Solin, Arno; Kok, Manon; Wahlstrom, Niklas; Schon, Thomas B.; Sarkka, Simo

Modeling and Interpolation of the Ambient Magnetic Field by Gaussian Processes

Modeling and Interpolation of the Ambient Magnetic Field by Gaussian Processes

Arno Solin, Manon Kok, *Member, IEEE*, Niklas Wahlström, Thomas B. Schön, *Senior Member, IEEE*, and Simo Särkkä, *Senior Member, IEEE*

Abstract—Anomalies in the ambient magnetic field can be used as features in indoor positioning and navigation. By using Maxwell's equations, we derive and present a Bayesian non-parametric probabilistic modeling approach for interpolation and extrapolation of the magnetic field. We model the magnetic field components jointly by imposing a Gaussian process (GP) prior on the latent scalar potential of the magnetic field. By rewriting the GP model in terms of a Hilbert space representation, we circumvent the computational pitfalls associated with GP modeling and provide a computationally efficient and physically justified modeling tool for the ambient magnetic field. The model allows for sequential updating of the estimate and time-dependent changes in the magnetic field. The model is shown to work well in practice in different applications: we demonstrate mapping of the magnetic field both with an inexpensive Raspberry Pi powered robot and on foot using a standard smartphone.

Index Terms—Gaussian process, magnetic field, Maxwell's equations, mapping, online representation

I. INTRODUCTION

Magnetic material causes anomalies in the ambient magnetic field. In indoor environments, large amounts of such magnetic material are present in the structure of buildings and in furniture. Our focus is on building maps of the indoor magnetic field these structures are inducing. These maps are constructed by interpolating three-dimensional magnetic field measurements obtained using magnetometers. An illustration of a map obtained using our proposed method is available in Figure 1.

Magnetic maps of indoor environments can be used in indoor positioning and navigation applications (see, *e.g.*, [1]). In these applications, sensors providing position information that is accurate on a short time-scale—but drifts on longer time-horizons—are typically combined with absolute position measurements. For example, data from wheel encoders and inertial sensors—which give accurate position information only on a short time-scale—can be combined with ultra-wideband, Wi-Fi, or optical measurement equipment such as cameras (see, *e.g.*, [2, 3]). The downside of these sources of absolute position is that they typically rely on additional

Arno Solin is with the Department of Computer Science, Aalto University, 02150 Espoo, Finland, and with IndoorAtlas Ltd., Helsinki, Finland, e-mail: arno.solin@aalto.fi.

Manon Kok is with the Department of Engineering, University of Cambridge, CB2 1PZ Cambridge, United Kingdom, e-mail: mk930@cam.ac.uk.

Niklas Wahlström and Thomas B. Schön are with the Department of Information Technology, Uppsala University, SE-751 05 Uppsala, Sweden, e-mail: {niklas.wahlstrom, thomas.schon}@it.uu.se.

Simo Särkkä is with the Department of Electrical Engineering and Automation, Aalto University, 02150 Espoo, Finland, e-mail: simo.sarkka@aalto.fi.

(a) Interpolated magnetic field strength

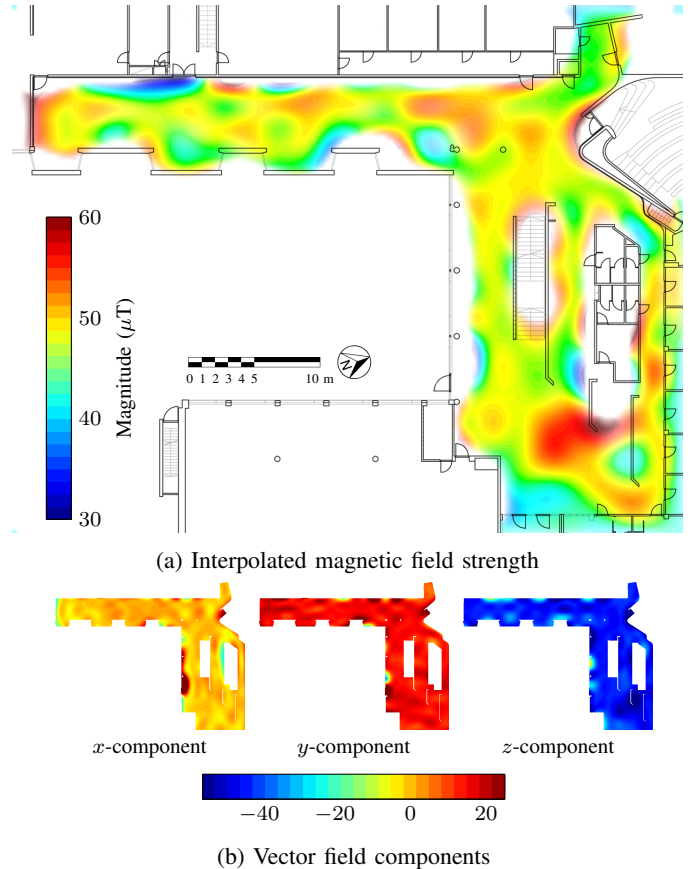x-component y-component z-component

(b) Vector field components

Fig. 1: Interpolated magnetic field of the lobby of a building at the Aalto University campus. The marginal variance (uncertainty) is visualized by the degree of transparency.

infrastructure or require certain conditions to be fulfilled such as line-of-sight. The advantage of using the magnetic field for positioning is that it can be measured by a small device, without additional infrastructure and without line-of-sight requirements. Furthermore, magnetometers are nowadays present in (almost) any inertial measurement unit (IMU) or smartphone. A requirement for localization using the ambient magnetic field as a source of position information is that accurate maps of the magnetic field can be constructed within reasonable computational complexity which is the focus of this work. This can also be regarded as a step towards simultaneous localization and mapping (SLAM) using magnetic fields in which localization is done while building the map (see, *e.g.*, [4, 5]).

We interpolate the magnetic field using a Bayesian non-parametric approach where prior knowledge about the properties of magnetic fields is incorporated in a Gaussian process (GP) prior. GPs (see, *e.g.*, [6, 7]) are powerful tools for Bayesian non-parametric inference and learning, and they provide a framework for fusing first-principles prior knowledge with quantities of noisy data. This has made them popular tools in signal processing, machine learning, robotics and control [8–10].

The contributions of this paper are three-fold. First, we model the ambient magnetic field using a Gaussian process prior in which we incorporate *physical knowledge* about the magnetic field. This extends the work by Wahlström et al. [11] by presenting an approach where the GP prior is a *latent (unobservable) magnetic potential* function. Second, we use a computationally efficient GP implementation that allows us to use the large amounts of data provided by the magnetometer. To circumvent the well-known computational challenges with GPs (see, *e.g.*, [7]), we rewrite the model in terms of a Hilbert space representation introduced by Solin and Särkkä [12]. We extend the approach to allow for modeling of the bias caused by the Earth magnetic field. Third, we use this method in combination with the sequential approach introduced in Särkkä et al. [13]. This allows for *online updating* of the magnetic field estimate. It also opens up the possibility to focus on the spatio-temporal problem in which the magnetic field can change over time, for instance due to furniture being moved around. An extensive evaluation of the proposed method is done using both simulated and empirical data. The simulation study and a small-scale experiment illustrate the feasibility and accuracy of the approach and allow for comparison with other methods. Experiments with a mobile robot and with a hand-held smartphone show the applicability to real-world scenarios.

This paper is structured as follows. The next section covers a survey of existing work, which also provides additional motivation for the approach. Section III provides a brief background of the properties of magnetic fields relevant to this work. The Gaussian process regression model is constructed in Section IV, which is then extended to explicit algorithms for batch and sequential estimation in the next section. Section VI covers the experiments. The experimental results and some additional comments regarding the methodology are discussed at the end of the paper.

II. RELATED WORK

Spatial properties of the magnetic field have been of interest in a large variety of research domains. For instance, the magnetic field has been extensively studied in geology (see, *e.g.*, [14]) but also in magnetospheric physics, geophysics, and astrophysics. In all of these domains, interpolation of the magnetic field is of interest (see, *e.g.*, [15–19] for examples of magnetic field interpolation in the respective areas).

In recent years interest has emerged in using the magnetic field as a source of position information for indoor positioning [20]. Feasibility studies have been conducted, focusing both on the time-varying nature of the magnetic field and on the amount of spatial variation in the magnetic field. Li et al. [21] report experiments showing that the magnetic field in a building typically shows large spatial variations and small time variations. This is also supported by the experimental study reported by Angermann et al. [22] in which significant anomalies of the ambient magnetic field are reported. These experiments give confidence that the magnetic field provides sufficient information for localization purposes. However, Li et al. [21] also report significant temporal changes in the magnetic field in the vicinity of mobile magnetic structures, in their case an elevator.

A number of approaches have been reported on building a map of the ambient magnetic field for indoor localization purposes. Le Grand and Thrun [23] propose a method to build a map of the magnetic field by collecting magnetometer data in a grid and linearly interpolating between these points. This map is subsequently used for localization with a particle filter combining magnetometer and accelerometer measurements from a smartphone. Robertson et al. [24] present a SLAM approach for pedestrian localization using a foot-mounted IMU. They use the magnetic field intensity which they model using spatial binning. Frassl et al. [25] discuss the possibility of using more components of the magnetic field (for instance the full three-dimensional measurement vector) in the SLAM approach instead. Vallivaara et al. [26, 27] present a SLAM approach for robot localization. They model the ambient magnetic field using a squared exponential GP prior for each of the magnetic field components. Wahlström et al. [11] incorporate additional physical knowledge by making use of Maxwell's equations resulting in the use of curl- and divergence-free GP priors instead.

As can be concluded, there exists a wide range of existing literature when it comes to modeling the ambient magnetic field. The amount of information that is used differs between the approaches. For instance, some approaches use full three-dimensional magnetic field vectors while others only use a one-dimensional magnetic field intensity. Furthermore, the amount of physical information that is included differs. In this paper, we build on the approach by Wahlström et al. [11] and use the full three-dimensional magnetometer measurements. We include physical knowledge in terms of the magnetic field potential.

As discussed above, GPs have frequently been used in modeling and interpolation of magnetic fields. GP regression has also successfully been applied to a wide range of applications (see, *e.g.*, [28–33]). Furthermore, it has previously been used for SLAM (see, *e.g.*, [34–38]). One of the challenges in using GPs is the computational complexity (see, *e.g.*, [7]), which scales cubically with the number of training data points. Considering the high sampling rate of the magnetometer and the fact that each observation contains three values, a large number of measurements is typically available for mapping. Because of these computational challenges, the data in Wahlström et al. [11] was downsampled.

Attempts to speed up GP inference have spawned a wide range of methods which aimed at bringing GP regression to data-intensive application fields. These methods (see [39] for a review) typically build upon reducing the rank of the Gram
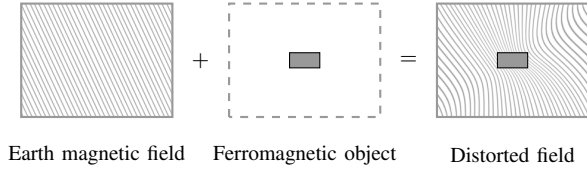
Fig. 2: A ferromagnetic object deflects the Earth's magnetic field and introduces distortions in the field.

(covariance) matrix and using the matrix inversion lemma to speed up matrix inversion. For stationary covariance functions, the spectral Monte Carlo approximation by Lázaro-Gredilla et al. [40], the Fourier features by Hensman et al. [41], or the Laplace operator eigenbasis based method introduced by Solin and Särkkä [12] can be employed. For uniformly spaced observations, fast Fourier transforms can provide computational benefits [42, 43]. As will be shown later on in this paper, the Laplace operator approach by Solin and Särkkä [12] falls natural to modeling of the magnetic field in terms of a magnetic field potential.

All approaches on mapping of magnetic fields discussed above assume that the magnetic field is constant over time. However, as shown by Li et al. [21] significant temporal changes in the magnetic field occur in the vicinity of mobile magnetic structures. For GP models evolving in time, spatio-temporal GP models (see, *e.g.*, [9]) can be solved efficiently using Kalman filtering methods [13, 44–47]. In this paper, we will take the approach of Särkkä et al. [13] to compose a spatio-temporal GP prior for the model and solve the inference problem by a sequential Kalman filtering setup. This allows for online estimation of the magnetic field estimate and can be used to allow for time variations in the magnetic field.

Combining models from physics with GPs has also been studied under the name *Latent force models* by Álvarez et al. [48, 49]. The connection of these models with spatio-temporal Kalman filtering was studied, for example, in the work by Hartikainen and Särkkä (see also [13, 50, 51]). However, the Kalman filtering approach itself dates back to Curtain and Pritchard [52] in the same way that GPs date back to O'Hagan [6].

III. THE AMBIENT MAGNETIC FIELD

On a macroscopic scale, magnetic fields are *vector fields*, meaning that at any given location, they have a direction and a strength (magnitude). These properties are familiar from everyday life: the force created by permanent magnets attracting and repelling ferromagnetic materials is used in various utensils, and the compass aligning itself with the direction of the Earth's magnetic field has proved invaluable for mankind during the past centuries. The Earth's magnetic field sets a background for the ambient magnetic field, but deviations caused by the bedrock and anomalies induced by man-built structures deflect the Earth's magnetic field. This makes the magnetic field vary from point to point, see Figure 2.

We describe the magnetic field with a function $\mathbf{H}(\mathbf{x})$, where $\mathbf{H} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. For each point in space $\mathbf{x}$, there will be an associated magnetic field $\mathbf{H}(\mathbf{x})$. Such a vector field can be
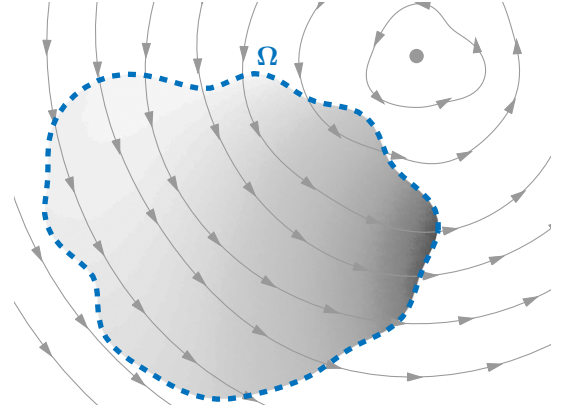

Fig. 3: Illustration of a vector field with non-zero curl. The vortex point makes it non-curl-free as the vector field curls around it. However, the subset Ω excludes the vortex point and the vector field is curl-free in this region. To this region a scalar potential φ can be associated, here illustrated with shading.

visualized with field lines, where points in space are associated with arrows. The principles under which the magnetic field is affected by structures of buildings are well known and governed by the very basic laws of physics (see, *e.g.*, [53–55]).

In this work, we make use of the fact that the magnetic field $\mathbf{H}$ is curl-free

$$\nabla \times \mathbf{H} = \mathbf{0} \tag{1}$$

provided that there is no free current (current in wires for example) in the region of interest (see [56] for more details). This assumption is valid in most indoor environments where the major source for variations in the ambient field is caused by metallic structures rather than free currents in wires.

One property of curl-free vector fields is that the line integral along a path P only depends on its starting point A and end point B, and not on the route taken

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

where $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$. This can be rewritten by interpreting φ as a scalar potential

$$\mathbf{H} = -\nabla\varphi. \tag{3}$$

Figure 3 illustrates the curl-free property and the scalar potential. Domain Ω is curl-free and has an associated scalar potential, while the entire domain is not curl-free due to the vortex point. Note that in a non-curl-free vector field no such scalar potential exists since a line integral around the swirl is non-zero. For the magnetic field $\mathbf{H}$, the swirl corresponds to a wire of free current pointing perpendicular to the plane, which we assume is not included in the region of interest.

The relation (3) is the key equation that we will exploit in our probabilistic model of the ambient magnetic field. We will choose to model the scalar potential φ instead of the magnetic field $\mathbf{H}$ directly. This implicitly imposes the constraints on the magnetic field that the physics is providing. This model will be explained in the next section.

IV. MODELING THE MAGNETIC FIELD USING GAUSSIAN PROCESS PRIORS

In this section, we introduce our approach to modeling and interpolation of the ambient magnetic field. We use a Bayesian non-parametric model in which we use knowledge about the physical properties of the magnetic field as prior information. We tackle the problem of interpolating the magnetic field using Gaussian process (GP) regression. In Section IV-A we first give a brief background on GPs. After this, we introduce the problem of modeling the magnetic field in Section IV-B. A commonly used GP model of the magnetic field is introduced in Section IV-C. In Section IV-D, we subsequently introduce our proposed method for modeling the magnetic field in which we encode the physical properties that were presented in the previous section.

A. Gaussian process regression

In GP regression [7] the model functions $f(\mathbf{x})$ are assumed to be realizations from a Gaussian random process prior with a given covariance function $\kappa(\mathbf{x}, \mathbf{x}')$. Learning amounts to computing the posterior process at some test inputs $\mathbf{x}_*$ given a set of noisy measurements $y_1, y_2, \ldots, y_n$ observed at $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, respectively. This model is often written in the form

$$f(\mathbf{x}) \sim \mathcal{GP}(0, \kappa(\mathbf{x}, \mathbf{x}')),$$
$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \tag{4}$$

where the observations y_i are corrupted by Gaussian noise $\varepsilon_i \sim \mathrm{N}(0, \sigma_{\mathrm{noise}}^2)$, for $i = 1, 2, \ldots, n$. Because both the prior and the measurement model are Gaussian, the posterior process will also be Gaussian. Hence, the learning problem amounts to computing the conditional means and covariances of the process evaluated at the test inputs.

Prediction of yet unseen process outputs at an input location $\mathbf{x}_*$ amounts to the following in GP regression: $p(f(\mathbf{x}_*) \mid \mathcal{D}) = \mathrm{N}(f(\mathbf{x}_*) \mid \mathbb{E}[f(\mathbf{x}_*)], \mathbb{V}[f(\mathbf{x}_*)])$. The conditional mean and variance can be computed in closed-form as (see [7])

$$\mathbb{E}[f(\mathbf{x}_*)] = \mathbf{k}_*^{\mathsf{T}} (\mathbf{K} + \sigma_{\mathrm{noise}}^2 \mathbf{I}_n)^{-1} \mathbf{y},$$
$$\mathbb{V}[f(\mathbf{x}_*)] = \kappa(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^{\mathsf{T}} (\mathbf{K} + \sigma_{\mathrm{noise}}^2 \mathbf{I}_n)^{-1} \mathbf{k}_*, \tag{5}$$

where $\mathbf{K}_{i,j} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{k}_*$ is an n-dimensional vector with the ith entry being $\kappa(\mathbf{x}_*, \mathbf{x}_i)$, and $\mathbf{y}$ is a vector of the n observations. Furthermore, due to Gaussianity, the marginal likelihood (evidence) of the covariance function and noise parameters can also easily be computed, allowing for Bayesian inference of the parameters as well [7].

The choice of a specific covariance function encodes the *a priori* knowledge about the underlying process. One of the most commonly used covariance functions, which will also frequently be used in the next sections, is the stationary and isotropic squared exponential (also known as exponentiated quadratic, radial basis function, or Gaussian). Following the standard notation from Rasmussen and Williams [7] it is parametrized as

$$\kappa_{\mathrm{SE}}(\mathbf{x}, \mathbf{x}') = \sigma_{\mathrm{SE}}^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\,\ell_{\mathrm{SE}}^2}\right), \tag{6}$$

where the hyperparameters σ_{SE}^2 and ℓ_{SE} represent the magnitude scale and the characteristic length-scale, respectively. These can be learned from data, for instance by maximizing the marginal likelihood.

B. Interpolation of magnetic fields

In this work we tackle the problem of interpolating the magnetic field to spatial locations from where we do not have any measurements. In other words, we will tackle the problem of predicting the latent (unobservable noise-free) magnetic field $\mathbf{f}(\mathbf{x}_*)$ (such that $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$) at some arbitrary location $\mathbf{x}_*$ given a set of measurements $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$ of the magnetic field. Here, the measurements $\mathbf{y}$ correspond to the H-field corrupted by i.i.d. Gaussian noise.

Two important things need to be noted with regard to the interpolation of the magnetic field. First, note that the measurements of the magnetic field are *vector-valued* (contrary to the scalar observations in (4)). This raises the question of how to deal with the different magnetic field components. They can either be treated separately as will be done in Section IV-C, or a relation between the different components can be assumed, as is the case in the method we propose in Section IV-D. Secondly, note that the function describing the magnetic field is not zero-mean, contrary to the GP model in (4). Instead, its mean lies around a local Earth magnetic field. This depends on the location on the Earth but can also deviate from the Earth's magnetic field in indoor environments due to magnetic material in the structure of the building. The unknown mean can be modeled as an additional part of the covariance function $\kappa(\mathbf{x}, \mathbf{x}')$ [7].

An illustration of GP regression for magnetic fields is provided in Figure 4, where noisy readings of a magnetic field have been collected along route A–D (comprising $\mathcal{D}$), and the magnetic field along route D–E (comprising the prediction locations $\mathbf{x}_*$) needs to be inferred from the measurements. Each component varies around a local magnetic field due to magnetic material in the vicinity of the sensor. Different interpolation techniques can be used based on different prior knowledge that can be incorporated in the GP. Two different interpolation results are shown: one based on independent modeling of each vector field components (with shared hyperparameters) and another based on associating the GP prior with the scalar potential of the (curl-free) vector field. These are based on the models we will introduce in the coming two sections.

C. Separate modeling of the magnetic field components

The most straightforward approach to GP modeling of vector-valued quantities is to model each of the field components as an independent GP. This approach has been widely applied in existing literature (see, *e.g.*, [26, 27, 57–59]). For each of the three magnetic field components $d \in \{1, 2, 3\}$, this model can be written as

$$f_d(\mathbf{x}) \sim \mathcal{GP}(0, \kappa_{\mathrm{const.}}(\mathbf{x}, \mathbf{x}') + \kappa_{\mathrm{SE}}(\mathbf{x}, \mathbf{x}')),$$
$$y_{d,i} = f_d(\mathbf{x}_i) + \varepsilon_{i,d}, \tag{7}$$

where the observations $y_{d,i}, i = 1, 2, \ldots, n$, are corrupted by independent Gaussian noise with variance $\sigma_{\mathrm{noise}}^2$. The non-zero

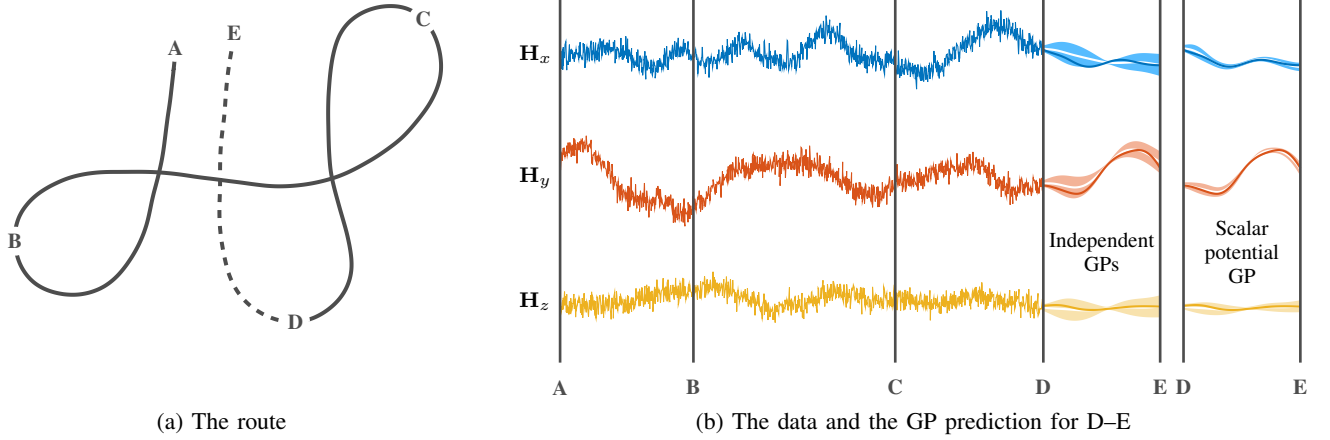(a) The route   (b) The data and the GP prediction for D–E

Fig. 4: A simulated example of the interpolation problem. (a) Training data has been collected along the route A–D, but the magnetic field between D–E is unknown. (b) The noisy observations of the magnetic field between A–D, and GP predictions with 95% credibility intervals. Both the independent GP modeling approach (with shared hyperparameters) and the scalar potential based curl-free GP approach are visualized. The simulated ground truth is shown by the solid lines.

mean of the magnetic field is handled by a constant covariance function (see [7])

$$\kappa_{\text{const.}}(\mathbf{x}, \mathbf{x}') = \sigma_{\text{const.}}^2, \tag{8}$$

where $\sigma_{\text{const.}}^2$ is a magnitude scale hyperparameter. The small-scale variation in the field is modeled by a squared exponential covariance function (6). Hence, the model (7) encodes the knowledge that the realizations are expected to be smooth functions in space with a constant shift from zero mean.

The model has four hyperparameters: two magnitude scale parameters ($\sigma_{\text{const.}}^2$ and σ_{SE}^2), a length-scale parameter (ℓ_{SE}), and a noise scale parameter (σ_{noise}^2). Assuming that the components are completely separate, each component has four hyperparameters to learn. The resulting model is flexible, as it does not encode any relation between the vector field components. In practice, this might lead to problems in hyperparameter estimation, with parameter estimates converging to local optima and magnetic field components behaving very differently with respect to each other. Therefore, the hyperparameters are often fixed to reasonable values—instead of learned from data (see, *e.g.*, [57]).

A more sensible approach for separate, but not completely independent, modeling of the magnetic field measurements, models them as realizations of three independent GP priors with joint learning of the shared hyperparameters (see also [57]). Note that for this model, the covariance in the GP posterior is independent of the outputs $\mathbf{y}$ and only depends on the input locations $\mathbf{x}$ (which are shared for all components in $\mathbf{y}$). Hence, calculating the marginal likelihood only requires inverting a matrix of size n (not $3n$). For this model, the expression for evaluating the log marginal likelihood function for hyperparameter optimization can be written as

$$\mathcal{L}(\boldsymbol{\theta}) = -\log p(\mathbf{y} \mid \boldsymbol{\theta}, \mathcal{D}) = \frac{3}{2} \log |\mathbf{K}_{\boldsymbol{\theta}} + \sigma_{\text{noise}}^2 \mathbf{I}_n|$$
$$+ \frac{1}{2} \text{tr} \left[\mathbf{y} (\mathbf{K}_{\boldsymbol{\theta}} + \sigma_{\text{noise}}^2 \mathbf{I}_n)^{-1} \mathbf{y}^{\mathsf{T}} \right] + \frac{3n}{2} \log(2\pi), \tag{9}$$

where $\mathbf{y} \in \mathbb{R}^{3 \times n}$ and $\mathbf{K}_{\boldsymbol{\theta}} \in \mathbb{R}^{n \times n}$.

Figure 4b shows the results of predicting the magnetic field behavior along the route D–E for the GP prior modeling the three magnetic field components separately but with joint learning of the shared hyperparameters. The colored patches show the 95% credibility intervals for the prediction with the mean estimate visualized by the white line. The simulated ground-truth (solid colored line) falls within the shown interval, and the model captures the general shape of the magnetic field variation along the path. The strengths of this model are that it is flexible and that the assumptions are conservative. The weaknesses on the other hand are evident: The model does not incorporate physical knowledge of the magnetic field characteristics. In the next section we will instead explore this knowledge by modeling the magnetic field as derivative measurements of a scalar potential.

D. Modeling the magnetic field as the gradient of a scalar potential

Following our choices in Section III, we assume that the magnetic field $\mathbf{H}$ can be written as the gradient of a scalar potential $\varphi(\mathbf{x})$ according to (3). Here, $\varphi : \mathbb{R}^3 \to \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^3$ is the spatial coordinate. We assume $\varphi(\mathbf{x})$ to be a realization of a GP prior and the magnetic field measurements $\mathbf{y}_i \in \mathbb{R}^3$ to be its gradients corrupted by Gaussian noise. This leads to the following model

$$\varphi(\mathbf{x}) \sim \mathcal{GP}(0, \kappa_{\text{lin.}}(\mathbf{x}, \mathbf{x}') + \kappa_{\text{SE}}(\mathbf{x}, \mathbf{x}')),$$
$$\mathbf{y}_i = -\nabla \varphi(\mathbf{x})\big|_{\mathbf{x}=\mathbf{x}_i} + \boldsymbol{\varepsilon}_i, \tag{10}$$

where $\boldsymbol{\varepsilon}_i \sim \text{N}(\mathbf{0}, \sigma_{\text{noise}}^2 \mathbf{I}_3)$, for each observation $i = 1, 2, \ldots, n$. The squared exponential covariance function (6) in (10) allows us to model the magnetic field anomalies induced by small-scale variations and building structures. The local Earth's magnetic field contributes linearly to the scalar potential as

$$\kappa_{\text{lin.}}(\mathbf{x}, \mathbf{x}') = \sigma_{\text{lin.}}^2 \mathbf{x}^{\mathsf{T}} \mathbf{x}', \tag{11}$$

where $\sigma_{\text{lin.}}^2$ is the magnitude scale hyperparameter. To simplify the notation in the next sections, we introduce the notation $\mathbf{f}(\mathbf{x})$ for the gradient field evaluated at $\mathbf{x}$.

Derivative measurements can straightforwardly be used in Gaussian process regression because nabla (∇) is a linear operator and Gaussianity is preserved under linear operations [7]. Hence, the model (10) allows us to learn a map of the magnetic field and make predictions at unseen locations using the methods described in Section IV-A. We learn the four hyperparameters of the model—the magnitude scale parameters ($\sigma_{\text{lin.}}^2$ and σ_{SE}^2), the length-scale parameter (ℓ_{SE}), and the noise scale parameter (σ_{noise}^2)—by maximizing the marginal likelihood.

The model (10)—where we place a GP prior on the scalar potential $\varphi(\mathbf{x})$ and the magnetometer measurements are derivative measurements—can in fact equivalently be written as a GP prior on the magnetic field $\mathbf{H}(\mathbf{x})$ as

$$\mathbf{H}(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, \sigma_{\text{const.}}^2 \, \mathbf{I}_3 + \mathbf{K}_{\text{curl}}(\mathbf{x}, \mathbf{x}')), \qquad (12)$$

and magnometer measurements being of the standard form (4). Here, $\mathbf{K}_{\text{curl}}(\mathbf{x}, \mathbf{x}')$ is the so-called curl-free kernel (see, *e.g.*, [60–62]). This kernel ensures that any sample drawn from this GP prior obeys the curl-free constraints (1). The equivalence of (10) and (12) is shown in Wahlström [56]. In Jidling et al. [63] a more general approach to force GP priors to obey linear constraints is presented. For our work, however, the model formulation through the scalar potential is crucial, as it will enable us to easily extend the model to an approximate form for efficient GP inference in the next section.

The second set of predictions for the route D–E in Figure 4b shows the interpolation outcome from the model (10). In comparison to the independent GP model, the scalar potential based GP prior provides additional information to the model by tying the vector field components to each other. This improves the estimates in terms of accuracy and makes the 95% credibility interval more narrow.

V. EFFICIENT GP MODELING OF THE MAGNETIC FIELD

Gaussian processes are convenient tools for assigning flexible priors to data—as we saw in the previous section. However, the main problem with the model in the previous section is its high computational cost. The approach scales as $\mathcal{O}(n^3)$ (recall that each observation is three-dimensional, meaning $3n$ becomes large very quickly). This computational complexity renders the approach more or less useless in practice, when the number of observations becomes large (say $n > 1000$).

This is a fundamental restriction associated with the naive formulation of GP models involving the inversion of the covariance matrix. Using special structure of the problem and/or approximative methods, this high computational cost can often be circumvented. In this section we present an approach which both uses the special differential operator structure and projects the model on a set of basis functions characteristic to the covariance function. We first present the method for spatial batch estimation, and then extend it to a temporal dimension as well.

Existing GP methods for mapping and interpolation of the magnetic field have been considering only batch estimation, where the data is first acquired and then processed as a batch. In this section, we aim to extend this to an online method, enabling the GP regression estimate of the magnetic field to be updated when new data is acquired. We denote such a data set as $\mathcal{D}_n = \{(\mathbf{x}_i, \mathbf{y}_i) \mid i = 1, 2, \ldots, n\}$, and thus $\mathcal{D}_i$ denotes all the data that has been observed up to time instance t_i.

Considering time as part of the data stream enables us to think of three distinctive setups for estimation of the magnetic field:

- **Batch estimation** of the magnetic field, where the data is first acquired and then the field is estimated at once.
- **Sequential updating** of the field estimate, where we assume all the measurements to be of the same static magnetic field.
- **Spatio-temporal estimation** of the time-dependent magnetic field, where we assume the field to change over time.

In the next sections we will present how these scenarios can be combined with the scalar potential based GP scheme without requiring to repeat the batch computations after each sample.

A. Reduced-rank GP modeling

A recent paper by Solin and Särkkä [12] presents an approach that is based on a series expansion of stationary covariance functions. The approximation is based on the following truncated series:

$$\kappa(\mathbf{x}, \mathbf{x}') \approx \sum_{j=1}^{m} S(\lambda_j) \, \phi_j(\mathbf{x}) \, \phi_j(\mathbf{x}'), \qquad (13)$$

where $S(\cdot)$ is the spectral density of the covariance function $\kappa(\cdot, \cdot)$, $\phi_j(\mathbf{x})$ is the jth eigenfunction of the negative Laplace operator and λ_j^2 is the corresponding eigenvalue. The efficiency of this approach is based on two properties: (i) the eigenfunctions are independent of the hyperparameters of the covariance function, and (ii) for many domains the eigenfunctions and eigenvalues can be solved beforehand in closed-form. Truncating this expansion at degree $m \ll n$ allows the GP regression problem to be solved with a $\mathcal{O}(nm^2)$ and the hyperparameters to be learned with a $\mathcal{O}(m^3)$ time complexity. The memory requirements scale as $\mathcal{O}(nm)$.

Our interest lies in modeling the magnetic field in compact subsets of $\mathbb{R}^3$, allowing us to restrict our interest to domains Ω comprising three-dimensional cuboids (rectangular boxes) such that $\mathbf{x} \in [-L_1, L_1] \times [-L_2, L_2] \times [-L_3, L_3] \subset \mathbb{R}^3$ (recall that a stationary covariance function is translation invariant). In this domain, we can solve the eigendecomposition of the Laplace operator subject to Dirichlet boundary conditions

$$\begin{cases} -\nabla^2 \phi_j(\mathbf{x}) = \lambda_j^2 \phi_j(\mathbf{x}), & \mathbf{x} \in \Omega, \\ \phi_j(\mathbf{x}) = 0, & \mathbf{x} \in \partial\Omega. \end{cases} \qquad (14)$$

The choice of the domain and boundary conditions is arbitrary, but for regression problems with a stationary covariance function the model reverts back to the prior outside the region of observed data, so the Dirichlet boundary condition

does not restrict the modeling if Ω is chosen suitably. This particular choice of domain and boundary conditions yields the following analytic expression for the basis functions and the corresponding eigenvalues:

$$\phi_j(\mathbf{x}) = \prod_{d=1}^{3} \frac{1}{\sqrt{L_d}} \sin\left(\frac{\pi n_{j,d}(x_d + L_d)}{2L_d}\right), \qquad (15)$$

$$\lambda_j^2 = \sum_{d=1}^{3} \left(\frac{\pi n_{j,d}}{2L_d}\right)^2, \qquad (16)$$

where the matrix $\mathbf{n} \in \mathbb{R}^{m \times 3}$ consists of an index set of permutations of integers $\{1, 2, \ldots, m\}$ (*i.e.*, $\{(1,1,1), (1,1,2), \ldots, (1,2,1), \ldots, (2,1,1), \ldots\}$). The basis functions are independent of the hyperparameters, and thus only need to be evaluated once.

The Laplace operator eigenbasis approximation method can be combined with the independent GP approach, the independent GPs with shared hyperparameters, and the scalar potential GP approach. Our presentation will be specific to the scalar potential model presented in Section IV-D, but a similar setup can be constructed for the other methods as well. The approximation method is even better suited for the scalar potential model because the approximation is based on the eigendecomposition of the Laplace operator. This eigenbasis falls naturally to the problem formulation in which the latent potential field is observed through gradients.

The covariance in our model (10) consists of a squared exponential and a linear covariance function. The former is stationary and the approach from [12] can straightforwardly be applied. The latter is not stationary, but there is also no need to approximate it. Consequently, we consider the approximation

$$\kappa(\mathbf{x}, \mathbf{x}') = \kappa_{\text{lin.}}(\mathbf{x}, \mathbf{x}') + \kappa_{\text{SE}}(\mathbf{x}, \mathbf{x}')$$
$$\approx \sigma_{\text{lin.}}^2 \mathbf{x}^\mathsf{T} \mathbf{x} + \sum_{j=1}^{m} S_{\text{SE}}(\lambda_j)\, \phi_j(\mathbf{x})\, \phi_j(\mathbf{x}'). \qquad (17)$$

Note that adding the linear covariance function, the approximation will no longer revert to zero on the domain boundary, but will instead revert to the scalar potential of the local Earth magnetic field.

The computational benefits come from the approximate eigendecomposition of the Gram (covariance) matrix, $\mathbf{K}_{i,j} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$ (see [12] for derivations and discussion). It can now be written out in terms of the basis functions and spectral densities: $\mathbf{K} \approx \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^\mathsf{T}$. The basis functions, which span the solution, are collected in the matrix $\mathbf{\Phi} \in \mathbb{R}^{n \times (3+m)}$, with the following rows

$$\mathbf{\Phi}_i = \left(\mathbf{x}_i^\mathsf{T}, \phi_1(\mathbf{x}_i), \phi_2(\mathbf{x}_i), \ldots, \phi_m(\mathbf{x}_i)\right), \qquad (18)$$

for $i = 1, 2, \ldots, n$. Accordingly, we define the corresponding measurement model matrix projecting the derivative observations onto the basis functions. Analogously, we define the matrix $\nabla\mathbf{\Phi} \in \mathbb{R}^{3n \times (3+m)}$ as the following block-row matrix:

$$\nabla\mathbf{\Phi}_i = \left(\nabla\mathbf{x}_i^\mathsf{T}, \nabla\phi_1(\mathbf{x}_i), \nabla\phi_2(\mathbf{x}_i), \ldots, \nabla\phi_m(\mathbf{x}_i)\right), \quad (19)$$

for $i = 1, 2, \ldots, n$. Similarly we define $\mathbf{\Phi}_*$ and $\nabla\mathbf{\Phi}_*$ as vectors evaluated at the prediction input location $\mathbf{x}_*$ defined

analogously to Equations (18) and (19), respectively. The matrix $\mathbf{\Lambda}$ is defined by

$$\mathbf{\Lambda} = \text{diag}(\sigma_{\text{lin.}}^2, \sigma_{\text{lin.}}^2, \sigma_{\text{lin.}}^2, S_{\text{SE}}(\lambda_1), S_{\text{SE}}(\lambda_2), \ldots, S_{\text{SE}}(\lambda_m)). \tag{20}$$

For three-dimensional inputs, the spectral density function of the squared exponential covariance function (6) is given by

$$S_{\text{SE}}(\omega) = \sigma_{\text{SE}}^2 \, (2\pi\ell_{\text{SE}}^2)^{3/2} \exp\left(-\frac{\omega^2 \ell_{\text{SE}}^2}{2}\right), \qquad (21)$$

where the hyperparameters σ_{SE}^2 and ℓ_{SE} characterize the spectrum.

B. Batch estimation

We first tackle the batch estimation problem which provides the approximative solution to the GP regression problem in Equation (5) for the scalar potential GP.

Following the derivations of Solin and Särkkä [12], predictions for interpolation and extrapolation of the magnetic field at yet unseen input locations $\mathbf{x}_*$ are given by:

$$\mathbb{E}[\mathbf{f}(\mathbf{x}_*)] \approx \nabla\mathbf{\Phi}_* ([\nabla\mathbf{\Phi}]^\mathsf{T}\nabla\mathbf{\Phi} + \sigma_{\text{noise}}^2 \mathbf{\Lambda}^{-1})^{-1} [\nabla\mathbf{\Phi}]^\mathsf{T} \text{vec}(\mathbf{y}),$$
$$\mathbb{V}[\mathbf{f}(\mathbf{x}_*)] \approx \sigma_{\text{noise}}^2 \nabla\mathbf{\Phi}_* ([\nabla\mathbf{\Phi}]^\mathsf{T}\nabla\mathbf{\Phi} + \sigma_{\text{noise}}^2 \mathbf{\Lambda}^{-1})^{-1} [\nabla\mathbf{\Phi}_*]^\mathsf{T}, \tag{22}$$

where $\text{vec}(\cdot)$ is the vectorization operator which converts a matrix to a column vector by stacking its columns on top of each other, such that the $3 \times n$ matrix is converted into a vector of size $3n$. The basis functions $\nabla\mathbf{\Phi}$ and $\nabla\mathbf{\Phi}_*$ need to be evaluated by Equation (19), and $\mathbf{\Lambda}$ by (20).

For this model, the expression for evaluating the log marginal likelihood function for hyperparameter optimization can be written as

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2}\log|\mathbf{K}_{\boldsymbol{\theta}} + \sigma_{\text{noise}}^2 \mathbf{I}_{3n}|+$$
$$\frac{1}{2}\text{vec}(\mathbf{y})^\mathsf{T}(\mathbf{K}_{\boldsymbol{\theta}} + \sigma_{\text{noise}}^2 \mathbf{I}_{3n})^{-1}\text{vec}(\mathbf{y}) + \frac{3n}{2}\log(2\pi), \quad (23)$$

where the quantities can be approximated by

$$\log|\mathbf{K}_{\boldsymbol{\theta}} + \sigma_{\text{noise}}^2 \mathbf{I}_{3n}| \approx (3n - m)\log\sigma_{\text{noise}}^2$$
$$+ \sum_{j=1}^{m}[\mathbf{\Lambda}_{\boldsymbol{\theta}}]_{j,j} + \log|\sigma_{\text{noise}}^2 \mathbf{\Lambda}_{\boldsymbol{\theta}}^{-1} + [\nabla\mathbf{\Phi}]^\mathsf{T}\nabla\mathbf{\Phi}|, \qquad (24)$$
$$\text{vec}(\mathbf{y})^\mathsf{T}(\mathbf{K}_{\boldsymbol{\theta}} + \sigma_{\text{noise}}^2 \mathbf{I}_{3n})^{-1}\text{vec}(\mathbf{y})$$
$$\approx \frac{1}{\sigma_{\text{noise}}^2}\big[\text{vec}(\mathbf{y})^\mathsf{T}\text{vec}(\mathbf{y}) - \text{vec}(\mathbf{y})^\mathsf{T}\nabla\mathbf{\Phi}(\sigma_{\text{noise}}^2 \mathbf{\Lambda}_{\boldsymbol{\theta}}^{-1}$$
$$+ [\nabla\mathbf{\Phi}]^\mathsf{T}\nabla\mathbf{\Phi})^{-1}[\nabla\mathbf{\Phi}]^\mathsf{T}\text{vec}(\mathbf{y})\big], \qquad (25)$$

where the only remaining dependency on the covariance function hyperparameters is in the diagonal matrix $\mathbf{\Lambda}$ defined through the spectral density in Equation (20). In a software implementation, Cholesky decompositions can be employed for numerical stability in the calculation of determinants and matrix inverses. For optimizing the hyperparameters, gradient based optimizers can be employed (see [12] for details on deriving the partial derivatives).

Algorithm 1 describes the step-by-step workflow for applying these equations in practice. The inputs for the method are

Algorithm 1: Algorithm for batch estimation of the scalar potential GP magnetic field with the reduced-rank approach.

Input: $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$, $\mathbf{x}_*$, Ω, m.
Output: $\mathbb{E}[\mathbf{f}(\mathbf{x}_*)], \mathbb{V}[\mathbf{f}(\mathbf{x}_*)]$.
1: Use Eq. (19) to evaluate the basis functions $\nabla\Phi$ from $\mathbf{x}_i$s and Ω.
2: Use Eq. (23) to optimize hyperparameters $\boldsymbol{\theta} = \{\sigma_{\text{lin.}}^2, \sigma_{\text{SE}}^2, \ell_{\text{SE}}, \sigma_{\text{noise}}^2\}$.
3: Use Eq. (19) to evaluate the basis functions $\nabla\Phi_*$ from $\mathbf{x}_*$s and Ω.
4: Solve the GP regression problem by Eq. (22).

Algorithm 2: Algorithm for sequential modeling of the scalar potential GP magnetic field estimate. Alternative (a) corresponds to the sequential model, and (b) to the spatio-temporal modeling approach.

Input: $\mathcal{D}_n$, $\mathbf{x}_*$, Ω, m, $\boldsymbol{\theta}$.
Output: $\mathbb{E}[\mathbf{f}(\mathbf{x}_*)], \mathbb{V}[\mathbf{f}(\mathbf{x}_*)]$.
1: Initialize $\boldsymbol{\mu}_0 = \mathbf{0}$ and $\boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda_\theta}$ from Eq. (20).
2: **for** $i = 1, 2, \ldots, n$ **do**
3:     Evaluate $\nabla\Phi_i$ by Eq. (19) from $\mathbf{x}_i$.
4:     **(a)** Perform an update by Eq. (26).
       (b) Perform an update by Eqs. (31–32).
5:     Evaluate the current prediction at $\mathbf{x}_*$ by Eq. (27).
6: **end for**

the data $\mathcal{D}$ (spatial points and the magnetic field readings), the test points $\mathbf{x}_*$ to predict at, the domain boundaries Ω, and the approximation degree parameter m (controlling the accuracy of the Hilbert space approximation, see Eq. (13)). The algorithm returns the marginal mean and variance of the predicted magnetic field at $\mathbf{x}_*$. The scalar potential could be returned instead by using Equation (18) in step three.

The computational complexity of this method scales as $\mathcal{O}(nm^2)$ for prediction and $\mathcal{O}(m^3)$ for evaluating the marginal likelihood during optimization, which makes it computationally appealing in comparison to the computational complexity of the naive full GP solution which is $\mathcal{O}(n^3)$. The memory requirements scale as $\mathcal{O}(nm)$.

C. Sequential estimation

Many applications require online (sequential) estimates of the magnetic field. The following formulation provides the *same* (within numerical precision) solution as the batch estimation solution (22) in the previous section. The inference scheme in the previous section is in practice the solution of a linear Gaussian estimation problem. Sequential solutions for this type of problems have been extensively studied, and this mathematical formulation is widely known as the *Kalman filter*. The connection between Kalman filtering and Gaussian process regression has recently been studied, for example, in [13, 46, 47].

Reformulation of a batch problem to a sequential algorithm is discussed (with examples) in the book by Särkkä [64]. Following this formulation (and notation to large extent) we may write the following recursion: Initialize $\boldsymbol{\mu}_0 = \mathbf{0}$ and $\boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda_\theta}$ (from the GP prior). For each new observation $i = 1, 2, \ldots, n$ update the estimate according to

$$
\begin{aligned}
\mathbf{S}_i &= \nabla\Phi_i\boldsymbol{\Sigma}_{i-1}[\nabla\Phi_i]^{\mathsf{T}} + \sigma_{\text{noise}}^2 \mathbf{I}_3, \\
\mathbf{K}_i &= \boldsymbol{\Sigma}_{i-1}[\nabla\Phi_i]^{\mathsf{T}}\mathbf{S}_i^{-1}, \\
\boldsymbol{\mu}_i &= \boldsymbol{\mu}_{i-1} + \mathbf{K}_i(\mathbf{y}_i - \nabla\Phi_i\boldsymbol{\mu}_{i-1}), \\
\boldsymbol{\Sigma}_i &= \boldsymbol{\Sigma}_{i-1} - \mathbf{K}_i\mathbf{S}_i\mathbf{K}_i^{\mathsf{T}}.
\end{aligned}
\quad (26)
$$

This means that for a test input location $\mathbf{x}_*$ we get predictions for the mean and the variance of the magnetic field which are given by

$$
\begin{aligned}
\mathbb{E}[\mathbf{f}(\mathbf{x}_*) \mid \mathcal{D}_i] &\approx \nabla\Phi_* \boldsymbol{\mu}_i, \\
\mathbb{V}[\mathbf{f}(\mathbf{x}_*) \mid \mathcal{D}_i] &\approx \nabla\Phi_* \boldsymbol{\Sigma}_i [\nabla\Phi_*]^{\mathsf{T}},
\end{aligned}
\quad (27)
$$

and conditional on the data observed up to observation i. Writing out the conditioning on $\mathcal{D}$ was stripped in the earlier sections for brevity.

Here we do not consider optimization of the hyperparameters $\boldsymbol{\theta}$. The marginal likelihood can be evaluated through the recursion, but in an online setting we suggest optimizing the hyperparameters with some initial batch early in the data collection and then re-optimizing them later on if necessary.

Algorithm 2 presents the scheme of how to apply the equations in practice. The inputs are virtually the same as for the batch algorithm, but now the hyperparameters $\boldsymbol{\theta}$ are considered known. The current estimate can be returned after each iteration loop. The computational complexity of the sequential modeling approach scales as $\mathcal{O}(m^3)$ per update and thus has a total computational complexity of $\mathcal{O}(nm^3)$, but a memory scaling (if intermediate results are not stored) of $\mathcal{O}(m^2)$.

D. Spatio-temporal modeling

The sequential model allows for extending the modeling to also track dynamic changes in the magnetic field without virtually any additional computational burden. Let the data $\mathcal{D}_n = \{(t_i, \mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$ now also comprise a temporal variable t which indicates the time when each observation was acquired.

We present the following spatio-temporal model for tracking changes in the ambient magnetic field. The spatio-temporal GP prior assigned to the scalar potential $\varphi(\mathbf{x}, t)$, depending on both location and time, is defined as follows:

$$
\varphi(\mathbf{x}, t) \sim \mathcal{GP}(0, \kappa_{\text{lin.}}(\mathbf{x}, \mathbf{x}') + \kappa_{\text{SE}}(\mathbf{x}, \mathbf{x}')\kappa_{\text{exp}}(t, t')), \quad (28)
$$

where the additional covariance function $\kappa_{\text{exp}}(t, t')$ defines the prior assumptions of the temporal behavior. This covariance function is defined through

$$
\kappa_{\text{exp}}(t, t') = \exp\left(-\frac{|t - t'|}{\ell_{\text{time}}}\right), \quad (29)
$$

where ℓ_{time} is a hyperparameter controlling the length-scale of the temporal effects. In the temporal domain, this model is also known as the Ornstein–Uhlenbeck process (see, *e.g.*, [7]). The assumption encoded into it is that the phenomenon is

continuous but not necessarily differentiable. Therefore it provides a very flexible means of modeling the changing ambient magnetic field. Also note that in Equation (28) the temporal effects are only associated with the anomaly component, the bias being tracked as a static component.

Following the derivations in Hartikainen and Särkkä [44], we can write down the dynamic state space model associated with the time evolution of the spatio-temporal GP prior model (28)

$$
\begin{aligned}
\mathbf{A}_i &= \mathrm{blkdiag}(\mathbf{I}_3, \mathbf{I}_m \exp(-\Delta t_i/\ell_{\text{time}})), \\
\mathbf{Q}_i &= \mathrm{blkdiag}(\mathbf{0}_3, \mathbf{I}_m[1 - \exp(-2\Delta t_i/\ell_{\text{time}})]),
\end{aligned}
\tag{30}
$$

where $\Delta t_i = t_{i+1} - t_i$ is the time difference between two consecutive samples and $\mathbf{0}_3$ denotes a 3×3 zero matrix.

For the time update (Kalman prediction step) we may thus write

$$
\begin{aligned}
\tilde{\boldsymbol{\mu}}_i &= \mathbf{A}_{i-1}\boldsymbol{\mu}_{i-1}, \\
\tilde{\boldsymbol{\Sigma}}_i &= \mathbf{A}_{i-1}\boldsymbol{\Sigma}_{i-1}\mathbf{A}_{i-1}^\mathsf{T} + \mathbf{Q}_{i-1},
\end{aligned}
\tag{31}
$$

and the modified measurement update (Kalman update step)

$$
\begin{aligned}
\mathbf{S}_i &= \nabla\boldsymbol{\Phi}_i\tilde{\boldsymbol{\Sigma}}_i[\nabla\boldsymbol{\Phi}_i]^\mathsf{T} + \sigma_{\text{noise}}^2\mathbf{I}_3, \\
\mathbf{K}_i &= \tilde{\boldsymbol{\Sigma}}_i[\nabla\boldsymbol{\Phi}_i]^\mathsf{T}\mathbf{S}_i^{-1}, \\
\boldsymbol{\mu}_i &= \tilde{\boldsymbol{\mu}}_i + \mathbf{K}_i(\mathbf{y}_i - \nabla\boldsymbol{\Phi}_i\tilde{\boldsymbol{\mu}}_i), \\
\boldsymbol{\Sigma}_i &= \tilde{\boldsymbol{\Sigma}}_i - \mathbf{K}_i\mathbf{S}_i\mathbf{K}_i^\mathsf{T}.
\end{aligned}
\tag{32}
$$

Algorithm 2 features the workflow of applying the method (option (b)). In practice the only additional input is including the temporal length-scale in $\boldsymbol{\theta}$. The computational costs are the same as for the sequential model.

VI. Experiments

The experiments in this paper are split into four parts. We first demonstrate the feasibility of the scalar potential approach with simulated data, where comparison to known ground truth is possible. After this we present a small-scale proof-of-concept demonstration of the approach. The third experiment is concerned with mapping the magnetic field in a building using a handheld smartphone. The final experiment uses an inexpensive mobile robot for online mapping and real-time tracking of the changing magnetic field. The methods were implemented in Matworks MATLAB® on an Apple MacBook Pro (3.1 GHz Intel Core i7, 16 GB RAM).

A. Simulated experiment

As a first part of our experimental validation of the model, we present a simulation study. This will be used to illustrate our method and to quantify its performance through Monte Carlo simulations. In the simulations, we assume that the magnetic field measurements can indeed be modeled using a scalar potential as argued in Section III. Hence, we simulate the magnetometer data from the GP that models the magnetometer measurements as gradients of a scalar potential field as discussed in Section IV-D. Because we are interested in simulating a large number of data points, simulating this data using a full GP approach is computationally very expensive. We therefore simulate the data using the computationally

efficient approach described in Section V-B. We use a large number of basis functions ($m = 4096$). This has been shown to be a good approximation of the true model in [12]. We will also show that this is a good approximation in Section VI-B based on experimental data.

The magnetometer is assumed to move in a three-dimensional volume with $x, y, z \in [-0.4, 0.4]$ such that $\mathbf{x} = (x, y, z)$. The training data used for training the GP is randomly uniformly distributed over this volume. A validation data set is used to assess the predictive power of the trained GP. This validation data is a three-dimensional meshgrid over the same volume and consists of $n_{\text{val.}} = 9261$ positions $\mathbf{x}_{\text{val.}}$ with a true magnetic field $\mathbf{f}_{\text{true}}(\mathbf{x}_{\text{val.}})$. The magnetic field is predicted at these points using the trained GP, leading to $\mathbf{f}_{\text{train.}}(\mathbf{x}_{\text{val.}})$, after which the quality of the GP solution can be assessed in terms of the root mean square error (RMSE). We set the domain Ω in the GP model to $L_1 = L_2 = L_3 = 0.5$ and simulate using the following hyperparameters $\sigma_{\text{const.}}^2 = 0.3, \sigma_{\text{SE}}^2 = 1, \ell_{\text{SE}} = 0.1$, and $\sigma_{\text{noise}}^2 = 0.04$ (see Sec. V-B and IV-D for more details on the notation).

We compare our proposed method with two other approaches. The first is the approach by Vallivaara et al. [26, 27], where the magnetic field is modeled using independent GPs for all three components. Each GP consists of a constant and a squared exponential kernel and has its own hyperparameters. The second approach, considered in Kemppainen et al. [57], models the magnetic field similarly but with shared hyperparameters. For details, see also Section IV-C.

In a first set of Monte Carlo simulations, we analyze the performance of the three different approaches depending on the number of (randomly distributed) training data points, that is in terms of the sparseness of the magnetometer data and the amount of interpolation that is needed for prediction. For all three approaches we use a large number of basis functions ($m = 4096$). They are hence expected to approach the performance of the full GP solution. To exclude problems with local minima in the hyperparameter optimization—which will be the topic of a second set of simulations—the hyperparameter optimization is started in the values used for simulating the data. The results from 30 Monte Carlo simulations are shown in Figure 5a. Naturally, the more data is used for training the GP, the smaller the RMSE becomes. The GP that models the magnetic field measurements as gradients of a scalar potential field, outperforms the other two approaches, independent of the amount of training data used. This can be understood from the fact that this approach incorporates most physical knowledge.

In a second set of Monte Carlo simulations, the sensitivity to the initialization of the hyperparameter optimization is analyzed for the three different methods. Only one simulated data set is used but the hyperparameter optimization is started in 30 randomly selected sets of hyperparameters $\boldsymbol{\theta}_0^{\mathrm{a}}$ for a varying length of the training data. The hyperparameters are assumed to lie around the estimates that are obtained using the same optimization strategy as above for 8000 data points. Hence, these sets of hyperparameters $\boldsymbol{\theta}_{\text{true}}^{\mathrm{a}}$ are known to results in small RMSE values as depicted in Figure 5a. The superscript 'a' on $\boldsymbol{\theta}_{\text{true}}^{\mathrm{a}}$ is used to explicitly denote that
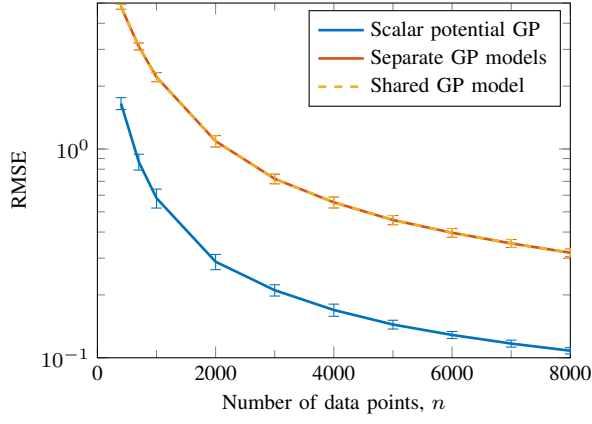
(a) Randomized over data sets
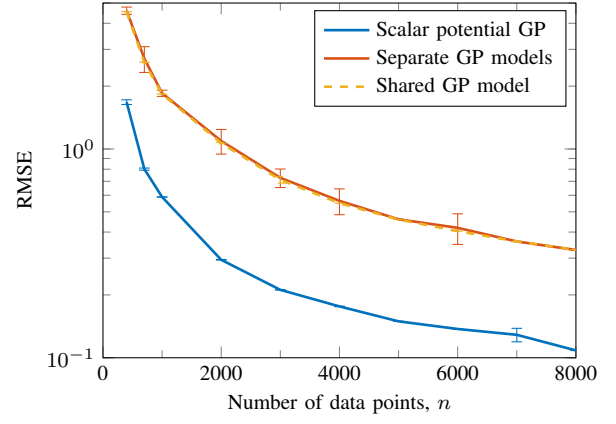
(b) Randomized over initializations

Fig. 5: The average RMSE (with standard deviations) from 30 Monte Carlo simulations as a function of the number of simulated data points used for training the GP. The constantly lower error for the Hilbert space scalar potential GP in comparison to the separate GP models and shared hyperparameter GP model is explained by the additional prior physical knowledge encoded into the model. In (b), the shared hyperparameter GP shows a slight advantage over the fully independent models.
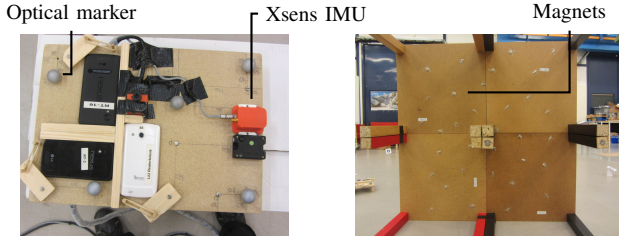
Fig. 6: Setup of the experiment: On left the sensor board that was used for collecting the data. The right side figure shows the magnetic environment used in the experiment shown from below comprising small magnets to ensure sufficient excitation of the magnetic field.

these sets of hyperparameters actually differ between the three different approaches. For the approach where the magnetic field components are modeled using independent GPs, each component results in a set of hyperparameters $\theta_{\text{true}}^{\text{a}}$. For simplicity, in this approach, $\theta_{\text{true}}^{\text{a}}$ is chosen to be the mean of these three sets of hyperparameters.

For each of the three approaches, the initial parameters θ_0 are then assumed to deviate from $\theta_{\text{true}}^{\text{a}}$ by at most 70% as

$$\theta_0 = \theta_{\text{true}}^{\text{a}} \left[1 + 0.7 \, \mathrm{U}(-1, 1) \right]. \tag{33}$$

The Monte Carlo simulation results are depicted in Figure 5b. As can be seen, the approach which models the three magnetic field components using separate GPs suffers most from local minima. Our proposed model using a scalar potential still outperforms the other two.

B. Empirical proof-of-concept data

To illustrate our approach using real data, we have performed an experiment where a number of sensors have been moved around in a magnetic environment. The sensor board used is shown in Figure 6. We use the magnetometer data from an Xsens MTi (Xsens Technologies B.V.,

http://www.xsens.com). Accurate position and orientation information is obtained using an optical system. These high-accuracy measurements were provided through the use of a Vicon real-time tracking system (Vicon Motion Systems Ltd., UK, http://www.vicon.com).

We obtain measurements while sliding the sensor board over a configuration of small tables. To ensure sufficient excitation, magnets have been placed in an irregular pattern underneath these tables as shown in Figure 6. Two different data sets have been collected. Both consist of approximately three minutes of data sampled at 100 Hz. One data set is used for training, while the second is for validation. The data of both the training and validation data sets are displayed in Figure 7a. To give an impression of the spatial variation of the magnetic field, the magnetic field intensity has been visualized through the colors of the data. Note that the magnetometer is calibrated such that it has a magnitude of one in a local undisturbed magnetic field.

The magnetometer inside the IMU measures the magnetic field at the different locations. The optical measurements are used for two purposes. First, the positions from the optical system are used as known locations in the GP approach. This is a fairly reasonable assumption due to the high accuracy of the measurements of the optical system. Second, the orientations estimated by the optical system are used to rotate the magnetometer measurements from the magnetometer sensor frame to the lab frame. This rotation of the magnetic field measurements is needed for any of the GP methods discussed in this work. To use the optical and magnetometer data together, they need to be time synchronized. This synchronization is done in post-processing by correlating the angular velocities measured by the optical system and by the gyroscope in the IMU.

We run Algorithm 1 for the training data set. The domain Ω in the GP model is set as $L_1 = L_2 = 0.6$ and $L_3 = 0.1$. The actual two-dimensional movement is performed in a rectangle of $80\,\text{cm} \times 100\,\text{cm}$ and is hence well within the domain Ω. The learned scalar potential and its marginal variance (uncertainty) are shown in Figure 7b. Because the training data covers

(a) Training data (left) and validation data (right)

(b) Learned scalar potential (left) and its uncertainty (right)

(c) The predicted magnetic field: $\mathbf{H}_x, \mathbf{H}_y$, and $\mathbf{H}_z$
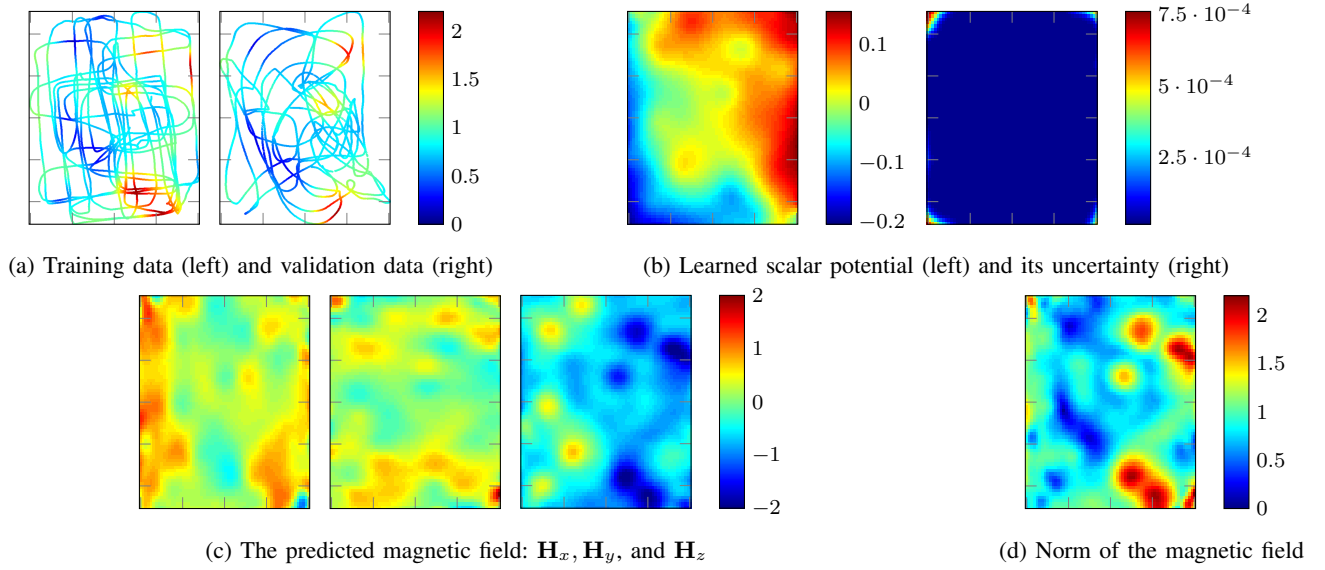
(d) Norm of the magnetic field

Fig. 7: Illustration of the magnetic field data and the results from the GP approach from Algorithm 1 for the experiment discussed in Section VI-B. In all figures, the y-axis is $-0.5, \ldots, 0.5$ m. The x-axis is $-0.4, \ldots, 0.4$ m. The units in the field surface plots are arbitrary due to normalization.

almost the whole displayed area, the learned scalar potential has very low uncertainty. It is also possible to compute the predicted magnetic field which is shown in Figure 7c. For completeness, the intensity of these predicted magnetic field measurements are shown in Figure 7d. Although this quantity is only indirectly related to the outcome of the GP approach, it is frequently used in the remaining sections because of its easy and intuitive visualization.

By predicting the measurements at the locations of the validation data set, it is also possible to compute the RMSE on the validation data. In Figure 8 we visualize the RMSE as a function of the number of basis functions used in Algorithm 1. We also compare to the RMSE from a full GP approach, that is using the same GP prior but without the Hilbert space approximation scheme to speed up the inference. To allow for comparison with a full GP approach—which suffers from a high computational complexity for large data sets—the data has been downsampled to 5 Hz. As can be seen, already for around $m = 1000$ basis functions, the quality of the estimates from Algorithm 1 approaches that of the full GP approach.

C. Mapping the magnetic field in a building

The third experiment was concerned with estimating a map of the magnetic environment inside a building by only using a smartphone for the data collection. The mapped venue is located on the Aalto University campus, and a floor plan sketch is shown in Figure 9. For practical reasons, we limited our interest to the lobby which is approximately 600 m^2 in size.

For the measurements, we used an Apple iPhone 4 and its built-in 9-dof IMU (3-axis AKM AK8975 magnetometer). All sensors were sampled at 50 Hz, and the data was streamed online to a laptop computer for processing and storing. The phone was held at waist-height and pointed towards the heading direction.
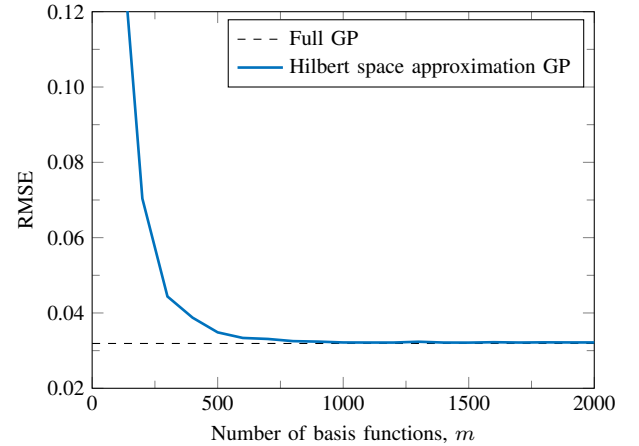

Fig. 8: RMSE as a function of the number of basis functions used in Algorithm 1 as compared to a validation data set. The approximative model approaches the full GP approach.

For reconstructing the walking path and phone orientation, we used a pedestrian dead-reckoning (PDR) approach developed at IndoorAtlas (IndoorAtlas Ltd., Finland, http://www.indooratlas.com), where only accelerometer and gyroscope readings were used—the path reconstruction thus being fully independent of the magnetometer readings. The alignment to the map and drift correction were inferred from a set of fixed points along the path during acquisition. Two sample paths are shown in Figure 9: one of the three training paths with similar routes and the validation path. The reconstructed paths were visually checked to match the 'true' walking paths.

We covered the walkable area in the lobby with three walking paths following the same route in each of them (see Fig. 9). The reconstructed paths were approximately 242, 253, and
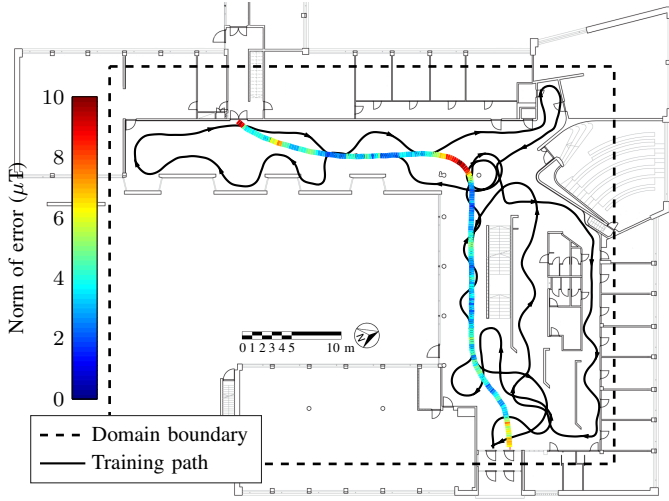
Fig. 9: A training (black) and validation (colored) free-walking path that was used in the experiment. Trajectories were collected by a mobile phone, and the magnetometer data was corrected for gravitation direction and heading using the inertial sensors in the device. The domain boundaries for the reduced-rank method are shown by the dashed line.
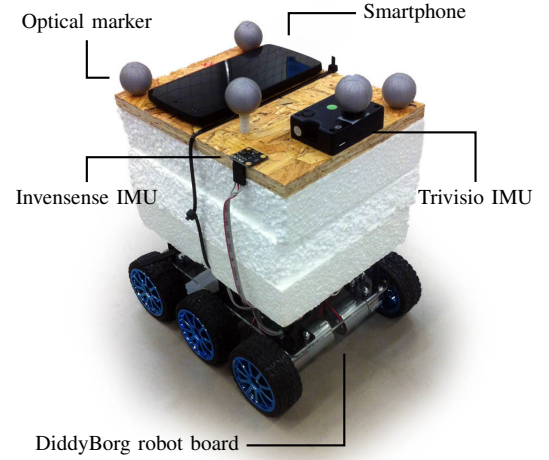

Fig. 10: The robot was built on a DiddyBorg robotics board and controlled by a Raspberry Pi single-board computer. The three sensors (Invensense, Trivisio, and smartphone) providing magnetometer readings were mounted on the top. The reference locations were provided by a Vicon optical tracking system.

302 meters long, respectively. The number of magnetometer data samples acquired along the paths were 9868, 10500, and 12335. Prior to each acquisition, the phone magnetometer was calibrated by a standard spherical calibration approach. The combined size of the training data set was $n = 32703$. For validation, we collected a walking path passing through the venue (length 54 m, $n = 2340$).

We considered a batch interpolation problem of creating a magnetic map of the lobby. The map was assumed static over time, and we applied Algorithm 1 to the training data with $m = 1024$ basis functions. The optimized hyperparameters were $\sigma_{\text{lin.}}^2 \approx 575\,(\mu\text{T})^2$, $\sigma_{\text{SE}}^2/\ell_{\text{SE}}^2 \approx 373\,(\mu\text{T})^2$, $\ell_{\text{SE}} \approx 1.87\,\text{m}$, and $\sigma_{\text{noise}}^2 \approx 5.53\,(\mu\text{T})^2$. The coefficients of the inferred linear (bias) model corresponding to the linear covariance function was $(-1.095, 12.995, -41.119)$.

Figure 1 shows the interpolated magnetic field magnitude ($\|\mathbf{f}\|$) and the vector field components. The component-wise maximum of the associated marginal standard deviation fields is visualized by the degree of transparency (with a standard deviation of $5\,\mu\text{T}$ being fully transparent). The overall shape of the estimate agreed even when the model was trained separately with each of the training paths. To most part, the strong fluctuations in the magnetic field are located near walls or other structures in the building. The strong magnetic field in the open area in the lower right part of the floorplan was identified to most likely be due to a large supporting structure on the lower floor-level. We also used the model for predicting the measurements along the validation path (see Fig. 9). The component-wise RMSEs were $(2.35\,\mu\text{T}, 3.05\,\mu\text{T}, 2.71\,\mu\text{T})$ and mean absolute errors $(1.72\,\mu\text{T}, 2.42\,\mu\text{T}, 2.03\,\mu\text{T})$. Figure 9 shows the norm of the error along the validation path. The measurement noise level of the magnetometer is in the magnitude of $1\,\mu\text{T}$ and the uncertainty in the PDR estimate contributes to the remaining variance.

D. Online mapping

Finally, we demonstrate the power of sequential updating and time-dependent magnetic field estimation. The mapping was performed by a lightweight and inexpensive mobile robot equipped with a magnetometer, and the task was to obtain an estimate of the magnetic environment of an indoor space by re-calculating the estimate in an online fashion. In the second part of the experiment the magnetic environment was abruptly changed during the experiment, and the aim was to catch this phenomenon by spatio-temporal modeling.

We used a robot for collecting the data (see Fig. 10). The robot was built on a DiddyBorg (PiBorg Inc., UK, http://www.piborg.org) robotics board, controlled by a Raspberry Pi 2 (model B) single-board computer (Raspberry Pi Foundation, UK, http://www.raspberrypi.org). For this example, we controlled the robot over Bluetooth with a joystick.

The robot was equipped with a 9-dof MPU-9150 Invensense IMU unit that was sampled at 50 Hz. The data were collected and stored internally on the Raspberry Pi. For additional validation, a Trivisio Colibri wireless IMU (TRIVISIO Prototyping GmbH, http://www.trivisio.com), sampled at 100 Hz, and a Google Nexus 5 smartphone (AKM AK8963 3-axis magnetometer), sampled at 50 Hz, were also mounted on the robot for checking the quality of the Invensense IMU data and ensure the repeatability of the experiment. To reduce disturbances caused by the robot, the sensors were mounted on an approximately 20 cm thick layer of Styrofoam. During post-processing the data, the sensor positions and alignments on the robot were corrected for.

High-accuracy location and orientation reference measurements were provided through the use of a Vicon real-time tracking system. The location measurements could alternatively be recovered by odometry and heading information

(a) Final interpolated magnitude map

(b) Vector field

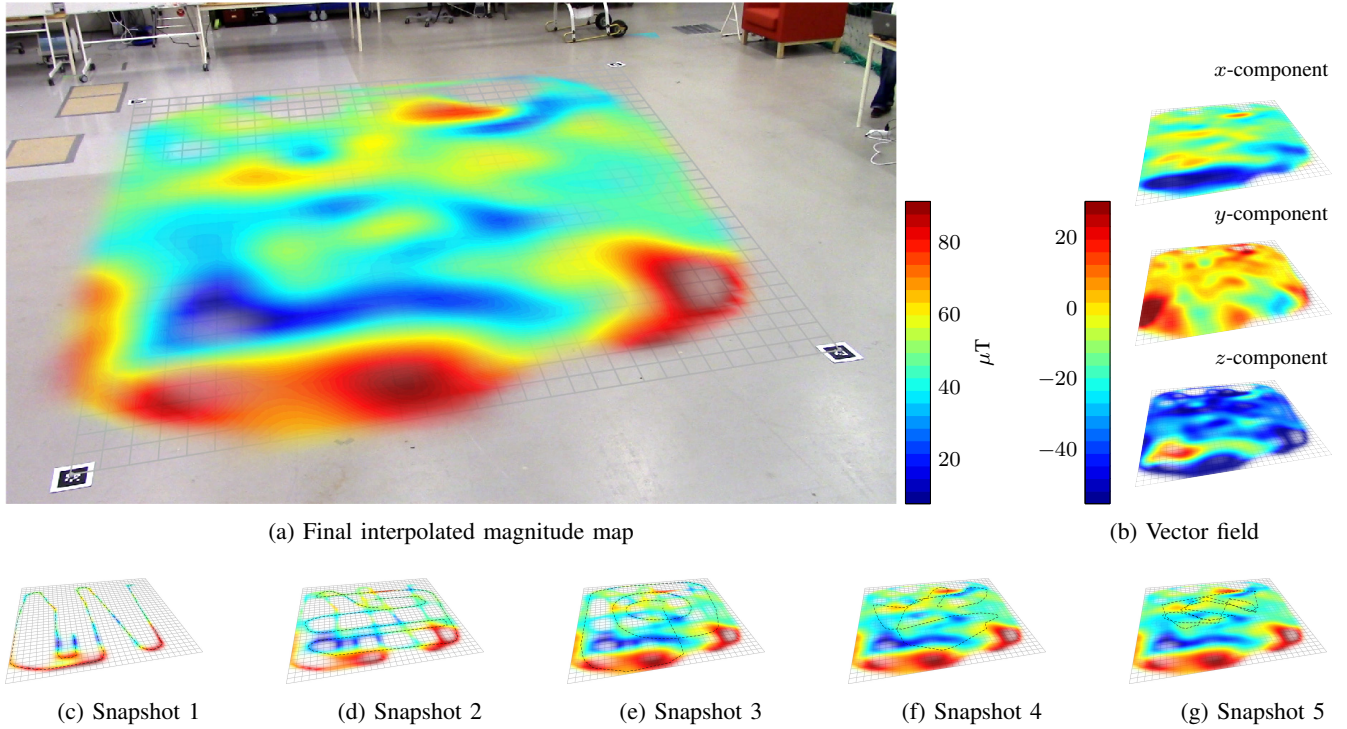(c) Snapshot 1   (d) Snapshot 2   (e) Snapshot 3   (f) Snapshot 4   (g) Snapshot 5

Fig. 11: The GP interpolation task the robot was faced with. The final interpolation outcome of the magnitude field is shown in (a), and the different vector field components are shown in (b). Snapshots along the temporally updating field estimate are shown in (c–g) together with the path travelled since the previous update. The marginal variance (uncertainty) is visualized by the degree of transparency.

provided by the robot, but the interest in this experiment was rather to focus on the interpolation of the magnetic field, not the path estimation.

The task was to map the magnetic field inside a marked region roughly 6 m × 6 m in size. The size of the region was limited by the field of view of the Vicon system. The magnetometers were calibrated in the beginning of the measurement session by rotating the robot around all of its axes. A standard spherical calibration approach was used. Due to limits in acquisition length of the Vicon system, we captured the data in parts, each roughly three minutes in length. The magnetic environment remained unchanged for the first five data sets (paths shown in Figs. 11c–11g), and later on changes in the field were initiated by bringing in large metallic toolbox shelves.

In the first part of the experiment, for interpolating the magnetic field we used the sequential reduced-rank scalar-potential approach presented in Algorithm 2. We assumed the magnetic environment to be stationary, and performed sequential updates in an online fashion. For practical reasons the calculations were done off-line, but the algorithm is fast enough for running in real-time online estimation. The rank of the approximation was fixed to $m = 1024$.

The length-scale, magnitude, and noise variance hyperparameters were learned from the first two data sets ($n = 17980$ vector valued observations) by maximizing with respect to marginal likelihood (see Alg. 1). The obtained values were $\ell_{\mathrm{SE}} \approx 0.32$ m, $\sigma_{\mathrm{SE}}^2/\ell_{\mathrm{SE}}^2 \approx 287\,(\mu\mathrm{T})^2$ and $\sigma_{\mathrm{noise}}^2 \approx 3.27\,(\mu\mathrm{T})^2$. The linear model magnitude scale parameter was fixed to

$\sigma_{\mathrm{lin.}}^2 = 500\,(\mu\mathrm{T})^2$. The noise model is not only capturing the sensor measurement noise, but the entire mismatch between the data and the model. This explains the rather large noise variance. We also checked, that the hyperparameter estimates remained stable when optimized using the rest of the data.

Figure 11 shows the results for the static magnetic field experiment. The estimate was updated continuously five times a second, and we show five snapshots of the evolution of the magnetic field estimate in Figures 11c–11g (vector field magnitude shown in figures). These snapshots also show the path travelled since the previous snapshot. The alpha channel acts as a proxy for uncertainty; the marginal variance of the estimate is giving the degree of transparency. The final magnitude estimate—after iterating through all the $n = 43029$ observations—is shown in Figure 11a together with the vector field components in Figure 11b.

The frontal part of the mapped region shows strong magnetic activity, whereas the parts further back do not show as strong fields. Inspection of the venue suggested metallic pipelines or structures in the floor to blame (or thank) for these features. In this particular case most parts of the effect is seen in the x-component. We repeated the reconstruction with data collected from the Trivisio and smartphone sensors, and the results and conclusions remained unchanged. As a supplementary file to this paper, there is a video[1] demonstrating the online operation which has been sped-up $50\times$.

---

[1] The supplementary video is available on YouTube: https://www.youtube.com/watch?v=enlMiUqPVJo

The last part of the experiment was dedicated to dynamical (time-dependent) modeling of the magnetic field. We used all the data from the first part of the experiment to train a sequential model and used that as the starting point for changing the field ($t = 0$). During acquisition of data while the robot was driving around, we brought in two metallic toolbox shelves: first a larger toolbox shelf on wheels (Fig. 12c) and then a smaller box (Fig. 12d). We acquired altogether some $300\,\mathrm{s}$ of data ($n = 15513$) of the changed environment.

For encoding the assumptions of a changing magnetic field, we used Algorithm 2. The additional hyperparameter controlling the temporal scale was fixed to $\ell_{\mathrm{time}} = 1$ hour, thus encoding an assumption of slow local changes. This choice is not restrictive, because the data is very informative about the abrupt changes.

Figure 12a shows the evolution of the magnetic field components for one fixed location (indicated by a cross marker in the figures). The two toolboxes induce clear changes in the local anomaly field, but the effects are restricted to the immediate vicinity of the boxes. Thus the spatio-temporal model only gains information about the changed field, when the robot passes by the location of interest. This effect is clearly visible around $t = 20\,\mathrm{s}$, and later on around $t = 70\,\mathrm{s}$ and $t = 130\,\mathrm{s}$. After this the estimate stabilizes and only drifts around for the remaining time. Even though the changes in the field components appear clear in Figure 12a, they are only around $2\,\mu\mathrm{T}$ and thus only account for a variation of about 2% in the scale of the entire field visualized in Figure 11b. Inducing more noticeable changes in the magnetic field would require moving around larger structures (say an elevator). Yet, even changes this small can be tracked by the modeling approach.

VII. DISCUSSION

In recent years, interest has emerged in mapping of the magnetic field by Gaussian processes for robot and pedestrian localization. This paper has aimed at presenting a new efficient method for mapping, but also at providing a study of best practices in using GPs in this context. Thus, we went through three different approaches for formulating GP priors for the magnetic field (independent, shared hyperparameters, and curl-free/scalar potential). These three models differ in the amount of prior knowledge encoded in the model, and the more information available in the prior, the better the interpolation and extrapolation capabilities in the model—as long as the data agrees with the assumptions. This was also demonstrated in Figure 4 and in Section VI-A.

The methods presented in this paper are related to the use of Gaussian random field priors in inverse problems [65, 66]. This connection has been explored from various points of views (*cf.* [67]). However, the machine learning [7] way of interpreting the GP priors indeed brings something new on the table— we are explicitly modeling the uncertainty in the field by using a stochastic model, which has interesting philosophical implications. The formulation of the prior through a GP covariance function provides both an intuitive and theoretically justified way of encoding the information.

The scalar potential approach is not the only way to build the model. It would also be possible to include disturbance
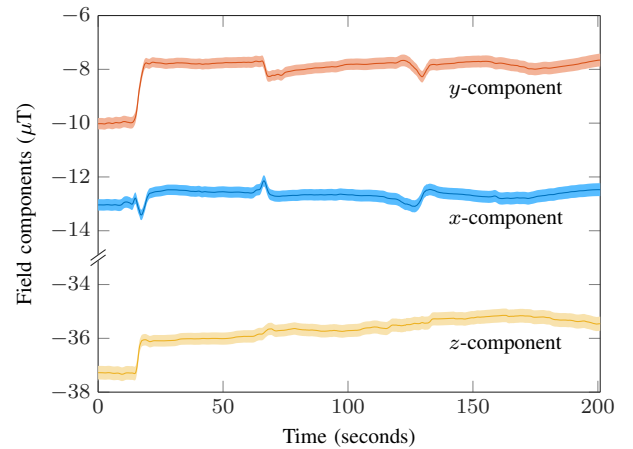
(a) Magnetic field estimate at a fixed location

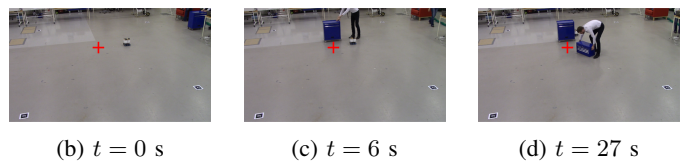(b) $t = 0$ s   (c) $t = 6$ s   (d) $t = 27$ s

Fig. 12: (a) Evolution of the magnetic field at one spatial location over a time-course of 200 seconds. (b–d) The blue metallic toolboxes are brought in at the beginning of the experiment. The abrupt changes in the field estimate corresponds to time instances when the robot has passed the toolboxes and gained information about the changed environment.

in the model, which would model the effect of free currents in the area or its boundaries. Furthermore, more complicated assumptions of the temporal time-changing behavior of the magnetic field could be included in the temporal covariance function. For example, various degrees of smoothness or periodicity could be included in the framework.

This paper has been considering the 'M' (mapping) part in SLAM. The 'L' (localization) part based on these maps is presented in Solin et al. [1]. The online mapping scheme presented in Algorithm 2 opens up for possibilities for simultaneously building the map and localization within the map. As seen in the experiments presented both in this work and in Solin et al. [1], this appears feasible and provides an interesting direction for further research.

VIII. CONCLUSION

Small variations in the magnetic field can be used as inputs in various positioning and tracking applications. In this paper, we introduced an effective and practically feasible approach for mapping these anomalies. We encoded prior knowledge from Maxwell's equations for magnetostatics into a Bayesian non-parametric probabilistic model for interpolation and extrapolation of the magnetic field.

The magnetic vector field components were modeled jointly by a Gaussian process model, where the prior was associated directly with a latent scalar potential function. This ensures the field to be curl-free—a justified assumption in free spaces. This assumption couples the vector field components and

additionally encodes the assumption of a baseline field with smooth small-scale variations. We also presented connections to existing formulations for vector-valued Gaussian process models.

In addition to constructing the model, we also presented a novel and computationally efficient inference scheme for interpolation and extrapolation using it. We built upon a Laplace operator eigenbasis approach, which falls natural to the formulation of the model. The inference scheme ensures a linear computational complexity with respect to the number of observations of the magnetic field. We also extended the method to an online approach with sequential updating of the estimate and time-dependent changes in the magnetic field.

We presented four experiments demonstrating the feasibility and practicality of the methods. A simulated experiment showed the benefit of including additional knowledge from physics into the model, and a simple proof-of-concept example demonstrated the strength of the approximation scheme in solving the model. Two real-world use cases were also considered: we mapped the magnetic field in a building on foot using a smartphone, and demonstrated online mapping using a wheeled robot.

REFERENCES

[1] A. Solin, S. Särkkä, J. Kannala, and E. Rahtu, "Terrain navigation in the magnetic landscape: Particle filtering for indoor positioning," in *Proceedings of the European Navigation Conference*, 2016, pp. 1–9.

[2] O. J. Woodman, "Pedestrian localisation for indoor environments," Ph.D. dissertation, University of Cambridge, United Kingdom, 2010.

[3] J. D. Hol, "Sensor fusion and calibration of inertial sensors, vision, ultra-wideband and GPS," Ph.D. dissertation, Linköping University, Sweden, 2011.

[4] J. J. Leonard and H. F. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," in *Proceedings of the International Workshop of Intelligent Robots and Systems (IROS)*, vol. 3. IEEE, 1991, pp. 1442–1447.

[5] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[6] A. O'Hagan, "Curve fitting and optimal design for prediction (with discussion)," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 40, no. 1, pp. 1–42, 1978.

[7] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[8] N. A. Cressie, *Statistics for Spatial Data*. New York: Wiley-Interscience, 1993.

[9] N. Cressie and C. K. Wikle, *Statistics for Spatio-Temporal Data*. Hoboken, NJ: John Wiley & Sons, 2011.

[10] M. P. Deisenroth, D. Fox, and C. E. Rasmussen, "Gaussian processes for data-efficient learning in robotics and control," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 2, pp. 408–423, 2015.

[11] N. Wahlström, M. Kok, T. B. Schön, and F. Gustafsson, "Modeling magnetic fields using Gaussian processes," in *Proceedings of the 38th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 3522–3526.

[12] A. Solin and S. Särkkä, "Hilbert space methods for reduced-rank Gaussian process regression," *ArXiv preprint arXiv:1401.5508*, 2014.

[13] S. Särkkä, A. Solin, and J. Hartikainen, "Spatiotemporal learning via infinite-dimensional Bayesian filtering and smoothing," *IEEE Signal Processing Magazine*, vol. 30, no. 4, pp. 51–61, 2013.

[14] M. N. Nabighian, V. J. S. Grauch, R. O. Hansen, T. R. LaFehr, Y. Li, J. W. Peirce, J. D. Phillips, and M. E. Ruder, "The historical development of the magnetic method in exploration," *Geophysics*, vol. 70, no. 6, pp. 33ND–61ND, 2005.

[15] A. Guillen, P. Calcagno, G. Courrioux, A. Joly, and P. Ledru, "Geological modelling from field data and geological knowledge: Part II. Modelling validation using gravity and magnetic data inversion," *Physics of the Earth and Planetary Interiors*, vol. 171, no. 1, pp. 158–169, 2008.

[16] P. Calcagno, J.-P. Chilès, G. Courrioux, and A. Guillen, "Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules," *Physics of the Earth and Planetary Interiors*, vol. 171, no. 1, pp. 147–157, 2008.

[17] F. Mackay, R. Marchand, and K. Kabin, "Divergence-free magnetic field interpolation and charged particle trajectory integration," *Journal of Geophysical Research: Space Physics*, vol. 111, no. A6, p. A06208, 2006.

[18] B. K. Bhattacharyya, "Bicubic spline interpolation as a method for treatment of potential field data," *Geophysics*, vol. 34, no. 3, pp. 402–423, 1969.

[19] V. Springel, "Smoothed particle hydrodynamics in astrophysics," *Annual Review of Astronomy and Astrophysics*, vol. 48, pp. 391–430, 2010.

[20] J. Haverinen and A. Kemppainen, "Global indoor self-localization based on the ambient magnetic field," *Robotics and Autonomous Systems*, vol. 57, no. 10, pp. 1028–1035, 2009.

[21] B. Li, T. Gallagher, A. G. Dempster, and C. Rizos, "How feasible is the use of magnetic field alone for indoor positioning?" in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012, pp. 1–9.

[22] M. Angermann, M. Frassl, M. Doniec, B. J. Julian, and P. Robertson, "Characterization of the indoor magnetic field for applications in localization and mapping," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012, pp. 1–9.

[23] E. Le Grand and S. Thrun, "3-axis magnetic field mapping and fusion for indoor localization," in *IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2012, pp. 358–364.

[24] P. Robertson, M. Frassl, M. Angermann, M. Doniec, B. J. Julian, M. Garcia Puyol, M. Khider, M. Lichtenstern, and L. Bruno, "Simultaneous localization and mapping for pedestrians using distortions of the local magnetic field intensity in large indoor environments," in *Proceedings of the International Conference*

on *Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2013, pp. 1–10.

[25] M. Frassl, M. Angermann, M. Lichtenstern, P. Robertson, B. J. Julian, and M. Doniec, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 913–920.

[26] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Röning, "Simultaneous localization and mapping using ambient magnetic field," in *Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2010, pp. 14–19.

[27] ——, "Magnetic field-based SLAM method for solving the localization problem in mobile robot floor-cleaning task," in *Proceedings of the 15th International Conference on Advanced Robotics (ICAR)*, 2011, pp. 198–203.

[28] S. T. O'Callaghan and F. T. Ramos, "Gaussian process occupancy maps," *The International Journal of Robotics Research*, vol. 31, no. 1, pp. 42–62, 2012.

[29] M. Smith, I. Posner, and P. Newman, "Adaptive compression for 3D laser data," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 914–935, 2011.

[30] S. Kim and J. Kim, "Hierarchical Gaussian processes for robust and accurate map building," in *Proceedings of Australasian Conference on Robotics and Automation*, 2015, pp. 117–124.

[31] F. Ramos and L. Ott, "Hilbert maps: Scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, 2016.

[32] R. Senanayake, L. Ott, S. O'Callaghan, and F. T. Ramos, "Spatio-temporal Hilbert maps for continuous occupancy representation in dynamic environments," in *Advances in Neural Information Processing Systems 29*. Curran Associates, Inc., 2016, pp. 3925–3933.

[33] T. Vidal-Calleja, D. Su, F. De Bruijn, and J. V. Miro, "Learning spatial correlations for Bayesian fusion in pipe thickness mapping," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 683–690.

[34] C. H. Tong, P. Furgale, and T. D. Barfoot, "Gaussian process Gauss–Newton for non-parametric simultaneous localization and mapping," *The International Journal of Robotics Research*, vol. 32, no. 5, pp. 507–525, 2013.

[35] B. Ferris, D. Hähnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *Proceedings of Robotics: Science and Systems*, 2006, pp. 303–310.

[36] S. Barkby, S. B. Williams, O. Pizarro, and M. V. Jakuba, "Bathymetric particle filter SLAM using trajectory maps," *The International Journal of Robotics Research*, vol. 31, no. 12, pp. 1409–1430, 2012.

[37] T. D. Barfoot, C. H. Tong, and S. Särkkä, "Batch continuous-time trajectory estimation as exactly sparse Gaussian process regression," in *Proceedings of Robotics: Science and Systems*, 2014, pp. 1–9.

[38] S. Anderson, T. D. Barfoot, C. H. Tong, and S. Särkkä, "Batch nonlinear continuous-time trajectory estimation as exactly sparse Gaussian process regression," *Autonomous Robots*, vol. 39, no. 3, pp. 221–238, 2015.

[39] J. Quiñonero-Candela and C. E. Rasmussen, "A unifying view of sparse approximate Gaussian process regression," *Journal of Machine Learning Research*, vol. 6, pp. 1939–1959, 2005.

[40] M. Lázaro-Gredilla, J. Quiñonero-Candela, C. E. Rasmussen, and A. R. Figueiras-Vidal, "Sparse spectrum Gaussian process regression," *Journal of Machine Learning Research*, vol. 11, pp. 1865–1881, 2010.

[41] J. Hensman, N. Durrande, and A. Solin, "Variational Fourier features for Gaussian processes," *Journal of Machine Learning Research*, vol. 8, no. 151, pp. 1–52, 2018.

[42] C. J. Paciorek, "Bayesian smoothing with Gaussian processes using Fourier basis functions in the spectralGP package," *Journal of Statistical Software*, vol. 19, no. 2, pp. 1–38, 2007.

[43] J. Fritz, I. Neuweiler, and W. Nowak, "Application of FFT-based algorithms for large-scale universal kriging problems," *Mathematical Geosciences*, vol. 41, no. 5, pp. 509–533, 2009.

[44] J. Hartikainen and S. Särkkä, "Kalman filtering and smoothing solutions to temporal Gaussian process regression models," in *Proceedings of the International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, 2010, pp. 379–384.

[45] S. Reece and S. Roberts, "An introduction to Gaussian processes for the Kalman filter expert," in *Proceedings of the 13th International Conference on Information Fusion (FUSION)*, 2010, pp. 1–9.

[46] M. Osborne, "Bayesian Gaussian processes for sequential prediction, optimisation and quadrature," Ph.D. dissertation, Department of Engineering Science, University of Oxford, United Kingdom, 2010.

[47] M. F. Huber, "Recursive Gaussian process: On-line regression and learning," *Pattern Recognition Letters*, vol. 45, pp. 85–91, 2014.

[48] M. A. Álvarez, D. Luengo, and N. D. Lawrence, "Linear latent force models using Gaussian processes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2693–2705, 2013.

[49] M. A. Álvarez and N. D. Lawrence, "Latent force models," in *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*, ser. JMLR W&CP, vol. 5, 2009, pp. 9–16.

[50] J. Hartikainen and S. Särkkä, "Sequential inference for latent force models," in *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, 2011, pp. 311–318.

[51] S. Särkkä and J. Hartikainen, "Infinite-dimensional Kalman filtering approach to spatio-temporal Gaussian process regression," in *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*, ser. JMLR W&CP, vol. 22, 2012, pp. 993–1001.

[52] R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory*. Springer–Verlag, 1978.

[53] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. New York: Wiley, 1999.

[54] J. Vanderlinde, *Classical Electromagnetic Theory*. Dordrecht: Springer, 2004.

[55] D. J. Griffiths and R. College, *Introduction to Electrodynamics*. Upper Saddle River, NJ: Prentice hall, 1999.

[56] N. Wahlström, "Modeling of magnetic fields and extended objects for localization applications," Ph.D. dissertation, Linköping University, Sweden, 2015.

[57] A. Kemppainen, J. Haverinen, I. Vallivaara, and J. Röning, "Near-optimal exploration in Gaussian process SLAM: Scalable optimality factor and model quality rating," in *Proceedings of the 5th European Conference on Mobile Robots (ECMR)*, 2011, pp. 283–289.

[58] J. Jung, T. Oh, and H. Myung, "Magnetic field constraints and sequence-based matching for indoor pose graph SLAM," *Robotics and Autonomous Systems*, vol. 70, pp. 92–105, 2015.

[59] A. Viseras Ruiz and C. Olariu, "A general algorithm for exploration with Gaussian processes in complex, unknown environments," in *Proceeding of the IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 3388–3393.

[60] E. J. Fuselier, Jr, "Refined error estimates for matrix-valued radial basis functions," Ph.D. dissertation, Texas A&M University, 2007.

[61] L. Baldassarre, L. Rosasco, A. Barla, and A. Verri, "Vector field learning via spectral filtering," in *Machine Learning and Knowledge Discovery in Databases*, ser. Lecture Notes in Computer Science. Springer, 2010, vol. 6321, pp. 56–71.

[62] M. A. Álvarez, L. Rosasco, and N. D. Lawrence, "Kernels for vector-valued functions: A review," *Foundations and Trends in Machine Learning*, vol. 4, no. 3, pp. 195–266, 2012.

[63] C. Jidling, N. Wahlström, A. Wills, and T. B. Schön, "Linearly constrained Gaussian processes," in *Advances in Neural Information Processing Systems (NIPS) 30*, 2017, pp. 1215–1224.
[64] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.
[65] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2004.
[66] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*. Springer, 2005.
[67] S. Särkkä, "Linear operators and stochastic partial differential equations in Gaussian process regression," in *Artificial Neural Networks and Machine Learning – ICANN 2011*, ser. Lecture Notes in Computer Science, vol. 6792. Springer, 2011, pp. 151–158.

Simo Särkkä received his M.Sc. (Tech.) degree (with distinction) in Engineering Physics and Mathematics, and D.Sc. (Tech.) degree (with distinction) in electrical and communications engineering from Helsinki University of Technology, Espoo, Finland, in 2000 and 2006, respectively.

Dr. Särkkä is an Associate Professor and Academy Research Fellow with Aalto University, Technical Advisor and Director of IndoorAtlas Ltd., and an Adjunct Professor with Tampere University of Technology and Lappeenranta University of Technology. His research interests are in multi-sensor data processing systems with applications in location sensing, health technology, machine learning, inverse problems, and brain imaging. He is a Senior Member of the IEEE.

Arno Solin received his M.Sc. (Tech.) degree (with distinction) in Engineering Physics and Mathematics and D.Sc. (Tech.) degree (with distinction) in computational science from Aalto University, Finland, in 2012 and 2016, respectively.

Dr. Solin is an Academy of Finland Postdoctoral Researcher with Aalto University, and Technical Advisor of IndoorAtlas Ltd. His research interests are in probabilistic inference for temporal, spatial, and spatio-temporal models with applications in sensor fusion for tracking and navigation, brain imaging, and machine learning problems.

Manon Kok received M.Sc. degrees in Applied Physics and in Philosophy of Science, Technology and Society (2009 and 2007, respectively), both from the University of Twente. From 2009 until 2011 she worked as a Research Engineer at Xsens Technologies. She received the PhD degree in Automatic Control from Linköping University in 2017.

Dr. Kok is currently a Postdoc in the Machine Learning Group of the Computational and Biological Learning Lab at the University of Cambridge. Her research interests are in the fields of probabilistic inference for sensor fusion, signal processing and machine learning.

Niklas Wahlström received the M.Sc. degree in Applied Physics and Electrical Engineering in 2010 and Ph.D. degree in automatic control in 2015, both from Linköping University, Sweden.

Dr. Wahlström is since 2016 a Researcher at the Department of Information Technology, Uppsala University. His research interests include machine learning, deep learning, sensor fusion, localization and mapping, especially using magnetic sensors.

Thomas B. Schön received the M.Sc. degree in Applied Physics and Electrical Engineering in 2001, the B.Sc. degree in Business Administration 2001 and the Ph.D. degree in Automatic Control in 2006, all from Linköping University.

Dr. Schön is Professor of the Chair of Automatic Control in the Department of Information Technology at Uppsala University. His main research interest is nonlinear inference problems, especially within the context of dynamical systems, solved using probabilistic methods, more specifically sequential Monte Carlo, particle MCMC and graphical models. He is active within the fields of machine learning, signal processing and automatic control. He is a Senior Member of the IEEE.