12-2018

A Compressed Sensing Approach to Detect Immobilized Nanoparticles Using Superparamagnetic Relaxometry

Sara Thrower

A COMPRESSED SENSING APPROACH TO DETECT IMMOBILIZED NANOPARTICLES USING SUPERPARAMAGNETIC RELAXOMETRY

by

Sara Lynn Loupot Thrower, B.S.

APPROVED:

_____

John D. Hazle, Ph.D.
Advisory Professor

_____

James Bankson, Ph.D.

_____

Robert Bast, M.D.

_____

David Thomas Alfonso Fuentes, Ph.D.

_____

Konstantin Sokolov, Ph.D.

APPROVED:

_____

Dean, The University of Texas MD Anderson Cancer Center
UTHealth Graduate School of Biomedical Sciences

A COMPRESSED SENSING APPROACH TO DETECT IMMOBILIZED NANOPARTICLES USING SUPERPARAMAGNETIC RELAXOMETRY

A

DISSERTATION

Presented to the Faculty of

The University of Texas

M. D. Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Sara Lynn Loupot Thrower

B.S, Texas A&M University, 2013

Houston, Texas

December, 2018

To my family

Judson, Phyllis, Darcy, and Scott

Acknowledgments

Abstract

Superparamagnetic relaxometry (SPMR) is an emerging technology that leverages the unique properties of biologically targeted superparamagnetic iron oxide nanoparticles to detect cancer. The use of ultra-sensitive sensors enables SPMR to detect tumors ten times smaller than current imaging methods. Reconstructing the distribution of cancer-bound nanoparticles from SPMR measurements is challenging because the inverse problem is ill posed. Current methods of source reconstruction rely on prior knowledge of the number of clusters of bound nanoparticles and their approximate locations, which is not known in clinical applications. In this work, we present a novel reconstruction algorithm based on compressed sensing methods that relies on only clinically feasible information. This approach is based on the hypothesis that the true distribution of cancer-bound nanoparticles consists of only a few highly-focal clusters around tumors and metastases, and is therefore the sparsest of all possible distributions with a similar SPMR signal. We tested this hypothesis through three specific aims. First, we calibrated the sensor locations used in the forward model to measured data, and found a 5% agreement between the forward model and the data. Next, we determined the optimal choice of the data fidelity parameter and investigated the effect of experimental factors on the reconstruction. Finally, we compared the compressed sensing-based algorithm with the current reconstruction method on SPMR measurements of phantoms. We found that when a multiple sources were reconstructed simultaneously, the compressed sensing approach was more frequently able to detect the second source. In a blinded user analysis, our compressed sensing-based reconstruction algorithm was able to correctly classify 80% of the test cases, whereas the current reconstruction method had an accuracy of 43%. Therefore, our algorithm has the potential to detect early stage tumors with higher accuracy, advancing the translation of SPMR as a clinical tool for early detection of cancer.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Superparamagnetic relaxometry (SPMR) is an emerging technology that utilizes the unique properties of biologically targeted superparamgnetic iron oxide nanoparticles (SPIOs) to detect cancer. The use of ultra-sensitive superconducting quantum interference devices (SQUIDs) with relaxation-based discrimination enables SPMR to specifically detect nanoparticles that have bound to as few as 15,000 cancer cells [1, 2]. This could mean detection of tumors 10 times smaller than current imaging methods (Figure 1.1) [3]. The localization of cell-bound nanoparticles models clusters as magnetic dipoles and requires solving the electromagnetic inverse problem, which has many possible solutions [4]. This problem is well-known to the field of magnetoencephalography (MEG), as there is still a gap in knowledge regarding how to determine the true distribution of dipole sources from the multitude of possible distributions that create nearly identical magnetic field patterns. Prior studies have circumvented this problem by restricting the distribution to a single 2-dimensional plane, limiting the number of pre-determined dipoles, or by the use of multiple excitation coils [1, 5–7]. However, in clinical applications for which a 3-dimensional distribution of an unknown number of bound nanoparticle clusters at extended depths is required, these simplifications are infeasible. Consequently, the clinical translation of this technology depends on the development of a new method of reconstructing the distribution of cancer-bound nanoparticles that is suitable for use in clinical applications.
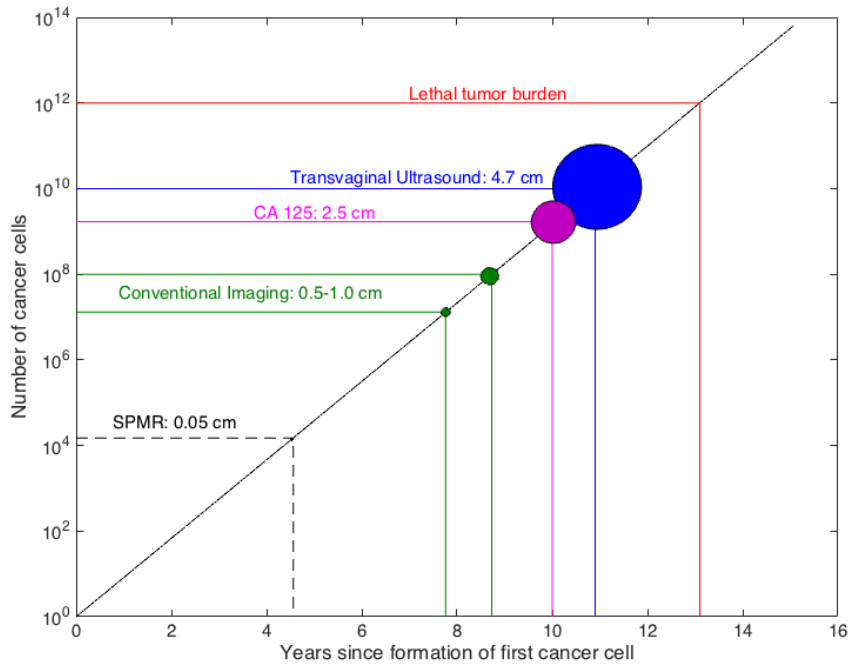
Figure 1.1: The number of cancer cells as a function of the years since the inception of the first cancer cell (black line), assuming monoexponential growth with a doubling time of 120 days. Also shown are the clinical limits of current methods of cancer detection based on tumor diameter, assuming a spherical tumor consisting of 20% tumor cells by volume with a tumor cell density of 1 million cells per microliter. The circles indicate the scale of the tumor size relative to the blue circle. No circle is shown for the lethal tumor burden because it would be larger than the graph. The circle for SPMR is shown, but is approximately the same as the line thickness.

Our-long term goal is to translate SPMR into the clinic as a tool for the early detection, diagnosis, and staging of disease. The objective of this project is to develop an algorithm that can reconstruct the distribution of cancer-bound nanoparticles in three dimensions quickly, reliably, and with only minimal prior information that is realistically available in these applications. Our central hypothesis is that the true distribution of cancer-bound nanoparticles consists of a only few highly-focal clusters around tumors and metastases, and therefore represents the most sparse distribution of all of the possible distributions. Based on this hypothesis, we have developed a reconstruction algorithm using compressed sensing methods to identify the maximally sparse distribution of magnetic

dipoles given SPMR measurements of the residual magnetic field. We expect this new algorithm will enable SPMR to provide more complete information about the location and number of cancer-bound particles than is currently possible.

1.1 Hypothesis

The hypothesis of this work is that a sparse reconstruction algorithm - based on the physics of SPMR optimized for the geometry of the device - will reconstruct distributions of up to two clusters of bound particles more reliably than the current reconstruction method without requiring prior information regarding the number or location of the clusters.

1.2 Specific aims

We plan to test this hypothesis with the following specific aims:

- **Specific Aim 1: Develop an experimentally informed forward model.** We hypothesize that physics models, in conjunction with environmental parameters such as the geometry and noise characteristics of the detector array, can predict the magnetic field from a given point source within 2% of the measured data. First, we will derive an analytical expression to describe the signal returned by the SQUIDs from a magnetic dipole distribution given the arrangement of the array of second-order gradiometer pickup coils. We will then tune the location of the pickup coils within the model to best fit measured data. The completion of this aim will provide a calibrated signal model that can then be used to reconstruct an unknown distribution of magnetic dipole sources from measurements of the magnetic field.

- **Specific Aim 2: Evaluate the accuracy of the reconstruction algorithm through simulations and measured phantom data.** A sparse reconstruction algorithm will be implemented to recover the bound particle distribution from measurements of the SPMR signal. We will then perform a series of *in silico* trials and simulations to evaluate how the reconstruction algorithm is affected by noise in the measured signal and the choice of data fidelity parameter, the number of samples and stage positions, and the extent and discretization of the field of view. This thorough characterization of the algorithm with respect to potential variables will inform future design of experimental protocols to maximize the chances of a quality reconstruction. We will also characterize the accuracy of the algorithm as a function of source strength, location, and number of sources in a study using bound nanoparticle phantoms.

- **Specific Aim 3: Compare the developed algorithm to the current state-of-the-art with respect to the ability to identify clinically-inspired distributions of two bound particle clusters.** Finally, we will compare the ability of our algorithm to the current state of the art in a blinded classification study designed to simulate the distribution of bound particles in pre-clinical small animal models with and without tumors. Based on our experience, we expect that our algorithm will correctly determine the true number of clusters in an unknown sample than the current state-of-the-art-algorithm, especially for distributions of two sources.

1.3 Dissertation organization

In Chapter 2, I will present the clinical motivation and potential applications of SPMR technology. I will then review the history and development of SPMR and describe the SPMR process. In Chapter 3 a theoretical basis of the physics and mathematics underlying the SPMR measurement and reconstruction process will be presented, followed by a

thorough description of the reconstruction algorithm. Chapter 4 presents our derivation of the forward model and presents the procedure to calibrate the model to the geometry of the sensors based on measured data. An analysis of the noise in the data and an assessment of the model-data agreement is also provided. Chapter 5 includes the *in silico* tests that characterize the numerical stability of the model as a function of signal to noise ratio, method of parameter selection, and choice of field of view and pixel size. In Chapter 6, the accuracy of the reconstructed location and strength on a set of phantom data containing up to two sources of bound particles is determined, and a blinded classification study comparing the ability of three algorithms to classify an unknown source distribution as containing zero, one, or two sources. Finally, I present the overall conclusions of the study and directions for future work.

Chapter 2

Background

2.1 Motivation

It is well known that early detection of cancer improves patient outcomes [8]. The introduction of mammography led to an estimated 15% decrease in breast cancer mortality, and endoscopic screening decreased colorectal mortality by 14% between 1975 and 2000 [9, 10]. Clinical trials suggest that screening with low dose CT may reduce lung cancer mortality rates by as much as 20% [11]. However, there are still many cancers for which there is no recommended screening method.

One potential application of SPMR is to fill the need for a highly specific and sensitive test for the early detection of ovarian cancer. In 2012, there were an estimated 239,000 new cases of ovarian cancer worldwide and an estimated 152,000 deaths from the disease [12]. In terms of deaths per new case, ovarian cancer is three times deadlier than breast cancer worldwide. Known to some as the "silent killer", ovarian cancer's danger is due to the fact that it is often found in an advanced stage after the cancer has already metastasized [13]. When diagnosed at this point, the 5-year survival rate is less than 30%, whereas when diagnosed while still locally confined the survival rates is higher than 90% [14]. The drastic difference in outcomes between early and late stage at diagnosis indicates that early detection could have a significant impact on overall mortality rate.

However, screening average-risk, asymptomatic women is currently not recommended [15]. This is largely due to evidence from two large clinical trials that current screening methods do not decrease overall mortality, or deaths from the disease. The first of these trials was the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial which enrolled over 78,000 women in the United States from 1993 to 2001 [16]. The study randomly assigned half of the women to undergo annual screening consisting of transvaginal ultrasound (TVU) and assessment of levels of the CA125 biomarker based on a cutoff value of 34 U/ml for four to six years. The other half of the women were assigned to the control arm which received usual care from their primary care provider. For women in the screening arm, results of screening tests were managed by the participant's primary care provider [17]. After 15 years of follow-up, there was no significant difference in the mortality rate between the women in the two groups [18]. Furthermore, screening did not detect more early-stage cancers than usual care. On the flip side, 3.2% of the participants in the screening group underwent unnecessary surgery, of which 15% had complications.

The second trial was the UK Collaborative Trial of Ovarian Cancer Screening (UKC-TOCS) recruited over 200,000 women in the UK between 2001 and 2005 [19]. In this trial, women were assigned to one of three study arms: annual screening with TVU only, annual multimodality screening, or a no screening control group, in a 1:1:2 ratio. Women in the multimodality screening group received annual CA125 assessment by the Risk of Ovarian Cancer Algorithm (ROCA) with TVU as a second-line test. The women in the TVU only group received annual TVU at their regional care center with TVU from a specialized care center as a second line test. Despite the improvements over the PLCO trial, the results of this study also failed to show a significant difference in overall mortality between any of the treatment groups [20]. However, this study did show a reduction in unnecessary surgeries in the multimodality arm (1%) versus the TVU only arm (3%) and the PLCO trial (3%), as well as an increase in the percentage of early stage tumors detected. This indicates that multiple levels of tests may improve the overall specificity and sensitivity of

a screening program. Without the benefit of reduced overall mortality, subjecting patients to screening can lead to unnecessary treatment and anxiety. Potential complications from surgery as well as the removal of reproductive organs without medical indication pose a risk to patients that must be considered [21–23]. Without a clear and demonstrated benefit to mortality, screening for ovarian cancer is not worth these risks or cost.

These recent advances in our understanding of the biology of ovarian cancer may explain the failure of prior screening programs to reduce overall mortality. Since the completion of these clinical trials, there has been mounting evidence that most ovarian cancers can be classified as one of two types. Type I cancers, which makeup about 20% of ovarian carcinomas, are classified as low-grade disease, which is slow growing and has low metastatic potential. However, approximately 75% of ovarian carcinomas are classified as Type II, which represent fast growing, highly malignant, tumors that are quick to metastasize [**?**]. Since these trials, studies have shown that most of the early stage tumors found with screening by TVU or CA125 are of the Type I class [24,25]. Due to the low-metastatic potential of these tumors, early detection may not confer a large benefit to survival. In contrast Type II tumors, which are responsible for the majority of ovarian cancer deaths, were more likely to be missed by conventional screening methods [26]. Increasing evidence that high grade (Type II) carcinomas actually originate in the distal end of the fallopian tube may explain why they are missed by TVU, which in the screening protocols was used to visualize the ovary [27–29]. Mathematical models and modern genetic analysis of Type II tumors suggest that this subclass of tumors is extremely quick to metastasize, and may progress to late stage while the primary tumor is only 5-10 mm in diameter [30–34]. Based on the highest levels of cellular CA125 secretion reported in the literature and the widely accepted 34 U/ml cutoff for a positive result, models predict that a tumor must reach a minimum of 1.3 cm in diameter to be detectable by the CA125 test [35]. An analysis of serial TVU tests indicates that a tumor must reach a diameter of 2.7 cm to be detectable by TVU [36]. Taken together, it is possible that the cancers had metastasized prior to
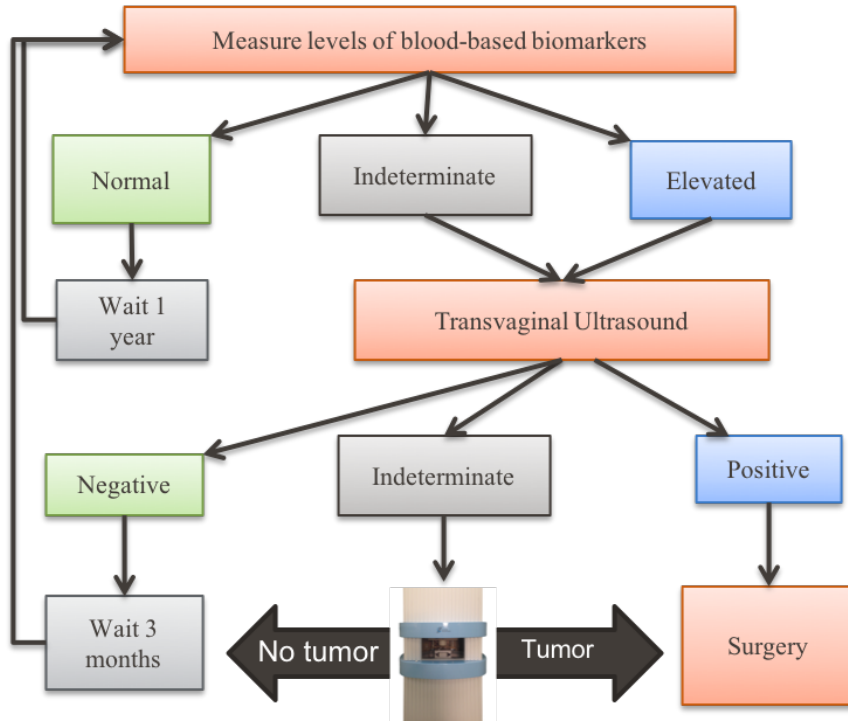
Figure 2.1: Potential role of SPMR as a third-line screen

being detectable by either method at which point surgery to remove the primary tumor is unlikely to result in a significant benefit to overall survival.

In light of these findings, perhaps detection of a small tumor volume, rather than early stage, is a more appropriate target. It is becoming clear that ovarian cancer does not follow the natural progression of other tumors which start from a precursor lesion, grow to a locally invasive tumor, and then spread to distant organs. Instead of looking for a single primary lesion, we may be better served by screening tools that are able to identify small, multi-focal lesions containing as few as 100 million cells. This level of sensitivity may be possible with an improved CA125 assay, but a lack of spatial information will limit the overall usefulness of any blood-based assay. Due to its high sensitivity, specificity, and spatial information, SPMR is an ideal candidate for this task. In light of the benefits of a multi-modal screening procedure demonstrated by the UKCTOCS trial, SPMR has been proposed as a potential third-line screening method to improve the sensitivity and specificity of the multimodality screening procedure used in the UKCTOCS trial [37]. As shown in Figure 2.1, patients

with elevated risk based on the ROCA analysis and an unsatisfactory TVU result would be referred for further analysis with SPMR. The SPMR results would then help elucidate whether the elevated ROCA result is due to normal CA125 fluctuations, or the presence of disease too small or outside the visual scope of TVU. Along with the overall tumor burden, SPMR will be able to determine whether the disease is concentrated to a single primary tumor, indicating the need for surgery, or has spread from the fallopian tube to the ovary and surrounding peritoneum, indicating the need for systemic chemotherapy or radiation.

2.2   The history of SPMR

Investigations into the unique relaxation properties of superparamagnetic materials began as far back as the 1940s. These early theoretical investigations found practical applications along side the development of superparamagnetic iron oxide nanoparticles (SPIONs). Through the 1990s, researchers developed models to link the theory of superparamagnetic relaxation to new experimental data [38–41]. The first practical application demonstrated the measurements of the magnetic relaxation process could be used to distinguish between antibody-labeled particles immobilized by interactions with antigens attached to a substrate and those unbound in solution [42]. This work led to the development of the Magnetic Relaxation Immunoassay (MARIA) [43, 44]. Others applied magnetic relaxometry to the characterization of nanoparticles in suspension and quantification of nanoparticle uptake by cells [45, 46].

The first device designed for measurements of large (potentially *in vivo*) samples was proposed in 1999, which employed a planar gradiometer pickup coil and an applied field along the x axis [47]. This device demonstrated the ability to produce spatially resolved magnetic field patterns from dried nanoparticles [48], suggesting that the technology could transition from an assay to an imaging modality. The first *in vivo* demonstrations of magnetic relaxometry measurements followed shortly, showing the ability to spatially resolve

non-specific uptake of nanoparticles in the liver of a mouse [49]. All the work until this point relied on the change in the magnetic field for analysis. No attempts had yet been made to reconstruct the magnetic moment.

The SPMR experimental setup and process used in this work was originally introduced in 2005 by Flynn, et al [1]. It was the first to solve the inverse problem using a least-squares solver to determine the strength and location of a pre-determined number of dipoles, based on the theory developed by others [50]. It has since been used for diagnosing transplant rejection using T-cells, leukemia cells from a magnetic needle biopsy, and for breast cancer detection in cells and in small animals. [51–54]. Meanwhile, work continued on MARIA, expanding it to bead-based substrates and further refining the fit of the decay curve to characterize particle dynamics such as aggregation and size distribution [45, 55].

The next phase of MRX research took on the inverse problem. Baumgarten *et al* proposed a two step approach that built on the work of Flynn, *et al* [5]. The first step used a Levenberg-Marquardt minimization of a single dipole to identify the depth of the plane that contained the distribution. The second step designed a system matrix based on the forward model calculated at set grid points across the field of view, and minimized the norm of the difference between the modeled and detected magnetic fields with Tikhonov regularization to determine the 2D distribution of particles. They studied simulations of 2D distributions of nanoparticles and phantom studies of 3D distributions and found that the strength of the source and the general location could be reconstructed, but the resolution of the physical shape and extent of the phantom was poor. Significant blurring occurred even in the noiseless simulations due to the nature of the minimum norm estimation. This approach was also tested on coil phantoms designed to simulate extended distributions of particles, rather than point sources, and ex-vivo tissue samples [56, 57]. The same group used a truncated singular value decomposition (TSVD) approach to track a moving subject with time-resolved reconstruction [58]. However, these studies also showed poor spatial resolution. Shortly after, another group proposed to overcome the ill-posedness

of the inverse problem by modifying the excitation process using gradient fields to excite only one voxel at a time [59]. This allows a direct mapping of the reconstructed magnetic moment to each voxel. The results showed 4 mm resolution of a single dipole, but required an hour long scan.

Another approach proposed replacing the single pair of Helmholtz coils with multiple excitation coils arranged on all sides of the phantom [60]. The coils can be excited individually, or in combination at different strengths, to excite a select volume of the object of interest. Both simulation studies and measurements of nanoparticle phantoms showed that this approach could produce good reconstructions of extended distributions of nanoparticles [6, 61, 62]. Optimization of the excitation coil pattern and imposing a non negativity constraint further improved on these results [63, 64]. However, the depth penetration of the field produced by a coil is proportional to its diameter. This might limit the feasibility of using small coils for human applications, where particles may be as much as 20 cm from the nearest excitation coil. Also, the voxel size used in these studies tends to be large, on the order of centimeters, with on the order of 1000 voxels per field of view. Clinical applications would require much finer mesh sizes, with more voxels per field of view. Even with multiple excitation coils, the current reconstruction methods for magnetic relaxometry could not compete with the resolution and depth penetration of up and coming modality of magnetic particle imaging (MPI) [65].

Recently, a new method for reconstruction of relaxometry measurements from a single excitation coil was proposed [7]. Along with a novel approach to pre-processing of the relaxation curve to determine the initial field values, this work built on the dipole fitting proposed by Flynn, et al. [1]. Instead of a single reconstruction with a pre-determined number of dipoles, they proposed calculating the reconstruction multiple times with a various number of pre-determined dipoles each time, and from there determining which model fit best. For each quantity of dipoles, the reconstruction was also run multiple times with different initial conditions to improve the chances that a global minimum was found.

This algorithm, along with that of Flynn, et al, is described in detail in Chapter 6.

2.3 Measurement process and work-flow

(a) (b)
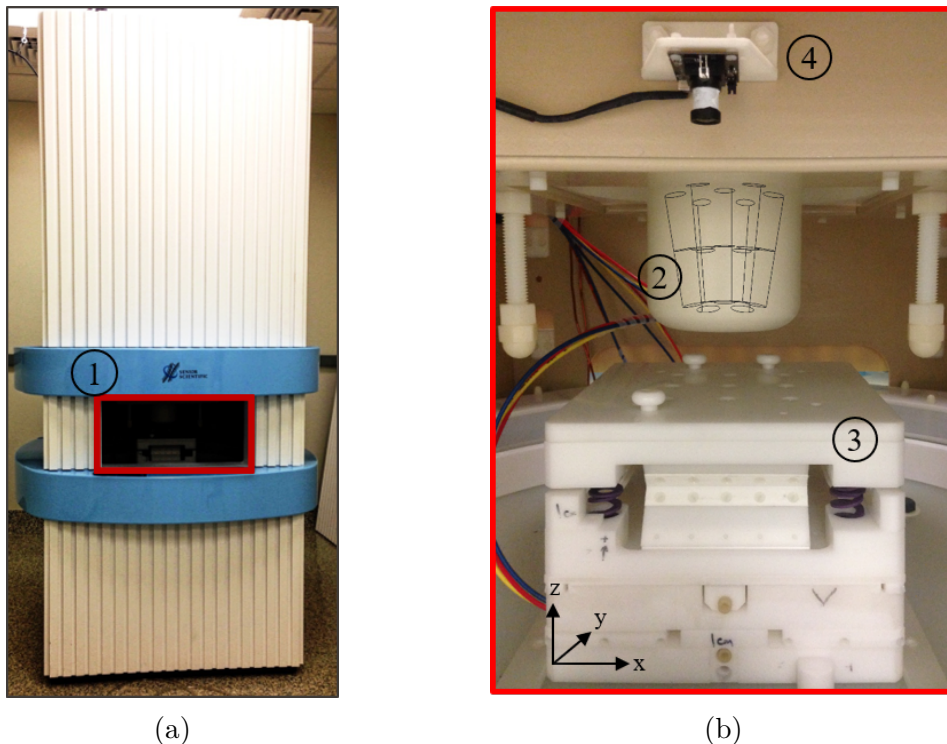
Figure 2.2: The MagSense™device, used for the superparamagnetic relaxometry (SPMR) measurements in this work. The device (a) consists of a pair of Helmholtz coils (1), which apply a homogeneous magnetic field across the region outlined in red. Part (b) shows a zoomed-in view of the region in the red box in (a). A sketch of the array of superconducting quantum interference devices (SQUIDs) that are located within the cryostatic dewar is denoted by (2). The sample is placed on a non-magnetic stage (3) which can be translated in three dimensions. A small camera (4) is included to allow for registration of the nanoparticle distribution with photographs of the subject.

The SPMR measurements in this work were conducted using the MagSense™device, shown in Figure 2.2 (Imagion Biosciences, LLC., San Diego, California, USA). It consists of a pair of Helmholtz coils (1), which apply a homogeneous magnetic field across a sample positioning system (SPS) (3) located between the coils. An array of seven second-order gradiometer coils, is located inside a cryogenic dewar above the SPS. The gradiometer

array, originally designed for use in magnetoencephalography (MEG), consists of a central gradiometer aligned parallel to the applied field surrounded by six gradiometers at a slight angle away from the central gradiometer (2). The gradiometer coils are wound in a series of three loops, each offset by 4 cm. The center loop consists of twice as many turns as the bottom and top loops, and is wound counter to the other loops, which gives the coil its second-order gradiometric features. Rather than measuring the absolute value of the flux at a point, the gradiometric design effectively measures the change in the flux as a function of distance along the gradiometer. This inherently cancels the effects of the environment by only measuring local changes near the bottom coil of the gradiometer, which leads to improved sensitivity in disturbed environments.

The gradiometric pickup coils convert the magnetic flux into electrical current, which is converted to voltage by the low-temperature SQUID (LTS) circuitry. The circuit is controlled by LabView (National Instruments, Austin, Texas) software on a computer located outside of the measurement room. The use of LTS sensors provides the sensitivity of $2 \times 10^{-14} \, \mathrm{T\,Hz^{-1/2}}$ to $5 \times 10^{-14} \, \mathrm{T\,Hz^{-1/2}}$ necessary to detect the very small magnetic fields produced by immobilized nanoparticles [1]. It also requires the sensor array to be contained in a cryogenic dewar to maintain an operating temperature of $4.2 \, \mathrm{K}$ [**?**]. This limits the minimum distance between the sensor array and the sample to approximately 2 cm. Other sensors have been investigated for use in SPMR, such as atomic magnetometers [54,66] and fluxgate magnetometers [67,68]. The voltage is recorded by software developed by Imagion Biosystems using LabWindows$^{\mathrm{TM}}$/CVI by National Instruments.

The essence of the SPMR measurement process is shown in Figure 2.3. Each measurement starts with a magnetic field applied across the sample (1) along the z axis of the system. This pulse aligns the magnetic domains of both those nanoparticles that are free to rotate (shown in blue in Figure 2.3) and those that have been immobilized, for instance by the interaction of antibodies with a targeted substrate (shown in red in Figure 2.3). After a certain amount of time, typically 1-2 seconds, the field is turned off. Within milliseconds,
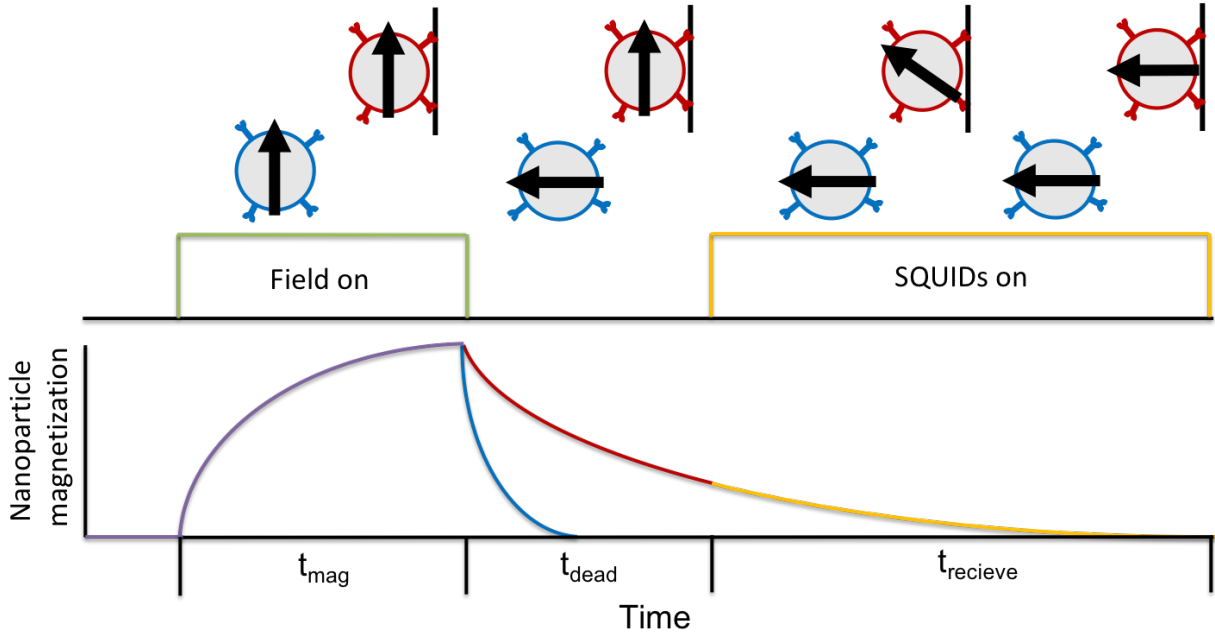
Figure 2.3: A illustrative example of the superparamagnetic relaxometry (SPMR) process. First, a magnetic field is applied for a length of time (t_{mag}), during which both the immobilized nanoparticles (red circles, top) and those that are free to rotate (blue circles, top) align their magnetic moments (arrows) with the applied field, resulting in a net magnetization effect in the direction of the applied field (bottom). During the few milliseconds t_{dead} between when the field has been removed but the superconducting quantum interference devices (SQUIDs)s have not yet been turned on, the unbound particles relax back to a zero-field configuration via Brownian motion (blue curve). The magnetization of the bound particles (red curve) decays only slightly during this time period, because they must reach zero-field configuration through the much slower Néel relaxation process. The SQUIDs are then turned on for a length of time $t_{receive}$ to record the decay of the residual magnetization of the immobilized particles (orange curve). The entire sequence lasts approximately 3 seconds.

the unbound particles have returned to a zero-field configuration (2) through physical rotation via Brownian processes [69]. Those that are unable to physically rotate return to a zero field configuration by switching the polarity of their magnetic domain through a deterministic process known as Néel relaxation [70]. After a short delay to allow for the decay of induced magnetization in the device housing and unbound particles, the SQUIDs are turned on. The SQUIDs then measure the decay of the remaining field due to only immobilized particles for a period of just over 2 seconds. The entire pulse sequence takes

a total of 3 seconds. Generally, multiple pulses are recorded and averaged to improve the signal to noise ratio.

The SPS is used to move the sample to multiple positions under the sensor array. SPMR measurements are taken at each position of the stage to effectively increase the number of sensor locations. This allows the system to resolve multiple dipole sources. In general, $4n$ sensor locations are required to resolve n dipoles with the current reconstruction algorithm. The effect of the number of stage positions on the reconstruction is investigated further in Chapter 5.
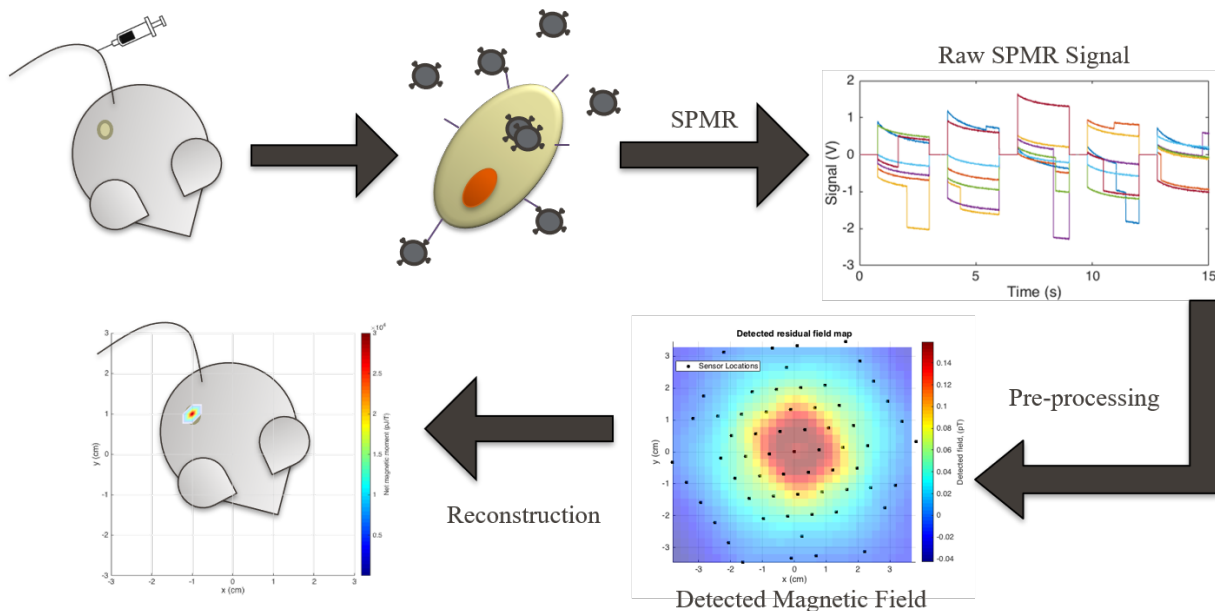
Figure 2.4: The work flow for pre-clinical SPMR measurements. First, the biologically targeted nanoparticles are injected into the tail vein or tumor bed of a mouse. Then, the particles bind to the cancer cells via interactions with the surface antigens or are taken up and aggregate inside the cell. Then SPMR measurements are recorded, and the raw data is uploaded to the server. The raw curves are then filtered and averaged, and the decay curves are fit to determine the magnitude of the residual field at each sensor location. The 3D distribution of immobilized nanoparticles is then reconstructed from the 2D residual field map, and overlaid on an image of the subject.

An example work flow for a pre-clinical SPMR experiment is depicted in Figure 2.4. First, SPIONs labeled with cancer-specific antibodies are injected into the tail vein or directly in the tumor of a tumor-bearing mouse. There the nanoparticles are immobilized,

either through interactions with cell-surface antigens or after being taken up by the cell and aggregating in the vacuoles. The subject is then measured with SPMR, and the magnetization decay curves are recorded and stored on a server. The raw decay curves are then filtered, averaged, and fit to determine the magnitude of the residual field at each of the sensor locations. The 3D distribution of the dipole moment of the immobilized particles is then reconstructed from the 2D map of the residual magnetic field, and potentially registered to photographs of the subject taken at the time of the SPMR measurement. Although reported in units of magnetic moment (pJ/T), without further calibration the reconstruction is only a relative representation of the immobilized particles. The conversion from the reconstructed relative distribution to the absolute number of immobilized particles, and then to the number of cancer cells, requires further knowledge of the specific properties of the nanoparticles, cells, and SPMR system, which is discussed in the following chapter.

Chapter 3

Theory

The Biot Savart Law describes the magnetic field as a function of distance from a magnetic dipole source. This relationship is used in SPMR to relate the signal detected by the SQUID pickup coils to the location and magnitude of the bound nanoparticles. In order to determine the location of bound particles from the detected magnetic field, the magnetic inverse problem must be solved. This problem is ill-posed: there are more unknown parameters than known. To employ state-of-the-art methods to attempt to solve this problem, the non-linear Biot Savart relationship must be made into a linear form, $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. Under some assumptions, this can be done.

The notation in this work is as follows. A simple letter (a, A) indicates a scalar value. Vectors are denoted by bolded lower case letters, or with a vector symbol $(\vec{a}, \boldsymbol{a})$. Normal vectors are labeled with a hat $(\hat{\boldsymbol{a}})$, so that $\boldsymbol{a} = a\hat{\boldsymbol{a}}$. Matrices are indicated by bold capital letters $(\mathbf{A})$. Notation is summarized in Table 3.1.

Table 3.1

Element	Example
Scalar	a
Vector	$\vec{a}, \boldsymbol{a}$
Matrix	$\mathbf{A}$

3.1 Physics of magnetic nanoparticle relaxation

Superparamagnetic iron oxide nanoparticles (SPIONs) are small, spherical particles (less than 40 nm) that behave as magnetically single-domain particles and thus exhibit superparamagnetic properties [71]. A superparamagnetic nanoparticle has a magnetic moment $\mu_p = M_s V_p$ where M_s [$\mathrm{A\,m^{-1}}$] is the saturation magnetization of the nanoparticle (typically larger than the saturation magnetization of the bulk material), and V_p [$\mathrm{m^3}$] is the volume of the particle. In this work, we approximate that a cluster of particles behaves as a single dipole with a moment equivalent to the vector sum of the moments of the nanoparticles, neglecting the effect of dipole-dipole interactions [72].

Consider a small volume containing a uniform distribution of n nanoparticles with identical core volumes as a single dipole. The magnitude of the apparent magnetization (M) of the cluster is the vector sum of the individual dipole moments of the nanoparticles each pointing in a direction $\hat{r}$, as in Equation 3.1. When all of the the particles are aligned such that $\sum_{i=1}^{n} \hat{r}_i = n$, then the net magnetization is $M = n M_s$, and the apparent magnetic moment of the cluster is $\mu = n M_s V_p$. However this is configuration is energetically unfavorable. Much like tiny magnets, in the absence of an external magnetic field a collection of non-interacting nanoparticles will tend to align their dipoles randomly to minimize the magnitude of the net magnetization density, $M = \mu / n V_p$.

$$M = M_s \sum_{i=1}^{n} \hat{r}_i \tag{3.1}$$

The particles can reorient their dipole moment through one of two processes. The first is to reverse the polarity of the dipole internally, without moving the particle itself. This process, called Néel relaxation, requires an amount of energy (anisotropy energy, E_A [J]) proportional to the magnetic anisotropy (K) and volume of the particle $E_A = K V_p$ [70]. Magnetic anisotropy, K [$\mathrm{J/m^3}$], accounts for the local variations in the shape and magnetocrystalline structure of the particles. Due to this anisotropy, the magnetic moment

of the particle will have a preferred axis of orientation within the particle, along a so-called "easy axis" at which point its energy is locally minimized [73]. For iron oxide particles with uniaxial anisotropy, the probability of randomly overcoming this barrier with thermal energy from the environment at temperature T [K] is given by the Néel relaxation time τ_N, Equation 3.2 [70]. In equation 3.2, τ_0 is the characteristic damping time constant, typically taken to be 10×10^{-10} s, and $k_B = 1.38 \times 10^{-23}\,\mathrm{J\,K^{-1}}$ is the Boltzman constant [1].

$$\tau_N = \tau_0 \exp \frac{KV_p}{k_B T} \tag{3.2}$$

The second method is through rotation of the nanoparticle itself. The Brownian relaxation time, in Equation 3.3, characterizes the rate of change in overall magnetization due to physical movement of the nanoparticles [69]. This is determined by the viscosity of the local environment, η, and the hydrodynamic diameter of the particle, V_h. The hydrodynamic diameter is a measure of the local mobility of the particle. This can be restricted by the addition of surface-bound molecules and proteins, but also interactions with other nanoparticles, substrates, or cells, that the particle has bound to.

$$\tau_B = \frac{3\eta V_h}{k_B T} \tag{3.3}$$

An effective relaxation time τ_{eff}, defined in Equation 3.4 considers the effect of both mechanisms. The value of τ_{eff} closely follows the smaller of the two components. Since τ_N is exponentially related to the particle volume, and τ_B only linearly dependent on particle volume, $\tau_{eff} \approx \tau_B$ for small particles. However, even with a small particle core, as the hydrodynamic diameter of the particle increases, possibly through interactions with other particles ($V_h \approx V_p$) or binding to cells ($V_h \to \inf$), $\tau_{eff} \approx \tau_N$.

$$\tau_{eff} = \frac{\tau_N \tau_B}{\tau_N + \tau_B} \tag{3.4}$$

The presence of an external magnetic field ($\vec{H}$) will reduce the relative energy asso-

ciated with the dipole orientation in the direction of the field and the nanoparticles will accordingly align their dipoles with the applied field. After a period of time, the effects of thermal fluctuations will reach equilibrium at the new energy level resulting in a constant net magnetization parallel to the applied field. The net magnetization at equilibrium (M_H) depends on the degree to which the energy barrier is decreased. Unless the applied field is very strong, only a portion of the dipoles will align with the field, and the net magnetization will be less than the saturation magnetization as described by Equation 3.5, where $\mu_0 = 4\pi \times 10^{-7} \text{N/A}^2$ is the magnetic permeability of free space.

$$M_H = M_s L\left(\frac{\mu_0 H M_s V_p}{k_B T}\right)$$
$$L(x) = \coth(x) - \frac{1}{x} \tag{3.5}$$

The Langevin function, $L(x)$, which goes from -1 to 1, describes the fraction of saturation magnetization that is induced by an applied field. Figure 3.1 shows the magnetized fraction M_H/M_s) of iron oxide particles with diameters of 20 nm, 25 nm, or 30 nm, and saturation magnetization density per unit mass $(M_s/\rho_{Fe_3O_4})$ of $73.9\,\text{Am}^2/\text{kg}$ by a field from 1 and 100 Gauss at 293 K [2]. From this example we can see that for an applied field of 40 Gauss, only about 70% of the 25 nm particles are magnetized.

Assuming the nanoparticle are isotropically oriented when the field is applied, the time it takes to reach equilibrium magnetization via the Néel and Brownian processes is described in Equations 3.6 and 3.7, respectively [50]. The combined effect (τ_{mag}) is still described as in Equation 3.4, calculated with $\tau_{N,mag}$ and $\tau_{B,mag}$.

$$\tau_{mag,N} = \tau_N \left(1 - \frac{0.82\mu_0 H M_s}{K}\right) \tag{3.6}$$

$$\tau_{mag,B} = \tau_B \left(1 + 0.21\left(\frac{\mu_0 H M_s V_p}{k_B T}\right)\right)^{1/2} \tag{3.7}$$
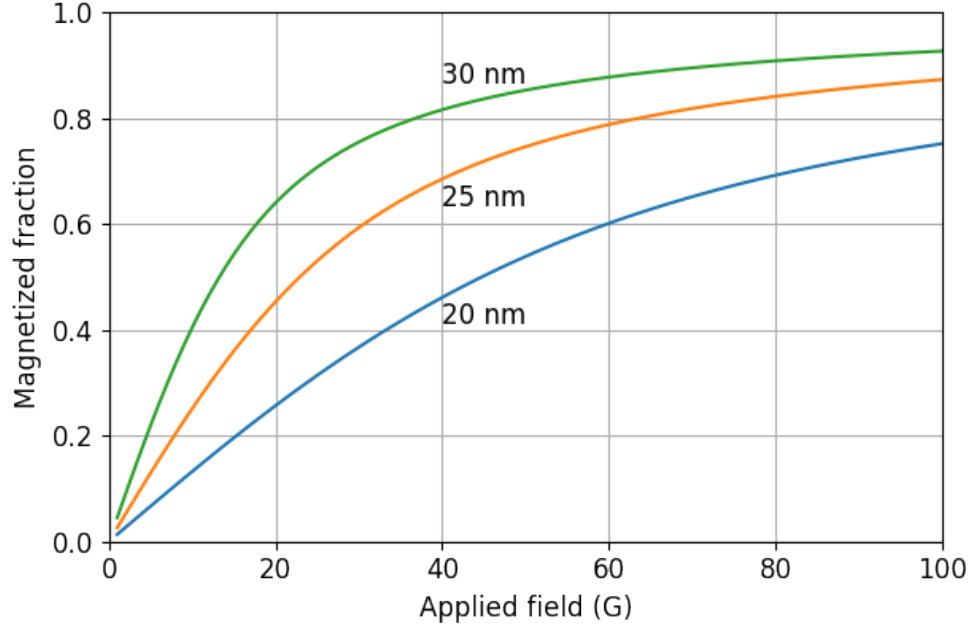
Figure 3.1: The fraction of 20 nm, 25 nm, and 30 nm diameter nanoparticles that are magnetized after an external field is applied for long enough that the system has reached equilibrium (M_H/M_s).

The magnetization induced in the nanoparticles M_i after a field $\vec{H}$ has been applied for a time t_{mag} is given by Equation 3.8, where τ_{mag} is the effective relaxation time calculated with $\tau_{N,mag}$ and $\tau_{B,mag}$ [74]:

$$M_i = M_H \left[1 - \exp\left(\frac{-t_{mag}}{\tau_{mag}(H, V_p, V_h)} \right) \right] \tag{3.8}$$

When the field is removed, the decay of the residual magnetization, M_r, can be described by Equation 3.9, where τ_{eff} is the effective decay in the absence of an applied field, Equation 3.4.

$$M_r = M_i \exp\left(\frac{-t}{\tau_{eff}(V_p, V_h)} \right) \tag{3.9}$$

The decay of the induced magnetization density after an applied field (equation 3.9) is what the SQUIDs measure. If we revisit our collection of identical nanoparticles uniformly distributed throughout a small volume, we can calculate the apparent magnetic moment

22

of the cluster through $\mu = nMV_p$, as in Equation 3.10.

$$\mu(t) = nV_pM_sL\left(\frac{\mu_0HM_sV_p}{k_BT}\right)\left[1 - \exp\left(\frac{-t_{mag}}{\tau_{mag}(H, V_p, V_h)}\right)\right]\exp\left(\frac{-t}{\tau_{eff}(V_p, V_h)}\right) \quad (3.10)$$

In reality, the nanoparticles in a sample are not identical, but will have a range of core and hydrodynamic diameters. The distributions of core diameters can typically be described by a log-normal distribution with a mean volume V_m and standard deviation σ_V [2]. The proportion of particles that have a core volume V_p is described by Equation 3.11. The hydrodynamic diameter, V_h, is effected by binding state and the properties of the local environment, but can be simplified to two general conditions: when the particle is unbound and free to move in solution and when it is immobilized ($V_h \in [V_p, \infty]$).

$$P_P(V_p, V_m, \sigma_V) = \frac{1}{\sqrt{2\pi}\sigma_V V_p}\exp\frac{-ln^2(\frac{V_p}{V_m})}{2\sigma_V^2} \quad (3.11)$$

We account for the variation of particle sizes in our physics model by integrating over the distribution of core volumes for the fraction of particles in each binding state. In a total population of n particles, consider that a fraction n_N are immobilized and therefore relax primarily through the Néel mechanism, and a fraction n_B are free in solution, and therefore relax primarily via Brownian motion, such that $n = n_N + n_B$. Equation 3.12 describes the decay over time of the magnitude of the apparent moment of a collection of bound and unbound nanoparticles with a distribution of core sizes.

$$\mu(t) = \int_{V_P} P(V_p)V_pM_sL\left(\frac{\mu_0HM_sV_p}{k_BT}\right)$$
$$\left\{n_N\left[1 - \exp\left(\frac{-t_{mag}}{\tau_{mag,N}(H, V_p)}\right)\right]\exp\left(\frac{-t}{\tau_N(V_p)}\right) + \right.$$
$$\left. n_B\left[1 - \exp\left(\frac{-t_{mag}}{\tau_{mag,B}(H, V_p)}\right)\right]\exp\left(\frac{-t}{\tau_B(V_p)}\right)\right\}dV_p \quad (3.12)$$
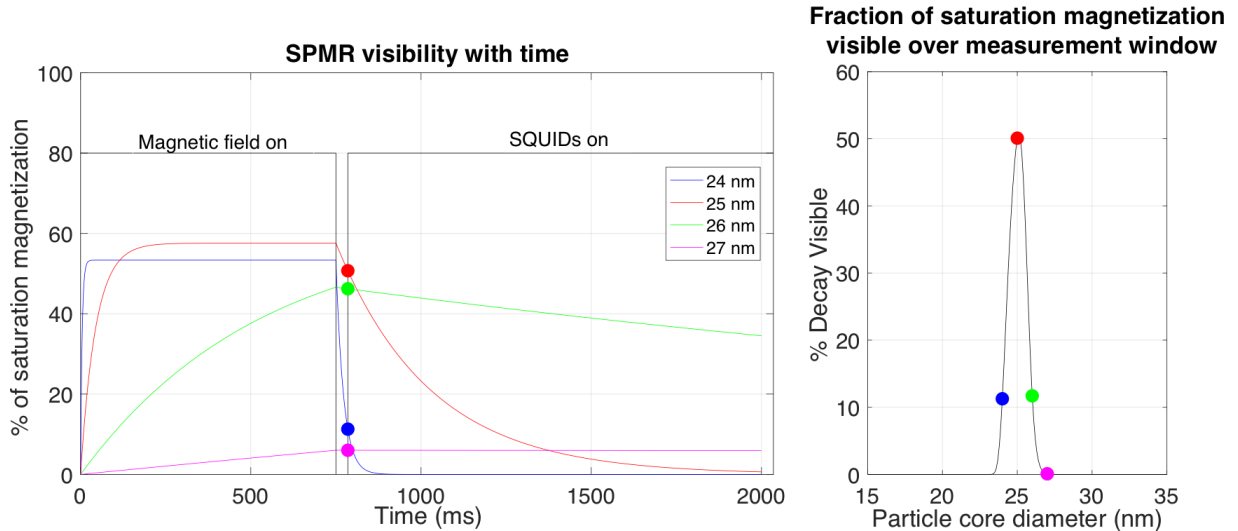
Figure 3.2: Left: The net magnetization (as a percentage of the saturation magnetization) of particles ranging from 24 nm to 27 nm in diameter during a superparamagnetic relax-ometry (SPMR) measurement sequence that consists of a 49 Gauss magnetic field applied for 750 ms, followed by a 35 ms dead time, and 2,250 ms SQUID collection window. Right: The change in residual magnetization, as a percentage of the saturation magnetization, that occurs over the SQUID measurement window. This value should be maximized for optimal SPMR results.

Note that the magnitude of the net magnetic moment of a cluster of nanoparticles (μ) is different from the inherent dipole moment of a single nanoparticle, (μ_p). While μ_p is a constant physical property of a nanoparticle, μ is the resultant vector sum of many dipole moments μ_p. It is essentially a population average based on the arrangement of the dipole moments of the individual nanoparticles, which can be altered by external forces such as temperature, surface or environmental modifications, or the application and removal of magnetic fields. The magnitude of the net magnetic moment of a cluster of particles (μ) will simply be referred to as the "moment" or "magnetic moment". However, it should be clarified that this is a population average effect due to an external force, not a physical property, and is highly dependent on many factors, as demonstrated by Equation 3.12.

The total magnetization of immobilized particles that is detectable within the SQUID measurement window is highly dependent on the diameter of the magnetic core of the particles. The time course of the net magnetization as a percentage of the saturation magnetization during an applied magnetic field for 750 ms, a dead time of 35 ms, and a

measurement period of 2,215 ms is shown in Figure 3.2 for particles with a core size of 24, 25, 26 and 27 nm, assuming only Néel relaxation. Optimal particle sizes will reach a high magnetization within the time frame of the applied magnetic field, then have minimal signal loss during the dead time before the SQUIDs are turned on. The 24 nm particles quickly reach their magnetic saturation, although it is minimal due to their small volume. However, during the dead time between when the field is turned off and the SQUIDs are turned on, the magnetization decays significantly, leaving only a small fraction of the initial magnetization visible to the sensors. What little magnetization is left decays quickly during the measurement period. The 25 nm particles take slightly longer to reach full magnetization than the 24 nm particles, but reach a higher final magnetization, and still do so well within the time frame of the applied field. When the field is removed, the particles decay more slowly lose much less of their magnetization before the SQUIDs are turned on, and almost all of their decay is contained within the measurement window. As particle size increases to 26 nm, the particles do not quite reach saturation during the applied magnetic field. They lose very little of their magnetization during the dead time, and continue to decay so slowly that only a fraction of their total magnetization decays within the measurement time window. Had the field been applied for longer, these particles would have reached a final saturation magnetization higher than the 25 nm particles, but the measurement window would have to be increased significantly in order to collect the entire decay curve. Even larger particles, 27 nm in diameter and above, hardly magnetize at all before the field is removed, and then hardly decay at all during the measurement window. Ultimately, the magnitude of the signal collected by the SQUIDs is proportional to the change in the magnetization within the measurement time frame. An ideal signal would then need to start as high as possible, lose as little magnetization as possible during the dead time, and then decay as much as possible within the SQUID measurement window. The change in magnetization (as a percentage of the saturation magnetization) within the measurement window is plotted as a function of particle diameter on the right side of Figure

25

3.2. From this we can see that particles with 25 nm are an optimal choice for a 49 Gauss field applied for 750 ms, followed by a 35 ms dead time and 2,250 ms measurement window.

Finally, we will evaluate the validity of collapsing the distribution of hydrodynamic diameters into two cases: $V_h \in [V_p, \infty]$. Clearly, this simplification is valid for very small $V_h \approx V_p$ and very large $V_h >> V_p$. However, we should evaluate how values of hydrodynamic diameter in between these cases effects the relaxation properties of the particle, to determine at what point this simplification breaks down. As hydrodynamic diameter increases, τ_B increases, and the particles begin to behave as if they are immobilized. Therefore, it is important that the particles used for SPMR have a small enough hydrodynamic diameter to not appear immobilized when they are in fact free in solution. Figure 3.3 illustrates the fraction of saturation magnetization that decays within the SQUID measurement window due to Brownian relaxation as a function of hydrodynamic diameter. This must be small to ensure that unbound particles are not contributing to the overall SPMR signal. From here we can see that the particles begin to contribute to the overall SPMR signal above hydrodynamic diameters of approximately 400 nm. This is well above the hydrodynamic diameter of the particles we use, which tends to be between 50 nm and 100 nm, depending on the surface modifications.

3.2   The magnetic forward problem

As any magnetic dipole, the cluster of particles with a net magnetic moment produces a magnetic field. The Biot Savart Law (Equation 3.13) defines the magnetic field $\vec{B}$ at a location $\vec{r} = [x, y, z]$ from a dipole $\vec{\mu}$. In SPMR, we measure the decay of the residual net magnetic moment of a cluster of particles by measuring the decay of the magnetic field it produces.

$$\vec{B}(\vec{r}, \vec{\mu}) = \frac{\mu_0}{4\pi} \left[\frac{3(\vec{\mu} \cdot \vec{r})\vec{r}}{|r|^5} - \frac{\vec{\mu}}{|r|^3} \right] \tag{3.13}$$

26
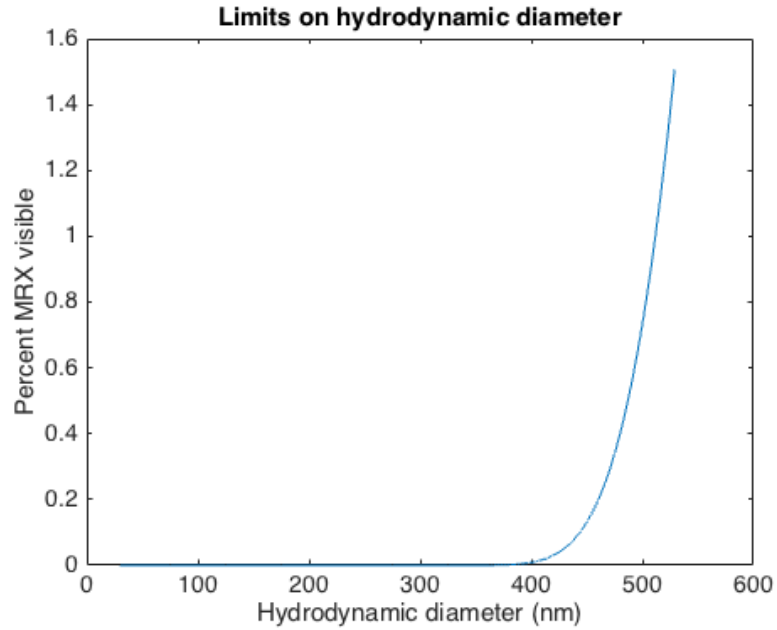
Figure 3.3: The magnetization as a percentage of the saturation magnetization that decays within the SPMR measurement window as a function of the hydrodynamic diameter of the particles. Particles with hydrodynamic diameters above 400 nm will appear immobilized regardless of binding state.
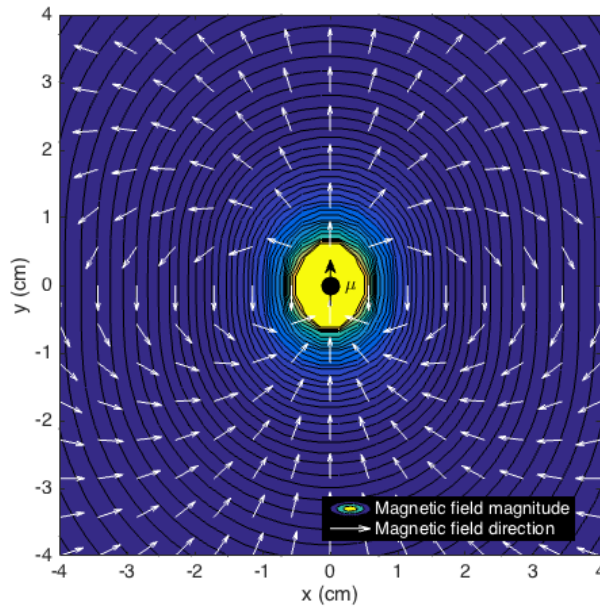
Figure 3.4: The magnetic field from a dipole source of unit strength

The magnetic flux density, or simply magnetic field, $\vec{B}(\vec{r}, \vec{\mu})$ created from a magnetized volume $\vec{\mu} = V\vec{M}(\vec{r})$ is given by the Biot Savart Law. Immediately after the field is turned
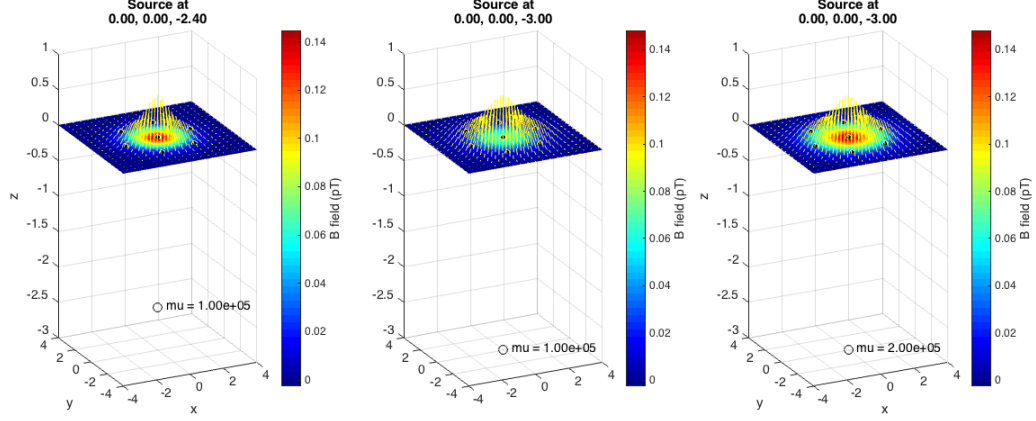
Figure 3.5: Magnetic field lines (yellow) from a magnetic dipole source (white) at the plane of the SQUIDs (black). The figures on the left and right show how the field from a weak source close to the detector plane is very similar to the field from a strong source farther away. The center is a weak source at a far distance for comparison.

off, the effective moment of a cluster of particles is parallel to the applied field, $\vec{\mu} = \mu\hat{\mu} = \mu\hat{z}$, which simplifies the dot product:

$$\vec{\mu} \cdot \vec{r} = \mu_i r_i = \mu z$$

which then allows us to factor out the magnitude of the magnetic dipole:

$$\vec{B}(\vec{r}, \vec{\mu}) = \frac{\mu_0}{4\pi} \left[\frac{3z\vec{r}}{|r|^5} - \frac{\hat{z}}{|r|^3} \right] \mu \qquad (3.14)$$

Now, $\vec{B}$ is separable into independent functions that depend on the magnitude of the magnetic moment and the distance $\vec{r}$ between the location of the dipole $(x_{dip}, y_{dip}, z_{dip})$ and the point of measurement. It is the magnetic flux (the integral of the magnetic field over the area of the coil) through the pickup coils that is converted to a voltage in the SQUID sensor and recorded out by the electronics. For simplicity, here we assert that the change in the magnetic field over the area of the pickup coil is small, and can be linearly approximated by the value of the field at a measurement point at the center of the coil

$(x_{det}, y_{det}, z_{det})$. We will revisit this approximation in Chapter 4.

$$\vec{B}(\vec{r}, \vec{\mu}) = \mu \left[\vec{a}(\vec{r})\right] \tag{3.15}$$

for

$$\vec{r} = (x_{det} - x_{dip})\hat{x} + (y_{det} - y_{dip})\hat{y} + (z_{det} - z_{dip})\hat{z} \tag{3.16}$$

where

$$\vec{a}(\vec{r}) = \frac{\mu_0}{4\pi} \left[\frac{3z\vec{r}}{r^5} - \frac{\hat{z}}{r^3}\right] \tag{3.17}$$

where

$$r = \sqrt{(x_{det} - x_{dip})^2 + (y_{det} - y_{dip})^2 + (z_{det} - z_{dip})^2} \tag{3.18}$$

$$z = \sqrt{(z_{det} - z_{dip})^2} \tag{3.19}$$

The pickup coils only detect the component of the magnetic field parallel to their axis, so the detected field is then:

$$B(\vec{r}, \mu) = \vec{B}(\vec{r}, \mu) \cdot \hat{r}_{det} = \mu \left[\vec{a}(\vec{r}) \cdot \hat{r}_{det}\right] = \mu \left[a(\vec{r})\right] \tag{3.20}$$

When multiple dipoles are present, the magnitude of the detected field (b) is the sum of the magnetic fields from each of the magnetic dipoles.

$$b(\vec{r}) = \sum_{j} B(\vec{r}_{det} - \vec{r}_j, \mu_j) \tag{3.21}$$

We then assume a discretized field of view with j voxels, each containing a magnetic moment μ_j a distance $\vec{r}_{ij}$ from sensor i. To construct the linear problem, define a vector $\boldsymbol{b} \in \mathbb{R}^{n \times 1}$ of the field detected by each sensor i (Equation 3.21) where n is the number of sensors, a vector $\boldsymbol{x} \in \mathbb{R}^{p \times 1}$ of the magnetic moment (μ) in each voxel j where p is the number of voxels, and a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$ with elements $a(\vec{r}_{ij})$ (Equation 3.17) where $\vec{r}_{ij}$ is the vector

from voxel j (dip) to detector i (det), as defined in Equations 3.16 through 3.19.

$$
\begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a(\vec{r}_{1,1}) & \cdots & a(\vec{r}_{1,p}) \\ \vdots & \ddots & \vdots \\ a(\vec{r}_{n,1}) & \cdots & a(\vec{r}_{n,p}) \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_j \\ \vdots \\ \mu_p \end{bmatrix} \qquad \Leftrightarrow \boldsymbol{b} = \mathbf{A}\boldsymbol{x} \tag{3.22}
$$

Since the location of the detectors with respect to the field of view grid points is defined, $\boldsymbol{x}$ is the only unknown. Columns of the matrix $\mathbf{A}$ may be interpreted as 'distance' vectors from the magnetic moment contained in pixel j to each sensor, whereas the rows can be interpreted as the sensitivity of a sensor to a dipole of unit strength in each voxel. To illustrate this, Figure 3.6 shows the elements of the first row of the matrix $\mathbf{A}$ corresponding to the voxels along the plane of $y = 0$ plotted corresponding to the location of the voxels they represent. The location of the center of the bottom coil of the associated gradiometer is denoted by a black dot. The values of $a_{i,j}$ have been normalized to the maximum value of the row to portray their values relative to each other.
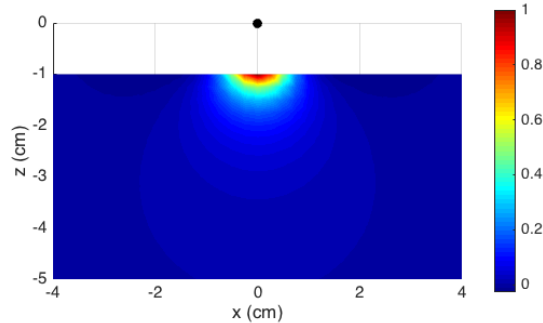
Figure 3.6: The elements of the first row of $\mathbf{A}$, $a(\vec{r})$, reshaped and plotted along $y = 0$ as a function of $\vec{r}$, demonstrating that the rows of $\mathbf{A}$ can be interpreted as a map of the sensitivity of a given detector (location denoted by a black point) to a dipole moment in any voxel. The values are normalized to the maximum value to portray their values relative to each other.

An analysis of the null space of $\mathbf{A}$ can help demonstrate the ill-posedness of the system

in Equation 3.22. According to the Rank and Nullity theorem, the dimension of a matrix is the sum of its rank and null space. The dimension of is the number of columns in a matrix, and the rank is the number of columns that are linearly independent. Naturally, the maximum rank of a matrix is the number of rows. The null space of a matrix is the set of all vectors $\boldsymbol{n}$ that satisfy Equation 3.23.

$$\mathbf{A}\boldsymbol{n} = \boldsymbol{0} \tag{3.23}$$

Typically, the null space is described as a span of a vector space N for which any linear combination of the vectors in N is a solution to Equation 3.23. The importance of this is that for a true source distribution $\boldsymbol{x}$ for which $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, we can find a second source distribution n such that $\mathbf{A}(\boldsymbol{x} + \boldsymbol{n}) = \boldsymbol{b}$. This is simply a mathematical way of saying that the detected field pattern is not unique to a single source distribution, as we have said before.

We will use a simplified example to show how the null space of the matrix $\mathbf{A}$ can identify all of the combinations of source distributions with identical magnetic field patterns. Consider a magnetic dipole with a magnetic moment μ located a distance z beneath the central sensor $(\vec{r}_{det} - \vec{r} = [0, 0, z])$. The magnitude of the magnetic field from this dipole at the center of the sensor is given by Equation 3.24.

$$
\begin{aligned}
B(\vec{z}, \mu) &= \mu \left(\vec{a}(\vec{z}) \cdot \hat{z} \right) \\
&= \mu \left(\frac{\mu_0}{4\pi} \left[\frac{3z^2}{z^5} - \frac{1}{z^3} \right] \right) \\
&= \mu \left(\frac{\mu_0}{2\pi z^3} \right)
\end{aligned} \tag{3.24}
$$

If three dipoles all lie along a line parallel to the axis of the central sensor, the linear system

in Equation 3.22 becomes

$$b = \begin{bmatrix} \dfrac{\mu_0}{2\pi z_1^3} & \dfrac{\mu_0}{2\pi z_2^3} & \dfrac{\mu_0}{2\pi z_3^3} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \tag{3.25}$$

To find the relationship between the three dipole moments required to produce a single magnetic field pattern, we set $b = 0$ and solve by getting the augmented matrix to reduced row echelon form.

$$\left[\begin{array}{ccc|c} \dfrac{\mu_0}{2\pi z_1^3} & \dfrac{\mu_0}{2\pi z_2^3} & \dfrac{\mu_0}{2\pi z_3^3} & 0 \end{array} \right] \tag{3.26}$$

$$\left[\begin{array}{ccc|c} 1 & \dfrac{z_1^3}{z_2^3} & \dfrac{z_1^3}{z_3^3} & 0 \end{array} \right] \tag{3.27}$$

The solution says that any combination of μ_1, μ_2, μ_3, and z_1, z_2, z_3, that satisfy Equation 3.28 will produce a net zero magnetic field pattern. Therefore, they can be added to any other source distribution without changing the magnetic field pattern.

$$\mu_1 + \frac{z_1^3}{z_2^3}\mu_2 + \frac{z_1^3}{z_3^3}\mu_3 = 0 \tag{3.28}$$

3.3   SPMR data acquisition and analysis

SPMR data is acquired by exposing the sample to a 50 Gauss magnetic field for 0.75 s. After a delay of 35 ms once the field is turned off, the residual magnetic field from the nanoparticles is sampled by an array of seven SQUID detectors at 1000 Hz for 2.25 s. Multiple such pulses are collected at each of several stage positions. Using several stage positions effectively increases the number of sensors available for use in the reconstruction. In this section, we will first outline the current method of preprocessing and reconstruction, known as Multi-Source Analysis (MSA). Our group has improved both the preprocessing

32

and reconstruction portions of the analysis. The majority of this work focuses on the reconstruction task, but it is always used in conjunction with our preprocessing method, which is presented in this section.

3.3.1   MSA

MSA is the current method of SPMR source reconstruction, and is described in [1]. The preprocessing algorithm first detects and corrects any flux jumps above a given threshold. Flux jumps are local discontinuities in the decay curve that would otherwise corrupt the fit of the curve, and therefore the estimation of the residual field. These jumps are an artifact caused by a saturation and subsequent resetting of the SQUIDs [77]. It then applies a filter to eliminate the 60 Hz component of the noise and its harmonics. Then, the initial value of the residual field due to immobilized nanoparticles at any given sensor immediately after the applied field is removed is determined by fitting the filtered relaxation curve with two models. First, the portion of the decay curve extending from 135 ms after the applied field is removed to the end of the measurement period is fit with Equation 3.29, based on the work of Chantrell *et al* [39], to determine the arbitrary DC offset, $A1$, which is due to the gradiometric properties of the pickup coils. Then, the first 200 ms of the curve, which correspond to 35 ms after the applied field is removed to 235 ms after the applied field is removed, is fit with Equation 3.30 - which accounts for the decay due to Néel relaxation of the bound particles, parameterized by a decay constant of 1/63 ms - to determine the parameters $Ae1$ and $Ae2$. The value t is the time since the field was removed in milliseconds. The decay constants of 2300 and 63 were determined experimentally and are assumed to be unchanged. Finally, the initial value of the residual field ($b(t = 0)$) is determined using Equation 3.31.

$$b(t) = A1 + A2 \left[ln(1 + \frac{2300}{t}) \right] \tag{3.29}$$

$$b(t) = Ae1 + Ae2 \left[e^{t/63} \right] \tag{3.30}$$

$$b(t = 0) = Ae1 + Ae2 - A1 \tag{3.31}$$

The reconstruction method uses a least-squares method to fit the location and strength of a user-determined number of dipoles to the detected field. According to the Magnetic Superposition Model [39, 50], the magnetic field at a point $\boldsymbol{s}_j$ due to a distribution of n distinct dipoles with moments m_i located at points $\boldsymbol{r}_i$ can be represented as a sum of the field induced from each dipole (f). For each dipole there are four unknown variables, $\boldsymbol{r} = [x, y, z]$ and m. Therefore, as long as the magnetic field is measured at at least $4n$ locations $\boldsymbol{b} \in \mathbb{R}^{\geq 4n}$ Equation 3.34 can be solved for the location and moment of each of the n dipoles, where f represents the Biot Savart Law in Equation 3.13.

$$b(\boldsymbol{s}) = \sum_{i=1}^{n} f(\boldsymbol{s} - \boldsymbol{r}_i, m_i) \tag{3.32}$$

$$f(\boldsymbol{r}, m) = \frac{\mu_0}{4\pi} \frac{1}{r^3} \left[3(\boldsymbol{m} \cdot \hat{\boldsymbol{r}})\hat{\boldsymbol{r}} - \boldsymbol{m} \right] \tag{3.33}$$

$$\min_{\boldsymbol{r}, m} \sum_{j=1}^{n} \left[b_j - \sum_{i=1}^{p} f(\boldsymbol{r}_j - \boldsymbol{r}_i, m_i) \right]^2 \tag{3.34}$$

Equation 3.34 is then solved using the Levenberg-Marquardt method for non-linear minimization [75], which requires an initial estimate of the location and strength of each source as a starting point. One downside of this approach is that the final result can be highly dependent on the appropriateness of the user-supplied initial conditions. If the initial conditions are far from the true solution, the algorithm may converge to a local minimum that is close to the initial conditions but far from the true solution. There is no straightforward way to determine an appropriate initial condition without prior knowledge

of the true solution, and there is no way to determine whether the result of the Levenberg-Marquardt algorithm is indeed the true solution or just a local minimum. In addition, the algorithm is limited to solving for exactly the number of user-supplied dipole moments and there is no straightforward way to determine whether the number of dipoles is an accurate representation of the true distribution.

3.3.2 MRXImage preprocessing

In this work, we used an improved method of preprocessing, that has been shown to reduce the standard deviation between repeated measurements [76]. The purpose of the preprocessing is to transform the raw decay curves into values representative of the residual magnetic field due to immobilized particles. This is done in a three step process:

- Detect and remove flux jumps

- Determine the slope of the decay curve, which is proportional to the magnitude of the residual magnetic field

- Average repeated sample measurements and subtract the average of the background measurements

The details of this process are thoroughly explained by Stefan, et al ([76]), and are briefly outlined here.

Flux jump correction

The first step is to remove flux jumps. Consider a flux jump located at time point t_j along a decay curve. If the value of the decay curve before the flux jump is $p(t < t_j) = f(t)$ then the value after the flux jump is offset by the height of the flux jump for the remainder of the curve: $p(t \geq t_j) = f(t) + u$. If the location of the jump is known, the height of the jump can be determined by simply taking the difference of the value just before and just after the jump:

$$u = p(t_j) - p(t_{j-1}) \tag{3.35}$$

To detect the location of the jump, we define a small time window just before and just after a potential jump location t_j.

$$p(t) = \begin{cases} p^+(t) & \text{if } t_j - \Delta t < t < t_j \\ p^-(t) & \text{if } t_j \leq t < t_j + \Delta t \end{cases} \tag{3.36}$$

If there is no jump, then $u = 0$ and the decay curve will fit the same functional form in both windows: $p^+(t) = p^-(t) = f(t)$. Additionally, the functions will be continuous as $t \to t_j$.

$$\lim_{t \to t_j^+} p^+(t) = \lim_{t \to t_j^-} p^-(t) = f(t_j) \tag{3.37}$$

The failure of this condition indicates the presence of a flux jump within the time window $2\Delta t$ around t_i. For each time point $t \in (\Delta t, T]$, the measured decay curve is fit within the window $t^- = t_i - \Delta t \to t_i$ and $t^+ = t_i \to t_i + \Delta t$ to a series of basis functions that approximate the expected shape of the decay curve to the find the underlying functions p^- and p^+, respectively. Then $u(t_i)$ is calculated by taking the difference of each function evaluated at t_i. The window is then shifted by one time point and the process is repeated to determine the value of u for each time point in the data such that $\boldsymbol{u} = [u(t)], \Delta t < t < T$. Simply calculating $u(t)$ produces a good estimation of the flux function if the flux jump is located at time point t. However, values of $u > 0$ also occur when the jump exists within the time window t^+ or t^-. To account for this, we consider that $\boldsymbol{u}$ can be approximated by a convolution of the true jump heights and locations v with the waveform of the Heaviside function h_i:

$$h_i = \begin{cases} 0 & i < T/2 \\ 1 & i \geq T/2 \end{cases} \tag{3.38}$$

$$\boldsymbol{u} \approx \boldsymbol{v} * (\boldsymbol{h} * \boldsymbol{h}) \tag{3.39}$$

The true jump heights $\hat{\boldsymbol{v}}$ can be solved for according to Equation 3.40, where λ is a tuning parameter that is calibrated to the data using cross validation.

$$\hat{\boldsymbol{v}} = \arg\min_{v} \left\{ \frac{1}{2} ||\boldsymbol{u} - \boldsymbol{v} * (\boldsymbol{h} * \boldsymbol{h})||_2^2 + \lambda ||\boldsymbol{v}||_1 \right\} \tag{3.40}$$

Finally, the corrupted decay curve is corrected by subtracting the corrected jump function $\boldsymbol{v}$.

Estimating the magnitude of the residual magnetic field

Once the flux jumps have been corrected, the decay curve can be analyzed to estimate the magnitude of the residual magnetic field due to immobilized nanoparticles [76]. In theory, the decay curve can be modeled by Equation 3.12. In practice however, we do not have accurate values for each of the parameters such as M_s, T, and η. In the analysis of the decay curve we assume that the unbound particles do not contribute to the measured decay, and therefore the shape of the decay curve is independent of the number of bound particles. From Equations 3.12 and 3.14, we can see that the measured decay curve ($f(t)$) is proportional to the number of bound particles times a characteristic decay $g(t)$. The inherent measurement properties of the SQUIDs adds a random DC offset c. The analysis of the decay curves aims to find a value α which is proportional to the number of bound particles and thus the detected field from the bound particles (Equation 3.41).

$$f(t) = \alpha g(t) + c \tag{3.41}$$

First, we eliminate the offset by taking the derivative with respect to time. Then we estimate α_{est} that minimizes the difference between the derivative of the measured $(f_i' = f'(t_i))$ and modeled $(g_i' = g'(t_i))$ curves over all $i = 1...T$ time points.

$$\alpha_{est} = \underset{\alpha}{\operatorname{argmin}} \frac{1}{2} \sum_{i=1}^{T} (\alpha g_i' - f_i')^2$$
$$\alpha_{est} = \frac{\sum_{i=1}^{T} f_i' g_i'}{\sum_{i=1}^{T} g_i'^2} \tag{3.42}$$

If we define weights $w_i := \frac{\sum_{i=1}^{T} g_i'}{\sum_{i=1}^{T} g_i'^2}$, then we get Equation 3.43.

$$\alpha_{est} = \sum_{i=1}^{T} f_i' w_i \tag{3.43}$$

We then estimate the derivative of the measured decay curve (f_i') at any time point i by evaluating the local change in the value of the curve around t_i. We determine f_i' by fitting the measured curve within a window from t_i to $t_i + \Delta t$ with a set of basis functions $\Psi(t)$ to account for contamination of the decay curve from $60\,\mathrm{Hz}$ power line oscillations and harmonics and other noise. The fit of the measured curve to the N_b basis functions $(p(t))$ is described in Equation 3.44.

$$p(t) = \sum_{k=1}^{N_b} a_k \Psi_k(t) \tag{3.44}$$

The weight of each basis function a_k can be found by minimizing the difference between the model and the data over $M > N_b$ time points, as in Equation 3.45.

$$a_1, \ldots, a_{N_b} = \underset{a_1,\ldots,a_{N_b}}{\operatorname{argmin}} \sum_{i=1}^{M} (f_i - p(t_i))^2 \tag{3.45}$$

Consider the $N_b = 12$ basis functions:

$$\Psi_1(t) = 1, \qquad \Psi_2(t) = t, \qquad \Psi_3(t) = t^2, \qquad \Psi_4(t) = t^3$$
$$\Psi_5(t) = \sin(60\pi t), \quad \Psi_6(t) = \cos(60\pi t), \quad \Psi_7(t) = \sin(120\pi t), \quad \Psi_8(t) = \cos(120\pi t),$$
$$\Psi_9(t) = \sin(180\pi t), \quad \Psi_{10}(t) = \cos(180\pi t), \quad \Psi_{11}(t) = \sin(240\pi t), \quad \Psi_{12}(t) = \cos(240\pi t)$$

The first four represent a third-order polynomial fit to the decay curve, and the remainder account for the first four harmonics of the 60Hz power line oscillations. The derivative of f is approximated by the second basis function, $\Psi_2(t) = t$ and the weight a_2. We collect the weights a_k and the measured decay at each time point f_i into vectors $\boldsymbol{a} := [a_1, \dots, a_{N_b}]$ and $\boldsymbol{f} := [f_1, \dots, f_T]$, and the basis functions into a matrix $\boldsymbol{\Psi} := [\Psi_1^\top(t), \dots, \Psi_{N_b}^\top(t)]$. The contribution of each basis function to the fit can be found according to 3.46, considering Equations 3.44 and 3.45.

$$\boldsymbol{\Psi} \boldsymbol{a} = \boldsymbol{f}$$
$$\boldsymbol{a} = \boldsymbol{\Psi}^\dagger \boldsymbol{f}$$

$$(3.46)$$

Recall that the derivative is associated with the second basis function. Then, $f'(0) \approx a_2 = \boldsymbol{\Psi}_2^\dagger \boldsymbol{f}$ where $\boldsymbol{\Psi}_2^\dagger$ is the second row of the inverse of $\boldsymbol{\Psi}$.

We find the derivative at time point t_i by defining $\boldsymbol{\Psi}_i := [\Psi_1^\top(\Delta_i), \dots, \Psi_{N_b}^\top(\Delta_i)]$ at the time points $\Delta_i = t_i, \dots, t_M$. We then build a convolution matrix $\mathbf{D}^{(T-M) \times T}$, where $\boldsymbol{\Psi}_{2i}^\dagger$ is the second row of the matrix $\boldsymbol{\Psi}_i^\dagger$:

$$\mathbf{D} = \begin{bmatrix} \boldsymbol{\Psi}_{2i}^\dagger & 0 & \dots & 0 \\ 0 & \boldsymbol{\Psi}_{2i}^\dagger & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \boldsymbol{\Psi}_{2i}^\dagger \end{bmatrix}$$

$$(3.47)$$

Then $\boldsymbol{f}' = \mathbf{D} \boldsymbol{f}$, or equivalently, the derivative at each time point $t_i \in i = 1, \dots, T - M$ is a convolution of the kernel of the basis function $\boldsymbol{d}$ with the decay curve: $\boldsymbol{f}' = \boldsymbol{d} * \boldsymbol{f}$. From Equation 3.43, we define a vector of weights at each time point, $\boldsymbol{w} := [w_1, \dots, w_{T-M}]$. In this work, we define the weight at time point i, $w_i := w(t_i)$, according to Equation 3.48. This has been shown to closely fit our data for parameter values of $\tau = 0.07$ and $\beta = 0.1$ [76].

$$w_i = \exp\left[\left(-\frac{t_i}{\tau}\right)^\beta\right]$$

$$(3.48)$$

We can then find α_{est} according to Equation 3.49.

$$\alpha_{est} = \boldsymbol{f}'^\top \boldsymbol{w} \qquad (3.49)$$

Data averaging, detection of outliers, and background subtraction

The preprocessing described in Section 3.3.2 is used to find a value α_{est} for each decay curve measured by each detector which is proportional to the total magnetic field at the location of that detector from the magnetic moments of the bound particle clusters. After preprocessing, the data is returned as a matrix with one row for each pulse, and one column for each detector. When multiple pulses are collected at a single detector location, they are averaged after the removal of outliers, described below.
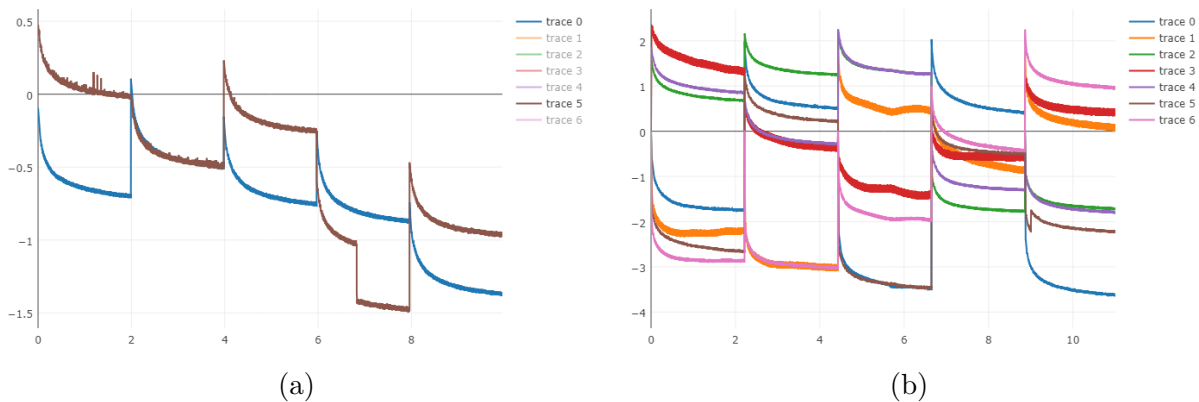
Figure 3.7: The decay curves from five measurement pulses are shown. In (a), the data acquired by an outer ring sensor (in brown) shows high frequency contamination in the last half of the first two pulses and the first part of the third pulse. In the same acquisition, the data collected by the central sensor (in blue) showed no high frequency contamination. In (b), the third pulse in particular shows the manifestation of low frequency contamination, which affects the data from every sensor to varying degrees.

In the course of recording SPMR data, it is possible that the data is corrupted by outside interference. This interference manifests in the data as either a high or low frequency component. High frequency interference manifests as small regions of spikes in the decay curve, as shown in Figure 3.7a. It is likely due to radio frequency interference from sources such as hand held radios and cell phones. Low frequency interference causes the decay curve

to be distorted, as shown in Figure 3.7b. This type of contamination is due to changes in the local magnetic field, potentially due to movement of a large ferromagnetic object (such as a metal cart or wrench) nearby. In the current preprocessing and reconstruction method (MSA), the user must manually identify and exclude pulses that exhibit features that indicate the pulse has been contaminated. In this work, we recognize that the contamination of the pulse only matters insofar as it affects the estimation of α_{est}. Therefore, all of the measured pulses are preprocessed according to Sections 3.3.2 and 3.3.2 without manual identification of bad pulses. Then, all the values of α_{est} at a given sensor location are averaged, and any values of α_{est} outside of two standard deviations of the mean are excluded. Then the remaining values of α_{est} are averaged to obtain the final estimate of the initial field value at that sensor location, which we will call b. The values at each sensor location are then arranged in a vector $\boldsymbol{b} = [b_1, \ldots, b_M]$, which is input into the reconstruction algorithm to recover the distribution of bound particle clusters.

3.4 The MRXImage Reconstruction Algorithm

It would seem that the solution to the linear system $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ derived in Section 3.2 is trivial, with a simple linear inverse solve $\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}$. However, since $n << p$ the problem is ill-posed. This means there are an infinite number of exact solutions. For example, a simple inverse solution may result in a solution in which there is a source under each detector proportional to the strength of the signal received from that detector. Clearly this is a feasible solution, but it is unlikely to be the true solution. There have been several approaches to SPMR reconstruction developed in the past. The simplest approach is MSA, described above, in which the number of nanoparticle clusters is defined, and then a Levenberg-Marquardt algorithm is used to find the strength and location of each source given the detected field [1]. However, this is either tedious or not feasible when the number of bound particle clusters is unknown. Expanding on this method, the multi-start spatio-

temporal algorithm (MSST), solves the least-squares problem using multiple different initial conditions, and repeats this for a range of quantities of bound particle clusters, typically 1 to 5 [7]. Although computationally more time consuming, this method can reconstruct multiple dipole distributions without prior knowledge of the number of clusters to find. More information and some results from MSST are included in the Appendices of this work. Here we present a new approach that can perform the reconstruction in a timely manner and without prior knowledge of the number of particle clusters or their approximate locations.

3.4.1 l_1 norm minimization

In the MRX application, we can assume that the true solution will be sparse, meaning only a few voxels will have a non-zero contribution to the net moment. Therefore, we hypothesized that a compressed sensing approach such as the sparsity averaging reweighting algorithm (SARA) proposed by Carrillo et. al. [78]. We can use this approach to find the minimum number of sources that still solve Equation 3.22 exactly. The l_p norm can be used to define sparsity with parameter p.

$$l_p(\vec{x}) = \left[\sum_i x_i^p \right]^{1/p} \tag{3.50}$$

When p=2, the l_p norm is the magnitude of the vector $\boldsymbol{x}$. When p=1, l_p is the sum of the components of $\boldsymbol{x}$. When p=0, l_p is the number of non-zero entities of $\boldsymbol{x}$. Ideally, to find the sparsest solution that satisfies the equality constraint, it would be best to minimize the l_0 norm. Since the l_0 norm is not convex, it is difficult to compute explicitly. Instead, we must maximize sparsity by minimizing the L1 norm of $\boldsymbol{x}$ while applying the condition that $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. In some cases, it can be shown that the minimum L1 norm solution is also the sparsest solution [79].

3.4.2 Mathematical consideration of noise

Recall that we model the magnetic field from a distribution of dipole sources as a linear system, which can be denoted $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}^*$. In reality, the detected signal is the sum of the field induced by the magnetic moment of immobilized particles, plus some uncertainty in the measured field which is approximately normally distributed with zero mean $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2)$.

$$\boldsymbol{b} = \boldsymbol{b}^* + \boldsymbol{\epsilon} \tag{3.51}$$

This uncertainty term is the sum of several sources including spurious signals in the electronics, magnetic fields in the environment (such as radios or cell phones), and uncertainty added during the background subtraction, smoothing, and curve fitting described in Section 3.3. In the MRXImage package, this uncertainty is handled in the tolerance term of the minimization criteria:

$$||\mathbf{A}\boldsymbol{x} - \boldsymbol{b}|| \leq ||\boldsymbol{\epsilon}|| = \lambda \tag{3.52}$$

The choice of the value for the tolerance parameter λ is discussed in more depth in Section 5.1.

3.4.3 CVX

The core of the MRXImage algorithm solves the optimization problem described by Equation 3.53. In order to incorporate both the positivity constraint and the limit on the uncertainty in the detected field we can use a convex optimization framework, which is implemented with the CVX package for MATLAB [80, 81].

$$\min_{x} ||\boldsymbol{x}||_1 \quad \text{such that} \quad \begin{cases} ||\mathbf{A}\boldsymbol{x} - \boldsymbol{b}|| \leq \lambda \\ \\ \boldsymbol{x} \geq 0 \end{cases} \tag{3.53}$$

CVX is a wrapper for translating which allows the user to describe the optimization in simple MATLAB syntax, and then translates the system into the appropriate form for the solvers according to the disciplined convex programming rule set [82]. The user can select from either the freely available SeDuMi or SDPT3 optimization algorithms, or the commercial versions MOSEK and Gurobi if the appropriate licenses are obtained [83–86]. The MATLAB implementation of Equation 3.53 using CVX is presented below.

```
cvx_begin

    variable x(n);

    minimize(norm(x,1));

    subject to

        x >=0;

        norm(A*x-b)<= lambda;

cvx_end
```

Figure 3.8 shows a geometric interpretation of Equation 3.53 for a simplified vector $\boldsymbol{x} \in \mathbb{R}^2$ with one component representing the signal and one component the noise. In Figure 3.8a, when $\lambda = 0$, the reconstruction exactly matches the measured signal, but the noise component of the reconstruction (n_r) is large. As λ increases (Figure 3.8b), both components of the reconstruction and the distance e_r between the measured signal (s_m) and reconstruction (x_r) decreases. For some value of λ (Figure 3.8c), the error (e_r) reaches a minimum. As n_r goes to zero (Figure 3.8d), the decrease in $||x_r||_1$ is due to decreasing s_r. In Figure 3.8e, the noise component of the reconstruction (n_r) goes to zero, and the sparsity is maximized. Finally, in Figure 3.8f, the magnitude of the remaining component decreases and reaches zero when $\lambda = ||x_m||$.
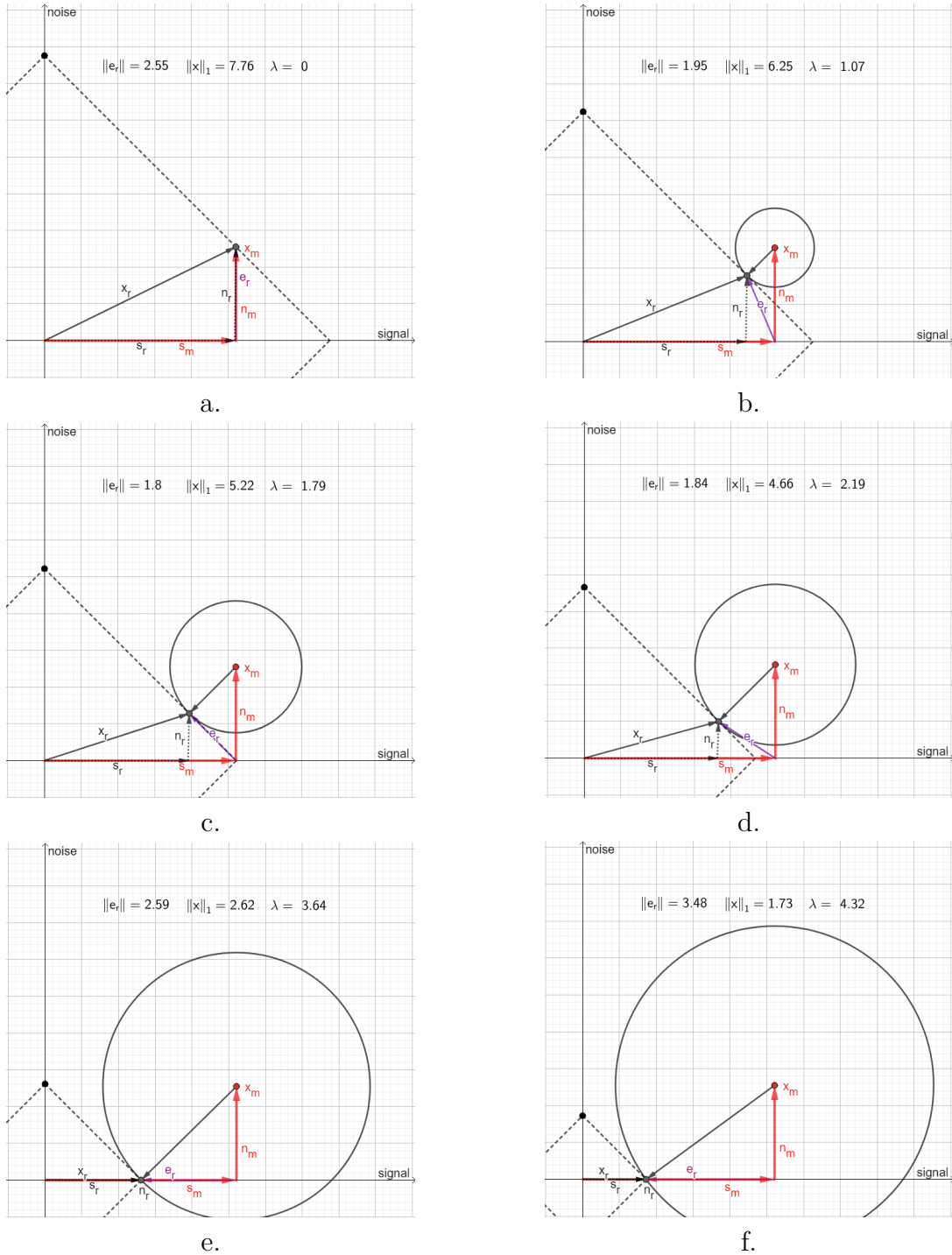
Figure 3.8: An illustration of the optimization function in two dimensions. The measured vector x_m and the reconstructed vector x_r are simplified to two components representing signal (s_m and s_r) and noise (n_m and n_r). The vector e_r is the distance between the measured signal (s_m) and the reconstruction (x_r). The function $||x||_1 = c$ is depicted by a dotted line where c is the minimum value within the constraint $||\mathbf{A}x_r - \mathbf{A}x_m|| < \lambda$ depicted by a circle. For this example, $\mathbf{A} = [1, 1]$.

Within CVX, Equation 3.53 is translated into the standard form of a primal second order cone problem with $p + 1$ inequality constraints [87, 88]. Here, we present the log-barrier method, which is one approach this optimization problem. Equation 3.53 can be written as

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} \quad & \frac{1}{2}||\mathbf{A}\boldsymbol{x} - \boldsymbol{b}||^2 - \epsilon \leq 0 \\ & -x_i \leq 0 \quad i = 1...p \end{aligned} \tag{3.54}$$

where $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{c} = \mathbf{1}^{p \times 1}$, $\mathbf{A} \in \mathbb{R}^{n \times p}$ is the system matrix, and $\boldsymbol{b} \in \mathbb{R}^{n \times 1}$ is the detected field at each sensor. For simplicity, we denote the left hand side of the first constraint as f_ϵ, and the positivity constraints as $f_{+,i}$, along with the gradient and Hessian in Equation 3.55, wherein $\boldsymbol{e}_i$ is the elementary vector corresponding to element i.

$$\begin{aligned} f_\epsilon &:= \tfrac{1}{2}||\mathbf{A}\boldsymbol{x} - \boldsymbol{b}||^2 - \epsilon & f_{+,i} &:= -x_i \\ \nabla f_\epsilon &= \mathbf{A}^T(\mathbf{A}\boldsymbol{x} - \boldsymbol{b}) & \nabla f_{+,i} &= -\boldsymbol{e}_i \\ \nabla^2 f_\epsilon &= \mathbf{A}^T \mathbf{A} & \nabla^2 f_{+,i} &= 0 \end{aligned} \tag{3.55}$$

A solution $\boldsymbol{x}^*(\tau)$ of Equation 3.54 is optimal if it satisfies the Karush-Kuhn-Tucker (KKT) conditions, in Equation 3.56.

$$\begin{aligned} \nabla f_o + \nu_\epsilon \nabla f_\epsilon + \sum_{i=1}^{p} \nu_{+,i} \nabla f_{+,i} &= 0 \\ f_\epsilon &\leq 0 \\ f_{+,i} &\leq 0, \quad i = 1 \dots p \\ \nu_\epsilon f_\epsilon 0 &= 0 \\ \nu_{+,i} f_{+,i} &= 0 \\ \nu_\epsilon &\geq 0 \\ \nu_{+,i} &\geq 0 \end{aligned} \tag{3.56}$$

The inequality constraints can be incorporated into the objective function through the log-

barrier function. The log-barrier function enforces the constraints by becoming large when the solution is approaching the boundary defined by the constraints. As $\tau \to \infty$, Equation 3.57 more closely approximates Equation 3.54.

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \boldsymbol{c}^T \boldsymbol{x} + \frac{1}{\tau} \left[-\log(-f_\epsilon) + \sum_i -\log(-f_{+,i}) \right] \tag{3.57}$$

Denoting the objective function as $f_o \in \mathbb{R}$, the gradient ($\nabla f_o \in \mathbb{R}^p$) and Hessian ($\nabla^2 f_o \in \mathbb{R}^{p \times p}$) of f_o are defined in Equation 3.58.

$$
\begin{aligned}
f_o &:= \boldsymbol{c}^T \boldsymbol{x} + \frac{1}{\tau} \left[-\log(-f_\epsilon) + \sum_i -\log(-f_{+,i}) \right] \\
\nabla f_o &= \boldsymbol{c} + \frac{1}{\tau} \left[-f_\epsilon^{-1} \nabla f_\epsilon - \sum_i f_{+,i}^{-1} \nabla f_{+,i} \right] \\
\nabla^2 f_o &= \frac{1}{\tau} \left[f_\epsilon^{-2} \nabla f_\epsilon (\nabla f_\epsilon)^T - f_\epsilon^{-1} \nabla^2 f_\epsilon + \sum_i f_{+,i}^{-2} \nabla f_{+,i} (\nabla f_{+,i})^T \right]
\end{aligned}
\tag{3.58}
$$

To simplify the Hessian in Equation 3.58, define $\mathbf{F}_*$ as a diagonal matrix with the elements of f_* along the diagonal. Furthermore, let $\mathbf{F}_*^{-1}$ be a diagonal matrix with elements of f_*^{-1} along the diagonal. Also define the residual between the modeled and the detected field $\boldsymbol{r} = \mathbf{A}\boldsymbol{x} - \boldsymbol{b}$. The Hessian in Equation 3.58 can then be written as

$$\nabla^2 f_o = \frac{1}{\tau} \left[\mathbf{F}_\epsilon^{-2} \mathbf{A}^T \boldsymbol{r} \boldsymbol{r}^T \mathbf{A} - \mathbf{F}_\epsilon^{-1} \mathbf{A}^T \mathbf{A} + \mathbf{F}_+^{-2} \right] \tag{3.59}$$

Starting from a feasible solution $\boldsymbol{x}$, which is one that satisfies the inequality constraints in Equation 3.55, we seek a Newton step in the direction $\Delta \boldsymbol{x}$ towards the minimum of the objective function. Given that the gradient of the objective function is zero at its minimum, we can move towards the optimal value of $\boldsymbol{x}^*(\tau)$ by taking steps $\boldsymbol{x}^{k+1}(\tau) = \boldsymbol{x}^k(\tau) + \alpha \Delta \boldsymbol{x}$, where $0 < \alpha < 1$ and $\Delta \boldsymbol{x}$ satisfies Equation 3.60.

$$\nabla^2 f_o(\boldsymbol{x}) \Delta \boldsymbol{x} = -\nabla f_o(\boldsymbol{x}) \tag{3.60}$$

After each step k, the solution is evaluated for optimality. One metric used to evaluate

the optimality of the solution is the Newton decrement $(\lambda_N(\boldsymbol{x}))$, defined in Equation 3.61. When the Newton decrement is less than a pre-defined tolerance on the optimality of $\boldsymbol{x}^*$, then $\boldsymbol{x}^*(\tau) := \boldsymbol{x}^k$ is returned.

$$\lambda_N(\boldsymbol{x}) = \sqrt{\nabla f_o(\boldsymbol{x})^T \nabla^2 f_o(\boldsymbol{x})^{-1} \nabla f_o(\boldsymbol{x})} \tag{3.61}$$

When the Newton iterations have converged to $\boldsymbol{x}^*(\tau)$, then τ is increased by a factor $\mu > 1$, and Newton's method is used to find the next value of $\boldsymbol{x}^*(\mu\tau)$, starting from the point $\boldsymbol{x}^*(\tau)$. This is repeated until τ is large enough that Equation 3.57 adequately approximates the original objective.

Defining the optimality of $f(\boldsymbol{x}^*(\tau))$

Through an analysis of the Lagrangian of Equation 3.53, we can show that the difference between the optimal value of the log-barrier objective function at any value of τ $(f_0(\boldsymbol{x}^*(\tau)))$ and the optimal value of the original objective function $p^* = f(\boldsymbol{x}^*)$ is simply n/τ, where n is the number of inequality constraints [87]. First, we restate the optimization problem 3.57 in a more general form.

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & f_i(\boldsymbol{x}) \leq 0 \quad i = 1...n \end{aligned} \tag{3.62}$$

The Lagrangian of Equation 3.62, with Lagrangian multipliers $\nu_i > 0$ is given in Equation 3.63.

$$L(\boldsymbol{x}, \nu) = f_0(\boldsymbol{x}) + \sum_{i=1}^{n} \nu_i f_i(\boldsymbol{x}) \tag{3.63}$$

The Lagrange dual function is defined as a function of ν for which the gradient of the Lagrangian with respect to $\boldsymbol{x}$ is zero, and is a lower bound on the objective function of

Equation 3.62.

$$g(\nu) = \inf_x L(\boldsymbol{x}, \nu) = \{L(\boldsymbol{x}, \nu) | \nabla_x L(\boldsymbol{x}, \nu) = 0\} \tag{3.64}$$

$$\nabla_x L(\boldsymbol{x}, \nu) = \nabla f_0(\boldsymbol{x}) + \sum_{i=1}^n \nu_i \nabla f_i(\boldsymbol{x}) \tag{3.65}$$

The difference between $f_0(\boldsymbol{x})$ and $g(\nu)$ is called the duality gap. When the duality gap is zero, $f_0(\boldsymbol{x}) = p^*$ is a global minimum of Equation 3.62.

We can also cast the system 3.62 in the log-barrier form in terms of the log-barrier function $\phi(\boldsymbol{x}) = -\sum_{i=1}^n \log(-f_i(\boldsymbol{x}))$.

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad f_0(\boldsymbol{x}) + \frac{1}{\tau}\phi(\boldsymbol{x}) \tag{3.66}$$

The value $\boldsymbol{x} = \boldsymbol{x}^*(\tau)$ that optimizes Equation 3.66 satisfies Equation 3.67.

$$\begin{aligned} 0 &= \nabla f_0(\boldsymbol{x}) + \frac{1}{\tau}\nabla\phi(\boldsymbol{x}) \\ &= \nabla f_0(\boldsymbol{x}) + \sum_{i=1}^n \frac{-1}{\tau f_i(\boldsymbol{x})}\nabla f_i(\boldsymbol{x}) \end{aligned} \tag{3.67}$$

Setting $\nu_i(\tau) = \frac{-1}{\tau f_i(\boldsymbol{x}^*(\tau))}$, Equation 3.67 satisfies Equation 3.65 for $x = x^*(\tau)$. Given that the value of the primal objective function is bounded below at all points by $g(\nu)$, then the duality gap can be determined by Equation 3.68.

$$\begin{aligned} f_0(\boldsymbol{x}^*) &\geq g(\nu(\tau)) \\ &= L(\boldsymbol{x}^*(\tau), \nu(\tau)) \\ &= f_0(\boldsymbol{x}^*(\tau)) + \sum_{i=1}^n \frac{-1}{\tau f_i(\boldsymbol{x}^*(\tau))} f_i(\boldsymbol{x}^*(\tau)) \;\; = f_0(\boldsymbol{x}^*(\tau)) - \frac{n}{\tau} \end{aligned} \tag{3.68}$$

Equation 3.68, demonstrates that the optimal value of the objective function in Equation 3.57 for any value of $\tau > 0$ is at most n/τ suboptimal, and as $\tau \to \infty$, $f_0(\boldsymbol{x}^*(\tau)) \to f_0(\boldsymbol{x}^*(\tau))$. This relationship is extremely valuable for defining stopping criteria for the log-barrier method of convex optimization. For a solution that is within a desired tolerance ϵ of the true global minimum, simply choose to stop when $\tau \leq n/\epsilon$. Evaluating the optimality

conditions for the solution at $\boldsymbol{x}^*(\tau), \nu(\tau)$ results in the modified KKT conditions (Equations 3.69 through 3.72). The only difference between the modified KKT conditions (for inequality constrained problems) and the standard KKT conditions for equality constrained problems is in the complementary slackness condition. For the standard KKT conditions, we have $\nu_i f_i(x) = 0$. In our previous analysis, we found that at $\boldsymbol{x}^*(\tau)$, $\nu_i = -1/\tau f_i(\boldsymbol{x}^*(\tau))$. Rearranging, we get the modified complementary slackness condition in Equation 3.71.

$$f_i(\boldsymbol{x}) \leq 0 \tag{3.69}$$

$$\nu_i \geq 0 \tag{3.70}$$

$$-\nu_i f_i(\boldsymbol{x}) = \frac{1}{\tau} \tag{3.71}$$

$$\nabla f_0(\boldsymbol{x}) + \sum_{i=1}^{n} \nu_i \nabla f_i(\boldsymbol{x}) = 0 \tag{3.72}$$

The primal feasibility condition, Equation 3.69, comes from the fact that $\log(-f_i(x))$ is undefined for $f_i(x) > 0$. Equation 3.70 follows from 3.69 and the choice of $\tau > 0$. Equations 3.71 and 3.72 come from the definition of $\nu(\tau)$ from Equation 3.67. Because this problem is convex, the modified KKT conditions are necessary and sufficient to show that for a given value of τ, $f_0(\boldsymbol{x}^*(\tau))$ is a unique minimum of Equation 3.66. As $\tau \to \infty$, $f_0(\boldsymbol{x}^*(\tau)) \to f_0(\boldsymbol{x}^*)$, where $\boldsymbol{x}^*$ minimizes Equation 3.62. This is apparent in the modified slackness condition. As $\tau \to \infty$, $1/\tau \to 0$, and the modified KKT conditions more closely approximate the standard KKT conditions.

The solvers implemented by CVX seek to satisfy Equation 3.53 in as few operations as possible. Each performs this optimization in a slightly different way, for instance changing the step size selection, applying pre-conditioning matrices, or storing factorizations of the system to calculate the Newton step, each with the goal reach the most accurate solution in the fewest number of iterations. A detailed account of the differences between the approaches of the solvers available in CVX is given by Cai and Toh in memorial of Jos Sturm, the author of SeDuMi [89]. Because the system is convex, any solver will progress

toward the same global minimum. The choice of solver will only affect how close to the true minimum the optimization converges and how quickly it gets there [89]. We have tried both of the two free solvers in CVX, and found this to be true. However, due to ongoing open-source development of SeDuMi as part of the CVX package, we chose to use it in this work, and recommend using it for future work.

3.4.4 Bias correction

The matrix $\boldsymbol{A}$ in equation 3.22 is biased towards solutions nearest the detectors. The bias correction of [90] is applied to correct for this. The first row of the matrix before and after the bias correction is applied is shown in Figure 3.9. Define a weighting matrix, $\boldsymbol{W}$

$$\boldsymbol{W} = \mathrm{diag}(\|a_j\|^{-1}) \tag{3.73}$$

where a_j are the columns of $\boldsymbol{A}$. Then,

$$\boldsymbol{A}\vec{x} = \vec{b} \rightarrow \boldsymbol{A}\boldsymbol{W}\boldsymbol{W}^{-1}x = b \rightarrow \hat{\boldsymbol{A}}\hat{x} = b \tag{3.74}$$

where $\hat{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{W}$ is now column unit normalized and $\hat{x} = \boldsymbol{W}^{-1}x$. Equation 3.74 is put into the algorithms presented in the subsequent section to solve for $\hat{x}$, then the result is corrected to recover x.

$$x = \boldsymbol{W}\hat{x} \tag{3.75}$$
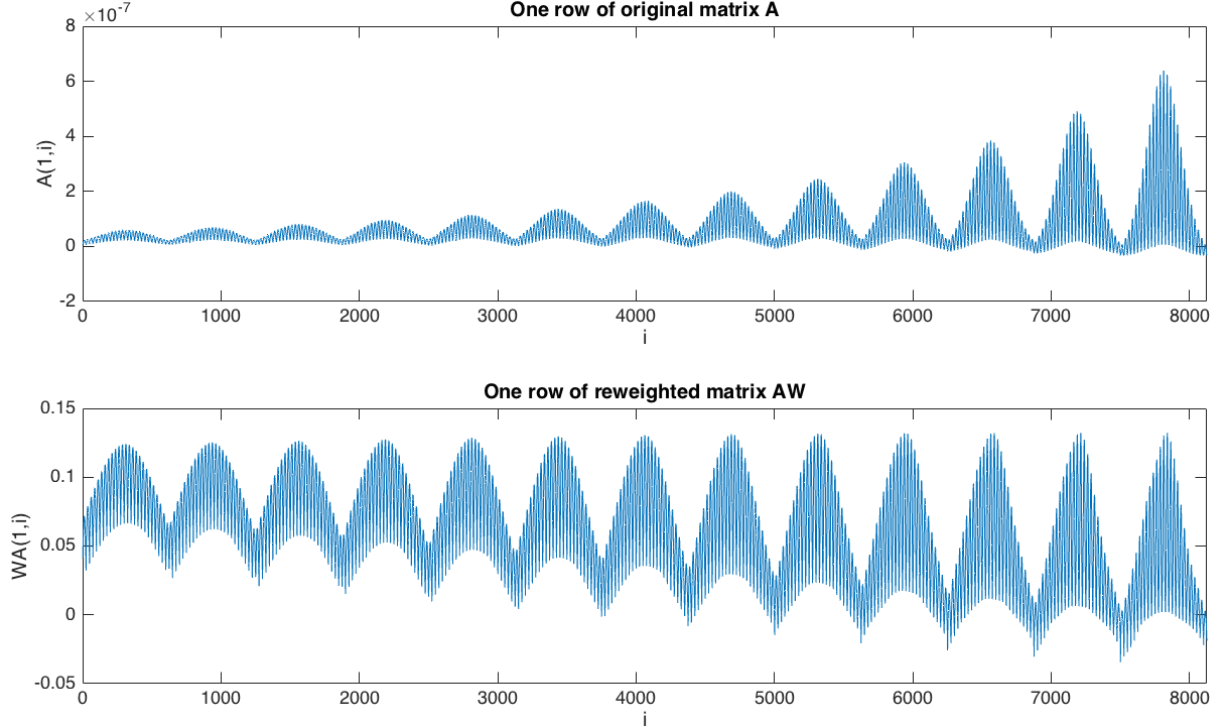
Figure 3.9: The first row of the system matrix, before and after applying the column-norm bias correction. The voxel index is listed along the x-axis starting from a corner of the field of view on the furthest plane from the sensors, continuing along each row and column of that plane, and then progressing to the next closest plane to the sensors. The voxels furthest to the right are located on the plane closest to the sensors.

To demonstrate the effect of column normalization for bias correction, we simulated a single dipole source at the center of the field of view, 3.5 cm below the central sensor, with a strength of 2×10^4 pJ T^{-1}. We calculated the signal vector b using the forward model considering a single stage position and no added noise. We then performed the reconstruction with and without the column normalization. The data fidelity parameter was set to 1×10^{-4} to account for rounding error. The results of the reconstruction are shown in Figure 3.10. Both reconstructions had a residual $\|\mathbf{A}\boldsymbol{x} - \boldsymbol{b}\|$ equal to 1×10^{-4}, but the reconstruction without the column normalization had a smaller L1 norm ($\|x_W\|_1 = 5.7 \times 10^{-3}$) than the column normalized reconstruction ($\|x_A\|_1 = 2 \times 10^4$). This demonstrates the ill-posedness of the problem. With the un-normalized system matrix, the algorithm was able to find a solution that satisfied the residual requirement using a collection of small dipole

moments located very close to the detectors. Since the L1 norm is the total magnitude of the voxels in the solution, multiple small dipoles near the detector can account for the detected field and have a total magnitude smaller than a single large dipole farther from the detectors. When the columns are un-normalized, the voxels near the detectors require a much smaller moment to produce the same field as a large moment in a voxel farther from the detectors. Therefore, to minimize the total moment, the solution is weighted more heavily towards voxels near the detectors. By normalizing the columns of the system matrix, each voxel is given an equal weight regardless of its location, so the solution is no longer biased towards the voxels closest to the detector.
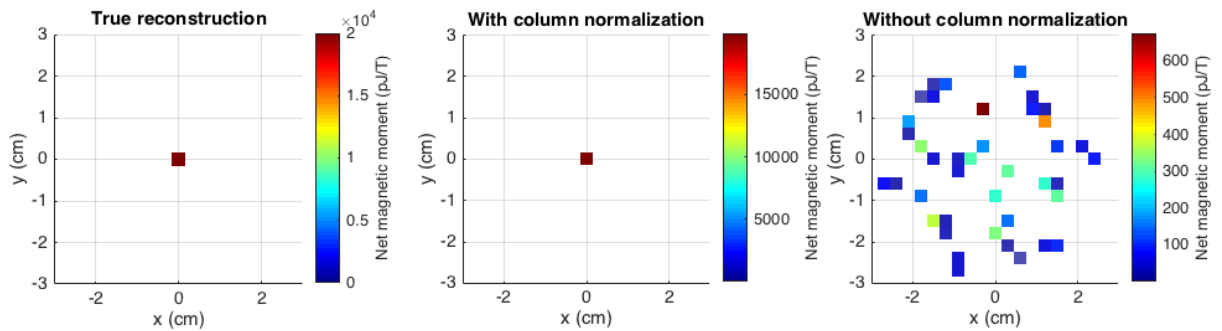
Figure 3.10: The reconstruction with (right) and without (middle) the column normalization for depth bias correction. Without the normalization, the solution is confined to the voxels closest to the detectors, whereas after the column normalization, the reconstruction closely resembles the true source distribution (left).

Chapter 4

Calibrating the forward problem

The success of the reconstruction algorithm is inherently dependent on the accuracy of the forward model. In this case, the forward model assumes that the location of the magnetic field measurements is precisely known, so that any uncertainty in the measurements is from noise alone. In order to best satisfy this assumption, we take a data-driven approach to defining the sensor locations. We start with the locations provided in the drawings of the system, then use a sensitivity map to find the sensor location parameters that best account for the measured field from a source at a known location. Then we can better identify where an unknown source is given the known sensor locations.

Inside the dewar, six gradiometers are arranged in a hexagonal pattern around a central gradiometer. The six outer sensors are tilted slightly so that the bottom coils sit on the surface of an imaginary sphere. The location of the gradiometers is defined by four parameters, shown in Figure 4.1: in (a), the angle of rotation between the central and outer gradiometers, ϕ, the radius of the imaginary sphere, h, in (b) the rotation of the array around the z axis, θ, and in (c) the offset in the (x,y) plane between the center of the stage and the center of the central gradiometer, d.

Each SQUID sensor is made up of two components: a pickup coil and the SQUID circuitry. A second-order gradiometer pick-up coil, shown in Figure 4.1a, converts the magnetic field into an electrical current [91]. The gradiometer consists of three pickup coils
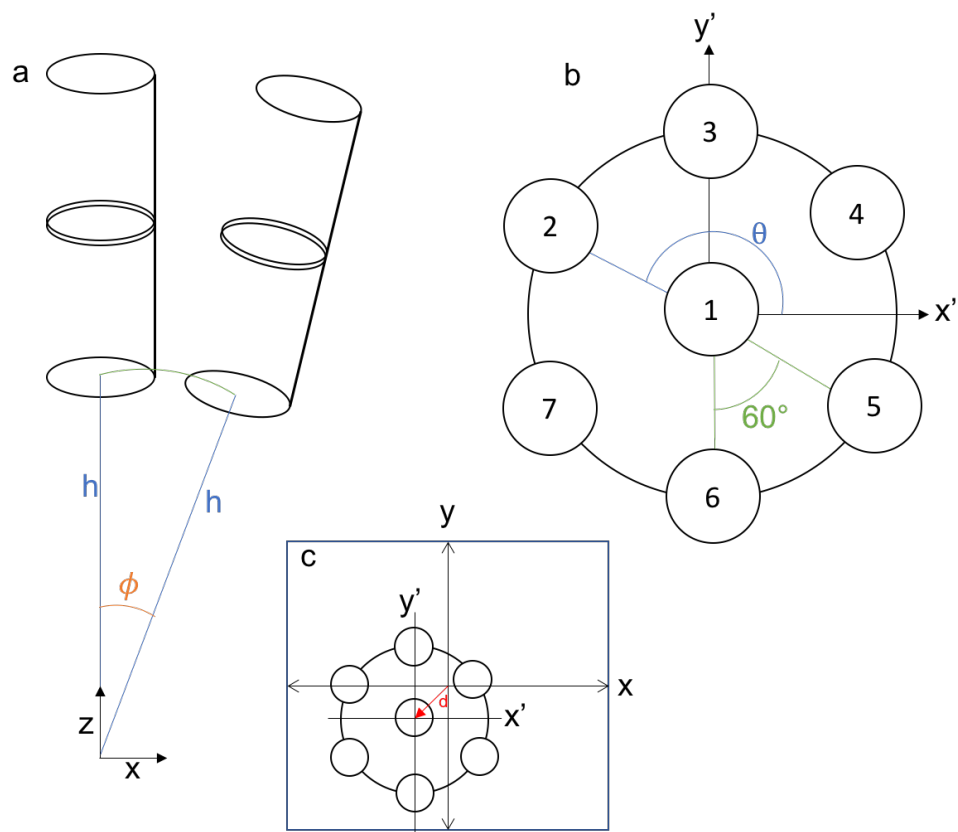
Figure 4.1: Illustration of the sensor parameters

spaced 4 cm apart. The coils consist of the equivalent of one turn at the bottom and top, and two turns in the middle wound counter to the top and bottom turns. This gives the gradiometer its spatial-gradient characteristic. The relationship between induced current and magnetic flux through each coil as described in Equation 4.1. The current induced by the magnetic field flows from the gradiometer to the second component of the sensor, the SQUID circuit. This highly sensitive, near zero resistance circuitry converts the extremely small induced currents into the voltage that is recorded by the software.

$$I = \Phi_1 - 2\Phi_2 + \Phi_3 \tag{4.1}$$

The flux through each loop (Φ) is integral of the component of the magnetic field (b) perpendicular to the loop over the area of the loop. To calculate the flux, we approximate the integral with a summation over N segments of a loop with centers at r_i and area a_i, as in Equation 4.2.

$$\Phi(\vec{r}) = \int_A b(\vec{r}) \cdot dA \approx \sum_{i=1}^{N} b(\vec{r_i}) a_i \tag{4.2}$$

Using the technique described in [92], we discretize the coil into equal segments and calculate the magnetic field in each segment, as shown in Figure 4.2. Figure 4.3 shows the convergence of the flux calculation with an increasing number of segments using both a radial and equal discretization.

To accurately calculate the magnetic field at each point within the loop, the forward model requires the precise position and orientation of the center of each loop of each pickup coil relative to the source location to correctly correlate the detected field values to the source dipole. First, we define the location of the center of the bottom coil of the central sensor (Sensor 1) to be a distance h above the origin, at the coordinates $(0, 0, h)$. The middle and top coils are then located at $(0, 0, h + 4$ cm$)$ and $(0, 0, h+8$ cm$)$, respectively. From there, the location of the coils in Sensor 2 are found by applying a rotation of ϕ
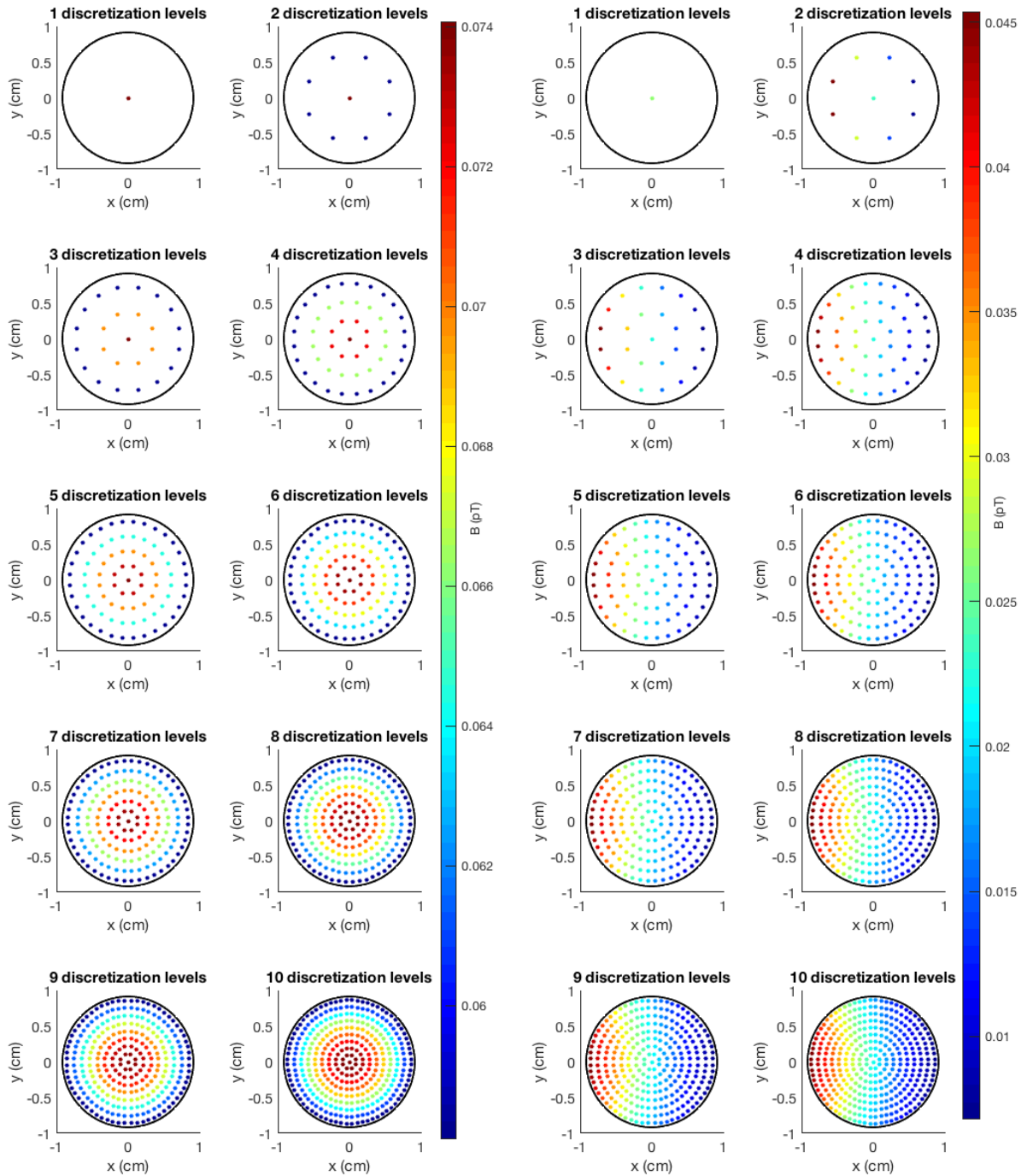
Figure 4.2: The magnetic field calculated at each segment of the discretized sensor coil using the equal discretization scheme for the central (left two columns) and outer (right two columns) sensors

57
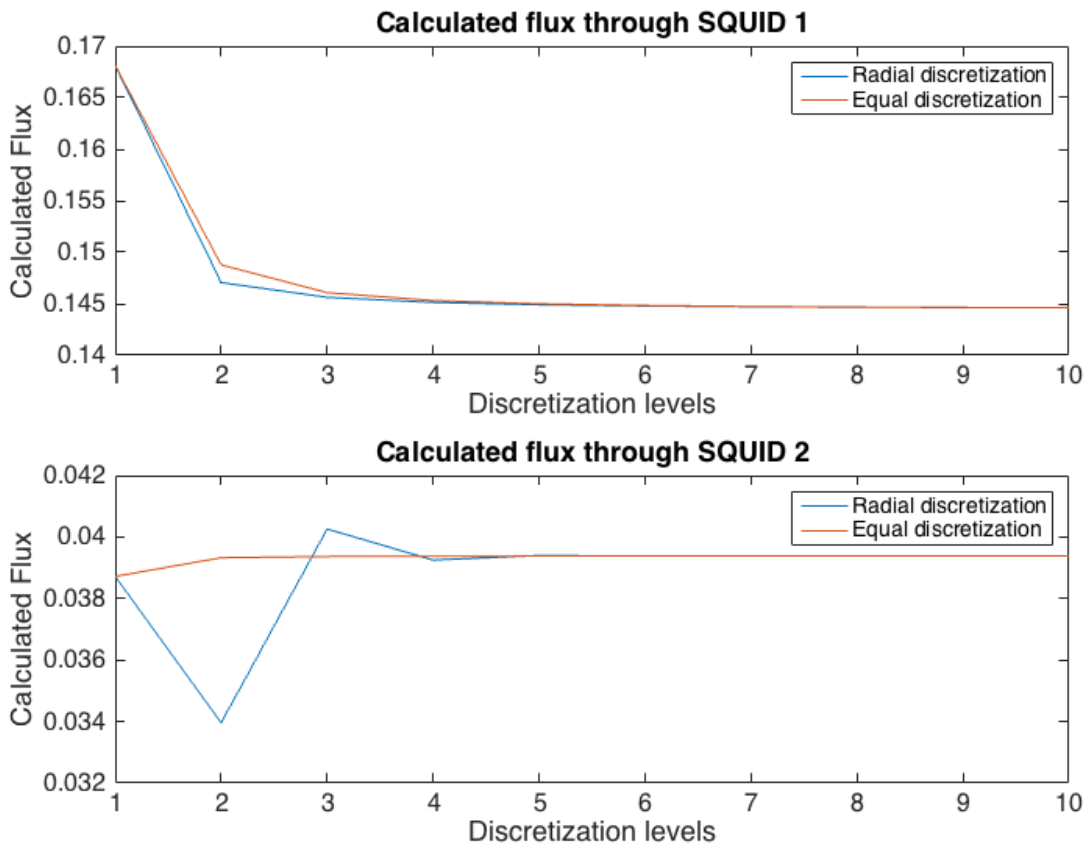
Figure 4.3: The calculated flux converges as the number of segments increases for both the radial and equal discretization methods.

around the y axis to the coils of Sensor 1. Then the location of the coils in sensor 3 are found by rotating the coils in Sensor 2 by $\pi/3$ around the z axis. Similarly, the location of Sensors 4 through 7 are found by applying a rotation of $\pi/3$ to the previous sensor's location. Once the sensor locations relative to the central sensor are found, the array is aligned with the stage coordinate system by rotating the entire sensor array around the z axis by an additional θ, and applying a shift $d = (d_x, d_y)$ to account for the displacement between the central sensor and the center of the stage. Finally, the origin is moved to the plane of the bottom coil of the central sensor by subtracting $(0, 0, h)$ from each sensor location.

Given the voltage (V) recorded by a sensor from a source with a known dipole strength μ at a known location $(\vec{r}_s)$, we can solve for the sensor location parameters ϕ, θ, d, and h. The voltage recorded by the system is a function of the flux through each coil of a sensor multiplied by a calibrated constant to convert from Volts to magnetic flux (C) (Equation 4.3).

$$V = C\left[\Phi_1(\mu, \vec{r}(\phi, \theta, h, \vec{d}) - \vec{r}_s) - 2\Phi_2(\mu, \vec{r}(\phi, \theta, h+4, \vec{d}) - \vec{r}_s) + \Phi_3(\mu, \vec{r}(\phi, \theta, h+8, \vec{d}) - \vec{r}_s)\right]$$

(4.3)

Where

$$\vec{r}(\phi, \theta, h+\delta, d) = \begin{bmatrix} \cos(\phi)\cos(\theta) & -\sin(\theta) & \sin(\phi)\cos(\theta) \\ \cos(\phi)\sin(\theta) & \cos(\theta) & \sin(\phi)\sin(\theta) \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ h+\delta \end{bmatrix} - \begin{bmatrix} d_x \\ d_y \\ h \end{bmatrix}$$

(4.4)

4.1 Methods

To find the unknown parameters, we measured the magnetic field from a single point source phantom (Figure 4.4) at 81 points in the x-y plane at two depths 0.9 cm apart, for a total of 162 source locations. The distance from the source at a given depth to the

origin along the z-axis was unknown, but because the increment in depth was known, the starting position in z relative to the origin at the plane of the sensors could be included in the optimization as an additional unknown parameter. Specifically, the distance along the z-axis from the plane of the bottom coil to the source with the 0.9 cm spacer in place was defined as z. Then when the spacer was removed, the depth of the source was $z + (-0.9)$ cm. The exact strength of the source was also unknown, but because it was constant over all of the data, it too could be solved for in the optimization. Altogether our system of equations consisted of 1,134 data points (81 x-y locations × 2 depths × 7 sensors) and 7 unknown parameters $(\phi, \theta, d_x, d_y, h, z$ and $\mu)$. The parameters were fit to the measured field (V_{ij}) at each sensor located at $\vec{r}_i$ and each source location $\vec{r}_j$ by the constrained optimization in Equation 4.5. Because only one source strength was used, and the mass of particles was unknown, we set $C = 1$ for this calibration. The value of C can be calibrated by measuring a coil with a known magnetic moment. Using $C = 1$, the parameter μ found in the optimization will be scaled by a factor of C. However, the other parameters will be unaffected because only the magnitude of the detected signal is affected by C, and the other parameters are determined by the shape of the field.

$$\min \sum_{i=1}^{7} \sum_{x_j, y_j} (V_{ij} - b(\mu, \vec{r}_i(\phi, \theta, h, \vec{d}) - [x_j, y_j, z]))^2 \quad \text{subject to} \quad \begin{cases} 0 < \phi \leq \pi/4 \\ 0 \leq \theta \leq 3\pi/2 \\ 10 \text{ cm} \leq h \leq 30 \text{ cm} \\ -6 \text{ cm} \leq z \leq -1 \text{ cm} \\ 0 \leq \mu \leq 1 \times 10^8 \text{pJ/T} \\ -4 \text{ cm} \leq d_x \leq 4 \text{ cm} \\ -4 \text{ cm} \leq d_y \leq 4 \text{ cm} \end{cases}$$

$$(4.5)$$

Where

$$b(\mu, \vec{r}_i(\phi, \theta, h, \vec{d}) - [x_j, y_j, z]) = \frac{1}{a} \Big[\Phi(\mu, \vec{r}(\phi_i, \theta + \pi/3(i-2), h, \vec{d}) - [x_j, y_j, z])$$

$$-2\Phi(\mu, \vec{r}(\phi_i, \theta + \pi/3(i-2), h + 4, \vec{d}) - [x_j, y_j, z]) + \Phi(\mu, \vec{r}(\phi_i, \theta, h + 8, \vec{d}) - [x_j, y_j, z]) \Big]$$

$$\phi_i = \begin{cases} 0 & i = 1 \\ \phi & i \neq 1 \end{cases} \qquad (4.6)$$
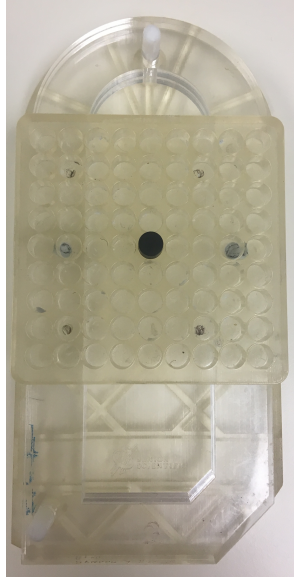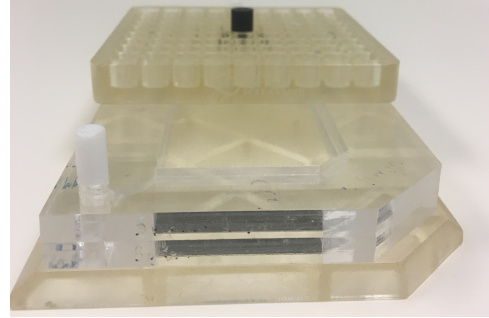
(a)   (b)

Figure 4.4: Phantom configuration (a) with both spacers (b)

A series of 30 background pulses with no source present was taken before each set of data at a given depth. The background field was subtracted from each data point before optimization. Twenty measurements of the magnetic field were recorded at each source location. The stage was not moved from its position at the origin at any point during data collection. At each depth, the source was started at the upper-left-most position in the grid phantom, which corresponds to (-3.6 cm, 3.6 cm) in (x, y). From there, it was moved to the right one position (0.9 cm) at a time until it reached (3.6 cm, 3.6 cm). Then it was moved back to the left-most side of the grid and down one position (0.9 cm) in y to (-3.6

Table 4.1

Parameter	Description	Lower Bound	Upper Bound
ϕ	Sensor rotation angle	0	$\pi/4$
θ	Dewar rotation angle	0	$3\pi/2$
h	Distance to axis of rotation	10 cm	30 cm
z	Depth of the phantom source	-6 cm	-1 cm
m	Magnetic moment of the phantom source	0	$1 \times 10^8 \, \mathrm{pJ\,T^{-1}}$
d_x	Offset between stage and sensor array in x	-4 cm	4 cm
d_y	Offset between stage and sensor array in y	-4 cm	4 cm

cm, 2.7 cm), and then across to the right. This was repeated until the source reached the bottom-most-right position at (3.6 cm, -3.6 cm). After all 81 points had been measured with one spacer, the spacer was removed and the process was repeated with no spacer.

The raw data was preprocessed using our flux jump correction method followed by a curve fitting procedure to determine the magnitude of the magnetic field represented by each decay curve, described in Section 3.3.2 [76]. Thirty measurements of the magnetic field were output from the preprocessing for each source location. From each group of 30, outliers -defined as data points outside of +/- 2 standard deviations of the sample mean- were removed, and the remaining initial field values were averaged to get a single estimate of the residual magnetic field at each sensor for each source location. Then, the data was fit to the forward model using the trust region reflective algorithm included in the MATLAB function `lsqnonlin`to find the detector location parameters which best accounted for the measured data. The optimization was bound by physically reasonable parameter limits, listed in Table 4.1. To ensure that the solution was indeed a global minimum, the optimization algorithm was run with 1000 iterations of starting conditions randomly sampled from a uniform distribution within the bounds of the sensor location parameter values.

The optimization was conducted with the midpoint approximation as well as the discretized model. In order to keep the computational time reasonable, the discretized forward problem used a sensor area discretization of 12 equal area samples and one sample

in the center with half the area of the others, as shown in Figure 4.5. In addition to the discretized approach, the optimization was run using the midpoint approximation. The midpoint approximation method is identical to the calculation described above, but uses only a single discretization step. In this way, the magnetic field at the center of the coil is used as an approximation of the average of the field across the entire area of the coil, which significantly reduces the computational time required to calculate the forward problem. Figure 4.2 shows the effect of this approximation (equivalent to one discretization level) relative to the discretized model for increasing levels of discretization. If the discretized model does not substantially improve the fit of the model to the data, it would be highly preferable to use the midpoint approximation approach due to the significant reduction in computational time required to recalculate the forward model for every set of data.
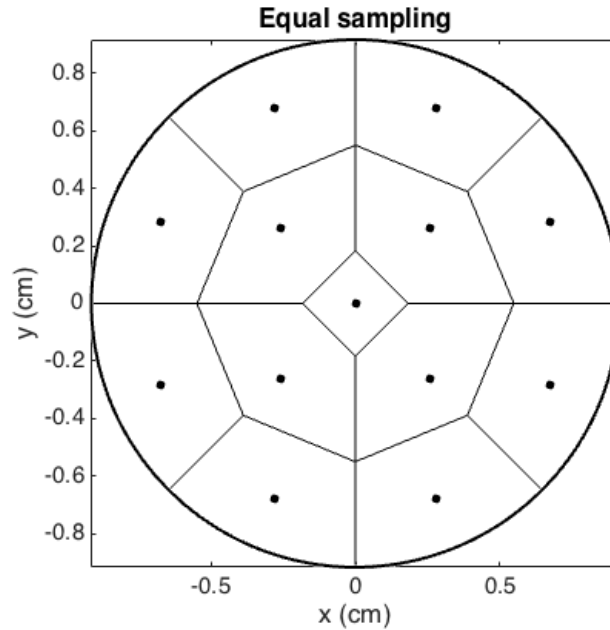
Figure 4.5: Discretization of the gradiometer coils for the forward problem optimization

The model/data agreement was characterized according to Equation 4.7, as well as the average absolute difference over all of the sensor locations.

$$\% \text{ error} = \frac{||B_{i,modeled} - B_{i,detected}||}{||B_{i,detected}||} \times 100 \qquad (4.7)$$

Parameter	From drawings	From midpoint optimization	From discretized optimization
ϕ	6.5°	6.5°	7.18°
θ	148°	130°	130°
h	20.3 cm	19.1 cm	17.2 cm
d_x	0 cm	0.19 cm	0.19 cm
d_y	0 cm	0.24 cm	0.24 cm
z	-	-3.18 cm	-2.87 cm
m	-	$1.33 \times 10^5 \, \text{pJ T}^{-1}$	$1.16 \times 10^5 \, \text{pJ T}^{-1}$

Table 4.2: Initial and optimized forward model parameters

4.2 Results

4.2.1 Optimized sensor parameters

The optimized parameter results of the midpoint approximation for which the optimization resulted in a local minimum are shown in Figure 4.6. By looking at the 2D histogram values, we can see to which parameter value most of the optimizations converged, and confirm that the modes of these distributions belong to the same solution. For example, any mode on a histogram can be followed across the column to identify the cluster of solutions that correspond to complementary values of the other parameters. From this analysis, we find the global optimum sensor location parameter values, listed in Table 4.2 along with the parameters derived from the system drawings. The strength of the phantom source was found to be $1.33 \times 10^5 \, \text{pJ T}^{-1}$ and the depth with the spacer in place was found to be -3.12 cm below the lowest coil of Sensor 1. The magnetic field predicted using the new model parameters is visibly a good representation of the actual measured data, as shown in Figures 4.7 and 4.8. The percent agreement between the model and the data, as defined by Equation 4.7 was 4.53%.
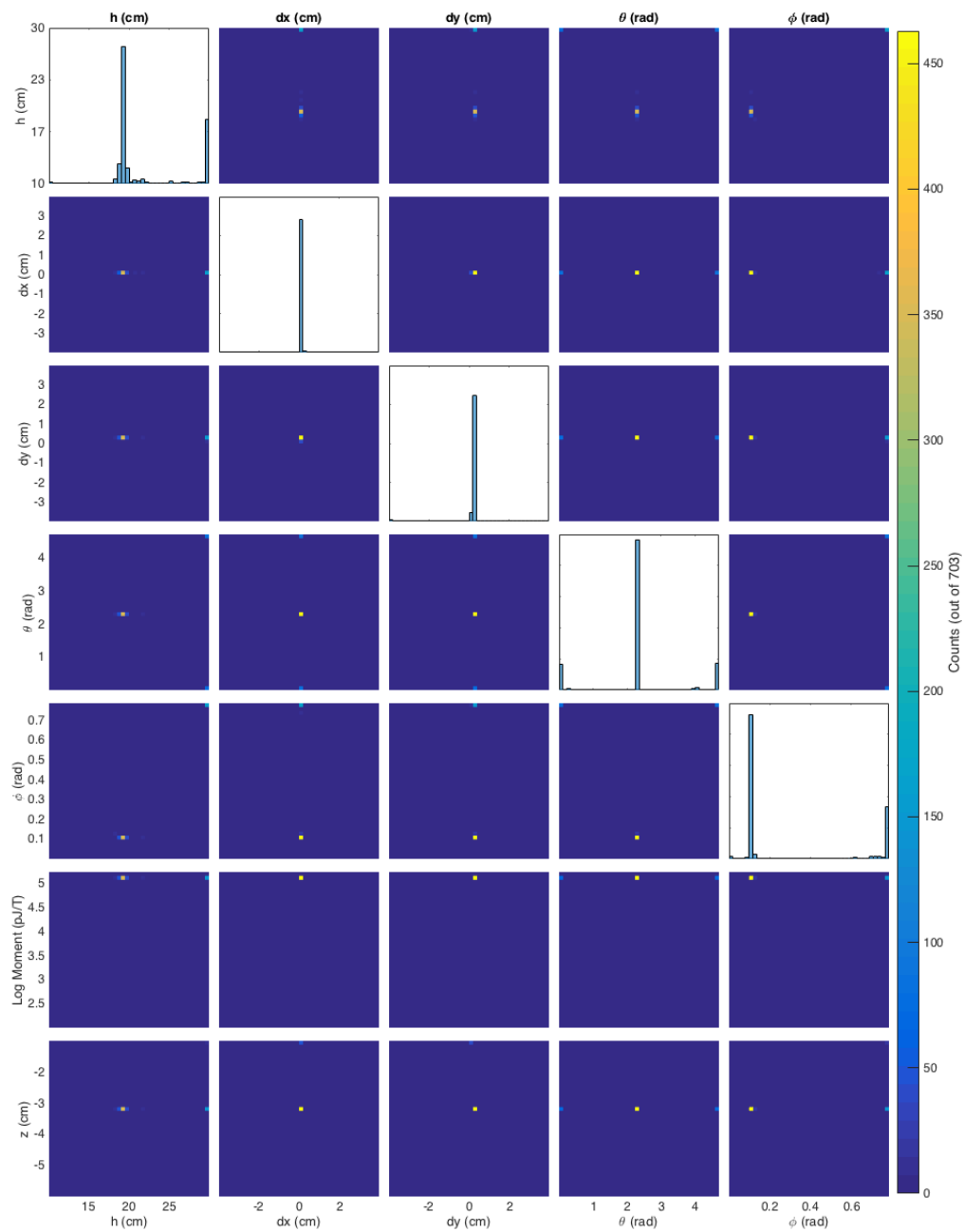
Figure 4.6: Distribution of optimal parameter values from optimizations of the midpoint model which converged to a local minimum, or for which the change in the solution or residual was less than the specified tolerance of 1×10^{-6} 1D histograms shown on the diagonal of the distribution of a single variable.
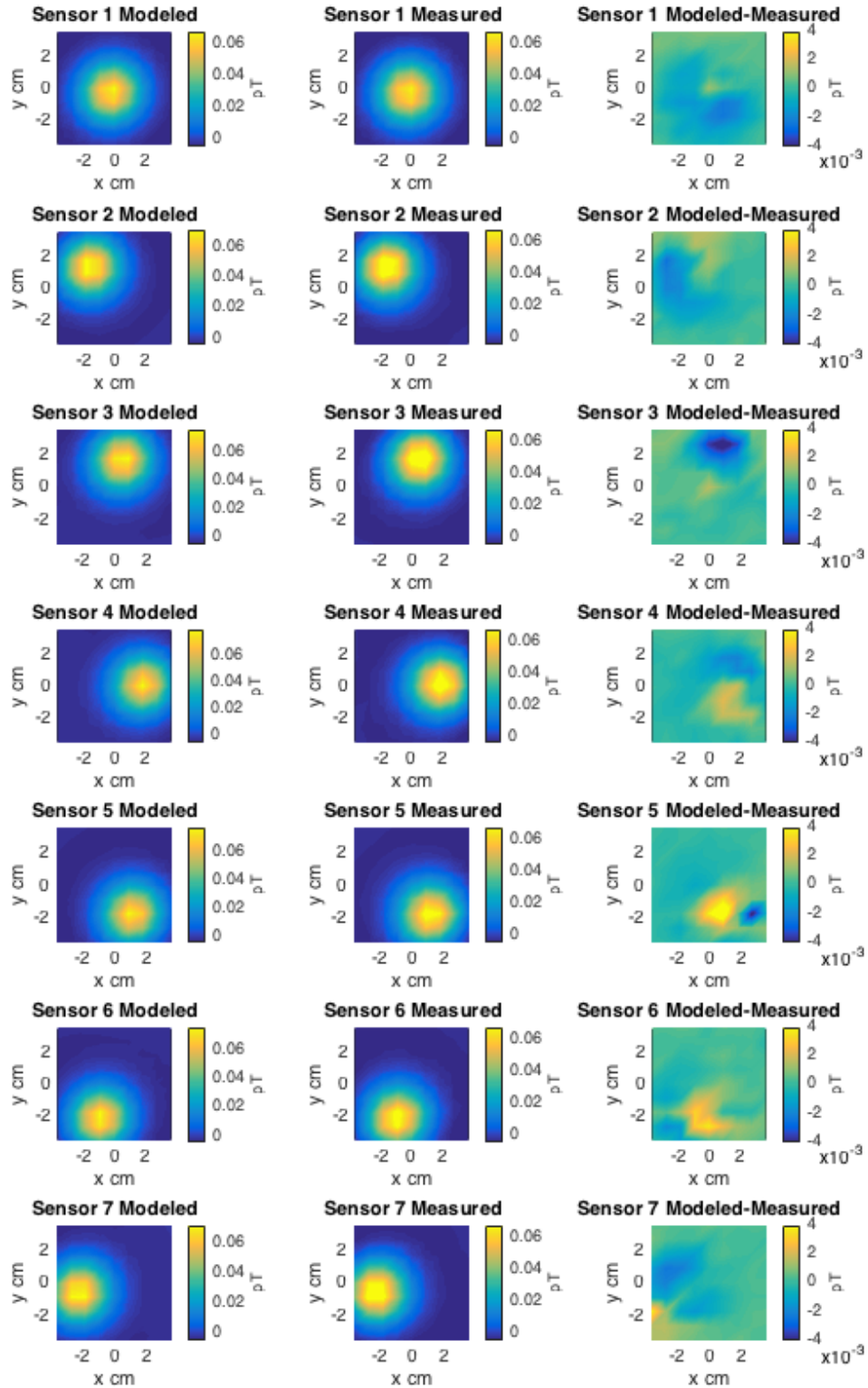
Figure 4.7: The modeled field using the midpoint model optimized parameters (left column), the measured data (middle column) and the residual between the two (right column) for each sensor (rows), as a function of the source location with the spacer in place.

66
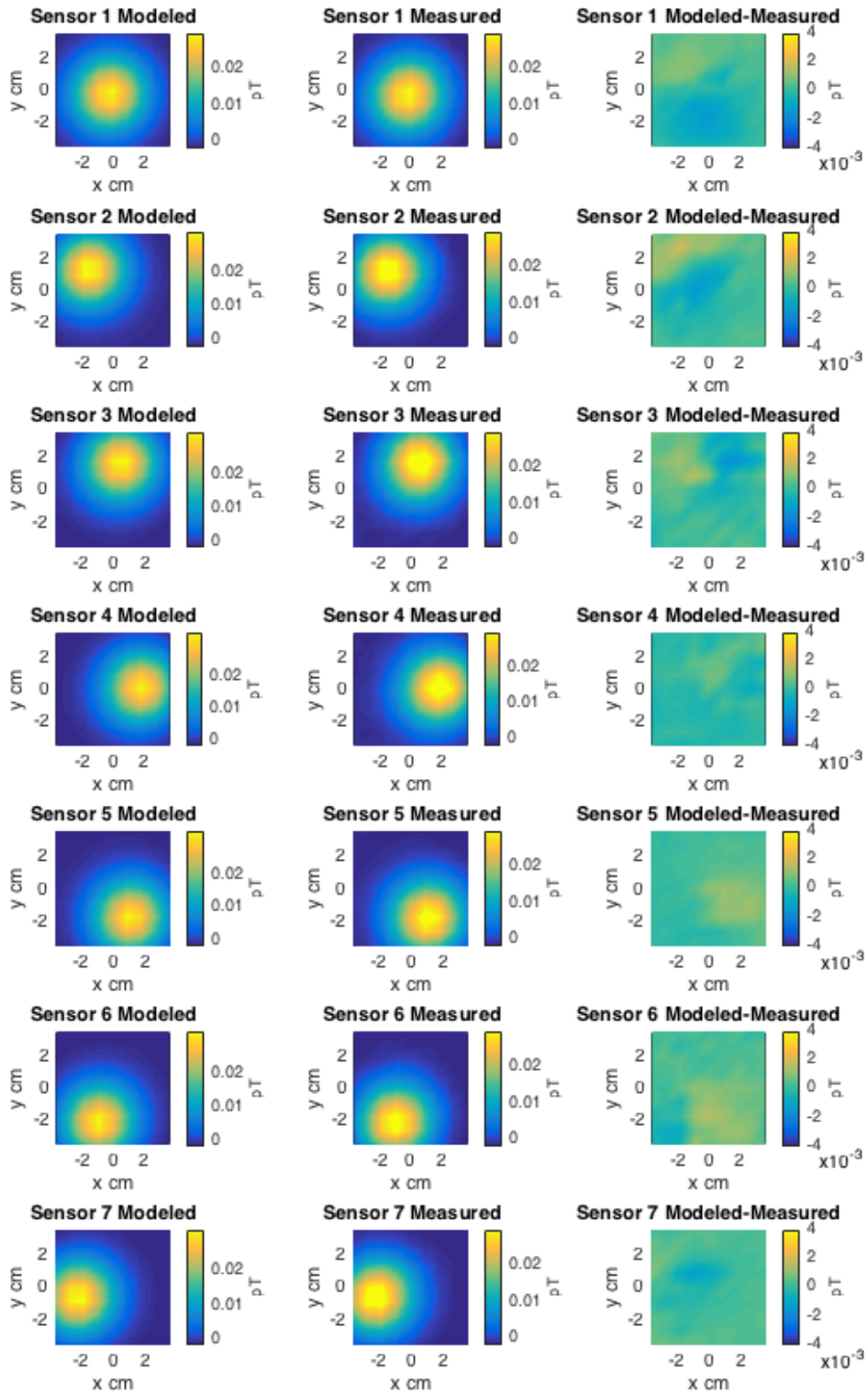
Figure 4.8: The modeled field using the midpoint model optimized parameters (left column), the measured data (middle column) and the residual between the two (right column) for each sensor (rows), as a function of the source location with no spacer in place.
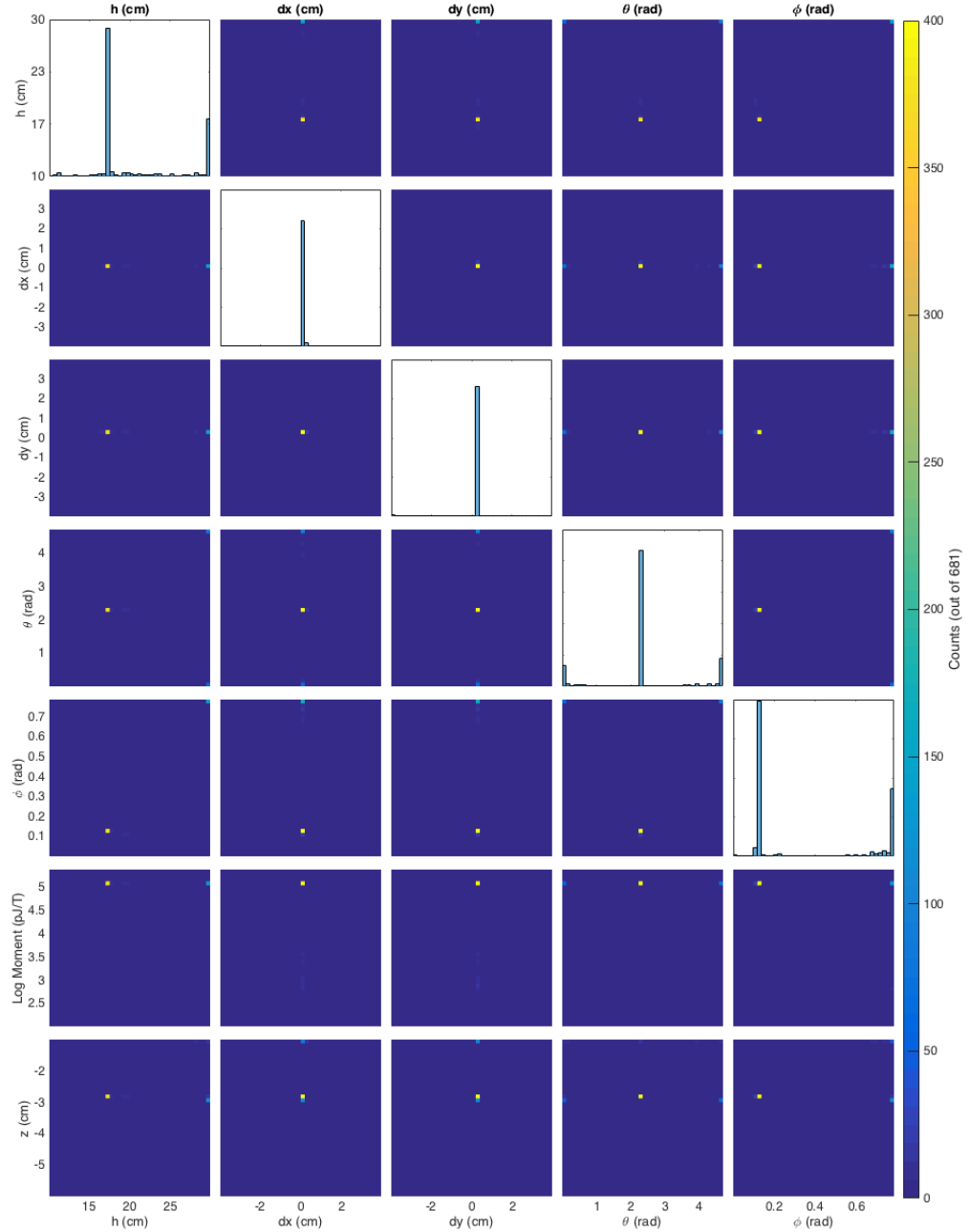
Figure 4.9: Distribution of optimal parameter values from optimizations of the discretized model which converged to a local minimum, or for which the change in the solution or residual was less than the specified tolerance of 1×10^{-6} 1D histograms shown on the diagonal of the distribution of a single variable.

Surprisingly, the optimal parameters found using the discretized sensor model, listed in the right hand column of Table 4.2, do not agree with the technical drawings as well as those found using the midpoint approximation. The biggest differences are in the pa-

rameters ϕ and h, which are related through the arc angle $\alpha = h \sin \phi$. These parameters determine the normal vector of and location of the coils of Sensors 2 through 7, located on the outer ring of the array. This has a large influence on the reconstructed depth, and therefore also the strength, of the source. Accordingly, these were also different for the discretized model. The discretized model had a percent error of 19.3%.

We can tell from Figure 4.9 that it is unlikely that the discrepancy is due to the optimization finding a local minimum because the optimal values from optimizations which converged to a local minimum are tightly clustered around a small range of values. Still, the minimum residual values for this set of parameters is comparable than that found using the midpoint approximation, indicating that the parameters found using the midpoint approximation fit the measured data about as well as those found using the fully discretized model. From Figures 4.10 and 4.11 we can see that the optimized discretized model tends to overestimate the detected field values, especially towards the center of the field of view where the signal is strongest.

From this data we can also evaluate the agreement between the model and the data, both with and without the discretization. The agreement between the measured data and what is predicted by the model is directly related to the maximum quality of the reconstruction and limit of detection in that we must allow the reconstruction to differ from the detected field to an amount necessary to account for disagreement between the model and the detected data, which will inherently limit the accuracy of the reconstruction. To assess the difference between the two models, the optimal sensor location parameters, source strength and depth found in each optimization were used to calculate the expected magnetic fields for each known source position. The measured field values were plotted against the expected field values from each model, and fit with a line, shown in Figure 4.12. For perfect model/data agreement, the slope and intercept of the line would be 1 and 0, respectively, and all of the data points would lie along the line, resulting in a R^2 of 1. For the discretized model, the average absolute difference between the modeled and measured

69
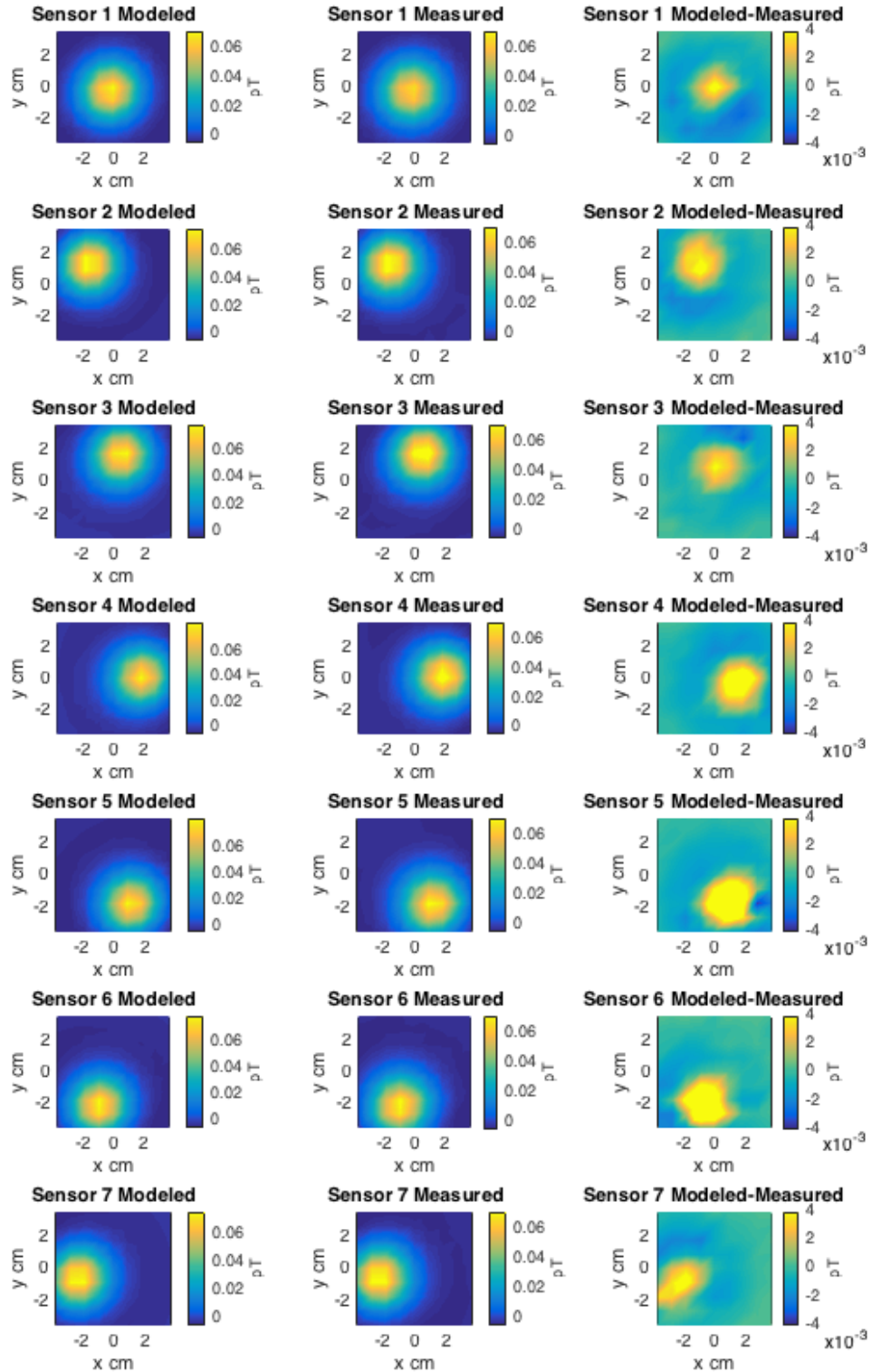
Figure 4.10: The modeled field using the discretized model optimized parameters (left column), the measured data (middle column) and the residual between the two (right column) for each sensor (rows), as a function of the source location with the spacer in place.

70
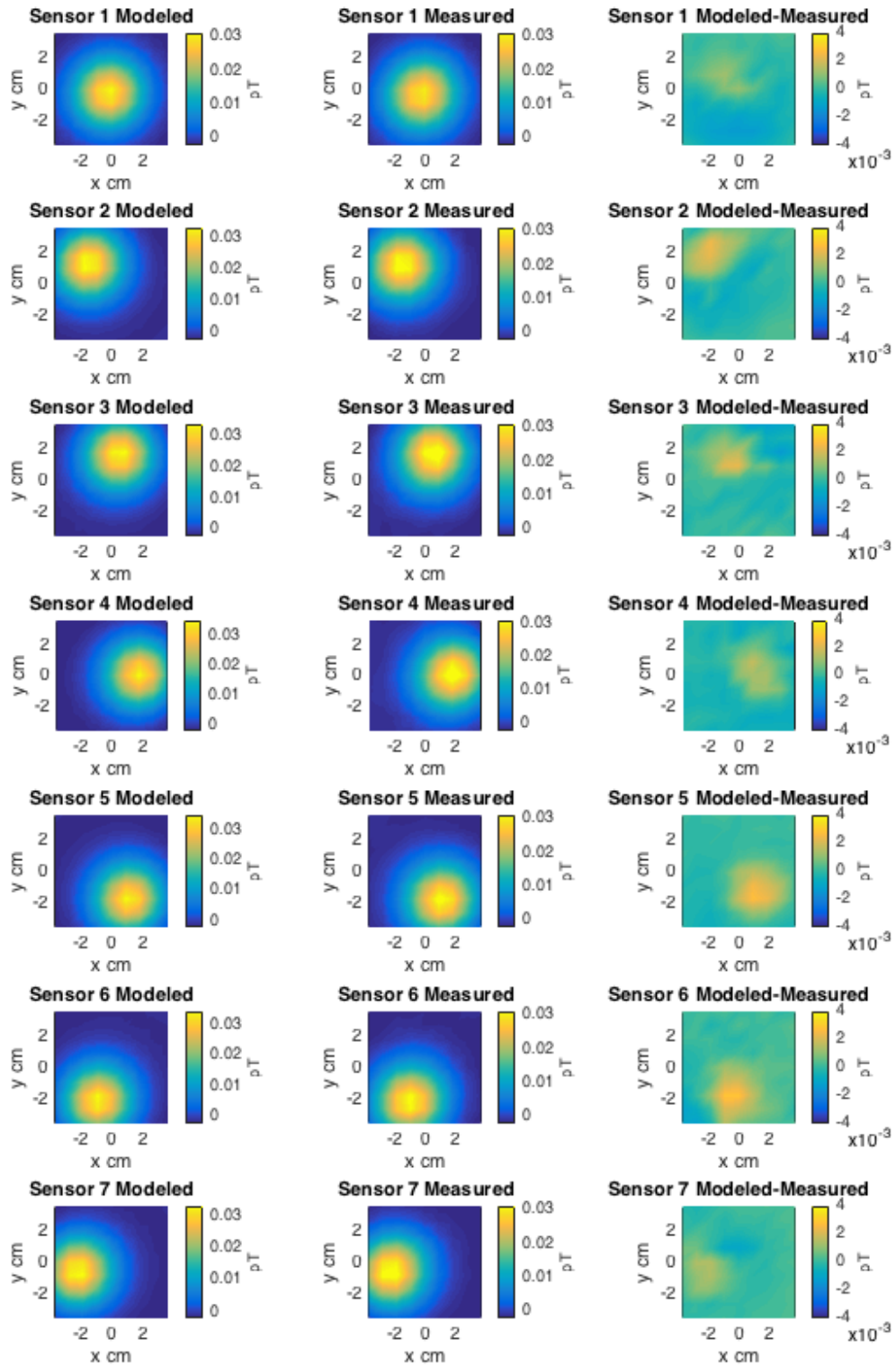
Figure 4.11: The modeled field using the discretized model optimized parameters (left column), the measured data (middle column) and the residual between the two (right column) for each sensor (rows), as a function of the source location with no spacer in place.
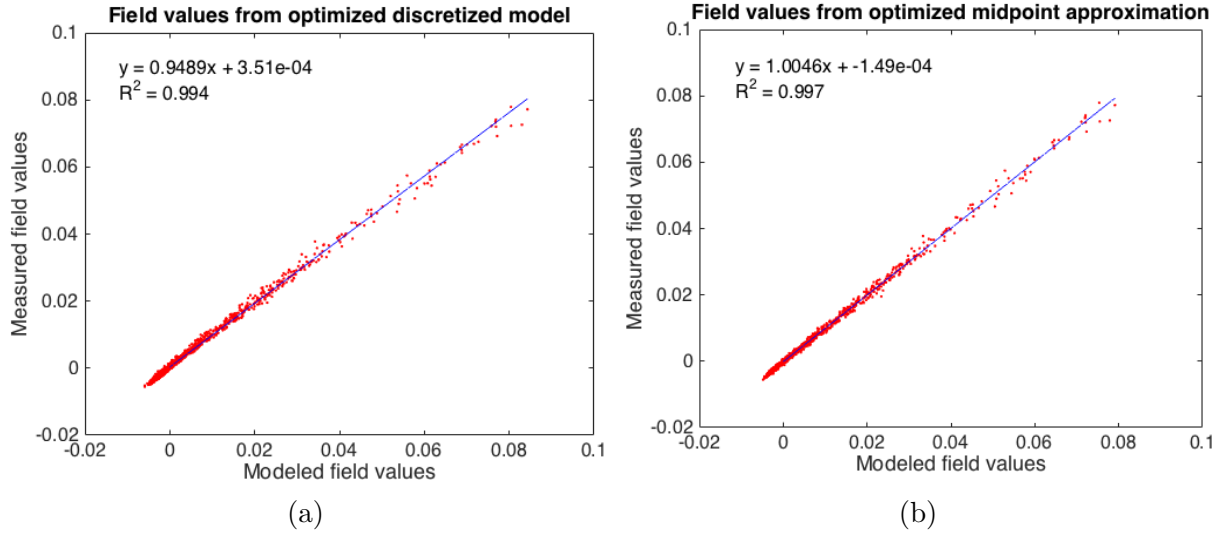
Figure 4.12: Model/data agreement using (a) discretized sensor model and (b) midpoint approximation to the sensor model

values was 7.1×10^{-4} pT, and the slope of the linear fit between modeled and measured data was 0.95, indicating that the measured values tended to be lower than expected from the model. The R^2 of the linear fit was 0.994. When the optimized midpoint approximation was used, the average difference between measured and modeled values was 4.9×10^{-4} pT, an improvement of more than 30%. The slope of the linear fit increased to 1.00, and the R^2 increased to 0.997. This substantial improvement from a change in the sensor tilt angle of less than 2 degrees emphasizes the importance of having an accurate description of the sensor location parameters.

4.2.2 Noise analysis

We can also use this calibration data to characterize the noise in the data. There are two assumptions we make regarding the noise characteristics of the data that must be tested. The first is that the noise is independent of the signal strength, and the second is that it is approximately normally distributed. We used the data from the calibration phantom study to test these hypotheses. First, we computed the Pearson's linear correlation coefficient (Equation 4.8, in which $\bar{x}$ is the mean of a vector x) to determine linear correlation between the mean and standard deviation of the measurements of each sample location [93]. The results for each sensor are shown in Figure 4.13. Then, we subtracted the mean of each data set from each of the data points in the set, and fit the resulting distribution to a normal curve, as shown in Figure 4.14.

$$\text{corr}(x, y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{1/2}} \tag{4.8}$$

4.3 Conclusions

In this chapter, we calibrated the parameters that determine the location of the sensors in the forward model to data collected from a single point source located at each of 162 positions that spanned the field of view, and quantified the agreement between our model and the data. To better understand the benefit gained from discretizing the sensor coils to better estimate the magnetic flux, we conducted the study twice: once using the discretized forward model, and once using the field at the center of the coil to approximate the field over the entire coil. The benefit of the midpoint approximation method is that it is significantly less computationally expensive because it requires only one calculation of the field per sensor coil, whereas the discretized model must calculate the field at each discrete location in the coil which in this case was 13. To ensure that the optimization
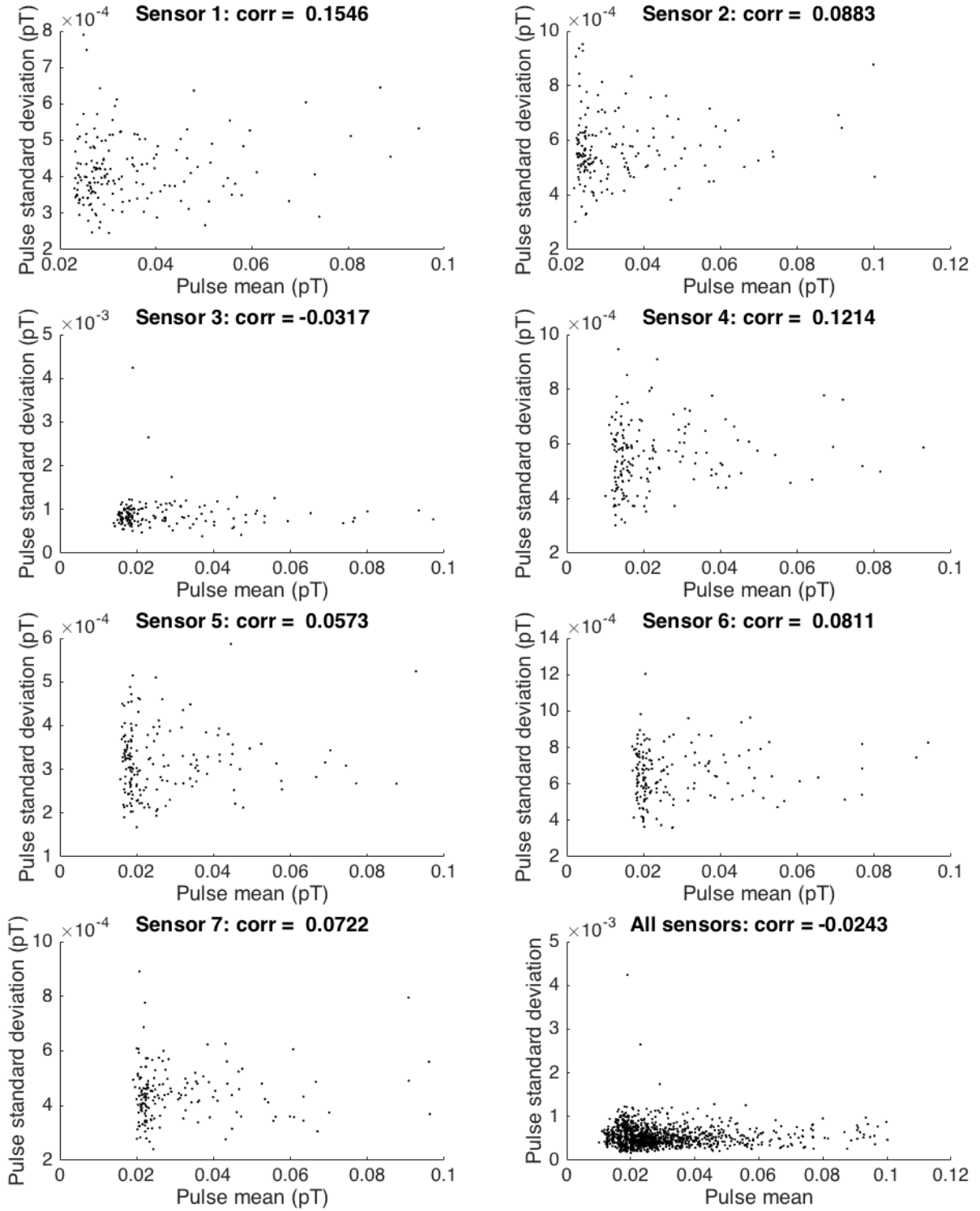
Figure 4.13: The correlation coefficient for the mean and standard deviation of each measurement (left) and the distribution of the difference from the mean (right) for each sensor.
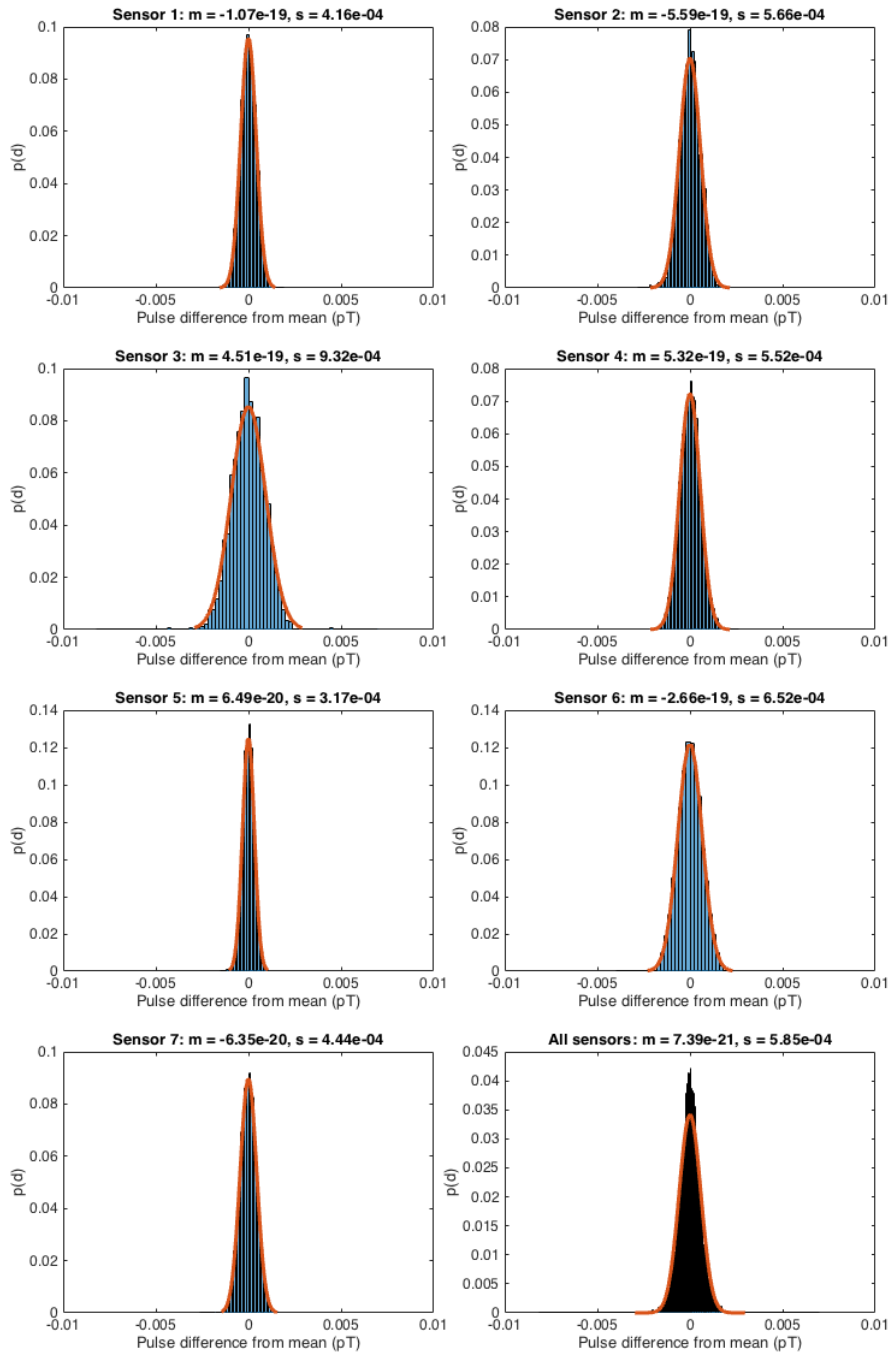
Figure 4.14: Distribution of pulse deviation from the mean for all pulses in the calibration data set

did not find a local minimum or a non-physical value, a bounded optimization using a trust-region-reflective algorithm was performed with 1000 starting conditions randomly selected from a uniform distribution of physically feasible values. The optimal parameter values found using the midpoint approximation were in good agreement with those listed in the technical drawings of the sensor array, but the the discretized model found a slightly different tilt angle for the sensors around the outer ring of the array. Additionally, the optimized midpoint approximation model actually fit the detected data better than the discretized model. This result was surprising because we would expect that discretizing the area of the pickup coil would lead to a more accurate estimation of the magnetic flux, which would in turn result in a better fit to the detected data. While unexpected, this result is encouraging because the superior model is also the more computationally convenient method.

Chapter 5

In Silico Studies of the Reconstruction Algorithm

In this chapter, we characterize the behavior of the reconstruction algorithm as a function of experimental and reconstruction design parameters through simulations. Before testing the algorithm experimentally, it is important that we have a good understanding of how both experimental parameters such as the number of stage positions and number of pulses collected at each stage position, and the reconstruction parameters such as the field of view and the data fidelity parameter, affect the end results of the reconstruction. As described in Chapter 3, the data fidelity parameter is crucial to the final reconstruction. Therefore, we focus on the selection of the data fidelity parameter throughout the chapter, and evaluate it repeatedly in conjunction with the experimental parameters. First, we review in detail the role of the data fidelity parameter in balancing the noise and accuracy of the reconstruction. Then, we present three potential methods of determining the data fidelity parameter as a function of the data. Finally, we evaluate through simulations the interaction of the data fidelity parameter and experimental parameters of the number of stage positions and the number of pulses per stage position. The work presented in this chapter will help provide a theoretical foundation on which to base future experimental methods for optimal results. In the following chapter, we will see these principles applied to measured data to compare with the current method of reconstruction.

5.1 Theory and motivation

As introduced in Chapter 3, the balance between the sparsity of the reconstruction and the degree to which it accounts for the measured data is determined by the choice of the parameter λ in Equation 3.53. As shown in Figure 3.8, λ is bounded between 0 and $||\boldsymbol{b}||$. If $\lambda \geq ||\boldsymbol{b}||$, then the optimal solution to Equation 3.53 is simply $\boldsymbol{0}$, because there is no smaller possible value for the l1-norm than $||\boldsymbol{0}||_1 = 0$, and it is in the domain of the constraint: $||\mathbf{A}\boldsymbol{0} - \boldsymbol{b}|| = ||\boldsymbol{b}|| \leq \lambda$. As λ decreases, the reconstruction becomes less sparse and better fits the measured data. As $\lambda \to 0$, the reconstruction begins to fit the noise, and may become infeasible in cases with low signal to noise ratio (SNR). When $\lambda = 0$, a solution $\boldsymbol{x}^*$ is only feasible if it satisfies $\mathbf{A}\boldsymbol{x}^* = \boldsymbol{b}$. If there is any uncertainty in $\boldsymbol{b}$ such that it is not identical to the signal generated by the true solution $\mathbf{A}\boldsymbol{x} \neq \boldsymbol{b}$, then it is clear that if a solution $\boldsymbol{x}^*$ to Equation 3.53 exists, it must be different from the true solution. The value of λ is bounded below by zero because it is the upper limit on a norm and a norm must be positive. Therefore, at some value of λ between $||\boldsymbol{b}||$ and 0, there is an optimal reconstruction that best approximates the true source distribution without over fitting the noise. However, the exact value of this optimal λ cannot be determined *a priori*, and must be estimated as best as is possible. Here, we consider a method of selecting a suitable value for λ using the chi-squared metric. As discussed previously, the parameter λ accounts for the inherent mismatch between the model and the measured data. This discrepancy is due to low frequency environmental shifts in the detected magnetic field, uncertainty introduced during preprocessing and initial field estimation, and inaccuracies in the forward model. While every effort is taken to minimize these uncertainties, they cannot be eliminated entirely. Due to the random nature of the uncertainty, it cannot be known explicitly, but it can be characterized and estimated to the best of our ability. Here, we will describe how we determine a value of λ that is likely to achieve an adequate balance between an accurate and noisy solution.

The simplest approach is to recognize that, at least mathematically, the parameter λ should be equal to the norm of the difference between the detected field and the field calculated from the model of the source distribution. Ideally, the difference between the detected and modeled field should be only the noise. However, since the noise is random, we cannot know it explicitly except for in simulations. It is useful to explore the effect of the parameter on the reconstruction using simulations where the true signal and the true noise is known exactly. In this scenario, we would expect that the difference between the simulated source distribution and the reconstructed source distribution would be minimized when $\lambda = ||\boldsymbol{b} - \boldsymbol{b}^*||$, where $\boldsymbol{b}^*$ is the field calculated from the forward model of the simulated source distribution, and $\boldsymbol{b} = \boldsymbol{b}^* + \boldsymbol{n}$ is the signal model plus a known noise vector $n \in \mathcal{N}(0, \sigma)$.

However for real data acquisition, the true noise vector is not explicitly available, and the upper bound on the norm of the error in the detected field must be estimated from the data. One approach uses the properties of the chi-squared distribution to estimate a likely upper bound on the residual norm [78, 94]. The chi-squared distribution is defined in Equation 5.1 for a random variable x_i from a normal distribution with expected value $E(x_i)$ and variance σ^2.

$$\chi^2(k) = \sum_{i=1}^{k} \left(\frac{x_i - E(x_i)}{\sigma_i} \right)^2 \tag{5.1}$$

Substituting b_i for the field value observed at sensor location $i = 1 : k$ and an expected value of $\boldsymbol{a_i x}$, where $\boldsymbol{a_i}$ is the i-th row of the matrix $\mathbf{A}$ results in Equation 5.2.

$$\chi^2(k) = \sum_{i=1}^{k} \left(\frac{b_i - \boldsymbol{a_i x}}{\sigma_i} \right)^2 \tag{5.2}$$

Under the assumption that the noise is consistent across all sensors, σ can be factored out and the remaining sum is equivalent to the square of the norm of the residual.

$$\chi^2(k) = 1/\sigma^2 ||\mathbf{A}\boldsymbol{x} - \boldsymbol{b}||^2 \tag{5.3}$$

As a property of the chi-squared distribution, the probability of observing a value of χ^2 greater than two standard deviations $(\sqrt{2k})$ above the mean (k) of the distribution is small (less than 5%). Therefore, we want to select a value for the tolerance parameter that is greater than or equal to this limit.

$$\chi^2(k) = 1/\sigma^2 ||\mathbf{Ax} - \boldsymbol{b}||^2 \leq k + 2\sqrt{2k}$$
$$||\mathbf{Ax} - \boldsymbol{b}||^2 \leq \sigma^2(k + 2\sqrt{2k}) \tag{5.4}$$
$$||\mathbf{Ax} - \boldsymbol{b}|| \leq \sqrt{\sigma^2(k + 2\sqrt{2k})} = \lambda$$

For application to SPMR, k is defined as the number of sensor locations. There are two ways to consider the value used for σ. One way is to consider it to represent the expected standard deviation of the measured data. This approach does not take into account that the $\boldsymbol{b}$ in Equation 5.4 is an average of multiple pulses, and therefore is a "worst case scenario" estimate of the error. A more accurate approach would be to consider σ to be the standard error on the mean, or an estimate of the distance between the mean of the measurements and the true value of the magnetic field. This approach takes into consideration the number of pulses that are collected at each stage position by dividing the variance of the pulses by the number of pulses collected at each stage position, N, as in Equation 5.5.

$$SE = \frac{\sigma}{\sqrt{N}} \tag{5.5}$$

We will investigate both methods, defined in Equations 5.6 and 5.7 in which σ is the standard deviation of the measurements, k is the number of sensor locations, and N is the number of measurements per sensor location.

$$\lambda_\sigma = \sqrt{\sigma^2(k + 2\sqrt{2k})} \tag{5.6}$$

$$\lambda_{SE} = \sqrt{\frac{\sigma^2}{N}\left(k + 2\sqrt{2k}\right)} \tag{5.7}$$

In this section, we will investigate these three methods for defining the data fidelity parameter λ which balances the sparsity of the reconstruction and the agreement with the measured data. The first comes from the mathematical description of λ as the Euclidean distance between the true field ($\boldsymbol{b}^*$) and the simulated noisy detected field ($\boldsymbol{b}$), which we will call $\lambda_{\Delta \boldsymbol{b}}$ (Equation 5.8). However, since the error in the detected field is not explicitly known from real measurements, it is only feasible to calculate this in simulations. Additionally, as demonstrated theoretically in Figure 3.8, the reconstruction at $\lambda = \lambda_{\Delta \boldsymbol{b}}$ may not be the reconstruction closest to the true solution (x_t).

$$\lambda_{\Delta \boldsymbol{b}} = ||\boldsymbol{b} - \boldsymbol{b}^*|| \tag{5.8}$$

The other two are derived from the chi-squared function. The first, defined in Equation 5.6, uses the standard deviation of the pulses to approximate the error in the detected field, and is an upper bound estimate of the optimal choice of λ, since it does not consider that the magnetic field used for reconstruction is an average of multiple measurements. Given multiple assumptions that are implicit in the use of the chi-squared function - such as normality in the error and identical uncertainty from each detector - it may be useful in applications with real data. The second, defined in Equation 5.7, accounts for the number of pulses at each sensor location. The benefit of λ_{SE} is that it considers all of the factors that could impact the reconstruction: the number of stage positions, the number of pulses per stage position, and the noise. However, it relies on many simplifying assumptions which real data may not strictly adhere to. Ergo, it may tend to under-approximate the best choice of λ when applied to real measurement data, which could result in excessively noisy solutions and a potential reduction in sensitivity.

As described above, the parameter is inherently dependent on the number of stage positions (through k), the noise in the data (through σ), and the number of pulses per stage position (through N). The following analysis is divided into two sections. The first will look at the effect of the number of stage positions for a constant number of pulses per stage

position. The second will investigate the effect of the number of pulses per stage position while keeping the number of stage positions constant. Both analyses will be conducted for distributions of one and two sources, both over a range of SNR levels from 0 to 20 dB.

5.2 Methods

The general procedure for the following simulation studies is as follows. A baseline detected field ($\boldsymbol{b}_b$) was calculated from a baseline source distribution according to the forward model presented in Chapter 4. The true magnetic field $\boldsymbol{b}^*$ and true source distribution $\boldsymbol{x}_t$ were determined by scaling the baseline source strength and detected field by a factor of C to achieve the desired SNR. SNR_{dB} is defined in Equation 5.9 wherein $\boldsymbol{b}$ is the average value of N measurements of the magnetic field with standard deviation σ at each of $i = 1 \ldots M$ sensor locations. The baseline source distributions and the corresponding baseline magnetic fields are shown in Figure 5.1.

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{\boldsymbol{b}}{\sigma} \right) \tag{5.9}$$

$$C = \frac{\sigma 10^{\frac{\text{SNR}_{\text{dB}}}{10}}}{\bar{b}_b} \tag{5.10}$$

A single measurement of the magnetic field was simulated by adding a noise vector randomly sampled from a normal distribution with zero mean and a standard deviation of 4.9×10^{-4} pT. The final noise-corrupted measured field, $\boldsymbol{b}$, was simulated by taking the average of N simulated measurements where N is the number of pulses per stage position. To accurately simulate the procedure for measured data, the standard deviation of the measurements was estimated by subtracting $\boldsymbol{b}$ from each of the measurement vectors and calculating the standard deviation of the result.
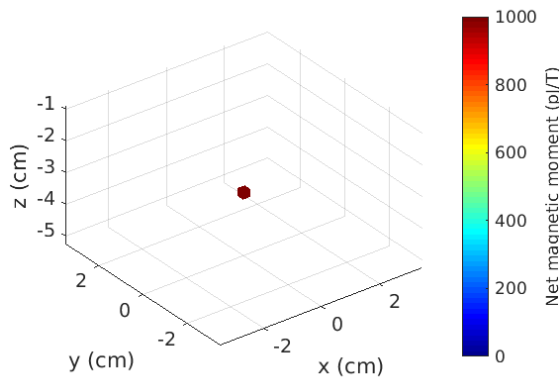
(a) The simulated single source distribution
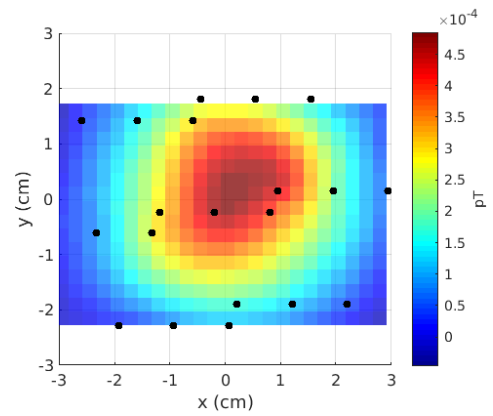
(b) The simulated magnetic field from the single source distribution
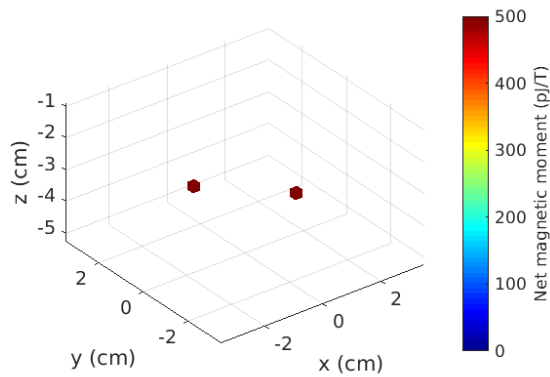
(c) The simulated two-source distribution
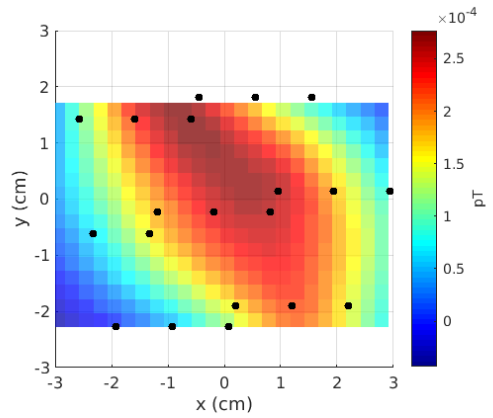
(d) The simulated magnetic field from the two-source distribution

Figure 5.1: The baseline single and two-source distributions and the corresponding magnetic field detected at three stage positions used in the simulation studies.

5.3 Effect of number of stage positions

In most linear systems, the stability of a solution in the presence of noise is dependent on the condition of $\mathbf{A}$, that is the independence of the information in the rows (determined by the sensor locations) and the columns (determined by the voxel locations). Others have shown the effects of the extent of the field of view as well as the number of sensors and voxels on the quality of the solution from a 2-dimensional TSVD reconstruction through simulations of a single point source [95]. This work showed that for an overdetermined system, a grid of fewer voxels improved the condition of the matrix, and thus the quality of the reconstruction. Groups working with multiple excitation coils have also investigated optimizing the system matrix by adjusting the location and pattern of the sensors, the excitation coils, and the voxels [6, 60–63, 96]. Here, we will see if these principles are conserved for a system with much fewer sensor locations, a 3 dimensional field of view, and a sparse reconstruction method.

All three of the values of λ that we investigated increase as a function of the number of stage positions. This is necessary because the domain of feasible values of λ has an upper bound of $||\boldsymbol{b}||$, which in increases as the number of non-zero elements increases. The actual difference between the detected field and the true field increases approximately linearly with the norm of the signal vector, assuming the added noise is random, and with $1/\sqrt{s}$ for s stage positions. According to Equations 5.6 and 5.7, the values of λ_σ and λ_{SE} go as $\sqrt{7s + 2(\sqrt{14s})}$ with s stage positions. This means that the values of λ_{SE} and λ_σ increase more slowly with the number of stage positions than the magnitudes of the signal vector and corresponding noise vector. To explain this relationship, consider that as the number of stage positions increases, more information is available for the reconstruction, but benefit is tempered by the increase in the magnitude of the noise vector. However, the fact that λ_σ and λ_{SE} both overestimate and increase more slowly than the norm of the error vector means that they should more closely approximate the norm of the error vector

with additional stage positions.

The results are reported in most cases in terms of SNR_{dB} so that they can be generalized for systems with different noise characteristics to arrive at a final conclusion with respect to minimal detectable source strength. To account for the scaling of the source strength, the error in the reconstructed distribution ($\boldsymbol{x}^*$) was normalized to the magnitude of the true source distribution, as in Equation 5.11.

$$\text{Error}_x = \frac{||\boldsymbol{x}^* - \boldsymbol{x}_t||}{||\boldsymbol{x}_t||} \tag{5.11}$$

The error metric in Equation 5.11 is useful mathematically for deriving generalized reconstruction guarantees and error bounds, but it may not accurately quantify what an end user of SPMR would consider to be error. For example, if a dipole of the exact strength of the true solution is reconstructed even one pixel away from the true solution, the Error_x will be equal to $\sqrt{2}$. A user however, may consider that to be a pretty accurate solution because the moment accuracy is 100% and the location accuracy is one pixel. To better capture how a user would evaluate a reconstruction we defined a new metric called the region of interest accuracy. The ROI accuracy is a ratio of the sum of the reconstructed voxels within a designated ROI around the true source to the total strength of the true source, as in Equation 5.12. The results in terms of Error_x and further justification for using ROI accuracy are presented in 8.1.1.

$$\text{ROI accuracy} = \frac{\sum\limits_{x^* \in \text{ROI}} x^*}{\sum x_t} \tag{5.12}$$

The ROIs used to calculate the ROI accuracy for the single source and two-source distributions are shown in Figure 5.2. All ROIs were defined from -4 cm to -2 cm in z, inclusive, to avoid the noise at the top and bottom limits of the field of view. The extent in x and y was chosen to be large enough to wholly encompass any voxels near the intended source position and be of the same size without overlapping each other. For the single

source case, the ROI consisted of a 1 cm $\times$ 1 cm $\times$ 2 cm region that extended from -0.5 cm to 0.5 cm in x and y. For the two-source distribution, the ROI was split into two parts, one for each source. Each was the same size as the single source ROI, but shifted to center at (-1 cm, 1 cm) and (1 cm, -1 cm) in (x, y).

(a) Single source ROI (b) Two-source ROI
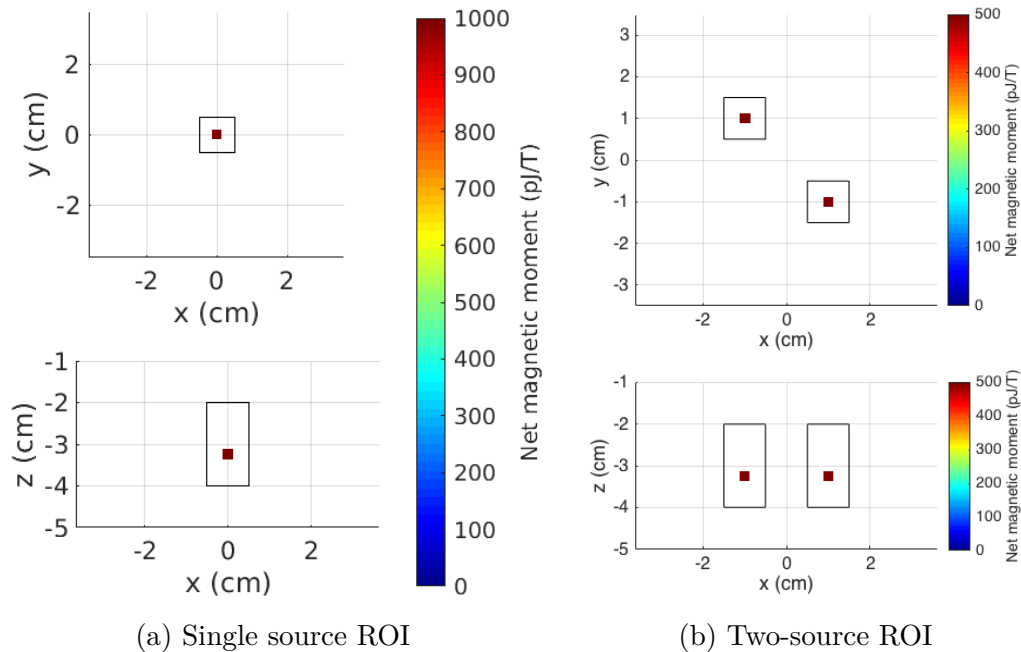
Figure 5.2: The regions of interest (ROI) around the true source location used in the calculation of the ROI accuracy for the single source (a) and two source (b) distributions.

5.3.1 Methods

First, the one and two-source distributions used in Section 5.4 were reconstructed with no additional noise for values of λ between 0 and $||\boldsymbol{b}||$. When no noise is added $\sigma = 0$, and therefore $\lambda_\sigma = \lambda_{SE} = 0$. The reconstructions were repeated with 1 to 9 stage positions, shown in Figure 5.3. The stage positions are listed in order in Table 5.1. These stage positions were selected because they result in a nearly even distribution of sensor locations across the field of view, and require only 0.5 cm interval movements. The ROI accuracy was calculated for each reconstruction as a function of λ. Finally, the accuracy was evaluated as a function of stage position at a constant value of $\lambda = 1.4 \times 10^{-4}||\boldsymbol{b}||$, to

account for roundoff error.

Stage position	x cm	y cm	z cm
1	0	0	0
2	-1	0	0
3	1	0	0
4	0	1	0
5	0	-1	0
6	-0.5	0.5	0
7	-0.5	-0.5	0
8	0.5	-0.5	0
9	0.5	0.5	0

Table 5.1: The stage positions used in the *in silico* study, listed in the order that they are added to the sequence.
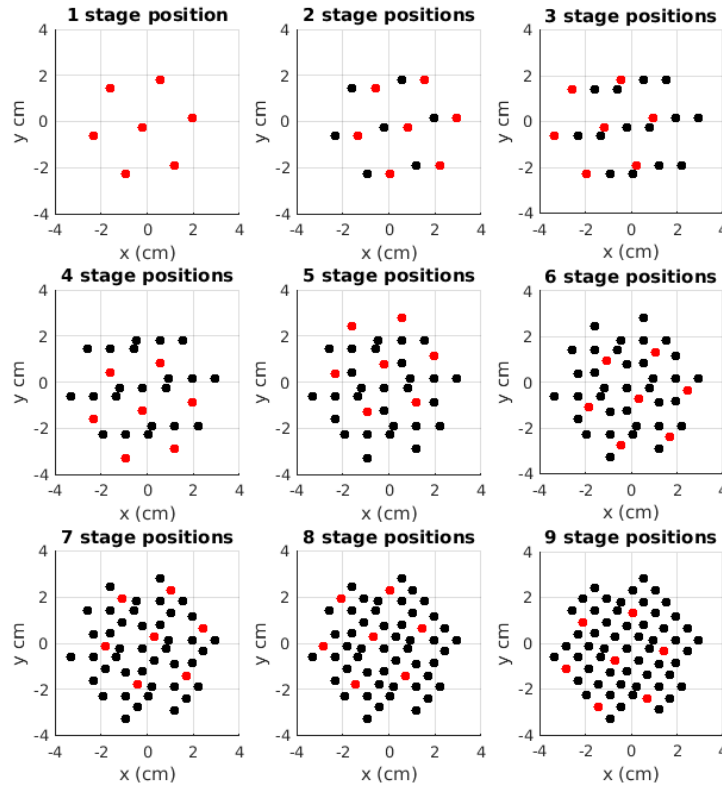
Figure 5.3: The 9 stage positions used in the investigation of stage positions, with the additional stage position shown in red.

Noise was added to the data at the SNR levels used in Section 5.4. An average over 30 pulses was reconstructed with 40 values of λ between $1 \times 10^{-4}||\boldsymbol{b}||$ and $||\boldsymbol{b}||$, plus λ_{SE} ,

λ_σ , and $\lambda_{\Delta b}$. The ROI accuracy was calculated for each reconstruction using the ROIs defined in Section 5.4.

5.3.2 Results

The ROI accuracy of the single source (solid lines) and two-source (dashed lines) reconstructions with no added noise is shown in Figure 5.4 as a function of $\lambda/||\boldsymbol{b}||$ for one to nine stage positions. The reason for plotting the accuracy as a function of λ normalized by the norm of the signal vector, rather than λ, is that it reduces the effects of the change in the length of the signal vectors as more stage positions are added and the differences in the magnitudes of the signal from the single and two-source distributions. These effects are only aesthetic and do not effect the overall results.
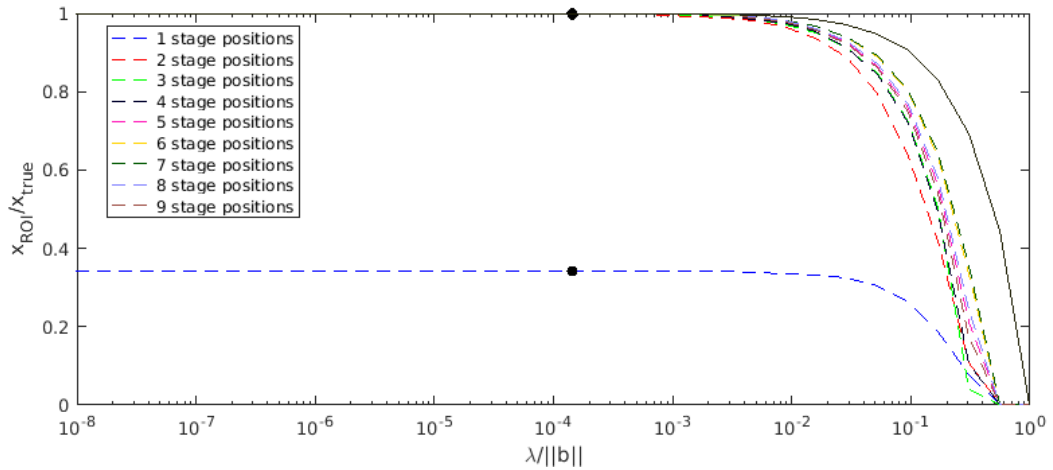
Figure 5.4: The ROI accuracy of the reconstruction of simulated data of a single source located at the center of the field of view (solid lines) and a two-source distribution (-1 cm, 1 cm) and (1 cm, -1 cm) from the center of the field of view, all 3.25 cm below the lowest coil of the central gradiometer, measured at one to nine stage positions, with no added noise, as a function of the signal-normalized data fidelity parameter.

For the single source reconstruction, the ROI accuracy is almost exactly linear with the data fidelity parameter (accounting for the semi-log scale), indicating that the data fidelity parameter is only changing the magnitude of the reconstructed source, and not the location of the source. When plotted against $\lambda/||\boldsymbol{b}||$, the curves overlap, indicating

that the number of stage positions has no effect on the reconstruction of a single point source. For the two-source distribution, the ROI accuracy is not linear for all values of the normalized data fidelity parameter. Additionally, the reconstruction from one stage position has a maximum accuracy of less than 40%, whereas the maximum accuracy of the reconstructions with two or more stage positions is 100%. This indicates that regardless of the SNR, a distribution of more than a single source cannot be adequately reconstructed using a single stage position. The dropoff in accuracy towards the higher values of λ varies across stage positions, rather than being identical as in the single source case. The decrease in accuracy occurs at the highest λs for the reconstructions with 6 and 7 stage positions. Finally, we see that the ROI accuracy reaches zero before the single source reconstructions, and before the limit of $\lambda = ||\boldsymbol{b}||$.
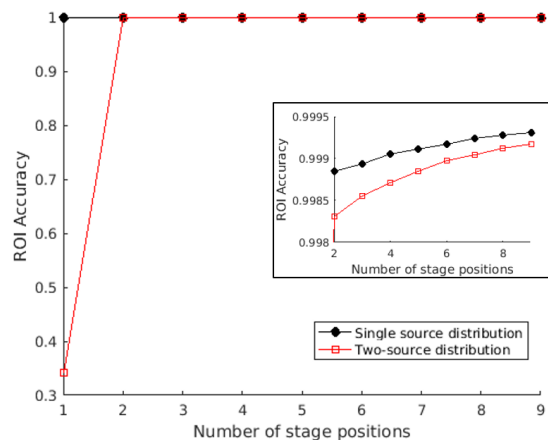
Figure 5.5: The ROI accuracy for reconstructions of measurements at 1 to 9 stage positions of a single source (black circles) and two-source (red squares) distribution in the presence of no noise. A data fidelity parameter of $\lambda = 1 \times 10^{-4}||\boldsymbol{b}||$ was used to account for numerical rounding error. The inset shows a zoomed-in view of the accuracy for two or more stage positions.

The accuracy of the single and two-source reconstructions at the value of $\lambda = 1 \times 10^{-4}||b||$ (plotted as black dots in Figure 5.4) is shown in Figure 5.5. This clearly shows that the accuracy two-source reconstruction greatly improves with a second stage position, but shows only incremental improvement with each additional stage position. The reconstruction of a single source improves only slightly with each additional stage positions (see

inset of Figure 5.5).
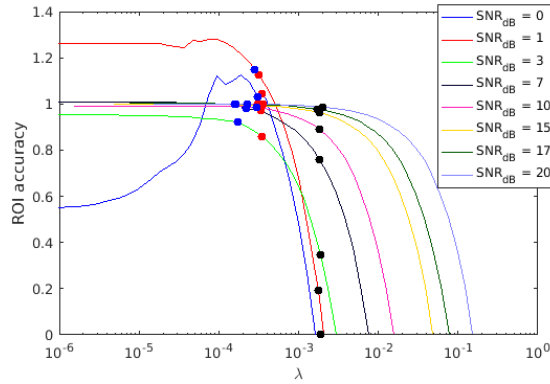
(a) Single source distribution measured at one stage position
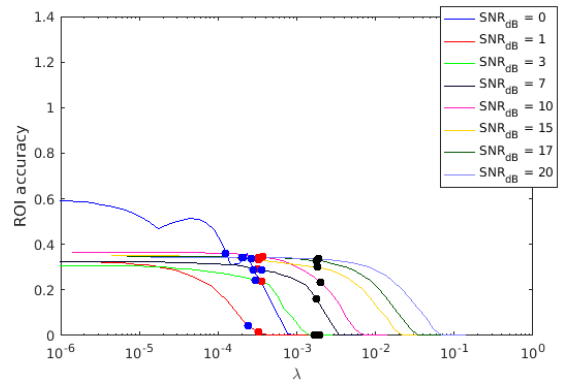
(b) Two-source distribution measured at one stage position
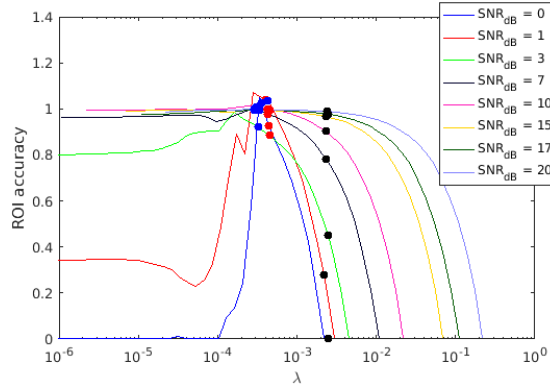
(c) Single source distribution measured at two stage positions
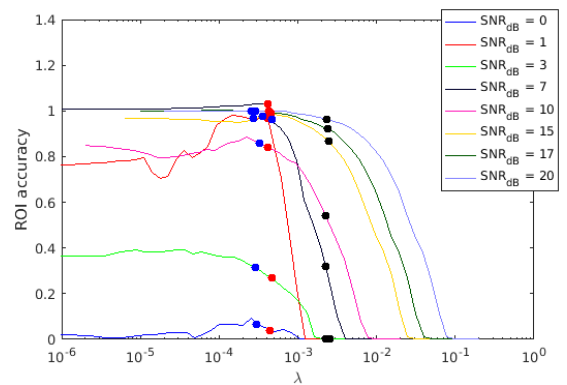
(d) Two-source distribution measured at two stage positions

Figure 5.6: The ROI accuracy as a function of λ for reconstructions of simulated measurements taken at one (top row) and two (bottom row) stage positions of distributions containing a single source (left column) and two sources (right column), with various levels of SNR. The reconstructions at $\lambda_{\Delta b}$, λ_{SE} , and λ_{σ} are denoted by blue, red, and black dots, respectively.

These results hold when noise is added. Figure 5.6 shows the ROI accuracy of the reconstructions at various levels of SNR as a function of λ for the single source (left column) and two-source (right column) distributions when one stage position (top row) and two stage positions (bottom row) are used. The reconstructions at $\lambda_{\Delta b}$, λ_{SE} , and λ_{σ} are denoted by blue, red, and black dots, respectively. These are plotted with the actual value of λ rather than the normalized value because all of the curves in a single graph are from simulations with the same number of stage positions. This also shows how the actual values of $\lambda_{\Delta b}$,

λ_{SE}, and λ_σ remain consistent across all SNR, because the signal is scaled to vary the SNR while the noise is constant. We can see that for simulations with high SNR, the curves resemble the shape of those from the simulations with zero noise, only they are shifted horizontally by a factor of $||\boldsymbol{b}||$. For simulations with low SNR, the curves tend to show high variability for low values of λ, then regain linearity as λ increases. The early part of these curves indicates the values of λ for which the reconstruction is fitting the noise. Values of ROI accuracy greater than 1 indicate that the reconstruction is trying to compensate for the noise by reconstructing a source slightly larger than the true source and slightly further from the detectors, but still within the ROI. For the single source simulations in (a) and (c), the ROI accuracy of the reconstructions at $\lambda_{\Delta\boldsymbol{b}}$ and λ_{SE} tends to be close to 1, especially for the cases with SNR greater than 7 dB. For these values of λ we see a slight improvement with the second stage position, especially for values of SNR less than 7 dB. In both cases, the reconstructions at λ_σ overestimate the optimal value of λ. These reconstructions lie within the region of the curve for which the reconstruction has converged to a single voxel, and the excess error allowed by λ is accounted for by decreasing the magnitude of the source. At high values of SNR, specifically above 10 dB, the magnitude of the source is large enough that this does not greatly affect the accuracy, but for SNR less than 10 dB this overestimation of λ severely decreases the accuracy of the reconstruction.

The accuracy of the reconstructions of the two-source simulations confirm that the improvement in the ROI accuracy between one and two stage positions holds independent of λ and SNR. As seen in Figure 5.5, in general, the highest achievable value of ROI accuracy is 0.4, even for the highest levels of SNR. Of course this is only serves to confirm what one would expect, since there is no reason that introducing additional noise should improve the accuracy of the reconstruction. When a second stage position is used, the reconstructions with high SNR more than double in accuracy. However little, if any, improvement is seen for the lowest values of SNR, most notably in the case of 0 SNR, which hardly reaches an ROI accuracy of 0.1. It is unexpected that the accuracy of the reconstruction with an SNR

(a) One source: $\lambda = \lambda_\sigma$

(b) Two sources: $\lambda = \lambda_\sigma$

(c) One source: $\lambda = \lambda_{SE}$

(d) Two sources: $\lambda = \lambda_{SE}$

(e) One source: $\lambda = ||b * -b||$
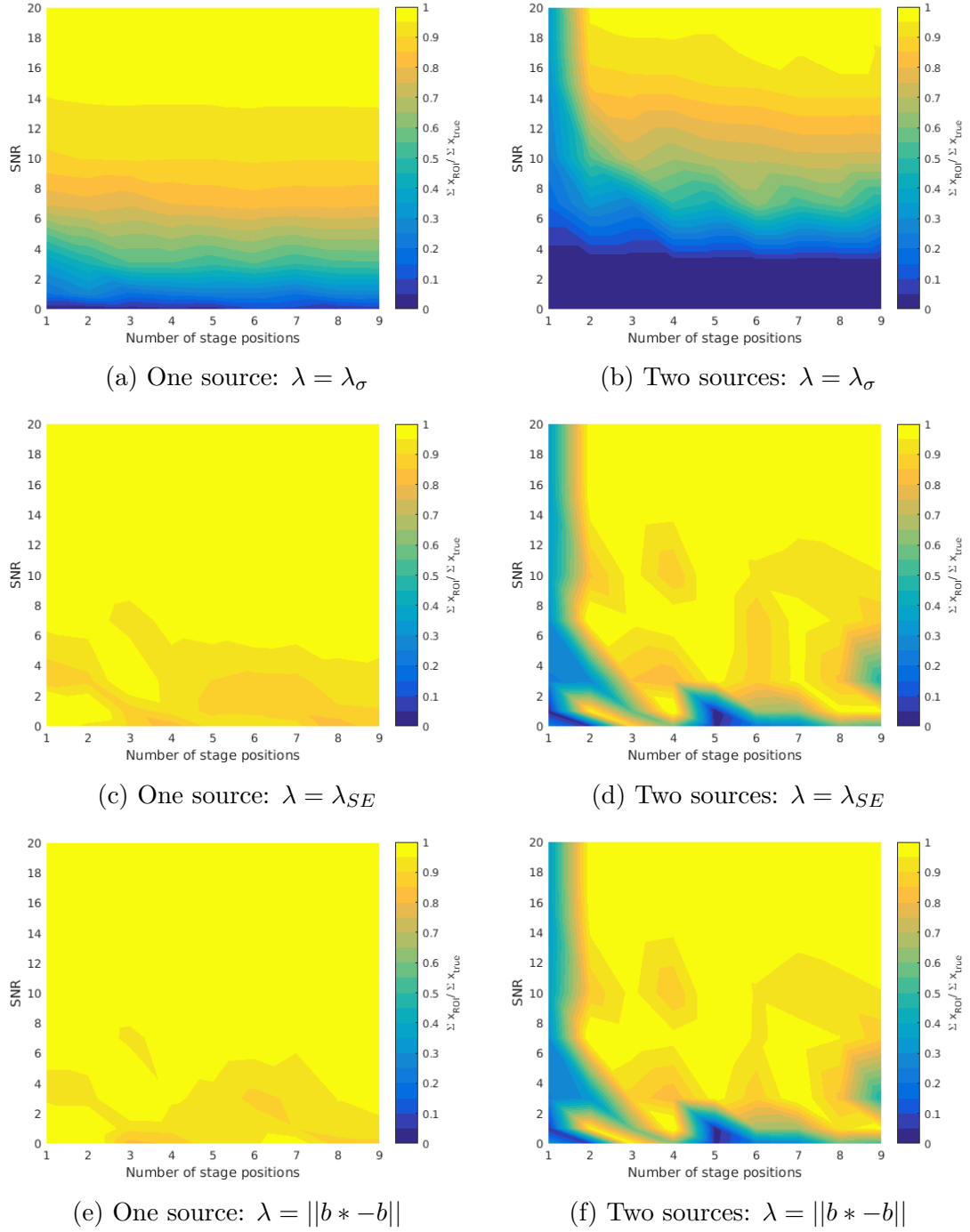
(f) Two sources: $\lambda = ||b * -b||$

Figure 5.7: The ROI accuracy of reconstructions of a single source (left column) and two-source (right column) distributions, as a function of the number of stage positions and SNR. Reconstructions are of data averaged over 30 simulated pulses, with three choices of λ: (a and b) λ_σ , (c and d) λ_{SE} , and (e and f) $\lambda_{\Delta\boldsymbol{b}}$.

of 1 saw more benefit from the second stage position than the reconstruction with an SNR of 3. This may be due to the nature of the noise being randomly sampled, and perhaps the average of the samples from the noise distribution taken for the simulation with SNR of 1 happened to be smaller than those taken for the SNR of 3. Using 30 samples in each simulation was designed to minimize the probability of such an occurrence, but it could not be avoided entirely. Repeating the simulations with identical noise added would be able to test this. Additionally, the cause of this unexpected result may become clear upon observation of the reconstructions along each curve.

For all three choices of λ, the accuracy of the single source reconstructions was fairly constant across all the stage positions, whereas the two-source cases showed a large increase in accuracy between one and two stage positions. Above two stage positions, the accuracy increased slightly for additional stage positions for the two-source cases, but stayed approximately constant for the single source cases. For both one and two sources, λ_{SE} resulted in higher accuracy than λ_σ . In fact, for the single source cases, the minimum accuracy of reconstructions at λ_{SE} was 80% at all levels of SNR, and 100% for SNR over 10 across all stage positions. In contrast, reconstructions at λ_σ reached 80% accuracy only for SNR above 7 dB. Reconstructions of the two-source distribution at λ_σ were less than 40% accurate below 7 dB, regardless of the number of stage positions. When at least two stage positions were used, the two-source reconstructions with λ_{SE} were 80% accurate or better above 3 dB. The accuracy was highly variable below 3 dB across all stage positions.

5.3.3   Stage positions and the conditioning of A

Much of the work in the literature on designing and quantifying the ability of a system matrix to reconstruct a solution as a function of the number of sources or the noise is based on the conditioning number of the matrix. However, this quantity is only applicable to systems in which the relationship between the reconstruction and the signal vector can be represented as $\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}$. This is true in methods that are based on minimizing the least-

squares difference between the reconstructed and measured signals, such as the TSVD and minimum norm methods In these cases, smaller values of the conditioning number correspond to more stable reconstructions in the presence of noise. The condition number, $\kappa(\mathbf{A}) = ||\mathbf{A}^{-1}||||\mathbf{A}||$, in Equation 5.13 describes the relationship between the true solution and signal $(\boldsymbol{x}, \boldsymbol{b})$ and reconstructed solution and detected signal $(\boldsymbol{x}^*, \boldsymbol{b}^*)$. Such a metric would allow for the $\mathbf{A}$ matrix to be optimized as a function of number and location of stage positions, voxel size, and extent of the field of view, without explicit studies of each possible configuration.

$$\frac{||\boldsymbol{x} - \boldsymbol{x}^*||}{||\boldsymbol{x}||} = \kappa(\mathbf{A})\frac{||\boldsymbol{b} - \boldsymbol{b}^*||}{||\boldsymbol{b}||} \tag{5.13}$$

We calculated the condition number of the $\mathbf{A}$ matrix for each of the 1 through 9 stage positions. We then plotted the error in the reconstruction, defined by the left hand side of Equation 5.13, versus the error in the detected signal, defined by the fraction on the right hand side of Equation 5.13. If the conditioning number can be used as a measure of the quality of the reconstruction in the presence of noise, then the error in the reconstruction should be linear to the error in the detected field with a slope equal to the condition number of the matrix. Since this should hold for any source distribution, we examine both the one and two-source distributions.

Figure 5.8 shows the error in the reconstruction of the one (a) and two-source (b) distributions as a function of the error in the detected field for one through nine stage positions. Figure 5.8 (c) shows a plot of the slope of the error in $\boldsymbol{x}$ versus the error in $\boldsymbol{b}$ as a function of stage position for the one (dashed line) and two (solid line) source distributions, as well as the conditioning of the system matrix (blue). We see that the change in the solution relative to the change in the signal decreases with an increasing number of stage positions. However, this relationship is not accurately described by the condition number of the system matrix, which increases as the number of stage positions increases.
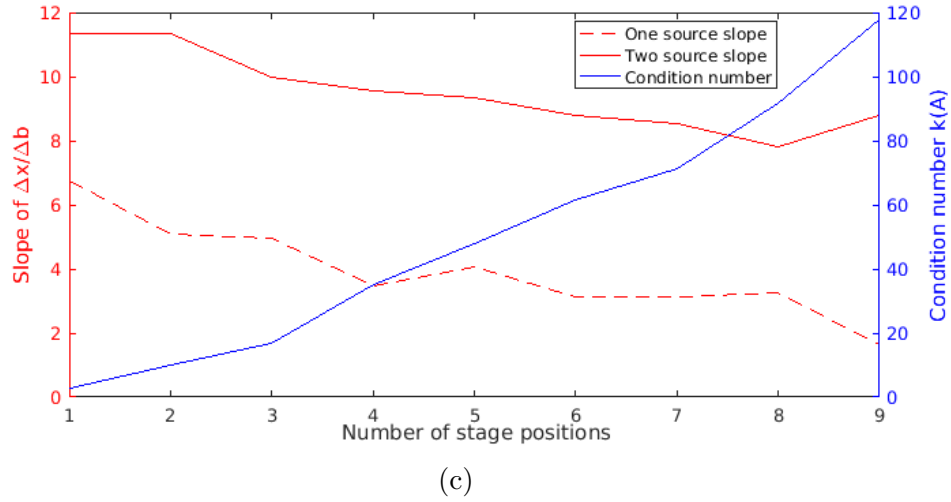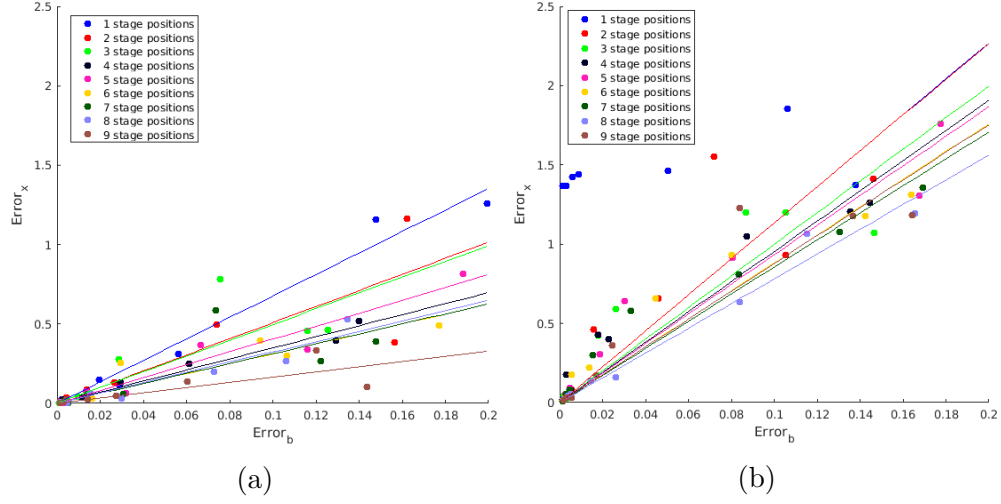
Figure 5.8: The effect of the number of stage positions on the condition number $\kappa(\mathbf{A})$. Parts (a) and (b) show the linear relationship between the error in the reconstruction ($\text{Error}_x = \frac{\|\boldsymbol{x} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}\|}$) and the error in the signal ($\text{Error}_b = \frac{\|\boldsymbol{b} - \boldsymbol{b}^*\|}{\|\boldsymbol{b}\|}$) for reconstructions of simulated measurements of a single (a) and two-source (b) distribution (see Figure 5.2) taken at between 1 and 9 stage positions with various levels of added Gaussian noise. In (c), the relationship between the change in the reconstruction and the change in the measured field is not described by the conditioning number.

95

5.4 Effect of the number of pulses

5.4.1 Methods

The single source simulations followed the procedure in 5.2 with a baseline dipole with a magnitude of $1 \times 10^4 \, \text{pJ T}^{-1}$ located at the center of the field of view (x,y = 0), 3.25 cm below the bottom coil of the central gradiometer. The baseline two-source distribution consisted of one $500 \times 10^3 \, \text{pJ T}^{-1}$ dipole at (-1 cm, 1 cm -3.25 cm) and one $500 \times 10^3 \, \text{pJ T}^{-1}$ dipole at (1 cm, -1 cm $-3.25 \, \text{cm}$). The simulated SPMR measurement protocol consisted of 1, 5, 10, 15, 20, 25 or 30 pulses (also called samples) collected at each of three stage positions in (x cm, y cm, z cm): (0, 0, 0), (-1, 0, 0) and (1, 0, 0). For the single pulse scenario, the standard deviation of the noise vector was used to estimate σ because the typical procedure results in $\lambda_{SE} = \lambda_\sigma = 0$ when the N=1. The noise-corrupted signal was reconstructed using 40 values of λ logarithmically distributed between 0 and $||\boldsymbol{b}^*||$, plus the values $\lambda_{\Delta \boldsymbol{b}}$, λ_{SE} and λ_σ .

5.4.2 Results

Figure 5.9 shows the three estimates of λ as a function of the number of pulses per stage position. The mean over the eight levels of SNR from 0 to 20 dB is plotted for the one source (a) and two source (b) scenarios, with error bars representing the standard deviation over SNR. It can be seen that λ_{SE} provides a reliable upper bound to the true error, and decreases as the number of pulses per stage position increases, allowing the reconstruction to more closely fit the measured data, whereas λ_σ overestimates the upper bound of the error and is constant even as the number of pulses increases.

The ROI accuracy for each of the three choices of λ as a function of SNR and number of pulses is shown in Figure 5.10. The top row shows the ROI accuracy for the reconstructions at λ_σ of the single (left column) and two-source (right column) simulations. It can be seen that, except for some variation due to the random sampling of the noise, the accuracy is not
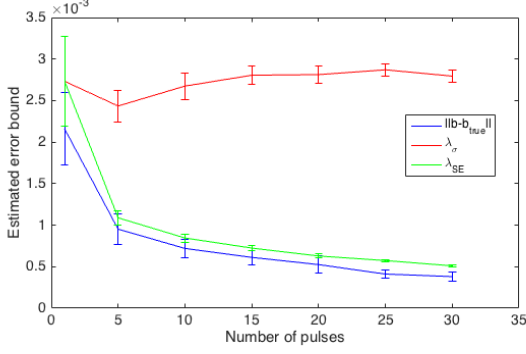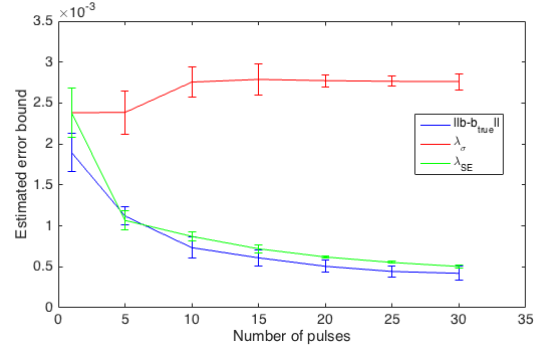
(a) Single source simulation   (b) Two-source simulation

Figure 5.9: The mean and standard deviation of the true error in the measured magnetic field vector ($\lambda_{\Delta b}$), and values of λ_σ and λ_{SE} for simulated cases of one (a) and two (b) sources across signal to noise ratios of 0 to 20 dB. The simulations consisted of 1 to 30 pulses taken at each of three stage positions (x cm, y cm, z cm): (0, 0, 0), (-1, 0, 0) and (1, 0, 0). The noise vector was randomly sampled from a normal distribution with zero mean and a standard deviation of 4.9×10^{-4} according to the results in Chapter 4. The simulated field and dipole were scaled to achieve the desired signal to noise ratio.

improved by increasing the number of pulses collected per stage position. This is because λ_σ does not account for the number of samples, but only considers it implicitly in the estimation of the noise. With more samples, the estimation of the mean and of the standard deviation between the measurements improves, which is likely why we see the accuracy slightly decrease with the number of samples. This may happen if the true standard deviation is being underestimated by the small sample sizes. These results emphasize how the data fidelity parameter ultimately determines the accuracy of the reconstruction. By not allowing the parameter to decrease as the estimation of the mean improves, none of the benefit of the additional pulses is seen in the reconstruction.

For both the single and two-source simulations, the ROI accuracy for the reconstructions at λ_{SE}, closely resembles that at $\lambda_{\Delta b}$. This suggests that λ_{SE}, which can be calculated based on the available data, is a suitable approximation of $\lambda_{\Delta b}$, which cannot be known *a priori* because it depends on the true signal. For both λ_{SE}, and $\lambda_{\Delta b}$, the accuracy of the reconstructions based on only a single measurement closely resembles that of the reconstructions at λ_σ. This is expected since for a single pulse, λ_{SE} is equal to λ_σ
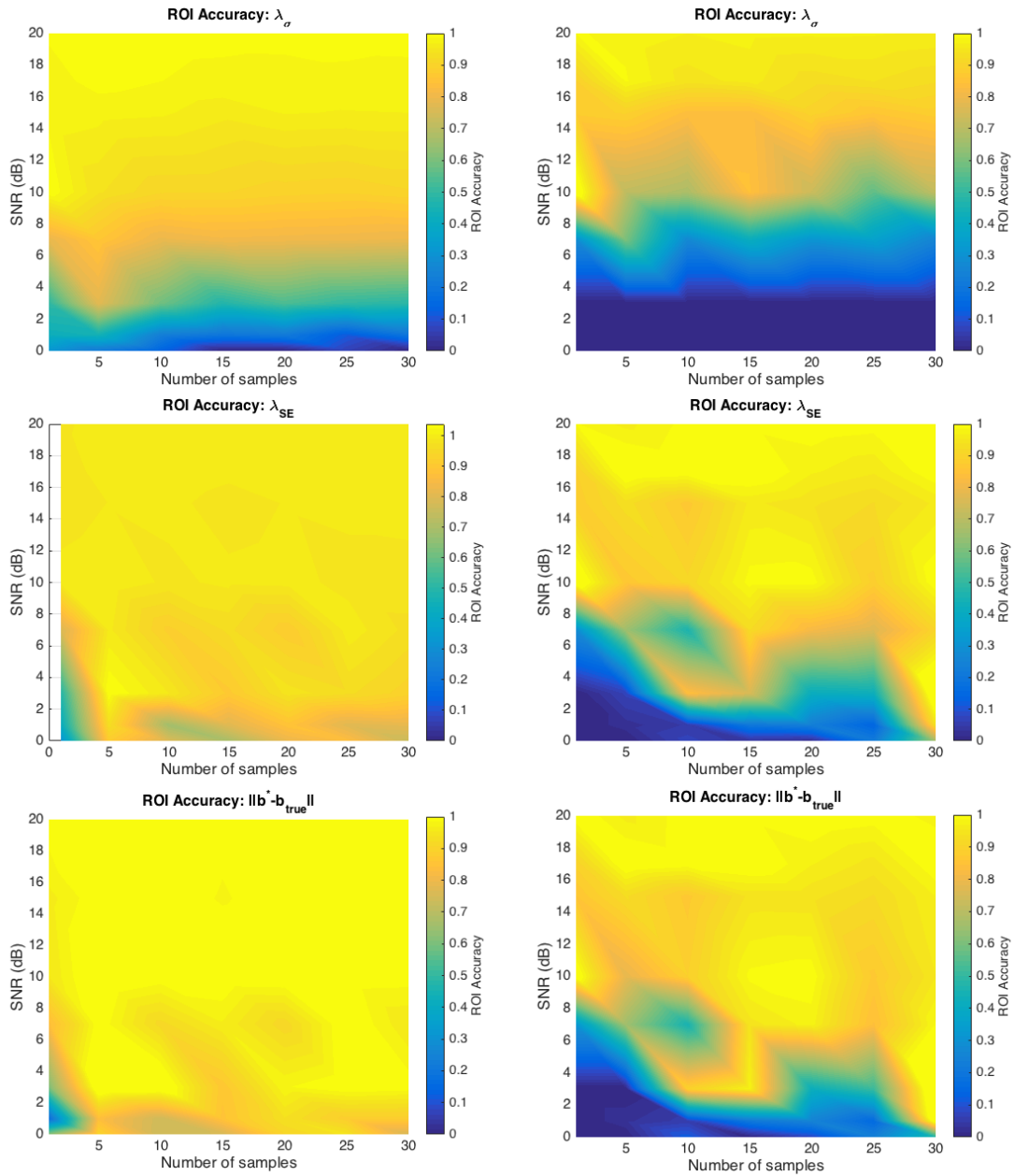
Figure 5.10: The ROI accuracy of the reconstructions of simulations consisting of a single (left column) and two-source (right column) distribution as a function of the number of samples (pulses per stage position), at three values of λ: λ_{SE}, λ_{σ}, and $\lambda_{\Delta \boldsymbol{b}}$ for values of signal to noise ratio (SNR) from 0 dB to 20 dB. SNR was achieved by increasing the strength of the simulated dipole(s) with a constant noise level.

. For five or more pulses, the accuracy is high (> 0.8) for nearly all of the single-source reconstructions, except for those that used only a single pulse at the lowest levels of SNR. For the two-source simulations, the accuracy was consistently high at SNR levels greater than 10 dB, and improved with additional pulses as SNR decreased. This implies that the extra time required to collect additional pulses (above 5 or 10) may not improve the accuracy of the reconstruction of a single source. However when multiple sources are present, additional pulses can vastly improve the accuracy of the reconstruction in cases of low SNR.

To look at the relationship between the of the number of pulses and λ, the ROI accuracy as a function of λ for the reconstructions of 10 pulses and 30 pulses. The results for the single source distribution are presented in Figure 5.11. The accuracy of the reconstructions at $\lambda_{\Delta b}$, λ_{SE} , and λ_σ are denoted by blue, red, and black dots, respectively. Figure 5.12 shows the reconstructions at four values of λ (one within the first few values of λ, one at which the Euclidean distance between the truth and the reconstruction is minimized, and the values of $\lambda_{\Delta b}$, λ_{SE} , and λ_σ) at three levels of SNR (1 dB, 7 dB, and 17 dB) which represent a small, intermediate, and strong source, respectively. As in the previous sections, the curves at high SNR follow the same trend where the accuracy is linear to the data fidelity parameter. This is consistent for both 10 and 30 pulses per stage position, indicating that at high SNR (greater than 10 dB) acquiring 30 pulses per stage position confers little benefit over 10 pulses per stage position for a single source distribution. At lower levels of SNR, we see that the accuracy of the reconstructions with 30 pulses is less sensitive to the choice of data fidelity parameter than the reconstructions with 10 pulses for small parameter values. This means that if the parameter is chosen manually within this region, the reconstructions from an average of 10 pulses could vary greatly based on the choice of parameter. When the reconstruction is based on the average of 30 pulses, the reconstruction is less dependent on the choice of the parameter. This is especially true for SNR levels of 3 dB and below. However, at λ_{SE} , the accuracy of the reconstruction is similar between reconstructions of 10 and 30 pulses, as shown Figure 5.10. This emphasizes
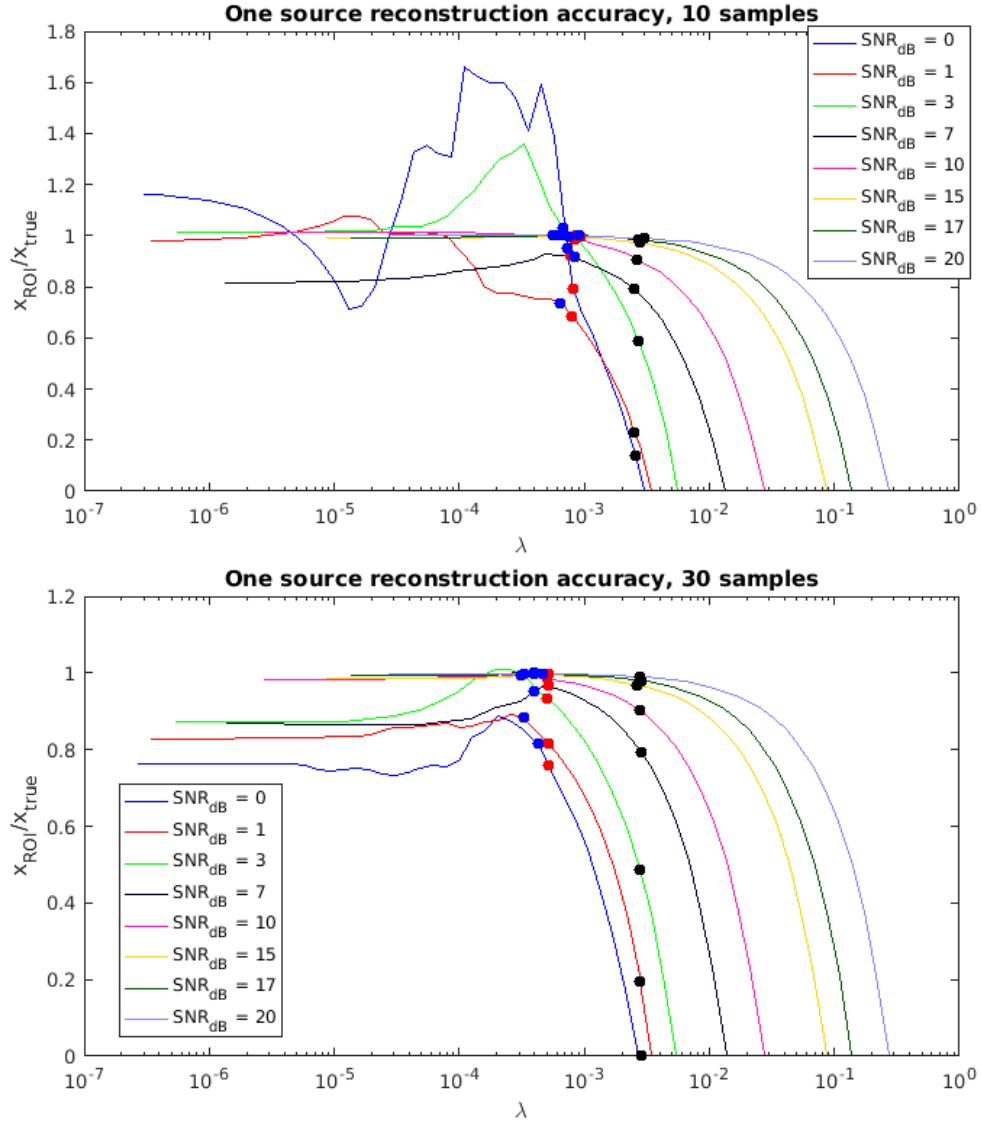
Figure 5.11: The ROI accuracy (Equation 5.12) of the reconstruction of a simulated single source distribution measured at three stage positions using 10 pulses per stage position (top) and 30 pulses per stage position (bottom) for signal to noise ratios from 0 to 20 dB. The error from reconstructions using $\lambda_{\Delta b}$, λ_{SE} , and λ_{σ} are denoted by blue, red, and black dots, respectively.

the benefit of explicit parameter selection based on the data over manual selection based on empirical evaluation of the reconstructions.
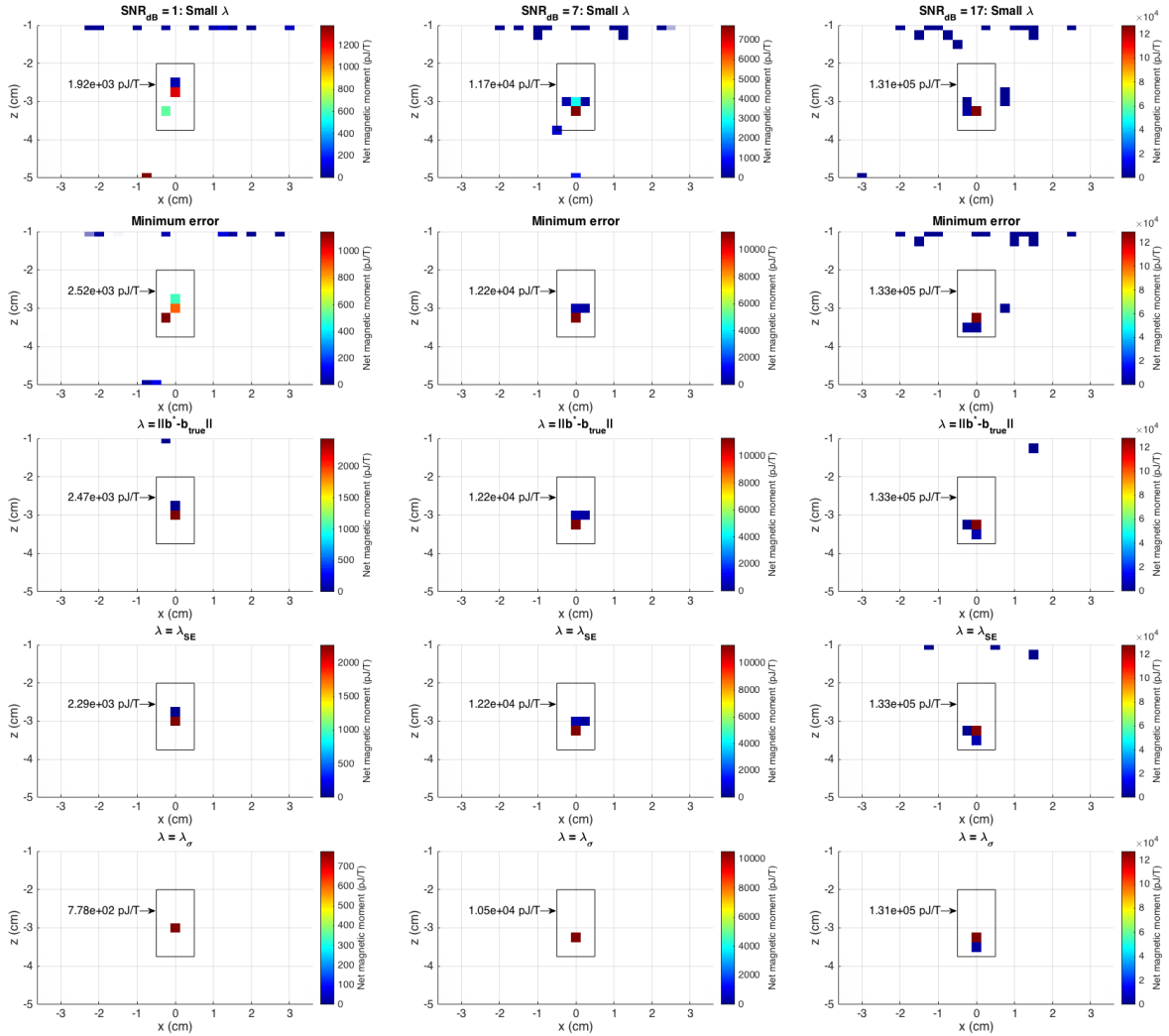
Figure 5.12: Selected reconstructions of simulated measurements of a single source consisting of 10 pulses, as presented in Figure 5.11 for SNR levels of 1 (left column), 7 (middle column), and 17 (right column) dB. The first row shows a reconstruction with a small value of λ. The second row shows the reconstruction with the minimum distance from the true solution. The third row shows the reconstruction with $\lambda_{\Delta b}$. The fourth and fifth rows show the reconstructions with λ_{SE} and λ_{σ}, respectively. The total moment within an region of interest drawn as a box is shown to the left of the reconstructed source. For comparison, the true source strength for SNRs of 1, 7, and 17 dB, were $3.35 \times 10^3 \, \mathrm{pJ \, T^{-1}}$, $1.33 \times 10^4 \, \mathrm{pJ \, T^{-1}}$, and $1.33 \times 10^5 \, \mathrm{pJ \, T^{-1}}$, respectively.

Figures 5.12 and 5.13 provide further insight into the effect of the choice of parameter on the reconstruction of a single source at various levels of SNR. Figure 5.12 shows

the reconstructions based on the average of 10 simulated pulses per stage position. For reference, the true source distributions are shown in Figure 5.2. Recall that for this study, the SNR was achieved by increasing the strength of the simulated source in the presence of a constant noise level ($n \in \mathcal{N}(0, \sigma)$). For our device, we determined (in Chapter 4) $\sigma = 4.9 \times 10^{-4}$ pT. At this noise level, the SNR levels of 1, 7 and 17 correspond to a source strength of 3.35×10^3 pJ T^{-1}, 1.33×10^4 pJ T^{-1}, and 1.33×10^5 pJ T^{-1}, respectively. The reconstructions at an SNR of 1 demonstrate the effect of increasing λ that was described in Chapter 3. When λ is small, there is a substantial number of non-zero voxels at the very top and very bottom of the field of view. This is due to the algorithm meeting the constraint of reconstructing noisy data within a small tolerance. As λ increases, this noise disappears, and the source within the ROI increases to compensate for the detected field. For values of $\lambda > \lambda_{\Delta b}$, the only non-zero voxels are within the ROI, and their values begin to decrease with increasing λ. It is interesting to note that the best reconstruction (defined as the minimum Euclidean distance between the true source distribution and the reconstructed source distribution) is the closest reconstruction of the total magnitude of the simulated source, but is not free of noise. This indicates that simply selecting the value of λ for which there is no noise may result in an under estimation of the true source strength. However at this level of SNR, the values of λ_{SE} and λ_{σ} also underestimate the true magnitude of the source. The error was substantially worse in the latter case, which underestimated the true source strength by nearly 40%. While the magnitude of the source was highly variable with each choice of λ, at each value shown here, the source was reconstructed within one voxel of its true location. This indicates that at low SNR, either due to high uncertainty in the detected field or low source strength, even if the location of a source is correctly detected, the magnitude of the source may be underestimated.

At higher levels of SNR, we see the same pattern of first a reduction in the noise at the top and bottom of the field of view accompanied by an increase in the source within the ROI, followed by a reduction of the magnitude of the source within the ROI with increasing

λ. However, the variation in the magnitude of the source as a function of λ is much less dramatic than in the case with SNR of 1. In fact, for the case with an SNR of 7, the value of λ which resulted in the minimum Euclidean distance from the true solution was equal to $\lambda_{\Delta b}$ and λ_{SE} . At both SNR levels of 7 and 17, the maximum valued voxel corresponded to the voxel containing the true simulated source across all of the values of λ shown in Figure 5.15. Only the noise and the overall strength of the reconstructed source varied with λ. For the high SNR case (SNR of 17), the accuracy is higher across all of the λs sampled, and the noise persists through λ_{SE} .

When the reconstruction is based on the average of 30 pulses, shown in Figure 5.13, the accuracy of each reconstruction increases. In contrast to the reconstructions based on only 10 pulses, the location of the maximum voxel corresponds to the true location of the simulated source even at the SNR level of 1. Additionally, the nonzero voxels outside of the ROI are substantially reduced for all but the smallest values of λ. Finally, the reconstruction seems to be more spread out, or less sparse, around the location of the simulated source than for the reconstructions from 10 pulses.

The accuracy of the reconstruction of the two-source simulation, presented on the left of Figure 5.10, is shown as a function of λ for 10 pulses per stage position and 30 pulses per stage position in Figure 5.14. Overall, the curves follow the same general trend that wee have seen in the previous plots. However, the difference the extra pulses make is much more apparent in the two-source case than for the single source reconstruction. With only 10 samples per stage position, only the highest levels of SNR achieve ROI accuracy levels near 1. The accuracy of the mid-range SNR levels hover between 0.8 and 0.9 for $\lambda = \lambda_{SE}$ and $\lambda_{\Delta b}$. At an SNR of 7 dB, the accuracy is not better than 0.6, and for the SNRs of 0 and 1, the accuracy is less than 0.2 for all values of λ. When 30 pulses are averaged, the accuracy at all levels of SNR above zero increases to above 0.8 for $\lambda = \lambda_{SE}$ and $\lambda_{\Delta b}$. Even the accuracy of the reconstruction in zero SNR is greater than 0.6 at $\lambda = \lambda_{SE}$. As before, the accuracy at λ_σ is lower than λ_{SE} at all levels of SNR, and falls within the region of
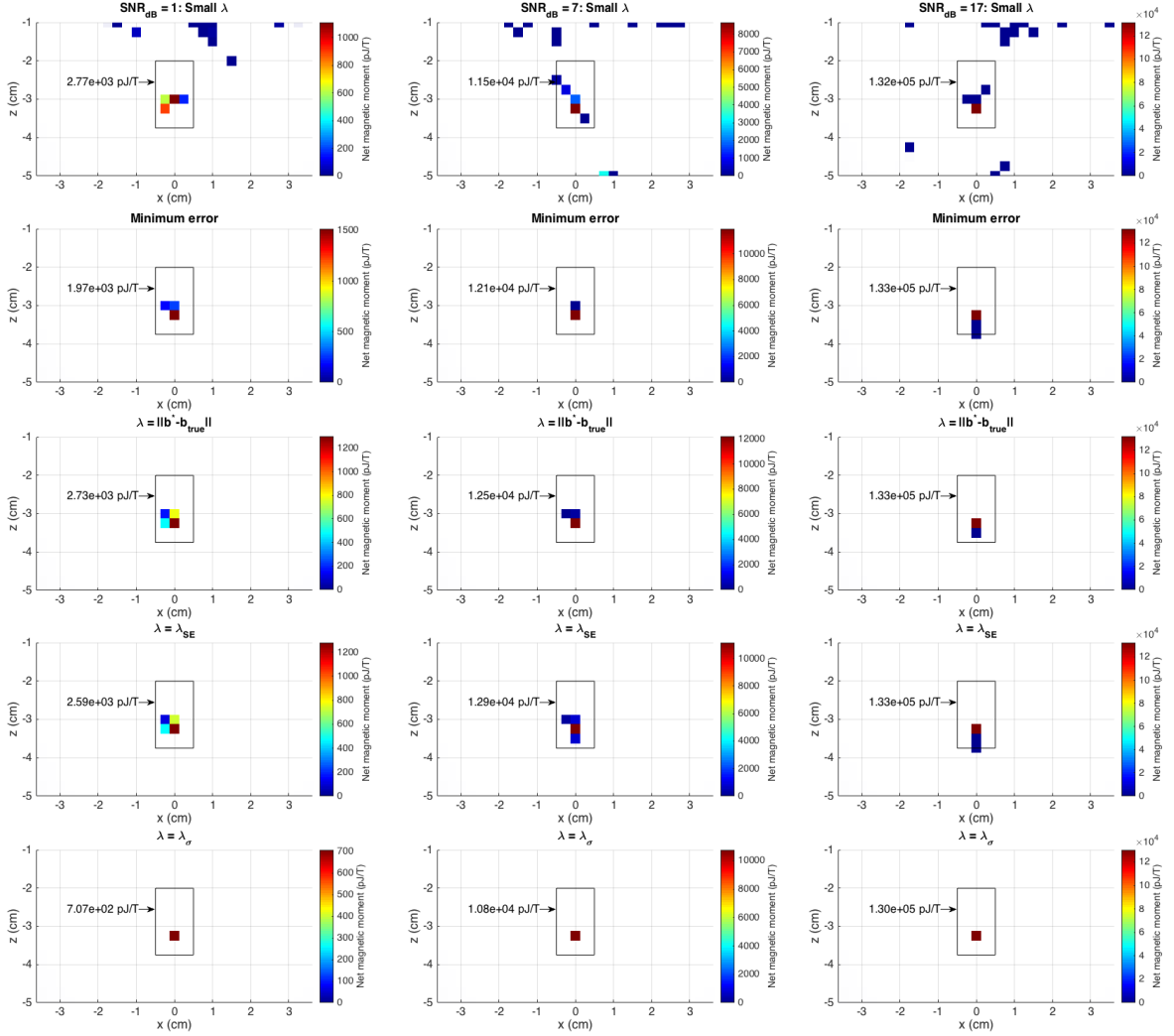
Figure 5.13: Selected reconstructions from the 30-sample data presented in Figure 5.11 for SNR levels of 1 (left column), 7 (middle column), and 17 (right column) dB. The first row shows a reconstruction with a small value of λ. The second row shows the reconstruction with the minimum distance from the true solution. The third row shows the reconstruction with $\lambda_{\Delta b}$. The fourth and fifth rows show the reconstructions with λ_{SE} and λ_{σ}, respectively. The total moment within an region of interest drawn as a box is shown to the left of the reconstructed source. For comparison, the true source strength for SNRs of 1, 7, and 17 dB, were $3.35 \times 10^3 \, \text{pJ T}^{-1}$, $1.33 \times 10^4 \, \text{pJ T}^{-1}$, and $1.33 \times 10^5 \, \text{pJ T}^{-1}$, respectively.

104

Figure 5.14: The ROI accuracy for the two-source reconstruction using 10 pulses per stage position (top) and 30 pulses per stage position (bottom) for signal to noise ratios from 0 to 20 dB. The accuracy of the reconstructions using $\lambda_{\Delta b}$, λ_{SE} , and λ_{σ} are denoted by blue, red, and black dots, respectively.

the curve for which the location of the reconstruction is no longer changing but only the magnitude of the reconstructed source is decreasing with λ.
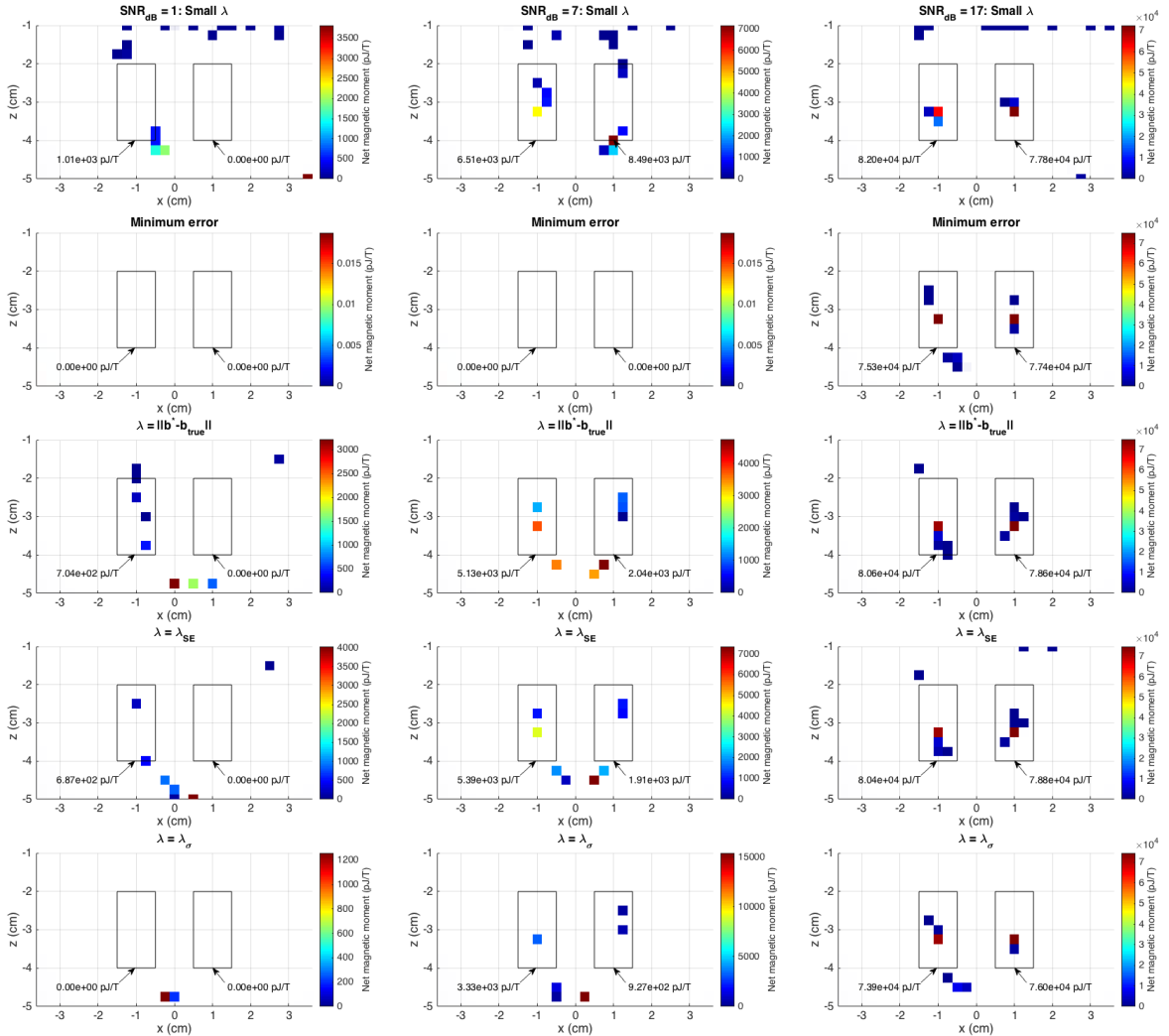
Figure 5.15: Selected reconstructions from the 10-sample data presented in Figure 5.14 for SNR levels of 1 (left column), 7 (middle column), and 17 (right column) dB. The first row shows a reconstruction with a small value of λ. The second row shows the reconstruction with the minimum distance from the true solution. The third row shows the reconstruction with $\lambda_{\Delta b}$. The fourth and fifth rows show the reconstructions with λ_{SE} and λ_{σ} , respectively. The total moment within an region of interest drawn as a box is shown to the left of the reconstructed source. For comparison, the true source strength for SNRs of 1, 7, and 17 dB, were $2 \times 10^3 \, \mathrm{pJ \, T^{-1}}$, $7.9 \times 10^3 \, \mathrm{pJ \, T^{-1}}$, and $7.9 \times 10^4 \, \mathrm{pJ \, T^{-1}}$, respectively.

These results can be better understood by examining the reconstructions at selected points along these curves. The reconstructions based on the average of 10 pulses per stage position are shown in Figure 5.15. For reference, the true simulated source distribution is

106

shown in Figure 5.2. For the reconstruction with low SNR, the algorithm fails to reconstruct any non-zero component of the source within the right-hand ROI at any value of λ. We also see that the reconstruction with the minimum Euclidean error is all zeros, which further reinforces the shortcomings of using this as an error metric. For the values of $\lambda_{\Delta b}$, λ_{SE}, and λ_σ, we see that the largest component of the reconstruction is placed below and in the center of the ROIs. This is an artifact of the algorithm attempting to find the fewest number of sources that can reconstruct the detected field within a given allowance of error for the noise. In this case, we see that the difference between the field from a single source towards the bottom of the field of view and the detected field is less than the difference between the field from the true distribution and the detected field. As the SNR increases, the concentration of sources at the bottom of the field of view is split into two, with smaller components showing up inside the ROI. Eventually at the highest level of SNR, the sources begin to match the true source distribution.

When the reconstructions are based on the average of 30 pulses per stage position, the detected magnetic field is better defined, and can no longer be approximated with a single source at the bottom of the field of view for all values of λ, even in low SNR scenarios. This explains the dramatic increase in accuracy between the cases with 10 pulses and 30 pulses that we see in Figure 5.14. For the reconstructions with SNR of 1, there is at least some non-zero signal reconstructed within the ROI for $\lambda_{\Delta b}$ and λ_{SE}. This is further improved for each sampled value of λ as the SNR increases. In fact, only the reconstructions at λ_σ show any substantial non-zero component below the ROIs. Additionally, it is interesting to note that while the magnitude of the source within the ROI was underestimated in every reconstruction of the single source simulation, the reconstructions of the two-source simulations in some cases overestimate the total magnitude of the source. This too is due to the sparsity enforcing term in the optimization, which tends to favor a single stronger source farther from the detectors over multiple sources closer to the detectors. Overall, these results emphasize the benefit of repeated measurements when reconstructing multiple

107
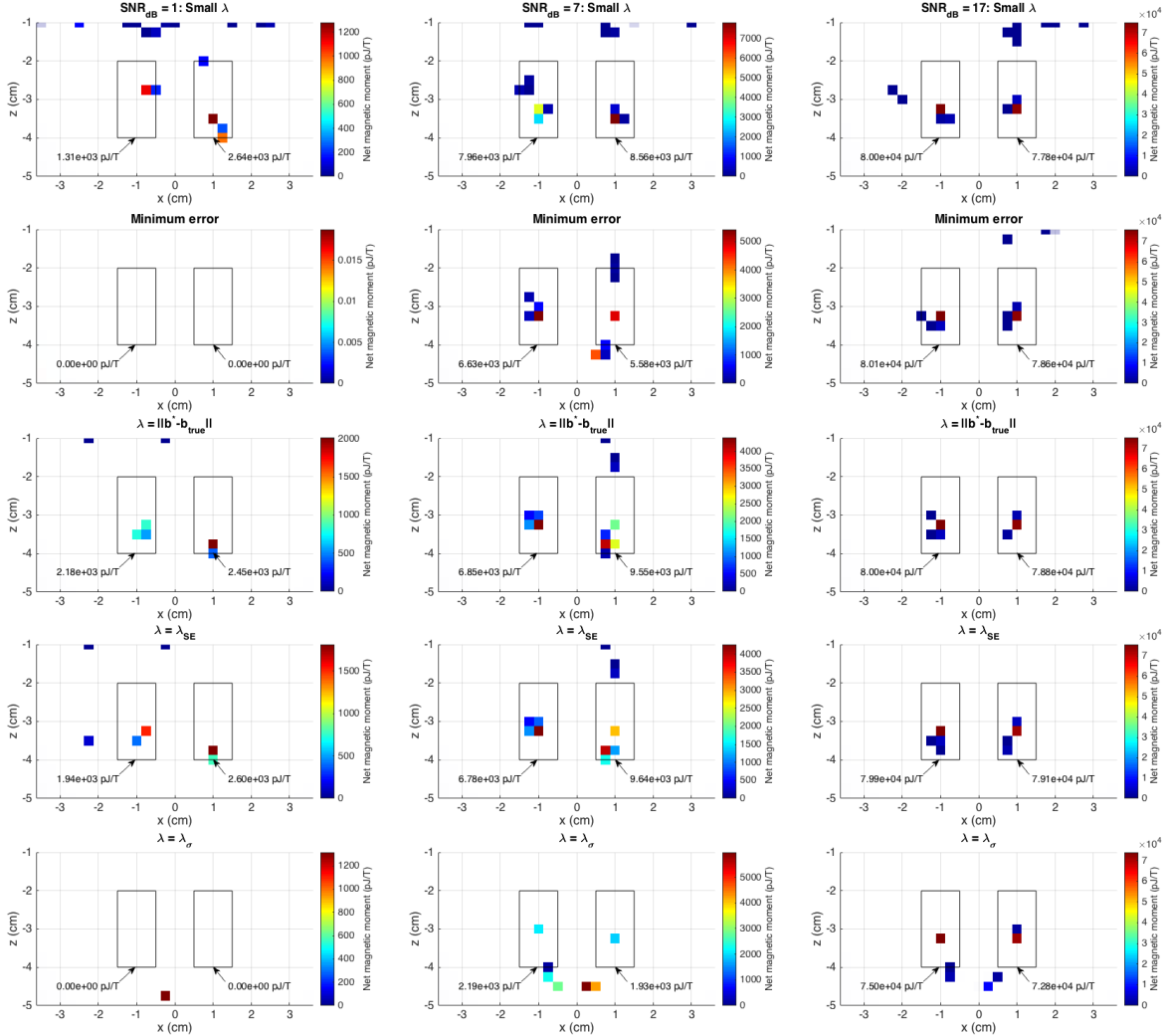
Figure 5.16: Selected reconstructions from the 30-sample data presented in Figure 5.14 for SNR levels of 1 (left column), 7 (middle column), and 17 (right column) dB. The first row shows a reconstruction with a small value of λ. The second row shows the reconstruction with the minimum distance from the true solution. The third row shows the reconstruction with $\lambda_{\Delta b}$. The fourth and fifth rows show the reconstructions with λ_{SE} and λ_{σ} , respectively. The total moment within an region of interest drawn as a box is shown to the left of the reconstructed source. For comparison, the true source strength for SNRs of 1, 7, and 17 dB, were $2 \times 10^3 \, \mathrm{pJ \, T^{-1}}$, $7.9 \times 10^3 \, \mathrm{pJ \, T^{-1}}$, and $7.9 \times 10^4 \, \mathrm{pJ \, T^{-1}}$, respectively.

sources.

5.5   Conclusions

In this chapter, we characterized the behavior of the reconstruction algorithm as a function of the experimental and reconstruction parameters. Specifically, we investigated the effects of the number of stage positions and number of pulses per stage position together with the choice of the data fidelity parameter on the accuracy of the reconstruction of simulated data of a single and two-source distribution over a range of SNR levels. We found that for reconstructions of a single source, the accuracy of the reconstruction does not greatly improve beyond five pulses collected at a single stage position for SNR levels greater than zero. However, if the number of distinct clusters of nanoparticles is unknown at the time of SPMR measurement, at least two stage positions should be used, with at least 15 pulses collected per stage position, to ensure that multiple distinct sources can be reconstructed if present.

From these results, it may seem clear that λ_{SE} is the superior choice to estimate the true error in the data, but these results must be taken with a grain of salt. First, one must consider that these results are based on simulations, in which there is no error in the forward model and the noise is exactly normally distributed. In real data, as demonstrated in Chapter 4, there is some error in the forward model, and the noise in the data is only approximately normally distributed. The simulations and the error metric also assume the same levels of uncertainty across all of the sensors, which is an approximation of the true characteristics of the sensor array. Future work should improve on the approach presented here by taking into account the variation in uncertainty characteristics among the sensors. Considering these factors, although λ_{SE} is a better estimation of the true error in simulations, it will likely underestimate the true error when applied to real data. This could result in overly noisy reconstructions in which small sources can be lost. Therefore,

λ_σ is used for the reconstruction of measured data throughout the remainder of this work, unless otherwise specified.

The biggest takeaway from our simulation studies of stage positions is that even under perfect conditions, at least two stage positions are required to reconstruct more than a single source. The fact that this can be demonstrated even without added noise indicates that it is a numerical limit imposed by the physics and math that define the system. One explanation of this limit comes from the current method of reconstruction. Each source consists of 4 unknown values that must be solved for: the position in x, y, and z, and the magnetic moment. Therefore, reconstruction algorithm solves a system of k equations, one for each sensor, for $4n$ unknowns, n is the number of sources. In order for a unique solution to be determined, the number of unknown variables must be less than the number of equations. For the system geometry used here, each stage position adds 7 equations to the system. To reconstruct two sources, there must be at least 8 equations, or sensor locations, which is why we see the improvement after two stage positions. The same limitations apply to our algorithm. Although the reconstruction is theoretically limited to a number of sources less than or equal to the number of rows in the system matrix, the limit is decreased due to the sparsity constraint. With measurements at only seven locations, the fields produced by a distribution of one source and two sources may look very similar, especially if the two sources are close together. In cases when the field from a single source looks very close to the field from two sources, our algorithm will always favor the single source reconstruction due to the nature of the L1 minimization. Indeed, we saw that decreases in the accuracy of the reconstruction were due to a concentration of the moment in a single source towards the center of the field of view. As the number of sampled locations increases, the differences between the single source field and the two-source field are more pronounced, which encourages the algorithm to more closely reconstruct the correct distribution.However, at low SNR these differences can potentially be lost in the noise. This is likely why we see high variability in the accuracy of the two-source distribution

at low SNR, especially for fewer stage positions, while the single source reconstructions are accurate even at low SNR.

In our study of the number of pulses per stage position, we found little benefit in acquiring more than 5 pulses per stage position for reconstruction only a single source at the center of the field of view. However, we simulated SNR levels only down to zero, at which point the magnitude of the uncertainty is equal to the magnitude of the signal. However, it is possible that for sources that produce magnetic field values less than the magnitude of the uncertainty in the measurements, collecting additional pulses at each stage position would improve the accuracy. Future work should be done to extend the results presented here to even lower SNR values to investigate this possibility. For reconstructions with multiple sources, the benefit of additional pulses is much more pronounced. This is because these pulses improve the definition of the shape of the detected magnetic field, which further distinguishes it from the field produced by a single source. When there is high uncertainty in the detected field, the sparsity enhancing nature of the algorithm will tend to account for the field from two sources with a single source located between and below the two true sources. With additional pulses, the mean detected field becomes closer to the true magnetic field, and can no longer be adequately approximated with a single source. In this study, we saw the accuracy of the two-source reconstructions approximately double when using 30 pulses instead of 10 at low SNR. Small sources that were not reconstructed within the ROI at any value of λ from only 10 pulses were found within the ROI when 30 pulses were used. Therefore, in cases in which multiple small sources are to be detected, it is highly advisable to collect at least 30 pulses per stage position. In our simulations based on the measurement uncertainty of our system and environment, the difference between 10 and 30 pulses per stage position represented the difference successfully detecting two sources of $2.0 \times 10^3 \, \mathrm{pJ \, T^{-1}}$, and accurately reconstructing two sources of $7.9 \times 10^3 \, \mathrm{pJ \, T^{-1}}$.

In a clinical setting, these results will influence to the sensitivity and specificity of the reconstruction. For small sources (low SNR), it is likely that a single source can be recon-

structed with reasonably high accuracy. However, distributions of multiple small sources may be reconstructed as a single larger source, likely between the true source locations, at increased depth, and stronger. The clinical impact of this type of error will depend on the application. For the early detection of ovarian cancer, it could lead to mistaking dispersed, multifocal disease for a large, locally confined tumor. This could lead to unnecessary surgery when systemic treatment would have been more appropriate. As shown here, several stage positions can improve the ability of the algorithm to distinguish distributions of multiple sources. Collecting multiple pulses at each stage position will further increase the accuracy by increasing the SNR. Clinically, it is impossible to know prior to measurement whether the true distribution of nanoparticles will be largely concentrated in a single location, or if the particles will collect in multiple focal regions. Therefore, it is important that any experimental procedure include multiple stage positions to increase the number of sources that can be found, and enough pulses to reach levels of at least 7 dB given the noise characteristics of the device and environment. Additional stage positions and pulses will likely improve the reconstruction, but come at the cost of increasing the scan time and complexity. All of these factors should be taken into account when designing an experimental or clinical protocol.

Optimizing the design of the system matrix prior to an experiment could increase the chances that the algorithm will successfully recover the true bound particle distribution. In our simulations, we saw that the ratio of the error in the solution to the error in the signal did decrease with increasing stage positions. This is a positive result which indicates that the error in the solution becomes less sensitive to errors in the signal as the number of stage positions is increased. However, the condition number of the system matrix was not an accurate characterization of this relationship. This is likely because for our algorithm, the reconstruction cannot be related to the signal by $\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}$. Most of the theory that has been developed for recovery guarantees for compressed sensing systems starts with the assumption that the system matrix satisfies a condition on the Restricted Isometry

Property (RIP) for a given number of sources [97]. It is possible to generate a matrix that satisfies a certain condition on the RIP, but has only been done for a small set of general types of matrices, such as random matrices generated by Gaussian, Bernoulli, and random Fourier ensembles [98]. However, it is computationally infeasible to certify that an already constructed matrix, such as our system matrix, satisfies the RIP [99, 100]. The nullspace property is another metric that can provide necessary and sufficient conditions to guarantee recovery of a sparse solution. Unfortunately, calculating this property of a given matrix is also computationally intractable [100]. The final property that is commonly used is the mutual coherence of a matrix. This measures the maximum degree to which any two columns in the system matrix are linearly dependent [101]. The mutual coherence of a matrix is easy to compute, but is a weak bound on the recovery guarantee because it typically underestimates the true potential of a system. For example, the literature estimates that at most s sources can be recovered from a dictionary with a coherence of m [102]:

$$ s < \frac{1}{2} \left(1 + \frac{1}{m} \right) $$

Most of the system matrices in our application have coherence values of > 0.9 on a scale from 0 to 1. This translates to guarantees of recovering at least one source, which is only slightly helpful. Recently others have demonstrated that system matrices with elements defined by a parameterized function such as ours can be optimized by minimizing the coherence of the matrix over the parameter space [103]. Future work should be done to apply this approach to optimize our sensor configuration. To make the most of this approach, we would need to allow the sensors to be moved independently rather than as an array. However, some improvement may be gained by optimizing over the stage positions alone.

Chapter 6

Phantom validation studies

6.1 Phantom validation studies

Finally, we can apply the calibrated forward model and what we have learned from the *in silico* characterization of the inverse model to validate the reconstruction method in phantom data, and compare it to the current method of reconstruction MSA. First, we extend the methods of the simulation studies in the previous chapter to single source phantom data to validate our choice of λ. Then, we compare MRXImage to MSA on SPMR measurements of one and two sources. Finally, we compare the algorithms in a pre-clinical tumor detection study in which a user is asked to determine the presence or absence of a tumor without prior knowledge of the true source distribution.

6.1.1 Single source titration

Methods

We conducted a single source titration study with two aims. The first was to determine an appropriate choice for the data fidelity parameter. The second was to compare the reconstructions with MRXImage to that of the current reconstruction method, MSA. For this study, a single point phantom consisting of between 0.1 µg and 1000 µg of nanoparticles dried on the tip of a cotton swab was placed in the center of the field of view. The

background measurement consisting of an average of 30 pulses collected with no source present was subtracted from the data before reconstructing. The sample measurements consisted of 15 pulses collected at each of three stage positions in (x, y): (0 cm, 0 cm), (-1 cm, 0 cm), and (1 cm, 0 cm). We repeated these measurements four times over the course of three days. An illustrative example of the field detected at each of the sensor locations with a source of $500\,\mu$g nanoparticles at the center of the field of view is shown in Figure 6.1.
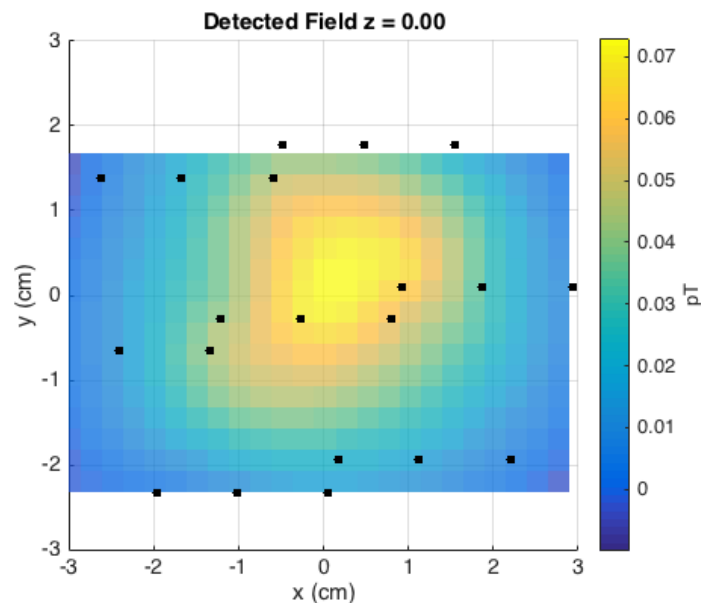
Figure 6.1: The residual magnetic field detected by the sensors (black dots) at three stage positions from a 500 μg source located at the center of the field of view.

To determine the best choice of parameter, a single replicate of the source at each location was reconstructed using 40 values of λ logarithmically spaced between 0 and $||\mathbf{b}||$, plus λ_{SE} and λ_σ, as in Chapter 5. The reconstruction was done on a field of view (FOV) spanning from -3 to 3 cm in x and y, and from -5 cm to -1 cm in z, divided into $2\times2\times2$ mm voxels. This voxel size was chosen because it was the smallest size that was both evenly divisible into the FOV and did not incur memory challenges calculating the A matrix. An ROI analysis was used to compare the voxel-based reconstruction returned by MRXImage

with the location and strength of the point source returned by MSA. All ROIs were defined from -4 cm to -2 cm in z, inclusive, to avoid the noise at the top and bottom limits of the field of view. The extent in x and y was defined to be large enough to wholly encompass any voxels near the intended source position and be of the same size without overlapping each other. The ROI for the source at (-0.9 cm, 0.9 cm) was defined from -1.5 cm to -0.5 cm in x and 1.5 cm to 0.5 cm in y. The ROI for the source at the center was defined from -0.5 cm to 0.5 cm in x and y, inclusive. The ROI for the source at (0.9 cm, -0.9 cm) was defined from 1.5 cm to 0.5 cm in x, and -1.5 cm to -0.5 cm in y. The magnitude of the source was defined as the sum of all of the voxels within an ROI, and the location was the centroid of the non-zero voxels within an ROI. We evaluated the reconstructed moment per mass of nanoparticles as well as the distance between the reconstructed source and the expected source location as a function of λ.

All replicates of the data were then reconstructed with both algorithms. For the MSA reconstruction, a single source was designated and the initial condition was set to 0, 0, -4 cm, and $1 \times 10^4 \, \mathrm{pJ} \, \mathrm{T}^{-1}$ for x, y, z, and magnetic moment, respectively. All of the dipole parameters were fit with the model. The optimized sensor location parameters found in Section 3.2 using the midpoint approximation method were used for both algorithms. The MRXImage reconstructions were performed and analyzed as described above. The parameter value λ_σ was used.

While the true magnetic moment of the sources was unknown, the mass of particles contained in each phantom was well defined, so the reconstructions were evaluated based on the linearity of the moment with mass, and the variation between multiple measurements of the same phantom. Since the mass of nanoparticles and the magnetic moment are both distributed along a logarithmic scale, we applied a weighted log-log fit (Equation 6.1) with a weighting factor of y^{-2} to the reconstructed moment (y) and the mass of nanoparticles (x) to determine the slope (m), intercept (b), and R-squared value for each algorithm.

$$y = 10^{m \log x + b} \tag{6.1}$$

We used the Aikaike's Informative Criteria (AIC) to determine whether a single relationship between the mass of nanoparticles and reconstructed magnetic moment adequately described the data at all three source locations, or if the relationship was different for each source location. We also determined whether the error in the reconstructed location was dependent on the position of the source by an ordinary two-way analysis of variables (ANOVA) on the mass of the nanoparticles and the location of the source, followed by a Tukey's multiple comparisons test. Finally, we examined how the L0-reweighting affected the total magnetic moment within the ROI.

Finally, we investigated the effect of voxel size on the accuracy of the reconstructed magnetic moment and location. We reconstructed one replicate of the single source titration with each source placed at the center of the field of view and at (0.9 cm, -0.9 cm) in (x,y) using cubic voxels with side lengths of 1.5 mm, 2 mm, 2.5 mm, 3 mm and 4 mm.

Results

To continue the study of the choice of λ from the previous section, one trial of the titration with the source located at the center of the field of view was reconstructed with 42 values of λ. Because the true magnetic moment is unknown, it is not possible to calculate Error_x or the ROI accuracy for this data set. Instead, we rely on the assumption that the moment per mass should be constant. Figure 6.2 shows the total moment reconstructed within an ROI around the true source location per mass of nanoparticles as a function of λ for each mass of nanoparticles in the titration, with λ_σ denoted by a point. Presented this way, we can define the optimal value of λ to be the one for which the reconstructed magnetic moment per mass of nanoparticles is constant for every mass of nanoparticles. In Figure 6.2, the solid lines represent the total reconstructed moment per mass of nanoparticles within the ROI corresponding to the actual location of the phantom, and the dashed lines

117

represent the total reconstructed moment per mass within the other ROIs. This allows us to evaluate whether the reconstructed noise at a given value of λ could be interpreted as a false positive source.

The curves in Figure 6.2 closely resemble those from the simulation study. Small values of λ had little effect on the total reconstructed moment within the ROI, because at this point, changing λ mostly affects the noise towards the top of the field of view. For sources of 1 µg of particles or less, we see some unstable behavior for the mid-range values of λ. However, past a certain value we see the same steady decrease in the magnitude of the reconstructed moment for all of the source strengths. When the 0.1 µg source was located at the center of the field of view, no substantial moment was reconstructed within the ROI for any value of λ, indicating that the choice of λ would not have affected the detectability of this source. When it was located at (0.9 cm, -0.9 cm), a substantial moment was reconstructed within the target ROI for only a few values of λ between λ_{SE} and λ_{σ}. Interestingly, at one value of $\lambda_{SE} < \lambda < \lambda_{\sigma}$, the reconstructed moment per mass was within the range of that of the other source strengths. However, this data is from a single measurement and therefore could be due to chance. Further work should be done to see if this result is reproducible.

From these results, we see that λ_{σ} leads to better agreement between source strengths with respect to the reconstructed moment per mass than λ_{SE}. The mean and standard deviation of the reconstructed moment per mass across all source strengths at $\lambda = \lambda_{SE}$ was $(723 \pm 356)\,\mathrm{pJ\,T^{-1}\,µg^{-1}}$, $(1120 \pm 966)\,\mathrm{pJ\,T^{-1}\,µg^{-1}}$, and $(994 \pm 560)\,\mathrm{pJ\,T^{-1}\,µg^{-1}}$ at (-0.9 cm, 0.9 cm), (0 cm, 0 cm), and (0.9 cm, -0.9 cm), respectively. The mean and standard deviation of the reconstructed moment per mass across all source strengths at $\lambda = \lambda_{\sigma}$ was $(577 \pm 176)\,\mathrm{pJ\,T^{-1}\,µg^{-1}}$, $(735 \pm 406)\,\mathrm{pJ\,T^{-1}\,µg^{-1}}$, and $(733 \pm 268)\,\mathrm{pJ\,T^{-1}\,µg^{-1}}$ at (-0.9 cm, 0.9 cm), (0 cm, 0 cm), and (0.9 cm, -0.9 cm), respectively. We also see an increased number of false positive sources at values of $\lambda < \lambda_{\sigma}$

The error in the reconstructed location as a function of λ is plotted on the left of
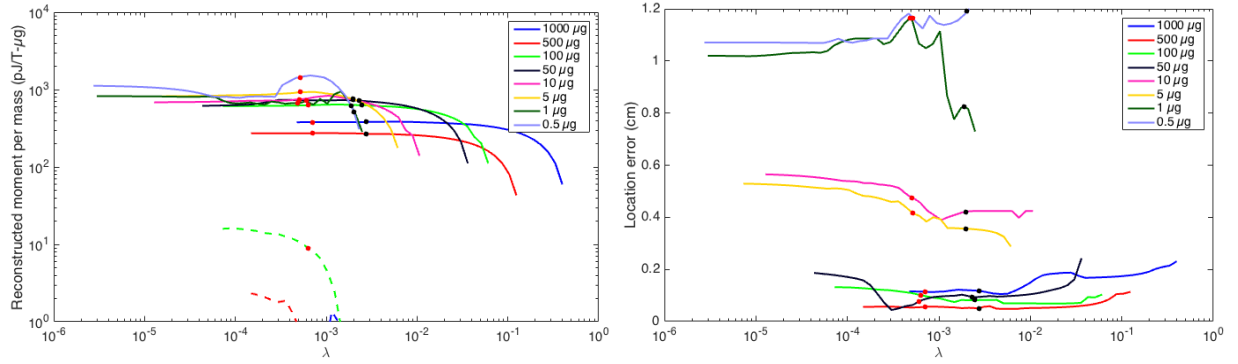
(a) Source at (-0.9 cm, 0.9 cm)
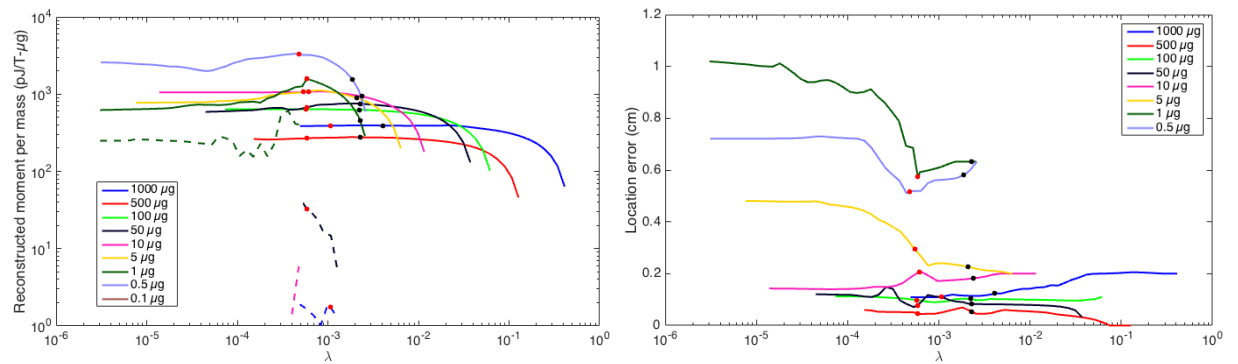
(b) Source at (0 cm, 0 cm)
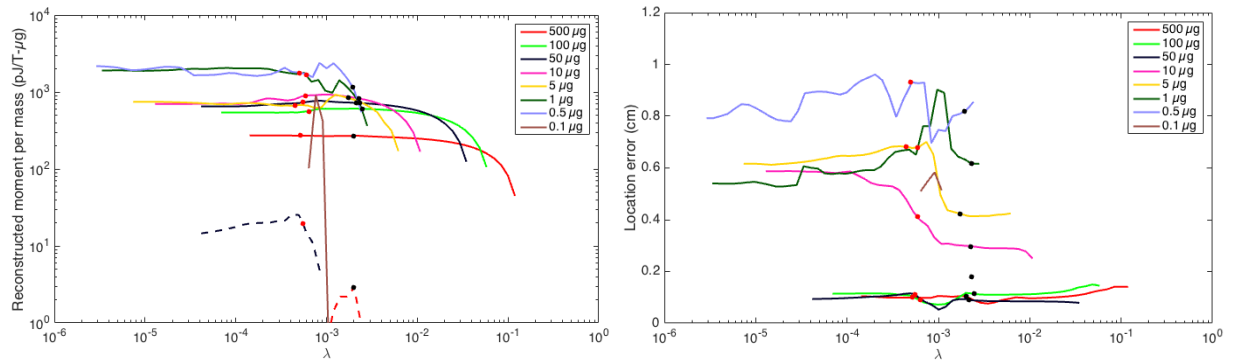
(c) Source at (0.9 cm, -0.9 cm)

Figure 6.2: The moment per mass of nanoparticles (left) and error in the location (right) reconstructed by MRXImage within an ROI around the true location of the source (solid lines) and two locations without a source (dashed lines) for a single source placed at each of three locations. The value λ_σ is denoted by a black dot and the value of λ_{SE} is denoted by a red dot. The reconstruction of the 0.1 µg source is not shown in (a) because it was not measured at this location and in (b) because there was no substantial moment reconstructed within the ROI for any value of λ. The 1000 µg source was not measured at the location in (c). For clarity, only values of reconstructed moment per mass greater than 1 are shown.

119

Figure 6.2. Again we see that the relationship between error in reconstructed location and λ found in our simulation study resembles the results from our phantom study. Especially for the stronger ($> 5\,\mu g$) sources, the choice of λ has little impact on the location of the reconstructed source. For the smaller sources, the error starts high due to the tendency of the reconstruction towards the top of the field of view, and then decreases and levels out as the reconstructed source collects in the center of the field of view. The average error in the reconstructed location was slightly lower at λ_σ (0.32 cm) than at λ_{SE} (0.37 cm), but a paired t-test determined that this difference was not statistically significant.

We chose to use λ_σ for the remainder of the phantom studies in this section, because it had a fewer false positives, showed better agreement in the reconstructed moment per mass across source strengths at each source location, and had a smaller average error in the estimated location of the source across all source strengths. However, future work should be done to see if the results of the following studies are changed when using a different value of λ.
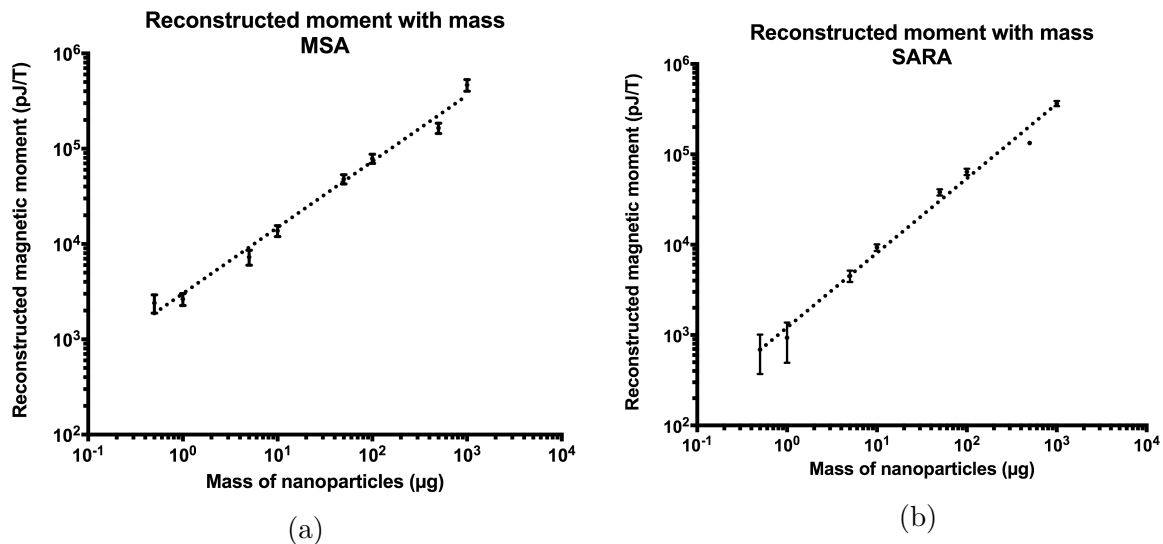
Figure 6.3: The moment reconstructed by MSA (a) and MRXImage (b) as a function of mass of nanoparticles contained in a single source placed at the center of a field of view. The error bars represent the standard deviation over four measurements.

Figure 6.3 compares the reconstructed moment as a function of the mass of nanopar-

ticles contained in the phantom returned by MSA (A) along with that reconstructed by MRXImage (B) . The error bars denote the standard deviation over all four trials. The dotted line shows the log-log fit of the data, which had a slope of 0.83 and a y-intercept of 3.1 for MRXImage, and 0.69 and 3.5 for MSA. The AIC indicated a 99.8% chance that the relationship between the mass of particles contained in the phantom and the reconstructed moment was constant regardless of source location for both algorithms.

The 0.1 µg source was considered undetectable for both algorithms. MSA reconstructed it at a depth of -1.3 cm, which is impossible because that would put the source inside the dewar. For MRXImage, the parameter λ_σ was greater than the norm of the detected field, indicating that the SNR was too low to adequately distinguish the signal from noise. The standard error of the magnetic moment reconstructed by MRXImage for the 0.5 µg, 1 µg, and 5 µg sources was 46%, 47%, and 14%, respectively. The standard error in the magnetic moment reconstructed with MSA was 17%, 11%, and 13% for the same sources. For sources with 10 µg of nanoparticles or more, the standard error in the reconstructed moment was less than 10% for both algorithms.

Although the slopes were not the same, the moment reconstructed by MRXImage and MSA were linearly related. In fact, when plotted against one another, it can be clearly seen that the two are linearly related with a slope of 1.2, intercept of -420 pJ/T, and R^2 of 0.9997, as shown in Figure 6.4. This consistent linear relationship indicates that the difference between the magnitude of the magnetic moment reconstructed by the algorithms is likely due to differences in the pre-processing that converts the raw decay curves into initial field values which are then reconstructed. The slope of 1.2 is likely due to the conversion from volts to pT that is applied in MSA but not accounted for in MRXImage, or the difference in how the initial field value is derived from the decay curve. The offset of 420 pJ/T is likely the result of differences in how the preprocessing algorithms calculate the DC offset of the decay curves.

The error in the reconstructed location at each position is shown in Figure 6.5 a) for
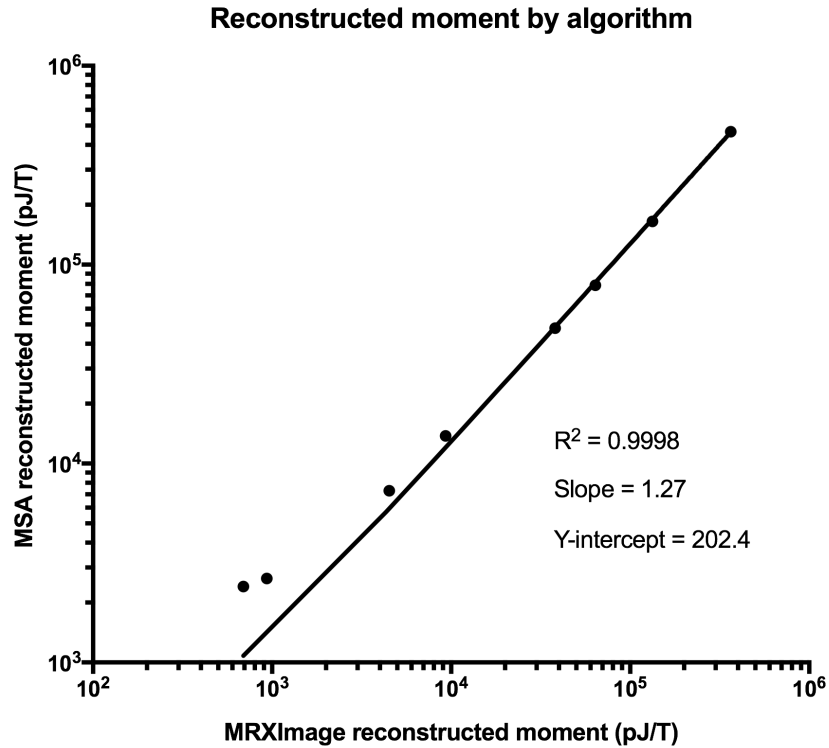
Figure 6.4: Comparison of the average reconstructed moment of four measurements of each source strength reconstructed by MRXImage and MSA, showing a linear relationship between the two of 1.27*MRXImage+202 = MSA
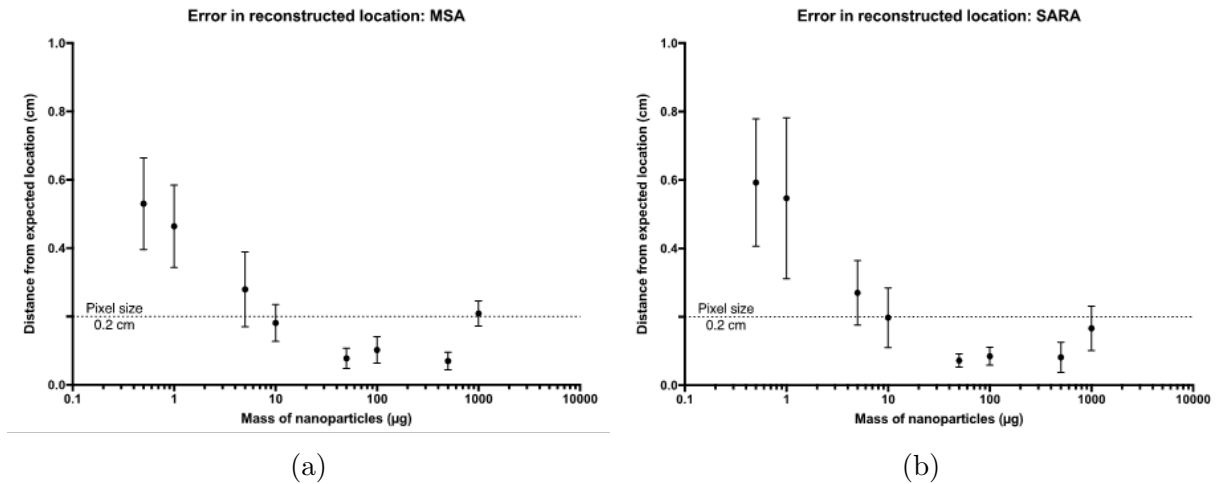
Figure 6.5: The mean distance between the true location and the location reconstructed by MSA (a) and MRXImage (b) for a single source placed at each of three positions within the field of view.

MSA and b) for MRXImage. Overall, the average error in the reconstructed location was similar between the two algorithms as a function of mass of nanoparticles. Both MRXImage and MSA reconstructed sources containing between 10 μg and 500 μg of nanoparticles within 2 mm, approximately 1 voxel for MRXImage, of the true source location. Both algorithms had increasing error with decreasing mass of nanoparticles below 10 μg. The error also increased for the reconstruction of the phantom containing 1000 μg using both algorithms. A two-way ANOVA showed no signficant difference in the reconstructed error as a function of the source location for MSA (p=0.41) and MRXImage (p=0.93). Over all of the reconstructions, the maximum error was 6.1 mm (about 3 voxels) for MRXImage and 7 mm for MSA. For both algorithms, the largest error in the location was in the reconstructed depth of the source. In fact, both algorithms reconstructed all of the phantoms within 2 mm of their true location in the x-y plane.

Figure 6.6 shows the effect of the voxel size on the magnetic moment and location of the source reconstructed by MRXImage. Across all voxel sizes, the reconstructed location of the source varied less than 1 mm. The reconstructed moment varied less than 2% for sources with at least 5 μg, and less than 4% for all sources.

6.1.2 Multiple source detection

In the previous section, we quantified the expected uncertainty in moment and location as a function of source strength and position when a single source is present with both algorithms. In this section, we will test the hypothesis that our algorithm will more accurately reconstruct distributions of multiple sources than the current algorithm, MSA. To do so we measured pairs of phantoms with ratios of 2, 5, 10, 50 and 100. The stronger of the pair was located at (-0.9 cm, 0.9 cm), and the weaker at (0.9 cm, -0.9 cm), with the same SPMR procedure used in the single source test. We reconstructed each distribution in the same way as with the single source distributions. To evaluate the reconstruction, we compared the error in the reconstructed moment of each source to that of when the same
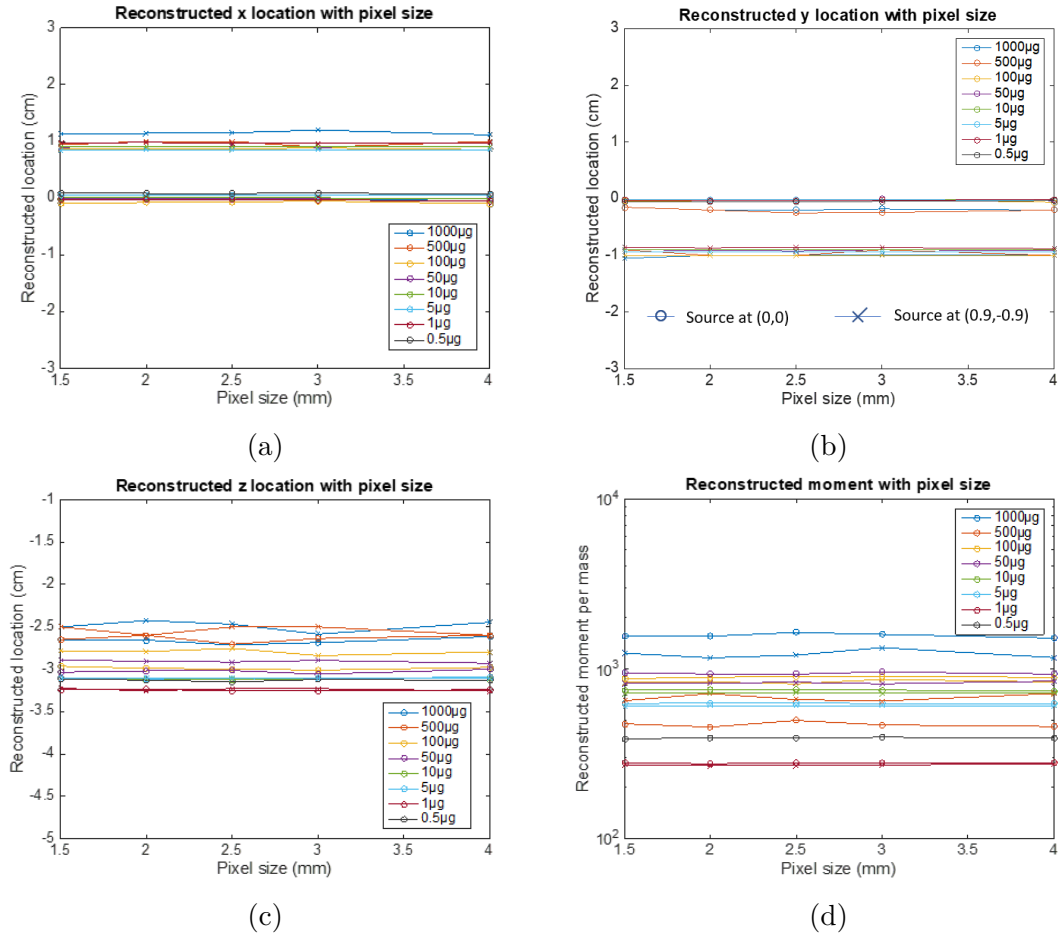
Figure 6.6: The reconstructed location of the source in x (a), y (b), and z (c) and the magnetic moment (d) for a source containing between 0.5 µg and 1000 µg of nanoparticles, reconstructed with MRXImage using cubic voxels with side lengths between 1.5 mm and 4 mm. The reconstructed location of the source when placed at the center of the field of view is denoted by 'o' and the '×' denote the source was placed at (0.9 cm, -0.9 cm).

source was measured in the same location in the single source trial. We normalized the errors by the expected standard deviation from the single source trial, so that each error is with respect to the expected error for the same source measured alone. We call this quantity Error$_\sigma$. Equation 6.2 describes the Error$_\sigma$ of a magnetic moment m_i measured in the two-source configuration for a source strength i with a mean magnetic moment $\bar{m}_i$ and standard deviation σ_i when measured alone.

$$\text{Error}_\sigma = \frac{m_i - \bar{m}_i}{\sigma_i} \tag{6.2}$$

Multiple source results

The error in the reconstructed moment as a function of source strength is shown in Figure 6.7 for MSA (left column) and MRXImage (right column) reconstructions of the weaker (top row) and stronger (bottom row) of each pair of sources. The black '×'s denote when the source was not found within the designated ROI. As shown in the left column of Figure 6.7, MSA failed to reconstruct 13 out of the 46 sources that contained at least 0.5 µg of nanoparticles, and failed to reconstruct all but two sources containing at least 100 µg of nanoparticles within two standard deviations of the expected moment from the single source trials. For the 36 sources between 0.5 µg and 100 µg of nanoparticles, MSA reconstructed only 11 within two standard deviations of their single source measurements. In every case in which the 5 µg source was the stronger of the two, and in half of the cases in which the 10 µg source was the stronger of the two, neither source was found. While MSA did not detect the 0.1 µg source in any of the cases when it was measured alone, it did detect it in one of the two-source distributions.

As shown on the right column of Figure 6.7, MRXImage missed only two sources with at least 0.5 µg of nanoparticles. MRXImage reconstructed all but two sources with at least 50 µg of nanoparticles within one standard deviation of the moment when measured alone. All of the sources with 5 µg or 10 µg of nanoparticles were detected in the two-
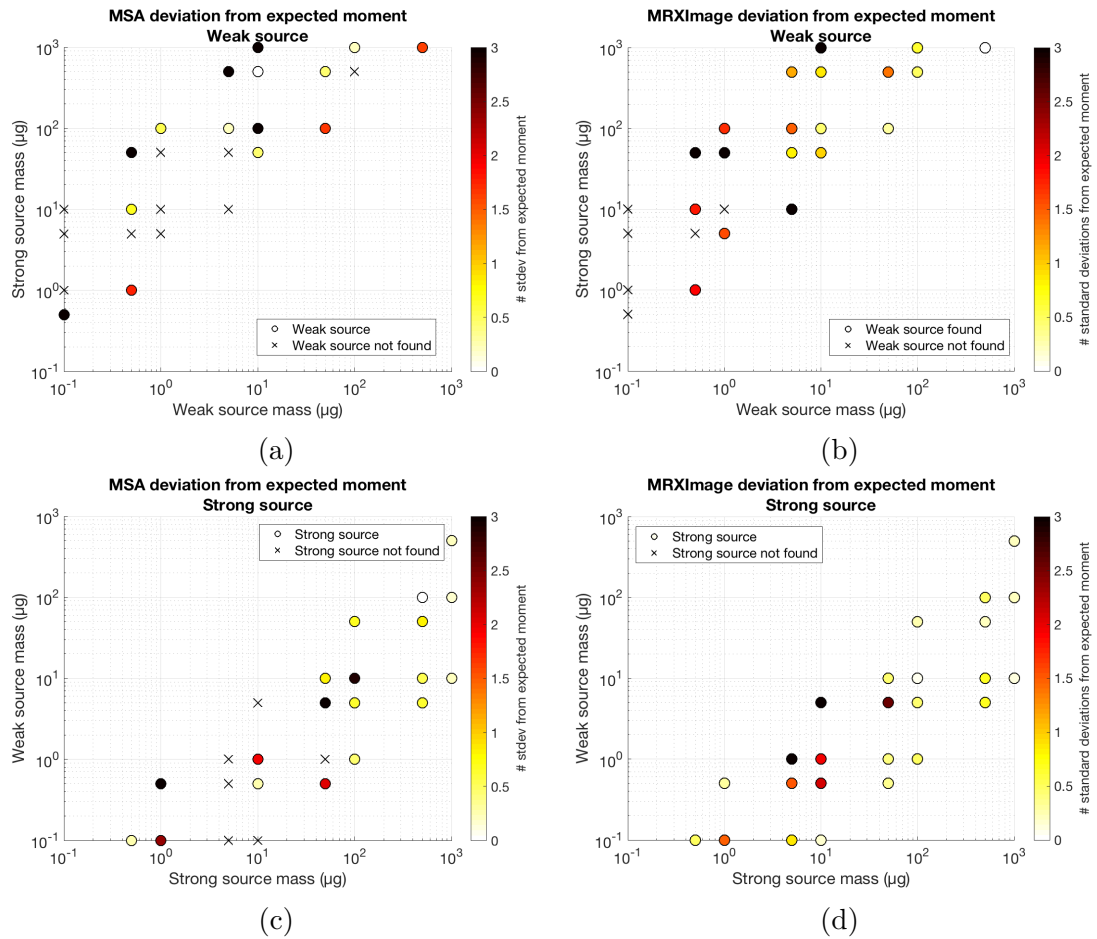
Figure 6.7: Error in the reconstructed magnetic moment for distributions of two sources, using MSA and MRXImage. The sources that were not found within the limits of the region of interest (ROI) are denoted by an 'x'.

source configuration, but the errors were higher than what was seen when the sources were measured alone. The 0.5 µg and 1 µg sources were each missed in one case, and when detected, the error in the reconstructed magnetic moment was higher than expected. For the cases in which the 0.1 µg-source was present, which was undetectable in the single source test, the larger source was found within the expected error, and the 0.1 µg source was not detected. For both algorithms, the errors tended to be higher for sources were similar in strength.
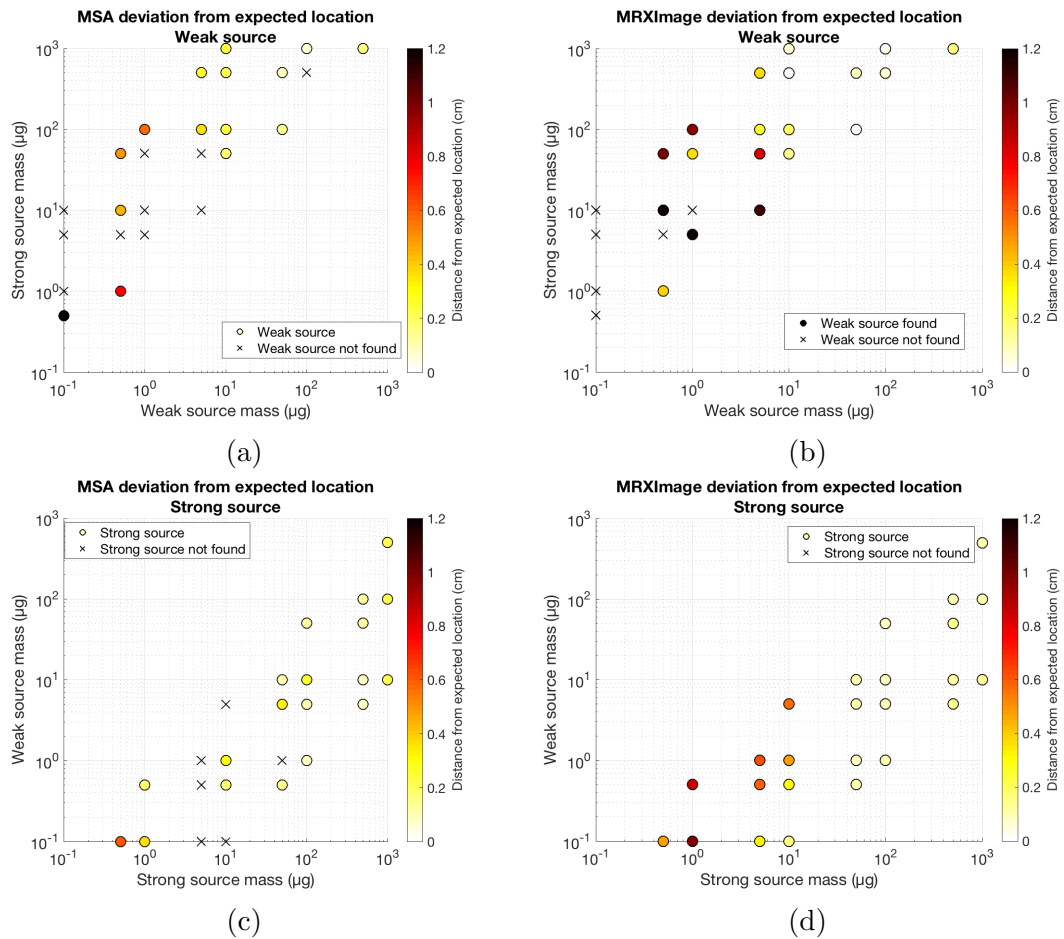
Figure 6.8: Error in the reconstructed location for distributions of two sources, using MSA and MRXImage. The sources that were not found within the limits of the region of interest (ROI) are denoted by an 'x'.

The error in the reconstructed location for each pair of sources is shown in Figure 6.8 for MSA (left column) and MRXImage (right column). Of the sources that MSA

reconstructed within the ROI, all but three sources with at least 10 μg of nanoparticles within 2 mm of their true location. Of the 18 sources between 0.1 μg and 10 μg, ten were reconstructed within 8 mm of their true location. MRXImage reconstructed all but two sources with at least 10 μg of nanoparticles within 2 mm of their expected location. Sources with smaller amounts of nanoparticles were reconstructed with a maximum error of 1.3 cm from their expected location, or not found at all.
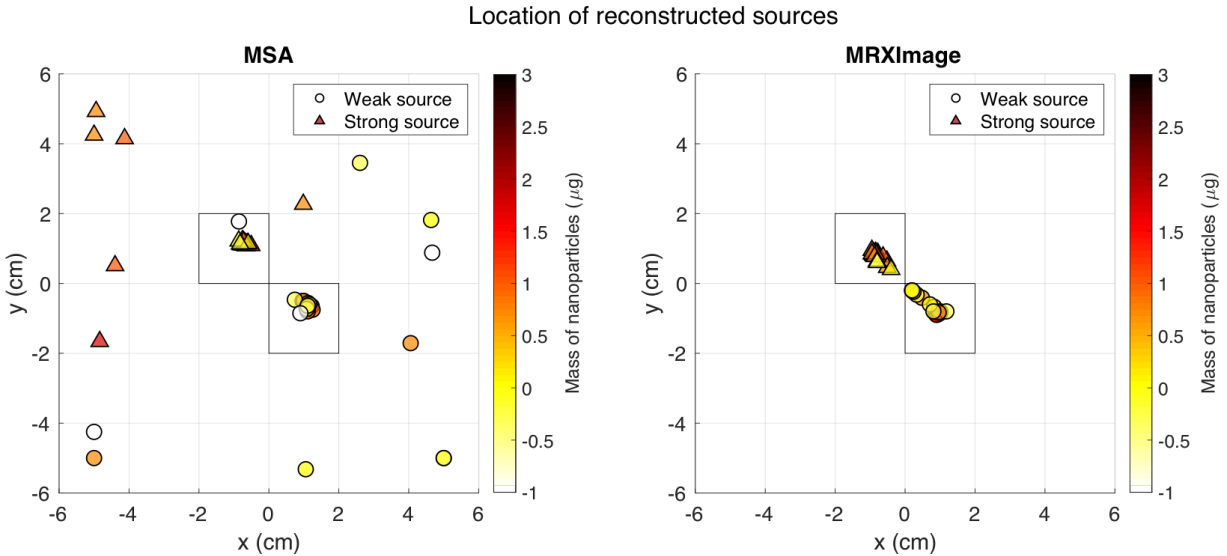
Figure 6.9: The reconstructed locations of the strong (triangle) and weak (square) sources

Errors in the MSA reconstructed location of the sources tended to be scattered across the field of view, resulting in many sources being reconstructed outside of the ROI, as shown in Figure 6.9. Errors in the reconstructed location with MRXImage tended to cause the sources to move towards the center and bottom of the field of view. For both algorithms, large errors in the reconstructed location did not necessary correspond to large errors in the reconstructed moment.

6.2 Blinded tumor detection study

In order for SPMR technology to safely move into the clinic, it must first be shown to work in small animal models. However, SPMR of mice introduces some challenges due to the much smaller anatomy than humans. The biggest challenge is the close proximity of the liver to the tumor. It is well known that a large fraction of intravenously injected SPIO particles are non-specifically taken up by macrophages and ultimately end up in the liver [citation]. Therefore, it is important that the reconstruction algorithm is able to detect a small tumor signal in close proximity to a large signal from the liver. We tested three SPMR reconstruction algorithms in a phantom study designed to simulate the scenario of detecting a small amount of tumor-bound particles in close proximity to a large amount of non-specific binding in the liver. The goal of this study was to compare the ability of three reconstruction algorithms to accurately determine whether a second source was present without prior knowledge of the true number of sources.

6.2.1 Methods

6.2.2 Experimental procedure

Clusters of immobilized particles in the tumor and liver were simulated by drying various amounts of SPIO particles on the tip of a cotton swab in zero field. The phantoms were positioned within the field of view according to the approximate location of the liver and potential tumor sites on the left and right flank of the mouse, as shown in Figure 6.10. Out of a total of 60 test cases, 6 contained no phantoms. In 20 cases, a single phantom was placed at the at the location of the liver (Figure 6.10b) to simulate a mouse with no tumor. In 34 cases two sources were present. Of the two, the phantom containing more particles was always located at the liver position, and the second phantom was located at the position of the left flank in 17 cases and the right flank in 17 cases (Figure 6.10c and d). The number of cases with a given mass of particles used in the single source cases are

listed in Table 6.1 and in Table 6.2 for the two source cases.
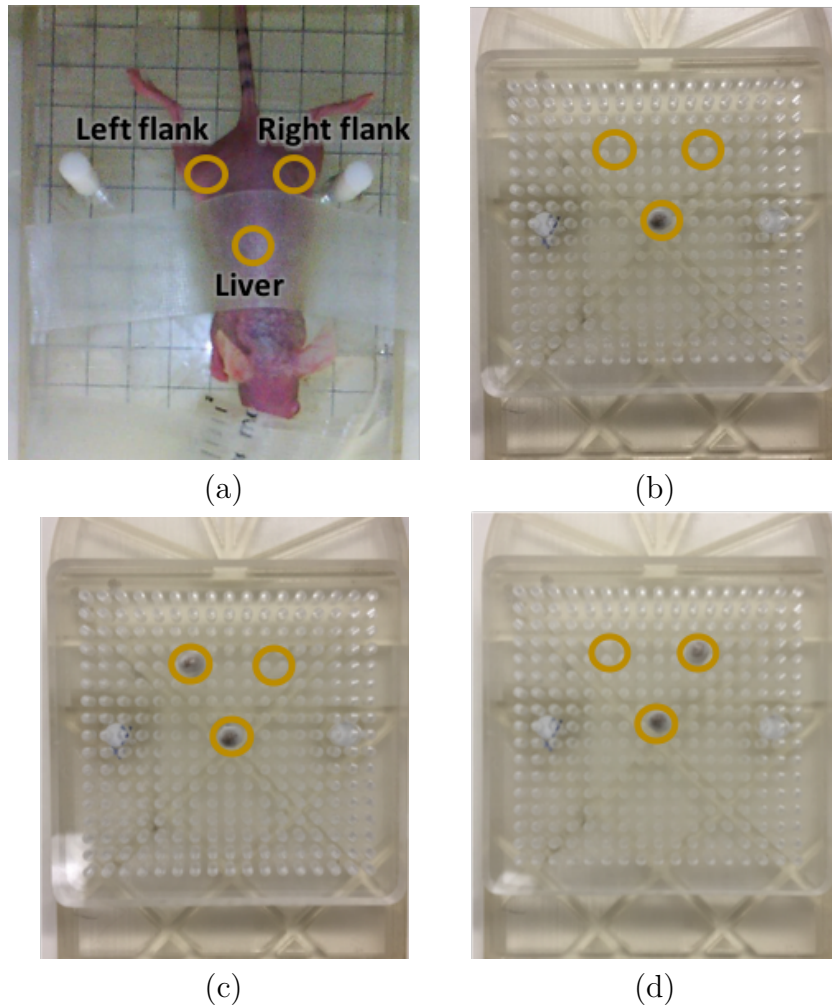
Figure 6.10: The arrangement of the phantoms used in the tumor detection study is based on a small animal model in which after intravenous injection, most of the particles collect in the liver and a only small number bind in the tumor located on the left or right flank. The user was asked to determine whether each case contained zero (not shown), one (b), or two sources with the second source to the left (c) or right (d).

The SPMR data for the test set was collected over five days. Background measurements consisted of 50 SPMR pulses collected with no source present. Background measurements were collected at the start of each day, and repeated after every five test cases. For analysis, each test case was assigned the most recent background measurement prior to when the data was collected. The 60 test cases each containing zero, one or two sources were collected in a random order. For each test case, the stage was manually positioned at

Source mass (μg)	Number of test cases
0.39	3
0.78	3
1.56	3
3.13	3
6.25	2
12.5	2
25	2
50	2

Table 6.1: The number of single source test cases in the detection study with a given mass of nanoparticles.

	Large source mass		
	50 μg	25 μg	12.5 μg
6.25 μg	2	0	0
3.13 μg	4	2	0
1.56 μg	4	4	4
0.78 μg	2	4	4
0.39 μg	0	0	4

Small source mass (row label)

Table 6.2: The number of two source cases in the detection study with each combination of large source mass (columns) and small source mass (rows) of nanoparticles. For each pair of large and small sources, half were located on the left and half on the right. The test set contained a total of 34 cases containing two sources.

six locations in (x cm, y cm): (0, 0), (0, 1), (-1, 1), (-1, 0), (-1, -1), (0, -1). The location of the sensors with respect to the field of view for these stage positions according to the calibration conducted shortly before the data was collected is shown in Figure 6.11. A total of 10 SPMR pulses were collected at each of the six stage positions.
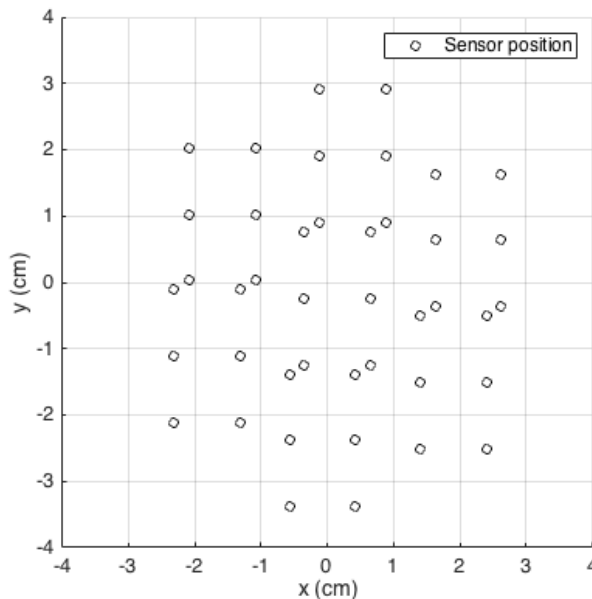
Figure 6.11: The location of the sensors (black circles) relative to the field of view for the six stage positions used in the detection study.

The data was reconstructed by users who had no part in the data collection and initially no knowledge of the true phantom configuration in the test cases. One user was assigned MSA and one MRXImage. Each algorithm consists of a pre-processing method and a reconstruction method. The users were instructed to classify each test case as having no sources, one source, two sources with the second to the left, or two sources with the second to the right. Before being provided the test set, each user was given a training set that consisted of 14 cases each with one or two sources along with the true configuration for each case. This was meant to allow the user to become familiar with the software and formulate appropriate criteria for determining the number of sources in each case, and therefore the results of the training set were not recorded.

The initial blinded analysis with MRXImage and MSA are referred to as MRXImage-B6 and MSA-B, respectively. Each analysis included all six stage positions. Per usual protocol in MSA, the user manually selected bad pulses to be excluded based on visual inspection. The default initial conditions for the least-squares optimization corresponding to $[x_0, y_0, z_0, m_0] = [0 \text{ cm, } 0 \text{ cm, -4 cm, } 1 \times 10^4 \text{ pJ T}^{-1}]$ were used for both sources. The sensor tilt angle and arc length for MSA were set to according to measured values of 6.5° and 2.3 cm (corresponding to h = 20.3 cm), respectively. A dewar rotation angle of 148° was used according to measurements performed at the time of installation.

The MRXImage reconstructions were performed on a field of view that spanned from -3.9 cm to 3.9 cm in x, -4.2 cm to 3.6 cm in y, and -5 cm to -1 cm in z, divided into $3 \text{ mm} \times 3 \text{ mm} \times 2.5 \text{ mm}$ voxels. The data fidelity parameter was set to λ_σ (Equation 5.6). MRXImage calculated the forward model using a dewar rotation angle of 144.8, a sensor tilt angle of 7.86°, an arc length of 2.2 cm (corresponding to h = 16.2 cm), and a displacement of 0.34 cm in x and 0.21 cm in y found from the optimization procedure described in Chapter **??** performed one month before the data for this study was collected. The largest difference between the two sets of model parameters is in the tilt of the sensors (and therefore h). This will have the strongest effect on the reconstructed depth of the source, which shouldn't change the results of the classification study since only the presence or absence of a source is being tested, not its location. Based on the technical drawings of the sensors and results of subsequent forward problem calibrations, the sensor tilt angle of 6.5° used in MSA was likely more accurate than the 7.86° used in MRXImage, and the dewar rotation angle of 144.8° used in MRXImage was more accurate than the 148° used in MSA. To ensure that the slightly shifted dewar angle did not adversely affect the results of MSA, we repeated the study using a dewar angle of 144.8°. The results of this analysis are included in Appendix 10.3, but saw no improvement in the results indicating that the original parameters could to be considered correct for the purposes of this study.

Two additional analyses were performed with MRXImage. First, the data was re-

analyzed by the same blinded user with only the first four of the six stage positions. This analysis is referred to as MRXImage-B4. Then, the data was re-analyzed by a different blinded user with sensor location parameters optimized for the discretized forward model but calculated the matrix **A** using the midpoint approximation. Specifically, values of 15.73 cm for h (corresponding to an arc length of 2.15 cm), a dewar rotation angle of 143.9°, and 0.24 cm for the displacement in y, and the first four stage positions were used. All other reconstruction parameters remained the same, and this analysis is referred to as MRXImage-E4. The four analyses presented here are summarized in Table 6.3.

Name	Algorithm	Stage Positions	Parameters	User
MSA-B	MSA	6	Correct	A
MRXImage-B6	MRXImage	6	Correct	B
MRXImage-B4	MRXImage	4	Correct	B
MRXImage-E4	MRXImage	4	Incorrect	C

Table 6.3: Summary of the analyses conducted for the blinded study

Analysis of results

The overall results of each analysis are presented in a 4×4 confusion matrix in which each column represents the true classification of each case, and the column represents the classification based on the reconstruction. The benefit of a 4×4 confusion matrix is that it presents all of the results in a way that is organized and easy to understand and allows us to calculate the overall accuracy (cases correctly classified out of the total number of cases) of the algorithm. However, it is not conducive to calculating important metrics which apply only to binary classification tasks. For this reason, we must compute the precision and recall of each class separately. We do so by calculating the 2×2 confusion matrices for each class using a one-versus-all approach [104]. This approach, for a given class A, compares cases classified as "A" with those classified as any other class, or "Not A". One effect of this simplification is that there will be a large number of true negatives ("Not A" cases correctly classified as "Not A") in each class, which may skew some metrics. Therefore, we chose to

evaluate the precision and recall metrics, because they are not affected by the number of true negatives. The precision of a test (Equation 6.3) with respect to class "A" is defined as the number of cases that were correctly classified as "A" (true positives) out of all of the cases that were classified as "A" (true positives plus false positives). Precision is also known as positive predictive value, or the probability that a positive test result is correct. Recall (Equation 6.4 is defined as the number of cases that were correctly classified as "A" (true positives) out of all of the cases that truly belong to class "A" (true positives plus false negatives). Recall is also called sensitivity, or the probability that a test will detect a truly positive case. A diagnostic test needs to have both a high precision and recall. In the case of cancer diagnosis, a test with high recall will reliably detect disease in patients that truly have disease, whereas a test with high precision will reliably identify patients who are truly healthy. The consequences of a test with low recall would be that patients with cancer could go undetected, while a test with low precision could result in unnecessary treatment of healthy patients.

$$precision = \frac{TP}{TP + FP} \tag{6.3}$$

$$recall = \frac{TP}{TP + FN} \tag{6.4}$$

We also evaluate the performance of each algorithm as a classifier using Cohen's Kappa statistic (Equation 6.5) and Cohen's Weighted Kappa statistic (Equation 6.7) [105, 106]. In our case, this statistic measures how much better a classifier performs than would be expected from a trivial classifier. For instance, given a data set in which only one case out of 100 was classified as "true", a classifier that simply labeled everything "false" would have 99% accuracy. However, the kappa statistic for this classifier would be zero. Given that our test set has multiple classes with an unequal number of cases in each class, this statistic provides a valuable single measure of the accuracy of a classifier. See Appendix

8.1.1 for further discussion of the kappa statistic.

$$\kappa = 1 - \frac{1 - p_o}{1 - p_e} \tag{6.5}$$

$$p_e = \frac{1}{N^2} \sum_k n_{ka} n_{kT} \tag{6.6}$$

In Equation 6.5, p_o is the observed accuracy of the classifier, and p_e is the expected accuracy, which is determined by Equation 6.6, where n_{ka} is the number of times rater a predicted category k, n_{kT} is the true number of items in category k, for a total of N items.

$$\kappa_w = 1 - \frac{\sum_{i=1}^{k} \sum_{j=1}^{k} w_{ij} o_{ij}}{\sum_{i=1}^{k} \sum_{j=1}^{k} w_{ij} e_{ij}} \tag{6.7}$$

Another benefit of the kappa statistic is that it can also account for the severity of classification errors. Take for example a case in which there are two sources, with the second to the left. A classification of zero sources and a classification of two sources with the second to the right are both wrong. However, a classification of zero sources would have more severe clinical consequences than finding both with one in the wrong location. The weighted kappa statistic would allow the zero source classification to be penalized more strongly than the location error. The penalties are defined by a weighting matrix, where the elements are rated with the severity of the error. The weighted kappa k_w is defined in Equation 6.7 for elements o_{ij} and e_{ij} of the observed and expected confusion matrices, respectively. The weighting matrix we used is shown in Table 6.4. When the off diagonal elements are 1 and the diagonal elements are 0, the weighted kappa is equal to the kappa statistic.

There are some limitations to the kappa statistic. It should not be used to compare performance on different data sets, or project to other measures of performance. Here, we are simply using it to compare the classification of this data set between different algorithms. The kappa statistic reported here should not be used to predict the specific

	0	*1*	*L*	*R*
0	0	1	2	2
1	1	0	1	1
L	2	1	0	0.5
R	2	1	0.5	0

Table 6.4: The weighting matrix used in the calculation of the weighted kappa statistic

performance of any algorithm on future trials.

6.2.3 Results

MSA

The results for MSA-B are detailed in Table 6.5. MSA correctly classified 26 out of the 60 cases, resulting in an overall accuracy of 43.3%. The kappa score for MSA was 0.19 and the weighted kappa score was 0.27, indicating that the reconstruction classified the cases slightly better than random chance. There were no cases in which a source was found when there was in fact no source present, and only 2 cases no source was found when at least one source was present. This resulted in a recall of 1 and a precision of 0.75 for the zero source classification. In the cases in which a single source was present, MSA correctly identified all of the sources that contained at least 1.6 µg of nanoparticles, and 2/3 of the cases with sources containing 0.4 µg and 0.8 µg of nanoparticles. MSA did not misclassify any cases containing a single source as containing two sources. The recall for the one-source classification was 0.9 and the precision was 0.36. The largest contribution to the low overall accuracy was because most (32 out of 34) of the cases containing two sources were classified as having only one source. This resulted in a recall and precision of 0.06 and 1, respectively, for both the Right and Left classifications.

MSA-B		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**6**	2	0	0
	1	0	**18**	16	16
	L	0	0	**1**	0
	R	0	0	0	**1**

Table 6.5: The overall classification results for MSA-B

MRXImage

The blinded analysis with MRXImage using all six stage positions and the correct sensor location parameters, MRXImage-B6, performed the best of the blinded analyses. The detailed results are presented in Table 6.6. MRXImage-B6 correctly classified 48 out of the 60 cases, for an overall accuracy of 80%, a kappa of 0.72 and a weighted kappa of 0.74. Most notably, it had 100% accuracy in determining the presence or absence of a source, in that it correctly classified all six cases that contained no source, and in no case did it classify a case containing one or more sources as having no source. This resulted in perfect (100%) precision and recall for the zero source cases. Of the single source cases, MRXImage-B6 misclassified two as having two sources, with one to the left and one to the right, resulting in a recall score of 90%. The analysis misclassified nine cases with two sources as having only one source, leading to a precision score of 67%. MRXImage-B6 correctly identified 14 out of 17 cases with two sources and the second source to the left. The three cases that were misclassified were classified as having only one source contained 1.56 µg of nanoparticles or less. The precision and recall for the cases with a second source to the left were 88% and 82%, respectively. The performance on two source cases in which the second source was to the right was slightly worse than when the second source was on the left. Six of these 17 cases were classified as containing only one source. In all but one case in which the second source was missed, the second source contained 0.78 µg of particles or less. In one case the algorithm missed the second source containing 3.13 µg. In

one case, the second source was found on the wrong side. The precision and recall for the two-source cases with the second to the right were 91% and 59%, respectively.

MRXImage-B6		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**6**	0	0	0
	1	0	**18**	3	6
	L	0	1	**14**	1
	R	0	1	0	**10**

Table 6.6: The overall classification results for MRXImage-B6

The effect of 4 versus 6 stage positions

The data was also analyzed with MRXImage using only the first 4 stage positions. The 4x4 confusion matrix for MRXImage-B4 is presented in Table 6.7. The overall accuracy of the analysis with four stage positions was 73.3%. The kappa score and weighted kappa score were 0.63 and 0.66, respectively. In comparison to the analysis with all six stage positions, there was an increase in the number of single source cases misclassified as containing two sources, and a few additional cases with two sources misclassified as having only one source. This indicates that the additional stage positions help determine the number of sources, which could reduce the number of false positives in the case of only one source, and false negatives in the case that there are two sources. However, the use of fewer stage positions did not reduce the accuracy of the algorithm for distinguishing the presence of at least one source from the absence of any source.

The effect of inaccurate forward model parameters

Finally, we can evaluate the effect of having inaccuracies in the forward model on the end classification result. This analysis is possible due to the fact that after the first blinded analysis with MRXImage errors were discovered in the forward problem parameters.

MRXImage-B4		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**6**	0	0	0
	1	0	**15**	2	8
	L	0	5	**15**	1
	R	0	0	0	**8**

Table 6.7: The overall classification results for MRXImage-B4

This might explain why after an initial analysis, it was thought that two of the stage positions were corrupted, and that excluding them would improve results. However, as shown in the previous section, excluding stage positions reduced the performance of all of the algorithms. It is likely that the residual field patterns that prompted the exclusion of the stage positions were actually due to errors in the forward problem rather than errors in the stage positioning. Specifically, the initial analysis used the parameters from the fit of the discretized forward model, but then did not calculate the system matrix using the discretization method. Specifically, the parameter h was set to 15.73 cm instead of 16.1 cm, and the dewar angle was 2.51 rad rather than 2.58 rad. There was also a typo in the y displacement, which was 0.2410 cm rather than 0.2041 cm. The error in the displacement would have shifted the result, but shouldn't have affected the classification. The errors in the dewar angle and h shifted the modeled sensor locations as shown in Figure 6.12.

The classification performance measures for the analysis with the incorrect forward model parameters (grey) and the corrected parameters (black) are shown in Figure 6.13. Errors in the sensor locations of just a few millimeters decreased the accuracy of the classification by more than 10%. Even with only four stage positions and errors in the forward model, the overall accuracy was still higher than that for MSA with all six stage positions.

A summary of the performance metrics for each analysis is presented in Figure 6.14. For simplicity, the binary-based precision and recall scores presented here are an average score over the four classifications for each algorithm. Overall, MRXImage had the highest
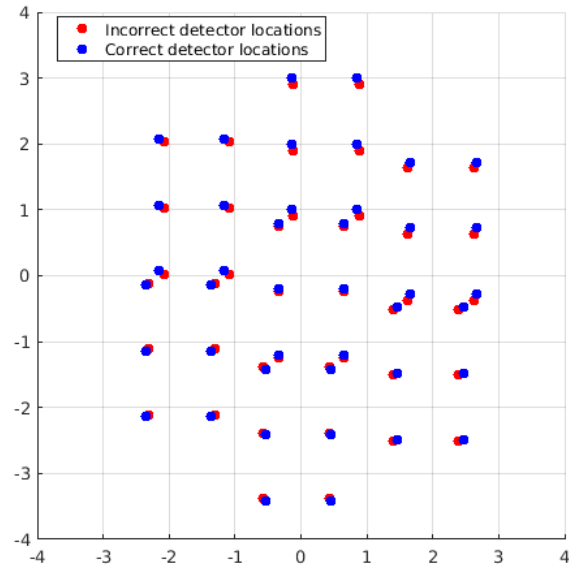
Figure 6.12: Errors in the forward problem parameters caused a shift in the sensor locations for one analysis of the blinded study with MRXImage. The corrected forward model parameters showed improved classification results.
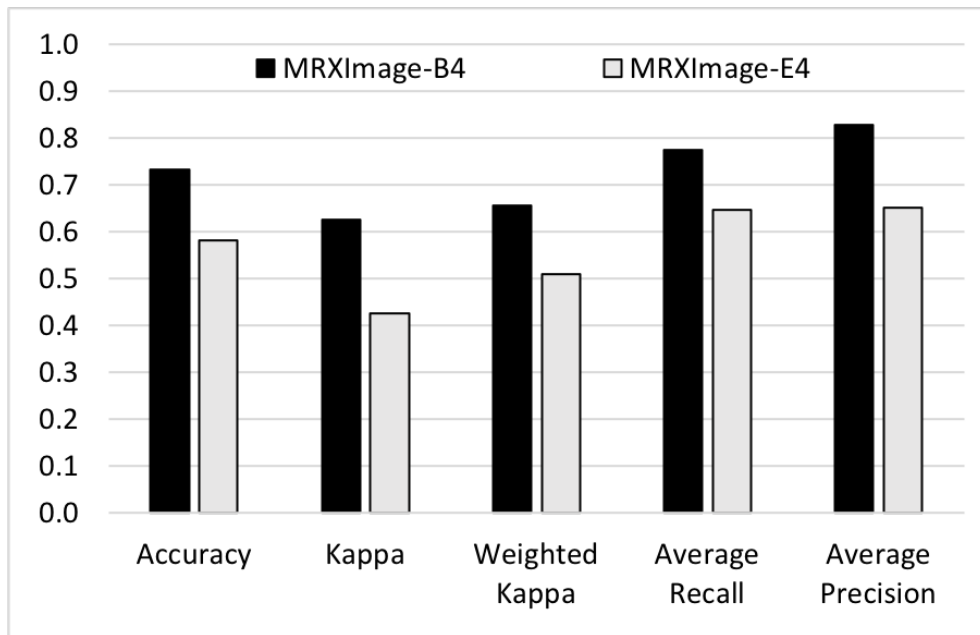
Figure 6.13: The classification performance of MRXImage with incorrect (grey) and correct (black) forward model parameters.
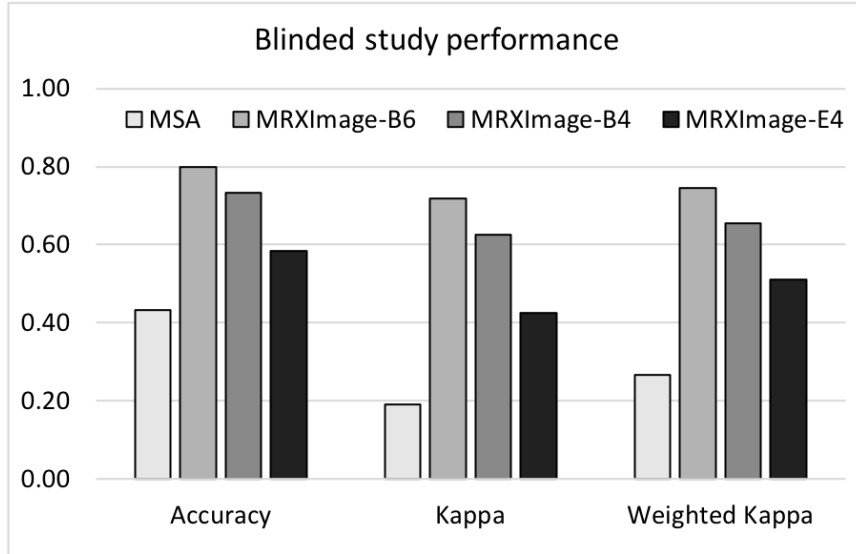
Figure 6.14: The kappa statistic (Eq. 6.5), weighted kappa (Eq. 6.7), and overall accuracy of the blinded classification study analyzed with MSA (MSA-B), and three analyses with MRXImage (MRXImage-B6, MRXImage-B4, and MRXImage-E4)

score for each of the performance metrics. The weighted kappa was slightly higher than the kappa statistic for all three algorithms. This indicates that there was a higher occurrence of less severe errors, such as misclassifying the location of a source than of more severe errors, such as missing the presence of both sources. The high average precision score for MSA-B highlights why binary classification metrics such as precision and recall may not be appropriate measures for multi-class data. MSA had perfect precision for the two-source cases, because by classifying all but one case on each side as having only one source (negative) it avoided having any false positives. In turn, this resulted in a low precision for the single source class. However, since the class-wise averaging considers the performance of each class equally, the high precision for the two-source classes twice outweighed the low score for the single source class, resulting in a artificially high class-average precision score. MSA also had a higher overall accuracy relative to its kappa scores than the accuracy of MRXImage relative to its kappa scores. This further demonstrates the ability of the kappa statistics to adjust for imbalances in the classifiers resulting in a more intuitive measure of performance.

6.3 Conclusions

In this section, we presented the results of a virtual clinical model designed to simulate the detection of tumors embedded on the left or right flank of a mouse model. We found that the current reconstruction method accurately determined the absence of a source, and detected all of the cases with single sources with at least 1.6 μg of nanoparticles. However, it was almost completely unable to identify a second source, misclassifying all but two out of 34 two-source distributions as having only a single source. This has serious implications for clinical applications, in which these missed cases represent false negatives. This could result in a high number of cancers going undetected or being under-staged, leading to failures in treatment and recurrence. The new reconstruction method performed much better on the cases with two sources. MRXImage had the better performance by all measures of the three blinded analyses of the algorithms with an accuracy of 80%. MRXImage found false positive sources in only two of the 20 cases with one source, but none when there was no source present. MRXImage correctly classified 25 of the 32 two-source cases as having two sources, showing a vast improvement over MSA.

In a meta-analysis we investigated how the number of stage positions, method of classification given a reconstruction, and errors in the forward model affected the overall classification performance of the algorithms. We found that an objective analysis based on simple criteria on the clusters of the reconstruction from MRXImage resulted in approximately the same performance as a human reader. This indicates that MRXImage is largely unaffected by reader choices, and that the classification could potentially be easily automated for standardized outcomes. We also found that increasing the number of stage positions from 4 to 6 improved the accuracy of the classification by 5 to 7% across all of the algorithms. This indicates that although simulation studies found little to no benefit for two-source reconstructions above 2 stage positions, that when the complexity of real data is considered, more stage positions can indeed improve the results. Further investi-

gation is warranted to better determine how the performance of the algorithms changes as a function of stage positions with each combination of 1 to 6 stage positions and different source strengths. Finally, we found that even small errors in the forward model can have a large detrimental impact on the overall performance of MRXImage.

Chapter 7

Conclusions and Future Work

Superparamagnetic relaxometry (SPMR), could potentially improve the early detection of cancer. However, progress toward translation into the clinic is currently impeded by the inability of the current reconstruction method to reliably distinguish multiple clusters of cancer-bound nanoparticles within a sample. In this work, we demonstrated that a compressed sensing approach to the reconstruction of SPMR data improved the detection of multiple sources over the current method of reconstruction.

7.1   Theory and measurement methods

The physics of nanoparticle relaxation is highly complex and is still an active area of research. Throughout this work, we apply simplified models of this complex system to allow us to analyze the feasibility of our methods to a first approximation. We make one such simplification by treating clusters of immobilized nanoparticles as a single dipole moment and assume no interactions between the particles or neighboring clusters. While this assumption is valid for small clusters of highly immobilized particles, such as the cotton swab phantoms used in this study, it may not be true in all cases. For example, in *in vivo* models, the nanoparticles will likely be dispersed throughout an organ such as the liver, which cannot be said to resemble a point source. If the particles are traveling

through a tight cluster of lymph nodes, there may be interaction between the particles in neighboring lymph nodes. In either case, the condition that the distribution of particles resembles a sparse field of point sources may not be adequate. Faced with this scenario, it is likely that the reconstruction algorithm presented here would identify the centroid of neighboring clusters or the volume containing the nanoparticles. Future work should be done to experimentally determine the physical limits of these assumptions through a spatial resolution study with point sources and the development of new phantoms containing bound nanoparticles distributed throughout a volume. Modifications to the optimization problem, such as the use of a total variation term, may extend the application of the algorithm to such cases and should be investigated.

We also assume that the particles are either "bound" and thus relax entirely through the Néel process, or free to rotate, and thus relax entirely through Brownian processes. As demonstrated in Chapter 3, this approximation is suitable for the work presented here. However, it is possible that multiple modes of decay could be detected and isolated from the decay curve. This would open the possibility to determine not just whether particles are immobilized, but by what mechanism. For instance, the portion of the residual field due to aggregated particles could be distinguished from the portion due to cancer-cell bound particles due to the subtle differences in the decay constants. Whether this is possible given the current sensitivity of the device should be determined in future experiments. If it is possible, then the values of α_{est} from each decay mode could simply be reconstructed separately, resulting in a separate "cell-bound" distribution and an "aggregation" distribution. The potential for this has been investigated by other groups in the context of magnetic particle imaging, and applying these methods to SPMR would be straightforward and worthwhile [107].

7.2 The forward model

In Chapter 4, we presented a procedure for calibrating the location of the sensors used in the forward model to the data. This process is necessary because it is impossible to measure the exact location of the gradiometer coils once they have been placed inside the cryogenic dewar. Using the data-calibrated sensor locations, we found that our forward model agreed with the measured data within 2%. The data we collected for the calibration also allowed us to evaluate the variation in the measured data, as well as the average agreement between the model and the data. We found that the variation between values of repeated measurements was not dependent on the magnitude of the measurements. We also found that the variation in repeated measurements was different for each sensor. The average deviation of the measured field values from the field values predicted by the model over all of the sensors was 4.9×10^{-4} pT. In our simulation studies, we used this knowledge of the measurement characteristics to model measured data.

In the validation study, we showed that using incorrect sensor parameters can decrease the overall performance of the algorithm. Therefore, it is important that these values are well known when any data set is analyzed. Most of the sensor location parameters are physical dimensions that are determined when the gradiometer array is built and therefore can be validated with manual measurements of the array and should not change over time. We were unable to verify the measurements explicitly through physical measurement because it would require removing the probe stack from the cryogenic dewar, which is prohibited by the manufacturer. Therefore, we relied on the physics model to determine the values implicitly from measurements. This also acts as a secondary validation of the forward model, since the optimal parameters agreed with the drawings that we had been provided by the manufacturer. The only sensor location parameter that can change after the array is constructed is the dewar rotation angle. The name may be a misnomer, since even though the dewar can not move in the housing structure of the MRX device,

we have found that the gradiometer array can rotate within the dewar after installation. Improvements to prevent this in the future are currently in development, but this further emphasizes the importance of this calibration procedure. The data collection for the entire procedure requires several days, so it is not feasible to do on a frequent basis. However, for future studies we recommend a much shorter procedure to spot check the forward model is undertaken before every study. If the spot check identifies a possible error between the model and the data, the full calibration procedure can be undertaken to determine the cause of the discrepancy and recalibrate the forward model if necessary. In addition to diagnosing and correcting errors, the calibration procedure should be performed immediately after installation and any major hardware changes to establish an initial set of sensor location parameters and verify that the forward model and data agree with the technical drawings of the gradiometer array.

The calibration procedure presented here can easily be applied to any gradiometer geometry. For example, in the future parallel gradiometers may be employed in place of the rotated array described in this work. In this case, the sensor location parameters used in this work no longer be able to adequately describe the location of each sensor. Instead, new parameters will need to be derived that describe the new geometry of the array. Then, the location of the gradiometers within the forward problem would need to be adjusted according to the new geometry. Once these changes are made, the calibration procedure and reconstruction algorithm can be applied without further modification.

7.3   Sensitivity to parameters

We then used our validated forward model to perform simulated experiments to investigate the effect that choices of experimental design and reconstruction parameters had on the overall reconstruction accuracy. The experimental design parameters we tested included the use of one to nine stage positions and taking between 1 and 30 pulses per

stage position. For each experimental design parameter we tested, we also examined how the reconstruction was affected by the choice of data fidelity parameter. We tested these conditions on simulations of one and two dipole sources.

We found that for a single dipole source, the use of more than one stage position conferred only marginal increases in the accuracy of the reconstruction. For the simulation of two sources, the addition of a second stage position more than doubled the accuracy of the reconstruction. This result follows from the math behind the least-squares fit of the forward model to locate a known number of sources. Given that only one dipole is present, there are four unknown variables to solve for: the moment, and three dimensions of location. Therefore, measurements at only four sensor positions are needed to reconstruct the source. Since each stage position provides seven measurement locations, a single stage position is sufficient to reconstruct a single source. Each additional source introduces four additional unknowns, and therefore requires four additional sensor locations. This is why for the two source distribution, which would require at least eight measurement locations, we see a substantial loss in accuracy when the reconstruction is based on only one stage position (which provides only seven sensor locations). The accuracy is restored when the number of sensor locations is more than 8, in our case by adding a second stage position. While this intuition is based on reconstructing the sources from a typically over-determined least-squares fit of the forward model, it is interesting that the limits are not overcome by using the under-determined compressed sensing approach.

Future work should be done to extend these preliminary findings. First, it cannot be assumed that these results hold for distributions of more than two dipoles. While it is mathematically feasible to solve for n dipoles with at least 4n sensor locations, the reconstruction may not return the correct distribution. By design, the algorithm aims to find the fewest number of dipoles that can account for the detected magnetic field within the limits imposed by the data fidelity parameter. This means that if the algorithm can account for the field detected at the sensor locations within the limits imposed by the

data fidelity parameter with a single dipole, it will always return this solution. The only way to avoid this is to sample the magnetic field at locations that best distinguish the true magnetic field pattern from one that can be fit with a single dipole. Without prior knowledge of the true dipole locations, this is an impossible task. However, each additional sensor location improves the probability that these differences are captured.

It is also possible that the sensor locations that provide the best chance to reconstruct multiple sources can be determined mathematically. In this work, we attempted to use the conditioning number as a measure of the ability of the sensing matrix $\mathbf{A}$ to accurately reconstruct the true dipole distribution. However, we found that it did not serve as an appropriate measure of propagation of error from the measurement to the reconstruction as it should, and therefore is unlikely to be a reliable measure of quality. This is likely due to the fact that this behavior relies on the sensing matrix to satisfy the relationship $\boldsymbol{x} = \mathbf{A}^{\dagger}\boldsymbol{b}$, which does not hold for our sparse reconstruction approach. Other metrics for evaluating the potential information in a sensing matrix have been proposed in the literature, but only apply to sufficiently incoherent matrices, of which ours is not. Future work should be done to determine how to optimize the sensing matrix. Potential directions include minimizing the mutual information between sensor locations or using adaptive grid spacing to reduce coherence.

Additionally, this work considered only one possible set of stage positions. Further work should be done to confirm that these results hold for different choices of stage positions, or for these stage positions in a different order. The pattern of sensor locations for these stage positions is dependent on the rotation of the sensor array within the dewar, and therefore will vary between devices. It is important that these findings are confirmed using other sensor array rotation angles before the results are considered applicable beyond the specific setup used here. It is highly likely that the location of the sensors is as or more important than the number of sensor locations for cases with at least 4 times as many sensor locations as potential dipoles. Therefore, while the overall conclusions should be

consistent, the specific values of accuracy may be different for a different sensor geometry.

We used simulations to determine the effect of the choice of the data fidelity parameter λ on the accuracy of the reconstruction. We simulated measurements of a single source and two sources using three stage positions, and between 1 and 30 pulses per stage position over a range of SNR from 0 dB to 20 dB. We reconstructed these simulations using 40 values for the data fidelity term λ which balances the sparsity and the accuracy of the solution. We investigated three values of λ specifically. The true value of the error, $\lambda_{\Delta \boldsymbol{b}}$, can only be known in simulations because it is dependent on the true value of the magnetic field from the simulated source, but is a helpful reference point for evaluating other values of λ. In particular, we investigated two methods of determining λ from the measured data, λ_{SE} and λ_{σ} . While both are based on the chi-squared statistic, λ_{SE} adjusts for the number of pulses that are averaged to determine the mean value of the magnetic field. Therefore, as the number of pulses per stage position increases and the estimation of the true magnetic field increases, λ_{SE} accordingly reduces the allowed error in the reconstruction. In our simulations, λ_{SE} closely approximated $\lambda_{\Delta \boldsymbol{b}}$ and produced much more accurate reconstructions than λ_{σ} , especially in simulations with more than 15 pulses per stage position. However, in real measurements for which no more than 15 pulses were collected per stage position, the measurement uncertainty is not identical between the sensors, λ_{SE} tends to underestimate the error, leading to noisy and inaccurate solutions. Therefore, we used λ_{σ} for our phantom studies which slightly overestimates the error in simulations, but is a better approximation of the real uncertainty in measurements.

Future work should be done to further refine the method for determining the value of λ. Currently, the calculation uses the measurements from all of the sensors to estimate the uncertainty and then applies that uncertainty to all of the sensors. However, from the results in Chapter **??**, we know that the sensors do not have the same uncertainty characteristics. Therefore, it may be beneficial to account for the uncertainty of each sensor separately when determining λ. This should be validated through simulations and

phantom studies, as was done here, for both λ_{SE} and λ_σ . Additionally, it is possible that the constraints in the optimization equation could be separated by sensor, rather than combined into a single parameter. This work considered only a single parameter for simplicity, with the aim to show the feasibility of the overall method and maintain generality with the literature. However, there is no reason that going forward the algorithm could not be tailored to be more specific to this application.

Our simulation studies also aimed to determine the appropriate number of pulses that should be collected per stage position. We found that for reconstructing a single source with a magnitude at least as large as the magnitude of the uncertainty in the measurements, there was little benefit to collecting more than ten pulses at a single stage position. However, there may be benefit to additional pulses per stage position when the magnitude of the sources is less than the magnitude of the uncertainty of the measurements. For the characteristic uncertainty of our system, this would correspond to a single source less than $3.35 \times 10^3 \, \mathrm{pJ\,T^{-1}}$ no more than 3 cm below the gradiometers, but may be more or less for other systems depending on the uncertainty of repeated measurements. Additional simulations should be done to extend the results to values of SNR less than zero.

For distributions of multiple sources, we found a substantial benefit to additional pulses for SNR values less than 10. Reconstructions of two sources of $2.0 \times 10^3 \, \mathrm{pJ\,T^{-1}}$ that were undetected within an ROI around the true source location, were successfully detected when 30 pulses were used. For sources of $7.9 \times 10^3 \, \mathrm{pJ\,T^{-1}}$, the accuracy of the reconstruction doubled when 30 pulses were used rather than only 10. These results indicate that while a single source may be reliably detected with only a few pulses per stage position, the chance of successfully reconstructing distributions containing multiple sources is greatly improved with additional pulses per stage position. This should be considered carefully when designing future experiments for which the number of bound nanoparticle clusters is unknown. Future work should be done to confirm these results in distributions of more than two sources, and for distributions of sources at different depths or of different strengths

within the same field of view.

Future work should be done to evaluate other reconstruction parameters such as the size of the voxels, the layout of thee voxels, and the extent of the field of view. These parameters are visited briefly in Chapter 6, but due to the design of this study and the sparse nature of the reconstruction, these parameters are not expected to substantially affect the results. However, it is likely that these parameters, in conjunction with the data fidelity parameter and experimental parameters presented here, would have an effect on the spatial resolution of the algorithm - that is, the distance between two sources at which the algorithm reconstructs a single source. It is also likely that these parameters will have an important role in the ability of the algorithm to distinguish distributions containing more than two sources. These important interactions are therefore left to future work.

7.4 Performance

Finally, we validated the reconstruction algorithm on measurements of phantoms, and compared it against the current method of reconstruction, MSA. In our phantom studies, we validated the choice of λ_σ as the data fidelity parameter, then confirmed that the reconstruction algorithm was comparable to the current method for reconstructing a single source, and superior for reconstructing multiple sources. We confirmed these results through a user-blinded tumor detection study to approximate how this improvement would affect the detection of tumors in pre-clinical trials.

The first phantom studies aimed to validate that our compressed sensing reconstruction algorithm, MRXImage, produced results in good agreement with what is currently used for SPMR studies, MSA, on reconstructions of a single source for which MSA is known to perform well. First, we confirmed the choice of λ_σ as the data fidelity parameter through a similar analysis as was applied to the simulation studies. We found that, contrary to the results of the simulation studies, λ_σ resulted in better reconstructions, meaning more

consistent values of magnetic moment per mass of nanoparticles and fewer false positive sources, than λ_{SE} . This result may be surprising, but as discussed in Chapter 5 is likely a better estimate of the true error due to the number of pulses used in this study and the variation of the uncertainty in the measurements between sensors. Based on these studies, we recommend the use of λ_{σ} in future experimental work. However, as discussed in Section 7.3 there are still improvements that can be made to how this parameter is defined, specifically accounting for the sensor-specific uncertainties individually.

A linear titration of mass of nanoparticles dried on the tip of a cotton swab confirmed that the magnetic moment reconstructed by both algorithms is linear with the mass of nanoparticles. The value of the magnetic moment reconstructed by MSA was found to be higher than that of MRXImage by a factor of 1.27. This is likely due to a combination of a scaling factor that is included in MSA to convert the Volts output by the SQUIDs to magnetic field and differences in the way the value of the magnetic field is determined from the decay curves in the preprocessing algorithms. This difference can be easily accounted for by calibrating the signal to a known mass of immobilized nanoparticles, and has no impact on the final goal of the reconstruction algorithm as long as it is known and accounted for. We also found that both algorithms have similar accuracy in locating the dipole moment as a function of mass of nanoparticles. For both algorithms, the $0.1\,\mu g$ source was unable to be reconstructed. For MSA, this means that it was either reconstructed with a magnetic moment of zero or located outside of the physically feasible limits of the system. For MRXImage, the value of λ_{σ} was greater than the magnitude of the detected field, indicating that the SNR was too low to reliably reconstruct a source. Sources between $0.5\,\mu g$ and $1\,\mu g$ were able to be reconstructed by both algorithms, but the accuracy of both the location and the strength of the source was low. In these cases, the majority of the error came from the sources being reconstructed closer to the detectors and weaker than expected. Sources greater than $1\,\mu g$ were reconstructed within 10% and 5 mm of their expected strength and location, respectively. Sources greater than $5\,\mu g$ were reconstructed

within 5% and 3 mm of their true strength and location, respectively. Going forward, this information can be used in conjunction with the expected number of bound nanoparticles per cell to determine the appropriate dose of nanoparticles for animal and clinical trials.

The motivation behind developing a new approach to reconstructing SPMR data was the poor performance of the current algorithm when reconstructing distributions of more than one source. The two-source phantom studies and the blinded tumor detection study show that MRXImage improves our ability to distinguish multiple sources in the same measurement. When measured alone, MSA could reliably detect any source greater than 0.5 µg of dried nanoparticles. When two sources were present, MSA missed nearly one third of the sources with at least 0.5 µg of dried nanoparticles, while MRXImage missed only two. For the sources that were successfully detected, MRXImage reconstructed their strength more accurately. The location accuracy was high for both algorithms for sources with at least 10 µg of nanoparticles. For smaller sources that were reconstructed by both algorithms, MSA had slightly higher location accuracy. To summarize, for measurements of two sources located approximately 2 cm apart, MRXImage reconstructed sources containing at least 10 µg of nanoparticles within 5% and 2 mm of their expected strength and location, respectively, and successfully detected most sources with between 1 µg and 10 µg of nanoparticles. MSA was able to accurately reconstruct most sources with at least 10 µg of nanoparticles, but missed two thirds of the sources with between 1 µg and 10 µg of nanoparticles. These results indicate that when multiple clusters of bound nanoparticles are possible, for example in *in vivo* models, MRXImage is more likely to successfully reconstruct both sources.

To test how the algorithm is likely to perform in pre-clinical scenarios, we performed a user-blinded classification study. This study was designed to mimic the scenario in which a mouse has been injected with targeted nanoparticles with the aim to detect the presence or absence of a tumor. Based on our preliminary experience, we expect that a large portion of the nanoparticles will be taken up by macrophages and collect in the liver, resulting in

a large signal approximately 2 cm from the potential tumor site. For cases containing zero or one source, MRXImage and MSA had similar results. Each misclassified two one-source cases. MSA classified two cases as having no source, and MRXImage classified them as having two sources. The clinical implications of missing a source are much greater than finding two sources when there is only one. Out of the 34 cases containing two sources, MSA correctly classified only two, while MRXImage correctly classified 24. This result further strengthens thee conclusion that for multiple source distributions, MRXImage is more reliable that the current method of reconstruction.

To tie the work together, we also evaluated the effect of using incorrect forward problem parameters (from Chapter 4), and fewer stage positions (from Chapter 5) on the overall classification accuracy. When only four stage positions were used for the reconstruction instead of six, the accuracy was decreased by 7%. The reduction in stage positions did not reduce the ability of the algorithm to distinguish the presence of at least one source from the absence of any sources. However, the ability of the algorithm to identify the correct number of sources was reduced. This implies that for applications in which only the presence or absence of any source is needed and the correct number of sources is not necessary, six stage positions would not provide any benefit over four. This may be the case for applications where tissues samples are being measured *ex vivo* to determine whether any disease is present. However, if the number of sources is important, such as when samples are being measured *in vivo* and may be in close proximity to an area with non-specific binding, then additional stage positions may be beneficial. We also found that errors in the sensor locations of only a few millimeters can reduce the overall accuracy of the algorithm by more than 10%. This further emphasizes the importance of calibrating the forward model and confirming the accuracy of the current model before conducting any experiment.

7.5   Summary and conclusion

The overall goal of this study was to develop and validate a novel SPMR reconstruction algorithm that improved upon the limitations of the current algorithm, including the requirement for prior knowledge of the number and approximate location of the clusters of nanoparticles, and the poor performance of the algorithm for reconstructions of two sources. In this work, we introduce, characterize, and validate, a compressed sensing approach to the reconstruction of the distribution of immobilized nanoarticles from SPMR measurements. We first derive the algorithm from the underlying physics of the nanoparticle relaxation through the SPMR process, and describe how the pulse sequence and physical characteristics of the nanoparticles and their environment can affect the overall reconstructed magnetic moment. These relationships are vital to understanding how one translates the magnitude of the reconstructed moment that is returned by the algorithm to the desired endpoints of mass of bound nanoparticles, or number of cancer cells. Future work on the analysis of the relaxation curves may provide further insight into the source of the signal, enabling the distinction of cell-bound particles from aggregated particles. These fields could then be reconstructed separately to produce a spatial distribution of each interaction type.

We then present a method to calibrate the forward model to measured field values. With our method, we found that the calibrated forward model can predict the measured field values within 2%. Additionally, this analysis provides us valuable information about the uncertainty between repeated measurements for each sensor, which is used later for simulating measurements. The importance of having an accurate forward model is demonstrated in the blinded tumor detection study, where we show a reduction in accuracy of more than 10% when the forward model parameters are incorrect. Future work will be needed to adapt the forward model used here to new sensor array geometries. With the appropriate modifications to the forward model, the reconstruction algorithm will be applicable to any arrangement of sensors.

Through simulation studies, we find that we can determine an appropriate value for the data fidelity parameter that balances the sparsity of the reconstruction with how well it fits the measured data. We also determine that for reconstructions of a single dipole moment, only a few pulses at a single stage position are needed to achieve high accuracy. However when two sources are present, at least two stage positions are required regardless of the noise in the data, and that the accuracy of reconstructions in low SNR conditions can be as much as doubled by increasing the number of pulses per stage position. Future work should be done to extend these results to SNR conditions less than 0 dB, and for different source distributions, including those with more than two sources. There is a significant amount of work to be done regarding the mathematical characterization of the sensing matrix, which is largely coherent and therefore most common methods of analysis are not applicable.

Finally, we found that our algorithm matched the performance of the current algorithm on measurements of a single phantom, but repeatedly outperformed the current algorithm on reconstructions of two sources. For reconstructions of a single phantom, we found that the objectively determined data fidelity parameter provided accurate reconstructions. We also found that for both algorithms, single phantoms measured on our system with less than $0.5\,\mu g$ of immobilized nanoparticles could not be distinguished from background. Phantoms with between $0.5\,\mu g$ and $10\,\mu g$ of immobilized nanoparticles could be detected but were reconstructed closer to the detectors and weaker than expected. Phantoms containing $10\,\mu g$ or more immobilized particles were reconstructed within 5% and 2 mm of their expected strength and location, respectively. This knowledge will be useful when designing dosages for future *in vivo* experiments. For measurements of two phantoms, MRXImage repeatedly out performed MSA on both known distributions and blinded analyses. Most notably, MRXImage more than doubled the accuracy of MSA in a blinded classification study based on tumor detection. This means that the current limitations facing the progress of small animal SPMR studies, namely the inability to reliably distinguish

two sources and the need for *a priori* knowledge of the distribution of bound particles, can be overcome with our new reconstruction algorithm. Future work is still needed to extend these results to smaller sources, volumetric sources, and distributions of more than two sources. Additionally, while the algorithm has been thoroughly validated in phantom studies, it has yet to be validated in *in vivo* scenarios. In contrast to the phantom studies presented here, the bound nanoparticles may not be in tight point source-like clusters in a small animal model. It is currently unknown how the sparse reconstruction will handle sources as they approach the limit of sparse clusters. However, even in this case MRXImage should still out perform MSA, which also assumes singular clusters of bound nanoparticles. Additionally, adjustments can be made to the optimization performed by MRXImage, such as an additional term on the total variation of the solution, to allow for more volumetric reconstructions.

In conclusion, we have developed a reconstruction algorithm that overcomes the limitations of the current algorithm that were impeding the progress of *in vivo* studies. While there is still more work to be done, what is presented in this work constitutes a significant advancement in SPMR technology that is an important piece to the translation of this technology into the clinic. With the help of MRXImage, SPMR is one step closer to reducing the high mortality rate of ovarian cancer through safe and effective early detection.

Appendix A: Iterative reweighting for L0 approximation

8.1 Iterative sparsity

Because the l_1 norm is not truly a measure of sparsity, the optimal l_1 solution may consist of many very small but non-zero elements. This may result in the major components slightly underestimating the true strength of the signal. To minimize this effect, through an iterative regularization procedure the small elements in the L1-optimized result can be forced to zero, while the larger elements are made increasingly accountable for the detected field. We introduce a px1 reweighting vector Φ, initialized to all 1's, to the minimization term in 3.53.

$$\min_x ||\Phi x||_1 \qquad \text{such that} \qquad \begin{cases} ||Ax - b|| \leq \lambda \\ \\ x \geq 0 \end{cases} \tag{B.1}$$

Thus, there is initially no change to the system. After the l_1 solution is found, Φ is redefined for the next iteration following Equation B.2.

$$\Phi^{i+1} = \frac{1}{x + \epsilon} \tag{B.2}$$

In Equation B.2 for large values of x^i, the corresponding element of Φ^{i+1} is made small, while where x^i is small (relative to the factor ϵ) Φ^{i+1} is large. In the next iteration of the optimization, small values of x are further minimized while the larger components are left to satisfy the tolerance conditions. The iteration is stopped when there is negligible change ($||x^i - x^{i-1}|| < 1e - 6$) in the solution vector x between iterations. The full algorithm, with bias correction and sparsity iteration, is shown in Algorithm 1.

Algorithm 1 CVX

Require: A, b

Require: Parameters λ, ϵ

   $G = \mathrm{diag}(\frac{1}{||a_j||_2})$

   $\hat{A} = AG$

   $\Phi^1 = \vec{1}$

   k=1

   while $||x^k - x^{k+1}||$ is small **do**

      minimize $||\Phi^k x||_1$ subject to $||\hat{A}x - b|| \leq \lambda$ and $x \geq 0$

      $\Phi^{k+1} = \frac{1}{x+\epsilon}$

      k++

   end while

   return $x = G^{-1}x$

8.1.1 Effect on reconstructed moment

The reconstruction of a single point source phantom containing 50 µg of dried nanoparticles located in the center of the field of view before and after reweighting to approximate the L0 norm is shown in Figure B.1. The reconstruction, seen from above, shows a reduction in the number of non-zero voxels far from the location of the true source after just one reweighting iteration. Little change is made by subsequent iterations beyond the first.

Figure B.2 shows the total moment in the reconstruction of a single point source titration at the center of the field of view after the first (L1, denoted by an x) and fifth (L0 approximation, denoted by a +) reweighting step. A ratio paried t-test showed no significant difference in the reconstructed moment after reweighting (p = 0.253), and a correlation coefficient of 0.9998. Therefore, we can conclude that there is no benefit gained for the cost of the additional computational time required to recompute the reconstruction with the reweighting for single source distributions.

161
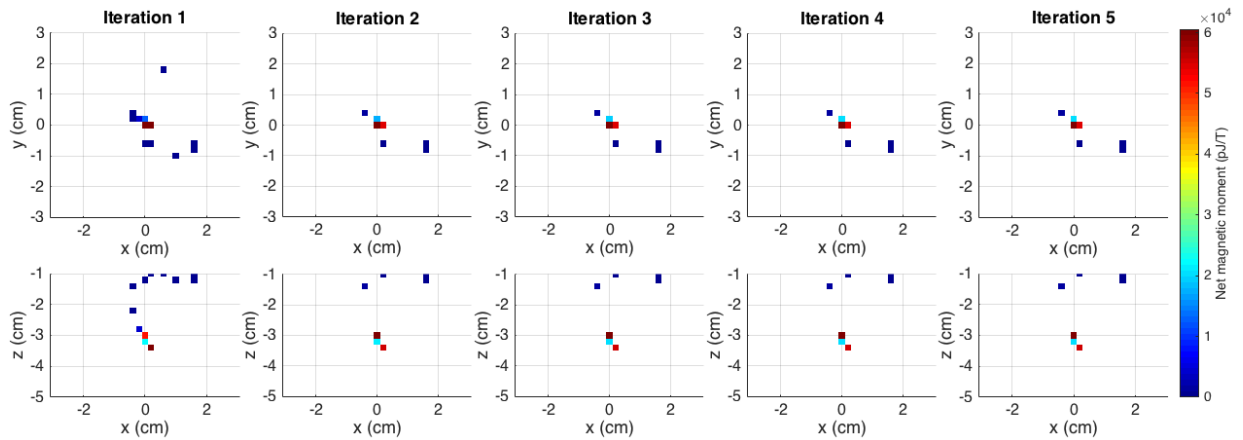
Figure B.1: The SARA reconstruction of a single source containing 50 μg of immobilized nanoparticles placed at the center of the field of view after each sparsity-enhancing reweighting step.
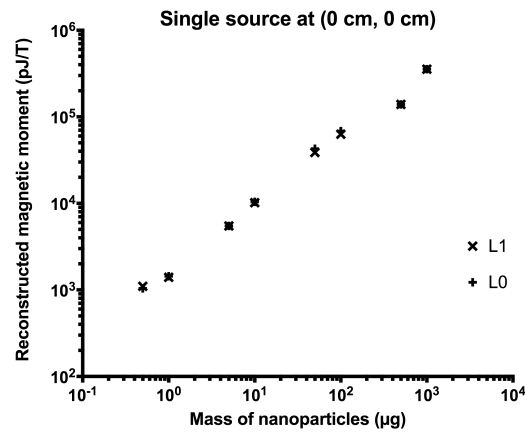
Figure B.2: The reconstructed moment of a single source located at the center of the field of view containing 0.5 μg of nanoparticles before (L1, denoted by an x) and after (L0 approximation, denoted by a +) five additional reweighting steps to increase the sparsity of the solution and better approximate the L0 norm.

As can be seen Figure B.2, the extra reweighting iterations did not change the total moment reconstructed within the ROI by MRXImage. This is because the reweighting steps simply compress the moment contained in the largest voxels of the L1 solution into a single voxel that best approximates the centroid of the L1 solution. To illustrate this effect the reconstruction of the 50 μg source at the initial L1 solution, and then after each reweighting step, is shown in Figure B.1. For this reason, only the L1 solution will be considered for this work, unless otherwise noted.

Appendix B: Analysis of simulation results with the normalized error metric

To avoid the drawbacks of the Error_x metric discussed in Section 5.2 and demonstrated below, we chose to present the results of the simulation studies in terms of ROI accuracy within the main body of the text. For those readers who are interested to see how the results of the analysis using the Error_x metric, they are presented here.

The normalized error (Equation 5.11) as a function of λ for a single source reconstruction using 10 and 30 pulses is shown in Figure C.1. Selected reconstructions of this data are presented in Figures 5.12 and 5.13. The shape of the error curve as a function of λ follows the same general trend at all levels of SNR, to varying degrees. The curves start out flat at low values of SNR, at which point the reconstruction has a high degree of noise (see the first row of Figures 5.12 and 5.13). As the voxels near the top of the field of view adjust to satisfy the data fidelity constraint, Error_x stays relatively constant because the true solution is zero for these voxels. The second phase of the pattern, the sparsity begins to increase and the reconstructed moment moves from many voxels at the top of the field of view toward a cluster of a few voxels at the center. The voxels within the ROI become increasingly accountable for reproducing measured field. At low levels of SNR, the Error_x becomes highly unstable, as the non-zero voxels in the reconstructed solution may or may not correspond to the single non-zero voxel in the true solution. Eventually, the reconstructed solution condenses to the single non-zero voxel in the true solution. Past this point, the Error_x steadily increases as the permitted deviation from the detected field is compensated for by decreasing the value of the single voxel.

For all but the smallest value of SNR, λ_σ selects a sparse solution, for both 10 and 30 pulses. The reconstructions at 10 SNR and above have mostly converged sparse solutions

at the values of λ less than or equal to λ_{SE}. For these SNRs, the value of λ_{SE} is fairly close to the minimum Error$_x$. The value of λ_σ, while sparse, is well into the region in which the magnitude of the solution is decreasing to compensate for the allowed error, resulting in increased Error$_x$. For the 7 db SNR case, λ_{SE} corresponded with the minimum Error$_x$ reconstruction for 10 pulses per stage position, but underestimated the minimum Error$_x$ solution for 30 pulses per stage position, resulting in a reconstruction that was stretched over multiple voxels. However, the solution at λ_{SE} better estimated the total moment of the true solution than the reconstruction with the minimum Error$_x$, which further highlights the weaknesses of the Error$_x$ measure.

The difference between 10 and 30 sample trials was most evident at 3 dB SNR. With only 10 pulses per stage position, Error$_x$ is highly unstable and greater than 1 at λ_{SE}, indicating that none of the non-zero components of the solution correspond to the non-zero voxel in the true distribution. At λ_σ, the solution has reached a sparse solution, but the error is still very high. With 30 pulses, the solution has converged to a low Error$_x$, sparse configuration at λ_{SE}, and a lower Error$_x$ at λ_σ. This implies that additional pulses per stage position may improve the reconstruction at lower levels of SNR.

Much of the shape of the curves is due to the characteristics of the Error$_x$ metric, making it hard to draw specific conclusions about the reconstructions. During the first phase of the trend, Error$_x$ is higher for the lower SNRs because the largest voxel of the reconstruction is slightly closer to the detectors than the true solution. Due to the definition of Error$_x$, it will always be greater than 1 until the reconstructed solution puts at least some non-zero moment in the voxel corresponding to the true solution. Because of this, Error$_x$ is a measure of both the source component of the reconstruction and the noise at the top of the field of view. As the SNR increases, the reconstruction puts at least some component of the solution in the corresponding voxel, which explains why the early values of Error$_x$ are less than 1. For very high SNR the majority of the reconstructed solution is contained in voxels corresponding to non-zero components of the true solution, even for

very small values of λ, so the Error$_x$ is largely a measure of the noise at the top of the field of view.

The normalized error in the reconstruction, defined in Equation 5.11, of two dipoles as a function of λ at each SNR is shown in Figure C.2 for 10 and 30 pulses per stage positions. Corresponding plots for the remaining pules per stage position are included in Appendix **??**. The error in the reconstructions with values of λ equal to the actual distance between the true and detected field, λ_{SE}, and λ_σ, are denoted by blue, red, and black dots, respectively. The reconstructions at selected values of λ are shown in Figures 5.15 and 5.16, for 10 and 30 pulses, respectively.
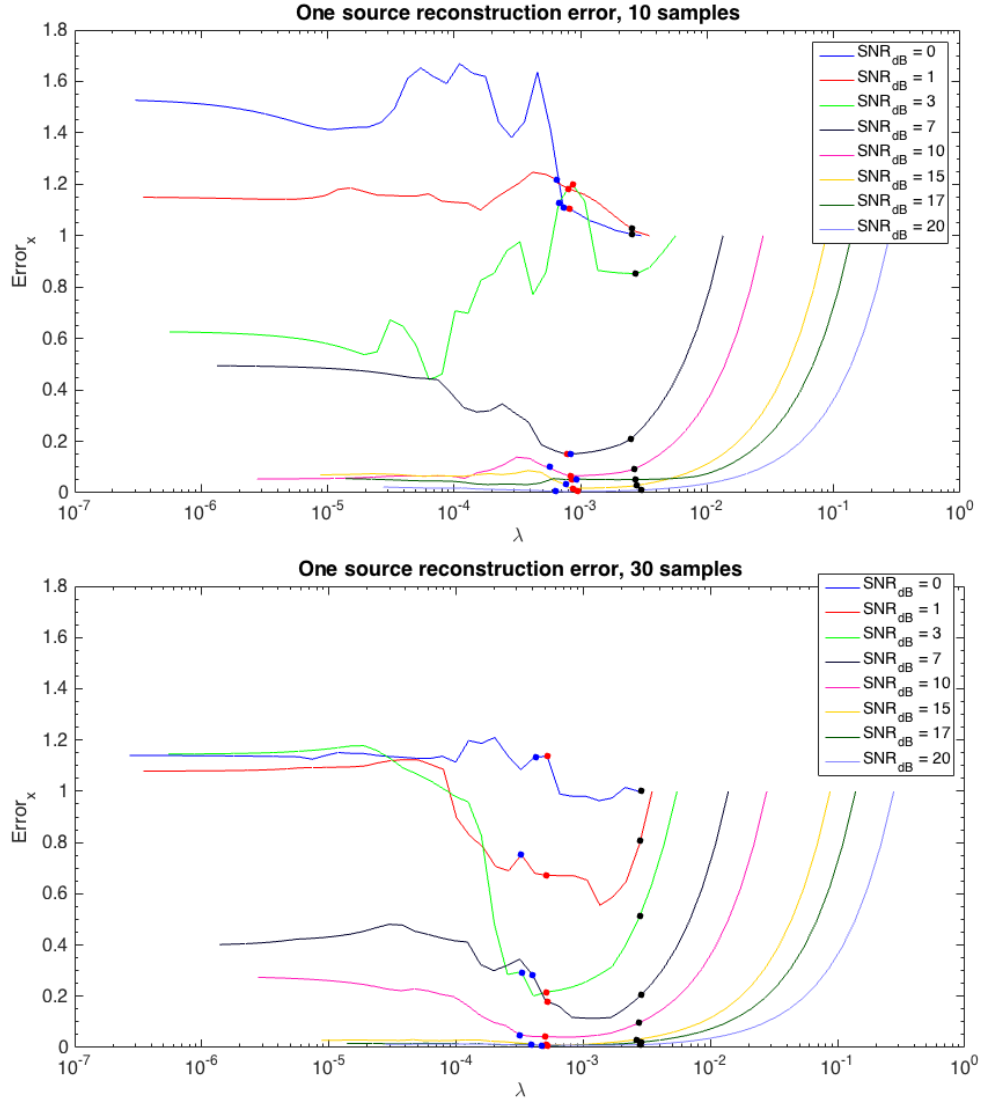
Figure C.1: The normalized error (Error_x) in the single source reconstruction using 10 pulses per stage position (top) and 30 pulses per stage position (bottom) for signal to noise ratios from 0 to 20 dB. The error from reconstructions using $\lambda = ||b_t - b^*||$, λ_{SE}, and λ_σ are denoted by blue, red, and black dots, respectively.
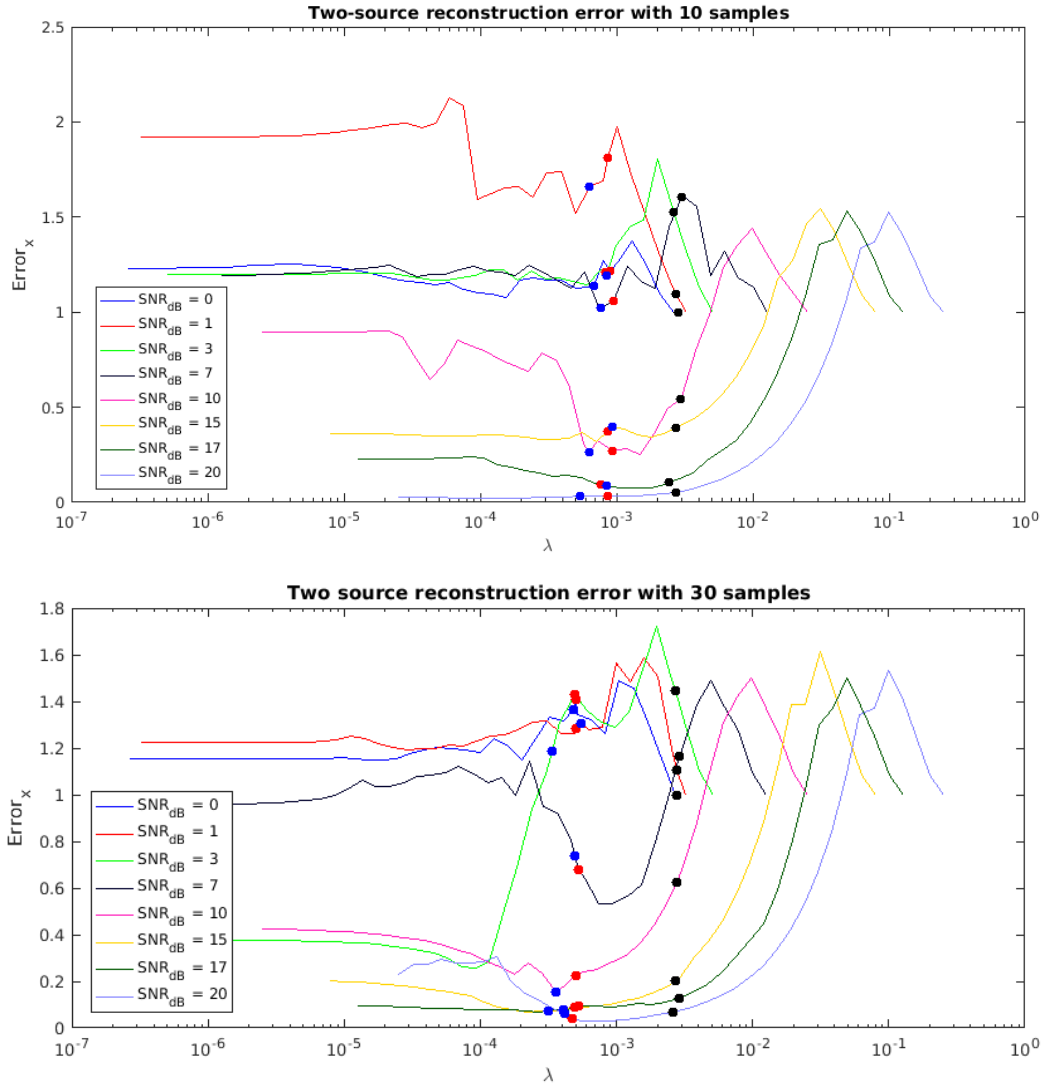
Figure C.2: The normalized error (Error_x) in the two-source reconstruction using 10 pulses per stage position (top) and 30 pulses per stage position (bottom) for signal to noise ratios from 0 to 20 dB. The error from reconstructions using $\lambda_{\Delta b}$, λ_{SE}, and λ_σ are denoted by blue, red, and black dots, respectively.

Appendix C: Cohen's Kappa Statistic

10.1   Cohen's kappa statistic

Cohen's kappa, κ, is a measure of agreement between two classifiers. In our study, one classifier was the reconstruction algorithm and the other was the true classification. The benefit of Cohen's kappa is that it provides a measure of classification accuracy corrected for the number of cases in each class. It does so by defining a ratio of the observed error to the expected error, based on the distribution of cases in each class. The observed error is the number of cases that are misclassified by the classifier. Mathematically speaking, the observed error is the number of cases in class j that were assigned to class $i \neq j$.

$$E_o = \sum_{i,j \in i \neq j} n_{i,j} \tag{B.1}$$

The expected error is defined as the expected number of cases in class j that will be assigned to assigned to class $i \neq j$. Mathematically, this is defined as the probability $(p(T_j))$ that a case truly belongs to class j times the probability $(p(C_i))$ that it will be assigned to class $i \neq j$ times the total number of cases, N.

$$E_e = \sum_{i,j \in i \neq j} p(C_i) p(T_j) N \tag{B.2}$$

The probability that a case will be assigned to class i is defined as the number of cases that were assigned to class i divided by the total number of cases. The probability that a case truly belongs to class j is equal to the number of cases truly in class j divided by the total number of cases.

$$p(C_i) = \frac{\sum_j n_{i,j}}{N}$$
$$p(T_j) = \frac{\sum_i n_{i,j}}{N} \tag{B.3}$$

The kappa statistic κ is then defined as the ratio of the errors subtracted from perfect accuracy (1).

$$\kappa = 1 - \frac{E_o}{E_e} \tag{B.4}$$

10.2 Cohen's weighted kappa statistic

In some cases, it may be reasonable to penalize some classification errors more than others. For instance, in our scenario classifying a source as being in the wrong location is not as bad as missing the source completely. We can account for these differences by including a weighting value $w_{i,j} \in \mathbb{R}^+$ to the penalty of assigning a case to class i when it is truly of class j.

$$E_{o,w} = \sum_{i,j \in i \neq j} w_{i,j} n_{i,j}$$
$$E_{e,w} = \sum_{i,j \in i \neq j} w_{i,j} p(C_i) p(T_j) N \tag{B.5}$$

The value of $w_{i,j} = 0$ when $i = j$ because there is no penalty for correct classification. Higher weighting values incur higher penalties. When $w_{i,j} = 1$ for all $i \neq j$, κ_w is equal to κ.

$$\kappa_w = 1 - \frac{E_{o,w}}{E_{e,w}} \tag{B.6}$$

10.3 Example cases

The values of κ and κ_w can range from -1 to 1. Positive values indicate that the classifier performed better than expected based on the distribution of cases in each class. Negative values indicate that performance was worse than expected. A value of 0 indicates that the accuracy of the classifier was no better than if every case were assigned to the same class. To further illustrate, we calculated the accuracy and kappa values for a variety of possible scenarios, included below in Table B.1.

Description	Accuracy	Kappa	Weighted Kappa
All correct	1	1	1
All classified as 0	0.1	0	0
All classified as 1	0.33	0	0
All classified as L	0.28	0	0
All classified as R	0.28	0	0
True 0 classified as 1. True 1, L, R classified as 0	0	-0.14	-0.13
True 0 classified as R. True 1, L, R classified as 0	0	-0.13	-0.20
All off by one source	0	-0.34	-0.07
True R classified as 0. True 0, 1, L classified as R	0	-0.30	-0.38
True 1, L and R classified as 1. True 0 correct	0.43	0.18	0.24
True 0 classified as 1. True 1, L, R correct.	0.9	0.86	0.85
True R classified as L. True 0, 1, L correct.	0.43	0.21	0.62

Table B.1: The number of single source test cases in the detection study with a given mass of nanoparticles.

Appendix D: Blinded study full results

Sample number	Source 1 mass (μg)	Source 2 mass (μg)	True clas-sification	MSA	MRXImage B6	MRXImage B4	MRXImage E4
1	50	1.56	**R**	1	R	1	R
2	0	N/A	**0**	1	0	0	0
3	12.5	1.56	**R**	1	R	R	R
4	25	0.78	**R**	1	1	1	1
5	50	1.56	**L**	1	L	L	L
6	3.13	N/A	**1**	1	1	1	1
7	50	1.56	**R**	0	R	1	1
8	1.57	N/A	**1**	1	1	1	1
9	12.5	1.56	**L**	1	L	L	L
10	25	N/A	**1**	0	1	L	L
11	0	N/A	**0**	0	0	0	0
12	25	0.78	**L**	1	L	L	L
13	0.78	N/A	**1**	L	1	1	0
14	25	3.13	**R**	1	1	L	1
15	12.5	0.78	**R**	1	R	R	1
16	25	3.13	**L**	L	L	L	L
17	1.56	N/A	**1**	1	1	1	L
18	25	0.78	**R**	1	1	1	1
19	1.56	N/A	**1**	1	1	1	1
20	50	3.13	**L**	1	L	L	L
21	25	1.56	**L**	0	1	1	1
22	25	N/A	**1**	L	R	1	1
23	12.5	0.39	**L**	1	L	L	1
24	0	N/A	**0**	0	0	0	0
25	0.78	N/A	**1**	0	1	1	1
26	0	N/A	**0**	0	0	0	0
27	3.13	N/A	**1**	1	1	1	1
28	25	0.78	**L**	1	1	1	1
29	50	3.13	**R**	1	R	R	R
30	25	1.56	**R**	1	R	1	1

Sample number	Source 1 mass (μg)	Source 2 mass (μg)	True classification	MSA	MRXImage B6	MRXImage B4	MRXImage E4
31	12.5	0.39	**R**	0	1	R	1
32	0	N/A	**0**	0	0	0	0
33	50	N/A	**1**	1	1	L	1
34	12.5	0.39	**R**	1	1	1	1
35	3.13	N/A	**1**	1	1	1	1
36	50	0.78	**L**	0	1	L	1
37	50	N/A	**1**	1	1	L	1
38	12.5	0.78	**L**	1	L	L	1
39	50	1.56	**L**	1	L	L	L
40	12.5	0.39	**L**	1	L	L	L
41	0.39	N/A	**1**	0	1	1	0
42	0	N/A	**0**	0	0	0	0
43	50	3.13	**L**	1	L	L	L
44	50	6.26	**L**	1	L	L	L
45	12.5	N/A	**1**	L	1	L	1
46	25	1.56	**R**	1	R	1	R
47	0.39	N/A	**1**	0	1	1	0
48	50	0.78	**R**	1	1	1	1
49	12.5	0.78	**R**	1	L	R	L
50	50	6.25	**R**	1	R	R	R
51	12.5	1.56	**L**	1	L	L	L
52	25	1.56	**L**	1	L	L	R
53	6.25	N/A	**1**	1	1	1	1
54	12.5	1.56	**R**	R	R	R	1
55	12.5	N/A	**1**	1	L	L	1
56	0.39	N/A	**1**	0	1	1	0
57	6.25	N/A	**1**	1	1	1	1
58	0.78	N/A	**1**	0	1	1	0
59	50	3.13	**R**	1	R	R	1
60	12.5	0.78	**L**	1	L	L	L

A

MSA-B No source		True classification	
		0	*1, R, or L*
Reconstructed	*0*	**5**	10
	1, R, or L	1	**44**

B

MSA-B One source		True classification	
		1	*0, R, or L*
Reconstructed	*1*	**11**	29
	0, R, or L	9	**11**

C

MSA-B Two sources Left		True classification	
		L	*0, 1, or R*
Reconstructed	*L*	**1**	3
	0, 1, or R	16	**40**

D

MSA-B Two sources Right		True classification	
		R	*0, 1, or L*
Reconstructed	*R*	**1**	0
	0, 1, or L	16	**43**

Table A.1: MSA One-vs-all results

MRXImage-B6 No source		True classification	
		0	*1, R, or L*
Reconstructed	*0*	**6**	0
	1, R, or L	0	**54**

A

MRXImage-B6 One source		True classification	
		1	*0, R, or L*
Reconstructed	*1*	**18**	9
	0, R, or L	2	**31**

B

MRXImage-B6 Two sources Left		True classification	
		L	*0, 1, or R*
Reconstructed	*L*	**14**	2
	0, 1, or R	3	**41**

C

MRXImage-B6 Two sources Right		True classification	
		R	*0, 1, or L*
Reconstructed	*R*	**10**	1
	0, 1, or L	7	**42**

D

Table A.2: MRXImage-B6 One-vs-all results

MRXImage-B4 No source		True classification	
		0	*1, R, or L*
Reconstructed	*0*	**6**	0
	1, R, or L	0	**54**

A

MRXImage-B4 One source		True classification	
		1	*0, R, or L*
Reconstructed	*1*	**15**	10
	0, R, or L	5	**30**

B

MRXImage-B4 Two sources Left		True classification	
		L	*0, 1, or R*
Reconstructed	*L*	**15**	6
	0, 1, or R	2	**37**

C

MRXImage-B4 Two sources Right		True classification	
		R	*0, 1, or L*
Reconstructed	*R*	**8**	0
	0, 1, or L	9	**43**

D

Table A.3: MRXImage-B4 One-vs-all results

MRXImage-E4 No source		True classification	
		0	*1, R, or L*
Reconstructed	*0*	**6**	5
Reconstructed	*1, R, or L*	0	**49**

A

MRXImage-E4 One source		True classification	
		1	*0, R, or L*
Reconstructed	*1*	**13**	16
Reconstructed	*0, R, or L*	7	**24**

B

MRXImage-E4 Two sources Left		True classification	
		L	*0, 1, or R*
Reconstructed	*L*	**11**	3
Reconstructed	*0, 1, or R*	6	**40**

C

MRXImage-E4 Two sources Right		True classification	
		R	*0, 1, or L*
Reconstructed	*R*	**5**	1
Reconstructed	*0, 1, or L*	12	**42**

D

Table A.4: MRXImage-E4 One-vs-all results

Appendix E: Additional blinded study results

12.1   MSST

The MSST algorithm developed by Huang et. al. improves on both the pre-processing and reconstruction methods of MSA [7]. First, it requires the user to manually exclude bad pulses in order to improve the overall signal average. Then, it uses a second-derivative approach to detect flux jumps and corrects them using a Taylor series expansion around the first data point prior to the jump. Once the bad pulses have been excluded and the flux jumps have been corrected, the algorithm determines the decay constant by fitting the strongest signal with either Equation 3.29 or 3.30 to determine the appropriate decay constant. Using this decay constant, it fits all of the decay curves with either Equation 3.29 or Equation 3.30 to determine the constants A1 and A2, or Ae1 and Ae2, respectively. Finally, the magnetic field map is determined using the ratio of the constant A2 or Ae2 at each sensor to that of the strongest, or reference, sensor. This avoids any inaccuracies that are encountered from extrapolating back to t=0 to determine the initial field. Additionally, using this approach it is unnecessary to fit the 60 Hz noise.

The values which represent the magnetic field perpendicular to each detector are then expressed as a vector $B^\perp$. The same forward model used in MSA is used to relate the detected field at a given sensor to the distribution of dipole sources, which can be expressed as a product of a non-linear relationship of the location of the dipole sources and a linear relationship of the moment of the dipole sources, as in Equation 3.15. In each iteration of MSST, the location of each dipole is first determined using a nonlinear minimization, followed by a linear solve for the moment at each determined location. The location is determined using a multistart downhill simplex method to perform a non-linear

minimization of the percent variance explained, or PVE, defined in Equation B.1, where $B_i^\perp$ is the calculated magnetic field perpendicular to the pickup coils and $\tilde{B}_i^\perp$ is the measured magnetic field [108].

$$\text{PVE} = \frac{\sum\limits_{i=1}^{M}(B_i^\perp - \tilde{B}_i^\perp)^2}{\sum\limits_{i=1}^{M}(\tilde{B}_i^\perp)^2} \times 100\% \tag{B.1}$$

Each iteration of the nonlinear minimization results in a matrix $\mathbf{G}^{M\times N}$ in which each element in row $i \in [1, M]$ and column $j \in [1, N]$ represents the portion of the magnetic field detected at sensor i from dipole j at the most recent estimate of dipole $j's$ location. Each element in the vector $\mathbf{P}^{N\times 1}$ is the magnitude of the magnetic moment for dipole j. The vector $\mathbf{P}$ is then determined using the singular vector decomposition of $\mathbf{G}$ (Equation B.2) with the regularized form of $\mathbf{S}$ where the diagonal elements $s_i = 1/(\lambda_i + \kappa\lambda_1)$.

$$\mathbf{G} = \mathbf{USV}^T \tag{B.2}$$

$$\mathbf{P} = \mathbf{VS}^{-1}\mathbf{U}^T\mathbf{B}^\perp \tag{B.3}$$

The reconstruction process is repeated for each value of N, typically $N \in [1, 5]$. The N for which the chi-squared cost function in Equation B.4 is minimized. The chisquared function determines the balance between under-fitting and over-fitting the data because a minimum will be reached when the degrees of freedom ($df = M - 4N$) in the denominator decreases more rapidly than the difference between the model and measured data in the numerator as more dipoles are added to the model.

$$\chi_r^2 = \frac{1}{df}\sum_{i=1}^{M}\frac{(B_i^\perp - \tilde{B}_i^\perp)^2}{\sigma_i^2} \tag{B.4}$$

MSST improves on MSA insofar that it automates the process of solving the least squares optimization with multiple initial conditions and number of dipoles and provides

a quantitative method of determining the true number of dipoles in the data. In return, the computational time required is much larger than that of MSA. The largest time requirement is the manual selection of "bad" pulses, which also is very subjective. After the preprocessing, the computational time is proportional to the product of the number of choices of initial conditions, the number of Monte Carlo iterations, and the maximum number of dipoles to be solved for. The time required to reconstruct a signal from start to finish can quickly become quite substantial because what is essentially the same reconstruction algorithm that is run once with MSA must be repeated for each combination of number of dipoles, choice of initial conditions, and Monte Carlo iteration in MSST.

MSST

MSST-B6 analysis showed an improvement over MSA-B at an overall accuracy of 61.6%, kappa score of 0.47 and a weighted kappa score of 0.56. The no source classification had a specificity of 0.67 and a sensitivity of 0.96. This was the only analysis that misclassified any cases with no source. Of the 20 cases that contained a single source, MSST-B6 misclassified two (10%) as having no source and seven (35%) as having a second source to the right, resulting in a sensitivity of 0.55 and a specificity of 0.8. The sources that were missed both contained the smallest mass of nanoparticles (4 µg of all of the sources tested. The cases that were misclassified as having two sources contained anywhere between 1.6 µg to 50 µg. Of the 34 cases that contained two sources, 22 (65%) were correctly classified. In 6 cases (18%) the second source was missed, and in 6 cases (18%) the second source was found on the wrong side. Sources with 6.25 µg of nanoparticles or more were always found when another source was present. For the classification of the second source to the left the sensitivity was 0.53 and the specificity was 0.97. For the classification of the second source to the right, the sensitivity was 0.76 and the specificity was 0.72. The detailed results for MSST-B6 are shown in Table B.1.

MSST-B6		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**4**	2	0	0
	1	2	**11**	3	3
	L	0	0	**9**	1
	R	0	7	5	**13**

Table B.1: The overall classification results for MSST-B

12.2 MSA with a dewar angle of 144.8

The results for the analysis with MSA using a dewar angle of 144.8 are detailed in Table B.2. MSA correctly classified 18 out of the 60 cases, resulting in an overall accuracy of 30.0%. The kappa score for MSA was 0.04 and the weighted kappa score was 0.12, indicating that the reconstruction classified the cases only slightly better than random chance. In one instance a source was found when there was in fact no source present, and in 10 cases no source was found when at least one source was present. This resulted in a sensitivity of 1 and a specificity of 0.96 for the zero source classification. In the cases in which a single source was present, MSA detected all of the sources that contained at least 1.6 µg of nanoparticles, and 66% of the cases with 0.4 and 0.8 µg. MSA did not misclassify any cases containing a single source as containing two sources, but did misclassify two cases with one source as having no sources. The sensitivity for the one-source classification was 0.9 and the specificity was 0.2. The largest contribution to the low overall accuracy was because most (32 out of 34) of the cases in which two sources were present, only one source was found. This resulted in a sensitivity and specificity of 0.06 and 1, respectively, for both the Right and Left classifications.

MSA-B		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**5**	2	2	2
	1	1	**11**	14	14
	L	0	3	**1**	0
	R	0	0	0	**1**

Table B.2: The overall classification results for MSA with a dewar angle of 144.8°

12.3   Objective versus subjective analysis

This appendix includes a few additional analyses of the blinded study data that may be of interest. First we will look at the effect subjective versus objective analysis to classify the reconstruction. In MRXImage-O6, the reconstructions from MRXImage-B6 were analyzed by a set of objective criteria based on the centroid and total magnitude of clusters of non-zero voxels in the reconstruction. A source was considered valid if its location in x and y was between -3 cm and 3 cm, if its depth was more than 2 cm below the detectors ($z < -2$), and if its strength was greater than 3.3×10^{-3} pJ/T. If there were two valid sources, the second source was designated the one with the more positive y coordinate. Thee second source was determined to be on the left if its x coordinate was negative and on the right if its x coordinate was positive. The results of the objective analysis are presented in Table B.3.

MRXImage-O6		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**6**	1	0	0
	1	0	**17**	2	5
	L	0	0	**13**	0
	R	0	2	2	**12**

Table B.3: The overall classification results for MRXImage-O6

The overall accuracy was the same for the subjective and objective analyses at 80%, but different cases were misclassified between the two methods. The kappa value was 0.72 for both analyses, but the weighted kappa was slightly higher for the objective analysis (0.76) than for the subjective analysis (0.74) indicating that the errors in the objective classification were less severe than in the subjective classification. The average recall was the same for both analyses (0.83), but the average precision was higher for the subjective analysis. Overall, these results indicate that the classification results with MRXImage are not strongly dependent on the user.

12.4 The effect of 4 versus 6 stage positions

In addition to MRXImage-B4, two additional analyses were done with both 4 and 6 stage positions. These include the subjective blinded analyses with MSST and MRXImage (MSST-B6/4 and MRXImage-B6/4) as well as an objective analysis with MRXImage (MRXImage-O6/4). To examine the effect of the number of stage positions on the overall performance, we can compare the analyses based on four stage positions to their six-stage position counterparts. The 4x4 confusion matrices of MSST-B6, MRXImage-B6, and MRXImage-O6 were presented previously, in Tables B.1, 6.6, and B.3, respectively. The 4x4 confusion matrix for MSST-B4, MRXImage-B4, and MRXImage-O4 are presented in Tables B.4, 6.7, and B.5, respectively.

MSST-B4		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**4**	1	0	0
	1	2	**8**	6	3
	L	0	4	**8**	0
	R	0	7	3	**14**

Table B.4: The overall classification results for MSST-B4

MRXImage-O4		True classification			
		0	*1*	*L*	*R*
Reconstructed	*0*	**6**	1	0	0
	1	0	**17**	3	6
	L	0	1	**13**	3
	R	0	1	1	**8**

Table B.5: The overall classification results for MRXImage-O4

Across all performance measures and all algorithms, using only four stage positions reduced the quality of the classification. The overall accuracy was decreased by 5 to 7 percentage points. The weighted kappa value was most heavily impacted, decreasing by 8 to 14 percentage points, indicating that the errors were not only more frequent, but also more severe.

Bibliography

[1] Flynn E, Bryant H. A biomagnetic system for in vivo cancer imaging. *Physics in medicine and biology.* 2005;50(6):1273.

[2] De Haro LP, Karaulanov T, Vreeland EC, Anderson B, Hathaway HJ, Huber DL, Matlashov AN, Nettles CP, Price AD, Monson TC, Flynn ER. Magnetic relaxometry as applied to sensitive cancer detection and localization. *Biomedical Engineering/Biomedizinische Technik.* 2015;60(5):445–455.

[3] Hori SS, Gambhir SS. Mathematical model identifies blood biomarker–based early cancer detection strategies and limitations. *Science translational medicine.* 2011;3(109):109ra116–109ra116.

[4] Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics.* 1993;65(2):413.

[5] Baumgarten D, Liehr M, Wiekhorst F, Steinhoff U, Münster P, Miethe P, Trahms L, Haueisen J. Magnetic nanoparticle imaging by means of minimum norm estimates from remanence measurements. *Medical & biological engineering & computing.* 2008;46(12):1177.

[6] Liebl M, Steinhoff U, Wiekhorst F, Haueisen J, Trahms L. Quantitative imaging of magnetic nanoparticles by magnetorelaxometry with multiple excitation coils. *Physics in medicine and biology.* 2014;59(21):6607–20.

[7] Huang MX, Anderson B, Huang CW, Kunde GJ, Vreeland EC, Huang JW, Matlashov AN, Karaulanov T, Nettles CP, Gomez A, Minser K, Weldon C, Paciotti G, Harsh

M, Lee RR, Flynn ER. Development of advanced signal processing and source imaging methods for superparamagnetic relaxometry. *Physics in medicine and biology.* 2017;62(3):734.

[8] Drescher CW, Hawley S, Thorpe JD, Marticke S, McIntosh M, Gambhir SS, Urban N. Impact of screening test performance and cost on mortality reduction and cost-effectiveness of multimodal ovarian cancer screening. *Cancer prevention research (Philadelphia, Pa).* 2012;5(8):1015–24.

[9] Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JDF, Feuer EJ. Effect of screening and adjuvant therapy on mortality from breast cancer. *New England Journal of Medicine.* 2005;353(17):1784–1792.

[10] Edwards BK, Ward E, Kohler BA, Eheman C, Zauber AG, Anderson RN, Jemal A, Schymura MJ, Lansdorp-Vogelaar I, Seeff LC, van Ballegooijen M, Goede SL, Reis LAG. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer.* 2010;116(3):544–573.

[11] National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011;2011(365):395–409.

[12] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer.* 2015;136(5):E359–E386.

[13] Jasen P. From the "silent killer" to the "whispering disease": ovarian cancer and the uses of metaphor. *Medical History.* 2009;53(4):489–512.

[14] Howlader N, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse S, Kosary C, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis D, Chen H, Feuer E, Cronin K. SEER Cancer Statistics Review, 1975-2014. Bethesda, MD; 2017. posted to the SEER web site.

[15] Henderson JT, Webber EM, Sawaya GF. Screening for ovarian cancer: updated evidence report and systematic review for the US Preventive Services Task Force. *Jama.* 2018;319(6):595–606.

[16] Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, Fogel R, Gelmann EP, Gilbert F, Hasson MA, Hayes RB, Johnson CC, Mandel JS, Oberman A, O'Brien B, Oken MM, Rafla S, Reding DJ, Rutt W, Weissfeld JL, Yokochi LA, Gohagan J. Design of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Controlled clinical trials.* 2000;21(6):273S–309S.

[17] Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, Reding DJ, Greenlee RT, Yokochi LA, Kessel B, Crawford ED, Church TR, Andriole GL, Weissfeld JL, Fouad MN, Chia D, O'Brien B, Ragard LR, Clapp JD, Rathmell JM, Riley TL, Hartge P, Pinsky PF, Zhu CS, Izmirlian G, Kramer BS, Miller AB, Xu JL, Prorok PC, Gohagan J, Berg CD. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening randomized controlled trial. *Jama.* 2011;305(22):2295–2303.

[18] Pinsky PF, Yu K, Kramer BS, Black A, Buys SS, Partridge E, Gohagan J, Berg CD, Prorok PC. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15 years follow-up. *Gynecologic oncology.* 2016;143(2):270–275.

[19] Menon U, Gentry-Maharaj A, Hallett R, Ryan A, Burnell M, Sharma A, Lewis S, Davies S, Philpott S, Lopes A, Godfrey K, Oram D, Herod J, Williamson K, Seif M, Scott I, Mould T, Woolas R, Murdoch J, Dobbs S, Amso NN, Leeson S, Cruickshank

D, Mcguire A, Campbell S, Fallowfield L, Singh N, Dawnay A, Skates SJ, Parmar M, Jacobs IJ. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *The lancet oncology.* 2009;10(4):327–340.

[20] Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, Amso NN, Apostolidou S, Benjamin E, Cruickshank D, Crump DN, Davies S, Dawnay A, Dobbs S, Fletcher G, Ford J, Godfrey K, Gunu R, Habib M, Hallett R, Herod J, Jenkins H, Karpinskyj C, Leeson S, Lewis S, Liston W, Lopes A, Mould T, Murdoch J, Oram D, Rabideau DJ, Reynolds K, Scott I, Serif MW, Sharma A, Singh N, Taylor J, Warburton F, Windschwendter M, Williamson K, Woolas R, Fallowfield L, Mcguire A, Campbell S, Parmar M. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *The Lancet.* 2016;387(10022):945–956.

[21] Evans EC, Matteson KA, Orejuela FJ, Alperin M, Balk EM, El-Nashar S, Gleason JL, Grimes C, Jeppson P, Mathews C, Wheeler TL, Murphy M. Salpingo-oophorectomy at the time of benign hysterectomy: a systematic review. *Obstetrics & Gynecology.* 2016;128(3):476–485.

[22] Lowder JL, Oliphant SS, Ghetti C, Burrows LJ, Meyn LA, Balk J. Prophylactic bilateral oophorectomy or removal of remaining ovary at the time of hysterectomy in the United States, 1979-2004. *American journal of obstetrics and gynecology.* 2010;202(6):538–e1.

[23] Rice MS, Hankinson SE, Tworoger SS. Tubal ligation, hysterectomy, unilateral oophorectomy, and risk of ovarian cancer in the Nurses' Health Studies. *Fertility and sterility.* 2014;102(1):192–198.

[24] National Academies of Sciences Engineering and Medicine. *Ovarian cancers: evolving paradigms in research and care.* Washington, DC: National Academies Press; 2016.

[25] Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *The Lancet.* 2014;384(9951):1376–1388.

[26] Temkin SM, Miller EA, Samimi G, Berg CD, Pinsky P, Minasian L. Outcomes from ovarian cancer screening in the PLCO trial: Histologic heterogeneity impacts detection, overdiagnosis and survival. *European Journal of Cancer.* 2017;87:182–188.

[27] Salvador S, Gilks B, Köbel M, Huntsman D, Rosen B, Miller D. The fallopian tube: primary site of most pelvic high-grade serous carcinomas. *International Journal of Gynecological Cancer.* 2009;19(1):58–64.

[28] Seidman JD, Cho KR, Ronnett BM, Kurman RJ. Surface epithelial tumors of the ovary. In: *Blaustein's pathology of the female genital tract.* Springer; 2011. p. 679–784.

[29] Kroeger Jr PT, Drapkin R. Pathogenesis and heterogeneity of ovarian cancer. *Current opinion in obstetrics & gynecology.* 2017;29(1):26.

[30] Olivier R, Van Beurden M, Lubsen M, Rookus M, Mooij T, Van De Vijver M, Van't Veer L. Clinical outcome of prophylactic oophorectomy in BRCA1/BRCA2 mutation carriers and events during follow-up. *British journal of cancer.* 2004;90(8):1492.

[31] Callahan MJ, Crum CP, Medeiros F, Kindelberger DW, Elvin JA, Garber JE, Feltmate CM, Berkowitz RS, Muto MG. Primary fallopian tube malignancies in BRCA-positive women undergoing surgery for ovarian cancer risk reduction. *Journal of Clinical Oncology.* 2007;25(25):3985–3990.

[32] Brown PO, Palmer C. The preclinical natural history of serous ovarian cancer: defining the target for early detection. *PLoS medicine.* 2009;6(7):e1000114.

[33] Lee JY, Yoon JK, Kim B, Kim S, Kim MA, Lim H, Bang D, Song YS. Tumor evolution and intratumor heterogeneity of an epithelial ovarian cancer investigated using next-generation sequencing. *BMC cancer.* 2015;15(1):85.

[34] Chien J, Neums L, Powell AF, Torres M, Kalli KR, Multinu F, Shridhar V, Mariani A. genetic evidence for early Peritoneal spreading in Pelvic high-grade serous cancer. *Frontiers in oncology.* 2018;8:58.

[35] Lutz AM, Willmann JK, Cochran FV, Ray P, Gambhir SS. Cancer screening: a mathematical model relating secreted blood biomarker levels to tumor sizes. *PLoS medicine.* 2008;5(8):e170.

[36] Botesteanu DA, Lee JM, Levy D. Modeling the dynamics of high-grade serous ovarian cancer progression for transvaginal ultrasound-based screening and early detection. *PloS one.* 2016;11(6):e0156661.

[37] Mathieu KB, Bedi DG, Thrower SL, Qayyum A, Bast Jr R. Screening for ovarian cancer: imaging challenges and opportunities for improvement. *Ultrasound in Obstetrics & Gynecology.* 2018;51(3):293–303.

[38] Yasumori I, Reinen D, Selwood PW. Anisotropic behavior in superparamagnetic systems. *Journal of Applied Physics.* 1963;34(12):3544–3549.

[39] Chantrell R, Hoon S, Tanner B. Time-dependent magnetization in fine-particle ferromagnetic systems. *Journal of magnetism and magnetic materials.* 1983;38(2):133–141.

[40] Berkov DV, Kötitz R. Irreversible relaxation behaviour of a general class of magnetic systems. *Journal of Physics: Condensed Matter.* 1996;8(9):1257–1266.

[41] Kötitz R, Weitschies W, Trahms L, Semmler W. Investigation of Brownian and Neel relaxation in magnetic fluids. *Journal of magnetism and magnetic materials.* 1999;201(1-3):102–104.

[42] Weitschies W. Determination of relaxing or remnant nanoparticle magnetization provides a novel binding specific technique for the evaluation of immunoassays. *Pharm Pharmacol Lett.* 1997;7:5–8.

[43] Kotitz R, Matz H, Trahms L, Koch H, Weitschies W, Rheinltinder T, Semmler W, Bunte T. SQUID based remanence measurements for immunoassays. *IEEE Transactions on Applied Superconductivity.* 1997;7(2):3678–3681.

[44] Chemla YR, Grossman HL, Poon Y, McDermott R, Stevens R, Alper MD, Clarke J. Ultrasensitive magnetic biosensor for homogeneous immunoassay. *Proceedings of the National Academy of Sciences of the United States of America.* 2000;97(26):14268–72.

[45] Eberbeck D, Wiekhorst F, Steinhoff U, Trahms L. Aggregation behaviour of magnetic nanoparticle suspensions investigated by magnetorelaxometry. *Journal of Physics: Condensed Matter.* 2006;18(38):S2829–S2846.

[46] Hilger I, Kießling A, Romanus E, Hiergeist R, Hergt R, Andrä W, Roskos M, Linss W, Weber P, Weitschies W, Kaiser WA. Magnetic nanoparticles for selective heating of magnetically labelled cells in culture: preliminary investigation. *Nanotechnology.* 2004;15(8):1027–1032.

[47] Schambach J, Warzemann L, Weber P. SQUID gradiometer measurement system for magnetorelaxometry in a disturbed environment. *IEEE Transactions on Applied Superconductivity.* 1999;9(2 PART 3):3527–3530.

[48] Warzemann L, Schambach J, Weber P, Weitschies W, Kötitz R. LTS SQUID gradiometer system for in vivo magnetorelaxometry. *Superconductor Science and Technology.* 1999;12(11):953–955.

[49] Romanus E, Hückel M, Groß C, Prass S, Weitschies W, Bräuer R, Weber P. Magnetic nanoparticle relaxation measurement as a novel tool for in vivo diagnostics. *Journal of Magnetism and Magnetic Materials.* 2002;252(1-3 SPEC. ISS.):387–389.

191

[50] Eberbeck D, Hartwig S, Steinhoff U, Trahms L. Description of the magnetization decay in ferrofluids with a narrow particle size distribution. *Magnetohydrodynamics.* 2003;39(1):77–83.

[51] Flynn ER, Bryant HC, Bergemann C, Larson RS, Lovato D, Sergatskov Da. Use of a SQUID array to detect T-cells with magnetic nanoparticles in determining transplant rejection. *Journal of Magnetism and Magnetic Materials.* 2007;311(1 SPEC. ISS.):429–435.

[52] Adolphi NL, Huber DL, Jaetao JE, Bryant HC, Lovato DM, Fegan DL, Venturini EL, Monson TC, Tessier TE, Hathaway HJ, Bergemann C, Larson RS, Flynn ER. Characterization of magnetite nanoparticles for SQUID-relaxometry and magnetic needle biopsy. *Journal of Magnetism and Magnetic Materials.* 2009;321(10):1459–1464.

[53] Adolphi NL, Butler KS, Lovato DM, Tessier TE, Trujillo JE, Hathaway HJ, Fegan DL, Monson TC, Stevens TE, Huber DL, Ramu J, Milne ML, Altobelli Sa, Bryant HC, Larson RS, Flynn ER. Imaging of Her2-targeted magnetic nanoparticles for breast cancer detection: Comparison of SQUID-detected magnetic relaxometry and MRI. *Contrast Media and Molecular Imaging.* 2012;7(3):308–319.

[54] Johnson C, Adolphi NL, Butler KL, Lovato DM, Larson R, Schwindt PDD, Flynn ER. Magnetic relaxometry with an atomic magnetometer and SQUID sensors on targeted cancer cells. *Journal of Magnetism and Magnetic Materials.* 2012;324(17):2613–2619.

[55] Eberbeck D, Bergemann C, Wiekhorst F, Steinhoff U, Trahms L. Quantification of specific bindings of biomolecules by magnetorelaxometry. *Journal of nanobiotechnology.* 2008;6:4.

[56] Wiekhorst F, Baumgarten D, Haberkorn W, Steinhoff U, Haueisen J, Bär M, Trahms L. A physical phantom modeling extended magnetic nanoparticle distributions in bio-

logical systems. In: *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany.* Springer; 2009. p. 293–296.

[57] Baumgarten D, Liehr M, Haueisen J. Imaging of magnetic nanoparticle loaded rat organs using a 16 channel micro SQUID system. In: *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany.* Springer; 2009. p. 366–369.

[58] Baumgarten D, Haueisen J. A Spatio-Temporal Approach for the Solution of the Inverse Problem in the Reconstruction of Magnetic Nanoparticle Distributions. *IEEE Transactions on Magnetics.* 2010 aug;46(8):3496–3499.

[59] Sarangi S, Tan IC, Brazdeikis A. Magnetic imaging method based on magnetic relaxation of magnetic nanoparticles. *Journal of Applied Physics.* 2009;105(9):1–6.

[60] Crevecoeur G, Baumgarten D, Steinhoff U, Haueisen J, Trahms L, Dupré L. Advancements in magnetic nanoparticle reconstruction using sequential activation of excitation coil arrays using magnetorelaxometry. *IEEE Transactions on Magnetics.* 2012;48(4):1313–1316.

[61] Liebl M, Steinhoff U, Wiekhorst F, Coene A, Haueisen J, Trahms L. Quantitative reconstruction of a magnetic nanoparticle distribution using a non-negativity constraint. *Biomedical Engineering / Biomedizinische Technik.* 2013 jan;58(1):24–25.

[62] Coene A, Crevecoeur G, Liebl M, Wiekhorst F, Dupré L, Steinhoff U. Uncertainty of reconstructions of spatially distributed magnetic nanoparticles under realistic noise conditions. *Journal of Applied Physics.* 2014;115(17).

[63] Hoo NK, Klemm M, Supriyanto E, Baumgarten D. Effects of excitation coil configurations in magnetorelaxometry imaging of magnetic nanoparticles. *IECBES 2014, Conference Proceedings - 2014 IEEE Conference on Biomedical Engineering and Sci-*

ences: "Miri, Where Engineering in Medicine and Biology and Humanity Meet". 2015;(December):715–718.

[64] Liebl M, Wiekhorst F, Eberbeck D, Radon P, Gutkelch D, Baumgarten D, Steinhoff U, Trahms L. Magnetorelaxometry procedures for quantitative imaging and characterization of magnetic nanoparticles in biomedical applications. *Biomedical Engineering / Biomedizinische Technik.* 2015;60(5):427–443.

[65] Gleich B, Weizenecker J. Tomographic imaging using the nonlinear response of magnetic particles. *Nature.* 2005;435(7046):1214.

[66] Knappe S, Sander TH, Kosch O, Wiekhorst F, Kitching J, Trahms L. Cross-validation of microfabricated atomic magnetometers with superconducting quantum interference devices for biomagnetic applications. *Applied Physics Letters.* 2010;97(13):3–5.

[67] Ludwig F, Heim E, M??uselein S, Eberbeck D, Schilling M. Magnetorelaxometry of magnetic nanoparticles with fluxgate magnetometers for the analysis of biological targets. *Journal of Magnetism and Magnetic Materials.* 2005;293(1):690–695.

[68] Ludwig F, Heim E, Eberbeck D, Schwarz K, Trahms L, Schilling M. Comparison and calibration of fluxgate and squid magnetorelaxometry techniques for the characterization of magnetic core-shell nanoparticles. *IEEE Transactions on Magnetics.* 2009;45(10):4857–4860.

[69] Brown Jr WF. Thermal fluctuations of a single-domain particle. *Physical Review.* 1963;130(5):1677.

[70] Néel L. Théorie du traînage magnétique des ferromagnétiques en grains fins avec applications aux terres cuites. *Ann géophys.* 1949;5:99–136.

[71] Hornyak GL, Tibbals HF, Dutta J, Moore JJ. *Introduction to nanoscience and nanotechnology.* CRC press; 2008.

[72] Eberbeck D, Wiekhorst F, Steinhoff U, Trahms L. Aggregation behaviour of magnetic nanoparticle suspensions investigated by magnetorelaxometry. *Journal of Physics: Condensed Matter.* 2006;18(38):S2829.

[73] Bedanta S. Supermagnetism in magnetic nanoparticle systems. Duisburg: Universität Duisburg-Essen, Fakultät für Physik; 2006.

[74] Wiekhorst F, Steinhoff U, Eberbeck D, Trahms L. Magnetorelaxometry assisting biomedical applications of magnetic nanoparticles. *Pharmaceutical research.* 2012;29(5):1189–1202.

[75] Levenberg K. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics.* 1944;2(2):164–168.

[76] Stefan W, Mathieu K, Thrower S, Fuentes D, Kaffes C, Sovizi J, Hazle J. Automated algorithms for improved pre-processing of magnetic relaxometry data. In: *Medical Imaging 2018: Physics of Medical Imaging.* vol. 10573. International Society for Optics and Photonics; 2018. p. 105733R.

[77] Prêle D, Piat M, Sipile L, Voisin F. Operating Point and Flux Jumps of a SQUID in Flux-Locked Loop. *IEEE Transactions on Applied Superconductivity.* 2016;26(2):1–5.

[78] Carrillo RE, McEwen J, Wiaux Y. Sparsity Averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging. *Monthly Notices of the Royal Astronomical Society.* 2012;426(2):1223–1234.

[79] Donoho DL. For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution. *Communications on pure and applied mathematics.* 2006;59(6):797–829.

[80] Grant M, Boyd S. CVX: Matlab Software for Disciplined Convex Programming, version 2.1; 2014. http://cvxr.com/cvx.

[81] Grant M, Boyd S. Graph implementations for nonsmooth convex programs. In: Blondel V, Boyd S, Kimura H, editors. *Recent Advances in Learning and Control.* Lecture Notes in Control and Information Sciences. Springer-Verlag Limited; 2008. p. 95–110. http://stanford.edu/~boyd/graph_dcp.html.

[82] Grant M, Boyd S, Ye Y. Disciplined convex programming. In: *Global optimization.* Springer; 2006. p. 155–210.

[83] Sturm JF. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software.* 1999;11(1-4):625–653.

[84] Toh KC, Todd MJ, Tütüncü RH. SDPT3—a MATLAB software package for semidefinite programming, version 1.3. *Optimization methods and software.* 1999;11(1-4):545–581.

[85] Mosek A. The MOSEK optimization software. *Online at http://www mosek com.* 2010;54(2-1):5.

[86] Gurobi Optimization L. Gurobi Optimizer Reference Manual; 2018.

[87] Boyd S, Vandenberghe L. *Convex optimization.* Cambridge university press; 2004.

[88] Lobo MS, Vandenberghe L, Boyd S, Lebret H. Applications of second-order cone programming. *Linear algebra and its applications.* 1998;284(1-3):193–228.

[89] Cai Z, Toh KC. Solving second order cone programming via a reduced augmented system approach. *SIAM Journal on Optimization.* 2006;17(3):711–737.

[90] Gorodnitsky IF, George JS, Rao BD. Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm. *Electroencephalography and clinical Neurophysiology.* 1995;95(4):231–251.

[91] Fagaly R. Superconducting quantum interference device instruments and applications. *Review of scientific instruments.* 2006;77(10).

[92] Beckers B, Beckers P. A general rule for disk and hemisphere partition into equal-area cells. *Computational Geometry.* 2012;45(7):275–283.

[93] Gibbons JD, Chakraborti S. Nonparametric statistical inference. In: *International encyclopedia of statistical science.* Springer; 2011. p. 977–979.

[94] Candes EJ, Romberg JK, Tao T. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics.* 2006;59(8):1207–1223.

[95] Eichardt R, Baumgarten D, Petković B, Wiekhorst F, Trahms L, Haueisen J. Adapting source grid parameters to improve the condition of the magnetostatic linear inverse problem of estimating nanoparticle distributions. *Medical & biological engineering & computing.* 2012;50(10):1081–1089.

[96] Baumgarten D, Braune F, Supriyanto E, Haueisen J. Plane-wise sensitivity based inhomogeneous excitation fields for magnetorelaxometry imaging of magnetic nanoparticles. *Journal of Magnetism and Magnetic Materials.* 2015;380:255–260.

[97] Candès EJ. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique.* 2008;346(9-10):589–592.

[98] Nguyen TL, Shin Y. Deterministic sensing matrices in compressive sensing: a survey. *The Scientific World Journal.* 2013;2013.

[99] Bandeira AS, Dobriban E, Mixon DG, Sawin WF. Certifying the restricted isometry property is hard. *IEEE transactions on information theory.* 2013;59(6):3448–3450.

[100] Tillmann AM, Pfetsch ME. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory.* 2014;60(2):1248–1259.

[101] Donoho DL, Huo X. Uncertainty principles and ideal atomic decomposition. *IEEE transactions on information theory.* 2001;47(7):2845–2862.

[102] Fuchs JJ. On sparse representations in arbitrary redundant bases. *IEEE transactions on Information theory.* 2004;50(6):1341–1344.

[103] Obermeier R, Martinez-Lorenzo JA. Sensing matrix design via mutual coherence minimization for electromagnetic compressive imaging applications. *IEEE Transactions on Computational Imaging.* 2017;3(2):217–229.

[104] Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Information Processing & Management.* 2009;45(4):427–437.

[105] Cohen J. A coefficient of agreement for nominal scales. *Educational and psychological measurement.* 1960;20(1):37–46.

[106] Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin.* 1968;70(4):213.

[107] Rahmer J, Halkola A, Gleich B, Schmale I, Borgert J. First experimental evidence of the feasibility of multi-color magnetic particle imaging. *Physics in Medicine & Biology.* 2015;60(5):1775.

[108] Huang MX, Aine C, Davis L, Butman J, Christner R, Weisend M, Stephen J, Meyer J, Silveri J, Herman M, Lee RR. Sources on the anterior and posterior banks of the central sulcus identified from magnetic somatosensory evoked responses using Multi-Start Spatio-Temporal localization. *Human Brain Mapping.* 2000;11(2):59–76.

Vita

Sara Lynn Loupot was born to Judson Emile Loupot and Phyllis McGrath Loupot in College Station, Texas. Her younger sister is Darcy Annell Loupot. She received her primary and secondary education in the College Station Independent School District in Texas, and graduated Summa Cum Laude from A&M Consolidated High School in 2009. She then matriculated in Texas A&M University to pursue her Bachelor of Science degree. In 2013 she received her BS in Nuclear Engineering and began a Doctor of Philosophy degree in Medical Physics at The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences. In 2015, Sara married Scott Thrower.

Permanent address as of December 2018:

13826 Oakwood Ln

Sugar Land, Texas, USA 77498