Magnetic Field Mapping Using Spatial Derivative Measurements with Gaussian Process Regression

J.M. Beurskens

Master of Science Thesis

TU Delft Delft University of Technology

Delft Center for Systems and Control

Magnetic Field Mapping Using Spatial Derivative Measurements with Gaussian Process Regression

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

J.M. Beurskens

January 9, 2025

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Abstract

Indoor localization is an active research area since traditional localization methods, such as global navigation satellite systems (GNSS), are ineffective in indoor environments. One promising solution is to make use of anomalies in the magnetic field caused by ferromagnetic materials within buildings. By constructing a magnetic field map in an indoor location, the spatial variability in the magnetic field can be utilized to aid in inferring the location. Gaussian process regression (GPR) is frequently used to model the magnetic field and provide magnetic field predictions as well as the corresponding predictive uncertainty. In addition to mapping an entire building or room, the existing literature has also been studying mapping in close proximity to the sensor, as this provides valuable information for odometry purposes. Magnetic field measurements are obtained using sensors known as magnetometers. Magnetometer arrays, which are typically two- or three-dimensional fixed structures that contain multiple magnetometers, are used to construct these local maps as they are capable of measuring the magnetic field at multiple locations simultaneously. However, the time complexity of GPR scales cubically with the number of data points. When these magnetometer arrays are used, the size of the data increases rapidly. As a result, GPR can quickly become computationally intractable in combination with arrays. In this thesis, building on prior work that demonstrated the use of spatial derivatives of arrays for odometry purposes, we instead analyze how effective the information on the array can be approximated using the spatial derivative of the magnetic field to alleviate some of the computational burden. We analyze this by examining the performance of using the spatial derivative computed from different array configurations in terms of the number of magnetometers, their spacing, and noise levels, and comparing it with using all magnetometer measurements on the array. Through simulations, we show that the spacing is a key factor in determining the spatial derivative, and that more magnetometers and less noise result in better estimates. It is shown that the spatial derivative is not an effective approximation to the full information on the array. At least for the arrays considered in this thesis, it did not reduce the computational burden sufficiently to justify the resulting loss in map quality. In addition to examining which array configurations perform best with spatial derivatives, we also evaluate which configurations yield accurate maps for existing odometry methods that rely on arrays. We show that the array configuration plays an important role and requires deliberate selection for optimal performance.

Table of Contents

Preface

This master's thesis is the last part of my Master of Science graduation project within the Delft Center for Systems and Control (DCSC) department at Delft University of Technology. This thesis explores the potential of approximating the information on magnetometer arrays with the spatial derivative and analyzes which array configurations produce the highest quality magnetic field maps.

The idea of doing my thesis on this subject came after some suggestions from my supervisor Thomas Edridge. Our discussion really helped shape the project's direction. The topic of using spatial derivatives to improve the magnetic field map using Gaussian processes aligned with my interests and the objectives of the research group.

I would like to thank my supervisors dr. Manon Kok and ir. Thomas Edridge from the DCSC department for their valuable feedback, guidance and support.

Chapter 1

Introduction

Satellite navigation systems, such as global navigation satellite systems (GNSS) [22], are widely used and valued for their affordability and ease of use in outdoor localization. However, they struggle to provide accurate and reliable locations in indoor and underground environments [14]. GPS signals are significantly attenuated as they pass through obstacles, rendering them ineffective in such environments. Indoor localization offers considerable benefits for a large number of applications, for example, navigation for rescue/emergency teams and navigation in large buildings such as airports and shopping centers [24]. Numerous indoor localization techniques have been studied, including infrared-based methods [1], Bluetooth-based methods [25], radio frequency identification (RFID) [36], and optical-based methods [41]. However, these methods typically require additional infrastructure in order to function. Besides the deployment costs, most of these techniques are also dependent on not having too much interference between the sender and receiver, i.e., to some extent they should have a clear line of sight. Examples that do not require expensive extra equipment are Wi-Fi [20] and mobile camera-based techniques [17]. Wi-Fi is already available in virtually all buildings and can be measured using mobile devices, such as a mobile phone. However, the accuracy and reliability of the Wi-Fi-based method can deteriorate if there is no clear line of sight between the sender and receiver. Furthermore, the quality and temporal variability of the Wi-Fi also influences the accuracy of the method. Camera-based methods only require a mobile camera, such as a mobile phone. However, it has the downside that it forces the user to ensure that their phone is recording their surroundings. Despite the variety of indoor localization techniques available, they still suffer from limitations related to additional infrastructure or line-of-sight requirements. Consequently, indoor localization remains an area of active research [14, 28].

An alternative indoor localization approach that makes use anomalies in the magnetic field has been explored in recent years [15, 32]. These anomalies give rise to spatial variations in the magnetic field due to the presence of ferromagnetic materials in building structures and furniture [32]. These variations provide the magnetic field with information about the location. By constructing a global map, e.g. of a building, that links locations to the magnetic field at any desired position, predictions of the magnetic field can be made and compared to

the measurements to obtain location-specific information. It is called a global map because it maps the whole area of interest.

A major benefit of the magnetic field-based localization approach is that, unlike many of the previously mentioned methods, no expensive additional equipment is required. The magnetic field can be measured by magnetometers which are small, affordable sensors already integrated into most smartphones nowadays. However, creating a magnetic field map does require access to accurate locations where the corresponding measurements were taken, for which expensive infrastructure would be necessary, undermining the method's advantage. To address this, simultaneous localization and mapping (SLAM) can be employed, allowing the map to be constructed dynamically while navigating the area drift-free, where the location and map are jointly estimated [19, 37]. Neither does this approach suffer from line-of-sight requirements, as the magnetic field penetrates obstacles such as walls and furniture.

The magnetic field can be utilized not only for obtaining absolute location estimates through global magnetic field maps, but also for obtaining changes in location and orientation, called magnetic field odometry. From here on, position and orientation are together called *pose*. Unlike global mapping, magnetic field odometry relies on local maps instead of global maps. In order to obtain these local maps, magnetometer arrays are used. A magnetometer array is a fixed structure, typically a two- or three-dimensional, composed of multiple magnetometers, which allows for measuring the magnetic field at multiple locations simultaneously. This enables the creation of local maps that specify how the magnetic field changes locally. This local information on how the magnetic field changes can be utilized to obtain estimates of the pose changes. This process is very similar to visual odometry systems, as highlighted in [30]. Similarly to how a camera can capture an image of the surrounding environment, an array of magnetometers produces an image-like measurement of the magnetic field. By comparing two consecutive 'images', pose changes can be estimated based on the observed shift.

To create a magnetic field map, whether for global or local maps, and be able to make magnetic field predictions at different locations, an interpolation method has to be used. Such an interpolation method enables the estimation of the magnetic field at locations where no direct measurements were taken. The interpolation method provides a continuous representation of the magnetic field based on finite measurements. Two different techniques are commonly used for this purpose, both capable of incorporating physical knowledge of how the magnetic field evolves. Typically either Gaussian process regression (GPR) [19, 32, 35, 37] is used or a polynomial model [29, 30]. Both methods can account for the curl- and divergence-free properties of the magnetic field, as formulated by Maxwell in 1865 [23].

In this thesis, the magnetic field is modeled using Gaussian processes (GPs), a supervised machine learning method [27]. GPs can capture highly complex and nonlinear functions, even with limited data, and have been shown to produce accurate maps in [19, 32, 35, 37]. As GPs provide a probabilistic framework for their regression tasks, they provide a full probability distribution of their predictions. Hence, the actual predicted value is obtained as well as the associated probabilistic uncertainty. Since localization methods combine data from multiple sensors using probabilistic filtering techniques such as extended Kalman filters (EKFs) and particles filters, the predictive uncertainty allows one to weight the influence of the magnetic field prediction accordingly. However, GPs have the downside that the computational complexity scales cubically with the number of data points.

One particular method that uses these arrays for magnetic field odometry, which will be

referred to as the derivative-based odometry method, was first introduced by Vissière et al. in 2007 [39]. By using the spatial derivative of the magnetic field, derived from the array of magnetometers, the velocity can be estimated. When combined with accelerometers and gyroscopes, this approach forms a magnetic-field-aided inertial navigation system (INS), significantly reducing the error growth rate compared to an INS without magnetic odometry [2, 3, 7, 42]. A more recent method, inspired by this approach that also relies on the spatial information provided by the array, compares two consecutive measurements from the array to infer the change in pose [16, 29, 30]. However, unlike the derivative-based method, this method utilizes all spatial information from the array rather than reducing it to the spatial derivative. First, a local magnetic field model is learned from all measurements on the array at a certain time instant, then at the next time instant the new measurements are compared to the model created at the previous time step to infer the change in pose. Unlike a global map, this local map only models the magnetic field based on the measurements of the array at a single time instant. From now on, this approach will be referred to as the model-based odometry approach. These two magnetic odometry methods illustrate that the spatial information captured by the array provides useful information for odometry and localization purposes in general.

A promising class of techniques that is currently being studied, involves incorporating these magnetic odometry methods into SLAM. Note that the focus of this thesis is not on SLAM, but the results are useful for it. SLAM relies not only on location estimates from the map but also heavily on odometry data, typically from accelerometers and gyroscopes, to track changes in pose. For this technique to be effective, so-called *loop closures* are essential, which involve revisiting previously mapped regions to correct pose estimates that have drifted away over time. However, for the system to work properly, accurate odometry information is needed to ensure reliable loop closures and increase the path length between these closures [16]. Including magnetic odometry methods into SLAM has the potential to improve odometry precision, leading to more reliable loop closures and improved overall performance. A magnetic-field-aided INS based on the model-based odometry method was tested using real-world data in [16] and has the potential to be implemented into SLAM. An alternative technique currently under active investigation does not explicitly implement a magnetic odometry method in SLAM, but uses all magnetometers on the array to create the global map. The odometric information could potentially be inherently captured this way in the global map, eliminating the need for a separate odometry method.

In summary, there are two methods that make use of magnetic-field maps for indoor localization. The first is global localization, which relies on a global map and provides location estimates in the area of interest. The second is magnetic odometry, which uses a local map to estimate the pose change. In contrast to the global localization, the magnetic odometry methods must use magnetometer arrays to be useful. Although the global localization method does not necessarily need an array, it would likely benefit from one when combined with SLAM because of its odometric properties. Since both of these two methods can make use of arrays to construct maps, this thesis analyzes how arrays can be used effectively and efficiently to create magnetic field maps. Specific attention will be paid to analyzing how accurately information from the array can be approximated using a single measurement in combination with the spatial derivative. Using the spatial derivative reduces the number of data points compared to using all magnetometers on the array. Since GPs scale cubically with the number of data points, the algorithm runs into computational problems very

quickly. Reducing the number of data points by using the spatial derivative could alleviate some of the computational burden. The idea of approximating the array measurements with the spatial derivative arises from the derivative-based method, as they already approximate the array measurement with the spatial derivative. Approximating the information on the array is particularly relevant when the array is used to create a global map, as the number of data points increases quickly with the size of the array. This is, for instance, the case where the array is incorporated into SLAM to create global maps. The following research question is considered in this thesis:

> *How can magnetic field measurements from a magnetometer array be utilized effectively and efficiently to create magnetic field maps with Gaussian process regression?*

. To address this question, three sub-questions will be examined:

- How does the array configuration influence the accuracy with which the measurements of the magnetometer array can be approximated using a single measurement combined with the spatial derivative?
- How do different methods of computing the spatial derivative compare in terms of accuracy and computational efficiency?
- How does using a single measurement combined with the spatial derivative affect the accuracy and computational efficiency of magnetic field mapping compared to using all measurements on the array?
  - What is the outcome of this comparison for global maps?
  - What is the outcome of this comparison for local maps?

The first sub-question aims to find the reliability with which the magnetometer array measurements can be approximated using the spatial derivative from a two-dimensional array for different array configuration. Several variables will be analyzed: the number of magnetometers on the array, the spacing between the magnetometers and the noise level present in the measurements. The goal is to find how these variables interact to produce accurate derivatives and to identify the effect of the individual variables. This is particularly directly useful for the derivative-based odometry method, as it directly computes and uses the derivative. Hence, using an array that can accurately and reliably compute the spatial derivative is crucial. However, having access to the variance of the spatial derivative estimate is also essential to use them as measurements to GPs. This allows the derivatives to be used for mapping purposes as well. The second sub-question will investigate three distinct techniques for computing the spatial derivative. Their properties will be compared across the different array configurations. Both the accuracy of the methods as well as the time required to compute the derivative will be considered. The goal of the final sub-question is to examine whether there is a difference in map quality between the local odometry map and the global map for either using the spatial derivative or all measurements. The aim of this question is to examine whether the spatial derivative is relatively more suitable for odometry applications or for global mappings.

This thesis consists of four additional chapters. Chapter 2 provides the necessary background on GPR, magnetic field mapping and modeling, and magnetometer arrays. It establishes the foundation for the subsequent research. Chapter 3 covers the methods and experimental setups used in this research. All results are based on simulations and no real-world experiments are conducted. The chapter begins by introducing the three methods for computing the spatial derivative from the array. Subsequently, it explains the approach for including spatial derivative measurements into GPR for magnetic field mapping. Finally, the simulation setups for the different experiments are described. This includes the setup for evaluating the variance of the spatial derivative estimates for the different array configurations, as well as the setups for assessing the map quality in terms of accuracy and computational efficiency. The latter will be done for both the global and local odometry maps. In Chapter 4, the results of the experiments that are needed to answer the research questions are presented and illustrated with figures. Among other things, it shows which array configurations provide the most accurate spatial derivative and how well the spatial derivative performs at creating local and global maps. Finally, Chapter 5 discusses the results of Chapter 4 and draws the necessary conclusions to answer the research questions. In addition, suggestions are also given for future work.

Chapter 2

Background

This chapter provides the background information essential to understand the methods and results discussed in this thesis. Furthermore, the two odometry methods and simultaneous localization and mapping (SLAM) are explained in slightly more detail to highlight the potential contributions of this thesis for the different applications. The chapter begins with a general introduction to Gaussian process regression (GPR) in Section 2-1. Section 2-2 gives a short introduction to magnetic fields and how they can be measured. The mapping and modeling of the magnetic field using GPR is subsequently presented in Section 2-3. Together, these first three sections make clear how magnetic field maps can be constructed using GPR. Subsequently, we shift the focus towards using magnetometer arrays and the spatial derivative. Section 2-4 shows the properties of the spatial derivative inherited by the curl- and divergence-free properties. Finally, Section 2-5 introduces magnetometer arrays and magnetic odometry methods and examines how the results of this thesis could contribute to those odometry methods.

2-1 Gaussian process regression

GPR is a supervised machine learning method [27]. Machine learning methods are used to try to find a latent function that describes unknown behavior or phenomena. To achieve this, data sets with input-output pairs are required as examples. In contrast to many machine learning methods that start by assuming that the underlying function can be captured by a to-be-determined model with a certain amount of parameters, Gaussian processes (GPs) instead are non-parametric. An advantage of GPs is that they allow incorporating prior knowledge about the form of the latent function, such as smoothness, to constrain possible realizations. Even with small amounts of data, highly complex and nonlinear functions can be captured rather accurately with GPs. As GPs provide a probabilistic framework for their regression tasks, they provide a full probability distribution of their predictions. Hence, the actual predicted value is obtained as well as the associated probabilistic uncertainty. However, GPs have the downside that the computational complexity increases rapidly as the data set

grows. The time complexity scales cubically with respect to the number of data points. This poses challenges in scalability. The main goal of this thesis is to determine the extent to which the spatial derivative can approximate the full information from the magnetometer array, thereby reducing the computational cost.

GPs are stochastic processes for which the input space is divided into a range of discrete input values. Each input point in this range is considered a separate random variable. This implies that a Gaussian process is completely defined by its mean and its covariance function. The mean and covariance functions are defined as

$$\begin{aligned} m(\mathbf{x}) &= \mathbb{E}[f(\mathbf{x})] \\ k(\mathbf{x}, \mathbf{x}') &= \mathbb{E}\big[\big(f(\mathbf{x}) - m(\mathbf{x})\big)\big(f(\mathbf{x}') - m(\mathbf{x}')\big)\big], \end{aligned} \tag{2-1}$$

where $\mathbf{x}$ is the input and $f(\mathbf{x})$ the output of the latent function. and a Gaussian process as

$$f(\mathbf{x}) \ \sim \ \mathcal{GP}\big(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\big). \tag{2-2}$$

The same notation as used in Rasmussen's book on Gaussian processes [27] is used here.

The covariance function of a multivariate Gaussian distribution specifies the correlation between the input random variables. A high covariance between two random variables means that they are strongly correlated, and as a result, the two corresponding outputs will have very similar values. In the case of GPs, the covariance function is designed by the engineer to incorporate prior knowledge about the system and is called the kernel function. The kernel function is used to measure the similarity between inputs. It should be able to evaluate any possible input pair, ensuring that GPs can give distributions for all input locations. For this reason, GPs can be interpreted to produce distributions over functions. The number of input points can be made arbitrarily large and, as the limit goes to infinity, a function is obtained.

2-1-1 Prior Distribution

The kernel function has to be chosen manually and contains prior knowledge of how the inputs are correlated with each other. This puts constraints on the possible function realizations that one can obtain from a GP. The kernel must therefore be chosen with care. One typical choice for a kernel is the squared exponential (SE) kernel which is given by

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(- \frac{||\mathbf{x} - \mathbf{x}'||^2}{2l^2} \right). \tag{2-3}$$

This kernel has a high correlation between points close to each other which decreases exponentially as the distance increases. By considering a certain number of test input points, X_*, and a kernel matrix, $K(\mathbf{x}, \mathbf{x}')$, the possible output points can be inferred from the multivariate normal distribution. Note that no observations are included yet and therefore the distribution is called the prior distribution. The possible values, $\mathbf{f}_*$ are given by the normal distribution

$$\mathbf{f}_* \ \sim \ \mathcal{N}\big(\mathbf{0}, K(X_*, X_*)\big). \tag{2-4}$$

Note that the mean of this distribution is set to zero. It is common to set it to zero, but not necessary. It is possible to use a fixed function or even a fixed basis function as the mean. For more details, see Section 2.7 of [27].

The plot on the left of Figure 2-1 shows a typical example of a prior distribution with an SE kernel. The kernel function should be chosen so that potential samples from the prior distribution represent what the function outputs are expected to look like.

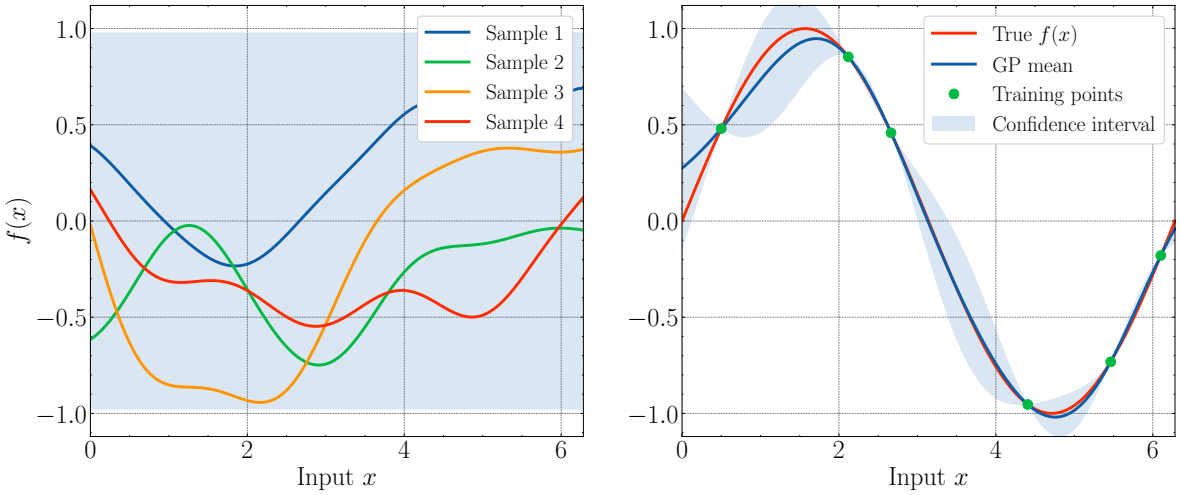

Figure 2-1: Left: A plot showing four samples drawn from a typical prior distribution with a SE kernel. The shaded region represents the confidence interval with twice the standard deviation. Right: A plot showing the mean of the posterior distribution in dark blue after three data points have been observed. The shaded region represents the confidence interval with twice the standard deviation. The underlying function is displayed by the red line.

2-1-2 Posterior Distribution

The goal of a Gaussian process is to learn the underlying probability distribution that describes the data. If one has training data consisting of training inputs, $\mathbf{X}$, and training outputs, $\mathbf{f}$, the joint distribution of the training and test outputs is given by

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X,X) & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\right). \tag{2-5}$$

To arrive at the distribution of the test output points, the joint distribution must be conditioned on the training data. The conditional distribution of a multivariate normal distribution can be found by explicitly working out the law of conditional probability in PDF form

$$p(\mathbf{f}_*|\mathbf{f}) = \frac{p(\mathbf{f}_*, \mathbf{f})}{p(\mathbf{f})}. \tag{2-6}$$

This gives rise to the following normal distribution

$$\begin{aligned} \mathbf{f}_*|X_*, X, \mathbf{f} \sim \mathcal{N}(&K(X_*,X)K(X,X)^{-1}\mathbf{f}, \\ &K(X_*,X_*) - K(X_*,X)K(X,X)^{-1}K(X,X_*)) \end{aligned} \tag{2-7}$$

and is called the posterior distribution. Note that $\mathbf{f}_*$ is also conditioned on X_* and X since these input locations are required to evaluate the covariance matrices. Naturally, those influence the distributions. Also note that this posterior distribution is valid if the mean is set

to zero. If the mean is set to some function, the equation will look slightly different. The predictions of the test output points are given by the mean of the distribution of Eq. (2-7) and the uncertainty is given by the variances on the diagonal of the covariance matrix. The plot on the right in Figure 2-1 shows the mean and confidence interval of the posterior distribution after three observations. As the input moves further away from the training points, the uncertainty increases quickly.

Noisy Observations

Unlike what was omitted in the previous sections, the training outputs very often suffer to some degree from noise. The actual function values, $\mathbf{f}$, are unknown and cannot be measured directly. Instead, a measurement corrupted with noise is available: $\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\epsilon}$. It is usually assumed that this noise is zero mean white noise and has an indepentend and identically distributed (IID) Gaussian distribution with a variance of $\sigma_n^2 I$. This model can be written as

$$\begin{aligned} f(\mathbf{x}) &\sim \mathcal{GP}\big(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\big) \\ \mathbf{y} &= f(\mathbf{x}) + \boldsymbol{\epsilon} \end{aligned} \tag{2-8}$$

The noise term also gives rise to an alteration in the distribution of $\mathbf{y}$ and receives an additional variance term on the diagonal of the covariance matrix as

$$\mathbf{y} \ \sim \ \mathcal{N}\big(\mathbf{0}, K(X, X) + \sigma_n^2 I\big). \tag{2-9}$$

The joint distribution between $\mathbf{y}$ and $\mathbf{f}_*$ is given by

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \ \sim \ \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right). \tag{2-10}$$

This joint distribution gives rise to the following posterior distribution

$$\begin{aligned} \mathbf{f}_* | X_*, X, \mathbf{y} \ \sim \ \mathcal{N}\Big(&K(X_*, X)\big(K(X, X) + \sigma_n^2 I\big)^{-1}\mathbf{y}, \\ &K(X_*, X_*) - K(X_*, X)\big(K(X, X) + \sigma_n^2 I\big)^{-1} K(X, X_*)\Big) \end{aligned} \tag{2-11}$$

2-1-3 Hyperparameters

The squared exponential kernel in Eq. (2-3), like any other kernel, is subject to several hyperparameters. In this case, the signal variance σ_f and the lengthscale l are the hyperparameters. These are the settings that can be modified to change the behavior of the kernel. The signal variance indicates how much variance is present in the output data. This is the variance that can be observed when there are no observations available or nearby, see Figure 2-1. The lengthscale controls how fast the correlation between points farther apart from each other decreases. This determines how nonlinear the data are. A small lengthscale leads to a rapid decrease in the correlation between points that are farther apart, as can be verified in Eq. (2-3). This would result in a flexible and nonlinear approximation of the function. Larger lengthscales give rise to less spatially varying functions.

Determining suitable values can be a challenging task. It is crucial to find hyperparameters that represent the latent function. A common method to find optimal values is to maximize the marginal likelihood. The marginal likelihood represents the probability of generating the observed values $\mathbf{y}$ given the input data X and can be computed by

$$p(\mathbf{y}|X) = \int p(\mathbf{y}|\mathbf{f}, X)p(\mathbf{f}|X)\mathrm{d}\mathbf{f}. \tag{2-12}$$

By maximizing the marginal likelihood with respect to the hyperparameters, one tries to find the hyperparameters that maximize the probability of observing $\mathbf{y}$ given the data. Fortunately, for GPs the marginalization over the function values $\mathbf{f}$ is analytically tractable. Typically, the marginal log-likelihood is maximized for optimization purposes and is given by

$$\log p(\mathbf{y}|X, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^\intercal(K(X,X) + \sigma_n^2 I)^{-1}\mathbf{y} - \frac{1}{2}\log|K(X,X) + \sigma_n^2 I| - \frac{n}{2}\log 2\pi, \tag{2-13}$$

where $\boldsymbol{\theta}$ denotes the hyperparameters and the kernel function K depends on these hyperparameters. The first term on the right-hand side puts a penalty on how well the data are fitted and the second term on the complexity of the function. This ensures that there is a trade-off between the fit and the complexity of the function. Note, however, that the objective function is nonlinear and as a result it is possible that multiple local optima exist. It is advisable to start the optimization from multiple initialization points to capture the global optimum. Sometimes the algorithm will find multiple values for the hyperparameters that all have very similar performance with respect to the objective function. In that case, the engineer has to decide which one suits the underlying function best.

To increase optimization efficiency, it is useful to include the partial derivatives of Eq. (2-13) with respect to the hyperparameters. Depending on which kernel function is used, the derivative will differ. In general, the partial derivative is given by

$$\begin{aligned}\frac{\partial}{\partial\theta_j}\log p(\mathbf{y}|X,\boldsymbol{\theta}) &= \frac{1}{2}\mathbf{y}^\intercal K^{-1}\frac{\partial K}{\partial\theta_j}K^{-1}\mathbf{y} - \frac{1}{2}\mathrm{tr}(K^{-1}\frac{\partial K}{\partial\theta_j}) \\ &= \frac{1}{2}\mathrm{tr}\Big((\boldsymbol{\alpha}\boldsymbol{\alpha}^\intercal - K^{-1})\frac{\partial K}{\partial\theta_j}\Big),\end{aligned} \tag{2-14}$$

where $\boldsymbol{\alpha} = K^{-1}\mathbf{y}$ and $\mathrm{tr}(\cdot)$ is the trace operator.

2-2 Magnetic Fields

The magnetic field is a three-dimensional vector field with a direction and a magnitude. The Earth generates a magnetic field that changes both spatially and temporally due to variations in the Earth's crust, core convection, Earth's rotation, and solar wind interactions [43]. However, on a small scale these variations are very small or they change slowly over a very long period of time. Therefore, in this thesis, the assumption is made that, at least locally at the scale of a building, the Earth's magnetic field can be modeled as a vector field with constant direction and magnitude. We denote the magnetic field with the function $\mathbf{H}(\mathbf{x})$, where $\mathbf{H} : \mathbb{R}^3 \to \mathbb{R}^3$ and where $\mathbf{x}$ denotes the location in space.

Maxwell formulated the theory of the electromagnetic field in 1865 [23]. He described how electric and magnetic fields are influenced by changes in the fields and by electric charges

and currents. When the magnetic and electric fields are stationary and no free currents are present, the magnetic field exhibits the property of being curl- and divergence-free. For both properties to hold simultaneously, it is assumed that only external magnetic fields are considered, excluding any internal fields within magnetic objects. The curl- and divergence-free properties are respectively given by

$$\begin{aligned} \nabla \times \mathbf{H} &= 0, \\ \nabla \cdot \mathbf{H} &= 0. \end{aligned} \tag{2-15}$$

Throughout this thesis, the assumption is made that the curl-free property holds at all time and hence no free currents are present. This property is used to incorporate physical knowledge of the magnetic field into the GP model. The divergence-free property is not imposed in this thesis.

Anomalies in the magnetic field, which are caused by man-made structures that contain ferromagnetic materials, induce deviations from the constant magnetic field of the Earth [31]. This variability in the magnetic field between different locations can be used as features in indoor localization. By constructing a magnetic field map, the measurements of the magnetic field can be compared to the map to obtain location-specific information. These measurements are collected using small sensors known as magnetometers, which provide noisy measurements of the magnetic field at a specific location denoted as

$$\mathbf{y} = \mathbf{H}(\mathbf{x}) + \epsilon, \tag{2-16}$$

where $\mathbf{y} \in \mathbb{R}^3$ denotes the noisy measurement and $\epsilon \in \mathbb{R}^3$ the measurement noise, which is assumed to be zero mean white noise with an IID Gaussian distribution. We also assume that the measurements are bias-free.

In addition to using single magnetometers to measure the magnetic field, magnetometer arrays can also be used to measure at several points simultaneously. Since the primary goal of this thesis is to investigate how magnetic field measurements from magnetometer arrays can be utilized effectively and efficiently to create magnetic field maps, magnetometer arrays play a central role. A magnetometer array is a fixed structure, typically two- or three-dimensional, composed of multiple magnetometers. In this thesis, we assume that the exact spacing between the magnetometers is known. An array allows for measuring the magnetic field at several locations simultaneously, providing information that can be used for odometry purposes. So far, the literature on magnetic field localization has focused mainly on two-dimensional arrays. These arrays are usually arranged in a flat, grid-like configuration. Therefore, in this thesis, we will look into those two-dimensional arrays in a grid-like configuration as well. However, some studies related to the derivative-based odometry method use three-dimensional arrays [2, 3, 8, 9]. In addition, there are other applications, such as the localization of magnetic objects, that use three-dimensional arrays [13, 21, 33]. Insights from these applications regarding derivative estimation are used in this thesis.

2-3 Magnetic Field Mapping using GPR

To map the magnetic field using GPR, a dataset with spatial coordinates as inputs and the corresponding magnetic field measurements as outputs is required. The spatial coordinates

can take any dimension from one-dimensional to three-dimensional. GPs can handle input data of any dimension, provided that the kernel can measure the similarity between any two inputs in terms of a scalar value. Regarding the magnetic field output, either the magnitude of each measurement vector can be taken or the vector components can be considered separately [15]. Using vector components provides more detailed information and is likely to result in more accurate estimates. Furthermore, if the vector is used, it is still possible to retrieve the orientation, whereas this information is lost when the magnitude is considered [15]. In this thesis, the three components of the magnetic field are utilized. Additionally, only two-dimensional spatial coordinates will be considered, with the z-components of the field always set to zero.

2-3-1 Independent Component Modeling

A simple extension to incorporate vector outputs, as discussed in [31], is to model each component of the magnetic field separately without any correlation between them. This is a reasonable assumption, but also in [31] a method is introduced that does include cross-correlations between the magnetic field components. This method will be discussed in Section 2-3-2. To model the components separately, the kernel function can be extended as follows

$$k(\mathbf{x}, \mathbf{x}') = \begin{bmatrix} \sigma_{f_x}^2 \exp\left(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l_x^2}\right) & 0 & 0 \\ 0 & \sigma_{f_y}^2 \exp\left(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l_y^2}\right) & 0 \\ 0 & 0 & \sigma_{f_z}^2 \exp\left(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l_z^2}\right) \end{bmatrix} \tag{2-17}$$

where σ_{f_i} denotes the signal variance of the i^{th} magnetic field component and l the length-scale of the i^{th} components [40]. Since each component is modeled by a separate GP, it is possible to use separate hyperparameters for each output. This can be beneficial if there is a clear distinction between the behavior in different directions. The optimization algorithm can be three independent times to determine the hyperparameters. However, in [31] it argued that independently determining the hyperparameters might yield suboptimal solutions, as the parameters are even more inclined to converge to local optima and the magnetic field components might show very different behavior. Therefore, they suggested sharing the parameters between the components and doing a single optimization run that takes into account all components.

2-3-2 Incorporating Physical Properties into Magnetic Field Models

Instead of modeling each component of the magnetic field with a separate GP, a different approach to improving the map involves including prior knowledge in the kernel function. In 2013, Wahlström et al. [40] studied how the well-understood characteristics of magnets and their magnetic fields can be included in GPs. They exploited the divergence- and curl-free properties of the magnetic field.

In [31] it is described how the curl-free property can be included in the model. A curl-free vector field does not exhibit circulation around a fixed point in the domain. This indicates

that any line integral of the vector field between two arbitrary points A and B in the domain depends only on the difference between the values of the two points. The route taken to get from point A to B has no influence, as shown in

$$\int_P \mathbf{H}(\mathbf{x}) \cdot \mathrm{d}\mathbf{x} = \varphi(A) - \varphi(B), \tag{2-18}$$

where $\varphi : \mathbb{R}^3 \to \mathbb{R}$. φ can be considered to be scalar potential, analogous to how, for example, potential energy is the scalar potential of the gravitational field. Taking the gradient of Eq. (2-18) gives rise to the relationship between the magnetic field and the scalar potential, shown in

$$\mathbf{H} = -\nabla\varphi. \tag{2-19}$$

This relationship can be used to impose constraints on the potential output functions of the GP, as demonstrated in [31]. By initializing a GP with the scalar potential, $\varphi(\mathbf{x})$, as its realization and using spatial coordinates as input, the magnetic field vector is obtained by taking the negative gradient of $\varphi(\mathbf{x})$. This can be represented in a model as

$$\begin{aligned} \varphi(\mathbf{x}) &\sim \mathcal{GP}(0, k_{\text{lin.}}(\mathbf{x}, \mathbf{x}') + k_{\text{SE}}(\mathbf{x}, \mathbf{x}')), \\ \mathbf{y}_i &= -\nabla\varphi(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_i} + \epsilon_i. \end{aligned} \tag{2-20}$$

The $k_{\text{lin.}}$ was added to model the linear increase in the scalar potential due to the constant influence of the Earth's magnetic field. It is given by

$$k_{\text{lin.}}(\mathbf{x}, \mathbf{x}') = \sigma^2_{\text{lin.}} \mathbf{x}^\intercal \mathbf{x}'. \tag{2-21}$$

Since Gaussian distributions, and therefore also GPs, are closed under linear operations, and the gradient ∇ is a linear operator, it can incorporated into the GP. This gives rise to a special kernel function that ensures that the GP outputs possess the curl-free property. This model was computed to be

$$\mathbf{H}(\mathbf{x}) \;\sim\; \mathcal{GP}(\mathbf{0}, \sigma^2_{\text{lin.}}\mathbf{I}_3 + K_{\text{curl}}(\mathbf{x}, \mathbf{x}')), \tag{2-22}$$

where

$$K_{\text{curl}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(I_3 - \left(\frac{\mathbf{x} - \mathbf{x}'}{l}\right)\left(\frac{\mathbf{x} - \mathbf{x}'}{l}\right)^\intercal \right) \exp\left(- \frac{||\mathbf{x} - \mathbf{x}'||^2}{2l^2} \right). \tag{2-23}$$

More detail on how linear operations are incorporated into GPs will follow in Section 3-1. In [31] they implemented this method and showed that significant improvements can be made compared to a standard GP that uses only a squared exponential kernel. More specifically, predictions were shown to improve in terms of accuracy and the 95% confidence interval became narrower. This modeling techniques, or a variation of it, has also been applied in other magnetic field localization and SLAM implementations [19, 29, 31, 37]

Additionally, it is possible to incorporate the divergence-free property into the GP prior [29]. However, this approach, specifically in combination with GPs, is less commonly implemented. Consequently, we decided to solely use the curl-free property throughout this thesis. However, the following section in this chapter still considers the divergence-free property as well to provide a comprehensive background. An alternative approach is to model the magnetic field using a polynomial model [16, 29, 30], however, this method is not explored in this thesis.

2-4 Constraining Spatial Derivative of Magnetic Field

By including the curl- and divergence-free properties in the magnetic field model, it puts constraints on the spatial derivative. These constraints reduce the degrees of freedom of the Jacobian and, as a result, also the minimum number of magnetometers required to compute the spatial derivative. Assuming that the magnetic field is stationary and that no electric charges and currents are present, the magnetic field is curl- and divergence-free as shown in Eq. (2-15).

2-4-1 First-Order Spatial Derivative

The divergence property yields that the trace of the Jacobian must be zero, as can be directly observed from the definition

$$\begin{aligned} \nabla \cdot \mathbf{H}(\mathbf{x}) &= 0 \\ \frac{\partial H_1}{\partial x_1} + \frac{\partial H_2}{\partial x_2} + \frac{\partial H_3}{\partial x_3} &= 0. \end{aligned} \tag{2-24}$$

This constraint reduces the number of independent terms in the Jacobian from nine to eight, as shown in [10]. The curl-free property implies that the Jacobian matrix has to be symmetric, as can also be inferred from the definition

$$\begin{aligned} \nabla \times \mathbf{H} &= \mathbf{0} \\ \begin{bmatrix} \frac{\partial H_3}{\partial x_2} - \frac{\partial H_2}{\partial x_3} \\ \frac{\partial H_1}{\partial x_3} - \frac{\partial H_3}{\partial x_1} \\ \frac{\partial H_2}{\partial x_1} - \frac{\partial H_1}{\partial x_2} \end{bmatrix} &= \mathbf{0}. \end{aligned} \tag{2-25}$$

This equation shows that the symmetric terms have to be equal. This reduces the number of independent terms in the Jacobian even further from eight to five [10].

Since there are only five independent terms in the Jacobian, it is possible to reconstruct an entire column of the Jacobian and only compute the values for the other two [33]. In other words, an array that is positioned on a two-dimensional plane is already sufficient to determine the full Jacobian. In [33] they equivalently showed that the Jacobian only has five independent terms due to the curl- and divergence-free properties and it can be expressed as

$$J = \begin{bmatrix} \frac{\partial H_x}{\partial x} & \frac{\partial H_x}{\partial y} & \frac{\partial H_x}{\partial z} \\ \frac{\partial H_y}{\partial x} & \frac{\partial H_y}{\partial y} & \frac{\partial H_y}{\partial z} \\ \frac{\partial H_z}{\partial x} & \frac{\partial H_z}{\partial y} & \frac{\partial H_z}{\partial z} \end{bmatrix} = \begin{bmatrix} j_{xx} & j_{xy} & j_{xz} \\ j_{xy} & j_{yy} & j_{yz} \\ j_{xz} & j_{yz} & -j_{xx} - j_{yy} \end{bmatrix}, \tag{2-26}$$

where J denotes the Jacobian, x, y and z the spatial coordinates and j a single element of the Jacobian matrix. The first index of j represents the magnetic field component that is

considered, while the subsequent indices specify the coordinate with respect to which the derivative is taken. This three-dimensional Jacobian in Eq. (2-26) confirms that, for instance, the last column column can be written in terms of the components of the first two. This shows that the spatial derivative can be computed using an array that lies in the x-y plane. These constraints on the shape of the Jacobian matrix can be utilized to get better Jacobian estimates that satisfy the behavior of the magnetic field.

However, note that this thesis only considers the curl-free property. The effect of this choice on the Jacobian is further explained in Section 3-2.

2-4-2 Higher-Order Spatial Derivatives

The effect of the curl- and divergence-free properties can also be applied to higher-order spatial derivatives. In [4] it was shown that using the curl- and divergence-free property, higher-order spatial derivatives can also be derived from a two-dimensional planar arrangement of magnetometers. Since this thesis does not take into account the divergence-free property, only the curl-free property is considered here.

The effect of the curl-free property on higher-order spatial derivatives is analyzed in [4]. They showed that any derivative of order n is equal to the derivative that uses any permutation of the set of n derivative terms and the considered magnetic field component. Consider a function ρ from $\{1, 2, \ldots, n\}$ to the three spatial coordinates $\{x, y, z\}$, and σ any permutation on $\{1, 2, \ldots, n\}$, with $n > 1$, then this property is given by

$$\frac{\partial^{(n-1)} H_{\rho(n)}}{\partial\rho(1)\partial\rho(2)\ldots\partial\rho(n-1)} = \frac{\partial^{(n-1)} H_{\rho(\sigma(n))}}{\partial\rho(\sigma(1))\partial\rho(\sigma(2))\ldots\partial\rho(\sigma(n-1))}. \tag{2-27}$$

They used the curl-free property from Eq. (2-25) and Schwarz's theorem to prove this by induction on n. For a first-order derivative, with $n = 2$, Eq. (2-25) immediately showed that the property holds. Then Schwarz's theorem showed that if it holds for $n - 1$ it also holds for n, which completed the proof. The property in Eq. (2-27), simply put, means that the indices of any element in a certain derivative matrix can be rearranged without changing the variable. For instance, the elements $j_{zyx} = j_{xyz}$ are equal.

2-5 Magnetic Odometry Methods

The aim of this thesis of using the spatial derivative as an approximation of the magnetometer array originates from two magnetic odometry methods: the derivative-based [2] and the model-based [30] odometry methods, which were briefly introduced in Chapter 1. These two methods differ, among other things, in the information they use from the array. The model-based odometry method uses all of the magnetometers on the array to determine the change in pose. The derivative-based method, on the other hand, computes the spatial derivatives from the information on the array and uses that to estimate the change in position. Essentially, the derivative-based method approximates the information on the array, whereas the model-based method does not. However, the derivative-based method does not necessarily use this spatial derivative as an approximation to save computational resources; the method simply requires the spatial derivative. Therefore, the main objective of this thesis is to analyze whether

the spatial derivative could serve as an approximation to the complete information on the array for mapping purposes using GPR. This approximation could, for instance, be utilized for the model-based odometry method or SLAM implementations that include arrays. The derivative-based and model-based odometry methods are explained in more detail here to give more context on how arrays have been used before and to illustrate how the findings from this thesis could potentially be useful for these methods.

Both methods are odometry techniques that rely on the spatial information captured by the array. Odometry is the process of estimating the change in pose of a moving object, typically based on motion data. Typically inertial measurement units (IMUs) are used to measure acceleration and angular velocity. The integration of these variables would give rise to this pose change, though it accumulates errors over time due to sensor errors [18]. Magnetic field odometry works by analyzing how the magnetic field changes over time and how this relates to the spatial information captured by the array. This also provides an estimate of the pose change.

2-5-1 Derivative-Based Odometry Method

The derivative-based odometry method introduced in [39] derived a differential equation that can explicitly compute the velocity based on the spatial derivative:

$$\frac{d\mathbf{H}}{dt} = -\boldsymbol{\Omega} \times \mathbf{H} + \frac{d\mathbf{H}}{d\mathbf{r}}\mathbf{v}, \tag{2-28}$$

where $\boldsymbol{\Omega}$ denotes the angular velocity, $\mathbf{r}$ the position vector and $\mathbf{v}$ the velocity. In Eq. (2-28), $\mathbf{v}$ is the only unknown. The time derivative of $\mathbf{H}$ can be straightforwardly determined using a finite difference method, $\boldsymbol{\Omega}$ is measured by the gyroscope and the spatial derivative of $\mathbf{H}$ can be determined from the array.

A traditional inertial navigation system (INS) only utilizes the measurements from the accelerometer and gyroscope to retrieve the pose estimate. Typically, state-space representations in combination with an extended Kalman filter (EKF) or observer are used to model the dynamics and take the measurements into account. In [5, 6, 10, 39, 42], magnetic field odometry is integrated into an INS by including Eq. (2-28) into the state-space representation. The spatial derivative is either used as a separate state or an input to the state-space model. In either case, the spatial derivative needs to be computed from the array. Eq. (2-28) links the magnetic field measurements to the velocity, ensuring observability of the velocity. Without including the magnetic field odometry in the INS the velocity would not be observable. Integration of the velocity then gives rise to the desired estimate of the position. In case there is enough spatial variability in the magnetic field, this magnetic field-aided INS has a position error that grows linearly with time.

The accuracy of this odometry method is largely dependent on the accuracy with which spatial derivative can be derived. An important aspect influencing the accuracy of the derivative is the calibration of the magnetometers. One can imagine that since the distance between the magnetometers is small, it results in a similarly small difference in the magnetic field. As a result, even minor calibration errors can significantly affect the measured difference in magnetic field, leading to a large impact on the accuracy of the spatial derivative estimate.

Therefore, in their following studies, they also put effort into further exploring calibration schemes [11, 12].

Another important aspect is related to the configuration of the array, including the shape of the array, the number of magnetometers, the spacing between the magnetometers and the noise levels. All of these variables influence the accuracy of the spatial derivative estimates. This is related to the first research sub-question and is investigated in this thesis. While the derivative-based odometry method typically utilizes three-dimensional arrays, this thesis focuses mainly on two-dimensional arrays. Nevertheless, the results still give a general understanding of how the different parts of the array configuration influence the derivative estimates. Furthermore, it also emphasizes the importance of a suitable array configuration.

2-5-2 Model-Based Odometry Method

An alternative approach, inspired by the derivative-based odometry method, is the model-based odometry technique [16, 29, 30]. Instead of directly evaluating the velocity from Eq. (2-28), this method first constructs a local magnetic field map based on the measurements from the array at a single time instant. The measurements at the next time step are then compared to this local map to obtain an estimate of the change in pose [29, 30]. A polynomial model is used to create the local map, and either a maximum likelihood estimator or a nonlinear least squares method is used to compare the measurements at the next time step to the local map, resulting in an estimate of the change in pose. In essence, this method relies on the partial overlap between two consecutive series of measurements to infer the change in pose. An advantage of this method compared to the derivative-based method is that an estimate of the change in orientation is also obtained.

In [16], this method is integrated into a magnetic field-aided INS. This INS utilizes an error state Kalman filter (ESKF) to account for the non-Euclidean space in which the quaternions live. The state vector is augmented with the coefficients of the polynomial model, which are updated iteratively during the time updates. When the new magnetic field measurements come in, the measurement update is executed, correcting the error state. The nominal state is then updated, and the error state is reset, leading to updated model coefficients. At the same time, this measurement update ensures that the pose error is updated and corrected based on the difference between the new magnetic field measurements and the expected measurements. This method showed a location error of less than three meters in most cases after a two-minute long trajectory using an array of 30 magnetometers. This implementation into an INS demonstrates the functionality of magnetic field odometry and the importance of the investigation of magnetometer arrays.

To the best of our knowledge, all studies regarding the model-based odometry method use two-dimensional arrays so far. However, they do create models that are capable of making predictions in all three dimensions. This is possible because the curl- and divergence-free properties of the magnetic field impose constraints that allow the spatial derivative to be determined from measurements that lie in a two-dimensional plane. In other words, a two-dimensional array is enough to be able to tell how the magnetic field will evolve in all three directions. This concept was explained in Section 2-4. Because of this effect, in [16] they evaluate the performance of the magnetic field-aided INS for all three spatial directions using a two-dimensional array.

The accuracy of this odometry method in terms of location drift depends mainly on the accuracy with which the local map can make predictions. A higher quality map would provide more accurate estimates of the magnetic field, resulting in more precise pose change estimates and reduced drift over time. A small contribution of this thesis to this odometry method lies in testing different array configurations for the quality of the local map. Despite two key differences between the typical implementation of this odometry method and the methods used in this thesis, the results may still provide valuable insight into how the configuration of the array influences the local mapping accuracy. The first key difference is that in [16] they use a polynomial model to represent the magnetic field, and this thesis focuses on GPR. And secondly, in [16] the divergence-free property is also included in the model and magnetic field, whereas in this thesis only the curl-free property is considered. Nevertheless, this thesis provides results that could be beneficial for this odometry method. Specifically, the map quality of local maps for different array configurations when using all magnetometers on the array is analyzed in this thesis. Finding an effective array configuration that produces high-quality local maps could improve the performance of this odometry method.

2-5-3 Simultaneous Localization and Mapping (SLAM)

A potential application of these odometry methods and magnetometer arrays is in SLAM. SLAM solves the problem of having to create a map required for indoor localization by jointly estimating the global map and pose while exploring the area of interest [19, 37]. It uses odometry data, typically from accelerometers and gyroscopes, to track its pose in the short term and corrects it for drift by using the global map constructed along the way. Incorporation of the magnetic odometry methods using magnetometer arrays is actively being studied and has the potential to improve the overall performance.

Alternatively, instead of explicitly incorporating a magnetic odometry method that utilizes the array into SLAM, the possibility of using all magnetometer measurements from the array to create the global map is being studied. However, since the time complexity of GPs scales cubically with the number of data points, using all magnetometers on the array could quickly cause computational problems. By approximating the information on the array with a single measurement and the spatial derivative, the computational burden can be reduced while still capturing meaningful spatial information. This thesis explores the implementation of this approximation and whether it is effective. Not only could the spatial derivative be used as an approximation to information on the array for this particular SLAM application, but it is applicable to any mapping application that uses arrays of measurements and suffers from computational issues. The results of this thesis will shed light on whether the spatial derivative is an effective and efficient approximation.

Chapter 3

Methods

This chapter presents the methods required for implementing and using the spatial derivative for magnetic field mapping. In Section 3-1, we analyze how previous studies have incorporated derivatives in Gaussian process regression (GPR) and we build upon this approach to include the spatial derivative into the curl-free Gaussian process (GP) model. Section 3-2 then introduces how three existing derivative computation methods can be utilized to compute the spatial derivative from a magnetometer array; the finite difference method, the linear least squares (LLS) method and the local GP method. Finally, the simulation experiments designed to answer the research questions are described in Section 3-3. Two types of simulation are presented: the first assesses the accuracy of spatial derivative computation for the three derivative estimation methods across various array configurations, while the second assesses map quality using both the derivative and the full kernel.

3-1 Incorporating Spatial Derivatives into GPR

In order to make prediction of the magnetic field using GPR based on the spatial derivative, the derivative has to be included into the GP prior. Incorporating the derivative into the prior ensures that the model accounts for not only the field's values but also the derivative. First, linear operators are introduced and their application to GPs. Subsequently, based on the existing literature, it is shown how derivatives can be included into GPs. This is then used to incorporate the spatial derivative into the curl-free model.

3-1-1 Linear Operators

Linear operators are a more general form of functions that take functions as input and output functions and have been studied extensively, for instance in [34] in combination with GPR. Following the notation of [34], the application of the linear operator $\mathcal{L}_x$ to the function $f(x)$ will be denoted as $g(x) = \mathcal{L}_x f(x)$, where the subscript x indicates the variable on which the operator is applied. It is assumed that the input domain of the original function has the same

dimensionality as the functions produced by the operator. The output dimensionality may change.

Gaussian processes are closed under linear operators, and therefore a GP remains a GP after a linear operator has been applied [27]. Consider a GP that outputs a scalar and is denoted by

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu_f, K_{ff}(\mathbf{x}, \mathbf{x}')). \tag{3-1}$$

By considering the linear operator $g(x) = \mathcal{L}_{\mathbf{x}} f(\mathbf{x})$, the mean and covariance function of the GP prior change according to the operator. According to [34], the kernel functions that describe the covariance between $g(x)$ and $f(x)$ are obtained as follows

$$\begin{aligned} K_{gf}(\mathbf{x}, \mathbf{x}') &= \mathcal{L}_{\mathbf{x}} K_{ff}(\mathbf{x}, \mathbf{x}') \\ K_{fg}(\mathbf{x}, \mathbf{x}') &= K_{ff}(\mathbf{x}, \mathbf{x}') \mathcal{L}_{\mathbf{x}'}^{\intercal} \\ K_{gg}(\mathbf{x}, \mathbf{x}') &= \mathcal{L}_{\mathbf{x}} K_{ff}(\mathbf{x}, \mathbf{x}') \mathcal{L}_{\mathbf{x}'}^{\intercal}. \end{aligned} \tag{3-2}$$

and the GP prior mean is given by

$$\mu_g = \mathcal{L}_{\mathbf{x}} \mu_f \tag{3-3}$$

If the operator is applied from the right, it means that it operates on the second argument. This gives rise to the GP prior for $g(\mathbf{x})$ as

$$g(\mathbf{x}) \sim \mathcal{GP}(\mathcal{L}_{\mathbf{x}} \mu_f, \mathcal{L}_{\mathbf{x}} K_{ff}(\mathbf{x}, \mathbf{x}') \mathcal{L}_{\mathbf{x}'}^{\intercal}). \tag{3-4}$$

3-1-2 Applying the Derivative Operator on the Curl-Free GP Model

The information on linear operators in Section 3-1-1 can be used to obtain a GP that models the spatial derivative of the curl-free GP from Eq. (2-22). That is because differentiation can be represented by a linear operator as

$$\mathcal{D}_{\mathbf{x}} f(\mathbf{x}) = \nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}, \tag{3-5}$$

where $\mathcal{D}_{\mathbf{x}}$ is the derivative operator and n is the dimensionality of the input.

To obtain a GP prior that not only models the magnetic field $\mathbf{H}(\mathbf{x})$, but also the spatial derivative $\mathbf{H}_d(\mathbf{x})$, a joint distribution of the magnetic field and the spatial derivative must be derived. This distribution can be obtained by applying the derivative operator on the curl-free GP from Eq. (2-22). This yields the following distribution

$$\begin{aligned} \begin{bmatrix} \mathbf{H}(\mathbf{x}) \\ \mathbf{H}_d(\mathbf{x}) \end{bmatrix} &\sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathcal{D}_{\mathbf{x}} \mathbf{0} \end{bmatrix}, \begin{bmatrix} K_{\text{curl}}(\mathbf{x}, \mathbf{x}') & K_{\text{curl}}(\mathbf{x}, \mathbf{x}') \mathcal{D}_{\mathbf{x}'}^{\intercal} \\ \mathcal{D}_{\mathbf{x}} K_{\text{curl}}(\mathbf{x}, \mathbf{x}') & \mathcal{D}_{\mathbf{x}} K_{\text{curl}}(\mathbf{x}, \mathbf{x}') \mathcal{D}_{\mathbf{x}'}^{\intercal} \end{bmatrix} \right) \\ &= \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, K_{\text{curl,d}}(\mathbf{x}, \mathbf{x}') \right), \end{aligned} \tag{3-6}$$

where $\mathbf{H}_d(\mathbf{x}) \in \mathbb{R}^9$ is defined according to Eq. (3-5) as

$$\mathbf{H}_d(\mathbf{x}) = \mathcal{D}_{\mathbf{x}} \mathbf{H}(\mathbf{x}) = \begin{bmatrix} \frac{\partial \mathbf{H}(\mathbf{x})}{\partial x} \\ \frac{\partial \mathbf{H}(\mathbf{x})}{\partial y} \\ \frac{\partial \mathbf{H}(\mathbf{x})}{\partial z} \end{bmatrix}. \tag{3-7}$$

The derivation of $K_{\text{curl,d}}(\mathbf{x}, \mathbf{x}')$ is provided in Appendix A. Later in the thesis it is useful if sub-matrices of $K_{\text{curl,d}}(\mathbf{x}, \mathbf{x}')$ are defined. Therefore, the following definitions are introduced

$$\begin{aligned} K^{KD}_{\text{curl,d}}(\mathbf{x}, \mathbf{x}') &= \begin{bmatrix} K_{\text{curl}}(\mathbf{x}, \mathbf{x}') & K_{\text{curl}}(\mathbf{x}, \mathbf{x}')\mathcal{D}^{\intercal}_{\mathbf{x}'} \end{bmatrix} \\ K^{DK}_{\text{curl,d}}(\mathbf{x}, \mathbf{x}') &= \begin{bmatrix} K_{\text{curl}}(\mathbf{x}, \mathbf{x}') \\ \mathcal{D}_{\mathbf{x}} K_{\text{curl}}(\mathbf{x}, \mathbf{x}') \end{bmatrix} \end{aligned} \tag{3-8}$$

3-2 Derivative Estimation

An important aspect when computation a spatial derivative, is the choice of the derivative estimation method. To examine the impact of these methods on the approximation of the information on the array, three different numerical differentiation techniques are considered. Each method is evaluated based on its accuracy in computing the first-order derivative and the time required to perform the computation. The computational efficiency plays an essential role, as the goal is to approximate the information on the array to significantly reduce the computational time.

This thesis solely uses the curl-free property of the magnetic field to generate and model the data. Consequently, the Jacobian consists of six independent terms instead of five when the divergence-free property was also included, as shown in Section 2-4-1. We also chose to align the array with the x-y plane. As a result, we are able to directly obtain estimates of the components of the first two columns of the Jacobian. However, since the divergence-free property is not included, the final column cannot be fully written in terms of the components of the first two columns. That is because the divergence-free property is required to determine the j_{zz} component. Hence, from this setup, all but one component can be determined from a two-dimensional array. The Jacobian is represented by

$$J = \begin{bmatrix} j_{xx} & j_{xy} & j_{xz} \\ j_{xy} & j_{yy} & j_{yz} \\ j_{xz} & j_{yz} & j_{zz} \end{bmatrix}, \tag{3-9}$$

where the off-diagonal terms are symmetric relative to the diagonal due to the curl-free property. This property is included in all three following techniques.

By incorporating the curl-free property into the GP model, which was described in Section 2-3-2 [31], the model takes these constraints on the Jacobian into account. As a result, part of the structure of the Jacobian is already included in the model. Therefore, the amount of additional information provided by the Jacobian is slightly reduced. Furthermore, including the curl-free property reduces the minimum number of magnetometers required to compute the derivative, which already suggests that it contains less information.

3-2-1 Finite Difference Methods

The most ubiquitous way to determine derivatives is through finite differences. Depending on the number of magnetometers, different orders of finite difference schemes can be employed.

To derive the coefficients of the finite difference schemes, the Taylor series expansion is utilized

$$f(x_0 + ih) = \sum_{n=0}^{\infty} \frac{(ih)^n}{n!} f^{(n)}(x_0), \tag{3-10}$$

where h denotes the spacing between measurement points and $i \in \mathbb{Z}$. For instance, the expansion at $x_{-1} = x_0 - h$ and $x_1 = x_0 + h$ yields

$$\begin{aligned} f(x_{-1}) &= f(x_0) - hf'(x_0) + \frac{h^2}{2!} f''(x_0) - \frac{h^3}{3!} f^{(3)}(x_0) + \frac{h^4}{4!} f^{(4)}(x_0) - \mathcal{O}(h^5), \\ f(x_1) &= f(x_0) + hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \frac{h^3}{3!} f^{(3)}(x_0) + \frac{h^4}{4!} f^{(4)}(x_0) + \mathcal{O}(h^5), \end{aligned} \tag{3-11}$$

where $x_i = x_0 + ih$, for $i = \ldots, -1, 0, 1, \ldots$. By linearly combining the Taylor series expansions of the function f at different points, various orders of the approximation of $f'(x_0)$ can be obtained. Note that for this method, as well as for the subsequent two methods, a measurement at the locations where the first-order derivative will be determined, $f(x_0)$, is required. Therefore, in this thesis, where the spatial derivative is derived from the array, the location x_0 is chosen based on whether a magnetometer is present at the center of the array. If so, x_0 is selected at the center location. Otherwise, x_0 is selected at the magnetometer location positioned to the top right of the center.

Solving a finite difference equation yields only the derivative of one component of the magnetic field in a single direction, meaning it provides just one element of the Jacobian. Since there are six independent terms present in the Jacobian, see Eq. (3-9), six separate finite difference equations must be solved to estimate the full Jacobian. However, since this thesis focuses mainly on two-dimensional arrays and mapping, the z-direction is ignored most of the time, leaving five components to be estimated from five equations. Each finite difference equation uses magnetometers aligned in a specific direction to determine the spatial derivative along that axis. As a result, not all magnetometers are used when computing the spatial derivative with this method. Only those forming a cross pattern on the array are used. As the number of magnetometers increases, the finite difference method uses a smaller proportion of the available magnetometers.

For a two-by-two magnetometer array, where x_0 is the top-right location, the only applicable finite difference formulas are the well-known forward and backward difference. In this thesis, the backward difference is used

$$\begin{aligned} \frac{f(x_0) - f(x_{-1})}{h} &= f'(x_0) - h\frac{f''(x_0)}{2!} + h^2\frac{f^{(3)}(x_0)}{3!} - h^3\frac{f^{(4)}(x_0)}{4!} + \mathcal{O}(h^4) \\ &= f'(x_0) + \mathcal{O}(h), \end{aligned} \tag{3-12}$$

providing a first-order approximation of $f'(x_0)$. Three out of four magnetometers are used to find the Jacobian matrix, except for the j_{zz} component.

The three-by-three array has access to three magnetometers along each in-plane axis and can combine Taylor series expansions of $f(x_{-1})$, $f(x_0)$ and $f(x_1)$, where in this case x_0 is indeed the middle magnetometer. This results in the central difference approximation

$$\begin{aligned} \frac{f(x_1) - f(x_{-1})}{2h} &= f'(x_0) + h^2\frac{f^{(3)}(x_0)}{3!} + \mathcal{O}(h^4) \\ &= f'(x_0) + \mathcal{O}(h^2), \end{aligned} \tag{3-13}$$

which provides a second-order approximation of $f'(x_0)$. This array configuration makes use of five out of nine magnetometers on the array. Interestingly, due to the symmetry of the central difference formula, all terms involving odd powers of h cancel out. This is why in the first part of Eq. (3-13), $\mathcal{O}(h^4)$ is added instead of $\mathcal{O}(h^3)$. This cancellation due to symmetry will have effects on the results, even for the least squares approach introduced later.

The four-by-four array can combine Taylor series expansions of $f(x_{-2})$, $f(x_{-1})$, $f(x_0)$ and $f(x_1)$, where x_0 is the magnetometer to the top right from the center of the array. It gives rise to the following third-order approximation

$$\begin{aligned}\frac{2f(x_1)+3f(x_0)-6f(x_{-1})+f(x_{-2})}{6h} &= f'(x_0)+h^3\frac{f^{(4)}(x_0)}{12}+\mathcal{O}(h^4)\\ &= f'(x_0)+\mathcal{O}(h^3).\end{aligned} \tag{3-14}$$

Finally, the five-by-five array can combine Taylor series expansions of $f(x_{-2})$, $f(x_{-1})$, $f(x_0)$, $f(x_1)$ and $f(x_2)$. Together they yields a fourth-order central difference formula

$$\begin{aligned}\frac{-f(x_2)+8f(x_1)-8f(x_{-1})+f(x_{-2})}{12h} &= f'(x_0)-h^4\frac{f^{(5)}}{30}+\mathcal{O}(h^6)\\ &= f'(x_0)+\mathcal{O}(h^4),\end{aligned} \tag{3-15}$$

where due to the symmetry of the array, all odd powers of h again cancel out.

These are the four finite difference formulas used to compute the first-order derivative in the finite difference method.

In this case, the curl-free property is also included to improve the estimate. The property constrains the Jacobian in such a way that the matrix must be symmetric. By averaging the derivative pairs that should be equal, some of the noise can be mitigated. This is done for j_{xy} and j_{yx}, for j_{xz} and j_{zx}, and for j_{yz} and j_{zy}.

3-2-2 Linear Least Squares Method

An alternative approach also revolves around the use of the Taylor series expansion. In [33] they use the Taylor series expansion of the magnetic field to solve for the first- and second-order spatial derivative matrices. However, they make use of the so-called rotating modulation method. This method requires one or more magnetometers on a spinning disk. In this thesis a fixed array is considered, and hence this method cannot be applied. Instead, the Taylor series expansion is tailored to an LLS problem. This fixed array requires more magnetometers to compute the spatial derivative, which will be covered later in this section.

The n^{th}-order Taylor series expansion of the magnetic field around the origin, expressed in vector form, is given by

$$\mathbf{H}(\mathbf{x}) \approx \mathbf{H}(\mathbf{x}_0)+\mathcal{D}_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)+\frac{1}{2}\Big(\mathcal{D}^2_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)\Big)(\mathbf{x}-\mathbf{x}_0)^{\otimes 2}+\cdots+\frac{1}{n!}\Big(\mathcal{D}^n_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)\Big)(\mathbf{x}-\mathbf{x}_0)^{\otimes n}, \tag{3-16}$$

where $\mathbf{H}(\mathbf{x})$ denotes the magnetic field function evaluated at location $\mathbf{x}$, $\mathbf{x}_0$ again denotes the location of the magnetometer assigned as the middle magnetometer, $\mathcal{D}^n_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0) \in \mathbb{R}^{3\times 3^n}$ denotes the n^{th}-order spatial derivative of the magnetic field where the derivatives of the

different components get stacked in a long row matrix, and where $(\mathbf{x}-\mathbf{x}_0)^{\otimes n}$ denotes n subsequent evaluations of the Kronecker products as $(\mathbf{x}-\mathbf{x}_0)\otimes(\mathbf{x}-\mathbf{x}_0)\otimes\cdots\otimes(\mathbf{x}-\mathbf{x}_0)$.

By capturing enough data points with the array at sufficiently varying locations and organizing the measurements and positions into matrices, a full-rank system of equations can be formed, which can be solved for the elements of the derivative matrices using LLS. The only term in Eq. (3-16) that is nonlinear in the derivative elements is $(\mathbf{x}-\mathbf{x}_0)^{\otimes i}$. Fortunately, this term represents the relative distances between the magnetometers on the array and is assumed to be a known variable and can be evaluated. Therefore, Eq. (3-16) is a linear expression in the derivative elements, which are embedded in the $\mathcal{D}^i_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)$ terms. To express the collected data in the standard linear least-squares form, all N magnetometer measurements are first arranged into the following matrices

$$\begin{bmatrix} \sum_{i=1}^{n}\frac{1}{i!}\mathcal{D}^i_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)(\mathbf{x}_1-\mathbf{x}_0)^{\otimes i} \\ \sum_{i=1}^{n}\frac{1}{i!}\mathcal{D}^i_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)(\mathbf{x}_2-\mathbf{x}_0)^{\otimes i} \\ \vdots \\ \sum_{i=1}^{n}\frac{1}{i!}\mathcal{D}^i_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)(\mathbf{x}_N-\mathbf{x}_0)^{\otimes i} \end{bmatrix} \approx \begin{bmatrix} \mathbf{H}(\mathbf{x}_1) \\ \vdots \\ \mathbf{H}(\mathbf{x}_N) \end{bmatrix} - \begin{bmatrix} \mathbf{H}(\mathbf{x}_0) \\ \vdots \\ \mathbf{H}(\mathbf{x}_0) \end{bmatrix}, \tag{3-17}$$

where the subscript $\mathbf{x}_i$ denotes the specific magnetometer index. Subsequently, the Sympy Python library was used to symbolically perform the operations in the left-hand side (LHS) of Eq. (3-17) and to obtain the LLS expression. Fortunately, the LHS of the equation is independent of the measurements themselves and depends only on the relative distances between the magnetometers that are time-invariant. As a result, the LHS only has to be calculated once using the Sympy library, rather than at each time step when new measurements arrive. The right-hand side of Eq. (3-17) has to be computed every time step, but its computational cost is negligible compared to solving the LLS.

Similarly to the finite difference method, various orders of approximation can be used. This can be done by including more or fewer higher-order derivative terms in the Taylor series expansion. This method requires assigning symbolic variables to all spatial derivative matrices, which will be estimated using LLS. To include the curl-free property, it must also be correctly incorporated into the higher-order spatial derivatives. While Eq. (3-9) showed this for the Jacobian, it can also be extended to higher-order derivatives. For instance, the second-order spatial derivative, $\mathcal{D}^2_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0)$, can be derived using Eq. (2-27) to be

$$\begin{aligned} \mathcal{D}^2_{\mathbf{x}}\mathbf{H}(\mathbf{x}_0) = \mathbf{G} &= \begin{bmatrix} \mathbf{G_x} & \mathbf{G_y} & \mathbf{G_z} \end{bmatrix}, \\ \mathbf{G_x} = \frac{\partial J}{\partial x} &= \begin{bmatrix} g_{xxx} & g_{xxy} & g_{xxz} \\ g_{xxy} & g_{xyy} & g_{xyz} \\ g_{xxz} & g_{xyz} & g_{xzz} \end{bmatrix}, \\ \mathbf{G_y} = \frac{\partial J}{\partial y} &= \begin{bmatrix} g_{xxy} & g_{xyy} & g_{xyz} \\ g_{xyy} & g_{yyy} & g_{yyz} \\ g_{xyz} & g_{yyz} & g_{yzz} \end{bmatrix}, \\ \mathbf{G_z} = \frac{\partial J}{\partial z} &= \begin{bmatrix} g_{xxz} & g_{xyz} & g_{xzz} \\ g_{xyz} & g_{yyz} & g_{yzz} \\ g_{xzz} & g_{yzz} & g_{zzz} \end{bmatrix}, \end{aligned} \tag{3-18}$$

where g denotes the individual elements of the Hessian, $\mathbf{G}$. Similarly, the elements of higher-order spatial derivatives can be derived. Naturally, it is also possible to include the divergence-free property in the higher-order spatial derivatives, but this is not covered in this thesis; see [4] for more details.

Minimum Required Number of Magnetometers

Higher-order approximations require more measurements for the LLS method to be full-rank. The number of independent terms in a n^{th}-order derivative matrix can be derived using the property of Eq. (2-27), which is derived from the curl-free property. This means that any permutation of the indices of any element in a certain derivative matrix gives rise to an element that is identical. For instance, the elements $j_{zyx} = j_{xyz}$ are equal, since zyx is a permutation of xyz and vice versa. Given this property, the number of independent elements for an n^{th}-order derivative matrix is equal to the number of n+1-combinations with repetition of the set $\{x, y, z\}$ with size $k = 3$ where the ordering does not matter. Using combinatorics theory the number of independent variables of a n-order derivative matrix using the curl-free property is given by

$$\binom{n+k}{n+1} = \frac{(n+k)!}{(n+1)! \cdot (k-1)!} = \frac{(n+3)!}{(n+1)! \cdot 2} = \frac{(n+3)(n+2)}{2}, \tag{3-19}$$

which indeed gives rise to six independent variables for the first-order derivative.

Since a Taylor series expansion consists of a sum of derivatives of increasing order, the total number of unknown parameters for the LLS problem of an n^{th}-order approximation corresponds to the sum of the number of independent terms for each order. To compute the number of magnetometers required to solve the LLS problem the reference measurement $\mathbf{H}(\mathbf{x}_0)$ from Eq. (3-17) must also be taken into account. This is achieved by running the summation from the zero$^{\text{th}}$-order instead of the first, effectively adding the additional three terms. This gives rise to the following number of magnetometers for an n^{th}-order approximation

$$\left\lceil \frac{\sum_{i=0}^{n} \frac{(i+3)(i+2)}{2}}{3} \right\rceil = \left\lceil \sum_{i=0}^{n} \frac{(i+3)(i+2)}{6} \right\rceil, \tag{3-20}$$

where $\lceil \cdot \rceil$ denotes the ceiling function. Since each magnetometer provides three measurements, the summation was divided by three. The ceiling function was used to round the number up to the nearest whole number, since of course it is not possible to use a fractional magnetometer. However, note that the minimum number of required magnetometers does not mean that that number is optimal. Since there is always noise present in the sensor measurements, it is beneficial to have more sensors active so that noise errors can be averaged out.

Also note that Eq. (3-20) includes derivative elements that are related to the z-direction and z-component as well. The minimum number of magnetometers required to determine the spatial derivative for a two-dimensional mapping problem is slightly lower. For a two-dimensional map with the curl-free property active, all derivative elements that contain two or more z's are unnecessary. For instance, j_{zz} and j_{xzz} do not provide useful information. A single z in the derivative element is still useful as it specifies how the z-component of the magnetic field evolves in a certain direction. However, as soon as a second z is introduced,

it must specify how the z-component, or any derivative of the z-components, changes in the z-direction. This is no longer applicable in two dimensions. Therefore, the number of useful and independent terms in an n^{th}-order approximation can be reduced by the number of n-1-combinations from the set of three $\{x, y, z\}$. The n-1-combinations effectively fix two of the indices of the derivative elements to be z. n-1-combinations with repetition from a set of size $k = 3$ is

$$\binom{n+k-2}{n-1} = \frac{(n+k-2)!}{(n-1)! \cdot (n+k-2-n+1)!} = \frac{(n+1)!}{(n-1)! \cdot 2} = \frac{n(n+1)}{2}. \tag{3-21}$$

Subtracting this equation from the summation in Eq. (3-20) gives rise to the minimum number of magnetometers required to solve an n^{th}-order approximation for the useful derivative elements, where the LLS problem solves for all independent and useful derivative elements.

$$\left\lceil \frac{\sum_{i=0}^{n} \frac{(i+3)(i+2)-i(i+1)}{2}}{3} \right\rceil = \left\lceil \frac{\sum_{i=0}^{n}(2i+3)}{3} \right\rceil. \tag{3-22}$$

Not coincidentally, the number of independent and useful terms in this equation for an n^{th}-order spatial derivative, $2n + 3$, is equal to the number of independent terms in an n^{th}-order spatial derivative for a magnetic field that possesses both the curl- ánd divergence-free property. This illustrates that the divergence-free property exactly constrains the derivative elements that are related to the out-of-plane direction. However, the derivation of this number for a vector field free of curl and divergence is not provided in this thesis. Table 3-1 shows the minimum number of magnetometers required to obtain a full-rank LLS problem for different orders of approximation. If fewer magnetometers are present than the minimum required number, infinitely many solutions are possible, often resulting in a sub-optimal solution. Therefore, the results will later show that it is always preferable to apply an order of approximation to an array that has the specified minimum number of magnetometers.

Table 3-1: Minimum required number of magnetometers to solve an n^{th}-order approximation for all independent and useful derivative elements according to Eq. (3-22)

Order of approx.	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}
Min. magnetometers	3	5	8	12	16

3-2-3 Local GP Method

The final method to estimate the spatial derivative is to make use of GPs. Using the joint distribution of the magnetic field and the spatial derivative, similar to the distribution from Eq. (3-6), but now adapted for a single array, the spatial derivative can be estimated. Instead of adding a spatial derivative for each magnetic field measurement into the joint distribution, only one spatial derivative is added for all measurements on the array. The spatial derivative serves as test data and the magnetometer measurements serve as training data. It is assumed that the noise on the measurements of the magnetic field is zero mean white noise with an independent and identically distributed Gaussian distribution, as stated in Section 2-2. This

joint distribution is given by

$$\begin{bmatrix} \mathbf{y}_1(\mathbf{x}_1) \\ \mathbf{y}_2(\mathbf{x}_2) \\ \vdots \\ \mathbf{y}_N(\mathbf{x}_N) \\ \mathbf{H}_d(\mathbf{x}_0) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, K_{\text{local,d}} \right), \tag{3-23}$$

$$K_{\text{local,d}} = \left[\begin{array}{cccc|c} K_{\text{c}}(\mathbf{x}_1,\mathbf{x}_1)+\Sigma & K_{\text{c}}(\mathbf{x}_1,\mathbf{x}_2) & \dots & K_{\text{c}}(\mathbf{x}_1,\mathbf{x}_N) & K_{\text{c}}(\mathbf{x}_1,\mathbf{x}_0)\mathcal{D}_{\mathbf{x}_0}^{\intercal} \\ K_{\text{c}}(\mathbf{x}_2,\mathbf{x}_1) & K_{\text{c}}(\mathbf{x}_2,\mathbf{x}_2)+\Sigma & \dots & K_{\text{c}}(\mathbf{x}_2,\mathbf{x}_N) & K_{\text{c}}(\mathbf{x}_2,\mathbf{x}_0)\mathcal{D}_{\mathbf{x}_0}^{\intercal} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ K_{\text{c}}(\mathbf{x}_N,\mathbf{x}_1) & K_{\text{c}}(\mathbf{x}_N,\mathbf{x}_2) & \dots & K_{\text{c}}(\mathbf{x}_N,\mathbf{x}_N)+\Sigma & K_{\text{c}}(\mathbf{x}_N,\mathbf{x}_0)\mathcal{D}_{\mathbf{x}_0}^{\intercal} \\ \hline \mathcal{D}_{\mathbf{x}_0}K_{\text{c}}(\mathbf{x}_0,\mathbf{x}_1) & \mathcal{D}_{\mathbf{x}_0}K_{\text{c}}(\mathbf{x}_0,\mathbf{x}_2) & \dots & \mathcal{D}_{\mathbf{x}_0}K_{\text{c}}(\mathbf{x}_0,\mathbf{x}_N) & \mathcal{D}_{\mathbf{x}_0}K_{\text{c}}(\mathbf{x}_0,\mathbf{x}_0)\mathcal{D}_{\mathbf{x}_0}^{\intercal} \end{array}\right]$$

where $\mathbf{x}_i$ denotes the locations of the magnetometers on the array, Σ the diagonal covariance matrix of the white noise with an indepentend and identically distributed (IID) Gaussian distribution, N is the number of magnetometers on the array, $\mathbf{x}_0$ is the location of the magnetometer that is assigned as the middle magnetometer. Although $\mathbf{x}_0$ is the same as one of the $\mathbf{x}_i$'s, it is named differently since it depends on the size of the array which ones are equal. K_{c} is shorthand notation for K_{curl} to save some space. By conditioning the derivative $\mathbf{H}_d(\mathbf{x}_0)$ on all magnetic field measurements $\mathbf{y}_i$ using Eq. (2-11), the posterior distribution of the spatial derivative is obtained. This distribution again provides an estimate of the mean as well as the uncertainty of the estimate in terms of the covariance matrix. Essentially, this method uses the same data as would be used for constructing a local magnetic field map, but instead of making a prediction of the field it predicts the spatial derivative, hence the name local GP method. Like the LLS method, this method also utilizes all magnetometers present on the array.

Ultimately, the spatial derivative predictions are meant to be used as training data for constructing the magnetic field map. In noisy GPR, it is necessary to provide a covariance matrix of the noise in the measurements. Fortunately, this local GP method immediately provides this covariance matrix along with the prediction. In contrast, the other two methods do not offer this directly, so the noise in their estimates must be determined from the data.

This method incorporated the curl-free property through the use of the kernel K_{curl} that is being used. However, since the GP model depends on the hyperparameters, these have to be specified for the method to function properly, adding a layer of complexity. That said, in case the spatial derivatives are to be used for magnetic field mapping using GPR, determining these hyperparameters would be necessary regardless. However, in the case of this thesis, since all data are simulated using certain hyperparameters, which will be specified in Section 3-3, the exact same ones will be applied to predict the spatial derivatives. In other words, the exact same model that was used to generate the data is used to predict the spatial derivatives. This might give this method a slight advantage over the other techniques. This advantage would, however, not carry over to real-world data.

In contrast to the other two methods, no order of approximation has to be specified for this technique. This has the advantage that the most accurate estimate is always obtained, whereas for the other two, an optimal order of approximation must be selected.

3-3 Simulation Setups

In this section, setups for several simulation experiments are explained. However, first, the array configurations that are considered are introduced. Next, we describe the setup for determining the variance of the spatial derivative estimate error for the different derivative estimation methods. Subsequently, the simulation setups used to evaluate the map quality across the various array configurations are outlined. These setups are evaluated for two distinct cases. The first case utilizes all magnetometers on the array to create the map, which will be referred to as the full kernel. The second case uses only the magnetometer in the center of the array in combination with the spatial derivative estimate at that location, which will be referred to as the derivative kernel. If there is no center magnetometer available, the neighboring magnetometer to the top right is used along with the derivative estimate at that location. Both kernels include the curl-free property.

All data is generated by sampling from GP priors. The prior from Eq. (3-6), either with or without the derivative part, is used for this purpose. The hyperparameters were set to fixed values for all simulations, except for the measurement noise. The values $\sigma_{\text{lin.}} = 15\,\mu\text{T}$, $\sigma_f = 5\,\mu\text{T}$ were set to the same values used in [30], and the lengthscale was set to $l = 0.5$. The measurement noise $\sigma_y \in \mathbb{R}$, which acts on each of the axis of the magnetometer identically, differs per simulation setup.

All simulations and other code implementations were carried out in Python. Less computationally demanding tasks were performed on a 2017 HP ZBook Studio G4 laptop (Intel Core i7-7700HQ @ 2.80 GHz, 8 GB RAM), while most of the tasks, particularly the more demanding ones, such as generating the data and making predictions with GPR, were done in the cloud. Several cloud services were used. The cloud service provided by the TU Delft was used, as well as the cloud services of Google and IBM. Throughout the process, virtual machines with different specifications were used. Typically, depending on the specific task, virtual machines with 16 or 8 vCPUs were used with 64 or 32 GB of RAM.

3-3-1 Array Configuration

This thesis focuses mainly on two-dimensional arrays. To further limit the possible configurations, only square arrays are considered, as their symmetrical configuration appears to be beneficial for deriving the spatial derivative. Several of those square, two-dimensional array configurations are studied to determine the effect of three different variables: the spacing of magnetometers, the number of magnetometers, and the noise level of the magnetometers. Spacing refers to the physical distance between adjacent magnetometers on the array. Since it is impractical to plot all these three variables against each other and would require immense amounts of data, the decision was made to compare the spacing against the number of magnetometers and against the noise level. These are probably the most interesting combinations. That is because the optimal spacing to accurately estimate the derivative is not immediately obvious. A spacing that is too small causes the noise to have a large impact on the small difference in the magnetic field being measured. A spacing that is too large, on the other hand, results in a significant change in the function between measurements, making it difficult to determine the derivative. In contrast, it is obvious that more magnetometers and less noise would give rise to better derivative estimates. Hence, plotting the number of magnetometers

against the noise level would likely not give very interesting insights into which configuration provides the most accurate derivatives.

Another important consideration for how well the spatial derivative can be computed with a certain configuration, is how rapidly the field changes. This variable is parameterized by the lengthscale l in GPs and specifies the distance in the input space over which a significant change in the output can be observed. Hence, a small lengthscale results in a rapidly changing field and a large value in a more gradually changing field. For instance, if the field is rapidly changing, a much smaller spacing is desired to ensure the field does not vary too much between the magnetometers; otherwise, the derivative estimate would be compromised. Through simulation, it was confirmed that the ratio between the lengthscale and the spacing determines the relative accuracy with which the spatial derivative can be determined, which is explained in Appendix B in Section B-1. That is, the relative accuracy of the derivative estimate is the same if the lengthscale and the spacing both scale with the same factor. To avoid the need to consider this lengthscale as a separate variable, a new variable is introduced as the lengthscale over the spacing, which will be referred to as l_s. This dimensionless variable ensures that the results can be applied to any lengthscale. The physical interpretation is equal to the average number of magnetometers per lengthscale. A small value means that there are very few magnetometers relative to the lengthscale (larger spacing), whereas a large value means that there are many magnetometers relative to the lengthscale (smaller spacing).

Similarly, the impact of the noise on the quality of the signal is always in relation to the signal variance. To express this relation, the noise-to-signal ratio (NSR) ratio is used. When the noise is small relative to the signal, i.e. a small NSR, a more accurate estimate of the derivative can be obtained than if the NSR is large. The NSR is computed as

$$\text{NSR} = \frac{P_{\text{noise}}}{P_{\text{signal}}}, \tag{3-24}$$

where P_{signal} is the power of the signal and P_{noise} the power of the noise. To compute the power of a signal, one also has to take into account the mean of the signal. However, in the case of magnetic field mapping, since only the anomalies, or deviations from the constant Earth's magnetic field, are of interest, the mean of the signal does not provide any additional information. Neither does it for computing the spatial derivative. Therefore, if the mean was subtracted from the signal, the amount of relevant information would be preserved for this application. As a result, the power of the useful part of the signal can be considered to be equal to the zero-mean equivalent of the signal. For a signal that has a zero mean, the power is equal to the variance. Since the measurement noise is modeled as white noise, the noise signal is a zero-mean signal as well. Therefore, the NSR can be computed as

$$\text{NSR} = \frac{\sigma_y^2}{\sigma_f^2} \tag{3-25}$$

Similarly to the dimensionless variable l_s, the NSR also ensures that the results of this thesis, as long as the NSR match, can be applied to any combination of signal and noise variance. Section B-2 in Appendix B shows that if the NSR stays the same, the derivative can be computed with the same relative accuracy. Note, though, that if the standard deviation of the signal increases, the signal's magnitude scales equally. As a result, the standard deviation of the derivative error also scales by the same factor. However, since both the signal's magnitude

and the standard deviation of the derivative error scale equally, the relative accuracy of the derivative estimation remains the same.

Several array configurations are considered. The variable values that will be considered are given in Table 3-2. In the simulations and result, l_s is compared with the number of magnetometers and the measurement noise.

Table 3-2: Considered Variables Values for Array Configurations

Variable	Values
Spacing [m]	$0.01, 0.013, 0.02, 0.03, 0.05, 0.07, 0.1, 0.12, 0.15, 0.17, 0.2, 0.25, 0.3, 0.357$
l_s [-]	$50, 38.5, 25, 16.7, 10, 7.1, 5, 4.2, 3.3, 2.9, 2.5, 2, 1.7, 1.4$
Number of mags.	$2 \times 2, 3 \times 3, 4 \times 4, 5 \times 5$
σ_y^2 [µT]	$0.1^2, 0.3^2, 0.5^2, 0.7^2, 0.9^2$
NSR [-]	$0.040\%, 0.36\%, 1.0\%, 1.96\%, 3.23\%$

3-3-2 Variance of Spatial Derivative Estimate Error

In order to use the spatial derivative as training data for GPR, the variance of the derivative estimates must also be provided for all the desired array configurations. To achieve this, a simulation experiment was conducted by placing one hundred arrays within a random magnetic field in a grid-like structure. For each of the hundred arrays, the spatial derivative was estimated. Then these derivative estimates were compared to the true derivatives to calculate the variance of the derivative error. This process was repeated 70 times to assess the uncertainty in the variance estimate. By performing this analysis across all array configurations and all derivative estimation methods, the most effective array configuration and method can be identified. Figure 3-1 shows an example of the setup to determine the variance of the derivative estimate. It only shows the magnitude of the magnetic field, but to determine the variance of course the individual components are used.

We generated all data by sampling from the GP priors at the desired locations. This is done by first drawing samples from a standard normal distribution with mean zero and unit variance as the same size as the desired kernel from Eq. (3-6) using the Numpy library in Python. By applying the Cholesky decomposition to the kernel matrix and multiplying the samples from the standard normal distribution with the decomposed L of the kernel, the samples from the kernel are obtained. Before applying the Cholesky decomposition, jitter of size $1e^{-8}$ is added to the kernel to ensure positive definiteness and numerical stability. This process of drawing samples from a kernel matrix K is given by

$$\begin{aligned} Z &\sim \mathcal{N}(\mathbf{0}, \mathbf{I_K}) \\ \mathbf{L}\mathbf{L}^T &= \text{chol}(\mathbf{K} + 1e^{-8}\mathbf{I}) \\ \mathbf{s} = \mathbf{L}\mathbf{Z} &\sim \mathcal{N}(\mathbf{0}, \mathbf{K}), \end{aligned} \tag{3-26}$$

where $\mathbf{Z}$ are the samples from the standard normal distribution, $\mathbf{s}$ are the desired samples and $\mathbf{I_K}$ the identity matrix with the same size as $\mathbf{K}$.

For each of the arrays in Figure 3-1 the spatial derivative is evaluated using one of the proposed derivative estimation methods, where the magnetic field measurements are corrupted with a

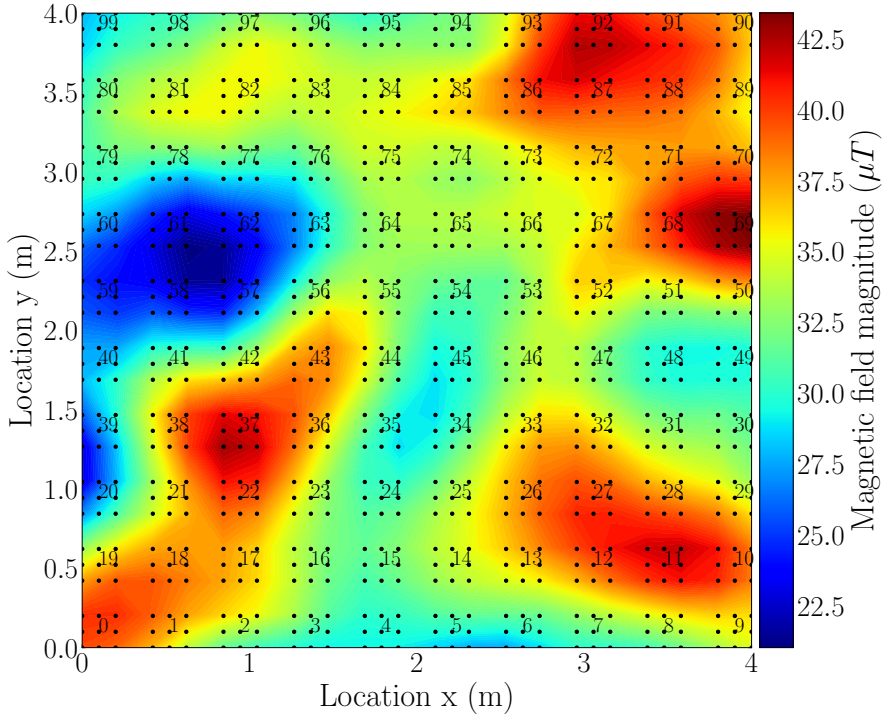

Figure 3-1: Example of the magnitude of a generated magnetic field with one hundred magnetometer arrays spread over the area in a grid. These magnetic field maps are used to estimate the variance of the derivative estimate. Each black dot represents a magnetometer, with each number indicating an array.

certain amount of white noise. These results are then used to determine the variance of the derivative estimate error for that particular map as

$$\sigma_{d,i}^2 = \frac{\sum_{i=1}^{100}(\hat{\mathbf{H}}_{d,i} - \mathbf{H}_{d,i})^2}{100 - 1}, \tag{3-27}$$

where $\sigma_{d,i}$ denotes the standard deviation of the i^{th} component of the spatial derivative, $\mathbf{H}_{d,i}$ the i^{th} component of the spatial derivative and $\hat{\mathbf{H}}_{d,i}$ the corresponding estimate. The standard deviation of all components together in a vector is indicated by $\boldsymbol{\sigma}_d \in \mathbb{R}^9$. By repeating this process for 70 different randomly generated maps, a variance of the variance of the derivative estimate error is obtained, providing an indication of the uncertainty of the variance. In addition to the randomly generated magnetic field, the noise realizations also differ between the 70 iterations.

This procedure is applied to all the desired different array configurations and all derivative estimation methods, giving rise to several different plots. Note, however, that in the case of the local GP method, the variance of the derivative estimate is already provided by the GP. Since the model used to generate the magnetic field is identical to the GP model used to make the derivative estimate, the variance provided by the GP is also identical to if one were to manually determine the variance using Eq. (3-27). This was verified experimentally, as can be observed in Appendix D.

Additionally, a preliminary experiment is also performed on a three-dimensional cubic array to examine how the results from the two-dimensional array might transfer to three-dimensional

arrays. This experiment also evaluates the variance of the derivative estimate error for several numbers of magnetometers and follows the exact same procedure as for the two-dimensional arrays. The experiment is limited to the local GP method since including the variance can be computed analytically. The other methods were excluded because of the large computational demand and time required to manually compute the variance, especially for three-dimensional arrays.

3-3-3 Map Quality: Local Odometry and Global Maps

Once the variance of the derivative estimate error has been determined, the spatial derivative can be used as training data for GPR. Two scenarios are analyzed. First, the derivative and full kernel are applied to a global map with a fixed trajectory. The derivative kernel uses a single measurement at the center of the array along with the spatial derivative, while the full kernel uses all magnetometers on the array. The comparison between the two kernels aims to reveal whether the reduction in computational cost justifies the loss in map accuracy. In the second scenario, both kernels are applied to a local map, where only a single time instant of measurements from an array is used to make predictions in the neighborhood. This process mimics how the magnetic field is used for the model-based odometry method, since the method relies on making predictions in close proximity to the single set of measurements from the array. Comparison of the two kernels should provide insight into whether the derivative kernel might be more suitable for odometry or global mapping purposes. However, since the number of data points for the local map is significantly lower compared to a global map, it is less likely to encounter computational issues. Therefore, it is less common to need a reduction in computational burden, which reduces the number of potential use cases for the derivative kernel for odometry purposes. Only in the case that the update frequency is really high or if little computational power is available could the derivative kernel potentially be useful.

Global Map

In case of the global map, the map quality for the derivative and the entire kernel is evaluated for locations far away and close to the training data. The map will be evaluated for the different array configurations to determine which configurations provide the most accurate and reliable predictions while considering computational cost. For each considered configuration, 70 simulations will be run to account for the variability in the measurement noise. Both the trajectory and the map are fixed between the simulations. Hence, the only difference between the simulations are the realizations of the measurement noise. The trajectory and map were kept constant between the simulations to avoid too much variability and ensure more consistent and comparable results. Not only is it important to analyze whether the predictions are close to the true magnetic field, but also how accurate the uncertainty is. Therefore, the map was selected so that test data that are near and far from training data is considered. Figure 3-2 shows the trajectory and magnetic field map used to evaluate the two kernels. At the first 43 time instances, arrays are present on the global map and serve as training data, and the last 57 points are single points and serve as test data.

The magnetic field values and spatial derivatives in Figure 3-2 are all generated using Eq. (3-26) and Eq. (3-6). To make predictions using the full kernel, the training locations $\mathbf{X}_{\text{all}}$ and

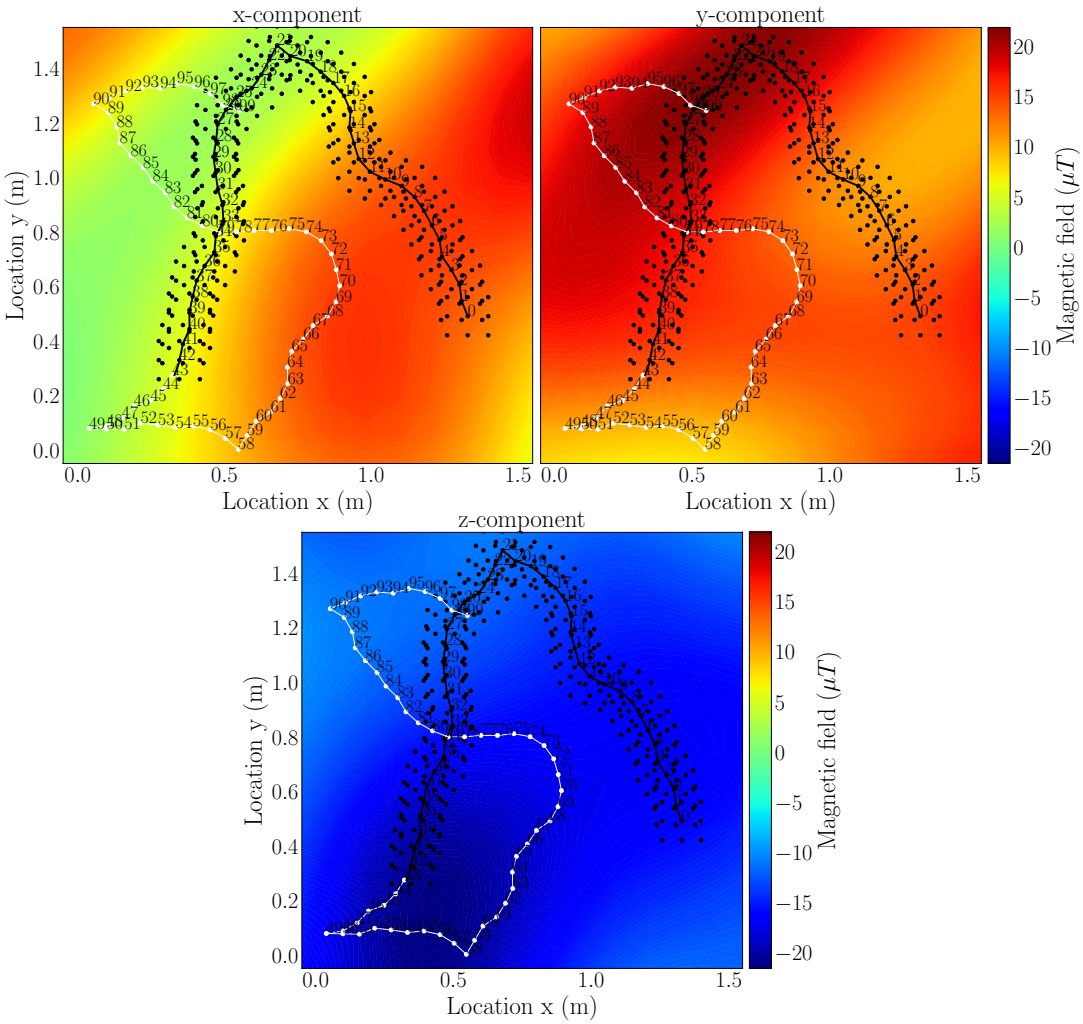

Figure 3-2: The trajectory of the array and the magnetic field used for evaluating the derivative and the full kernel. The black dots indicate the magnetometers used to obtain the training data and the white dots the test data. This particular instance uses an array of three by three and a spacing of seven centimeters, $l_s = 7.1$.

test locations $\mathbf{X}^*$ are introduced as follows

$$\mathbf{X}_{\text{all}} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_{n_a} \end{bmatrix}, \quad \mathbf{X}_i = \begin{bmatrix} \mathbf{x}_{i,1} \\ \vdots \\ \mathbf{x}_{i,n_m} \end{bmatrix}, \quad \mathbf{X}^* = \begin{bmatrix} \mathbf{x}_1^* \\ \vdots \\ \mathbf{x}_{n^*} \end{bmatrix}, \tag{3-28}$$

where $\mathbf{X}_{\text{all}}$ denotes the locations of the magnetometers on all arrays, $\mathbf{X}_i$ the locations of the i^{th} array, n_a the number of arrays, $\mathbf{x}_{i,j}$ the location of the j^{th} magnetometer on the i^{th} array, n_m the number of magnetometers on the array, $\mathbf{x}_i^*$ the i^{th} test data location and n^* the number of training points. The training data of all magnetometers on the array of the

magnetic field and the test data are introduced as

$$\mathbf{Y}_{\text{all}} = \begin{bmatrix} \mathbf{Y}_{\text{all},1} \\ \mathbf{Y}_{\text{all},2} \\ \vdots \\ \mathbf{Y}_{\text{all},n_a} \end{bmatrix}, \quad \mathbf{Y}_{\text{all},i} = \begin{bmatrix} \mathbf{H}(\mathbf{x}_{i,1}) + \boldsymbol{\epsilon}_y \\ \vdots \\ \mathbf{H}(\mathbf{x}_{i,n_m}) + \boldsymbol{\epsilon}_y \end{bmatrix}, \quad \mathbf{F}^* = \begin{bmatrix} \mathbf{H}(\mathbf{x}_1) \\ \vdots \\ \mathbf{H}(\mathbf{x}_{n^*}) \end{bmatrix}, \tag{3-29}$$

where $\mathbf{Y}_{\text{all}}$ denotes the magnetic field measurements of all magnetometers, $\mathbf{Y}_{\text{all},i}$ the magnetic field measurements of the i^{th} array, $\boldsymbol{\epsilon}_y \sim \mathcal{N}(0, \sigma_y^2 \mathbf{I}_3)$ is the magnetometer measurement noise and $\mathbf{F}^*$ the predictions of the magnetic field. The joint distribution of the training and test data is given by

$$\begin{bmatrix} \mathbf{Y}_{\text{all}} \\ \mathbf{F}^* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K_{\text{curl}}(\mathbf{X}_{\text{all}}, \mathbf{X}_{\text{all}}) + \mathbf{I}_{n_a} \otimes \sigma_y^2 \mathbf{I}_3 & K_{\text{curl}}(\mathbf{X}_{\text{all}}, \mathbf{X}^*) \\ K_{\text{curl}}(\mathbf{X}^*, \mathbf{X}_{\text{all}}) & K_{\text{curl}}(\mathbf{X}^*, \mathbf{X}^*) \end{bmatrix}\right). \tag{3-30}$$

The notation of using vectors of locations as inputs to the kernels, rather than single locations, indicates that the kernel is evaluated for all possible pairs of locations between the two arrays, resulting in a matrix where each entry corresponds to a kernel evaluation between a pair of locations.

The predictions of the magnetic field at the test locations are now determined by substituting the kernels from the joint distribution into the posterior distribution from Eq. (2-11). This provides the posterior distribution of $\mathbf{F}^*$ conditioned on $\mathbf{Y}_{\text{all}}$.

To make predictions using the derivative kernel, the spatial derivative for each array is computed using the noisy measurements. These will be included in the training data. The test locations $\mathbf{X}^*$ are equivalent as for the full kernel, but the training locations $\mathbf{X}_{\text{der}}$ are given by

$$\mathbf{X}_{\text{der}} = \begin{bmatrix} \mathbf{x}_{1,0} \\ \vdots \\ \mathbf{x}_{n_a,0} \end{bmatrix}, \tag{3-31}$$

where $\mathbf{x}_{i,0}$ denotes the magnetometer that is assigned as the middle one on the i^{th} array. The training data including the magnetic field as well as the corresponding spatial derivative are introduced as

$$\mathbf{Y}_{\text{der}} = \begin{bmatrix} \mathbf{H}(\mathbf{x}_{1,0}) + \boldsymbol{\epsilon}_y \\ \mathbf{H}_d(\mathbf{x}_{1,0}) + \boldsymbol{\epsilon}_d \\ \mathbf{H}(\mathbf{x}_{2,0}) + \boldsymbol{\epsilon}_y \\ \mathbf{H}_d(\mathbf{x}_{2,0}) + \boldsymbol{\epsilon}_d \\ \vdots \\ \mathbf{H}(\mathbf{x}_{n_a,0}) + \boldsymbol{\epsilon}_y \\ \mathbf{H}_d(\mathbf{x}_{n_a,0}) + \boldsymbol{\epsilon}_d \end{bmatrix}, \tag{3-32}$$

where $\mathbf{Y}_{\text{der}}$ denotes the measurement data including the magnetic field and the spatial derivative at the middle magnetometer of each array and $\boldsymbol{\epsilon}_d \sim \mathcal{N}(0, \boldsymbol{\sigma}_d^2)$ the noise on the spatial derivative estimates. The joint distribution of the training and test data is given by

$$\begin{bmatrix} \mathbf{Y}_{\text{der}} \\ \mathbf{F}^* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K_{\text{curl,d}}(\mathbf{X}_{\text{der}}, \mathbf{X}_{\text{der}}) + \mathbf{I}_{n_a} \otimes \boldsymbol{\sigma}_{y,d}^2 & K_{\text{curl,d}}^{\text{DK}}(\mathbf{X}_{\text{der}}, \mathbf{X}^*) \\ K_{\text{curl,d}}^{\text{KD}}(\mathbf{X}^*, \mathbf{X}_{\text{der}}) & K_{\text{curl}}(\mathbf{X}^*, \mathbf{X}^*) \end{bmatrix}\right), \tag{3-33}$$

where $\boldsymbol{\sigma}^2_{y,d} = \text{diag}(\sigma_y^2\mathbf{I}_3, \boldsymbol{\sigma}_d^2)$, K_{curl} is defined in Eq. (2-23), $K_{\text{curl,d}}$ in Eq. (3-6), and $K^{\text{DK}}_{\text{curl,d}}$ and $K^{\text{KD}}_{\text{curl,d}}$ in Eq. (3-8). Equivalent to the full kernel, the predictions for the derivative kernel are determined using Eq. (2-11) to compute the posterior distribution of $\mathbf{F}^*$ conditioned on $\mathbf{Y}_{\text{der}}$.

Local Odometry Map

In addition to comparing the derivative kernel with the full kernel for the global map, the comparison will also be made for a local map. Whereas the focus of the global map is more on evaluating the map quality for test points that are further away from the training data, the local map solely analyzes test points that are close or even within the range of the array. This is a better indicator of odometry performance. The following simulation experiment is designed with the model-based odometry approach in mind to evaluate how well the different configurations and kernels are capable of making predictions that are relevant for this kind of odometry approach.

Equivalent to the global map, the local map will be evaluated for the different array configurations. For each configuration, multiple simulations will be run to obtain an average map quality. In contrast to the global map, though, not only will the measurement noise vary between simulations, but the map will as well. This approach ensures that the results are not biased towards a specific spatial derivative. In the case of the global map, many different spatial derivatives are considered because many array locations are included in a single simulation, which makes it reasonable to use only a single map. However, the local map focuses on one array with only a single spatial derivative. This could give rise to biased results because certain configurations might be better suited for that particular derivative. Hence, the decision was made to not fix the map for all simulations. Given the increased variability between simulations compared to the global map, a larger number of simulations per configuration is desired. Therefore, 200 simulations are conducted per configuration instead of 70.

This experiment is inspired by at what locations the model-based odometry method, which was introduced in Section 2-5-2, has to evaluate the model [29, 30]. This method creates a local magnetic field map based solely on the measurement of all magnetometers on the array at the current time step. Then at the next time step it compares the next measurements to the local map and finds the pose change that is most likely. In this process, it evaluates the local map several times at the location estimates of all magnetometers on the array as they progress to shift towards the final estimate. Therefore, this experiment will use a model based on a single set of measurements from an array to evaluate the magnetometer locations of a shifted version of this array. To obtain sufficient test data, six shifted versions of the array in different directions are considered per simulation.

Since the quality of the local map is heavily influenced by the distance between the center of the local map and the prediction locations, several distances will be analyzed separately. From here on, those distances will be referred to as the step size, in line with the meaning of this distance when the local map is used for odometry. Similarly to how the spacing between the magnetometers was made dimensionless by dividing the lengthscale by the spacing, the step size is also made dimensionless. This new variable is introduced as

$$l_{\text{ss}} = \frac{l}{\text{step size}}. \tag{3-34}$$

One can imagine that if the magnetic field changes rapidly in value, a smaller step size should provide similar results as when the field changes less quickly and a larger step size is used. By making the variable in this way dimensionless, the effect of the lengthscale can be neglected. The physical meaning of l_{lss} is the number of steps that fit within a lengthscale. Table 3-3 shows the step sizes that will be considered in the experiment.

Table 3-3: Considered Step Sizes for Local Map Quality

Variable	Values
Step size [m]	$0.01, 0.02, 0.05, 0.075, 0.1, 0.125$
l_{ss} [-]	$50, 25, 10, 6.7, 5$

Figure 3-3 shows an example of a local map used for this experiment. The three components of the magnetic field are shown in the three plots. A single three-by-three magnetometer array is located at the center of the map, indicated by the black dots, which specifies the training data. The white dots indicate the test data and specify five shifted versions of the array in five different directions with a distance equal to the step size. The five test arrays are evenly spaced along the circle around the original array. To ensure that the exact same test locations are not used for all simulations, the positions of the test arrays on the circle differ per simulation, while keeping the spacing constant.

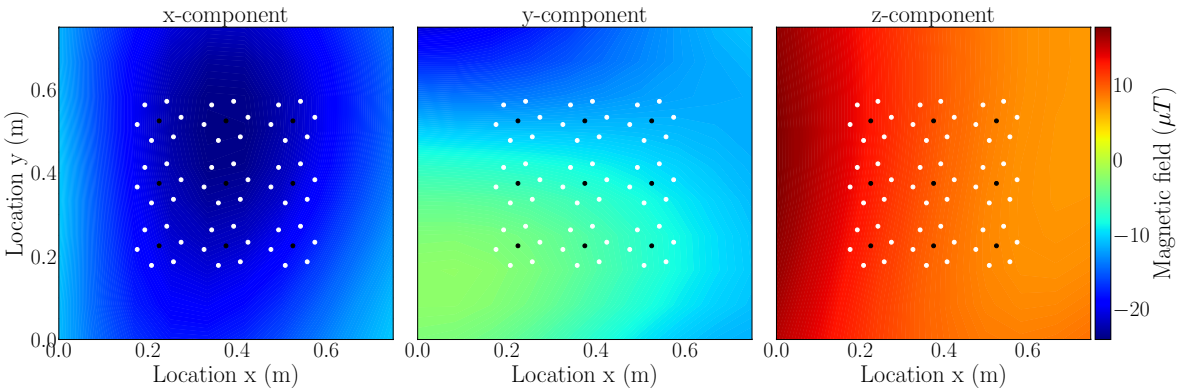

Figure 3-3: An example of a simulation used for the local map. The black dots indicate the array used as training data and the white dots the test data. This particular simulation uses an array of three by three with a spacing of 15 centimeters, $l_{\text{s}} = 3.3$. The test points specify shifted version of the array in different directions with a distance equal to the step size. The step size in this particular simulation is equal to five centimeters, $l_{\text{ss}} = 10$

The procedure for making predictions for the local map is almost identical to that of the global map. The only difference lies in the number of training and test data points. Therefore, only the locations and magnetic field vectors are briefly introduced here. Using the full kernel, the training locations $\mathbf{X}_{\text{all}}$ and test locations $\mathbf{X}^*$ for the local map are introduced as

$$\mathbf{X}_{\text{all}} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{n_m} \end{bmatrix}, \quad \mathbf{X}^* = \begin{bmatrix} \mathbf{X}_1^* \\ \vdots \\ \mathbf{X}_{n_a^*} \end{bmatrix}, \quad \mathbf{X}_i^* = \begin{bmatrix} \mathbf{x}_{i,1}^* \\ \vdots \\ \mathbf{x}_{i,n_m}^* \end{bmatrix} \tag{3-35}$$

where n_a^* is the number of test arrays, which in this experiment is fixed to five, and where $\mathbf{x}_{i,j}^*$ is the location of the j^{th} magnetometer on the i^{th} test array. The training data of all

magnetometers on the array of the magnetic field and the test data are introduced as

$$\mathbf{Y}_{\text{all}} = \begin{bmatrix} \mathbf{H}(\mathbf{x}_1) + \boldsymbol{\epsilon}_y \\ \vdots \\ \mathbf{H}(\mathbf{x}_{n_m}) + \boldsymbol{\epsilon}_y \end{bmatrix}, \quad \mathbf{F}^* = \begin{bmatrix} \mathbf{F}_1^* \\ \vdots \\ \mathbf{F}_{n_a^*}^* \end{bmatrix}, \quad \mathbf{F}_i^* = \begin{bmatrix} \mathbf{H}(\mathbf{x}_{i,1}^*) \\ \vdots \\ \mathbf{H}(\mathbf{x}_{i,n_m}^*) \end{bmatrix}, \tag{3-36}$$

where $\mathbf{x}_i$ denotes the i^{th} magnetometer on the array.

The joint distribution of the training and test data is equivalent to the one given for the global map in Eq. (3-30), but using Eq. (3-35) and Eq. (3-36). The predictions of the magnetic field at the test locations are now again determined by substituting the kernels from the joint distribution into the posterior distribution from Eq. (2-11).

For the derivative kernel the training locations $\mathbf{X}_{\text{der}}$ and the training data $\mathbf{Y}_{\text{der}}$ are given by

$$\mathbf{X}_{\text{der}} = \begin{bmatrix} \mathbf{x}_0 \end{bmatrix}, \quad \mathbf{Y}_{\text{der}} = \begin{bmatrix} \mathbf{H}(\mathbf{x}_0) + \boldsymbol{\epsilon}_y \\ \mathbf{H}_d(\mathbf{x}_0) + \boldsymbol{\epsilon}_d \end{bmatrix}, \tag{3-37}$$

where $\mathbf{x}_0$ denotes the middle magnetometer on the array. The joint distribution, required to compute the prediction at the test locations, is equivalent to the one used for the global map in Eq. (3-32), but uses the training locations and data defined here.

Chapter 4

Results

In this chapter, the results of the experiments introduced in Section 3-3 are presented. First, Section 4-1 provides the results on the variance of the derivative estimate error for the different array configurations and derivative estimation methods. These results will show which derivative estimation method provides the most accurate derivative estimate and for which array configuration they are obtained, addressing the first research sub-question and part of the second. This needs to be addressed, as it is crucial to use an accurate spatial derivative in the derivative kernel before examining whether it can serve as an approximation to the full kernel. In addition, results on the computational efficiency of the derivative estimation methods are discussed, addressing the last part of the second research sub-question. Section 4-2 subsequently presents the results of the experiment assessing the map quality for the derivative and the full kernel, evaluated for both the global and local odometry map. The derivative and the kernel will be compared to each other in terms of mapping quality and computation time to examine whether the derivative kernel is a viable solution to alleviate part of the computational burden. This answers both parts of the third research sub-question.

The results on the spacing are measured in the dimensionless variable l_s. However, to more clearly link these results to reality, some of the results will be translated back to their original variable by using typical values for the lengthscale. Typical lengthscale values depend to a large extent on characteristics of the building and on how close the magnetic field is measured to the floor. Closer measurements result in greater fluctuations in the field and a smaller lengthscale, and vice versa. In the literature, there seems to be a distinction between studies that use smaller and larger lengthscales, for instances in [30, 26] smaller lengthscales are used and in [19, 38] larger ones. Therefore, the main results on l_s will also be translated into spacing for both scenarios. For the smaller lengthscales, a representative value of $l_0 = 0.15$ is used, which is used in [30]. For the larger lengthscales, a value of $l_1 = 1.2$ is used, which is in the range of values used in [19, 38]. To quickly express an l_s value in terms of spacing for l_0 and l_1, the following notation will be used: $l_s = 1(l_0 : 0.15\,\text{m},\ l_1 : 1.2\,\text{m})$, where in this example l_0 gives rise to a spacing of 0.15 m and l_1 to a spacing of 1.2 m.

Similarly, to more clearly link the results for the measurement noise to real-world scenarios, we will focus specifically on an noise-to-signal ratio (NSR) of 1%. This is also, for example,

the NSR used in [37]. To the best of our knowledge, this is a typical NSR value. In this thesis, an NSR of 1% translates to a measurement noise of $\sigma_y = 0.5\,\mu\text{T}$.

4-1 Spatial Derivative Estimates

Several experiments were conducted that considered different array configurations and derivative estimation methods to evaluate the accuracy of the spatial derivative estimate. This section will first examine and analyze the result from these experiments. The setups of these simulation experiments were presented in Section 3-3-2. The computational efficiency of the different derivative estimation methods will then be discussed.

4-1-1 Variance of Spatial Derivative Estimate Error

This section focuses on with which accuracy the spatial derivative estimate can be determined for the three derivative estimation methods that were introduced in Section 3-2. These methods will be analyzed and compared with each other for the various array configurations. In practice, it is possible to construct a magnetometer array of any desired size and shape. These different configurations heavily influence the accuracy of the derivative estimate. A higher accuracy of the derivative estimate is beneficial for the derivative kernel and yields a higher quality map. This shows the importance of the array configuration. In Section 3-3-2 the simulation experiment that is used to determine the variance of the derivative estimate error is introduced. This experiment simulates random magnetic fields, and derivatives are estimated for magnetometer arrays placed in these fields. The variance of the estimation error is computed by comparing the true derivative to the estimates. This approach identifies the accuracy of the derivative estimation methods for various configurations. As the three different methods yield similar results for the various configurations, not all results are presented for each method to prevent repetition. Since finite difference method will turn out to provide less accurate derivative estimates and the local Gaussian process (GP) method has more uncertainty regarding how well the results translate to real-world data, the main focus will be on the linear least squares (LLS) method.

The first-order spatial derivative consists of nine terms, or six independent terms when the curl-free property is applied. Five of these terms are important when the derivative is used for creating a two-dimensional magnetic field map, as the zz-component plays no role. Due to the curl-free property, different components have different variances. For instance, the xy/yx-component generally has a lower variance than the xx- and yy-components, since it has access to more data. Specifically, it can use how the x-component changes in the y-direction, as well as how the y-component changes in the x-direction. This gives rise to a more accurate estimate for the xy/yx-components. Furthermore, the variance of the zz-components is nearly constant across all variables, as there is no information available about magnetic field changes in the z-direction. As a consequence, the derivative estimates always yield zero, resulting in a variance of approximately $300\,\mu\text{T}^2$ for all variables. Therefore, estimates for any derivative component with a variance greater than $300\,\mu\text{T}^2$ are even less accurate than always estimating zero. However, in the main sections of this thesis, instead of showing the variance of the derivative estimation error for all or one of the components of the derivative, a combined term is introduced. This combined term is the mean of the variances of the six

relevant components for creating a two-dimensional map: the xx-, xy-, yx-, yy-, zx- and zy-components. This saves a significant amount of space and has the advantage that it better represents how well the spatial derivative can be used for creating a map as it considers all relevant components. The confidence interval/standard deviation of this new mean variance term is taken as the mean of the confidence intervals/standard deviations of the individual components. Figures showing the variances for all the individual components will be included in Appendix C.

Variance Using the Linear Least Squares Method

This section focuses on the LLS method, which is described in Section 3-2-2. First, the results on the effect of the number of magnetometers on the array are presented, followed by the results on the effect of the measurement noise. Figure 4-1 shows the variance of the combined components of the spatial derivative error when the LLS method is used. The measurement noise was set to $\sigma_y = 0.5\,\mu\text{T}$, which equals a NSR of 0.01.

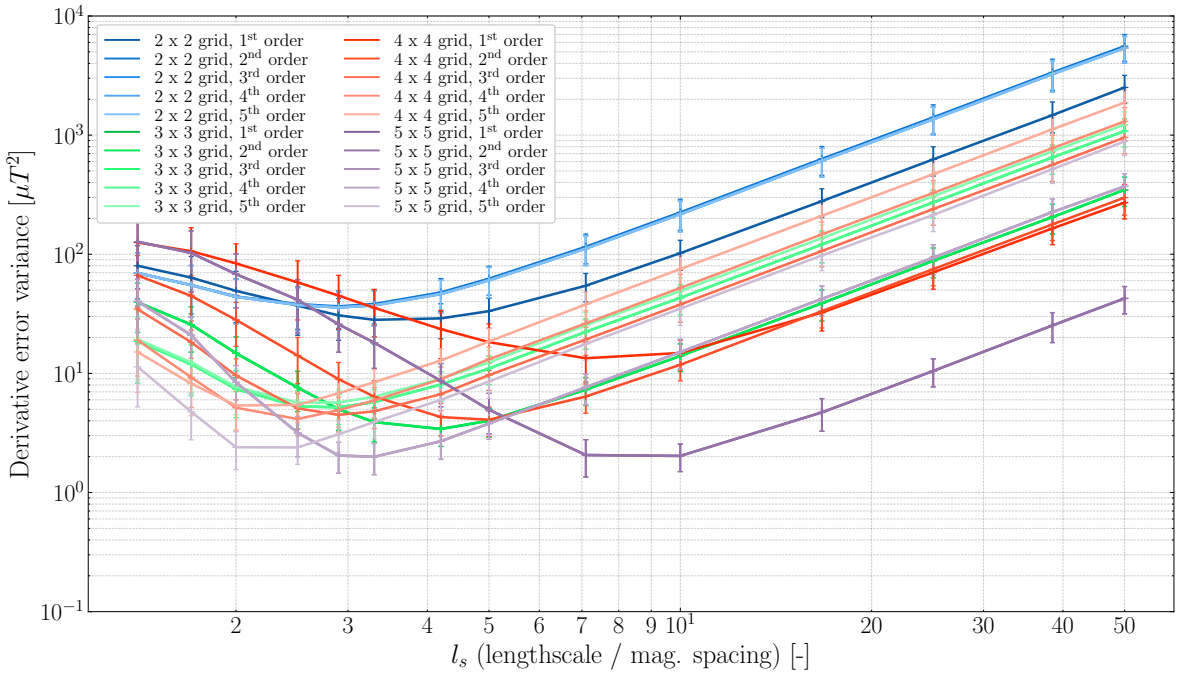

Figure 4-1: Relationship between the mean variance of the combined components of the spatial derivative estimate error and l_s using the LLS method. The relationship is plotted for different number of magnetometers on the array and the different order of approximations of the LLS method. The color of each bar denotes the number of magnetometers on the array, and the shade of the color the order of approximation of the LLS method. The bars on top of the lines indicate the 95% confidence interval. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%).

When first inspecting the general trend of this figure, it is clear that neither small nor large values of l_s give rise to the best estimate with the lowest error variance. For spacings $l_s > 10$ the variance continues to increase, and the same holds true for $l_s < 2.5$. When l_s is large, that is, the spacing is small relative to the lengthscale, the noise negatively influences the

estimate of the spatial derivative. This is because, when the spacing is small relative to the lengthscale, the difference in magnetic field measured by the magnetometers, which is used to compute the derivative, is small compared to the measurement noise. As a result, the noise has a relatively large impact on this small difference in magnetic field, preventing an accurate computation of the spatial derivative. However, very small values of l_s, which means a large spacing relative to the lengthscale, do not provide the most accurate estimates either. In this case, the magnetometers are spaced too far apart, meaning that the measured differences in the magnetic field are no longer representative of the local derivative. Instead, they serve more as separate measurements with less correlation with each other. The optimal spacings for all the different magnetometers on the array lie in a range from $l_s = 2.5$ ($l_0 : 0.06\,\text{m},\ l_1 : 0.48\,\text{m}$) to $l_s = 10$ ($l_0 : 0.015\,\text{m},\ l_1 : 0.12\,\text{m}$). The lowest overall variance is found with a five-by-five array at $l_s = 3.3$ ($l_0 : 0.045\,\text{m},\ l_1 : 0.33\,\text{m}$) with a variance of $2\,\mu\text{T}^2$. However, variances at $l_s = 2.9$, $l_s = 7.1$ and $l_s = 10$ are very similar.

When examining the effect of the order of the LLS method, several points stand out. Higher-order approximations provide better estimates when more magnetometers are present and l_s is on the lower end. For instance, if the number of magnetometers is only two by two, there are insufficient measurements to accurately evaluate the higher-order spatial derivatives, as can be observed from Table 3-1. As a result, the higher-orders provide almost identical results compared to the lower-orders. However, if more magnetometers are available on the array, the higher-order terms can be uniquely determined and start to provide more accurate results of the first-order derivative. Since now a higher-order model can be fitted, which better represents the local behavior of the function, the accuracy of the first-order spatial derivative estimate also improves. However, the degree to which higher-order approximations improve or even deteriorate accuracy also depends on l_s. As can be observed from the figure, if l_s is small (approximately $l_s < 5$), the higher-orders do improve the accuracy in case there are sufficient measurements. But if l_s grows (approximately $l_s > 5$), the higher-order approximations start performing worse than the lower-orders. This can be attributed to two factors. First, since the spacing is small and the measurement of the magnetic field values are quite close to each other, the evolution of the field can already be approximated rather well with lower-order approximations. Second, since the relative impact of the noise increases as the spacing becomes smaller, higher-order approximations start to overfit to the noise as they are more complex models. The lower-orders suffer less from overfitting, and hence perform relatively well compared to the higher-orders. Figure C-1 in Appendix C shows the variances of the derivative estimates for all relevant components of the Jacobian.

By reducing the data in Figure 4-1 so that only the order of approximation is shown that provides the best result, Figure 4-2 is obtained, where the annotations denote the order of approximation used for that particular data point. This provides a clearer line graph to compare the number of magnetometers.

The optimal order of approximations in Figure 4-1 are in line with the minimum required number of magnetometers from Table 3-1. Almost none of the data points indicate an optimal order of approximation that is higher than the corresponding number of magnetometers on the array would allow for, based on Table 3-1. Only the first three l_s values of the two-by-two array indicate a second-order approximation, while five magnetometers are required. However, if we examine the first three l_s values for the blue lines in Figure 4-1, we can observe that the higher-order approximations yield virtually the same variance as the first-order. Overall, this suggests that the theoretical results from Table 3-1 indeed apply to this particular setup.

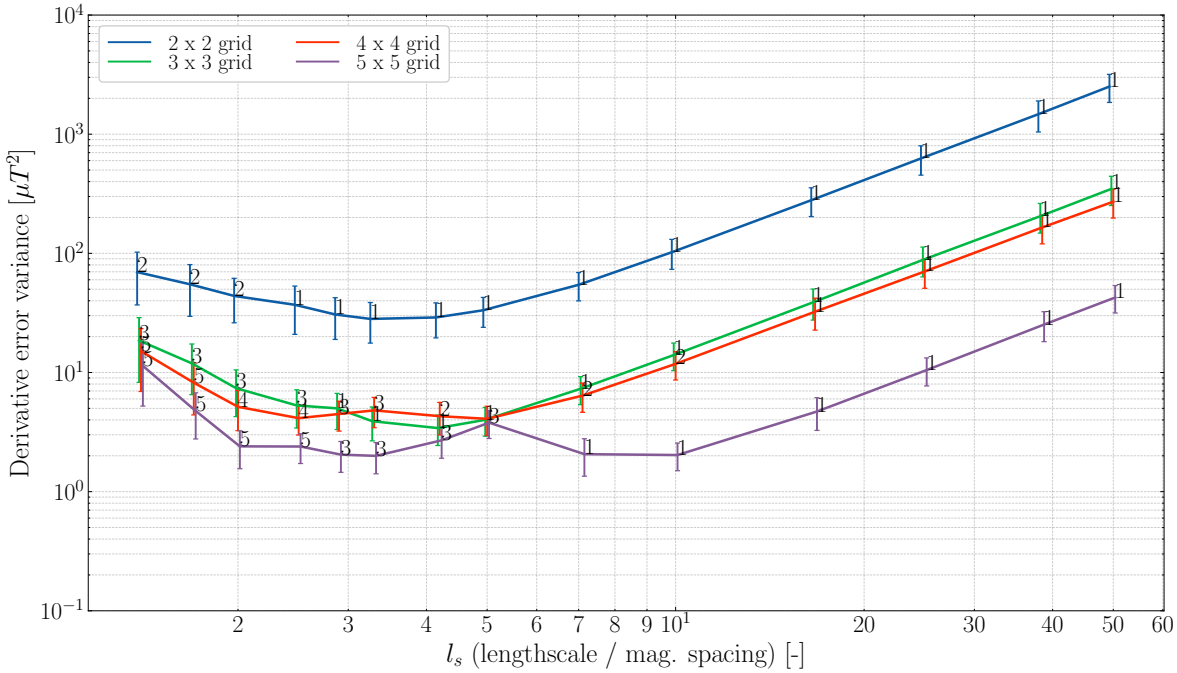

Figure 4-2: Relationship between the mean variance of the combined of the spatial derivative estimate error and l_s using LLS method with the order of approximation that gave rise to the lowest variance. The relationship is plotted for different number of magnetometers on the array. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%). To ensure that the confidence intervals do not overlap, some of the lines are shifted to the right, but they do still correspond to the same l_s

Figure 4-2 shows that larger grids provide better estimates of the spatial derivative. Especially the grids that have a magnetometer at the center of the array, so the three-by-three and five-by-five grids, perform relatively well compared to the others. This is probably due to the symmetry that allows for more effective utilization of the magnetometers to compute the derivative. Additionally, arrays with more magnetometers generally yield better estimates. Particularly at smaller spacings of $l_s \geq 7.1$, using more magnetometers yields notably more accurate estimates, despite the fact that the same order is used across the different arrays in this region. This can be attributed to arrays with more magnetometers being better at averaging out the noise, which plays a more prominent role at smaller spacings. Apart from the fact that more magnetometers provide more accurate estimates, the range for which accurate estimates are obtained is significantly larger as well. This can again partially be attributed to the fact that more magnetometers are better at averaging out the noise, but also that it can use larger order of approximations at smaller spacings. Figure C-2 in Appendix C shows the mean variance for the individual components for the order that gave rise to the lowest variance.

Figure 4-3 shows the mean variance of the combined components of the spatial derivative error for different measurement noise values when the LLS method is used. The number of magnetometers was set to five by five. The noise value 0.1 µT, 0.3 µT, 0.5 µT, 0.7 µT and 0.9 µT correspond to the following NSR values 0.040%, 0.36%, 1.0%, 1.96% and 3.23%, respectively.

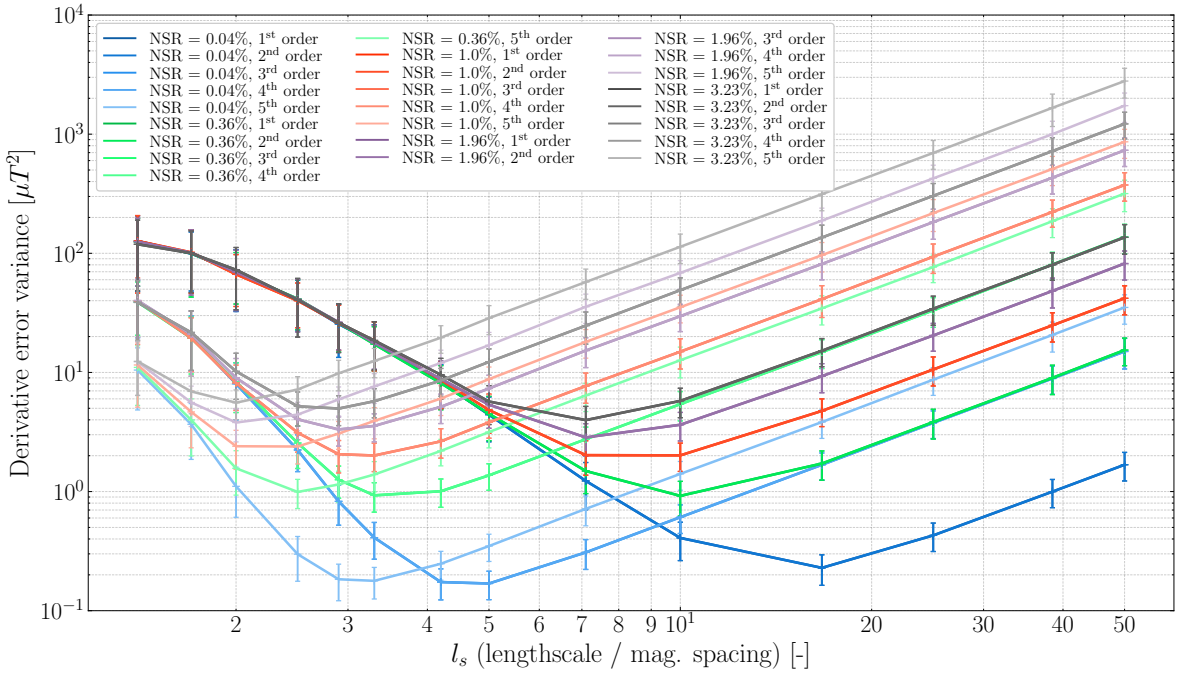

Figure 4-3: Relationship between the variance of the xx-component of the spatial derivative estimate error and l_s using the LLS method. The relationship is plotted for different measurement noise level and different order of approximations of the LLS method. The color of each bar in the figure denotes the measurement noise, and the shade of the color the order of approximation of the LLS method. The bars on top of the lines indicate the 95% confidence interval. The number of magnetometers for all configurations was set to five by five.

While some similar observations can be drawn from Figure 4-3, it also provides new insights. Specifically, the plot verifies that the noise has a relatively large impact on the derivative accuracy at smaller spacings. For instance, at $l_s = 1.4$ a larger noise value does not influence the variance much, but at $l_s = 50$ an increased noise level significantly deteriorates the derivative estimate. Furthermore, the plot also illustrates that higher-order approximations suffer more from noisy measurements than lower-order approximations. This can be clearly observed at intermediate values of l_s, for instance, at $l_s = 4.2$. There, the blue lines show that for small noise values higher-order approximations provide significantly better estimates, whereas if the noise increases, see, for instance, the black lines, the higher- and lower-orders perform equally well. In fact, the lower-order approximations are not much influenced by noise. Only if l_s increases, the impact of the noise becomes apparent for lower-order approximations. Ultimately, the optimal order of approximation varies not only with the spacing and number of magnetometers, but also with the measurement noise.

Another interesting observation is that the first- and second-order, as well as the third- and fourth-order approximations, yield the exact same variance. This is also evident in Figure 4-1 for array configurations with a magnetometer in the center of the array. The reason that two consecutive orders provide the exact same derivative estimate in the case of a symmetric array configuration, relates to a phenomenon that was described in Section 3-2-1. There, it was shown that using the Taylor series expansion with a symmetric setup, the even-order spatial derivatives dropped out of the approximation. Since the LLS method is also based on

the Taylor series expansion, applying a first-order LLS method means the second-order does not contribute in (3-16). Therefore, every odd-order LLS method produces the same spatial derivative estimate as the subsequent even-order. As a result, in the case of a symmetric array, the extra computational power required for an even-order LLS method is never justified when compared to the preceding odd-order. Figure C-3 in Appendix C shows the variance for all individual components.

By reducing the data in Figure 4-3 so that the values of the orders that resulted in the lowest variance are shown, Figure 4-4 is obtained. The two local minima in Figure 4-4 that are present for all noise levels are caused by the fact that the different orders yield the best results for different spacings. The figure also clearly illustrates the effect of the measurement noise on the optimal l_s. As the noise level decreases, the optimal l_s value shifts toward larger values, indicating that less noise gives rise to a tighter optimal spacing. This is because, in the absence of noise and with perfect measurement, a spacing that goes to zero in the limit would produce the best derivative estimate. Therefore, by decreasing the noise and its impact, the optimal spacing becomes tighter. Figure C-4 in Appendix C shows the variance for all individual components.

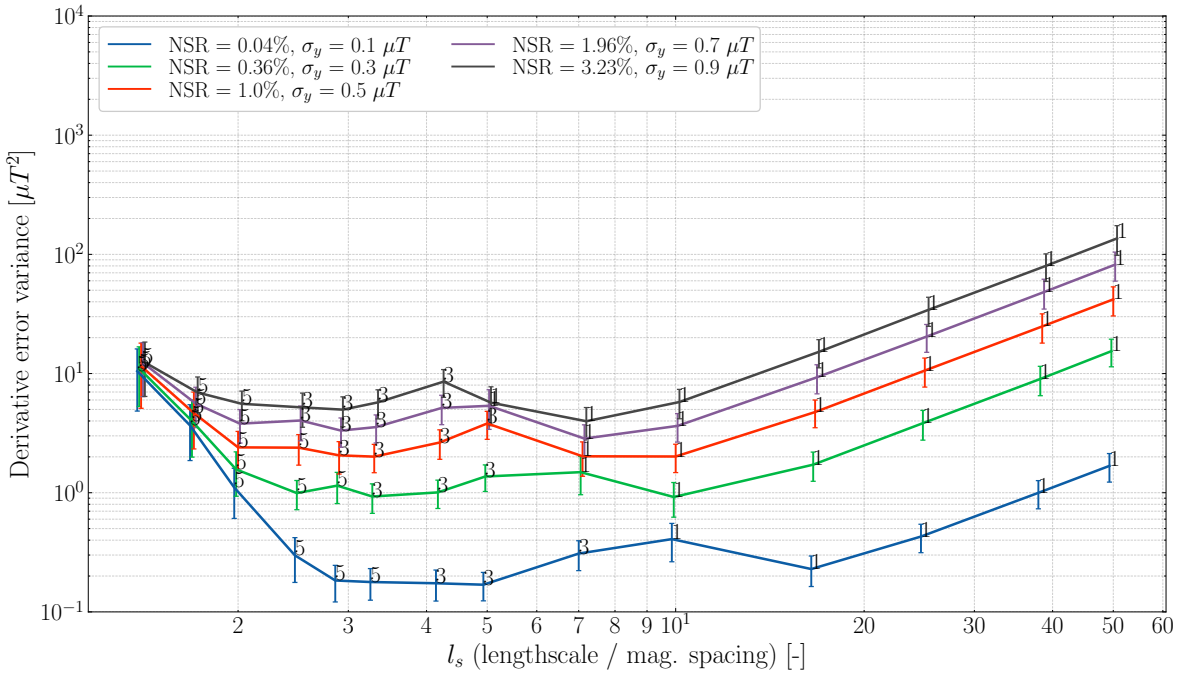

Figure 4-4: Relationship between the mean variance of the combined components of the spatial derivative estimate error and l_s using the LLS method with the order of approximation that gave rise to the lowest variance. The number of magnetometers for all configurations was set to five by five.

Overall, the results of the LLS method indicate that arrays with the highest number of magnetometers and the lowest noise levels can provide the most accurate derivative estimates. Hence, to get the most accurate derivative estimate, one should use as many magnetometers with the lowest possible noise. However, in practice, there is always a trade-off between performance and costs. Symmetric arrays with a magnetometer in the center perform relatively well. These arrays with up to five by five magnetometers achieve their lowest variance at an

optimal spacing that generally lies in a wide range from $l_s = 2.5$ ($l_0 : 0.06\,\text{m},\ l_1 : 0.48\,\text{m}$) to $l_s = 10$ ($l_0 : 0.015\,\text{m},\ l_1 : 0.12\,\text{m}$). The exact optimal spacings depend on the number of magnetometers and the level of noise: increasing the number of magnetometers or decreasing the noise level shifts the optimal spacing towards tighter spacings and vice versa. To select a suitable spacing given a certain number of magnetometers, Figure 4-1 and Figure 4-2 should be consulted. If the measurement noise of the magnetometers deviates from the NSR=1% used in those figures, the spacing can be adjusted according to Figure 4-3 and Figure 4-4. When doing so, however, it should be taken into account that those figures are based on arrays with five by five magnetometers. Arrays with fewer magnetometers are more sensitive to noise, and the impact of the noise on the optimal spacing is more pronounced. Selecting the optimal spacing in practice is challenging, as these results do not perfectly translate to real-world conditions. Real-world magnetic field measurements are likely to be noisier due to sensor errors [18], for instances due to biases. In addition, assumptions regarding the curl-free property and known distances between magnetometers might not hold entirely in practice, introducing additional noise. Therefore, it is advisable to select a spacing closer to the larger spacings of the optimal range, as more noise tends to shift the optimal spacings toward that direction.

When selecting an order of approximation to use for a certain chosen array, we again suggest to initially base it on the number of magnetometers in Figure 4-1 and Figure 4-2. If the NSR deviates significantly from the NSR=1% used in those figures, any possible alterations can be made again based on Figure 4-3 and Figure 4-4.

As a general recommendation, a symmetric array of five by five with a spacing of $l_s = 2.9$ ($l_0 : 0.052\,\text{m},\ l_1 : 0.41\,\text{m}$) is a reliable choice that performs well for a wide range of noise levels using the LLS method.

Variance Using the Finite Difference Method

Another commonly used method for computing derivatives is the finite difference method, described in Section 3-2-1. In this thesis, due to its widespread use, it serves mainly as the baseline for comparing the other methods. Figure 4-5 shows the mean variance of the combined components of the derivative estimate error using the finite difference method for different number of magnetometers. This figure shows that the optimal spacing for the finite difference method is located at $l_s = 2$ ($l_0 : 0.075\,\text{m},\ l_1 : 0.6\,\text{m}$) with a grid of five by five magnetometers and a fourth-order approximation. However, the difference with a second-order central difference scheme at an l_s of 2.9 is rather small. Interestingly, for most values of l_s the second-order central difference scheme performs better than higher-orders, which is again due to the fact that higher-order approximations tend to overfit to the noise if l_s is large.

Compared to the LLS method in Figure 4-2, the finite difference method performs worse for practically all spacings. Only for $l_s < 1.7$ is the variance quite similar. Moreover, the LLS method performs remarkably better, especially at small spacings. This can be explained by the fact that measurement noise plays a large role for small spacings and the LLS method uses all magnetometers on the array and is hence better at averaging out the noise. The finite difference method only uses magnetometers that form a cross on the array. This also explains why, especially for arrays with many magnetometers, the LLS method performs

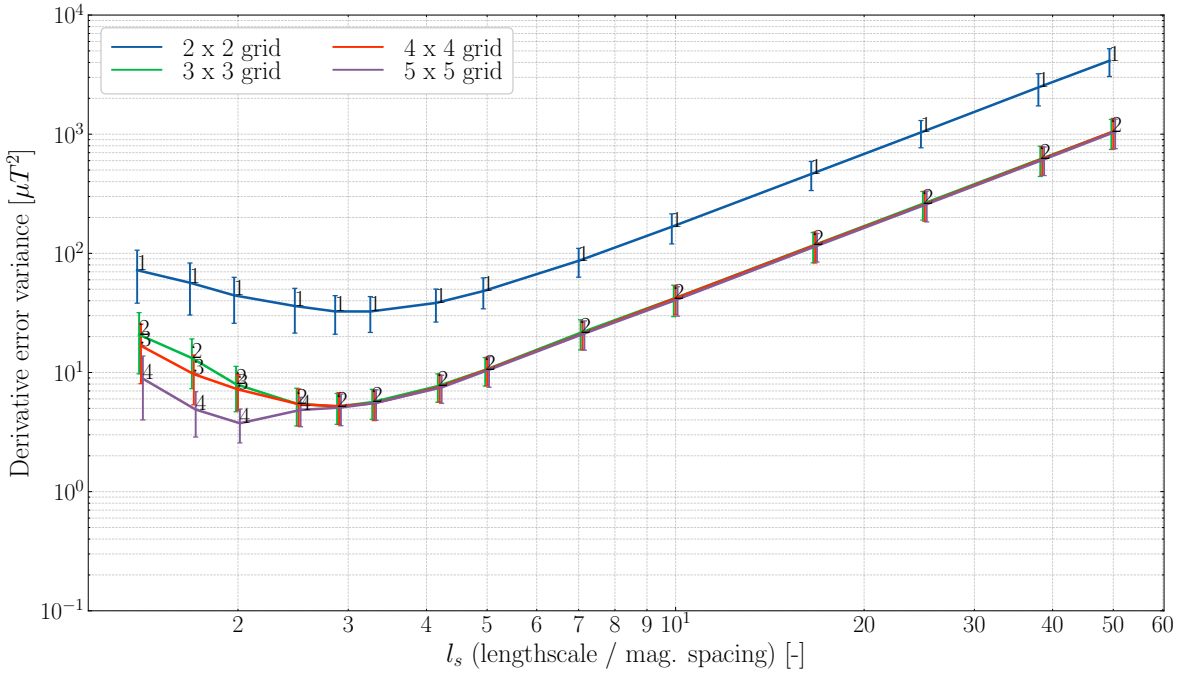

Figure 4-5: Relationship between the mean variance of the combined components of the spatial derivative estimate error and l_s using the finite difference method with the order of approximation that gave rise to the lowest variance. The annotations denote the order of approximation that resulted in the lowest variance. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%).

better compared to the finite difference method. For the finite difference method, larger arrays have many unused magnetometers, whereas for smaller arrays, a larger portion of the available magnetometers is used. Similar comparison observations can be drawn from the figure showing the different noise values. Therefore, these results are shown in Figure C-5 in Appendix C.

Unless an array with $l_s < 1.7$, corresponding to spacings larger than (l_0 : 0.088 m, l_1 : 0.71 m), is already available, it is better to apply the LLS method instead of the finite difference. If this finite difference method is implemented, it is optimal to use an array with five by five magnetometers, a fourth-order approximation and an $l_s = 2$ (l_0 : 0.075 m, l_1 : 0.6 m). However, since a three-by-three array with a smaller spacing of $l_s = 2.9$ (l_0 : 0.052 m, l_1 : 0.41 m) and a second-order approximation is significantly smaller and does not perform much worse, this might be a good option as well.

Variance Using the Local GP Method

Finally, the local GP method, described in Section 3-2-3, is analyzed and compared with the other two methods. The uncertainties, which are provided along the predictions in Gaussian process regression (GPR), depend solely on the kernel and relative input locations, as can be observed in Eq. (2-7), meaning that the variance remains the same across simulations. Therefore, it is not necessary to run multiple simulations in order to determine the variance, as each simulation would yield the same variance. Due to this reduced computational burden,

the decision was made to include even larger magnetometer arrays in the analysis. Arrays of up to nine by nine magnetometers are included. It was experimentally verified for arrays with up to five by five magnetometers that the variance provided by the GPR was practically identical to if the variances were computed manually by running 70 simulations of Figure 3-1. These results are shown in Appendix D.

Figure 4-6 presents the variance for arrays with different numbers of magnetometers using the local GP method. This method also shows that array configurations with a magnetometer at the center of the array perform relatively well. In fact, the variances corresponding to arrays with a central magnetometer closely follow those of arrays with one additional magnetometers, while notably outperforming the one with one fewer. Additionally, this figure also demonstrates that increasing the number of magnetometers helps to average out the measurement noise, increasing the optimal l_s value. This not only shifts the optimal value to a smaller spacing, but also widens the range of l_s values for which an accurate estimate can be computed.

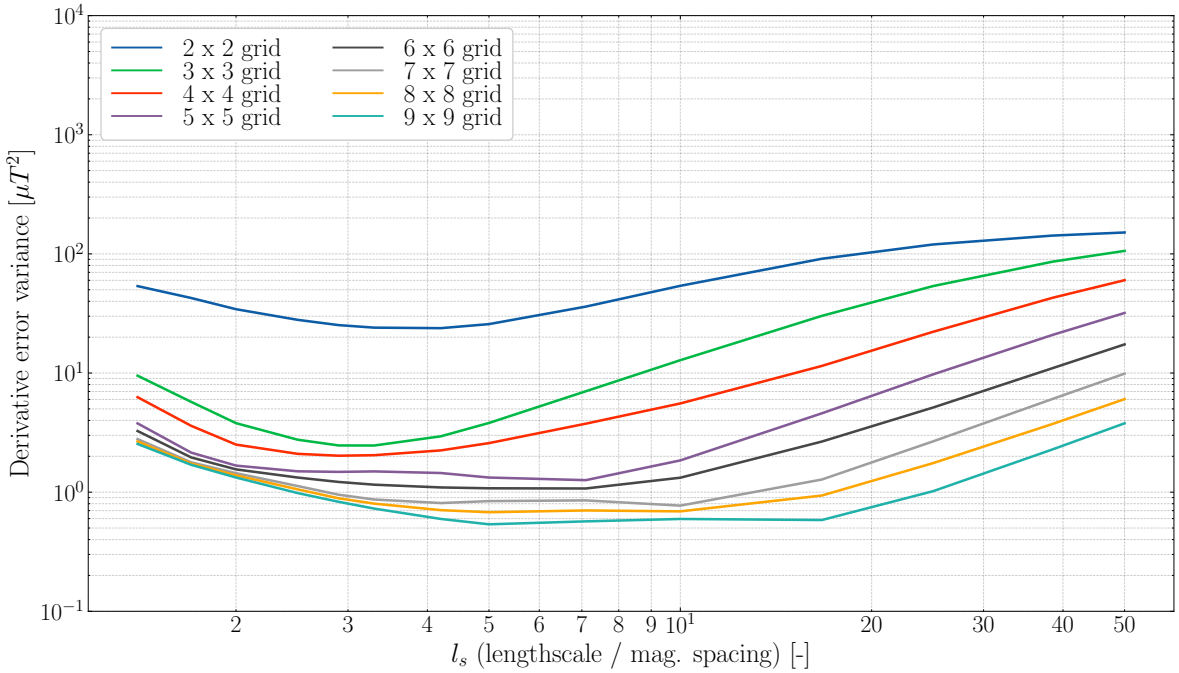

Figure 4-6: Relationship between the mean variance of the combined components of the spatial derivative estimate error and l_s using the local GP method. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%).

Compared to the LLS method in Figure 4-2 the local GP method in Figure 4-6 produces variances that are generally slightly lower. In addition to taking the curl-free property into account, as in the other two methods, this technique also takes the whole model into account. Hence, it is aware of the hyperparameters l, σ_f, $\sigma_{\text{lin.}}$ and the amount of measurement noise σ_y. This is probably the reason why the performance is better. The information on the noise probably gives the local GP method an advantage over the LLS method at small spacings. The local GP method appears to converge to a variance of approximately 200 for $l_s > 50$, which is equal to the variance obtained when always predicting zero. The variance of the LLS method for $l_s > 50$ in Figure 4-2, on the other hand, appears to grow unbounded

and does not converge. The other hyperparameters, particularly l and σ_f, describe how the magnetic field evolves over space. This information is especially useful for arrays with a larger spacing, as it helps characterize how the magnetic field evolves between the measurements. The variance using the local GP method performs significantly better at larger spacings compared to the LLS method. As a consequence, the local GP method typically reaches its optimal spacings at larger spacings compared to the LLS method. For instance, for a 3×3 grid, the local GP method has an optimum at approximately $l_s = 3$, whereas the LLS method at approximately $l_s = 4$. This shows that due to the awareness of the model the local GP method can capture the spatial derivative more accurately at larger spacings. To what extent these advantages are transferred to real-world data is unknown. The range of optimal spacings of the local GP method is similar to the range of the LLS method, but it is slightly shifted toward larger spacings. When considering arrays up to five by five, the optimal spacings lie in an approximate range from $l_s = 2$ ($l_0 : 0.075\,\text{m},\ l_1 : 0.6\,\text{m}$) up to $l_s = 7$ ($l_0 : 0.021\,\text{m},\ l_1 : 0.17\,\text{m}$). Arrays with more magnetometers have a wider range and a smaller optimal spacing. The figure illustrating the variance for various values of the measurement noise is shown in Figure C-6 in Appendix C.

Effect of Hyperparameter Uncertainty on Local GP Method

To better understand what the effect is of using hyperparameters that do not exactly match the underlying model, a simulation experiment is conducted in which the hyperparameters are varied. The hyperparameters of the underlying model that are used to generate the magnetic field are kept constant. Only the hyperparameters associated with the local GP method are varied from one array to another. The following hyperparameters are changed: σ_f, l, σ_y and $\sigma_{\text{lin.}}$. Each hyperparameter is perturbed using a uniform distribution from -0.4 to 0.4 times its original value. This results in a mean absolute perturbation of 20% for each hyperparameter. Figure 4-7 shows the comparison of the variance of the derivative estimate error between varying the hyperparameters and using the underlying true hyperparameters.

The figure illustrates that the variance based on the varying hyperparameters is higher than the corresponding variance using the true hyperparameters. The largest increase in variance is observed for larger spacings. The variance increases for larger spacings because the larger distance between the magnetometers requires the GP to rely more on the model to accurately predict the magnetic field and spatial derivative. However, the accuracy of this model decreases due to the varying hyperparameters. In the case of smaller spacings, the model plays a less crucial role. Because of the fact that the modeling errors result in a less accurate derivative estimate at larger spacings, the optimal spacings tend to shift to smaller spacings, albeit very slightly. Figure C-7 in Appendix C shows the comparison for all different components of the spatial derivative.

When comparing the variances with the varying hyperparameters from Figure 4-7 to the variances obtained using the LLS from Figure 4-2, it can be observed that the difference between the local GP method and the LLS method has become even smaller. The advantage that the local GP had compared to the LLS method, particularly at larger spacings, is diminished. However, despite the variable hyperparameters, the local GP method still provides lower variances than the LLS method.

Overall, if the accuracy with which the hyperparameters can be determined is similar to the

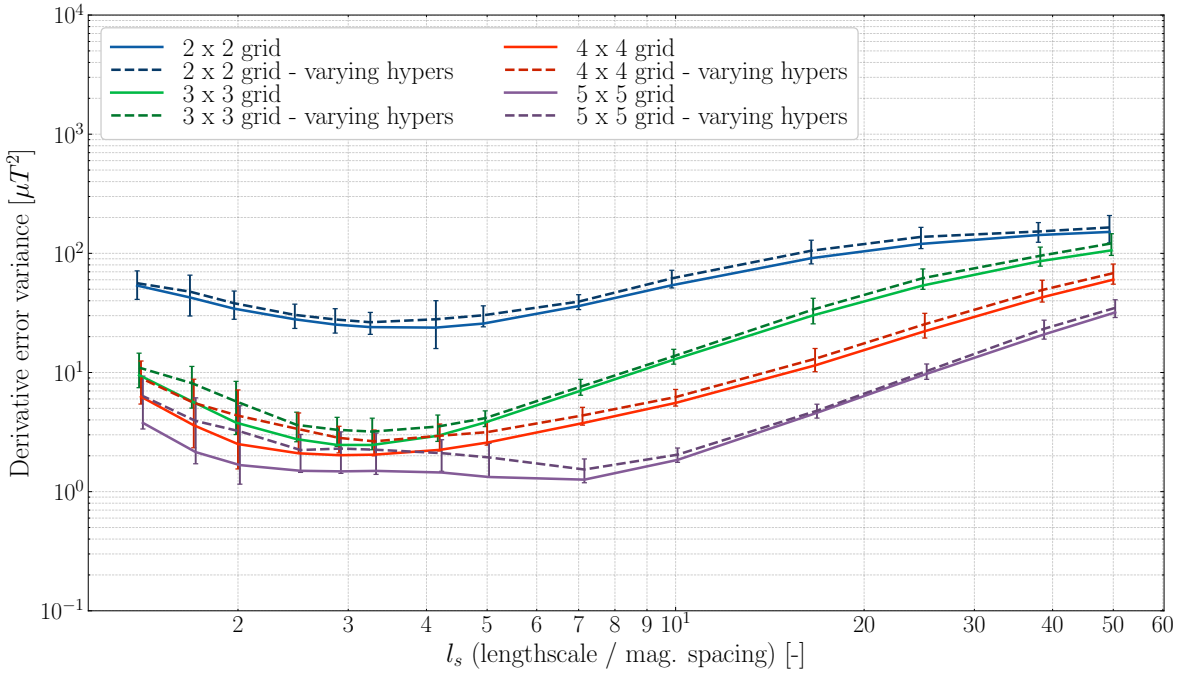

Figure 4-7: Comparison of the mean variance of the spatial derivative estimate error between using varying hyperparameters and true hyperparameters using the local GP method for arrays with different number of magnetometers. Only the confidence interval for the simulation with varying hyperparameters is shown, the other simulation with true hyperparameters is based on the analytic variance. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%).

uncertainty used in this experiment or better, the local GP method appears to be the better option. However, in practice, more uncertainties and errors are present, such as sensor errors and modeling assumptions. To what extent these factors exactly influence the local GP and LLS methods remains unknown.

In practice, selecting an optimal spacing for the local GP method should be done similarly to the LLS method. However, Figure 4-7 showed that errors in modeling assumptions and hyperparameters shift the optimal spacing to smaller spacings, since larger arrays rely more on the model. This phenomenon is specific for the local GP method, as it makes more assumptions on the model than the LLS method. Increased noise levels still shift the optimal spacing to larger spacings, as can be observed in Figure C-7 in Appendix C. As a general recommendation, a symmetric array of five-by-five with a spacing of $l_s = 6$ ($l_0 : 0.025\,\text{m}$, $l_1 : 0.2\,\text{m}$) is a reliable choice, even if the hyperparameters and model do not represent the underlying magnetic field very accurately. However, if the hyperparameters represent the underlying model well, a larger spacing of $l_s = 2.5$ ($l_0 : 0.06\,\text{m}$, $l_1 : 0.48\,\text{m}$) is also a good recommendation.

Three-Dimensional Cubic Array Configuration

All the results so far are related to two-dimensional arrays. However, at least some of the studies on the derivative-based odometry method used three-dimensional arrays [2, 3, 8, 9]. To examine how the results relate to three-dimensional arrays, an experiment was simulated on three-dimensional cubic arrays using the local GP method. Note that the three-dimensional

arrays used in the derivative-based odometry method are not structured in a cubic configuration but more in a tetrahedron-like structure. Therefore, neither do these results directly apply to those studies. Figure 4-8 compares the mean variance of the combined spatial derivative components for cubic arrays to planar arrays. It illustrates that, while the overall shape of the variance for the cubic structure is still very similar to that of the planar structure, the optimal spacing is slightly smaller. The reason for this decrease is probably that significantly more magnetometers are present and that, even though the extra magnetometers are at different heights, the noise can be averaged out better, resulting in a smaller optimal spacing. This idea is supported by the fact that most of the improvements of the derivative estimate lie in the region of smaller spacings, where noise plays a larger role.

The cubic structure affects the variance of certain components of the derivative more than others. Figure C-8 in Appendix C illustrates this. All elements that include at least one z in the index show a significant improvement. The additional magnetometers that extend into the z-direction now allow us to compute the zz-component, whereas with the planar structure this was impossible before. Additionally, the xz- and yz-components also show a great improvement because of the extra information on how the x- and y-components change in the z-direction.

Overall, based on this preliminary experiment, extending the planar grid structure to a three-dimensional cubic structure does not fundamentally change how the array configuration influences the accuracy of the derivative estimate. In other words, the results from a two-dimensional array transfer probably transfer reasonably well to a three-dimensional array. However, the increased number of magnetometers and improved capability of averaging out the noise result in a smaller optimal spacing. Therefore, even though some of the studies related to the magnetic field odometry method utilize three-dimensional array, the results from this thesis are still useful. When applying the results to a three-dimensional cubic array, one should use a spacing that is slightly smaller than the optimal spacing for the corresponding two-dimensional array. However, if not a cubic structure is utilized but a more sparse three-dimensional array, a similar optimal spacing is likely to apply as for the two-dimensional array. That is because a sparse three-dimensional array does not have the benefit of the increase in the number of magnetometers. Therefore, it is not necessarily better at averaging out the noise.

4-1-2 Computational Efficiency of Derivative Estimation Methods

When implementing the derivative kernel, the derivative must be computed repeatedly. Therefore, computational time also plays an important role, especially in combination with an iterative algorithm such as simultaneous localization and mapping (SLAM) or any odometry method. Since this thesis focuses on the efficient use of magnetometer arrays, the computational efficiency of the derivative estimation methods is also analyzed. Figure 4-9 shows the computation times for the various methods. The calculations were performed in Python on a virtual machine in the IBM cloud with 8 vCPUs. Since the programming language and code are not optimized for high-speed computations, only parts of the estimation process are considered in the computation time. For the local GP method, the time required to fill the kernels is ignored, only the computation of the inverse and the matrix operations required to compute the predictions and uncertainty as specified in Eq. (2-7) are considered. These

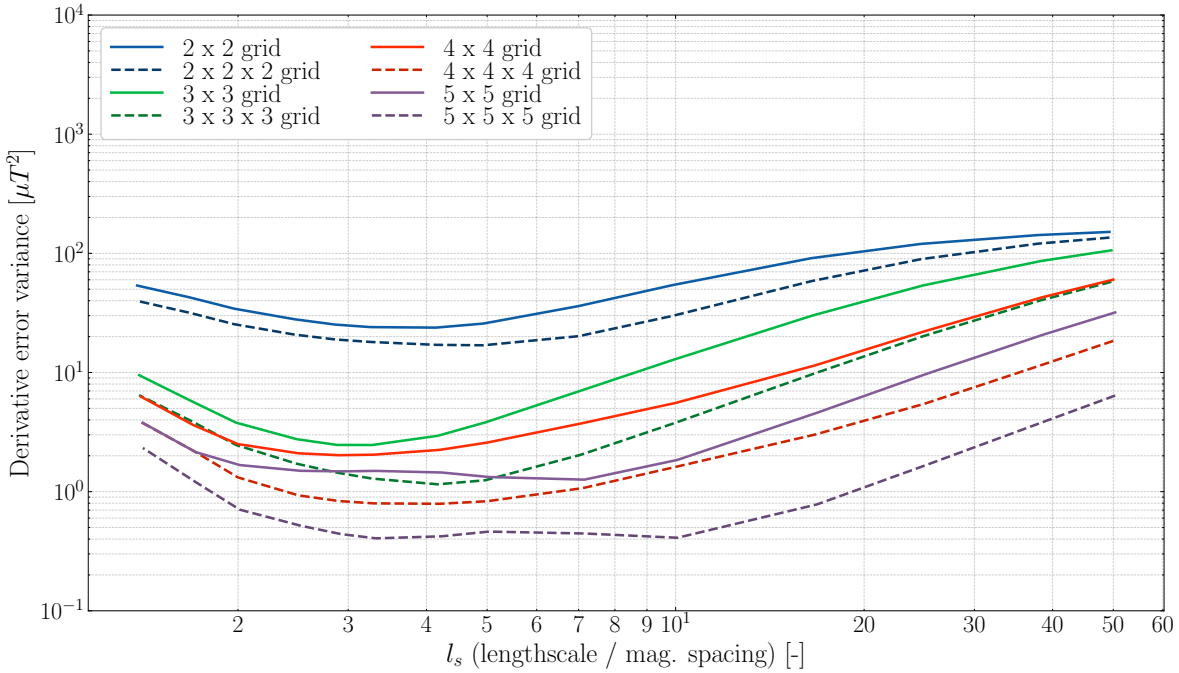

Figure 4-8: Comparison of the mean variance of the spatial derivative estimate error between two- and three-dimensional arrays using the local GP method. The relationship is plotted for different number of magnetometers. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%).

functions are implemented using the Numpy library and are highly optimized and should be the bottleneck when this method is implemented for high-speed computations, specifically the inverse computation. Furthermore, since this method is likely to be implemented within an iterative algorithm where the derivative must be computed at each time step, the kernels have only to be computed once beforehand. After all, the relative distances between the magnetometers remain the same from time step to time step, resulting in the same kernels. This also justifies excluding the filling of the kernels from the computation time. For the LLS method, the bottleneck is related to solving the system of linear equations. The whole left-hand side of Eq. (3-17) is independent of the measurements itself and can be computed once beforehand using Sympy and is therefore ignored in the computation time. The right-hand side differs from time step to time step and is included. The computation time of the finite difference method includes the evaluation of the equations in Section 3-2-1. The reason that some bars for orders are missing for the finite difference method for configurations with a certain number of magnetometers is that there are not sufficient magnetometers to evaluate the higher orders.

The figure shows that there is a clear uptrend in computation time as the number of magnetometers and the order of approximation increase. The finite difference method is significantly faster than the other two methods, as it only has to do basic arithmetic operations such as addition, subtraction, multiplication, and division. It is typically between ten and one hundred times faster. Depending on the order, the LLS and local GP method exhibit a more similar computational time. For higher-order approximation, the LLS method generally has a computational time within the same order of magnitude as the local GP method, whereas

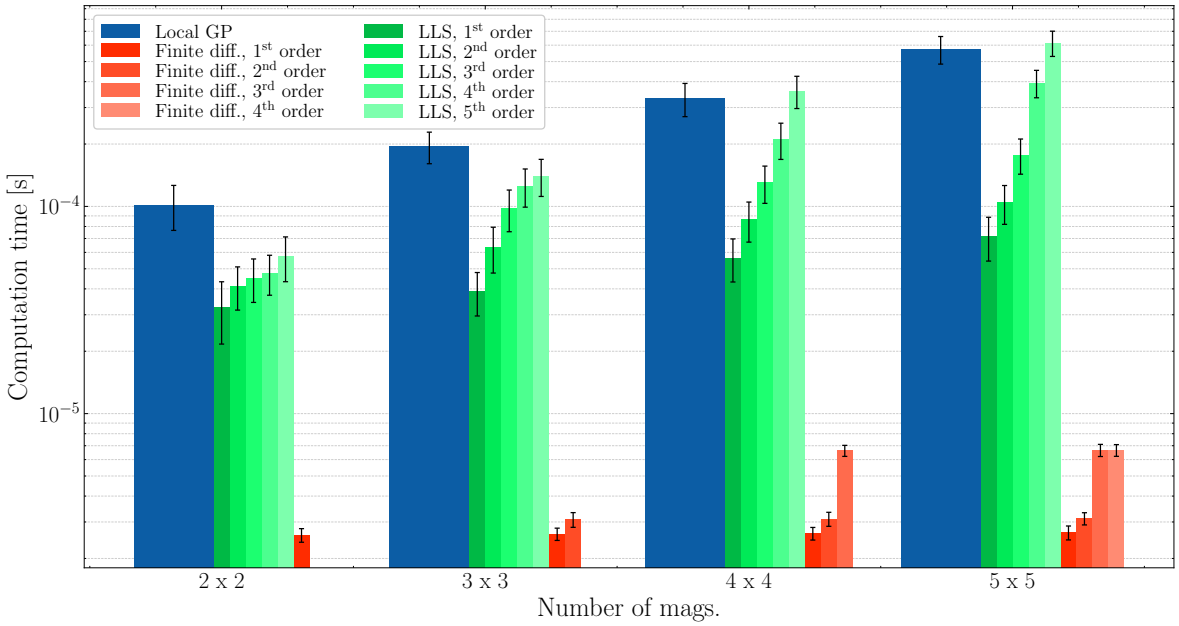

Figure 4-9: Computation times for the local GP, the finite difference and the LLS method. The finite difference and LLS method also make the distinction between the different order of approximations.

for lower orders, it is considerably faster.

Although the finite difference method generally has a higher variance compared to the LLS and local GP methods, the simplicity of implementation and low computational time make it a possible solution in scenarios where these factors outweigh the need for more accurate estimates. The difference in computation time between the local GP and LLS method are smaller. If computational resources are very limited, it is probably beneficial to opt for a lower-order approximation using the LLS method. This approach offers a shorter computation time compared to the local GP method and the higher-order LLS approximations. Then one should select a relatively small spacing, depending on the exact number of magnetometers and noise level, in an approximate range from $l_s = 6$ ($l_0 : 0.025\,\text{m}$, $l_1 : 0.2\,\text{m}$) to $l_s = 15$ ($l_0 : 0.01\,\text{m}$, $l_1 : 0.08\,\text{m}$), as can be observed in Figure 4-2 and Figure 4-4. While the resulting variance is slightly worse, it is comparable to variances at larger spacings when larger-order approximations are used, but with the benefit of a lower computation time. Although, even for the method that requires the most computational power, the computation time is still under one millisecond. In case computation resources are not very limited, it is probably best to use a higher-order LLS method or the local GP method.

4-2 Map Quality: Derivative Kernel and Full Kernel

The goal of this thesis is to study how magnetometer arrays can be used effectively and efficiently to create magnetic field maps. The main focus is on comparing the derivative kernel and the full kernel. Using the results from Section 4-1, this section evaluates the global and local map qualities achieved with both kernels and explores which configurations yield

the highest quality maps. The simulation experiments to assess the map quality for the global and local maps were described in Section 3-3-3. Section 4-2-1 examines the global map quality and Section 4-2-2 the local map quality.

To quantify the map quality, the mean standardized log loss (MSLL) performance metric is utilized. The MSLL is a metric typically used for probabilistic regression, as it can take into account the predictive variance. That is, it not only evaluates the accuracy of the predictions, but also how well the model quantifies the uncertainty [27]. Therefore, it can be used for GPR. It is given by

$$\text{MSLL} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{2}\log 2\pi\sigma_{*,i}^2 + \frac{(f_{*,i} - \bar{f}(\mathbf{x}_{*,i}))^2}{2\sigma_{*,i}^2}\right), \tag{4-1}$$

where $\sigma_{*,i}^2 = \mathbb{V}[f_{*,i}]$ is the predictive variance of the GPR. Since the MSLL is defined as a loss function, a lower value denotes a better score and a better prediction.

4-2-1 Global Map Quality

We want to compare the accuracy and computational efficiency of the derivative kernel with the full kernel. To that end, both kernels are evaluated for the global map. The global map represents a predefined magnetic field map with a predefined trajectory, with the data selected such that test locations are considered that are both far and near from the training data. This evaluation contributes to answering research the third sub-question.

Global Map Quality Using Derivative Kernel

Figure 4-10 shows the results on using the derivative kernel on the global map from Figure 3-2 for arrays with multiple number of magnetometers using the LLS method. The figure shows the map quality where for each array configuration the order of approximation was used that gave rise to the lowest mean variance of the spatial derivative. The map quality is given in terms of MSLL.

The figure shows that the results for the x- and y-components are very similar and do not appear to change significantly between each other. This similarity between the x- and y-component will be apparent throughout the remainder of this thesis, as the array moves in the xy-plane and hence there is no fundamental difference between the x- and y-direction that would give rise to different behavior. However, the z-direction does cause a different behavior, since the array neither extends nor moves along this axis. The figure shows that the map quality of the z-component is generally significantly worse. The performance of each component for a two-dimensional map is mainly determined by the accuracy with which the spatial derivative of that particular component can be determined in the x- and y-directions. For instance, the map quality of the z-component of the magnetic field depends on the estimate accuracy of the z-component in the x- and y-directions, i.e. the xz- and yz-components of the Jacobian. Similarly, the map quality of the x- and y-components of the magnetic field depends on the xx- and xy-components, and the yy- and xy-components, respectively. Since in general the xx-, xy- and yy-components have better estimates than the xz- and yz-components, as can be observed in Figure C-2 in Appendix C, the map quality for the x- and y-components of the magnetic field is better in general. Especially the xy-component of the Jacobian has

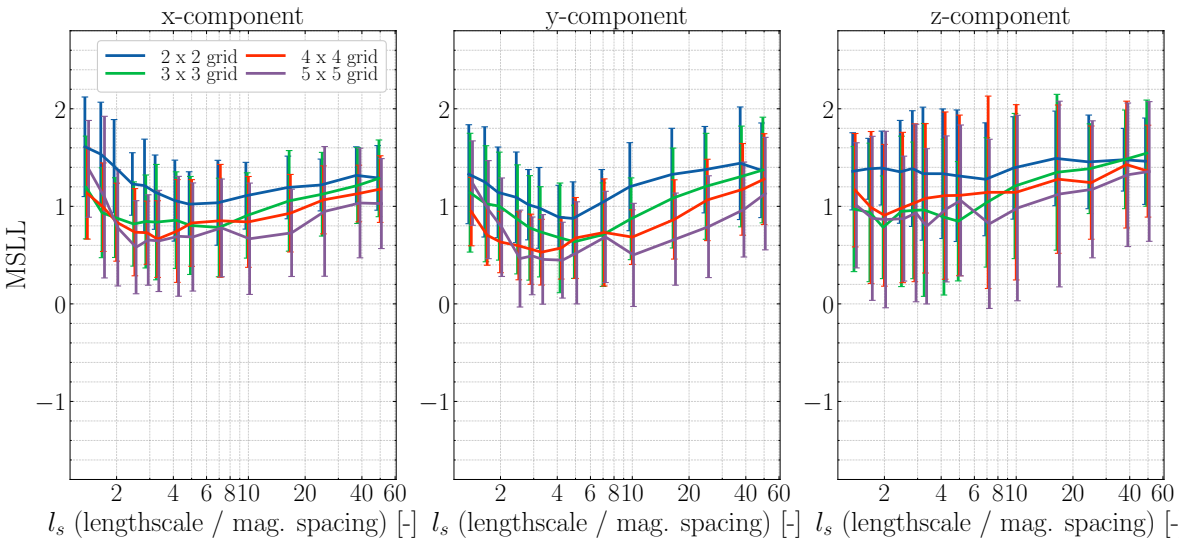

Figure 4-10: Map quality using the derivative kernel for different number of magnetometers and l_s values using the LLS method. The different plots indicate the quality for the different magnetic field components. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%). To prevent the overlap of the confidence interval some of the plotted lines are shifted slightly to the right.

relatively accurate estimates. This is because the xy-component has twice as much data due to the fact that the curl-free property implies that the xy- and yx-components have to be equal. As a result, the xy-component is the main contributor to good estimates of the x- and y magnetic field components. This leads to the fact that using the derivative kernel, a two-dimensional array is slightly better at predicting the x- and y-components of the magnetic field when only the curl-free assumption is used.

Since the only underlying variable that differs between simulations when using the derivative kernel is the accuracy with which the spatial derivative can be estimated, it is reasonable to assume that this is the only variable that causes differences in the map quality. As a result, it makes sense that the map quality follows a pattern very similar to the variance of the derivative estimate error present in the corresponding figures. This can be observed when comparing Figure 4-10 with Figure C-2 in Appendix C. Similarly to the figures that illustrate the variances of the derivative estimate error, the map quality figure also shows that increasing the number of magnetometers improves the map quality. Furthermore, larger arrays with more magnetometers also tend to allow for a wider range of effective l_s values for which a good performance is obtained. For a grid of five by five, the best performance is obtained roughly for l_s values between 2.5 (l_0 : 0.06 m, l_1 : 0.48 m) and 10 (l_0 : 0.015 m, l_1 : 0.12 m). While arrays with fewer magnetometers also reach their optimal values approximately within this range, their effective intervals are typically narrower.

The map quality using the derivative kernel for different values of measurement noise is shown in Figure 4-11. Similarly to the map quality figures regarding the number of magnetometers, the x- and y-component perform relatively similarly, whereas the z-component performs notably worse. Furthermore, the patterns and trends are again very similar to the corresponding derivative error variance plots in Figure C-4 in Appendix C. Less noise significantly improves

the map quality at each spacing and also slightly shifts the optimum l_s value toward smaller spacings.

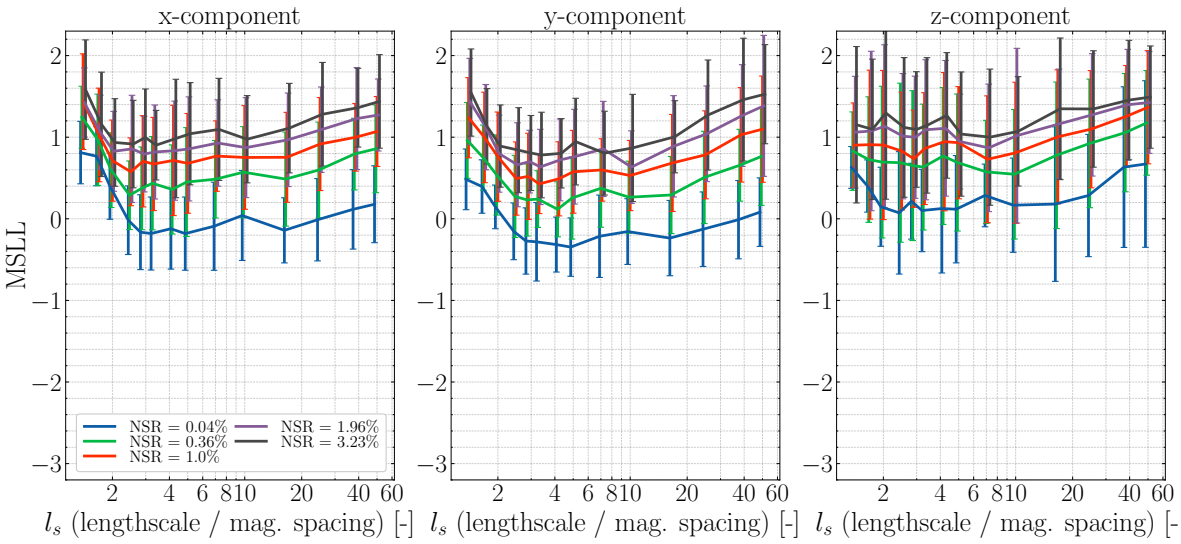

Figure 4-11: Map quality using the derivative kernel for different measurement noise and l_s values using the LLS method. The different plots indicate the quality for the different magnetic field components. The number of magnetometers for all configurations was set to five by five. To prevent the overlap of the confidence interval some of the plotted lines are shifted slightly to the right.

Global Map Quality Using Full Kernel

Figure 4-12 illustrates the map quality across arrays with different numbers of magnetometers when using the full kernel. It shows that arrays with more magnetometers outperform the smaller ones, and that small l_s values notably result in better performance than large values. The optimal l_s value for all grids is located at the largest spacing at $l_s = 1.4$. This is because the arrays become so large that the training points start to overlap with the test data. The x- and y-components of the magnetic field give almost identical performance, whereas again the z-component performs slightly worse. This again arises from the fact that the curl-free property is incorporated in the GP model and that the array moves in the xy-plane, resulting in relatively less information on how the z-component evolves.

When comparing the map quality using the derivative kernel from Figure 4-10 to the full kernel from Figure 4-12, it is clear that the map quality using the full kernel outperforms the derivative kernel for each combination of l_s and the number of magnetometers. This is in line with expectations as the derivative kernel serves as an approximation to the full kernel.

Figure 4-13 illustrates the map quality when the full kernel is used for different values of measurement noise. It again clearly shows that the full kernel consistently outperforms the derivative kernel from Figure 4-11, indicating that, also for different measurement noise values, the derivative kernel serves as an approximation.

Overall, this section indicates that the full kernel has a large advantage compared to the derivative kernel at large spacings, as it can measure the magnetic field close to the test

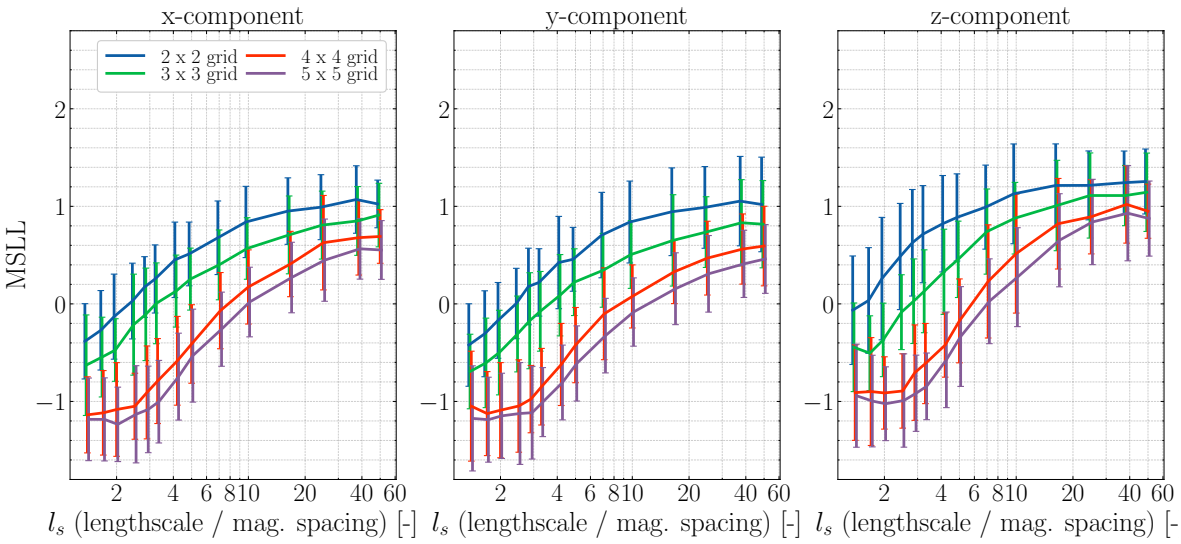

Figure 4-12: Map quality using the full kernel for different number of magnetometers and l_s values using the LLS method. The different plots indicate the quality for the different magnetic field components. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu\text{T}$ (NSR=1%). To prevent the overlap of the confidence interval some of the plotted lines are shifted slightly to the right.

data. This provides a lot more relevant information than the pure local information that the derivative kernel uses. However, whether the higher map quality of the full kernel outweighs the extra required computational power compared to the derivative kernel is analyzed in Section 4-2-3. Furthermore, the optimal spacings that provide the highest map quality all appear to be equal to the spacings that resulted in the lowest variance in Section 4-1-1. When using a magnetometer array, it is more effective to select the configuration based on the optimal findings in Section 4-1-1, as the results in this section are significantly noisier but yield the highest map quality for the same configurations.

4-2-2 Local Odometry Map Quality

Since magnetometer arrays have shown potential for odometry purposes, the local odometry map is also analyzed for the full kernel and derivative kernel. The local odometry map will predict the magnetic field at the magnetometer locations on arrays that are shifted in different directions with a certain step size. This comparison aims to examine whether the derivative kernel is better suited for global or local maps, or whether there is no significant difference between the two. The main reason to opt for the derivative kernel is due to its favorable computational time compared to the full kernel. However, since local maps consist of significantly fewer data points compared to global maps, the computational burden of odometry is significantly lower compared to global mapping. As a result, the derivative kernel becomes relevant for local maps only if the update frequency is high or if little computational power is available. Apart from the comparison between the derivative kernel and the full kernel to assess whether the derivative kernel could effectively reduce the computational burden in specific scenarios, the individual results per kernel in this section provide useful information as well. Specifically, these results will show which array configuration is likely to yield the

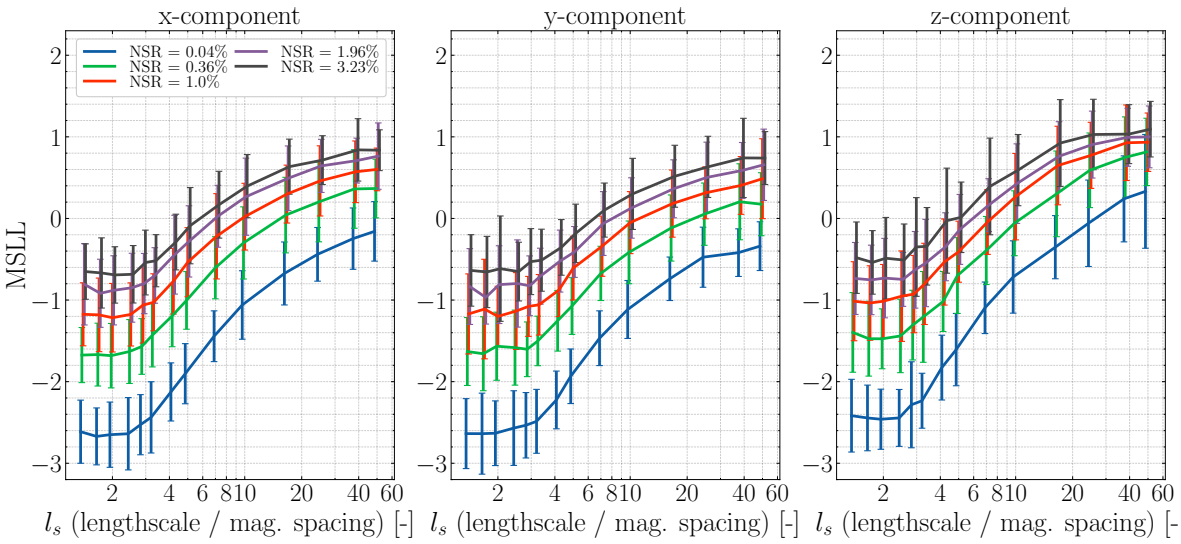

Figure 4-13: Map quality using the full kernel for different measurement noise and l_s values using the LLS method. The different plots indicate the quality for the different magnetic field components. The number of magnetometers for all configurations was set to five by five. To prevent the overlap of the confidence interval some of the plotted lines are shifted slightly to the right.

best odometry performance for both kernels. While these results related to the derivative kernel are valuable, the results on which configuration performs best for the full kernel are equally significant, since odometry using all magnetometers has already been implemented in [16, 29, 30].

Note that the results of the derivative kernel in this section are based on the local GP derivative estimation method, described in Section 3-2-3. We made this decision for practical reasons. However, the results using the LLS method, described in Section 3-2-2, would be very similar and the conclusions would not have been different.

Local Map Quality Using Derivative Kernel

Figure 4-14 shows the local map quality using the derivative kernel for arrays with different numbers of magnetometers and different step sizes relative to the lengthscale l_{ss}. Only the x-component is shown, since the y-component has almost identical performance and the z-component slightly worse, as expected by now. Three different factors play an important role for this local map quality: the size of the array, the accuracy with which the spatial derivative can be determined, and the step size. The size of the array is crucial here because the test points are shifted versions of the original array. Therefore, if the size of the array is larger, then part of the test points lie further away from the middle magnetometer, where the spatial derivative is determined. The derivative is only accurate within a limited range around that point. As a result, the figure illustrates that as the spacing increases, the array becomes larger and the performance deteriorates. This can be observed, for instance, for the five-by-five array with a spacing of $l_s < 10$. Similarly, as the array consists of more magnetometers, which give rise to a larger array, the performance decreases compared to arrays with fewer

magnetometers for spacings of approximately $l_s < 10$.

However, arrays with more magnetometers perform better at smaller spacings of approximately $l_s > 10$. At smaller spacings, the array size does not differ as much between arrays with different numbers of magnetometers, which reduces its effect. Therefore, the second factor, the accuracy with which the spatial derivative can be determined, starts to dominate as the spacing decreases. As a result, the arrays with more magnetometers perform better at smaller spacings as they provide a more accurate derivative estimate.

Finally, the step size also influences the performance. A larger step size overall means that the test points lie further away from the middle magnetometer, which decreases the performance. Additionally, it can be observed that the difference in performance between arrays with different numbers of magnetometers is smaller at small step sizes, specifically at small spacings of $l_s > 10$. This is because as the step size decreases, the test points lie closer to the middle magnetometer. This results in the fact that the prediction is more based on the measured magnetic field itself and not so much on the spatial derivative estimate anymore. This also leads to the optimal spacing shifting very slightly towards smaller spacings as the step size decreases. That is because the contribution of the spatial derivative becomes smaller and hence it is more favorable to have a very small spacing such that the test points lie close to the middle magnetometer. The optimal spacing that is relevant for determining the spatial derivative becomes less important.

Figure C-9 in Appendix C illustrates the local map quality for various NSR values. Lower noise values improve the map quality, but barely change the optimal spacing.

Overall, this derivative kernel achieves the highest local map quality scores for five-by-five arrays with a spacing in the range from $l_s = 16.7$ ($l_0 : 0.009\,\text{m},\ l_1 : 0.072\,\text{m}$) to $l_s = 50$ ($l_0 : 0.003\,\text{m},\ l_1 : 0.024\,\text{m}$) depending on the step size. For smaller step sizes, the best performance is obtained at the smaller spacings of the range, whereas in the case of larger step sizes, the optimum shifts towards the larger spacings. Therefore, when the derivative kernel would be implemented in a model-based odometry approach, it is best to select the spacing of the array based on the range of step sizes that is expected to occur the most in that particular application.

However, the performance of the derivative kernel for this model-based odometry scenario is not optimal due to two conflicting effects. The optimal spacing to determine the spatial derivative is typically on the larger side, as can be observed in Figure 4-6, which, in general, should provide the highest quality maps. However, this leads to moderately large array sizes, which, in turn, reduce the quality of the local map due to the increased distance between the spatial derivative and the test points. As a result, these effects prevent that the optimal spacing for computing the spatial derivative also provides the highest map quality for the model-based odometry approach. Whether the derivative kernel is still a suitable choice also depends on the performance of the full kernel and the computational time, which will be analyzed next.

Local Map Quality Using Full Kernel

Figure 4-15 shows the local map quality using the full kernel for arrays with different numbers of magnetometers and different step sizes. Whereas in Section 4-2-1 the global map quality

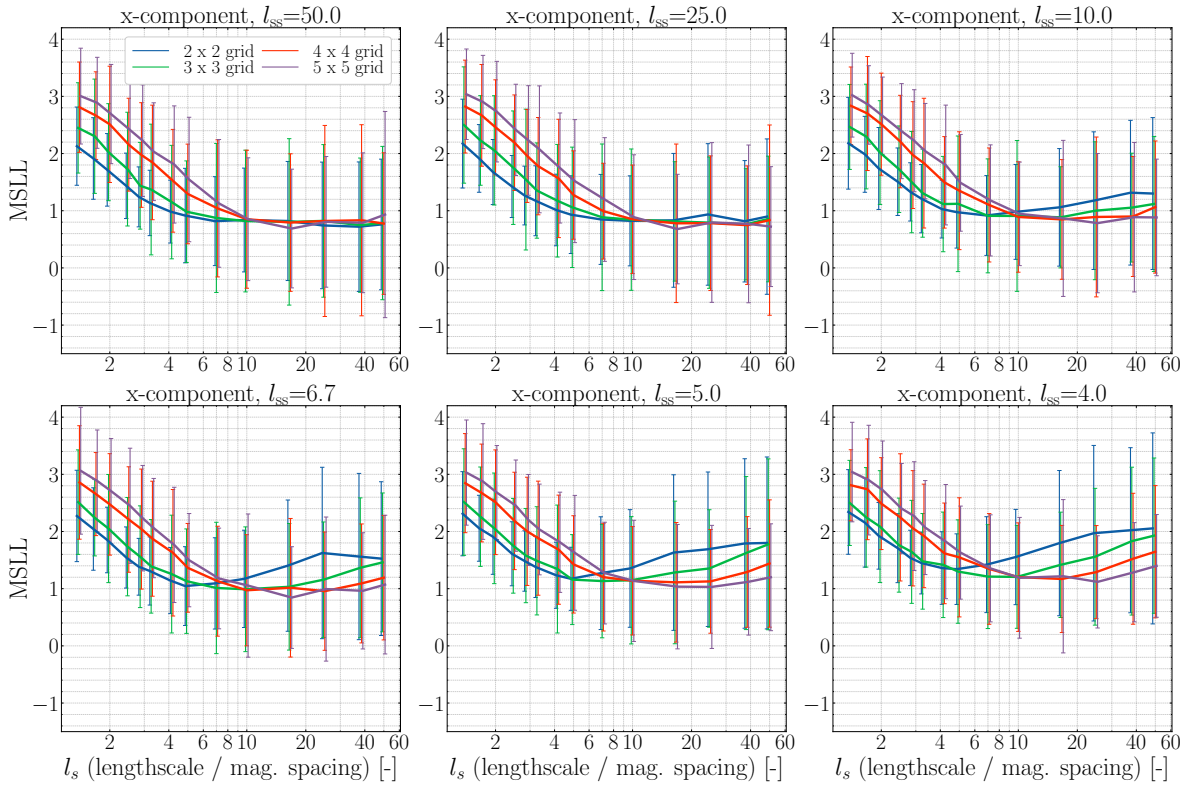

Figure 4-14: Local map quality using the derivative kernel for arrays with different numbers of magnetometers and l_s values using the local GP method. The different plots indicate the quality for the different relative step sizes l_{ss}. To prevent the overlap of the confidence interval some of the plotted lines are shifted slightly to the right.

using the full kernel really benefits from large spacings as it results in a much broader coverage of the area, the local map quality does not. The local map quality is more dependent on finer spacings, where the best performance is achieved when most of the test points nicely fall within the range of the array. There is always this trade-off between covering more test points within the array by choosing a larger spacing and keeping the resolution of the measurement high by choosing a smaller spacing. In other words, larger spacings have the advantage that the size of the array is larger and the step size is relatively small compared to the size of the array, which ensures that most test points fall within the range of the array. Smaller spacings, however, give rise to a higher resolution and have the advantage that the test points that fall within the span of the array can be predicted more accurately. This trade-off affects arrays differently depending on the number of magnetometers.

This trade-off can be observed in Figure 4-15. For instance, the three-by-three array achieves the highest map quality at spacings that are approximately equal to the step size. For example, at a step size of $l_{ss} = 6.7$ it has the highest map quality at a spacing of $5 < l_s < 7.1$. Arrays with more magnetometers achieve the best performance at spacings that are smaller than the step size. For instance, at the same step size of $l_{ss} = 6.7$, the five-by-five array performs best at a spacing of approximately $l_s = 10$. This can be explained by the fact that the outer rim of arrays with more magnetometers contains a lower proportion of the total number of magnetometers compared to arrays with fewer magnetometers. The outer rim of the test

arrays is always the first region to extend beyond the range of the training array. For arrays with fewer magnetometers the proportion of this outer rim is larger. Therefore, it is more favorable to have a larger spacing such that, because of the relatively smaller step size, the test points extend slightly less over the array's range. In this case, the higher coverage that a larger spacing offers, carries more weight compared to having a larger resolution. Arrays with more magnetometers benefit more from a smaller spacing and a higher resolution than the higher coverage since the proportion of magnetometers on the outer rim is smaller.

Apart from the difference in optimal spacing between arrays with different numbers of magnetometers, Figure 4-15 also shows that arrays with more magnetometers provide higher quality local maps compared to arrays with fewer magnetometers. This holds true for all step sizes. As the number of magnetometers increases, the proportion of the test array that falls within the range of the training array increases. Additionally, the figure also shows that as the step size decreases, the optimal spacing becomes smaller. For instance, whereas at a step size of $l_{\text{ss}} = 5$ a five-by-five array has an optimal spacing of approximately $l_s = 7.1$, at a step size of $l_{\text{ss}} = 25$ the optimal spacing is approximately $l_s = 25$. This is because as the step size decreases, it becomes possible to use smaller spacings as most of the test points fall within the range of an array with such a smaller spacing. The smaller spacings increase the resolution of the training points, and hence the map quality. Equivalently, the optimal map quality also increases as the step size decreases.

Figure **??** in Appendix C illustrates the local map quality for various NSR values. Lower noise values improve the map quality and very slightly shift the optimal spacing to smaller spacings. However, the difference is very small and should play a minor role in determining the optimal spacing.

Overall, the results illustrate that the highest local map quality is obtained for five-by-five arrays with a very small spacing of $l_s = 50$ and an equally small step size of $l_{\text{ss}} = 50$. Naturally, even more magnetometers and smaller spacing and step sizes would result in even higher map qualities. However, in practice, the step size cannot be freely selected and is dependent on the specific application and the available computational power. In the event that a frequently occurring range of step sizes is known, or the performance of a certain range of step sizes calls for improvement, the spacing between the magnetometers can be selected based on the results in Figure 4-15. When all step sizes are considered, a five-by-five array with a spacing in the range from approximately $l_s = 10$ ($l_0 : 0.015\,\text{m},\ l_1 : 0.12\,\text{m}$) to $l_s = 16.7$ ($l_0 : 0.009\,\text{m},\ l_1 : 0.072\,\text{m}$) yields the highest overall map quality.

When comparing the map qualities of the full and the derivative kernel for the equivalent combinations of array configuration and step size, the full kernel consistently outperforms the derivative kernel. However, in a practical implementation of a magnetic field odometry method, the step size would not be a constant variable, as it would vary depending on the speed of the considered entity as well as the computational speed of the algorithm and/or design choices. The optimal spacing of both the full kernel and the derivative kernel depends on the step size. However, the step size has a significantly larger impact on the optimal spacing for the full kernel than it does for the derivative kernel. The full kernel really has to match the spacing to the step size to ensure that the majority of the test points fall nicely within the range of the array. As a result, the derivative kernel has a relative advantage over the full kernel when the step sizes vary significantly between iterations. However, which kernel is ultimately preferable also depends on the computational efficiency of both methods,

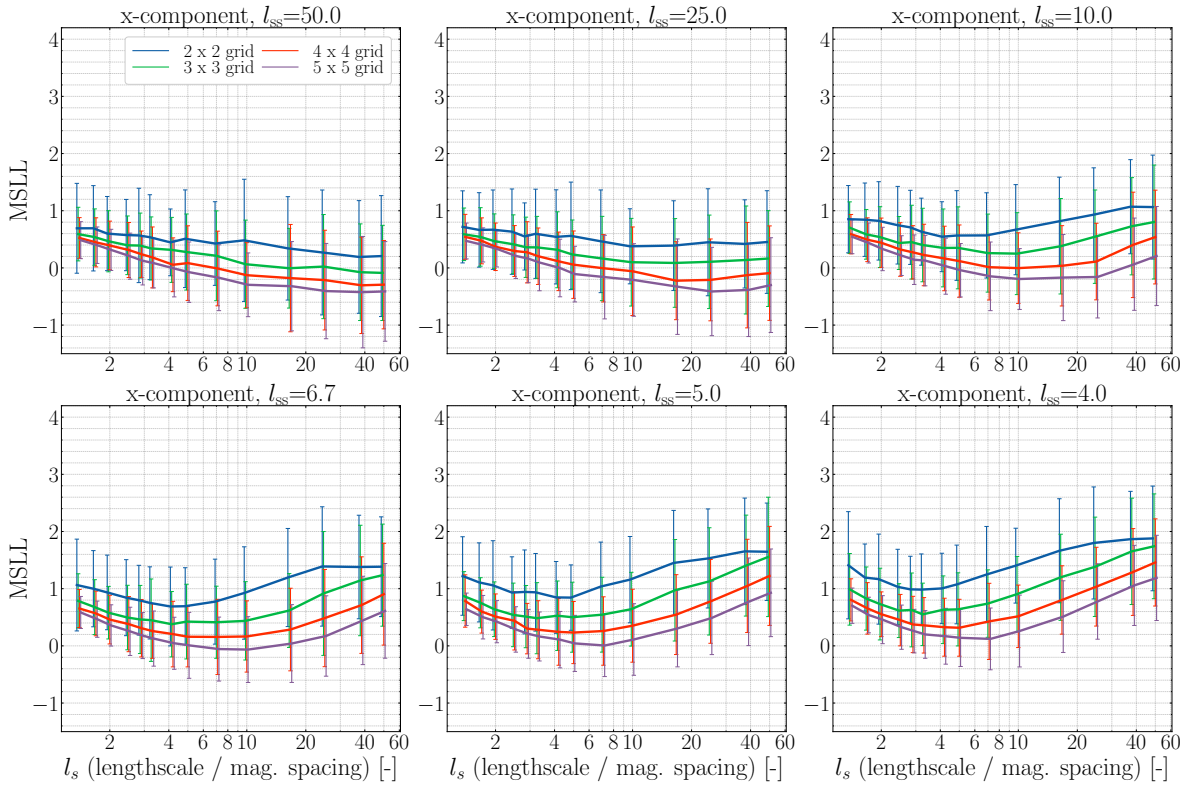

Figure 4-15: Local map quality using the full kernel for arrays with different numbers of magnetometers and l_s values using the local GP method. The different plots indicate the quality for the different relative step sizes l_{ss}. To prevent the overlap of the confidence interval some of the plotted lines are shifted slightly to the right.

which will be analyzed next.

4-2-3 Computational Efficiency: Curl-Free Derivative Kernel and Curl-Free All Kernel

The theoretical time complexity of GPs is $\mathcal{O}(N^3)$ [27], where N is the number of data points. In the case of the full kernel, the number of data points is $N = 3m^2n$, where m is the number of magnetometers along each axis and n the number of times the array takes a series of measurements. This gives rise to a time complexity of $\mathcal{O}(m^6n^3)$. The derivative kernel has $N = 12n$ data points, which gives rise to a time complexity of $\mathcal{O}(n^3)$

To assess the computational efficiency of the two kernels in practice, 700 simulations were run for each kernel, each using a single set of measurements from an array to predict the magnetic field at one location. The filling of the kernels is again excluded from the computation time, because the implementation with Python is not the most time-efficient way. Only the inverse and matrix operations described in Eq. (2-7) are taken into account, as they are efficiently implemented using the Numpy library. Figure 4-16 illustrates the mean computation time for the arrays with different numbers of magnetometers and kernels. As expected, the computation time for the derivative kernel is independent of the number of magnetometers. This is because the number of data points does not increase with the number of magnetometers

because each time a single measurement and the spatial derivative are used. This results in 12 data points, 9 from the spatial derivative and 3 from the magnetic field. In contrast, the computation time of the full kernel does grow quickly as the number of magnetometers increases. It should be noted that for the 2×2 grid, the computation times for both kernels are very similar, which is expected since the full kernel then also uses 12 data points. This allows for a quick comparison between the derivative and the full kernel with an equal computation time.

First, when comparing the global map quality between the derivative kernel in Figure 4-10 and the full kernel in Figure 4-10 for the 2×2 grid, it is evident that the full kernel consistently has a better optimal score than the derivative kernel. Specifically, the full kernel with a grid of two by two achieves an optimal MSLL of -0.2 at large spacings of $l_s = 1.4$, whereas the derivative kernel with the most magnetometers only approximately reaches an optimal MSLL of 0.4. This demonstrates that for global maps, at least for the map used in this simulation experiment, a large array with only four magnetometers provides better map quality than a more densely packed array with magnetometers whose information is approximated using the derivative kernel. With fewer, widely spaced magnetometers, the measurements effectively serve as individual measurements, providing completely new information and yielding more spatial information per unit of computational power. Furthermore, even if more computational power is available, the full kernel has more flexibility to optimally utilize this power. In addition to using the extra computational power by increasing the path length and collecting more data, the full kernel also has the possibility of using an array with more magnetometers. The derivative kernel can only increase the path length.

These conclusions probably do not depend on the measurement noise. Comparing the global map quality for the different kernels across the different measurement noise values in Figure 4-11 and Figure 4-13, illustrates that the amount of improvement of the map quality for lower measurement values at the optimal spacing is very similar between the full and the derivative kernel. The improvement in the full kernel is even slightly larger than in the derivative kernel. As a result, a different noise level should not alter the conclusion that an array with a widely spaced grid of 2×2 provides a higher map quality than a more densely packed array with more magnetometers whose information is approximated with the derivative kernel.

In case of the local map quality, it is no longer beneficial for the full kernel to utilize an array that spans a very large range with few magnetometers. In addition, step sizes must be taken into account. When examining the 2×2 grid for the full kernel in Figure 4-15 it turns out that a spacing of approximately $l_s = 5.1$ performs well across all considered step sizes, achieving a mean map quality of approximately 0.7 with a worst MSLL of 1.05 at a step size of $l_{\text{ss}} = 4$. In comparison, the derivative kernel in Figure 4-14 performs optimally with spacings of approximately $l_s = 25$, yielding a mean map quality of approximately 0.94 and a worst MSLL of 1.2 at a step size of $l_{\text{ss}} = 4$. This demonstrates that also for local maps when various step sizes are considered, the full kernel yields higher-quality maps than the derivative kernel for the same computational power. The performance of the derivative kernel can be improved by increasing the number of magnetometers to obtain a better derivative estimate. However, the amount of improvement this would yield is unknown.

Initially, one might expect that for the local map the derivative kernel could be useful when many but very noisy magnetometers are available. In the case where the full kernel only used a two-by-two array with an NSR of 3.23% to save computational power, it would be sensitive to

the measurement noise as few magnetometers are used. If the derivative kernel now used a five-by-five array with the same NSR it could still provide a reasonable accurate spatial derivative according to Figure 4-4. However, it appears that even in this scenario the full kernel yields a higher local map quality for all step sizes for a certain spacing. The corresponding figures and some additional elaboration are shown in Section C-1 in Appendix C. If the derivative kernel has access to even more magnetometers and/or if the measurement noise is larger, it is possible that the derivative kernel would outperform the full kernel eventually.

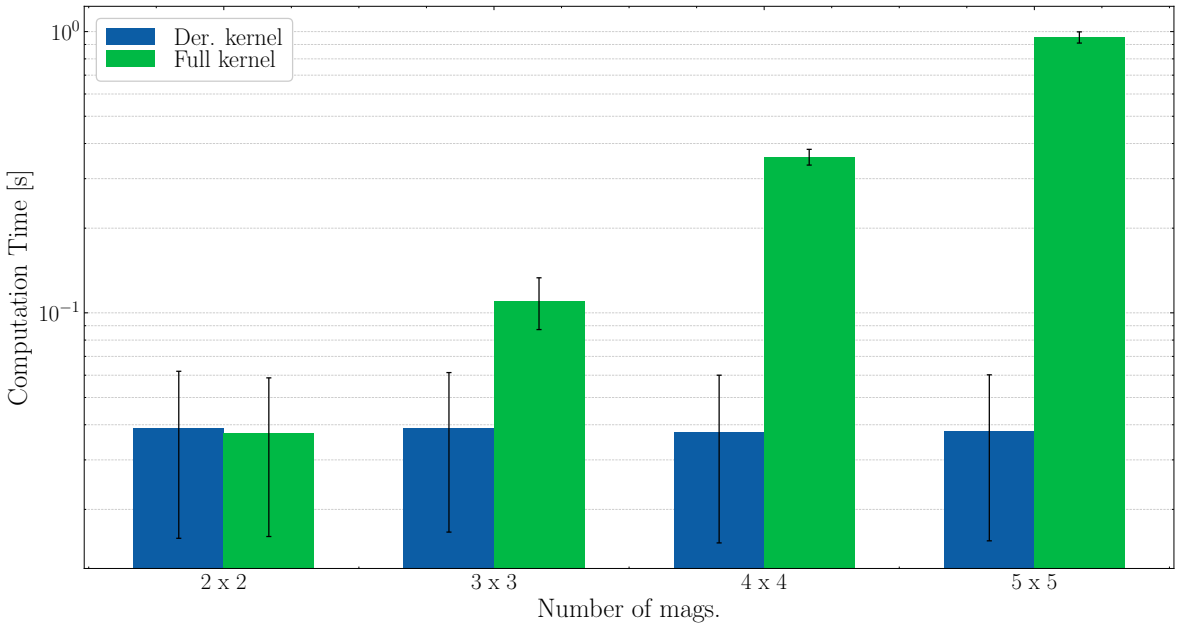

Figure 4-16: Computation times for making a single prediction with the derivative kernel and the full kernel for different number of magnetometers.

Chapter 5

Conclusion

Magnetometer arrays are a promising tool for creating magnetic field maps and odometry purposes, both for local and global maps. However, since the time complexity of Gaussian processes (GPs) scales cubically with the number of data points, using all magnetometers on the array to create maps can be extremely computationally demanding. Therefore, in this thesis, we analyzed whether the spatial derivative could be used as an approximation to the full information on the array. In addition, we also examined which array configurations are likely to yield the best performance for the derivative- and model-based odometry methods. This gave rise to the research question of this thesis: *How can magnetic field measurements from a magnetometer array be utilized effectively and efficiently to create magnetic field maps with Gaussian process regression?*. This thesis addresses this research question by analyzing the optimal array configuration for global and local maps. Two ways of including the measurement were compared to each other: using all magnetometers on the array to create the map, also referred to as the full kernel, and using a single magnetometer plus the spatial derivative, referred to as the derivative kernel. The derivative kernel approximates the information from the array with the spatial derivative and a single measurement. This approach significantly reduces the number of data points for Gaussian process regression (GPR), thereby also lowering the computational cost.

The derivative kernel requires the computation of the spatial derivative from the array. A crucial aspect to computing this spatial derivative is the array configuration. This is covered in the first research sub-question: *How does the array configuration influence the accuracy with which the spatial derivative can be computed?*. In this thesis, three aspects of the array configuration were analyzed: the number of magnetometers, the spacing between the magnetometers and the measurement noise. First of all, the results showed that arrays with more magnetometers generally provided more accurate derivative estimates than arrays with fewer. Specifically, arrays with a magnetometer in the center performed relatively well because of the symmetrical distribution of the measurements. In terms of spacing, neither very small nor very large spacings yielded the most accurate derivatives. Small spacings lead to noise-dominated measurements, while large spacings caused the magnetometers to be too far apart, leading to measurements that are no longer representative of the local field changes.

Generally, for arrays up to five by five the optimal spacings lie in a range of approximately $2 < l_s < 7$ for the local GP and $2.5 < l_s < 10$ for the linear least squares (LLS) method. The exact optimal spacing depends on the number of magnetometers and the level of noise: increasing the number of magnetometers or decreasing the noise level shifts the optimal spacing towards tighter spacings. Furthermore, the results showed that arrays with more magnetometers generally performed better for all spacings. The arrays with more magnetometers span a larger area and are better at averaging out the noise, which improves the derivative estimates. As a result, the range of spacings that produce accurate estimates is larger for arrays with more magnetometers. In practice, however, it is advisable to use spacings that are on the higher side of the range, as the smaller spacings are more prone to noise and errors, and in practice it is likely that more measurement errors are present due to calibration and modeling errors. The analysis of the measurement noise confirmed that noise plays a more prominent role for smaller spacings. Lower noise levels slightly shifted the optimal spacing to tighter spacings and improved the overall derivative estimate. Additionally, although the results are mainly based on two-dimensional arrays, it was shown with a preliminary experiment that using a three-dimensional array does not fundamentally change the effects of the array configurations described so far. Since more magnetometers are used in a three-dimensional array, the optimal spacing shifted to smaller spacings, but the underlying relationships remain unchanged. Therefore, the results of two-dimensional arrays probably transfer quite well to three-dimensional arrays when the tighter spacings are taken into account.

The second research sub-questions states: *How do different methods of computing the spatial derivative compare in terms of accuracy and computational efficiency?*. The LLS, local GP and finite difference methods were analyzed. The local GP method yielded the most accurate estimates. The results were better than the LLS method for all variables. The fact that the local GP method showed better performance is partially because it utilizes the exact same model that was used to generate the data, making it aware of the lengthscale and measurement noise. However, based on an experiment in which the hyperparameters varied with a mean absolute perturbation of 20% with respect to the true ones, the results indicated that the estimates were still marginally more accurate compared to the LLS method. Both the LLS and the local GP method significantly outperformed the finite difference method. The finite difference method showed decent performance for large spacings that is comparable to the performance of the other two methods. However, as the spacing became smaller, the performance quickly deteriorated due to the limited number of magnetometers, making it more susceptible to noise. In terms of computational efficiency, the local GP and LLS methods are of the same order of magnitude where the LLS method is typically a little faster, except for large arrays with a high order of approximation. Then the LLS method is slightly slower. The finite difference method is one to two orders of magnitudes faster than the other two. Overall, if computational resources are very limited, the finite difference method might be an option. If not, then the local GP and LLS method are more suitable. Among these, if the hyperparameters can be determined with a fairly accurate accuracy in the entire area of interest, the local GP is likely the better choice. However, if a more straightforward implementation and setup is desired, the LLS method may be the preferred choice.

The third research sub-question addresses: *How does using a single measurement combined with the spatial derivative affect the accuracy and computational efficiency of magnetic field mapping compared to using all measurements on the array?*. The first part of this sub-question focuses on: "What is the outcome of this comparison for global maps?". The results showed

that the derivative kernel approximates the information captured by the array which is used in its entirety by the full kernel. This holds for all combinations of array configurations considered. In particular, array configurations that performed well in estimating the spatial derivative also yielded the highest map quality. Naturally, the full kernel provided a higher map quality by explicitly using all magnetometer measurements. Excluding the required time to compute the spatial derivative, the computational efficiency of the derivative kernel is constant for all arrays with different numbers of magnetometers. The full kernel using a two-by-two grid has an almost identical computation time as the derivative kernel. However, using their optimal spacings, the full kernel with a two-by-two grid has a significantly better global map quality than the derivative kernel with a five-by-five grid. The highest map quality for the full kernel is generally obtained for the largest considered spacing of $l_s = 1.4$ and as a result uses the four magnetometers as highly valuable individual measurements, which measure the magnetic field closely to the test points. But even for smaller spacings that do not come close to overlapping the test point, does the full kernel yield equal or higher map quality. This shows that in the case of computational constraints, it is more favorable to use fewer magnetometers or just collect more data points instead of using the spatial derivative.

Finally, the derivative and full kernel were also applied to a local map to answer the second part of the third research sub-question: *What is the outcome of this comparison for local maps?*. Since for the local odometry map the full kernel no longer benefits from an array with very large spacings, the gap in map quality between the derivative kernel and the full kernel is reduced. Additionally, since the optimal spacing for the full kernel depends more on the step size, a compromise must be made to select a spacing that performs reasonably well for all possible step sizes. This trade-off leads to a spacing that has sub-optimal performance for most step sizes. In this case, a five-by-five array with a spacing of approximately $10 < l_s < 16.7$ yields the best overall performance. The derivative kernel does not suffer as much from the variable map quality due to different step sizes as the full kernel and produces the highest overall map quality for a five-by-five array with a spacing in the range of $16.7 < l_s < 50$. Nevertheless, for a two-by-two grid with optimal spacing, the full kernel still yields a higher local map quality compared to the derivative kernel with optimal configuration. Although the difference in map quality between the full and derivative kernels is less than for the global map, it is still more advantageous to use the full kernel with a grid with fewer magnetometers in case of computational constraints.

To answer the main research question, two approaches for utilizing the magnetic field measurements to create magnetic field maps were analyzed in this thesis: the full kernel and the derivative kernel. The derivative kernel served as an approximation to the full kernel while reducing the computational load. However, neither for local odometry nor global mapping did the derivative kernel prove to be a better alternative to the full kernel. It did not sufficiently reduce the computational burden to justify the resulting loss in map quality. The full kernel provided higher map qualities for an equivalent computation time. More spatially distributed measurements were found to be more useful than additional local information. The optimal array configuration varied between kernel and map types. Naturally, more magnetometers generally led to better map quality. For the derivative kernel, the best results for the global were achieved with arrays that included a magnetometer in the center, producing the most accurate spatial derivatives and the highest map quality. Typically, the optimal spacing became smaller as the number of magnetometers increased or the noise decreased. In case of the local map, the spacing influences the location of the test points which shifts the optimal spacing to

smaller spacings compared to the global map. For the full kernel, the best global map quality was obtained with very large spacings, as each measurement contributed completely new and relatively uncorrelated information. For the local map quality, the full kernel showed that the optimal spacing depends on the step size, suggesting that the spacing should be chosen based on the most common range of step sizes. Increasing the number of magnetometers decreased the optimal spacing to find a new balance in the trade-off between achieving sufficient overlap between the test and training points and maintaining a high resolution.

The results related to the accuracy with which the spatial derivative can be determined for different array configurations can be valuable to the derivative-based odometry method, as this method directly computes the derivative. More accurate derivative estimates should reduce the drift rate. Additionally, incorporating the results of the full kernel using the local map into the model-based odometry method should provide modest improvements. This is because this method relies on evaluating the magnetic field at the next time step based on the local model. More accurate predictions could reduce the drift rate.

While this thesis provides valuable insights into the use of magnetometer arrays, specifically the use of the spatial derivative, several opportunities for further research can be expanded upon. The first suggestion is to examine whether the derivative kernel might yield higher-quality maps in a three-dimensional setup relative to the full kernel. The spatial derivative probably still provides the same amount of information relative to the larger space, whereas the information of an array with four magnetometers has become significantly more sparse. To accurately compute the spatial derivative in three dimensions, however, the array configuration should also probably be extended into a three-dimensional structure. Furthermore, the optimal array configurations for a three-dimensional array would also have to be studied.

The second suggestion concerns the determination of the spatial derivative. Instead of relying on computing the spatial derivative at each time step again, an extended Kalman filter (EKF) can be used to filter the derivative [42]. This requires the dynamics of the spatial derivative, which naturally includes the second-order spatial derivative. Fortunately, an estimate of this derivative can also be directly estimated from the LLS method. In [42] it was demonstrated using simulations that filtering the spatial derivative significantly improved the accuracy of the spatial derivative, as well as the position estimation.

The final suggestion is to analyze the effect of incorporating the divergence-free property into the model and data along with the curl-free property. This thesis solely included the curl-free property in the models, and as a result also in the generated simulation data. The potential effects of the divergence-free property on optimal array configurations and map qualities remain unknown.

Appendix A

Curl-Free Derivative Kernel Derivation

This appendix provides additional information to Section 3-1-2 on how the curl-free derivative kernel is obtained.

In order to use the spatial derivative as a measurement for Gaussian process regression (GPR) the joint distribution of the magnetic field and the spatial derivative must be derived. This distribution is given by

$$\begin{aligned}\begin{bmatrix}\mathbf{H}(\mathbf{x})\\ \mathbf{H}_d(\mathbf{x})\end{bmatrix} &\sim \mathcal{N}\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}K_{\text{curl}}(\mathbf{x},\mathbf{x}') & K_{\text{curl}}(\mathbf{x},\mathbf{x}')\mathcal{D}_{\mathbf{x}'}^{\mathsf{T}}\\ \mathcal{D}_{\mathbf{x}}K_{\text{curl}}(\mathbf{x},\mathbf{x}') & \mathcal{D}_{\mathbf{x}}K_{\text{curl}}(\mathbf{x},\mathbf{x}')\mathcal{D}_{\mathbf{x}'}^{\mathsf{T}}\end{bmatrix}\right)\\ &= \mathcal{N}\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, K_{\text{curl,d}}(\mathbf{x},\mathbf{x}')\right).\end{aligned} \tag{A-1}$$

The mean of this distribution is a vector of zeros as the mean offset is captured by the curl-free kernel. The elements of $K_{\text{curl,d}}(\mathbf{x},\mathbf{x}')$ are obtained by applying the derivative operator on the curl-free kernel. For a clear overview, the kernel of the curl-free Gaussian process (GP) is given here again

$$K_{\text{curl}}(\mathbf{x},\mathbf{x}') = \sigma_{\text{lin.}}^2 I_3 + \sigma_f^2\left(I_3 - \left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)^{\mathsf{T}}\right)\exp\left(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\right). \tag{A-2}$$

A-1 First-Order Derivative of Curl-Free Kernel

Since reporting the full derivative of the curl-free kernel with respect to all components of $\mathbf{x}$ is too extensive, they are given with respect to a single arbitrary component. As a result, instead of giving $\mathcal{D}_{\mathbf{x}}K_{\text{curl}}(\mathbf{x},\mathbf{x}')$, $\frac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_p}$ is computed, where $\mathbf{x}_p$ denotes the $p^{\text{th}} \in \{0,1,2\}$ component of $\mathbf{x}$. To compute the derivative of matrices, first the $(i,j)^{\text{th}}$ component is given to subsequently extend them to the full matrix derivative. To simplify the process of computing the full derivative of the curl-free kernel in Eq. (A-2), the derivatives of the

individual components are computed. The $(i,j)^{\text{th}}$ component of the first relevant matrix term is given by

$$\left(\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)^{\mathsf{T}}\right)_{i,j} = \left(\frac{\mathbf{x}_i-\mathbf{x}'_i}{l}\right)\left(\frac{\mathbf{x}_j-\mathbf{x}'_j}{l}\right)^{\mathsf{T}}, \tag{A-3}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ denote the $i^{\text{th}} \in \{0,1,2\}$ and $j^{\text{th}} \in \{0,1,2\}$ components of $\mathbf{x}$ respectively.

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}_p}\left(\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)^{\mathsf{T}}\right)_{i,j} &= \frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\mathbf{x}_i-\mathbf{x}'_i}{l}\right)\left(\frac{\mathbf{x}_j-\mathbf{x}'_j}{l}\right)^{\mathsf{T}} \\ &= \begin{cases} \dfrac{\mathbf{x}_j-\mathbf{x}'_j}{l^2}, & \text{if } p=i, \\ \dfrac{\mathbf{x}_i-\mathbf{x}'_i}{l^2}, & \text{if } p=j, \\ 0, & \text{otherwise,} \end{cases} \\ &= \delta_{p,i}\frac{\mathbf{x}_j-\mathbf{x}'_j}{l^2} + \delta_{p,j}\frac{\mathbf{x}_i-\mathbf{x}'_i}{l^2},\end{aligned} \tag{A-4}$$

where $\delta_{p,i}$ is the Kronecker delta function. This element-wise derivative can be written in matrix form as

$$\frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)^{\mathsf{T}} = \mathbf{e}_p\left(\frac{\mathbf{x}-\mathbf{x}'}{l^2}\right)^{\mathsf{T}} + \frac{\mathbf{x}-\mathbf{x}'}{l^2}\mathbf{e}_p^{\mathsf{T}}, \tag{A-5}$$

where $\mathbf{e}_p$ is the standard basis column vector with 1 in the p^{th} position and 0 elsewhere.

The derivative of the scalar exponent term can be computed to be

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}_p}\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big) &= \frac{\partial}{\partial \mathbf{x}_p}\exp\Big(-\frac{(\mathbf{x}-\mathbf{x}')^T(\mathbf{x}-\mathbf{x}')}{2l^2}\Big) \\ &= -\frac{\mathbf{x}_p-\mathbf{x}'_p}{l^2}\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big).\end{aligned} \tag{A-6}$$

Based on the derivatives of these individual components in Eq. (A-5) and Eq. (A-6), the derivative of $K_{\text{curl}}(\mathbf{x},\mathbf{x}')$ with respect to $\mathbf{x}_p$ can be computed using the product rule as

$$\begin{aligned}\frac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_p} =& \frac{\partial}{\partial \mathbf{x}_p}\sigma_f^2\left(I_3\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big) - \Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)\Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)^{\mathsf{T}}\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big)\right) \\ =& \sigma_f^2\left(-I_3\frac{\mathbf{x}_p-\mathbf{x}'_p}{l^2}\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big) - \left(\mathbf{e}_p\Big(\frac{\mathbf{x}-\mathbf{x}'}{l^2}\Big)^{\mathsf{T}} + \frac{\mathbf{x}-\mathbf{x}'}{l^2}\mathbf{e}_p^{\mathsf{T}}\right)\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big)\dots\right. \\ &\left. + \Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)\Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)^{\mathsf{T}}\frac{\mathbf{x}_p-\mathbf{x}'_p}{l^2}\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big)\right) \\ =& \frac{\sigma_f^2}{l^2}\left(-I_3(\mathbf{x}_p-\mathbf{x}'_p) - \mathbf{e}_p(\mathbf{x}-\mathbf{x}')^{\mathsf{T}} - (\mathbf{x}-\mathbf{x}')\mathbf{e}_p^{\mathsf{T}} + \Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)\Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)^{\mathsf{T}}(\mathbf{x}_p-\mathbf{x}'_p)\right)\cdot\dots \\ &\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big).\end{aligned} \tag{A-7}$$

This provides the first part of the $K_{\text{curl,d}}(\mathbf{x}, \mathbf{x}')$ kernel. When taking the derivative with respect to $\mathbf{x}'_p$ one can observe in Eq. (A-2) that the difference between taking the derivative with respect to $\mathbf{x}'_p$ or $\mathbf{x}_p$ is a minus sign. As a result, the derivative with respect to $\mathbf{x}'_p$ is equal to the derivative with respect to $\mathbf{x}_p$ up to a minus sign

$$\frac{\partial K_{\text{curl}}(\mathbf{x}, \mathbf{x}')}{\partial \mathbf{x}'_p} = -\frac{\partial K_{\text{curl}}(\mathbf{x}, \mathbf{x}')}{\partial \mathbf{x}_p}. \tag{A-8}$$

A-2 Second-Order Derivative of Curl-Free Kernel

The only remaining derivative to compute is the second-order derivative of the curl-free kernel with respect to both $\mathbf{x}_p$ and $\mathbf{x}'_q$, where $\mathbf{x}'_q$ denotes the $q^{\text{th}} \in \{0, 1, 2\}$ component of $\mathbf{x}'$. To simplify this derivation, the derivative of individual components in Eq. (A-7) are computed first.

$$\frac{\partial(\mathbf{x}_p - \mathbf{x}'_p)}{\partial \mathbf{x}'_q} = -\delta_{p,q} = -\mathbf{e}_p^\intercal \mathbf{e}_q. \tag{A-9}$$

The next term is a matrix term again for which the derivative of the $(i, j)^{\text{th}}$ component is computed first.

$$\frac{\partial\Big(\mathbf{e}_p(\mathbf{x} - \mathbf{x}')^\intercal\Big)_{i,j}}{\partial \mathbf{x}'_q} = \frac{\partial(\mathbf{e}_p)_i(\mathbf{x}_j - \mathbf{x}'_j)}{\partial \mathbf{x}'_q} = -(\mathbf{e}_p)_i \delta_{q,j} = -\delta_{p,i}\delta_{q,j}. \tag{A-10}$$

In matrix form this derivative is given by

$$\frac{\partial \mathbf{e}_p(\mathbf{x} - \mathbf{x}')^\intercal}{\partial \mathbf{x}'_q} = -\mathbf{e}_p \mathbf{e}_q^\intercal. \tag{A-11}$$

Similarly, the derivative of the term $(\mathbf{x} - \mathbf{x}')\mathbf{e}_p^\intercal$ is given by

$$\frac{\partial(\mathbf{x} - \mathbf{x}')\mathbf{e}_p^\intercal}{\partial \mathbf{x}'_q} = -\mathbf{e}_q \mathbf{e}_p^\intercal. \tag{A-12}$$

The derivative of Eq. (A-5) with respect to $\mathbf{x}'_q$ instead of $\mathbf{x}_p$ is slightly different

$$\frac{\partial}{\partial \mathbf{x}'_q}\left(\frac{\mathbf{x} - \mathbf{x}'}{l}\right)\left(\frac{\mathbf{x} - \mathbf{x}'}{l}\right)^\intercal = -\mathbf{e}_q\left(\frac{\mathbf{x} - \mathbf{x}'}{l^2}\right)^\intercal - \frac{\mathbf{x} - \mathbf{x}'}{l^2}\mathbf{e}_q^\intercal. \tag{A-13}$$

The derivative of Eq. (A-6) with respect to $\mathbf{x}'_q$ instead of $\mathbf{x}_p$ is equal up to a factor of minus one. Utilizing these derivative terms, the derivative of Eq. (A-7) can be computed using the

product rule as

$$\begin{aligned}\frac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_p \mathbf{x}'_q} =& \frac{\sigma_f^2}{l^2}\Bigg(I_3 \mathbf{e}_p^\intercal \mathbf{e}_q + \mathbf{e}_p \mathbf{e}_q^\intercal + \mathbf{e}_q \mathbf{e}_p^\intercal - \Big(\mathbf{e}_q \Big(\frac{\mathbf{x}-\mathbf{x}'}{l^2}\Big)^\intercal + \frac{\mathbf{x}-\mathbf{x}'}{l^2}\mathbf{e}_q^\intercal\Big)(\mathbf{x}_p - \mathbf{x}'_p)\dots \\ & - \Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)\Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)^\intercal \mathbf{e}_p^\intercal \mathbf{e}_q\Bigg) \exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big)\dots \\ & + \frac{\sigma_f^2}{l^2}\Bigg(-I_3(\mathbf{x}_p - \mathbf{x}'_p) - \mathbf{e}_p(\mathbf{x}-\mathbf{x}')^\intercal - (\mathbf{x}-\mathbf{x}')\mathbf{e}_p^\intercal + \Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)\Big(\frac{\mathbf{x}-\mathbf{x}'}{l}\Big)^\intercal (\mathbf{x}_p - \mathbf{x}'_p)\Bigg)\cdot \dots \\ & \frac{\mathbf{x}_q - \mathbf{x}'_q}{l^2}\exp\Big(-\frac{||\mathbf{x}-\mathbf{x}'||^2}{2l^2}\Big).\end{aligned} \tag{A-14}$$

Finally, using the derivatives from Eq. (A-7), Eq. (A-8) and Eq. (A-14), the covariance of the full joint distribution between the magnetic field and the spatial derivative is given by

$$\begin{aligned}K_{\text{curl,d}}(\mathbf{x},\mathbf{x}') &= \left[\begin{array}{c|c} K_{\text{curl}}(\mathbf{x},\mathbf{x}') & K_{\text{curl}}(\mathbf{x},\mathbf{x}')\mathcal{D}_{\mathbf{x}'}^\intercal \\ \hline \mathcal{D}_{\mathbf{x}} K_{\text{curl}}(\mathbf{x},\mathbf{x}') & \mathcal{D}_{\mathbf{x}} K_{\text{curl}}(\mathbf{x},\mathbf{x}')\mathcal{D}_{\mathbf{x}'}^\intercal \end{array}\right] \\ &= \left[\begin{array}{c|ccc} K_{\text{curl}}(\mathbf{x},\mathbf{x}') & -\dfrac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_0} & -\dfrac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_1} & -\dfrac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_2} \\ \hline \dfrac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_0} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_0 \mathbf{x}'_0} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_0 \mathbf{x}'_1} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_0 \mathbf{x}'_2} \\ \dfrac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_1} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_1 \mathbf{x}'_0} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_1 \mathbf{x}'_1} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_1 \mathbf{x}'_2} \\ \dfrac{\partial K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_2} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_2 \mathbf{x}'_0} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_2 \mathbf{x}'_1} & \dfrac{\partial^2 K_{\text{curl}}(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}_2 \mathbf{x}'_2} \end{array}\right].\end{aligned} \tag{A-15}$$

Appendix B

Dimensionless Variables

B-1 Lengthscale-To-Spacing Ratio

This section will show that as long as the l_s ratio stays remains the same, the relative accuracy with which the spatial derivative can be determined also remains constant. This has the advantage that the result from this thesis can be applied to any magnetic field with any particular lengthscale. To show that the relative accuracy indeed remains the same, two different simulations are run. One involves determining the variance of the derivative error estimate for a lengthscale of $l = 0.5$ and the standard magnetometer spacings that are shown in Table 3-2 using the analytical solution from the Gaussian process (GP). This gives rise to the standard l_s ratios that are also depicted in Table 3-2. The other simulation used a lengthscale and magnetometers spacings that are half the size. This ensures that the l_s ratios remain the same, while the magnetic field varies more quickly over shorter distances. Figure B-1 illustrates the results, where the dashed lines indicate the larger lengthscale and the solid lines the smaller lengthscale. The figure shows that the variances between the simulations with the different lengthscales have the same shape but that size is exactly four times larger. However, one also has to take into account the magnitude of the spatial derivatives. Decreasing the lengthscale by a factor of two causes the magnetic field to change more rapidly, resulting in the magnitude of the spatial derivative increasing by a factor of two. As a consequence, the variance of the spatial derivative increases by a factor of four. Since both the variance of the spatial derivative itself and the estimate error scale by the same factor, the relative accuracy stays constant.

B-2 Noise-To-Signal Ratio

Similar to Section B-1, this section will show that as long as the noise-to-signal ratio (NSR) remains the same, the relative accuracy with which the spatial derivative can be determined also remains the same. This makes sure that the results from this thesis concerning the different measurement noise values also apply to magnetic fields with different signal variances.

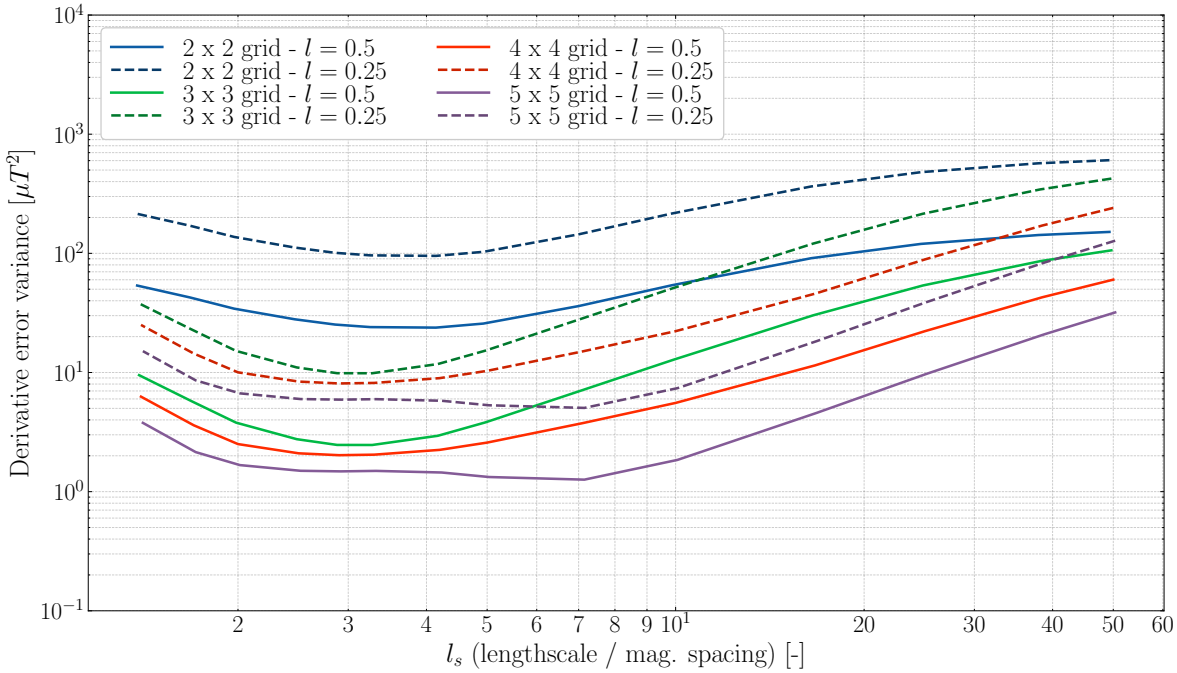

Figure B-1: Comparison between the effect of different lengthscales while keeping the l_s constant. The mean variance of the derivative estimate error is computed analytically using the local GP method. The darker and dashed lines indicate the simulation with the smaller lengthscale and the lighter and solid lines the larger lengthscale.

Two simulations with different measurement noise variances and signal variances were run, while keeping the NSR constant to verify this. These results are very similar to the analysis of l_s in Section B-1 and the plots are, in fact, identical. Figure B-2 shows the results, where the darker and dashed lines indicate the simulations with $\sigma_y = 1\mu T$ and $\sigma_f = 10\mu T$ and the lighter and solid lines the simulations with $\sigma_y = 0.5\mu T$ and $\sigma_f = 5\mu T$, both resulting in a NSR of 1%.

When examining the figure, it can be seen that the shapes of variances for both simulations are exactly the same. Since the standard deviation of the signal, represented by the dashed lines, scales by a factor of two, the magnitude/amplitude of the magnetic field signal scales by the same factor. As a consequence, the magnitude/amplitude of the spatial derivative also scales by this factor. This results in the fact that the variance of the derivative estimate error scales by a factor of $2^2 = 4$. This is the factor with which the variance of the dashed lines scales compared with solid lines. As the variance of the spatial derivative itself and the variance of the derivative estimate error scale by the same amount, the relative accuracy of the estimate remains the same.

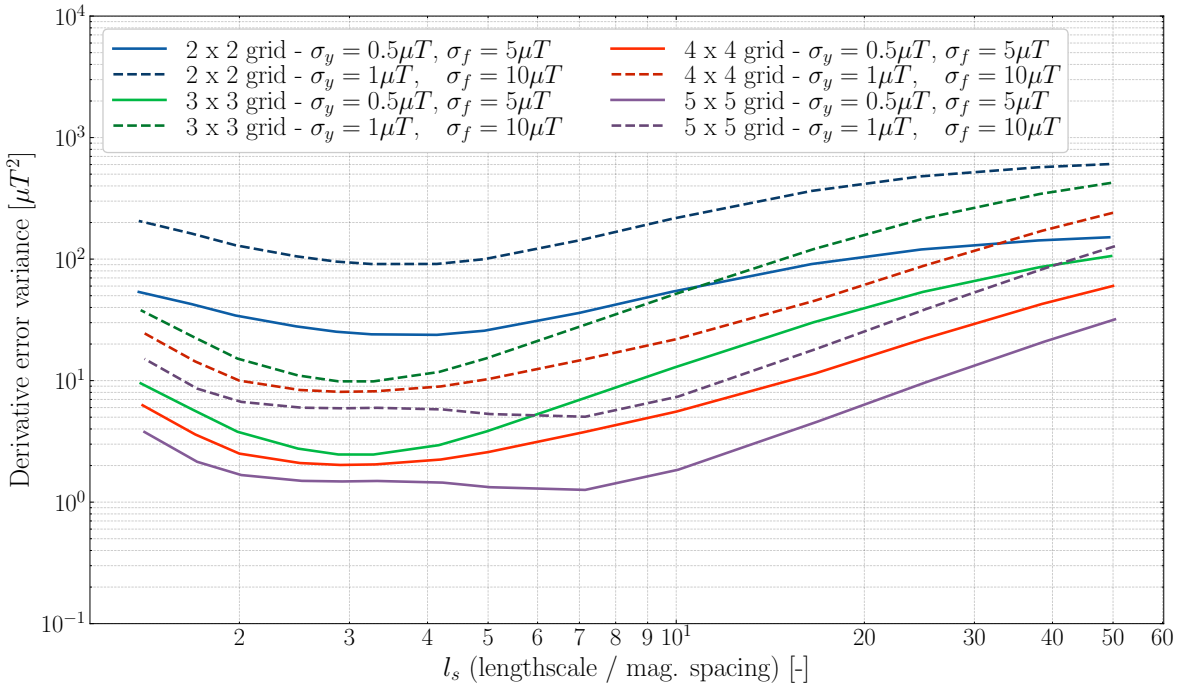

Figure B-2: Comparison between the effect of different values of the measurement noise variance and the signal variance while keeping the NSR constant. The mean variance of the derivative estimate error is computed analytically using the local GP method. The darker and dashed lines indicate the simulations with a $\sigma_y = 1\mu T$ and a $\sigma_f = 10\mu T$ and the lighter and solid lines the simulations with a $\sigma_y = 0.5\mu T$ and a $\sigma_f = 5\mu T$. All simulations have an NSR of 1%.

Appendix C

Additional Results

This appendix provides additional results that were too specific and not relevant enough for the main sections of the thesis.

Figure C-1 shows the variance of the spatial derivative estimate error for all individual components of the spatial derivative for arrays with different numbers of magnetometers using the linear least squares (LLS) method.

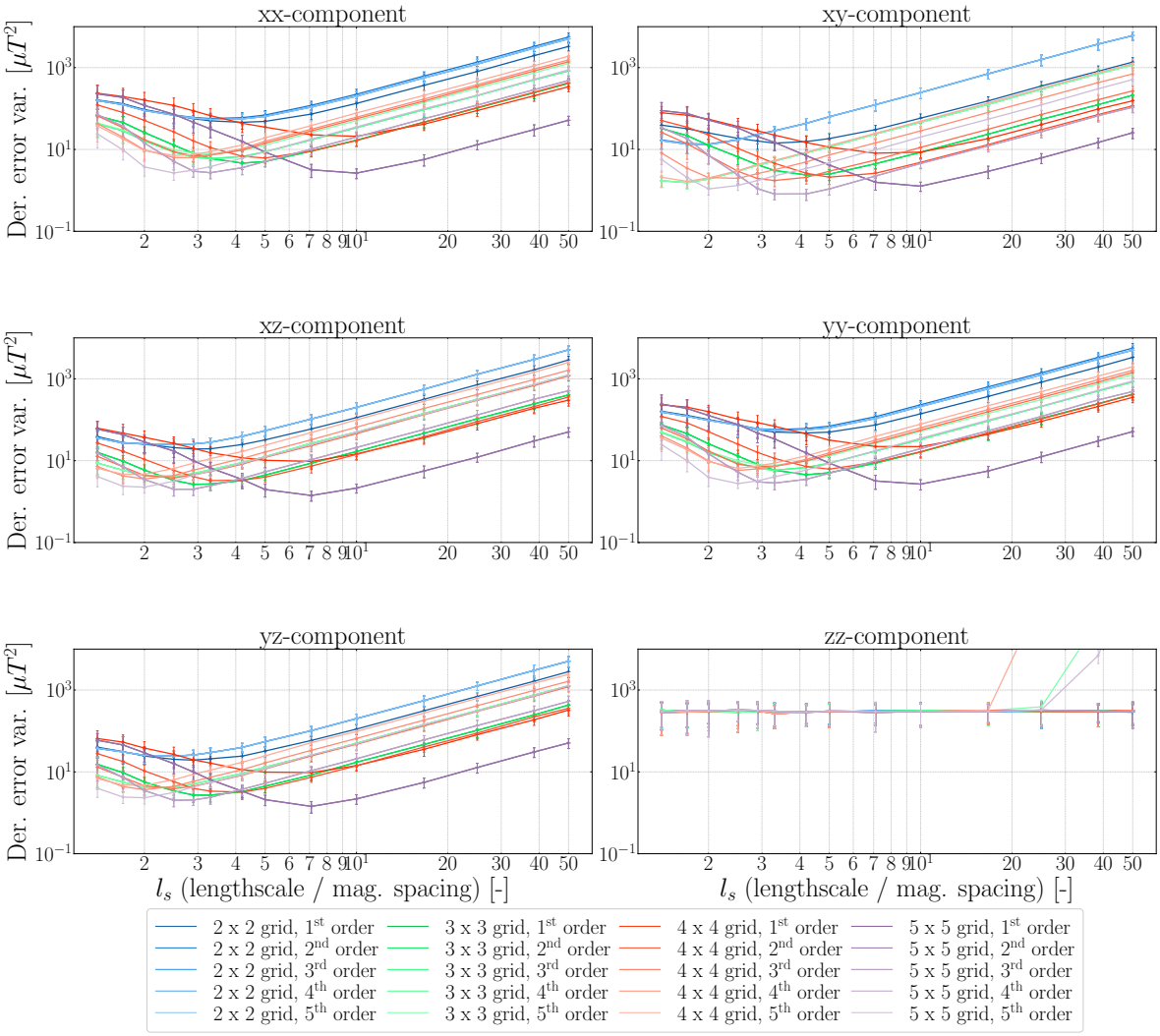

Figure C-1: Relationship between the variance of the spatial derivative estimate error and l_s using LLS method for all components. The relationship is plotted for the different order of approximations and the different number of magnetometers on the array. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu T$ (noise-to-signal ratio (NSR)=1%).

Figure C-2 shows the variance of the spatial derivative estimate error for all individual components of the spatial derivative for arrays with different numbers of magnetometers using the LLS method.

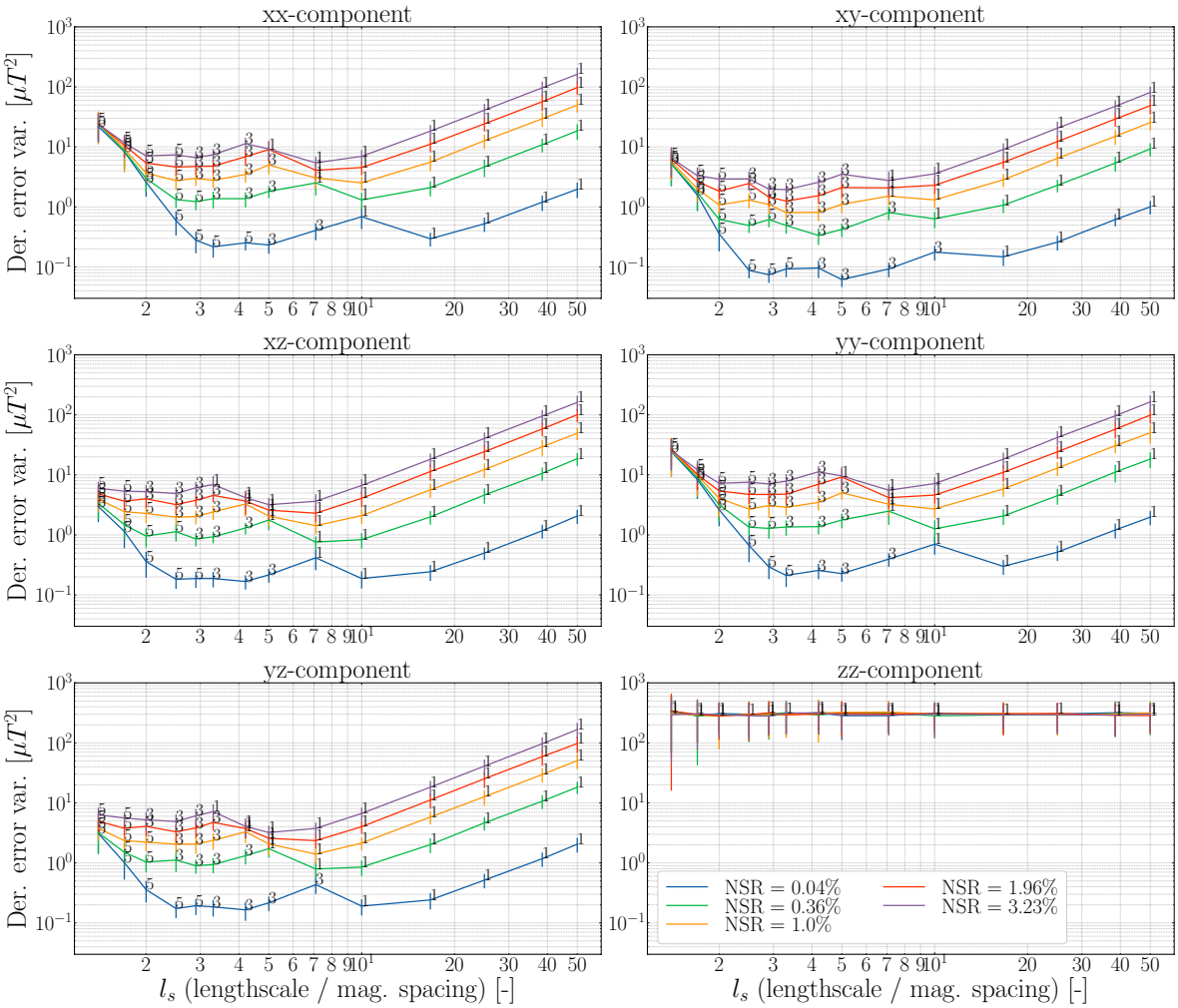

Figure C-2: Relationship between the variance of the spatial derivative estimate error and l_s using LLS method with the order of approximation that gave rise to the lowest variance. The relationship is plotted for different measurement noise values. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu T$ (NSR=1%).

Figure C-3 shows the variance of the spatial derivative estimate error for all individual components of the spatial derivative for different noise levels using the LLS method.

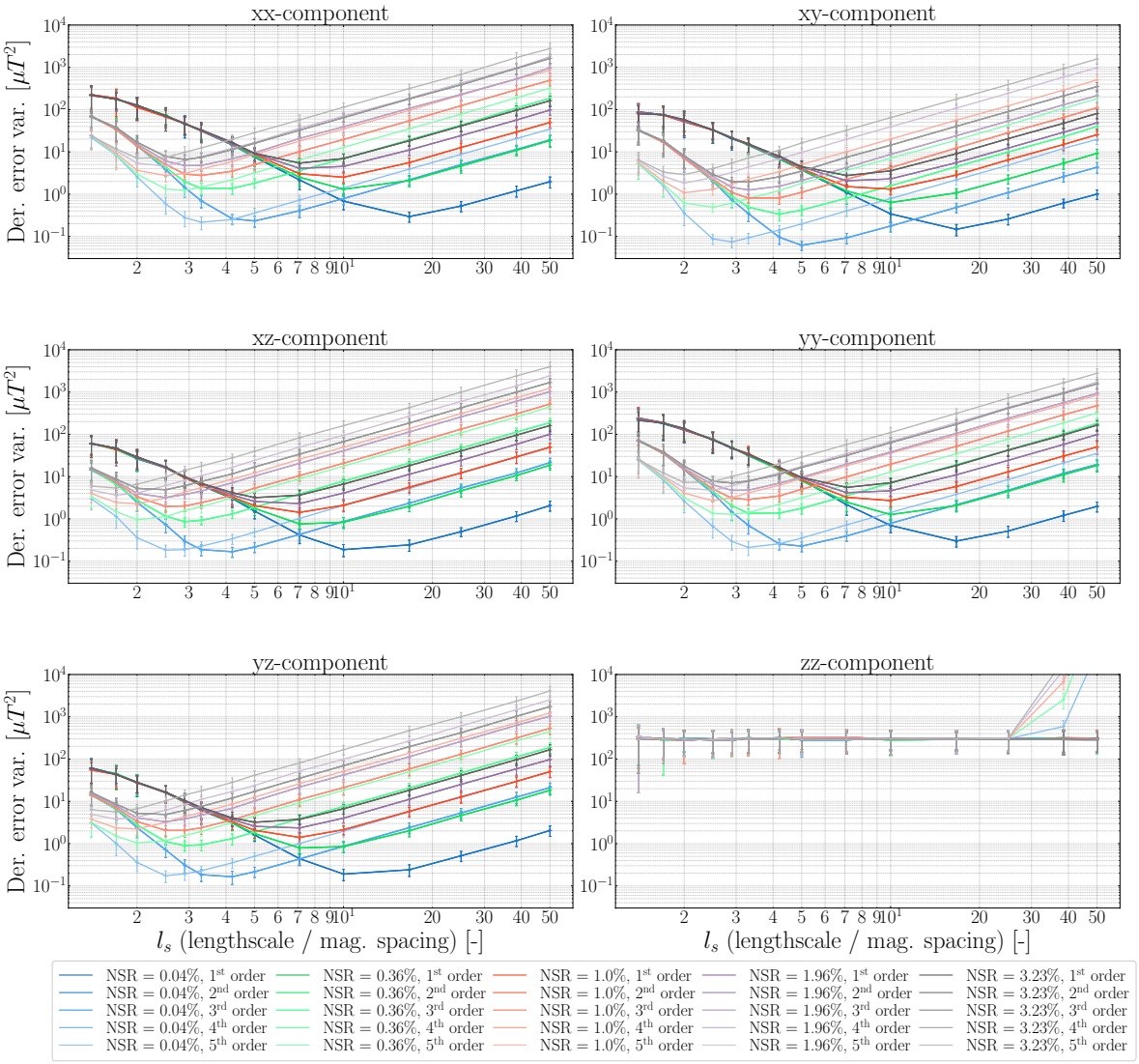

Figure C-3: Relationship between the variance of the spatial derivative estimate error and l_s using LLS method with the order of approximation that gave rise to the lowest variance. The relationship is plotted for different number of magnetometers on the array. The grid size for all configurations was set to five by five.

Figure C-4 shows the variance of the spatial derivative estimate error for all individual components of the spatial derivative for different noise levels using the LLS method.

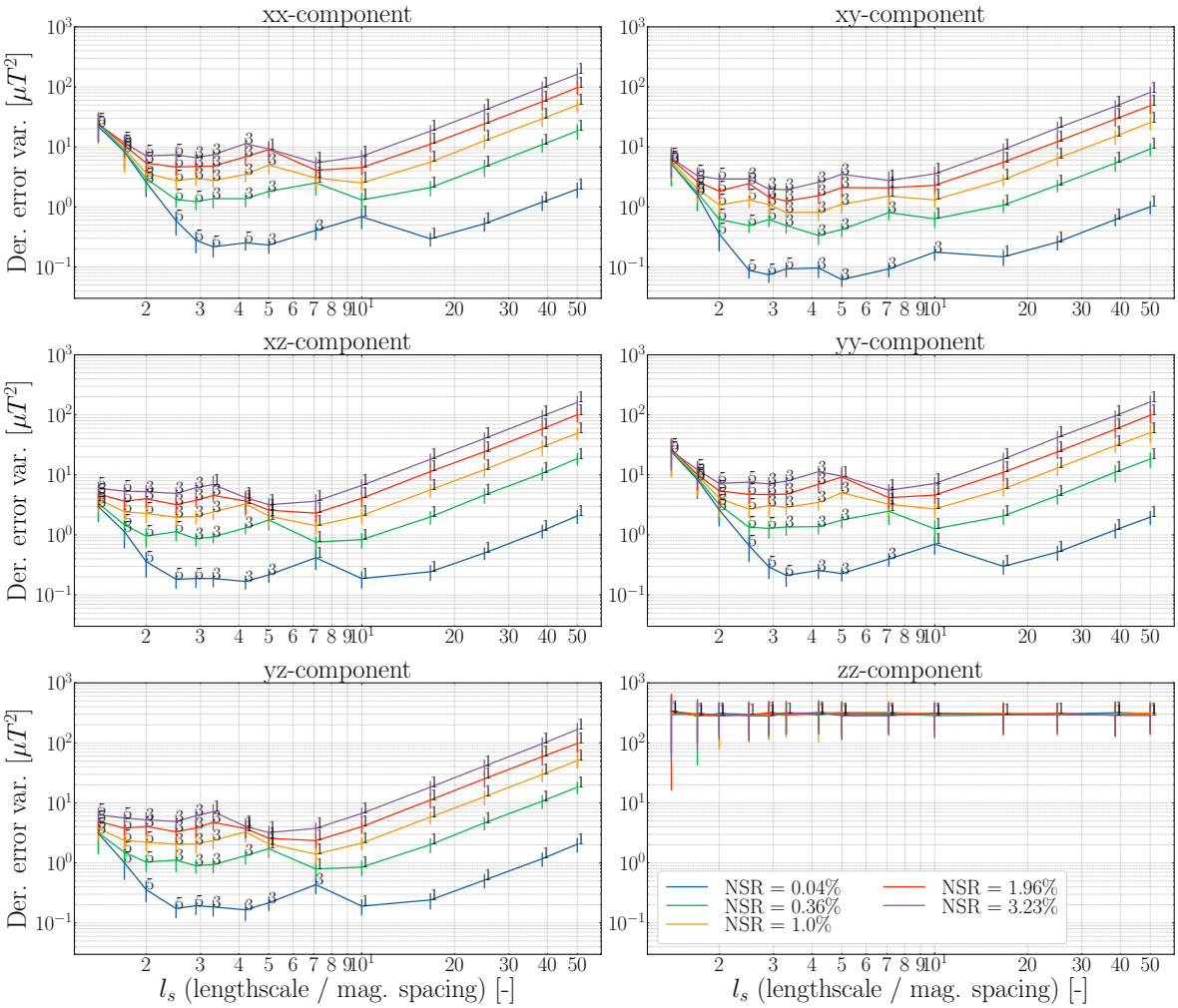

Figure C-4: Relationship between the variance of the spatial derivative estimate error and l_s using LLS method with the order of approximation that gave rise to the lowest variance. The relationship is plotted for different number of magnetometers on the array. The grid size for all configurations was set to five by five.

Figure C-5 shows the mean variance of the spatial derivative estimate error for different noise levels using the finite difference method.

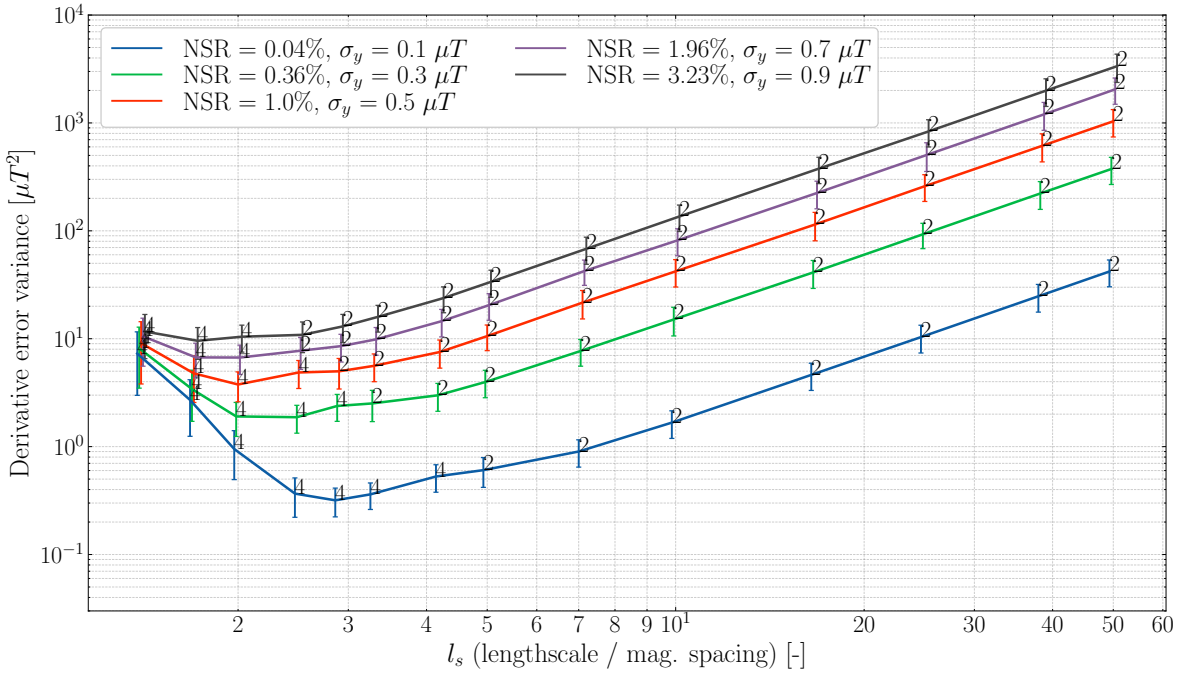

Figure C-5: Relationship between the mean variance of the spatial derivative estimate error and l_s using the finite difference method with the order of approximation that gave rise to the lowest variance. The grid size for all configurations was set to five by five.

Figure C-6 shows the mean variance of the spatial derivative estimate error for different noise levels using the local Gaussian process (GP) method.

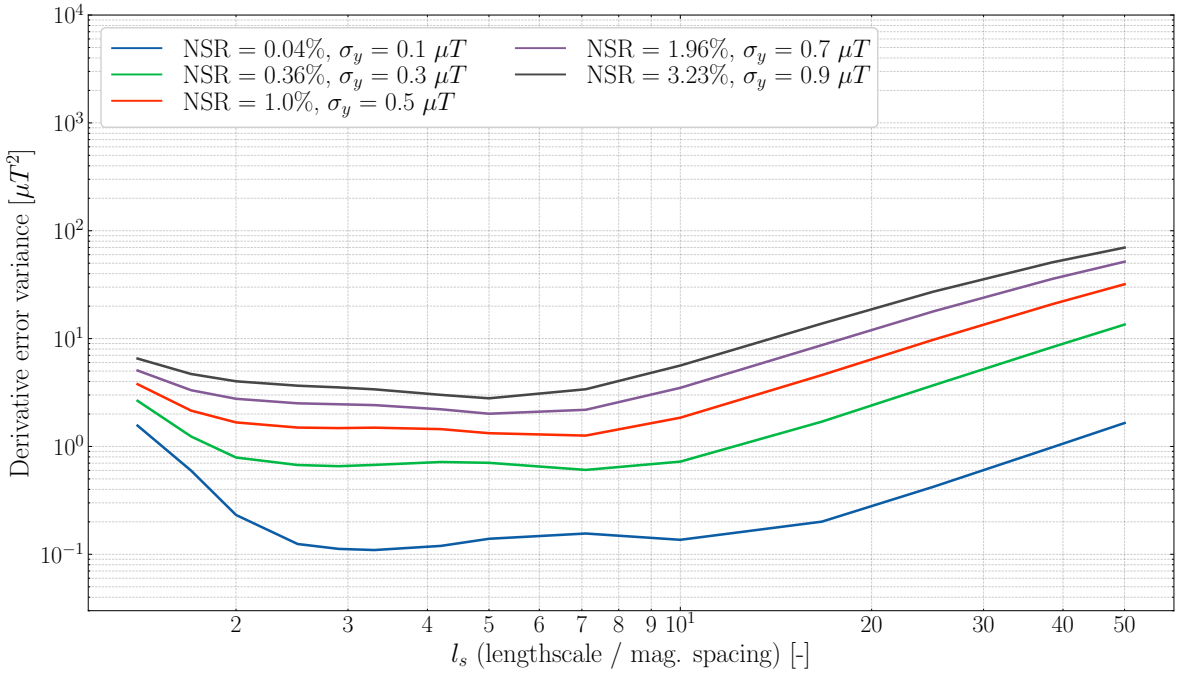

Figure C-6: Relationship between the analytical mean variance of the spatial derivative estimate error and l_s using the local GP method for different noise levels. The grid size for all configurations was set to five by five.

Figure C-7 shows the comparison between using varying hyperparameters and true hyperparameters for all relevant components of the spatial derivative.

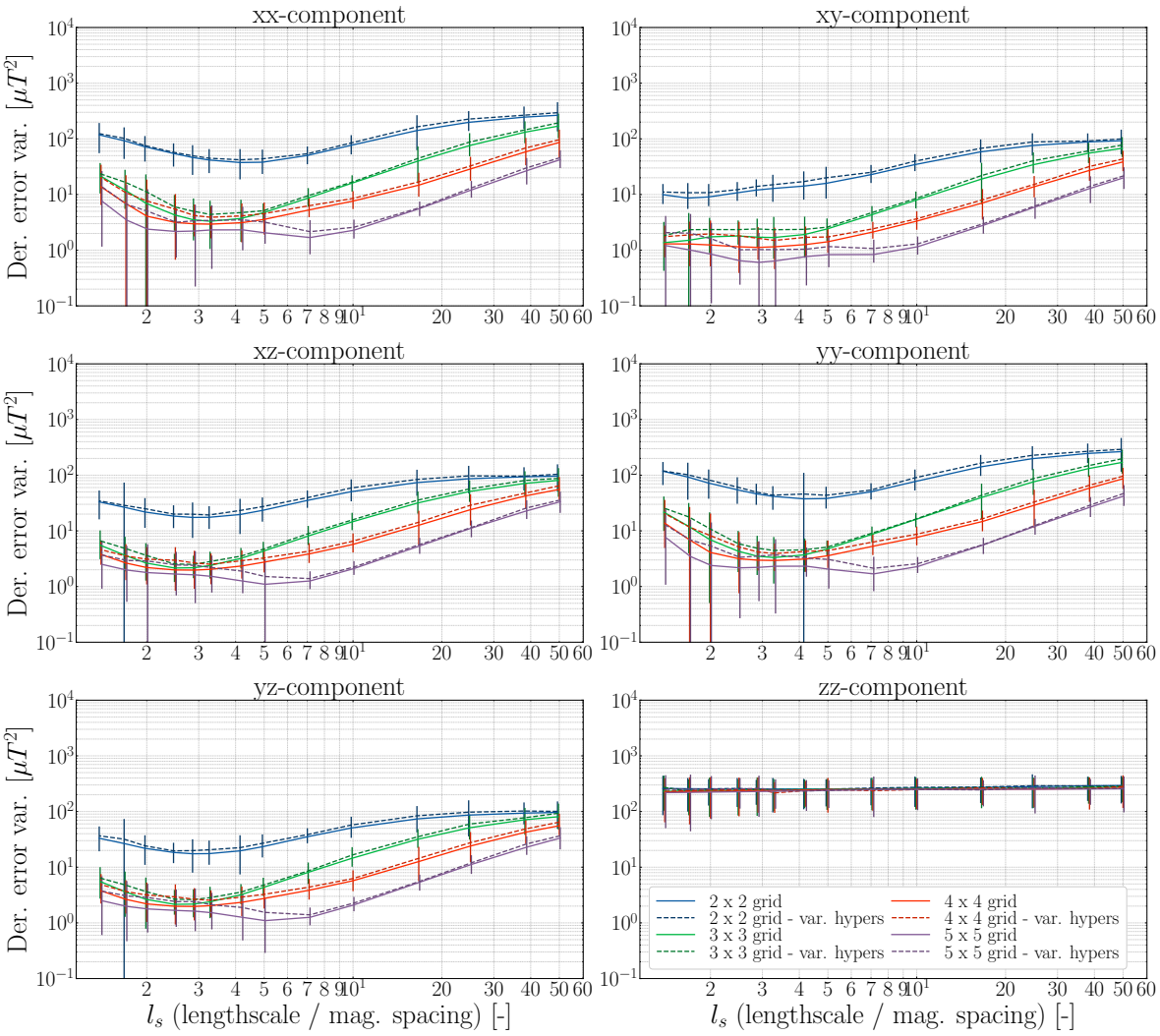

Figure C-7: Comparison of the variance of the spatial derivative estimate error between for all components using varying hyperparameters and true hyperparameters using the local GP method for arrays with different numbers of magnetometers. Only the confidence interval for the simulation with varying hyperparameters is shown, the other simulation with true hyperparameters is based on the analytic variance. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu T$ (NSR=1%).

Figure C-8 shows the comparison between using a three-dimensional cubic structure and using a two-dimensional array when using the local GP method for different numbers of magnetometers.

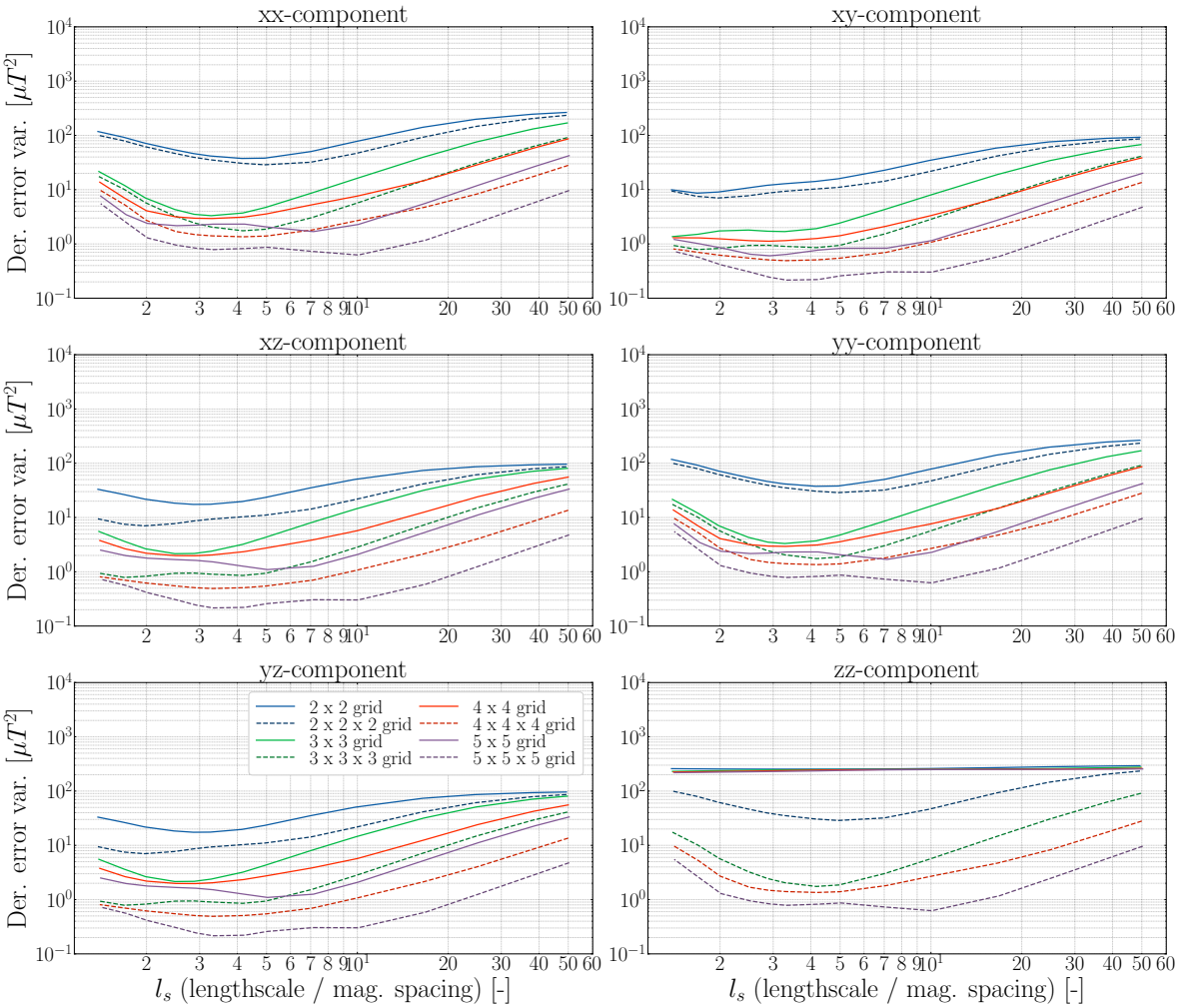

Figure C-8: Comparison of the variance of the spatial derivative estimate error between two- and three-dimensional arrays using the local GP method for all components. The relationship is plotted for different grid sizes. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu T$ (NSR=1%).

Figure C-9 shows the local map quality using the derivative kernel for different levels of measurement noise. It shows that the measurement noise does not significantly shift the optimal spacing. We would expect that the optimal spacing shifts to larger spacings as the noise increases. This effect can perhaps be observed for larger step sizes, as for those the influence of the spatial derivative is larger. However, the differences are small and too noisy to conclude for certain.

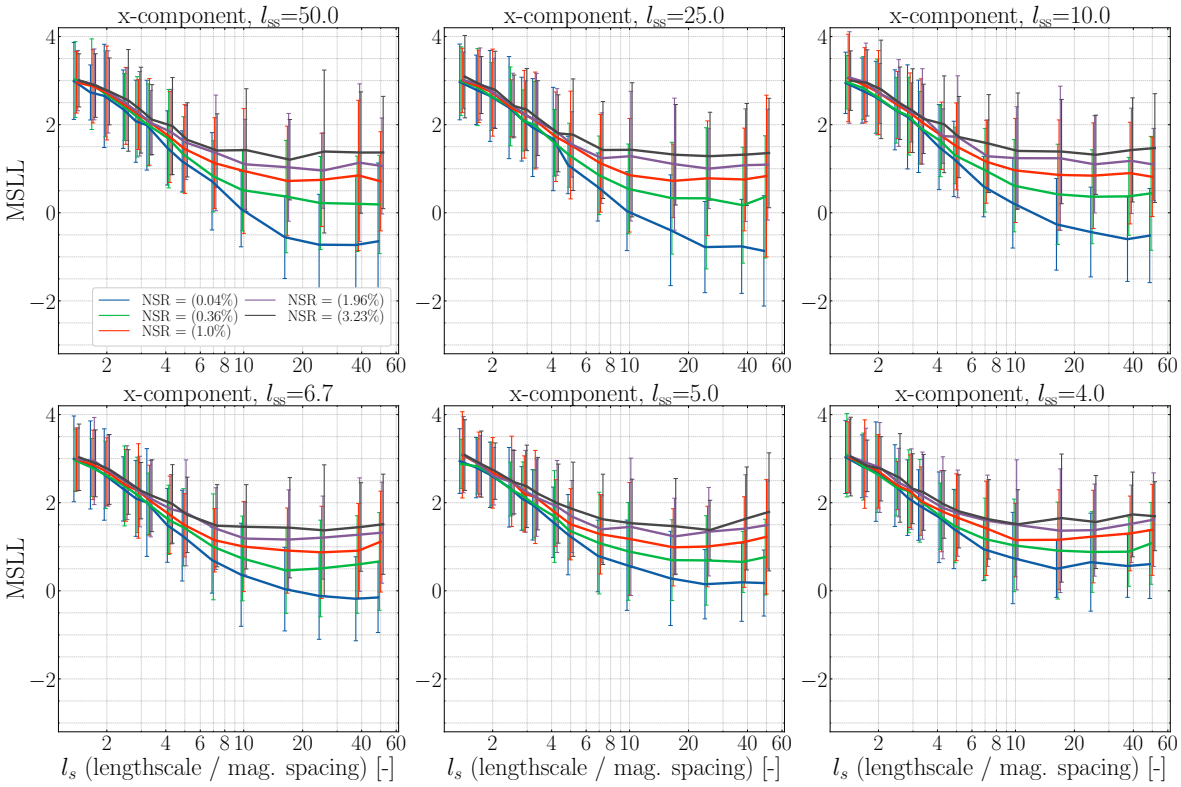

Figure C-9: Local map quality using the derivative kernel for arrays with different noise levels and l_s values using the local GP method. The different plots indicate the quality for the different relative step sizes l_{ss}. The grid size for all configurations was set to five by five.

Figure C-10 shows the local map quality using the full kernel for different levels of measurement noise. It shows that the measurement noise shifts the optimal spacing to slightly smaller spacings. This effect can mainly be observed for larger step sizes.

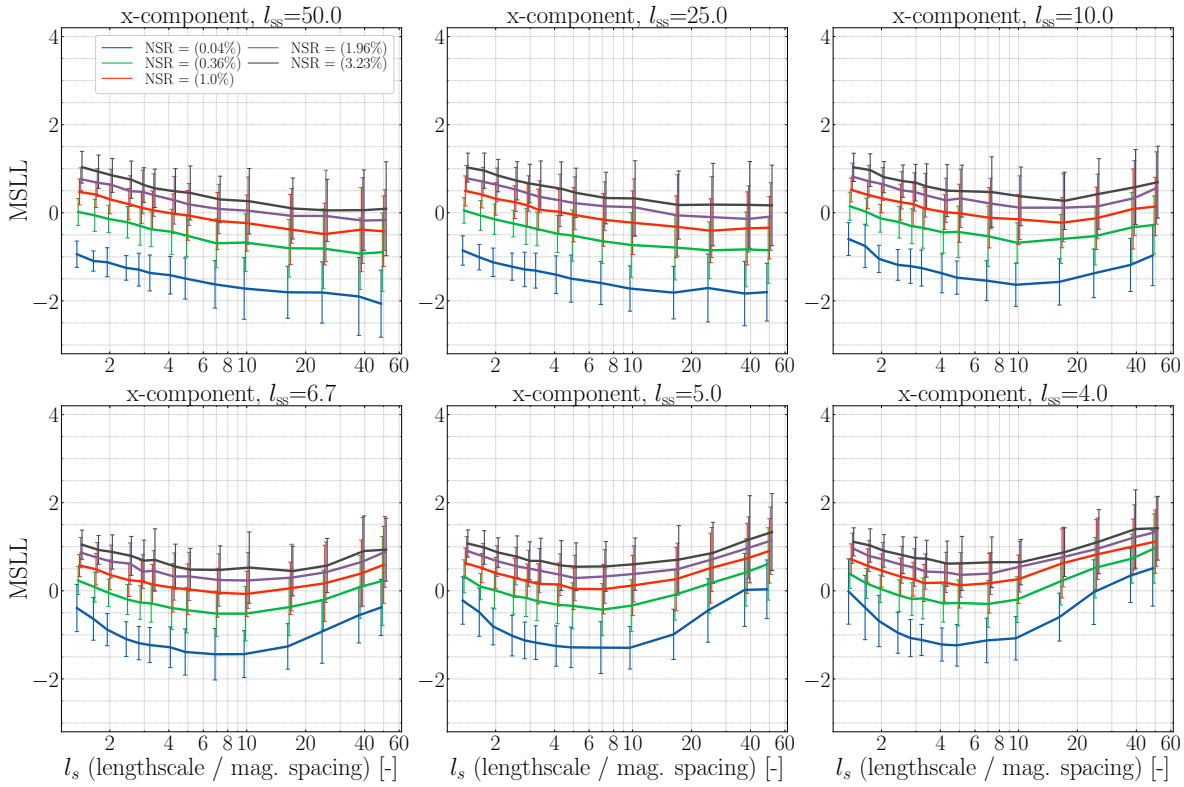

Figure C-10: Local map quality using the full kernel for arrays with different noise levels and l_s values using the local GP method. The different plots indicate the quality for the different relative step sizes l_{ss}. The grid size for all configurations was set to five by five.

C-1 Local Map Quality with High Measurement Noise

If we consider magnetometers with high measurement noise and fix the full kernel to use an array of two by two magnetometers, the derivative kernel has a relative advantage as it can use many more magnetometers and average out some of the noise. Figure C-11 shows the local map quality using the full kernel with a two-by-two array. If we consider the black line with an NSR of 3.23%, an optimal spacing that performs well across all step sizes is $l_{ss} = 4$. This yields a mean MSLL of approximately 1.17.

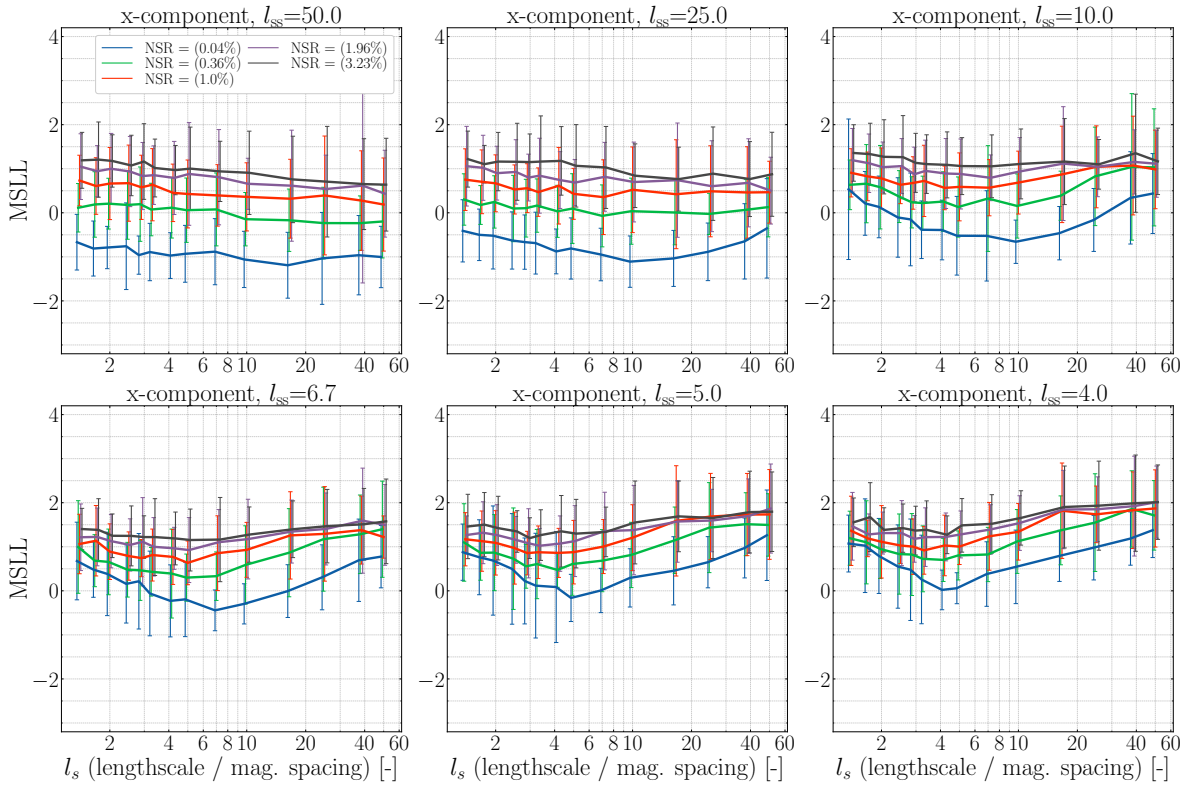

Figure C-11: Local map quality using the full kernel for arrays with different noise levels and l_s values using the local GP method. The different plots indicate the quality for the different relative step sizes l_{ss}. The grid size for all configurations was set to two by two.

When considering the derivative kernel with a five-by-five array and an NSR of 3.23% from Figure C-9, the optimal spacing across the step sizes is approximately given by $l_s = 25$. This yields a mean MSLL of approximately 1.37. The difference in performance is still about the same as for the comparison with the smaller measurement noise in Section 4-2-3. This shows that the difference in the number of magnetometers between the full and derivative kernels should be significantly greater before the derivative kernel outperforms the full kernel.

Appendix D

Local GP Method: Analytical and Manual Variances

Since in the main content of this thesis the spatial derivative estimate using the local Gaussian process (GP) method is evaluated based on the analytical variance provided by the GP, it is necessary to show that this analytical variance is indeed equal to the variance that is computed manually based on 70 simulations demonstrated by Figure 3-1. Figure D-1 shows the comparison between the manually computed variance and the analytical variance. When comparing the dotted to the solid lines, one can observe that the variances are indeed virtually identical.

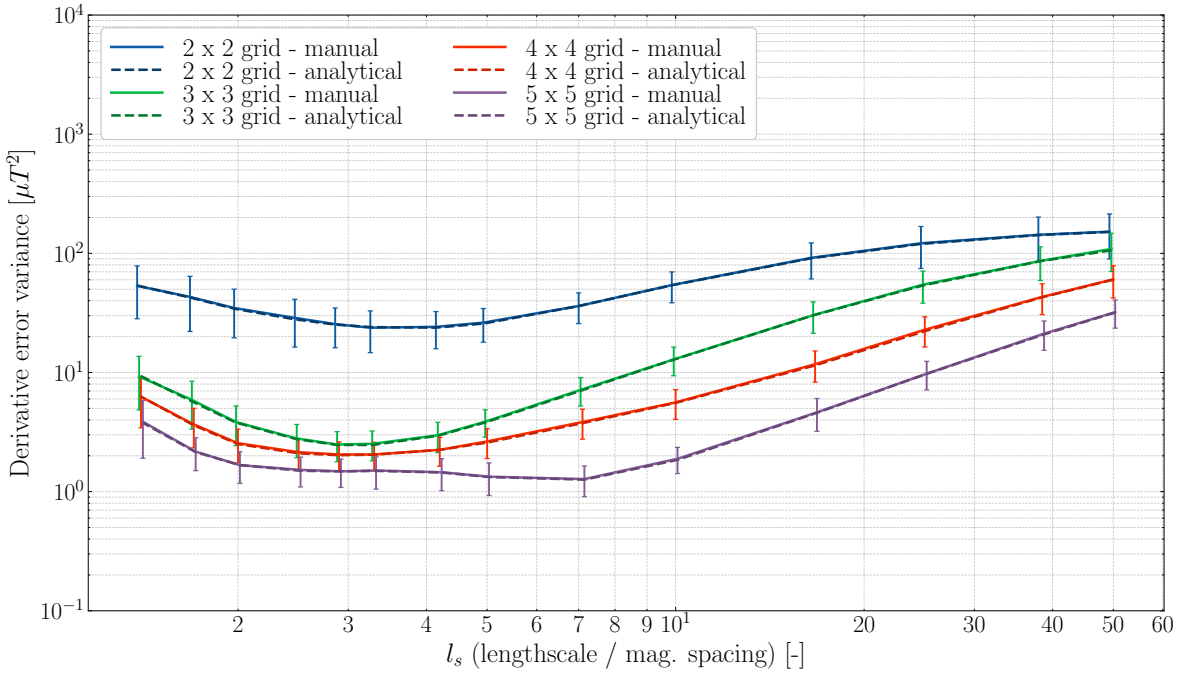

Figure D-1: Comparison between the analytically and manually computed mean variances of the spatial derivative using using the local GP method. The darker and dashed lines denote the analytical variance and the solid and lighter lines the manual variance. The variance of the measurement noise is set to $\sigma_y^2 = 0.5^2\,\mu T$ (noise-to-signal ratio (NSR)=1%).

Note, however, that the analytical and manual variances are equal only because of the fact that the experimental data is generated using the same model as was used for making the spatial derivative predictions. If the hyperparameters are not exactly the same, discrepancies will arise.

Bibliography

[1] Taiga Arai, Takahiro Yoshizawa, Takuya Aoki, Keiichi Zempo, and Yukihiko Okada. Evaluation of indoor positioning system based on attachable infrared beacons in metal shelf environment. In *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–4, 2019.

[2] David Caruso, Alexandre Eudes, Martial Sanfourche, David Vissière, and Guy le Besnerais. Robust indoor/outdoor navigation through magneto-visual-inertial optimization-based estimation. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4402–4409, 2017.

[3] David Caruso, Martial Sanfourche, Guy Le Besnerais, and David Vissière. Infrastructureless indoor navigation with an hybrid magneto-inertial and depth sensor system. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8, 2016.

[4] Charles-Ivan Chesneau. *Magneto-Inertial Dead-Reckoning in inhomogeneous field and indoor applications*. PhD thesis, Université Grenoble Alpes, 2019.

[5] Charles-Ivan Chesneau, Mathieu Hillion, Jean-François Hullo, Guillaume Thibault, and Christophe Prieur. Improving magneto-inertial attitude and position estimation by means of a magnetic heading observer. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8, 2017.

[6] Charles-Ivan Chesneau, Mathieu Hillion, and Christophe Prieur. Motion estimation of a rigid body with an ekf using magneto-inertial measurements. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–6, 2016.

[7] Eric Dorveaux, Thomas Boudot, Mathieu Hillion, and Nicolas Petit. Combining inertial measurements and distributed magnetometry for motion estimation. In *Proceedings of the American Control Conference (ACC)*, pages 4249–4256, 2011.

[8] Eric Dorveaux, Thomas Boudot, Mathieu Hillion, and Nicolas Petit. Combining inertial measurements and distributed magnetometry for motion estimation. In *Proceedings of the American Control Conference*, pages 4249–4256, 2011.

[9] Eric Dorveaux and Nicolas Petit. Presentation of a magneto-inertial positioning system: navigating through magnetic disturbances. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–2, 2011.

[10] Eric Dorveaux and Nicolas Petit. *Magneto-inertial navigation: principles and application to an indoor pedometer*. PhD thesis, MINES ParisTech, 2012.

[11] Eric Dorveaux, David Vissière, Alain-Pierre Martin, and Nicolas Petit. Iterative calibration method for inertial and magnetic sensors. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with the 28th Chinese Control Conference (CCC)*, pages 8296–8303, 2009.

[12] Eric Dorveaux, David Vissière, and Nicolas Petit. On-the-field calibration of an array of sensors. In *Proceedings of the American Control Conference (ACC)*, pages 6795–6802, 2010.

[13] Cedric Fischer, Thomas Quirin, Christophe Chautems, Quentin Boehler, Joris Pascal, and Bradley J. Nelson. Gradiometer-based magnetic localization for medical tools. *IEEE Transactions on Magnetics*, 59(2):1–5, 2023.

[14] Anum Hameed and Hafiza Anisa Ahmed. Survey on indoor positioning applications based on different technologies. In *Proceedings of the 12th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, pages 1–5, 2018.

[15] Janne Haverinen and Anssi Kemppainen. Global indoor self-localization based on the ambient magnetic field. *Robotics and Autonomous Systems*, 57(10):1028–1035, 2009.

[16] Chuan Huang, Gustaf Hendeby, Hassen Fourati, Christophe Prieur, and Isaac Skog. Mains: A magnetic-field-aided inertial navigation system for indoor positioning. *IEEE Sensors Journal*, 24(9):15156–15166, 2024.

[17] Alex Kendall, Matthew Grimes, and Roberto Cipolla. Posenet: A convolutional network for real-time 6-dof camera relocalization. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2938–2946, 2015.

[18] Manon Kok, Jeroen D. Hol, and Thomas B. Schön. 2017.

[19] Manon Kok and Arno Solin. Scalable magnetic field slam in 3d using gaussian process maps. In *Proceedings of the 21st International Conference on Information Fusion (FUSION)*, pages 1353–1360, 2018.

[20] Qiyue Li, Wei Li, Wei Sun, Jie Li, and Zhi Liu. Fingerprint and assistant nodes based wi-fi localization in complex indoor environment. *IEEE Access*, 4:2993–3004, 2016.

[21] Yongying Li, Hui Cheng, Zahrah Alhalili, Guiying Xu, and Guo Gao. The progress of magnetic sensor applied in biomedicine: A review of non-invasive techniques and sensors. *Journal of the Chinese Chemical Society*, 68(2):216–227, 2021.

[22] Thyagaraja Marathe, Ali Broumandan, Ali Pirsiavash, and Gérard Lachapelle. Characterization of range and time performance of indoor gnss signals. In *Proceedings of the European Navigation Conference (ENC)*, pages 27–37, 2018.

[23] James Clerk Maxwell. *A dynamical theory of the electromagnetic field.* The Society, 1865.

[24] Dennis Odijk and Frank Kleijer. Can gps be used for location based services at schiphol airport, the netherlands? In *Proceedings of the 5th Workshop on Positioning, Navigation and Communication (WPNC)*, pages 143–148, 2008.

[25] Dieter Oosterlinck, Dries F. Benoit, Philippe Baecke, and Nico Van de Weghe. Bluetooth tracking of humans in an indoor environment: An application to shopping mall visits. *Applied Geography*, 78:55–65, 2017.

[26] Mostafa Osman, Frida Viset, and Manon Kok. Indoor slam using a foot-mounted imu and the local magnetic field. In *Proceedings of the International Conference on Information Fusion (FUSION)*, pages 1–7, 2022.

[27] Carl Edward Rasmussen and Christopher K.I Williams. *Regression*, page 7–31. MIT press, 2006.

[28] Ali Raza, Lazar Lolic, Shahmir Akhter, and Michael Liut. Comparing and evaluating indoor positioning techniques. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8, 2021.

[29] Isaac Skog, Gustaf Hendeby, and Fredrik Gustafsson. Magnetic odometry - a model-based approach using a sensor array. In *Proceedings of the 21st International Conference on Information Fusion (FUSION)*, pages 794–798, 2018.

[30] Isaac Skog, Gustaf Hendeby, and Felix Trulsson. Magnetic-field based odometry – an optical flow inspired approach. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8, 2021.

[31] Arno Solin, Manon Kok, Niklas Wahlström, Thomas B. Schön, and Simo Särkkä. Modeling and interpolation of the ambient magnetic field by gaussian processes. *IEEE Transactions on Robotics*, 34(4):1112–1127, 2018.

[32] Arno Solin, Simo Särkkä, Juho Kannala, and Esa Rahtu. Terrain navigation in the magnetic landscape: Particle filtering for indoor positioning. In *Proceedings of the Conference on European Navigation Conference (ENC)*, pages 1–9, 2016.

[33] Yangyi Sui, Keith Leslie, and David Clark. Multiple-order magnetic gradient tensors for localization of a magnetic dipole. *IEEE Magnetics Letters*, 8:1–5, 2017.

[34] Simo Särkkä. Linear operators and stochastic partial differential equations in gaussian process regression. In Timo Honkela, Włodzisław Duch, Mark Girolami, and Samuel Kaski, editors, *Proceedings of the International Conference on Artificial Neural Networks and Machine Learning (ICANN)*, pages 151–158, 2011.

[35] Ilari Vallivaara, Janne Haverinen, Anssi Kemppainen, and Juha Röning. Simultaneous localization and mapping using ambient magnetic field. In *Proceedings of the International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 14–19, 2010.

[36] Arnaud Vena, Isabelle Illanes, Lucie Alidieres, Brice Sorli, and François Perea. Rfid based indoor localization system to analyze visitor behavior in a museum. In *Proceedings of the IEEE International Conference on RFID Technology and Applications (RFID-TA)*, pages 183–186, 2021.

[37] Frida Viset, Rudy Helmons, and Manon Kok. An extended kalman filter for magnetic field slam using gaussian process regression. *Sensors*, 22(8):2833, 2022.

[38] Frida Viset, Rudy Helmons, and Manon Kok. Distributed multi-agent magnetic field norm slam with gaussian processes. In *Proceedings of the International Conference on Information Fusion (FUSION)*, pages 1–8, 2023.

[39] David Vissière, Alain Martin, and Nicolas Petit. Using distributed magnetometers to increase IMU-based velocity estimation into perturbed area. In *Proceedings of the 46th IEEE Conference on Decision and Control (CDC)*, pages 4924–4931, 2007.

[40] Niklas Wahlström, Manon Kok, Thomas B. Schön, and Fredrik Gustafsson. Modeling magnetic fields using gaussian processes. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3522–3526, 2013.

[41] Ke Wang, Ampalavanapillai Nirmalathas, Christina Lim, Kamal Alameh, and Efstratios Skafidas. Optical wireless-based indoor localization system employing a single-channel imaging receiver. *Journal of Lightwave Technology*, 34(4):1141–1149, 2016.

[42] Makia Zmitri, Hassen Fourati, and Christophe Prieur. Magnetic field gradient-based ekf for velocity estimation in indoor navigation. *Sensors*, 20(20):5726, 2020.

[43] Teti Zubaidah, Bulkis Kanata, Paniran, Misbahuddin, Rosmaliati, Made Sutha Yadnya, and Susilawati Riskia. Earth magnetic fields evolution over nusa tenggara region from declination and inclination changes on lombok geomagnetic observatory. In *Proceedings of the 2nd International Conference on Applied Electromagnetic Technology (AEMT)*, pages 85–91, 2018.

Glossary

List of Acronyms

EKF	extended Kalman filter
ESKF	error state Kalman filter
GP	Gaussian process
GPR	Gaussian process regression
GNSS	global navigation satellite systems
IID	indepentend and identically distributed
INS	inertial navigation system
IMU	inertial measurement unit
LHS	left-hand side
LLS	linear least squares
MSLL	mean standardized log loss
RFID	radio frequency identification
SLAM	simultaneous localization and mapping
NSR	noise-to-signal ratio

List of Symbols

ϵ	Measurement noise
$\boldsymbol{\Omega}$	Angular velocity
$\boldsymbol{\epsilon}_d$	Measurement noise of the spatial derivative estimate
$\boldsymbol{\epsilon}_y$	Measurement noise of magnetometer
$\boldsymbol{\sigma}_d$	Standard deviation of spatial derivative estimate
Σ	White noise covariance matrix

$\sigma_{\text{lin.}}$	Magnitude scale hyperparameter
σ_f	Signal variance
σ_y	Standard deviation measurement noise
φ	Scalar potential function
$\mathbf{F}^*$	Magnetic field predictions
$\mathbf{G}$	Hessian
$\mathbf{H}(\mathbf{x})$	Magnetic field function
$\mathbf{H}_d(\mathbf{x})$	Spatial derivative function of the magnetic field
$\mathbf{r}$	Position
$\mathbf{v}$	Velocity
$\mathbf{x}$	Location
$\mathbf{X}^*$	Test locations
$\mathbf{x}_0$	Location of middle magnetometer
$\mathbf{X}_{\text{all}}$	Training locations full kernel
$\mathbf{X}_{\text{der}}$	Training locations derivative kernel
$\mathbf{y}$	Magnetic field measurement
$\mathbf{Y}_{\text{all}}$	Magnetic field measurements for full kernel
$\mathbf{Y}_{\text{der}}$	Magnetic field and spatial derivative measurements for derivative kernel
$\mathcal{D}_{\mathbf{x}}$	Derivative operator
h	Magnetometer spacing
J	Jacobian
$K_{\text{curl,d}}(\mathbf{x}, \mathbf{x}')$	Curl-free derivative kernel function
l	Lengthscale
l_0	Small representative lengthscale
l_1	Small representative lengthscale
l_s	Lengthscale over spacing
N	Number of magnetometers
P_{noise}	Power of noise
P_{signal}	Power of signal
x	First spatial coordinate
y	Second spatial coordinate
z	Third spatial coordinate
$\mathrm{K}_{\text{local,d}}$	Curl-free derivative kernel for local GP
$\mathrm{K}_{\text{curl}}(\mathbf{x}, \mathbf{x}')$	Curl-free kernel function
l_{ss}	Lengthscale over step size
t	Time