MODELING MAGNETIC FIELDS USING GAUSSIAN PROCESSES

Niklas Wahlström, Manon Kok, Thomas B. Schön, Fredrik Gustafsson

Division of Automatic Control, Linköping University, Linköping, Sweden
E-Mail: {nikwa,manko,schon,fredrik}@isy.liu.se

ABSTRACT

Starting from the electromagnetic theory, we derive a Bayesian nonparametric model allowing for joint estimation of the magnetic field and the magnetic sources in complex environments. The model is a Gaussian process which exploits the divergence- and curl-free properties of the magnetic field by combining well-known model components in a novel manner. The model is estimated using magnetometer measurements and spatial information implicitly provided by the sensor. The model and the associated estimator are validated on both simulated and real world experimental data producing Bayesian nonparametric maps of magnetized objects.

***Index Terms*—** magnetic field, Gaussian processes, Maxwell's equations, divergence-free, curl-free

(a) Estimated shape of table (b) Real shape of table

Fig. 1: Estimated magnetic content in a table turned upside down.

1. INTRODUCTION

The magnetic field has for a long time been used in navigation for providing seafarers and merchants as well as orienteers and migrating birds with heading information. In indoor environments this navigation task is challenged by magnetic distortions caused by the ferromagnetic structure in buildings. However, these distortions can also provide position information using a magnetic map of the environment, either by navigating within a precomputed map or by performing simultaneous localization and mapping (SLAM). This requires good models of the magnetic field which will be investigated more deeply in this work by addressing the electromagnetic theory. The relation between the magnetic sources and their induced magnetic field is well understood and was already formulated by Maxwell [1, 2]. However, little work has been done incorporating this knowledge into a statistical framework suitable for estimating magnetic fields in complex magnetic environments based on noisy data. We present a Bayesian nonparametric model (a particular Gaussian process) capable of modeling the magnetic field as well as the magnetic sources, see Figure 1. Our model exploits the divergence- and curl-free properties of the magnetic field inherited by the electromagnetic theory.

Gaussian processes [3] have previously been used for modeling magnetic fields in an indoor environment to enable SLAM [4, 5]. Their navigation platform is equipped with a three-axis magnetometer and the positioning is aided with odometry. However, in contrast to our work, the model of the magnetic map does not incorporate knowledge from Maxwell's equations and is not able to estimate the location of the magnetic sources. The same model has been investigated further in [6]. Gaussian processes have recently also been used for SLAM in a scalar potential field [7]. However, we consider multiple vector fields rather than one scalar potential field in this work. Also in [8] the magnetic disturbances are used to improve IMU-based position estimation. That work uses the restrictions induced by the electromagnetic theory, but does not construct any magnetic map and does not localize the magnetic sources. Chung et al. [9] uses four tilt-compensated magnetometers to accomplish indoor localization. The magnetic map is captured in advance consisting of a collection of magnetic signatures. The localization is performed using magnetic map fingerprints, where the performance is enhanced by the multiplicity of the magnetometers. Also fusion of magnetic field anomalies and laser has been investigated [10].

This work only addresses the modeling aspects of the magnetic field and the magnetic sources. The localization problem we consider separately [11], but the model is also suitable for being used in the applications presented above. The contributions of this work are:

- A Gaussian process model which in a novel manner exploits the divergence- and curl-free properties of the magnetic field.
- The model enables the magnetic field and the magnetic sources to be estimated jointly.
- Interference with both magnetic field measurements and spatial information is possible.

The divergence- and curl-free properties of a vector field have previously been used for estimating fluid flows using Gaussian processes [12]. However, to the best authors' knowledge this has previously not been used in modeling magnetic fields.

2. MAGNETIC FIELDS

A magnetic field is a mathematical construction used for describing forces induced by magnetic materials and electric currents. For each point in space the magnetic field can be described using a vector and as such it is a vector field. There are two different, but closely related ways to describe the magnetic field, denoted with the symbols $\mathbf{B}$ and $\mathbf{H}$, where boldface denotes vector-valued quantities.

This work is supported by the Swedish Foundation for Strategic Research under the project Cooperative Localization and the Linneaus Center CADICS funded by the Swedish Research Council. High accuracy reference measurements are provided through the use of the Vicon real-time tracking system courtesy of the UAS Technologies Lab, Artificial Intelligence and Integrated Computer Systems Division (AIICS) at the Department of Computer and Information Science (IDA), Linköping University, Sweden. http://www.ida.liu.se/divisions/aiics/aiicssite/index.en.shtml

These fields can not be any two arbitrary vector fields, but need to obey physical laws, which in their most general form are described by Maxwell's equations [2]. By assuming absence of free currents and time-dependent effects, these equations will reduce to

$$\nabla \cdot \mathbf{B} = 0, \tag{1a}$$

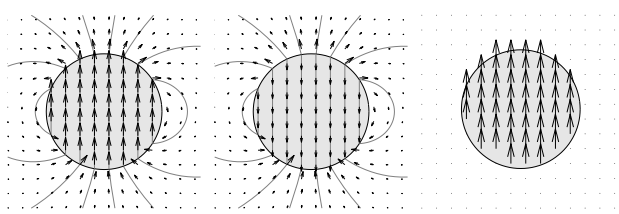
$$\nabla \times \mathbf{H} = \mathbf{0}, \tag{1b}$$

which means that the $\mathbf{B}$-field is divergence-free and the $\mathbf{H}$-field is curl-free. Further, these two fields are coupled as

$$\mathbf{M} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{H}, \tag{2}$$

where $\mathbf{M}$ is the magnetization describing our magnetic environment and μ_0 is the vacuum permeability, which is a physical constant having the value $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{V\,s\,A^{-1}\,m^{-1}}$. These fields will be illustrated with the following example. More details on the derivation can be found in [2].

Example 2.1 (Uniformly magnetized sphere) *Consider a sphere with a uniform permanent magnetization as depicted in Figure 2c. By solving* (1) *and* (2) *for this special geometry we will end up in a dipole field outside the sphere as depicted in Figure 2a and 2b. Note that the* $\mathbf{B}$*- and the* $\mathbf{H}$*-field will be identical (up to the proportional constant* μ_0*) outside the sphere, which follows directly from* (2) *using* $\mathbf{M} = \mathbf{0}$*. However, inside the sphere the* $\mathbf{B}$*- and the* $\mathbf{H}$*-field will be aligned in opposite directions in order to ensure that the* $\mathbf{B}$*-field is divergence-free (no sources or sinks) and that the* $\mathbf{H}$*-field is curl-free (no swirls).*



(a) $\mathbf{B}$-field (b) $\mathbf{H}$-field (c) $\mathbf{M}$-field

Fig. 2: The $\mathbf{B}$-, $\mathbf{H}$- and $\mathbf{M}$-field of a uniformly magnetized sphere. The $\mathbf{B}$-field is here normalized with μ_0.

By using (1)-(2) and prior knowledge of the magnetic environment, a number of things can be concluded, which will be used in Section 4 when modeling the magnetic fields:

1. **Additional information** In all non-magnetic materials the magnetization is equal to zero, $\mathbf{M} = \mathbf{0}$. This is especially true in locations where we measure the magnetic field, since air is non-magnetic. Due to physical constraints the sensor cannot be inside a magnetic material and we know that $\mathbf{M} = \mathbf{0}$ at these positions. This additional information will be used in our framework as an extra measurement. Sensor fusion with other sensors such as camera and laser range sensor, providing even richer information of where $\mathbf{M} = \mathbf{0}$, is possible. However, this is not considered in this work.

2. **External field** Most environments of interest consist of an external homogeneous field $\mathbf{B}_0$ or $\mathbf{H}_0$, usually the earth magnetic field or a slight deformation of it. Due to the linearity of the field equations (1), this external field can be superimposed throughout all space, where (2) gives the relation $\mathbf{B}_0 = \mu_0\mathbf{H}_0$. We will therefore later model the $\mathbf{B}$- and the $\mathbf{H}$-field to have a common, but unknown mean.

3. **Smoothness** If $\mathbf{M}(\mathbf{x}) = \mathbf{0}$ in a neighborhood of $\mathbf{x}$, the field equations (1) will ensure that $\mathbf{B}$ and $\mathbf{H}$ are infinitely continuously differentiable at $\mathbf{x}$. This gives the magnetic field a "smooth" character and the magnetic field at $\mathbf{x}_1$ will be very similar to the magnetic field at $\mathbf{x}_2$ if $\mathbf{x}_1$ and $\mathbf{x}_2$ are close. This property motivates us to employ Gaussian processes in modeling these fields, as will be explained in the next section.

3. GAUSSIAN PROCESSES

A Gaussian process (GP) [3] is a stochastic process suitable for modeling spatially correlated measurements. GPs can be seen as a distribution over functions

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\big(\boldsymbol{\mu}(\mathbf{x}), \mathrm{K}(\mathbf{x}, \mathbf{x}')\big), \tag{3}$$

where the process is uniquely defined with its mean function $\boldsymbol{\mu}(\mathbf{x})$ and covariance function $\mathrm{K}(\mathbf{x}, \mathbf{x}')$.

The GP is a generalization of the multivariate Gaussian probability distribution in the sense that the function values evaluated for a finite number of inputs $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are normally distributed

$$\begin{bmatrix} \mathbf{f}(\mathbf{x}_1) \\ \vdots \\ \mathbf{f}(\mathbf{x}_N) \end{bmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \mathrm{K}), \quad \text{where} \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}(\mathbf{x}_1) \\ \vdots \\ \boldsymbol{\mu}(\mathbf{x}_N) \end{bmatrix} \tag{4a}$$

$$\text{and} \quad \mathrm{K} = \begin{bmatrix} \mathrm{K}(\mathbf{x}_1, \mathbf{x}_1) & \cdots & \mathrm{K}(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & & \vdots \\ \mathrm{K}(\mathbf{x}_N, \mathbf{x}_1) & \cdots & \mathrm{K}(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix}. \tag{4b}$$

3.1. Mean function

In this work we will consider a constant, but unknown mean function $\boldsymbol{\mu}(\mathbf{x}) = \boldsymbol{\beta}$, where we put a Gaussian prior on the mean

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\big(\boldsymbol{\beta}, \mathrm{K}(\mathbf{x}, \mathbf{x}')\big), \quad \text{where} \quad \boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \sigma_\beta^2 \mathrm{I}). \tag{5a}$$

By integrating out the parameter $\boldsymbol{\beta}$, this can be reformulated as a zero mean GP

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\big(\mathbf{0}, \mathrm{K}(\mathbf{x}, \mathbf{x}') + \sigma_\beta^2 \mathrm{I}\big). \tag{5b}$$

3.2. Vector-valued covariance functions

The covariance function (a.k.a. kernel) is the crucial component when modeling using a GP. This function encodes the assumptions we make on the functions to be learned. For modeling smooth functions (as desired in Item 3 in Section 2) with scalar output the most common choice is the squared exponential (SE) covariance function

$$\mathrm{K}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 e^{-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2l^2}}, \tag{6}$$

where σ_f is the expected amplitude and l the expected length-scale of the function we want to learn. This covariance function can be extended for learning functions with multiple outputs as presented below. Learning functions with multiple outputs has recently attracted more attention. A review can be found in [13], which discusses different kernels for learning vector-valued functions.

3.2.1. Diagonal squared exponential covariance function

The most obvious extension of (6) to multiple outputs is to model each component $f_i(\mathbf{x})$ separately using a scalar SE covariance functions resulting in a diagonal SE kernel

$$\mathrm{K}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 e^{-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2l^2}} \cdot \mathrm{I}_{n_\mathbf{y}}, \tag{7}$$

where $n_{\mathbf{y}}$ is the dimension of the output. The kernel (7) can be extended to have different hyperparameters l and σ_f for each output dimension. This kernel was used by [4] and [5] in modeling the magnetic field of an indoor environment. However, this kernel does not allow for the possibility of modeling correlations between the different components $f_i(\mathbf{x})$. Especially, it does not produce functions which necessarily obey the field equations (1). This is made possible by the two covariance functions introduced below.

3.2.2. Divergence- and curl-free covariance functions

A kernel for learning divergence-free vector fields was first introduced by [14]. Based on the scalar SE kernel (6), this kernel reads

$$\mathrm{K}_{\mathbf{B}}(\mathbf{x},\mathbf{x}') = \sigma_f^2 e^{-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2l^2}} \cdot \left(\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)^{\mathsf{T}} + \left(n_{\mathbf{y}} - 1 - \frac{\|\mathbf{x}-\mathbf{x}'\|^2}{l^2}\right)\mathrm{I}_{n_{\mathbf{y}}} \right), \tag{8}$$

which ensures that all functions sampled from a GP with such a kernel will be divergence-free. Similarely, [15] introduced a kernel for learning curl-free vector fields, where the extension of (6) reads

$$\mathrm{K}_{\mathbf{H}}(\mathbf{x},\mathbf{x}') = \sigma_f^2 e^{-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2l^2}} \left(\mathrm{I}_{n_{\mathbf{y}}} - \left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)\left(\frac{\mathbf{x}-\mathbf{x}'}{l}\right)^{\mathsf{T}} \right). \tag{9}$$

The interested reader can refer to [15, 12, 16] for more analysis and discussion on these two kernels.

3.3. Regression

GPs are also capable of handling noisy measurements $\mathbf{y}_k$ of the function $\mathbf{f}(\mathbf{x}_k)$. We consider the measurement model

$$\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k) + \mathbf{e}_k, \qquad \mathbf{e}_k \sim \mathcal{N}(\mathbf{0}, \Sigma), \tag{10}$$

where $\mathbf{e}_k$ has the interpretation of being measurement noise. Our objective is to use a set of measurements together with their corresponding inputs $\{\mathbf{x}_k, \mathbf{y}_k | k = 1, \dots, N\}$ to learn the function values for other test inputs $\mathbf{f}_* = \begin{bmatrix}\mathbf{f}(\mathbf{x}_1^*)^{\mathsf{T}} & \dots & \mathbf{f}(\mathbf{x}_{N_*}^*)^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}$. In the same manner as in (4), the joint distribution for the measurements $\mathbf{y} = \begin{bmatrix}\mathbf{y}_1^{\mathsf{T}} & \dots & \mathbf{y}_N^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}$ and the test output $\mathbf{f}_*$ is

$$\begin{bmatrix}\mathbf{y}\\ \mathbf{f}_*\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\mathrm{K}(\mathrm{X},\mathrm{X}) + \mathrm{I}_N \otimes \Sigma & \mathrm{K}(\mathrm{X},\mathrm{X}_*) \\ \mathrm{K}(\mathrm{X}_*,\mathrm{X}) & \mathrm{K}(\mathrm{X}_*,\mathrm{X}_*)\end{bmatrix}\right), \tag{11}$$

where $\otimes$ denote the Kronecker product,

$$\mathrm{K}(\mathrm{X},\mathrm{X}_*) = \begin{bmatrix}\mathrm{K}(\mathbf{x}_1,\mathbf{x}_1^*) & \dots & \mathrm{K}(\mathbf{x}_1,\mathbf{x}_{N_*}^*) \\ \vdots & & \vdots \\ \mathrm{K}(\mathbf{x}_N,\mathbf{x}_1^*) & \dots & \mathrm{K}(\mathbf{x}_N,\mathbf{x}_{N_*}^*)\end{bmatrix} \tag{12}$$

and similarly for the other matrices $\mathrm{K}(\mathrm{X},\mathrm{X})$, $\mathrm{K}(\mathrm{X}_*,\mathrm{X})$ and $\mathrm{K}(\mathrm{X}_*,\mathrm{X}_*)$. From the joint Gaussian distribution $p(\mathbf{y},\mathbf{f}_*)$ in (11) the conditional distribution $p(\mathbf{f}_*|\mathbf{y})$ can easily be computed as

$$\mathbf{f}_*|\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{f}_*}, \Sigma_{\mathbf{f}_*}), \tag{13a}$$

$$\boldsymbol{\mu}_{\mathbf{f}_*} = \mathrm{K}_*^{\mathsf{T}}\mathrm{K}_{\mathbf{y}}^{-1}\mathbf{y}, \qquad \Sigma_{\mathbf{f}_*} = \mathrm{K}_{**} - \mathrm{K}_*^{\mathsf{T}}\mathrm{K}_{\mathbf{y}}^{-1}\mathrm{K}_*, \tag{13b}$$

where $\mathrm{K} = \mathrm{K}(\mathrm{X},\mathrm{X})$, $\mathrm{K}_* = \mathrm{K}(\mathrm{X},\mathrm{X}_*)$, $\mathrm{K}_{**} = \mathrm{K}(\mathrm{X}_*,\mathrm{X}_*)$ and $\mathrm{K}_{\mathbf{y}} = \mathrm{K}(\mathrm{X},\mathrm{X}) + \mathrm{I}_N \otimes \Sigma$.

3.4. Estimating hyperparameters

The hyperparameters of the covariance function $\mathrm{K}(\mathbf{x},\mathbf{x}')$ and the measurement noise covariance matrix Σ can be estimated from the data $\{\mathbf{x}_k, \mathbf{y}_k | k = 1, \dots N\}$, which makes the learning of the function values $\mathbf{f}_*$ completely data driven in the sense that no tuning parameters are needed. This will be accomplished by maximizing the log marginal likelihood $\log p(\mathbf{y}, |\mathrm{X}, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denote the hyperparameters of $\mathrm{K}(\mathbf{x},\mathbf{x}')$ and Σ. From (11) we have that $\mathbf{y}|\mathrm{X},\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \mathrm{K}_{\mathbf{y}})$, which gives

$$\log p(\mathbf{y}|\mathrm{X},\boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^{\mathsf{T}}\mathrm{K}_{\mathbf{y}}^{-1}\mathbf{y} - \frac{1}{2}\log|\mathrm{K}_{\mathbf{y}}| - \frac{n_{\mathbf{y}}N}{2}\log 2\pi. \tag{14}$$

Following [3], the gradient of the log marginal likelihood w.r.t. the hyperparameters can be computed as

$$\frac{\partial}{\partial \boldsymbol{\theta}_j}\log p(\mathbf{y}|\mathrm{X},\boldsymbol{\theta}) = \frac{1}{2}\operatorname{tr}\left((\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathsf{T}} - \mathrm{K}_{\mathbf{y}}^{-1})\frac{\partial \mathrm{K}_{\mathbf{y}}}{\partial \boldsymbol{\theta}_j}\right), \tag{15}$$

where $\boldsymbol{\alpha} = \mathrm{K}_{\mathbf{y}}^{-1}\mathbf{y}$. This enables an efficient gradient based optimizing routine for maximizing (14). In this work the BFGS method [17] has been used.

4. MODELING

The GP framework will now be combined with the electromagnetic theory to construct a model for jointly estimating the $\mathbf{B}$- and the $\mathbf{M}$-field using a three-axis magnetometer. We assume that the measurements of the magnetic field are corrupted with Gaussian noise

$$\mathbf{y}_{\mathbf{B},k} = \mathbf{f}_{\mathbf{B}}(\mathbf{x}_k) + \mathbf{e}_{\mathbf{B},k}, \qquad \mathbf{e}_{\mathbf{B},k} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathrm{I}_3), \tag{16a}$$

where $\mathbf{y}_{\mathbf{B},k}$ is a three-axis magnetometer measurement transformed into world coordinates and $\mathbf{f}_{\mathbf{B}}(\mathbf{x}_k)$ is a function being equal to the $\mathbf{B}/\mu_0$-field (the $\mathbf{B}$-field normalized with μ_0) at location $\mathbf{x}_k$. As discussed in Item 1 in Section 2 we also know that the $\mathbf{M}$-field is zero at location $\mathbf{x}_k$, where the measurement $\mathbf{y}_{\mathbf{B},k}$ was collected. This information is incorporated by considering a noise free measurement $\mathbf{y}_{\mathbf{M},k} = \mathbf{0}$ with the following measurement equation

$$\mathbf{y}_{\mathbf{M},k} = \mathbf{f}_{\mathbf{M}}(\mathbf{x}_k) = \mathbf{f}_{\mathbf{B}}(\mathbf{x}_k) - \mathbf{f}_{\mathbf{H}}(\mathbf{x}_k), \tag{16b}$$

where $\mathbf{f}_{\mathbf{M}}(\mathbf{x}_k)$ and $\mathbf{f}_{\mathbf{H}}(\mathbf{x}_k)$ are functions corresponding to the $\mathbf{M}$- and the $\mathbf{H}$-field and where we have used the coupling given by (2). Note that this coupling is the key equation for our model since it enables us to jointly estimate the $\mathbf{B}$-field as well as the $\mathbf{M}$-field in contrast to prior work.

We put this into a statistical framework by considering $\mathbf{f}_{\mathbf{B}}$ and $\mathbf{f}_{\mathbf{H}}$ (and consequently also $\mathbf{f}_{\mathbf{M}}$ via (16b)) to be GPs. Following the discussion in Item 2 in Section 2 we consider $\mathbf{f}_{\mathbf{B}}$ and $\mathbf{f}_{\mathbf{H}}$ to have a common constant mean function (corresponding to the earth magnetic field) and we use the covariance functions given in (8) and (9) to preserve the divergence- and curl-free properties of $\mathbf{f}_{\mathbf{B}}$ and $\mathbf{f}_{\mathbf{H}}$ according to the field equations (1). This gives

$$\mathbf{f}_{\mathbf{B}} \sim \mathcal{GP}\big(\boldsymbol{\beta}, \mathrm{K}_{\mathbf{B}}(\mathbf{x},\mathbf{x}')\big), \quad \mathbf{f}_{\mathbf{H}} \sim \mathcal{GP}\big(\boldsymbol{\beta}, \mathrm{K}_{\mathbf{H}}(\mathbf{x},\mathbf{x}')\big), \tag{16c}$$

$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \sigma_\beta^2\mathrm{I}_3), \tag{16d}$$

where we have used a Gaussian prior on the unknown mean $\boldsymbol{\beta}$.

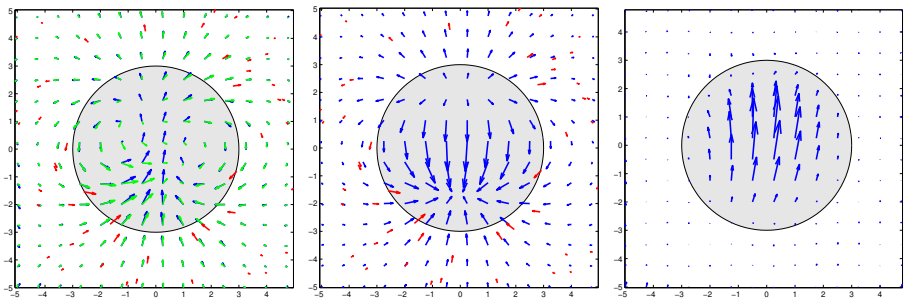
The model (16) can be reformulated into the standard model description outlined in Section 3

$$\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k) + \mathbf{e}_k, \tag{17a}$$

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}\big(\mathbf{0}, \mathrm{K}(\mathbf{x},\mathbf{x}')\big), \qquad \mathbf{e}_k \sim \mathcal{N}(\mathbf{0}, \Sigma), \tag{17b}$$

by augmenting the measurements and the noise covariance matrices

$$\mathbf{y}_k = \begin{bmatrix}\mathbf{y}_{\mathbf{B},k}\\ \mathbf{y}_{\mathbf{M},k}\end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix}\sigma_n^2\mathrm{I}_3 & 0\\ 0 & 0\end{bmatrix} \tag{17c}$$



(a) Estimated $\mathbf{B}$-field (b) Estimated $\mathbf{H}$-field (c) Estimated $\mathbf{M}$-field

Fig. 3: Estimated fields induced by a uniformly magnetized sphere (see Example 2.1) using our proposed kernel (17) (blue) and the SE kernel (7) (green) together with the training data (red).

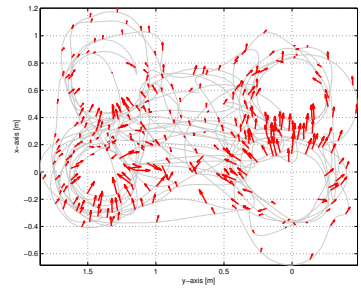


Fig. 4: The training data in the real world experiment seen from above together with the trajectory that the magnetometer has followed.

as well as the outputs of the functions that we want to learn

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \mathbf{f_B}(\mathbf{x}) \\ \mathbf{f_M}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathrm{I}_3 & 0_3 \\ \mathrm{I}_3 & -\mathrm{I}_3 \end{bmatrix} \begin{bmatrix} \mathbf{f_B}(\mathbf{x}) \\ \mathbf{f_H}(\mathbf{x}) \end{bmatrix} \sim \mathcal{GP}(\mathbf{0}, \mathrm{K}), \quad (17d)$$

where $\mathrm{K} = \mathrm{K}(\mathbf{x}, \mathbf{x}')$. The augmented function $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^6$ will then have the covariance function

$$\mathrm{K} = \begin{bmatrix} \mathrm{K_B} + \sigma_\beta^2 \mathrm{I}_3 & \mathrm{K_B} \\ \mathrm{K_B} & \mathrm{K_B} + \mathrm{K_H} \end{bmatrix}, \quad (17e)$$

where the relation $\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, \mathrm{K}) \Rightarrow \mathrm{C}\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, \mathrm{CKC}^\mathsf{T})$ has been used as well as (5) to reformulate this as a zero mean GP. Finally, we encode $\boldsymbol{\theta} \triangleq \begin{bmatrix} \log \sigma_f^2 & \log l^2 & \log \sigma_\beta^2 & \log \sigma_n^2 \end{bmatrix}$, where the logarithm ensures the positiveness of σ_f^2, l^2, σ_β^2 and σ_n^2.

5. RESULTS

The ability of the proposed model to model magnetic fields will be evaluated by using a simulated data set as well as a real world data set. The results will be reported in this section.

5.1. Simulated experiment

The setup with a uniformly magnetized sphere presented in Example 2.1 is used to estimate the $\mathbf{B}$-, $\mathbf{H}$- and $\mathbf{M}$-field given in Figure 2. Consider a sphere centered at the origin with radius 3 m having a uniform magnetization of $\mathbf{M} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^\mathsf{T} \mathrm{A\,m^{-1}}$. In total $N = 50$ training inputs are chosen from a region outside the sphere and inside a square with dimension $10\,\mathrm{m} \times 10\,\mathrm{m}$ aligned with the xy-plane, which also is centered at the origin. For each training input the corresponding training output is computed using the true field perturbed with Gaussian noise having a standard deviation of $\sigma_n = 0.01$. The test inputs are chosen from a grid xy-plane with an interval of $0.75\,\mathrm{m}$. The estimated magnetic fields at these test inputs is then compared with the true magnetic fields. Both the SE kernel (7) and the proposed kernel (17) are applied to the data, where the hyperparameters for each kernel are estimated as described in Section 3.4. The results are given in Figure 3.

Both the SE kernel and the proposed kernel (17) are able to reproduce the character of the true $\mathbf{B}$-field as given in Figure 2a. By comparing the estimated $\mathbf{B}$-field with the true $\mathbf{B}$-field, the proposed covariance function is only slightly better with a root mean square error of $0.33\,\mathrm{A\,m^{-1}}$, whereas the corresponding number for the SE covariance function is $0.38\,\mathrm{A\,m^{-1}}$. However, the great advantage with the proposed covariance function is its ability to estimate the $\mathbf{M}$-field as shown in Figure 3c, which resembles the true $\mathbf{M}$-field in Figure 2c. Both the location of the magnetic source and the direction of its magnetization are correctly captured. By using the relation (2), also the $\mathbf{H}$-field can be estimated as presented in Figure 3b, which can be compared with the true $\mathbf{H}$-field in Figure 2b.

5.2. Real world experiment

A real world experiment was conducted in a magnetic environment consisting of a table with metal frame turned upside down as displayed in Figure 1b. A three-axis magnetometer has been used to measure the magnetic field at various locations around that table and the position and the orientation of the magnetometer unit were measured using an optical reference system (Vicon). The magnetometer measurements were then transformed into world coordinates using the orientation provided by the reference. This data has been downsampled to 2 Hz to reduce the number of data points. Together with the position estimate from the reference this comprises the training data as displayed in Figure 4. For this dataset the hyperparameters have not been estimated but rather tuned to $\sigma_f = 0.3$, $l = 0.15$, $\sigma_b = 1$ and $\sigma_n = 0.3$ for reasons discussed below. In Figure 1a the region of the $\mathbf{M}$-field which exceeds 30% of the maximal estimated $\mathbf{M}$-field is displayed. This estimated magnetic map has visual similarities with the real table in Figure 1b. All four table legs can be distinguished as well as the frame on which the table top is attached.

The proposed GP (as any other stationary GP) is restricted to using the same set of hyperparameters for all data. This is problematic in environments which have different characteristic length scales and signal amplitudes in different regions in space. When estimating the hyperparameters in the proposed manner using data collected in such environments, the result might not be sound. The hyperparameters have therefore been considered as tuning parameters.

6. CONCLUSION AND FUTURE WORK

We have introduced a Bayesian nonparametric model for jointly estimating both the magnetic field and the magnetic sources. The model is based on a vector-valued stationary Gaussian process (GP) with a covariance function exploiting the divergence- and curl-free properties of the magnetic field derived from the electromagnetic theory. The model has been compared with a component-wise GP proposed by [4] for modeling magnetic fields. In the comparison only a small improvement in estimation performance could be reported. However, the great advantage of the proposed method is its ability to also model the magnetic sources in a nonparametric manner, which has been illustrated using both simulated and real world data.

In future work we will extend our nonparametric model to handle more complex environments. One promising idea is to use a multiplicity of GPs governed by a hierarchical Dirichlet process [18]. Others ideas are to use a new concept of multiresolution Gaussian processes [19] or an infinite mixture of Gaussian process experts [20]. Our final goal is a full SLAM framework.

7. REFERENCES

[1] J.C. Maxwell, "A dynamical theory of the electromagnetic field," *Philosophical Transactions of the Royal Society of London*, vol. 155, pp. 459–512, 1865.

[2] J.D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, Inc, 3rd edition, 1998.

[3] C.E. Rasmussen and C.K.I. Williams, *Gaussian processes for machine learning*, MIT press Cambridge, MA, 2006.

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

[5] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Roning, "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)*, Tallin, Estonia, June 2011, IEEE, pp. 198–203.

[6] 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 European Conference on Mobile Robots*, Örebro, Sweden, September 2011, pp. 283–290.

[7] J. Murphy and S. Godsill, "Simultaneous localization and mapping for non-parametric potential field environments," in *Proceedings of the 15th International Conference on Information Fusion*, Singapore, July 2012, ISIF.

[8] D. Vissière, A.P. Martin, and N. Petit, "Using magnetic disturbances to improve IMU-based position estimation," in *Proceedings of the European Control Conference*, Kos, Greece, July 2007, pp. 2853–2858.

[9] J. Chung, M. Donahoe, C. Schmandt, I.J. Kim, P. Razavai, and M. Wiseman, "Indoor location sensing using geo-magnetism," in *Proceedings of the 9th International Conference on Mobile systems, Applications, and Services*, Bethesda, USA, June 2011, ACM, pp. 141–154.

[10] H. Zhang and F. Martin, "Robotic mapping assisted by local magnetic field anomalies," in *Proceedings of the IEEE Conference on Technologies for Practical Robot Applications (TePRA)*, Woburn, USA, April 2011, pp. 25–30.

[11] M. Kok, N. Wahlström, T.B. Schön, and F. Gustafsson, "MEMS-based inertial navigation based on a magnetic field map," in *Proceedings of the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013.

[12] I. Macêdo and R. Castro, "Learning divergence-free and curl-free vector fields with matrix-valued kernels," Tech. Rep., Technical report, Instituto Nacional de Matematica Pura e Aplicada, 2008.

[13] M. A. Álvarez, L. Rosasco, and N.D. Lawrence, "Kernels for vector-valued functions: A review," *Found. Trends Mach. Learn.*, vol. 4, no. 3, pp. 195–266, March 2012.

[14] F.J. Narcowich and J.D. Ward, "Generalized hermite interpolation via matrix-valued conditionally positive definite functions," *Mathematics of Computation*, vol. 63, no. 208, pp. 661–688, 1994.

[15] E.J. Fuselier Jr, *Refined error estimates for matrix-valued radial basis functions*, Ph.D. thesis, Texas A&M University, 2006.

[16] L. Baldassarre, L. Rosasco, A. Barla, and A. Verri, "Vector field learning via spectral filtering," *Machine Learning and Knowledge Discovery in Databases*, pp. 56–71, 2010.

[17] J. Nocedal and S.J. Wright, *Numerical optimization*, Springer verlag, 1999.

[18] Y.W. Teh, M.I. Jordan, M.J. Beal, and D.M. Blei, "Hierarchical Dirichlet processes," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1566–1581, 2006.

[19] E. Fox and D.B. Dunson, "Multiresolution Gaussian processes," in *Advances in Neural Information Processing Systems 25*, pp. 746–754. 2012.

[20] C.E. Rasmussen and Z. Ghahramani, "Infinite mixtures of Gaussian process experts," *Advances in neural information processing systems*, vol. 2, pp. 881–888, 2002.