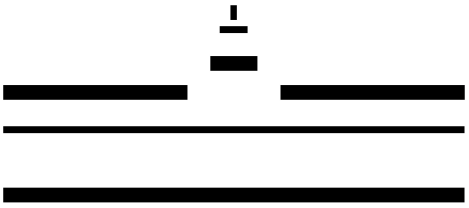


Westfälische Wilhelms-Universität Münster

Fachbereich Mathematik und Informatik

Masterarbeit in Mathematik

Beamformer inverse analysis in presurgical epilepsy diagnosis

Eingereicht von Frank Neugebauer
Münster, 14.03.2016

Gutachter:
PD Dr. Carsten Wolters
Prof. Dr. Martin Burger

Abstract

We compare the performance of the variance based beamformer SAM to the kurtosis based beamformer SAM(g_2). The analysis of kurtosis is extended with a segmentation method to bias for or against signal frequency. The beamformers are tested in simulations with different noise strengths, signal frequencies, and two dipole sources. Simulations are done for EEG and MEG in a realistic 6 compartment model with a correct and a 3 compartment leadfield. In this simulations kurtosis has shown to be better suited to localize sources when noise is not very low. The segmentation methods have worked well to separate high and low frequency sources. The model error for the 3 compartment leadfield was moderate for the EEG but high for the MEG.

Contents

Contents **1**

1 Introduction **3**

2 Beamformer **5**
2.1 Data Model 5
2.2 Linearly Constrained Minimum Variance 8
2.2.1 Optimal Filter Design 8
2.2.2 LCMV Problem 9
2.2.3 Solution to the LCMV Problem 9
2.2.4 LCMV Localization 11
2.2.5 The Neural Activity Index and Noise 12
2.2.6 Correlated Sources 12
2.3 Covariance Matrix Error and Finite Integration 14
2.4 Synthetic Aperture Magnetometry (SAM) 22

3 Kurtosis and Beamforming **25**
3.1 Kurtosis and g_2 25
3.2 Metrics for Brain Research 27
3.3 SAM(g_2) 28

4 Test and Discussion **29**
4.1 Model 29
4.2 Software 30
4.3 Methods 30
4.3.1 Grid 30
4.3.2 Signal 30
4.3.3 Noise 31
4.3.4 Window Length 32
4.3.5 Trials 32
4.3.6 Visualisation and Interpretation 32
4.4 Noise test 32
4.4.1 Results 33
4.5 Frequency 35

4.5.1 Results 36
4.6 Multiple Dipoles 38
4.6.1 Results 38
4.7 3 Department Model 41
4.7.1 Results 41
4.8 MEG Differences 43
4.8.1 Results 43

5 Summary and Discussion **46**

6 Outlook **48**

Chapter 1

Introduction

Epilepsy is a neurological disease characterized by epileptic seizures. These seizures can vary in length and intensity, but their unpredictability and consequences reduce the life quality of affected persons enormously. Electroencephalography (EEG) and magnetoencephalography (MEG) measure the activity in the brain without the need to open the head. These methods can be used by doctors to confirm epilepsy and localize the part of the brain that starts the seizures. If treatment with drugs does not help, removal of these parts can be necessary. To localize the source traditionally the spikes have to be marked by hand so averaging can cancel the effects of noise. Marking the spikes, however, is a difficult and tiring work when hours of activity have to be searched. Beamformer approaches differ from traditional approaches as they work on raw data without the need for averaging and marking. This is done by spatial filtering to separate possible sources and analysing activity by statistical means. At first, variance was used to describe source strength. In recent times the outlier test kurtosis was introduced as an alternative approach to the variance. In practice the EEG/MEG measurements are segmented into small time samples and beamforming is applied to each segment to localize possible sources. Perfect localization would be optimal, but if the existence of spikes correspond to high beamformer output values, this would already be a great help for doctors. The goal of this thesis is to introduce different methods of kurtosis analysis and to compare them to the classic variance approach. This can all be done by variations of the SAM beamformer. To do this the following topics are discussed:

- First the general approach to a spatial filter and the vector beamformer is given. A detailed derivation of the Linear Constraint Minimum Variance (LCMV) beamformer is given as it sets the foundation to the other approaches.
- As all beamformers use the covariance matrix to determine source activity, the estimation error is discussed.
- The Synthetic Aperture Magnetometry (SAM) approach is described.

- The outlier test kurtosis is introduced and two methods to change kurtosis behaviours to frequency are explained. The application of kurtosis in beamforming is discussed.
- To show the performance of the beamformers different tests are evaluated. These tests include noise strength, signal frequency, and two dipoles. The model error is done for both the EEG and MEG.
- At the end the results are summarized and an outlook is given.

Chapter 2

Beamformer

Beamformers are spatial filtering techniques that can be used to solve the EEG/MEG inverse problem. Instead of signal amplitude they use statistical methods to describe source strength. The basis of spatial filtering and covariance localization will be derived in this chapter as a wide introduction to beamforming in brain research. We will use the term "sensor" for both EEG electrodes, MEG magnetometers or gradiometers, or a combination of those. The beamformer is independent of the measurement technique as long as the leadfield is appropriate. The whole chapter is directly based on Stephanie Sillekens' diploma thesis [16] in its structure and content.

2.1 Data Model

In this section the general idea of the beamformer and the covariance as source strength is derived as done by Van Veen et al [19] and the review by Stephanie Sillekens [16].

Let there be N sensors and let q be a 3×1 source location. Let $\mathbf{H}(\mathbf{q})$ represents a $N \times 3$ solution matrix to the forward problem of a single dipole at location $\mathbf{q}$ with strength 1 in direction x, y and z, respectively. $\mathbf{H}(\mathbf{q})$ is called the leadfield matrix to $\mathbf{q}$. It represents the physical properties of the head and the sensors. In practice $\mathbf{H}$ is of course dependent on the head model and can never be fully realistic, but for the theory a perfect leadfield will be assumed. If $\mathbf{x}$ is a $N \times 1$ vector representing the noise free measurement of a dipole at location $\mathbf{q}$ and $\mathbf{m}(\mathbf{q})$ is the 3×1 moment of the source, then $\mathbf{x} = \mathbf{H}(\mathbf{q})\mathbf{m}(\mathbf{q})$.

Because the brain can be considered a linear medium the potential at the sensors is the superposition from all active sources. If we assume that $\mathbf{x}$ is only composed of the potentials due to K active dipole sources at locations $\mathbf{q}_i$, $i = 1 \dots K$, and noise, then

$$\mathbf{x} = \sum_{i=1}^{K} \mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i) + \mathbf{n}, \tag{2.1}$$

where $\mathbf{n}$ is the $N \times 1$ additive measurement noise.

The essential concept of beamformers is to treat the moments $\mathbf{m}(\mathbf{q}_i))$ as random variables and describe them in terms of mean and covariance. Every measurement is treated as a realization of these random variables.

The moment mean vector $\overline{\mathbf{m}}\ (\mathbf{q}_i)$ and the data covariance matrix $\mathbf{Cov}(\mathbf{q}_i)$ are denoted as

$$\overline{\mathbf{m}}(\mathbf{q}_i) = \mathrm{E}[\mathbf{m}(\mathbf{q}_i)], \tag{2.2}$$

$$\mathbf{Cov}(\mathbf{q}_i) = \mathrm{E}[\mathbf{m}(\mathbf{q}_i) - \overline{\mathbf{m}}(\mathbf{q}_i))(\mathbf{m}(\mathbf{q}_i) - \overline{\mathbf{m}}(\mathbf{q}_i))^T]. \tag{2.3}$$

Three basic assumptions are that the noise $\mathbf{n}$ is of zero mean, that is $\mathrm{E}[\mathbf{n}] = 0$, uncorrelated with the dipole moments with a $N \times N$ covariance matrix $\mathbf{Q}$, and that the moments associated with different dipoles are uncorrelated, this is

$$\mathbf{Cov}(\mathbf{q}_i, \mathbf{q}_k) = \mathrm{E}[(\mathbf{m}(\mathbf{q}_i) - \overline{\mathbf{m}}(\mathbf{q}_i))(\mathbf{m}(\mathbf{q}_k) - \overline{\mathbf{m}}(\mathbf{q}_k))^T] = 0 \ \forall i \neq k. \tag{2.4}$$

Section 2.2.6 deals with the problem of correlated sources, as this is the most problematic assumption for practical purposes.

With the assumptions the mean of the data vector $\mathbf{x}$ is

$$\begin{aligned}
\overline{\mathbf{x}} &= \mathrm{E}[\mathbf{x}] \\
&= \mathrm{E}\left[\sum_{i=1}^{K} \mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i) + \mathbf{n}\right] \\
&= \sum_{i=1}^{K} \mathrm{E}\left[\mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i)\right] + E\left[\mathbf{n}\right] \\
&= \sum_{i=1}^{K} \mathrm{E}\left[\mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i)\right] \\
&= \sum_{i=1}^{K} \mathbf{H}(\mathbf{q}_i)\mathrm{E}\left[\mathbf{m}(\mathbf{q}_i)\right] \\
&= \sum_{i=1}^{K} \mathbf{H}(\mathbf{q}_i)\overline{\mathbf{m}}(\mathbf{q}_i).
\end{aligned} \tag{2.5}$$

Note that $\mathbf{H}(\mathbf{q}_i)$ is deterministic as it is a function of the location of $\mathbf{q}_i$ and not

of the moment's realisation. The data covariance matrix $\mathbf{C}(\mathbf{x})$ is given by

$$
\begin{aligned}
\mathbf{C}(\mathbf{x}) =& \mathrm{E}\left[(\mathbf{x}-\overline{\mathbf{x}})(\mathbf{x}-\overline{\mathbf{x}})^T\right] \\
=& \mathrm{E}\Bigg[\left(\left(\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i)+\mathbf{n}\right)-\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)\overline{\mathbf{m}}(\mathbf{q}_i)\right) \\
& \left(\left(\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)\mathbf{m}(\mathbf{q}_i)+\mathbf{n}\right)-\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)\overline{\mathbf{m}}(\mathbf{q}_i)\right)^T\Bigg] \\
=& \mathrm{E}\Bigg[\left(\left(\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i))\right)+\mathbf{n}\right) \\
& \left(\left(\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i))\right)+\mathbf{n}\right)^T\Bigg] \\
=& \mathrm{E}\Bigg[(\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i)))(\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i))^T \\
& +\sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i))\mathbf{n}^T \\
& +\sum_{i=1}^{K}\mathbf{n}(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i))^T\mathbf{H}(\mathbf{q}_i)^T+\mathbf{n}\mathbf{n}^T\Bigg] \\
=& \mathrm{E}\left[\sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{H}(\mathbf{q}_i)(\mathbf{m}(\mathbf{q}_i)-\overline{\mathbf{m}}(\mathbf{q}_i))(\mathbf{m}(\mathbf{q}_j)-\overline{\mathbf{m}}(\mathbf{q}_j))^T\mathbf{H}(\mathbf{q}_j)^T\right]+\mathbf{Q} \\
=& \sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{H}(\mathbf{q}_i)\underbrace{\mathbf{Cov}(\mathbf{q}_i,\mathbf{q}_j)}_{=0 \text{ for } i\neq j}\mathbf{H}(\mathbf{q}_j)^T+\mathbf{Q} \qquad (2.6) \\
=& \sum_{i=1}^{K}\mathbf{H}(\mathbf{q}_i)\mathbf{C}(\mathbf{q}_i)\mathbf{H}(\mathbf{q}_i)^T+\mathbf{Q}. \qquad (2.7)
\end{aligned}
$$

As it can be seen in (2.7) the covariance matrix $\mathbf{C}$ gives a good indicator of the influence of the source at $\mathbf{q}_i$. As a scalar is more convenient to measure strength, $\mathbf{C}(\mathbf{q}_i)$ is replaced by it's trace. The source strength of $\mathbf{q}_i$ is therefore estimated as

$$\mathrm{Var}(\mathbf{q}_i) = tr(\mathbf{C}(\mathbf{q}_i)). \qquad (2.8)$$

In case of one dipole with a fixed direction this yields the true source strength:
The source moment $\mathbf{m}(\mathbf{q})$ can be divided into a dipole direction $\mathbf{d}(\mathbf{q})$ and a time dependent source strength $\mathbf{s}(t) = \mathbf{s}_q(t)$. Then,

$$\mathbf{m}(\mathbf{q}) = \mathbf{d}(\mathbf{q})\mathbf{s}(t) \qquad (2.9)$$

and the source covariance is given by ([14])

$$\mathbf{Cov}(\mathbf{q}) = \mathbf{Var}(\mathbf{s})\mathbf{d}\mathbf{d}^T \tag{2.10}$$

and $\mathbf{Var}(\mathbf{s})$ is the only non-zero eigenvalue. Because the trace of a matrix is invariant to basis transformation, this yields

$$tr(\mathbf{Cov}(\mathbf{q})) = \mathbf{Var}(\mathbf{s}). \tag{2.11}$$

2.2 Linearly Constrained Minimum Variance

Before the covariance matrix for a source location $\mathbf{q}_i$ can be estimated, the source moment $\mathbf{m}(\mathbf{q}_i)$ has to be acquired. This is done by weighting the data according to the source's influence to gain spatial samples. In this section the concept of the Linearly Constrained Minimum Variance (LCMV) beamformer is derived as done by van Veen et al. [19] and the review of Sillekens [16]. First the concept of an optimal spatial filter will be discussed. As this will show to be impractical, the LCMV filter will be derived. Then the localization with the LCMV beamformer is described.

2.2.1 Optimal Filter Design

The signal at each location in the brain consists of the three component dipole moment $\mathbf{m}(\mathbf{q}_i)$ which provides information about the strength and direction of the dipole. So for each location a filter with three components is constructed to estimate these moments. The filter focusing on location $\mathbf{q}_0$ is expressed as the $N \times 3$ matrix $\mathbf{W}(\mathbf{q}_0)$. Then the 3×1 filter output $\widehat{\mathbf{m}}$ is given by the inner product of $\mathbf{W}(\mathbf{q}_0)$ and

$$\widehat{\mathbf{m}} = \mathbf{W}(\mathbf{q}_0)^T\mathbf{x}. \tag{2.12}$$

An ideal narrow-band spatial filter passes the signal from a specific location and fully suppresses every other signal. Mathematically, it satisfies

$$\mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}) = \begin{cases} \mathbf{I} & \text{if } \mathbf{q} = \mathbf{q}_0 \\ \mathbf{0} & \text{if } \mathbf{q} \neq \mathbf{q}_0 \end{cases}, \mathbf{q} \in \Omega \tag{2.13}$$

where Ω represents the volume of the brain. This can be divided into the linear response constraint

$$\mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}) = \mathbf{I} \tag{2.14}$$

and the zero response constraint

$$\mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}_s) = \mathbf{0} \text{ for } \mathbf{q}_0 \neq \mathbf{q}_s. \tag{2.15}$$

If (2.13) is satisfied, the filter output of a noise free measurement is $\widehat{\mathbf{m}} = \mathbf{m}(\mathbf{q}_0)$, which is exactly the dipole moment at the location of interest.

If $N \geq 6$ and the columns of $\mathbf{H}(\mathbf{q}_0)$ and $\mathbf{H}(\mathbf{q}_s)$ are linearly independent, it is mathematically possible to satisfy (2.13). However, only a limited number of

source locations other than $\mathbf{q}_0$ can be completely blocked by the filter, depending on its degree of freedom. Each column of $\mathbf{W}(\mathbf{q}_0)$ only has N degrees of freedom. The unit constraint (2.14) uses three of these and each independent null uses additional three, so only $N/3-1$ locations can be optimally filtered. As this is far to few for most practical purposes, a non-ideal filter has to be designed to enable performance for more locations. One basic approach is the LCMV filter discussed next.

2.2.2 LCMV Problem

The filter output for a location $\mathbf{q}_i$ will consist of the desired moment $\mathbf{m}(\mathbf{q}_i)$ and the influence of other sources $\mathbf{q}_j$. Fulfilling the linear response constraint (2.14) is needed so the desired moment is not suppressed. As the variance of $\mathbf{m}(\mathbf{q}_i)$ will be used as source measurement, minimizing the influence of other sources means to minimize the trace of the estimated covariance matrix.

The LCMV problem is therefore mathematically posed as

$$\min_{\mathbf{W}(\mathbf{q}_0)} tr\big(\mathbf{C}(\widehat{\mathbf{m}})\big) = \min_{\mathbf{W}(\mathbf{q}_0)} tr\big(\mathbf{C}(\mathbf{W}^T(\mathbf{q}_0)\mathbf{x}\big) \text{ subject to } \mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}_0) = \mathbf{I}. \tag{2.16}$$

Using (2.12) and considering that $\mathbf{W}$ is not a random quantity, $\mathbf{C}(\widehat{\mathbf{m}})$ can be written as (see [14])

$$\mathbf{C}(\widehat{\mathbf{m}}) = \mathbf{W}^T(\mathbf{q}_0)\mathbf{C}(\mathbf{x})\mathbf{W}(\mathbf{q}_0). \tag{2.17}$$

Therefore, the LCMV problem can also be expressed as

$$\min_{\mathbf{W}(\mathbf{q}_0)} tr\big(\mathbf{W}^T(\mathbf{q}_0)\mathbf{C}(\mathbf{x})\mathbf{W}(\mathbf{q}_0)\big) \text{ subject to } \mathbf{W}^T(\mathbf{q}_0)\mathbf{H}(\mathbf{q}_0) = \mathbf{I}. \tag{2.18}$$

2.2.3 Solution to the LCMV Problem

The solution to (2.18) can be obtained using Lagrange multipliers. Let $\mathbf{L}$ be a 3×3 matrix of Lagrange multipliers. For readability the terms $(\mathbf{q}_0)$ and $(\mathbf{x})$ are omitted. As $\mathbf{W}^T\mathbf{H}$ is a 3×3 matrix, 9 Lagrange multipliers are needed. Let those be denoted as $\lambda_{1,1}\dots\lambda_{1,3}, \lambda_{2,1}\dots\lambda_{2,3}, \lambda_{3,1}\dots\lambda_{3,3}$ and let Λ be the 3×3 matrix of them. Note that $\lambda_{i,j}$ is the multiplier for $(\mathbf{W}^T\mathbf{H}-\mathbf{I})_{j,i}$ to account for the later matrix multiplication. Again for readability $\mathbf{W}$ and Λ are written for

the arguments of $\mathbf{L}$ instead of the single variables. The Lagrangian is

$$\begin{aligned}\mathbf{L}(\mathbf{W},\Lambda) =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W}\big) + \sum_{i,j=1}^{3} \lambda_{j,i}(\mathbf{W}^T\mathbf{H}-\mathbf{I})_{i,j} \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W}\big) + \sum_{i=1}^{3}\sum_{j=1}^{3} \lambda_{j,i}(\mathbf{W}^T\mathbf{H}-\mathbf{I})_{i,j} \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W}\big) + \sum_{i=1}^{3}((\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda)_{i,i} \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W}\big) + tr\big((\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda\big) \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda\big). \end{aligned} \tag{2.19}$$

Substituting Λ with 2Λ and using that the trace is invariant against transposition for any square matrix, (2.19) can be rewritten as

$$\begin{aligned}\mathbf{L}(\mathbf{W},\Lambda) =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H}-\mathbf{I})2\Lambda\big) \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda + (\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda\big) \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda + \Lambda^T(\mathbf{H}^T\mathbf{W}-I)\big) \end{aligned} \tag{2.20}$$

and further expanded to

$$\begin{aligned}\mathbf{L}(\mathbf{W},\Lambda) =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H}-\mathbf{I})\Lambda + \Lambda^T(\mathbf{H}^T\mathbf{W}-I)\big) \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + (\mathbf{W}^T\mathbf{H}\Lambda-\Lambda) + (\Lambda^T\mathbf{H}^T\mathbf{W}-\Lambda^T)\big) \\ =& tr\big(\mathbf{W}^T\mathbf{C}\mathbf{W} + \mathbf{W}^T\mathbf{C}\mathbf{C}^{-1}\mathbf{H}\Lambda + \Lambda^T\mathbf{H}^T\mathbf{W} \\ & + \Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\Lambda - \Lambda - \Lambda^T - \Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\Lambda\big) \\ =& tr\big((\mathbf{W}^T\mathbf{C} + \Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{C})(\mathbf{W} + \mathbf{C}^{-1}\mathbf{H}\Lambda) - \Lambda - \Lambda^T \\ & - \Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\Lambda\big) \\ =& tr\big((\mathbf{W}^T + \Lambda^T\mathbf{H}^T\mathbf{C}^{-1})\mathbf{C}(\mathbf{W} + \mathbf{C}^{-1}\mathbf{H}\Lambda) - \Lambda - \Lambda^T \\ & - \Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\Lambda\big) \\ =& tr\big((\mathbf{W} + \mathbf{C}^{-1}\mathbf{H}\Lambda)^T\mathbf{C}(\mathbf{W} + \mathbf{C}^{-1}\mathbf{H}\Lambda) - \Lambda - \Lambda^T \\ & - \Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\Lambda\big). \end{aligned} \tag{2.21}$$

Here $\mathbf{C}$ is assumed to be invertible in the presence of noise. As $\mathbf{C}$ is a symmetric matrix, it is $\mathbf{C} = \mathbf{C}^T$ and thus $\mathbf{C}^{-1} = (\mathbf{C}^{-1})^T$. Since $\mathbf{C}$ is a covariance matrix, it is positive semi-definite and as only the first term is a function of $\mathbf{W}$ the minimum of $\mathbf{L}(\mathbf{W},\Lambda)$ is attained by setting the first term to zero, that is

$$\mathbf{W} = -\mathbf{C}^{-1}\mathbf{H}\mathbf{L}. \tag{2.22}$$

Λ is now obtained by substituting $\mathbf{W}$ in the constraint $\mathbf{W}^T\mathbf{H} = \mathbf{I}$ to get

$$-\Lambda^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H} = \mathbf{I} \tag{2.23}$$

or

$$\Lambda^T = -(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}. \tag{2.24}$$

Substituting (2.24) into (2.22) yields the solution

$$\begin{aligned}
\mathbf{W}^T =& \Big(-\mathbf{C}^{-1}\mathbf{H}\mathbf{L}\Big)^T \\
=& \Big(-\mathbf{C}^{-1}\mathbf{H}\big(-(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\big)^T\Big)^T \\
=& \Big(\mathbf{C}^{-1}\mathbf{H}\big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\big)^T\Big)^T \\
=& (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}(\mathbf{C}^{-1}\mathbf{H})^T \\
=& (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T(\mathbf{C}^{-1})^T \\
=& (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1},
\end{aligned}$$

or, writing all arguments again,

$$\mathbf{W}(\mathbf{q}_0)^T = \Big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{C}(\mathbf{x})^{-1}\mathbf{H}(\mathbf{q}_0)\Big)^{-1}\mathbf{H}(\mathbf{q}_0)^T\mathbf{C}(\mathbf{x})^{-1}. \tag{2.25}$$

2.2.4 LCMV Localization

Using the filter derived in the previous subsection, (2.25), the estimated strength or variance at a location $\mathbf{q}_0$ is the minimum of the cost function in (2.16) or (2.18).

Substituting (2.25) into (2.18) leads to

$$\begin{aligned}
\widehat{\mathrm{Var}}(\mathbf{q}_0) =& tr\Big(\mathbf{W}^T\mathbf{C}\mathbf{W}\Big) \\
=& tr\Big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{C}(\mathbf{C}^{-1})^T\mathbf{H}\big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\big)^T\Big) \\
=& tr\Big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})\big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\big)^T\Big) \\
=& tr\Big(\big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\big)^T\Big) \\
=& tr\Big((\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\Big)
\end{aligned}$$

or with the arguments

$$\widehat{\mathrm{Var}}(\mathbf{q}_0) = tr\Big(\big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{C}(\mathbf{x})^{-1}\mathbf{H}(\mathbf{q}_0)\big)^{-1}\Big). \tag{2.26}$$

For localization (2.26) is calculated as a function of the locations $\mathbf{q}_i$ within the brain volume. Regions of large variance are presumed to have substantial neural activity. The result of (2.26) is referred to as neural activity or "spatial spectrum" of neural activity. This approach does not require specification of the number of dipole sources in the brain.

However, it should be mentioned that it is up to interpretation what values should be considered high. As noise leads to a spread of the signal, separation of a spread signal and more than one dipole is a task left open for the researcher.

2.2.5 The Neural Activity Index and Noise

Of course, noise has an influence on the beamformer performance and is often a significant part of the estimated spatial spectrum. Especially brain noise that appears concentrated or non-uniformly distributed will interfere with localization. To handle the influence of noise, it is useful to consider a measurement $\mathbf{x}$ consisting entirely of spatial (sensor) white noise with a mean of $\mu = 0$ and standard deviation $\sigma = 1$, $\mathbf{C}(\mathbf{x}) = \mathbf{I}$. In this case, (2.26) simplifies to

$$\widehat{\mathrm{Var}}(\mathbf{q}_0) = tr\Big(\big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{H}(\mathbf{q}_0)\big)^{-1}\Big). \tag{2.27}$$

Hence, the noise spatial spectrum just depends on the leadfield matrices $\mathbf{H}(\mathbf{q}_0)$. For locations $\mathbf{q}_0$ far from any sensor, the elements of $\mathbf{H}(\mathbf{q}_0)$ are generally quite small, so $\big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{H}(\mathbf{q}_0)\big)^{-1}$ will have large elements, resulting in a large value for $\widehat{\mathrm{Var}}(\mathbf{q}_0)$. Therefore, noise generally has a dome shaped spatial spectrum corresponding to the deepness of the location. Locations of sensors and unusual head geometry can of course alter this shape significantly. While the LCMV spatial spectrum is not linear, the Woodbury matrix identity states that that it contains an additive noise component of the form

$$tr\Big(\big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{Q}^{-1}\mathbf{H}(\mathbf{q}_0)\big)^{-1}\Big). \tag{2.28}$$

By normalizing the estimated spatial spectrum the influence of noise should be reduced. The normalized term is named the "neural activity index" and is defined as

$$\widehat{\mathrm{Var}_N}(\mathbf{q}_0) = \frac{tr\Big(\big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{C}(\mathbf{x})^{-1}\mathbf{H}(\mathbf{q}_0)\big)^{-1}\Big)}{tr\Big(\big(\mathbf{H}(\mathbf{q}_0)^T\mathbf{Q}^{-1}\mathbf{H}(\mathbf{q}_0)\big)^{-1}\Big)}. \tag{2.29}$$

The neural activity index requires knowledge of the noise covariance matrix. If noise is assumed to be spatial white noise that is uncorrelated between channels, then $\mathbf{Q} = \sigma^2\mathbf{I}$.

2.2.6 Correlated Sources

The moments associated with distinct dipoles were assumed to be uncorrelated (2.4). This is problematic in practice because high correlation is not unusual. Auditory sources for example fire at the same time on both sides of the brain. This leads to a strong reduction of the estimated activity for the correlated sources.Using that (2.6) applies for correlated sources, it can be derived that the output variance at location $\mathbf{q}_k$ is given by

$$
\begin{aligned}
\widehat{\mathrm{Var}}(\mathbf{q}_k) &= tr\Big(\mathbf{W}(\mathbf{q}_k)^T\mathbf{C}(\mathbf{x})\mathbf{W}(\mathbf{q}_k)\Big) \\
&= tr\Bigg(\sum_{i=1}^{K}\sum_{n=1}^{K}\mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_i)\mathbf{Cov}(\mathbf{q}_i,\mathbf{q}_n)\mathbf{H}(\mathbf{q}_n)^T\mathbf{W}(\mathbf{q}_k) \\
&\quad + \mathbf{W}(\mathbf{q}_k)^T\mathbf{Q}\mathbf{W}(\mathbf{q}_k)\Bigg) \\
&= tr\Bigg(\underbrace{\mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_k)}_{=\mathbf{I}\ (2.14)}\mathbf{Cov}(\mathbf{q}_k,\mathbf{q}_k)\underbrace{\mathbf{H}(\mathbf{q}_k)^T\mathbf{W}(\mathbf{q}_k)}_{=\mathbf{I}} \\
&\quad + \sum_{\substack{n=1\\n\neq k}}^{K}\underbrace{\mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_k)}_{=\mathbf{I}}\mathbf{Cov}(\mathbf{q}_k,\mathbf{q}_n)\mathbf{H}(\mathbf{q}_n)^T\mathbf{W}(\mathbf{q}_k) \\
&\quad + \sum_{\substack{i=1\\i\neq k}}^{K}\mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_i)\mathbf{Cov}(\mathbf{q}_i,\mathbf{q}_k)\underbrace{\mathbf{H}(\mathbf{q}_k)^T\mathbf{W}(\mathbf{q}_k)}_{=\mathbf{I}} \\
&\quad + \sum_{\substack{i=1\\i\neq k}}^{K}\sum_{\substack{n=1\\n\neq k}}^{K}\mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_i)\mathbf{Cov}(\mathbf{q}_i,\mathbf{q}_n)\mathbf{H}(\mathbf{q}_n)^T\mathbf{W}(\mathbf{q}_k) \\
&\quad + \mathbf{W}(\mathbf{q}_k)^T\mathbf{Q}\mathbf{W}(\mathbf{q}_k)\Bigg) \\
&= tr\Bigg(\mathbf{C}(\mathbf{q}_k) + \sum_{\substack{i=1\\i\neq k}}^{K}\mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_i)\mathbf{C}(\mathbf{q}_i)\mathbf{H}(\mathbf{q}_i)^T\mathbf{W}(\mathbf{q}_k) \\
&\quad + \sum_{\substack{i=1\\i\neq k}}^{K}\big(\mathbf{Cov}(\mathbf{q}_k,\mathbf{q}_i)\mathbf{H}(\mathbf{q}_i)^T\mathbf{W}(\mathbf{q}_k) + \mathbf{W}(\mathbf{q}_k)^T\mathbf{H}(\mathbf{q}_i)\mathbf{Cov}(\mathbf{q}_i,\mathbf{q}_k)\big) \\
&\quad + \sum_{\substack{n,i=1\\n,i\neq k\\n\neq i}}\big(\mathbf{W}(\mathbf{q}_k)^T\big(\mathbf{H}(\mathbf{q}_i)\mathbf{Cov}(\mathbf{q}_i,\mathbf{q}_n)\mathbf{H}(\mathbf{q}_n)^T \\
&\quad + \mathbf{H}(\mathbf{q}_n)\mathbf{Cov}(\mathbf{q}_n,\mathbf{q}_i)\mathbf{H}(\mathbf{q}_i)^T\big)\mathbf{W}(\mathbf{q}_k)\big) \\
&\quad + \mathbf{W}(\mathbf{q}_k)^T\mathbf{Q}\mathbf{W}(\mathbf{q}_k)\Bigg). \qquad (2.30)
\end{aligned}
$$

The cross terms on the right-hand side of (2.30) are not positive for all $\mathbf{W}(\mathbf{q}_k)$. By exploiting the correlation between sources the estimated variance can thus be much lower than the real variance. For full correlation even $tr\big(\mathbf{W}(\mathbf{q}_k)^T\mathbf{C}(\mathbf{x})\mathbf{W}(\mathbf{q}_k)\big) = 0$ hold. That means that the beamformer is blind to full correlated sources, which is a well known phenomenon of LCMV in radar

and sonar [21].

[15] have shown that the reductions in the reconstructed signal intensities is small for sources with a medium degree of correlation (less than 20% for two correlated sources (correlation coefficients of 0.7-0.8) can still be reconstructed if their intensities are strong enough to overcome the signal cancellation. This can be achieved by manipulating the sensors to weight the signals. For auditory sources only one side of the head can be measured, for example.

2.3 Covariance Matrix Error and Finite Integration

From (2.26) it is obvious that the spatial filters and therefore the source reconstruction depends directly on the covariance matrix. This stays true for the SAM filter as SAM described in section 2.4.

As the real data covariance matrix is unknown, the covariance matrix used is a data driven estimation. In this section the error gained by finite integration time in estimating the covariance matrix is derived analogously to Sillekens [16] revision of Brookes et al [2]. Even though Brookes focuses on the MEG, there is no reason the calculations do not hold for the EEG as well.

In this section we assume a correct leadfield matrix $\mathbf{H}$ and one active source with a forward solution $\mathbf{h}$ and variance ρ^2. Noise is assumed to be of zero mean and uncorrelated to the signal and among sensors with a covariance matrix $\mathbf{Q} = \sigma^2\mathbf{I}$. Noise is assumed to be independent in time. Furthermore, the measurement is assumed to be of zero mean. With (2.7) and (2.8) the true covariance matrix is

$$\mathbf{C}_0 = \rho^2\mathbf{h}\mathbf{h}^T + \sigma^2\mathbf{I}. \tag{2.31}$$

The first term represents the real source power, the second the additive sensor noise power. The data covariance matrix is in practice estimated as

$$\mathbf{C}_{ij} = \frac{1}{T}\sum_{l=1}^{T}(\mathbf{x}_i(t_l))(\mathbf{x}_j(t_l)), \tag{2.32}$$

where T is the number of recorded samples and $\mathbf{x}_i(t_l), \mathbf{x}_j(t_l)$ represents the lth measurement of the ith or jth sensor at time point t_l.

Let the true covariance matrix be denoted as $\mathbf{C}_0$ and $\mathbf{C}$ be its estimation. We assume that the difference of the matrices is only due to finite integration time and define the covariance matrix error as

$$\Delta\mathbf{C} = \mathbf{C} - \mathbf{C}_0. \tag{2.33}$$

Under our assumptions $\mathbf{x}_i(t_l)$ is given by

$$\mathbf{x}_i(t_l) = \mathbf{x}_{0i}(t_l) + \mathbf{n}_i(t_l), \tag{2.34}$$

where $\mathbf{x}_{0i}(t_l)$ is the noise free measurement and $\mathbf{n}_i$ the additional sensor noise. Substituting (2.34) into (2.32) yields

$$\mathbf{C}_{ij} = \frac{1}{T}\sum_{l=1}^{T}\big(\mathbf{x}_{0i}(t_l) + \mathbf{n}_i(t_l)\big)\big(\mathbf{x}_{0i}(t_l) + \mathbf{n}_j(t_l)\big). \tag{2.35}$$

By expanding product and sum it can be shown that

$$\mathbf{C}_{ij} = \mathbf{C}_{0ij} + \Delta\mathbf{C}_{ij}, \tag{2.36}$$

where $\mathbf{C}_{0ij} = \frac{1}{T}\sum_{l=1}^{T}\mathbf{x}_{0i}(t_l)\mathbf{x}_{0j}(t_l)$ holds.

For a sufficiently large T the estimation on the right hand side is assumed to be equal to the analytical $\mathbf{C}_{0ij}$ from (2.33) due to the sampling theorem.

The remaining terms of (2.35) are $\Delta\mathbf{C}_{ij} = A_1 + A_2 + A_3$ with

$$A_1 = \frac{1}{T}\sum_{l=1}^{T}\mathbf{x}_{0i}(t_l)\mathbf{n}_j(t_l), \tag{2.37}$$

$$A_2 = \frac{1}{T}\sum_{l=1}^{T}\mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l), \tag{2.38}$$

$$A_3 = \frac{1}{T}\sum_{l=1}^{T}\mathbf{n}_i(t_l)\mathbf{n}_j(t_l). \tag{2.39}$$

$$\tag{2.40}$$

With the assumptions $\Delta\mathbf{C}_{ij}$ tends to zero in the limit of infinite integration time. Brookes assumes that the A_k are uncorrelated, but Sillekens [16] has noted that for $i = j$ $A_i = A_j$ holds. To estimate the error, the cases $i = j$ and $i \neq j$ must separated.

Let $i \neq j$.
In this case A_1 is uncorrelated to A_2 and A_3:

$$\begin{aligned}
\mathrm{Cov}(A_1, A_2) =& \mathrm{E}[A_1 A_2] - \mathrm{E}[A_1]\mathrm{E}[A_2] \\
=& \mathrm{E}[A_1 A_2] - \mathrm{E}\left[\frac{1}{T}\sum_{l=1}^{T} \mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)\right]\mathrm{E}[A_2] \\
=& \mathrm{E}[A_1 A_2] - \left(\frac{1}{T}\sum_{l=1}^{T} \mathrm{E}[\mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)]\right)\mathrm{E}[A_2] \\
=& \mathrm{E}[A_1 A_2] - \left(\frac{1}{T}\sum_{l=1}^{T} \mathrm{E}[\mathbf{x}_{0j}(t_l)]\underbrace{\mathrm{E}[\mathbf{n}_i(t_l)]}_{=0}\right)\mathrm{E}[A_2] \\
=& \mathrm{E}[A_1 A_2] - 0 \\
=& \mathrm{E}\left[\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T} \mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)\mathbf{x}_{0i}(t_r)\mathbf{n}_j(t_r)\right] \\
=& \left(\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T} \mathrm{E}[\mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)\mathbf{x}_{0i}(t_r)\mathbf{n}_j(t_r)]\right) \\
=& \left(\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T} \mathrm{E}[\mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)\mathbf{x}_{0i}(t_r)]\underbrace{\mathrm{E}[\mathbf{n}_j(t_r)]}_{=0}\right) \\
=& 0
\end{aligned} \tag{2.41}$$

and

$$\begin{aligned}
\mathrm{Cov}(A_1, A_3) =& \mathrm{E}[A_1 A_3] - \mathrm{E}[A_1]\mathrm{E}[A_3] \\
=& \mathrm{E}[A_1 A_3] - 0\mathrm{E}[A_3] \\
=& \mathrm{E}\left[\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T} \mathbf{x}_{0i}(t_l)\mathbf{n}_j(t_l)\mathbf{n}_i(t_r)\mathbf{n}_j(t_r)\right] \\
=& \mathrm{E}\left(\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T} \mathrm{E}[\mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)\mathbf{x}_{0i}(t_r)\mathbf{n}_j(t_r)]\right) \\
=& \mathrm{E}\left(\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T} \mathrm{E}[\mathbf{x}_{0j}(t_l)\mathbf{n}_i(t_l)\mathbf{x}_{0i}(t_r)]\underbrace{\mathrm{E}[\mathbf{n}_j(t_r)]}_{=0}\right) \\
=& 0.
\end{aligned} \tag{2.42}$$

A_2 and A_3 can be shown to be uncorrelated analogously. In this case

$$\mathrm{Var}(\Delta\mathbf{C}_{ij}) = \mathrm{Var}(A_1) + \mathrm{Var}(A_2) + \mathrm{Var}(A_3). \tag{2.43}$$

holds.

Let now be $i = j$.
Then $A_1 = A_2$ and A_1 is uncorrelated with A_3:

$$\begin{aligned}
\mathrm{Cov}(A_1, A_3) =& \mathrm{E}[A_1 A_3] - \mathrm{E}[A_1]\mathrm{E}[A_3] \\
=& \mathrm{E}[A_1 A_3] - 0\mathrm{E}[A_3] \\
=& \mathrm{E}\left[\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T}\mathbf{x}_{0i}(t_l)\mathbf{n}_i(t_l)\mathbf{n}_i(t_r)^2\right] \\
=& \left(\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T}\mathrm{E}[\mathbf{x}_{0i}(t_l)\mathbf{n}_i(t_l)\mathbf{n}_i(t_r)^2]\right) \\
=& \left(\frac{1}{T^2}\sum_{r=1}^{T}\sum_{l=1}^{T}\underbrace{\mathrm{E}[\mathbf{n}_i(t_l)]}_{=0}\mathrm{E}[\mathbf{x}_{0i}(t_l)\mathbf{n}_i(t_r)^2]\right) \\
=& 0. \qquad (2.44)
\end{aligned}$$

in this case $\Delta\mathbf{C}_{ii} = 2A_1 + A_3$ holds and (2.43) must be rewritten to

$$\begin{aligned}
\mathrm{Var}(\Delta\mathbf{C}_{ij}) =& \mathrm{Var}(2A_1 + A_3) \\
=& \mathrm{Var}(2A_1) + \mathrm{Var}(A_3) \\
=& 4\mathrm{Var}(A_1) + \mathrm{Var}(A_3) \\
=& 2(\mathrm{Var}(A_1) + \mathrm{Var}(A_2)) + \mathrm{Var}(A_3) \qquad (2.45)
\end{aligned}$$

for $i = j$.

By using the Pearson product-moment correlation coefficient the terms in (2.37),(2.38), and (2.39) can be further expanded. Let $\widehat{\cdot}$ indicate empirical quantities and $||\cdot||_F$ the Frobenius norm. Then it is

$$\begin{aligned}
A_3 =& \frac{1}{T}\sum_{l=1}^{T}\mathbf{n}_i(t_l)\mathbf{n}_j(t_l) \qquad (2.46) \\
=& \widehat{\mathrm{Cov}(\mathbf{n}_i, \mathbf{n}_j)} \cdot 1 \\
=& \widehat{\mathrm{Cov}(\mathbf{n}_i, \mathbf{n}_j)} \frac{\sqrt{\sum_{l=1}^{T}\mathbf{n}_i(t_l)^2}\sqrt{\sum_{l=1}^{T}\mathbf{n}_j(t_l)^2}}{T\frac{1}{T}\sqrt{\sum_{l=1}^{T}\mathbf{n}_i(t_l)^2}\sqrt{\sum_{l=1}^{T}\mathbf{n}_j(t_l)^2}} \\
=& \widehat{\mathrm{Cov}(\mathbf{n}_i, \mathbf{n}_j)} \frac{||\mathbf{n}_i||_F||\mathbf{n}_j||_F}{T\widehat{\mathrm{Std}}(\mathbf{n}_i)\widehat{\mathrm{Std}}(\mathbf{n}_j)} \\
=& \frac{\widehat{r}(\mathbf{n}_i, \mathbf{n}_j)||\mathbf{n}_i||_F||\mathbf{n}_j||_F}{T}. \qquad (2.47)
\end{aligned}$$

If T is sufficiently large, the Frobenius norm can be estimated by

$$||\mathbf{n}_i||_F = \sqrt{\sum_{l=1}^{T} \mathbf{n}_i(t_l)^2} \approx \sqrt{\sigma_i^2 T} = \sigma_i \sqrt{T}. \tag{2.48}$$

where σ_i represents the level of the uncorrelated noise at sensor i.

Since $\mathbf{n}_i$ and $\mathbf{n}_j$ are uncorrelated random processes, $r(\mathbf{n}_i, \mathbf{n}_j) \to 0$ holds in the limit to infinite integration. For finite integration times the standard deviation of the correlation coefficient $\mathrm{Std}((\widehat{r}(\mathbf{n}_i, \mathbf{n}_j))$ is given by [1]

$$\mathrm{Std}((\widehat{r}(\mathbf{n}_i, \mathbf{n}_j)) = \frac{1}{\sqrt{T}} \text{ for } i \neq j \tag{2.49}$$

and

$$\mathrm{Std}((\widehat{r}(\mathbf{n}_i, \mathbf{n}_j)) = \frac{2}{\sqrt{T}} \text{ for } i = j. \tag{2.50}$$

Note that the correlation coefficient in (2.49) is defined as $\frac{\widehat{\mathrm{Cov}(\mathbf{n}_i, \mathbf{n}_j)}}{\mathrm{Std}(\mathbf{n}_i)\mathrm{Std}(\mathbf{n}_j)}$ and not $\frac{\widehat{\mathrm{Cov}(\mathbf{n}_i, \mathbf{n}_j)}}{\widehat{\mathrm{Std}(\mathbf{n}_i)}\widehat{\mathrm{Std}(\mathbf{n}_j)}}$ as in (2.47).

The autocorrelation coefficient in (2.50) is defined analogously. Combining (2.46)-(2.50) gives

$$\mathrm{Std}(A_3) \approx \frac{\sigma_i \sigma_j}{\sqrt{T}} \text{ for } i \neq j \tag{2.51}$$

and

$$\mathrm{Std}(A_3) \approx \sqrt{\frac{2}{T}} \sigma_i^2 \text{ for } i = j. \tag{2.52}$$

This is shown by

$$\begin{aligned}
\mathrm{Std}(A_3) =& \sqrt{\mathrm{Var}}(A_3) \\
=& \sqrt{\mathrm{Var}\left(\frac{\widehat{r}(\mathbf{n}_i, \mathbf{n}_j)||\mathbf{n}_i||_F||\mathbf{n}_j||_F}{T}\right)} \\
\approx& \sqrt{\mathrm{Var}\left(\frac{\widehat{r}(\mathbf{n}_i, \mathbf{n}_j)\sigma_i\sqrt{T}\sigma_j\sqrt{T}}{T}\right)} \\
=& \sqrt{\frac{\sigma_i^2\sigma_j^2T^2}{T^2}\mathrm{Var}(\widehat{r}(\mathbf{n}_i, \mathbf{n}_j))} \\
=& \sigma_i\sigma_j\sqrt{\mathrm{Var}(\widehat{r}(\mathbf{n}_i, \mathbf{n}_j))} \\
=& \sigma_i\sigma_j\mathrm{Std}(\widehat{r}(\mathbf{n}_i, \mathbf{n}_j)) \\
=& \begin{cases} \frac{\sigma_i\sigma_j}{\sqrt{T}}, & i \neq j \\ \sqrt{\frac{2}{T}}\sigma_i^2, & i = j. \end{cases}
\end{aligned} \tag{2.53}$$

The standard deviations of A_1 and A_2 are derived similarly. A_1 can be written as

$$\begin{aligned} A_1 =& \frac{1}{T}\sum_{l=1}^{T}\mathbf{x}_{0i}(t_l)\mathbf{n}_j(t_l) \\ =& \widehat{\mathrm{Cov}}(\mathbf{x}_{0i},\mathbf{n}_j) \\ =& \widehat{r}(\mathbf{n}_i,\mathbf{n}_j)\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\widehat{\mathrm{Std}}(\mathbf{n}_j) \\ =& \widehat{r}(\mathbf{n}_i,\mathbf{n}_j)\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\frac{||\mathbf{n}_j||_F}{\sqrt{T}}. \end{aligned} \tag{2.54}$$

$\widehat{\mathrm{Std}}(\mathbf{x}_{0i})$ is given by

$$\widehat{\mathrm{Std}}(\mathbf{x}_{0i}) = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(\mathbf{x}_{0i}(t_l))^2} \approx \rho^2\mathbf{h}_i, \tag{2.55}$$

with $\mathbf{h}$ representing the forward solution to the fixed dipole at the ith sensor. Analogous to equation (2.49), it is

$$\mathrm{Std}(\widehat{r}(\mathbf{x}_{0i},\mathbf{n}_j)) = \frac{1}{\sqrt{T}}. \tag{2.56}$$

Together with (2.48) and (2.54) this concludes to

$$\begin{aligned} \mathrm{Std}(A_1) =& \sqrt{\mathrm{Var}(A_1)} \\ =& \sqrt{\mathrm{Var}\left(\widehat{r}(\mathbf{x}_{0i},\mathbf{n}_j)\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\frac{||\mathbf{n}_j||}{\sqrt{T}}\right)} \\ \approx& \sqrt{\mathrm{Var}\left(\widehat{r}(\mathbf{x}_{0i},\mathbf{n}_j)\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\frac{\sigma_j\sqrt{T}}{\sqrt{T}}\right)} \\ =& \sqrt{\widehat{\sigma}^2(\mathbf{x}_{0i})\sigma_j^2\mathrm{Var}(\widehat{r}(\mathbf{x}_{0i},\mathbf{n}_j))} \\ =& \widehat{\mathrm{Std}}(\mathbf{x}_{0i})\sigma_j\sqrt{\mathrm{Var}(\widehat{r}(\mathbf{x}_{0i},\mathbf{n}_j))} \\ =& \widehat{\mathrm{Std}}(\mathbf{x}_{0i})\sigma_j\frac{1}{\sqrt{T}} \\ =& \frac{\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\sigma_j}{\sqrt{T}}. \end{aligned} \tag{2.57}$$

The standard deviation of A_2,

$$\mathrm{Std}(A_2) \approx \frac{\widehat{\mathrm{Std}}(\mathbf{x}_{0j})\sigma_i}{\sqrt{T}}, \tag{2.58}$$

is derived analogously.

As the noise levels were assumed to be equal at all channels, that is $\sigma_i = \sigma_j = \sigma$, substituting of (2.53),(2.57), and (2.58) into (2.43) yields

$$\begin{aligned}
\mathrm{Var}(\Delta\mathbf{C}_{ij}) =& \mathrm{Var}(A_1) + \mathrm{Var}(A_2) + \mathrm{Var}(A_3) \\
=& \big(\mathrm{Std}(A_1)\big)^2 + \big(\mathrm{Std}(A_2)\big)^2 + \big(\mathrm{Std}(A_3)\big)^2 \\
\approx& \left(\frac{\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\sigma}{\sqrt{T}}\right)^2 + \left(\frac{\widehat{\mathrm{Std}}(\mathbf{x}_{0j})\sigma}{\sqrt{T}}\right)^2 + \left(\sqrt{\frac{1}{T}}\sigma^2\right)^2 \\
=& \frac{\sigma^2}{T}\left((\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + (\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sigma^2\right)
\end{aligned} \tag{2.59}$$

for $i \neq j$.

Substituting into (2.45) yields

$$\begin{aligned}
\mathrm{Var}(\Delta\mathbf{C}_{ij}) =& 2\mathrm{Var}(A_1) + 2\mathrm{Var}(A_2) + \mathrm{Var}(A_3) \\
=& 2\big(\mathrm{Std}(A_1)\big)^2 + 2\big(\mathrm{Std}(A_2)\big)^2 + \big(\mathrm{Std}(A_3)\big)^2 \\
\approx& 2\left(\frac{\widehat{\mathrm{Std}}(\mathbf{x}_{0i})\sigma}{\sqrt{T}}\right)^2 + 2\left(\frac{\widehat{\mathrm{Std}}(\mathbf{x}_{0j})\sigma}{\sqrt{T}}\right)^2 + \left(\sqrt{\frac{2}{T}}\sigma^2\right)^2 \\
=& \frac{\sigma^2}{T}\left(2(\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + 2(\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + 2\sigma^2\right) \\
=& \frac{2\sigma^2}{T}\left((\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + (\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sigma^2\right)
\end{aligned} \tag{2.60}$$

for $i = j$.

$\mathrm{Var}(\Delta\mathbf{C}_{ij})$ for $i = j$ and $i \neq j$ only differ by a factor of two which can be disregarded.

To derive a scalar value for the error we use Jensen's inequality and $\mathrm{E}[\Delta\mathbf{C}_{ij}] = 0$.

Combining the results above, it is

$$\begin{aligned}
\mathrm{E}[||\Delta\mathbf{C}||_F] =& \mathrm{E}\left[\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\Delta\mathbf{C}_{ij}^2}\right] \\
\leq& \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{E}[\Delta\mathbf{C}_{ij}^2]} \\
=& \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Var}(\Delta\mathbf{C}_{ij})} \\
=& \sqrt{\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\mathrm{Var}(\Delta\mathbf{C}_{ij}) + \sum_{i=1}^{N}\mathrm{Var}(\Delta\mathbf{C}_{ii})} \\
\approx& \frac{\sigma}{\sqrt{T}}\Bigg(\sum_{i=1}^{N}\sum_{j=1}^{N}\left((\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + (\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sigma^2\right) \\
&+ \sum_{i=1}^{N}\left((\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + (\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sigma^2\right)\Bigg)^{\frac{1}{2}}. \qquad (2.61)
\end{aligned}$$

The summations in this expression can be done separately, so $||\mathbf{C}||_F$ can be rewritten to

$$\begin{aligned}
\mathrm{E}[||\Delta\mathbf{C}||_F] \lessapprox& \frac{\sigma}{\sqrt{T}}\Bigg(\sum_{i=1}^{N}\sum_{j=1}^{N}\left((\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + (\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sigma^2\right) \\
&+ \sum_{i=1}^{N}\left((\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + (\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sigma^2\right)\Bigg)^{\frac{1}{2}} \\
=& \frac{\sigma}{\sqrt{T}}\Bigg(\sum_{i=1}^{N}\sum_{j=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}\sigma^2 \\
&+ \sum_{i=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + \sum_{i=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + \sum_{i=1}^{N}\sigma^2\Bigg)^{\frac{1}{2}} \\
=& \frac{\sigma}{\sqrt{T}}\Bigg(N\sum_{i=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + N\sum_{i=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + N^2\sigma^2 \\
&+ \sum_{i=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0i}))^2 + \sum_{i=1}^{N}(\widehat{\mathrm{Std}}(\mathbf{x}_{0j}))^2 + N\sigma^2\Bigg)^{\frac{1}{2}} \\
=& \sigma\sqrt{\frac{N+1}{T}\left(2\sum_{i=1}^{N}\left(\left(\mathrm{Std}(\mathbf{x}_{0i})\right)^2\right) + N\sigma^2\right)}. \qquad (2.62)
\end{aligned}$$

To conclude,

$$\begin{aligned}
\mathrm{E}[||\Delta\mathbf{C}||_F] \lessapprox & \sigma\sqrt{\frac{N+1}{T}\left(2\sum_{i=1}^{N}\left(\left(\mathrm{Std}(\mathbf{x}_{0i})\right)^2\right)+N\sigma^2\right)} \\
=&\sigma\sqrt{\frac{N+1}{T}\left(2\sum_{i=1}^{N}\left(\rho^2\mathbf{h}_i^2\right)+N\sigma^2\right)} \\
=&\sigma\sqrt{\frac{N+1}{T}\left(2\rho^2||\mathbf{h}||_F^2+N\sigma^2\right)} \\
=&\sigma\sqrt{\frac{N+1}{T}\left(2N\sigma^2\mathrm{SNR}+N\sigma^2\right)} \\
=&\sigma\sqrt{\frac{N(N+1)\sigma^2}{T}(2\mathrm{SNR}+1)} \\
=&\sigma^2\sqrt{N(N+1)\frac{2\mathrm{SNR}+1}{T}}.
\end{aligned} \tag{2.63}$$

For large $N \gg 1$ the limit reduces to

$$\begin{aligned}
\mathrm{E}[||\Delta\mathbf{C}||_F] \lessapprox & \sigma^2\sqrt{N^2\frac{2\mathrm{SNR}+1}{T}} \\
=&\sigma^2 N\sqrt{\frac{2\mathrm{SNR}+1}{T}}.
\end{aligned} \tag{2.64}$$

This formula gives an estimation of the expected upper boundary of the covariance error. Increasing the noise level σ^2 increases the error, which was to be expected. Increasing the number of sensors increases the cross talk and therefore the errors. Unexpectedly, the covariance matrix error for increases with the SNR. The relative covariance error $\frac{||\Delta\mathbf{C}||_F}{\mathrm{SNR}}$, however, decreases because

$$\frac{||\Delta\mathbf{C}||_F}{\mathrm{SNR}} \approx \frac{\sigma^2 N\sqrt{\frac{2\mathrm{SNR}+1}{T}}}{\mathrm{SNR}} \sim \frac{\sqrt{\mathrm{SNR}}}{\mathrm{SNR}} = \frac{1}{\sqrt{\mathrm{SNR}}}. \tag{2.65}$$

Increasing the number of time samples decreases the error. As T is the only parameter which can be strongly influenced, large time samples are surely to be preferred.

2.4 Synthetic Aperture Magnetometry (SAM)

In this section the Synthetic Aperture Magnetometry (SAM) approach to beamforming is described as done by Sillekens [16]. SAM differs to LCMV in the computation of the direction. While LCMV constructs a filter with a vector output from which source direction can be estimated, SAM computes the direction before constructing the filter. The filter output is then the moment of the

source in the precomputed direction and therefore a scalar value.

In this case the filter $\mathbf{W}(\mathbf{q})$ is no longer a $N \times 3$ matrix but a $N \times 1$ vector.

Let the dipole orientation at a location $\mathbf{q}$ be given by the 3×1 vector $\mathbf{e} = \mathbf{e}(\mathbf{q})$, then the dipole moment at $\mathbf{q}$ can be expressed as $\mathbf{m}(\mathbf{q}) = a\mathbf{e}$ for a scalar value a. Omitting the arguments $(\mathbf{q})$ and $(\mathbf{x})$ in this section, the $N \times 3$ leadfield matrix $\mathbf{H}(\mathbf{q})$ reduces to a $N \times 1$ matrix

$$\mathbf{h} = \mathbf{H}\mathbf{e} \tag{2.66}$$

and the spatial filter in (2.25) simplifies to

$$\mathbf{w}^T = \frac{\mathbf{h}^T\mathbf{C}^{-1}}{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}}. \tag{2.67}$$

Like in (2.27) the estimated spatial spectrum of the data can be written as

$$\widehat{\mathrm{Var}}(\mathbf{q}) = \frac{1}{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}}. \tag{2.68}$$

Unlike in the LCMV approach this is a scalar value, so the trace does not need to be considered.

As in the LCMV approach the projected sensor noise can be described. It is again assumed that the noise is distributed with zero mean, variance σ^2, and uncorrelated for all sensors. Then the noise covariance matrix $\mathbf{Q}$ can be written as

$$\mathbf{Q} = \sigma^2\mathbf{I}. \tag{2.69}$$

The noise spatial spectrum is now obtained by applying the spatial filter $\mathbf{w}$ to the noise covariance matrix (2.69)

$$\eta^2 = \mathbf{w}^T\sigma^2\mathbf{I}\mathbf{w} = \sigma^2\mathbf{w}^T\mathbf{w}. \tag{2.70}$$

In order to write (2.70) in full detail, note that

$$\begin{aligned}
\mathbf{w} = (\mathbf{w}^T)^T &= (\mathbf{h}^T\mathbf{C}^{-1}(\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h})^{-1})^T \\
&= \underbrace{((\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h})^{-1})^T}_{=((\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h})^T)^{-1}}(\mathbf{h}^T\mathbf{C}^{-1})^T \\
&= \frac{\mathbf{C}^{-1}\mathbf{h}}{(\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h})^T} \\
&= \frac{\mathbf{C}^{-1}\mathbf{h}}{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}}
\end{aligned} \tag{2.71}$$

and therefore

$$\eta^2 = \sigma\frac{\mathbf{h}^T\mathbf{C}^{-1}}{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}}\frac{\mathbf{C}^{-1}\mathbf{h}}{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}} = \sigma^2\frac{\mathbf{h}^T\mathbf{C}^{-2}\mathbf{h}}{(\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h})^2}. \tag{2.72}$$

Using (2.68) and (2.72) the pseudo-Z notation is given as

$$\widehat{\mathrm{Var}}_N = \frac{1}{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}} \frac{(\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h})^2}{\sigma^2\mathbf{h}^T\mathbf{C}^{-2}\mathbf{h}} = \frac{\mathbf{h}^T\mathbf{C}^{-1}\mathbf{h}}{\sigma^2\mathbf{h}^T\mathbf{C}^{-2}\mathbf{h}}. \tag{2.73}$$

The next step is to interpret (2.73) as a function of the dipole orientation $\mathbf{e}$. As $\mathbf{H} = \mathbf{He}$, it is

$$\widehat{\mathrm{Var}}_N(\mathbf{e}) = \frac{\mathbf{e}^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{He}}{\sigma^2\mathbf{e}^T\mathbf{H}^T\mathbf{C}^{-2}\mathbf{He}} \tag{2.74}$$

$$= \frac{\mathbf{e}^T\mathbf{Ae}}{\sigma^2\mathbf{e}^T\mathbf{Be}}, \tag{2.75}$$

with $\mathbf{A} = \mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}$ and $\mathbf{B} = \mathbf{H}^T\mathbf{C}^{-2}\mathbf{H}$. As (2.74) is the generalized Rayleigh Quotient, the dipole direction maximizing the pseudo-Z value is given by an eigenvector $\mathbf{e}_0$ corresponding to the biggest eigenvalue of $\mathbf{B}^{-1}\mathbf{A}$ and

$$\widehat{\mathrm{Var}}_N(\mathbf{e}_0) = \max_{||\mathbf{e}||=1} \widehat{\mathrm{Var}}_N(\mathbf{e}). \tag{2.76}$$

It should be noted, that the direction is invariant to the assumed noise strength σ^2 and thus it only serves as a scaling factor.

Chapter 3

Kurtosis and Beamforming

In chapter 2 the source strength was defined by its variance. In this chapter excess kurtosis is introduced as an alternative definition for source strength.

If the measurement is composed only of white noise and the source signal, the signal can be characterised simply by not being noise. Thus, a statistical outlier test like excess kurtosis can be used to estimate the source strength. Kurtosis will be described in the first section of this chapter. Then the sliding window will be introduced to extend the analytical possibilities of kurtosis as done by Prendergast et al. [11]. In section 3.3 kurtosis will combined with the SAM filter as first done by Robinson et al [13] and Kirsch et al. [8] and extended by Prendergast et al. [11].

3.1 Kurtosis and g_2

Kurtosis is a statistical method to measure the "tailedness" of a distribution in comparison to the normal distribution. The section is based on Westfall [20] and DeCarlo [3].

Let f be a distribution with expected value μ. The variance σ^2 is defined as $\mathrm{E}[(f-\mu)^2]$. This means that the variance is higher the more f differs from its mean. As $x^2 < x$ for $x < 1$ the variance favours larger deviation. The fourth moment μ_4 is analogously defined as $\mu_4 = \mathrm{E}[(f-\mu)^4]$ and can be considered a stronger weighted version of the variance. By squaring the variance and comparing it to the fourth moment a measurement of high deviation from the expected value is gained.

Kurtosis is defined as the ratio of μ_4 and σ^4, that is

$$KURT(f) = \frac{\mu_4}{\sigma^4} = \frac{E[(f-\mu)^4]}{E[(f-\mu)^2]^2}. \tag{3.1}$$

With Jensen's inequality the fourth moment is equal or larger than the squared variance. $KURT$ is therefore bound from below by 1. As the deviation gets

higher, the fourth moment grows faster than the variance and the quotient is not bound from above.

For every normal distribution $\mathcal{N}(\mu, \sigma^2)$ it is $\mu_4 = 3\sigma^4$, so

$$KURT(\mathcal{N}(\mu, \sigma^2)) = \frac{3\sigma^4}{(\sigma^2)^2} = 3. \tag{3.2}$$

For a better comparison to the normal distribution the 'excess kurtosis' is adjusted so the normal distribution has a value of zero. It is defined as

$$\gamma_2(f) = \frac{\mu_4}{\sigma^4} - 3 = \frac{E[(f-\mu)^4]}{E[(f-\mu)^2]^2} - 3. \tag{3.3}$$

Today both definitions are used and often simply referred to as kurtosis. In this work kurtosis will always mean the sample excess kurtosis described next.

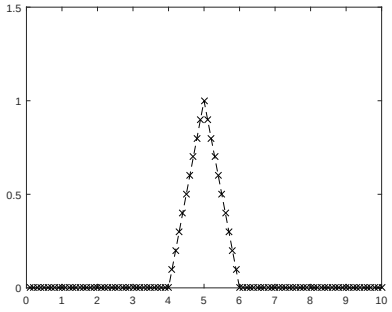
In practice only a sample of data can be measured. Then the distribution of the sample is derived and a discreet version of the excess kurtosis is applied. For a sample $x = (x_t)_{t=1}^T$ the sample (excess) kurtosis is defined as

$$g_2(x) = \frac{\frac{1}{T}\sum_{t=1}^{T}(x_t - \overline{x})^4}{(\frac{1}{T}\sum_{t=1}^{T}(x_t - \overline{x})^2)^2} - 3, \tag{3.4}$$

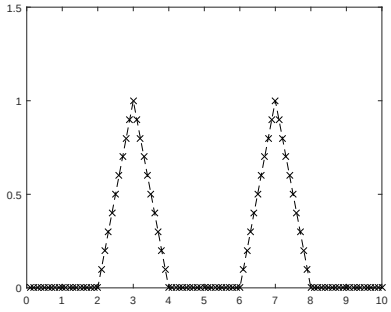
where $\overline{x}$ is the mean of x.

Note that the sample kurtosis is not a measure of distributions but of the data itself. If the data is supposed to be normally distributed, kurtosis can be used to test for outliers as high kurtosis is gained through sample values away from $\overline{x}$ [9]. This can lead to the data looking spiky, even though the its distribution may not. This is probably the reason why kurtosis is described as a measurement of peakedness in many articles. Westfall [20], however, has shown that the description of kurtosis as peakedness is false as the center has only a small influence on kurtosis and minimal changes in the center shape of the distribution can drastically change kurtosis.

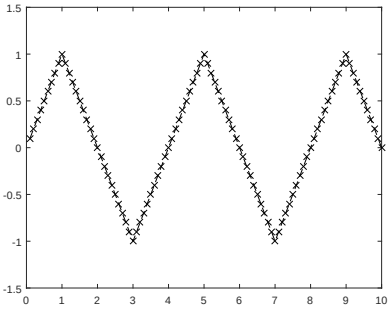
Therefore, a clear separation of sample kurtosis and kurtosis should be kept in mind.



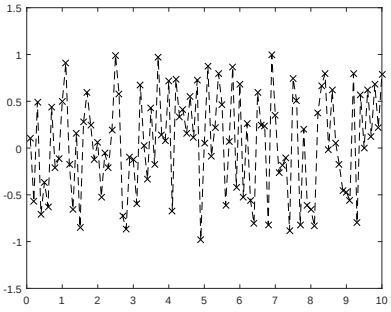
(a) 1 spike yielding a $g_2$value of 4.44



(b) 2 spikes yielding a $g_2$value of 0.16



(c) Oscillation yielding a $g_2$value of -1.08



(d) Noise between 1 and -1 with a $g_2$value of -1.09

Figure 3.1: g_2 values for different examples

3.2 Metrics for Brain Research

As it can be seen in figure 3.1 kurtosis favours high spikes and punishes high frequency. This can be a disadvantage in brain research when sources with a low amplitude and high frequency should be separated from high amplitude, low frequency sources. Segmentation of the measured data into small epochs could solve this problem but would negate the positive effects of a long measurement time discussed in section 2.3. Segmentation of the filtered data, however, can combine both positive effects and change the bias of kurtosis to favour frequency. This method was introduced by Harpaz et al. [4] is described in this section.

A sliding window of length n is a separation of a sample into overlapping segments of length n. In practice the last segments are reduced to a length smaller than n so sample length is not required to be a multiple of n. The g_2 value of each interval is computed individually and then combined to one value with a chosen metric.

To find the highest spike regardless of frequency, the maximum of the values is taken. In mathematical terms we get for a sample $x = (x_k)_{k=1}^K$ and a window of length n

$$g_{2max}^n = \max_k(\ g_2(x(1+kn/2)), \ldots, g_2(x(1+(k/2+1)n))). \tag{3.5}$$

If frequent, possibly weak spikes should be favoured over strong, single spikes, the sum metric can be used. By adding the values of the intervals, frequency is strongly favoured. Note that only positive values are added, so baseline and noise will not reduce signal strength. Using the same notation as above,

$$g_{2\text{sum}}^n = \sum_k\ [g_2((x(1+kn/2), \ldots, x(1+(k/2+1)n)))]_+. \tag{3.6}$$

The window length n should be chosen to separate the spikes of the source. However, as the window gets shorter kurtosis values decrease as spikes stop to be significant outlier. This will lead to a bias to high values that cannot be satisfied by realistic data and eventually to no detection of any activity. The high sample rate of EEG and MEG helps to reduce this effect as even small time intervals have a lot of time sample points, but the problem still persists. Harpaz et al. [4] suggest that a window length with one spike per window is optimal, but as spikes do not occur regularly and timing is often unknown, optimal window length must be gained heuristically.

3.3 SAM(g_2)

SAM(g_2) combines the filter of the SAM approach (2.67) with the g_2 formula (3.4). It was described and tested by Robinson[13] et al. and Kirsch et al. [8]. Using a filter $\mathbf{w}$ an estimation of the of the dipole moment $\mathbf{m}$ in direction $\mathbf{e}$ at $\mathbf{q}$ is gained by

$$\widehat{\mathbf{m}}(\mathbf{q}) = \mathbf{w}(\mathbf{q})^T\mathbf{x}. \tag{3.7}$$

Using this for all locations $\mathbf{q}$ and for all samples $\mathbf{x} = \mathbf{x}(1)\ldots\mathbf{x}(T)$, SAM virtual sensors are constructed that estimate the source waveforms. At each location the spike activity can now be evaluated with g_2. For a location $\mathbf{q}$ the estimated epileptic activity is

$$activity(\mathbf{q}) = g_2(\mathbf{w}(\mathbf{q})^T\mathbf{x}). \tag{3.8}$$

As tested by Harpaz et al [4] and described in section 3.2, the performance of g_2 and therefore SAM(g_2) can be changed by a sliding window. Using the notation as in chapter 3.2, the activity at location $\mathbf{q}$ with a sliding window of length n is

$$activity_{sum}^n(\mathbf{q}) = g_{2\text{sum}}^n(\mathbf{w}(\mathbf{q})^T\mathbf{x}) \tag{3.9}$$

or

$$activity_{max}^n(\mathbf{q}) = g_{2\text{max}}^n(\mathbf{w}(\mathbf{q})^T\mathbf{x}) \tag{3.10}$$

for the sum and the maximum metric, respectively.

Chapter 4

Test and Discussion

In this chapter SAM, SAM(g_2), and the metric beamformer are tested with synthetic EEG and MEG data. Therefore, the model and software used in the simulations is described and general information to the test environments are given. Then the results are shown and discussed.

4.1 Model

For simulating the sources a realistic head model is used. After written informed consent had been obtained, MR images of a healthy 25-year-old male were taken and analysed. Using T1- and T2-weighted images the head was separated into skin, skull compacta, skull spongiosa, cerebrospinal fluid (CSF), cerebral gray (GM) and white matter (WM), cerebellar gray (cGM) and white matter (cWM), brainstem, eyes, and neck muscles. These surfaces were combined into one mesh. As the modelling is not the focus of this work, the exact procedure will not be mentioned here, but can be found in [12]. For more information on the model [10] and [7] are recommended. The conductivities used for simulating the sources are given in figure 4.1. Anisotropy was not used in the simulations.

	6 Compartment	**3 Compartment**
Skin	0.43	0.43
Compacta	0.007	0.01
Spongiosa	0.025	0.01
CSF	1.79	0.33
GM	0.33	0.33
WM	0.14	0.33

Figure 4.1: Conductivities for the 6 and 3 compartment model

4.2 Software

The synthetic data was simulated using SimBio [17]. The SimBio code provides algorithms to solve the EEG/MEG inverse and analytical and quasi-analytical solutions to the forward problem. The subtraction FEM method of SimBio was used to generate the leadfields and dipole output.
All beamformer algorithms were implemented in MATLAB[6]. Visualisation of the data was done in MATLAB, too.

4.3 Methods

4.3.1 Grid

For source locations a $1325 \times 1715 \times 1175$ mm^3 grid with a width of 5 mm was created. Every point was tested whether it was inside the pial surface and rejected, if it it was not. This lead to 8694 source locations for which leadfields were computed.

4.3.2 Signal

If not stated otherwise, all measurements were simulated using EEG and 80 sensors. Location of the sensors are shown in figure 4.3 The signal was assumed to be measured at 1200 Hz, that is 1200 samples per second. The dipole strength was set to 0, 10, 0, -5, 0 mAm over a time of five samples. This leads to a roughly realistic shape seen in figure 4.2. The dipoles were simulated in the 6 compartment model. This means that the model error was zero for the 6 compartment model. The source dipoles are not grid points to prevent an inverse crime and the dipole used in all tests has a minimal distance to the grid of 2.1mm. Achieving this localization error will be denoted as optimal or perfect.

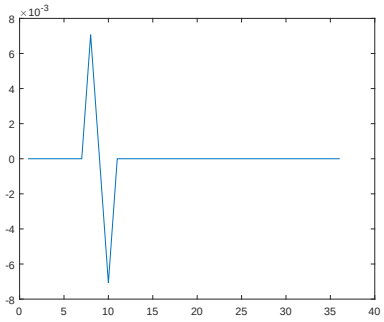


Figure 4.2: A spike measured at the 80th EEG sensor

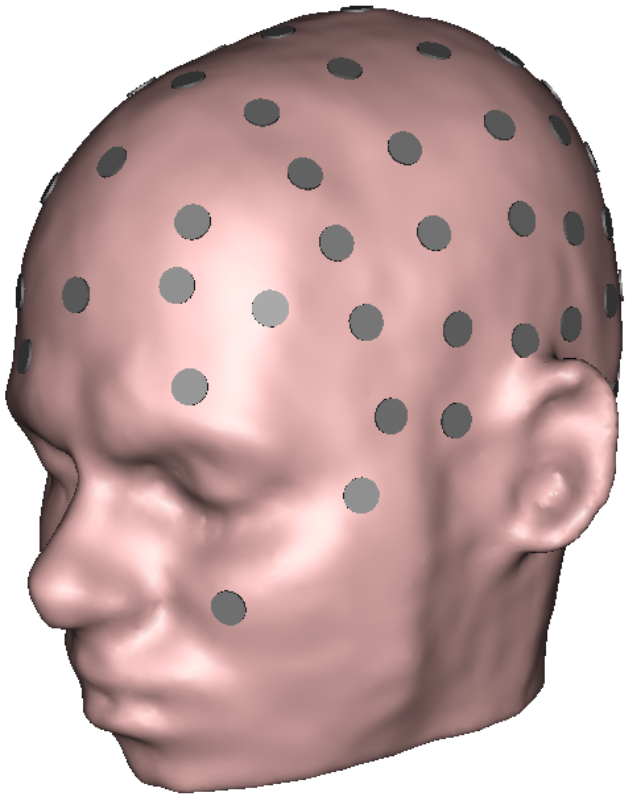


Figure 4.3: The EEG sensors on the head

4.3.3 Noise

Sensor noise was created using the MATLAB random generator randn() that creates normally distributed numbers. For every sensor a vector $\mathbf{n}$ of noise was created with

$$\begin{aligned} |\text{mean}(\mathbf{n})| &\leq 0.1, \\ |\text{Var}(\mathbf{n}) - 1| &\leq 0.1, \\ |g_2(\mathbf{n})| &\leq 0.1, \end{aligned}$$

practically fulfilling the assumptions made in chapter 2. This vector was then multiplied with the square root of the noise strength to gain noise of proper variance.

4.3.4 Window Length

The window length for the metric beamformers was set to 1/4, 1/2, 1, 2 seconds, that is 300, 600, 1200, 2400 samples, to test for optimal length. For better readability in the legends they are denoted as sum(window length) and max(window length) instead of the notation used in chapter 3.

4.3.5 Trials

To see the influence of noise every test had 10 trials per parameter. The location error given in the result sections is the mean of these trials. Values in the visualisation are the mean of the trials, too. The single trial value did not change much for the given figures.

4.3.6 Visualisation and Interpretation

As described in chapter 2 the filter of the beamformer is not perfect and the output value is always a combination of the actual activity at a point and non-filtered activity of other locations. Thus, there can't be a criterion which value represents an active source, but all values and locations must be interpreted together to separate noise and source activity. The global maximum of the output will be the reference for the localization error in the analysis. For visualisation the values of one trial are linearly mapped to the MATLAB default colormap *parula* which can be seen in figure 4.4. This means that every plot has its own scale and only qualitative distribution of power but not numeric values can be seen.

The dipole used will be marked with a red arrow pointing to the location of the dipole in the direction of the dipole moment. In section 4.6 another dipole will be added and displayed in magenta.

When SNR is measured it is always the SAM-beamformer SNR, that is the variance of the source divided by the variance of the noise. It was rounded to one digit after the decimal point in dB.

Figure 4.4: Colorbar for all brain plots from low to high activity

4.4 Noise test

To test how robust to noise SAM, SAM(g_2), SAM($g_{2\text{sum}}$), and SAM($g_{2\text{max}}$) are, a 1 minute EEG with 16 activations of one dipole has been created. The activations have been regularly spread over the time. Noise was scaled to have a variance of 10^{-13} to 10^{-1}V^2 resulting in SNR between 8,3dB and -3,7dB.

4.4.1 Results

The localization error is shown in fig 4.5. The differences in sliding window length are shown in fig 4.6.

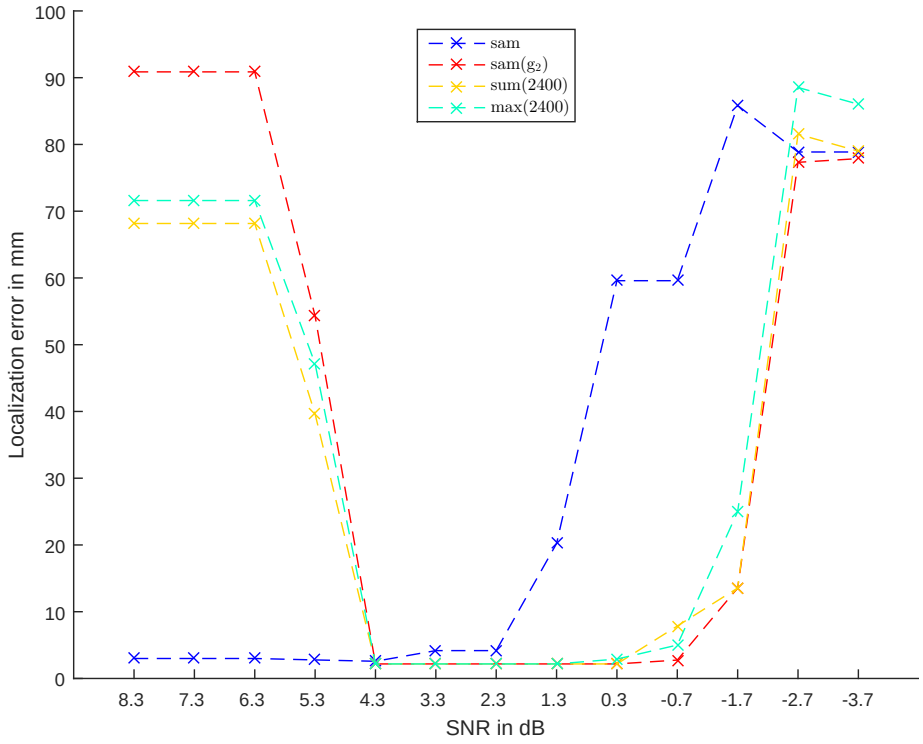


Figure 4.5: Localization error at different noise levels

At SNR between 8.3dB and 5.3dB g_2 approaches had large errors due to differences in noise. While the source was localized right in 4 and 3 trails for SAM(g_2) and the metrics, other trials had errors bigger than 70 mm. SAM's error was low to minimal, choosing the right voxel or an adjacent one.

While the poor performance of the kurtosis beamformer at high SNR seems counterintuitive at first, it shows both the performance of the beamformer in regard to noise and the difference between variance and kurtosis. At high SNR the difference between a grid point's leadfield and the measurement cannot be explained by noise. Therefore, the true source's activity is mainly filtered for the location. The filter output is only a measurement of low source activity and noise. This activity increases the variance of the output but does not form significant outliers. Therefore, SAM can detect this activity, but kurtosis values are too low to function reliably. To handle this problem a finer grid can be used to decrease the differences in leadfields. [5]

At medium SNR more grid point's leadfield can partially explain the measurement. By "picking" the suitable part of the measurement and declaring the rest as noise, the filter output contains the source activity at the price of adding noise. By measuring variance SAM rewards containing noise if amplitude can be high. In contrary to variance kurtosis favours outlier and therefore rewards low noise more than amplitude.

This explains both the poor performance of SAM at medium SNR and the wide spread of the kurtosis peaks shown in figure 4.7.

Between 4.3dB and 0.3 g_2 approaches performed very good with only minor error at higher noise. The classic SAM(g_2) performed best, followed by the sum metric. The maximum metric was a little less robust to noise and showed higher error at -0.7dB. At noise level larger than this, all approaches gained large error.

The length of the sliding window had some impact on the error gain at the border between middle and large noise and broadness of the kurtosis. Larger windows performed better in regard to noise. The good performance in the middle noise levels correspond well to the values of kurtosis measured. For classic g_2, values at low error were greater than 6, up to 1944 at level 1.3dB, and otherwise smaller than 0.1. The same proportions were measured for the maximum metric and the sum metric, even though not as distinct.

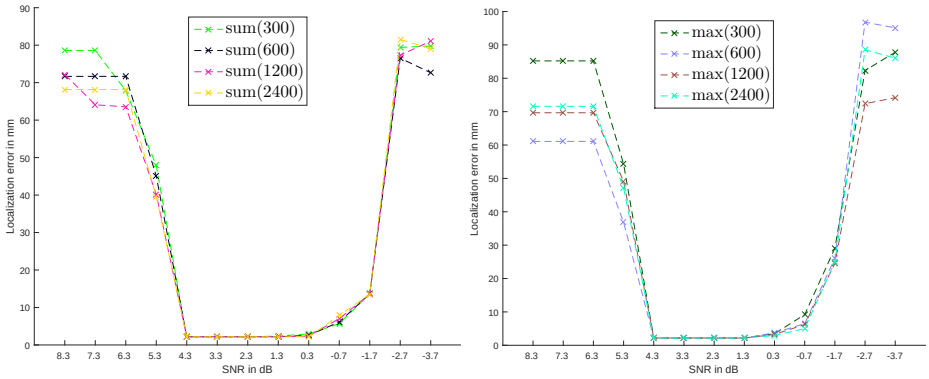


Figure 4.6: Localization error at different noise for different window lengths

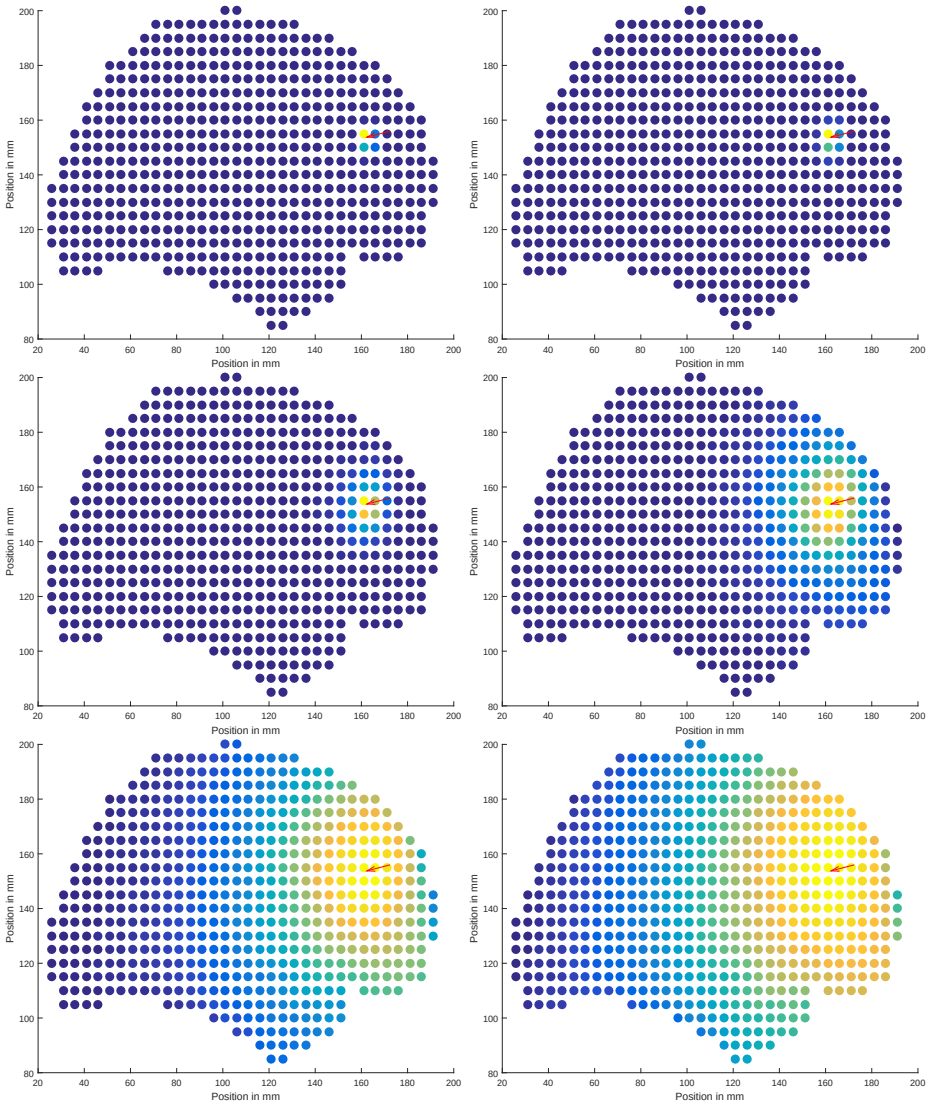


Figure 4.7: Comparison between the mean g_2 values at a SNR of 4.3 dB (top-left) to -0.7dB (down-right)

4.5 Frequency

In the next test the influence of signal repetition should be explored. At a noise level of $10^{-6}V^2$ (1.3 dB in section 4.4) the dipole will fire regularly 1,2,4,8, and 16 times per trial. Furthermore, 120 and 240 repetitions are simulated. These are frequencies of 2 and 4 spikes per second and should test whether high frequencies yield low kurtosis, as their mean increases.

4.5.1 Results

The localization errors are shown in figure 4.8 and figure 4.9. One signal has a rounded variance of $1,2 * 10^{-6}$ and 16 signals a variance of $18,7 * 10^{-6}$. At a noise level of $10^{-6}\mathrm{V}^2$ one signal has a SNR of similar to 16 signals at a level of $10^{-5}\mathrm{V}^2$ and both SAM and SAM(g_2) performed accordingly compared to figure 4.5.

For the metric beamformers, however, signal repetition is more important than noise. With a larger window lengths the sum metric performed only slightly worse than classic g_2, but short windows lead to high errors. As the metric is weighted towards repetition, this was to be expected. The maximum metric had relatively low errors at all window lengths, but again larger windows performed better and classic g_2 was more precise. As the metric was constructed so signal repetition is not weighted, this is somewhat disappointing.

At high frequencies the performance of the g_2-beamformer is practically the same with larger windows having a clearer peak. Note that in the figures the lines overlap and are therefore hard to separate. SAM was not able to localize the source correctly at high frequencies, making it again worse in comparison.

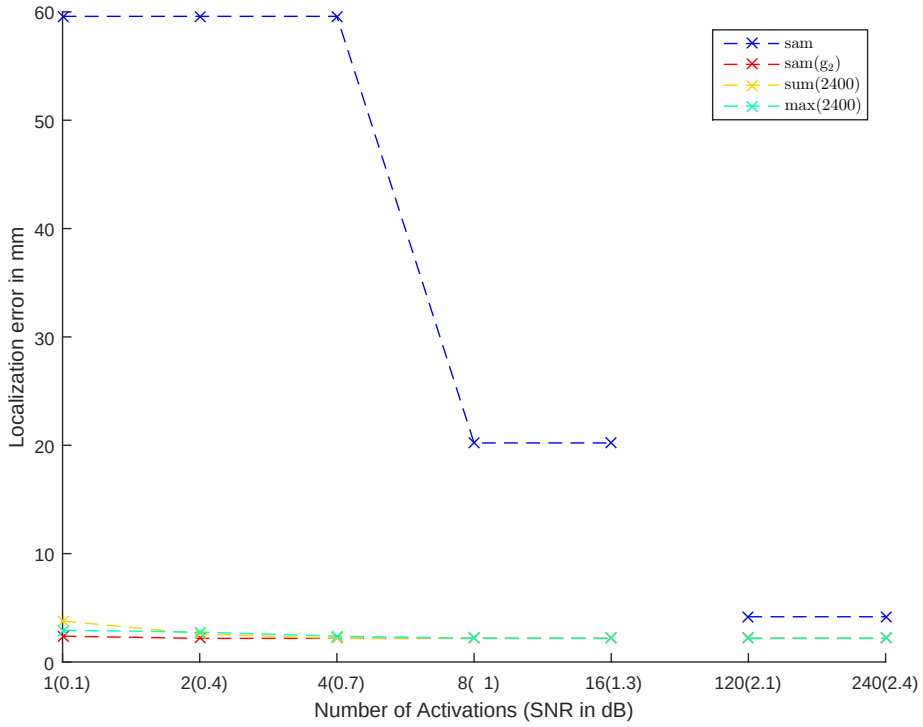


Figure 4.8: Localization error at different frequencies

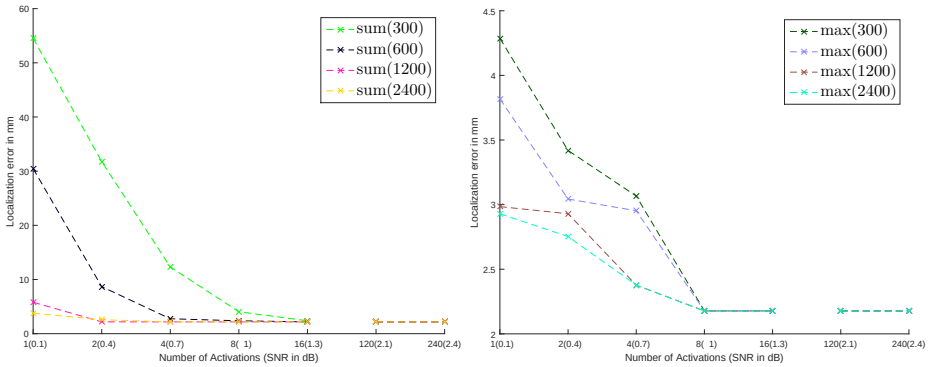


Figure 4.9: Localization error at different frequencies for different window lengths

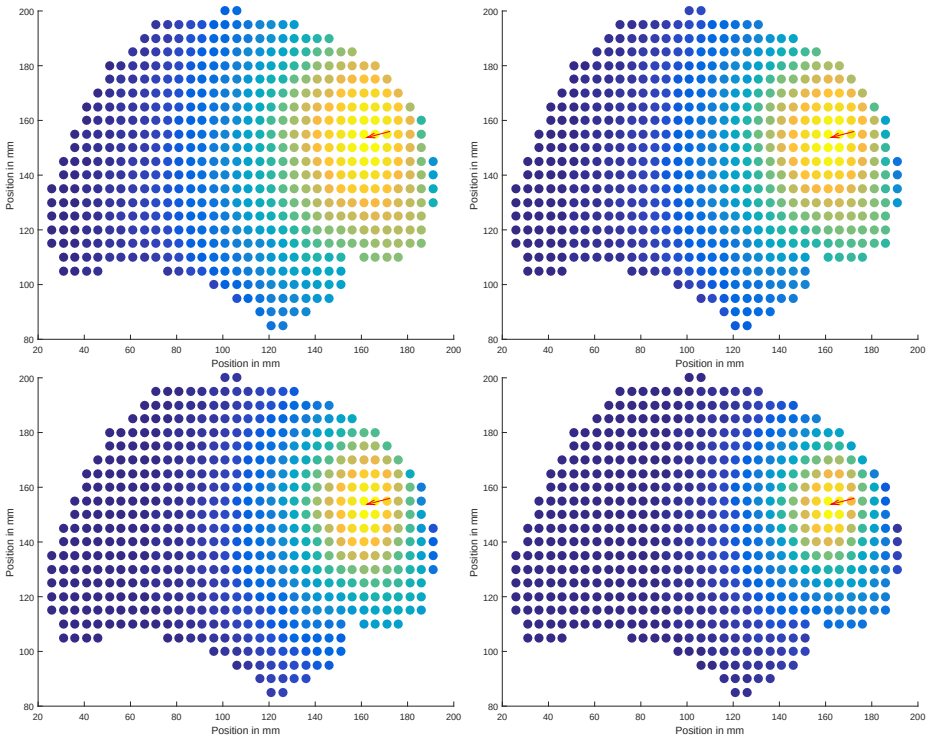


Figure 4.10: Comparison between the mean g_2 values at different frequencies from 1 signal (top-left) to 8 signals (down-right)

4.6 Multiple Dipoles

As introduced by Harpaz [4], the maximum and sum metrics are constructed to separate different active dipole sources. To test the performance at two different sources another dipole is added to the signal. As the second dipole is a deeper source its measurement was weaker. Thus, the maximum metric should favour the first one. The dipoles are not active at the same time and the mean of the signal is low, so correlation between the signals is weak. The noise level is again set to $10^{-6}\mathrm{V}^2$.

First, both dipoles fire 16 times in 1 minute. To test the possibilities of the sliding window the time differences between the signals are set to 50 samples (0.0417 seconds) and 2250 samples (1.8750 seconds).

In a second run the first dipole's fire rate is reduced to 8 signals to test how frequency and fire rate are weighted by the beamformers.

4.6.1 Results

The results of SAM(g_2) can be seen in figure 4.11. At the same signal frequency both dipoles are seen and reach a local maximum. Interestingly, lowering the frequency strengthened the signal and the second dipole was only weakly visible at 8 to 16 signal fires. At 16 activations the g_2 value at 50 samples difference were 1944 for the first and 1773 for the second. At 8 and 16 activations the value of the second dipole was 1772, remaining the same. The first dipole doubled its strength to 3578. So while the dipoles may not be clearly visible in the plot, looking for local maxima with high kurtosis would still localize both targets.

The time differences between the signals did not change the performance for SAM(g_2) as both the filter and kurtosis are not influenced by signal timing as long as signals are not timely correlated.

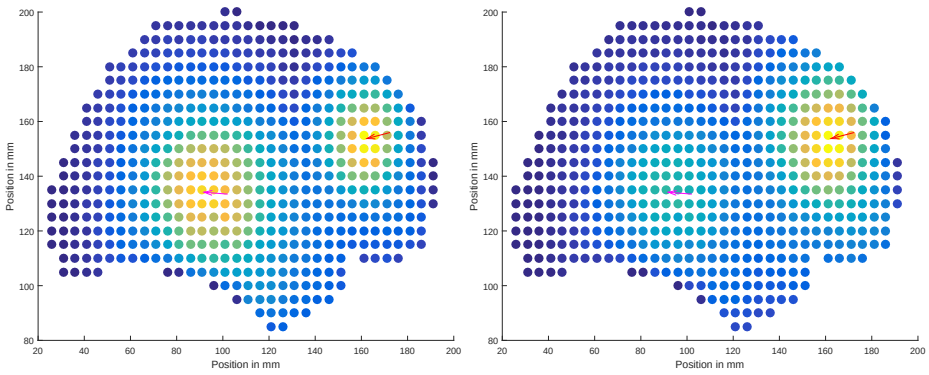


Figure 4.11: Comparison of g_2 values at 16 to 16 signals (left) and 8 to 16 signals (right)

At the same frequency the sum metric performed worse than SAM(g_2), spreading the kurtosis value over a bigger area. After reducing the fire rate the large 2 second window localized the high frequency target correctly, even though the peak of values was still broad. At this window length the sum metric performed as the counterpart to normal g_2, which was exactly what was hoped for.

Shorter windows had much broader peaks, spreading high values over large areas. The 0.25 second window could not separate the signals, marking a broad area between the source locations. Unexpectedly, the short time difference between signals yielded better outcomes, contradicting Prendergast's [11] one-spike-per-window criterion about optimal window length at least for spikes from different locations.

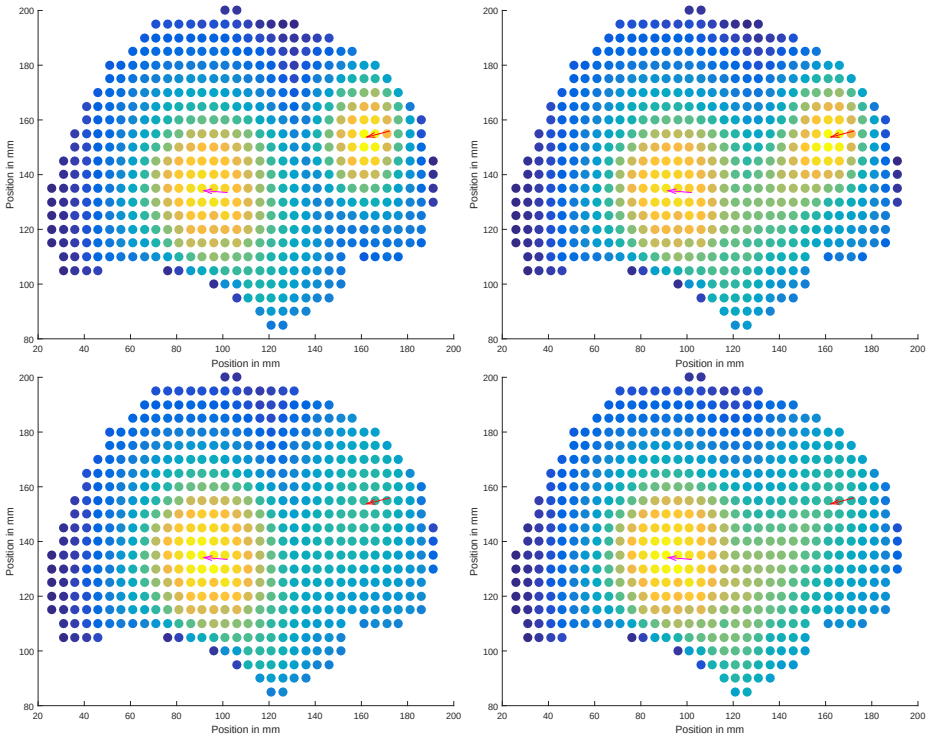


Figure 4.12: Comparison of sum metric at the large window. 16 to 16 signals on the top and 16 to 8 signals on the bottom. 50 samples difference on the left, 2250 samples on the right.

The maximum metric with a large window performed unbiased in comparison, marking both spots at all frequencies and time differences. Reducing the fire rate broadened the area around the first (red) dipole, but as both locations are clearly visible the maximum metric can be valuable to determine possible locations. Shorter windows yielded much broader areas, making the large window favourable. The peaks in the values, however, were at the correct voxels even at short windows. This continues the trend of the maximum to be more robust to short windows, but large windows are still preferable.

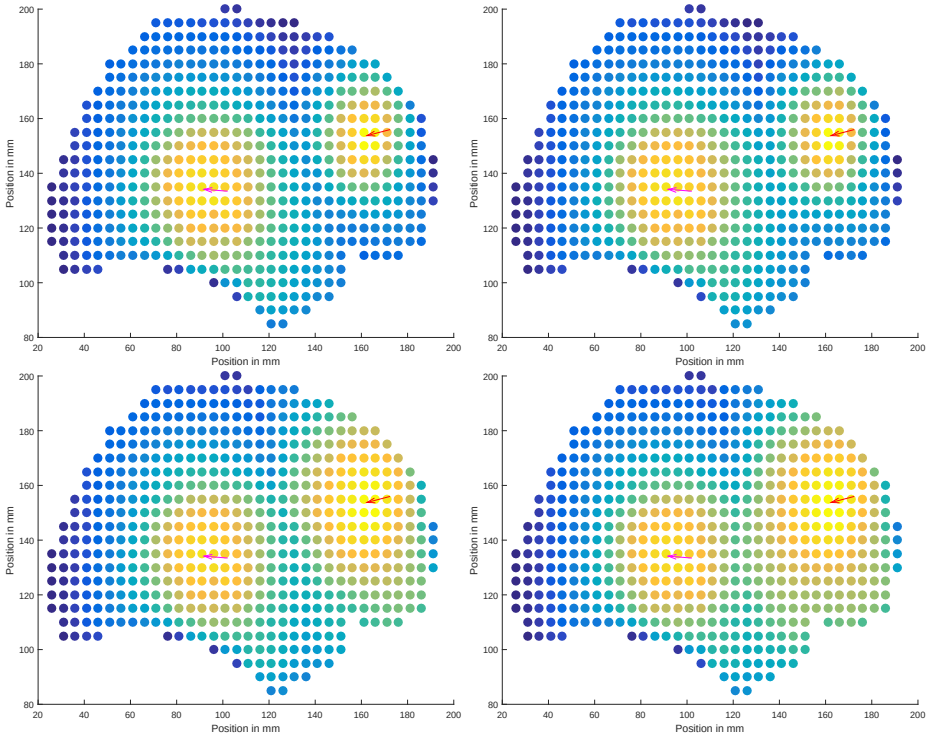


Figure 4.13: Comparison of maximum metric at the large window. 16 to 16 signals on the top and 16 to 8 signals on the bottom. 50 samples difference on the left, 2250 samples on the right.

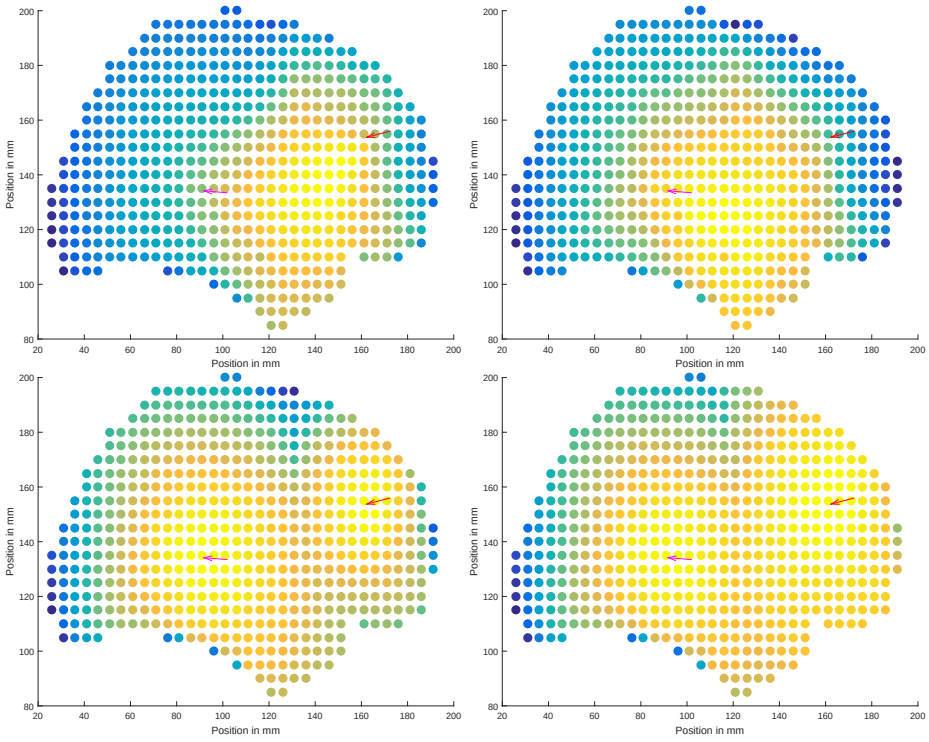


Figure 4.14: g_2 values of the sum and maximum metric at the 1/4s window and 2250 samples time difference. The sum metric is shown in the first row, the maximum in the second. Values for 16 to 16 signals on the left, 8 to 16 signals on the right.

4.7 3 Department Model

In the previous tests the same model was used to generate the signals as well as the leadfields. As this yields no model error, the influence of modelling should be tested. For this, the first test 4.4 is evaluated again using leadfields computed with the three compartment model described in section 4.1. Grid, signal, and noise are reused.

4.7.1 Results

The localization error can be seen in figures 4.15 and 4.16. The algorithms performed similar with the three compartment leadfield, but got less robust to noise. While the source was localized correctly sometimes, none was reliable at any noise level. With small differences in noise the peak changed to adjacent voxel, adding small errors between 2,5 mm to 5,3 mm. As in the first test window length had little influence, but short windows did comparatively better than for

the correct lead field. At very high or low noise levels the errors were lower, but still too high for practical purposes. Kurtosis values are about the same for the leadfields. If the aim is only to search for epileptic activity without the need for correct localization, the model error is insignificant. To help choosing spikey segments out of long records these results can still be helpful.

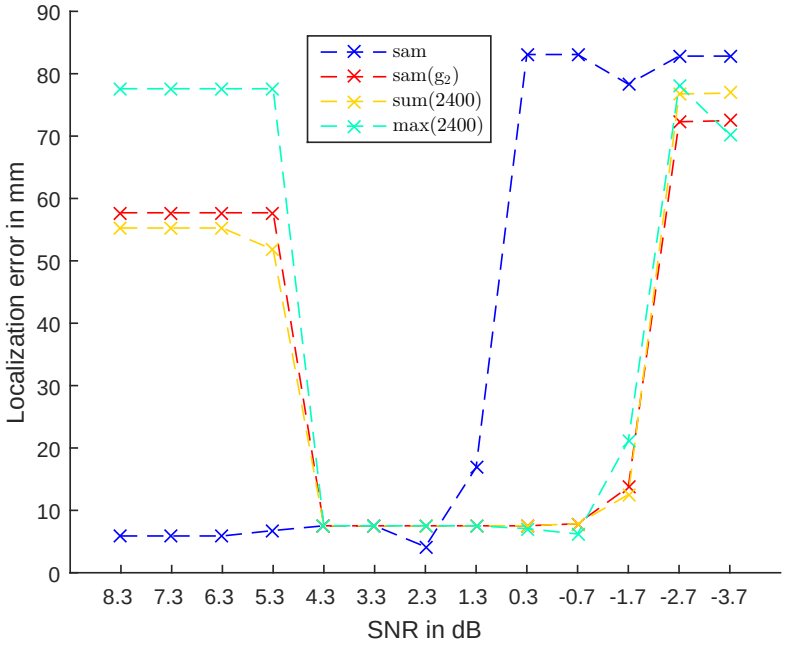


Figure 4.15: Localization error for the 3C model

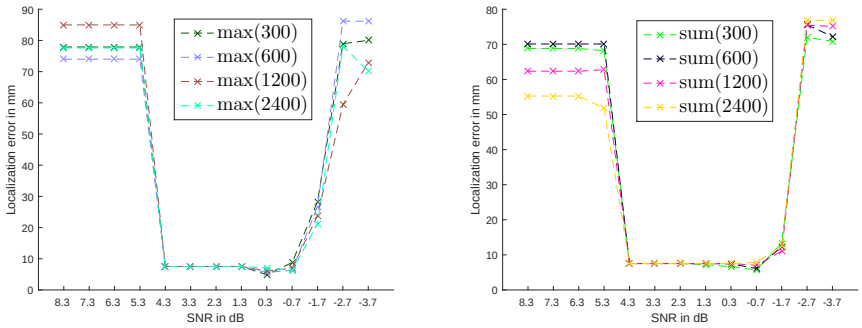


Figure 4.16: Localization error for the 3C model at different window lengths

4.8 MEG Differences

To test the differences between EEG and MEG the leadfields were computed for the MEG as well. The MEG sources and leadfields were computed for 273 gradiometers. Noise was computed as for the EEG, the grid points were not changed.
The signal amplitude was about ten times higher than for the EEG. This led to a higher signal strength of about 1000 times the EEG strength. At similar SNR the ratio of amplitude was thus less favourable for the MEG. Noise strength was between $10^{-10}\mathrm{T}^2$ and $10^{1}\mathrm{T}^2$.

4.8.1 Results

The localization error of the MEG analysis can be seen in figures 4.17 and 4.18. The MEG performed very similar to the EEG at similar SNR. At 0.2dB, 1.2dB, and 2.2dB the MEG kurtosis peak was steeper than for the EEG. For SAM(g_2) about 10 grid points had over 95% of the maximum kurtosis value at 0.2 dB for the MEG but about 24 at 0.3dB for the EEG.

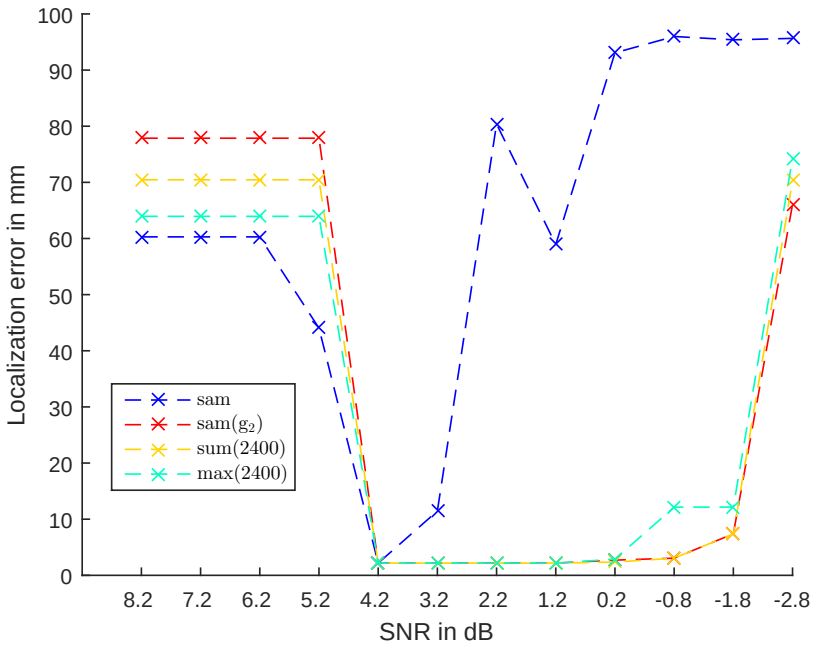


Figure 4.17: Localization error for the MEG

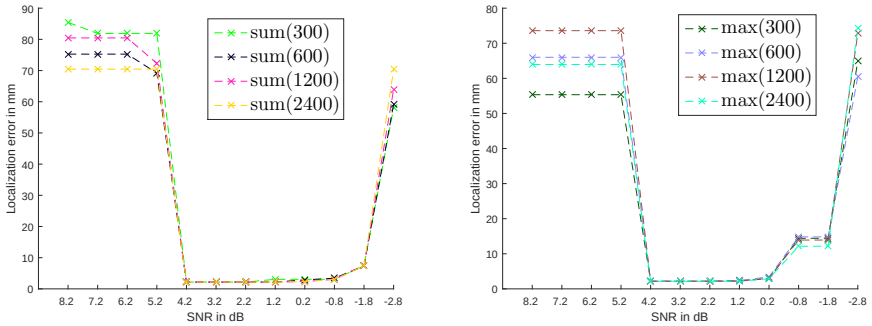


Figure 4.18: Localization error for the MEG at different window lengths

Using the three compartment model to compute the leadfields the same test was again evaluated. The results are shown in figures 4.19 and 4.20. The kurtosis beamformer performed again similarly to the EEG with an additional error of about 5 mm. SAM, however, was not able to localize the source at all. This contradicts the work of Steinsträter et al. [18] who found the MEG to be more robust to deficiencies in forward modelling. As only one source location was used in this work and Steinsträter has looked for the location with the highest leadfield error, this might just be luck at choosing the dipole location and orientation. This should be studied in future works especially for combining EEG and MEG.

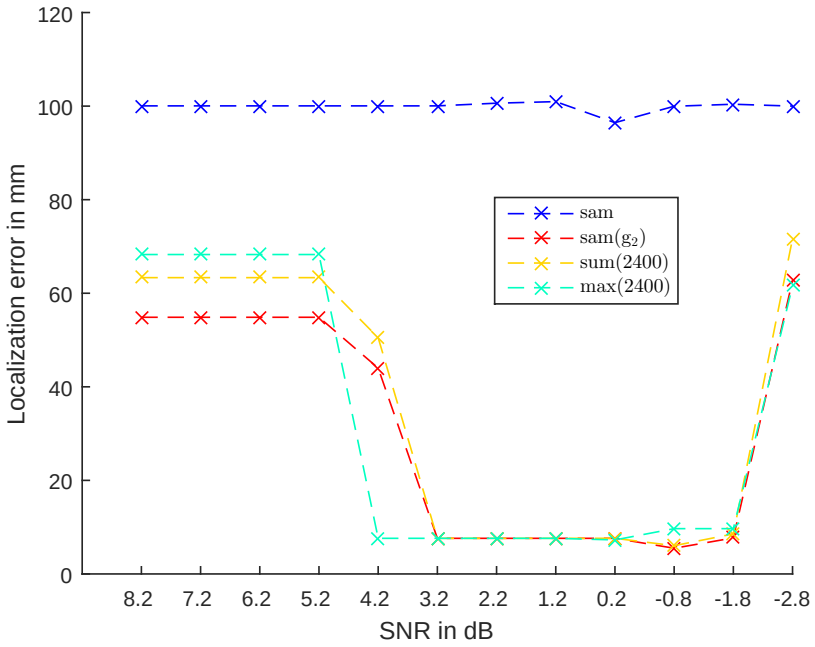


Figure 4.19: Localization error for the MEG 3C leadfield

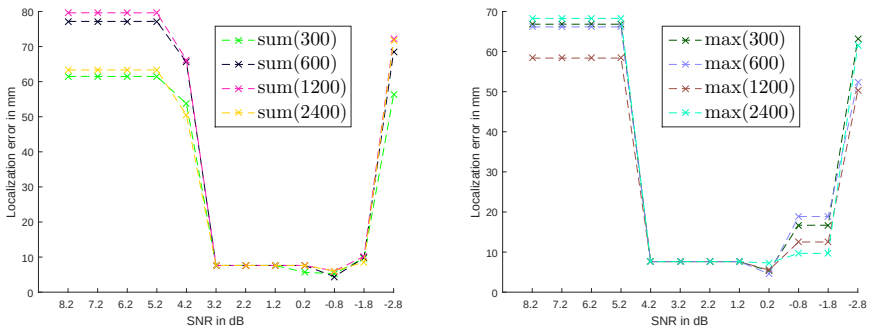


Figure 4.20: Localization error for the MEG 3C leadfield at different window lengths

Chapter 5

Summary and Discussion

The aim of this thesis was to compare the performance of different beamformer approaches. The approaches used the same filter and differed only in the analysis of the filter output. The methods were variance, g_2, maximum(g_2), and sum(g_2).

The kurtosis approaches have shown better performance when noise was not very low. This should be explained by the distance of the grid points to the source as the filter worked to well to detect activity. Using a finer grid should remove this disadvantage of kurtosis. Classic g_2 was more robust to noise than the metrics, but when the window length was large, this effect was rather low. Higher noise yielded higher kurtosis values and a much broader peak, but localisation was still correct when kurtosis values were high.
The variance based SNR is a good indicator for the performance of SAM and SAM(g_2) as the noise and frequency tests showed. Signal frequency has shown to strengthen the signal, but one spike is enough when SNR is sufficient.
This was not true for the metrics. As the variance of source and noise is estimated over the whole time and both metrics use further segmentation, this was somehow to be expected. The sum metric is biased to high frequent spikes and its failure to localize at low frequency is less important. The maximum metric, however, should not favour frequency and analyse signal amplitude. Its failure at low frequency is disappointing as it contradicts the concept.
When two dipoles were activated SAM could not localize any source. SAM(g_2) could localize both sources when signal frequency was the same. It highly favoured the stronger, less frequent source in the second trial as was to be expected by the work of Prendergast [11]. The sum metric could change this bias to favour the weaker, more frequent source. Therefore, the usefulness of the approach seems to be proven. The maximum metric could localize both spikes in the trials, making it useful as non-biased method for location.
Window length should be rather large, as short windows blur the results and can even merge two sources into one. The best analysis is to combine all three g_2 methods and look for medically plausible source locations, as false positive results have not shown for sufficient large window lengths.

The model error for both EEG and MEG was about 5 mm for the kurtosis

based beamformer, changing the peak of kurtosis to adjacent grid points. As kurtosis values stayed at the same level, it can surely be used to detect the existence of spikes even at imperfect localization. The model error for SAM was severe at the MEG, contradicting the normal notion that the MEG is more robust to model errors. As only one source position has been evaluated, this could, however, be due to this specific location and not a general trend.

In conclusion, kurtosis beamforming can surely help to point to epileptic activity in noisy data. Perfect localization needs a very good modelling of the head and a fine grid, but the three compartment model can be used to help doctors by marking time samples with high activity.

Chapter 6

Outlook

The relationship between grid width and SNR should be researched. Furthermore, a new definition of SNR containing signal frequency should be found to describe the performance of the metric beamformer and distinguish multiple dipoles. As the direction of the filter optimizes SNR, this might help to stabilize the kurtosis beamformer as Prendergast already suggested [11]. If kurtosis is used to describe source strength the filter minimising the kurtosis instead of the variance might yield better and more coherent results.

The combination of EEG and MEG shows good results for the classical approach and beamformer should follow to use the strength of both measurements. The similarity of EEG and MEG in our simulation makes hope that this combination can be done without many problems.

As theory's worth lies in the application, the next step should be to test the methods on real data. Only when the methods can actually help to find and localize epileptic spikes, they can be used in clinical context to help patients and doctors.

List of Figures

3.1 g_2 values for different examples 27

4.1 Conductivities for the 6 and 3 compartment model 29
4.2 A spike measured at the 80th EEG sensor 30
4.3 The EEG sensors on the head 31
4.4 Colorbar for all brain plots from low to high activity 32
4.5 Localization error at different noise levels 33
4.6 Localization error at different noise for different window lengths . 34
4.7 g_2 values at different noise levels 35
4.8 Localization error at different frequencies 36
4.9 Localization error at different frequencies for different window lengths 37
4.10 g_2 values at different frequencies 37
4.11 g_2 at two dipoles with different signal frequencies 38
4.12 Comparison of sum metric at large window 39
4.13 Comparison of maximum metric at large window 40
4.14 Sum and maximum metric short window 41
4.15 Localization error for the 3C model 42
4.16 Localization error for the 3C model at different window lengths . 42
4.17 Localization error for the MEG 43
4.18 Localization error for the MEG at different window lengths . . . 44
4.19 Localization error for the MEG 3C leadfield 45
4.20 Localization error for the MEG 3C leadfield at different window lengths 45

Bibliography

[1] J. Bendat and A. Piersol. Random data: Analysis and measurement procedures. *John Wiley and Sons*, 1986.

[2] Matthew J. Brookes, Jiri Vrba, Stephen E. Robinson, Claire M. Stevenson, Andrwe M. Peters, Gareth R. Barnes, Arjan Hillebrand, and Peter G. Morris. Optimising experimental design for meg beamformer imaging. *NeuroImage*, 2008.

[3] Lawrence T. DeCarlo. On the meaning and use of kurtosis. *Psychological Methods*, 2, 1997.

[4] Yuval Harpaz, Stephen E. Robinson, Mordekhay Medvedovsky, and Abraham Goldstein. Improving the excess kurtosis (g_2) method for localizing epieptic sources in magnetoencephalogric recordings. *Clin Neurophysiol*, 2014. `http://dx.doi.org/10.1016/j.clinph.2014.09.002`.

[5] Arjan Hillebrand, Krish D. Singh, Ian E. Holloday, Paul L. Furlong, and Gareth R. Barnes. A new approach to neuroimaging with magnetoencephalography. *Human Brain Mapping*, 25, 2005.

[6] The MathWorks Inc. MATLAB version R2015b (8.6.0.267246) , 2015.

[7] A.M. Janssen, S.M. Rampersad, F. Lucka, B. Lanfer, S. Lew, Ü. Aydin, C.H. Wolters, D.F. Stegeman, and T.F. Oostendorp. The influence of sulcus width on simulated electric fields induced by transcranial magnetic stimulation. *Physics in Medicine and Biology*, 58(14):4881, 2013.

[8] Heidi E. Kirsch, Stephen E. Robinson, Mary M. Mantle, and Srikantan S. Nagarajan.

[9] John H. Livesey. Kurtosis provides a good omnibus test for outliers in small samples. *Clinical Biochemistry*, 40, 2007.

[10] F. Lucka, S. Pursiainen, M. Burger, and C.H. Wolters. Hierarchical bayesian inference for the eeg inverse problem using realistic fe head models: depth localization and source separation for focal primary currents. *NeuroImage*, 61(4):1364–1382, 2012.

[11] Garreth Prendergast, Gary G.R. Green, and Mark Hymers. A robust implementation of a kurtosis beamformer for the accurate identification of epileptogenic foci. *Clinical Neurophysiology*, 124, 2012. `http://dx.doi.org/10.1016/j.clinph.2012.09.024`.

[12] S.M. Rampersad, A.M. Janssen, F. Lucka, Ü. Aydin, B. Lanfer, S. Lew, C.H. Wolters, D.F. Stegeman, and T.F. Oostendorp. Simulating transcranial direct current stimulation with a detailed anisotropic human head model. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):441–452, 2014.

[13] Stephen E. Robinson, S.S. Nagarajan, M Mantle, V. Gibbons, and Heidi Kirsch. Localization of interictal spikes using sam(g_2) and dipole fit. *Neurol Clin Neurophysiol*, 2004.

[14] N. Schmitz. Vorlesungen über Wahrscheinlichkeitstheorie. *Teubner Verlag*.

[15] Kensuke Sekihara, Srikantan S. Nagarajan, David Poeppel, and Alec Marantz. Performance of an meg adaptive-beamformer technique in the presence of correlated neural activities: Effects on signal intensity and time-course estimates. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 49, 2002.

[16] Stephanie Sillekens. Influence of volume conduction on beamformer source analysis in the human brain. 2008.

[17] SimBio Development Group. SimBio: A generic environment for bionumerical simulations. online, `https://www.mrt.uni-jena.de/simbio, accessedMar9`, 2016.

[18] Olaf Steinsträter, Stephanie Sillekens, Markus Junghoefer, Martin Burger, and Carsten H. Wolters. Sensitivity of beamformer source analysis to defiencies in forward modeling. *Human Brain Mapping*, 31, 2010.

[19] Barry D. Van Veen, Wim van Drongelen, Moshe Yuchtman, and Akifumi Suzuki. Localization of Brain Electrical Activity via Linearly Constrained Minimum Variance Spatial Filtering. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 44, 1997.

[20] Peter H. Westfall. Kurtosis as Peakedness, 1905-2014. R.I.P. *Am Stat.*, 2014.

[21] Bernard Widrow, Kenneth M. Duvall, Richard P. Gooch, and William C. Newman. Signal cancellation phenomena in adaptive antennas: Causes and cures. *IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION*, 30, 1982.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit über Quellenanalyse mit Beamforming selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken – auch elektronischen Medien – entnommen sind, wurden unter Angabe der Quelle kenntlich gemacht.

Münster, 14.03.2016

Ich erkläre mich mit einem Abgleich der Arbeit mit anderen Texten zwecks Auffindung von Übereinstimmungen sowie mit einer zu diesem Zweck vorzunehmenden Speicherung der Arbeit in eine Datenbank einverstanden.

Münster, 14.03.2016