Master's Thesis in Biomedical Engineering

# 3D Finite Difference Time Domain Simulations and Analysis of Thrombolytic Treatment for Microwave Measurements of Stroke Patients

Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden
Master Thesis SSYX04

Antti Stålnacke
anttis@student.chalmers.se

March 14, 2014

# Abstract

Stroke is a big issue for today's health care and society. It causes large direct costs for health care, large indirect costs for society and is also the fourth largest cause of disease burden measured in *disability-adjusted life years* (DALYs). Stroke means that the blood supply to parts of the brain is cut off or limited leading to reductions in cognitive or physical functions. Stroke is most often due to a blood clot, a so called *ischemic stroke* (IS), but can also in about 15% of patients be due to a bleeding, a so called *intracerebral hemorrhage* (ICH). The treatment paths differ between these two conditions and an earlier diagnosis than what's possible today would be beneficial. Medfield Diagnostics is developing a microwave based diagnostic tool for early diagnosis of stroke. This tool relies on a classification algorithm and one purpose of this thesis is to increase the understanding of how this algorithm handles microwave based data related to stroke. The other purpose is to develop a variable 3D microwave model of patients with ICH for *finite difference time domain* (FDTD) simulations to investigate how the ICH effects the transmissions and reflections of microwaves and also classification results.

The thesis consist of two parts. Part one is to look at data from patients with an IS while they receive a clot dissolver. The goal here being to see a correlation between the *National Institute of Health Stroke Scale* (NIHSS) score and a value output by the classification algorithm over time. The second part consists of using an FDTD solver to simulate microwave measurements on a large number of simulated patients. These simulations are then analyzed to observe how the classification algorithm handles different cases. In total 1000 patients are simulated of which 500 represented healthy patients and 500 are patients with ICH. The parameters that can be varied are head size , hair thickness, bleeding position and bleeding size.

A correlation between NIHSS and output of the classification algorithm could not be seen with any of the approaches used in this thesis. Therefore a number of other approaches are suggested in the future work section. The simulation model and the resulting simulations showed promising results for being a valid model of the clinical measurement of the patient. The size of the bleeding together with the thickness of the hair had a large impact on the outcome of the classification. The best classification was also expected to improve with added number of patients. It was found that this was not true for patient sets smaller than 200 but true once the set became larger and the best classification for the whole data set had a maximum *area under curve* (AUC) value of 0.93.

# Acknowledgments

Good work is seldom carried out in solitude and this is no exception. I would like to thank my supervisor Stefan Candefjord for his daily support and inspiration. I would like to thank my examiner Andreas Fhager for providing the FDTD-solver used in my thesis and for helping out when I got lost in the world of Debye parameters and Cole-Cole models. I would like to thank the people at Medfield Diagnostics: Dag Jungenfelt, Fabian Wenger, Lennart Nilsson, Christoffer Sundström, Jasila Prabahar, Tobias Olsson and Samuel Lunde for providing a nice work environment. A special thanks goes out to Johan Klingstedt and Gustav Mårtensson for taking the time to explain the classifier in detail and for answering questions on a wide array of topics. I would like to thank Mahmood Qaiser for providing me with the head model. The computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at Chalmers Centre for Computational Science and Engineering (C3SE) for which I'm thankful since it would have taken around 37 years to do the computations on my laptop.

This thesis is not only a product of the last six months, it is also the final chapter of my education at Chalmers Univeristy of Technology. Therefore I would also like to thank the friends that I have studied, laughed, played, danced, climbed and lived with for the last six years. I have realized that you can learn as much from studying with a group of dedicated people as you can from lectures.

Finally I want to thank my family. For providing me with an upbringing void of boundaries and limitations and for providing support when I have needed it.

# Contents

# Chapter 1

# Introduction

## 1.1 Background

### 1.1.1 Burden and current treatment of stroke

A stroke is defined as an acute neurological deficit with a vascular basis, meaning that the blood supply to parts of the brain is cut off or limited leading to reductions in cognitive/mental or physical functions. To be defined as a stroke the symptoms also have to persist for more than 24 hours [1]. Stroke is the second cause of death after ischemic heart disease and a major cause of disability world wide. It is the fourth largest cause of disease burden measured in *disability-adjusted life years* (DALYs) in the whole population and the second largest cause of both lost DALYs and death in the age group 45–69 years [2].

Worldwide stroke accounts for around 4% of the total direct health care costs [3] and the indirect costs due to extensive rehabilitation and loss of labor are also significant. For Sweden Ghatnekar *et al.* [4] investigated the costs associated with stroke during the full lifetime of patients and found the total cost, including indirect costs (social services, outpatient visits, rehabilitation, production loss etc.), to be 12.3 billion Swedish kronor (SEK) per year. For the direct part of this cost 45% was found to be social services like institutionalized living. Only around 15% of the direct cost is the first, acute, hospital visit. Persson *et al.* found the total costs to be 3.7 billion SEK per year when only taking into account the cost during the first year after the stroke [5]. Due to this novel ways of treating and diagnosing stroke are of both great interest and importance.

There are two main types of stroke. The first type is *ischemic stroke* (IS), which is due to a blood clot blocking a vein or artery, and thereby the blood

supply to parts of the brain. The second type is stroke due to *intracerebral hemorrhage* (ICH).This is due to the rupture of a vein or artery in the brain leading to a loss of blood supply to parts of the brain. The percentages reported for the two main subtypes varies slightly between different sources but a span of 72–90% for IS and 10–20%[1] ICH shows that IS is more common [1], [6], [7] .

Differentiating between an IS and an ICH is important since the paths of treatment differ greatly [1]. One of the most important differences is the use of *thrombolytic treatment* (TT), often called clot dissolvers. In the case of an ischemic stroke TT has been found to be beneficial if administered within 4.5 hours after stroke onset since it dissolves the clot that is responsible for the stroke. The use of TT for patients with an IS is also suggested by The Swedish National Board of Health and Welfare (Socialstyrelsen) [8], [9]. However if TT is given to a patient with an ICH it has severe negative effects due to increased bleeding.

Despite the fact that the majority of strokes are IS only around 1–8% [10] of all stroke patients receive TT. One of the main reasons for this is that the differentiation between IS and ICH cannot be made within 4.5 hours. There are several reasons for this but one main issue is that the diagnostic tools available today are *computer tomography* (CT) and *magnetic resonance imaging* (MRI). Both tools require the patient to reach the hospital before an ICH can be ruled out and the time from stroke onset to reaching the hospital often exceeds 4.5 hours. Special stroke ambulances equipped with a CT have been developed but they are expensive and not suitable for other kinds of ambulance missions beside stroke [11].

### 1.1.2 Medfield Diagnostics' solution

Cheaper and more mobile tools for diagnosis are needed if more people are to receive TT, thereby reducing both the costs and suffering associated with stroke. A diagnostic method that has garnered a lot of interest is microwave based techniques. The main advantages are that the devices can be made portable and cheap, compared to CTs or MRIs, and also that electromagnetic radiation in the microwave range is non–ionizing. Medfield Diagnostics is a medical start up founded in 2005 and they are developing a head-mounted *microwave measurement system* (MMS) for stroke diagnosis. The goal is to develop a device that can be used in the ambulance for differentiating between IS and ICH at an early stage; thereby making an

[1]The lower interval for IS does not sum to 100% since the paper that reported it ([6]) lists more subtypes than just IS and ICH.

earlier admission of TT possible for patients diagnosed with IS. If the patient is diagnosed with ICH the ambulance can for example be rerouted to a hospital that is well equipped to to deal with strokes due to ICH.

The MMS equipment developed by Medfield Diagnostics has undergone several clinical trials. The data from these trials has also been used to develop a classification algorithm with the goal of being able to distinguish IS from ICH. The classification algorithm is based on machine learning and *singular value decomposition* (SVD). The algorithm has been developed at Chalmers University of Technology and some of the work leading up to the classifier is covered in [12], [13]. However, exactly what physiological information that is the basis for the classifier decision and how it relates to the underlying physics of the microwave measurement is not fully understood. Medfield Diagnostics is interested in increasing their knowledge in this area since it would be valuable in the future development of the classification algorithm and could also be useful for choosing the direction of future clinical trials.

## 1.2 Purpose

The project contains two parts. The first part is to look at clinical data on patients undergoing TT. The goal of this part is to evaluate if the output of the classifier correlates well with the *National Institute of Health Stroke Scale* (NIHSS) score of the patient over time and to see if further understanding of what the classifier bases it's classification on can be gained by having data on the same patient with a clot and with at least partially or fully restored blood-flow. NIHSS is explained further in section 2.1.1 of this thesis.

The other part of the project is to create a variable dielectric model of a patient with ICH. The validity of this model will then be checked in *finite difference time domain* (FDTD) simulations. These simulations will then be used as input for the classification algorithm. The goal here is not to replicate clinical cases in simulations. The goal is rather to create a large data base of simulations from which conclusions about important parameters and their effect on the classification can be drawn. Also the goal of the project is not to do detailed hardware modeling for example of the antennas so assumptions will have to be made to keep the model within the scope of this project.

## 1.3 Aims

### 1.3.1 Thrombolysis analysis

- Find out how well the severity of the stroke over time correlates with an output of the classification algorithm used by Medfield Diagnostics during thrombolytic treatment. This is done by comparing the NIHSS score of a patient with the distance output by the classifier over time

- Detect and filter out movement artifacts to possibly improve the correlation between the NIHSS score and the distance output by the classifier

- Increase knowledge of how this classification algorithm acts when applied to stroke diagnosis by having access to data on the same patient with a clot and with a partially or fully restored blood-flow.

- Use the first data points in time of the thrombolysis data to study the drift with time of the distance outputted by the classification algorithm. The goal being to try to differentiate between drift due to the dissolving clot and drift due to noise. Movement artifacts and drift in the equipment for example.

### 1.3.2 Simulations

- Construct a variable, realistic dielectric model of a head with an ICH and an antenna array. The parameters that are to be varied are: Head size, hair thickness, bleeding position and bleeding size

- Try to draw conclusions about the connections between clinical data and the simulations

- Study how the classifier handles different head sizes, hair thicknesses, bleeding positions etc.

- Evaluate if the classifier can separate bleedings in the right hemisphere from bleedings in the left hemisphere

- Investigate how different levels of noise affect the classification

## 1.4 Delimitations

- **Modifying the classification algorithm**
  Even though the goal of the project is increased knowledge of how the classifier acts and performs in this application modifying the fundamental classification algorithm falls outside the scope of the project.

- **Making new measurements**
  The goal of the project is to analyze existing data. To gather new data apart from the simulations is not intended to be a part of the project.

# Chapter 2

# Theoretical Background

## 2.1 Stroke

Neurons cannot store glucose and do not cope well with anaerobic metabolism [14]. Due to this they are reliant on a high and constant supply of oxygen and glucose via the blood stream. The cerebral blood flow (blood flow to the brain) at rest is around 50 ml per minute per 100 grams of grey matter and 20 ml per minute per 100 grams of white matter [15]. In total the cerebral blood flow accounts for 15% of the cardiac output meaning that the brain is a very energy and oxygen demanding organ for it's size.

As mentioned a stroke is defined as an acute neurological deficit with a vascular basis, leading to reductions in cognitive or physical functions. The symptoms also have to persist for more than 24 hours for it to be labeled a stroke. If the symptoms persist for less than 24 hours it is labeled as a *transient ischemic attack* (TIA) [1]. In this report acute stroke will refer to the lack of blood supply to parts of the brain while stroke will refer to the symptoms due to the acute stroke. This means that the stroke can persist after the acute stroke is over.

### 2.1.1 NIHSS

The National Institute of Health Stroke Scale (NIHSS) is a scoring system used to evaluate the severity of a stroke by asking the patient questions and assessing the patient's motor control. The minimum score is 0, which indicates a healthy patient, and the maximum score is 42, extremely severe stroke, and the scale has been found to correlate with the outcome of the stroke meaning that lower score indicates a better recovery over time [16]. However the NIHSS score seems to weigh deficiencies in anterior (front of the

brain, involved in motor control, speech and reasoning etc.) circulation more heavily than deficiencies in posterior (back of the brain, visual processing) circulation. It has also been found that a zero on the NIHSS score does not exclude neurological damage [17]. An example of the factors included and their respective scores can be found in [18].

## 2.2 Microwave Theory

### 2.2.1 Electromagnetic waves

Electromagnetic radiation is a form of energy consisting of two alternating fields, an electric field and a magnetic field [19]. An illustration can be seen in Figure 2.1.

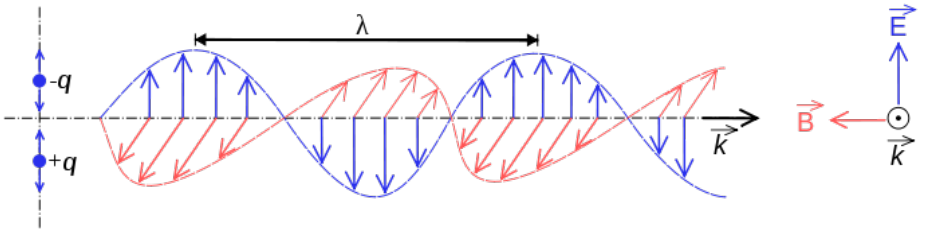


**Figure 2.1:** An electromagnetic wave propagating through a medium with it's alternating magnetic and electric waves in phase and perpendicular to each other. The wavelength $\lambda$ is also marked out. Figure is in the public domain. Retrieved from: `http://upload.wikimedia.org/wikipedia/commons/3/35/Onde_electromagnetique.svg` on 2013-10-01.

If traveling through a lossless medium the electromagnetic wave will self propagate. The magnetic component gives rise to the next electric component which gives rise to the next magnetic component and so forth.

Electromagnetic radiation comes at many different frequencies and depending on the frequency they are called different things, from gamma rays, through x-rays, through visible light, through microwaves to long radio waves. An illustration of the different frequencies and their respective names can be seen in Figure 2.2.

The frequency range of interest in this report is 100 MHz up to around 2 GHz since this is the frequency range being investigated by Medfield Diagnostics. Frequency is related to wavelength through Equation 2.1 where $\lambda$ is the wavelength, $c$ is the speed of light and $f$ is the frequency.

$$\lambda = c/f \tag{2.1}$$

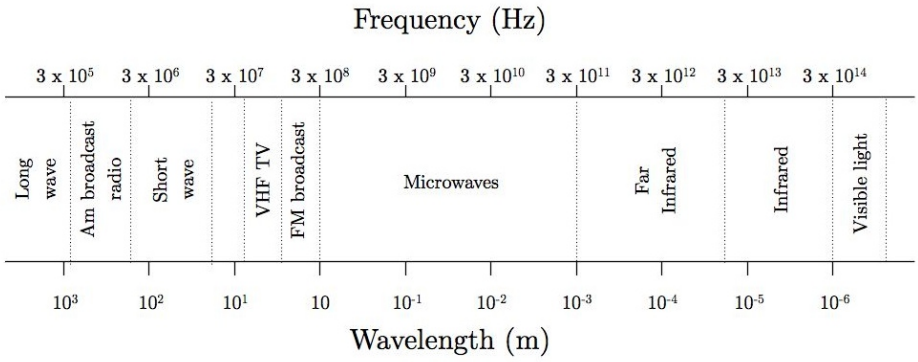


**Figure 2.2:** A chart of the different frequency regions of the electromagnetic spectrum and their names. Adapted from *Microwave Engineering* by David M. Pozar [19].

| Tissue | Wavelength [m] |
|---|---|
| Air | 0.30 |
| Blood | 0.037 |
| White matter | 0.048 |
| Grey matter | 0.037 |
| Fat | 0.13 |

**Table 2.1:** Wavelengths of microwaves at 1 GHz for a few different media [20].

Electromagnetic waves travel at different speeds in different media. In air the wavelengths corresponding to 100 MHz to 2 GHz are 15 cm to 3 m. In denser materials, like water or brain, the wavelength is shorter [19]. The wavelength of electromagnetic radiation at 1 GHz for some tissue types is given in Table 2.1.

### 2.2.2 Maxwell's equations

In standard circuit theory the assumption that the wavelength of the signal or current is much larger than the physical dimensions of the device is made. Due to this the phase of the signal or current is assumed to be the same everywhere in the devise meaning that the voltage in a wire for example is the same in the whole length of the wire at each time instant [19].

In the case of higher frequencies this assumption no longer holds since the wavelength of the current is in the order of the physical dimensions of

the device. This means that components have to be considered as *distributed elements* meaning that the phase of the current or voltage, *i.e.* the measured current or voltage, can vary along the length of a wire at a certain time instant.

Due to this standard circuit theory is no longer valid and Maxwell's equations have to be used, see equations 2.2(a-d). These equations are usually referred to as Faraday's law (2.2a), Ampère's law (2.2b) and Gauss' laws (2.2(c-d)).

$$\nabla \times \bar{\mathcal{E}} = \frac{-\partial \bar{\mathcal{B}}}{\partial t} - \bar{\mathcal{M}} \tag{2.2a}$$

$$\nabla \times \bar{\mathcal{H}} = \frac{\partial \bar{\mathcal{D}}}{\partial t} + \bar{\mathcal{J}} \tag{2.2b}$$

$$\nabla \cdot \bar{\mathcal{D}} = \rho \tag{2.2c}$$

$$\nabla \cdot \bar{\mathcal{B}} = 0 \tag{2.2d}$$

With the following notation:

$\bar{\mathcal{E}}$ is the electric field, [V m$^{-1}$]

$\bar{\mathcal{H}}$ is the magnetic field, [A m$^{-1}$]

$\bar{\mathcal{D}}$ is the electric flux density, [C m$^{-2}$]

$\bar{\mathcal{B}}$ is the magnetic flux density, [Wb m$^{-2}$]

$\bar{\mathcal{M}}$ is the magnetic current density, [V m$^{-2}$]

$\bar{\mathcal{J}}$ is the electric current density, [A m$^{-2}$]

$\rho$ is the electric charge density, [C m$^{-3}$]

In free space the equations 2.3(a,b) connect the magnetic flux density to the magnetic field and the electric flux density to the electric field.

$$\bar{\mathcal{B}} = \mu_0 \bar{\mathcal{H}} \tag{2.3a}$$

$$\bar{\mathcal{D}} = \epsilon_0 \bar{\mathcal{E}} \tag{2.3b}$$

Where $\mu_0 = 4\pi \times 10^{-7}$ [H m$^{-1}$] is the permeability of free-space, and $\epsilon_0 = 8.854 \times 10^{-12}$ [F m$^{-1}$] is the permittivity of free space [19].

### 2.2.3 Phasor notation and the effect of the media

The four equations 2.2(a-d) together with the constitutive relations in Equation 2.3 are enough to describe an electromagnetic wave with arbitrary time dependence traveling through free space. In most cases however we can assume that the wave is sinusoidal and at steady state. In this case phasor notation can be used meaning that the waves are represented as sinusoidals whose amplitude, frequency and phase do not vary with time. Under these assumptions all field quantities can be written as complex vectors with an implied $e^{j\omega t}$ time dependence where $\omega$ is the angular frequency, $\omega = 2\pi f$ rad/s. The time derivatives in equations 2.2(a-d) can now be replaced by $j\omega$ and we get [19]:

$$\nabla \times E = -j\omega B - M, \tag{2.4a}$$

$$\nabla \times H = j\omega D + J, \tag{2.4b}$$

$$\nabla \cdot D = \rho, \tag{2.4c}$$

$$\nabla \cdot B = 0. \tag{2.4d}$$

To account for the wave traveling through media other than free space we also need to expand the constitutive relations to:

$$B = \mu H \tag{2.5a}$$

$$D = \varepsilon E \tag{2.5b}$$

$$J = \sigma E \tag{2.5c}$$

Where $\mu$ is now the permeability of the material and $\epsilon$ is the permittivity of the material. We also have to add $\sigma$ which is the conductivity of the material since this is no longer zero, as in free space. These three values are called the dielectric properties of the material and often vary with frequency. Values for the permittivity and conductivity of different biological tissues at different frequencies can be found in *The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues* by Gabriel *et al.* [21]. The observant reader might note that the permeability is omitted in the work of Gabriel *et al.* The reason for this is that the permeability of organic tissues is very close to the permeability of free space [22].

In general the permittivity of organic tissue decreases with increasing frequency and does so in three steps designated the $\alpha$-, $\beta$-, $\gamma$-dispersions. In the lower frequency span, 0-10 kHz, the $\alpha$-dispersion, which is associated with ionic diffusion processes at the cell membrane dominates. From 10 kHz to 100 MHz the $\beta$-dispersion, associated with the polarization of the cell membrane dominates. At frequencies larger than 100 MHz the change in permittivity is mainly associated with the polarization of the water molecules in the tissue, the so called $\gamma$-dispersion [22]. The conductivity of biological tissues generally increases with increasing frequency [21], [22].

The fundamental idea underlying microwave based stroke detection is that the stroke gives rise to a change in the dielectric parameters in the area of the brain affected by the stroke and its surrounding regions. In the case of an ICH a volume of blood is formed in the brain and there is a significant difference in the dielectric properties of blood compared to brain tissue [21]. In the case of an IS the clot gives rise to an ischemia and it has been found that the ischemia and the effects it has on the tissue gives rise to changes in the dielectric parameters of the tissue [23], [24]

The propagation of a wave in a general lossy media can be described by the propagation constant $\gamma$ seen in Equation 2.6 and the wave impedance $\eta$ seen in Equation 2.7.

$$\gamma = j\omega\sqrt{\mu\epsilon'}\sqrt{1 - j\frac{\sigma}{\omega\epsilon'}} \tag{2.6}$$

$$\eta = \frac{j\omega\mu}{\gamma} \tag{2.7}$$

Where $\epsilon'$ is the real part of the permittivity, $\mu$ is the permeability and $\sigma$ is the conductivity. The propagation constant describes what happens to the amplitude and phase of a wave as it travels through the media and the wave

impedance is the ratio of the transverse electric and magnetic components of the wave as it travels through the media [19]. The propagation constant and the wave impedance will arise in the solutions of Maxwell's equations and connect these solutions to the dielectric parameters of the material.

At the borders of materials with different dielectric properties electromagnetic radiation in the microwave spectrum will act in a similar way to electromagnetic radiation in the visible light range traveling from one material to another. Some part of the radiation will be reflected at the interface of the two media while some is transmitted.

Apart from the amplitude the phase of the electromagnetic wave can also be affected by traveling through a media due to taking different paths through the media. The phase is a relative property meaning that it can only be defined relative to another wave of the same frequency. Most often the phase of the wave at the sending antenna is compared to the phase of the wave at the receiving antenna thus extracting more information about the dielectric properties than just looking at the amplitude. An illustration of a phase shift can be seen in Figure 2.3.

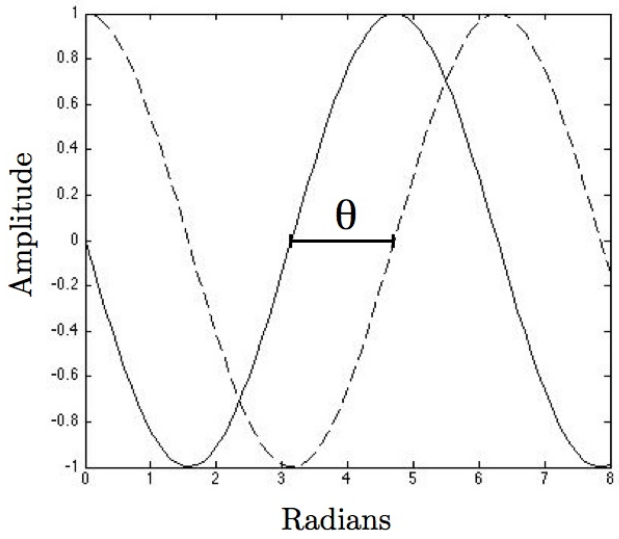


**Figure 2.3:** Illustration of a $\theta$ phase shift between two waves. Adapted from *Detection of traumatic epidural and subdural hematomas in brain phantoms using microwave technology* by Tora Dunås and Madeleine Kildal Schilliger [25].

### 2.2.4 Medfield Diagnostics' use of microwaves

Electromagnetic waves in the frequency range of 100 MHz to 2 GHz are emitted from an antenna located close to the head of the patient. These waves then propagate through the air, and through the different tissues of the head of the patient and back out to a receiving antenna on the other side of the head. This propagation is governed by the dielectric properties of the different media that the waves encounter and the geometries of different media also affect the waves due to giving rise to different boundaries between media. At each interface some part of the wave is transmitted and some part is reflected and passed on. Different frequencies travel through different paths giving rise to different received amplitudes and phases at the receiving antenna giving information about both the tissues in the head of the patient and their geometries. An illustration can be seen in Figure 2.4.

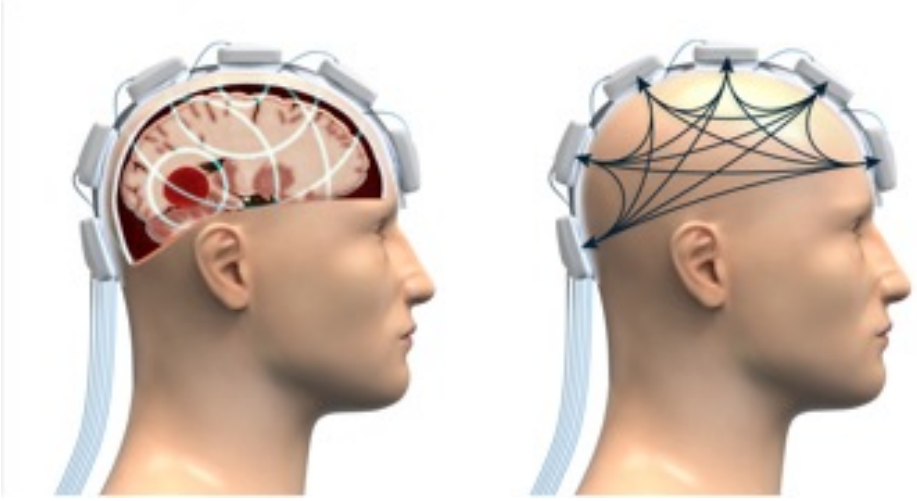

**Figure 2.4:** Illustration of the transmitting and receiving antennas on the head of the patient and the electromagnetic radiation that they emit. Used with permission from Medfield Diagnostics.

### 2.2.5 S-Parameters

A common way of illustrating the connection between the sending and receiving antenna in the summary above is to think about the measurement object, *i.e.* the head of the patient, the hair and the air surrounding it, as a two port network where the receiving antenna is the first port and the sending antenna is the second port. The connection between these ports can then

be described by so called S-parameters. Traditionally the S–parameters are defined as the ratio of the voltage at the receiving port over the voltage at the transmitting port [19] as in Equation 2.8.

$$S_{ij} = \frac{V_i}{V_j} \tag{2.8}$$

Where $S_{ij}$ is the S–parameter when $i$ is the receiving port and port $j$ is the transmitting port. $V_i$ is the voltage on receiving port $i$ and $V_j$ is the voltage on transmitting port $j$. This can however also be generalized to phase by looking at the phase of the receiving port over the phase at the transmitting port.

The full two port network can then be described by the matrix in Equation 2.9.

$$\begin{bmatrix} b_i \\ b_j \end{bmatrix} = \begin{bmatrix} S_{ii} & S_{ij} \\ S_{ji} & S_{jj} \end{bmatrix} \begin{bmatrix} a_i \\ a_j \end{bmatrix} \tag{2.9}$$

Where $a_i$ is the incident signal on port $i$ while $a_j$ is the incident signal on port $j$. $b_i$ is the signal coming out of port $i$ and $b_j$ is the signal coming out of port $j$. $S_{ii}$ is then the amount of signal coming back to port $i$ when port $i$ is transmitting. $S_{ij}$ is the amount of signal out of port $i$ when port $j$ is transmitting. $S_{ji}$ and $S_{jj}$ are interpreted in the same way. In an antenna array these S–parameters are often gathered for all antenna combinations and for a frequency span of interest. An example of S–parameters can be seen in Figure 2.5. This means that the effect that a system has on electromagnetic radiation can be described by it's S-parameters [19].

### 2.2.6 S-Parameters applied to the measurement system

The ports of the systems studied in this report are antennas and in this case the S-parameters can be understood in the following way.

The magnitude of the reflection basically shows at what frequencies the antenna actually radiates. The antenna in the example in Figure 2.5 has resonances and radiation peaks around 0.3 and 1 GHz where the reflection is -20 dB. This means that the reflection is 100 times smaller than the transmission, *i.e.* it transmits 99% of the power fed to it.

The magnitude of the transmission shows how much is actually transmitted between the two antennas at different frequencies, in Figure 2.5 most power is transmitted in around 0.3 to 1 GHz.

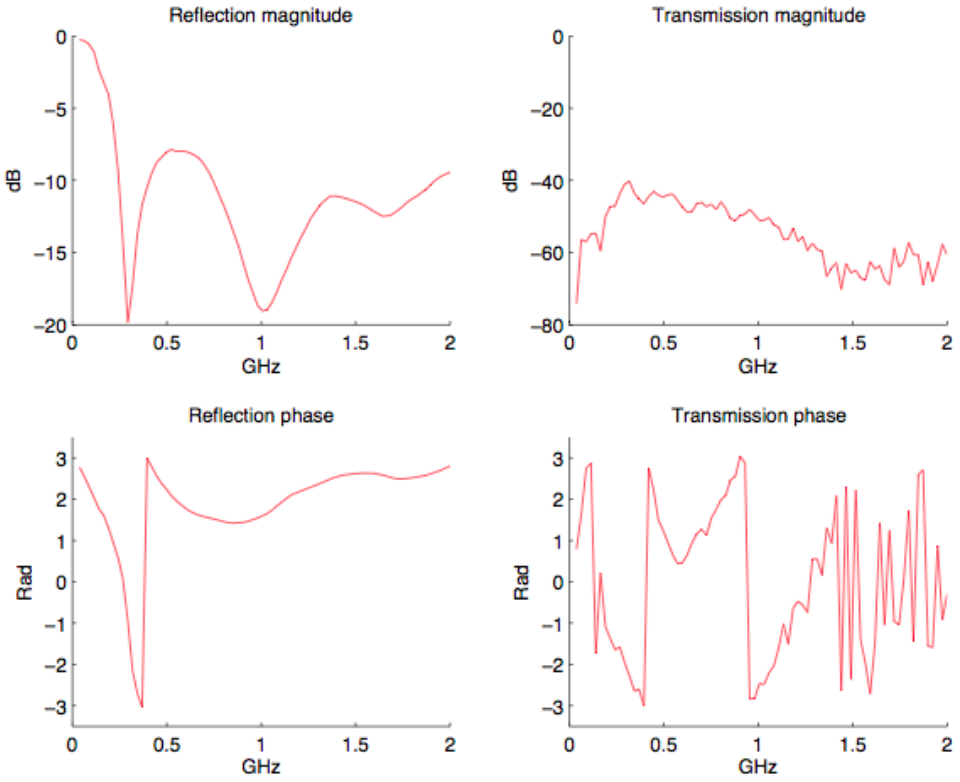


**Figure 2.5:** S–Parameters from example measurement on water filled tank. The upper left figure shows the magnitude of the complex signal received at antenna 0 when antenna 0 is transmitting *i.e.* the magnitude of the reflection. The lower left figure shows the phase of the same signal *i.e.* the phase-shift of the reflection. In the upper right figure we see the magnitude of the transmission from antenna 1 to antenna 2 and in the lower right figure we see the phase-shift of the same transmission. The phase has been wrapped to not exceed $\pm\pi$.

The phase gives indications about the paths traveled by the microwaves. If the phase is smooth except for doing a few jumps of $2\pi$[1] it indicates that all the microwaves have traveled along the same path for all frequencies. The change in phase is only due to the smooth change in electric length of the path experienced by the different frequencies. The electric length $l_e$ is

[1]The wrapping of the signal to $\pm\pi$ is performed to better visualize the small changes in phase that are not due to the frequency dependent electrical length. Removing the inclination of a linear regression of the frequency curve, or parts of it, from the frequency curve is another way to achieve the same thing.

the length that an electromagnetic wave experiences along a transmission line and is given by[26]:

$$l_e = \frac{L}{\lambda} \tag{2.10}$$

Where $L$ is the physical length of the transmission line and $\lambda$ is the wavelength of the electromagnetic wave.

This smooth phase can be seen in the phase of the reflection in Figure 2.5. The observant reader might note that the curve does not go as a straight line with frequency as Equation 2.10 indicates. This is due to non-ideal effects but the smoothness of the curve is preserved. In the case of the phase of the transmission we can see more and more ripples the further out we go from 1 to 2 GHz. This indicates that the path used varies a lot with frequency therefore the phase of the transmitted wave also varies a lot with frequency.

### 2.2.7 Effective medium approximations

*Effective medium approximations* (EMAs) deals with the problem of approximating the permittivity, conductivity or other material parameters of a macroscopically inhomogeneous media [27]. Many materials fall into this category. For example a metal-dielectric composite consisting of metallic and dielectric grains ordered in some random fashion.

A number of different EMAs have been proposed and developed over the years, some more general and some for specific conditions. The EMA used in this report is a version of Bruggemanns formula for spheres of two different kinds of materials which can be seen in Equation 2.11 [28].

$$\delta_1 \frac{\alpha_1 - \alpha_m}{\alpha_1 + (n-1)\alpha_m} + \delta_2 \frac{\alpha_2 - \alpha_m}{\alpha_2 + (n-1)\alpha_m} = 0 \tag{2.11}$$

Where $\delta_1$ and $\delta_2$ are the volume fractions of the two materials, $\alpha_1$, $\alpha_2$ and $\alpha_m$ is the material parameter to be approximated for material one, material two and the mixture and $n$ is the dimensionality (1D, 2D or 3D).

## 2.3 FDTD and the forward problem

The forward, or direct, problem is the problem of calculating the propagation of electromagnetic radiation through a body or object with known dielectric properties and geometry. For very simple geometries it is possible to solve Maxwell's equations analytically and thereby arrive at the answer [29]. If one wants to solve the forward problem for a more realistic case

(The propagation of electromagnetic radiation through the human head for example.) solving Maxwell's equations becomes very hard or even impossible. In these situations one has to resort to numerical methods. One such method is the Finite-Difference Time-Domain (FDTD) method.

In 1966 Kane S. Yee developed a time-dependent solution of Maxwell's equations in the differential form (Seen in Equation 2.2(a-d)) [30]. Yee's solution was to use central difference approximations for both the space and time derivatives and to stagger the discretization of the magnetic and electric field by half a step in both space and time. This approach leads to a second order accurate numerical solution to Maxwell's equations which is generally referred to as the Yee-algorithm or FDTD [31].

Central difference approximations of Faraday's and Ampère's laws lead to six equations, one in each Cartesian direction for the electric and magnetic field respectively. These equations can then be rearranged so that the magnetic or electric field along one dimension for a certain time instant only relies on the magnetic and electric fields of previous time instants. The interested reader is referred to *Introduction to the Finite-Difference Time-Domain (FDTD) Method for Electromagnetics* by S. D. Gedney [31] for the mathematical derivations.

So if the initial values of the fields are known at all points in the discrete space the update equations provide an explicit recursive way to calculate the fields at future time points for all points in the discrete space [31].

Central difference approximations of Gauss' laws together with the discrete forms of Faraday's and Ampère's laws can be used to show that Gauss' laws still hold meaning that no false charge is created [31].

What about stability then? An analysis of this is provided in *Introduction to the Finite-Difference Time-Domain (FDTD) Method for Electromagnetics* by S. D. Gedney [31] page 48-52. For the purpose of this paper however it is enough to state that the final stability criterion on the time step allowed becomes:

$$\Delta t = \frac{1}{c_{max}} \frac{1}{\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}}} \tag{2.12}$$

Where $c_{max}$ is the speed of light in the medium with the highest speed of light in the simulation space and $\Delta x, \Delta y, \Delta z$ are the length of the edges of one cell in the simulation space. In practice this means that the wave is restricted to only propagate a third of the distance along the cubic cell diagonal [31].

### 2.3.1 Absorbing boundary conditions and the PML

A problem with the FDTD method is how to handle the edges of the simulation space since we want to simulate something in free air. We want to simulate a infinite unbounded region with a finite dimensional grid. The first approach might be to ignore the problem. This however leads to the boundaries acting as a box of perfectly conducting material giving rise to large reflections back into the simulation space. The second solution might be to make the simulation space very large to avoid reflections due to them not making it back during the simulation time. This has the severe drawback of increasing simulation time since the FDTD method does not differentiate between empty space and other materials. The only plausible solution is to add a boundary that absorbs any wave traveling towards it. These kinds of boundaries are called absorbing boundary conditions (ABC).

The current standard ABC is the perfectly matched layer (PML) which was pioneered by Jeanne-Pierre Bérenger in the mid nineties [32]. The PML is such named since it absorbs an arbitrarily polarized, broad-band electromagnetic wave impinging on a PML half space without reflection [31]. These properties means that the PML is a very good truncation of the computational space in a FDTD simulation. The more interested reader is referred to chapter six in *Introduction to the Finite-Difference Time-Domain (FDTD) Method for Electromagnetics* by S. D. Gedney [31] and to *Perfect matched Layer (PML) for Computational Electromagnetics* by J. P. Bérenger [33].

## 2.4 Medfield R10 and correlation of measurements

The R10 system (see Figure 2.6) consists of a computer used to operate the system and also responsible for gathering and storing the data, a vector network analyzer (VNA), a microwave switch, cables connecting the antennas to the switch, eight patch antennas and a headband keeping the antennas in place on the head of the patient.

The R10 system gathers the real and imaginary parts of the S-parameters at a number of frequencies for all its antenna combinations. Each measurement in time, or time sample, from a patient is represented by a complex matrix with the dimensions $m,n$. Where $m$ is the number of samples in frequency and $n$ is the number of antenna combinations, see Figure 2.7. So the S-parameter of an antenna combination can be found along the column representing that antenna combination.

If several consecutive measurements are made there will be $p$ matrices

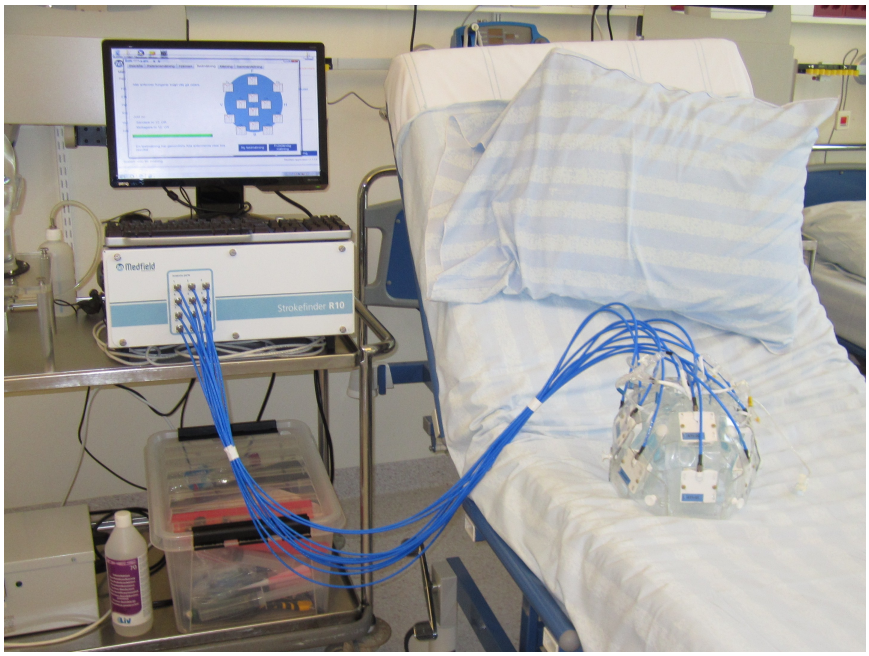


**Figure 2.6:** Picture of Medfield Diagnostics' R10 system. This picture shows an older version with a twelve antenna helmet instead of the eight antenna headband used in this study.

of this kind, one for each time sample, so the whole data set for one patient can be seen as a three dimensional matrix with the dimensions $m$,$n$,$p$ where $m$ represents frequency, $n$ antennas and $p$ time. See Figure 2.8

The $m \times n$ matrix for a time instance $p$ can be seen as a point in $m \cdot n$ dimensional space (put all columns or rows after each other into one long vector). Let $A_p$ denote the $m \times n$ matrix at time instant $p$ and let $A_{p+1}$ denote the $m \times n$ matrix at time instant $p + 1$. In this case the change in the measurement from time instance $p$ to $p + 1$ can be represented by the matrix $A_{diff}$ according to Equation 2.13.

$$A_{diff} = A_p - A_{p+1} \tag{2.13}$$

The matrix $A_{diff}$ can then be seen as the vector from the point represented by $A_p$ to the point represented by $A_{p+1}$.

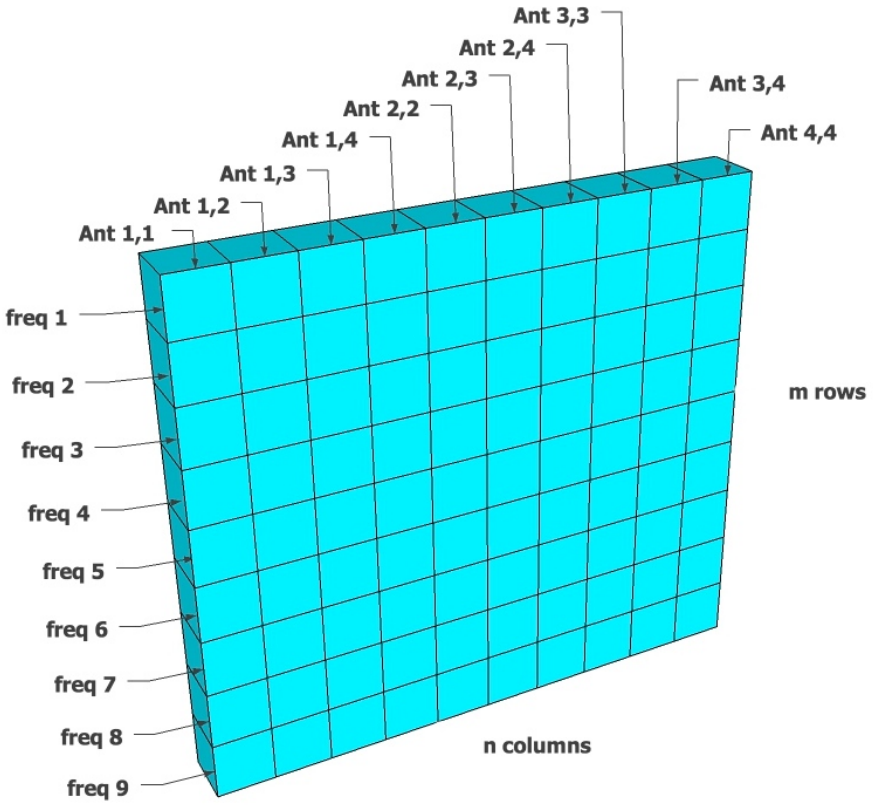


**Figure 2.7:** Illustration of one measurement or time sample from a patient. In this case with 4 antennas and 9 frequency samples.

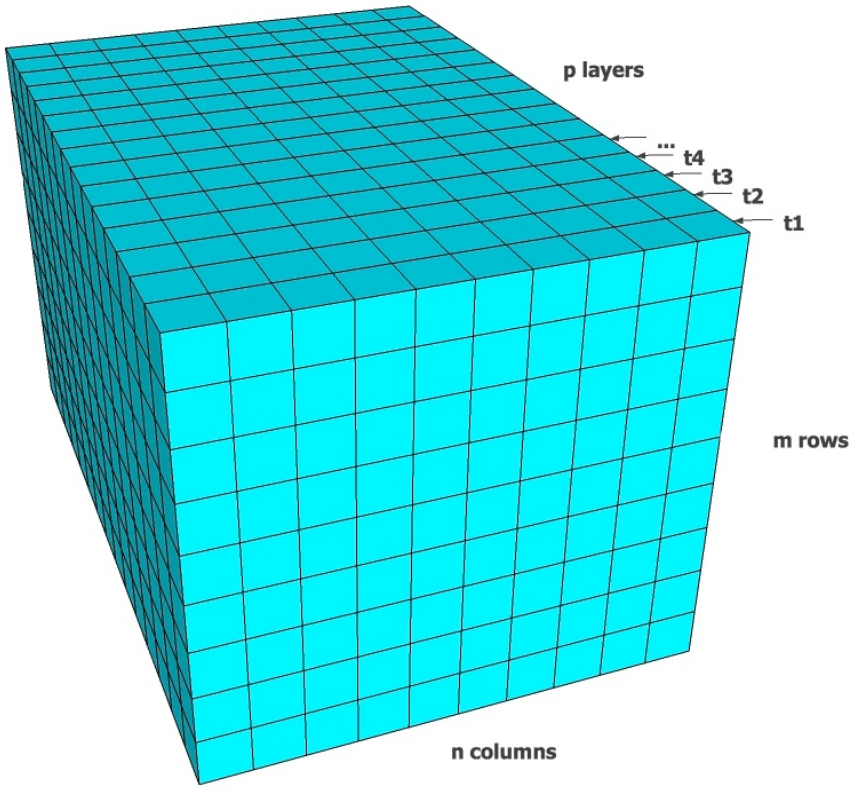


**Figure 2.8:** Illustration of a full measurement of a patient with 15 time samples.

## 2.5 Classification and pattern recognition

Humans are very apt at recognizing patterns and sorting things based on them. We can separate happy faces from sad ones or an apple from a banana only based on smell. Due to this it is not surprising that machines and algorithms are being trained to act on the same patterns and in the same way as humans do.

### 2.5.1 One dimensional example

To illustrate this the separations of atlantic cod from atlantic salmon at a fish processing plant will be covered. See Figure 2.9 and Figure 2.10. How would one go about if one wanted to build a machine that could classify which species a certain fish belong to?

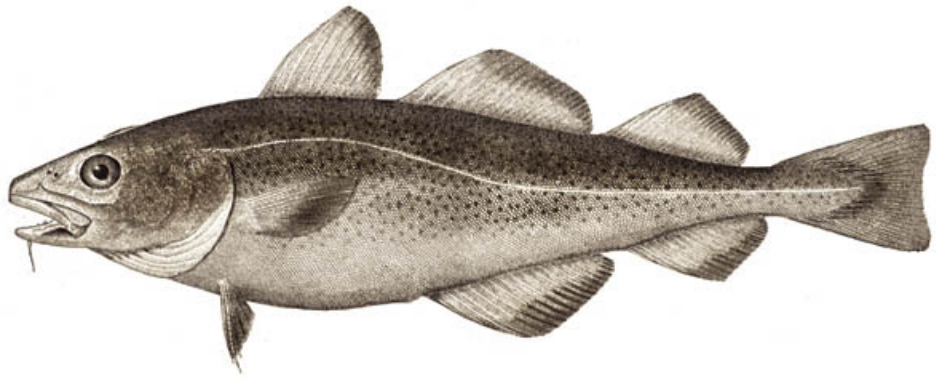

**Figure 2.9:** Illustration of cod. Figure is in the public domain. Retrieved from: `http://upload.wikimedia.org/wikipedia/commons/a/a3/Atlantic_cod.jpg`. Access date: 2014-02-13

Measurable parameters, and decision boundaries for these parameters are needed that set the two species apart. One such parameter could be weight. Weighing 20 samples of each fish might give the result seen in Figure 2.11. Salmon weighs more than cod on average but a threshold cannot be set that separates salmon from cod simply based on weight.

### 2.5.2 Sensitivity, Specificity, ROC and AUC

At this point it is appropriate to introduce specificity, sensitivity, the *receiver operating characteristics curve* (ROC) and the *area under curve* (AUC). As

**Figure 2.10:** Illustration of salmon. Figure is in the public domain. Retrieved from: `http://upload.wikimedia.org/wikipedia/commons/archive/0/06/20091123150247%21Salmo_salar_%28crop%29.jpg`. Access date: 2014-02-13

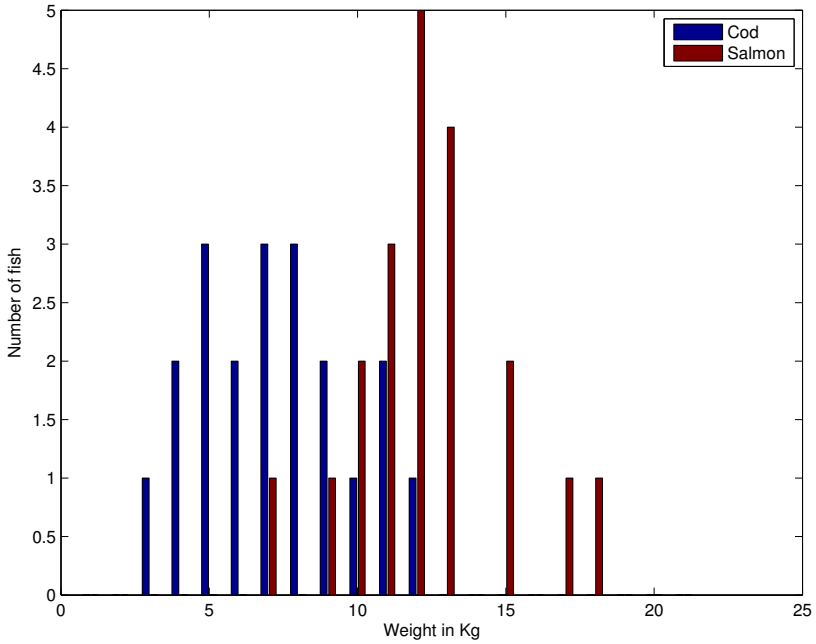


**Figure 2.11:** Weights of fishes from the twenty samples of each species. Red is salmon and blue is cod.

mentioned a threshold that achieves a perfect separation between salmon and cod based on weight cannot be found. Independent of weight-threshold some of the samples of at least one fish will be miss-classified.

If salmon is considered the positive result sensitivity is specified as [34]:

$$Sensitivity = \frac{Number\ of\ true\ positives}{Total\ number\ of\ positives} \tag{2.14}$$

Meaning how many of the samples classified as salmon are actually salmon. The specificity is similarly defined, if cod is considered the negative result, as:

$$Specificity = \frac{Number\ of\ true\ negatives}{Total\ number\ of\ negatives} \tag{2.15}$$

Meaning how many of the samples classified as cods are actually cods. If the threshold is set to 13 kg, meaning that a fish that weighs 13 kg or more is considered a salmon, everything classified as salmon will be salmon. Giving a sensitivity of 1. However this is payed for by a fairly low specificity. Out of the 32 fishes classified as cod due to weighing less than 13 kg only 20 are actually cod and 12 were salmon. This gives a specificity of 0.625. Setting the threshold to 10 kg gives a sensitivity of 0.81 and a specificity of 0.89. Some of the sensitivity has been traded for specificity. The optimal choice for this trade-off depends on application and the requirements set on sensitivity and specificity.

As mentioned a trade-off has to be made between sensitivity and specificity and a way to illustrate this trade-off, as the weight threshold is moved, is a *receiver operating characteristics* (ROC) curve. The ROC-curve plots the sensitivity on the $y-axis$ and one minus the specificity on the $x-axis$. One minus the specificity is also known as the *False Positive Rate*. The ROC-curve illustrates what happens with sensitivity as the threshold value is swept [34]. In the fish case the threshold value is swept from the weight of the lightest fish to the weight of the heaviest fish. The resulting ROC-curve can be seen Figure 2.12.

A general measure of the quality of a classifier independent of the specific threshold chosen is the area under the ROC-curve (AUC). If AUC is one a perfect classification can be made since both sensitivity and specificity are 1 for some threshold value. The larger the AUC value is the better the classification potential of the classifier. The AUC value for the ROC-curve in Figure 2.12 is 0.9415.

### 2.5.3 Two dimensional example

What if the sensitivity or specificity provided by a weight based threshold is not enough? In that case more parameters have to be introduced. The

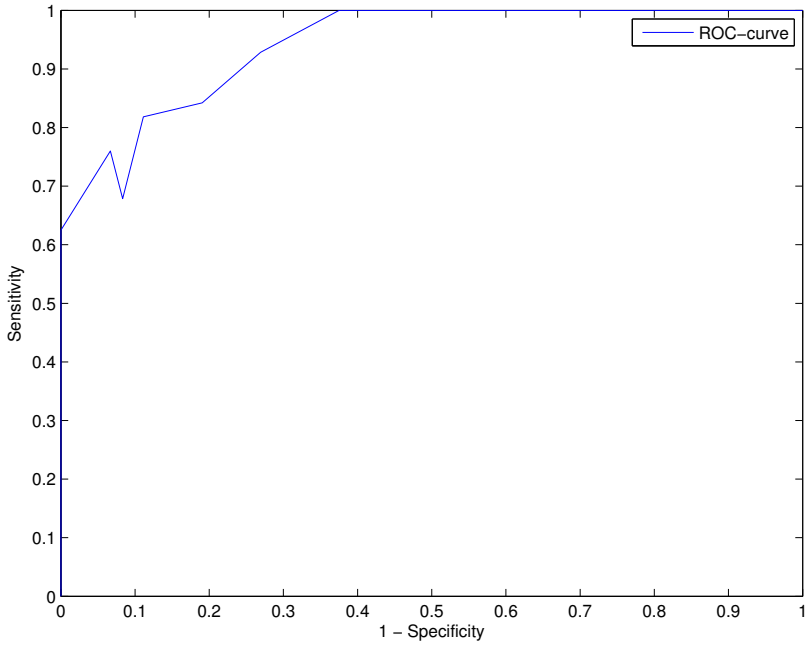


**Figure 2.12:** ROC-curve for classification of Salmon and Cod based on weight.

skin and scales of the salmon are more red than the skin and scales of the cod. This can be used as a parameter. We set up a camera that takes a RGB-picture of each fish. The red values of each pixel are then summed assuming that the background is the same for each fish. Measuring the amount of red combined with the weight of the forty hypothetical fish might give something similar to Figure 2.13

With the inclusion of amount of red an almost perfect classification of these 20 cods and 20 salmon can be achieved. The suggested threshold line has the equation:

$$b = -0.04w + 1 \tag{2.16}$$

Where $b$ is the brightness and $w$ is the weight. The weight and brightness of each fish that reaches the separation stage is measured, if the measurement pair ends up above the line in Figure 2.13 it is classified as salmon and if it ends up below it is classified as cod.

In the example above there were two measurements or two $features$ for

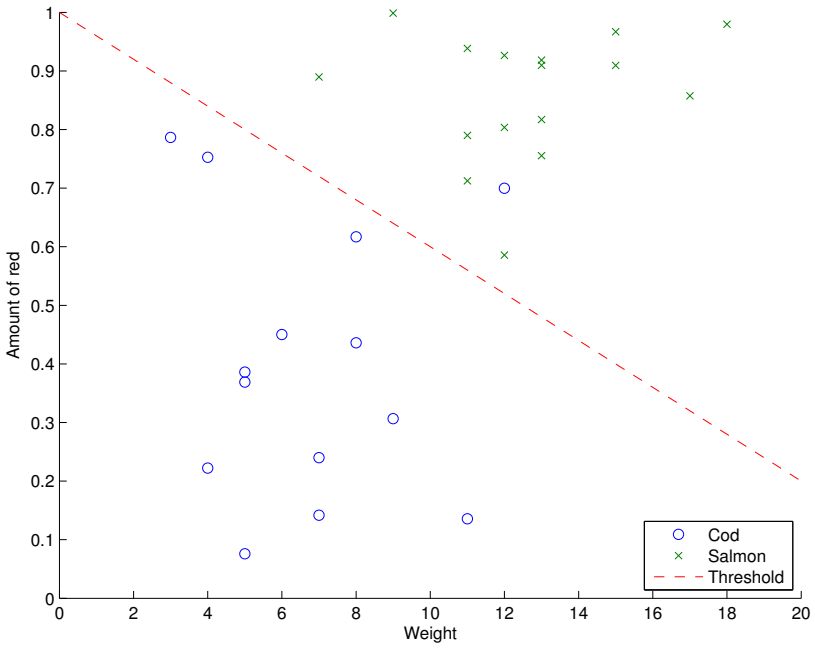


**Figure 2.13:** Scatter plot of weight and brightness of forty samples. A suggested threshold line is also presented.

each sample. This means that the *feature space* is two dimensional and is represented by weight on one axis and brightness on the other. These features have also been chosen manually. Meaning that someone has specified what the machine is to look for. These low dimensional feature spaces are easy to visualize and work well if it is possible to find a few key features that separate the classes well [34].

### 2.5.4 Classification in higher dimensions

If it is not possible to manually select a few features or find patterns that yield a satisfactory classification one often has to resort to gathering a lot of data on each sample. This means that the feature space is of high dimensionality and hard to visualize. To cope with high dimensionality pattern recognition has to be employed. In the low dimensional fish case the twenty samples of cod and twenty samples of salmon can be considered the *training set*. The measurements from these samples was the information that was

available in trying to separate the two species by some kind of threshold. Training sets are used in pattern recognition as well but instead of having two features for each sample there might be thousands of features for each sample. It is then up to the pattern recognition algorithm to separate the classes as well as possible in the hyperdimensional feature space based on the training set.

**LDA and PCA** One common classification algorithm is *Linear Discriminant Analysis* (LDA). The goal of LDA is to project the high dimensional data onto a one dimensional subspace. This projection is also done in a way that maximizes the between class scatter over the within class scatter [35]. A two dimensional example of this can be seen in Figure 2.14.

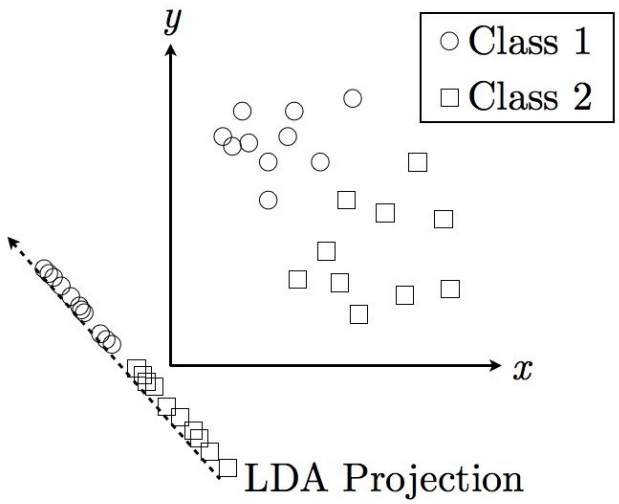


**Figure 2.14:** Example of 2 dimensional LDA. The principle is the same in higher dimensions but it is harder to visualize. Adapted from *Classification of High Dimensional Signals with Small Training Sample Size* by Yinan Yu [35].

The "within class scatter matrix" $\mathbf{S}_W$ and the "between class scatter matrix" $\mathbf{S}_B$ are defined in Equation 2.17 [35].

$$\boldsymbol{S}_W = \frac{1}{N_1}\sum_{i=1}^{N_1}(\boldsymbol{x}_{1,i}-\boldsymbol{\mu}_1)(\boldsymbol{x}_{1,i}-\boldsymbol{\mu}_1)^T + \frac{1}{N_2}\sum_{j=1}^{N_2}(\boldsymbol{x}_{2,j}-\boldsymbol{\mu}_2)(\boldsymbol{x}_{2,j}-\boldsymbol{\mu}_2)^T \quad (2.17)$$

$$\boldsymbol{S}_B = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \tag{2.18}$$

Where $N_1$ and $N_2$ is the number of samples of class one and two respectively, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are their respective means over all samples and $\boldsymbol{x}$ represents a sample vector. The goal is to find a vector $\boldsymbol{w}$ that maximizes:

$$\frac{\boldsymbol{w}^T \boldsymbol{S}_B \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{S}_w \boldsymbol{w}} \tag{2.19}$$

When $\boldsymbol{S}_w$ has full rank the vector $\boldsymbol{w}$ will be the eigenvector corresponding to the largest eigenvalue of the matrix $\boldsymbol{F} = \boldsymbol{S}_W^{-1} \boldsymbol{S}_B$. When we have found this vector $\boldsymbol{w}$ we can simply multiply each sample with it to convert the samples from the higher dimension down to a point on a one-dimensional line. On this line the scatter within the classes has also been minimized while the scatter between classes has been maximized [35].

What if $\boldsymbol{S}_w$ is not of full rank? In that case a dimension reduction step has to be employed before the LDA. A very common dimension reduction method is *Principal Component Analysis* (PCA). The goal of PCA is to reduce the dimensionality of a sample while conserving as much of the descriptive information as possible. Given an $m \times p$ matrix $\boldsymbol{x}$, where the $m$ rows represent measurements from one sample and the $p$ columns are different samples, the PCA can be calculated in the following way [36]:

- Remove the mean of $\boldsymbol{x}$ creating $\bar{\boldsymbol{x}}$
- Calculate the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}}$ of $\bar{\boldsymbol{x}}$
- Calculate the eigenvalues and eigenvectors for $\boldsymbol{C}_{\boldsymbol{x}}$. The eigenvectors will form a orthonormal basis for the space spanned by the original samples in $\boldsymbol{x}$ and the eigenvalues show how much the data varies in each orthonormal direction.
- Choose a desired number of dimensions to keep, $n$, based on the application. Form a *feature matrix* by inserting the eigenvectors corresponding to the $n$-largest eigenvalues as columns in the feature matrix.
- Multiply the transpose of the feature matrix with $\bar{\boldsymbol{x}}$. This will give a $n \times p$ matrix where each sample has had its dimension reduced by $m - n$.

PCA can be thought of as analogous to fitting a polynomial to a number of data points and then only storing the coefficients of the polynomial instead

of the data points. Some data points will probably be outliers and not described by the polynomial but it still preserves the main trends in the data. Analogously PCA is not loss-less as there is loss of information about weakly varying dimensions but it conserves the main trends.

### 2.5.5 HDLSS and subspace distance based classification

The original classification problem related to stroke, to separate bleedings from clots based on clinical data, is a so called *High Dimensionality with Low Sample Size* (HDLSS) problem. This means that our *feature space* is of high dimensionality, we have thousands of data points from each measurement/patient. At the same time it is hard to acquire training data since it all has to be gathered in clincal studies meaning that we have few training samples. HDLSS leads to a number of problems [35]:

- The global distribution of each class in this multidimensional space is impossible to grasp. This means that it is difficult to establish a statistical model with high confidence levels based on few samples of training data.
- Outliers are hard to detect.
- The covariance matrix is hard or impossible to estimate properly since there are not enough samples to catch the covariance structure.
- Information redundancy. Many of the dimensions measured may contain redundant information and the classifier has to account for this.
- Overfitting is hard to detect.

This means that methods like LDA are not appropriate for classifying this kind of data. The interested reader is referred to [37]–[39] for a more thorough review of the HDLSS problem and some suggested solutions.

The solution to the HDLSS problem chosen in this thesis is a so called subspace distance classifier. The foundation of this kind of classifier is described in [12] and will also be covered in this section.

Suppose that we have $C$ classes were each class $c_i (i = 1, \cdots, C)$ has a probability density function in our $n$-dimensional space $\mathcal{R}^n$. A discriminant rule divides this space $\mathcal{R}^n$ into $C$ separate regions $R_1, \cdots, R_C$ ($\cup R_i = \mathcal{R}^n$. This rule allocates each test sample $x_t$ to $c_i$ if $x_t \in R_i$. To be able to achieve a perfect separation the probability density functions of the different classes

have to be orthogonal in this $n$-dimensional space. So how do we set up this discriminant rule?

Assume that a measurement $x_c$ from class $c$ can be written as a linear combination of a number of orthonormal vectors with weights and some noise.

$$x_c = \sum_{k=1}^{m_c} u_{c,k} a_k + e \tag{2.20}$$

Where $x_c$ is the measurement $m_c$ is the number of dimensions spanned by the class, $k$ is the dimension, $u$ is an orthonormal base vector, $a$ is a weight and $e$ is random noise assumed to be Gaussian with mean zero and some variance. The noise is also assumed to be the same for all classes.

The subspace of class $c_i$ can then be said to be the space $U_c$ spanned by it's orthonormal base vectors $u_{c,k}$ and can be written as:

$$U_c = [u_{c,1}, \cdots, u_{c,m_c}] \tag{2.21}$$

We assume that $m_c$ is finite while the noise exists in all dimensions.

If we now wish to test what class a sample $x$ belongs to we can can project it onto each subspace and calculate the projected distance for each class:

$$d_i = \parallel U_i U_i^T x \parallel^2 \tag{2.22}$$

The sample $x$ is then classified as belonging to the the class $c$ which gives the largest $d_i$. This classification is maximum likelihood.

This has assumed that we know the orthonormal base vectors of each class and thereby the subspace of each class. What if we don't?

In that case we gather training data on each class and for training data matrices $X_c$ from these:

$$X_c = [x_{c,1}, \cdots, x_{c,n_c}] \tag{2.23}$$

Where $x_{c,n}$ is a training measurement from class $c$ and $n_c$ represents the number of training samples available for class $c$. In this case:

$$\frac{1}{n_c} \sum_{k=1 \ldots n_c} X_{c,k} X_{c,k}^T = \lim_{n_c \to \infty} L_c L_c^T + \sigma_e^2 I_d \tag{2.24}$$

$L_c$ is then numerically a $d \times n_c$ matrix that spans a certain subspace and when $n_c \to \infty$, $L_c$ and $U_c$ will span the same subspace.

This means that we can approximate the subspace $U_c$ with the approximation $\hat{U}_c$ by taking the singular value decomposition of the matrix $X_c$ and extracting the $m_c$ first dimensions out of $\hat{U}_c$. In Matlab notation:

$$[\hat{U}_c, \sim, \sim] = svd(X_c) \tag{2.25}$$

$$\hat{U}_c = \hat{U}_c(:, 1 : m_c) \tag{2.26}$$

We can then use Equation 2.22 replacing $U_i$ with $\hat{U}_i$ as calculated above and thereby classify each sample based on the subspaces approximated using training data.

## 2.6 Random walks and Polya's Recurrence Theorem

Polya's Recurrence theorem states the following: Let $p(d)$ be the probability that a random walk on a $d$-dimensional lattice ever returns to it's origin. Then, we have that $p(d) = 1$ for $d < 3$ while $p(d) < 1$ for $d \geq 3$. So for the one and two dimensional case the random walk is recurrent while it is transient for dimensions larger than 2. This mean that for dimensions 1 and 2 the random walk is guaranteed to pass through its origin a number of times if left to wander long enough. This is not the case for dimensions larger than two however. It is instead likely that it will drift further away from its origin with time and this divergence will be faster the higher the dimension of the lattice. The proofs for the different dimensions are available in *An Introduction to Random Walks From Pólya to Self-Avoidance* by Michael Kozdron [40].

# Chapter 3

# Method

## 3.1 Tuning of the classifier

The classifier described in section 2.5.5 was applied to microwave measurements gathered by the equipment described in section 2.4. This was done with the goal of separating bleedings from healthy or clots from healthy meaning that there were two classes of interest. There were a number of parameters that affected the classification that could be set before the classification started and they are listed below.

### 3.1.1 Variable parameters

**scramble** If this was turned on the training data in the two sets was scrambled. This was simply done by randomly moving the raw data part from one training sample to another training sample. The goal was to create a data set that was impossible to classify. This data set was then used as a check to verify that the classifier was working as intended.

**fvec** This parameter defined the frequency span of interest.

**preprocessing** The preprocessing consisted of normalizing the measurement matrix of each patient and then taking the natural logarithm of it.

**internal parameters** Two internal parameters in the classifier were also varied. One of them was intended to remove noise and the other was intended to remove large variations that were not due to a clot or a bleeding.

### 3.1.2 Steps of the classification, simplified

- The first step of the classification was preprocessing of the raw data for each patient. The matrix containing the raw data for each patient was normalized using Matlab's `norm` command.
- The natural logarithm of the data matrix was calculated
- The $m \times n$ preprocessed matrix was reformatted into a vector of length $m \cdot n$.
- The classifier used k-fold cross validation and at this point the subset that was to be classified was left out.
- An $X_c$ matrix as in Equation 2.25 for each of the two classes was formed. The $m \cdot n$ vectors for the patients of the bleeding class formed the columns in the matrix related to bleeding and the $m \cdot n$ vectors originating from patients with clot formed the columns of the matrix related to clot.
- The singular value decomposition was calculated as in Equation 2.25 for each matrix. This was done to extract the orthonormal directions and to see to what extent each subset varied in each direction.
- The singular value decomposition output the $\hat{U}_c$ matrices in Equation 2.25.
- The subset that was to be classified was multiplied onto the two matrices $\hat{U}_1$ and $\hat{U}_2$. This lead to a projection of each sample to be classified onto the spaces spanned by the two matrices.
- The length of each projection was then calculated and the difference of length one minus length two was calculated. This difference in length was the final output of the classifier. If this length was positive the sample classified belonged to $c_1$ and if it was negative it belonged to $c_2$.

## 3.2 Thrombolysis analysis

### 3.2.1 Included data

The data analysis was performed on patient data gathered at Sahlgrenska Univeristy Hospital (SU) and Södra Älvsborg Hospital (SÄS). The data

| Patient# | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Hospital | SÄS | SÄS | SÄS | SÄS | SÄS | SÄS | SÄS | SÄS |
| Gender | F | M | F | M | F | F | M | F |
| Age (years) | 79 | 64 | 64 | 57 | 85 | 81 | 88 | 32 |
| Measurement time (m) | 174 | 184 | 134 | 178 | 182 | 180 | 165 | 173 |

**Table 3.1:** Table of patients included in the thrombolysis analysis.

was gathered as part of Medfield Diagnostics ongoing clinical study MF03 using Medfield Diagnostics' Stroke Finder R10 system with eight antennas described in section 2.4

Before data gathering started eye glasses, ear rings and other object were removed from the head of the patient and a measurement was performed with an empty system to verify that the system was working as intended. Only patients above the age of 18 with a suspected stroke were included in the study.

In this data analysis only patients receiving thrombolytic treatment were of interest. There were eight patients receiving thrombolytic treatment included in the study so far and these made up the data set. More information on the patients can be seen in Table 3.1.

The R10 system was assumed to be reciprocal meaning that the signal from antenna $a$ to antenna $b$ is the same as the signal from antenna $b$ to antenna $a$. This gave a total of 36 possible antenna combinations including reflections. The S-parameters were gathered in the frequency range 0–2 GHz with a sampling interval of 7.5 MHz. In the case of the thrombolysis patients these measurements were performed continuously every 40 seconds, which is the minimum measurement time of the R10 system. The total measurement time for each patient can be seen in Table 3.1.

### 3.2.2 Analysis

Each patient was analyzed separately. Firstly the correlation of each measurement in time was compared to the following measurement in time for all patients and for all time instances. The correlation was simply calculated using Matlab's `corr2` function between the $m \times n$ matrix for time instance $p$ and $p + 1$ for all $p$ and for all patients.

A threshold of 0.8 for the correlation was chosen based on visual inspection of the signal and if the correlation dropped below 0.8 when comparing two consecutive measurements, for example $p$ and $p+1$, it was assumed that

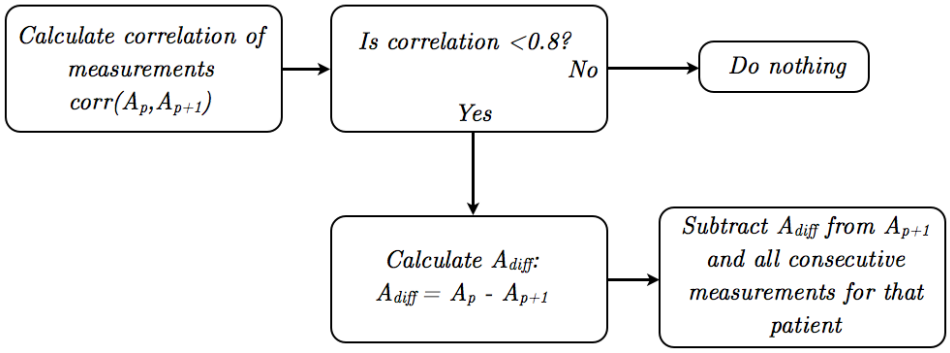


**Figure 3.1:** Illustration of the correction algorithm used.

the change was due to motion artifacts. To correct for this the $A_{diff}$ matrix was calculated for these two time samples as explained in section 2.4. This matrix was then removed from the $p + 1$ measurement and all consecutive measurements for that patient. See Figure 3.1 for an illustration of the procedure. This approach to correction assumed that the errors were additive since the later measurements for a patient might be corrected several times.

Previous experiences at Medfield Diagnostics had shown that the classifier needed around 15-20 training samples of each class to be able to distinguish between the two classes. Due to this the first and last 20 samples of each patient was used as training data for that patient. The first 20 samples represented stroke (High NIHSS) and the last 20 healthy (Low NIHSS). When the classifier had been trained the rest of the measurements for each patient were fed to the classifier. The distance to the two classes created by the training sets, and the difference between these distances were recorded and analyzed.

The analyses were performed with two different parameter sets for the classifier. The first set has given the best classification results in the clinical study. The second set was used to illustrate the effect of changing an internal parameter in the classifier which was supposed to remove the largest variations. It was suspected that the S-parameters for the reflection would look different to those of the transmission and therefore these were analyzed separately.

The data was analyzed by looking at the absolute maximum drop in NIHSS and the absolute maximum drop in the distance output by the classifier for each patient. The theory was that a large drop in NIHSS for a patient should coincide with a larger drop in the distance output by the

classifier. The correlation between these two measures was therefore calculated both for the original data and after attempts had been made to remove the motion artifacts as described above.

## 3.3 Simulations

### 3.3.1 Head model

The foundation of each simulation was a tissue model of the human head provided by Mahmood Qaiser of Med Tech West who had removed artifacts from a model retrieved from Brain Web [1]. This head model contained labels for tissue types: Background, cerebrospinal fluid (CSF), grey brain matter, white brain matter, fat, muscle, skin and skull. This original head model (See Figure 3.2) was of size $181 \times 217 \times 181$ elements giving a spatial resolution of around 1 $mm^3$ per voxel.

### 3.3.2 Modifications to the head model

A layer with a label representing hair was added, an area labeled bleeding was introduced if the simulation was supposed to represent a patient with a bleeding and the model was stretched in the $x$, $y$ and $z$ directions. See Table 3.2 for a list of all variable parameters in the model and their ranges. All voxel coordinates given below are with respect to the simulation space which was $210 \times 240 \times 200$ voxels.

The layer representing hair was added to the outside of the head model. See Figure 3.3. The hair covered the upper head and upper part of the forehead for the front of the head and down to the bottom of the ear for the back of the head. The hair for each patient was of constant thickness. This thickness could be between 0 to 5 mm. The value for the hair thickness for each patient was drawn from an uniform distribution.

A sphere representing blood with a radius between 9 and 33 mm was added. This corresponded to a volume of 10 to 150 ml. The sphere was placed so it only replaced grey matter, white matter or CSF. See Figure 3.3. The coordinates and radius of the sphere were all drawn from uniform distributions.

The stretching of the heads in each the of the $x$, $y$ and $z$ dimensions was drawn individually from normal distributions. The mean and variance of these normal distributions were chosen so the distributions fit the first percentile (smallest) of female heads up to the 99th percentile (largest) of

[1] `http://brainweb.bic.mni.mcgill.ca/brainweb/`

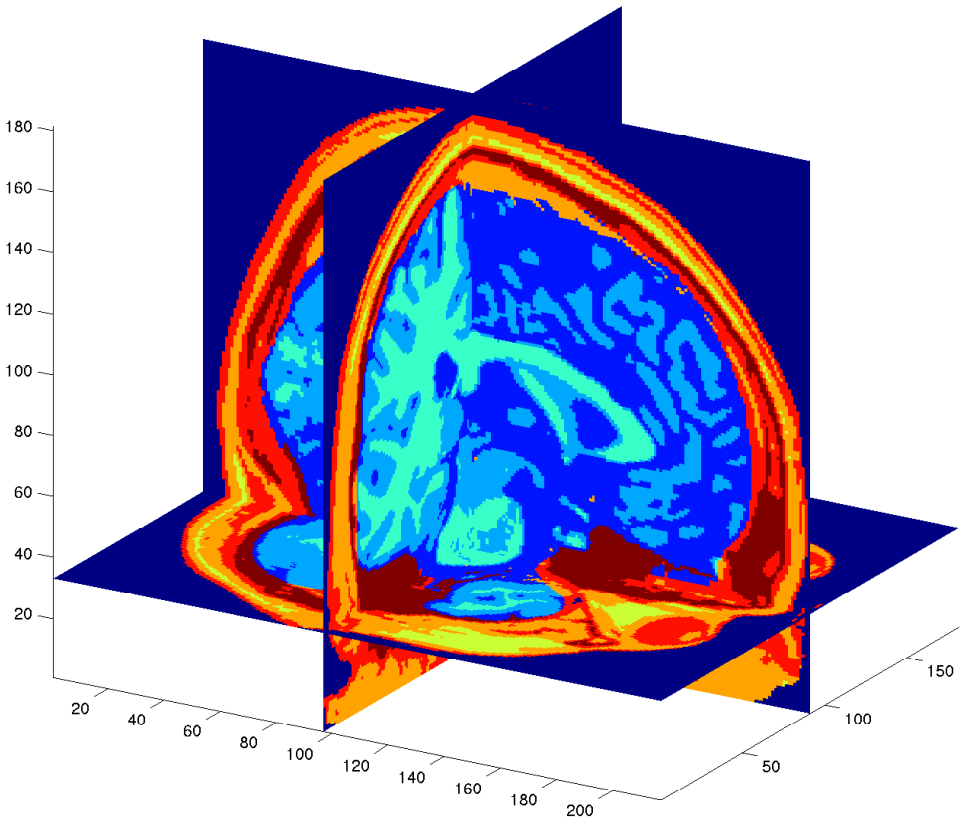


**Figure 3.2:** The original head model used in the simulations. The colors indicate different tissue labels.

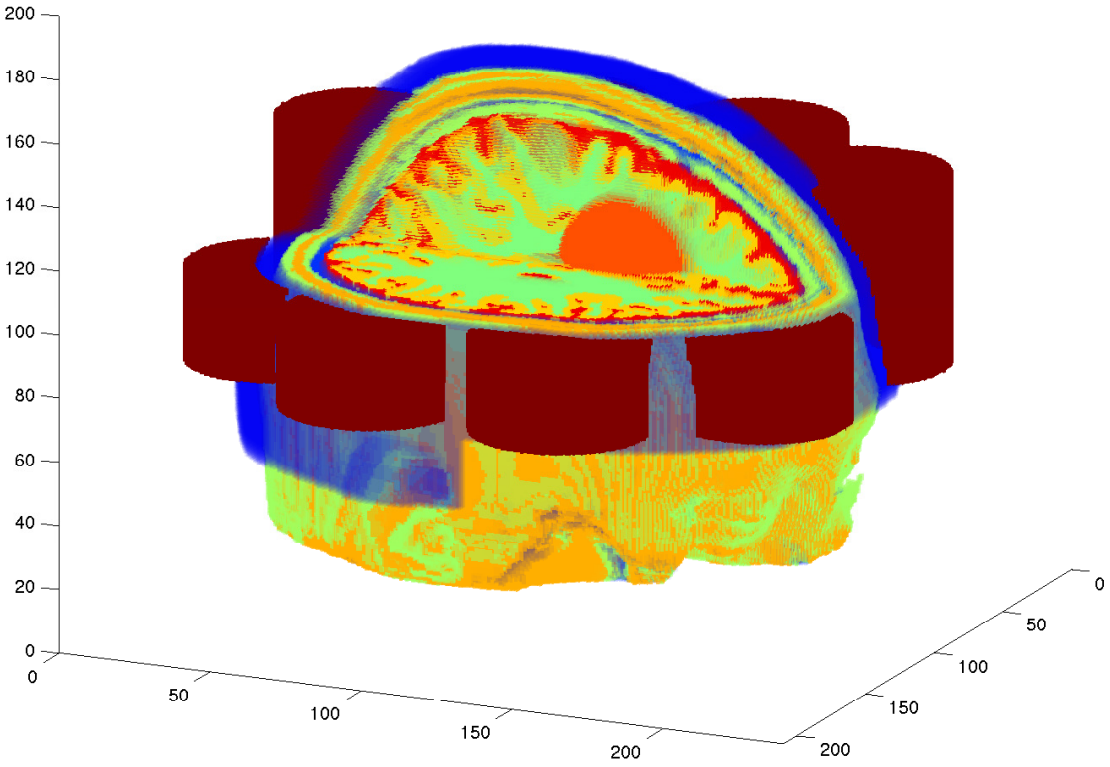


**Figure 3.3:** Rendering of the head of a simulated patient. The red cylinders are the water bags in which the antennas were placed. The blue layer is the hair that was added to the head and the orange sphere is the bleeding.

| Parameter | Distribution | Values/Range |
|---|---|---|
| Stretch in $x$ | Normal | $\mu = 0.855$, $\sigma = 0.0451$ |
| | | (Head size $x$: $\mu = 150$ mm, $\sigma = 8$ mm) |
| Stretch in $y$ | Normal | $\mu = 0.895$, $\sigma = 0.0537$ |
| | | (Head size $y$: $\mu = 188$ mm, $\sigma = 12$ mm) |
| Stretch in $z$ | Normal | $\mu = 0.738$, $\sigma = 0.0602$ |
| | | (Head size $z$: $\mu = 123.5$ mm, $\sigma = 10$ mm) |
| Hair thickness | Uniform | 0-5 pixels (0-5 mm) |
| Bleeding position in $x$ | Uniform | Dependent on bleeding size |
| Bleeding position in $y$ | Uniform | Dependent on bleeding size |
| Bleeding position in $z$ | Uniform | Dependent on bleeding size |
| Bleeding radius | Uniform | 9-33 pixels (Volume: 10-150 ml) |

**Table 3.2:** List of the variable parameters in the simulations and their respective ranges. $\mu$ is the mean and $\sigma$ is the variance for each parameter. Head size in $z$ is measured from the ear canal.

male heads in each of the three directions. The percentiles for the different directions for male and female heads was taken from *The Measure of Man and Woman: Human Factors in design, Revised Edition* by Alvin R. Tilley [41]. The tails of the distributions were cut off at the 1st and 99th percentile to make sure no head got to small or to large for the simulations space. This means that the smallest percentile of female heads and the largest percentile of male heads were not included in the simulations.

### 3.3.3 Dielectric properties of tissues

Once the model had been modified it was fed to a Matlab script that converted the 3D matrix of labels into two three dimensional matrices; one representing the permittivity of each voxel and one representing the conductivity of each voxel. The conductivity and permittivity at 1 GHz was used since this was the resonance frequency of the antennas used. The conductivity and permittivity of each tissue type except hair was taken from a web page maintained by Andreuccetti *et al.*[20]. Dielectric parameters of hair were not readily available and therefore an approximation using Bruggemanns formula (Section 2.2.7) was made. Hair was assumed to be long cylinders of keratin embedded in air. Nails also consist of keratin and the dielectric properties of nails were available. Bruggemanns formula as stated in Equation 2.11 is valid for spheres. Therefore the EMA was calculated for a mixture consisting of spheres of air and keratin in 2D. Optimal packing

| Label | Relative permittivity at 1 GHz [Unitless] | Conductivity [S/m] |
|---|---|---|
| Background/Air | 1 | 0 |
| CSF | 68.4 | 2.45 |
| Gray matter | 52.3 | 0.98 |
| White matter | 38.6 | 0.62 |
| Fat | 5.44 | 0.0535 |
| Muscle | 54.8 | 0.98 |
| Skin | 40.9 | 0.90 |
| Bone | 12.3 | 0.16 |
| Blood | 61.1 | 1.58 |
| Hair | 8.30 | 0.09 |
| Water | 78.0 | 0.20 |

**Table 3.3:** Conductivities and permittivities for the tissues and water bags in the simulations.

of cylinders give a packing density of 90% [42]. The hair was assumed to not be optimally packed and the volume fractions were assumed to be 20% air and 80% keratin. These spheres in 2 dimensions were then assumed to represent cylinders when extended to the third dimension. Both the conductivity and permittivity of hair was calculated in this way. The dielectric properties of all materials in the simulation can be seen in Table 3.3. The original source of the dielectric measurements is the work by Gabriel *et al.* [21]. These two tensors, one with the conductivity values and one with the permittivity, were then fed as input data to the FDTD solver.

### 3.3.4 FDTD simulations

Simulations were carried out using an FDTD-solver provided by Andreas Fhager at the Signals and Systems department of Chalmers Univeristy of Technology [43]. The solver was used to simulate different patients cases. The simulation space was $210 \times 240 \times 200$ voxels where the first dimension, $x$, represented right to left, the second, $y$, represented back to front and the third, $z$, represented down to up. Each voxel was set to represent 1 $mm^3$.

Eight dipole antennas with a length of 21 mm and a width of 0.2 mm were added to the simulation space in a circle around the imported head model. Each antenna had a water bag surrounding it to steer the transmission into the head of the patient instead of out into the open air. The placement of the water bags and antennas were constant for all measurements. They were placed so that the largest head did not come in contact with the antennas

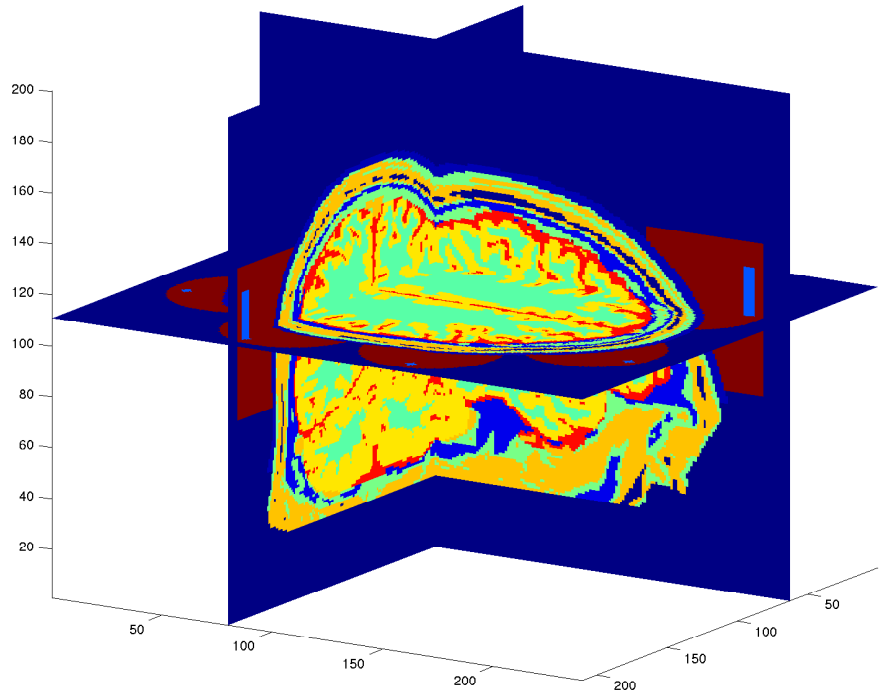


**Figure 3.4:** An example of a simulation space. The dark red is water bags and the light blue is representations of antennas.

and so that the smallest head still was in contact with the water bags but without the water bags being in contact with each other. Also there was at least five voxels of empty space between the water bags and the edges of the simulation space. The edges of the simulation space consisted of PML. An example of a simulation space with antennas and water bags can be seen in Figure 3.4. The simulations were carried out in the frequency range of 0 to 3 GHz. The conductivity and permittivity was assumed to be constant throughout the frequency range and no dispersion model was used.

In total 1000 simulations were carried out of which 500 simulated patients with a bleeding and 500 simulated healthy patients. No bleeding was placed in the healthy patients but head sizes and hair thicknesses still varied.

The simulations output both time domain data on transmissions and reflections of all antennas as well as frequency domain transmissions and reflection *i.e.* S-parameters.

### 3.3.5 Analysis

Firstly the distributions of the parameters that could be varied for each patient were checked to make sure that everything had been simulated correctly. After that the bulk of the analysis was carried out by classifying either all or different subsets of the simulations using the classifier described in section 2.5.5 and 3.1.

### 3.3.6 Investigating the effect of set sizes

The effect of number of patients on the results of the classification was studied. This was done by extracting subsets of patients from the 1000 simulated patients. The sizes of the patient sets were 20, 40, 60, 100, 140, 200, 400, 600, 800, 1000. Half of the simulated patients in each set were healthy and half had a bleeding. Each set except the set containing all 1000 patients was drawn randomly from the full set of 1000 patients. This randomized subset was then classified and a full parameter sweep of the internal parameters of the classifier was preformed to see how they affected the results of the classification. The frequency span was most often divided into two, 0 to 1.5 GHz and 1.5 GHz to 3 GHz. The classifier was either given only one of the two frequency sets or both. Preprocessing was also turned on and off to be able to study the effects of it. In total ten randomized subsets of each size except 1000 were created and classified as described above. For each subset the maximum AUC found by the parameter sweep of internal parameters and frequency span was recorded together with the mean of the AUC for all possible parameter combinations. When the optimal internal parameters had been found a run with higher frequency resolution was performed for the whole data set consisting of 1000 patients.

### 3.3.7 Investigating the effects of varying parameters

It was possible to extract the heads with the largest or smallest values for any of the parameters that could vary in the simulations. Subsets with large and small bleedings, large, small and midsized heads, and thick and thin hair were all classified separately to see how these effects affected the classification.

For the largest and smallest bleedings the bleeding class consisted of the $n$ largest or smallest bleedings in the data set. The healthy class consisted of equally many healthy patients.

For the thickest and thinnest hair both bleeding patients and healthy patients were selected based on their hair thickness and the $n$ patients with

| Region | $z$-limits | Number of bleedings | Total number of patients |
|---|---|---|---|
| Top of brain | $z > 112$ | 93 | 186 |
| Middle of brain | $89 > z > 100$ | 91 | 182 |
| Bottom of brain | $z < 76$ | 95 | 190 |

**Table 3.4:** Limits for the $z$ coordinates for bleedings from different sections of the brain.

thickest or thinnest hair from each class were used. The case was the same for head size. Head size was defined as the sum of the stretch in $x$, $y$ and $z$ and the $n$ smallest or largest heads could be extracted for each class. $n$ patients from the middle of the head size scale could also be extracted for each class.

Sets based on the $z$ position of the bleeding was also created. The goal was to have 100 bleedings from the top of the brain, 100 bleedings from the middle of the brain and 100 bleedings from the bottom of the brain. The limits on $z$ that gave these subsets are illustrated in Figure 3.5 and are given in Table 3.4. Each data set was filled with as many healthy patients as it contained bleedings. The healthy patients were drawn on random.

A subset of data only consisting of bleeding patients was also created. In this subset each patient was labeled based on the location of the bleeding. Either right hemisphere or left hemisphere. The classifier was then assigned the task to separate the patients with a bleeding in the left hemisphere from the patients with a bleeding in the right.

### 3.3.8 Adding noise

Normally distributed noise was added in some of the runs. This was done by adding a matrix consisting of random complex numbers to the measurement matrix of each patient. Each element in this complex matrix had it's imaginary and real part drawn from a normal distribution with mean zero and varying standard deviation generated by Matlab's `randn` command. A new matrix was generated for each patient. The standard deviations ($\sigma$) used can be seen in Table 3.5. An illustration of these noise levels in relation to the signal strength can be seen in Figure 3.6. The noise was mainly added to data sets consisting of the largest bleedings since the effect of the noise should be more apparent on a data set that is easy to classify when it it noiseless.

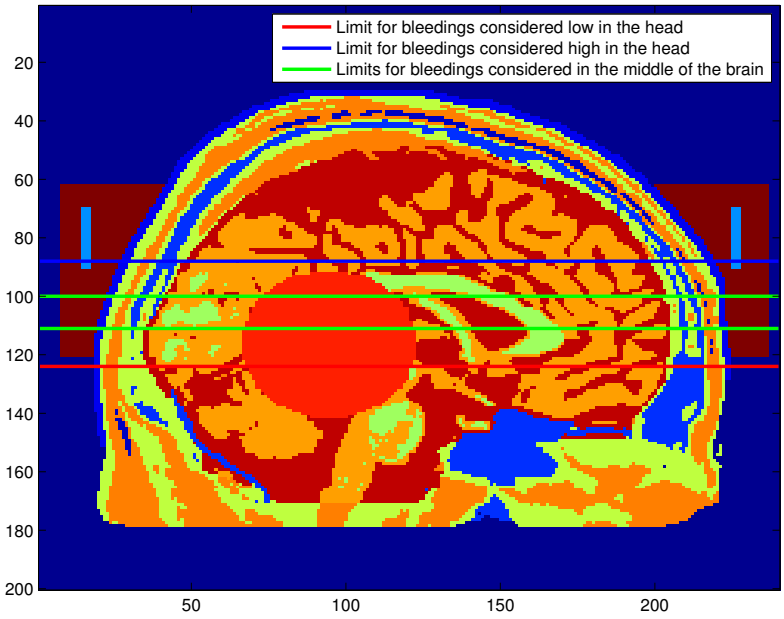


**Figure 3.5:** Illustration of the limits for the separation of bleedings considered high, in the middle or low in the brain. Bleedings with their center above the blue line were considered high. Bleedings with their center between the green lines were considered to be in the middle part of the brain and bleedings below the red line were considered to be low in the brain. The bleeding depicted in the figure would not belong to any of the classes since it's center does not lie in one of the three regions. Antennas and water bags have been added as reference to visualize their location relative to the limits.

| | |
|---|---|
| $\sigma_1 = 0$ (No noise) | $\sigma_7 = 2.2 \cdot 10^{-3}$ (-53 dB) |
| $\sigma_2 = 1 \cdot 10^{-5}$ (-100 dB) | $\sigma_8 = 5 \cdot 10^{-3}$ (-46 dB) |
| $\sigma_3 = 1 \cdot 10^{-4}$ (-80 dB) | $\sigma_9 = 1 \cdot 10^{-2}$ (-40 dB) |
| $\sigma_4 = 2 \cdot 10^{-4}$ (-74 dB) | $\sigma_{10} = 1 \cdot 10^{-1}$ (-10 dB) |
| $\sigma_5 = 5 \cdot 10^{-4}$ (-66 dB) | $\sigma_{11} = 1 \cdot 10^{0}$ (0 dB) |
| $\sigma_6 = 1 \cdot 10^{-3}$ (-60 dB) | |

**Table 3.5:** The different standard deviations used for the noise.

### 3.3.9 Higher frequency resolution

The data set consisting of the 100 largest bleedings was also classified with different levels of noise and a better frequency resolution. Instead of dividing the frequency interval of the simulation (0-3 GHz) into two frequency spans it was divided into six frequency spans. This was done to be able to analyze the effect of the noise on different frequency spans. A full sweep of the internal parameters in the classifier was preformed for each frequency band.

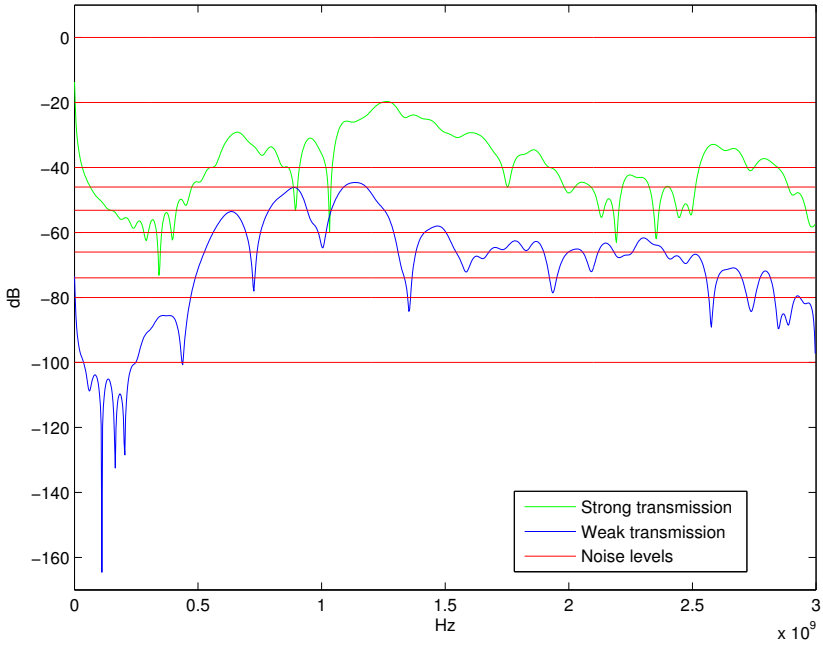


**Figure 3.6:** The different levels of noise in relation to the transmission with the maximum amplitude peak (Green line) and the transmission with the maximum amplitude dip (Blue line).

# Chapter 4

# Results

## 4.1 Thrombolysis analysis

Figure 4.1 shows the absolute drops in NIHSS and distance output by the classifier. The results were similar indipendent of the settings of the classifier or correction algorithm.

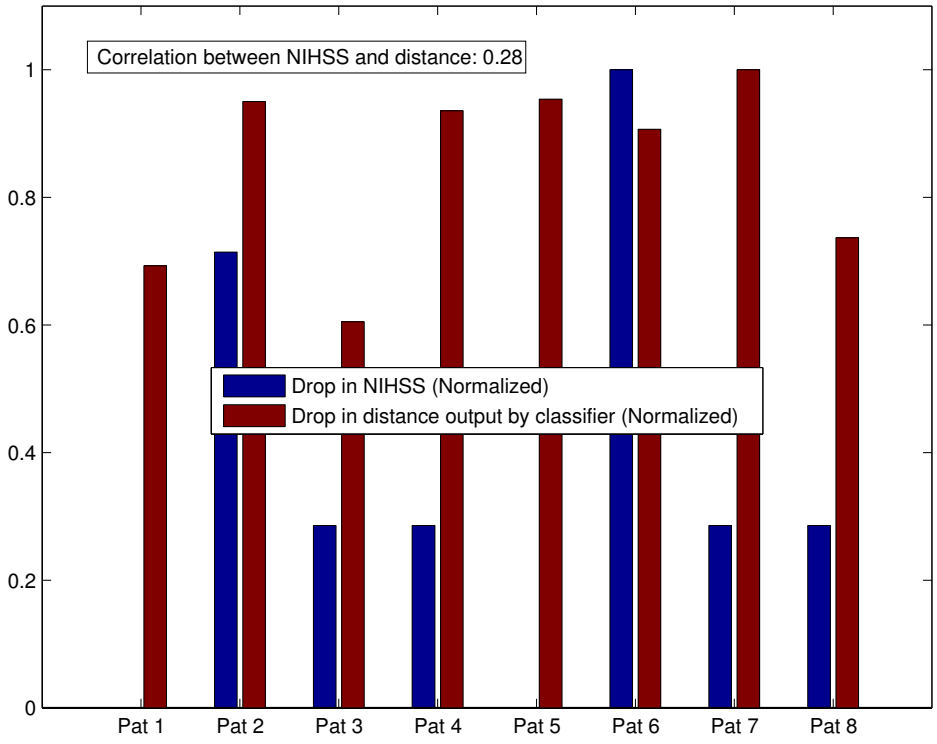


**Figure 4.1:** Figure of maximum absolute drop in NIHSS and distance output by the classifier for each patient. The correlation between the two measures is also presented. The analysis was performed with the classifier set to not remove large variations and correction was turned on.

### 4.1.1 Single patient results

Figure 4.2 to 4.5 show results on a single patient level and illustrate what happens with the measurements from one patient when the parameters change and correction is turned on and off. The distance in clot decreases with time while the distance in healthy increases leading to a decreasing difference in distances for all patients and settings. NIHSS decreases with time as well. Sharp jumps corresponding to drops in correlation can be seen in the distance if neither correction nor setting the classifier to remove large variations is turned on (Figure 4.2). Most of the sharp jumps are removed by turning either of the two on (Figure 4.4 and 4.3). Figure 4.5 shows the distances output by the classifier for patient 5 when we look at the S-parameters of the reflection instead of the transmission. Jumps in the two distance belonging to each class can still be seen but they lie more or less on top of each other. This means that the change of difference in distances is very small.

Turning scrambling on for the transmissions without correction produces results similar to the ones seen for the reflections in Figure 4.5. The distances to the two classes are basically the same for all time instances and the difference in distances is very close to 0. Just like the reflections the distances show some sharp jumps and some slow drifts. The scale is however in the thousands as for the other transmissions and not in the tens as for the reflection.

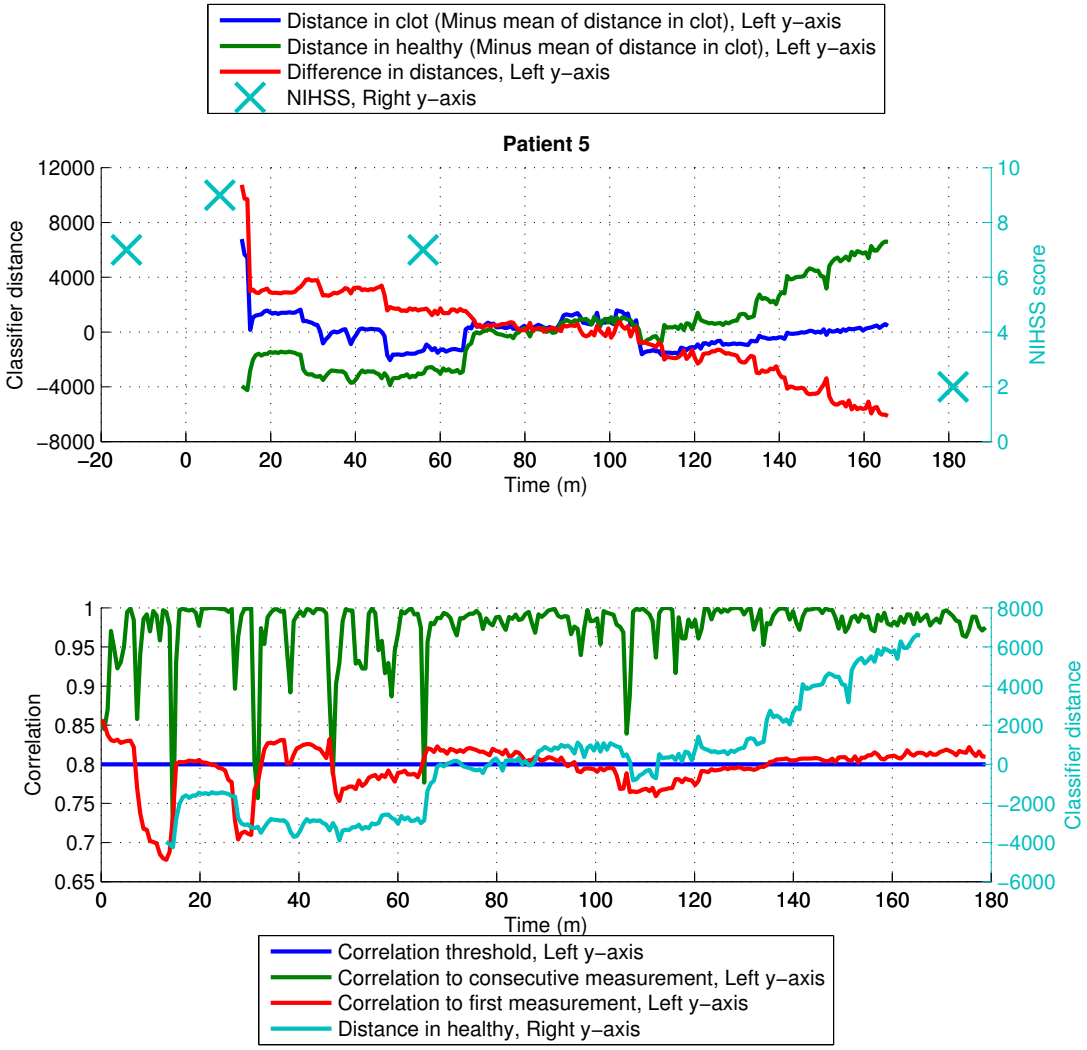


**Figure 4.2:** Figure of the measurement for patient 5. Correction was turned off, the classifier was not set to remove large variations and only the S-parameters of the transmission were considered. The mean of the distance in clot has been removed both from the distance in clot and the distance in healthy to increase the visual content of the figure. This does not affect the difference of the distances.

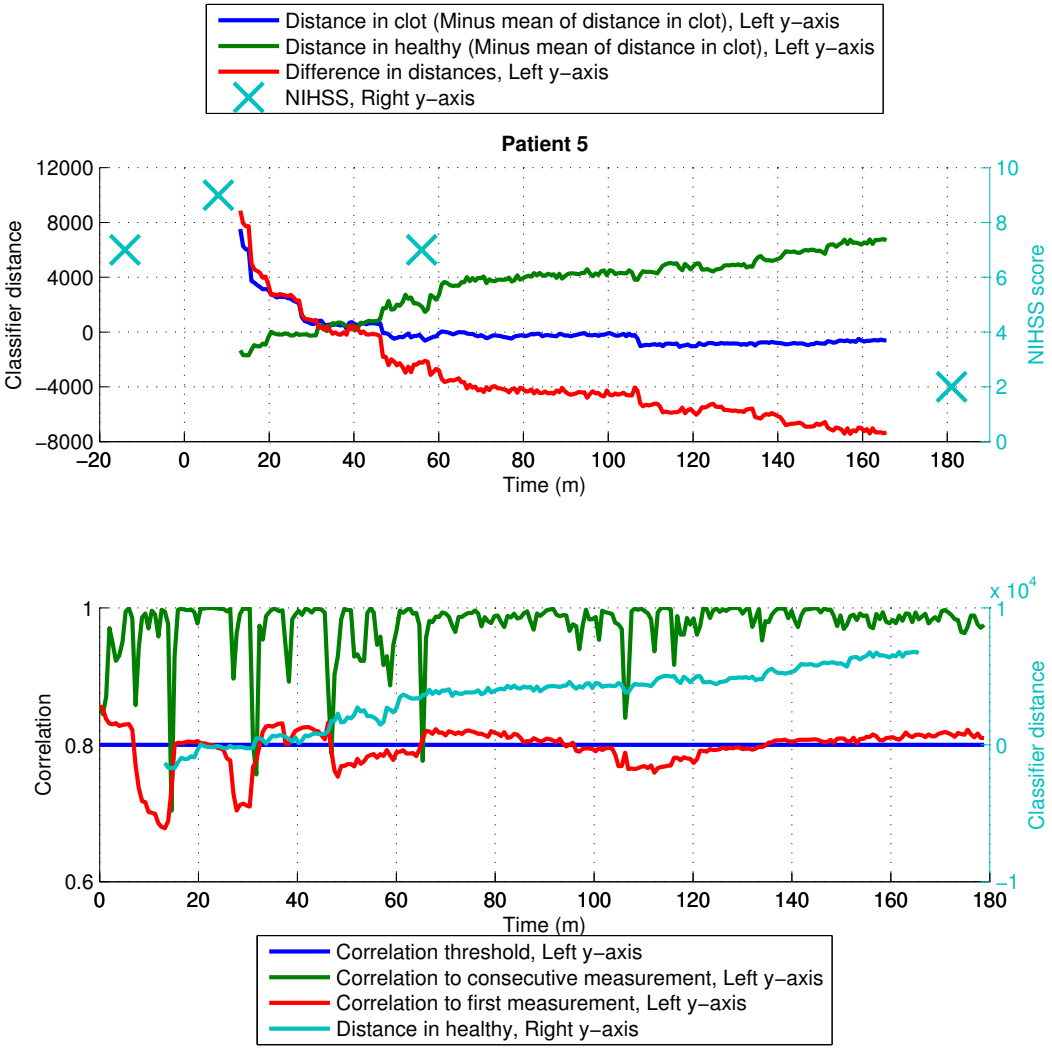


**Figure 4.3:** Figure of the measurement for patient 5. Correction was turned on, the classifier was not set to remove large variations and only the S-parameters of the transmission were considered. The mean of the distance in clot has been removed both from the distance in clot and the distance in healthy to increase the visual content of the figure. This does not affect the difference of the distances.

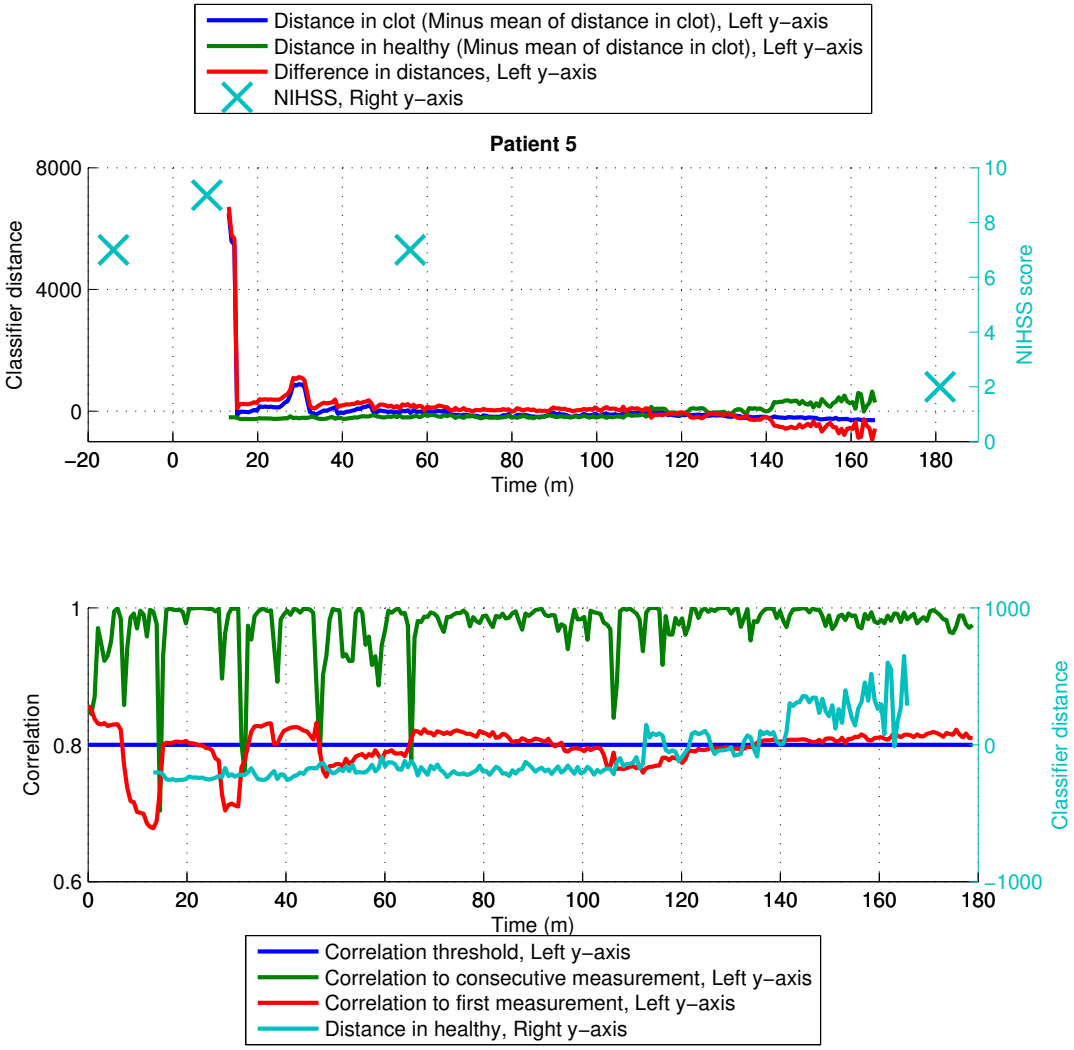


**Figure 4.4:** Figure of the measurement for patient 5. Correction was turned off, the classifier was set to remove large variations and only the S-parameters of the transmission were considered. The mean of the distance in clot has been removed both from the distance in clot and the distance in healthy to increase the visual content of the figure. This does not affect the difference of the distances.

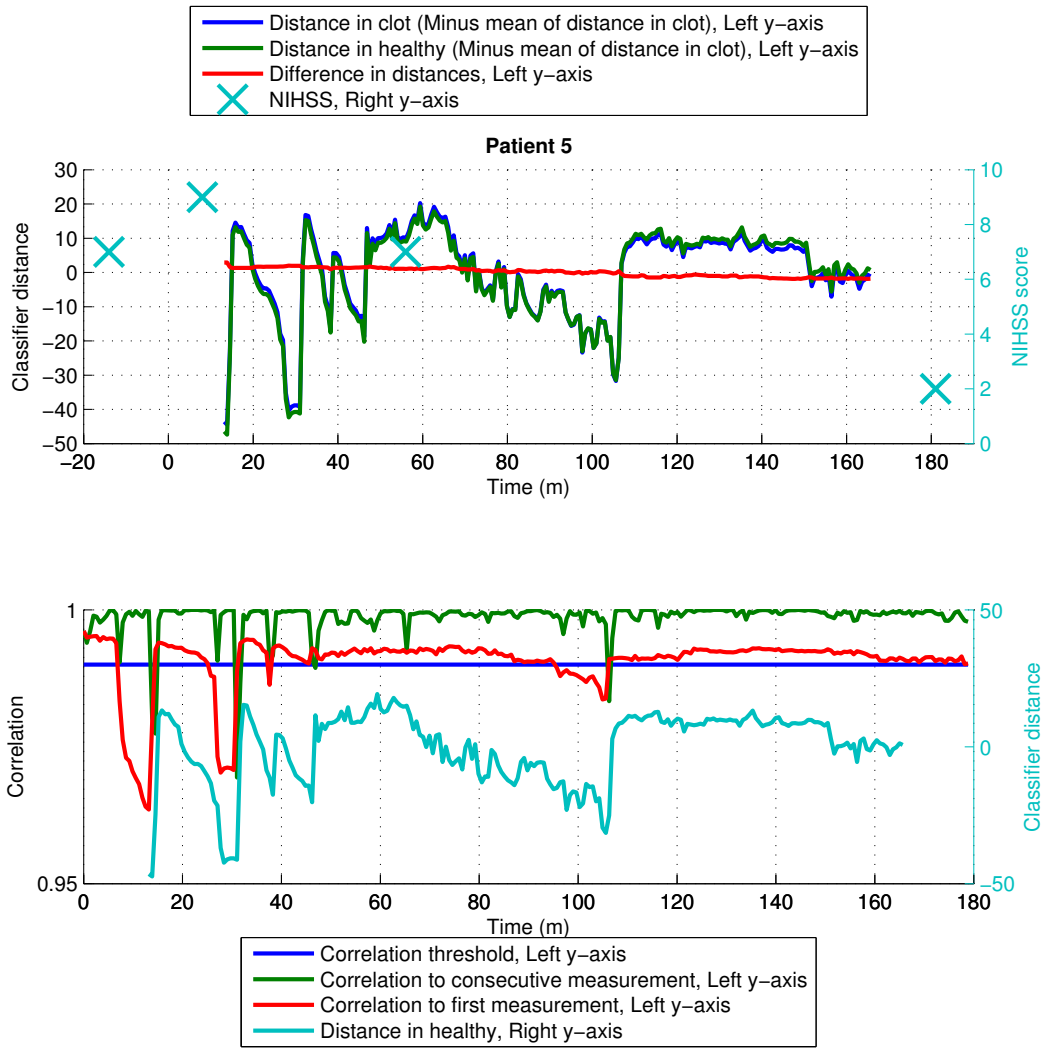


**Figure 4.5:** Figure of the measurement for patient 5. Correction was turned off, the classifier was set to not remove large variations, and only the S-parameters of the reflection were considered. The mean of the distance in clot has been removed both from the distance in clot and the distance in healthy to increase the visual content of the figure. This does not affect the difference of the distances. Observe the change of scale in both the lower subplot and in the absolute value of the distances output by the classifier.

## 4.2 Simulations

The variable parameters in the simulation were examined to see if they followed the expected behaviors. The approximate distributions can be seen in Figure 4.6.

The maximum and mean AUC for different sizes of patient sets when preprocessing was turned off can be seen in Figure 4.7. Turning preprocessing on lead to the smaller data sets performing slightly better while the bigger (>200) performed worse.

The results from the runs where subsets based on bleeding size, head size, hair thickness and bleeding position had been extracted are presented in Table 4.1.

The effects of adding different levels of noise can be seen in Figure 4.8. The noise was added to data sets consisting of the largest bleedings.

Figure 4.9 shows results for the data set consisting of the 100 largest bleedings and 100 healthy patients. The frequency span of the simulation (0-3 GHz) was divided into six frequency bands and the maximum and mean AUC for the parameter sweep is presented for each noise level.

For frequency the best classification was achieved for the lower half of the frequency spectrum. The run with higher frequency resolution on the whole data set consisting of 1000 patients showed that the optimal frequency was between 200-800 MHz and the highest AUC attained was 0.93.

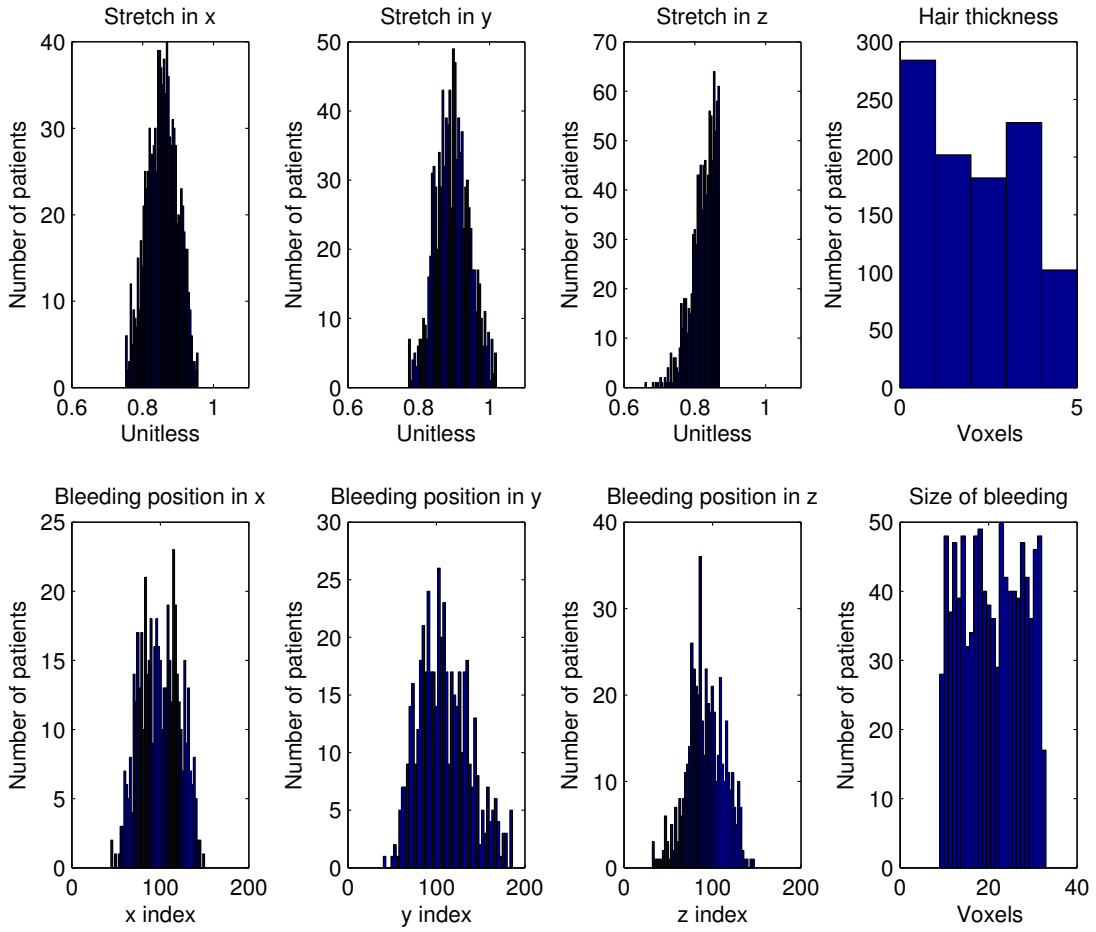


**Figure 4.6:** Distributions of the variable parameters in the model.

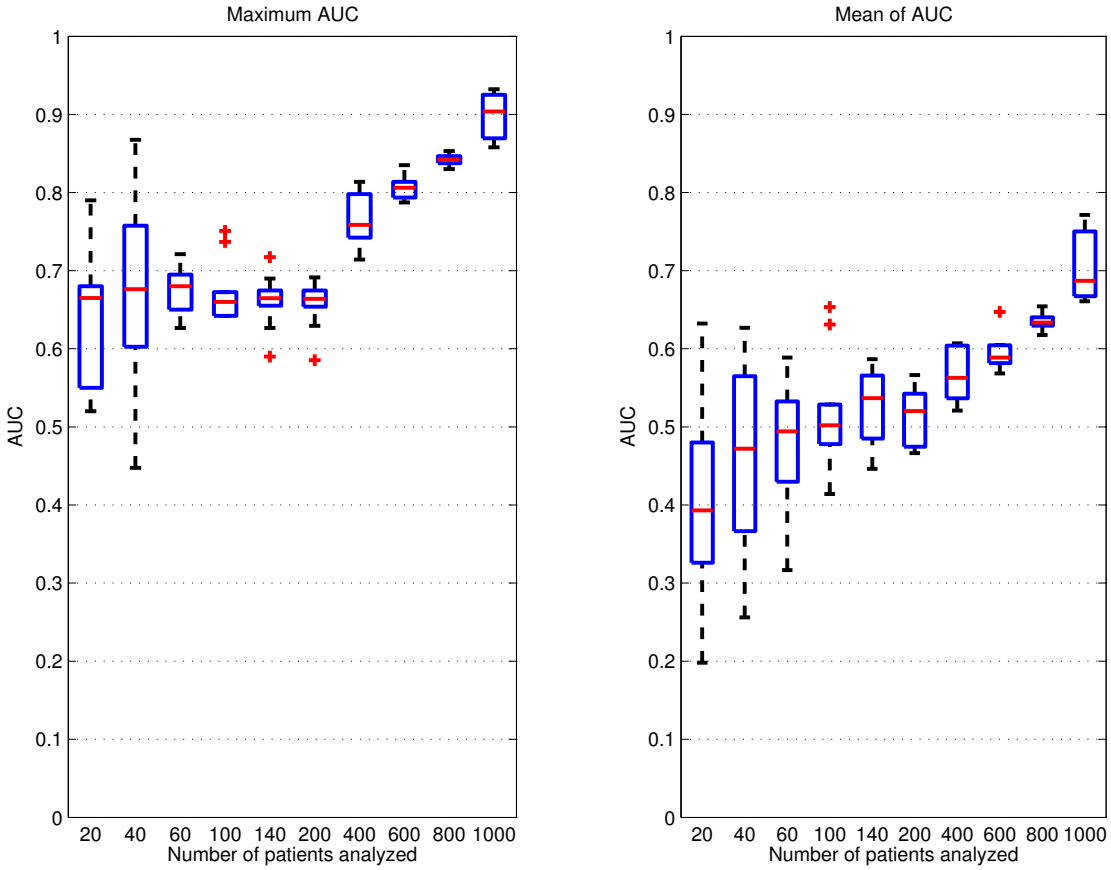


**Figure 4.7:** The change in maximum AUC and mean AUC with increasing patient sets. Each box plot consists of ten random subsets of the size specified on the x-axis. On each box, the central red line marks the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers. Outliers (points outside of $\pm$ 2.7 $\sigma$) are plotted individually. The plot above shows the results when preprocessing was turned off.

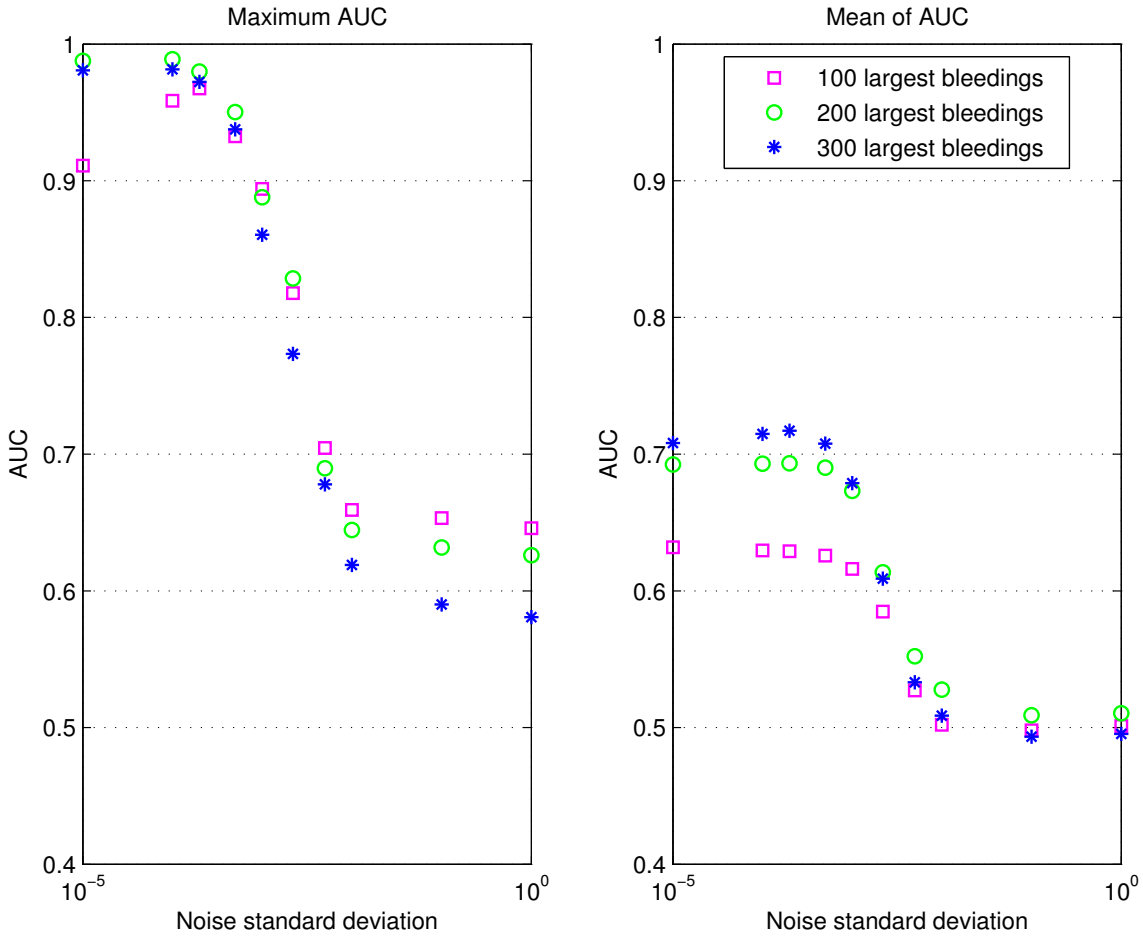


**Figure 4.8:** Change in maximum and mean AUC with increasing noise for different sizes of data sets. Data sets consisting of the largest bleedings were used since they were easy to classify when noiseless therefore showing a clearer change with increasing noise Each data set contains as many healthy patients as it does bleedings. The $x$-axis is plotted in log scale.

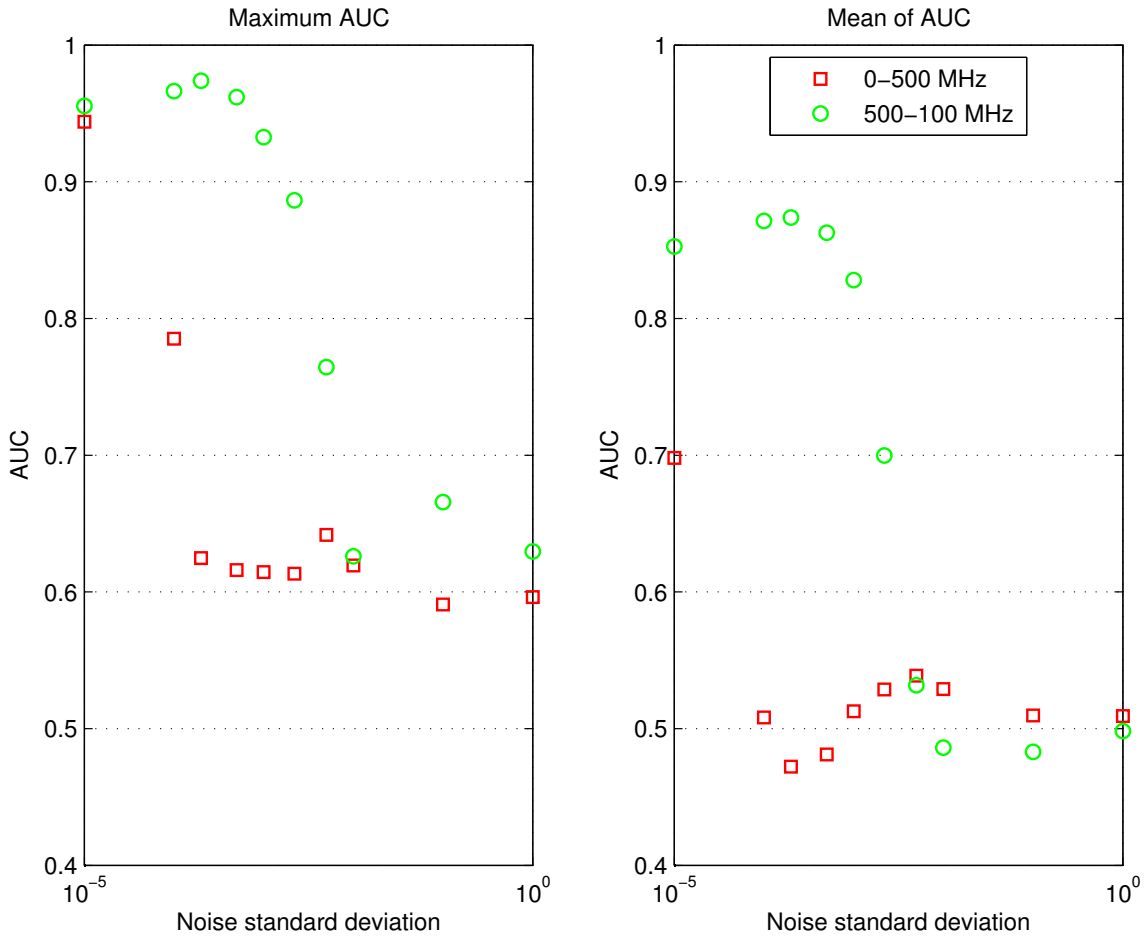


**Figure 4.9:** Change in maximum and mean AUC with increasing noise for two frequency bands. The data set consisted of the 100 largest bleedings and 100 healthy patients.The $x$-axis is plotted in log scale.

| Subset | Maximum AUC | Mean AUC | Total number of patients |
|---|---|---|---|
| 100 smallest heads | 0.77 | 0.62 | 200 |
| 100 middle heads | 0.79 | 0.59 | 200 |
| 100 largest heads | 0.78 | 0.53 | 200 |
| 100 smallest bleedings | 0.60 | 0.46 | 200 |
| 100 largest bleedings | 0.94 | 0.64 | 200 |
| 100 thickest hair | 0.73 | 0.56 | 200 |
| 100 thinnest hair | 0.82 | 0.61 | 200 |
| Bleedings on left versus bleedings on right | 0.83 | 0.62 | 500 (Only bleedings) |
| Bleedings high in brain | 0.81 | 0.63 | 186 |
| Bleedings in middle of brain | 0.81 | 0.58 | 182 |
| Bleedings low in brain | 0.64 | 0.56 | 190 |

**Table 4.1:** Results from the classification based on different subsets.

# Chapter 5

# Discussion

## 5.1 Thrombolysis analysis

### 5.1.1 Correlation between NIHSS and distance

All patients display a decreasing distance in the clot subspace, an increasing distance in the healthy subspace and a decreasing difference in distance over time. These changes are due to the fact that the classifier has been trained on the first and last 20 samples and is told that these belong to different classes and therefore tries to separate them. Therefore the drop itself is not of interest but rather if the size of the drop correlates with the size of the drop in NIHSS for each patient.

No correlation can be seen between the maximum absolute drop in distance output by the classifier and the maximum absolute drop NIHSS score of the patient. This was true irregardless of what parameter set that was used or if the correction algorithm was turned on or off (See Figures **??** to 4.1). One reason for this might be that the data set is simply to small. If the correlation is weak eight patients might not be enough to detect it. It might also be naive to think that the physiological changes in the brain during thrombolysis treatment can be described by two classes. One representing "disease" and one "healthy". It is possible that more classes are needed to describe the change over time as the clot dissolves. Gathering training data for these hypothetical classes poses a large problem though since the patient would have to be monitored very closely during the thrombolysis treatment to determine which phase the patient is for each measurement.

### 5.1.2 Single patient results and noise removal

Figure 4.2 illustrate the results for one patient when correction is turned off and with the classifier set to not remove large variations. The change in clot distance and healthy distance over time are suspected to come from two sources. One being the dissolving of the clot and return of blood-flow to the affected area, which is the change that is desirable to detect. The other being noise, due to movement of the patient affecting the antennas and cables and different kinds of drift in the equipment. The movement artifacts is suspected to act on a much faster timescale than the dissolving of the clot. The noise due to movement can be seen as a random step in the highly dimensional feature space and as described in section 2.6 these steps will not have a mean value of zero. Therefore a regular low pass filter cannot be employed.

Both the correction algorithm (Figure 4.3) and setting the classifier to remove large variations (Figure 4.4) seem to remove most of the sudden jumps in the difference in distances output by the classifier.

The version relying on a threshold has two main drawbacks, both stemming from the fact that it relies on a threshold. This threshold has to be set manually introducing an artificial parameter that is based on visual inspection in this report and which is not understood from a physical viewpoint or derived from a deeper understanding of the system. The threshold also acts as a hard limit meaning that if we are just above the threshold nothing is removed but as soon as we fall below it everything is removed. An adaptive threshold might be implementable alleviating some of the issues with this approach. The threshold approach has the advantage of being easy to understand however and it is also easy to detect if it is working as intended.

The solution to remove the movement noise through internal parameters of the classifier also has some drawbacks and benefits. One could argue that this is also an arbitrarily chosen threshold but this choice is at least based on extensive experience with the Medfield Diagnostics Strokefinder R10. The drawback of this solution is that it is harder to understand exactly what happens since it is unclear exactly what information is removed by the classifier.

The threshold based correction (Figure 4.3) still leaves a clear declining trend in the difference in distances. It is unclear if this trend is due to drift in the equipment, motion artifacts not corrected for due to not dropping below 0.8 or due to erroneous changes introduced by the correction algorithm. Setting the classifier to remove large variations removes more of the big jumps but does not seem to improve the correlation.

A few conclusions can be drawn from Figure 4.5. The clot distance and healthy distance basically overlap. This indicates that the two training sets contain the same information since the projected distance is the same for both of them. This in turn means that the reflection measurements from the beginning of the measurement are the same as those from the end and no drift has occurred.

Scrambling the training sets also produce overlapping distances. There are still movement up and down of the two overlapping distances. If it is assumed that this movement happens in dimension that are not fully covered by the subspace of the training sets and we further assume that the training sets contain the variation that we want to detect, i.e. the change from clot to healthy, it could be argued that the movements we see when the classes are scrambled is all due to noise. The problem is that we only see one distance in this multidimensional space which makes it hard to correct for. The distance that the classifier outputs takes all these dimensions into account and weighs them to try to find the most informative dimensions to base this distance on. However it still does not provide us with information on how every dimension in this multidimensional space has varied.

## 5.2 Simulations

The goal of the simulations is to create a controlled environment which makes it possible to vary specific parameters to see how this affects the classification. Optimally the simulations should correspond to clinical data meaning that the classifier should give the same output for the simulation based on clinical data as it does for the clinical case itself. This was not tested due to lack of time. In hindsight it might also not be achievable since a lot of the variability in the data is suspected to be due to cable movements, antenna characteristics, and other measuring equipment which was not be modeled or simulated.

Simulations can still however give valuable insight on which parameters that effect the measurement the most and what kind of changes in the measurements these parameters can give rise to.

Overall the simulations seem to produce results that are relatable to the clinical case. The transmission and reflection curves, of which two transmission curves can be seen in figure 3.6, look similar to those attained in the patient measurements. Also as expected smaller data sets are harder to classify than larger ones, noise affects the classification in a negative way and smaller bleedings are harder to detect than larger ones.

### 5.2.1 Distributions of variable parameters

Figure 4.6 shows the expected behaviors for the stretch in $x$ and $y$, they look like normal distributions with their tails cut off. Stretch in $z$ however looks to be skewed towards larger values. After inspecting the setup file for the simulations it was discovered that the mean for the stretch in $z$ had been entered incorrectly which lead to a larger stretch in $z$ than intended. This should however not affect the quality of the results to much. The stretch in $z$ should affect the transmission the least since it does not directly increase the distance that the signal has to travel trough the skull. It however means that the simulations no longer represent the first to 99th percentile of female and male heads.

Hair thickness and size of bleeding look like uniform distributions as expected. The position of the bleeding should also be a uniform distribution but it does not look like one. Inspection of the setup file showed that they were indeed uniform and the reason for their distribution is the shape of the brain. Each voxel in the brain was equally likely to receive a bleeding. However the brain is ellipsoid so the distribution of pixels is not uniform when looked at from the front for example. Due to this the distributions do not look uniform.

### 5.2.2 The effect of set size

Figure 4.7 shows that both the mean of the AUC for all parameter combinations as well as the maximum AUC increase with the size of the data set. The variance in the results also decrease with the size of the data set. This could be due to the data set increasing but the fact that there were only 1000 patients available probably also plays a part. When 20 patients are to be drawn from 1000 ten times it is quite likely that all the sets contain unique patients. However when 800 patients is to be drawn ten times all the sets are bound to contain at least 600 patients that are also part of other sets leading to less variation. Another interesting result is that the median of maximum AUC stays around 0.67 for patient sets smaller than 400 after which AUC increases with the size of the patient set. This indicates that the classifier needs a number of samples before it can start to pick out strong patterns in the data and actually improve the classification when given more data. This turning point is most likely a function of the dimensionality and variability of the data fed to the classifier. In the case of these simulations the turning point seems to be somewhere between 200-400 patients. The maximum AUC of around 0.87 in these runs is also a promising result for

the possible viability of Medfield's equipment. It could be further increased by increasing the frequency resolution which gave a classification of 0.93 around 200-800 MHz. This optimal frequency band probably depends on the antennas and on the coupling of them to the head. It is however promising that a good classification can be gained in a fairly broad frequency band since at least the upper parts of it should be realizable in lab experiments.

### 5.2.3 The effect of noise

Figure 4.8 indicates that a small gain in the maximum AUC of the classification can actually be gained from adding small amounts of noise. This rise is more apparent for the smaller data set of the 100 largest bleedings. The figure also highlights the size of the bleeding as an important factor for the outcome of the classification. If only the 200 largest bleedings are considered an AUC of 0.98 is attainable. The 200 largest bleedings fall in a range of 80-150 ml however which is quite large. The classification copes well with noise up to a standard deviation of $2 \cdot 10^{-4}$ (The fourth point from the left in Figure 4.8, third red line from the bottom in Figure 3.6) which corresponds to a noise floor of -74 dB after which it falls with increasing noise. With a noise floor of -60 dB (Sixth point in Figure 4.8, fifth red line from the bottom in Figure 3.6) the AUC has dropped to around 0.85-0.9. This noise level should be possible to realize in the lab.

The amplitude of the transmission is the weakest in the frequency range of 0-500 MHz (See Figure 3.6) and as expected the classification based on this frequency band is corrupted by noise of lower levels than the classification based on a higher frequency band as indicated by Figure 4.9.

### 5.2.4 Effect of single parameters on the classification

The different subsets listed in 4.1 show some trends. The first and most important is that larger bleedings are easier to classify than smaller ones. This is expected since larger bleedings should affect the overall conductivity and permittivity of the simulated head more than a small bleeding. The second trend is hair. Thicker hair seems to have a negative effect on classification. This is also expected. The conductivity and permittivity of hair differs significantly from the conductivity and permittivity of both the water bags on one side and the head on the other. This means that large parts of the signal are reflected at these two interfaces. Also the hair itself is a poor conductor of electromagnetic radiation. The last trend is that bleedings low in the brain are hard to classify. This can be explained by the fact that these

bleedings are not in the path between any two antennas. This means that the signal has to bounce and diffract a lot to reach the bleeding and then make it back to the receiving antenna. Most likely only a small portion of the radiation takes the path via the bottom of the brain. This means that the change in the signal between a healthy patient and a patient with a bleeding in the bottom of the brain probably is minute. This highlights the importance of placing antennas in a way that puts as much of the brain as possible in the line of sight between two antennas.

The fact that it is possible to classify bleedings on the right side of the brain from bleedings on the left with an AUC of 0.83 is also interesting. A desired feature in the final product is to be able to tell the position of the clot or at least which hemisphere it is in since it is of importance to the clinicians. This result indicates that it should be possible to say something about the position of the clot.

### 5.2.5 Preprocessing

Initially a large part of the simulations were carried out using the preprocessing described in section 3.1. Turning this off however yielded AUC values that were on average 0.15-0.2 larger than with preprocessing turned on. Why the preprocessing leads to a reduction in the AUC is unclear. One reason could be that the method of preprocessing used in this thesis might have been developed for using data from both the transmissions and reflection. The reflected signal is often several orders of magnitude stronger than the transmitted signal which means that it could be reasonable to normalize the individual measurements if they are to be used in the same analysis. In this thesis however we only looked at transmissions and in this case the normalization in the preprocessing seems to remove essential information. All results presented in this report have been gathered with the preprocessing turned off.

# Chapter 6

# Conclusions

## 6.1 Thrombolysis analysis

No correlation could be seen between the NIHSS and the distance output by the classification algorithm using the approaches in this report. Using the first 20 and last 20 samples as training data for each patient is probably not a good approach as the results from the simulation indicate that a much larger training set is needed before the classification algorithm starts outputting stable results. Using the first and last measurements as training data also creates an inherent change in distance from one class to the other which is hard to separate from the change in distance due to the dissolving clot.

## 6.2 Simulations

The results from the simulation indicate that it is a good tool for modelling stroke patients with a bleeding. This is indicated by the fact that the S-parameters look like expected and the results from the classifications. Classification of subsets of the data are in line with expected results. Smaller clots are harder to classify than large ones, thicker hair causes problems for the classifier and a larger training set gives a better classification. There seems to be a turning point between 100–200 samples of each class after which the classification improves with increased training sets. The highest AUC attained for the whole data set was 0.93 and the preprocessing seems to play a very important role in the classification.

## 6.3 Future work

### 6.3.1 Thrombolysis analysis

A lot remains to be understood in the chain of events behind the thrombolysis analysis experiments. How does the NIHSS-score relate to the physiological state of the brain, how does the thrombolysis affect the physiological state of the brain? What of this does the R10 system detect and how does that relate to the distance it outputs? At what timescales do these effects act? Since around 100 measurements seems to be needed in each class and the measurement from each patient consists of between 100-250 measurements it is probably not possible to analyze each patient separately. Two other approaches are possible. One is to pool the first and last $n$ measurements from the thrombolysis patients and use these as the training sets. Thereby attaining more training samples for each class. The other approach is to use the existing measurements on patients with clot as one training set and gather measurements from healthy volunteers for the other training set.

Thought also has to be put into if one training class of patients with clot (or from the beginning of the measurement) and one class of healthy patients (or from the end of the measurement) is enough to detect the physiological changes in the brain during thrombolysis treatment. More classes might be needed.

### 6.3.2 Simulations

- Redo simulations with frequency dependent conductivity and permittivity and study effects.

- Develop ways to approximate cables, other types of antennas and different kinds of noise. The possibility to investigate movement artifacts would also be valuable.

- Currently the bleeding replaces the original tissues of the brain. A model in which the bleeding instead deforms the brain would be more realistic. This model would make it possible to study midline shift for example which is an important indicator for the clinicians.

- Replace the uniform distribution of bleedings position with a distribution that reflects the probability of bleedings at different positions in the brain.

- Investigating the similarity of clinical cases and simulations that try to model the clinical case.

# Bibliography


[1] K. Muir, "Stroke," *Medicine*, vol. 37, no. 2, pp. 109–114, 2009.

[2] K. Strong, C. Mathers, and R. Bonita, "Preventing stroke: saving lives around the world," *Lancet Neurol*, vol. 6, no. 2, pp. 182–187, 2007.

[3] G. Donnan and M. Fisher, "Stroke," *Lancet*, vol. 371, no. 9624, p. 1612, 2008.

[4] O. Ghatnekar, U. Persson, E.-L. Glader, and A. Terént, "Cost of stroke in Sweden: an incidence estimate," *International Journal of Technology Assessment*, vol. 20, no. 3, p. 375, 2004.

[5] J. Persson, J. Ferraz-Nunes, and I. Karlberg, "Economic burden of stroke in a large county in Sweden," *BMC Health Services Research*, vol. 12, no. 1, 2012.

[6] A. G. Thrift, H. M. Dewey, R. A. Macdonne, J. J. McNeil, and G. A. Donnan, "Incidence of the major stroke subtypes: Initial findings from the North East Melbourne stroke incidence study (NEMESIS)," *Stroke; a journal of cerebral circulation*, vol. 32, no. 8, pp. 1732–1738, 2001.

[7] H. van der Worp and J. Gijn, "Acute ischemic stroke," *The New England Journal of Medicine*, vol. 357, no. 6, pp. 572–579, 2007.

[8] W. Hacke, T. Machnig, D. Schneider, R. von Kummer, and N. Wahlgren, "Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke," *The New England Journal of Medicine*, vol. 359, no. 13, pp. 1317–1329, 2008.

[9] Socialstyrelsen, *Nationella riktlinjer för strokesjukvård 2009*, Web and print, 2009.

[10] K. Fassbender, C. Balucani, S. Walter, S. R. Levine, A. Haass, and J. Grotta, "Streamlining of prehospital stroke," *Lancet Neurology*, vol. 12, no. 6, pp. 585–596, 2013.

[11] W. Stomp, "Specialized stroke ambulance features CereTom portable CT scanner to reduce time to treatment," 2012, [Online]. Available: http://www.medgadget.com/2012/06/specialized-stroke-ambulance-features-ceretom-portable-ct-scanner-to-reduce-time-to-treatment.html.

[12] M. A. Khorshidi, T. McKelvey, and M. Persson, "Classification of microwave scattering data based on a subspace distance with application to detection of bleeding stroke," *IEEE*, p. 301, 2009.

[13] M. A. Khorshidi, "Classification of microwave scattering data with application to detection of bleeding stroke," Master's thesis, Chalmers University of Technology, 2009.

[14] D. R. Tomlinson and N. J. Gardiner, "Glucose neurotoxicity," *Neuroscience*, vol. 9, no. 1, pp. 36–45, 2008.

[15] C. Gilkes and P. Whitfield, "Intracranial pressure and cerebral blood flow," *Surgery*, vol. 25, no. 12, pp. 530–535, 2007.

[16] K. W. Muir, C. J. Weir, G. D. Murray, C Povey, and K. R. Lees, "Comparison of neurological scales and scoring systems for acute stroke prognosis," *Stroke*, vol. 27, no. 10, pp. 1817–1820, 1996.

[17] S. Martin-Schild, K. C. Albright, J. Tanksley, V. Pandav, and E. B. Jones, "Zero on the NIHSS does not equal the absence of stroke," *Ann Emerg Med*, vol. 57, no. 1, pp. 42–45, 2011.

[18] National Institutes of Health, "NIH stroke scale," [Online]. Available: http://www.ninds.nih.gov/doctors/NIH_Stroke_Scale.pdf.

[19] D. M. Pozar, *Microwave Engineering*, 4:th. John Wiley and Sons, Inc., 2012.

[20] D. Andreuccetti, R. Fossi, and C. Petrucci, "An internet resource for the calculation of the dielectric properties of body tissues in the frequency range 10 hz - 100 ghz," Based on data published by C.Gabriel et al. in 1996., 1997, [Online]. Available: http://niremf.ifac.cnr.it/tissprop/.

[21] S Gabriel, R. W. Lau, and C Gabriel, "The dielectric properties of bilogical tissues: III. Parametric models for the dielectric spectrum of tissues," *Physics in medicine and biology*, vol. 41, no. 11, p. 2271, 1996.

[22] R Pethig, "Dielectric properties of body tissues," *Clinicalphysics ad sphysiological measurement: an official journal of the Hospital Physicists' Association*, vol. 8, no. 4A, pp. 5–12, 1987.

[23] S. Y. Semenov, R. H. Svensson, and G. P. Tatsis, "Microwave spectroscopy of myocardial ischemia and infarction. 1. experimental study," *Annals of Biomedical Engineering*, vol. 28, no. 1, pp. 48–54, 2000.

[24] S. Semenov, M. Taran, G.P.Tatsis, A. Starostin, and R. Svensson, "Dielectric properties of canine acute and chronic myocardial infarction at a cell relaxation spectrum. 1. experiments," in *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering'*, vol. 1, 1997, pp. 202–205.

[25] T. Dunås and M. Kildal Schilliger, "Detection of traumatic epidural and subdural haematomas in brain phantoms using microwave technology," Master's thesis, Chalmers University of Technology, 2013.

[26] R Schmitt, *Electromagnetics explained: a handbook for wireless/RF, EMC, and high-speed electronics*. Boston: Newnes, 2002.

[27] D Stroud, "The effective medium approximations: Some recent developments," *Superlattices and Microstructures*, vol. 23, no. 3, pp. 567–573, 1998.

[28] R. Landauer, "Electrical conductivity in inhomogenous media," in *AIP Conference Proceedings*, 1978.

[29] A. Fhager, "Microwave tomography," PhD thesis, Chalmers University of Technology, 2006.

[30] K. Yee, "Numerical solution of initial boundary value problem involving Maxwell's equation in isotropic media," *IEEE Transactions on Antennas and Propagation*, vol. 14, no. 3, pp. 302–307, 1966.

[31] S. D. Gedney, *Introduction to the finite-difference time-domain (FDTD) method for electromagnetics*. Morgan and Claypool, 2011.

[32] J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 114, pp. 195–200, 1994.

[33] J. P. Berenger, *Perfectly Matched Layer (PML) for Computational Electromagnetics*. Morgan and Claypool, 2007.

[34] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley, 2000.

[35] Y. Yu, *Classification of High Dimensional Signals with Small Training Sample Size*. Chalmers University of Technology, 2013.

[36] L. I. Smith, "A tutorial on principal components analysis," [Online]. Available: `http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf`.

[37] J. R. Sarunas and A. K. Jain, "Small sample size effects in statistical pattern recognition," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 252–264, 1991.

[38] J. S. M. P. Hall and A. Neeman, "Geometric representation of high dimension, low sample size data," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 3, pp. 427–444, 2005.

[39] D. C. Hoyle, "Automatic PCA dimension selection for high dimensional data and small sampe size," *Journal of Machine Learning Research*, vol. 9, pp. 2733–2759, 2008.

[40] M. Kozdron. (1998). An introduction to random walks from Pólya to self-avoidance.

[41] A. R. Tilley, *The Measure of Man and Woman: Human Factors in design, Revised Edition*. Wiley, 2002.

[42] J. H. Conway, *Sphere packings, lattices and groups*. New York: Springer, 1999.

[43] A. Fhager, S. K. Padhi, and J. Howard, "3D image reconstruction in microwave tomography using efficient FDTD model," *IEEE Antennas and Wireless Propagation Letters*, vol. 1, pp. 1353 –1356, 2009.