

**DOTTORATO DI RICERCA IN
FISICA**

Ciclo XXVIII

Settore Concorsuale di afferenza: 02/D1
Settore Scientifico disciplinare: FIS/07

TITOLO TESI

**MICROWAVE BREAST CANCER IMAGING:
SIMULATION, EXPERIMENTAL DATA,
RECONSTRUCTION AND CLASSIFICATION**

Presentata da: Massimiliano Grandi

**Coordinatore Dottorato:
Chiar.mo Prof.
Gastone Castellani**

**Relatore:
Chiar.mo Prof.
Renato Campanini**

**Correlatori:
Chiar.mo Prof. Nico Lanconelli
Dott. Simone Masetti, Ph.D**

Esame finale anno 2016

*A Deborah, Francesco e Gaia
E a mia madre*

”Femmine un giorno e poi madri per sempre
nella stagione che stagioni non sente”
(F. De André)

Contents

1	Basic foreword on breast: anatomy, cancer pathology and screening	2
1.1	Breast Cancer	2
1.2	Breast Anatomy	3
1.3	Pathology	3
1.4	Breast density assessment	5
1.5	Screening: motivation and modalities	6
1.5.1	Mammography	7
1.5.2	MRI	9
1.5.3	Ultrasound	10
1.5.4	Other combined methods	11
2	Electromagnetism review and antenna theory	13
2.1	Maxwell's equations	13
2.2	Time Harmonic Field	15
2.3	Antenna characteristics	18
2.3.1	S-parameters definition	20
3	Microwave imaging	22
3.1	Theoretical basis	22
3.2	Dielectrical properties of breast tissues: an open issue	23
3.3	MWI reconstruction methods	25
3.3.1	Tomographic method	26
3.3.2	Linear methods	28
3.4	The TR-MUSIC algorithm	30
3.4.1	MUSIC	32
3.4.2	MUSIC for the monostatic configuration	33
3.5	Multifrequency	34
3.6	Existing MWI systems	35
4	Synthetic data simulation	38
4.1	Numerical solutions	38
4.1.1	FDM	38
4.1.2	MoM	39
4.1.3	FEM	39
4.2	COMSOL modeling	40

4.2.1	Space dimensionality	40
4.2.2	Interface and study	41
4.2.3	The geometry	41
4.2.4	Assigning material to the objects	42
4.2.5	Mesh	43
4.2.6	Lumped port	44
4.2.7	Antenna modeling	45
4.3	MATLAB automation	47
4.4	Homogeneous breast model	48
4.5	A breast of intermediate complexity	48
4.6	A real breast modeling	49
4.6.1	Dataset definition	50
4.6.2	Adding the real phantom to COMSOL	53
4.6.3	Conformal cup	57
4.7	Performance evaluation and scalability	58
4.7.1	Running time	61
5	Reconstruction of simulated data	62
5.1	Cluttering removal and filtering	62
5.1.1	Monostatic filtering	62
5.1.2	Multistatic filtering	63
5.2	Evaluation of the reconstruction and metrics	65
5.2.1	Spatial metrics	65
5.2.2	Contrast metric	66
5.3	Homogeneous breast model in monostatic configuration	68
5.3.1	Working Frequency	68
5.3.2	Multifrequency	69
5.3.3	Tumor position	70
5.3.4	Tumor size	72
5.3.5	Tumor-background contrast ratio (TBR)	73
5.3.6	Breast size	75
5.4	Monostatic and multistatic configurations	76
5.4.1	Architectures comparison	76
5.4.2	Multifrequency	78
5.4.3	Noise robustness	79
5.4.4	Number of field points	79
5.5	Inhomogeneous breast	81
5.5.1	Imaging algorithm for extended target	85
5.5.2	Model description	85
5.5.3	Role of the frequency	86
5.5.4	The eigenvectors role	86
5.5.5	The role of field points and the monostatic approach	88
5.5.6	Scatterer inside the fibroglandular structure	88
5.5.7	The fibroglandular relative permittivity	90
5.5.8	The maximum eigenvalue	91

6	Experimental data	94
6.1	Experimental scenario	94
6.2	Instruments	96
6.2.1	Dipole	97
6.2.2	VNA	97
6.2.3	Personal Computer	98
6.3	Experimental setup in monostatic configuration	98
6.3.1	Air-Water configuration	99
6.3.2	Measures with oil as background medium	100
6.3.3	Measures in oil with an ethyl alcohol scatterer	102
6.4	Achille and Briseide: the improved phantoms	103
	Appendices	106
A	Classification	107
A.1	CNN	107
A.2	Digital Database for Screening Mammography	109
A.2.1	Dataset generation	110
A.3	Binary classification	110
A.3.1	Receiver Operating Characteristic (ROC)	111
A.4	Model selection	111
A.4.1	Training the model	112
A.5	Classification results	112

Introduction and motivation

One of the most promising and emerging medical imaging modalities is Microwave Imaging (MWI). MWI involves the illumination of the portion of the body under investigation and resultant backscattered signals is exploited to generate images of the internal structures of the body. The main characteristics that suggest its deep study are low-power, low-cost, tissue-penetrating, focus-able, non-ionizing imaging modality with also therapeutic capabilities. Most of this MWI research to date has been accomplished in simulation and laboratory studies, with only limited translational research into the clinical environment.

In this thesis we aim to develop an experimental phantom used to collect data and compare them with the results obtained in the simulations.

The work is organized as follow: a brief introduction on breast cancer (chapter 1) and on key-concept on electromagnetism and antenna theory (chapter 2) is given; chapter 3 is devoted to an introduction on MWI theoretical basis and a description on the state-of-the-art regarding algorithms and clinical translation. In chapter 4 the simulation process is detailed and in chapter 5 the reconstruction results are discussed. Chapter 6 shows an experimental setup suitable for collecting data and the resulting reconstruction images.

In appendix is briefly detailed a classification method based on a deep learning technique using a mammographic dataset.

Chapter 1

Basic foreword on breast: anatomy, cancer pathology and screening

In this chapter an introduction to breast cancer is provided, together with a basic foreword on methodologies to diagnose the disease by means of the commonly accepted screening procedures.

1.1 Breast Cancer

The human body is made up of trillions of cells. These cells are living, that is they grow, split into new cells and die in an ordinary cycle. Though there are many different types, cancer begins when cells start to generate new daughter cells out of control in a localized part of the body and this is due to a damaged DNA. When a cell is a normal cell and DNA gets damaged, the cell itself either repairs the damage or dies. In cancer cells, the damaged DNA is not repaired, but the cell does not die like it should. Instead, this cell goes on making new cells that the body does not need. Unfortunately these new cells all have the same damaged DNA as the parent cell does. Breast cancer is a malignant tumor that starts in the cells of the breast. A malignant tumor is a cluster of cancer cells that can grow into surrounding tissues or spread to other areas of the body. This spreading is also called metastasis disease. Independently where a cancer has spread, it is always named by the place where it started. For example, breast cancer that has spread to the lung is still called breast cancer, not lung cancer [57]. This disease occurs almost entirely in women, but men can get it, too.

In figure 1.1 some incidence¹ statistics is given for breast cancer in the United States: more than 230.000 new cases is the preview for the year 2015 that corresponds to about the 14% of all new cancer cases. The estimated deaths are more than 40.000, the 6.8% of the total occurring

¹Incidence is the number of new cases arising in the given period

deaths caused by cancer. The percentage survival rate at 5 years, between year 2005 and 2011, had been almost 90%.

Breast cancer is the second leading cause of cancer death in women after lung cancer and the chance that breast cancer will be responsible for a woman’s death around the 3%.

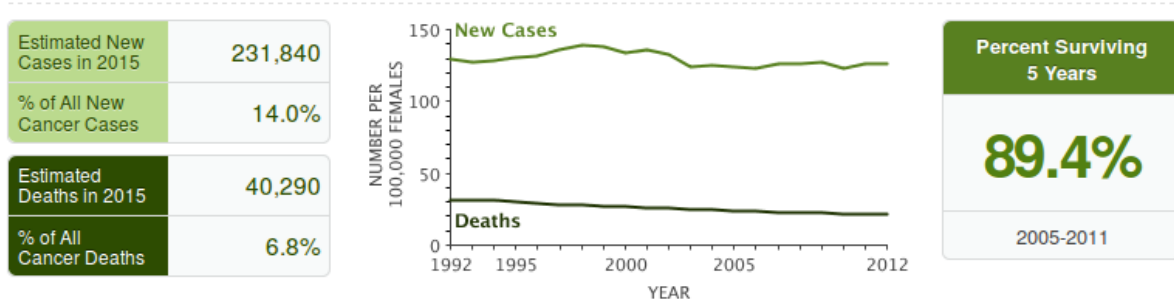


Figure 1.1: USA breast cancer statistics for year 2015. [1]

1.2 Breast Anatomy

To fully understand the difference with cancerous tissues it may help to have some basic knowledge about the normal structure of the mammal. Woman’s breast functionality is to produce milk to feed newborn babies, inside it there are, in general, 15 to 20 sections, called lobes where milk is produced. Each lobe is made of many smaller sections called lobules. Milk from lobe is then carried to the nipple through little vessels called ducts. Skin surrounding the internal is about 1 mm to 3 mm. Fibrous tissue and fat fill the spaces between the lobules and ducts. Fat can be found in three different regions: subcutaneous, just under the skin, retro-mammary in the back of the breast, and intra-glandular between the glandular structures. Minor presence have nerves, vascular and lymphatic tissue as well as a small number of lymph nodes inside the mammal.

Figure 1.2 shows a cross section in which all breast tissues are detailed.

1.3 Pathology

The most of initial development of the cancer occurs inside the glandular tissue, this can be either lobes or ducts. Regarding the starting location, we have an invasive lobular carcinoma (ILC) or invasive ductal carcinoma (IDC). About 10% of all invasive breast cancers are ILC while about 80% are IDC, as reported incidence based on the documentation of about 190,000 women with breast cancer in the Surveillance, Epidemiology, and End Results (SEER) program (www.seer.cancer.gov). The remaining 10% is cancer that originates in the pectoralis muscle, fat, or connective tissue in the breast and is referred to as sarcoma.

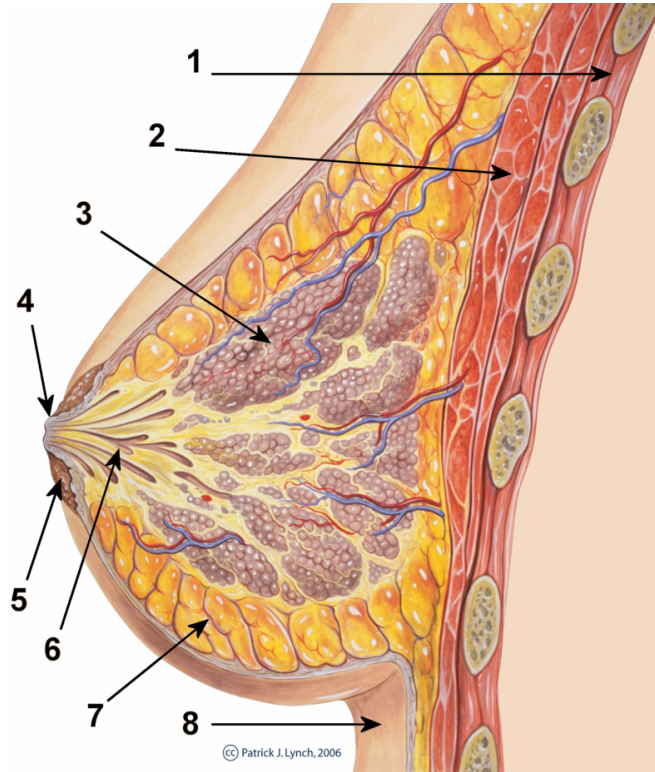


Figure 1.2: Cross sectional image of the normal breast [40]: 1. Chest wall 2. Pectoralis muscles 3. Lobules 4. Nipple 5. Areola 6. Milk duct 7. Fatty tissue 8. Skin

More precise taxonomy can be defined also if cancer remains confined, *in situ*, like ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS) or has attacked the surrounding tissue or even has spread away using lymph nodes and blood vessels to move away. In figure 1.3 are shown the evolution stages of a normal ductal cell transforming in a invasive ductal cancer cell.

Risk factor

There are many risk² factors to develop breast cancer, some are linked to items that cannot be prevented as age (higher probabilities are in women older than 55 years) or gender (male has a probability 100 times smaller to get the disease even if he could) or even race and ethnicity (African woman above 45 have a major risk). Also breast density affects the risk, this condition will be more detailed in the next section because it also affects the accuracy of screening examinations. Other risks are linked to the life style: drinking alcohol, getting hormones or not having a physical activity are considered risk factors. Unfortunately these are only risks, no one can tell if the disease will be got or when it will happen.

²A risk factor is anything that affects your chance of getting a disease, such as cancer.

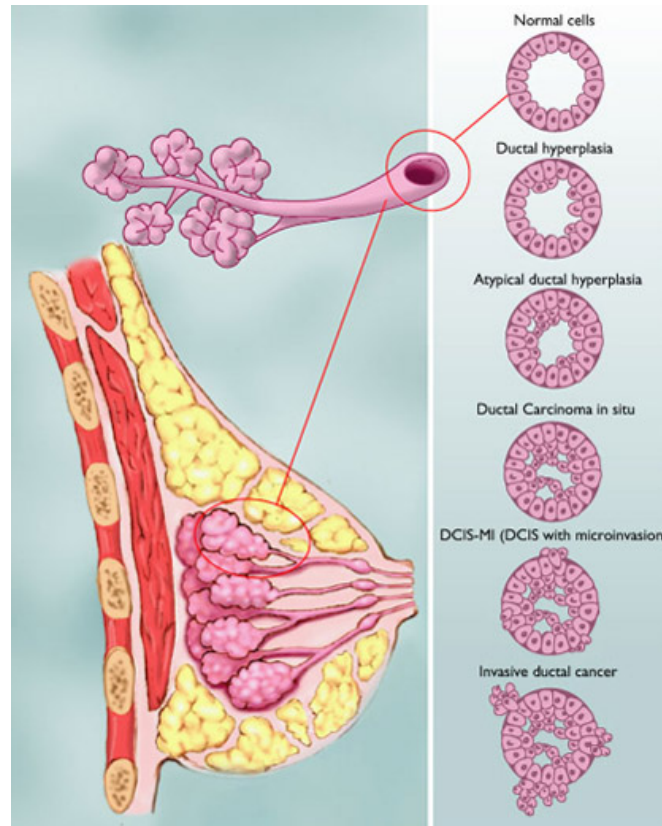


Figure 1.3: DCIS: Evolutionary representation of cell mutations [2]

1.4 Breast density assessment

The breast, as told before, is mostly made of fatty and fibro-glandular tissue with the relative proportions of the two varying in a wide range. Attempts to develop a reproducible quantitative method of assessing breast density began in the early 1980s. Visual estimation of the percentage of the breast occupied by breast tissue has been used frequently, with the most common method having six categories, where fibro-glandular tissue occupies $> 75\%$, $50\text{--}75\%$, $25\text{--}50\%$, $10\text{--}25\%$, $< 10\%$, or none of the breast [10]. The American College of Radiology (ACR) developed the Breast Imaging-Reporting and Data System (BI-RADS), which is becoming the standard of mammography-reporting terminology and assessment and recommendation categories [47]. The BI-RADS classifies mammographic density into four categories: Breast Density by Mammography

- BI-RADS category 1: breast with the density of fat,
- BI-RADS category 2: fatty breast with scattered fibroglandular densities,
- BI-RADS category 3: heterogeneously dense breast,
- BI-RADS category 4: dense breast.

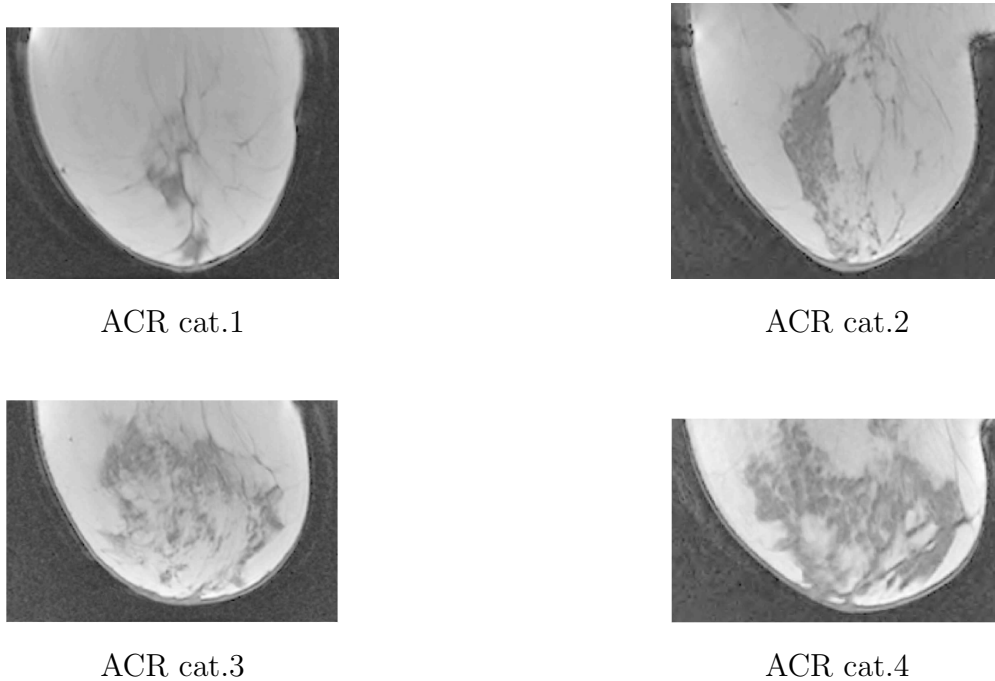


Figure 1.4: Representative mammograms shown from MRI sagittal slice. From (a) to (d) are displayed the increasing values of the four BI-RADS categories [66]

1.5 Screening: motivation and modalities

Breast cancer is sometimes found after symptoms appear, but many women with early breast cancer have no symptoms therefore getting the recommended screening tests before any symptom appears is extremely important. When diagnosing and monitoring cancer, the evaluation of radiological image data is an important source of objective information. Such evaluation includes the detection and classification of suspicious lesions as well as the classification of tumor tissue under treatment. Identifying and categorizing abnormality relies on a set of consistent criteria based on image features to which the human visual system will respond with particular sensitivity. Many imaging modalities help radiologists making their decisions on the basis of the appearance of a tumor mass while monitoring dynamic changes[57]. Some common methodologies are now detailed with a special interest for mammography and magnetic resonance imaging (MRI) because the former will be used in the chapter 7 for the classification stage and the latter will be used in chapter 6 to produce a realistic phantom for the simulated data. A summary of whole diagnosing techniques is even reported in figure 1.8.

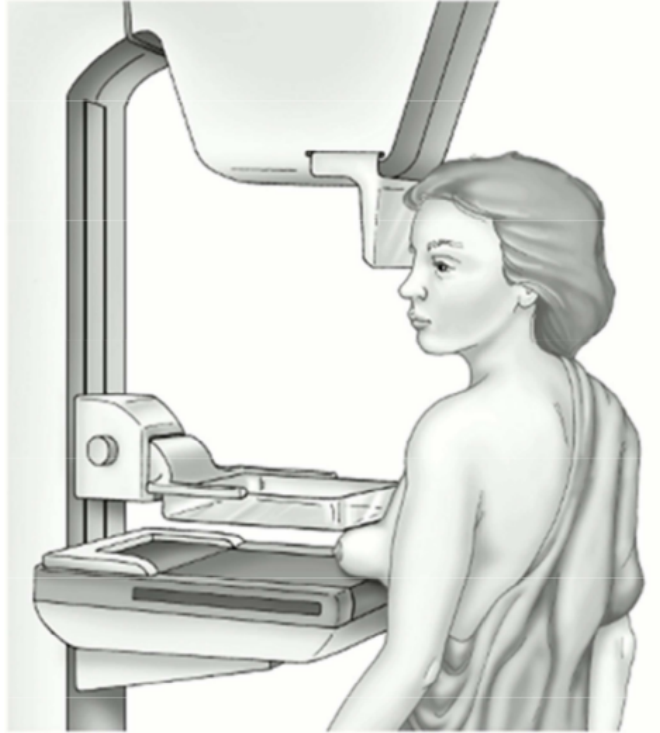


Figure 1.5: Mammography and breast compression [57]

1.5.1 Mammography

A mammogram is a 2D imaging method that makes use of X-rays to investigate breast disease. Early x-ray mammography (1950s) used a direct exposure film, this required a high dose of radiation having low contrast and poor diagnostic quality [11]. With continuous refinement and with the regulations dictated from ACR, mammography became the standard methodology for screening and finally the advent of full field digital mammography (FFDM) allowed a low-cost, low-dose, rapid acquisition and better quality imaging technology worldwide used up-to-day. In FFDM the x-rays do not go through tissue as easily as those used for routine bone x-rays (chest, arm, legs) and this helps improve the image quality. The goal of a screening mammogram is to find breast cancer when too small to be felt palpable by a human. Finding breast cancers early, that is before they have grown and spread, greatly improves a woman's chance for successful treatment and keeps her alive. As an accepted standard 2 pictures are taken for each breast: the mediolateral oblique (MLO) view and the craniocaudal (CC) view. In figure 1.6 is shown the way to obtain the two projections.

Mammography can also be used as a diagnostic tool to perform further evaluation of abnormalities identified by palpation or screening mammography. In this case additional projections and magnifications might be included.

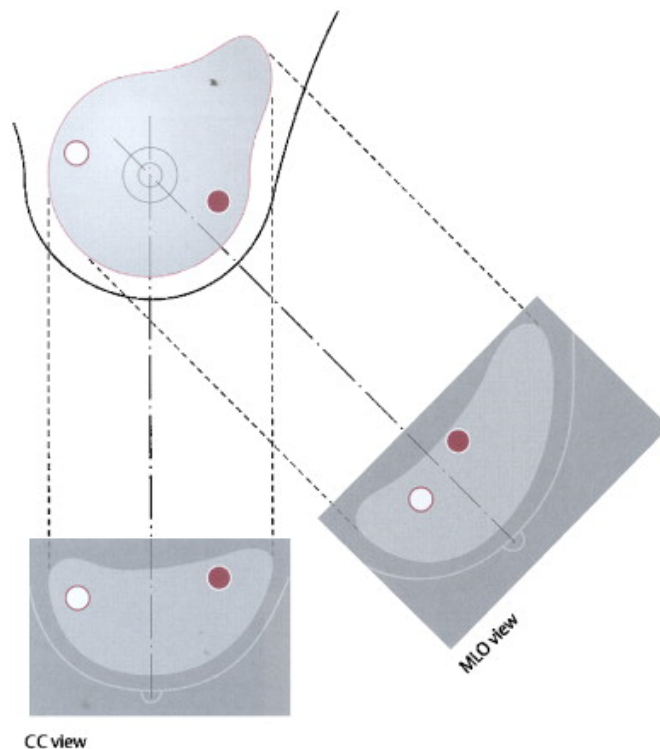


Figure 1.6: How are get CC and MLO views in mammography (<http://ars.els-cdn.com/content/image/>)

Drawbacks

Unfortunately one in four women will attend their first mammogram and never return again [9]. Clearly, the screening experience itself is discouraging many women from receiving the benefit of screening. One year after the single screening experience, the cancers found among these women will be just as large and just as lethal as the cancers found among women who never use screening. The simulation data suggest that this will double the risk of cancer death among these women. Many women find mammography unpleasant and stressful and this is mainly due to the breast compression. Moreover screening is not without its negative consequences and mammography suffers of common mistakes known as “false-positives” and ”false-negative” results.

False-negative results A false-negative mammogram appears normal even though breast disease is present. Overall, screening mammograms do not find about 1 in 5 breast cancers [citations needed]. False negatives occur more often among women with dense breasts. Breasts often become somewhat less dense with increasing of the age, so false negatives are more common among younger women. False-negative results can delay treatment and give a false sense of security for the woman.

False-positive results A false-positive mammogram looks abnormal but no disease is actually present. False-positive results are more common in women who are younger, have dense breasts, have had breast biopsies, have breast cancer in the family, or are taking estrogen. About half the women getting annual mammograms over a 10-year period will have a false-positive finding. The odds of a false-positive finding are highest for the first mammogram, and are lower on subsequent mammograms. Abnormal mammograms lead a follow-up treatment like biopsies and other medical interventions among women who subsequently are not found to have cancer. One widely quoted study, by Elmore [19], “estimate[d] that among women who do not have breast cancer, 18.6% will undergo a biopsy after 10 mammograms.” Women who have past mammograms available for comparison halves their odds of a false-positive finding.

1.5.2 MRI

Magnetic resonance imaging (MRI) is based on the physical principle of nuclear magnetic resonance (NMR). In 1971 Damadian showed that relaxation times in NMR experiments on rats differed between healthy and malignant tissue [17]. One year later the first magnetic resonance images were produced by Lauterbur. He showed that using magnetic gradients it was possible to locate the origin of a radio signal emitted by nuclear spins within an excited body [36]. From there onwards, the study of MRI has rapidly grown: image quality, tissue discrimination and imaging speed have largely improved meanwhile. The first in vivo MRI study of the breast has been reported by Ross in 1982 [51] showing the potential of breast MRI to detect lesions. However, image contrasts between normal tissue and malignant lesions were poor and only large lesions were detected: the sensitivity of breast MRI was comparable to that of mammography at a much higher costs. [41] The value of breast MRI changed radically when contrast agents based on gadolinium were for the first time injected intravenously to detect breast cancer [20]. These contrast agents are diamagnetic and shorten the T1 (spin-lattice) relaxation time of surrounding tissue [21]. This leads to an enhancement at T1-weighted images and signal changes that can also be observed over time.

Recommendations MRI is mainly used in women who have been diagnosed with breast cancer, to help measuring its the size and look for any other lesions in the breast. It also can be used to look at the opposite breast, to be sure that it does not contain any tumors. The American Cancer Society has advised women at high risk for breast cancer to have breast MRI screening along with a yearly screening mammography [26].

MRI is not recommended as a screening tool by itself because it can miss some cancers that a mammogram would detect (false-negative), moreover although MRI can find some cancers not seen on mammograms, it is also more likely to find something that turns out to be false-positive and

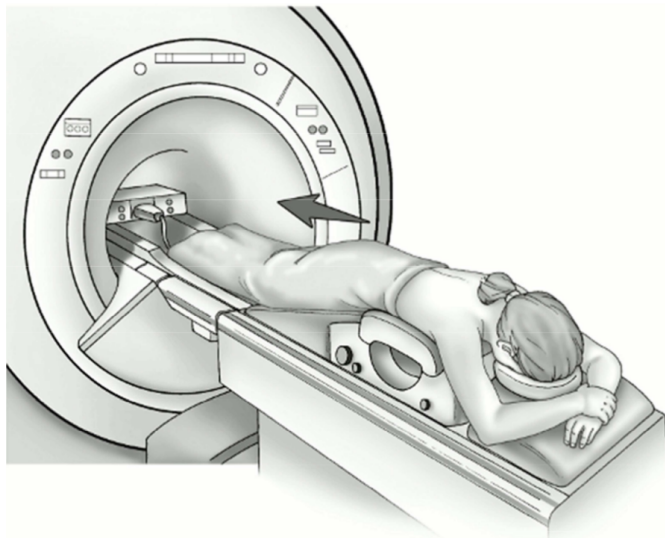


Figure 1.7: Breast MRI

that need a follow-up to check out the presence of cancer. This is why MRI is not recommended as a screening test for women at average risk of breast cancer, as it would result in unneeded biopsies and other tests in many of these women.[57]. Figure ?? depicts a typical magnetic resonance of the breast.

1.5.3 Ultrasound

Ultrasound is an imaging modality inherited from US Navy in 1950s formerly used to differentiate breast cysts from masses. It uses a small transducer probe coupled with an ultrasound gel spread on the skin. Sound waves beyond the human auditory threshold (i.e. > 20 KHz) are transmitted from probe to the body by means of the gel and picks up the echoes as they bounce back from body tissues. This echoes are converted into a visual image on a computer screen. This test is painless and does not expose to radiation. Breast ultrasound is most often used to evaluate breast problems that are found during a mammogram or physical exam, and to guide a needle biopsy of a suspicious mass. In someone with a breast tumor, it may also be used to look at lymph nodes under the arm. Although ultrasound is less sensitive than MRI (the ability to detect tumors is lower), it has become a valuable tool to use along with mammograms because it's widely available, non-invasive, and costs less than other options. As with any breast imaging, breast ultrasound depends on the level of skill and experience of the doctor interpreting the images and, if handheld transducer is used (recently some automated configurations have been made available), a good image optimization relies on the skill and experience of the person doing the scan [28].

1.5.4 Other combined methods

Microwave imaging and Dynamic Optical Breast Imaging (DOBI) are thought to support mammography in making decisions. Both are low cost, non-invasive, fast execution, without ionizing radiation and do not need the breast compression.

DOBI

DOBI diagnostic equipment is innovative, non-invasive, which does not require ionizing radiation emissions and that uses optical technology for the early detection of breast cancer in young women. It is a safe exam that uses light in the visible spectrum to illuminate the breast tissue and identify tumor lesions at an early stage.

It is a diagnostic method particularly innovative, able to detect changes of vascularization (angiogenesis) that surround the tumor lesion. Therefore, while mammography and breast ultrasound are diagnostic tests that study the changes in the shape and density of the tissues (morphological examinations).

Dobi, such as breast MRI, is a functional examination, that provides information about the changes that occur in the mammary gland for the presence of a specific lesion. It is a very rapid examination taking about 10 minutes to be completed with a good performance in differencing benign from malignant lesions.

Microwave imaging

MWI reached popularity amongst researcher when it had been demonstrated that there is an high contrast between healthy and malignant tissues on electrical permittivity and conductivity and this contrast can be higher with respect to x-ray imaging. Compared to the other methods here reported, MWI is not at the same stage even though many algorithms have been developed: unfortunately very low number of prototypes have been implemented and tested in a clinical environment.

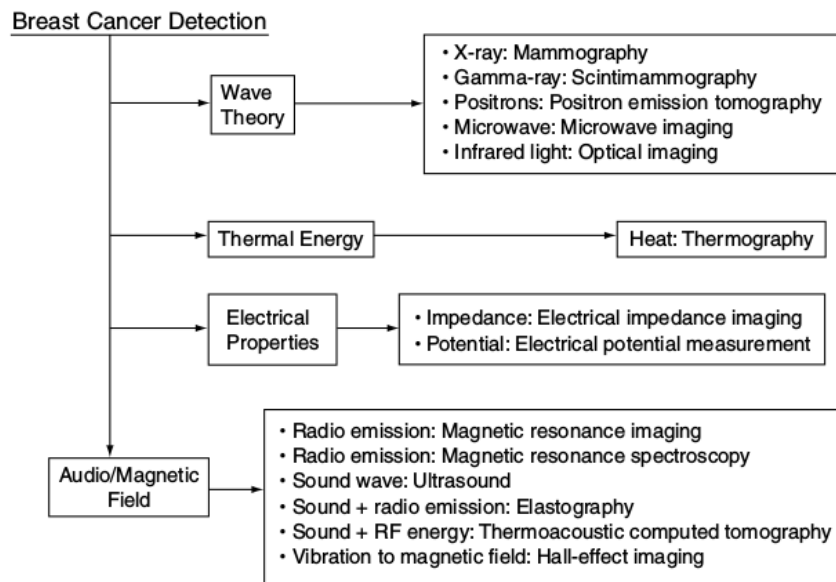


Figure 1.8: Summary of breast imaging techniques

Chapter 2

Electromagnetism review and antenna theory

2.1 Maxwell's equations

The electromagnetic field can be described by the well known Maxwell's equations which relate the fields with their sources, below reported in differential (local) form:

$$\nabla \times \vec{\mathcal{E}} = -\frac{\partial \vec{\mathcal{B}}}{\partial t}, \quad (2.1a)$$

$$\nabla \times \vec{\mathcal{H}} = \frac{\partial \vec{\mathcal{D}}}{\partial t} + \vec{\mathcal{J}}, \quad (2.1b)$$

$$\nabla \cdot \vec{\mathcal{D}} = \rho_e, \quad (2.1c)$$

$$\nabla \cdot \vec{\mathcal{B}} = 0 \quad (2.1d)$$

All quantities have a spatial and temporal dependence, hence $\vec{\mathcal{E}} = \vec{\mathcal{E}}(\vec{r}, t)$ with \vec{r} in meters e t in seconds. The $\vec{\mathcal{E}}$, $\vec{\mathcal{D}}$, $\vec{\mathcal{H}}$ and $\vec{\mathcal{B}}$, are the vector fields that describe the whole electromagnetic field and are named *electric field* (V/m volts per meter), *electric flux density* (C/m² Coulombs per square meter), *magnetic field* (A/m Ampere per meter), *magnetic flux density* (Wb/m² Weber per square meter). The sources are ρ *volumetric electric charge density* expressed in C/m³ (Coulomb per cubic meter) and $\vec{\mathcal{J}}$ the *current charge density* expressed in A/m² (Ampere per square meter).

It is known that the presence of a medium affects the electromagnetic field through three different ways, *electric polarization*, *magnetic polarization* and *electric conduction* but in the Maxwell's equation, that are valid for all electromagnetic phenomema, no information is given about the medium. It has been demonstrated that not all the twelve scalar quantities, corresponding to each

component of each field, can be solved using (2.1a-d) but only six. This suggests that other relations should be present for the vector fields. In general the electric flux density and the magnetic flux density are a function of the electric and magnetic fields:

$$\vec{\mathcal{D}} = F(\vec{\mathcal{E}}, \vec{\mathcal{H}}) \quad (2.2a)$$

$$\vec{\mathcal{B}} = G(\vec{\mathcal{E}}, \vec{\mathcal{H}}) \quad (2.2b)$$

These relations are called *constitutive equations* and give the remaining six equations needed to completely solve 2.1d. For most materials the operators F and G are linear, moreover the electric flux density depends only from electric field and magnetic flux density depends only from the magnetic field:

$$\vec{\mathcal{D}} = F(\vec{\mathcal{E}}) \quad (2.3a)$$

$$\vec{\mathcal{B}} = G(\vec{\mathcal{H}}) \quad (2.3b)$$

The electromagnetic properties of a medium are reflected in the following three equations:

$$\vec{\mathcal{D}} = \varepsilon \vec{\mathcal{E}} = \varepsilon_0 \varepsilon_r \vec{\mathcal{E}} \quad (2.4a)$$

$$\vec{\mathcal{B}} = \mu \vec{\mathcal{H}} = \mu_0 \mu_r \vec{\mathcal{H}} \quad (2.4b)$$

$$\vec{\mathcal{J}} = \sigma \vec{\mathcal{E}} \quad (2.4c)$$

The ε values is the electrical permittivity (F/m Farad per meter) and μ (H/m Henry per meter) is the magnetic permeability. These equations let us to make a classification of a medium based on the three quantities that appear in them. Because our interest is not on magnetic materials, the classification based on μ will not be exhaustive.

Spatial classification: if any of the quantity ε , σ , μ depends as a function of the spatial coordinates (i.e. $\varepsilon = \varepsilon(x, y, z)$), the medium is called *inhomogeneous*.

Classification based on $\vec{\mathcal{D}}$ and $\vec{\mathcal{B}}$ directions: if $\vec{\mathcal{D}}$ is parallel to the direction of $\vec{\mathcal{E}}$ and $\vec{\mathcal{B}}$ is parallel to the direction of $\vec{\mathcal{H}}$ the medium is called *isotropic* otherwise another form of the (2.2a-b) must be given in tensorial notation:

$$\begin{pmatrix} \vec{\mathcal{D}}_x \\ \vec{\mathcal{D}}_y \\ \vec{\mathcal{D}}_z \end{pmatrix} = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{pmatrix} \begin{pmatrix} \vec{\mathcal{E}}_x \\ \vec{\mathcal{E}}_y \\ \vec{\mathcal{E}}_z \end{pmatrix} \quad (2.5)$$

This can be expressed in a more compact way as:

$$\vec{\mathcal{D}} = \bar{\epsilon} \vec{\mathcal{E}} \quad (2.6a)$$

$$\vec{\mathcal{B}} = \bar{\mu} \vec{\mathcal{H}} \quad (2.6b)$$

Classification based on frequency dependence: if the electrical parameters depend on frequency, $\epsilon = \epsilon(\nu)$, $\sigma = \sigma(\nu)$ the medium is called *dispersive*.

Classification based on the value of the conductivity: if $\sigma > 0$ the medium is said *lossy*, if $\sigma \rightarrow \infty$ then the medium is called *Perfect Electric Conductor* (PEC).

2.2 Time Harmonic Field

If the electromagnetic field is in co-sinusoidal form containing a single frequency (also said to be time harmonic) then we can express it in the following way:

$$\vec{\mathcal{E}}(\vec{r}, t) = \mathbf{E}_0(\mathbf{r}) \cos(\omega t + \phi_0(\mathbf{r})) \quad (2.7)$$

and applying Euler's formula:

$$\vec{\mathcal{E}}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r}) \operatorname{Re}(e^{j(\omega t + \phi_0(\mathbf{r}))}) = \operatorname{Re}(\mathbf{E}(\mathbf{r}) e^{j\omega t}) \quad (2.8)$$

the quantity containing the amplitude and the initial phase depending only from the spatial coordinates

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) e^{j\phi_0} \quad (2.9)$$

is called *phasor*. This way of solving Maxwell's equations for $\mathbf{E}(\mathbf{r})$ is referred as working in the frequency domain, in fact applying the Fourier transform of (2.5) where ω has been replaced with the selected operating frequency ω_0 , we obtain:

$$\begin{aligned} \vec{\mathcal{E}}(\vec{r}, \omega) &= \operatorname{Re} \left(\int_{-\infty}^{+\infty} \mathbf{E}(\mathbf{r}) e^{j\omega_0 t} e^{-j\omega t} dt \right) = \operatorname{Re} \left(\mathbf{E}(\mathbf{r}) \int_{-\infty}^{+\infty} e^{j\omega_0 t} e^{-j\omega t} dt \right) \\ &= \operatorname{Re}(\mathbf{E}(\mathbf{r}) \delta(\omega - \omega_0)) = \operatorname{Re}(\mathbf{E}(\mathbf{r})) \end{aligned} \quad (2.10)$$

In the important case of the frequency (or spectral) domain, the Maxwell's equations can be

written as follows:

$$\nabla \times \mathbf{E}(\mathbf{r}) = -j\omega\mathbf{B}(\mathbf{r}), \quad (2.11a)$$

$$\nabla \times \mathbf{H}(\mathbf{r}) = j\omega\mathbf{D}(\mathbf{r}) + \mathbf{J}(\mathbf{r}), \quad (2.11b)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}) = \rho_e(\mathbf{r}), \quad (2.11c)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}) = 0 \quad (2.11d)$$

When we deal with dispersive media, the dielectric properties ε and σ depend on the frequency. In the frequency domain this leads to a complex-value of permittivity as we will demonstrate below. This dependence is modeled by several formulas, such as the Cole-Cole or Debye one. In this work the breast tissues will be modeled using the following Debye's formula [7]:

$$\varepsilon(\omega) = \varepsilon_\infty + \frac{\varepsilon_\infty - \varepsilon_s}{1 + j\omega\tau} = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + j\omega\tau} \quad (2.12)$$

ε_s represents the permittivity in presence of a static field ($\omega \rightarrow 0$) and ε_∞ corresponds to $\omega \rightarrow \infty$, τ is the *relaxing time*, a parameter characteristic of the medium of the order of the picoseconds. Moreover in a conducting medium, where the Ohm's law is valid, some induced currents may arise:

$$\mathbf{J}_i(\mathbf{r}) = \sigma\mathbf{E}(\mathbf{r}) \quad (2.13)$$

in this case the equation 2.11b for a linear, isotropic medium where 2.4 hold, becomes:

$$\nabla \times \mathbf{H}(\mathbf{r}) = j\omega\varepsilon\mathbf{E}(\mathbf{r}) + \sigma\mathbf{E}(\mathbf{r}) + \mathbf{J}_0(\mathbf{r}) \quad (2.14)$$

where the total current density $\mathbf{J}(\mathbf{r})$ has been replaced by the sum of the density impressed currents $\mathbf{J}_0(\mathbf{r})$ with the conduction density current given by the above Ohm's law. Now using some algebraic ability diving by μ_r and taking the curl on both sides in equation 2.11a, substituting the expression of $\nabla \times \mathbf{H}(\mathbf{r})$ given from equation 2.14, and recalling that $1/c^2 = \mu_0\varepsilon_0$

$$\begin{aligned} \nabla \times \frac{1}{\mu_r} \nabla \times \mathbf{E}(\mathbf{r}) &= -j\omega\mu_0 \left(j\omega\varepsilon\mathbf{E}(\mathbf{r}) + \sigma\mathbf{E}(\mathbf{r}) + \mathbf{J}_0(\mathbf{r}) \right) \\ &= \omega^2\mu_0\varepsilon_0 \left(\varepsilon_r + \frac{\sigma}{j\omega\varepsilon_0} \right) \mathbf{E}(\mathbf{r}) - j\omega\mu_0\mathbf{J}_0(\mathbf{r}) \\ &= \frac{\omega^2}{c^2} \left(\varepsilon_r + \frac{\sigma}{j\omega\varepsilon_0} \right) \mathbf{E}(\mathbf{r}) - j\omega\mu_0\mathbf{J}_0(\mathbf{r}) \end{aligned} \quad (2.15)$$

If we highlight the quantity present at the second row of 2.15

$$\varepsilon_c = \varepsilon_0 \left(\varepsilon_r + \frac{\sigma}{j\omega\varepsilon_0} \right) \quad (2.16)$$

calling it complex-value dielectric permittivity, this leads to rewrite 2.15 in the well known vector wave equation form:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) = \omega^2 \mu \varepsilon_c \mathbf{E}(\mathbf{r}) - j\omega \mu \mathbf{J}_i(\mathbf{r}) \quad (2.17)$$

Same considerations can be done to retrieve the vector wave equation for \mathbf{H} , that reads:

$$\nabla \times \nabla \times \mathbf{H}(\mathbf{r}) = \omega^2 \mu \varepsilon_c \mathbf{H}(\mathbf{r}) + \nabla \times \mathbf{J}_i(\mathbf{r}) \quad (2.18)$$

To uniquely solve these equations in an unbounded medium the Silver-Muller radiation conditions must be provided [49], in this case the solution of the vector wave equation can be expressed in the following integral form:

$$\mathbf{E}(\mathbf{r}) = j\omega \mu \int_V \mathbf{J}_0(\mathbf{r}') \bar{\bar{G}}(\mathbf{r}', \mathbf{r}) dV \quad (2.19a)$$

$$\mathbf{H}(\mathbf{r}) = - \int_V \nabla \times \mathbf{J}_0(\mathbf{r}') \bar{\bar{G}}(\mathbf{r}', \mathbf{r}) dV \quad (2.19b)$$

it is noteworthy that the volume integral is extended to regions where the impressed currents are present. The $\bar{\bar{G}}$ is the dyadic Green tensor as defined in [60]:

$$\bar{\bar{G}}(\mathbf{r}', \mathbf{r}) = -\frac{1}{4\pi} \left[\bar{\bar{I}} + \frac{1}{k^2} \nabla \nabla \right] \frac{e^{-jk|\mathbf{r}'-\mathbf{r}|}}{|\mathbf{r}'-\mathbf{r}|} \quad (2.20)$$

where $k = \omega \sqrt{\mu \varepsilon}$ is the wavenumber expressed in [m^{-1}] and is the solution of the vector wave equation where the excitation is provided from an elementary current source:

$$\nabla \times \nabla \times \bar{\bar{G}}(\mathbf{r}', \mathbf{r}) - k^2 \bar{\bar{G}}(\mathbf{r}', \mathbf{r}) = I \delta(\mathbf{r}' - \mathbf{r}) \quad (2.21)$$

In its general form the dyadic Green function has nine independent components (nonionic) because an impressed current coming along an axis gives rise to three distinct components of the field along the 3 coordinate axis.

The expression for the scalar Green function in 2-dimensions is:

$$G(\mathbf{r}', \mathbf{r}) = -\frac{i}{4} H_0^{(2)}(k|\mathbf{r}' - \mathbf{r}|) \quad (2.22)$$

where $H_0^{(2)}$ is the Hankel function of 0-order and type 2, while its expression in 3-dimensions

is:

$$G(\mathbf{r}', \mathbf{r}) = -\frac{1}{4\pi} \frac{e^{-jk|\mathbf{r}'-\mathbf{r}|}}{|\mathbf{r}'-\mathbf{r}|} \quad (2.23)$$

2.3 Antenna characteristics

Basically antennas are used to send and receive signals in form of electromagnetic waves. There are several types of antennas based on their particular use and there are common quantities that completely define their working behavior reported by standard IEEE publications[]. The most important parameter is the power density, the active power radiated per unitary surface observed in the far-field region, that is on a sphere whose radius is at least:

$$r_{\text{far-field}} > \frac{2D^2}{\lambda} \quad (2.24)$$

where D is the highest geometrical dimension of the antenna and λ the source wavelength. The total power radiated in a lossless medium is:

$$P_T = \int_{S_r} \frac{\|E\|^2}{2\eta} dS_r = \int_{4\pi} \frac{\|E\|^2}{2\eta} r^2 d\Omega \quad (2.25)$$

where η is the vacuum impedance. The quantity:

$$I_R(\theta, \phi) \equiv \frac{\|E\|^2}{2\eta} r^2 \quad (2.26)$$

is called radiation intensity and can also be seen as:

$$I_R(\theta, \phi) = \frac{dP_T}{d\Omega} \quad (2.27)$$

The radiation intensity is a continuous function defined on a compact set so it admits maximum and minimum, in particular there exists at least one maximum direction:

$$I_R(\theta_M, \phi_M) \quad (2.28)$$

We define the radiation function in this way:

$$f(\theta, \phi) = \sqrt{\frac{I_R(\theta, \phi)}{I_R(\theta_M, \phi_M)}} = \frac{\|\mathbf{E}(r, \theta, \phi)\|}{\|\mathbf{E}(r, \theta_M, \phi_M)\|} \leq 1 \quad (2.29)$$

This function is extremely useful to characterize in a compact way how an antenna radiates, in particular if:

$$f(\theta, \phi) = 0 \quad (2.30)$$

we say it is a zero-direction for the antenna, that is the antenna does not radiate in this direction. The radiation surface is a 3D plot that shows how the field strength is distributed along the space. Making a projection onto a plane containing the origin we obtain a 2D graph called radiation diagram and this is plotted onto a polar diagram that indicates the magnitude of the response in any direction. The center of the diagram is a point referred to as the origin. This is surrounded by a curve whose radius at any given point is proportional to the magnitude of the property measured in the direction of that point. In figure 2.1 the dipole radiation diagram, for the E-plane determined by the maximum irradiation direction and the polarization direction is displayed .

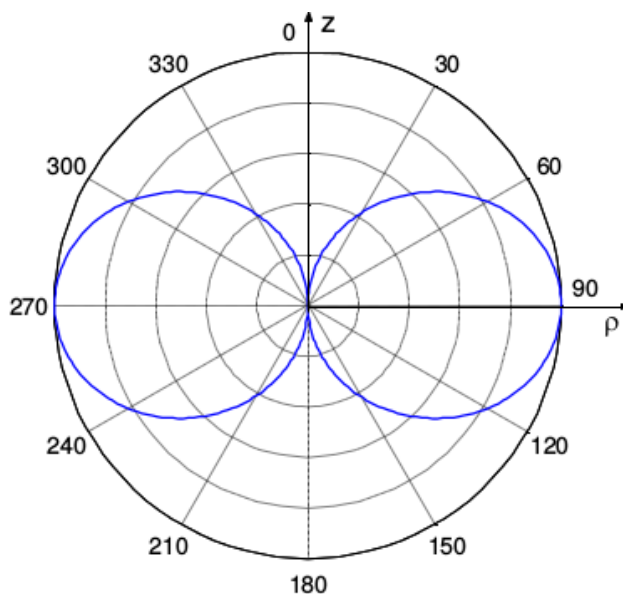


Figure 2.1: Half-wavelength dipole polar plot for the E-plane.

Directivity, efficiency and gain

While the transmitting power depends on the generator outside the antenna, the directivity function is an antenna characteristic quantity and is defined as:

$$d(\theta, \phi) = \frac{4\pi I_R(\theta, \phi)}{P_T} \quad (2.31)$$

The directivity, D , is simply the value of the directivity function computed along the maximum and measures the ability to radiate into a particular direction:

$$D = \frac{4\pi I_R(\theta_M, \phi_M)}{P_T} = \frac{4\pi I_R(\theta_M, \phi_M)}{\int_{4\pi} I_R(\theta, \phi) d\Omega} \geq 1 \quad (2.32)$$

For an omni-directional purpose it is desirable to have $D \simeq 1$, while for a point to point link is required $D \gg 1$.

The materials an antenna is made up, both conductors and dielectrics, are not ideal but rather lossy and the power radiated is always less than the feeding power coming from the source:

$$P_{\text{source}} = P_{\text{Joule_antenna}} + P_{\text{radiated}} \quad (2.33)$$

To characterize the losses, under the assumption that they are mainly due to the antenna, the transmission efficiency is introduced and defined as:

$$\eta_t = \frac{P_{\text{radiated}}}{P_{\text{source}}} = \frac{P_{\text{radiated}}}{P_{\text{Joule_antenna}} + P_{\text{radiated}}} \leq 1 \quad (2.34)$$

Efficiency does not depend on the directivity, that is an antenna can be very directive but with a very low efficiency so useless in practice.

The gain function is defined merging the above definitions:

$$g(\theta, \phi) = \frac{4\pi I_R(\theta, \phi)}{P_{\text{source}}} = \eta_t \cdot d(\theta, \phi) \quad (2.35)$$

the Gain is the maximum of the gain function:

$$G = \frac{4\pi I_R(\theta_M, \phi_M)}{P_{\text{source}}} = \eta_t \cdot D \quad (2.36)$$

2.3.1 S-parameters definition

In the antenna domain it is used to define the electromagnetic signal in term of scattering parameters (S-parameters) that refers to traveling waves instead of voltage and currents read in a port, so they substitute the standard H, Y and Z parameters.

In a one-port network we define the ratio from the incident wave a and the reflected wave b as

$$S_{11} = \frac{b}{a} = \Gamma \quad (2.37)$$

This is a complex value because it contains both amplitude and phase and is called reflection coefficient or even *return loss* if expressed in dB:

$$\text{return loss} = 10 \log_{10} |S_{11}| \leq 0 \quad (2.38)$$

This coefficient is useful to characterize the bandwidth of a transmitting antenna: it is commonly accepted to refer as operational bandwidth the frequency range in which the return loss is lower than -3 dB or -10 dB based on the scope.

For a two-port network, saying a the input signals and b the output signals we write:

$$b_1 = S_{11}a_1 + S_{12}a_2 \quad (2.39a)$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \quad (2.39b)$$

in this case the S-parameters are defined as follows:

$$S_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0}, \quad S_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0}, \quad S_{21} = \left. \frac{b_2}{a_1} \right|_{a_2=0}, \quad S_{22} = \left. \frac{b_2}{a_2} \right|_{a_1=0} \quad (2.40)$$

We can also write 2.39 in a matrix form

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (2.41)$$

What is true for two-port networks is similarly true for n-port networks and the S-parameters take the form of S_{mn} . In this case the number of measurements required for characterizing these more complex networks goes up as the square of the number of ports.

Finally the network is said reciprocal if the S-matrix is symmetric, that is if it holds:

$$S_{ij} = S_{ji} \quad (2.42)$$

Chapter 3

Microwave imaging

In this chapter the theoretical basis of MWI are detailed and some results involving the electrical properties of the breast tissues are given. Some reconstruction techniques are presented along with the one used to obtain our results. Finally some existing prototypes are illustrated.

3.1 Theoretical basis

Theoretical and experimental basis of microwave imaging relies upon that an inhomogeneity in the constitutive electromagnetic parameters of a material, electrical permittivity and conductivity, will cause an incident electromagnetic field to be scattered. Figure 3.1 shows how a basic MWI system works: the microwave imaging system is embedded inside a background medium which is characterized by the permittivity ϵ_{bg} and conductivity σ_{bg} and the circular area in the center of the antenna group represents the imaging domain. This domain is irradiated by a transmitting antenna and the total field may be measured by one or more antennas positioned outside the domain. When an object with contrast in the constitutive parameters is positioned inside the imaging domain, a scattered field will arise [61]. The response collected from the others receiving antennas is the superposition of the incident field and the scattered field

$$\mathbf{E}_{tot} = \mathbf{E}_{inc} + \mathbf{E}_{scat} \quad (3.1)$$

The process followed to produce the images starting from the collected data are summarized below:

- preliminary assessment of the scattering scenario;
- noise and artifacts reduction;
- image reconstruction;

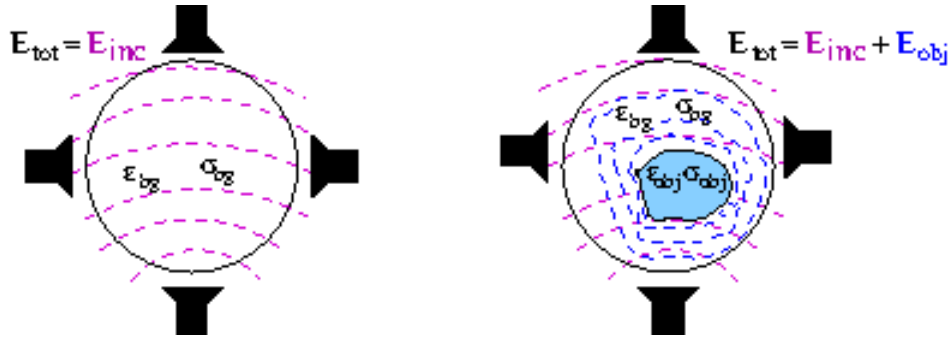


Figure 3.1: Basic schema on microwave imaging system works

3.2 Dielectrical properties of breast tissues: an open issue

The exact knowledge of the breast tissues dielectric properties is fundamental to produce a correct image but also permits the design of the devices involved in the measurement process along with the production of realistic phantoms, used in testing procedure to correctly mimic the microwaves propagation inside the tissues. Therefore along time several research groups efforts were devoted to study these properties.

The first systematic study can be ascribed to Chaudhary [13] around 1980. He measured ex-vivo the tumoral and normal breast dielectric properties at 25° in the range 3 MHz - 3 GHz obtaining that the contrast value between the two classes was almost 5:1 for the permittivity (see figure 3.2) and 4.7:1 for the conductivity.

In the 90th Campbell [12] made a fixed frequency study, 3.2 GHz, using a resonant cavity. Its results showed a sharp contrast between tumoral and fatty tissue, but almost null between benign and malignant tissues, concluding that the use of dielectric properties could not be conclusive to discriminate between them. Moreover healthy tissues had a wide range of variability, 9.8 to 46 for the relative permittivity and 7 - 49 S/m for the conductivity.

The following studies of Joines [32] and C. Gabriel [21, 22, 23, 24] highlighted that high permittivity values of the tissues were due to high concentration of water like in muscles and tumoral tissue. Low water content tissues like fat and normal tissue have lower permittivity. Joines measurement span in the radiofrequency (RF) range to the low MW range 50-900 MHz but are important because he attempt to separate glandular from the adipose tissue. The corresponding results stated a large dielectric property contrast between malignant and mammary tissue as much as 4:1 and 7:1 for permittivity and conductivity, respectively. Figure 3.3 show Joines results while figure 3.4 shows the Gabriel's ones.

Choi [15] research made in the range 0.5 - 30 GHz showed an high contrast between normal and tumoral tissue.

The most comprehensive study on this topics, performed ex-vivo, has been undoubtedly re-

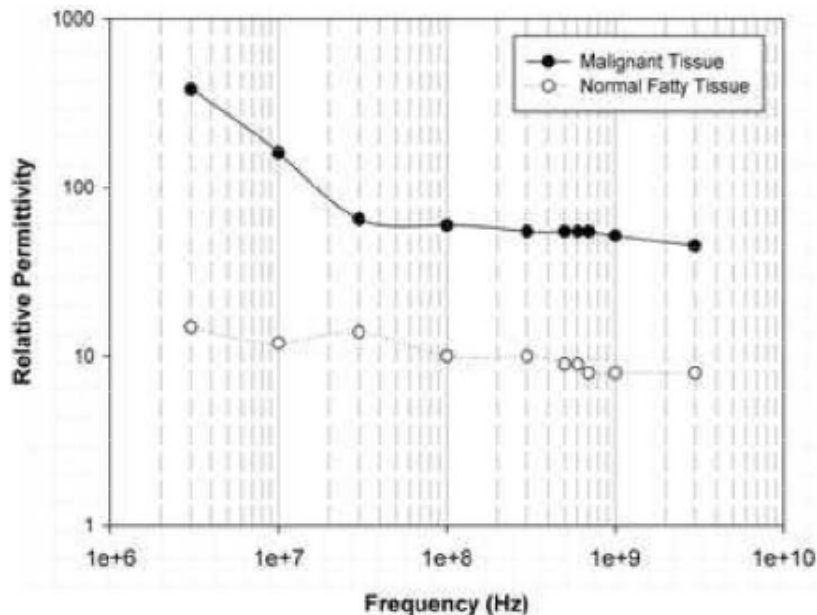


Figure 3.2: Relative permittivity of breast malignant and healthy tissue versus frequency in the Chaudhary’s study.

alized in 2007 by Lazebnik and made available through two articles [37, 38] still a reference for most MWI studies. The first study involved an high number of samples of normal tissue obtained from breast reduction, biopsy or mastectomy carried out short time before using an open-ended coaxial probe, the second one treated also tumoral tissue. The goal was to characterize the breast dielectric properties in the microwaves range 0.5-20 GHz and correlate the measurements with the specimen histopathological analysis. The tissue composition was quantified in terms of percentages of adipose, fibroconnective and glandular tissues. The study of distributions of sample tissue compositions demonstrates that the dielectric properties of breast tissue are primarily determined by the adipose content of each tissue sample. Final results of the analyses show that “the microwave-frequency dielectric-properties contrast between malignant breast tissues and normal adipose tissues is large, ranging up to a 10:1”; on the contrary the dielectric-properties contrast between malignant and normal fibroglandular breast tissues is considerably lower and no more than approximately 10%. Figure 3.5 shows the permittivity (left) and conductivity (right) values for low, medium and high adipose content tissue along with the tumoral one versus the microwave frequency obtained by Lazebnik.

Recently Meaney [44] put in serious discussion the procedure on how results of Lazebnik above described are obtained, in particular those stating the low contrast between low-adipose tissues and tumor. Its article demonstrate that an open ended dielectric probe can not give an accurate estimate of tissue properties over a heterogeneous sensing volume 2 to 3 mm below the surface of

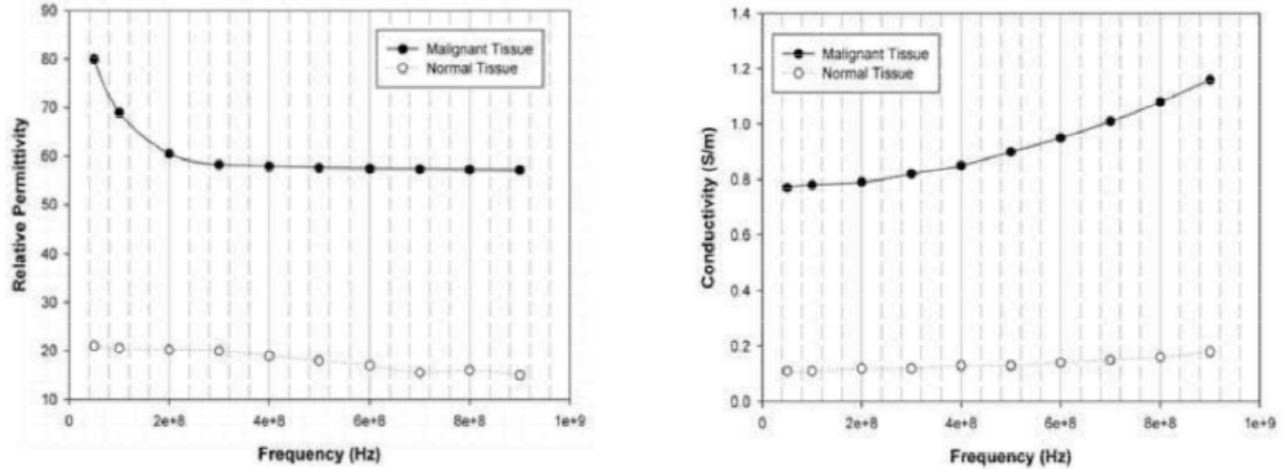


Figure 3.3: Joines results in frequency range 50-900 MHz for relative permittivity (left) and conductivity (right). Black line represents the tumor tissue values and white line is for normal tissue.

the probe because the material within the first few hundred microns exerts the dominant influence on the estimated properties and how the probe perceives them.

When the open-ended coaxial probe does not recover an appropriately averaged property estimate in the layered test configuration its fidelity is affected. If used to measure the dielectric properties of more randomly arranged heterogeneous mixtures of tissue like the breast, which commonly has infiltrations of fibroglandular tissue interwoven within a matrix of adipose cells, can produce incorrect values. Adipose tissue is more predominant in the breasts of many women, it is largely homogeneous and easily sampled. Fibroglandular breast tissue is more challenging. Joines results, even if conducted in a small range are more conclusive than that reached from Lazebnik where it is not return the compositional averages of the tissue specimens evaluate.

The exact contrast between normal and cancerous tissue is, at present, an open issue and more research is still to be followed.

3.3 MWI reconstruction methods

Research in MWI can be divided mainly into linear methods radar-like imaging and microwave tomography (MWT) that uses a non-linear approach. In both techniques, the scattered field is measured at several observation points outside the object under test. The linear methods are easier and do not take into account mutual scattering inside the breast tissues and their capabilities are limited to the detection and localization of breast abnormalities.

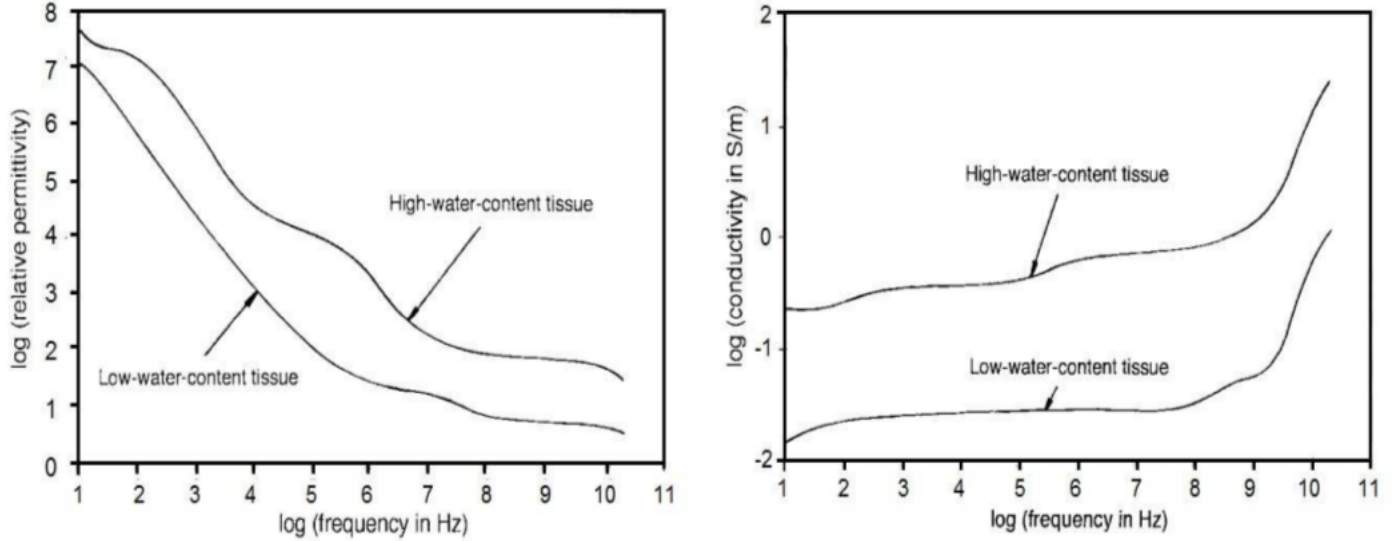


Figure 3.4: Comparison between high water content and low water content of relative permittivity (left) and conductivity (right) versus frequency obtained by Gabriel.

3.3.1 Tomographic method

In MWT image reconstruction, a nonlinear inverse scattering problem must be solved to retrieve a map of the dielectric properties of the target [46] either in 2D or 3D. In this situation the solution is iteratively adjusted so that the magnitude and phase of the measurements predicted using the transmitted waveforms as known inputs converge as closely as possible with those actually observed. This difference from actual measured data and the predicted ones is mathematically treated as the search of the minimum of a functional. It is therefore required a numerical solver, known in literature as forward solver, and a good starting point, an a priori knowledge that helps the optimal convergence and the time needed to produce an image can be critical. This implies a great importance of a preliminary assessment of the scattering scenario. Another known limitation is due to the resolution of an ill-posed problem and the cost function is not convex so it is subject to false solutions and some regularization procedures, such as Tikhonov, are required [49].

An example of a non-linear method based on the Gauss-Newton (GN) minimization algorithm is now briefly detailed.

GN method tries to estimate the spatial distribution of the electrical parameters inside the imaging domain using the complex square wave-number:

$$k^2(\mathbf{r}) = \omega^2 \mu_0 \varepsilon(\mathbf{r}) + j\omega \mu_0 \sigma(\mathbf{r}) \quad (3.2)$$

The wave-number distribution is obtained after solving:

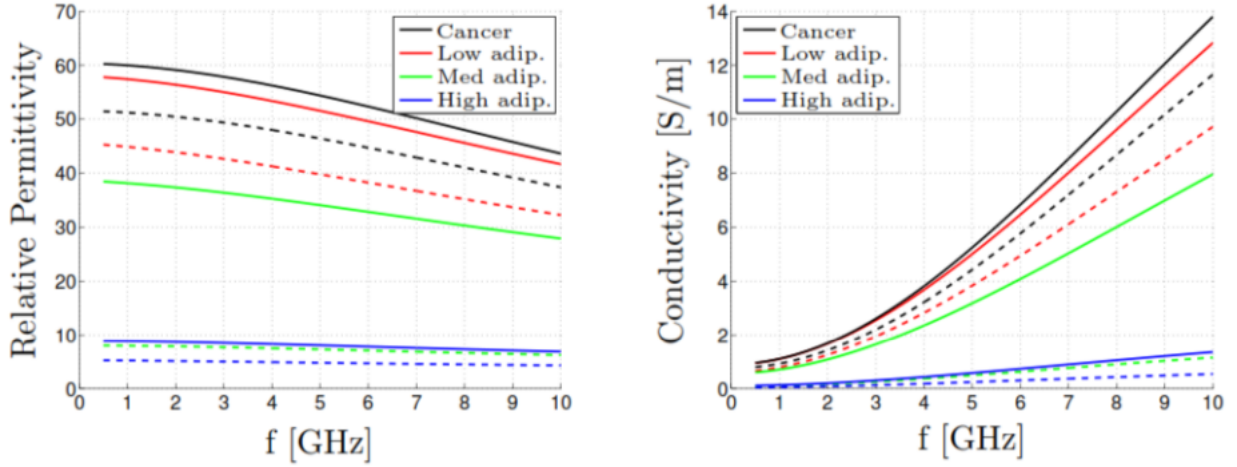


Figure 3.5: Permittivity and conductivity values obtained by Lazebnik [38]. The data represents the values for cancerous tissue (black line) and normal tissue obtained from cancer surgeries with low, medium and high content of adipose tissue (red, green, and blue lines, respectively). For each tissue type, the data for the 25 percentile is given by the dashed line and the data for the 75 percentile by the solid line.

$$\mathbf{k}^2 = \operatorname{argmin}\{\|\mathbf{E}^{\text{meas}} - \mathbf{E}^{\text{calc}}(\mathbf{k}^2)\|_2^2\} \quad (3.3)$$

where \mathbf{E}^{meas} is the measurements vector and \mathbf{E}^{calc} vector has been calculated from the forward solver. Image reconstruction is reached considering the log-phase variations and the difference is made between a situation with and without a target:

$$\mathbf{E}_{\text{lp}}^{\text{meas}} = \begin{bmatrix} \log|\mathbf{E}_{\text{target}}^{\text{meas}}| - \log|\mathbf{E}_{\text{empty}}^{\text{meas}}| \\ \text{phase}(\mathbf{E}_{\text{target}}^{\text{meas}}) - \text{phase}(\mathbf{E}_{\text{empty}}^{\text{meas}}) \end{bmatrix} \quad (3.4)$$

For calculated data we have:

$$\mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}^2) = \begin{bmatrix} \log|\mathbf{E}^{\text{calc}}(\mathbf{k}^2)| - \log|\mathbf{E}^{\text{calc}}(\mathbf{k}_0^2)| \\ \text{phase}(\mathbf{E}^{\text{calc}}(\mathbf{k}^2)) - \text{phase}(\mathbf{E}^{\text{calc}}(\mathbf{k}_0^2)) \end{bmatrix} \quad (3.5)$$

In this expression \mathbf{k}_0^2 represents the known distribution when the target is not present. Once defined this parameter, we deal with the following minimization problem:

$$\mathbf{k}^2 = \operatorname{argmin}\{\|\mathbf{E}_{\text{lp}}^{\text{meas}} - \mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}^2)\|_2^2\} \quad (3.6)$$

The solution is then obtained using a forward solver to compute the Maxwell's equation and the minimization is reached using the GN method. The non-linear expression of the field is written in term of the wavenumber distribution and can be locally treated as a first order Taylor's expansion:

$$\mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}_{n+1}^2) = \mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}_n^2) + \mathbf{J}(\mathbf{k}_n^2)(\mathbf{k}_{n+1}^2 - \mathbf{k}_n^2) \quad (3.7)$$

\mathbf{J} is the Jacobian matrix and n represents the iterations number. Now we can express the global minimization in term of many linear problems of local minimization. GN works in five steps:

1. the forward solver is used to calculate the electric field in the imaging domain starting from the parameters defining the wavenumber \mathbf{k}_n^2 and stopping if the convergence criterion is achieved or the maximum iteration number is reached;
2. calculate the Jacobian for the actual distribution \mathbf{k}_n^2 ;
3. finding the Newton's direction \mathbf{d}_n solving:

$$\mathbf{E}_{\text{lp}}^{\text{meas}} - \mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}_n^2) = \mathbf{J}(\mathbf{k}_n^2)\mathbf{d}_n \quad (3.8)$$

using the Tikhonov regularization algorithm, with λ regularization parameter:

$$(\mathbf{J}^T\mathbf{J} + \lambda\mathbf{I}) = \mathbf{J}^T(\mathbf{E}_{\text{lp}}^{\text{meas}} - \mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}_n^2)) \quad (3.9)$$

4. calculate the Newton step that satisfies:

$$\alpha_n = \text{argmin}\{\|\mathbf{E}_{\text{lp}}^{\text{meas}} - \mathbf{E}_{\text{lp}}^{\text{calc}}(\mathbf{k}_n^2 + \alpha_n\mathbf{d}_n)\|_2^2\} \quad (3.10)$$

5. update the parameters:

$$\mathbf{k}_{n+1}^2 = \mathbf{k}_n^2 + \alpha_n\mathbf{d}_n \quad (3.11)$$

An improvement of this algorithm is reached through the Conjugate Gradient (CG) algorithm, which makes use only of the Jacobian instead of the $\mathbf{J}^T\mathbf{J}$ computation, avoiding the huge matrix product.

Other algorithms such as genetic algorithm, distorted Born approximation, Rytov approximation are also used for non-linear imaging reconstruction purpose but are beyond the scope of this work. The interested reader is referred, for example, to [49].

3.3.2 Linear methods

As previously told linear imaging methods aim to generate a map describing the microwave energy after the scattering with breast tissues: in this approach a particular location in the volume to be measured is isolated, and data is taken for each individual voxel. The ability of these systems to detect malignancies relies on the increased backscatter caused by unhealthy tissue. The primary advantages of this approach are two-fold: there is no need for complex inversion techniques and

simple time-gating methods can be used to reduce noise and multiple scattering effects.

Delay-and-sum (DAS) algorithm. In a confocal breast imaging system, UWB pulses are generated by an antenna located on or near the surface of the breast. The backscattered waveform at that particular antenna location is then collected and stored in a computer. The set of N backscattered signals are then time-shifted to achieve coherent addition for a specific focal point within the breast. The focal point is then scanned throughout the breast by adjusting the relative amount of time-shift applied to each backscattered signal calculating the speed signal depending on the mean values of the electrical parameters. Figure 3.6 shows the working idea of DAS.

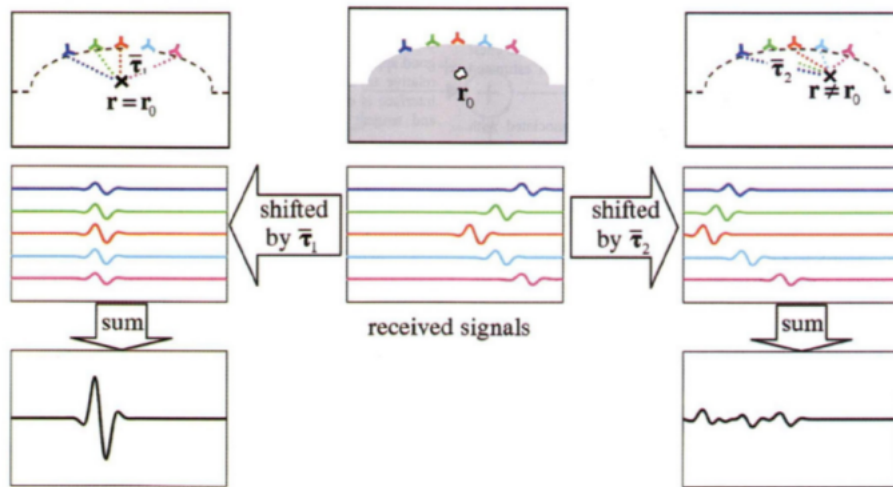


Figure 3.6: Simple delay-and-sum [59].

An improvement of DAS is suggested by Klemm [34] using a curved hemispherical antenna array where data are collected in a multistatic fashion. This signal contains not only information about the lesion but also other unneeded components coming from the skin and mutual coupling. To cancel this clutter the whole system is turned around its center and a second data set is acquired. Constant signal will appear at the same time and will be erased while tumor signals will appear at different time unless it is located in the center.

Other linear algorithms. Other algorithms falling in linear approach are the Delay-Multiply-And-Sum (DMAS): again a variation of the DAS algorithm where a multiplication step is added following the principle that multiplication can reduce the noise and increasing the signal providing a better resolution image.

Since the simple delay-and-sum beamformer described above does not have the ability to compensate for dispersion and other frequency-dependent propagation effects, Li developed a microwave imaging via space-time (MIST) beamforming system [39]. Like the confocal system, each

element of an antenna array both transmits an UWB pulse into the breast of a patient lying in the supine position and measures the backscattered waves. Just as in the delay-and-sum beamformer, the MIST beamformer performs spatial focusing by time-aligning these backscattered waves. Additionally, however, the MIST beamformer compensates for frequency-dependent propagation effects and optimally discriminates against artifacts and noise and this is achieved through the use of a frequency dependent filter in each channel [59].

As an alternative to MIST, Fear and Sill have developed a similar system that has been termed Tissue Sensing Adaptive Radar (TSAR). Like MIST, it seeks to address the shortcomings of the simple confocal system; however, there are a few fundamental differences between the two. For example, in the TSAR system, the patient lies prone and the pendulous breast is scanned from surrounding locations. Additionally, the TSAR system uses less complicated clutter reduction methods than MIST [59].

3.4 The TR-MUSIC algorithm

The reconstruction procedure, as told, involves the resolution of an inverse problem. We afford this problem using a linear technique called Time Reversal Multiple Signal Classification (TR-MUSIC) and developed by Devaney in [27] having now some renewed attention in MWI breast cancer community.

Time reversal

If we assume that the involved signals are expressible using the Fourier transform:

$$X(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) e^{-i\omega t} d\omega$$

having

$$\tilde{X}(\omega) = \int_{-\infty}^{+\infty} X(t) e^{-i\omega t} dt$$

this ensures we can perform all operations in the frequency domain.

Consider an array of N antennas, centered at the space points \mathbf{R}_j , $j = 1, \dots, N$ in which are embedded one or more scatterers (or targets) settled in positions \mathbf{X}_i , $i = 1, \dots, M$. Making the simplifying assumptions that the antennas are point sources and the targets are ideal point scatterers and neglecting all multiple scattering between targets, we obtain the expression of the field radiated by the j -th antenna and incident in the \mathbf{r} position:

$$E_{\text{inc}}^j(\mathbf{r}, \omega) = G(\mathbf{r}, \mathbf{R}_j; \omega) e_j(\omega) \quad (3.12)$$

where $G(\mathbf{r}, \mathbf{R}_j; \omega)$ is the 2 dimensional free space propagator whose expression in term of hankel function has been already defined [2.22](#).

The total scattered field, collected in the same position, due to all scatterers reads:

$$E_{\text{scatt}}^j(\mathbf{r}, \omega) = \sum_{m=1}^M e_j(\omega) G(\mathbf{r}, \mathbf{X}_m; \omega) \tau_m(\omega) G(\mathbf{X}_m, \mathbf{R}_j; \omega) \quad (3.13)$$

where $\tau_m(\omega)$ is the scattering strength of the m-th target and $e_j(\omega)$ is excitation amplitude of the j-th antenna. When all antennas are simultaneously excited using the voltage $e_j(\omega)$, the total incident and scattered fields are given by:

$$E_{\text{inc}}(\mathbf{r}, \omega) = \sum_{j=1}^N E_{\text{scatt}}^j(\mathbf{r}, \omega) = \sum_{j=1}^N G(\mathbf{r}, \mathbf{R}_j; \omega) e_j(\omega) \quad (3.14)$$

$$E_{\text{scatt}}(\mathbf{r}, \omega) = \sum_{j=1}^N E_{\text{scatt}}^j(\mathbf{r}, \omega) = \sum_{j=1}^N \sum_{m=1}^M \tau_m(\omega) G(\mathbf{r}, \mathbf{X}_m; \omega) e_j(\omega) G(\mathbf{X}_m, \mathbf{R}_j; \omega) \quad (3.15)$$

We assume that the field read to l-th antenna port is given only by the scattered field, so that:

$$v_l(\omega) = \sum_{j=1}^N E_{\text{scatt}}^j(\mathbf{R}_l, \omega) = \sum_{j=1}^N \sum_{m=1}^M G(\mathbf{R}_l, \mathbf{X}_m; \omega) \tau_m(\omega) G(\mathbf{X}_m, \mathbf{R}_j; \omega) e_j(\omega) \quad (3.16)$$

To express the above equation in a compact form, the N-dimensional Green vector function $g_m(\omega)$ is introduced: any vector element represent the green propagator from the m-th scatterer to any antenna:

$$g_m(\omega) = [G(\mathbf{R}_1, \mathbf{X}_m; \omega), G(\mathbf{R}_2, \mathbf{X}_m; \omega), \dots, G(\mathbf{R}_N, \mathbf{X}_m; \omega)]^T \quad (3.17)$$

the contribution of the m-th single scatterer referred to all antennas is given by the B-matrix:

$$B_m(\omega) = \tau_m(\omega) g_m(\omega) g_m(\omega)^T \quad (3.18)$$

summing up to all the scatterers contribution, we obtain:

$$\mathbf{A}(\omega) = \sum_{m=1}^M \tau_m(\omega) g_m(\omega) g_m(\omega)^T \quad (3.19)$$

and it is possible to re-write [3.16](#) as:

$$\mathbf{v}(\omega) = \mathbf{A}(\omega)\mathbf{e}(\omega) \quad (3.20)$$

$\mathbf{A}(\omega)$ is the so called *Multistatic response matrix* (MSR) and is the propagator matrix, while $\mathbf{e}(\omega)$ represents the excitation voltage of the antennas:

$$\mathbf{e}(\omega) = [e_1(\omega), e_2(\omega), \dots, e_N(\omega)] \quad (3.21)$$

From the MSR-matrix it is defined the time-reversal matrix (TRM) [27] as follows (the explicit dependence from ω is by now omitted):

$$\mathbf{T} = \mathbf{A}^\dagger \mathbf{A} = \mathbf{A}^* \mathbf{A} \quad (3.22)$$

\mathbf{A}^\dagger is the Hermitian matrix of \mathbf{A} and \mathbf{A}^* is the complex-conjugate, the last equality is simply derived from the symmetric behavior of the green function that reflects on the MSR. The TRM is Hermitian and non-negative hence it possesses a complete set of orthonormal eigenvectors having non-negative eigenvalues.

3.4.1 MUSIC

The MUSIC algorithm [56] is a way to extract the range of a self-adjoint linear operator, say L . In this case is well known [35] that L has a diagonal representation and the eigenvalues are all non-negative: let $\lambda_1 \geq \lambda_2 \dots \geq \lambda_N$ the eigenvalues defined in an ordered way and v_1, v_2, \dots, v_n the corresponding eigenvectors.

Suppose that $\lambda_{M+1} \dots \lambda_N$ are all zero, the corresponding eigenvectors $v_{M+1} \dots v_N$, span the *null space* of L . It can also happen that these eigenvalues are almost zero but below the noise level of the system described by L . Then these eigenvectors form a base for, or equivalently span, the *noise subspace* of L [14]. The projection operator P , defined over the noise subspace is thus defined:

$$P = \sum_{k>M} v_k \overline{v_k}^T \quad (3.23)$$

where the bar stands for the ordinary complex conjugate operation e the superscript T denotes the transpose of the eigenvector. The signal subspace is orthogonal to the noise subspace for an auto-adjoint operator, so any vector f is in the signal subspace if the value of its projection onto the noise subspace is zero

$$||P_{\text{noise}} f|| = 0 \quad (3.24)$$

or

$$\frac{1}{\|P_{\text{noise}}f\|} = \infty \quad (3.25)$$

The MUSIC algorithm makes use of the fact that the TRM is a projection operator onto the subspace of \mathcal{C}^N : the complex conjugates of the Green function vector span the signal subspace \mathcal{S} , while the noise subspace \mathcal{S} is spanned by all eigenvectors with null eigenvalue.

$$\langle \mathbf{G}_{m'}^*, \mathbf{V}_{m'} \rangle = \langle \mathbf{V}_{m'}^*, \mathbf{G}_m \rangle = 0 \quad (3.26)$$

where m' is the index over all the eigenvectors with null eigenvalues. Defining \mathbf{A}_k the *Steering matrix* as the MSR-matrix of the trial point \mathbf{r}_k inside the investigation domain:

$$\mathbf{A}_k = \mathbf{G}_k \tau_m \mathbf{G}_k^T \quad (3.27)$$

where $\mathbf{G}_k = [G(\mathbf{R}_1, \mathbf{R}_k), G(\mathbf{R}_2, \mathbf{R}_k), \dots, G(\mathbf{R}_N, \mathbf{R}_k)]^T$, finally making use of 3.25 we can define the pseudospectrum quantity:

$$\Psi = \frac{1}{|\mathbf{P}(\mathbf{A}_k)|^2} = \frac{1}{\left| \sum_{m'=M+1}^N \langle \mathbf{A}_k, \mathbf{V}_{m'} \rangle \right|^2} \quad (3.28)$$

The pseudospectrum value peaks in correspondence of the presence of a scatterer.

3.4.2 MUSIC for the monostatic configuration

In the monostatic case studied (in literature also called multi-monostatic [53]), a single antenna is moving on a circle and the reflected value of the signal is then collected on N discrete point \mathbf{R}_l . In this situation the MSR-matrix can not be obviously formed and the field value read at the antenna port of the l -th position is:

$$v_l(\omega) = E_{\text{scatt}}(\mathbf{R}_l, \omega) = \sum_{m=1}^M G(\mathbf{R}_l, \mathbf{X}_m; \omega) \tau_m(\omega) G(\mathbf{X}_m, \mathbf{R}_l; \omega) e_j(\omega) \quad (3.29)$$

then the vector of voltage output from the N antenna positions can be defined as

$$\mathbf{v}(\omega) = [v_1, v_2, \dots, v_N]^T \quad (3.30)$$

and can be expressed in term of *Monostatic Response Vector*

$$\mathbf{v}(\omega) = \mathbf{A}(\omega) e(\omega) \quad (3.31)$$

and $\mathbf{A}(\omega) = [A_1, A_2, \dots, A_N]^T$. In this case, instead of dealing with the projection over noise sub-space, we use the projection over the signal sub-space \mathcal{S} [58]:

$$|\mathbf{P}[\mathbf{A}_k(\omega)]|^2 = |(\mathbf{I} - \mathbf{Q})[\mathbf{A}_k(\omega)]|^2 \quad (3.32)$$

where $\mathbf{Q}[\cdot]$ is the projection operator over \mathcal{S} and $\mathbf{A}_k(\omega)$ is the *steering vector* defined as the Monostatic Response Vector in the trial points r_k belonging to the investigation domain.

In this case we can avoid the computation of the TRM since:

$$|(\mathbf{I} - \mathbf{Q})[\mathbf{A}_k(\omega)]|^2 = |\langle \mathbf{A}_k(\omega), \mathbf{A}(\omega) \rangle|^2 \quad (3.33)$$

and finally the pseudospectrum can be computed as:

$$\Psi(\mathbf{r}_k, \omega) = \frac{1}{1 - |\mathbf{P}(\mathbf{A}_k)|^2} = \frac{1}{1 - |\langle \mathbf{A}_k(\omega), \mathbf{A}(\omega) \rangle|^2} \quad (3.34)$$

This expression can easily be seen as the correlation between the two vectors, the collected signal vector and the steering vector, in the \mathbb{C}^N computed as standard inner products. Defining:

$$\cos(\eta) = \frac{|\langle \mathbf{A}_k(\omega), \mathbf{A}(\omega) \rangle|}{\|\mathbf{A}_k(\omega)\| \|\mathbf{A}(\omega)\|} \quad (3.35)$$

it is possible to rewrite 3.25 as follows:

$$\Psi(\mathbf{r}_k, \omega) = \frac{1}{1 - |\mathbf{P}(\mathbf{A}_k)|^2} = \frac{1}{1 - \|\cos(\eta)\|^2} \quad (3.36)$$

without having to deal with the TRM.

The resulting image is now straightforward: if the two vectors define the same direction in the complex hyper-space, the resulting cosine is near one and the denominator approximate to zero hence the psudospectrum peaks, if they are orthogonal the cosine is zero and the psudospectrum reach the value of 1.

3.5 Multifrequency

To improve the resolution of the image and to counteract emerging artifacts, a multiple frequencies version of MUSIC as in [3] is here adopted and the performance of three different combination methods are evaluated and compared to determine the more suitable:

Wide-band (WB) Music algorithm [65]:

$$\Psi_{WB} = \left[\sum_i \frac{1}{\Psi(\omega_i)} \right]^{-1} \quad (3.37)$$

Geometrical mean (GM) Music algorithm:

$$\Psi_{GM} = \sqrt[N]{\prod_i \Psi(\omega_i)} \quad (3.38)$$

Interferometric (I) Music algorithm [53]:

$$\Psi_{GM} = \left[\left| \log \left(\prod_i |\cos \eta| \right) \right| \right]^{-1} \quad (3.39)$$

3.6 Existing MWI systems

In literature some MWI systems are detailed and though used instruments, coupling medium, antenna type and arrangement are different, they share the same architecture that can be synthesized into three main components:

- a set of antennas operating as transmitters and receivers, placed in direct contact with the breast or in proximity by a means of coupling medium;
- front-end devices able to generate and capture the signal in the microwave range;
- back-end devices where data acquisition is stored and the image reconstruction is performed.

When the method of the reconstruction uses radar technique (UWB) these systems might adopt different antennas configuration and the ways the operate are mainly three:

- 1) monostatic: a single antenna is used as transmitter and receiver and is moved around the breast;
- 2) bistatic: two antennas are present and rotate around the breast;
- 3) multistatic: many antennas are concurrently present and transmit in turn.

Many groups worldwide have been developed different prototypes devoted to the early detection of breast cancer. In our objective to reproduce a full breast cancer MWI, we referred to all work actually available, but the most important one is devoted to the team from the Dartmouth college led by Meaney [43] where the system has been used to collect real data from a number of patients.

This system in its first edition, was able to operate in a frequency band ranging from 300 MHz to 1 GHz, later improved to 0.5-3 GHz and the image reconstruction uses the algorithm of Gauss-Newton in 2D. This system uses 32 acquisition channels along with 16 monopole antennas in vertical position that can be put either in RX mode or TX mode. The array is able to produce images from the inner position to outer position determined by the nipple. The array is linked to

the front-end electronics made by:

- heterodyne receiver;
- 200 kHz D/A board;
- function generator;
- mw source.

Data acquisition is performed by a canonical PC. The presence of a network switch guarantees the shift of the transmitting antenna. To obtain a multislice imaging of the breast the antennas can move along the vertical direction.

In the first edition the coupling medium used to minimize the mismatch with the mean permittivity of the breast was a saline solution then substituted with a glycerin-de ionized water solution at 88%: this choice also guarantees a low conductivity estimated in 0.1 S/m at 1 GHz without drawbacks when in contact with the human skin. Moreover tuning this value it is possible to best reproduce the four BI-RADS category (see section 1.4) of the breast density. The full system layout in the actual configuration is showed in figure 3.7.

An exhaustive treatment concerning existing prototypes, system setup, algorithm reconstruction and published results can be found in [25].

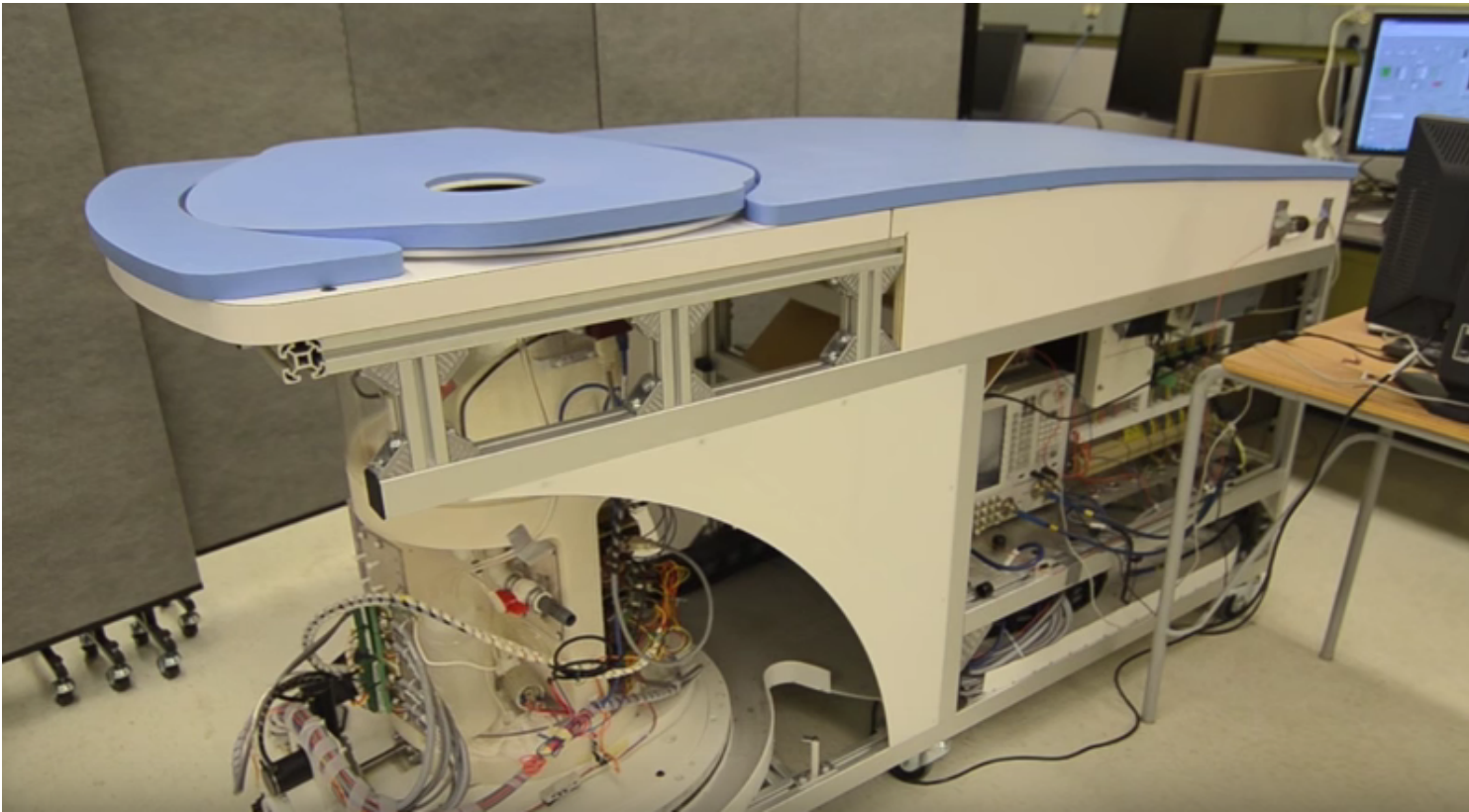


Figure 3.7: The MWI acquisition system of the Dartmouth college.

Chapter 4

Synthetic data simulation

In this chapter the simulation software is presented along with the numerical methods used to obtain the approximate solution. Then the procedure followed to realize a simulated scenario and how to manage it to perform a parametric study is detailed. Finally, the characteristic of all the models involved in the reconstruction stage are discussed and some performance indications to build efficient models are given.

4.1 Numerical solutions

Maxwell's equations can be solved analytically only in very few simple cases where we have the so called “closed form solutions” but their great ability to predict the behavior of the electromagnetic field is actually essential to design any experimental setup or for any needed outcome. The advent of computer technology was a strong impulse to study and develop numerical methods for their solution and the increasing capabilities of the hardware led to more versatility and accuracy in the solution methods. Figure 4.1 displays a summary of the most used numerical methods adopted in the computational electromagnetism.

Actually there are three major ways to perform electromagnetic simulations, that is the numerical computation of the electromagnetic field by means of the computers: the finite difference method (FDM), the finite elements method (FEM) and the method of moments (MoM).

4.1.1 FDM

The FDM is a numerical procedure for converting partial differential equations of a boundary-value problem into a set of algebraic equations that can be used to compute an approximate solution. Among the numerical methods, the FD method was the first to appear and its simplicity makes it very robust and efficient and it has been widely used in several engineering fields. Its application to

the analysis of electromagnetic problems was empowered by the development of a unique scheme for the discretization of the time - domain Maxwell's equations, which is known as the finite difference time - domain (FDTD) method. The basic idea of the finite difference method is to approximate the differentiation operators in the partial differential equations. This approximation can be obtained directly from the definition of the derivative:

$$f'(x) = \frac{df}{dx} = \frac{f(x + dx) - f(x)}{dx} \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} \quad (4.1)$$

In the one-dimensional case Δx represents the step of uniformly divided space.

4.1.2 MoM

MoM, instead of solving the Maxwell's equations in differential form, is intended to solve the so called EFIE/MFIE, electric or magnetic field integral equation, moving the governing equation of a boundary - value problem into a matrix equation that can be solved using computers. Its predominant use is in an open region such as wave scattering and antenna radiation.

4.1.3 FEM

FEM is a numerical procedure to convert partial differential equations into a set of linear algebraic equations to obtain approximate solutions to boundary-value problems. But, instead of approximating the differential operators, it approximates the solution of a partial differential equation [31]. The basic concept in the physical FEM, whose fundamental contribution is due to M.J. Turner and R.W. Clough [62], is the subdivision of the mathematical model into disjoint (non-overlapping) components of simple geometry called finite elements. The response of each element is expressed in terms of a finite number of degrees of freedom characterized as the value of an unknown function, or functions, at a set of nodal points [50]. The response of the mathematical model is then considered to be approximated by that of the discrete model obtained by connecting or assembling the collection of all elements. Unlike finite difference models, finite elements do not overlap in space. In the mathematical interpretation of the FEM, this property goes by the name disjoint support or local support. Because FEM is a discretization method, the number of degrees of freedom of a FEM model can be very high but is necessarily finite. They are collected in a column vector called u . This vector is generally called the degree of freedom (DOF) vector or state vector.

Below the FEM Solution Process Procedures:

1. divide structure into elements with nodes, this is the discretization or meshing step;
2. connect (union or assembly) the elements at the nodes to form an approximate system of equations for the whole structure (forming element matrices);

3. solve the system of equations involving unknown quantities at the nodes (e.g., electric field);
4. Calculate desired quantities (e.g., S-parameters) at selected elements.

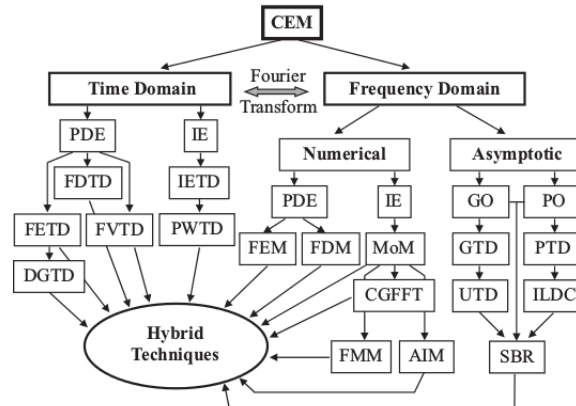


Figure 4.1: Computational methods for solving Maxwell’s equations

4.2 COMSOL modeling

The tools for numerical modeling are today essential for the researchers because they allow the optimization of the design process without complications and at a very low cost. This work, compared to the experimental one, permit the validation of the prototypes, easy customization of the need functionalities and has a great predictive power. COMSOL multiphysics has been used to solve the forward electromagnetic problem and this software belongs to FEM category. The choice has been made, among other electromagnetic simulation software, for its multi-physics environment, in fact future step of this research might also need to handle other fields such as heat transfer or plasma and to be ”modeled” in conjunction with microwaves propagation.

What we refer as the term ”model”? Commonly we deal with the term ”model” as an object that accurately reproduce, in scale, the original. In our sense *model* is intended in a mathematical point of view: a simplification of the original physical problem but that preserves its fundamental behavior under study, that is to model the statistical outcome of a coin toss we do not take in account gravity, friction and other physical aspects.

4.2.1 Space dimensionality

The process of modeling start in COMSOL by choosing the dimensionality of the problem, typical 2D or 3D as we used in our case, even if 2D axis symmetric can be set if the geometry allows it, that is geometry can be obtained revolving a cross section around an axis. In this case a real 3D geometry has a greatly simplified computational complexity meaning lower computer power, memory requirements and time to solve. Figure 4.2 shows the COMSOL user interface for the specification of the model dimensionality.

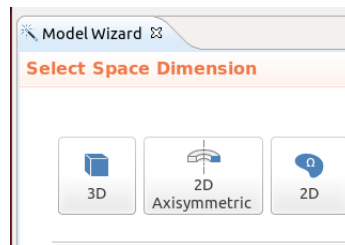


Figure 4.2: COMSOL user interface: space dimensionality of the model

4.2.2 Interface and study

The next choice involves the type of physics describing the scenario, this relates the model with the partial differential equations governing the phenomenon. COMSOL names **RF module** the case where the general field of electromagnetic waves are solved. Moreover, inside this interface one must choose the type of study to perform, in our particular case where the **E** field has to be evaluated, the natural choice has been *electromagnetic waves, frequency domain*. In this selection the electromagnetic field is solved using phasors whose theory has already been detailed in chapter 2.

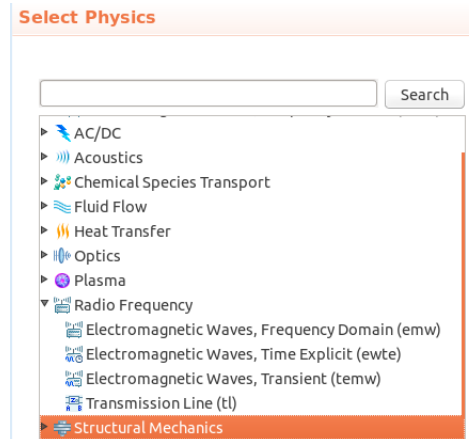


Figure 4.3: COMSOL choice about the type of physics and the related studies

4.2.3 The geometry

After defining the scope of the simulation, the setup of the geometry must be given, where for geometry we refer to all the real objects inside the scenario. In fig 4.4 is displayed a resulting geometry used to model a realistic breast.

After adding all the objects it is necessary to call the finalization process that consists in forming a union or an assembly, in this step each object has been divided into domains, boundaries

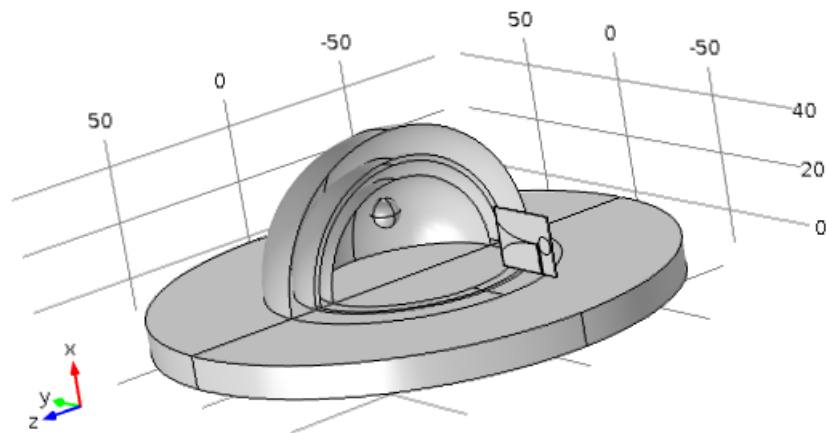


Figure 4.4: An example of geometry generated using the COMSOL internal CAD

or edges. This is particularly needed when objects are collocated inside each other or when there are intersections between them.

Commonly we have to deal with a possible infinite domain and suitable boundary conditions must be provided: it is therefore mandatory to add a particular material whose characteristic is to be total absorbing at any frequency. Clearly this material is not a real material and is known as perfectly matching layer (PML) and was firstly proposed by Berenger [8]; its job is basically to null the field at a finite distance reducing any kind of "non-physical" reflections and a common assumption is to set its thickness to be few wavelength of the radiation. Having to deal with the PML one must ensure that field really vanish at the end of the computational domain because this could affect the correctness of the solution. In fig 4.5 we can see a correct use of the PML where the field at the boundary is null, if the field should not vanish the solution could contain unwanted reflections due to the finite dimension of the simulated scenario.

4.2.4 Assigning material to the objects

After completing the addition of the objects it is necessary to specify the materials they are made up. It is possible to choose them from a material library or add a new material having specific properties. In fig 4.6 are shown the properties for the copper taken, for example, from the material library and linked to the surfaces constituting the antenna dipole. The green check marks on left assure that all the needed properties used by the study to solve the equations are correctly stated.

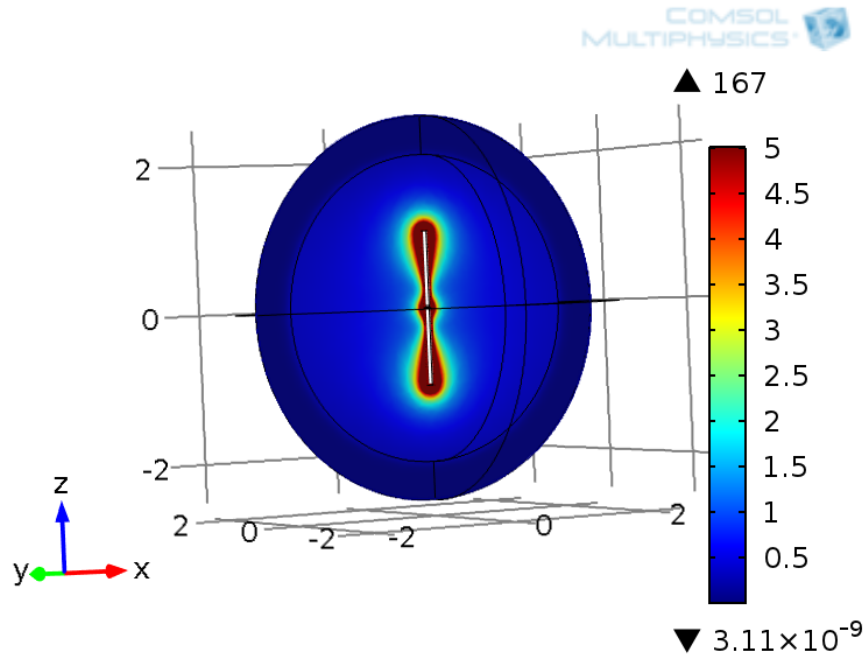


Figure 4.5: PML is correctly set: at the boundaries the field is null.

4.2.5 Mesh

The meshing procedure is one of the most important features we have to care about: meshing means dividing the domain into smaller sub-domains in which the solution can be approximated by elementary basis functions that can be constant, linear or even of higher order [16]. Depending on the element order of magnitude in the model, a finer or coarser mesh is required to resolve the solution. In general finer meshing is needed when the solution changes due to the geometry of the objects, that is a complicated geometry with several edges or fine details. A good practice is to remove such details if they do not influence the solution because they produce a lot of unnecessary mesh elements. Another important consideration is that to resolve a wave properly, it is necessary to use about 10 linear (or five 2nd order) elements per wavelength. For 2D problems the elements are triangles and for 3D they are tetrahedra, they are obtained from three nodes joined by three edges, these edges can be either straight lines or curvilinear. The user is requested to set the minimum and maximum element size for each domain or boundary, if the default meshing provided by the software is not adequate, otherwise the process is completely software-driven. An inadequate meshing leads, usually, to bad solutions of the problem and frequently the convergence is not guaranteed. Finally a good mesh represents the right trade-off between a greater number of meshing elements and the speed of convergence, so having a lower number of elements does not always correspond to a faster solution. In figure 4.7 can be seen the fine meshing on the antenna arms while a coarse mesh on the domain.

Material Contents					
Property	Name	Value	Unit	Property group	
✓ Relative permeability	mur	1	1	Basic	
✓ Electrical conductivity	sigma	5.998e7	S/m	Basic	
✓ Relative permittivity	epsilo	1	1	Basic	
Coefficient of thermal ex	alpha	17e-6[1/	1/K	Basic	
Heat capacity at constant	Cp	385[J/(k	J/(kg·K	Basic	
Density	rho	8700[kg	kg/m ³	Basic	
Thermal conductivity	k	400[W/(W/(m·	Basic	
Young's modulus	E	110e9[P	Pa	Young's modu	
Poisson's ratio	nu	0.35	1	Young's modu	

Figure 4.6: Specification of the electrical properties of the material of the object. The green check-mark assures that all the needed information are correctly set

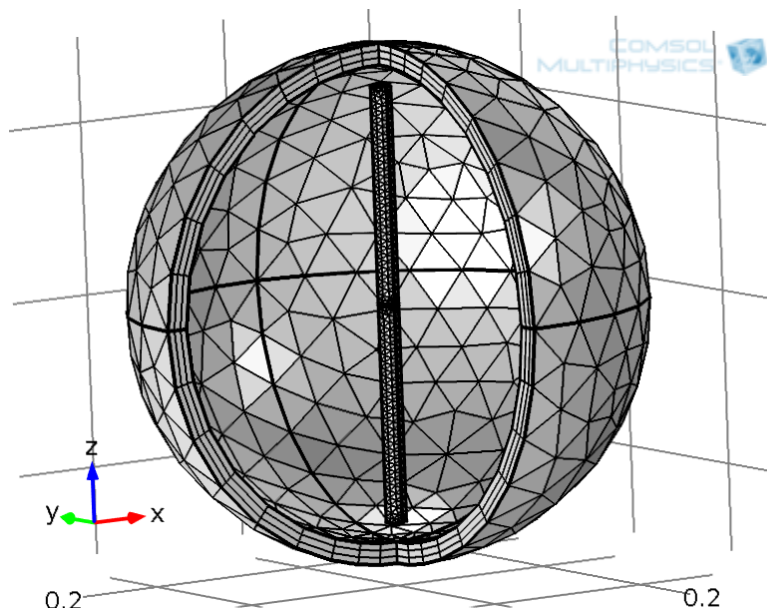


Figure 4.7: Meshing: different objects are meshed using different element size

4.2.6 Lumped port

A particular element one has to deal with in the *RF module* is the 'Lumped port'. Such element is used to simulate how active transmitters are fed. Concretely when the study does not involve the modes inside a transmission line, all the parts that contribute to the feeding can be neglected, they are indeed not added to the scenario, using the approximation of Lumped ports.

This port is not as accurate as the ordinary port in terms of calculating S-parameters, but it is easier to use. The lumped port must be applied between two metallic objects separated by a distance much smaller than the wavelength, that is a local quasi-static approximation must be justified. This is because the concept of port or gap voltage breaks down unless the gap is much smaller than the local wavelength. In transmission line theory voltages and currents are dealt with

rather than electric and magnetic fields, so the lumped port provides an interface between them. When a lumped port specified as an input port (the antenna is in TX mode) the software calculates the impedance, Z_{port} and S_{11} S-parameter for that port [16] which are given by the relations:

$$Z_{port} = \frac{V_{port}}{I_{port}} \quad (4.2a)$$

$$S_{11} = \frac{V_{port} - V_{in}}{V_{in}} \quad (4.2b)$$

4.2.7 Antenna modeling

Dipole antenna

Our work starts with modeling and characterizing a dipole antenna. This kind of antenna is mainly used in some MWI prototypes as described in chapter 3. This consideration and the ease with which it is possible to find on the market at a low price such antennas working in the microwave range, made the dipole our first choice. It is well known that the dipole is resonant in air at half-wavelength [6], in a coupling medium used to reduce the mismatch with mean properties of the tissues constituting the breast, this again holds true but the resulting wavelength is shortened by a factor of $\frac{1}{\sqrt{\epsilon}}$, the dependence of μ is ignored as we deal with non-magnetic materials.

Our implementation (see figure 4.8) has been realized with two thin metallic rods that have a sinusoidal voltage difference applied between them and a lumped port is used to simulate the feeding components. The standard dipole antenna is a narrow band antenna because it has only one resonant frequency in the range where we expect to work with (harmonics appear higher), this characteristic should make it not suitable for Ultra Wide Band (UWB) application that is the antenna can radiate efficiently in a broad range of frequencies. Another lacking characteristic is the directivity, in particular the radiation pattern is omni-directional causing much of the power to be radiated far from the investigating domain. In the following we will show that these features permit anyway a good quality of reconstruction either using simulated or experimental data.

Vivaldi antenna

To overcome the limiting features of the standard dipole many researchers considered uwb and directive antennas. A particular interest has been put on Vivaldi antennas [55, 52] whose characteristics are to be wide-band and moderately directive. We start from the COMSOL library Vivaldi antenna model designed to be used in air and whose geometry is displayed in figure 4.9: the tapered slot is patterned with a perfect electric conductor ground plane on the top of the dielectric substrate and a simple exponential function, $e^{0.044x}$ is adopted to create the tapered slot

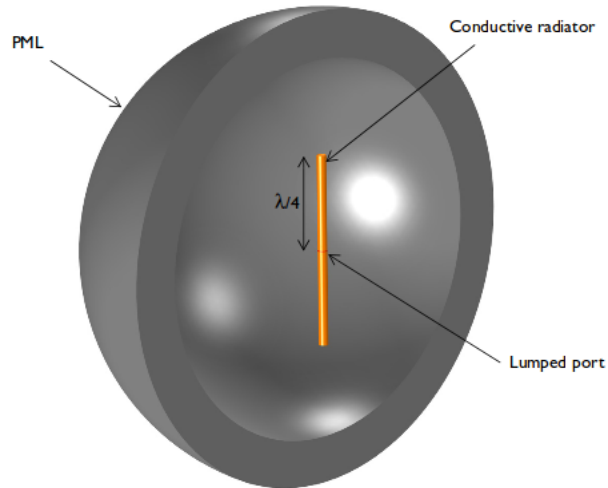


Figure 4.8: Dipole antenna geometry. The arm length is defined to be, in air, a quarter of the wavelength of the cosinusoidal impressed current. A spherical PML is used to limit the solution domain

curves. The tapering procedure aids to achieve a wide-band behavior. One end of the slot is open to air and the other end is finished with a circular slot. On the bottom of the substrate, the shorted 50Ω microstrip feed line is modeled as PEC surfaces. The entire modeling domain is bounded by a perfectly matched layer to stop the computation of the solution at finite distance.

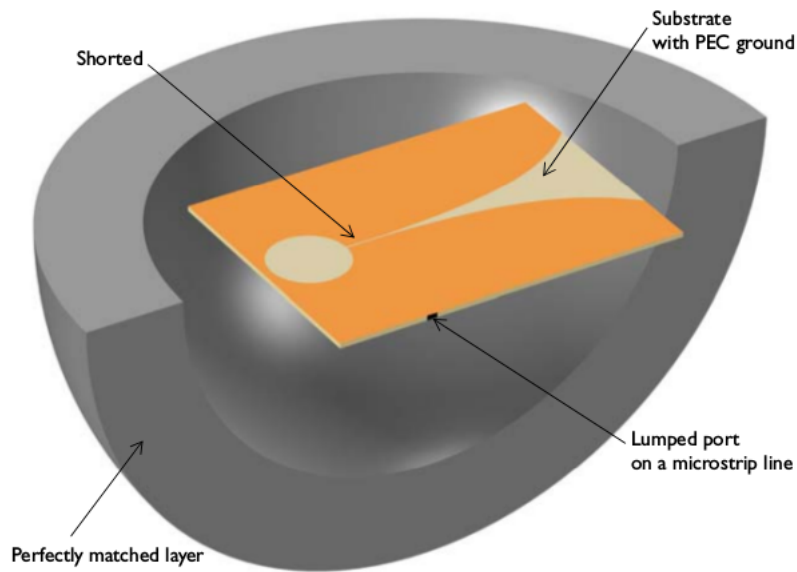


Figure 4.9: Vivaldi antenna. A spherical PML is used to limit the solution domain

Point source for 2D model

The point source has been implemented to speed-up the forward computation needed for the tomographic approach and to reach higher frequencies in the simulation designed for detecting the profile of extended targets. Given the high requirements of the non linear approach we decided to simplify the geometry moving toward a 2D model, because even if time might not matter like in linear approach, a 3D geometry at high frequency requires a very fine element meshing and keep the geometry fairly close to the realistic dimension of the breast and tumor is memory consuming and can easily exceeds 64 GByte of RAM. The point source is managed in COMSOL as an out-of-plane line of current that can be excited at a desired frequency. Figure 4.10 shows on the left the geometry of a 2D model: black spot is the line current out-of-plane, on the right the corresponding meshing to produce nodes and edges on which the Helmholtz equation must be solved.

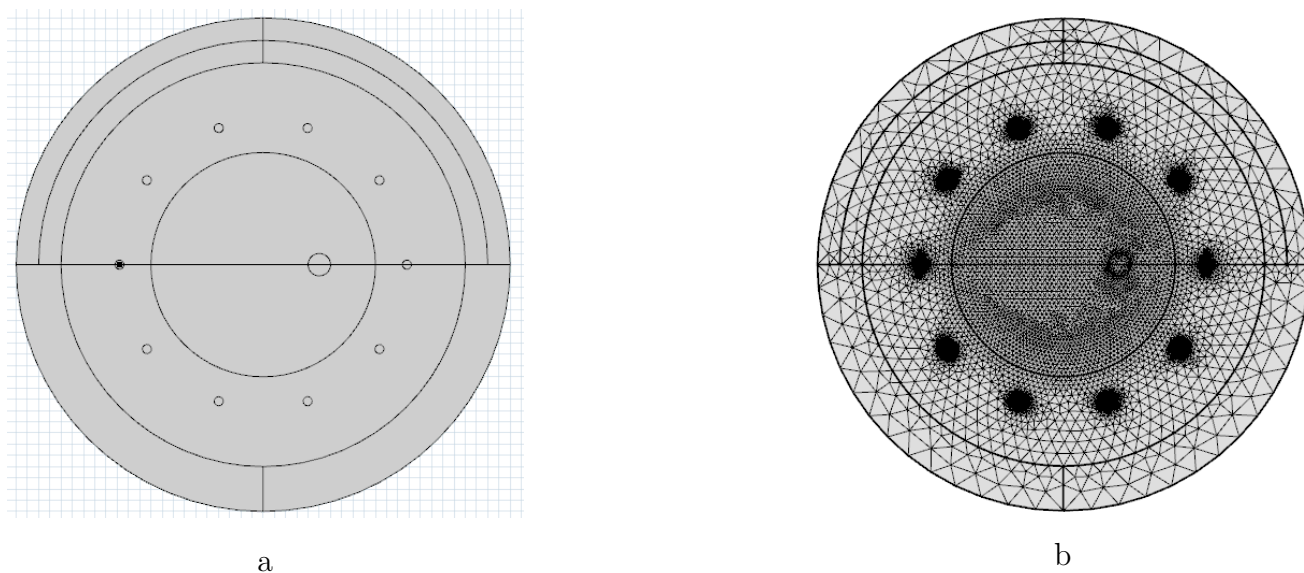


Figure 4.10: (a) geometry and (b) meshing for a 2D scenario with a point source.

4.3 MATLAB automation

COMSOL Multiphysics can be connected with MATLAB through the LiveLink interface. This interface makes available a scripting environment where it is possible to control the flow of programs such as implementing nested loops using *for* or *while* commands, making use of the conditional model settings with *if* or *switch* statements, or handling exceptions using *try* and *catch*. These are basic examples but all the object models can be navigated into. Moreover all the results can be easily accessed and extracted to be used in the MATLAB workspace or saved for future use. Using this feature MATLAB scripts have been built with the capability to generate several configurations

where the important parameters of the simulations were changed and the results stored for the post-processing activities.

4.4 Homogeneous breast model

The breast for 2D approach has been modeled with a cylinder with 50 mm radius having a skin layer 2 mm thick while the tumor is a cylinder of 2.5 mm radius. The height of the two cylinders is always kept at least 4 wavelength long to minimize the diffraction effects. The breast is also immersed in a coupling medium with fixed relative permittivity $\varepsilon_r = 10$ and very low conductivity $\sigma = 0.05 S/m$. The relative permittivity of the breast is modeled using the following Debye parameters 2.12 to take into account its variation versus the frequency used:

$$\varepsilon_\infty = 7, \varepsilon_{sup} = 10, \tau = 7 \text{ ps}, \sigma_b = 0.15 \text{ S/m} \quad (4.3)$$

so the mean ε_r in our range is about 10. The skin is represented in the geometry as the cylindrical layer before the internal section of the breast with fixed relative epsilon and $\varepsilon_r = 36$ and $\sigma = 4 S/m$. Debye parameters for the tumor are:

$$\varepsilon_\infty = 4, \varepsilon_{sup} = 54, \tau = 7 \text{ ps}, \sigma_t = 0.7 \text{ S/m} \quad (4.4)$$

corresponding to a mean ε_r in our range about 50. These values give an almost constant contrast factor $\varepsilon_{scat} : \varepsilon_{breast} = 5 : 1$. No inhomogeneities are considered other than the tumor. Table 4.1 summarize the geometrical and physical parameters used to model the homogeneous breast.

4.5 A breast of intermediate complexity

As it will be showed in the next section, the reconstruction of a tumor in an homogeneous medium is of good quality but this can not be said for a real breast model where the fibroglandular tissue has very similar electrical properties of the tumor tissue. To investigate the behavior of a scenario having an intermediate complexity a third simulation model has been introduced. In particular a fibroglandular island with a rugged profile is used. The profile is a 5-pointed star with the edges smoothed, the analytic polar equation of the curve expressed in mm is:

$$\rho = 25 \cdot (1 + 0.2 \cos(5\theta)) \quad (4.5)$$

At first, the model has been implemented following the identical COMSOL model used in the case of the homogeneous breast in 3-dimensions and detailed in section 4.4, but due to the high

Table 4.1: Summary of the parameters used for homogeneous breast modeling in COMSOL. The electrical parameters modeled with the Debye dispersion model are referred to the illumination frequency* $\nu = 2$ GHz

Parameter	Symbol	Value or expression
coupling medium radius	r_{domain}	100 mm
coupling medium permittivity	ε_{bck}	10
coupling medium conductivity	σ_{bck}	0.05 S/m
breast radius	r_b	50 mm
breast Debye parameters	$[\varepsilon_{\infty}, \varepsilon_{sup}, \tau, \sigma]$	[7, 10, 7 ps, 0.15 S/m]
breast relative permittivity*	ε_b	9.97
breast conductivity*	σ_s	0.18
skin thickness	l_{skin}	2 mm
skin relative permittivity	ε_s	36
skin conductivity	σ_s	4 S/m
tumor radius	r_s	2.5 mm
tumor Debye parameters	$[\varepsilon_{\infty}, \varepsilon_{sup}, \tau, \sigma]$	[4, 54, 7 ps, 0.7 S/m]
tumor relative permittivity*	ε_b	53.6
tumor conductivity*	σ_t	1.19 S/m
antenna radius	r_a	2 mm
illumination frequency range	ν	1-5 GHz

frequencies needed for perform the correct study that require too much memory, it has been converted in 2D. The resulting geometry is displayed in figure 4.11.

In this case the real behavior of the transmitting antenna is replaced with a line current point-source I_{exc} , and the S-parameters of the receiving antennas, if any, are replaced with integral mean of the \mathbf{E} over a suitable surface divided by the excitation current:

$$S_{n1} = \frac{\int_{\Omega_n} \mathbf{E}_z d\Omega_n}{I_{\text{exc}} \int_{\Omega_n} d\Omega} \quad (4.6)$$

4.6 A real breast modeling

After a very simple breast model, but with the advantage that the architecture can be easily replicated in an experimental phantom, where only the skin layer has been added other that the tumor in an homogeneous breast, a real breast model has been approached. This model is available from a public on-line repository (<http://uwcem.ece.wisc.edu/phantomRepository.html> -

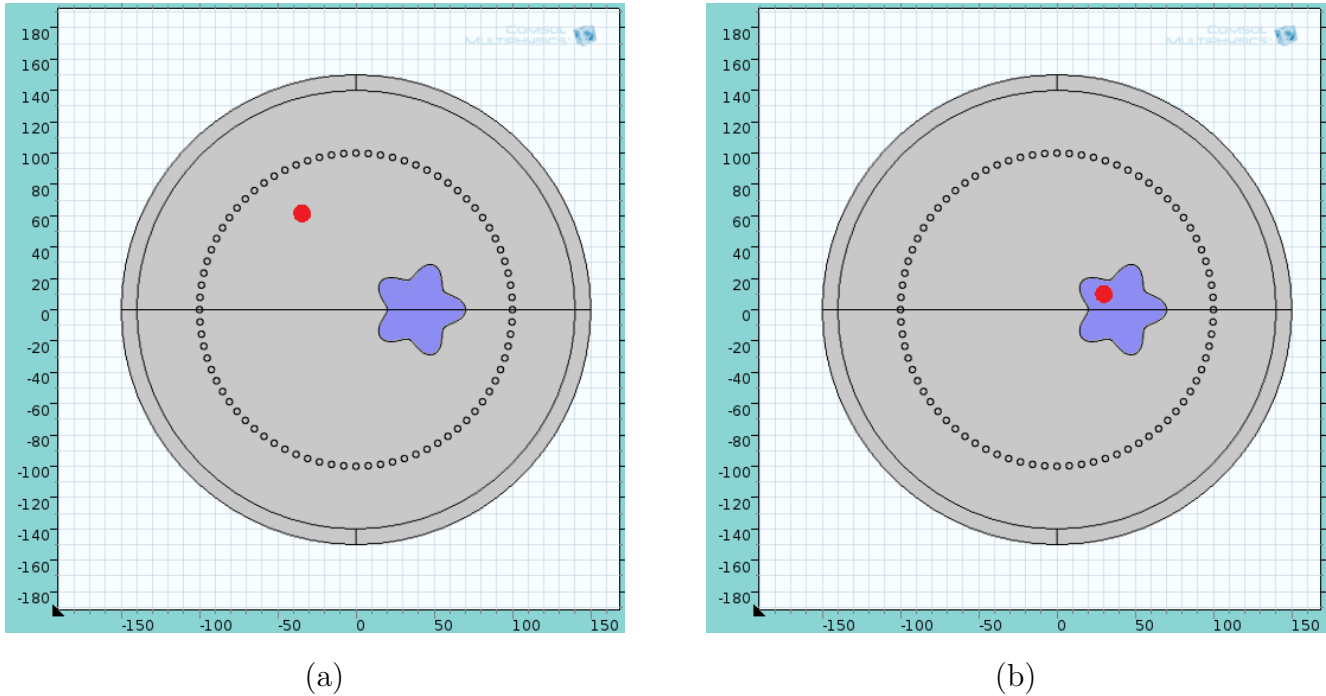


Figure 4.11: Model in 2D with a rugged fibroglandular structure depicted in blue and a scatterer in red: (a) the scatterer is outside the fibroglandular tissue (b) the scatterer is inside the fibroglandular tissue. The small surrounding circles are the line current out of plane. The skin layer has been removed. [Dimensions are in mm]

UWCEM) where a number of anatomically-realistic MRI-derived numerical breast phantoms can be retrieved. The breast tissues in these phantoms have the realistic ultra wide-band dielectric properties reported in [37] and used for simulation in the microwave range.

4.6.1 Dataset definition

The breast phantoms are derived from series of T1-weighted magnetic resonance images of patients in a prone position. Each phantom is comprised of a 3D grid of cubic voxels, where each voxel is $0.5 \text{ mm} \times 0.5 \text{ mm} \times 0.5 \text{ mm}$. The breast model includes a 1.5-mm-thick skin layer, a 1.5-cm-thick subcutaneous fat layer at the base of the breast, and a 0.5-cm-thick muscle chest wall. At least one phantom for each category breast density (see § 1.4) is provided and 3 text files are provided to describe the breast properties.

breastInfo.txt This files contains the number of the patient under examination, the geometrical dimension of the breast and the density assessment reported in the 'classification' row:

```
breast ID=012204
s1=316
s2=352
s3=307
classification=2
```

mtype.txt This file gives tissue type information for each voxel in the grid. The file has only a single column of data and each row represents a different voxel in the grid. The data file was generated using the following ordering of grid cell:

```
for k=1:s3
  for j=1:s2
    for i=1:s1
      .....
    end
  end
end
end
```

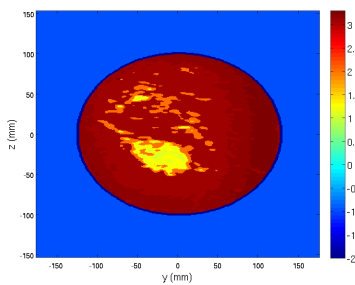
So the total number of voxels can be computed as $(s1 \times s2 \times s3)$.

Each row contains one of the media numbers given in Table 1. The coupling media number is assigned to all voxels outside the breast. The muscle media number is assigned to all voxels in the chest wall region. Media numbers associated with normal breast tissue are assigned to all voxels within the breast interior and the subcutaneous fat layer at the base of the breast. The normal breast tissue in the breast phantom is divided into seven tissue types, ranging from the highest-water-content fibroglandular tissue with the highest dielectric properties (media number = 1.1), to the lowest-water-content fatty tissue with the lowest dielectric properties (media number = 3.3). There is also a transitional region (media number = 2) with intermediate dielectric properties. Figure 4.12 shows the tissue type distribution for coronal, sagittal and transverse planes with the origin based on the center of the volume.

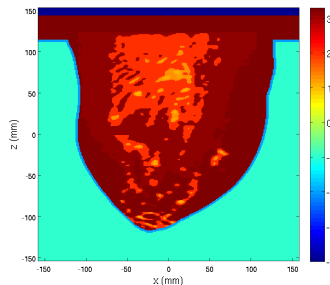
pval.txt This file is a single column file of the same length as mtype.txt where a value between 0 and 1 has been recorded. This value represents the fraction between the 2 lines in which the tissue type is placed. It can easily be seen as a weighted sum between the two values that define the range in which a tissue type is settled. As an example, suppose $pval = 0.12$ for a fibroglandular-2 tissue labeled as $mtype = 1.2$. The dielectric properties values associated fall in the region bounded from above by the group1-high red-dotted curve and from below by the group1-median black curve as can be seen in figure 4.13. The corresponding Debye parameters for the voxel are computed using the values reported in Table 4.3

Table 4.2: Media number for the breast tissues in MRI-derived numerical phantom

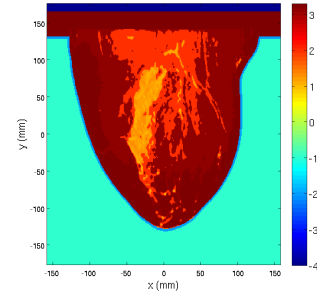
Tissue	Media number
Coupling medium	-1
Skin	-2
Muscle	-4
Fibroglandular-1	1.1
Fibroglandular-2	1.2
Fibroglandular-3	1.3
Transitional	2
Fatty-1	3.1
Fatty-2	3.2
Fatty-3	3.3



(a) Slice YZ



(b) Slice XZ



(c) Slice XY

Figure 4.12: Tissue distribution slice for a patient with classification ACR 2 extracted from an mtype.txt file.

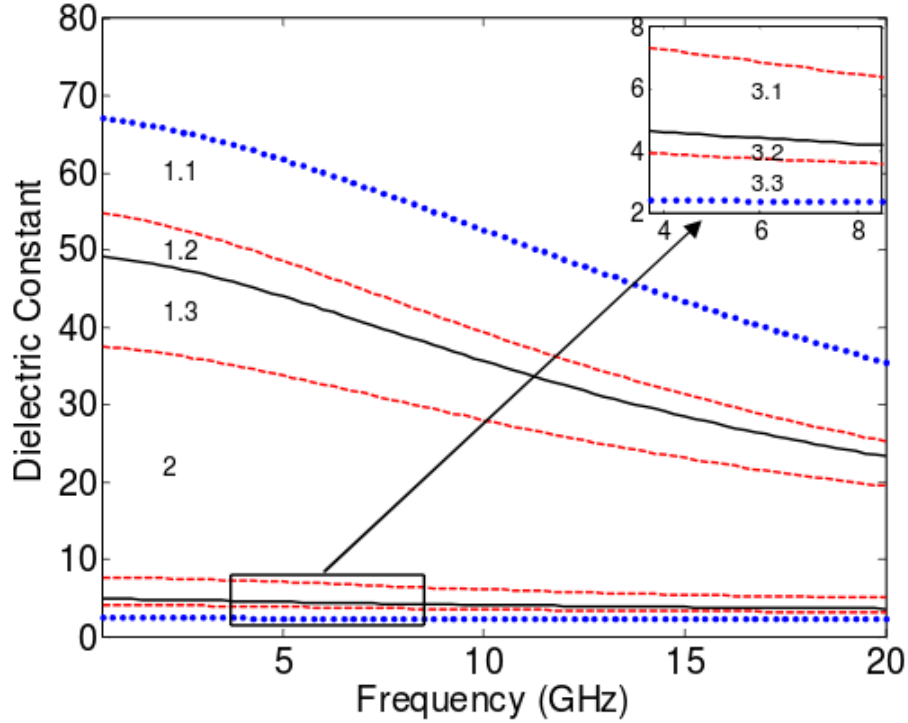


Figure 4.13: The dielectric properties of normal breast tissue. The seven tissue-type regions bounded by the curves in this graph are labeled with the media numbers reported in Table 4.3. From top to bottom, the curves correspond to the following eight cases: maximum, group1-high, group1-median, group1-low, group3-high, group3-median, group3-low, and minimum. (<http://uwcem.ece.wisc.edu/MRI/database/>)

$$\text{param} = pval * \text{upper_value} + (1 - pval) * \text{lower_value} \quad (4.7)$$

In the case we want to estimate ϵ_{∞}

$$\epsilon_{\infty} = 0.12 * 14.20 + (1 - 0.12) * 13.81 = 13.86 \quad (4.8)$$

Figure 4.14 shows the $pval$ distribution for coronal, sagittal and transverse plane with the origin based on the center of the volume.

4.6.2 Adding the real phantom to COMSOL

With the aid of matlab and using the three files above described, one file for any illumination frequency used in the simulation has been created, containing the relative permittivity and the conductivity for each voxel thus to mimic the exact behavior of the breast placed under illumination. COMSOL is non able to accept this kind of file as produced, in fact the standard procedure

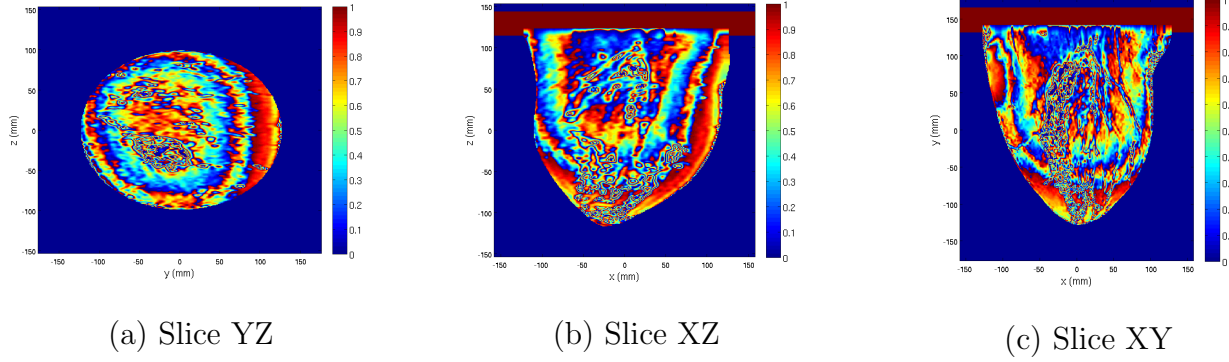


Figure 4.14: Pval slice distribution for a patient with classification ACR 2

Table 4.3: Single-pole Debye parameters from all the tissues in the mtype.txt file. Valid in the range 3GHz - 10GHz

	ϵ_{∞}	$\Delta\epsilon$	τ (ps)	σ (S/m)
minimum	2.309	0.092	13	0.005
group3-low	2.848	1.104	13	0.005
group3-median	3.116	1.592	13	0.050
group3-high	3.987	3.545	13	0.080
group1-low	12.99	24.40	13	0.397
group1-median	13.81	35.55	13	0.738
group-high	14.20	40.49	13	0.824
maximum	23.20	46.05	13	1.306
skin	15.93	23.83	13	0.831
muscle	21.66	33.24	13	0.886

to define the material properties of objects is to add a geometry, let COMSOL to define domains, boundaries and edges and then assign to those the desired materials. This cannot be done with a file which:

- (a) is defined in voxel coordinates;
- (b) has an enormous number of elements whose material has to be set

COMSOL, indeed, makes available the use of interpolation functions to overcome this limitation when dealing with materials whose properties change at very small scales.

The interpolation function can be activated providing a text spreadsheet file either in 2 or in 3 dimensions. In our case the file provided is even too much detailed, this situation prevents a fast computational time and to overcome it a down-sample of the files containing permittivity and conductivity distributions is made.

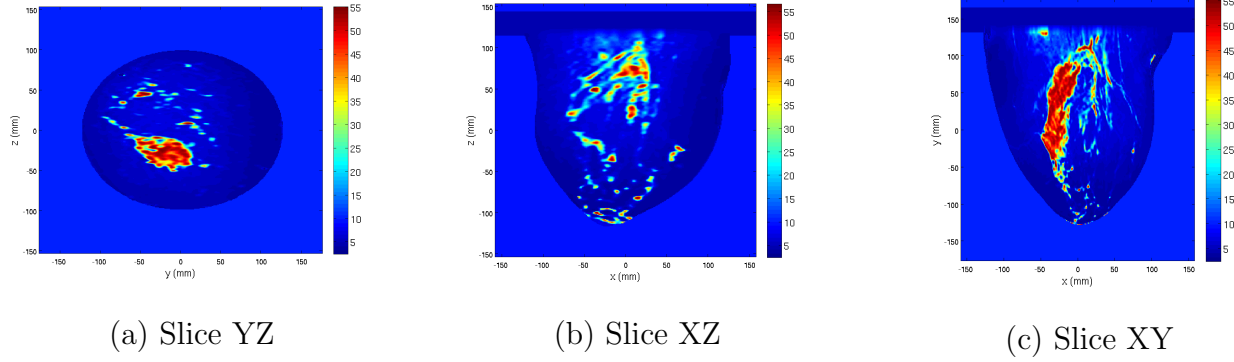


Figure 4.15: Epsilon slice for a patient with classification ACR 2 for an illumination frequency of 3 GHz

At this point a suitable geometry is needed to assign this material whose electrical properties are defined in term of two interpolation functions. The main idea is to perform a segmentation of the skin and use the pectoralis muscles line to close the surface, then create a stereo lithography (STL) file that COMSOL should import. STL files are text files that represent solid objects whose surface have been discretized by triangles and contain the 3 spatial coordinates for each of the 3 vertices and the information about the normal to the defined surface. Unfortunately the STL file generated in such way cannot be imported inside COMSOL even after some smoothing activities using standard software available on the Internet. Several issues happens, the most common is that after a correct import procedure the simulator can not finalize the geometry either as form union or assembly. More uncommon, but very tedious, is that during the angular spanning some antenna positions fail to converge. In figure 4.17 (a) the segmentation procedure to build the breast and visualized with Meshlab, (b) a large number of sharp surfaces are clearly visible after the import stage inside COMSOL.

Scatterer addition

All the MRI phantoms provided by UWCEM repository refer only to healthy breasts. To investigate the backscattered signal of unhealthy tissues a tumor has to be added and this is performed for any illumination frequency according to the Debye's formula.

The corresponding Debye's parameters are the same used in the Table 4.2. The 3-D shape chosen for the tumor is a sphere whose radius is $r_{\text{tumor}} = 5 \text{ mm}$.

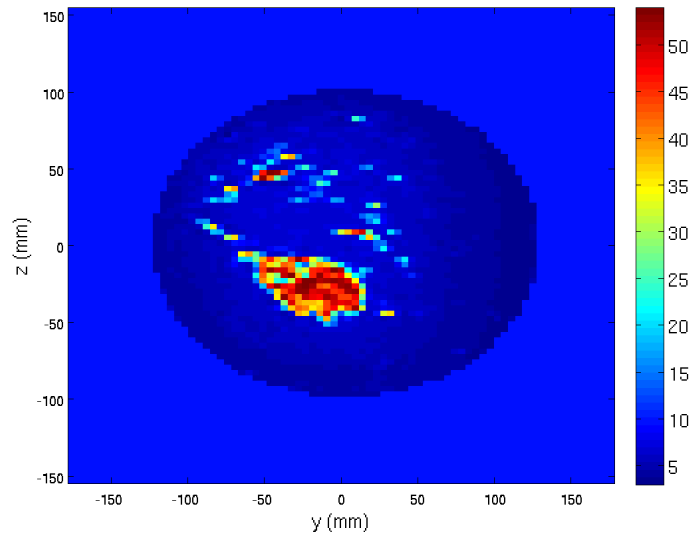


Figure 4.16: Relative permittivity slice down-sampled 4 times for a patient with classification ACR 2 using an illumination frequency of 3 GHz.

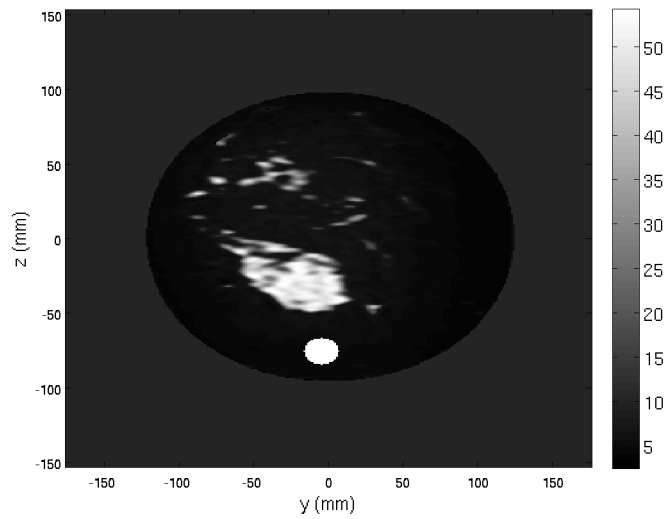


Figure 4.18: A coronal slice of the permittivity with a $r=5$ mm tumor added

This procedure is performed outside COMSOL so that the spatial distribution of the electrical parameters are imported when the tumor is already present. In figure 4.18 is shown a coronal slice of the permittivity taken approximately in the center of the breast.

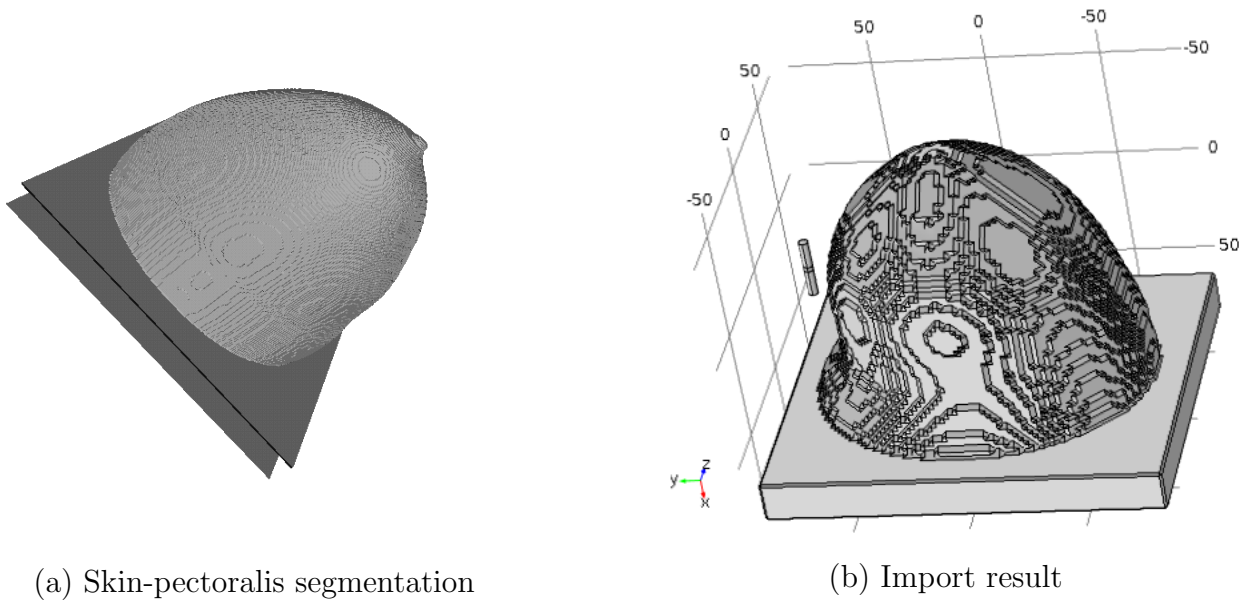


Figure 4.17: Geometry of a real breast import in COMSOL (b) after skin-pectoralis segmentation and some smoothing procedures (a)

4.6.3 Conformal cup

In order to activate a systematic study where the simulation parameters are modified we must ensure that convergence is reached when a new simulation is run. As told before there are some situations that stop the simulation with errors.

Moreover when moving an antenna around a real breast the distance between the antenna and the skin layer is not fixed and this is source of cluttering when performing the reconstruction. To overcome this situation a conformal cup can be considered as breast holder.

The cup has been designed using the information provided in [30] where an accurate study for the anatomical characterization of the breast is given by means of a number of parameters which describe breast shape, diameter and length.

Patients subject to computed tomography (CT) have been divided in four categories across four bra increasing cup sizes from 'A' to 'D'. The acquired data shows that the breast shape can be generally characterized by the breast effective diameter, which gradually decreases from posterior to anterior.

We referred to cup size 'C' whose profile is well described using third-order polynomial, below the values of the parameters:

$$a = -20.452, b = +24.397, c = -13.686, d = +13.585 \quad (4.9)$$

These parameters are referred to a profile normalized to 1 using as normalization factor the

mean value of the breast sagittal length $z_0 = 9 \text{ cm}$, so the function describing the cup becomes;

$$f(x) = \frac{1}{2} \cdot \left[a \cdot \left(\frac{x}{z_0} \right)^3 + b \cdot \left(\frac{x}{z_0} \right)^2 + c \cdot \left(\frac{x}{z_0} \right) + d \right] \quad (4.10)$$

Factor $\frac{1}{2}$ is due because we design half profile, the other is obtained by symmetry.

Now we have to find an analytical expression for a junction, this function has to have the vertical derivative with the x-axis intersection because we will obtain the full 3D cup by revolving along the x-axis. So that the choice falls upon an arc of a parabolic curve with the symmetry axis parallel to the x-axis:

$$x = a_1 y^2 + b_1 y + c_1 \quad (4.11)$$

Vertical derivative requires:

$$y_v = -\frac{b_1}{2a_1} = 0 \quad \Rightarrow \quad b_1 = 0 \quad (4.12)$$

The two unknown parameters a_1 and c_1 are computed by imposing the continuity of the function and the first derivative calculated in the junction point z_0 .

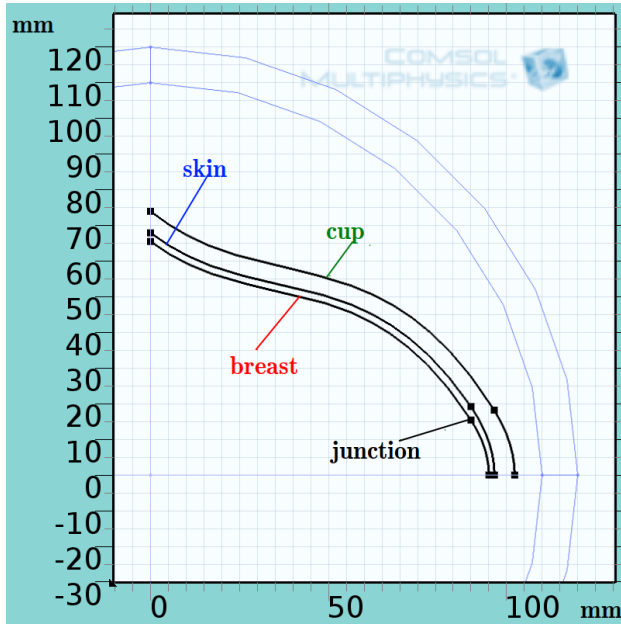
This definition can not be applied directly to the phantom model: the cup does not fit perfectly the COMSOL interpolated function that defines the material electric properties.

To overcome this limitation the skin is replaced inside the `mtype.txt` file and substituted with the coupling medium type, then a new 3rd-order polynomial function is used to define the skin line. Moreover the pectoralis muscles are not included in the breast line for this reason they are replaced into the `mtype.txt` by the coupling medium tissue value too. Muscles are modeled inside the simulation software. Finally another line is used for the cup definition using a best fit procedure to keep fixed its thickness. The 2D work-plane and the resulting full 3D model with a permittivity slice interpolation function added is shown in figure 4.19

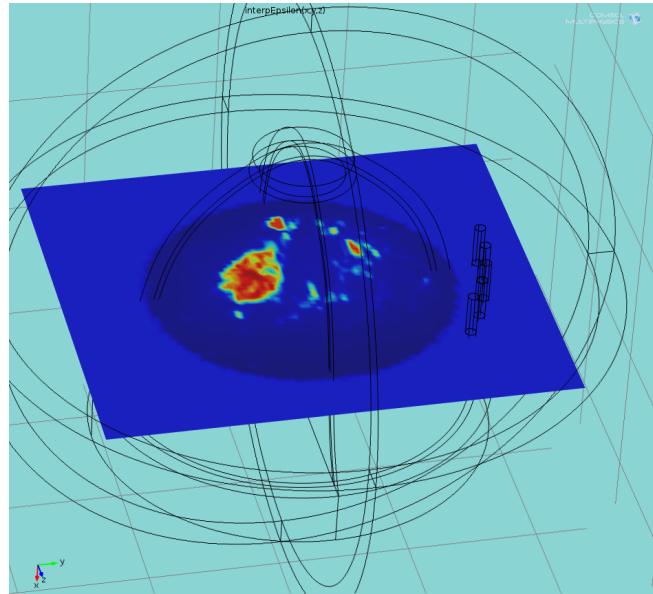
The whole procedure to define a real breast inside a containment cup model is summarized in figure 4.20.

4.7 Performance evaluation and scalability

In order to complete the described simulations two workstations have been employed, the less powerful is equipped with 32 GByte of RAM and an Intel Core I7-3770 at 3.40 GHz CPU with 4 cores, the more powerful has 64 GByte of ECC-RAM and can be expanded up to 256 GByte, the CPU is an Intel Xeon E5-2420V2 at 2.20 GHz with six cores but the mother board is dual CPU capable so another one can be hosted. Both workstations are equipped with NVidia GPU graphic



(a) 2D Workplane



(b) Revolved geometry

Figure 4.19: Breast MRI full model: (a) 2D work plane to define the breast internal area, skin depth and containment cup (b) wire-frame rendering of the resulting 3D simulated model along with a permittivity slice.

cards to speed-up the 3D rendering.

At the moment COMSOL does not include the GPUs support so it is not possible to benefit of their acceleration on calculus; this mainly due because the GPUs performances are not stable, so changing the physics, boundary conditions, geometry or whatever else can produce completely different results. Moreover the direct solvers used inside COMSOL are very useful to handle strongly coupled smaller size multiphysics problems, but they also require a lot of memory. Currently, dedicated general purpose GPUs have about 5-16 GB of internal memory. This corresponds to small problem sizes for the direct solvers, in the range where they do not yield a high scalability on GPUs.

As alternative COMSOL makes available the clustering option but the requirements for the Windows environment (costly Windows server licensing) or the skill needed in a Linux Red Hat/Debian/Ubuntu environment along with the suggested, but not mandatory, similar hardware characteristics has discouraged this implementation.

The main idea to overcome this limitation has been to implement a kind of load-balancing between the two workstations. The main simulation involved a parametric study where many values of the model are changed such as the tumor position inside the breast, the contrast factor, the tumor size. A computer program capable of producing a configuration file list, where each row

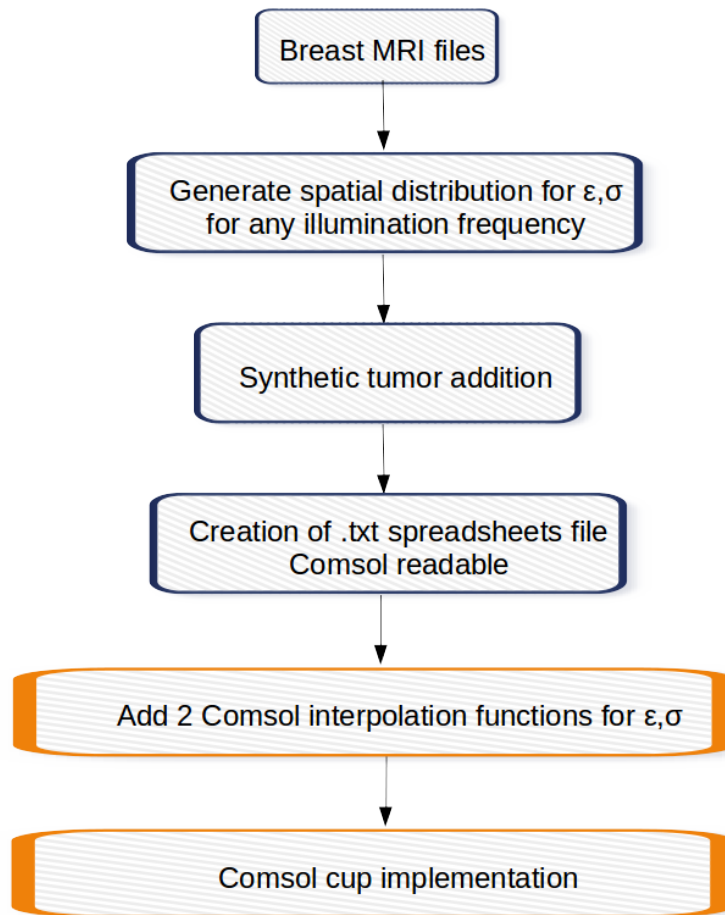


Figure 4.20: Schematic COMSOL real 3D unhealthy-breast model generation procedure

represents a configuration to be run, has been realized. The COMSOL model has been placed in a network share as well as the configurations file list and a MATLAB "looper" script is used to set the correct parameters inside the simulator through the LiveLink connection. To overcome the possible run of the same configuration a kind of locking method is used: any untouched configuration is labeled as 'TODO', when it is being working the label changes to 'LOCKED' and finally, after the results have been written, the status label is changed to 'DONE'. If errors occur the label is changed in 'COMSOL ERROR' and makes possible further investigations or re-run.

At the highest frequencies many 3D configurations exceed the limit of 32 GByte, in this case either a slow swapping on the local hard disk arise or the workstation entirely hangs. The looper accepts as input parameter the frequency range suitable for the workstation leaving to the other the "dirty" job.

Table 4.4: COMSOL 2D and 3D model performance comparison on Xeon Intel Xeon E5-2420V2 with 64 GByte of RAM. The operating frequency (ν), the running time, the memory needed and degree of freedom (DOF) solved for are also reported.

Model	Dim	ν	Time	Memory	DOF
Vivaldi	3D	1 GHz	4.30 min	6 GB	2.1 M
Vivaldi	3D	2 GHz	7.20 min	14 GB	2.8 M
Vivaldi	3D	3 GHz	23 min	40 GB	6.9 M
Dipole	3D	2 GHz	2 min		
5-points	2D	5 GHz	11 sec	1.6 GB	0.5 M
5-points	2D	10 GHz	22 sec	2.5 GB	1.0 M

4.7.1 Running time

The time needed to finish a single run varies with the operating frequency but is quite similar for both workstations. The single run corresponds to set an operating frequency, fixing the antenna position respect to the breast and wait COMSOL to calculate the solution for the \mathbf{E} field in the specified domain. The running time for the 3D homogeneous breast model illuminated by a dipole antenna working at 1 GHz is roughly 4 minutes and needs about 12 GByte of RAM. The same model run at 4 GHz is considerably slower and memory consuming: it takes 35 minutes and 50 GBytes of RAM are mandatory.

This 3D model could run at higher frequencies only after a rationale build of the domain and of the two cylinders mimic the breast and the tumor. To ensure that no diffraction effects arise, these cylinders are kept always at least four wavelength long but when the operating frequency increases the wavelength decreases accordingly. We can exploit this fact reducing the cylinders' heights and the coupling medium domain radius and make their geometrical dimensions dependent on the frequency: in these conditions the simulation can reach the limit 5 GHz where all the 64 GByte are allocated.

The model built to evaluate the feasibility to reconstruct a fibroglandular island needs an high number of field points and an high illumination frequency. The available hardware is not powerful enough to run such model and a 2D approach has therefore to be followed. Figure 4.4 shows elapsed time, memory requirements and number of degrees of freedom solved for a some 2D and 3D simulated models at some operating frequencies.

Chapter 5

Reconstruction of simulated data

In this chapter the simulated results are presented for the homogeneous and inhomogeneous breast. A comparison between monostatic and multistatic architecture based on homogeneous scenario is made. Denoising procedure and metrics to evaluate the reconstructed images used are detailed.

5.1 Cluttering removal and filtering

The scattered signal coming from the tumor is generally weaker than the antenna internal reflections and reflections coming from the skin layer and other internal tissues. Performing the reconstruction without a proper filtering procedure leads to inadequate results because the portion of clutter signal is the predominant in the whole signal.

Several clutter-rejection procedures have been proposed in literature: when working in time-domain a time-gating approach is followed to eliminate the first back-scattered signal because it is mainly due to the skin reflections. Other works rely on filtering methods that try to estimate the clutter signal at a given position and subtract it from the actual measurements. Alternatively, some procedures employ a differential scheme where two sets of measurements corresponding to two different rotations of the system are subtracted from each other but this is strongly dependent from the breast geometry and moreover it might subtract useful back-scattered signal coming from the tumor. When working in the frequency domain the pole splitting algorithm is applied to the radar signal [42] to remove the skin contribution and a subspace-based algorithm exploiting the singular value decomposition (SVD) has been used to remove generic artifacts [18].

5.1.1 Monostatic filtering

In our configuration we exploit the symmetry coming from the adopted central geometry in an homogeneous medium: if no tumor is added S_{11} is a constant complex value while, when a tumor id

added, a small perturbation in the signal appears. In the monostatic case the filtering procedure simply consists in subtracting the mean value of the signal collected at the N positions: this procedure removes the constant background signal and points out the tumor

$$S_{\text{clean}} = S_{\text{raw}} - \langle S \rangle \quad (5.1)$$

In figure 5.1 is shown the results of this operation: the difference enhances the small perturbation caused by presence of the scatterer at 180 deg. This 'clean' signal is later used to perform the reconstruction.

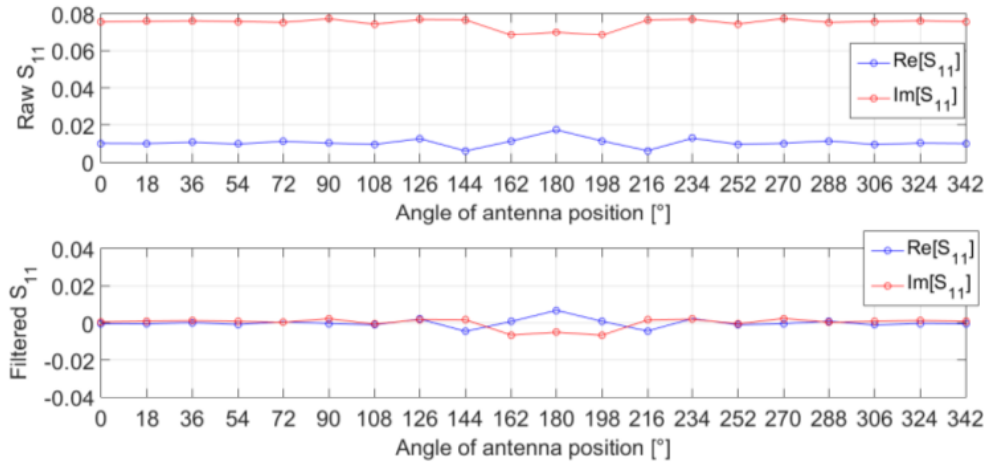


Figure 5.1: On top S_{11} raw collected signal, and on bottom same signal filtered with average trace subtraction. Blue line is the real part of the signal and red line is the imaginary part.

5.1.2 Multistatic filtering

The multistatic configuration can not permit a simple subtraction of the mean value of the S -matrix. In this case we have to filter out the unwanted components of skin reflection and antennas internal reflection such as in the monostatic case together with the mutual couplings typical to this configuration. This coupling is due to the direct incident field generated by the actual transmitting antenna towards the receiving ones, this field excites the standing antennas that might re-irradiate adding much more cluttering.

To find a suitable filter we make use again of the central symmetric geometry of the whole system. With the aid of figure 5.2 where 8 antennas are used for explanatory purpose, we see that the previous filter is equal to compute the average value on the main diagonal, that is the reflection coefficients S_{nn} , and subtract the value from it.

Now consider the S_{12} coefficient: its value should be the same for each first neighbors as it is clear if we think the value collected at each port as the "optical round trip path" of the signal,

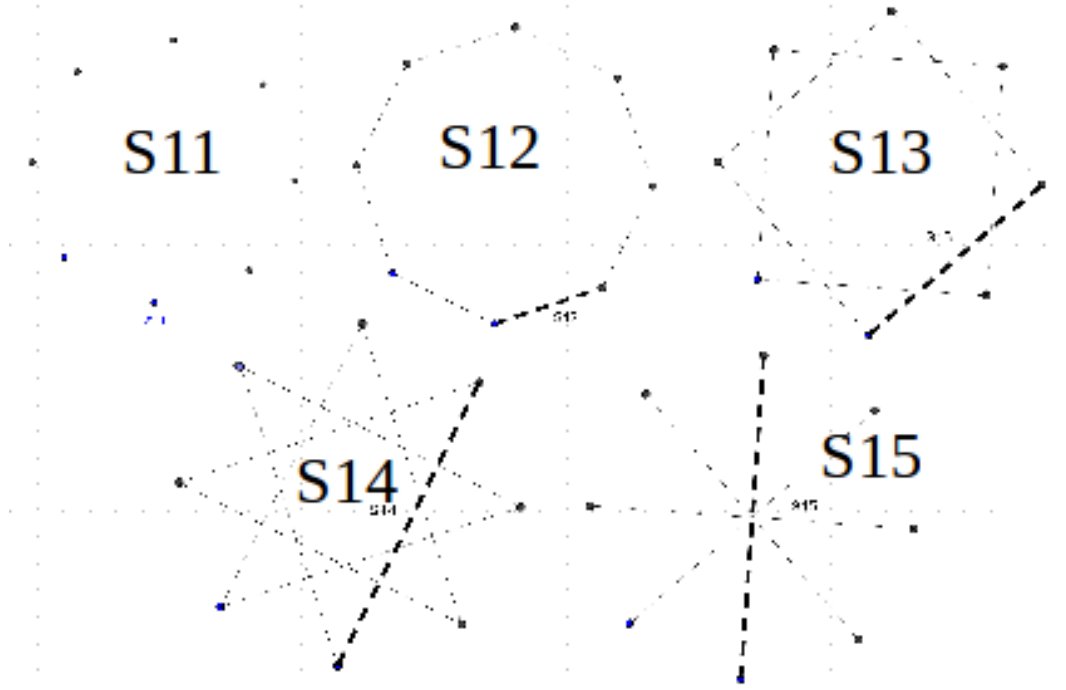


Figure 5.2: Independent elements in a multistatic central symmetric geometry

that is:

$$S_{12} = S_{23} = S_{34} = S_{45} = S_{56} = S_{67} = S_{78} \quad (5.2)$$

these equalities hold true also for the reversed order of the index, this guarantees the symmetry of the resulting S-matrix. The reasoning can be applied for all the remaining S-coefficients: figure 5.2 shows all the combinations for the 8 antennas system and figure 5.3 shows the form of resulting S-matrix. As can easily be seen **only** 5 independent elements appear in the form of a Toeplitz-symmetric matrix, that is a matrix whose diagonal elements are constant. The relation between an N-th order matrix and the number of the independent elements N_{ind} is given by:

$$\begin{cases} \text{for } N_{\text{even}} : & N_{\text{ind}} = \frac{N}{2} + 1, \\ \text{for } N_{\text{odd}} : & N_{\text{ind}} = \frac{N + 1}{2} \end{cases}$$

To estimate the equivalent of the $\langle S \rangle$ in the multistatic environment we make use of a linear operator, we called *diagonal-mean*: a Toeplitz-symmetric matrix whose elements are the average-value of the diagonal in which they stay

$$\langle S \rangle \simeq \text{diagonalMean}(S) \quad (5.4)$$

The effect produced by this filtering can be seen in 5.4, the cleaned matrix enhance the presence



Figure 5.3: Independent elements in a multistatic central symmetric geometry with 8 antennas: each color represent a different element

of a scatterer. It is noteworthy that, examining the scale value reported on the right of the figures, the filtered signal is two orders of magnitude greater than the raw collected one.

5.2 Evaluation of the reconstruction and metrics

MUSIC algorithm identifies the target position in the point corresponding to the maximum of the pseudospectrum and the reconstructed target should be represented by a Dirac delta function of infinite value, but in real reconstruction its form is a spread curve with a finite maximum.

In order to evaluate the performance achievable by the detection system different metrics are employed:

- Full-Width-Half-Maximum (FWHM);
- Spatial displacement (SD);
- Signal-to-Noise Ratio (SNR).

5.2.1 Spatial metrics

In many cases the reconstructed target shape is not symmetrical with respect to its maximum, thus we choose the centroid of the curve at Half Maximum (HM) as estimation of target position. The computation of the spatial metrics is performed in the maximum variance configuration obtained through the PCA (Principal Component Analysis) method.

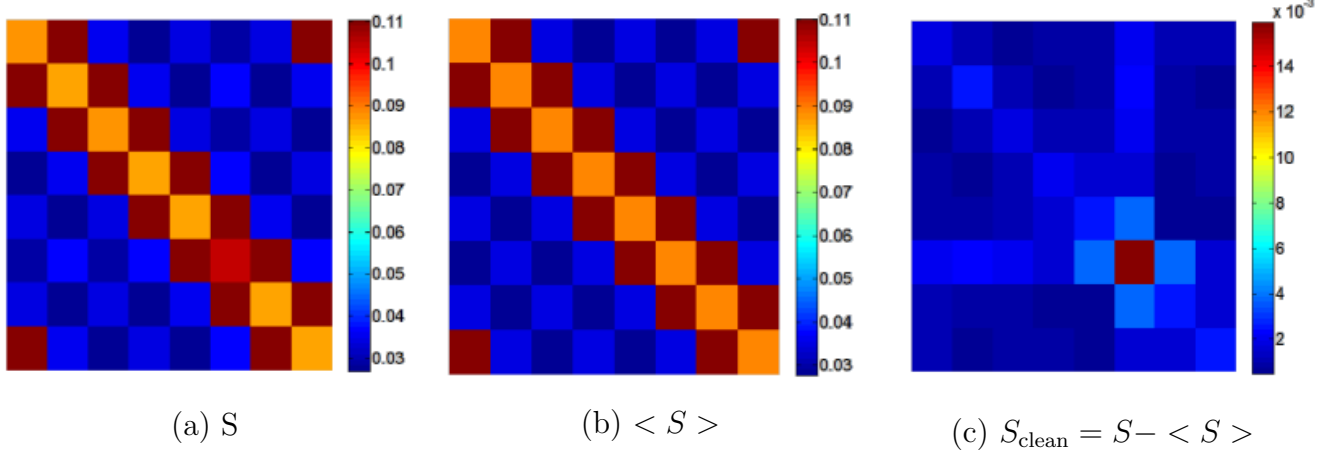


Figure 5.4: Multistatic Filtering procedure: (a) the raw collected S-matrix, (b) the filter matrix computed using the diagonalMean linear operator, (c) the clean signal used to perform the reconstruction

PCA is a statistical method mainly used to reduce the dimensionality of a problem retaining as much information as desired based on the variance [33]. Geometrically speaking is an orthogonal transformation based upon the computation of the eigenvalues and eigenvectors to change the basis such that the data are better defined and their structure can be highlighted.

In figure 5.5 is detailed how the spatial metrics are derived: the $X(x, y)$ point is the exact location of the target, while $C(x, y)$ represent the centroid of the spread curve cut at HM. The target spread expresses the degree of resolution reached and is the defined as:

$$\text{FWHM}_{\rho, \phi} = 2 \cdot \beta \sqrt{\lambda_{1,2}} \quad (5.5)$$

where $\lambda_{1,2}$ are the eigenvalue of the PCA transformation. For a confidence interval of 95%, $\beta = 1.96$.

The SD measures the error in tumor localization and accounts for the difference between the real position $X(x, y)$ of the tumor and the centroid $C(x, y)$ in the new coordinate system defined by PCA:

$$\text{SD}_{\rho, \phi} = X_{\rho, \phi} - C_{\rho, \phi} \quad (5.6)$$

5.2.2 Contrast metric

The standard deviation σ can be expressed as usual in term of the FWHM:

$$\sigma_{\rho, \phi} = \frac{\text{FWHM}_{\rho, \phi}}{2\sqrt{2 \ln(2)}} \quad (5.7)$$

by means of σ we define the useful area given by the signal and the background signal given

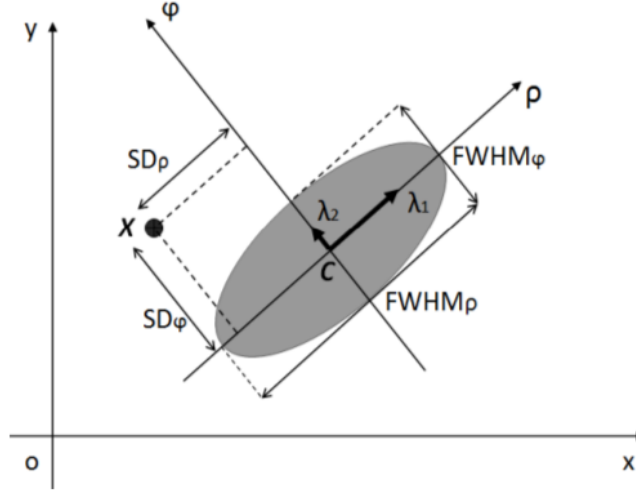


Figure 5.5: Spatial metrics definition using PCA

by the clutter. As reported in figure 5.6 the signal area is:

$$A_{\text{signal}}(\rho, \phi) \leq 3\sigma_{\rho, \phi} \quad (5.8)$$

while the clutter area is defined as:

$$3\sigma_{\rho, \phi} \leq A_{\text{noise}}(\rho, \phi) \leq 6\sigma_{\rho, \phi} \quad (5.9)$$

and the SNR is computed as the mean value of the pseudospectrum evaluated in the signal area over the mean value in the clutter area:

$$\text{SNR} = \frac{\frac{1}{N_s} \sum_{\mathbf{r}_i \in \text{signal}} \Psi(\mathbf{r}_i)}{\frac{1}{N_c} \sum_{\mathbf{r}_i \in \text{noise}} \Psi(\mathbf{r}_i)} \quad (5.10)$$

where N_s and N_c are respectively the number of points in the signal area and in the noise area.

As the spread occurs mainly in the radial direction ρ , in the following the ϕ -metrics will be neglected when displaying results.

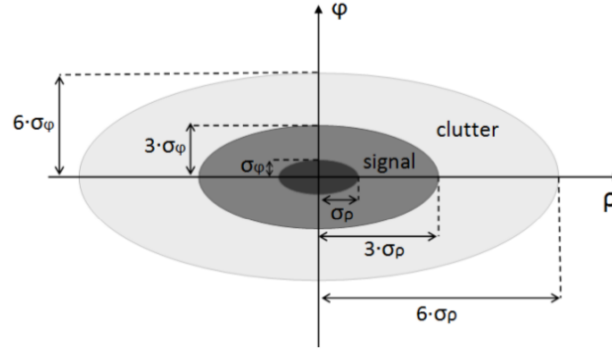


Figure 5.6: Spatial metrics definition using PCA

5.3 Homogeneous breast model in monostatic configuration

The breast model used in the simulation has been already detailed in the previous chapter and the operating conditions are summarized in table 4.1. In the following the reconstruction results are discussed changing some parameters that can affect the reconstruction in a real environment. Figure 5.7 shows the geometry model realized with COMSOL.

5.3.1 Working Frequency

We aim to identify and locate the scatterer varying the illumination frequency, to evaluate the reconstructions and finally to select the best frequency range. The investigated range is from $\nu = 1$ GHz to $\nu = 5$ GHz. The reconstructed pseudospectra, as a function of the varying frequency, are shown in figure 5.8.

Looking at the image series in this figure, it can be observed that as frequency increases the reconstructed spot size decreases, this effect is due to the improvement of the spatial resolution. However, at high frequencies some artifacts appear and we expect a corresponding worse SNR value, defined as the ratio between mean value in signal area and mean value in clutter area.

The metric values reported table 5.1 show an increasing resolution from lower frequencies to higher ones, in particular, in the range from 1 GHz to 2 GHz, the spread determined from the FWHM is almost halved, from 2 GHz to 5 GHz it decreases slowly.

Looking at SNR values, in the range [2, 4] GHz we have the higher ratio reaching the maximum value at 3 GHz; over 4 GHz clutter value becomes relevant. Finally we reported the Spatial Displacement (SD) metrics, that is error that the algorithm makes to reconstruct the exact location of the target and computed as the difference between the simulated position, chosen as the center of the scatterer, and the reconstructed one given by the imaging point with pseudospectrum highest

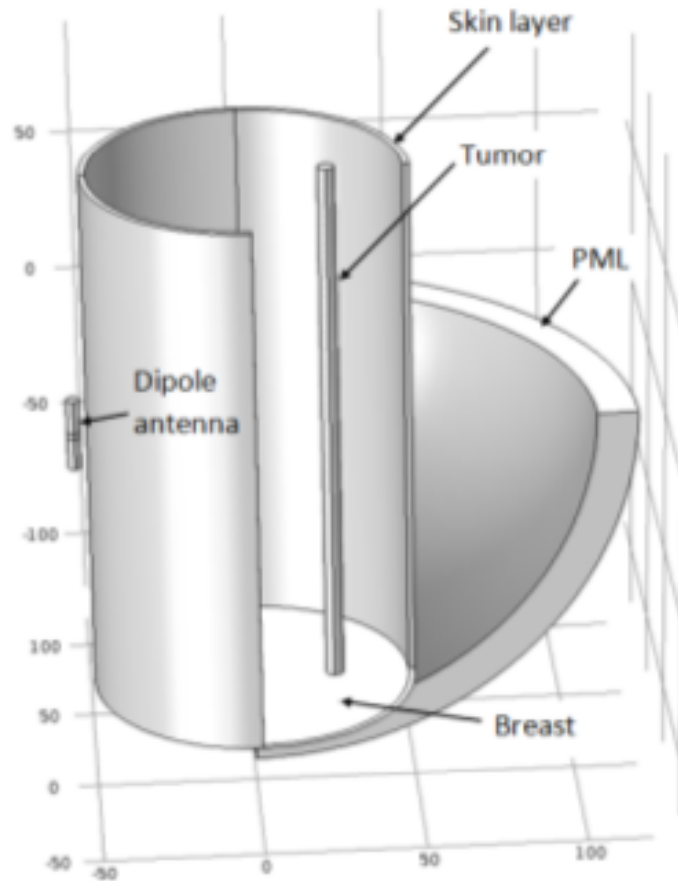


Figure 5.7: COMSOL model for the monostatic configuration. Dipole antenna, cylindrical breast and tumor, skin layer and PML are detailed

value. SD is always much lower than the resolution at any frequency.

Figures 5.9 and 5.10 show, respectively, the FWHM_ρ and SNR trend when the illumination frequency is changed.

5.3.2 Multifrequency

The single-frequency approach suffers from the trade-off between high resolution and artifact limitations. The reconstruction made at different frequencies can be combined to outperform the results of the single-frequency case. In this study we use the GM combination method whose definition has been given in 3.38.

Figure 5.11 show the resulting image from the combination in the range [2-4] GHz and table shows the corresponding metrics. Both FWHM and SNR reach better values than the corresponding single-frequency reconstruction. The GM-MUSIC metric errors are defined as the standard errors computed on ten measures after we have perturbed the system with white Gaussian noise

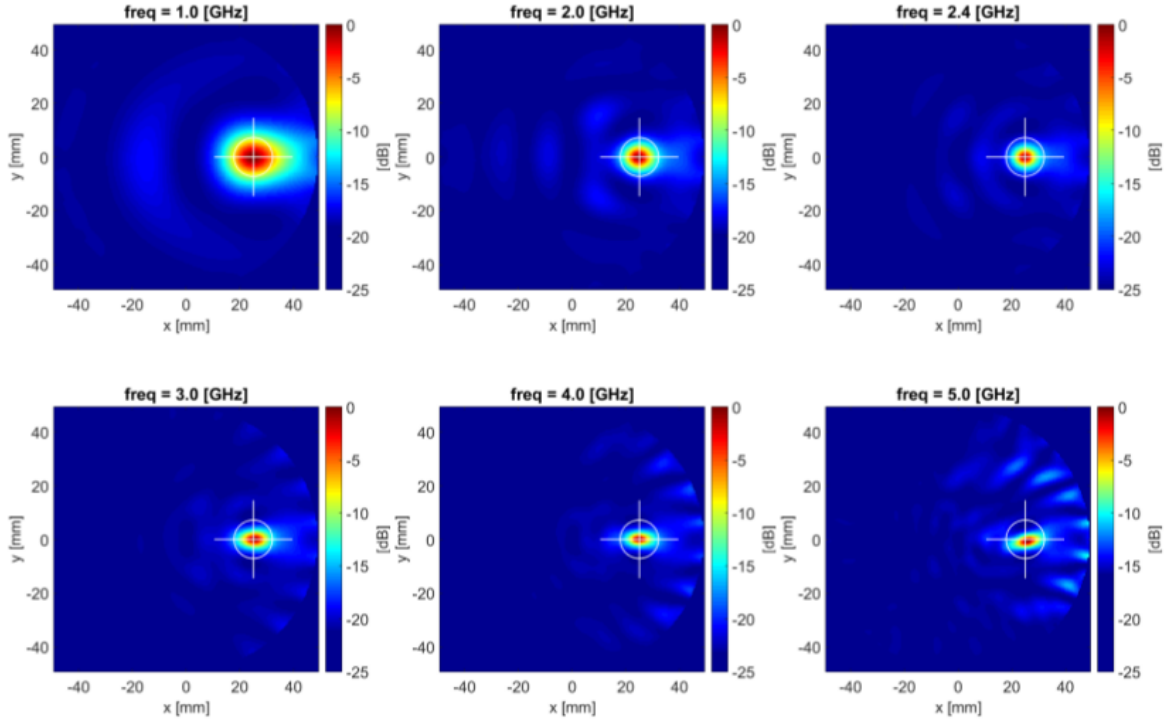


Figure 5.8: Pseudospectra frequency evolution from 1 GHz to 5 GHz

with SNR=10 dB. In the following study all results will be therefore given using this combination method.

5.3.3 Tumor position

The objective is to identify and locate the target varying its position. The study is performed investigating three positions:

near center: $X = (10, 0)$ mm;

half radius: $X = (25, 0)$ mm;

Table 5.1: Spatial and contrast metrics varying the illumination frequency in the range 1 GHz - 5 GHz

frequency	FWHM_ρ [mm]		SD_ρ [mm]		SNR [dB]	
1 GHz	18.2	± 1.2	1.4	± 0.7	4.3	± 0.6
2 GHz	12.9	± 0.5	1.0	± 0.3	4.6	± 0.5
3 GHz	10.1	± 0.3	0.7	± 0.2	6.5	± 0.2
4 GHz	10.0	± 0.6	0.5	± 0.3	5.6	± 0.2
5 GHz	9.8	± 0.3	1.1	± 0.1	3.6	± 0.3

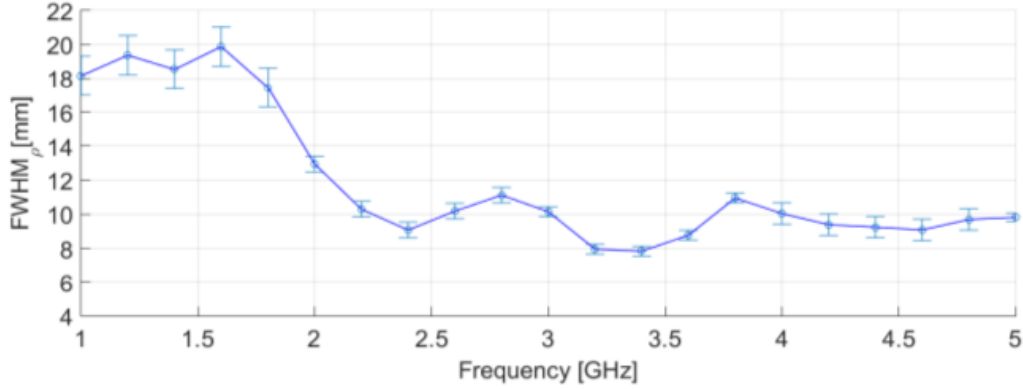


Figure 5.9: FWHM_ρ versus the illumination frequency from 1 GHz to 5 GHz with 0.2 GHz step.

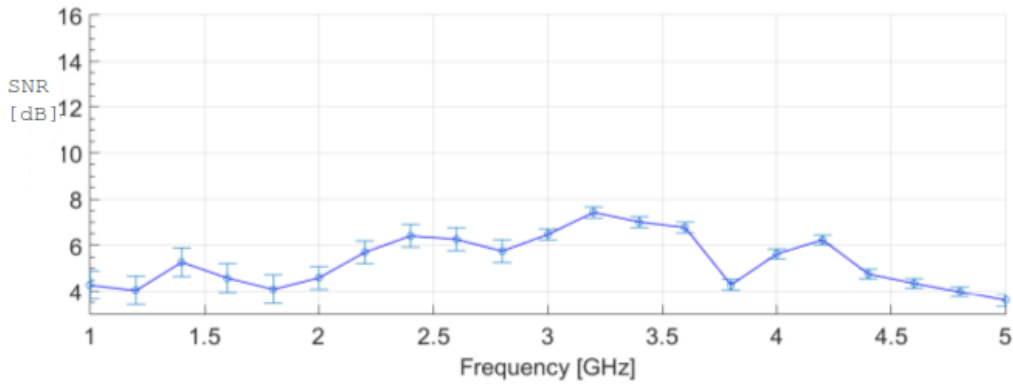


Figure 5.10: SNR_ρ versus the illumination frequency from 1 GHz to 5 GHz with 0.2 GHz step.

near skin: $X = (40, 0)$ mm.

The simulation is performed using the frequency in the optimal range $\nu = [2-4]$ GHz and the corresponding pseudospectrum is evaluated. Results are given using GM combination method. Figure 5.12 displays the images and table 5.3 shows the resulting metrics.

Metrics exhibit a performance degradation gradually moving from the inner to the outer position of the breast, this happens because the measure is affected by skin reflections near the boundary. The best position of the scatterer is near the center of the breast because a similar contribution is received from the moving antenna.

Table 5.2: Spatial and contrast metrics using GM combination method.

	FWHM _ρ [mm]	SD _ρ [mm]	SNR [dB]
GM	9 ± 1	0.8 ± 0.1	8.1 ± 0.1

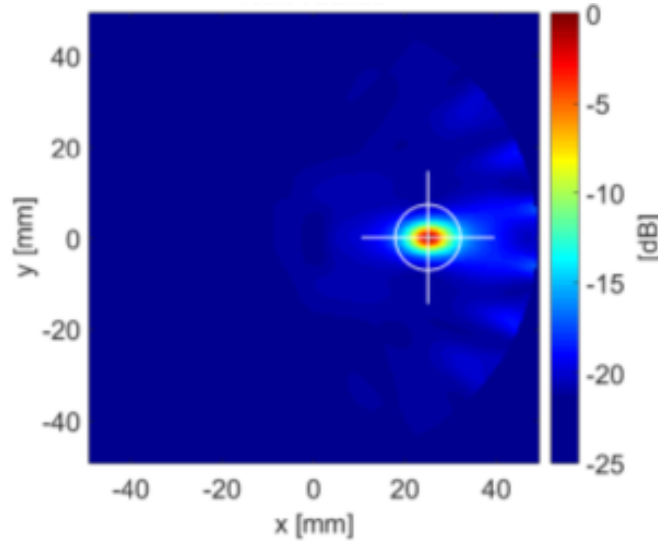


Figure 5.11: The resulting GM-MUSIC pseudospectrum combining frequency in $[2, 4]$ GHz range

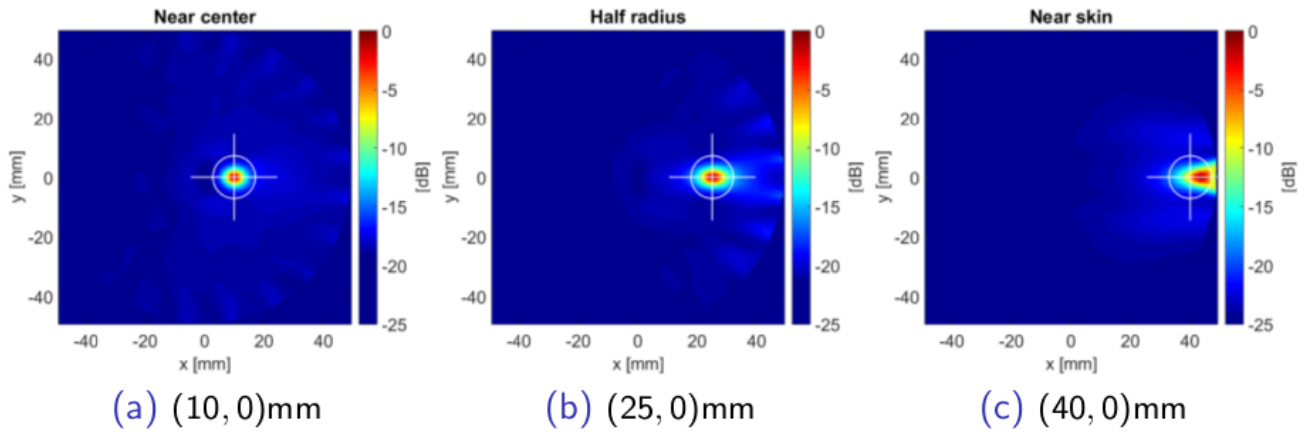


Figure 5.12: GM pseudospectrum changing the tumor positions. From left to right: (a)near center, (b)half radius, (c)near skin

5.3.4 Tumor size

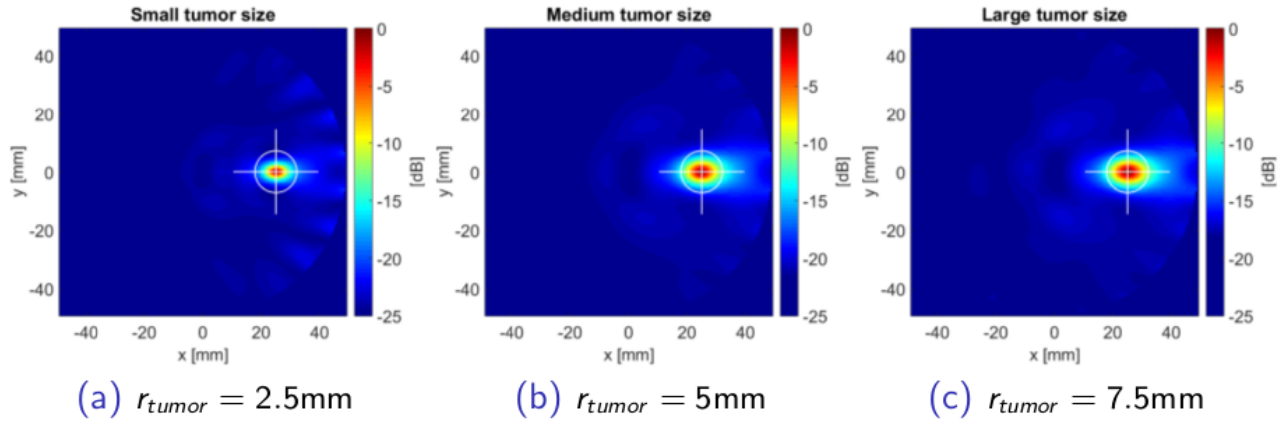
The goal is to identify and locate the scatterer varying its dimension. We want evaluate how the reconstruction is affected moving from a small tumor towards a larger one settled in the intermediate position between the skin and the center of the breast. We choose three different sizes:

- small tumor: $r = 2.5$ mm;
- medium tumor: $r = 5$ mm;
- large tumor: $r = 7.5$ mm.

Figure 5.13 shows the resulting reconstruction obtained with GM combination method in the frequency range $[2-4]$ GHz and table 5.4 summarizes the metrics results.

Table 5.3: Spatial and contrast metrics varying the tumor position using GM Music combination method

position	FWHM _{ρ} [mm]	SD _{ρ} [mm]	SNR [dB]
near center	6 ± 1	0.2 ± 0.1	9.5 ± 0.1
half radius	9 ± 1	0.8 ± 0.1	8.1 ± 0.1
near skin	11 ± 1	4.4 ± 0.5	4.1 ± 0.1

**Figure 5.13:** GM pseudospectrum changing the tumor size. From left to right: (a)small tumor, (b)medium size tumor, (c)large tumor

Obviously the FWHM increases with the tumor size since the FWHM is related to target size. We also note that when spread increases the signal reduces and the SNR decreases along with the tumor size.

5.3.5 Tumor-background contrast ratio (TBR)

In this study we investigate the effect of the TBR, that is how metrics are affected when the medium where the tumor is embedded changes. In a real scenario this corresponds at looking for a tumor in 3 different tissues. Inside the adipose tissue, inside an intermediate tissue or inside a fibroglandular tissue.

Table 5.4: Spatial and contrast metrics varying the tumor size.

tumor size	FWHM _{ρ} [mm]	SD _{ρ} [mm]	SNR [dB]
small	9 ± 1	0.8 ± 0.1	8.1 ± 0.1
medium	12 ± 1	0.4 ± 0.1	6.2 ± 0.1
large	13 ± 1	0.8 ± 0.1	5.4 ± 0.1

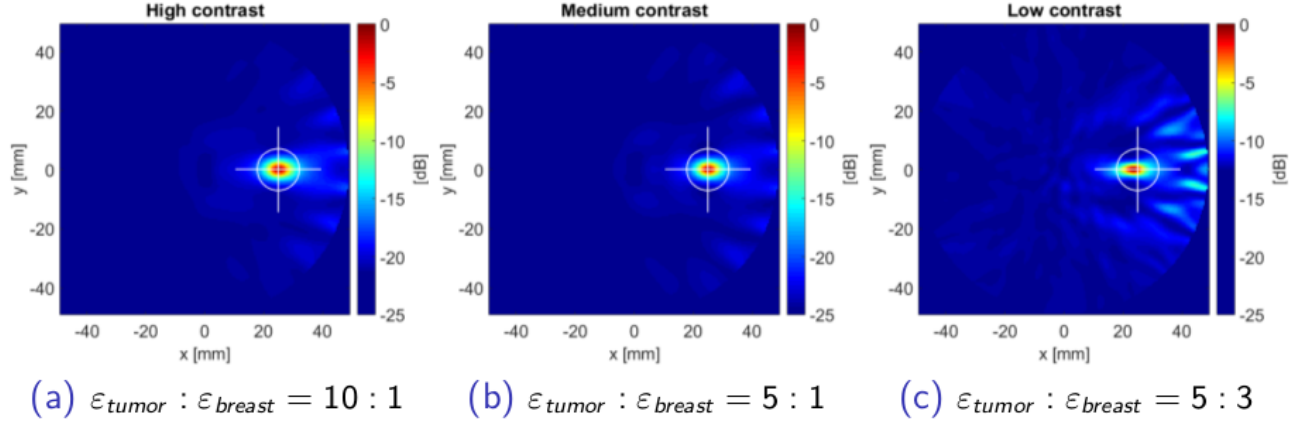


Figure 5.14: GM pseudospectrum for different TB-ratio values

To model these tissues in the frequency range needed to perform a multi-frequency combination the Debye dispersion model is again adopted. The following parameters are used to modify the tb ratio:

adipose:	$[\epsilon_{\infty} = 5.1, \epsilon_{sup} = 8.3, \tau =, 13 \text{ ps}, \sigma = 0.05 \text{ S/m}]$	$\bar{\epsilon} = 5$
intermediate:	$[\epsilon_{\infty} = 7, \epsilon_{sup} = 10, \tau =, 13 \text{ ps}, \sigma = 0.15 \text{ S/m}]$	$\bar{\epsilon} = 10$
fibroglandular:	$[\epsilon_{\infty} = 36, \epsilon_{sup} = 48, \tau =, 13 \text{ ps}, \sigma = 0.40 \text{ S/m}]$	$\bar{\epsilon} = 34$

which represents the changing background, while the Debye parameters of the tumor are kept fixed:

$$\text{tumor: } [\epsilon_{\infty} = 4, \epsilon_{sup} = 46, \tau =, 6 \text{ ps}, \sigma = 0.7 \text{ S/m}] \quad \bar{\epsilon} = 50$$

In the range [2-4] GHz permittivity with using the above Debye parameters can be considered almost constant and is referenced as $\bar{\epsilon}$ in the last column. Thus the TB-ratio decreases from fat to fibroglandular in the following manner:

tumor : fat = 10:1

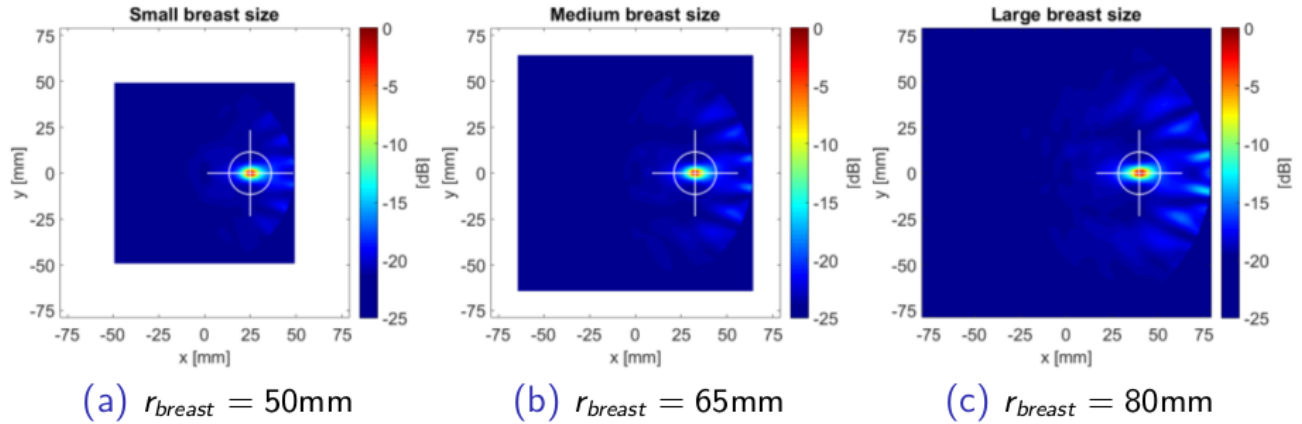
tumor : intermediate = 5:1

tumor : fibroglandular = 5:3

Figure 5.14 shows the pseudospectrum images obtained using GM combination method in the frequency range [2-4] GHz for a 2.5 mm tumor placed in the half-breast position and the corresponding metrics are reported in table 5.5. We can conclude that the algorithm performance is equal for high and medium contrasts but degrades for low values of the TBR.

Table 5.5: Spatial and contrast metrics varying TBR using GM Music combination method

TB ratio	FWHM $_{\rho}$ [mm]	SD $_{\rho}$ [mm]	SNR [dB]
10:1	9 \pm 1	0.8 \pm 0.1	8.1 \pm 0.1
5:1	9 \pm 1	0.6 \pm 0.1	8.1 \pm 0.1
5:3	9 \pm 1	-1.1 \pm 0.5	4.2 \pm 0.1

**Figure 5.15:** GM pseudospectrum for different breast size

5.3.6 Breast size

We decided to take into account the natural variability in breast size, therefore the algorithm has been tested changing this parameter to evaluate its performance in detection and localization. To simulate different size the breast radius r_b has been modified in following manner:

small breast size: $r_b = 50$ mm,

medium breast size: $r_b = 65$ mm,

large breast size: $r_b = 80$ mm.

the target location has been modify accordingly to remain in the half-radius position:

small breast size: $x=(25,0)$ mm,

medium breast size: $x=(32.5,0)$ mm,

large breast size: $x=(40,0)$ mm.

Figure 5.15 shows the resulting GM-MUSIC pseudospectrum of each breast size, computed using the frequency range [2-4] GHz and the relative metrics are specified in table 5.6 . From the study of the metrics we can conclude that the algorithm has same performance at different breast sizes in a perfectly homogeneous medium.

Table 5.6: Spatial and contrast metrics varying the breast size using GM Music combination method

	FWHM _{ρ} [mm]			SD _{ρ} [mm]			SNR [dB]		
small breast	9	\pm	1	0.8	\pm	0.1	8.1	\pm	0.1
medium breast	9	\pm	1	0.6	\pm	0.1	7.9	\pm	0.1
large breast	9	\pm	1	0.8	\pm	0.1	7.0	\pm	0.1

5.4 Monostatic and multistatic configurations

This section is devoted to a comparison between monostatic and multistatic (figure 5.16) approaches to gather information about their behavior in localization and detection in an homogeneous medium.

In this study the monostatic model is designed, an antenna is placed 2 mm far from the skin layer and used as transmitter and receiver at the same time, rotating over 20 different positions around the breast so that the angular step is 18° and S_{11} is numerically evaluated at each corresponding position.

In the multistatic scenario, 20 fixed antennas are settled in same positions where monostatic data have been previously collected: every antenna, in turn, acts as transmitter while the others are only receiving the scattered signal, thus to build up the full S-matrix. Both configurations are investigated along the frequency interval scanned from 1 GHz to 3 GHz with 0.2 GHz step.

All the following cases presented have been obtained by corrupting the scattered field with an additive white complex Gaussian noise with signal-to-noise ratio $\text{SNR} = 10$ dB.

Finally, we outline that the reconstruction in the multistatic case has been performed using only the eigenvector corresponding to the strongest not null eigenvalue of the TR matrix.

5.4.1 Architectures comparison

Both architectures have their advantages and drawbacks. Just to point out some of them, in monostatic there is a quite simple electronics and signal collection system, on the contrary multistatic has to deal with the multiple interconnections between the transmitting and the receiving elements and additional electronics devices, causing noise, must be introduced.

A simple multistatic configuration can also take advantage of the standing antennas and consequently no need to move any radiating parts. In the monostatic case a rotational stage around the breast has to be introduced, this increases the acquisition time due to the mechanical movements but permits to change the number of the data points collected. Monostatic has only one antenna so there are no interference between the radiating parts that causes mutual couplings but multistatic

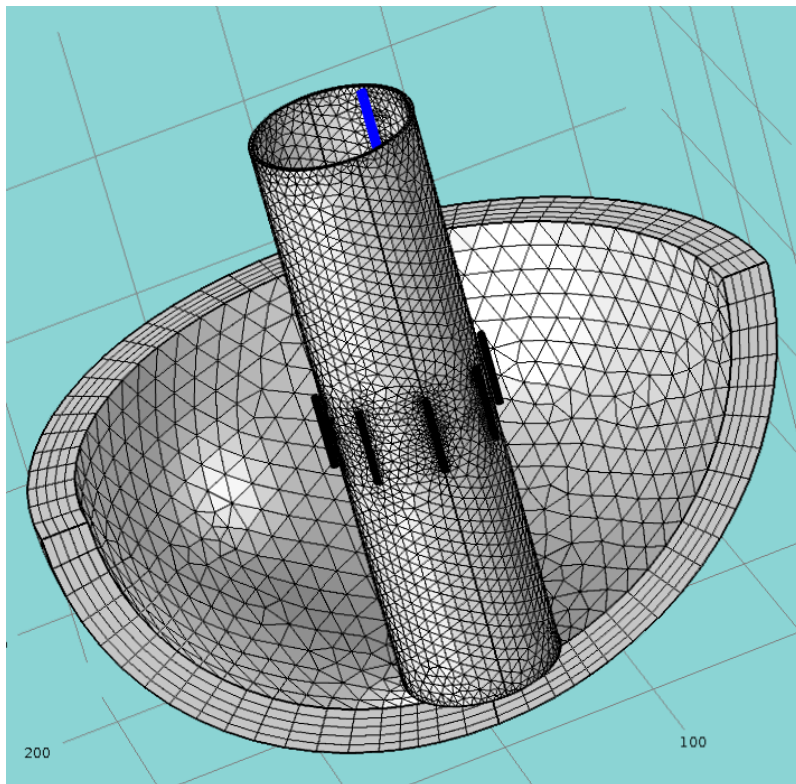


Figure 5.16: Multistatic architecture with meshing: the tumor is depicted in blue and around the breast modeled as cylinder are settled standing dipole antennas.

can exploit the indications coming from the full S-matrix.

Finally we outline that monostatic holds more choices in antenna design having no need to keep small the dimension of the antenna.

5.4.2 Multifrequency

We investigated the best combination method GM-Music, WB-Music and I-music defined in 3.5 changing the tumor position for both monostatic and multistatic approach to determine the one which performs better. The tumor coordinates are the same detailed in the previous section. To best aggregate data from now the inner position is named P1, the half-radius position P2 and the outer near-skin position P3. Figure 5.17 shows the pseudospectrum images for the three combination methods in the monostatic configuration using 11 frequencies in the range [1-3] GHz and 20 field points with tumor settled in P3, while figure 5.18 for the multistatic one under the same combination conditions.

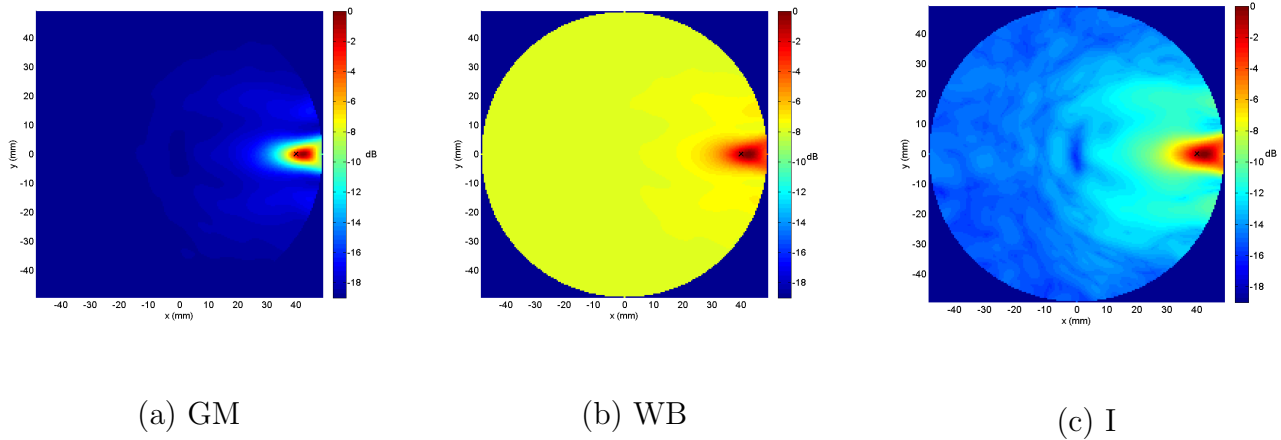


Figure 5.17: Monostatic pseudospectrum using different combination methods in P3 using 11 frequencies and 20 field points. (a) GM-Music (b) WB-Music (c) I-Music.

The metric results for all 3 combination methods are given in table 5.7 and 5.8.

Comparing these values we see that GM-Music combination method performs better than the other two methods and therefore its use in previous study and in the following is justified. The FWHM and SNR graphs for the two configurations as a function of the tumor position are displayed in figure 5.19.

For both configurations the FWHM improves moving from the outer position toward the inner position but monostatic performs better from 5% to 25%. Multistatic SNR is strongly dependent from the tumor position and the inner position gives best result while monostatic exhibits a weaker position dependence lightly increasing towards the skin.

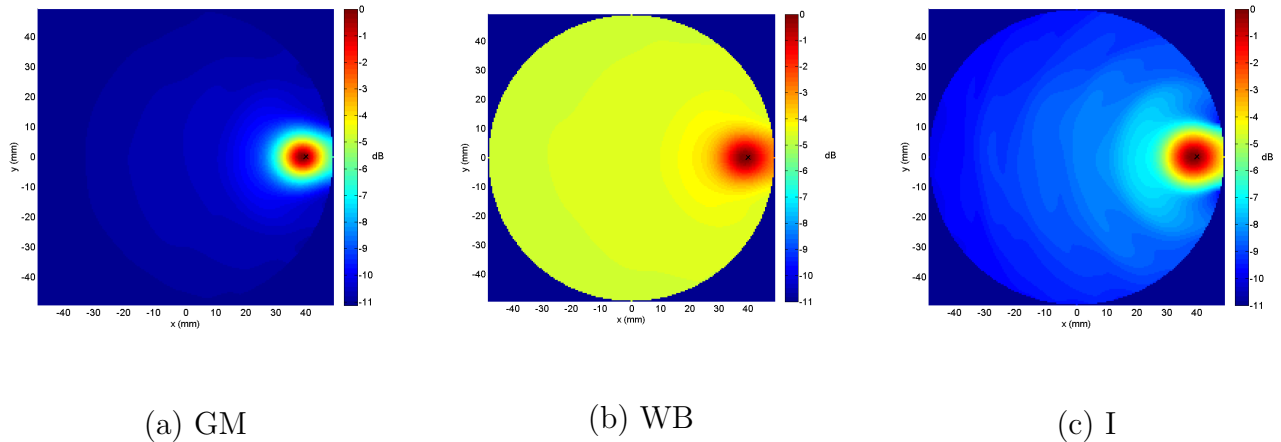


Figure 5.18: Multistatic pseudospectrum using different combination methods in P3 using 11 frequencies and 20 field points. (a) GM-Music (b) WB-Music (c) I-Music.

Table 5.7: Spatial and contrast metrics varying the tumor position for the three combination methods in monostatic configuration.

	GM		WB		I	
	FWHM [mm]	SNR [dB]	FWHM [mm]	SNR [dB]	FWHM [mm]	SNR [dB]
P1	7.7 ± 0.3	5.2 ± 0.1	7.6 ± 0.2	2.1 ± 0.1	13.0 ± 0.9	3.4 ± 0.4
P2	10.3 ± 0.3	5.4 ± 0.1	10.3 ± 2.3	2.1 ± 0.1	16.1 ± 0.7	3.4 ± 0.4
P3	13.9 ± 0.4	5.9 ± 0.1	13.8 ± 0.4	2.5 ± 0.1	17.6 ± 0.3	4.4 ± 0.1

5.4.3 Noise robustness

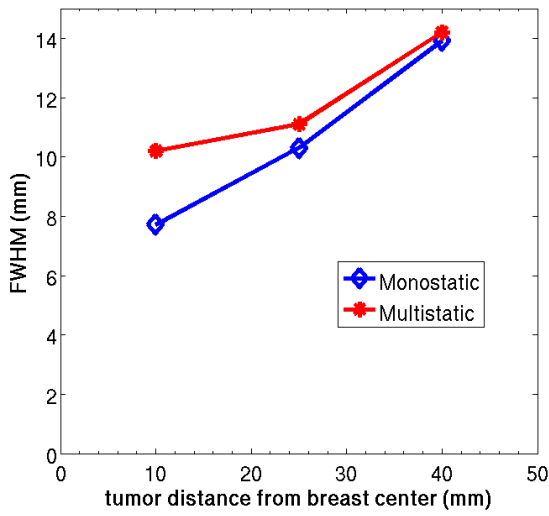
To evaluate the robustness of the two configurations to real noisy environment we have corrupted the data increasing the noise power. The results are summarized in table 5.9. Looking at figure 5.20 where data are displayed in a more explicit graphical way we note that SNR in the multistatic configuration exhibits more robustness to noise, while FWHM can be considered constant within the errors.

5.4.4 Number of field points

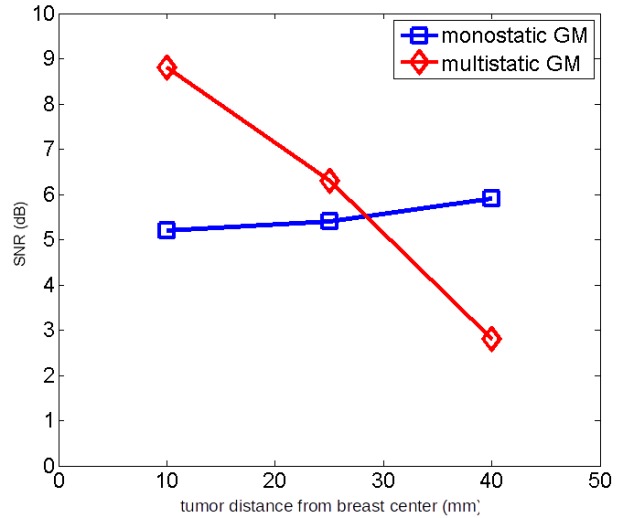
The goal of this study is evaluate an optimal field points number to be collected having anyway a good quality in the reconstruction metrics. This is a crucial point either in simulation, where adding a new collected point in multistatic and monostatic configuration means more time to accomplish the solution, or in real environment where only the multistatic configuration is affected. New antennas for the multistatic scenario, as detailed in the architecture comparison section, implies a more complicated electronics, system layout design and higher cost for cables and devices. We

Table 5.8: Spatial and contrast metrics varying the tumor position for the three combination methods in multistatic configuration.

	GM		WB		I	
	FWHM [mm]	SNR [dB]	FWHM [mm]	SNR [dB]	FWHM [mm]	SNR [dB]
P1	10.2 ± 0.1	8.8 ± 0.2	9.5 ± 0.1	3.7 ± 0.1	12.2 ± 0.1	4.5 ± 0.1
P2	11.1 ± 0.1	6.3 ± 0.4	10.6 ± 0.1	2.9 ± 0.1	13.2 ± 0.1	3.8 ± 0.1
P3	14.2 ± 0.1	2.8 ± 0.1	14.1 ± 0.1	1.2 ± 0.1	17.8 ± 0.1	3.0 ± 0.1



(a) FWHM



(b) SNR

Figure 5.19: FWHM comparison for monostatic and multistatic configuration versus the tumor position

lower the number of collected field points from 20 to 4. Table 5.10 summarizes the results while figure 5.21 holds the comparison plot. Monostatic metrics are strongly dependent from number of fields point and to higher number corresponds a better performance. For multistatic we have a great improvement in FWHM from 4 collected points to 8 collected points, but a slow increase from 8 to 20. Multistatic SNR increase slowly from 4 to 8 field points but decreases from 8 to 20. This is mainly due to the fact that a high number of antennas raises the mutual coupling and could be harmful to the reconstruction. Figure 5.21 shows the reconstruction images obtained with GM combination method for both configurations.

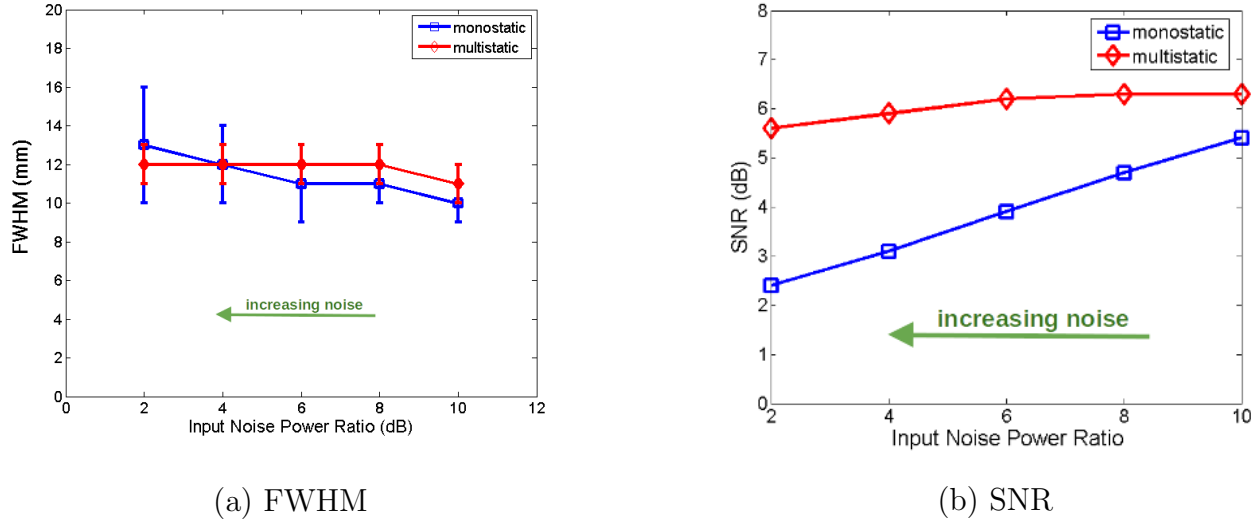


Figure 5.20: (a)FWHM and (b) SNR comparison for monostatic and multistatic configuration versus noise

Table 5.9: Noise robustness comparison for Monostatic and Multistatic configuration. Note: Noise Power Level (NPL) is the ratio between the signal power and the noise power, to an higher value of NPL corresponds a lower power of the added noise.

NPL	Monostatic		Multistatic	
	FWHM [mm]	SNR [dB]	FWHM [mm]	SNR [dB]
10 dB	10 ± 1	5.4 ± 0.1	11 ± 1	6.30 ± 0.04
8 dB	11 ± 1	4.7 ± 0.2	12 ± 1	6.30 ± 0.05
6 dB	11 ± 2	3.9 ± 0.2	12 ± 1	6.20 ± 0.07
4 dB	12 ± 2	3.1 ± 0.2	12 ± 1	5.9 ± 0.1
2 dB	13 ± 3	2.4 ± 0.2	12 ± 1	5.6 ± 0.2

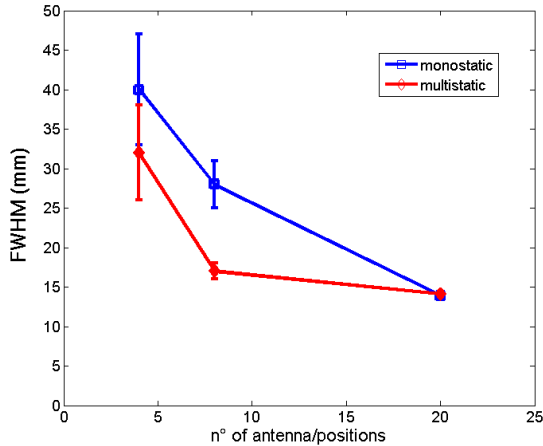
5.5 Inhomogeneous breast

As detailed in chapter four, we also have studied an inhomogeneous breast where, other than tumor, a fibroglandular island with a rugged profile is added (see figure 4.11). This approach has as primary goal the localization and detection of the tumor in a more realistic and challenging environment where the clutter signal produced by the fibroglandular structures is predominant over the tumor signal. The fibroglandular island is modeled as an extended target. The star shape has been chosen in compliance with the work of Hou-Solna and Zhao [29]. In their work the authors show that the TR-Music linear method along with a suitable propagator is capable to reconstruct the shape of an extended target embedded in a homogeneous background.

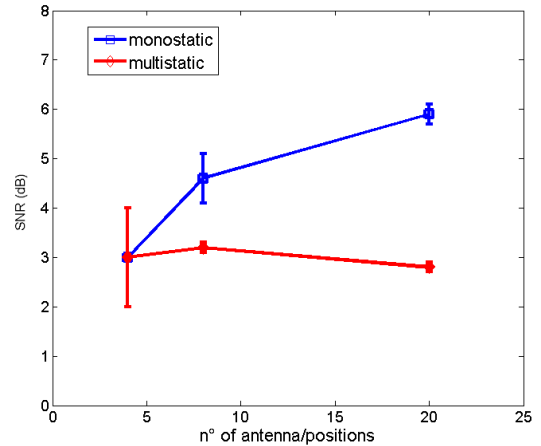
TR-Music states [27] that if the number of the point-like scatterers is less than the antennas

Table 5.10: Reconstruction metrics versus the number of field points collected

	Monostatic		Multistatic	
	FWHM [mm]	SNR [dB]	FWHM [mm]	SNR [dB]
4 f.p.	40 ± 7	3 ± 1	32 ± 6	3 ± 1
8 f.p.	28 ± 3	4.6 ± 0.5	17 ± 1	3.2 ± 0.1
10 f.p.	13.9 ± 0.4	5.9 ± 0.2	14.1 ± 0.1	2.8 ± 0.1



(a) FWHM



(b) SNR

Figure 5.21: (a) FWHM and (b) SNR comparison for monostatic and multistatic configuration versus the field points collected.

number, the resulting not null eigenvalues are equal to this number. A reconstruction of any scatterer can be performed using, individually, each corresponding eigenvector and making the scalar product with the steering vector of the imaged domain point.

For extended scatterers whose sizes are comparable to or larger than the resolution of the array the above analysis is not valid anymore. The response matrix has a more complicated structure. Even for a single extended scatterer, there will be many non-zero eigenvalues. In order to define a criterion to classify the dimension of the extended scatterer, we can refer to the metric of FWHM that is a measure of the support of the point spread function.

This is in general depending on the illumination wavelength and on the field points number. This means that a scatterer could be interpreted as "small" at lower frequencies and "extended" at higher frequencies. The number of the non-zero eigenvalues increases with MW frequency because the perceived size of the target get larger. Figure 5.23 shows clearly this behavior.

In this situation each eigenvector does not have clear physical interpretation and does not correspond to the illumination vector of a point on the target: an extended target can not be seen

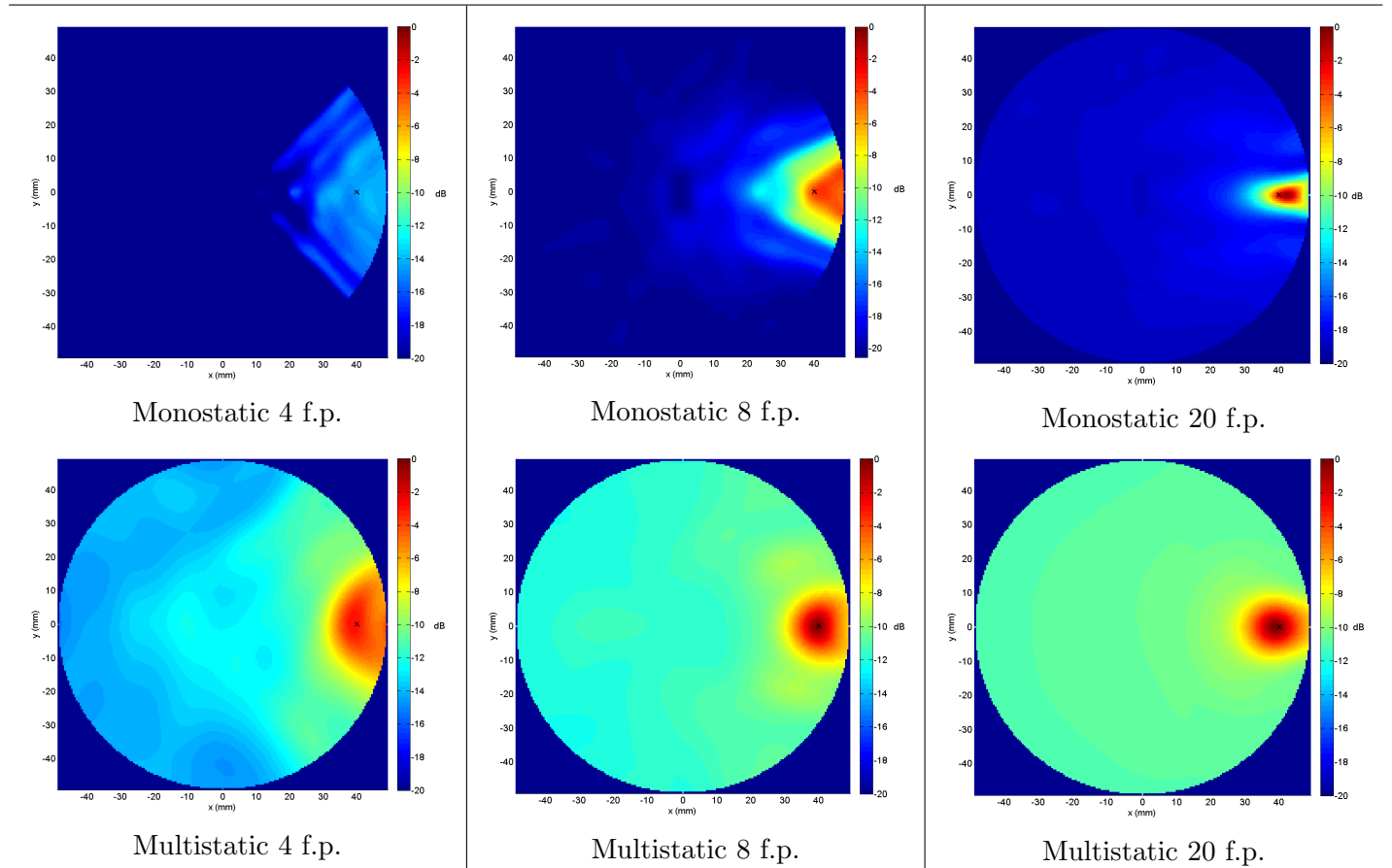
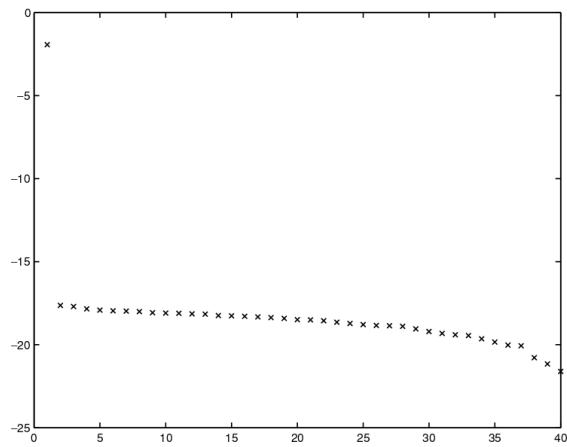
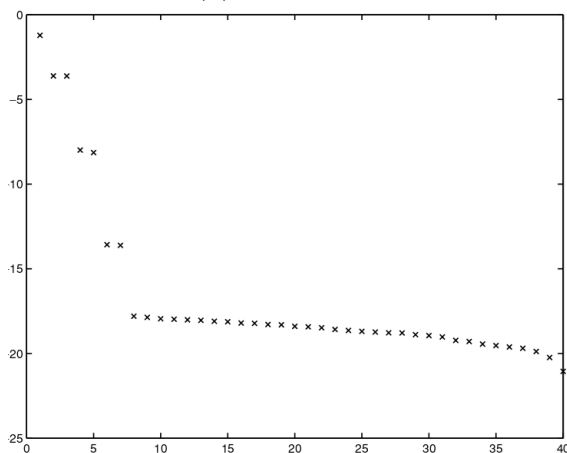


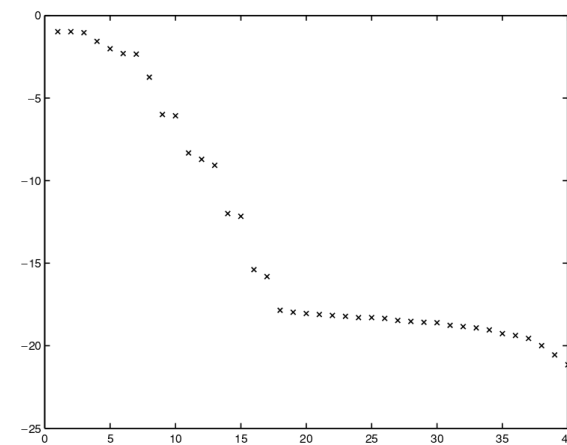
Figure 5.22: Monostatic (top row) and Multistatic (bottom row) reconstruction with increasing field points.



(a) point target



(b) small target



(c) extended target

Figure 5.23: Eigenspectrum examples in terms of wavelength compared to the target size: (a) point target, (b) small target (target size is a little fraction of the wavelength), (c) extended target (its size is comparable or larger than the wavelength). On x-axis the eigenvalue number on y-axis the log-magnitude [29].

as a collection of point targets. Considering the eigenvalues as an "ensemble", they can contribute to define the exterior boundary of the extended target, defining the signal subspace.

5.5.1 Imaging algorithm for extended target

This algorithm inherits the MUSIC (§ 3.4.1) framework to determine the TRM adding some changes we explain in the following.

- 1) TRM is computed starting from the simulated or experimental S-matrix data;
- 2) eigenvalues spectrum and corresponding eigenvectors are computed;
- 3) the MUSIC projection P is evaluated for each eigenvector having non-zero eigenvalue, using the free space propagator G defined in (2.22);
- 4) the strength of the i -th eigenvector is evaluated as the square root of its eigenvalue, λ_i and for each obtained pseudospectrum we define a weight w_i as follow:

$$w_i = \sqrt{\lambda_i} \quad (5.11)$$

- 5) we define a cutoff threshold to discard some eigenvectors: only the eigenvectors whose weight is over this value are taken into account to produce the final image;
- 6) the resulting pseudospectrum is computed as:

$$\Psi = \frac{1}{1 - \sum_i w_i P_i} \quad (5.12)$$

Actually we use an empirical method to define the threshold value by setting it at the 20% of the maximum value. The weighting method is justified by taking into account the definition of the TRM matrix. The eigenvalues are proportional to the square magnitude of scattered field, that is a measure of a scattered power. The square root is proportional to the amplitude of the scattered field.

5.5.2 Model description

An array of 80 antennas is placed on a circle line whose radius is 100 mm and each element is equally spaced from its neighbor. The background relative permittivity ε_b is set to 5 and the conductivity is set to 0.05 S/m. Within the breast there is a fibroglandular island whose polar equation is given by 4.5, the maximum value of its radius is 30 mm while the minimum is 20 mm and a circular scatterer whose radius is 5 mm. The fibroglandular and tumor dielectric parameters are modeled using the Debye formula (2.12) and already reported in table 4.1. In a first model where the tumor is outside the fibroglandular tissue, see Figure 4.11 (a) fibroglandular

dielectric parameters are set equal to those of the tumor. In the second model where the tumor is placed inside the fibroglandular, the fibroglandular dielectric parameters are instead fixed to a percentage of the tumor ones, ranging from 15% to 50%.

5.5.3 Role of the frequency

As previously told, the illuminating frequency determines how the system "sees" the target: when the frequency increases the corresponding wavelength decreases accordingly and the apparent size of the target becomes larger. As expected, a long tail in the eigenspectrum appears. Figure 5.24 shows the TRM eigenvalues spectrum when the frequency is increased from the lower 1 GHz value to the higher value 10 GHz with a corresponding long tail.

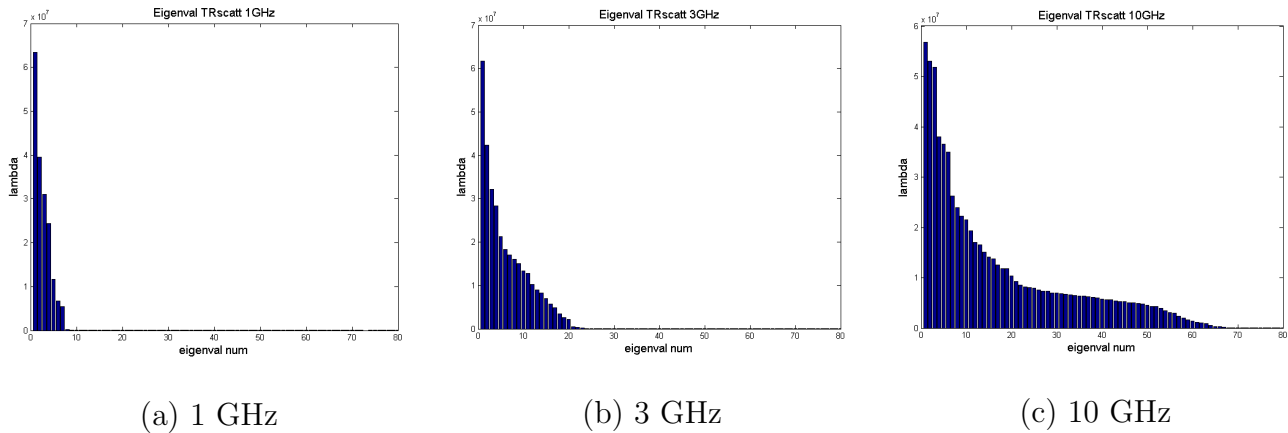


Figure 5.24: TRM eigenspectrum obtained with different illumination frequency. At higher frequency a long tail appears: (a) 1 GHz, (b) 3 GHz and (c) 10 GHz.

5.5.4 The eigenvectors role

As told, the reconstruction of the extended target should be performed combining the pseudospectrum obtained from several eigenvectors. Increasing this number the edges of the targets will appear with a better definition. The resulting pseudospectrum of our simulation can be examined in figure 5.25. With one eigenvector only the position of the tumor is visible with a low contrast. When adding more eigenvectors the edges become more clear and the dynamics gets better. The best reconstruction appears when 19 eigenvectors are concurrently used, after this value more clutter is introduced and the star edges become indistinct with a lower contrast. The resulting SNR is reported in Table 5.11: this value reflects the reconstruction and improves from one eigenvector up-to 19 doubling, then it becomes to decrease.

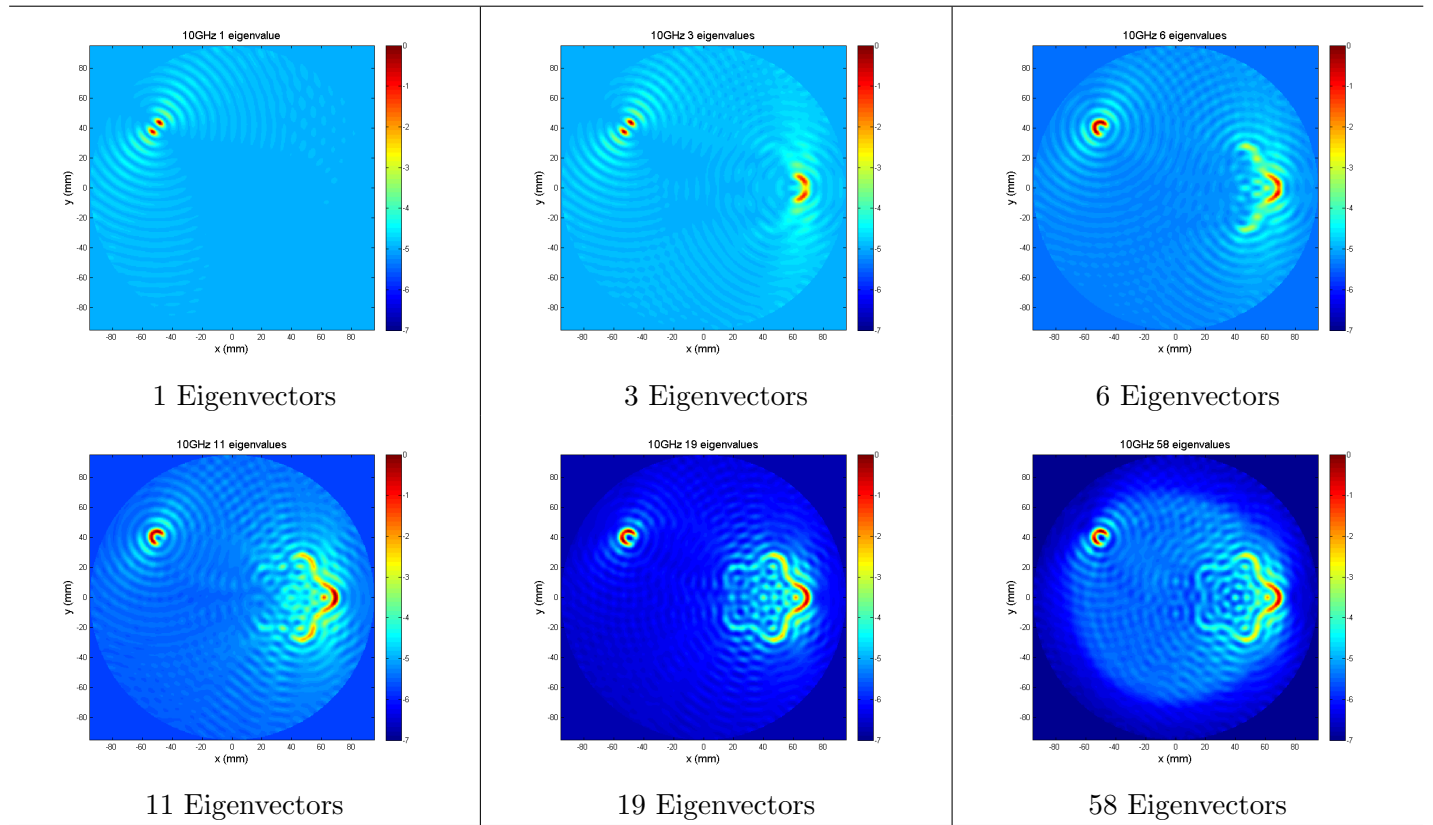


Figure 5.25: Reconstructed pseudospectrum obtained with 10 GHz illumination frequency and using an increasing number of eigenvectors computation.

Table 5.11: SNR evaluation corresponding to pseudospectrum reconstruction using a different number of eigenvectors.

N° Eigenvectors	SNR [dB]
1	3 ± 1
3	3 ± 1
6	4 ± 1
11	5 ± 1
19	6 ± 1
58	4 ± 1

5.5.5 The role of field points and the monostatic approach

We investigated how resolution is affected by the number of field points, in particular figure 5.26 shows the difference between the reference 80 field points of the model with an halved field points reconstruction when illumination frequency is 10 GHz in both simulations. In the latter case the threshold limits to 15 the number of eigenvectors used to compute the pseudospectrum corresponding to 37,5% of the total eigenvectors available. The optimal reconstruction with 80 antenna's uses 19 eigenvectors, that is the 24%. Visual inspection shows that the star edges are not sharp and more clutter appears if the number of fields points is reduced. This result appears in contrast with the previous study where is shown that multistatic performance suffers when the number of the antennas is increased. It is worth remarking that the illuminating sources are here modeled as ideal current sources. This suggests that a trade-off between field points number and number of antenna has to be afforded.

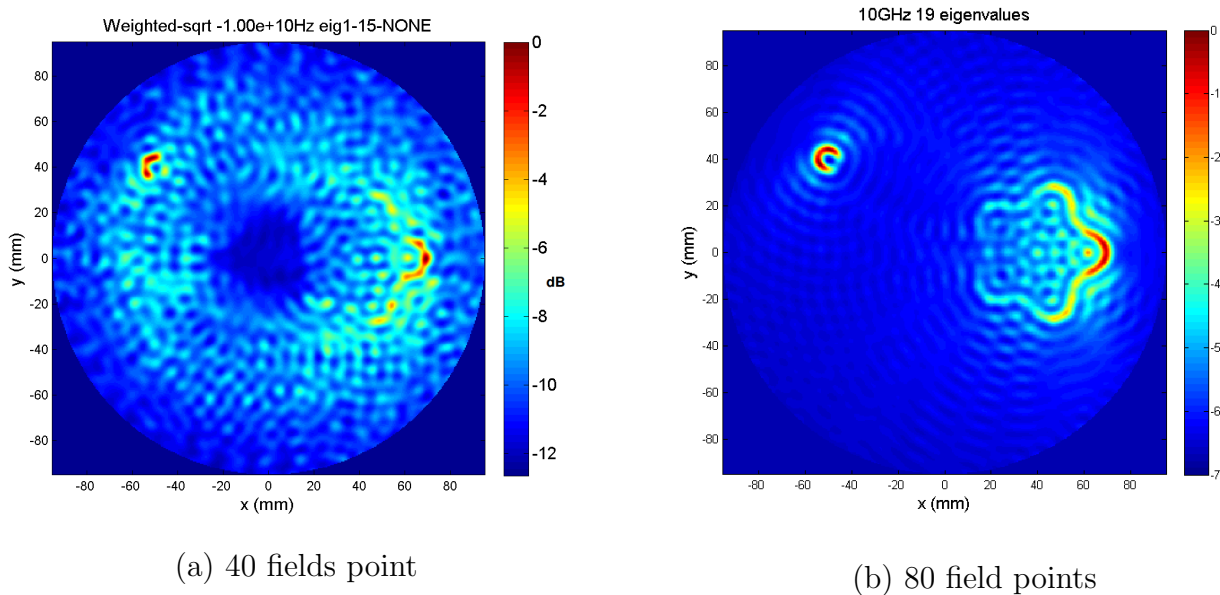


Figure 5.26: Reconstruction at 10GHz: (a) 40 field points and 15 eigenvectors, (b) 80 field points and 19 eigenvectors

Furthermore the monostatic configuration can not be used when rugged fibroglandular islands are added. Figure 5.27 is obtained using the monostatic approach by getting the S_{nm} value from the 80 field point S-matrix at 10 GHz and only clutter appears.

5.5.6 Scatterer inside the fibroglandular structure

We have settled the scatterer inside the fibroglandular structure to evaluate the resulting reconstruction. The parameters used in the model are the same as the previously exposed model except

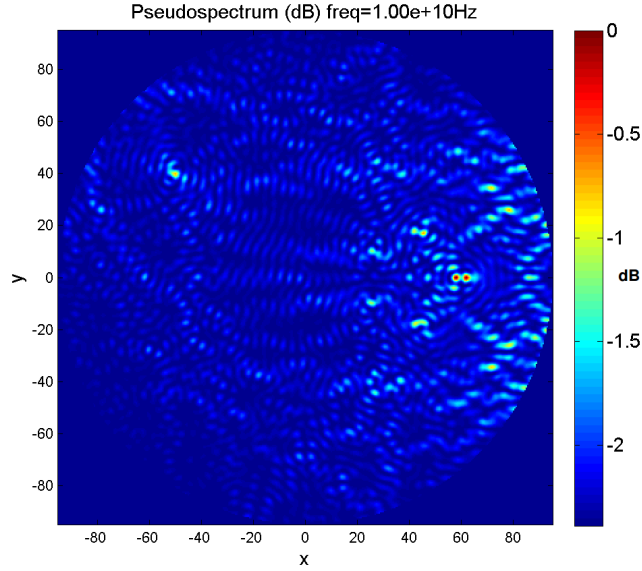


Figure 5.27: Monostatic reconstruction at 10 GHz: S_{11} is extracted from the 80 field points S-matrix.

for the dielectric parameters of the fibroglandular structure that has been changed from a minimum value of about 15% with respect to the tumor ($\varepsilon = 8$) up-to 50% ($\varepsilon = 25$).

Figure 5.28 shows the reconstruction varying the illumination frequency from 1 GHz to 5 GHz in the lower contrast case: increasing the frequency the star edges get sharper but the tumor is not detected.

Figure 5.29 shows the reconstruction obtained in the higher contrast situation: at 2 GHz the tumor is detected and at 10 GHz only the star edges are present. The GM-MUSIC combination shows both the fibroglandular tissue and the tumor inside it.

The range of frequency further explored is lower than the previous study and it has been fixed in the interval from 1.0 GHz to 1.5 GHz. The reason is in the resulting high value of the conductivity of the fibroglandular that, above this limit, extinguishes the tumor backscattered signal. We investigated the contrast limit in which the tumor signal can be revealed. Furthermore we evaluated the scattered power from the tumor and compared it with the power of the clutter signal due to the presence of the fibroglandular. In order to do this we defined the following scattering matrices:

S : full field matrix when both fibroglandular and tumor are present,

$S_{\text{no_tumor}}$: when only the fibroglandular is present,

S_{empty} : only homogeneous medium is present.

Figure 5.30 shows the COMSOL configuration whose solution allows the S-matrix.

The above S-matrices are individually obtained from COMSOL simulations. By difference we

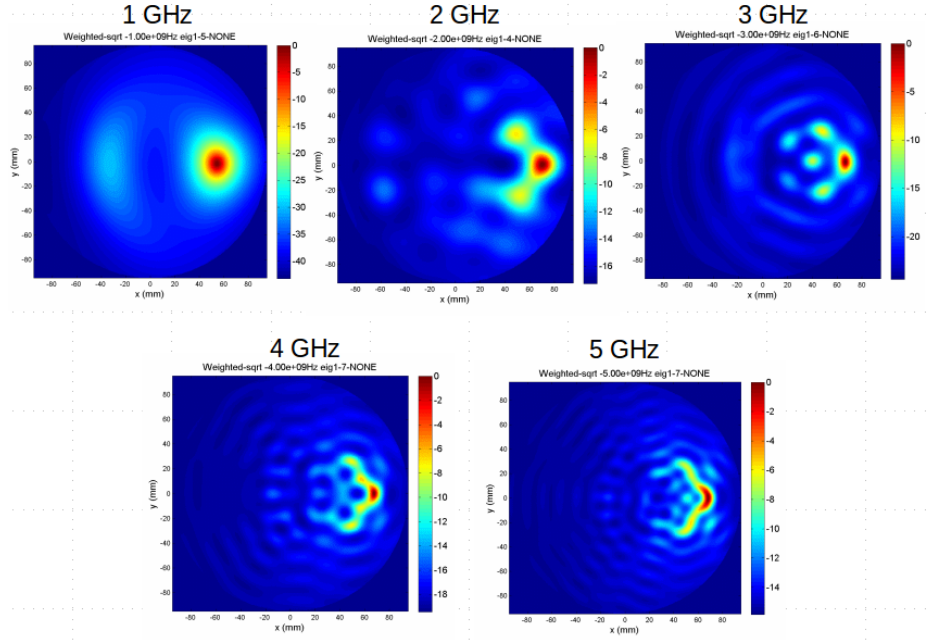


Figure 5.28: Reconstruction at different frequencies when tumor is inside the fibroglandular tissue: $\varepsilon_{\text{tumor}} = 50$, $\varepsilon_{\text{fbg}} = 25$, $\varepsilon_{\text{medium}} = 5$, $\sigma_{\text{fbg}} = 0.4 \sigma_{\text{tumor}}$. The tumor is not detected

define the tumor signal S-matrix and the clutter signal matrix:

$$S_{\text{scatt}} = S - S_{\text{no_tumor}} \quad (5.13)$$

$$S_{\text{clutter}} = S_{\text{no_tumor}} - S_{\text{empty}} \quad (5.14)$$

From these S-matrices, following the canonical procedure, we computed two time reversal matrices and the corresponding eigenspectra.

$$\begin{aligned} \text{TR}_{\text{scatt}} &= S_{\text{scatt}}^\dagger * S_{\text{scatt}} \Rightarrow \lambda_{\text{scatt}} \\ \text{TR}_{\text{clutter}} &= S_{\text{clutter}}^\dagger * S_{\text{clutter}} \Rightarrow \lambda_{\text{clutter}} \end{aligned} \quad (5.15)$$

As explained in § 5.5.1, the eigenvalues magnitude are a measure of the scattered power. The trend of the maximum eigenvalue magnitude is studied versus the illumination frequency and the fibroglandular relative permittivity.

5.5.7 The fibroglandular relative permittivity

We perform a new study at low frequency, because, as showed, only in this illuminating condition the tumor is detected, by changing the relative permittivity of the fibroglandular tissue.

Figure 5.31 shows the results: when the fibroglandular relative permittivity is low $\varepsilon_{\text{fbro}} = 8$ the tumor is correctly detected with a good $\text{SNR} = 8$ dB, increasing this value to 11 the detection

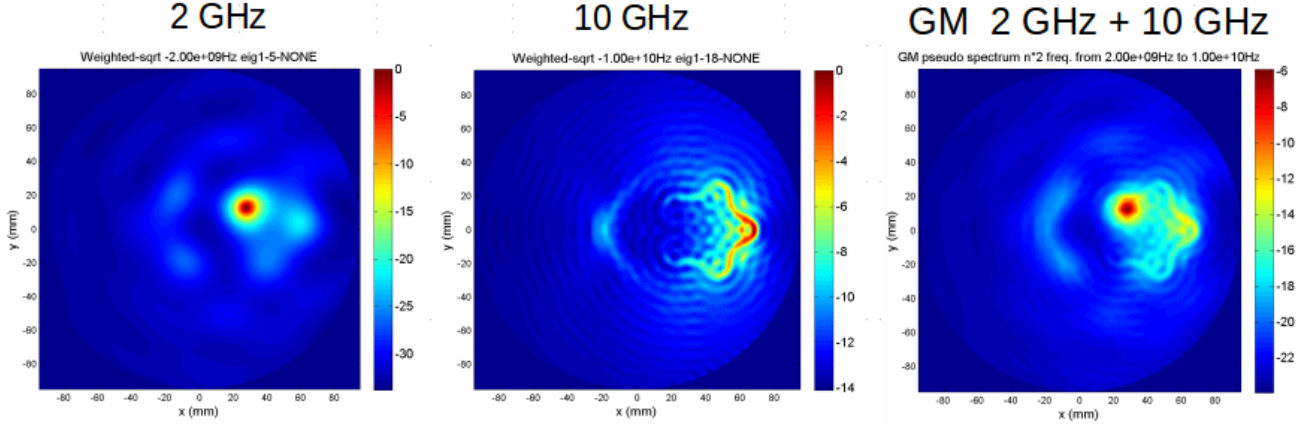


Figure 5.29: Reconstruction made at 2 GHz and 10 GHz when tumor is inside the fibroglandular tissue: $\varepsilon_{\text{tumor}} = 50$, $\varepsilon_{\text{fbg}} = 8$, $\varepsilon_{\text{medium}} = 5$, $\sigma_{\text{fbg}} = 0.15 \sigma_{\text{tumor}}$. The tumor is detected along with some edges the of fibroglandular structure.

is still performed but with lower SNR value and equals to 5 dB, when fibroglandular relative permittivity reaches the value of 16 the target is no longer detected and only cluttering appears.

5.5.8 The maximum eigenvalue

This study aims to quantify the variation of the magnitude of the maximum eigenvalue of the TR matrices defined in (5.15), that is considering clutter the signal emerging from the fibroglandular tissue. As told, this study is performed in the frequency range 1.0-1.5 GHz where the backscattered signal of the tumor can be revealed and varying the fibroglandular relative permittivity. Figure 5.32 helps to describe the results obtained. On the left we see the magnitude curve for the tumor eigenvalue: best performance are reached when $\varepsilon_{\text{fibro}} = [5 - 11]$ where we have maximum values at about 1.1-1.3 GHz, outside this permittivity range the magnitude decreases either at higher or at lower frequency value. When we move towards higher value of the permittivity the eigenvalue magnitude tends to rapidly lower and to be a constant in all frequencies interval. Examining the clutter eigenvalue behavior on the right graph we note that high values of $\varepsilon_{\text{fibro}} = [27 - 16]$ correspond to high eigenvalue magnitude, almost constant till the 1.3 GHz frequency then it decreases exponentially. Low values of $\varepsilon_{\text{fibro}} = [5 - 11]$ correspond to low and almost constant eigenvalue magnitude. We choose the illumination frequency of 1.3 GHz to show the trend of the magnitude versus the fibroglandular relative permittivity as shown in Figure 5.33: at low $\varepsilon_{\text{fibro}}$ the tumor eigenvalue dominates, until the crossing value at about 10 where clutter eigenvalue becomes stronger. This behavior is confirmed by the MUSIC pseudospectrum reconstruction.

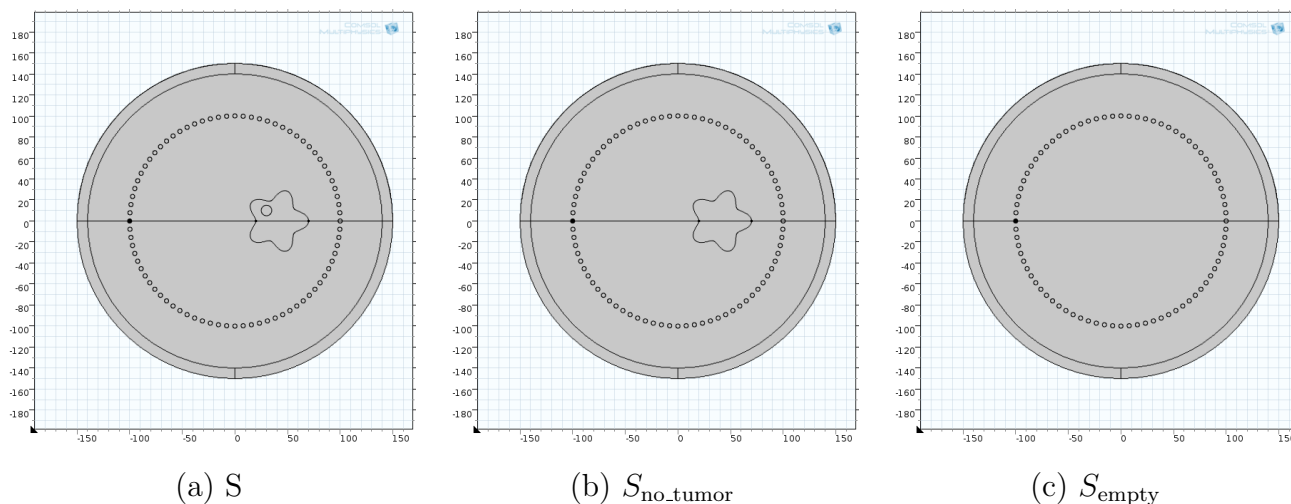


Figure 5.30: COMSOL simulation setup to obtain the S matrix in the corresponding scenario: (a) fibroglandular and tumor , (b) only fibroglandular and (c) homogeneous scenario

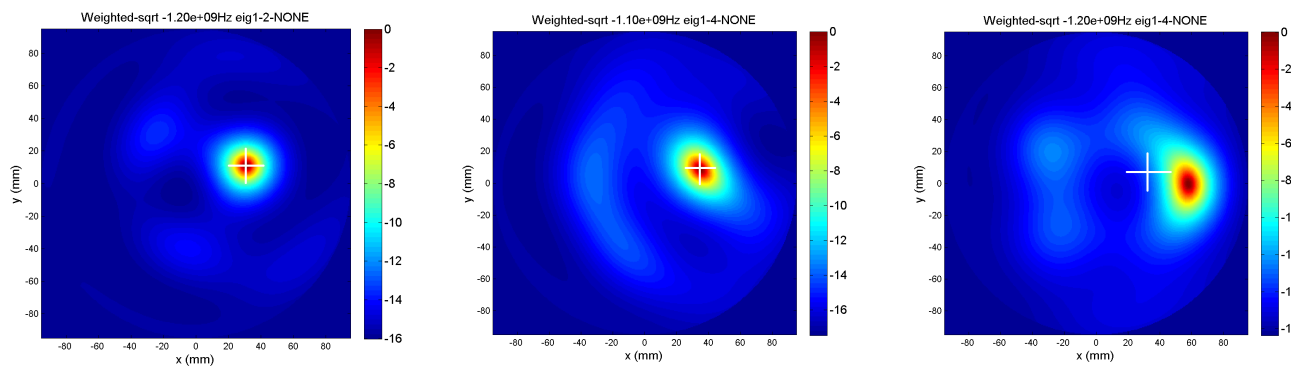
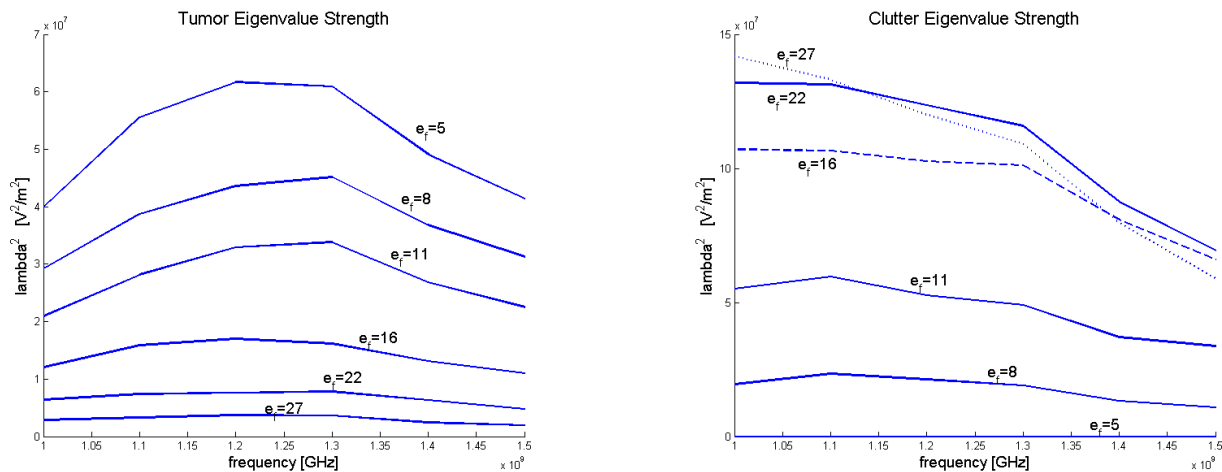


Figure 5.31: Reconstruction of the scenario changing the fibroglandular relative permittivity at low frequencies 1.1-1.2 GHz: (a) the target is detected with SNR = 8 dB, (b) the target is still detected with a lower SNR = 5 dB, (c) the target is not detected.



(a) tumor maximum eigenvalue

(b) clutter maximum eigenvalue

Figure 5.32: Maximum eigenvalue magnitude versus the illumination frequency. Each curve is obtained varying the fibroglandular relative permittivity: (a) TR_{tumor} , (b) $TR_{clutter}$.

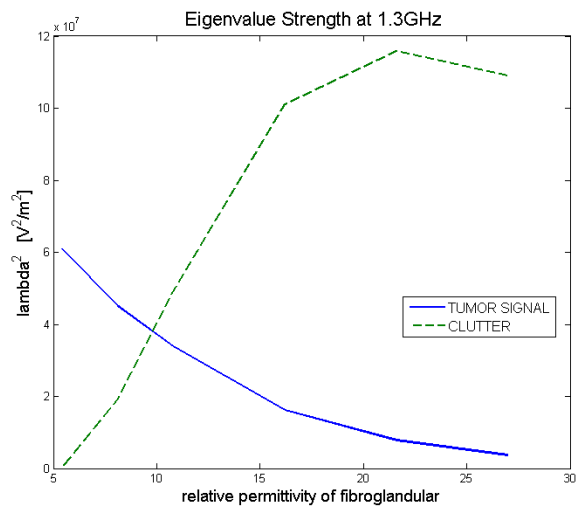


Figure 5.33: Tumor and clutter maximum eigenvalue magnitude versus the fibroglandular relative permittivity for a fixed illumination frequency at 1.3 GHz.

Chapter 6

Experimental data

The simulation analysis helped us to design an operating experimental setup. The study of the collected data helps to point out differences or similarities arising with the results obtained in the simulated framework. This chapter details how the experimental setup as been built: framework, materials and instruments. Then is presented the data acquisition procedure for the monostatic approach and an improvement for this framework is introduced. A second framework for multistatic approach is suggested.

6.1 Experimental scenario

The design of the experimental scenario mainly shares the same concepts described in chapter 3 and followed by all leading groups involved in breast MWI. As told before, to reduce the effects due to the mismatch between the mean permittivity of the breast internal tissue, the mammals have to be immersed in a cup filled with a matching fluid.

To perform the 2D reconstruction a sufficient height of the system must be kept, therefore it has been chosen a cylindrical plastic tank whose dimensions are 50 cm in height and 55 cm in diameter, used as container and limiting the scenario as PML does. Two more plastic cylinders have been used: the larger is 60 cm in height and 11 cm in diameter representing the breast. The cap of this cylinder is fixed on the bottom of the tank, so making quite easy the rotational stage. The second one is the tumor cylinder and has the same height of the breast cylinder, exactly how it happens in the simulation scenario, and its diameter is 15 mm having the lower end closed to avoid any leakage of liquid. To keep fixed the interaxial distance, $d = 33$ mm, between these cylinders two guides cable, similar to that shown in figure 6.1, are used.

Unfortunately this choice limits to only one scattering position in which tumor can be placed. A wooden board lies on the top of the tank and screwed on it. The big hole complete the rotational movement of the breast cylinder keeping fixed the distance with the transmitting antenna and its



Figure 6.1: The cable guide keeps the tumor cylinder in rest.

vertical axis perpendicular to the tank basis. Three other holes are realized on the board to dive the antenna into a suitable position.

Figure 6.2 shows a schematic top-view of the breast phantom. The thickness of the outer cylinder is about 2 mm, and should mimic the skin layer: though its relative permittivity is not exactly known, in this experimental condition is a very raw approximation, in fact plastics have a relative low permittivity (<http://www.engineeringtoolbox.com/>) that can be stated in the range of $2 \sim 6$ far from the value of 36 used in the simulation process. A corresponding low conductivity of the plastic with respect to the skin makes this setup a more favorable condition. Finally on the wood board, around the breast cylinder is placed a goniometric indicator with 1 mm precision used to manually read the angle of rotation.

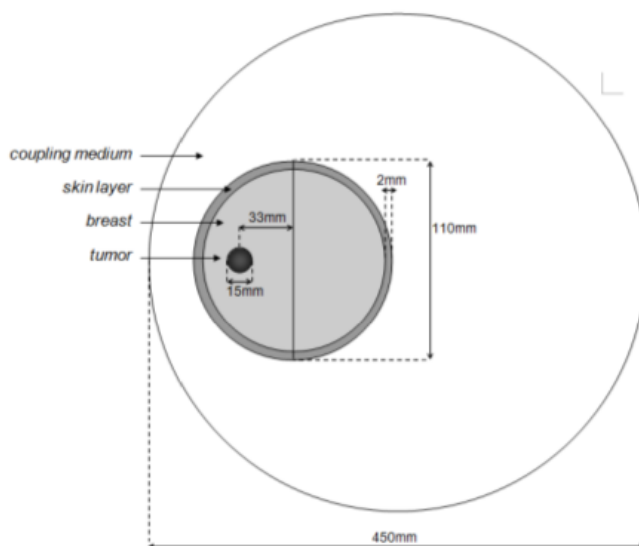


Figure 6.2: Schematic top-view of the phantom

This prototype, dubbed Archimede, has been fully developed and built by the Medical Imaging Group of the Physics and Astronomy Department of the University of Bologna. A detailed view

is in 6.3 (a) and the improved version (b), in the following detailed.



(a)



(b)

Figure 6.3: Full view of the monostatic imaging system: (a) Archimede first version, (b) Achille, the improved version.

6.2 Instruments

The instruments used for this monostatic data collection are a dipole antenna, an high precision coaxial cable designed for microwave range, a vector network analyzer (VNA) and a personal computer (PC).

6.2.1 Dipole

The dipole antenna is a ANT-24G-905-SMA developed by RS Solution. This is an outdoor antenna therefore suitable to be dived into liquid environments. An excerpt of its data sheet is reported in table 6.1 and figure 6.4 shows how is manufactured.



Figure 6.4: Dipole Antenna: ANT-24G-905-SMA

Table 6.1: Dipole antenna: excerpt of data-sheet

length	30 cm
Frequency (Center/Band)	2.4 GHz
Frequency Range	2.4 GHz ~ 2.5 GHz
Number of Bands	1
Gain	5 dB
connector type	SMA male-rp

6.2.2 VNA

High-frequency network analysis involves incident, reflected and transmitted waves moving along transmission lines. It is helpful to think of traveling waves along a transmission line in terms of a light-wave analogy: when light hits a material some part is reflected and some transmitted. VNA measures the proportion between the power transmitted and the power reflected in terms of the S-parameters already detailed in the section 2.3. A schematic block on how VNA accomplishes this task is given in figure 6.5.

Copper Mountain S5048 VNA has been used to collect and deliver data. It is two-ports fully vectorial, meaning that amplitude and phase of the signals are both collected: two reflection coefficients, S_{11} and S_{22} , and two transmission coefficients, S_{21} and S_{12} can be simultaneously measured. In table 6.2 are summarized its most important characteristics.

Prior to its use this device must be calibrated following the steps provided by the software. The two N-type ports are tested with three adapters: shorted, open and close. This operation has been performed using the Agilent 85032B/E calibration kit.

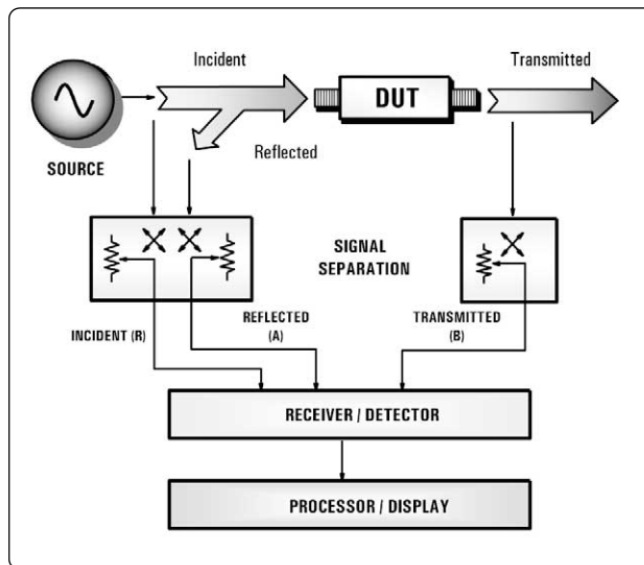


Figure 6.5: Schematic block of VNA operation.

Table 6.2: Copper Mountain S5048 characteristics: excerpt of data-sheet.

Frequency range	20 kHz to 4.8 GHz
Measured parameters	S_{11} , S_{21} , S_{12} , S_{22}
Sweep types	Linear frequency, log frequency, segment, power sweep
Dynamic range	120 dB (10 Hz)
Measurement speed	250 μ s per point at 85 dB dynamic range
Output power adjustment range	-50 dBm to +5 dBm

6.2.3 Personal Computer

The PC is directly linked to the VNA via an USB 2 connection and is used to hold the proprietary software deployed with the VNA and to record the data measured. It also holds the MATLAB reconstruction routines and the LABVIEW visual integrated environment along with the VNA virtual instruments driver to manage an automated acquisition in the improved version.

6.3 Experimental setup in monostatic configuration

Due to the simple phantom feature, the system has been only tested by changing the contrast factor between the breast and the inner tumor. The system is designed to replicate a typical monostatic scenario: the phantom is manually rotated for 360 degrees relative to the stationary dipole thus to simulate a circular antenna movement around the breast circumference as it happens in a real situation. For reading convenience, data are collected in 18 different positions corresponding at



Figure 6.6: Copper Mountain S5048 VNA: front view

20 degrees of angular step. The imaging domain radius corresponds to the breast radius, that is $r_{\text{breast}} = 55$ mm and the dipole turns around a circle whose radius is $r_{\text{dpos}} = 68$ mm.

6.3.1 Air-Water configuration

To test the system, a huge contrast between the breast permittivity ε_b and the tumor one ε_s was initially considered: the tank and breast cylinders has been left out-of-liquid and the tumor cylinder has been filled with water. In these experimental conditions the TB ratio is:

$$\varepsilon_b : \varepsilon_s = 80 : 1 \quad (6.1)$$

To apply the multi-frequency combination of the pseudospectrum using the GM-Music algorithm, other frequencies must be chosen. In particular we focus on the resonant frequencies visible in figure 6.7 where the reflection coefficient is showed in a log-magnitude graph versus a wide range of source frequency.

The first peak, below 1 GHz, is discarded due to the poor reconstruction resolution, limiting the working range in $[1, 4.2]$ GHz.

The reconstructions made at single frequencies show the same trend seen using synthetic data and the resulting metrics are shown in Table 6.3 where the mean value and the standard error are estimated making multiple acquisitions. The experimental results in air show that to lower frequencies corresponds larger FWHM, wider delocalization and higher contrast, while at high frequencies we encounter small FWHM, but less contrast due to emerging artifacts as it happened in the simulated scenario.

The best frequency range is $\nu = [2, 4]$ GHz, known result by simulations and confirmed by experimental measures. Using this information, the GM-MUSIC pseudospectrum is computed in

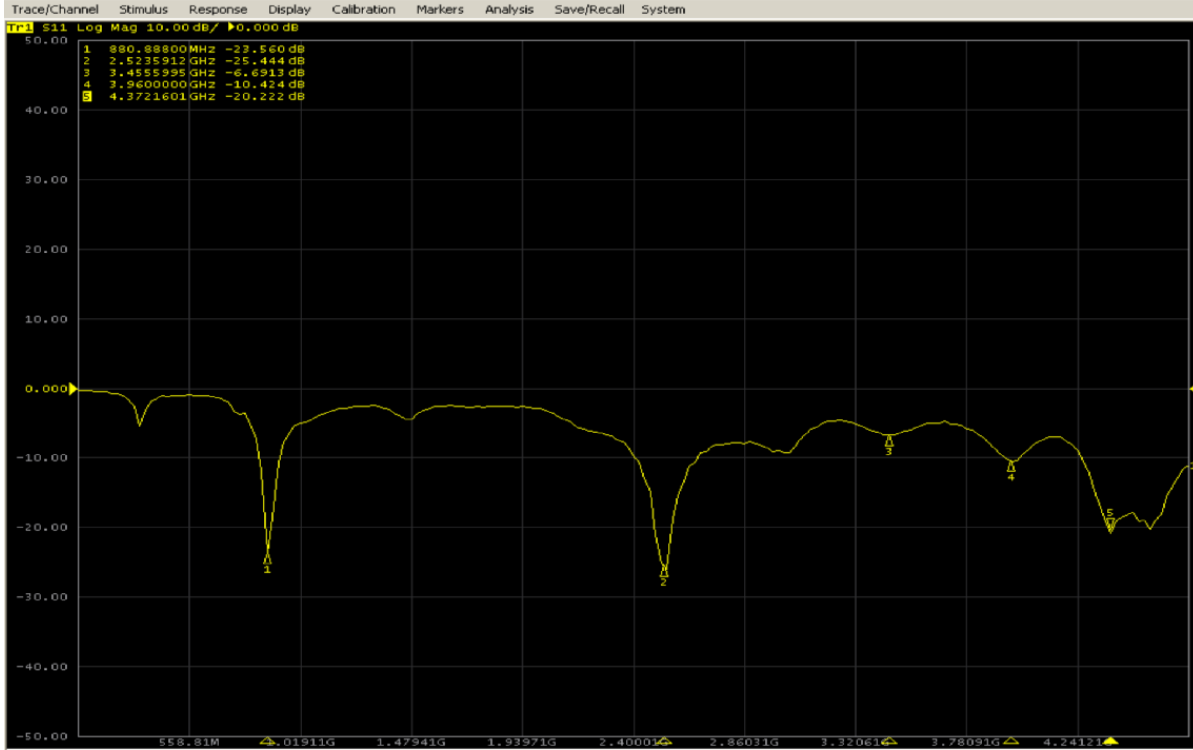


Figure 6.7: VNA return loss graph response for the dipole antenna. The resonant frequencies are used for the multi-frequency reconstruction.

the same range. The GM-Music metrics are reported in table 6.5, showing that for this algorithm lower contrast and smaller FWHM than the single-frequency case. Even in the experimental condition the localization metrics remains much smaller than spatial the resolution.

6.3.2 Measures with oil as background medium

After making calibration measures in air, we fill the phantom with seed oil whose relative permittivity can be variable in the range 2-3 depending on proportion and type of chosen seeds. To perform the assessment of the new scenario, the S11 signal produced by the water scatterer in the two different backgrounds is compared. Figure 6.9 shows the resulting return loss diagram.

The antenna behavior is affected by the presence of the background with a different relative permittivity and the operating frequency peak, 2.45 GHz in air, is now shifted to a lower frequency with a lower return loss. Other peaks appear at higher frequencies and should be useful for the multifrequency combination. To assess the correct value of the permittivity we perform the reconstruction by changing the corresponding wavenumber $k = k(\varepsilon)$ that appears in the Green function propagator defined in §2.22 tuning the relative permittivity along three different values: 2, 2.5 and 3. The reconstructed images are shown in figure 6.10 with the best localization and SNR using $\varepsilon_b = 2$, a value lower respect the mean value of a breast.

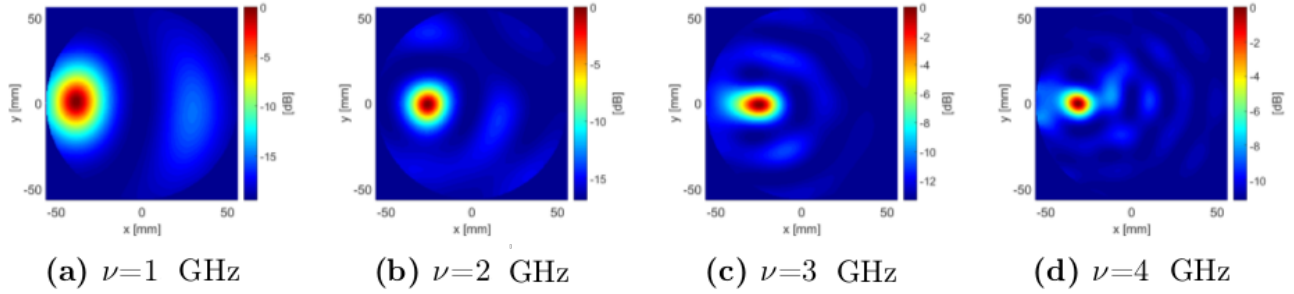


Figure 6.8: Air-water reconstructed images varying the illumination frequency in the VNA range.

Table 6.3: Spatial and contrast metrics varying the illumination frequency in the range 1 GHz - 4 GHz in air-water configuration.

frequency	FWHM _{ρ} [mm]	SD _{ρ} [mm]	SNR [dB]
1.0 GHz	21 ± 2	5 ± 1	5 ± 1
2.0 GHz	20 ± 2	1 ± 1	5 ± 1
3.0 GHz	15 ± 1	2 ± 1	4 ± 1
4.0 GHz	15 ± 2	-2 ± 1	3 ± 1

When using vegetables oil as coupling medium the castor oil, whose relative permittivity value is about 5, should be preferred. The assessment of this value allow us to define the TB ratio that becomes:

$$\varepsilon_t : \varepsilon_b = 40 : 1 \quad (6.2)$$

Clearly TBR magnitude does not still represents a real situation. We performed the single frequency reconstruction and some resulting images are displayed in figure 6.11, obtained using the same frequencies used to display figure 6.8.

The spatial and contrast metrics are shown in table 6.4 whose values represents the mean and standard error computed on repeated measures. Oil medium performs very badly at high frequencies and the artifact appearing at 4 GHz has greater value than the true signal while low frequencies are improved comparing to those in air. We choose the frequency range for that produces a useful reconstruction is 1.5-3.5 GHz having the experimental results in oil are similar to those in air due to the high TB ratio.

In this range the GM-MUSIC pseudospectrum combining single frequencies resulting metrics are reported in table 6.5. The GM-MUSIC has a similar contrast and larger FWHM compared to the best reconstruction in single frequency, but it exhibits better spatial displacement than the single-frequency.

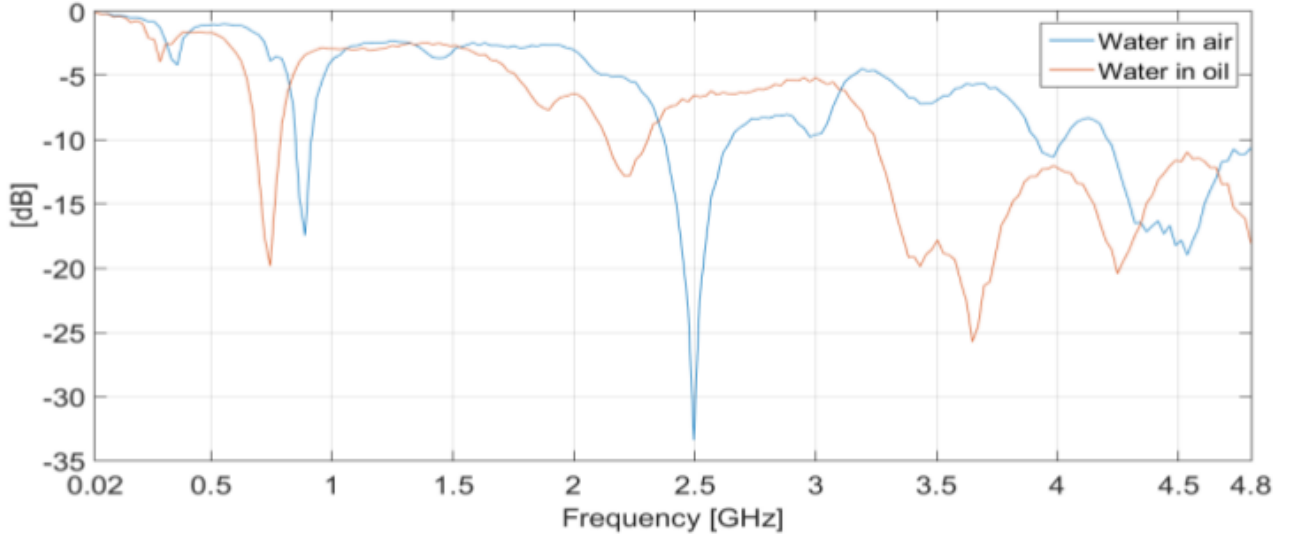


Figure 6.9: Return loss signal as a function of frequency. Red line: water scatterer in air, red line: water scatterer in oil.

Table 6.4: Spatial and contrast metrics varying the illumination frequency in the range 1 GHz - 4 GHz in oil-water configuration.

frequency	FWHM _{ρ} [mm]	SD _{ρ} [mm]	SNR [dB]
1.0 GHz	15 ± 2	-2 ± 1	7 ± 1
2.0 GHz	24 ± 2	-4 ± 1	4 ± 1
3.0 GHz	11 ± 1	1 ± 1	6 ± 1
4.0 GHz	21 ± 2	-22 ± 1	1 ± 1

6.3.3 Measures in oil with an ethyl alcohol scatterer

In order to further lower the TB ratio, the water in the scatterer cylinder in the ARCHIMEDE phantom has been replaced with a commercial solution of ethyl alcohol at 90%. The relative permittivity of ethyl alcohol mixture is equal $\varepsilon = 28$, by reference [63]. Now the TB ratio is.

$$\varepsilon_t : \varepsilon_b = 14 : 1 \quad (6.3)$$

We compute the GM-MUSIC pseudospectrum in the frequency range 1.5-3.5 GHz. The resulting images are summarized in figure 6.12 Table shows the metrics for the three experimental configurations along with a similar simulated case has been previously illustrated: $r_{\text{breast}} = 7.5$ cm located at half radius and with TBR = 10:1.

These data assert a good agreement between simulated and experimental data even if simulated

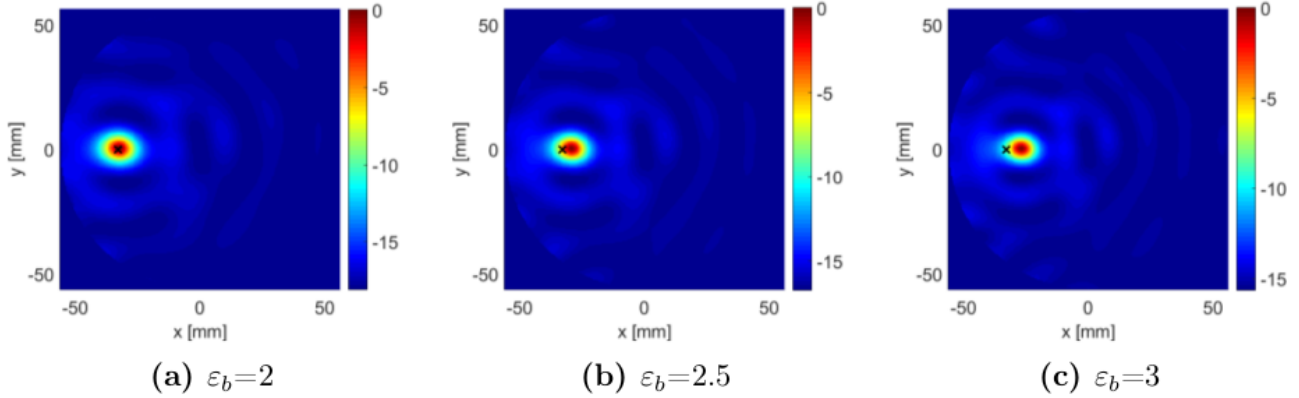


Figure 6.10: Air-water reconstruction at the illumination frequency 3 GHz, varying the background relative permittivity.

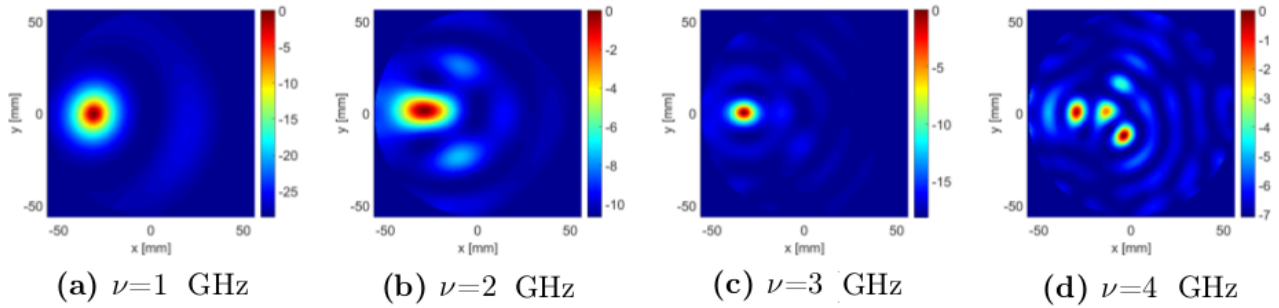


Figure 6.11: Oil-water reconstructed images varying the illumination frequency in the VNA range.

ones are slightly better in all metrics.

6.4 Achille and Briseide: the improved phantoms

The improved version of Archimede, named Achille, has been recently deployed: an alluminium cap is introduced on the top of the breast cylinder. This cap has been punched by several holes of different radius and position to ensure the full repetition of the simulated tests, varying tumor size and position and also to be able to add fibroglandular structures (see Figure 6.13).

Externally a wooden base is constructed where a TORNLABS CR1/M-Z8 motor, see figure 6.14, (with fine pitch worm gear that provides 360° continuous rotation) is fixed over a small pulley. A second pulley is joint to the alluminium cap and linked to the motor by a plastic belt to give the movement. This motorized system is fully supervised by a custom LABVIEW graphical interface installed on the PC: virtual instruments (VI) drive the rotational stage, collect VNA data and perform the reconstruction using the MATLAB routines.

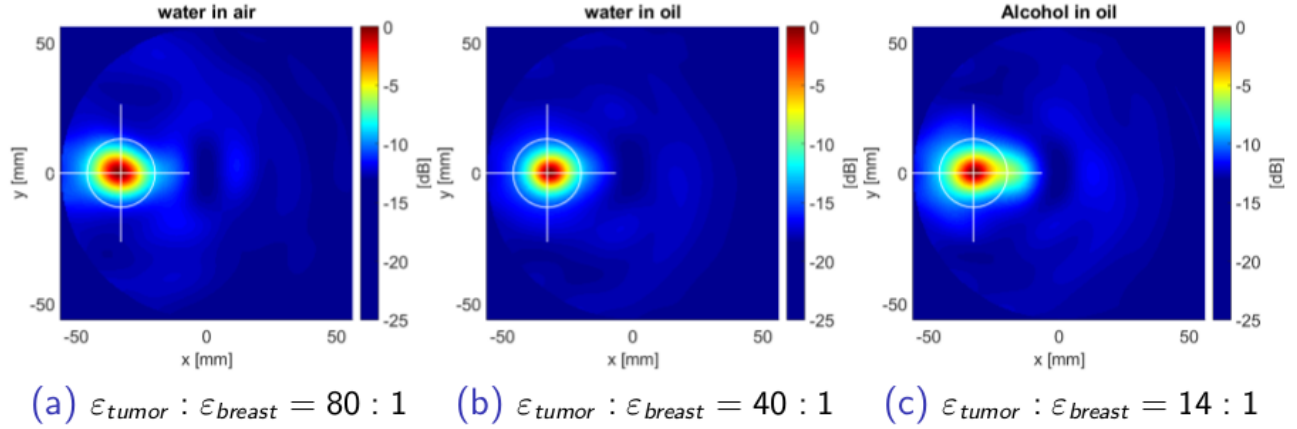


Figure 6.12: GM MUSIC presudospectrum for the three examined configuration.

Table 6.5: Spatial and contrast metrics for GM-MUSIC pseudospectrum in the three configurations. For comparison purpose the last row reports the value obtained in a similar simulated scenario having tumor radius equals to 7.5 mm and placed in half radius position. (TBR=10:1)

	FWHM _{ρ} [mm]	SD _{ρ} [mm]	SNR [dB]
water-air	18 ± 2	2 ± 1	4 ± 2
water-oil	15 ± 2	-2 ± 1	4 ± 2
alcohol-oil	17 ± 2	-2 ± 1	3 ± 2
Simulated	13 ± 1	0.8 ± 0.1	5.4 ± 0.1

Besides Achille, as myth tells, stands Briseide, our multistatic phantom. While the first is now acquiring data, the latter is in its very beginning. Twelve small cylinders are used to dive the dipole antennas into the tank and two knobs are used to raise and lower the antennas giving an appropriate longitudinal movement. A multichannel switch has also been added to the system instrumentation to manage the antennas shift. Briseide layout is finally displayed in figure 6.15.

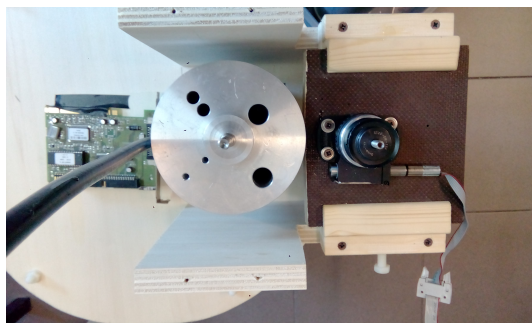


Figure 6.13: Achille top view: the holes host the tumor cylinders with different size and positions.



Figure 6.14: TORNLABS CR1/M-Z8 detailed view. It assures the automatic rotational stage in Achille phantom.

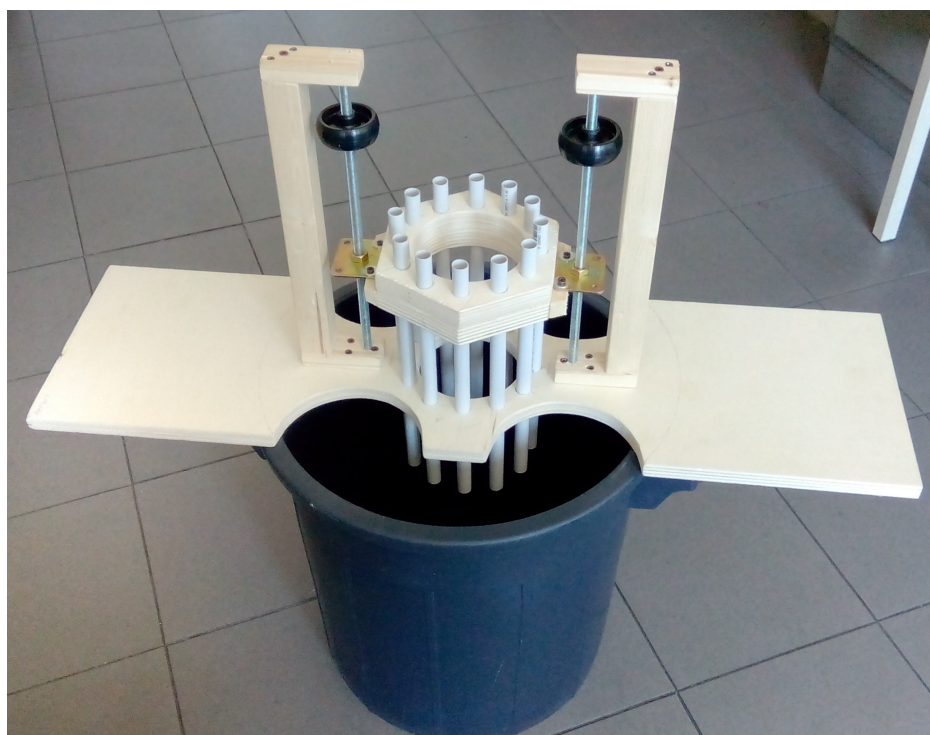


Figure 6.15: Briseide incoming phantom for the multistatic approach.

Appendices

Appendix A

Classification

The imaging system developed at the University of Bologna, department of physics and astronomy, presented in the previous chapter is in its early development, so that we were not able to test a classification technique based on the images gathered in real or clinic environment. In this chapter is presented an approach tested on a mammograms dataset and using a deep-learning technique, the Convolutional Neural Networks (CNN). This is an unsupervised methods belonging to the class of the artificial neural network, which aims to better mimic the human brain working model through high-level abstractions in data by using multiple processing layers (from which the term 'deep') with complex structures. The complete work, to which I have been co-advisor, can be found in [45].

A.1 CNN

CNN can be traced back to Neocognitron [20] in 1980 and their first use was in digit and object recognition and image interpretation. Since 2010 their use was impractical due to the high computational requirements which has been overcome with the Graphics Processing Unit (GPU) advent. They differentiate from the standard Artificial Neural Network (ANN) because its input layer is a map, that is a bi-dimensional structure instead the standard mono-dimensional vector (see Figure A.1). Theoretically an ANN fully connected and wide enough can learn raw images without features extraction, however following this approach many problems arise due to the high number of the parameters to train. Moreover traditional ANN uses a vector as input, this implies unrolling the 2D images and concatenate the rows or the columns in a resulting vector. Thus the topology of the input image will be ignored, losing the intrinsic spatial correlation of neighbor pixels.

Looking at Figure A.2, we see that one hidden layer consists of two different operation, namely a convolution and a sub-sampling. If we consider the input image I as a $N \times M$ matrix whose

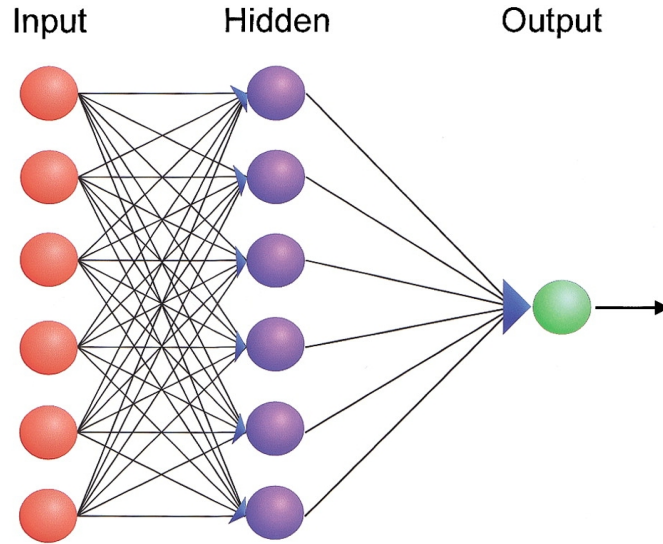


Figure A.1: Artificial neural network

coordinates, pixels, contain the corresponding intensity of the image at that point, we can define a filter as a transformation acting on the pixels intensity value of an image. The convolution is a filtering operation defined as:

$$O(N, M) = I(N, M) * K(N, M) = \sum_l \sum_j I(l, j) K(N - l, M - j) \quad (\text{A.1})$$

The K matrix acting on A is also known as *kernel*. If, as usual, the dimension of the kernel is lesser than the dimension of I we have a local filter, that is the intensity value of a pixel of the output matrix O depends only on a subregion of the input image I . We must take care because the convolution process is not well defined on the borders: in this case either a padding of the input image is needed or the convolution filter will be allowed to produce a smaller output image and, if the kernel has dimension $m \times n$, the resulting dimension the output matrix will be:

$$(M - m + 1) \times (N - n + 1) \quad (\text{A.2})$$

The CNN adopt the latter way. Again looking at Figure A.2, we see that the input image (32×32) is convolved with 6 kernels (5×5) providing 6 (28×28) new images called *Feature Maps*. Now the subsampling sublayer - also known as pooling - (2×2) reduces the dimension of the resulting features maps to (14×14) by substituting the region (2×2) with its mean and the new feature maps have rows and columns halved. The subsampled feature map then receives a weighting and trainable bias and finally propagates through an activation function. Some variants of this exist with as few as one map per layer or summations of multiple maps [5]. Last layer unrolls the feature maps and acts as standard ANN. Kernels are randomly initialized and are updated while

training through the back-propagation algorithm [48]. A suggested variation [64] is to substitute the subsampling stage with a non-linear max-pooling, that is instead of the mean the subsampling takes the max value of the considered region.

Convolutional Neural Networks

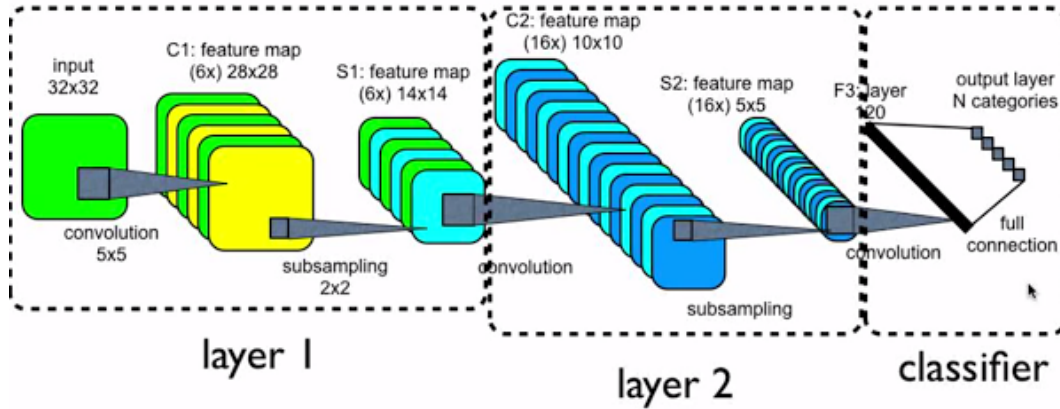


Figure A.2: Convolutional Neural Network

A.2 Digital Database for Screening Mammography

The Digital Database from Screening Mammography (DDSM) is an online resource of mammographic images maintained by the University of Florida. The purpose of the project is to provide a great database of mammograms in digital format to ease the development, evaluation and comparison of computer algorithms of Computer Aided Detection. The database consists of about 2600 cases, each of which contains four images, two for each breast (right MLO, CC right, MLO left, left CC), along with some information on the patient (age, breast density) and image information (type of scanner used for scanning, resolution space). Images that contain tumor lesions are associated to files with information regarding the location, the type and degree of evidence of the suspected region, thus to make a possible automated assessment of the detection algorithms performance. Unfortunately the format in which they have been saved is very heterogeneous: they come from an analog acquisition and later converted in digital format by means of three different scanners causing a very wide luminosity range. Furthermore some of them are damaged in due to dust, scratches and manual artifacts such as printed labels.

A.2.1 Dataset generation

The CNN training is made by choosing two crop types, one containing the lesions, said the positive set and one containing the normal tissue, said the negative set. Interesting lesions have 2.5 cm as maximum dimension thus to be referred the cancer as in "early stage". The crops are the input images for the CNN and their dimension has been chosen (64×64) pixels for computational issues. To ensure that the crop side is 2.5 cm a subsampling has been performed and the resolution of each scanner has been used to correctly compute window length L . For a scanner whose resolution $R = 43.5\mu m$.

1. From R Computing the pixels number to have a 2.5 cm crop side:

$$L = \frac{2.5 \text{ cm}}{R} = \frac{2.5 \text{ cm}}{43.5 \cdot 10^{-4}} = 575 \text{ pixel} \quad (\text{A.3})$$

2. update L to the nearest integer value divisible by 64:

$$L \rightarrow 576 \quad (\text{A.4})$$

Negative examples are extracted from the cases labeled as *normal* therefore without lesions. Crops are chosen to enclose at least the 80% of the mammal and some pre-processing like median filter for noise reduction has been performed. Masking with an adaptive threshold avoided the extraction by the background.

Positive examples are automatically extracted using the "overlay" text file that contains the ground-truth information, that is for cancer cases position, dimensions and morphology of the lesions. In our case only lesions in the range 1 cm - 2.5 cm are included. After applying the same preprocessing steps used for the negative cases, we center the lesion in this way:

1. creating a rectangular bounding box of the lesion taking the higher of the two dimension l
2. centering the bounding box in a square window with side l ;
3. centering the square window inside greater square window of side L , above computed;
4. cropping the window;
5. subsampling to generate the right size for the CNN input.

The corresponding dataset is made of 418 negative examples and 174 positive examples.

A.3 Binary classification

In general a medical test has the goal to tell if a situation is in a positive or negative condition (i.e. pregnant or not). When a case is correctly assessed is called a true positive (TP) if it belongs to the positive class and true negative (TN). False positive (FP) is a case belonging to the negative

class but associated to negative class and a false negative (FN) is a case belonging to the positive class but included in the positive one.

The ability of this test to correctly include a positive case in the positive class is called *sensitivity* and the ability to correctly include a negative case in the negative class is called *specificity*.

The measure of this ability is performed using the taxonomy just defined in particular:

$$\begin{aligned}\text{sensitivity} &= TPF = \frac{TP}{TP + FN} \\ \text{specificity} &= TNF = \frac{TN}{TN + FP}\end{aligned}$$

TPF is acronym for True Positive Fraction and TNF for True Negative Fraction. From these is possible define the False Negative Fraction (FNF) and the False Positive Fraction (FPF) as follows:

$$FNF = 1 - TPF$$

$$FPF = 1 - TNF$$

A.3.1 Receiver Operating Characteristic (ROC)

In a ROC curve the TPF is plotted as a function of the FPF for different cut-off points of a parameter (threshold). Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. The diagonal connecting (0,0) point with (1,1) represents a worthless system that does not distinguish between the two classes. Best ROC curve connects point (0,0) with point (0,1). The ROC curve can visually inspect the system performance in the operating point but lacks a quantitative index. The Area Under the Curve (AUC or A_z) is a quantitative index used to evaluate the performance of a classifier.

A.4 Model selection

Model selection involve the choice of the optimal architecture and the tuning of the system parameters and hyper-parameters. Regarding to CNN the architectural choice concerns:

- input dimension,
- number of layers (convolution + subsampling),
- filters number and dimension,
- subsampling type: average or max,

- classifiers: softmax (also known as multimultinomial logistic regression) or support vector machine (SVM).

The model parameters are:

- the values of the kernels
- the weights of fully-connected last layer.

The model hyper-parameters are:

- number of epochs for training,
- learning rate.

The architectural choice has been performed after a random shaking process similar to cross-validation. The dataset is divided in two more sets: training set containing the 80% of the examples and test set containing the remainder. The system is trained on the training set and the error is computed on the test set. The procedure is repeated 10-times with a random permutation of the examples and the final error is given averaging each iteration error.

Following these steps the resulting optimal architecture is thus defined:

- 1st layer: 5 kernels with (5×5) dimension,
- 2nd layer: subsampling with max-pooling technique,
- 3rd layer: 3 kernels with (3×3) dimension,
- 4th layer: subsampling with max-pooling technique,
- 5th layer: SVM classification.

and the hyper-parameters has been set in the following way:

- number of epochs = 3000,
- learning rate = 0.02.

A.4.1 Training the model

To perform a good training is important an accurate normalization process, in our case respect to the gray level of the images. We perform the normalization by subtracting from an image its mean value, obtaining pixels value distributed near zero. For each pixel $[(1,1), (1,2)...(64,64)]$ we calculate mean value and standard deviation using all images, then each pixel is normalized subtracting this mean value and dividing by the standard deviation.

A.5 Classification results

The idea to use a CNN to detect lesion in mammographic images was previously afforded in 1996 by Chan [54] using two approaches, with texture features and featureless: this CNN used only 1 layer and the crop dimension was 16×16 or 32×32 . Our approach is only featureless because does not

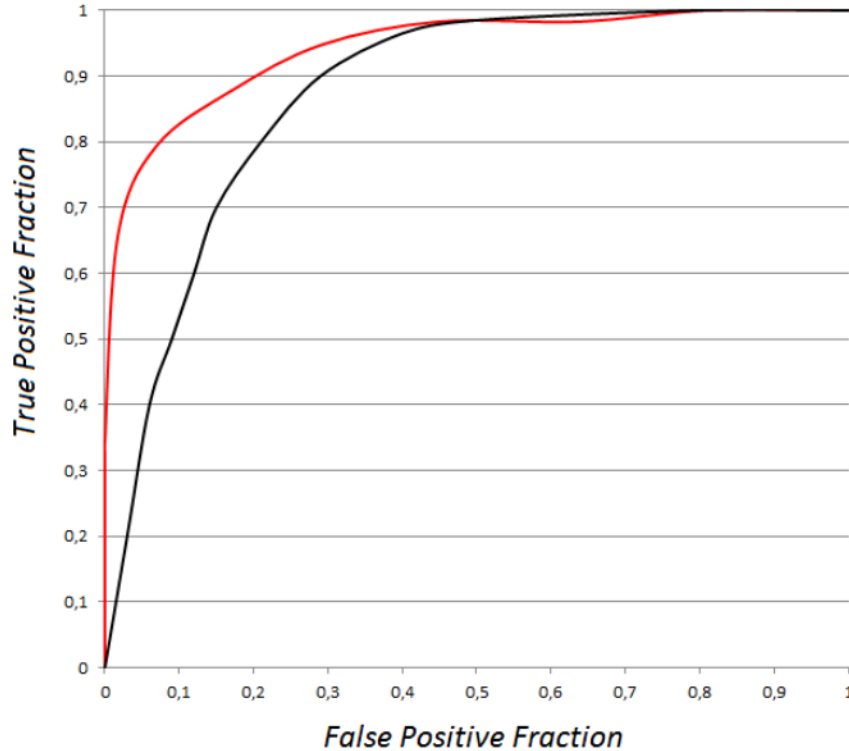


Figure A.3: ROC comparison between this work and Chan work.

Table A.1: A_z comparison. [54]

	A_z (featureless)	A_z (feature)
Our CNN	0.94 ± 0.01	/
Chan	0.83	0.87
Angelini (SVM)	0.973 ± 0.002	/

need further preprocessing steps. We add a second layer to the CNN and use larger crop, 64×64 , obtaining an $A_z = 0.94 \pm 0.01$. Figure A.3 shows the corresponding ROC curves, in our work approximately TPF=0.9 is reached when FPF = 0.2. Other works found in literature and working on the same dataset exhibit a slightly better behaviour, like in Angelini et.al. [4] in which DDSM is used for extracting 5000 negative and 1000 positive crops. The classification process is performed by means of a SVM in a featureless approach. They reach $A_z = 0.973 \pm 0.002$ and for a TPF=0.9 the corresponding FPF is 0.07. To improve our performance a higher number of examples should be preferable along with a better normalization process. The comparison between these works is in Table A.1.

Conclusions and future work

The process for building a breast cancer MWI phantom has been detailed. We develop some simulations using COMSOL starting from a simple scenario where a tumor is embedded in an homogeneous breast. This basic situation is gradually complicated adding a rugged fibroglandular structure and placing the investigating tumor outside and inside this tissue. Moreover a realistic model from MRI-derived phantom has also been simulated.

Two different approaches have been afforded inside the simulation process: monostatic and multistatic.

The reconstruction process has been performed using a linear approach and a novel multistatic filter for decluttering is also introduced.

An experimental phantom has been built from scratch for monostatic data acquisition: layout, instruments and methodology has been detailed. Experimental results are obtained varying the TB ratio.

The detection and localization in a homogenous scenario can be easily realized either in multistatic or monostatic configuration, and in latter case simulated results are comparable with those obtained in the experimental setup.

When fibroglandular tissue is added the detection becomes more challenging and can be reached only when the contrast between the fibroglandular tissue and the tumor is high. Therefore the exact knowledge of dielectric properties of the breast tissues are mandatory to further explore the MWI. In this case our results suggest not to follow the monostatic approach if used in combination with a linear reconstruction method. Multistatic approach has to deal with the mutual coupling between antennas and a trade-off with number of field points must be taken into account.

The results obtained using a simulated 3D MRI-derived phantom are less conclusive and have not been presented here.

The classification, shown in appendix as separate work, is performed using CNN, a leading technique in industrial pattern recognition, and realized on standard mammographic images due to the lacking of real microwave images. The classification results are better than those founded in literature using CNN in a different way, but further insight is needed to raise the performance to those achieved by means of other methods such SVM.

Future work We are already investigating an homogeneous breast from a full 3D point of view, that is using a 3D propagator for the point-source to evaluate the MUSIC steering vector. Mie propagator in 2D and 3D, suitable for extended targets, will also be implemented as soon as possible and some preliminary results not included in this work are already available.

In the meanwhile the non-linear approach for the inverse problem is being afforded for 2D simple situations.

The simulation step will be improved using other antenna types, more compact and with high gain and directivity, such as patch antenna.

As to the experimental stage, the phantom is now improved and the acquisition and reconstruction process is fully automated using LABVIEW. The newly introduced features allow changing the tumor size, position and number, moreover a big fibroglandular structure can be added.

The multistatic phantom is currently developed and soon will be operating, allowing a comparison with the already available simulated data.

Ringraziamenti (rigorosamente in italiano)

Quando si giunge alla conclusione di un cammino lungo, faticoso ma bello è naturale volgersi indietro e guardare quanto ci ha accompagnato.

Ringrazio la Polizia di Stato perché, pur non essendo obbligata, ha agevolato questa mia scelta.

Il Prof. Renato Campanini che mi ha permesso di affrontare il percorso che più mi si addiceva lasciando che le mie inclinazioni, mai state quelle della fisica delle particelle, potessero liberarsi in ambiti più consoni. Grazie per non avermelo mai fatto pesare!

Il prof. Nico Lanconelli che mi ha accolto sotto la sua guida, precisa, geniale ma anche umoristica e per avermi permesso di fare parte di un team coeso e determinato insieme al Prof. Simone Masetti che in questo ultimo periodo mi ha dedicato innumerevoli energie: dopo tanto rincorrersi finalmente limiamo insieme qualche pezzo di legno in cantina e finiamo a sorseggiare una birra al Number 10!

I ristoratori di via San Donato che hanno sempre fatto in modo di darci spunti di chiacchiera durante le pause per il pranzo. Mattia N. e Francesco P. ragazzi, ora Dottori, con cui ho avuto il piacere di lavorare durante il loro periodo di tesi.

Mio padre per aver tenuto in bella vista i suoi storici libri di fisica e analisi, Amaldi , Scorza Dragoni e altri che in giovane età mia hanno fatto incuriosire, consentendomi di intraprendere un percorso che non finisce oggi.

Luigi Z. con cui tante volte mi sono confrontato e che mi ha pungolato affinché, dopo il dolore, non desistessi.

Sara... perché tu sai cosa vuol dire.

Mia madre, tu oggi mio Angelo Custode, esattamente come lo sei stata da quando mi hai concepito. Ciao Mamma.

Voi miei figli, Francesco e Gaia, che avete saputo sostenermi e rimanere fantastici nonostante vi abbia sottratto molto tempo. A voi spetta cliccare sul pulsante "Deposita la tua tesi".

E ringrazio Te mia dolce Deborah: che nel silenzio hai permesso questo percorso. Che hai gioito delle mie emozioni e dei miei risultati. Che ti sei caricata tutto il peso della famiglia senza mai pretendere nulla in cambio. Che ti sei addormentata stremata. A te è sempre andato e sempre andrà il mio amore, camminando insieme, mano nella mano, nel viale di Villa delle Rose.

Bibliography

- [1] Cancer stat fact sheets: Female breast cancer, 2015.
- [2] Ductal carcinoma in situ, 2015.
- [3] Krishna Agarwal, Li Pan, Yew Kong Leong, Marcus Han, On Yu Chan, Xudong Chen, and Swee Ping Yeo. Practical applications of multiple signal classification. *International Journal of RF and Microwave Computer-Aided Engineering*, 22(3):359–369, 2012.
- [4] Enrico Angelini, Renato Campanini, Emiro Iampieri, Nico Lanconelli, Matteo Masotti, and Matteo Roffilli. Testing the performances of different image representations for mass classification in digital mammograms. *International Journal of Modern Physics C*, 17(01):113–131, 2006.
- [5] Itamar Arel, Derek C Rose, and Thomas P Karnowski. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. *Computational Intelligence Magazine, IEEE*, 5(4):13–18, 2010.
- [6] Constantine A Balanis. *Antenna theory: analysis and design*, volume 1. John Wiley & Sons, 2005.
- [7] Constantine A Balanis. *Advanced engineering electromagnetics*, volume 111. Wiley Online Library, 2012.
- [8] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.
- [9] Karen Blanchard, James A Colbert, Dhruv Puri, Joel Weissman, Beverly Moy, Daniel B Kopans, Emily M Kaine, Richard H Moore, Elkan F Halpern, Kevin S Hughes, et al. Mammographic screening: patterns of use and estimated impact on breast carcinoma survival. *Cancer*, 101(3):495–507, 2004.
- [10] NF Boyd, JW Byng, RA Jong, EK Fishell, LE Little, AB Miller, GA Lockwood, DL Tritchler, and Martin J Yaffe. Quantitative classification of mammographic densities and breast cancer

- risk: results from the canadian national breast screening study. *Journal of the National Cancer Institute*, 87(9):670–675, 1995.
- [11] Jerrold T Bushberg and John M Boone. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011.
- [12] AM Campbell and DV Land. Dielectric properties of female human breast tissue measured in vitro at 3.2 ghz. *Physics in medicine and biology*, 37(1):193, 1992.
- [13] SS Chaudhary, RK Mishra, Arvind Swarup, and Joy M Thomas. Dielectric properties of normal & malignant human breast tissues at radiowave & microwave frequencies. *Indian journal of biochemistry & biophysics*, 21(1):76, 1984.
- [14] Margaret Cheney. The linear sampling method and the music algorithm. *Inverse problems*, 17(4):591, 2001.
- [15] Jin Wook Choi, Jeiwon Cho, Yangsoo Lee, Jounghwa Yim, Byoungjoong Kang, Ki Keun Oh, Woo Hee Jung, Hee Jung Kim, Changyul Cheon, Hy-De Lee, et al. Microwave detection of metastasized breast cancer cells in the lymph node; potential application for sentinel lymphadenectomy. *Breast cancer research and treatment*, 86(2):107–115, 2004.
- [16] Comsol. *The Comsol RF user guide*, 2014.
- [17] Raymond Damadian. Tumor detection by nuclear magnetic resonance. *Science*, 171(3976):1151–1153, 1971.
- [18] Muhammad Adnan Elahi, Martin Glavin, Edward Jones, and Martin O’Halloran. Artifact removal algorithms for microwave imaging of the breast. *Progress In Electromagnetics Research*, 141:185–200, 2013.
- [19] Joann G Elmore, Mary B Barton, Victoria M Mocerri, Sarah Polk, Philip J Arena, and Suzanne W Fletcher. Ten-year risk of false positive screening mammograms and clinical breast examinations. *New England Journal of Medicine*, 338(16):1089–1096, 1998.
- [20] Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4):193–202, 1980.
- [21] Camelia Gabriel. Compilation of the dielectric properties of body tissues at rf and microwave frequencies. Technical report, DTIC Document, 1996.

- [22] Camelia Gabriel, Sami Gabriel, and E Corthout. The dielectric properties of biological tissues: I. literature survey. *Physics in medicine and biology*, 41(11):2231, 1996.
- [23] S Gabriel, RW Lau, and Camelia Gabriel. The dielectric properties of biological tissues: Ii. measurements in the frequency range 10 hz to 20 ghz. *Physics in medicine and biology*, 41(11):2251, 1996.
- [24] Sami Gabriel, RW Lau, and Camelia Gabriel. The dielectric properties of biological tissues: Iii. parametric models for the dielectric spectrum of tissues. *Physics in medicine and biology*, 41(11):2271, 1996.
- [25] Elisabetta Giacalone and Sessione. Sistemi a microonde dedicati all'imaging della mammella: prototipi sperimentali e risultati clinici.
- [26] EvaC Gombos, Jayender Jagadeesan, DanielleM Richman, and DanielF Kacher. Magnetic resonance imaging-guided breast interventions: Role in biopsy targeting and lumpectomies. *Magn Reson Imaging Clin N Am*, 23(4):547–561, 11 2015.
- [27] Fred K Gruber, Edwin A Marengo, and Anthony J Devaney. Time-reversal imaging with multiple signal classification considering multiple scattering between the targets. *The Journal of the Acoustical Society of America*, 115(6):3042–3047, 2004.
- [28] Regina J Hooley, Leslie M Scoutt, and Liane E Philpotts. Breast ultrasonography: state of the art. *Radiology*, 268(3):642–659, 2013.
- [29] Songming Hou, Knut Solna, and Hongkai Zhao. A direct imaging algorithm for extended targets. *Inverse Problems*, 22(4):1151, 2006.
- [30] Shih-Ying Huang, John M Boone, Kai Yang, Nathan J Packard, Sarah E McKenney, Nicolas D Prionas, Karen K Lindfors, and Martin J Yaffe. The characterization of breast anatomical metrics using dedicated breast ct. *Medical physics*, 38(4):2180–2191, 2011.
- [31] Jian-Ming Jin. *The finite element method in electromagnetics*. John Wiley & Sons, 2014.
- [32] William T Joines, Yang Zhang, Chenxing Li, and Randy L Jirtle. The measured electrical properties of normal and malignant human tissues from 50 to 900 mhz. *Medical physics*, 21(4):547–550, 1994.
- [33] Ian Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.
- [34] M Klemm, IJ Craddock, JA Leendertz, A Preece, and R Benjamin. Improved delay-and-sum beamforming algorithm for breast cancer detection. *International Journal of Antennas and Propagation*, 2008, 2008.

- [35] Serge Lang. *Algebra revised third edition*, volume 1. Springer Science and Media, 2002.
- [36] Paul C Lauterbur et al. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191, 1973.
- [37] Mariya Lazebnik, Leah McCartney, Dijana Popovic, Cynthia B Watkins, Mary J Lindstrom, Josephine Harter, Sarah Sewall, Anthony Magliocco, John H Booske, Michal Okoniewski, et al. A large-scale study of the ultrawideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries. *Physics in medicine and biology*, 52(10):2637, 2007.
- [38] Mariya Lazebnik, Dijana Popovic, Leah McCartney, Cynthia B Watkins, Mary J Lindstrom, Josephine Harter, Sarah Sewall, Travis Ogilvie, Anthony Magliocco, Tara M Breslin, et al. A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries. *Physics in Medicine and Biology*, 52(20):6093, 2007.
- [39] Xu Li, Shakti K Davis, Susan C Hagness, Daniel W Van der Weide, and Barry D Van Veen. Microwave imaging via space-time beamforming: Experimental investigation of tumor detection in multilayer breast phantoms. *Microwave Theory and Techniques, IEEE Transactions on*, 52(8):1856–1865, 2004.
- [40] P. J. Lynch and C. C. Jaffe. Breast normal anatomy cross-section, 2007.
- [41] Ritse M Mann. *The effectiveness of breast MRI in invasive lobular carcinoma*. [SI: sn], 2010.
- [42] Arash Maskooki, Erry Gunawan, Cheong Boon Soh, and Kay Soon Low. Frequency domain skin artifact removal method for ultra-wideband breast cancer detection. *Progress In Electromagnetics Research*, 98:299–314, 2009.
- [43] Paul M Meaney, Margaret W Fanning, Dun Li, Steven P Poplack, and Keith D Paulsen. A clinical prototype for active microwave imaging of the breast. *Microwave Theory and Techniques, IEEE Transactions on*, 48(11):1841–1853, 2000.
- [44] Paul M Meaney, Andrew P Gregory, Neil R Epstein, and Keith D Paulsen. Microwave open-ended coaxial dielectric probe: interpretation of the sensing volume re-visited. *BMC medical physics*, 14(1):3, 2014.
- [45] Mattia Neri.
- [46] Sima Noghianian, Abas Sabouni, Travis Desell, and Ali Ashtari. *Microwave Tomography*. Springer, 2014.

- [47] American College of Radiology. Breast imaging reporting and data system, 2003.
- [48] Rasmus Berg Palm. Prediction as a candidate for learning deep hierarchical models of data. *Technical University of Denmark*, 2012.
- [49] Matteo Pastorino. *Microwave imaging*, volume 208. John Wiley & Sons, 2010.
- [50] Quarteroni. Elementi finiti e metodi spettrali per equazioni iperboliche. In *Modellistica numerica per problemi differenziali*, UNITEXT, pages 241–280. Springer Milan, 2006.
- [51] Ronald J Ross, Jay S Thompson, Kyung Kim, and Ronald A Bailey. Nuclear magnetic resonance imaging and evaluation of human breast tissue: preliminary clinical trials. *Radiology*, 143(1):195–205, 1982.
- [52] Giuseppe Ruvio, Raffaele Solimene, and Max J Ammann. Evaluation of antenna types for rf breast cancer imaging using 2-layer planar tissue model. In *Microwave Conference (EuMC), 2010 European*, pages 212–215. IEEE, 2010.
- [53] Giuseppe Ruvio, Raffaele Solimene, Antonietta D’Alterio, Max J Ammann, and Rocco Pierri. Rf breast cancer detection employing a noncharacterized vivaldi antenna and a music-inspired algorithm. *International Journal of RF and Microwave Computer-Aided Engineering*, 23(5):598–609, 2013.
- [54] Berkman Sahiner, Heang-Ping Chan, Nicholas Petrick, Datong Wei, Mark Helvie, Dorit D Adler, Mitchell M Goodsitt, et al. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *Medical Imaging, IEEE Transactions on*, 15(5):598–610, 1996.
- [55] Sara M Salvador and Giuseppe Vecchi. Experimental tests of microwave breast cancer detection on phantoms. *Antennas and Propagation, IEEE Transactions on*, 57(6):1705–1712, 2009.
- [56] Ralph O Schmidt. Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions on*, 34(3):276–280, 1986.
- [57] American Cancer Society. Breast cancer detailed guide, 2015.
- [58] Raffaele Solimene, Giuseppe Ruvio, Angela Dell’Aversano, Antonio Cuccaro, Max J Ammann, and Rocco Pierri. Detecting point-like sources of unknown frequency spectra. *Progress In Electromagnetics Research B*, 50:347–364, 2013.
- [59] John Stang. A 3d active microwave imaging system for breast cancer screening. 2008.

- [60] Chen-To Tai. Dyadic green's functions in electrodynamic theory, 1971.
- [61] Rubæk Tonny. *Theory and computation of electromagnetic fields*. PhD thesis, 2007.
- [62] M Jon Turner. Stiffness and deflection analysis of complex structures. *Journal of the Aeronautical Sciences (Institute of the Aeronautical Sciences)*, 23(9), 2012.
- [63] Jeffries Wyman. The dielectric constant of mixtures of ethyl alcohol and water from-5 to 40. *Journal of the American Chemical Society*, 53(9):3292–3301, 1931.
- [64] Jianchao Yang, Kai Yu, Yihong Gong, and Tingwen Huang. Linear spatial pyramid matching using sparse coding for image classification. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1794–1801. IEEE, 2009.
- [65] Mehmet E Yavuz and Fernando L Teixeira. On the sensitivity of time-reversal imaging techniques to model perturbations. *Antennas and Propagation, IEEE Transactions on*, 56(3):834–843, 2008.
- [66] E Zastrow, SK Davis, M Lazebnik, F Kelcz, BD Van Veem, and SC Hagness. Database of 3d grid-based numerical breast phantoms for use in computational electromagnetics simulations. *Department of Electrical and Computer Engineering University of Wisconsin-Madison,[Online]*, 2008.