices, specifying for each data set the linear combination of common components.

In the previous chapter the coupled dipole model (CDM) [6] was presented. The CDM is a trilinear model in which multiple, related, MEG data sets (e.g. different conditions) are analysed simultaneously. It is assumed that certain sources and source time functions are equal or proportional across data sets, which relatively increases the signal-to-noise ratio (SNR). For each data set the coupling matrix specifies the combination of common sources and common source time functions that models the data in that particular data set. Thus, this matrix couples the common spatial to the common temporal components; hence, the name 'coupling matrix'. This matrix is designed by the user: the entries corresponding to the chosen combination are estimated from the data, while the remaining entries are set to zero. Although the CDM was shown to improve the stability of the Inverse Problem (IP) of MEG, this matrix design

is subject to the a priori information or assumptions of the user, because fewer or more assumptions about the similarity of the data sets are reflected in different coupling matrices. On the one hand, it is this exploitation of prior knowledge that is the most important asset of the CDM , although, on the other hand, this also causes subjectivity (see section 5.4).

In order to resolve this subjectivity, the current chapter presents the extended coupled dipole model (ECDM), that estimates the coupling matrices entirely. At the same time, the exploitation of prior information changes: only the dimension of the coupling matrix is chosen, while all entries are estimated from the data. An important advantage of the (E)CDM over most of the other component models cited is the statistical clarity within the maximum likelihood (ML) paradigm (cf. [24]). Moreover, whereas the identifiability constraints on the CDM parameters depend on the precise design of the coupling matrices, the identifiability of the ECDM parameters is guaranteed by constraints that only depend on the number of common components. In addition, this enables computing error bounds of estimated source parameters (cf. [24, 28, 41, 75, 76, 109]) by means of the Cramèr Rao bound for constrained parameters [100].

In the next section the model is explained and the maximum likelihood estimators for the ECDM parameters are presented. Furthermore, identifiability constraints for the parameters are given and the Constrained Cramèr Rao bound is computed for all parameters. Finally, an expression for the confidence region around the estimated source time functions is given. Computational details of this section are moved to the appendices. In the section on the results, two experimental applications of the ECDM are considered. In the first the CDM and the ECDM are compared and in the second the usefulness of the confidence regions is investigated. In the last section the results are discussed, and conclusions are drawn.

6.2   Methods

6.2.1   Model

In the ECDM multiple MEG data sets are modelled in one simultaneous analysis using a set of common stationary dipole sources and a set of common source time functions (STFs). This integrated model is, like the coupled dipole model, applicable when MEG data are measured in different though partly similar experimental conditions, such that in the various data sets (conditions) the same sources are activated and/or proportional STFs occur. An example would consist of different data sets in which the somatosensory cortex is activated using electric pulses on the median nerve in a stimulation frequency that varies across data sets. The so obtained MEG data sets will show certain similarities, though will be different.

In this chapter, the following notational convention is used: matrices are denoted by capital letters, either Greek or Latin (e.g. $\Omega, A, \mathbf{I}$); vectors are written either in bold face Latin or in lower case Greek letters (e.g. $\mathbf{a}, \xi$); scalars are in general denoted by lower case Latin letters; however, for indices the lower and upper cases of one letter are used, e.g. $p = 1, \ldots, P$; and finally, the $(n, m)^{th}$ entry of a matrix A is written as

$A_{n,m}$. For a full overview of all symbols used globally in this chapter paper, the reader is referred to the first appendix.

Suppose there are Q data sets, $q = 1, \ldots, Q$, each of which contains K_q trials. Following the Signal Plus Noise model for evoked response data, the measured response R_q^k in trial k of data set q is assumed to consist of two components: a fixed response R_q (that is, fixed within data set q) and additional noise E_q^k [70] (see section 1.3.1):

$$R_q^k = R_q + E_q^k. \tag{6.1}$$

The ECDM models the fixed responses, R_q, as linear combinations of the common spatial and temporal components, which are specified by coupling matrices. In the model equation for the fixed response, this is expressed as (see equation 5.1):

$$R_q = AC_qB. \tag{6.2}$$

Here $A \in \mathbb{R}^{I \times P}$ is the common dipole field matrix, $C_q \in \mathbb{R}^{P \times Z}$ is the coupling matrix for data set q and $B \in \mathbb{R}^{Z \times J}$ the common source time function matrix. I denotes the number of sensors, J the number of time samples, P the number of source time functions belonging to the common sources and Z the number of common source time functions. The orientations of the common sources in A and the common times series in B are normalised, while the magnitudes of activation are estimated in the coupling matrices C_q. Note that it may be assumed that B has full row rank and, in particular, that $Z \leq J$. To see this, assume $rank(B) = r$ and $r \leq J, r \leq Z$. Then there exists a $Z \times r$ matrix C_0, and a $r \times J$ matrix B_0, such that $B = C_0B_0$, and $rank(C_0) = rank(B_0) = r$. The model can then be written as $AC_qB = AC_q(C_0B_0) = A(C_qC_0)B_0$, where C_qC_0 is now a $P \times r$ matrix, functioning as coupling matrix and B_0 is the new common STF matrix with full row rank. The matrix A in (6.2) is parameterised by the nonlinear source location parameters, ξ, and by the linear source orientation parameters, η:

$$A = A(\xi, \eta). \tag{6.3}$$

If all common sources are single dipoles with fixed orientations, then P equals the number of common sources. The reason for this is that the forward field of a single dipole source is represented by one column in A and is parameterised by one location and one orientation. In case of symmetric sources, that is, sources consisting of two dipoles having symmetric locations and orientations with respect to the midsagittal plan, P equals twice the number of common sources. The forward field of a symmetric source is a set of two columns in A, one for the left and one for the right dipole because each dipole has its own amplitude function, and is parameterised by one location (the other is the mirrored location) and one orientation.

In the CDM the coupling matrices are designed by the user: the user specifies the dimension of the coupling matrices (i.e. the number of common components) and selects the zero entries. In the ECDM, the subjectivity of this zero selection is circumvented by estimating the entire coupling matrices, while the dimension is still set by the user. The first, immediate, consequence is the nonuniqueness of the solution to the model in equation (6.2):

$$R_q = AC_qB = A(C_qG^{-1})(GB) \tag{6.4}$$

for any nonsingular matrix G. Therefore constraints on the parameters will be necessary to obtain a unique solution. Moreover, whereas in the CDM the common temporal components bear a physiological meaning (e.g. contralateral STF), in the ECDM the direct physiological meaning of the matrices C_q and B has been changed into a more abstract one, depending on the choice of G, although the product $C_q B$ is still physiologically meaningful. This illustrates the trade-off between objectivity and interpretability in the ECDM and the CDM. And finally, the ECDM is more general than the CDM, because the ECDM solutions include all CDM solutions, though the converse is not true.

6.2.2 ML-estimators for the parameters

The model parameters to be estimated are

- the location parameters of the common sources, contained in the vector ξ

- the orientation parameters of the common sources, contained in the vector η

- the entire coupling matrices C_q, for $q = 1, \ldots, Q$

- the common STF matrix B.

In order to apply the maximum likelihood method to estimate these parameters, the likelihood of the data has to be derived. The noise is assumed to have a Gaussian distribution with a known spatiotemporal covariance following the single Kronecker product model [16]. This is expressed as (see section 1.4.3)

$$vec(E_q^k) \sim \mathcal{N}(0, T \otimes X) \tag{6.5}$$

where T and X are the temporal and spatial covariance matrices respectively, which are assumed to be known or estimated as described in the chapter 5. Furthermore, different trials are assumed to be independent, and different data sets even so (which are anyhow contained in different trials in practise). By these assumptions, X and T are independent of k and q. These matrices are decomposed as:

$$T^{-1} = W_T W_T^t \tag{6.6}$$
$$X^{-1} = W_X W_X^t. \tag{6.7}$$

The following definitions are convenient:

$$\tilde{R}_q^k := W_X^t R_q^k W_T \tag{6.8}$$

$$\tilde{R}_q := \frac{1}{K_q} \sum_{k=1}^{K_q} \tilde{R}_q^k \tag{6.9}$$

$$\tilde{E}_q^k := W_X^t E_q^k W_T \tag{6.10}$$

$$\tilde{E}_q := \frac{1}{K_q} \sum_{k=1}^{K_q} \tilde{E}_q^k. \tag{6.11}$$

Using these prewhitened matrices we define the arrays

$$\mathbf{R} = \begin{pmatrix} \sqrt{K_1}\tilde{R}_1 \\ \vdots \\ \sqrt{K_Q}\tilde{R}_Q \end{pmatrix} \;,\;\; \mathbf{A} = (\mathbf{I}_Q \otimes W_X^t A) \;,\;\; \mathbf{C} = \begin{pmatrix} \sqrt{K_1}C_1 \\ \vdots \\ \sqrt{K_Q}C_Q \end{pmatrix} \;,\;\; \mathbf{B} = BW_T$$

$$\text{and} \;\; \mathbf{E} = \begin{pmatrix} \sqrt{K_1}\tilde{E}_1 \\ \vdots \\ \sqrt{K_Q}\tilde{E}_Q \end{pmatrix} \tag{6.12}$$

so that the joint model of (6.1) for all q becomes

$$\mathbf{R} = \mathbf{ACB} + \mathbf{E}. \tag{6.13}$$

Note that the definition of $\mathbf{A}$ differs slightly from the one in the previous chapter. Matrix $\mathbf{A}$ is still parameterised by (ξ, η). Furthermore, since W_X and W_T are nonsingular matrices, an equivalent set of model parameters is given by $(\xi, \eta, \mathbf{C}, \mathbf{B})$. From (6.5) the distribution of $\sqrt{K_q}vec(\tilde{R}_q)$ is derived:

$$\sqrt{K_q}vec(\tilde{R}_q) \sim \mathcal{N}(vec(W_X^t A\sqrt{K_q}C_q BW_T), \mathbf{I}_{IJ}) \tag{6.14}$$

yielding the marginal likelihood for $\mathbf{R}$

$$L(\xi, \eta, \mathbf{C}, \mathbf{B}) = \frac{1}{(2\pi)^{\frac{IJQ}{2}}} e^{-\frac{1}{2}tr\left((\mathbf{R}-\mathbf{ACB})(\mathbf{R}-\mathbf{ACB})^t\right)} \tag{6.15}$$

and the log likelihood

$$l(\xi, \eta, \mathbf{C}, \mathbf{B}) = -\frac{IJQ}{2}\log(2\pi) - \frac{1}{2}tr((\mathbf{R} - \mathbf{ACB})(\mathbf{R} - \mathbf{ACB})^t). \tag{6.16}$$

The ML-estimators for $(\xi, \eta, \mathbf{C}, \mathbf{B})$ are defined as

$$(\hat{\xi}, \hat{\eta}, \hat{\mathbf{C}}, \hat{\mathbf{B}}) = \arg\max \; l(\xi, \eta, \mathbf{C}, \mathbf{B}). \tag{6.17}$$

The estimators are found be equating the corresponding first order derivatives of (6.16) to zero and solve for the parameters. Owing to the prewhitening, this is equivalent to the ordinary least squares method, discussed in section 1.3.2.

The expressions for the estimators are derived in three steps. In the first step, ξ, η and $\mathbf{B}$ are assumed to be known and $\mathbf{C}$ is estimated

$$\hat{\mathbf{C}} = (\mathbf{A^t A})^{-1}\mathbf{A^t R B^t}(\mathbf{BB^t})^{-1}. \tag{6.18}$$

$(\mathbf{A^t A})^{-1} = \mathbf{I}_Q \otimes (A^t X^{-1} A)^{-1}$ requires the inverse of a $P \times P$ matrix and $(\mathbf{BB^t})^{-1} = (BT^{-1}B^t)^{-1}$ requires a $Z \times Z$ inverse. Substituting (6.18) into (6.16) yields the concentrated log likelihood function $l(\xi, \eta, \mathbf{B})$

$$l(\xi, \eta, \mathbf{B}) = -\frac{IJQ}{2}\log 2\pi - \frac{1}{2}tr(\mathbf{R^t R}) + \frac{1}{2}tr(\mathbf{R^t A}(\mathbf{A^t A})^{-1}\mathbf{A^t R B^t}(\mathbf{BB^t})^{-1}\mathbf{B}). \tag{6.19}$$

In the next step $\mathbf{B}$ is estimated, under the assumption that (ξ, η) is known. First write

$$\mathbf{R^t A (A^t A)^{-1} A^t R} = U \Delta U^t \tag{6.20}$$

with $UU^t = U^t U = \mathbf{I}_J$, and $\Delta = \text{diag}(\mu_1, \mu_2, \ldots, \mu_J)$, with $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_J \geq 0$. Because U is orthogonal, there exists a $(Z \times J)$-matrix $B_0 = \mathbf{B}U$ such that

$$\mathbf{B} = B_0 U^t. \tag{6.21}$$

Now maximising the concentrated log likelihood in (6.19) is equivalent to

$$\max_{B_0} \ tr(\Delta B_0^t (B_0 B_0^t)^{-1} B_0). \tag{6.22}$$

In appendix 6.5.2 the optimal B_0 is determined:

$$\hat{B}_0 = G \begin{pmatrix} \mathbf{I}_Z & \mathbf{0} \end{pmatrix}, \tag{6.23}$$

with G any nonsingular $(Z \times Z)$-matrix. For $\mathbf{B}$ this implies

$$\hat{\mathbf{B}} = G U_Z^t \tag{6.24}$$

where U_Z denotes the matrix containing only the first Z columns of the matrix U defined in (6.20). The matrix G exhibits the nonuniqueness of the model, as noticed in equation (6.4). The imposed constraint is

$$G = \mathbf{I}_Z \ \text{ which is equivalent to } \ \mathbf{B}U_Z = \mathbf{I}_Z. \tag{6.25}$$

Note that the rows in $\hat{\mathbf{B}}$, the prewhitened STFs, are normalised by this constraint. Having estimated $\mathbf{B}$, the log likelihood function can again be concentrated, yielding $l(\xi, \eta)$:

$$l(\xi, \eta) = \frac{1}{2} \sum_{j=1}^{Z} \lambda_j (\mathbf{R^t A (A^t A)^{-1} A^t R}) \tag{6.26}$$

where $\lambda_j(A)$ denotes the jth biggest eigenvalue of matrix A.

In the third and last step, the source parameters (ξ, η) are estimated. In order to find these parameters, the expression in (6.26) has to be maximised using a nonlinear iterative algorithm (e.g. Marquardt).

Finally, the procedure of estimating ξ, η, $\mathbf{C}$ and $\mathbf{B}$ is run in reverse order:

1. Estimate the location and orientation parameters, $\hat{\xi}$ and $\hat{\eta}$, in a nonlinear iterative method maximising (6.26)

2. Derive $\hat{\mathbf{B}}$ from (6.24)

3. Calculate $\hat{\mathbf{C}}$ from (6.18)

As a remark, the reader is referred to an alternative to the first step which was presented in [23]. In that study the linear dependency of $\mathbf{A}$ on the orientation parameters η is exploited and only the locations in ξ are estimated in the (time consuming) nonlinear search. In this faster approach, the first step is replaced by the following scheme which is iterated until convergence of the cost function:

1. Obtain a starting value or update value for ξ using the nonlinear algorithm

2. Set the orientations in η to default values, e.g. $(1, 0, 0)$

3. Iterate until convergence of the maximisation of (6.26):

   a) Estimate the optimal matrix $\mathbf{B}$ and $\mathbf{C}$ using (6.24) and (6.18)

   b) Estimate the optimal orientations by

$$\hat{\eta} = \left(\sum_{j=1}^{J} (A(j))^t A(j) \right)^{-1} \sum_{j=1}^{J} (A(j))^t \mathbf{r_j} \tag{6.27}$$

(see appendix 6.5.3 for the meaning of $A(j)$ and $\mathbf{r_j}$)

Using this alternative approach, the second and third step of the initial procedure have become superfluous. For this fast approach the starting values for the location parameters are obtained by a global search.

6.2.3   Parameter testing

Due to noise in the data array $\mathbf{R}$ all estimated entries of $\hat{\mathbf{C}}$ in (6.18) will be nonzero with probability 1. Consequently, all common sources will have nonzero estimated activity in all data sets. To test the significance of the estimated activity of a certain source p, in a certain data set q, its estimated amplitude over time is tested against the null hypothesis of being zero. This amplitude is the vector-matrix product of the pth row in C_q and the matrix $\mathbf{B}$, $H_0 : ((\hat{\mathbf{C}}\hat{\mathbf{B}})_{p+(q-1)P;1}, \ldots, (\hat{\mathbf{C}}\hat{\mathbf{B}})_{p+(q-1)P;J}) = (0, \ldots, 0) \in \mathbb{R}^{1 \times J}$. Because the common moments and common STFs are normalised, the amplitude of the source activity over the entire time window is assessed by considering the $(p+(q-1)P)^{th}$ row in $\hat{\mathbf{C}}$. In other words, the null hypothesis can be written as a test on a subset of the ECDM parameters, $H_0 : (\hat{\mathbf{C}}_{p+(q-1)P;1}, \ldots, \hat{\mathbf{C}}_{p+(q-1)P;Z}) = (0, \ldots, 0) \in \mathbb{R}^{1 \times Z}$. In the sequel, the statistical distribution of these amplitudes and the other model parameters is derived first, and then the tests on the null hypotheses are described.

Using the ML paradigm, the estimators are asymptotically efficient [63, 95], i.e. the covariance of the estimated parameters approaches the Cramèr Rao lower bound (CRB) for infinitely many trials (observations). The calculation of this bound utilises the Fisher information matrix [46]:

$$\mathcal{I}(\theta) = \mathcal{E}\left[\left(\frac{\partial l(\theta)}{\partial \theta} \right)^t \frac{\partial l(\theta)}{\partial \theta} \right] = -\mathcal{E}\left[\frac{\partial^2 l(\theta)}{\partial \theta^t \partial \theta} \right], \tag{6.28}$$

where θ is the vector containing all model parameters and $l(\theta)$ is the log likelihood function. The Cramèr Rao inequality for the covariance of the estimated parameters $\hat{\theta}$ is given by [88]:

$$Cov(\hat{\theta}) \geq (\mathcal{I}(\theta^*))^{-1}, \tag{6.29}$$

where $A \geq B$ means that $A - B$ is positive semi-definite and θ^* denotes the true value of the parameter vector. Furthermore, the covariance matrix of an ML-estimated $\hat{\theta}_{ML}$, based on finitely many observations, is estimated by [63]:

$$\widehat{Cov}(\hat{\theta}_{ML}) = (\mathcal{I}(\hat{\theta}_{ML}))^{-1}. \tag{6.30}$$

For the ECDM the parameter vector θ is defined as

$$\theta = (vec(\mathbf{C}), \zeta, vec(\mathbf{B})) \tag{6.31}$$

with $\zeta = (\xi, \eta)$. The symmetric information matrix can then be partitioned as

$$\mathcal{I}(\theta) = \begin{pmatrix} \mathcal{I}_{CC} & \mathcal{I}_{C\zeta} & \mathcal{I}_{CB} \\ \mathcal{I}_{\zeta C} & \mathcal{I}_{\zeta\zeta} & \mathcal{I}_{\zeta B} \\ \mathcal{I}_{BC} & \mathcal{I}_{B\zeta} & \mathcal{I}_{BB} \end{pmatrix}. \tag{6.32}$$

From (6.16) and (6.28) an arbitrary entry of $\mathcal{I}(\theta)$ is given by

$$\mathcal{I}(\theta)_{l,l'} = tr(\frac{\partial(\mathbf{ACB})^t}{\partial\theta_l}\frac{\partial(\mathbf{ACB})}{\partial\theta_{l'}}) = \frac{\partial(vec(\mathbf{ACB}))^t}{\partial\theta_l}\frac{\partial vec(\mathbf{ACB})}{\partial\theta_{l'}}. \tag{6.33}$$

Therefore, the following derivatives are sufficient to compute all the submatrices in (6.32) (see [63] for differentiation with respect to matrices and vectors):

$$\frac{\partial vec(\mathbf{ACB})}{\partial vec(\mathbf{C})} = \mathbf{B}^t \otimes \mathbf{A} \tag{6.34}$$

$$\frac{\partial vec(\mathbf{ACB})}{\partial \zeta} = (\mathbf{B}^t\mathbf{C}^t \otimes \mathbf{I}_{IQ})\frac{\partial vec(\mathbf{A})}{\partial \zeta} \tag{6.35}$$

$$\frac{\partial vec(\mathbf{ACB})}{\partial vec(\mathbf{B})} = \mathbf{I}_J \otimes \mathbf{AC} \tag{6.36}$$

Using these derivatives, the information submatrices become

$$\begin{aligned} \mathcal{I}_{CC} &= \mathbf{BB}^t \otimes \mathbf{A}^t\mathbf{A} \\ \mathcal{I}_{C\zeta} &= (\mathbf{BB}^t\mathbf{C}^t \otimes \mathbf{A}^t)\frac{\partial vec(\mathbf{A})}{\partial \zeta} \\ \mathcal{I}_{CB} &= \mathbf{B} \otimes \mathbf{A}^t\mathbf{AC} \\ \mathcal{I}_{\zeta\zeta} &= (\frac{\partial vec(\mathbf{A})}{\partial \zeta})^t(\mathbf{CBB}^t\mathbf{C}^t \otimes \mathbf{I}_{IQ})\frac{\partial vec(\mathbf{A})}{\partial \zeta} \\ \mathcal{I}_{\zeta B} &= (\frac{\partial vec(\mathbf{A})}{\partial \zeta})^t(\mathbf{CB} \otimes \mathbf{AC}) \\ \mathcal{I}_{BB} &= \mathbf{I}_J \otimes \mathbf{C}^t\mathbf{A}^t\mathbf{AC} \end{aligned} \tag{6.37}$$

To compute the estimated covariance matrix (6.30) the inverse of the information matrix is required. However, the information matrix given by the submatrices in (6.37) is singular, which can be easily seen using the determinant formula for the partitioned

matrix [108]: for a nonsingular symmetric matrix $\mathbf{M} \in \mathbb{R}^{m \times m}$, a symmetric matrix $\mathbf{N} \in \mathbb{R}^{n \times n}$ and any matrix $\mathbf{K} \in \mathbb{R}^{m \times n}$ we have

$$\begin{vmatrix} \mathbf{M} & \mathbf{K} \\ \mathbf{K}^t & \mathbf{N} \end{vmatrix} = |\mathbf{M}||\mathbf{N} - \mathbf{K}^t \mathbf{M}^{-1} \mathbf{K}|. \tag{6.38}$$

Setting $\mathbf{M} = \mathcal{I}_{CC}$ and matrices $\mathbf{K}$ and $\mathbf{N}$ correspondingly, the term $\mathbf{N} - \mathbf{K}^t \mathbf{M}^{-1} \mathbf{K}$ becomes a block diagonal matrix having the following two matrices as blocks on the diagonal:

$$(\frac{\partial vec(\mathbf{A})}{\partial \zeta})^t (\mathbf{CBB^tC^t} \otimes (\mathbf{I}_{PQ} - \mathbf{A}(\mathbf{A^tA})^{-1}\mathbf{A^t}))\frac{\partial vec(\mathbf{A})}{\partial \zeta}$$

and

$$(\mathbf{I}_J - \mathbf{B^t(BB^t)^{-1}B}) \otimes \mathbf{C^tA^tAC}$$

of which the second is singular, because $\mathbf{I}_J - \mathbf{B^t(BB^t)^{-1}B}$ is a projection.

This singularity implies that the CRB cannot be calculated straightforwardly, which, indeed, characterises the overparametrisation (6.4). Sufficient identifiability constraints on the parameters for MEG data are [39, 40]:

1. normalised source orientations (P constraints)

2. tangential source orientations (P constraints)

3. $\mathbf{B}U_Z = \mathbf{I}_Z$, see (6.25) ($Z^2$ constraints)

The reader is referred to appendix 6.5.4 for the explicit formulas. In [100] the CRB under parametric constraints is derived. All constraints are written in one vector $\mathbf{c}(\theta) = \mathbf{0} \in \mathbb{R}^{(2P+Z^2) \times 1}$ of which the derivative is defined as

$$D(\theta) = \frac{\partial \mathbf{c}(\theta)}{\partial \theta}. \tag{6.39}$$

The matrix $D(\theta) \in \mathbb{R}^{(2P+Z^2) \times (PQZ+M+ZJ)}$ is discussed in appendix 6.5.4. Here M is the length of the source parameter vector ζ, i.e. $M = 3N + 3P$ (see appendix 6.5.4). $D(\theta)$ is assumed to be of full row rank. In other words, the linear approximations of the constraints are assumed to be linearly independent. Then there exists a matrix $V \in \mathbb{R}^{(PQZ+M+ZJ) \times (PQZ+M+ZJ-2P-Z^2)}$, having orthogonal columns, so that $DV = \mathbf{0}$, i.e. the columns of V span the null-space of D. The constrained Cramèr Rao bound (CCRB) is now given by [100]

$$\Sigma_\theta = V(V^t \mathcal{I}(\theta)V)^{-1}V^t. \tag{6.40}$$

This bound is based on the nonsingularity of $V^t \mathcal{I}(\theta)V$; the columns in V span the same space as the column (or row) space of $\mathcal{I}(\theta)$. In other words, the linear approximations to the constraints (the columns in $D(\theta)$) span the space perpendicular to the row space of the information matrix.

The first PQZ columns in $D(\theta)$, the derivative of $\mathbf{c}(\theta)$ with respect to $vec(\mathbf{C})$, are zero because there are no constraints on $\mathbf{C}$. Therefore we can choose V to be

$$V = \begin{pmatrix} I_{PQZ} & \mathbf{0} \\ \mathbf{0} & V_0 \end{pmatrix}, \tag{6.41}$$

where $V_0 \in \mathbb{R}^{(M+ZJ)\times(M+ZJ-2P-Z^2)}$ and the $\mathbf{0}$'s present null matrices of appropriate dimensions. Furthermore, V_0 has orthogonal columns such that

$$D \begin{pmatrix} \mathbf{0} \\ V_0 \end{pmatrix} = \mathbf{0}. \tag{6.42}$$

Σ_θ in (6.40) then becomes

$$\Sigma_\theta = \begin{pmatrix} I_{PQZ} & \mathbf{0} \\ \mathbf{0} & V_0 \end{pmatrix} \left[\begin{pmatrix} \mathcal{I}_{CC} \\ V_0^t \begin{pmatrix} \mathcal{I}_{\zeta C} \\ \mathcal{I}_{BC} \end{pmatrix} & \begin{pmatrix} \mathcal{I}_{C\zeta} & \mathcal{I}_{CB} \end{pmatrix} V_0 \\ V_0^t \begin{pmatrix} \mathcal{I}_{\zeta\zeta} & \mathcal{I}_{\zeta B} \\ \mathcal{I}_{B\zeta} & \mathcal{I}_{BB} \end{pmatrix} V_0 \end{pmatrix} \right]^{-1} \begin{pmatrix} I_{PQZ} & \mathbf{0} \\ \mathbf{0} & V_0^t \end{pmatrix}. \tag{6.43}$$

The derived CCRB is now used to test the significance of the estimated activity, as discussed above. For a certain source p in a particular data set q the null hypothesis is $H_0 : (\hat{\mathbf{C}}_{p+(q-1)P;1}, \ldots, \hat{\mathbf{C}}_{p+(q-1)P;Z}) = (0, \ldots, 0) \in \mathbb{R}^{1\times Z}$. To test such a subset of the estimated parameters, a Fisher F-test is performed [95], which is based on a linear approximation of the nonlinear model (appendix 6.5.5). Taking into account the constraints on the ECDM parameters, the $\mathcal{F}$-statistic in (6.85) is used to calculate the significance of the activity of the pth source in the qth data set over the entire time window, indicated by a p-value.

Although the activity of a certain source may be significant over the entire time window in a certain data set, its activity needs not be significant at each time sample. Hence, another interesting statistical feature is the confidence interval around the estimated amplitude for each source in each data set at all time instants. These STFs are contained in the matrix product $\mathbf{CB}$, which is a nonlinear function of the ECDM parameters. In [95], testing a nonlinear null hypothesis, $H_0 : ((\hat{\mathbf{C}}\hat{\mathbf{B}})_{p+(q-1)P;j} = 0 \in \mathbb{R}^{1\times 1}$, is performed through linear approximation. This same strategy is applied in the present study to compute the confidence regions around the STFs (appendix 6.5.6). The resulting $\mathcal{F}$-statistic in (6.89) determines the confidence region, (6.95).

6.3   Results

The ECDM was applied in two different experimental MEG designs, in order to examine its performance. In the first application, MEG data obtained in a Visual Evoked Field (VEF) experiment were studied. The same VEF data have been analysed also by the CDM in the previous chapter. This study allows for a comparison between the performances of the CDM and the ECDM. Secondly, the ECDM was applied to Somatosensory Evoked Field (SEF) data, which enabled investigation of the accuracy and importance of the significance tests and confidence regions.

6.3.1   VEF data

VEF data were obtained in three different experimental conditions; a 6' checkerboard pattern was shown to the subject in either the left or the right visual hemi-field, or full field. Using a 151 channel CTF Omega system, data of five subjects were acquired at

a sample rate of 625 Hz. Each data set contained 400 trials and the time window of analysis was set to 80-112 ms post-stimulus.

As outlined in chapter 5, describing the activity evoked by this check size in this time window requires two common sources and two common STFs. Although in the CDM the two STFs are given physiological meanings, contra- and ipsilateral STF, in the ECDM the common STFs are orthogonal and, therefore, lack such a direct physiological meaning. As model for the two sources a semi-symmetric dipole pair was employed, that is, two dipole sources having symmetric locations, though arbitrary orientations. This generated 3 (2×2) coupling matrices for each subject. In the CDM both hemi-field data sets were modelled by an ipsi- and contralateral source (with ipsi- and contralateral STFs respectively) and the full field data set was modelled by both sources having the contralateral STF. In the ECDM activity of both sources is estimated as a linear combination of the two orthogonal basic STFs.

In general, the estimated ECDM parameters are very close to the estimated CDM parameters. Over the five subjects, the average distance between the ECDM and the CDM estimated locations is 0.11 cm ($\pm$ 0.09 cm) and the average rotation between estimated left orientations equals 1.6° ($\pm$ 1.9°) and right 1.3° ($\pm$ 1.5°). Like the CDM solutions, the ECDM solutions, show nearly symmetric orientations. A slight difference was found in the shapes of the estimated STFs, i.e. the rows of **CB**; however, the magnitudes are comparable for subjects 1, 2, 4 and 5. For subject 3, the magnitude in the ECDM solution is approximately 1.9 times the magnitude in the CDM solution, while the shapes are similar. The explanation for this distinction lies in the difference between the estimated sources: the ECDM sources are located closer to the midsagittal plane than the CDM sources (ECDM: 0.4 cm; CDM: 0.8 cm), and the orientations (both ECDM and CDM) are mainly lateral (angle between orientation and pure lateral direction is 17° (left) and 9° (right) for the ECDM). In other words, the ECDM, which is less restrictive than the CDM, yields slightly cancelling sources, as was discussed in 5.1; two sources, close together, having opposite orientations are vulnerable for counterbalancing each other. The data of this subject were in any case the most troublesome of all, because separate analyses of all three data sets could not yield an interpretable solution (see section 5.3). Nevertheless, in general the ECDM and CDM solutions resemble each other reasonably for the uncomplicated data sets.

The results of subject 5 are representative for the type of shape (all subjects) and magnitude (subject 1,2,4,5) dissimilarities between the ECDM and the CDM. The results of the same subject were presented in chapter 5. In Table 6.1 the estimated source parameters are presented, and Figure 6.1 shows the estimated STFs together with the 95% confidence intervals. From this figure and table it is seen that the ECDM magnitude is slightly bigger than the CDM magnitude at most instants, and correspondingly, the CDM sources are slightly more superficial than the ECDM sources, although not significantly so.

6.3.2   SEF data

In circumstances where a certain source is not expected to be active in one (or some) of the data sets, the estimated coupling parameters in the ECDM will, nonetheless, be nonzero with probability 1. This is the case, for example, when the number of common

	position left (cm)			orientation left			orientation right		
	x	y	z	x	y	z	x	y	z
ECDM	-5.26	1.37	0.84	0.29	0.76	0.58	0.26	-0.96	0.10
CDM	-5.25	1.44	0.85	0.30	0.72	0.62	0.28	-0.95	0.12

Table 6.1: *Estimated source parameters of the semi-symmetric sources for subject 5. The positions of the sources are relative to the centre of the spherical volume conductor. The direction of the x-axis is forward, the y-axis is to the left and the z-axis points upwards.*

(a) Left hemi-field stimulation

(b) Right hemi-field stimulation

(c) Full field stimulation
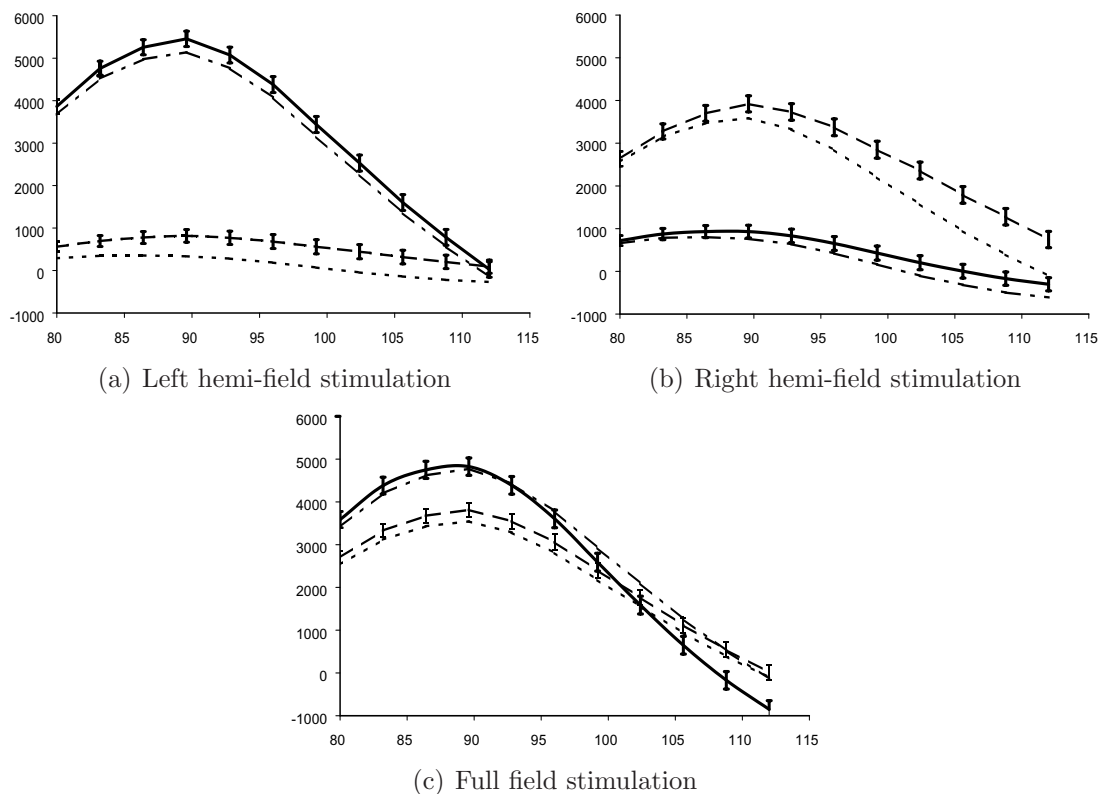
Figure 6.1: *Estimated STFs of the semi-symmetric sources resulting from both the ECDM and the CDM analyses of subject 5. The horizontal axis represents the time after stimulus in ms and the vertical axis shows the amplitude in nAcm. The solid (dashed) line presents the ECDM estimated STF of the right (left) source and the dot-dashed (dotted) line presents the CDM estimated STF of the right (left) source. The vertical error bars indicate a 95% confidence region around the ECDM amplitudes.*

sources is too large for one of the data sets; the entire coupling matrix is estimated, producing nonzero estimated activity for all common sources. To investigate the performance of the ECDM in cases of these expectedly *silent* sources, the model was applied to the N20 response in SEF data of the left and right median nerve. Typically, one would analyse these two data sets separately, using single source models. However, it

is a plausible assumption that the somatosensory cortices left and right have symmetric locations. Therefore, the ECDM can be applied assuming semi-symmetric sources. Nonetheless, the author wants to stress that the purpose of this SEF study is not to find the optimal SEF source localisation method, though rather to test the robustness and correctness of the proposed model.

The design of the ECDM for the two SEF data sets, SEF_L and SEF_R, used a semi-symmetric source pair. The two sources permit both contra- and ipsilateral activity. The number of STFs was varied from 1 to 4; one STF forces the ipsi- and contralateral amplitudes to be proportional, two STFs allow for different wave shapes. By extending the number of STFs even further to at most 4 (then $\mathbf{C}$ is square), the difference between the ECDM design and the separate analyses is reduced to only the assumption of symmetric source locations. The dimensions of the coupling matrices varied correspondingly from 2×1 to 2×4 for each data set.

Using a 151-channel CTF Omega system, data of five subjects were collected at a sample rate of 2083 Hz. For each subject, the frequency of the median nerve stimulation was set to 2 Hz for both data sets, each of which contained at least 500 trials. For each subject a narrow interval of 2 ms around the N20 peak was selected as time window of analysis, functioning both for the SEF_L and the SEF_R data set.

For all subjects the estimated contralateral activity clearly dominates the ipsilateral activity. Figures 6.2 and 6.3 show the estimated STFs together with the corresponding confidence intervals for subjects 3 and 5 for $Z = 2$. The p-values for both data sets, computed according to the above outlined tests, are reported in Table 6.2. As can be seen from the table, the p-values generally increase with Z, while the estimated source parameters (not shown) hardly change with Z. Moreover, not all p-values exceed the common 5% limit; in other words, some ipsilateral activity is estimated as significant. Examples of this significant activity are shown in Figures 6.3(a) and 6.3(b) ($Z = 2$). From these figures the importance of the confidence regions becomes clear: for all samples the activity balances on the verge of significance, yielding apparently a p-value just below 5 %. Finally, the p-values of the contralateral activity were calculated, yielding values ranging from 10^{-234} to 10^{-59}.

subject	Z=1 SEF_L	Z=2 SEF_L	Z=3 SEF_L	Z=4 SEF_L	Z=1 SEF_R	Z=2 SEF_R	Z=3 SEF_R	Z=4 SEF_R
1	10.97%	24.23%	38.60%	53.59%	0.00%	0.00%	0.00%	0.02%
2	1.14%	0.82%	1.67%	3.52%	6.46%	17.97%	31.55%	45.79%
3	33.73%	42.36%	62.51%	67.85%	3.98%	8.13%	15.90%	27.99%
4	30.13%	55.34%	55.26%	71.71%	4.78%	16.96%	26.26%	36.42%
5	2.36%	2.25%	5.50%	9.66%	0.34%	1.42%	3.32%	3.88%

Table 6.2: *P-values of the estimated ipsilateral activity in both the SEF data sets of the five subjects for the different values of Z.*

(a) SEF left median nerve   (b) SEF right median nerve

Figure 6.2: *ECDM estimated STFs of the semi-symmetric sources in the SEF data sets of subject 3. The horizontal axis represents the time after stimulus in ms and the vertical axis shows the amplitude in nAcm. The solid (dashed) line indicates the estimated amplitude of the left (right) source. The vertical error bars indicate a 95% confidence region around the amplitudes. The horizontal axis represents the post-stimulus time in ms and the vertical axis shows the amplitude in nAcm.*

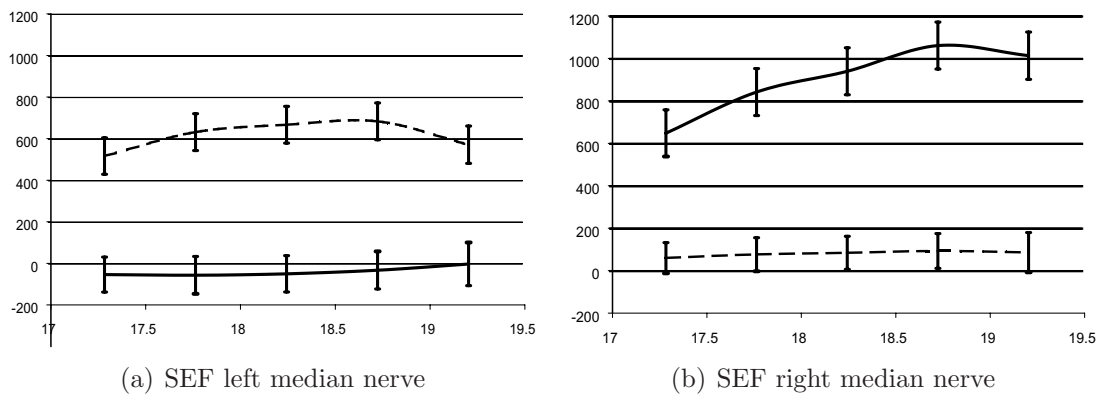(a) SEF left median nerve   (b) SEF right median nerve
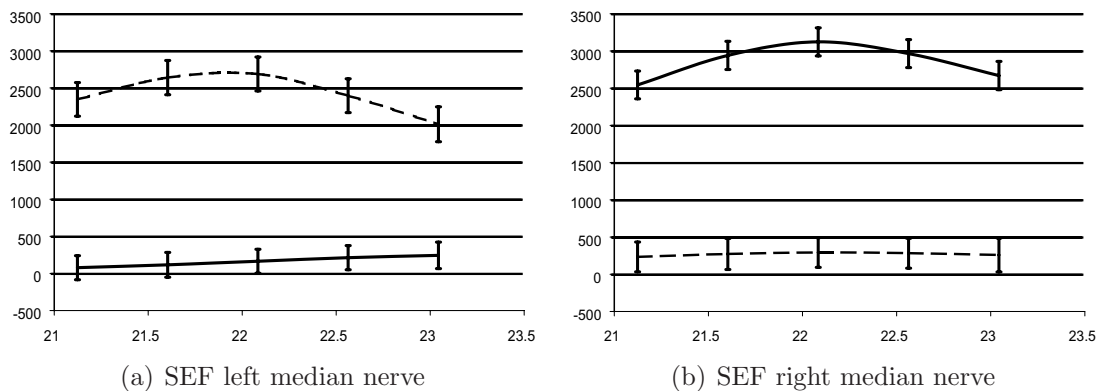
Figure 6.3: *ECDM estimated STFs of the semi-symmetric sources in the SEF data sets of subject 5. The horizontal axis represents the time after stimulus in ms and the vertical axis shows the amplitude in nAcm. The solid (dashed) line indicates the estimated amplitude of the left (right) source. The vertical error bars estimated amplitude of right (left) source. The vertical error bars indicate a 95% confidence region around the amplitudes. The horizontal axis represents the post-stimulus time in ms and the vertical axis shows the amplitude in nAcm.*

6.4   Discussion

The ECDM is an extension of the CDM, analysing multiple MEG data sets simultaneously, in which the subjectivity of the coupling design has been erased. The ECDM can be regarded as an extension of the CDM in the temporal domain: the wave shapes in **CB**, i.e. the estimated source time functions for the sources in all data sets, are more

general in the ECDM. Spatially, there is no noteworthy difference between the models.

The crucial enhancement by this temporal extension is the statistical transparency: identifiability constraints are independent of the user's design and the covariance of the parameters is accessible by using the Constrained Cramèr Rao bound. Another advantage of the ECDM over the CDM is the ease of design: only the dimension of the coupling matrices is chosen by the user. The price to be paid for this benefit is the loss of direct physiological interpretation of **C** and **B**; however, the product **CB** remains meaningful. Furthermore, the prior information cannot be exploited, which can be either a favour or a drawback, depending on the user's desire.

An alternative estimation method for the ECDM is the Generalised Multivariate Analysis Of Variance (GMANOVA) model [25, 49, 97, 108]. The GMANOVA (or growth curve) model, which is the trilinear extension ($Y = AXB$) of the common MANOVA model ($Y = AX$), was used to describe a single MEG/EEG data set in [24]. In that study, the spatial noise covariance was estimated simultaneously with the model parameters, having the advantage that the spatial covariance can capture modelling errors. In the more complicated ECDM, this benefit can not be achieved for the general spatiotemporal noise covariance, because the GMANOVA model admits this favour only for either the spatial or the temporal noise covariance. Hence, capturing modelling errors in the ECDM should be performed by iterative estimation of the noise covariance and the model parameters.

By applying the ECDM to the same VEF data as the CDM had been applied to, a direct comparison between estimated parameters was made. Spatially, no noteworthy differences were found. Regarding the temporal parameters, the extra freedom of the ECDM was manifested by more general wave shapes. For one of the five subjects somewhat different results were reported: the semi-symmetric sources were closer together in the ECDM and the magnitudes were higher. This directs to the phenomenon of cancelling sources as discussed in section 5.1. Apparently, the somewhat problematic data of subject 3 need a very strict model enabling a plausible solution: the separate analyses of the three data sets yield unlikely solutions; the ECDM yields a better parameter estimate which is still a little doubtful; and the CDM generates the most plausible solution, that is comparable with solutions of other subjects. In this kind of circumstances, model selection procedures [110] are helpful to decide on the correct model.

The estimated activity of supposedly *silent* sources was investigated in the SEF data analysis. The N20 response to a SEF stimulus has been reported arising from the contralateral SI area only [43, 69]. Activity of both contra- and ipsilateral sources was estimated by fitting a semi-symmetric source model to SEF data of left and right median nerve stimulation of five subjects. Computed p-values of estimated ipsilateral activity (Table 6.2) were generally high, although some of the values turned out to be significant. A closer look at the confidence interval for each time sample revealed that this significance can change from sample to sample. A few remarks can be made here. In general, the ordinary least squares method underestimates the standard errors of the model parameters, while the generalised least squares method, which is performed through prewhitening in the ECDM, yields a correct estimate of the standard error (and thus of the confidence intervals) (e.g. [42]). Nevertheless, it seems tenable from Figure 6.3 that the ECDM confidence intervals are still slightly underestimated. This

can be interpreted in line with a feature reported in [20] where it was shown that the source localisation suffers from comparable frequencies in source time functions and the temporal covariance. In such a case, not solely noise but also signal is removed by prewhitening. In a simulation study, not reported here, this phenomenon was studied and confirmed: the ratio between estimated standard errors and sample variances of the estimated amplitudes equalled approximately 0.70 for frequencies below 15 Hz and 0.92 for frequencies above 20 Hz. As a final remark, the liberal way of testing may have produced too optimistic confidence intervals; a correction for multiple tests (e.g. Bonferroni) might indicate no significance. Bearing these remarks in mind, existence of ipsilateral activity is neither approved nor disapproved by this ECDM study.

Summarising, the ECDM is an alternative to the CDM which is statistically more convenient, but does not allow for exploitation of prior information. It is the user's careful consideration whether to use the CDM or the ECDM.

6.5 Appendix

6.5.1 Dimensions and variables

The dimensions are defined as

- I - # sensors

- J - # time samples

- K_q - # trials data set q

- K - # trials all data sets

- M - # source parameters

- N - # common sources

- P - # source time functions of common sources

- Q - # data sets

- Z - # common STFs

and the variables as

- ζ - $(M \times 1)$ vector with all source parameters

- η - $(3P \times 1)$ vector with source orientation parameters

- ξ - $((M - 3P) \times 1)$ vector with source location parameters

- θ - $((PQZ + M + ZJ) \times 1)$ vector with all parameters

- $A(\xi, \eta)$ - $(I \times P)$ dipole field matrix of common sources

- $\mathbf{A}$ - $(IQ \times PQ)$ matrix $\mathbf{I}_Q \otimes A$

- B - $(Z \times J)$ common STF matrix, (STFs as rows)

- $\mathbf{B}$ - $(Z \times J)$ prewhitened STF matrix

- $\mathbf{c}(\theta)$ - $((2P + Z^2) \times 1)$ constraint vector

- C_q - $(P \times Z)$ coupling matrix data set q

- $\mathbf{C}$ - $(QP \times Z)$ stacked coupling matrices

- $D(\theta)$ - $((2P + Z^2) \times (PQZ + M + ZJ))$ derivative matrix of $\mathbf{c}(\theta)$

- E_q^k - $(I \times J)$ noise matrix trial k data set q

- $\mathbf{I}_n$ - $(n \times n)$ identity matrix

- R_q^k - $(I \times J)$ data matrix trial k data set q

- R_q - $(I \times J)$ model matrix data set q

- $\bar{R}_q$ - $(I \times J)$ average data matrix data set q

- $\mathbf{R}$ - $(IQ \times J)$ stacked average prewhitened data matrices

- T - $(J \times J)$ temporal covariance matrix

- U - $(J \times J)$ orthogonal matrix with singular vectors of $\mathbf{R^t A (A^t A)^{-1} A^t R}$

- U_Z - $(J \times Z)$ first Z columns of U

- X - $(I \times I)$ spatial covariance matrix

- Σ_θ - $((PQZ + M + ZJ) \times (PQZ + M + ZJ))$ constrained CRB matrix

6.5.2   Estimator for B_0

In this appendix the ML-estimator for B_0 is derived from equation (6.22). Define the matrix $P := B_0^t (B_0 B_0^t)^{-1} B_0 \in \mathbb{R}^{J \times J}$, which is a projection with $\mathrm{rank}(P) = \mathrm{rank}(B_0) = \mathrm{rank}(\mathbf{B}) = Z$. Therefore, P can be decomposed as $P = HH^t$, with $H \in \mathbb{R}^{J \times Z}$ having Z orthogonal columns. If the jth row vector of H is written as $\mathbf{h_j}^t \in \mathbb{R}^{1 \times Z}$, the trace in (6.22) becomes

$$tr(\Delta P) = tr(\Delta HH^t) = tr\left[\begin{pmatrix} \mu_1 \mathbf{h_1}^t \\ \vdots \\ \mu_J \mathbf{h_J}^t \end{pmatrix} \begin{pmatrix} \mathbf{h_1} & \cdots & \mathbf{h_J} \end{pmatrix}\right] = \sum_{j=1}^J \mu_j \mathbf{h_j}^t \mathbf{h_j} \qquad (6.44)$$

with $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_J \geq 0$. Now the trace of a projection matrix equals its rank, therefore

$$tr(P) = tr(HH^t) = \sum_{j=1}^J \mathbf{h_j}^t \mathbf{h_j} = Z. \qquad (6.45)$$

Defining the vectors $\mathbf{a} = (a_1, \ldots, a_J)$ with $a_j = \mathbf{h_j}^t \mathbf{h_j}$, and $\mu = (\mu_1, \ldots, \mu_J)$, the maximisation problem in (6.22) is reduced to

$$\max_{\mathbf{a}} \mathbf{a}^t \mu \quad \text{subject to} \quad \begin{cases} 0 \leq a_j \leq 1 \\ \sum_j a_j = Z \\ \mu_j \geq \mu_{j+1} \text{ for } j = 1, \ldots, J-1 \\ \mu_j \geq 0 \ \forall j \end{cases} \tag{6.46}$$

Obviously the optimal vector is given by $\hat{\mathbf{a}} = (1, \ldots, 1, 0, \ldots, 0)$, of which only the first Z elements equal 1. In case of less than Z nonzero μ_j this optimal solution for $\mathbf{a}$ is not unique. Yet, Z exceeding the number of nonzero μ_j, in (6.20), indicates overparametrisation; hence, it may be assumed that $\hat{\mathbf{a}}$ is unique. This implies for H

$$\mathbf{h_j h_j}^t = \begin{cases} 1 & \text{for} & 1 \leq j \leq Z \\ 0 & \text{for} & Z+1 \leq j \leq J \end{cases} \tag{6.47}$$

Writing

$$H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} \tag{6.48}$$

with $H_1 \in \mathbb{R}^{Z \times Z}$ and $H_2 \in \mathbb{R}^{(J-Z) \times Z}$, equation (6.47) implies that H_1 is orthogonal and H_2 equals the null matrix. The projection P then becomes

$$P = HH^t = \begin{pmatrix} \mathbf{I}_Z & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{6.49}$$

Recalling $P = B_0^t (B_0 B_0^t)^{-1} B_0$, the estimator for B_0 is

$$\hat{B}_0 = G \begin{pmatrix} \mathbf{I}_Z & \mathbf{0} \end{pmatrix}, \tag{6.50}$$

$\mathbf{0} \in \mathbb{R}^{Z \times (J-Z)}$ being the null matrix and $G \in \mathbb{R}^{Z \times Z}$ any nonsingular matrix.

6.5.3   Linear estimator for the source orientations

In this appendix an estimator is derived for the source orientations, which exploits the linear dependency of the model on the orientations. First, A is decomposed using the linear dependency in η: $W_X^t A(\xi, \eta) = \Phi(\xi)\Omega(\eta)$, where $\Phi(\xi)$ is the $(I \times 3P)$ prewhitened unit dipole field matrix in the three Euclidean moment directions, and $\Omega(\eta)$ is the $(3P \times P)$ block diagonal matrix containing P (3×1) moment blocks on the diagonal. Next, this decomposition is substituted in (6.15), which then is maximised with respect to η, or equivalently

$$\min_{\eta} \ tr((\mathbf{R} - (\mathbf{I}_Q \otimes \Phi\Omega)\mathbf{CB})(\mathbf{R} - (\mathbf{I}_Q \otimes \Phi\Omega)\mathbf{CB})^t) \tag{6.51}$$

for known ξ, $\mathbf{B}$ and $\mathbf{C}$. This quadratic function of η is minimum when the derivatives with respect to all the orientation parameters η_m, $m = 1, \ldots, 3P$, are zero, i.e.

$$tr(((\mathbf{I}_Q \otimes \Phi \frac{\partial \Omega}{\partial \eta_m})\mathbf{CB})^t (\mathbf{R} - (\mathbf{I}_Q \otimes \Phi\Omega)\mathbf{CB})) =$$

$$\sum_{j=1}^{J}((\mathbf{I}_Q \otimes \Phi \frac{\partial \Omega}{\partial \eta_m})\mathbf{Cb_j})^t (\mathbf{r_j} - (\mathbf{I}_Q \otimes \Phi\Omega)\mathbf{Cb_j}) = 0 \tag{6.52}$$

for all m, where $\mathbf{r_j}$ and $\mathbf{b_j}$ denote the jth column in $\mathbf{R}$ and $\mathbf{B}$ respectively. Defining

$$A_m(j) = (\mathbf{I}_Q \otimes \Phi \frac{\partial \Omega}{\partial \eta_m})\mathbf{Cb_j} \quad \text{for} \quad m = 1, \dots, 3P \tag{6.53}$$

$$A(j) = \begin{pmatrix} A_1(j) & A_2(j) & \cdots & A_{3P}(j) \end{pmatrix} \tag{6.54}$$

and noting that $\Omega = \sum_{m=1}^{3P} \frac{\partial \Omega}{\partial \eta_m} \eta_m$, equation (6.52) can be rewritten as

$$\sum_{j=1}^{J} (A_m(j))^t \mathbf{r_j} = \sum_{m'=1}^{3P} \sum_{j=1}^{J} (A_m(j))^t A_{m'}(j) \eta_{m'} \tag{6.55}$$

for all m. Finally, stacking the scalar equation (6.55) for all m as entries in one column, the following equation is obtained:

$$\sum_{j=1}^{J} (A(j))^t \mathbf{r_j} = \sum_{j=1}^{J} (A(j))^t A(j) \eta \tag{6.56}$$

and the estimator in (6.27), which requires a $3P \times 3P$ inverse, follows. An alternative expression for the linear system in (6.56) can be found in equation 5.19 and related equations.

6.5.4   Derivative matrix of the constraints

In this appendix, the derivative of the constraint vector $\mathbf{c}(\theta)$ with respect to all parameters is computed. Define the number of common sources to be N. Then the number of source location parameters is $3N$ and that of source orientations is $3P$, together yielding $M = 3N + 3P$. The constraints are

$$\eta_{3p-2}^2 + \eta_{3p-1}^2 + \eta_{3p}^2 = 1 \quad \text{for } p = 1, \dots, P \text{ (normality constraint)}$$
$$\xi_{3n-2}\eta_{3p-2} + \xi_{3n-1}\eta_{3p-1} + \xi_{3n}\eta_{3p} = 0 \quad \text{for } p = 1, \dots, P \text{ (tangentiality constraint)}$$
$$(\mathbf{B}U)_{z_1,z_2} = \delta_{z_1,z_2} \quad \text{for } z_1, z_2 = 1, \dots, Z$$

where n in the tangentiality constraints is chosen such that the pth STF corresponds to the nth source. Furthermore, $\delta_{i,j}$ denotes the Kronecker delta function (see equation (1.53)). The dimension of $D(\theta)$ in (6.39) is $(2P + Z^2) \times (PQZ + M + ZJ)$. The first $2P$ rows of $D(\theta)$ are straightforward. For the last Z^2 rows we have:

$$\frac{\partial(\mathbf{B}U)_{z_1,z_2}}{\partial B_{z,j}} = \delta_{z,z_1} U_{j,z_2} \tag{6.57}$$

$$\frac{\partial(\mathbf{B}U)_{z_1,z_2}}{\partial \zeta_m} = \mathbf{B}_{z_1}^t \frac{\partial U_{z_2}}{\partial \zeta_m} \tag{6.58}$$

where U_{z_2} and $\mathbf{B}_{z_1}$ denote the z_2^{th} and z_1^{th} column of U and $\mathbf{B}$, respectively. The last fraction in 6.58 is the derivative of an eigenvector of $\mathbf{R^t A(A^t A)^{-1} A^t R}$ with respect

to one of the parameters by which this matrix is parameterised. In [63] differentiation of eigenvectors with respect to the matrix is explained. The derivative with respect to matrix X of eigenvector u_0 of matrix X corresponding to eigenvalue λ_0 of matrix X is calculated by

$$du = (\lambda_0 \mathbf{I} - X)^+ (dX) u_0. \tag{6.59}$$

Here $(..)^+$ stands for the Moore-Penrose inverse. Substituting $X = \mathbf{R^t A (A^t A)^{-1} A^t R}$ which is a function of ζ_m, $u_0 = U_{z_2}$ with corresponding eigenvalue λ_{z_2} and applying (6.59) and the chain rule to (6.58) yields

$$\frac{\partial U_{z_2}}{\partial \zeta_m} = (\lambda_{z_2} \mathbf{I}_J - \mathbf{R^t A (A^t A)^{-1} A^t R})^+ \frac{\partial (\mathbf{R^t A (A^t A)^{-1} A^t R})}{\partial \zeta_m} U_{z_2}. \tag{6.60}$$

Employing the product rule, $d_X(X^{-1}) = -X^{-1}dX X^{-1}$ and (6.60), the matrix $D(\theta)$ can be calculated.

6.5.5   F-test on estimated coupling parameters

The F-test on the estimated parameters is derived in this appendix, following the derivation in [95]. First, the test is designed for unconstrained parameters and after that, the case of constrained parameters is described. The following vectors in $\mathbb{R}^{IJQ \times 1}$ are defined

$$\mathbf{r} = vec(\mathbf{R}) \tag{6.61}$$
$$\mathbf{f}(\theta) = vec(\mathbf{ACB}) \tag{6.62}$$
$$\varepsilon = \mathbf{r} - \mathbf{f}(\theta^*). \tag{6.63}$$

The model vector $\mathbf{f}(\theta)$ is expanded in a Taylor series around $\theta = \theta^*$, the true value of parameter θ

$$\mathbf{f}(\theta) \approx \mathbf{f}(\theta^*) + F(\theta)(\theta - \theta^*) \tag{6.64}$$

where

$$F(\theta) = \frac{\partial \mathbf{f}(\theta)}{\partial \theta}(\theta^*) \in \mathbb{R}^{IJQ \times n_\theta}. \tag{6.65}$$

n_θ is an abbreviation for the length of θ. For notational convenience the dependency of F on θ will be omitted in the sequel. Maximising the log likelihood function in (6.16) is equivalent to

$$\min_\theta \|\mathbf{r} - \mathbf{f}(\theta)\|^2 \approx \min_\theta \|\mathbf{r} - \mathbf{f}(\theta^*) - F(\theta - \theta^*)\|^2 = \min_\theta \|\varepsilon - F(\theta - \theta^*)\|^2. \tag{6.66}$$

The optimal value for θ is given by

$$\hat{\theta} - \theta^* = (F^t F)^{-1} F^t \varepsilon. \tag{6.67}$$

From (6.64) and (6.67) the following approximation can be made

$$\mathbf{f}(\hat{\theta}) - \mathbf{f}(\theta^*) \approx F(\hat{\theta} - \theta^*) = F(F^t F)^{-1} F^t \varepsilon \tag{6.68}$$
$$\mathbf{r} - \mathbf{f}(\hat{\theta}) \approx \mathbf{r} - \mathbf{f}(\theta^*) - F(\hat{\theta} - \theta^*) = \varepsilon - F(F^t F)^{-1} F^t \varepsilon. \tag{6.69}$$

Defining the projection

$$P_F = F(F^t F)^{-1} F^t, \tag{6.70}$$

this yields

$$\|\mathbf{f}(\hat{\theta}) - \mathbf{f}(\theta^*)\|^2 \approx \varepsilon^t P_F \varepsilon \sim \chi^2_{n_\theta} \tag{6.71}$$

$$\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2 \approx \varepsilon^t (I - P_F) \varepsilon \sim \chi^2_{IJQ - n_\theta} \tag{6.72}$$

where χ^2_l denotes the Chi-squared distribution with l degrees of freedom. Clearly $(I - P_F)\varepsilon$ and $P_F \varepsilon$ are independent, which leads to the following $\mathcal{F}$-statistic

$$F_{test} = \frac{\frac{\varepsilon^t P_F \varepsilon}{n_\theta}}{\frac{\varepsilon^t (I - P_F)\varepsilon}{IJQ - n_\theta}} = \frac{\|\mathbf{f}(\hat{\theta}) - \mathbf{f}(\theta^*)\|^2}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\theta}{n_\theta} \approx \frac{(\hat{\theta} - \theta^*)^t F^t F(\hat{\theta} - \theta^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\theta}{n_\theta}$$

$$\sim \mathcal{F}_{n_\theta,\ IJQ - n_\theta} \tag{6.73}$$

The above test statistic is for the entire parameter vector $\hat{\theta}$. The estimated covariance matrix of $\hat{\theta}$ is $(F(\hat{\theta})^t F(\hat{\theta}))^{-1}$, which is well-defined for unconstrained parameters. To test a subset of the estimated parameters, write

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \tag{6.74}$$

$$F = \begin{pmatrix} F_1 & F_2 \end{pmatrix} \tag{6.75}$$

and assume that the parameters in θ_2 have to be tested. Using the partitioning of F, P_{F_1} and P_{F_2} are defined similar to the projection in (6.70). This yields for the projector in the denominator of the test statistic

$$I - P_F = I - P_{F_1} - (I - P_{F_1})F_2[F_2^t(I - P_{F_1})F_2]^{-1}F_2^t(I - P_{F_1}). \tag{6.76}$$

In the numerator of the test statistic, $\hat{\theta} - \theta^*$ has to be replaced by $\hat{\theta}_2 - \theta_2^*$ and the inverse of the estimated covariance matrix for θ_2 has to be substituted for $F^t F$. This covariance matrix is given by the corresponding submatrix of $(F^t F)^{-1}$:

$$\widehat{COV}(\theta_2) = \begin{pmatrix} \mathbf{0} & I \end{pmatrix} \begin{pmatrix} F_1^t F_1 & F_1^t F_2 \\ F_2^t F_1 & F_2^t F_2 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ I \end{pmatrix} = (F_2^t(I - P_{F_1})F_2)^{-1}. \tag{6.77}$$

The numerator of the test statistic then becomes

$$(\hat{\theta}_2 - \theta_2^*)^t F_2^t (I - P_{F_1}) F_2 (\hat{\theta}_2 - \theta_2^*). \tag{6.78}$$

This expression can be written again as $\varepsilon^t P \varepsilon$ for a certain projection P, when the same partitioning is applied to (6.67):

$$\begin{pmatrix} \hat{\theta}_1 - \theta_1^* \\ \hat{\theta}_2 - \theta_2^* \end{pmatrix} = \begin{pmatrix} F_1^t F_1 & F_1^t F_2 \\ F_2^t F_1 & F_2^t F_2 \end{pmatrix}^{-1} \begin{pmatrix} F_1^t \\ F_2^t \end{pmatrix} \varepsilon \tag{6.79}$$

and the corresponding part for $\hat{\theta}_2 - \theta_2^*$ is selected:

$$\hat{\theta}_2 - \theta_2^* = (F_2^t(I - P_{F_1})F_2)^{-1} F_2^t(I - P_{F_1})\varepsilon \tag{6.80}$$

Substituting equation (6.80) in equation (6.78), the latter becomes of the form $\varepsilon^t P \varepsilon$ for $P = (I - P_{F_1})F_2(F_2^t(I - P_{F_1})F_2)^{-1}F_2^t(I - P_{F_1})$ Now the projection in (6.76) and the projection $(I - P_{F_1})F_2(F_2^t(I - P_{F_1})F_2)^{-1}F_2^t(I - P_{F_1})$ are orthogonal. In all, this yields the test statistic for the subset of n_{θ_2} parameters in θ_2

$$F_{subtest} = \frac{\frac{\varepsilon^t(I - P_{F_1})F_2(F_2^t(I - P_{F_1})F_2)^{-1}F_2^t(I - P_{F_1})\varepsilon}{n_{\theta_2}}}{\frac{\varepsilon^t(I - P_F)\varepsilon}{IJQ - n_\theta}}$$

$$= \frac{(\hat{\theta}_2 - \theta_2^*)^t F_2^t(I - P_{F_1})F_2(\hat{\theta}_2 - \theta_2^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\theta}{n_{\theta_2}} \sim \mathcal{F}_{n_{\theta_2}, \, IJQ - n_\theta}. \tag{6.81}$$

In case of constrained parameters the matrix $\mathcal{I} = F(\theta)^t F(\theta)$, is singular, whereas the matrix $V^t\mathcal{I}V$ is nonsingular, see (6.40). This can be interpreted as a parameter transformation (reduction):

$$\beta = V^t\theta, \tag{6.82}$$

since then $\mathcal{I}(\beta) = F(\beta)^t F(\beta) = V^t\mathcal{I}(\theta)V$. β is an unconstrained parameter of length $PQZ + M + ZJ - 2P - Z^2$. Therefore, the above derived tests, (6.73) and (6.81), can be applied to β and subsets of β. The analogous expression for (6.73) in terms of β is

$$\frac{(\hat{\beta} - \beta^*)^t V^t\mathcal{I}(\theta)V(\hat{\beta} - \beta^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\beta})\|^2} \frac{IJQ - n_\beta}{n_\beta}$$

$$= \frac{(\hat{\theta} - \theta^*)^t V(V^t\mathcal{I}(\theta)V)V^t(\hat{\theta} - \theta^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\beta}{n_\beta} \sim \mathcal{F}_{n_\beta, \, IJQ - n_\beta} \tag{6.83}$$

Hence, the difference made by the constraints is the substitution of a pseudo-inverse of Σ_θ for $(F^t F)^{-1}$ and the adaption of the number of degrees of freedom. Similarly, for a subset of parameters, β_2, the test statistic becomes

$$\frac{(\hat{\beta}_2 - \beta_2^*)^t([(V^t\mathcal{I}(\theta)V)^{-1}]_{\beta_2})^{-1}(\hat{\beta}_2 - \beta_2^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\beta})\|^2} \frac{IJQ - n_\beta}{n_{\beta_2}}. \tag{6.84}$$

where $[..]_{\beta_2}$ indicates the submatrix corresponding to the subset of parameters in β_2. Now the coupling parameters are left unchanged by the parameter transformation (6.82), see (6.41), hence we can take $\theta_2 = \beta_2$ containing only coupling parameters, and the resulting test statistic is

$$F_{coupling} = \frac{(\hat{\theta}_2 - \theta_2^*)^t([\Sigma_\theta]_{\theta_2})^{-1}(\hat{\theta}_2 - \theta_2^*)}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - n_\beta}{n_{\theta_2}} \sim \mathcal{F}_{n_{\theta_2}, \, IJQ - n_\beta} \tag{6.85}$$

with $n_\beta = PQZ + M + ZJ - 2P - Z^2$. To test the significance of a certain source in data set q, θ_2 should contain all coupling parameters in the corresponding row of C_q.

6.5.6   Confidence regions around estimated STFs

In this appendix the confidence region around the estimated STFs for each source in each data set is computed. The same notation is used as in appendix 6.5.5. In [95] testing the l-dimensional nonlinear hypothesis $\mathbf{h}(\theta) = \mathbf{h}^*$ is outlined using a linear approximation. Define

$$H(\theta) = \frac{\partial \mathbf{h}(\theta)}{\partial \theta} \in \mathbb{R}^{l \times (PQZ+M+ZJ)} \tag{6.86}$$

the derivative matrix of the hypothesis vector. Then approximately

$$\mathbf{h}(\theta) - \mathbf{h}(\theta^*) \approx H(\theta)(\theta - \theta^*). \tag{6.87}$$

For notational simplicity the dependence of H on θ is dropped in the sequel. The $l \times l$ covariance matrix of $\mathbf{h}(\theta) - \mathbf{h}(\theta^*)$ is then approximated by

$$\widehat{Cov}(\mathbf{h}(\theta) - \mathbf{h}(\theta^*)) \approx H\widehat{Cov}(\theta - \theta^*)H^t = H\Sigma_\theta H^t. \tag{6.88}$$

In all this yields the following test statistic

$$F_{hyptest} =$$

$$\frac{(\mathbf{h}(\hat{\theta}) - \mathbf{h}(\theta^*))^t (H\,\Sigma_{\hat{\theta}}\,H^t)^{-1}(\mathbf{h}(\hat{\theta}) - \mathbf{h}(\theta^*))}{\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2} \frac{IJQ - (PQZ + M + ZJ - 2P - Z^2)}{l}$$

$$\sim \mathcal{F}_{l,\ IJQ-(PQZ+M+ZJ-2P-Z^2)} \tag{6.89}$$

The estimated p^{th} STF at a certain time sample j in a certain data set q is a hypothesis vector $h_{p,q,j} \in \mathbb{R}^{1 \times 1}$. The derivative matrix $H \in \mathbb{R}^{1 \times (PQZ+M+ZJ)}$ vector. Furthermore, $h_{p,q,j}$ is the product of pth row in the qth coupling matrix and the jth column of the basic STF matrix. Hence, for given q, p and j

$$h_{q,p,j}(\theta) = (C_q B)_{p,j} = (C_q \mathbf{B} W_T^{-1})_{p,j} \tag{6.90}$$

$$\frac{\partial h_{q,p,j}(\theta)}{\partial (C_q)_{p,z}} = (\mathbf{B} W_T^{-1})_{z,j} \tag{6.91}$$

$$\frac{\partial h_{q,p,j}(\theta)}{\partial \mathbf{B}_{z,j'}} = (C_q)_{p,z}(W_T^{-1})_{j',j} \tag{6.92}$$

$$\tag{6.93}$$

with $z = 1, \ldots, Z$ and $j' = 1, \ldots, J$. Using these derivatives the vector $H(\theta)$ is determined. To compute the (1-α) confidence region for a certain $h_{q,p,j}(\theta)$ the threshold value F_{thres} is defined as

$$F_{thres} = \mathcal{F}^\alpha_{1,\ IJQ-(PQZ+M+ZJ-2P-Z^2)} \tag{6.94}$$

and the confidence region is then given by

$$\left\{ h \,\middle|\, |h_{q,p,j}(\hat{\theta}) - h| \le \sqrt{\frac{F_{thres}\|\mathbf{r} - \mathbf{f}(\hat{\theta})\|^2 (H\,\Sigma_{\hat{\theta}}\,H^t)}{IJQ - (PQZ + M + ZJ - 2P - Z^2)}} \right\}. \tag{6.95}$$

Chapter 7

Summary, discussion and conclusions

In this thesis, two main questions are addressed. The first is about the separation of the signal and noise components in measured MEG (and EEG) data and is discussed in chapters 2, 3 and 4. In the next two sections, 7.1 and 7.2, the findings of these three chapters are summarised, discussed and conclusions are drawn. The second major topic of this thesis is about combining multiple related data sets into one single model in order to improve the solution to the inverse problem. Chapters 5 and 6 present the coupled dipole model and its extension which allow for a simultaneous analysis of different data sets. A summary of these two chapters is presented in section 7.3 and the results are discussed in section 7.4.

7.1   Summary chapters 2, 3 and 4

The standard approach to average recorded MEG signals time-locked to the stimulus (section 1.3.1) in order to find the evoked response is invalid when this response varies from trial to trial. In chapter 2 a model that accounts for amplitude variation of the brain response is presented and maximum likelihood (ML) estimators for both the noise and the signal parameters are derived. In this model, the spatial and temporal correlations of the background noise are taken into account. The model is expressed by (see equation (2.2))

$$r_{ij}^k = \alpha^k r_{ij} + \varepsilon_{ij}^k, \tag{7.1}$$

where r_{ij} is the normalised brain response and α^k denotes the trial dependent amplitude. The ML estimator for the α^k, $k = 1, \ldots, K$ is found by solving by an eigenvalue problem.

This model was applied to two different kinds of experimental data. Firstly, SEF data of 9 healthy subjects were subjected to this model. From all but one subjects two data sets, one recorded during stimulation of the left median nerve and one for the right median nerve, were analysed. It appeared that the late component of the response to the SEF stimulus, ranging from 30 to 120 ms post-stimulus, showed a negative trend in 15 of the 17 data sets, and a slightly positive trend in the remaining 2 data sets. This indicates that the response weakens over trials, possibly indicating a habituation effect.

For the early component, ranging from 20 to 35 ms post-stimulus, no such systematic effect was found.

Secondly, this model was applied to MEG data containing epileptic spikes. The different polarities and magnitudes of the spikes occurring in data of an epilepsy patient can be regarded as varying amplitudes. Data epochs around the marked spikes were cut and used as trials. Regarding the spatiotemporal covariance of the residuals, this application shows that temporal nonstationarities diminish when the amplitude jitter is taken into account, whereas a clear peak in the variance is seen when this is omitted. This finding is in line with Figure 1.13 in the introduction, where the effect of amplitude variation was investigated.

The question whether similar temporal nonstationarities are present in the residuals of evoked field MEG data is addressed in chapter 3. Considering the trial-to-trial variation found in the SEF data of chapter 2, one would expect to find a nonstationary temporal pattern in the temporal covariance of such data. In order to study this effect and, moreover, to better understand the abstract, ML estimated matrix (see equation (1.64)), the temporal covariance of MEG residuals is parameterised in chapter 3.

For the spatial covariance, models have been derived already which can physiologically be interpreted by assuming background activity to be generated by randomly distributed dipolar sources having amplitude functions independent of the source locations [16]. This spatial covariance appears to be a function of sensor distance. In order to investigate whether the temporal covariance has a similar property, the ML estimated temporal covariance matrices of SEF data sets of several subjects are considered.

It appears that alpha activity plays an important role in the temporal covariance, as expected based on section 1.4.1 and Figure 1.11(a). Therefore, the background activity is modelled as the sum of an alpha component and an additional noise component. The simple ongoing alpha model assumes that the alpha component is one everlasting alpha wave, whereas the more advanced Poisson modulated alpha model (PoMAM) assumes that the alpha activity consists of separate waves. These waves all have one fixed duration, although the intermediate time is generated by a Poisson(λ) process. Furthermore, the trial onsets are assumed to be random. The model equation of the PoMAM is

$$\varepsilon_k(t) = \Omega 1_{[0,T_\alpha]}(t - t_{onset}) \sin\left(\omega t + \tau_k'\right) + \eta_k(t) \tag{7.2}$$

where Ω is the amplitude of the waves, T_α the duration, ω the frequency, t_{onset} the onset time of the wave and τ_k the random phase in each trial. The covariance of the alpha part of the PoMAM is a function of Ω, T_α, ω and λ. The covariance of the additional noise part is assumed to be exponentially decreasing with time lag, described by the parameters σ^2 (variance) and κ (covariance length).

In principle, the PoMAM is temporally stationary, though in practise this stationarity is disturbed by the baseline correction (BC). The BC is necessary in order to remove unwanted baseline shifts caused by external noise. The stationarity of the alpha part of the temporal covariance highly depends on the chosen length of the correction window. Taking one or a multiple of alpha periods as window turns out yielding the most stationary covariance structure after correction.

The PoMAM was fitted to ML estimated temporal covariance matrices of SEF data sets of 5 healthy subjects. The PoMAM matrices were calculated based on exactly

the same BC as taken in the ML estimation. These BC windows were not optimally chosen, thus allowing for a comparison of the predicted nonstationarities and the real nonstationarities in the ML estimated matrices. It turns out that, taking this BC into account correctly, the ML estimated matrices can be approximated up to an error of less than 1% in relative matrix power by the PoMAM.

This small error suggests that the temporal covariance is very stationary, indicating in turn, that no considerable trial-to-trial variations occur. However, it may be that the structure of the covariance model, the Kronecker product is too simple and forces the pattern to be stationary. Therefore, an extension of this model is examined in chapter 4, which may reveal nonstationarities.

The single Kronecker product (KP) model of equation (1.54) is extended to a sum of KP in chapter 4:

$$\Sigma = \sum_{n=1}^{N} T_n \otimes X_n. \tag{7.3}$$

Contrary to the single KP, this extension allows for describing multiple, independent phenomena in the ongoing background activity. Whereas the single KP model can be interpreted by assuming that background activity is generated by randomly distributed dipoles with certain spatial and temporal characteristics, the sum model can be interpreted physiologically by assuming a composite of such processes.

The terms in the sum are estimated through minimisation of the squared differences between the spatiotemporal sample covariance matrix and the model. Optimally, these terms should be estimated within the ML paradigm, though in practise that becomes rather intricate. Theoretically, by rearranging the elements of the covariance matrices, the LS estimators are found from an SVD of the rearranged sample covariance matrix. The best rank N approximation of this rearranged sample covariance matrix defines the estimators for $vec(T_n)$ and $vec(X_n)$. Taking enough terms in the sum into account, i.e. taking N large enough, the spatiotemporal sample covariance matrix is described exactly by the sum model. In practise though, the computation of this SVD is not feasible, due to the large dimension of the matrices. Therefore, an iterative system of estimators is derived to find T_n and X_n.

The terms in the estimated sum of KP can be rewritten, while the KP structure is left intact. In fact, this rewriting exploits the freedom in the representation of a best rank N approximation. Not all terms in the initially estimated sum of KP are interpretable as covariance matrices, due to their orthogonality. The rewriting can be used to transform the initially estimated sum into a sum of physiologically interpretable terms. In order to interpret a matrix as a covariance matrix, it should be positive (semi-)definite. Therefore, the cost function of this rewriting minimises the negative parts of the matrices in all KP terms in the sum.

The sum of KP model was applied to different kinds of data (SEF, VEF and AEF) for multiple subjects. It appears that the first KP roughly describes between 62% and 91% of the sample covariance matrix in terms of matrix power and the second between 1% and 12%. The sum of 2 KP explains between 67% and 93%. The higher the order of the term, the smaller the amount of explained power. Regarding this power distribution, taking into account more than 1 KP term in dipole localisation is not expected to yield a significant improvement. Considering the improvement already made by one

KP compared to no or only spatial correlations [20], a major further enhancement of the parameter accuracy is not expected for the second and higher order terms in the sum of KP. Moreover, in practise the incorporation of multiple terms requires an nonconventional amount of computer power.

For all data sets studied in chapter 4, the rewritten sum of 2 KP reveals two physiological processes in the background activity: focal alpha activity and a more widespread not frequency specific process. This separation is in line with the parametrisation of the noise in the temporal domain by the PoMAM in chapter 3, be it that the PoMAM does not distinguish different spatial patterns for the two components. The entireness of the separation varies somewhat over subjects, though is clearly visible in all data sets. This separation is not caused by a priori constraints of the rewriting, though results from minimising the cost function. The explanation for this identifiable separation may be that, amongst all background processes, the alpha process is the most prominent. Determining two independent processes in that case, could separate this eminent process from the other processes, which are combined in one large remainder process. To study this explanation, data with no background alpha activity or data with another dominant (pathological) background process should be studied.

7.2   Discussion on spatiotemporal covariance

The central question in the first part of this thesis is: what happens to the ongoing brain activity when a stimulus is presented to the subject? The SPN model of equation (1.24) gives a simple answer to this question: the stimulus evokes a fixed response additional to the ongoing background activity. However, as discussed in the introduction and shown in chapter 2 this answer is debatable. The response is likely to vary from trial to trial to some extent, especially when the response is assumed to be built by modulated background activity [65]. In chapters 3 and 4 the computed temporal covariances based on the SPN assumption are very stationary. Nevertheless, these results do not rule out possible response variations in the data for two reasons. Firstly, the trial-to-trial variations in the response may be very small and, therefore, be dominated by stationary processes in the covariance matrices in matrix power. Secondly, even though the estimated (sum of) KP may not contain nonstationarities, the part of the sample covariance matrix that is not accounted for by the KP model, i.e. the difference between the model and the sample covariance matrix, can contain nonstationarities. The latter can be investigated by fitting a Toeplitz matrix to the difference between the KP model and the sample covariance matrix in the spatiotemporal dimension. Thus, the simple SPN model needs to be advanced.

In fact, chapters 2, 3 and 4 all present an elaborated version of the simple SPN model. Whereas the SPN model simply states

$$r_{ij}^k = r_{ij} + \varepsilon_{ij}^k, \tag{7.4}$$

the models presented in those chapters either parameterise the response model r_{ij} or describe ε_{ij}^k and its statistical properties in further detail. Combining all the presented

approaches, one arrives at the generic model:

$$r_{ij}^k = r_{ij}(k) + \sum_{n=1}^{N} \sum_{p=1}^{P_n} \Psi(\zeta^{k,n})_{ip} S_{pj}^{k,n} + \eta_{ij}^k. \tag{7.5}$$

In this universal model, the response $r_{ij}(k)$ depends on trial k, thus accounting for response variation and η_{ij}^k is random white noise. The term $\sum_{n=1}^{N} \sum_{p=1}^{P_n} \Psi(\zeta^{k,n})_{ip} S_{pj}^{k,n}$ presents the background activity that is generated by N random dipole processes (chapters 3 and 4) [33]. Each of these processes consists of P_n randomly distributed dipoles with source parameters $\zeta^{k,n} = (\zeta_1^{k,n}, \ldots, \zeta_{P_n}^{k,n})$. The corresponding source time function of each dipole p, $S_{pj}^{k,n}$, is assumed to be independent of the location and orientation in $\zeta_p^{k,n}$, and of the source time functions at other locations p'. This assumption leads to a Kronecker product structure in the spatiotemporal covariance of the nth process, $T_n \otimes X_n$, as shown in [16]. The composite of N such independent processes then has a sum of N Kronecker products as covariance structure (chapter 4), because the sum of independent Gaussian variables again has the normal distribution. Each of these processes presents an ongoing process in the brain, for example alpha activity, pathological theta activity or beta waves. The entire noise component $\sum_{n=1}^{N} \sum_{p=1}^{P_n} \Psi(\zeta^{k,n})_{ip} S_{pj}^{k,n}$ may be interpreted as the so-called resting state network [34, 87]. Clearly, trial-to-trial variations are now accounted for by both the response component and the noise component. In matrix notation this generic model becomes

$$R^k = R(k) + \sum_{n=1}^{N} \Psi(\zeta^{k,n}) S^{k,n} + E^k. \tag{7.6}$$

The likelihood function of this general model is

$$f(R(k), T_n, X_n) =$$

$$\frac{\exp\left[-\frac{1}{2} \sum_{k=1}^{K} (vec(R^k - R(k)))^t (\sum_{n=1}^{N} T_n \otimes X_n)^{-1} vec(R^k - R(k))\right]}{(2\pi)^{IJK/2} |\sum_{n=1}^{N} T_n \otimes X_n|^{K/2}}. \tag{7.7}$$

Regarding the response model $R(k)$, different levels of parametrisation of trial-to-trial variations are possible. The least parametrisation is by keeping $R(k) = R$ fixed over trials. A first extension is by allowing amplitude jitter or habituation effects as presented in chapter 2, $r_{ij}^k = \alpha^k r_{ij}^k$. Also latency jitter can be taken into account by adding nonlinear parameters, $r_{i,j}^k = \alpha^k r_{i,(j-l_k)}^k$, with l_k the latency shift in trial k. Finally, when the responses differ in 'speed', i.e. when some are stretched in time compared to others, which can be called frequency jitter, the model can be further adapted to $r_{i,j}^k = \alpha^k r_{i,(a_k j - l_k)}^k$ where a_k is the nonlinear parameter denoting the relative rapidity of the response in the kth trial. An example of this frequency jitter could be epileptic spikes that vary in sharpness, i.e. time scale. For all the different levels of parametrisation of trial-to-trial variation, the normalised (unit) response r_{ij} is modelled as activity of a combination of dipolar, multi-polar and distributed sources. To obtain more insight in how to model the amplitude, latency and frequency jitter in the response model, one can cluster single trial responses. In this additional method of examining

responses of the brain to external stimuli, single trial responses are clustered like the epileptic spikes in chapter 2 [57, 106]. The different clusters will contain different categories of occurring responses in the data set, possibly yielding information about the relation between the ongoing pre-stimulus activity and the following response. This approach, though, is more likely to be successful when the signal-to-noise ratio in single trials is higher.

Considering the noise parameters, the general model allows for a further parametrisation of the spatiotemporal covariance matrix. Combining the findings of chapters 3 and 4 suggests a combination of the Poisson modulated alpha model and the sum of KP model. The two noise components in the PoMAM can be regarded as two independent background processes, thus leading to a model with $N = 2$. One can extend this idea by modelling the additional noise component as different processes, that is, by increasing N further. The temporal matrix associated with the alpha activity can be parameterised as in the PoMAM. When the distribution of the source parameters $\zeta^{k,n}$ of each process is known, also the spatial covariance matrices can be parameterised using the forward fields in $\Psi(\zeta^{k,n})$. If, for example, the alpha activity is assumed to be generated by an occipital symmetric dipole pair with source parameters ζ_α (see Figures 4.3(b), 4.5(b) and 4.7(b)) with forward field Ψ, the corresponding spatial covariance matrix is parameterised by $\Psi(\zeta_\alpha)\Psi(\zeta_\alpha)^t$.

Now estimating all the parameters in the ML sense has become rather complicated by the complex parametrisation. For example, the inverse and determinant of a sum of Kronecker products cannot be computed on smaller dimensions, as was the case for a single KP (see equation (1.55)). Hence, differentiation of the likelihood function with respect to the covariance matrices and equating the obtained derivatives to zero yield equations that are difficult to solve. This is the reason why in chapter 4 the LS paradigm is used in stead of the ML paradigm. Furthermore, the number of nonlinear parameters in both the model and noise components troubles the estimation method. Considering the required two step approach in the estimation of the parameters in the 'simple' SPN model presented in chapter 1, which was based on the averaged measured signal as response model with $N = 1$, the addition of all the (nonlinear) parameters in the generic model certainly requires multiple steps in the parameter estimation. A complicated system of alternating least squares at different levels in the estimation method will hopefully allow for the estimation of the generic model parameters to some degree, though the full parametrisation as described above will remain very difficult to estimate in practise, let alone optimal ML estimation of all parameters.

7.2.1   Stationarity of the residuals in two utmost cases

Regarding the mentioned theoretical concerns, a more practical approach is by studying the response and noise component in two 'extreme' situations. The first kind of data is, so-called, spontaneous data, recorded in absence of external stimuli. In that situation, the ongoing background, or spontaneous, activity is not influenced by any stimulus. These data allow for an investigation of only the noise components in model equation (7.6). The second utmost is data recorded during a *selective attention task* experiment. In such an experiment, the processing of external stimuli is modulated by requiring a certain response from the subject (e.g. pressing a button) for one target stimulus,

while the other stimuli do not require any reaction. Such a task, requiring some kind of reaction from the subject, usually leads to a variety in reaction times. Hence, data of a selective attention task are expected to generate a nonstationary temporal covariance under the simple SPN assumption, as discussed in section 1.4.2. Results of a preliminary investigation of these two kinds of data are discussed in this section. These two studies are helpful to give a scientific answer to the question what influence an external stimulus has on the ongoing background activity. Apart from the scientific answer, there are practical issues that need to be taken into account because, like the baseline correction, they can induce artificial, misleading temporal nonstationarities. Examples of these practical complications are considered in the next section, 7.2.2.

Spontaneous data

Spontaneous data should be the ultimate example of data that is temporally stationary on small time scale (less than 1 second) because of the absence of stimuli that disturb the ongoing background activity. No trial-to-trial variations are expected in such data, because there are no external stimuli that can cause time-locked reorganisation of the background activity. It is important to randomise the (artificially created) trial onsets, in order to avoid a locking between ongoing oscillations and trial onsets (e.g. exact 10 Hz oscillations in 1000 ms trials).

Spontaneous data of two subjects were recorded continuously in epochs of 10 seconds during 5 minutes at a sample rate of 312 Hz at 151 MEG sensors. Off-line the data were low-pass filtered at 50 Hz in order to remove high frequency noise. New data sets were created by selecting from each original epoch 9 new baseline corrected trials (epochs) with randomised onsets, each of which contained 1 second of data. As window for the baseline correction one alpha period was taken, corresponding to the alpha peak in the power spectrum. In all, this yielded new data sets consisting of 270 trials. These data sets were subjected to the sum of KP model presented in chapter 4. The resulting temporal matrices for subject SPON1 are plotted in Figure 7.1. The results of subject SPON2 are discussed in section 7.2.2.

The temporal structure of the spontaneous data of subject SPON1 is very stationary, as can be seen from Figure 7.1. In fact, to quantify the level of stationarity, one should fit a Toeplitz matrix to the estimated temporal covariance matrices. However, in general, large nonstationarities are clearly visible from colour plots. The data presented in Figure 7.1 do not show big colour fluctuations along the subdiagonals apart from the gradual increase due to the BC (see chapter 3). Hence, this example of spontaneous data confirms the expectation of having a stationary temporal covariance.

Selective attention VEF data

In this section a tentative investigation of the temporal nonstationarities in a selective attention experiment is presented. The stimuli in this experiment consisted of different visual patterns. After presentation of the target stimulus, the subject was asked to press a button, while the nontarget stimuli did not require a reaction from the subject. Line patterns with two different orientations (horizontal and vertical) and two different spatial frequencies (low and high) were presented full-field to the subject. In Figure 7.2

(a) First term in orthogonal sum (51.4%)

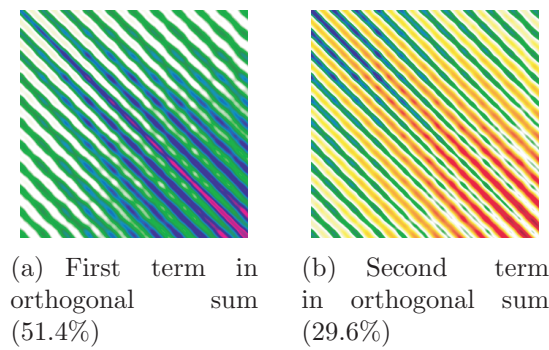(b) Second term in orthogonal sum (29.6%)

Figure 7.1: *Estimated temporal matrices in the spontaneous data of subject SPON1. Frames a and b show the first two orthogonally estimated temporal matrices. The time scale is 995 ms by 995 ms. The entries of the matrices are plotted in colour. The percentages show the relative matrix power of the sample covariance matrix explained by the two KP terms.*

similar stimuli are drawn. The MEG session consisted of four different experiments,

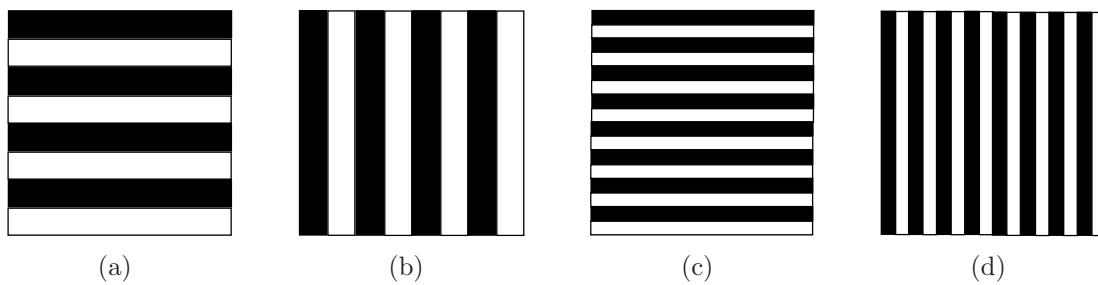(a)　　　　　　　(b)　　　　　　　(c)　　　　　　　(d)

Figure 7.2: *Line patterns similar to the visual stimuli used in the selective attention experiment. Figure a shows the horizontal pattern with low spatial frequency, figure b the vertical low spatial frequency, figure c the horizontal high spatial frequency and figure d the vertical high spatial frequency pattern.*

in each of which a different pattern functioned as the target stimulus. This stimulus is called the $(+,+)$ stimulus, while the other stimuli are called $(-,-)$, $(+,-)$ and $(-,+)$. The order of the signs is (orientation, frequency). For example, if Figure 7.2(a) were the target $(+,+)$ stimulus, Figures 7.2(b), 7.2(c) and 7.2(d) would be the $(-,+)$, $(+,-)$ and $(-,-)$ stimuli, respectively. In each experiment the four stimuli were presented in random order with an inter stimulus interval uniformly between 1900 and 2300 ms. The number of trials in each experiment was around 250 trials, yielding approximately 62 trials per stimulus. MEG data were recorded at a sample rate of 625 Hz at 151 MEG sensors for two subjects, ATT1 and ATT2.

By the low number of trials for each stimulus in each experiment (approx. 62) the analysis of the corresponding MEG residuals is hampered. This is one of the practical issues discussed in section 7.2.2. Therefore, the MEG residuals of the $(+,+)$ conditions in all four experiments were taken together in one $(+,+)$ residual data set. To create this

larger residual data set, the average of the (+,+)-trials was subtracted using the optimal BC window in each of the four experiments separately, and the resulting residuals were combined in the joint (+,+) residual data set consisting of approximately 250 trials. This number of trials should be enough to investigate nonstationarities in the residuals. Similarly, (+,-), (-,+) and (-,-) residual data sets were created.

The response to the stimuli in this experiment consists of two kinds of components, exogenous and endogenous. The former is the obligatory response from the (visual) cortex to the stimulus, while the latter is the response to the meaning of the stimulus, i.e. whether to attend to the stimulus and to press the button. By creating the big residual data sets, trial-to-trial variations of the exogenous components are mixed for the four line patterns. When, for example, the visual response to the thick horizontal line pattern in Figure 7.2(a) varies differently from trial to trial than the response to the thin horizontal line pattern in Figure 7.2(c), the big residual data sets contain a mixture of different nonstationary patterns stemming from the different stimuli. Nevertheless, if so, this mixture is present in all four joint residual data sets, and differences in stationarity between these joint residual data sets cannot be explained by this mixture. A true reason for differences in stationarity in the four residual data sets lies in different trial-to-trial variations of the endogenous components, which are not mixed by the creation of the joint residual data sets.

The data of subject ATT1 are discussed in section 7.2.2. The data of the second subject, ATT2, show interesting differences between for the (-,-) and (+,+) conditions. For both conditions, the first two orthogonally estimated KP terms were computed according to the method presented in chapter 4. In Figure 7.3 the so obtained temporal

(a) First orthogonal term (87.5%) in (-,-) condition

(b) Second orthogonal term (1.39%) in (-,-) condition

(c) First orthogonal term (88.7%) in (+,+) condition

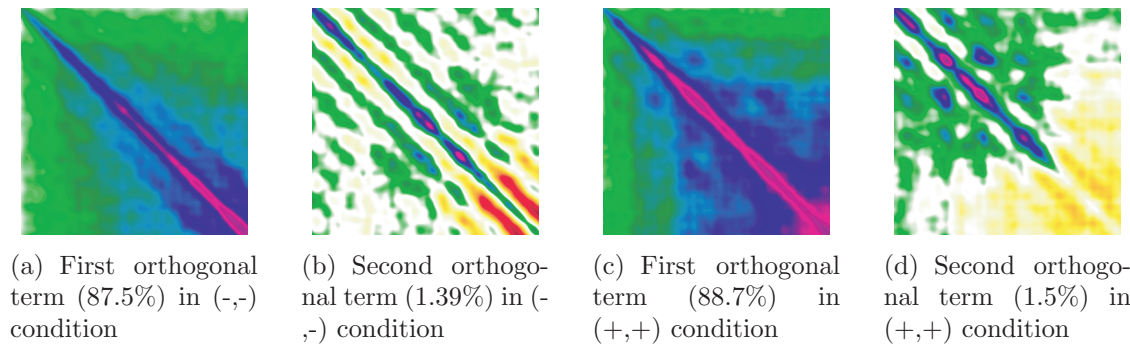(d) Second orthogonal term (1.5%) in (+,+) condition

Figure 7.3: *Estimated temporal matrices in the (-,-) and (+,+) residual data set of subject ATT2. Frames a and b show the first two orthogonally estimated temporal matrices in the (-,-) condition and frames c and d show the first two orthogonally estimated temporal matrices in the (+,+) condition. The time scale is 562 ms by 562 ms post-stimulus. The entries of the matrices are plotted in colour. The percentages show the relative matrix power of the sample covariance matrix explained by the KP terms.*

covariance matrices for both conditions are presented. The first matrix in the (-,-) condition in Figure 7.3(a) shows a very stationary pattern, apart from the gradual increase along the diagonal and subdiagonals caused by the BC. The first matrix of

the (+,+) condition in Figure 7.3(c) differs somewhat from that matrix, in having a clear nonstationary 'dot' between the time instants 125 ms and 250 ms post-stimulus. The latency of 125 ms roughly coincides with the onset of the first reported attention related component in this kind of data [2, 47]. The difference between the two conditions becomes bigger in the matrices of the second KP in Figures 7.3(b) and 7.3(d). Whereas in the (-,-) condition the temporal structure is rather stationary outside the interval (200, 400) ms post-stimulus, the (+,+) condition shows a dotted pattern in the interval (0, 400) ms post-stimulus, resembling the structure of the matrix in Figure 3.1(b), be it to a lesser degree. This difference in stationarity cannot be caused by a suboptimal BC, because the BC window was taken equal for both conditions. Hence, this figure suggests the presence of trial-to-trial variations in the (+,+) condition that are not present in the (-,-) condition. However, the (200, 400) ms interval post-stimulus in the (-,-) condition shows increased variance too, indicating that also in that condition the residuals are not purely stationary. In sum, both conditions seem to contain trial-to-trial variations. Examining the amplitude of the single trial signals and that of the averaged signal, one can conclude that the averaged response is in amplitude by far dominated by activity that does not build an average. This finding suggests, that the residuals contain important information about possible modulation of ongoing processes by the stimulus, and, therefore, that by simple averaging an important part of the data is discarded. To investigate this kind of data further, artificial nonstationarities caused by the creation of the big residual data sets need to be ruled out. This can be realised by recording more data in one experiment.

In conclusion, this small analysis of data of a selective attention experiment hints towards real response jitter, that may vary from condition to condition. However, the present study is far too small to infer conclusions about these nonstationarities in the residuals and more research is necessary to characterise this response variety.

7.2.2   Practical considerations regarding nonstationarities

When examining temporal nonstationarities to judge the presence of trial-to-trial variations in the response component, one needs to take into account some practical concerns. In this section, the number of trials and the shape of the frequency spectrum, which both can cause artifactual nonstationarities in the residuals, are discussed.

Number of trials

When only a small number of trials is available for estimating the spatiotemporal covariance (e.g. the sum of 2 KP), this may influence the stationarity features. For example, taking only 60 trials into account in stead of the 487 trials available of the 2SEF data set presented in chapter 4 yields a temporal pattern that is very different from the one presented in chapter 4, see Figures 7.5(a) and 7.5(b) versus Figures 7.5(c) and 7.5(d). The 60 trials yield a less stationary pattern than the 487 trials do.

Apparently, the sample of 60 trials is not big enough to describe the complete variety amongst the residuals. The complexity of the temporal properties of the residuals requires more trials in order to account for all possible variations in the background activity. Therefore, the temporal structure based on the 60 trials is biased, whereas the

487 trials contain enough variety to describe the possible phases and variations in the background oscillations.

This phenomenon has to be taken into account carefully when analysing small data sets. While a number of 60 trials may suffice considering the signal-to-noise ratio (e.g. experiments with high amplitude responses) to apply source localisation, this amount of data does not provide us with a complete overview of the properties of the residuals.

(a) subject SPON1   (b) subject SPON2
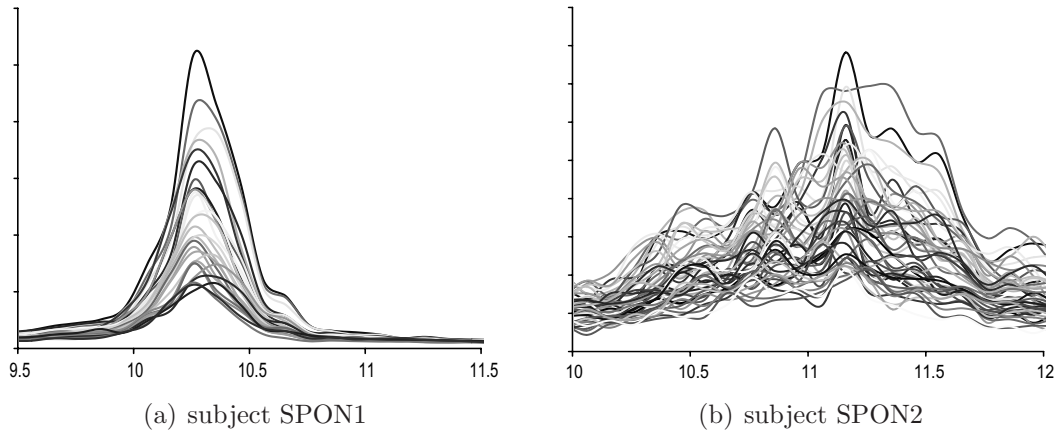
Figure 7.4: *The power spectra of the spontaneous data of subjects SPON1 and SPON2 around the alpha frequency (horizontal axis). In both figures a selection of channels is plotted that show the highest power in these frequencies. Figure a shows a peak for all channels at one and the same frequency, while figure b shows peaks at different frequencies for different channels.*

Frequency spectrum

Another practical issue that has influence on the temporal stationarity is the shape of the alpha peak in the frequency spectrum. Depending on the subject and the experimental circumstances, this peak is sharper or broader. This has a direct consequence on the optimal baseline correction (see section 3.2.3). In Figure 7.4 the power spectra of the two spontaneous data sets from section 7.2.1 are plotted.

There is a clear difference between the two spectra: in the data of subject SPON1 there is one sharp peak at 10.3 Hz, whereas a peak in the spectrum for subject SPON2 is not clear-cut. This implies for the second subject, that the BC window can not be chosen without having artificial nonstationarities introduced. To study the spatiotemporal covariance of spontaneous data, single epochs have to be created in the continuously measured signals. Cutting the data in 150 regular epochs of 2 seconds, the first two orthogonal KP were estimated for these spontaneous data sets of the two subjects. The resulting temporal matrices for the first 995 ms of each epoch are presented in Figure 7.6.

Clearly, the sharp alpha peak in the spectrum of subject SPON1 leads to a very stationary structure (as discussed in section 7.2.1), whereas the absence of a sharp peak for subject SPON2 leads to a less stationary structure. Randomising the trial onsets

(a) First orthogonal term for 487 trials

(b) Second orthogonal term for 487 trials

(c) First orthogonal term for 60 trials
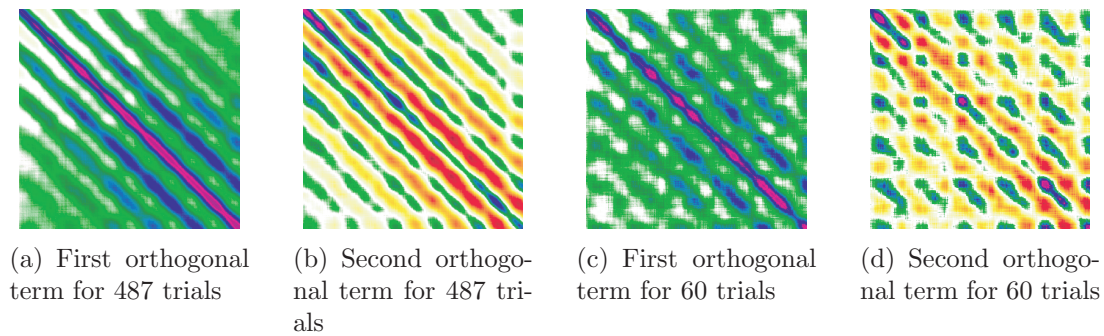
(d) Second orthogonal term for 60 trials

Figure 7.5: *Temporal covariance matrices of the 2SEF data set presented in chapter 4 for different numbers of trials. The time scale is 574 ms by 574 ms post-stimulus.*

(a) First orthogonal term for subject SPON1 (64.0 %)

(b) Second orthogonal term for subject SPON1 (18.7 %)

(c) First orthogonal term for subject SPON2 (81.6 %)

(d) Second orthogonal term for subject SPON2 (3.4 %)

(e) First orthogonal term for subject SPON2 (82.2%)

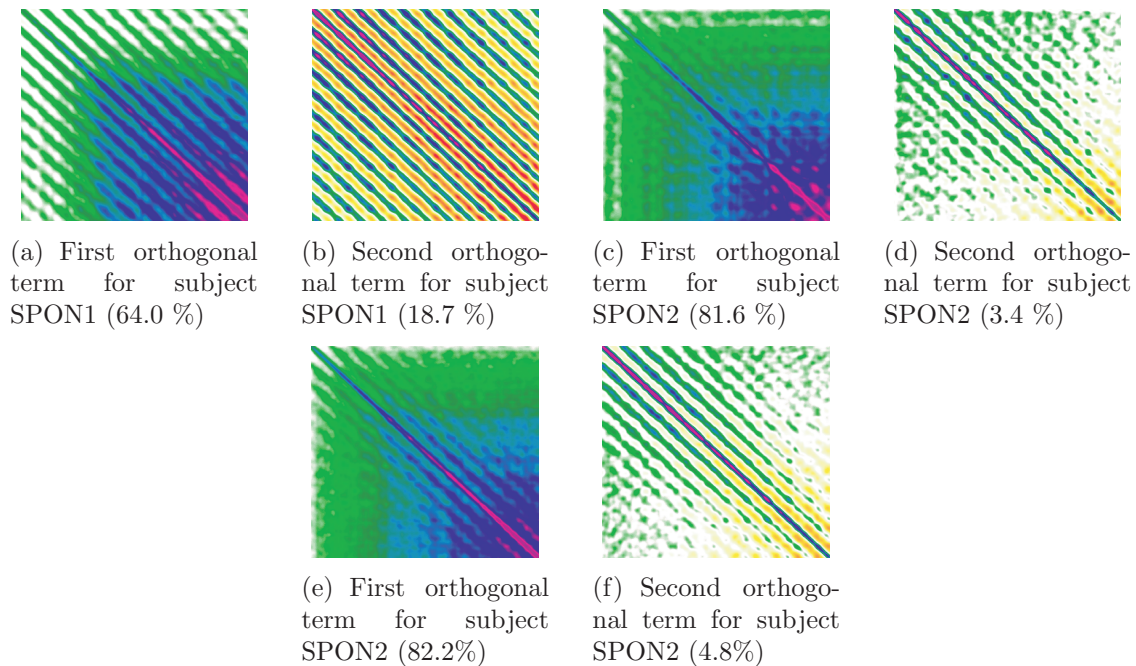(f) Second orthogonal term for subject SPON2 (4.8%)

Figure 7.6: *In figures a, b, c and d the first two orthogonally estimated temporal matrices for the spontaneous data sets based on 150 regular epochs of the two subjects of figure 7.4. For subject SPON2, the two matrices based on 270 epochs with randomised onsets are plotted in figures e and f. The length of the time window is 995 ms.*
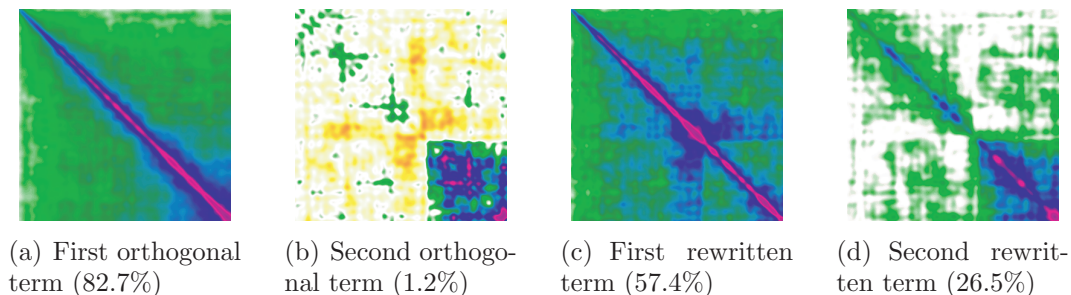

(a) First orthogonal term (82.7%)

(b) Second orthogonal term (1.2%)

(c) First rewritten term (57.4%)

(d) Second rewritten term (26.5%)

Figure 7.7: *Estimated temporal matrices in the (-,+) residual data set of subject ATT1. Frames a and b show the first two temporal matrices of the orthogonally estimated terms and frames c and d show the first two temporal matrices of the rewritten sum of two matrices. The time scale is 562 ms by 562 ms post-stimulus.*

and cutting more (smaller) epochs improves the stationarity of the two matrices in Figures 7.6(c) and 7.6(d), but the resulting temporal structure is still not as stationary as the structure for subject SPON1, see Figures 7.6(e) and 7.6(f).

Another property of the alpha peak is its height. For the data presented in chapter 4 it appeared that the rewritten sum of 2 KP consists of one term corresponding to alpha activity and one to other background activity. If there is (nearly) no alpha activity in the residuals present, due to suppression by a visual stimulus for example, this separation cannot occur due to the absence of alpha activity.

An example of this absent alpha activity is found in the (-,+) residual data set of subject ATT1 of the selective attention experiment, discussed in section 7.2.1. The temporal matrices of the sum of 2 KP and the rewritten sum for this data set are shown in Figure 7.7. In Figure 7.7(a) the first orthogonally estimated KP is shown. Apart from the BC effect, this matrix is very stationary and does not show alpha oscillations in the covariance. The second orthogonal term (Figure 7.7(b)) is very irregular and does not seem to posses a physiological meaning. This abstract pattern is likely to be caused by the orthogonality constraint, presented in chapter 4. Considering the amount of explained matrix power in the two terms, the second does not yield a major contribution. The rewritten terms (Figures 7.7(c) and 7.7(d)), although positive semi-definite, are both 'spoiled' by the irregular pattern. In all, this residual data set illustrates a possible negative effect of the artificial orthogonality constraint in the estimation of the sum of KP model as presented in chapter 4.

7.2.3   Conclusions about the validity of the Signal Plus Noise model

In the case of 'simple' MEG data, the Signal Plus Noise model seems to work adequately, even though external stimuli do influence the ongoing processes (e.g. closing eyes enhances alpha activity). Though in the case of more sophisticated experiments, like the presented selective attention task, the SPN seems not to separate adequately the recorded signals into a part influenced by the stimulus (the 'signal') and a part that is left unchanged by the external stimulus (the 'noise'). This suggests that real trial-to-trial variations on small time scale (less than 500 ms post-stimulus) are present in these data. This requires an adaptation of the SPN model in which response variation is modelled, thus accounting for modulation of ongoing processes by the stimulus [65, 89]. To better study this, more trials need to be recorded, in order to rule out the possibility of artifactual nonstationarities caused by too few trials.

7.3   Summary chapters 5 and 6

The presented coupled dipole model (CDM) and extended coupled dipole model (ECDM) are two models that can be used for the simultaneous analysis of related MEG/EEG data sets. By related data sets, data sets are meant that contain activity of the same sources or contain proportional source time functions. Both models are component models, that is, models that make use of a set of basic spatial components and a set of basic temporal components. The model for each data set under consideration is a linear

combination of these spatial and temporal components. The basic spatial components are given by a set of common dipole sources with normalised orientations, represented by a matrix A that contains the forward fields of these sources. The basic temporal components are unparameterised, normalised wave shapes that are gathered as rows in a basic matrix B. The SPN model for trial k in the qth data set is

$$R_q^k = AC_qB + E_q^k, \tag{7.8}$$

where C_q is the coupling matrix, that defines the linear combination of the basic components for the qth data set. Because A and B are normalised, the entries in C_q define the amplitude of the model. Furthermore, the spatiotemporal correlations are taken into account through prewhitening of the measured data arrays R_q^k by the matrices W_X and W_T. The joint, prewhitened SPN model for all averaged data sets is then given by

$$\mathbf{R} = \mathbf{ACB} + \mathbf{E} \tag{7.9}$$

where

$$\mathbf{R} = \begin{pmatrix} \sqrt{K_1}\tilde{R}_1 \\ \vdots \\ \sqrt{K_Q}\tilde{R}_Q \end{pmatrix} , \quad \mathbf{A} = (\mathbf{I}_Q \otimes W_X^t A) , \quad \mathbf{C} = \begin{pmatrix} \sqrt{K_1}C_1 \\ \vdots \\ \sqrt{K_Q}C_Q \end{pmatrix} , \quad \mathbf{B} = BW_T$$

$$\text{and} \quad \mathbf{E} = \begin{pmatrix} \sqrt{K_1}\tilde{E}_1 \\ \vdots \\ \sqrt{K_Q}\tilde{E}_Q \end{pmatrix} . \tag{7.10}$$

The numbers K_q indicate the number of trials in data set q for $q = 1, \ldots, Q$.

In the CDM the user defines for each data set which entries of the coupling matrix are nonzero and have to be estimated from the data. This selection allows the exploitation of prior knowledge or assumptions about the similarity of the different data sets, though it is suboptimal in terms of objectivity. Therefore, in the ECDM the coupling matrices are estimated entirely. The ECDM can be regarded as an extension of the CDM in the temporal domain: the wave shapes in $\mathbf{CB}$, i.e. the estimated source time functions for the sources in all data sets, are more general in the ECDM. Spatially, there is no noteworthy difference between the models, although the spatial parameters are influenced by the adaptation of the temporal parameters.

In chapter 5 the CDM is shown to improve the solution to the MEG inverse problem in two ways. Firstly, the simultaneous analysis appears to be less vulnerable to (local) minima in the separate cost functions that correspond to implausible solutions, because these minima occur at different solutions and, hence, do not produce a minimum in the joint cost function. Therefore, the CDM stabilises the solution to the IP in cases that the analyses of the different data sets separately do not (all) yield physiologically interpretable solutions. Secondly, by the incorporation of multiple data sets into the analysis, the SNR is relatively increased and the errors in the estimated model parameters in $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ drop. In a simulation study in chapter 5 using two dipoles in the visual cortex, the position error of the estimated sources is shown to drop by a factor up to 10 when 5 data sets are analysed simultaneously in stead of separately.

The difference made by the extension in the ECDM, is the removal of the subjectivity of the zero selection in the coupling matrices. In case the user is not certain about the similarity in the different data sets, the ECDM is more convenient. Only the dimension of the entirely estimated coupling matrices still needs to be chosen. The solution to the ECDM is not unique, because $\mathbf{C}$ and $\mathbf{B}$ are both unconstrained. Consequently, because $\mathbf{CB} = \mathbf{C}G^{-1}G\mathbf{B}$ for any nonsingular matrix G, $\mathbf{C}G^{-1}$ and $G\mathbf{B}$ constitute an equivalent solution. This nonuniqueness is solved by identifiability constraints on the matrix $\mathbf{B}$. Unfortunately, these constraints cause the direct physiological interpretation of $\mathbf{C}$ and $\mathbf{B}$ to disappear and be replaced by a more abstract one. This is the price to be paid for the increased objectivity of the extended model. Furthermore, due to noise in the measured signals, all entries of the coupling matrices will be estimated as nonzero; hence, estimated activity needs to be tested for statistical significance. Because of the statistical directness of the ECDM, this testing can be performed rather straightforwardly. On the contrary, testing the CDM parameters is more tedious and depends on the precise design of the coupling matrices.

In chapter 6 a comparison between the two models is made by applying both the CDM and the ECDM to the same VEF data. It appears that the estimated spatial parameters and the estimated product $\mathbf{CB}$ do not differ significantly between the two models for these data. Furthermore, the behaviour of the ECDM is tested in the case of overparametrisation by too many sources. Because of the entirely estimated coupling matrices, all sources will be estimated as active, though some of them will not be significantly active. It turns out that the activity of these supposedly *silent* sources is correctly detected as not significant in most cases. However, in some cases the straightforward test yields p-values slightly above the significance threshold. This may indicate the need for more conservative testing.

7.4   Discussion (extended) coupled dipole model

In the coupled dipole model more or fewer assumptions about the similarity of the data sets are reflected by selecting fewer or more zeroes, respectively, in the coupling matrices. Considering the extreme possibility of selecting no zeroes, the CDM has become equal to the ECDM. However, in that case, identifiability constraints are necessary, as shown in chapter 6. Apparently, the CDM is also subject to identifiability constraints in some situations. The higher the number of coupling parameters (i.e. the fewer zeroes in the coupling matrices), the more constraints necessary to obtain convergence of the CDM estimation algorithm. The examples studied in chapter 5 employed a number of coupling parameters that appeared to be small enough to let the estimation algorithm converge. In general, it is rather complicated to derive identifiability constraints for the CDM, due to the way of parametrisation. This consideration suggests that the ECDM can be regarded as a special case of the CDM, although the converse can also be defended because all CDM solutions are included in the ECDM solutions; hence, the name *extended* coupled dipole model.

7.4.1 MUSIC

The MUSIC algorithm [74] presented in section 1.3.3 can be applied to the simultaneous analysis of multiple data sets. The idea behind MUSIC is first to find an optimal projection matrix that minimises the cost function and then to scan the brain in order to find the optimal source locations. The first step only depends on the number of sources that is chosen to be fitted, while the second step is a 3D search requiring the solution to a generalised eigenvalue problem at each location.

In case of the ECDM an approach similar to the method described in section 1.3.3 is possible. The cost function for the ECDM, $\mathbf{R} = \mathbf{ACB} + \mathbf{E}$, is

$$\min_{\xi,\eta,\mathbf{C},\mathbf{B}} \|\mathbf{R} - \mathbf{ACB}\|_F^2 \tag{7.11}$$

with $\mathbf{R} \in \mathbb{R}^{IQ \times J}$, $\mathbf{A} = \mathbf{I}_Q \otimes \Psi \in \mathbb{R}^{IQ \times PQ}$, $\mathbf{C} \in \mathbb{R}^{PQ \times Z}$ and $\mathbf{B} \in \mathbb{R}^{Z \times J}$. Here, the matrix Ψ contains the prewhitened forward fields. For the first MUSIC step the trilinear model needs to be regrouped in two factors for estimating the projection matrix. This rewriting can be performed in two ways, I and II:

$$I : \quad \mathbf{R} = (\mathbf{AC})\mathbf{B} + \mathbf{E} \tag{7.12}$$
$$II : \quad \mathbf{R} = \mathbf{A}(\mathbf{CB}) + \mathbf{E}. \tag{7.13}$$

For the first possibility the dimension of the two regrouped terms are $IQ \times Z$ and $Z \times J$ and for the second possibility $IQ \times PQ$ and $PQ \times J$. It may be assumed that $PQ \geq Z$ without loss of generality, for a similar reasoning as $J \geq Z$ (see section 6.2.1). The estimators for the right hand term in the regrouped model in cases I and II are

$$I : \quad \hat{\mathbf{B}} = [(\mathbf{AC})^{\mathbf{t}}\mathbf{AC}]^{-1}(\mathbf{AC})^{\mathbf{t}}\mathbf{R} \tag{7.14}$$
$$II : \quad \widehat{\mathbf{CB}} = (\mathbf{A}^{\mathbf{t}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{t}}\mathbf{R} \tag{7.15}$$

and substituting these estimators yields the rewritten cost function

$$I : \quad \min_{\xi,\eta,\mathbf{C}} \|(\mathbf{I}_{IQ} - \mathbf{AC}[(\mathbf{AC})^{\mathbf{t}}\mathbf{AC}]^{-1}(\mathbf{AC})^{\mathbf{t}})\mathbf{R}\|_F^2 = \min_{\xi,\eta,\mathbf{C}} \|P_{\mathbf{AC}}^{\perp}\mathbf{R}\|_F^2 \tag{7.16}$$
$$II : \quad \min_{\xi,\eta} \|(\mathbf{I}_{IQ} - \mathbf{A}(\mathbf{A}^{\mathbf{t}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{t}})\mathbf{R}\|_F^2 = \min_{\xi,\eta} \|P_{\mathbf{A}}^{\perp}\mathbf{R}\|_F^2. \tag{7.17}$$

The ranks of the two possible projection matrices are

$$I : \quad rank(P_{\mathbf{AC}}^{\perp}) = IQ - Z \tag{7.18}$$
$$II : \quad rank(P_{\mathbf{A}}^{\perp}) = IQ - PQ. \tag{7.19}$$

Regarding these two ranks, the first will be the highest because $PQ \geq Z$. Hence, finishing the first MUSIC step by finding the optimal projection P that minimises the cost function,

$$I : \quad \hat{P}_I = \arg\min \|P_{\mathbf{AC}}^{\perp}\mathbf{R}\|_F^2 \tag{7.20}$$
$$II : \quad \hat{P}_{II} = \arg\min \|P_{\mathbf{A}}^{\perp}\mathbf{R}\|_F^2, \tag{7.21}$$

yields two possible lower bounds for the cost function of which the first will be the highest, i.e. the sharpest. Therefore, one could use this first MUSIC step to determine the number Z: by calculating this lower bound for different values of Z one can choose a value such that the lower bound is acceptably small.

Now the second step of the original MUSIC algorithm is to find the source parameters by spatial scanning. Scanning is possible because the joint cost function of all sources can be split into cost functions of each source separately, see equation (1.40). Such a splitting is also possible here, when rewriting II is chosen.

For rewriting II, the projection $\hat{P}_{II}$ is calculated in the first MUSIC step. In the second step, source parameters $\zeta = (\zeta_1, \ldots, \zeta_P)$ are found by minimising $\|\hat{P}_{II}\mathbf{A}\|_F^2$. Splitting $\hat{P}_{II}$ into Q blocks of I columns

$$\hat{P}_{II} = \left(\begin{array}{ccc} P^1 & \cdots & P^Q \end{array} \right) \tag{7.22}$$

and writing the forward field matrix Ψ as

$$\Psi = \left(\begin{array}{ccc} \Psi(\zeta_1) & \cdots & \Psi(\zeta_P) \end{array} \right) = \left(\begin{array}{ccc} \Psi_1 & \cdots & \Psi_P \end{array} \right) \tag{7.23}$$

where Ψ_p is the forward field matrix of the pth source, the cost function can be written as

$$\min_{\zeta_1,\ldots,\zeta_P} \|\hat{P}_{II}\mathbf{A}\|_F^2 = \min_{\zeta_1,\ldots,\zeta_P} \|\hat{P}_{II}(\mathbf{I}_Q \otimes \Psi)\|_F^2 = \min_{\zeta_1,\ldots,\zeta_P} \left\| \left(\begin{array}{ccc} P^1 & \cdots & P^Q \end{array} \right) \left(\begin{array}{ccc} \Psi & \mathbf{0} & \\ \mathbf{0} & \ddots & \mathbf{0} \\ & \mathbf{0} & \Psi \end{array} \right) \right\|_F^2$$

$$= \min_{\zeta_1,\ldots,\zeta_P} \left\| \left(\begin{array}{ccc} P^1\Psi & \cdots & P^Q\Psi \end{array} \right) \right\|_F^2 = \min_{\zeta_1,\ldots,\zeta_P} \sum_{q=1}^{Q} \|P^q\Psi\|_F^2 = \min_{\zeta_1,\ldots,\zeta_P} \sum_{q=1}^{Q} \sum_{p=1}^{P} \|P^q\Psi_p\|_F^2$$

$$= \min_{\zeta_1,\ldots,\zeta_P} \sum_{p=1}^{P} \left(\sum_{q=1}^{Q} \|P^q\Psi(\zeta_p)\|_F^2 \right). \tag{7.24}$$

This last expression resembles the original cost function in equation (1.40), though contains a summation over q in the cost function for each source. Hence, scanning can be performed similar to the original MUSIC scanning by computing the normalised cost function for each source ζ

$$Cost(\zeta) = \frac{\sum_{q=1}^{Q} \|P^q\Psi(\zeta)\|_F^2}{\|\Psi(\zeta)\|_F^2} \tag{7.25}$$

and selecting the P local minima. The similarity between this cost function and the original MUSIC cost function allows for the same approach for finding the orientation at each location, namely by solving a generalised eigenvalue problem. Hence, a 3D scanning over all locations suffices. After the source parameters have been estimated by this scanning, the matrix $\mathbf{A}$ can be computed, and the estimator for $\mathbf{CB}$ is given by $\widehat{\mathbf{CB}}$ in equation (7.15). Finally, the estimators $\hat{\mathbf{C}}$ and $\hat{\mathbf{B}}$ separately are obtained from the best rank-Z approximation of $\widehat{\mathbf{CB}}$.

In conclusion, the application of MUSIC for rewriting I allows for an objective determination of the number Z, while rewriting II enables the useful scanning method, circumventing the intensive nonlinear search procedure of the source location parameters described on page 104.

7.4.2   PARAFAC

An even higher lower bound for the cost function than by MUSIC can be obtained in a special case. When considering the CDM, rather than the ECDM, with purely diagonal coupling matrices, this model can be interpreted as a case of Parallel Factor Analysis (PARAFAC) or Canonical Decomposition (CANDECOMP) [10, 38]. The PARAFAC model equation for three-way arrays (e.g. space-time-condition) is

$$R_{i,j,q} = \sum_{p=1}^{P} a_{i,p} b_{j,p} c_{q,p}. \tag{7.26}$$

Relating this model to the CDM, the summation over p corresponds to the sources, $a_{i,p}$ is the unparameterised forward field on sensor i of source p, $b_{j,p}$ is the (normalised) source time function of source p at time sample j and $c_{q,p}$ is the amplitude (or coupling parameter) of source p in condition q. These factors are arranged in three, so-called, *loading* matrices, $A \in \mathbb{R}^{I \times P}$, $B \in \mathbb{R}^{J \times P}$ and $C \in \mathbb{R}^{Q \times P}$. The PARAFAC model is not restricted to three-way arrays but can also handle more general multi-way data arrays. The nice feature of PARAFAC is that the decomposition into the components is unique, when equation holds and the matrices A, B and C fulfill a certain rank constraint [10, 55]. However, when fitting this model to experimental data, equation will rarely hold without letting P increase beyond feasible limits. Nevertheless, for given CDM settings with diagonal coupling matrices, (e.g. simulation study 2 in chapter 5) fitting the PARAFAC model will yield the unparameterised forward fields in A, the coupling parameters in C and the unparameterised source time functions in B. This PARAFAC solution, then, yields a lower bound of the cost function in equation (7.11) which, in theory, is higher than the bounds given by the MUSIC projections in equations (7.20) and (7.21). The underlying reason is that the PARAFAC parametrisation is stricter than the MUSIC model, because in the MUSIC approaches I and II described above the coupling matrices in $\mathbf{C}$ are not parameterised as diagonal matrices.

   The PARAFAC model has been applied to space-time-frequency arrays of MEG data in [71], yielding a description of frequency specific behaviour over time and sensors. Another application of this model lies in the space-time-trial domain to investigate amplitude variations as in chapter 2. When it is assumed that all sources show the same source time function from trial to trial, though with varying amplitude, the PARAFAC decomposition can give an estimate of the amplitude variation over trials, as an alternative to the method presented in chapter 2. In that chapter, the normalised spatiotemporal pattern R was not parameterised any further, while the PARAFAC model approximates this matrix by a rank P matrix. In terms of source localisation, the PARAFAC model needs to be complemented by a nonlinear algorithm to find the source parameters that best fit the estimated unparameterised forward fields in the loading matrix A.

7.4.3   Conclusions about the (extended) coupled dipole model

The coupled dipole model is a flexible framework that can be transformed to different levels of parametrisation by changing the design of the coupling matrices. This way, it

is applicable to many different situations. The extended coupled dipole model allows for less prior assumptions, though is easier to estimate because of its statistical simplicity. The combination of the ECDM and the MUSIC approach is promising and needs to be examined in further detail. Different applications of the (E)CDM lie in analysing space-time-condition data arrays or space-time-subject arrays. The latter can be helpful to compute grand averages. Finally, by clustering epileptic spikes as in chapter 2, another application lies in analysing space-time-cluster data arrays, thus contributing to one of the major clinical application areas of Magnetoencephalography.

Bibliography

[1] A. Achim, S. Bouchard: *Toward a dynamic topographic components model*, Electroenc. Clin. Neurophys. 1997, 103: 381-385

[2] J.M.P. Baas, J.L. Kenemans, G.R. Mangun: *Selective attention to spatial frequency: an ERP and source localization analysis*, Clin. Neurophys. 2002, 113: 1840-1854

[3] S. Baillet, J.C. Mosher, R.M. Leahy: *Electromagnetic Brain Mapping*, IEEE Signal Proc. Mag. 2001, 18(6): 14-30

[4] M.F. Bear, B.W. Conners, M.A. Paradiso: *Neuroscience, Exploring the brain*, Williams & Wilkins, Baltimore, 1996

[5] F. Bijma, J.C. de Munck, H.M. Huizenga, R.M. Heethaar: *A mathematical approach to the temporal stationarity of background noise in MEG/EEG measurements*, NeuroImage 2003, 20(1): 233-243

[6] F. Bijma, J.C. de Munck, K.B.E. Böcker, H.M. Huizenga, R.M. Heethaar: *The coupled dipole model: an integrated model for multiple MEG/EEG data sets*, NeuroImage 2004, 23(3): 890-904

[7] F. Bijma, J.C. de Munck, H.M. Huizenga, R.M. Heethaar, A. Nehorai: *Simultaneous estimation and testing of sources in multiple MEG data sets*, IEEE Trans. Signal Proc. Spec. Issue Brain Imag. 2005, *to appear*

[8] F. Bijma, J.C. de Munck, R.M. Heethaar: *The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker Products*, NeuroImage 2005, *in revision*

[9] M.E. Brandt: *Visual and auditory evoked phase resetting of the alpha EEG*, Int. Journ. Psychophys. 1997, 26: 285-298

[10] R. Bro: *PARAFAC. Turorial and applications*, Chemomet. Intelligent Lab. Sys. 1997, 38: 149-171

[11] K.L. Chung: *Elementary probability theory with stochastic processes*, 3d edition, Springer Verlag, New York, 1979

[12] D. Cohen: *Magnetoencephalography: detecting of the brain's electrical activity with a superconducting magnetometer*, Science 1972, 175: 664-666

[13] B.W. Connors, Y. Amitai: *Making waves in the neocortex*, Neuron 1997, 18: 347-349

[14] R. Coppola, R. Tabor, M.S. Buchsbaus: *Signal to noise ratio and response variability measurements in single trial evoked potentials*, Electroenc. Clin. Neurophys. 1978, 44: 214-222

[15] J.C. de Munck, B.W. van Dijk, H. Spekreijse: *Mathematical dipoles are adequate to describe realistic generators of human brain activity*, IEEE Trans. Biomed. Eng. 1988, 35(11): 960-966

[16] J.C. de Munck, P.C.M. Vijn, F.H. Lopes da Silva: *A random dipole model for spontaneous brain activity*, IEEE Trans. Biomed. Eng. 1992, 39(8): 791-804

[17] J.C. de Munck: *A linear discretization of the volume conductor boundary integral equation using analytically integrated elements*, IEEE Trans. Biomed. Eng. 1992, 39(9): 986-990

[18] J.C. de Munck, B.W. van Dijk: *The spatial distribution of spontaneous EEG and MEG*, in: C. Uhl (ed.): *Analysis of neurophysiological brain functioning*, Springer Verlag Telos, 1999, 202-228

[19] J.C. de Munck, J.P.A. Verbunt D. van 't Ent, B.W. van Dijk: *The use of an MEG device as 3-D digitizer and motion detection system*, Phys. Med. Biol. 2001, 46: 2041-2052

[20] J.C. de Munck, H.M. Huizenga, L.J. Waldorp, R.M. Heethaar: *Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise*, IEEE Trans. Signal Proc. 2002, 50(7): 1565-1572

[21] J.C. de Munck, F. Bijma, P. Gaura, C. Sieluzycki, M.I. Branco, R.M. Heethaar: *A maximum likelihood estimator for trial to trial variations in noisy MEG/EEG data sets*, IEEE Trans. Biomed. Eng. 2004, 51(12): 2123-2128

[22] F. Di Russo, A. Martinez, M.I. Sereno, S. Pitzalis, S.A. Hillyard: *Cortical sources of the early components of the visual evoked potential*, Human Brain Mapping 2002, 15(2): 95-111

[23] A. Dogandžiç, A. Nehorai: *Localization of evoked electric sources and design of EEG/MEG sensor arrays*, Proc. 9th IEEE SP Workshop Stat. Signal Array Proc., Portland, OR, Sept. 1998: 228-231

[24] A. Dogandžiç, A. Nehorai: *Estimated evoked dipole responses in unknown spatially correlated noise with EEG/MEG arrays*, IEEE Trans. Signal Proc. 2000, 48(1): 13-25

[25] A. Dogandžiç, A. Nehorai: *Generalized multivariate analysis of variance*, IEEE Signal Proc. Mag. 2003, 20(5): 39-54

[26] J.R. Duann, T.P. Jung, W.J. Kuo, T.C. Yeh, S. Makeig, J.C. Hsieh, T.J. Sejnowski: *Single-trial variability in event-related BOLD signals*, NeuroImage 2002, 15(4): 823-835

[27] A.S. Field, D. Graupe: *Topographic component (parallel factor) analysis of multichannel evoked potentials: practical issues in trilinear spatiotemporal decompostion*, Brain Top. 1991, 3: 407-423

[28] M. Fuchs, M. Wagner, J. Kastner: *Confidence limits of dipole source reconstruction results*, Clin. Neurophys. 2004, 115: 1442-1451

[29] T. Gasser, J. Möcks, R. Verleger: *SELAVCO: a method to deal with trial-to-trial variability of evoked potentials*, Electroenc. Clin. Neurophys. 1983, 55: 717-723

[30] D.B. Geselowitz: *On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources*, IEEE Trans. Magn. 1970, 6: 346-347

[31] G.H. Golub, C.F. Van Loan: *Matrix computations*, 2nd edition, John Hopkins Univ. Press, Baltimore, 1990

[32] S.I. Gonçalves, J.C. de Munck, J.P.A. Verbunt, F. Bijma, R.M. Heethaar, F.H. Lopes da Silva: *In vivo measurement of the brain and skull resistivities using an EIT-based method and realistic models for the head*, IEEE Trans. Biomed. Eng. 2003, 50(6): 754-766

[33] R.P.P.P. Grasman: *Sensor array signal processing and the neuro-electromagnetic inverse problem in functional connectivity analysis of the brain* Ph.D. thesis, University of Amsterdam, 2004

[34] M.D. Greicius, B. Krasnow, A.L. Reiss, V. Menon: *Functional connectivity in the resting brain: A network analysis of the default mode hypothesis*, Proc. Nat. Acad. Sc. 2003, 100(1): 253-258

[35] D. Gutiérrez, A. Nehoari, C.H. Muravchik: *Estimating brain conductivities and dipole source signals with EEG arrays*, IEEE Trans. Biomed. Eng. 2004, 51(12): 2113-2122

[36] M. Hämäläinen, J. Sarvas: *Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data*, IEEE Trans. Biomed. Eng. 1989, 36(2): 165-171

[37] M. Hämäläinen, R. Hari, R.J. Ilmoniemi, J. Knuutila, O.V. Lounasmaa: *Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain*, Rev. Mod. Phys. 1993, 65(2): 413-497

[38] R.A. Harshman: *Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis*, Univ. California, UCLA Working Papers in Phonetics 1970, 16: 1-84

[39] H.M. Huizenga, P.C.M. Molenaar: *Equivalent source estimation of scalp potential fields contaminated by heteroscedastic and correlated noise*, Brain Top. 1995, 8: 13-33

[40] H.M. Huizenga, T.L. van Zuijen, P.C.M. Molenaar: *Simultaneous MEG and EEG source analysis*, Psys. Med. Biol. 2001, 46(7): 1737-1751

[41] H.M. Huizenga, D.J. Heslenfeld, P.C.M. Molenaar: *Optimal measurement conditions for spatiotemporal EEG/MEG source analysis*, Psychometrika 2002, 67(2): 299-313

[42] H.M. Huizenga, J.C. de Munck, L.J. Waldorp, R.P.P.P. Grasman: *Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model*, IEEE Trans. Biomed. Eng. 2002, 49(6): 533-539

[43] J. Huttunen, R. Hari, L. Leinonen: *Cerebral magnetic responses to stimulation of ulnar and median nerves*, Electroenc. Clin. Neurophys. 1987, 66: 391-400

[44] P. Jaśkowski, R. Verleger: *Amplitude and latencies of single-trial ERP's estimated by a maximum-likelihood method*, IEEE Trans. Biomed. Eng. 1999, 46: 987-993

[45] K. Jerbi, S. Baillet, J.C. Mosher, G. Nolte, L. Garnero, R.M. Leahy: *Localization of realistic cortical activity in MEG using current multipoles*, NeuroImage 2004, 22(2): 779-793

[46] S.M. Kay: *Fundamentals of statistical signal processing. Volume I: estimation theory*, Englewood Cliffs, NJ: Prentice-Hall, 1993

[47] J.L. Kenemans, A. Kok, F.T.Y. Smulders: *Event-related potentials to conjunctions of spatial-frequency and orientation as a function of stimulus parameters and repsonse requirements*, Electroenc. Clin. Neurophys. 1993, 88: 51-63

[48] J.L. Kenemans, J.M.P. Baas, G.R. Mangun, M. Lijffijt, M.N. Verbaten: *On the processing of spatial frequencies as revealed by evoked potential source modeling*, Clin. Neurophys. 2000, 111(6): 1113-1123

[49] C.G. Khatri, *A note on a MANOVA model applied to problems in growth curve*, Ann. Instit. Statist. Math. 1966, 18: 75-86

[50] W.E. Kincses, C. Braun, S. Kaiser, T. Elbert: *Modeling extended sources of event-related potentials using anatomical and physiological constraints*, Human Brain Mapping 1999, 8: 182-193

[51] W. Klimesch: *EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis*, Brain Research Rev. 1999, 29: 169-195

[52] W. Klimesch, B. Schack, M. Schabus, M. Doppelmayr, W. Gruber, P. Sauseng: *Phase-locked alpha and theta oscillations generate the P1-N1 complex and are related to memory performance*, Cogn. Brain Research 2004, 19(3): 302-316

[53] T.R. Knösche, E.M. Berends, H.R.A. Jagers, M.J. Peters: *Determining the number of independent sources of the EEG: a simulation study on information criteria*, Brain Top. 1998, 11(2): 111-124

[54] V. Kolev, J. Yordanova, M. Schürmann, E. Başar: *Event-related alpha oscillations in task processing*, Clin. Neurophys. 1997, 110: 1784-1792

[55] J.B. Kruskal: *Rank, decomposition and uniqueness for 3-way and N-way arrays*, in: R. Coppi, S. Bolasco (eds.): *Multiway data analyses*, North Holland, New York, 1989

[56] A.N. Langville, W.J. Stewart: *The Kronecker product and stochastic automata networks*, Journ. Comp. Appl. Math. 2004, 167: 429-447

[57] N.A. Laskaris, L.C. Liu, A.A. Ionnides: *Single-trial variability in early visual neuro-magnetic responses: an explorative study based on the regional activation contributing to the N70m peak*, NeuroImage 2003, 20(2): 765-783

[58] H. Laufs, K. Krakow, P. Sterzer, E. Eger, A. Beyerle, A. Salek-Haddadi, A. Kleinschmidt: *Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest*, Proc. Nat. Acad. Sc. 2003, 100(19): 11053-11058

[59] F.H. Lopes da Silva: *Neural mechanisms underlying brain waves: from neural membranes to networks*, Electroenc. Clin. Neurophys. 1991, 79: 81-93

[60] F.H. Lopes da Silva, A. van Rotterdam: *Biophysical aspects of EEG and Magnetoencephalogram generation*, in: F.H. Lopes da Silva, E. Niedermeyer (eds.): *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* 4th edition, Williams & Wilkins, Baltimore, 1999

[61] B. Lütkenhöner: *Dipole source localization by means of maximum likelihood estimation. I. Theory and simulations*, Electroenceph. Clin. Neurophys. 1998, 106: 314-321

[62] B. Lütkenhöner: *Dipole source localization by means of maximum likelihood estimation. II. Experimental evaluation*, Electroenceph. Clin. Neurophys. 1998, 106: 322-329

[63] J.R. Magnus, H. Neudecker: *Matrix differential calculus with applications in statistics and Econometrics*, revised edition, Wiley, New York, 1995

[64] J. Maier, G. Dagnelie, H. Spekreijse, B.W. van Dijk: *Principal Components Analysis for source localization of VEPs in man*, Vision Res. 1987, 27(2): 165-177

[65] S. Makeig, M. Westerfield, T.P. Jung, S. Enghoff, J. Townsend, E. Courchesne, T.J. Sejnowski: *Dynamic Brain Sources of Visual Evoked Responses*, Science 2002, 295: 690-694

[66] V. Mäkinen, H. Tiitinen, P. May: *Auditory event-related responses are generated independently of ongoing brain activity*, NeuroImage 2005, 24(4): 961-968

[67] E. Maris: *A resampling method for estimating the signal subspace of spatio-temporal EEG/MEG data*, IEEE Trans. Biomed. Eng. 2003, 50(8): 935-949

[68] J.E. Marsden, A.J. Tromba: *Vector Calculus*, Freeman And Company, New York, 1988

[69] F. Mauguière, I. Merlet, N. Forss, S. Vanni, V. Jousmäki, P. Adeleine, R. Hari: *Activation of a distributed somatosensory cortical network in the human brain. A dipole modelling study of magnetic fields evoked by median nerve stimulation. Part I: location and activation timing of SEF sources*, Electroenc. Clin. Neurophys. 1997, 104: 281-289

[70] C.D. McGillem, J.I. Aunon: *Analysis of event-related potentials*, in: A.S. Gevins, A. Remond (eds.): *Methods of analysis of brain electrical and magnetic signals, EEG Handbook*, Elsevier, Amsterdam, 1987, Chapter 5

[71] F. Miwakeichi, E. Martínez-Montes, P.A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, Y. Yamaguchi: *Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis*, NeuroImage 2004, 22(3): 1035-1045

[72] J. Möcks, T. Gasser, P.D. Puan, W. Köhler: *Trial-to-trial variability of single potentials: methodological concepts and results*, Int. Journ. NeuroScience 1987, 33: 25-32

[73] J. Möcks: *Topographic components model for event related potentials and some biophysical considerations*, IEEE Trans. Biomed. Eng. 1988, 35(6): 482-484

[74] J.C. Mosher, P.S. Lewis, R.M. Leahy: *Multiple dipole modeling and localization from spatio-temporal MEG data*, IEEE Trans. Biomed. Eng. 1992, 39(6): 541-557

[75] J.C. Mosher, M.E. Spencer, R.M. Leahy, P.S. Lewis: *Error bounds for EEG and MEG dipole source localization*, Electroenc. Clin. Neurophys. 1993, 86: 303-321

[76] C.H. Muravchik, A. Nehorai: *EEG/MEG error bounds for a static dipole source with a realistic head model*, IEEE Trans. Signal Proc. 2001, 49(3): 470-484

[77] G. Nolte, G. Curio: *On the calculation of magnetic fields based on multipole modeling of focal biological current sources*, Biophys. Journ. 1997, 73(3): 1253-1262

[78] G. Nolte: *The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors*, Phys. Med. Biol. 2003, 48: 3637-3652

[79] H. Obrig, H. Israel, M. Kohl-Bareis, K. Uludag, R. Wenzel, B. Müller, G. Arnold, A. Villringer, *Habituation of the visually evoked potential and its vascular response: implications for neurovascular coupling in the healthy adult*, NeuroImage 2002, 17: 1-18

[80] T.F. Oostendorp, J. Delbeke, D.F. Stegeman: *The conductivity of the human skull: results of in vivo and in vitro measurements*, IEEE Trans. Biomed. Eng. 2000, 47(11): 1487-1492

[81] P. Ossenblok, D. Reits, H. Speckreijse: *Check size dependency of the sources of the hemifield-onset evoked-potential*, Doc. Ophthalmologica 1994, 88(1): 77-88

[82] W.D. Penny, J. Stephan, J. Kiebel, J.M. Kilner, M.D. Rugg: *Event related brain dynamics*, Trends Neurosc. 2002, 25: 387-390

[83] B. Pfeiderer, J. Ostermann, N. Michael, W. Heindel: *Visualisation of auditory habituation by fMRI*, NeuroImage 2002, 17: 1705-1710

[84] G. Pfurtscheller, F.H. Lopes da Silva: *Event-related EEG/MEG synchronization and desynchronization: basic principles*, Clin. Neurophys. 1999, 110: 1842-1857

[85] D.T. Pham, J. Möcks, W. Köhler, T. Gasser: *Variable latencies of noisy signals: Estimation and testing in brain potential data*, Biometrika 1987, 74: 525-533

[86] D. Purves, G.J. Augustine, D. Fitzpatrick, L.C. Katz, A.S. Lamantia, J.O. Mc-Namara, S.M. Williams: *NeuroScience*, 2nd edition, Sinauer Associates, Sunderland, 2001

[87] M.E. Raichle, A.M. MacLeod, A.Z. Snyder, W.J. Powers, D.A. Gusnard, G.L. Shulman: *A default mode of brain function*, Proc. Nat. Acad. Sc. 2001, 98(2): 676-682

[88] C.R. Rao: *Linear Statistical Inference and its Applications*, 2nd edition, Wiley, New York, 1973

[89] D.S. Rizzuto, J.R. Madsen, E.B. Bromfield, A. Schulze-Bonhage, D. Seelig, R. Aschenbrenner-Scheibe, M.J. Kahana: *Reset of human neocortical oscillations during a working memory task*, Proc. Nat. Acad. Sc. 2003, 100(13): 7931-7936

[90] S.E. Robinson, J. Vrba: *Functional neuroimaging by synthetic aperture magnetometry*, in: T. Yoshimoto, M. Kotani, S. Kuriki, H. Karibe, N. Nakasato (eds.): *Recent Advances in Biomagnetism* Tohoku Univ. Press, Sendai, 1999, 302-305

[91] C.D. Saron, R.J. Davidson: *Visual evoked potential measures of interhemispheric transfer time in humans*, Behav. Neuroscience 1989, 103(5): 1115-1138

[92] J. Sarvas: *Basic mathematical and electromagnetic concepts of the biomagnetic invere problem*, Psys. Med. Biol. 1987, 32: 11-22

[93] M. Scherg, D. Von Cramon: *Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model*, Electroenc. Clin. Neurophys. 1985, 62: 32-44

[94] R.O. Schmidt: *Multiple emitter location and signal parameter estimation*, IEEE Trans. Antennas Prop. 1986, 34(3): 276-280

[95] G.A.F. Seber, C.J. Wild, *Nonlinear regression*, Wiley, New York, 1989

[96] K. Sekihara, F. Takeuchi, S. Kuriki, H. Koizumi: *Reduction of brain noise influence in evoked neuromagnetic source localization using noise spatial correlation*, Phys. Med. Biol. 1994, 39: 937-946

[97] M.S. Srivastava, C.G. Khatri, *An introduction to multivariate statistics*, North Holland, New York, 1979

[98] J. Steger, K. Imhof, J. Denoth, R.D. Pascual-Marqui, H.C. Steinhausen, D. Brandeis: *Brain mapping of bilateral visual interactions in children*, Psychophys. 2001, 38(2): 243-253

[99] D. Stirzaker: *Elementary probability*, Cambridge University Press, Cambridge, 1994

[100] P. Stoica, B.C. Ng: *On the Cramèr-Rao bound under parametric constraints*, IEEE Signal Proc. Letters 1998, 5(7): 177-179

[101] W.A. Truccolo, M. Ding, K.H. Knuth, R. Nakamura, L. Bressler: *Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity*, Clinical Neurophys. 2002, 113: 206-226

[102] B. Turetsky, J. Raz, G. Fein: *Representation of multi-channel evoked potential data using a dipole component model of intracranial generators: application to the auditory P300*, Electroenc. Clin. Neurophys. 1990, 76: 540-556

[103] K. Uutela, M. Hämäläinen, R. Salmelin: *Global optimization in the localization of neuromagnetc sources*, IEEE Trans. Biomed. Eng. 1998, 45(6): 716-723

[104] C.F. Van Loan: *The ubiquitous Kronecker product*, Journ. Comp. Appl. Math. 2000, 123: 85-100

[105] D. van 't Ent, J.C. de Munck, A.L. Kaas: *A fast method to derive realistic BEM models for E/MEG source reconstruction*, IEEE Trans. Biomed. Eng. 2001, 48(12): 1434-1443

[106] D. van 't Ent, I. Manshanden, P. Ossenblok, J.P.A. Verbunt, J.C. de Munck, D.N. Velis, F.H. Lopes da Silva: *Spike cluster analysis in neocortical localization related epilepsy achieves clinically significant source localization results in MEG*, Clin. Neurophys. 2003, 114(10): 1948-1962

[107] B.D. Van Veen, W. van Drongelen, M. Yuchtman, A. Suzuki: *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering*, IEEE Trans. Biomed. Eng. 2997, 44(9): 867-880

[108] E.F. Vonesh, V.M. Chinchilli, *Linear and nonlinear models for the analysis of repeated measurments*, Marcel Dekker, New York, 1997

[109] L.J. Waldorp, H.M. Huizenga, C.V. Dolan, P.C.M. Molenaar: *Estimated generalized least squares electromagnetic source analysis based on a parametric noise covariance model*, IEEE Trans. Biomed. Eng. 2001, 48(6): 737-741

[110] L.J. Waldorp, H.M. Huizenga, R.P.P.P Grasman, K.B.E. Böcker, J.C. de Munck, P.C.M. Molenaar: *Model selection in electromagnetic source analysis with an application to VEFs*, IEEE Trans. Biomed. Eng. 2002, 49(10): 1121-1129

[111] L.J. Waldorp: *Statistical model identification in electromagnetic source analysis*, Ph.D. thesis, University of Amsterdam, 2004

[112] K.M. Wang, H. Begleiter, B. Porjesz: *Trilinear modeling of event-related potentials*, Brain Top. 2000, 12(4): 263-271

[113] D.G. Wastell, D. Kleinman: *Potentiation of the habituation of human brain potentials* Biol. Psych. 1980, 10: 21-29

[114] I.S. Yetik, A. Nehorai, C. Muravchik, J. Haueisen: *Line-source modeling and estimation with magnetoencephalography*, IEEE Trans. Biomed. Eng. 2005, *to appear*

Samenvatting

Dit hoofdstuk bevat een gepopulariseerde samenvatting van dit proefschrift, getiteld Mathematische modellering van magneto-encefalografische data. *Een wetenschappelijke samenvatting is te vinden in de paragrafen 7.1 en 7.3.*

Magneto-encefalografie (MEG) is een techniek waarmee zeer zwakke magnetische velden worden gemeten aan de buitenkant van het hoofd. Deze magnetische velden worden veroorzaakt door neurale activiteit in het menselijk brein. De MEG-techniek is nauw verwant aan elektro-encefalografie (EEG) waarmee de electrische potentiaal op de hoofdhuid wordt gemeten. Beide technieken zijn heel gevoelig, dat wil zeggen, dat de gevolgen van kleine (neurale) stroompjes in de hersenen te volgen zijn.

Het doel van een MEG-meting is om te bepalen wat zich binnen in het hoofd afspeelt. In veel gevallen gaat het hierbij om het zogenaamde *lokaliseren van de bronnen* die ten grondslag liggen aan de gemeten hersenactiviteit. Voorbeelden zijn experimenten waarin functionele hersengebieden worden gelokaliseerd, zoals het visuele en het auditieve gebied. Deze functionele gebieden worden in dit soort experimenten geactiveerd door middel van een *stimulus*, zoals een plaatje of een pieptoontje. Een ander voorbeeld van bronlokalisatie is het bepalen van de lokatie van epileptogeen hersenweefsel. In dat geval worden MEG-data van een epilepsiepatiënt gemeten zonder een stimulus te presenteren. Focale epilepsie uit zich in kleine piekjes in de gemeten signalen. Door de onderliggende bron van deze gemeten piekjes te lokaliseren kan de plaats van het zieke, epileptogene hersenweefsel worden bepaald.

MEG-data worden gemeten in *ruimte* en in *tijd*. In ruimte worden MEG-signalen gemeten op 151 sensoren die zich bevinden in een helm, die over het hoofd geplaatst wordt (zie figuur 1.1 op pagina 2). In tijd worden deze signalen gemeten op achtereenvolgende tijdstippen, bijvoorbeeld duizend keer per seconde. In een meting waarin een functioneel gebied gelokaliseerd moet worden, wordt bovendien de stimulus herhaaldelijk gepresenteerd, bijvoorbeeld driehonderd keer. De reactie van de hersenen op de stimulus wordt dan driehonderd keer gemeten in ruimte en tijd. In geval van epilepsie wordt het MEG gemeten over een tijdsperiode van bijvoorbeeld een uur, waarin dan meerdere pieken optreden.

Om vervolgens op basis van de gemeten signalen de lokatie van de onderliggende bronnen te bepalen, wordt een aantal aannames gedaan. Zo wordt er een bronmodel en een hoofdmodel aangenomen. Het meest gebruikelijke bronmodel is de zogenaamde *dipoolbron*. Een dipoolbron is een bron in een punt met een bepaalde richting. Als hoofdmodel wordt gemakshalve voor MEG vaak een bolvorm aangenomen. Als alternatief kan men ook een meer realistisch model gebruiken, wat gebaseerd is op MRI beelden

van de hersenen (bijvoorbeeld figuur 1.5 op pagina 6). Deze verbetering in hoofdmodel is vooral voor EEG belangrijk; MEG is minder gevoelig voor (beperkte) onjuistheden in het hoofdmodel. Naast deze twee modellen, moet ook het aantal bronnen worden vastgesteld of aangenomen. Hoe meer bronnen, des te geavanceerder het model en des te beter het model de gemeten MEG-data kan beschrijven. De beperking hierbij is echter dat de lokalisatie moeilijker wordt naarmate er meer bronnen gelokaliseerd moeten worden. Veel bronnen maken de lokalisatiemethode instabiel en onbetrouwbaar. Daarom moet men altijd de balans zoeken tussen enerzijds genoeg bronnen om de gemeten signalen goed te kunnen verklaren en anderzijds niet te veel bronnen om de stabiliteit en oplosbaarheid te behouden.

De hersenen zijn voortdurend actief met het besturen van allerlei processen in het lichaam, ook wanneer een persoon (lichamelijk) in rust is. Deze activiteit kan worden beschouwd als achtergrondactiviteit. Wanneer een stimulus wordt gepresenteerd zullen de hersenen een reactie vertonen, de *respons*. Echter, naast deze opgewekte respons is er ook de genoemde achtergrondactiviteit. In experimenten waarin gebruik gemaakt wordt van een stimulus om een bepaald hersengebied te activeren, worden normaal gesproken de herhaalde metingen gemiddeld. Dat wil zeggen, van de driehonderd metingen wordt het gemiddelde berekend en dat gemiddelde wordt beschouwd als *de* respons van de hersenen op de stimulus. Het idee achter deze middeling is dat de respons van de hersenen op de stimulus elke keer hetzelfde is, terwijl de achtergrondactiviteit steeds wisselt. De achtergrondactiviteit is dan weer positief en dan weer negatief en verdwijnt zodoende wanneer het gemiddelde genomen wordt. Het gemiddelde signaal is dan de 'schone' opgewekte respons. Deze gedachtengang wordt het *Signaal Plus Ruis-model* genoemd: iedere meting is de som van een vast deel, het *signaal* ofwel de opgewekte respons, en een wisselend deel, de *ruis* ofwel de achtergrondactiviteit.

In hoofdstukken 2, 3 en 4 wordt de geldigheid van het Signaal Plus Ruis-model onderzocht. Het model is onjuist als zou blijken dat de achtergrondactiviteit verandert in reactie op de presentatie van de stimulus. Om dit te onderzoeken worden de zogenaamde *residuen* onderzocht. Het residu van elk van de driehonderd metingen is het verschil tussen het gemeten signaal en het gemiddelde signaal. Volgens het Signaal Plus Ruis-model is het residu bij benadering gelijk aan de achtergrondactiviteit en verandert niet wezenlijk van karakter in de tijd. Wanneer er sprake is van werkelijke invloed van de stimulus op de achtergrondruis, is dit zichtbaar in de residuen. Bijvoorbeeld wanneer de achtergrondactiviteit zich door de stimulus aanpast en steeds een specifiek patroon vertoont direct nadat de stimulus is gepresenteerd, dan zal dit specifieke patroon in het gemiddelde signaal zichtbaar worden. Dit patroon is dan niet aanwezig in de residuen omdat het in het gemiddelde zit. Wanneer men in dit geval het verloop van de residuen in de tijd bekijkt, is te zien dat voorafgaand aan de stimulus er activiteit te zien is die op het moment dat de stimulus wordt gepresenteerd opeens verdwijnt omdat de achtergrondactiviteit vanaf dat moment in het gemiddelde zit. Dit voorbeeld geeft aan dat de residuen informatie bevatten over de geldigheid van het Signaal Plus Ruis-model. Daarom worden deze residuen onderzocht in de genoemde hoofdstukken.

De verandering in tijd van de residuen wordt onderzocht door middel van de *tijdcovariantie*. Aan de tijd-covariantie is af te lezen of de residuen wezenlijk veranderen in tijd. Indien dit het geval is, is de basisgedachte van het Signaal Plus Ruis-model onjuist.

Uit het onderzoek in de hoofdstukken 2, 3 en 4 blijkt dat in geval van 'simpele' stimuli zoals een motorische prikkel of een pieptoontje, er geen aanleiding lijkt te zijn om het Signaal Plus Ruis-model niet te gebruiken. Wanneer men een meer complexe stimulus gebruikt, bijvoorbeeld meerdere stimuli waar verschillende aandacht aan besteed moet worden, lijkt het beter om het Signaal Plus Ruis-model uit te breiden tot een model dat de individuele responsies op de stimulus nader beschrijft. Mogelijke uitbreidingen van het vaste responsmodel zijn modellen waarbij iedere respons op de stimulus een andere amplitude (grootte) of latentie (reactietijd) vertoont. Amplitudevariatie is relatief eenvoudig te bepalen in de analyse van de gemeten signalen, terwijl variatie in reactietijd veel lastiger te bepalen is.

Naast het valideren van het Signaal Plus Ruis-model bevat de tijd-covariantie ook belangrijke informatie die gebruikt kan worden in de bronlokalisatie. De tijd-covariantie is in feite een groot getallenblok, een *matrix* genaamd, die van ieder tweetal tijdstippen A en B aangeeft hoe de gemeten signalen op die tijdstippen met elkaar samenhangen. Zo zullen de metingen op twee tijdstippen vlak na elkaar (bijvoorbeeld slechts vijf milliseconds na elkaar) min of meer hetzelfde signaal bevatten, terwijl de metingen op twee tijdstippen die verder van elkaar verwijderd zijn minder op elkaar zullen lijken. Het analogon in de ruimte is gegeven door de *ruimte-covariantie*. In de ruimte-matrix staat informatie over hoeveel de gemeten signalen op ieder tweetal sensoren in de MEG-helm op elkaar lijken. Net als in de tijd geldt ook hier: hoe dichter de sensoren bij elkaar liggen, hoe meer de gemeten signalen op elkaar lijken. Dit komt omdat sensoren die dicht bij elkaar zitten vooral activiteit van hetzelfde hersengedeelte zullen meten, terwijl sensoren die op heel verschillende plekken zitten meer activiteit van verschillende gedeeltes van de hersenen meten.

In feite hebben we te maken met een *ruimte-tijd-matrix*. Echter dat getallenblok is te groot om mee te rekenen. Daarom wordt deze matrix opgesplitst in een ruimte-matrix en een tijd-matrix. Deze splitsing is gebaseerd op de vereenvoudigende aanname dat de ruimtelijke verhoudingen niet veranderen in de tijd en dat de tijdsverhoudingen gelijk zijn voor alle sensoren in de MEG-helm. Hoewel de werkelijkheid complexer is, maakt deze aanname het rekenwerk een stuk gemakkelijker. In hoofdstuk 4 wordt nagegaan hoe goed deze vereenvoudigende splitsing de werkelijkheid beschrijft. Het blijkt dat voor bronlokalisatie deze splitsing goed genoeg werkt. Indien men echter geïnteresseerd is in de fysiologische eigenschappen van de achtergrondactiviteit is de hoeveelheid informatie in de aparte ruimte- en tijd-matrices slechts beperkt.

In de hoofdstukken 5 en 6 wordt onderzocht of het mogelijk is de bronnen in meerdere gemeten MEG-datasets tegelijk te lokaliseren. Deze simultane aanpak is met name interessant wanneer in verschillende datasets dezelfde bronnen geactiveerd zijn. Een voorbeeld is een tweetal experimenten waarin het visuele hersengebied op verschillende manieren wordt geactiveerd. Het lokaliseren van het visuele gebied wordt dan nauwkeuriger wanneer men beide datasets tegelijk analyseert, in plaats van een voor een. De achterliggende reden is dat meer gegevens (meer data) voor het vinden van een en dezelfde bron altijd een gunstig effect heeft op de nauwkeurigheid van de te bepalen positie.

Uit de studies in de genoemde hoofdstukken blijkt dat de onnauwkeurigheid in de geschatte positie in bepaalde gevallen wel met een factor 10 kan worden verkleind wan-

neer men vijf datasets aan een simultane analyse onderwerpt in plaats van vijf losse analyses. De gepresenteerde simultane analysemethode is heel flexibel, zodat de gebruiker gemakkelijk meer of minder aannames kan maken omtrent de gelijkheid van de bronnen in de verschillende datasets. Omdat deze aannames afhankelijk zijn van de informatie en keuzes van de gebruiker, maakt deze flexibiliteit het model tegelijkertijd subjectief. In hoofdstuk 6 wordt de simultane lokalisatiemethode daarom aangepast, zodat het niet mogelijk is heel specifieke (subjectieve) informatie in het model te stoppen. De gebruiker kan vervolgens kiezen of hij het meer objectieve model dan wel het model dat meer informatie toelaat, kiest.

Andere interessante toepassingen van deze simultane analysemethode zijn datasets van gelijke experimenten bij verschillende proefpersonen en datasets van epilepsiepatiënten. De toepassing op data van verschillende proefpersonen is interessant om meer algemeen het karakter en de lokatie van functionele hersengebieden te bepalen. De simultane analysemethode verbetert de nauwkeurigheid van de lokatie van epileptogeen hersenweefsel wanneer verschillende soorten epileptische pieken uit hetzelfde hersengebied aan deze nieuwe methode worden onderworpen.

Gearfetting

Yn dit haadstik wurdt dit proefskrift, mei de titel Matematyske modellearring fan magneto-ensefalografyske data, *gearfette op in wize dy't foar de bûtensteander begryplik is. In wittenskiplike gearfetting is te finen yn de paragrafen 7.1 en 7.3.*

Mei de magneto-ensefalografy (MEG) technyk kinne tige lytse magnetyske fjilden om 'e holle metten wurde. Dy magnetyske fjilden binne it gefolch fan neurale aktiviteit yn de minsklike harsens. MEG is nau besibbe mei elektro-ensefalografy, dêr't de elektryske potinsjaal op 'e hollehûd mei metten wurdt. Beide techniken binne wakker gefoelich, sadat tige lytse wearden en ferskillen waarnommen wurde kinne.

It doel fan in MEG-mjitting is fêst te stellen wat der yn 'e holle bart. Faak betsjut dat dat men de saneamde boarnen dy't de grûn foarmje foar de metten sinjalen, lokalisearje wol, bygelyks yn eksperiminten dêr't in funksjoneel harsensgebied yn lokalisearre wurdt. Yn dat soarte fan eksperiminten wurdt in gebied yn 'e harsens oantrune troch in oantruner fan bûtenôf, bygelyks in printsje foar it fisuéle gebied of in lûdsje foar it auditive gebied. Dizze oantruner soarget derfoar dat it gebied dêr't it om giet, aktyf wurdt. In oar foarbyld fan saneamde boarnelokalisaasje is it fêststellen fan de lokaasje fan in epileptogeen harsengebied. In dat twadde gefal wurdt it MEG metten fan in epilepsypasjint sûnder in oantruner te brûken. Fokale epilepsy is te werkennen oan lytse pykjes yn de MEG-sinjalen. Troch de boarne fan dizze pykjes te finen, kin de lokaasje fan it sike, epileptogene gebied fêststeld wurde.

MEG-sinjalen wurde metten yn tiid en romte. Yn romte betsjut dat dat de sinjalen op 151 sensoaren metten wurde, dy't harren plak ha yn de MEG-helm dy't om de holle komt by in MEG-mjitting (figuer 1.1 op side 2). Yn tiid wurdt it MEG metten op ferskillende tiidstippen fuort efterinoar, bygelyks tûzen kear yn 'e sekonde. Yn in mjitting dêr't in funksjoneel harsengebied mei lokalisearre wurde moat, wurdt fierder de oantruner ek nochris werhelle, bygelyks trijehûndert kear. De reaksje fan de harsens op de oantruner wurdt dan trijehûndert kear metten yn tiid en romte. By in epilepsymjitting wurdt net in oantruner fan bûtenôf brûkt, mar wurdt it MEG bygelyks ien oere lang metten. Yn sa'n oere ferskynt dan út en troch in pykje.

Foar it finen fan de lokaasjes fan de ûnderlizzende boarnen fan de metten sinjalen moat in tal fan oannames dien wurde. Sa moatte in model foar de boarne en in model foar de foarm fan de holle oannommen wurde. In gongber model foar de boarne is de saneamde dipoalboarne. In dipoalboarne is in boarne yn ien punt mei in beskate rjochting. As model foar de foarm fan 'e holle wurdt meastentiids in bolfoarm brûkt. Men kin as alternatyf ek in mear realistysk model brûke, dat makke wurde kin mei MR-bylden fan 'e harsens (figuer 1.5 op side 6). Sa'n krekter model is foar EEG wichtiger

as foar MEG; MEG is net tige gefoelich foar flaters yn it hollemodel. Njonken dizze twa modellen moat ek it tal fan boarnen fêststeld of oannommen wurde. Hoe mear boarnen, hoe moaier it model en hoe better it model de metten sinjalen beskriuwt. De begrinzing hjirby is lykwols dat de lokalisaasje fan de boarnen dreger wurdt al nei't der mear boarnen yn it model sitte. In soad boarnen soargje foar in ûnbetroubere en ynstabile lokalisaasjemetoade. Dêrom moat men altyd it lykwicht sykje tusken oan 'e iene kant genôch boarnen om de metten sinjalen goed ferklearje te kinnen, en oan 'e oare kant net te folle boarnen, sadat it stabyl en oplosber bliuwt.

De minsklike harsens binne oanientriedwei troch yn aksje. Se binne dwaande mei it oanstjoeren fan allegear prosessen yn it lichem, ek as de persoan fierder neat docht. Wannear't de oantruner oanbean wurdt, jouwe de harsens dêr in reaksje op, de saneamde *respûns*. Njonken dizze oanfitere respûns is der lykwols ek dy eftergrûnaktiviteit dy't oanientriedwei trochgiet. Yn eksperiminten dêr't in oantruner fan bûtenôf brûkt wurdt om in beskaat harsengebiet oan te fiterjen, wurde de werhelle mjittings meastentiids middele. Dat betsjut dat fan de trijehûndert mjittings it gemiddelde nommen wurdt en dat gemiddelde wurdt dan besjoen as de respûns op de oantruner. De gedachte efter dizze middeling is dat de reaksje fan de harsens op de oantruner hieltyd itselde is, wylst de eftergrûnaktiviteit hieltyd oars is. De eftergrûnaktiviteit is wikseljend posityf en negatyf en sadwaande is it gemiddelde dêrfan sawat nul. Dêrom sit yn it gemiddelde fan de trijehûndert mjittings allinnich noch de 'skjinne' oanfitere respûns fan 'e harsens. Dizze tinkwize wurdt ek wol it Sinjaal Plus Rûs-model neamd: eltse mjitting bestiet út in fêst part, it 'sinjaal' of de oanfitere respûns, en in wikseljend part, de 'rûs' of de eftergrûnaktiviteit.

Yn 'e haadstikken 2, 3 en 4 wurdt neigien oft it Sinjaal Plus Rûs-model jildich is. It model is ferkeard as bliken docht dat de eftergrûnaktiviteit feroaret nei de presintaasje fan 'e oantruner. Om dat út te finen wurde de saneamde residuën besjoen. It residu fan eltse mjitting fan 'e trijehûndert mjittings is it ferskil tusken dy mjitting en it gemiddelde sinjaal. Neffens it Sinjaal Plus Rûs-model is dat residu sawat gelyk oan 'e eftergrûnaktiviteit en feroaret it net wêzentlik yn 'e tiid. As de oantruner wier fan ynfloed is op de eftergrûnaktiviteit, dan is dat te sjen oan de residuën. As bygelyks troch de oantruner de eftergrûnaktiviteit him oanpast en nei de presintaasje fan 'e oantruner hieltyd in spesifyk patroan sjen lit, dan is krekt dat patroan te sjen yn it gemiddelde. Tagelyk is dat patroan dan net mear te finen yn de residuën, om't it yn it ôflutsen gemiddelde sit. As men yn sa'n gefal it ferrin fan 'e residuën besjocht, dan falt op dat der foardat de oantruner presintearre waard, aktiviteit wie, dy't nei de presintaasje ferdwûn is om't it fanôf dat stuit yn it gemiddelde bedarre is. Ut dit foarbyld docht wol bliken dat de residuën ynformaasje befetsje oer de jildichheid fan it Sinjaal Plus Rûs-model. Dêrom wurde yn de hjirboppe neamde haadstikken de residuën ûndersocht.

De feroaring fan 'e residuën yn 'e tiid wurdt besjoen oan 'e hân fan de tiid-kovariânsje. Oan de tiid-kovariânsje kin men sjen oft de residuën werklik yn 'e tiid feroarje. As dat wier sa is, is de oanname fan it Sinjaal Plus Rûs-model ûnterjochte. It docht bliken yn 'e haadstikken 2, 3 en 4 dat yn it gefal fan 'ienfâldige' oantruners sa as in motoryske prikkel of in lûdsje, der gjin oanlieding is om it Sinjaal Plus Rûs-model ôf te wizen. As der wat fernimstiger oantruners brûkt wurde, liket it better om it Sinjaal Plus Rûs-model út te wreidzjen sadat it model de yndividuéle reaksjes fan 'e harsens op

'e oantruner krekter beskriuwt. Mooglike útwreidings fan it fêste respûnsmodel binne modellen dêr't de reaksjes fan 'e harsens ferskillende amplitudes (gruttes) of ferskillende latinsjes (reaksjetiden) yn hawwe. Fariaasje yn amplitude kin frij maklik fêststeld wurde út de metten sinjalen, wylst fariaasje yn reaksjetiid tige lestich ynskat wurde kin.

Njonken it neigean oft it Sinjaal Plus Rûs-model goed is, befettet de tiid-kovariânsje ek belangrike ynformaasje foar it boarnelokalisearjen. De tiid-kovariânsje is yn wêzen in grut blok mei sifers, in saneamde matriks, dy't fan elts twatal tiidstippen oanjout hokfoar ferhâlding de mjittings op dy beide tiidstippen ha. Dizze sifers jouwe oan oft de mjittings op inoar likenje of net. Op twa tiidstippen flak efter inoar, bygelyks mei fiif millisekondes dertusken, sille de mjittings bot op inoar likenje, wylst mjittings fierder útinoar yn it algemien mear ferskillende aktiviteit sjen litte sille. Yn romte treffe wy it allyksa oan: de romte-kovariânsje. Yn de romte-matriks steane sifers dy't oanjouwe op hokfoar wize de mjittings op twa sensoaren yn de MEG-helm harren ferhâlde. Krekt as yn de tiid jildt hjir ek: hoe tichter de sensoaren byinoar sitte, hoe mear't de mjittings op inoar likenje. Dat komt om't sensoaren dy't ticht byinoar lizze aktiviteit út likernôch itselde harsengebied mjitte, wylst sensoaren fierder útinoar aktiviteit út ferskillende parten fan 'e harsens mjitte.

Ut soarte ha wy te krijen mei in romte-tiid-matriks. Dat grutte siferblok is lykwols te grut om mei te rekkenjen. Dêrom wurdt dat blok opdield yn in romte- en in tiid-matriks. Dizze opdieling is basearre op de ferienfâldigjende oanname dat de romtlike ferhâldingen net feroarje yn 'e tiid en dat de ferhâldingen yn tiid gelyk binne op alle sensoaren yn 'e MEG-helm. Hoewol't it lân der yn werklikheid wat komplekser hinne leit, makket dizze tinkwize it rekkenwurk in stik ienfâldiger. Yn haadstik 4 wurdt neigien foar hokfoar part dizze ferienfâldigjende opdieling de werklikheid feitlik beskriuwt. It docht bliken dat dizze opdieling presys genôch is, as boarnelokalisaasje it doel is. As men fan doel is om de aard fan 'e eftergrûnaktiviteit te ûndersykjen, is de ynformaasje yn dizze opdieling lykwols oan 'e krappe kant.

Haadstikken 5 en 6 befetsje in petear oer de mooglikheden fan it tagelyk lokalisearjen fan de boarnen yn in tal fan ferskillende MEG-mjittings. Dizze gelyktidige oanpak is benammen nijsgjirrich wannear't yn guon mjittings deselde boarnen aktyf binne. In foarbyld is in twatal eksperiminten dêr't it fisuéle harsengebied yn oanwakkere wurdt; yn it iene eksperimint wurdt it sa slim net oantrune as yn it oare eksperimint. Men kin dat gebied dan krekter lokalisearje as men beide mjittings tagelyk ûndersiket yn stee fan elts apart. De tinkwize dy't hjir efter sit is dat mear ynformaasje (mear mjittings) altyd in geunstich effekt hat op de krektens dêr't de posysje fan de boarne mei ynskat wurde kin.

Ut 'e stúdzjes yn dizze twa haadstikken docht bliken dat de flater yn de ynskatte posysje yn guon sitewaasjes wol tsien kear sa lyts wurde kin as men fiif mjittings tagelyk besjocht yn stee fan alle fiif apart. De gelyktidige metoade dy't yn dizze haadstikken presintearre wurdt is tige fleksibel, sa't de brûker maklik mear of minder oannames meitsje kin oangeande de gelikens fan de boarnen yn de ferskillende mjittings. Om't dizze oannames ôfhingje fan de ynformaasje en de kar fan 'e brûker, soarget dizze fleksibiliteit tagelyk foar subjektiviteit. Dêrom wurdt yn haadstik 6 de gelyktidige metoade sa oanpast dat de brûker net mear tige spesifike oannames of ynformaasje yn it model dwaan kin. Dêrmei wurdt it model dan objektiver. By eintsjebeslút moat de

brûker sels kieze oft er it mear objektive model brûke wol of dochs kieze wol foar it model dêr't mear ynformaasje yn kin.

Oare sitewaasjes dêr't it gelyktidige model yn brûkt wurde kin, binne gelikense eksperiminten by ferskillende persoanen en MEG-mjittings fan epilepsypasjinten. De earste tapassing is nijsgjirrich om yn mear algemiene sin de lokaasje en it karakter fan guon funksjonele harsengebieden te ûndersykjen. De nije, gelyktidige oanpak by epilepsypasjinten kin wichtich wêze om ferskillende soarten fan epileptyske pykjes út ien en itselde harsengebied krekter te lokalisearjen.

Publications

F. Bijma, J.C. de Munck, H.M. Huizenga, R.M. Heethaar: *A Mathematical Approach to the Temporal Stationarity of Background Noise in MEG/EEG measurements*, NeuroImage 2003, 20(1): 233-243

F. Bijma, J.C. de Munck, K.B.E. Böcker, H.M. Huizenga, R.M. Heethaar: *The Coupled Dipole Model: an integrated model for multiple MEG/EEG data sets*, NeuroImage 2004, 23(3): 890-904

F. Bijma, J.C. de Munck, H.M. Huizenga, R.M. Heethaar, A. Nehorai: *Simultaneous estimation and testing of sources in multiple MEG data sets*, IEEE Trans. Signal Proc. Spec. Issue Brain Imag. 2005, *in press*

F. Bijma, J.C. de Munck, R.M. Heethaar: *The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker Products*, NeuroImage 2005, *in press*

J.C. de Munck, F. Bijma, P. Gaura, C. Sieluzycki, M.I. Branco, R.M. Heethaar: *A maximum likelihood estimator for trial to trial variations in noisy MEG/EEG data sets*, IEEE Trans. Biom. Eng. 2004, 51(12): 2123-2128

S.I. Gonçalves, J.C. de Munck, J.P.A. Verbunt, F. Bijma, R.M. Heethaar, F.H. Lopes da Silva: *In vivo measurement of the Brain and Skull resistivities using an EIT based method and realistic models for the head*, IEEE Trans. Biomed. Eng 2003, 50(6): 754-767

Curriculum Vitae

Fetsje Bijma is geboren op 23 maart 1977 in Hijum. In 1995 behaalde zij haar VWO Atheneum diploma aan het CSG Comenius in Leeuwarden. Vervolgens studeerde ze *cum laude* af in theoretische wiskunde aan de Rijks*universiteit* Groningen in juni 2000. Na vrijwilligerswerk onder dak- en thuislozen in Londen begon zij op 1 april 2001 haar promotie-onderzoek Brain Imaging bij het VU Medisch Centrum. Dit onderzoek richtte zich met name op de statistische eigenschappen van achtergrondruis in MEG/EEG signalen. Verder hield zij zich bezig met het simultaan analyseren van meerdere, gerelateerde MEG/EEG datasets. In 2003 verbleef zij vier maanden aan de University of Illinois at Chicago voor een onderzoeksbezoek bij professor Arye Nehorai.

Fetsje besteedt niet alleen aandacht aan wetenschap, ze steekt ook energie in (klassieke) muziek. Naast koorzang houdt ze zich bezig met het bespelen van de fluit. Ze volgt lessen bij Ferenc Hutÿra. In september 2004 was zij, samen met Waldemar Moes (piano), te beluisteren in De Tamboer in Amsterdam. In 2005 staan er concerten gepland in juni (Amsterdam) en oktober (Soest).

Index