

## Large-scale magnetic field maps using structured kernel interpolation for Gaussian process regression

Menzen, Clara; Fetter, Marnix; Kok, Manon

**DOI**

[10.23919/FUSION52260.2023.10224210](https://doi.org/10.23919/FUSION52260.2023.10224210)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

Proceedings of the 26th International Conference on Information Fusion, FUSION 2023

**Citation (APA)**

Menzen, C., Fetter, M., & Kok, M. (2023). Large-scale magnetic field maps using structured kernel interpolation for Gaussian process regression. In *Proceedings of the 26th International Conference on Information Fusion, FUSION 2023* IEEE. <https://doi.org/10.23919/FUSION52260.2023.10224210>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

***Green Open Access added to TU Delft Institutional Repository***

***'You share, we take care!' - Taverne project***

**<https://www.openaccess.nl/en/you-share-we-take-care>**

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

# Large-scale magnetic field maps using structured kernel interpolation for Gaussian process regression

Clara Menzen, Marnix Fetter and Manon Kok

Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands

Email: c.m.menzen@tudelft.nl, marnixfetter@gmail.com, m.kok-1@tudelft.nl

**Abstract**—We present a mapping algorithm to compute large-scale magnetic field maps in indoor environments with approximate Gaussian process (GP) regression. Mapping the spatial variations in the ambient magnetic field can be used for localization algorithms in indoor areas. To compute such a map, GP regression is a suitable tool because it provides predictions of the magnetic field at new locations along with uncertainty quantification. Because full GP regression has a complexity that grows cubically with the number of data points, approximations for GPs have been extensively studied. In this paper, we build on the structured kernel interpolation (SKI) framework, speeding up inference by exploiting efficient Krylov subspace methods. More specifically, we incorporate SKI with derivatives (D-SKI) into the scalar potential model for magnetic field modeling and compute both predictive mean and covariance with a complexity that is linear in the data points. In our simulations, we show that our method achieves better accuracy than current state-of-the-art methods on magnetic field maps with a growing mapping area. In our large-scale experiments, we construct magnetic field maps from up to 40 000 three-dimensional magnetic field measurements in less than two minutes on a standard laptop.

**Index Terms**—Gaussian process regression, magnetic field maps, indoor localization, structured kernel interpolation.

## I. INTRODUCTION

Indoor positioning and navigation in indoor environments is an active and challenging field of research, see e.g. [1], [2]. Since the global positioning system (GPS) does not work properly indoors, existing technologies rely on e.g. WLAN [3] or ultra-wideband [4]. In recent years, a novel and promising approach uses the spatial anomalies of the ambient magnetic field that is present indoors, see e.g. [5]–[11]. Probabilistic algorithms for indoor localization with magnetic field measurements use e.g. an extended Kalman filter [12] or a particle filter [13] in combination with Gaussian process (GP) regression. The motivation to use GPs [14] is the fact that they can be used to construct a magnetic field map from measurements providing a mean and uncertainty information which are both crucial for probabilistic localization algorithms. However, full GP regression becomes intractable for a large number of data points  $N$ , so existing approaches for magnetic field mapping have downsampled the data [6], made maps only using data close to a position of interest [15] or have approximated the GP kernel in terms of a number of basis functions [16]. Each of these methods has downsides, which can impact the accuracy of the map: The first two methods do not use all the data and the latter relies on a sufficient number of basis functions  $M_{\text{bf}}$  to achieve a good approximation of the kernel

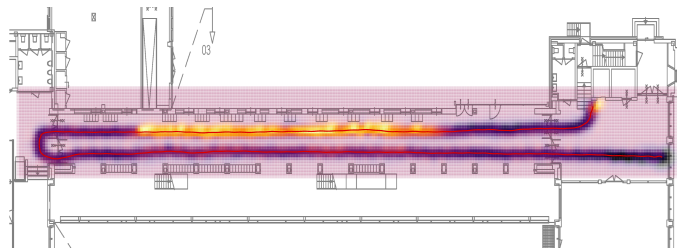


Fig. 1: Magnitude of the magnetic field computed by constructing a map from 21 931 measurements, where darker regions correspond to a higher magnitude. The red line is the walking path, along which measurements are collected.

function [16]. Inspired by the fact that the literature about approximate GPs offers numerous other approaches for large-scale GPs that overcome the aforementioned limitations, in this paper, we build on the SKI framework by [17] to construct magnetic field maps. In the SKI framework, the measurements are observed through  $M_{\text{ind}}$  inducing variables, where the inducing inputs are placed on a Cartesian grid. The structure of the inducing inputs naturally enables Kronecker structure in the corresponding kernel matrix, as well as structured kernel interpolation (SKI), i.e. approximation of kernel matrices by interpolation. Exploiting Kronecker algebra and the sparsity of the interpolation matrices in Krylov subspace methods, we can compute magnetic field maps in an efficient way. This allows us to compute large-scale magnetic field maps as illustrated in Fig. 1 that are computationally unfeasible for full GP regression on a regular laptop. To construct the map, we use magnetometer data in combination with positions and orientations that are assumed to be known. Based on previous work by e.g. [18], [19], we model the magnetic field with the scalar potential model that allows for incorporating physical knowledge into the GP prior. In this framework, we use SKI with derivatives (D-SKI) [20] to compute the predictive means with conjugate gradients. For the predictive variance, we adapt the LanczOs Variance Estimates (LOVE) algorithm [21] to the D-SKI framework. The associated computational complexity is  $\mathcal{O}((J + 2T)(3N + M_{\text{ind}}(M_{\text{ind}}^{(1)} + M_{\text{ind}}^{(2)} + M_{\text{ind}}^{(3)})))$ , where  $J$  and  $T$  are chosen based on the desired accuracy of the conjugate gradient and Lanczos tridiagonalization algorithm, respectively, and  $M_{\text{ind}}^{(d)}$  is the number of inducing inputs in the  $d$ th dimension.

## II. PROBLEM FORMULATION

We are interested in constructing large-scale magnetic field maps with GP regression. Similar to [19], [22], we assume the magnetic field to be curl-free and model the magnetic field measurements as the derivatives of a scalar potential  $\varphi$  on which we put a GP prior. Given  $n = 1, \dots, N$  3D positions  $\mathbf{p}$  at which magnetic field measurements have been collected, the GP model is given by

$$\begin{aligned} \varphi(\mathbf{p}) &\sim \mathcal{GP}(0, \kappa(\mathbf{p}, \mathbf{p}')), \\ \mathbf{y}_n &= -\nabla\varphi(\mathbf{p}_n) + \boldsymbol{\epsilon}_n, \quad \boldsymbol{\epsilon}_n \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I}_3), \end{aligned} \quad (1)$$

where  $\mathbf{y}_n \in \mathbb{R}^3$  contains the  $x$ -,  $y$ - and  $z$ - component of the magnetic field measurement,  $\kappa$  is the kernel function, and  $\sigma_y^2$  is the noise variance. We choose the kernel to be the squared exponential kernel, as in related literature, see e.g. [6], [23]–[26]. The squared exponential kernel is given by

$$\kappa(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|_2}{2\ell}\right), \quad (2)$$

where  $\sigma_f^2$  and  $\ell$  are the hyperparameters of the kernel, the signal variance, and the length scale.

Although we measure the Earth's magnetic field together with anomalies, we choose to only model the anomalies, because this model choice fits better into our approximation scheme. When also considering the local magnetic field as e.g. in [27], the kernel includes a linear term as well.

As the gradient operator is a linear operator given the linearity of differentiation [28], the predictive distribution of the three components of the magnetic field in a new location  $\mathbf{p}_*$  can be expressed in terms of a mean and a variance of  $\mathbf{f}_*$  given by

$$\begin{aligned} \mathbb{E}[\mathbf{f}_*] &= \partial^2(\mathbf{K}_{*,\mathbf{f}}) (\partial^2(\mathbf{K}_{\mathbf{f},\mathbf{f}}) + \sigma_y^2 \mathbf{I}_{3N})^{-1} \text{vec}(\mathbf{Y}^\top), \\ \mathbb{V}[\mathbf{f}_*] &= \partial^2(\mathbf{K}_{*,*}) - \\ &\quad \partial^2(\mathbf{K}_{*,\mathbf{f}}) (\partial^2(\mathbf{K}_{\mathbf{f},\mathbf{f}}) + \sigma_y^2 \mathbf{I}_{3N})^{-1} \partial^2(\mathbf{K}_{\mathbf{f},*}), \end{aligned} \quad (3)$$

where  $\mathbf{Y} = [\mathbf{y}_1^\top \mathbf{y}_2^\top \dots \mathbf{y}_N^\top] \in \mathbb{R}^{N \times 3}$  are all magnetic field measurements. The entries of  $\partial^2(\mathbf{K}_{\mathbf{f},\mathbf{f}})$ ,  $\partial^2(\mathbf{K}_{*,\mathbf{f}})$  and  $\partial^2(\mathbf{K}_{*,*})$  are computed block-wise in terms of  $3 \times 3$  blocks for each pair of positions with  $\nabla_{\mathbf{p}} \kappa(\mathbf{p}, \mathbf{p}') \nabla_{\mathbf{p}'}^\top$ ,  $\nabla_{\mathbf{p}} \kappa(\mathbf{p}, \mathbf{p}_*) \nabla_{\mathbf{p}_*}^\top$  and  $\nabla_{\mathbf{p}_*} \kappa(\mathbf{p}_*, \mathbf{p}'_*) \nabla_{\mathbf{p}'_*}^\top$ , respectively, where  $\nabla$  denotes the gradient that is taken w.r.t. to the vector specified in the subscript.

With (3) it is possible to predict the magnetic field in new locations. In practice, however, this is only possible for a small number of data points, since full GP regression generally scales cubically with  $N$ . In this case it even scales cubically with  $3N$ , because of the 3 derivatives. In this paper, we build on the SKI framework, described in the next section, to make large-scale magnetic field maps in an efficient way.

## III. SKI FRAMEWORK

The SKI framework is based on sparse approximations for GP regression, using a set of  $M_{\text{ind}}$  inducing inputs  $\mathbf{x}_u \in \mathbb{R}^D$ . In the context of magnetic field modeling, the inducing inputs are positions in  $\mathbb{R}^3$ . Based on the Nyström approximation [29]

of the kernel, the simplest formulation of the inducing input approach is the subset of regressors (SoR), which can be implemented similarly to the predictive distribution for full GP regression using an approximation to the kernel function [30], given by

$$\kappa_{\text{SoR}}(\mathbf{x}, \mathbf{x}') = \kappa(\mathbf{x}, \mathbf{x}_u) \mathbf{K}_{u,u}^{-1} \kappa(\mathbf{x}_u, \mathbf{x}'), \quad (4)$$

where  $\mathbf{K}_{u,u}$  denotes the covariance matrix of all the inducing inputs. The approximated kernel function results in new kernel matrices, which are then given by

$$\mathbf{K}_{\mathbf{f},\mathbf{f}} = \mathbf{K}_{\mathbf{f},u} \mathbf{K}_{u,u}^{-1} \mathbf{K}_{u,\mathbf{f}}, \quad (5a)$$

$$\mathbf{K}_{*,*} = \mathbf{K}_{*,u} \mathbf{K}_{u,u}^{-1} \mathbf{K}_{u,*}, \quad (5b)$$

$$\mathbf{K}_{*,\mathbf{f}} = \mathbf{K}_{*,u} \mathbf{K}_{u,u}^{-1} \mathbf{K}_{u,\mathbf{f}}. \quad (5c)$$

In the SKI framework [17], the inducing inputs are placed on a Cartesian grid, which is equispaced per dimension and of size  $M_{\text{ind}}^{(1)} \times M_{\text{ind}}^{(2)} \times \dots \times M_{\text{ind}}^{(D)}$ , for a total of  $M_{\text{ind}} = \prod_{d=1}^D M_{\text{ind}}^{(d)}$  inducing inputs.

Consequently, product kernels - here the squared exponential kernel is considered - decompose over the input dimensions. Thus,  $\mathbf{K}_{u,u}$  can be expressed as a Kronecker product of  $D$  matrices [17], [31], given by

$$\mathbf{K}_{u,u} = \bigotimes_{d=1}^D \mathbf{K}_{u,u}^{(d)}, \quad (6)$$

where  $\mathbf{K}_{u,u}^{(d)}$  is computed with a squared exponential kernel having a scaled signal variance  $\sigma_f^{2/D}$  [31].

In addition, in the SKI framework, the cross-covariance matrices  $\mathbf{K}_{\mathbf{f},u}$  and  $\mathbf{K}_{*,u}$  are approximated using sparse interpolation matrices,  $\mathbf{W}_{\mathbf{f}} \in \mathbb{R}^{N \times M}$  and  $\mathbf{W}_* \in \mathbb{R}^{N_* \times M}$ , such that

$$\mathbf{K}_{\mathbf{f},u} \approx \mathbf{W}_{\mathbf{f}} \mathbf{K}_{u,u} \quad \text{and} \quad \mathbf{K}_{*,u} \approx \mathbf{W}_* \mathbf{K}_{u,u}. \quad (7)$$

Each row of the interpolation matrices contains  $4^D$  interpolation weights for cubic interpolation [32] which is suggested in [17].

## IV. LARGE-SCALE MAGNETIC FIELD MAPS

Our goal is to compute magnetic field maps in 3D using magnetic field measurements as training data. In order to be able to use large data sets, we exploit mathematical formulations in the SKI framework adapted to magnetic field modeling to compute predictive means and variances in an efficient way. As the predictive distribution of the scalar potential model in (3) is based on the derivatives of the magnetic scalar potential, we approximate the elementwise computation  $\partial^2(\cdot)$  of the kernel matrices with D-SKI [20]. Using D-SKI,  $\partial^2(\cdot)$  can be simplified through differentiation of the interpolation scheme, such that it is

$$\partial^2(\mathbf{K}_{\mathbf{f},\mathbf{f}}) \approx (\partial \mathbf{W}_{\mathbf{f}}) \mathbf{K}_{u,u} (\partial \mathbf{W}_{\mathbf{f}})^\top, \quad (8a)$$

$$\partial^2(\mathbf{K}_{*,\mathbf{f}}) \approx (\partial \mathbf{W}_*) \mathbf{K}_{u,u} (\partial \mathbf{W}_{\mathbf{f}})^\top, \quad (8b)$$

where  $\partial \mathbf{W}_{\mathbf{f}} \in \mathbb{R}^{3N \times M_{\text{ind}}}$  and  $\partial \mathbf{W}_* \in \mathbb{R}^{3N_* \times M_{\text{ind}}}$ . Note that the first size of the interpolation matrices is multiplied by a

factor of 3 compared to (7) due to the three components of the magnetic field. As mentioned in the previous section, in SKI a cubic interpolation is advised [17], while in D-SKI a quintic interpolation scheme is used [20]. We use a cubic interpolation scheme for D-SKI, since preliminary experiments show that the approximation is sufficient for our application.

The predictive distribution in a new location  $\mathbf{p}_* \in \mathbb{R}^3$  of the scalar potential model for magnetic field modeling using D-SKI is given by

$$\mathbb{E}[\mathbf{f}_*] = (\partial \mathbf{W}_*) \mathbf{K}_{\mathbf{u}, \mathbf{u}} (\partial \mathbf{W}_f)^\top \mathbf{A}^{-1} \text{vec}(\mathbf{Y}^\top) \quad (9a)$$

$$\mathbb{V}[\mathbf{f}_*] = (\partial \mathbf{W}_*) \mathbf{K}_{\mathbf{u}, \mathbf{u}} (\partial \mathbf{W}_*)^\top - (\partial \mathbf{W}_*) \mathbf{C} (\partial \mathbf{W}_*)^\top \quad (9b)$$

The matrices  $\mathbf{A}$  and  $\mathbf{C}$  only depend on the training data and are defined as

$$\mathbf{A} := (\partial \mathbf{W}_f) \mathbf{K}_{\mathbf{u}, \mathbf{u}} (\partial \mathbf{W}_f)^\top + \sigma_y^2 \mathbf{I}_{3N}, \quad (10a)$$

$$\mathbf{C} := \mathbf{K}_{\mathbf{u}, \mathbf{u}} (\partial \mathbf{W}_f)^\top \mathbf{A}^{-1} (\partial \mathbf{W}_f) \mathbf{K}_{\mathbf{u}, \mathbf{u}}. \quad (10b)$$

A naive computation of the predictive mean and variance with (9a) and (9b) would require an inverse of a  $3N \times 3N$  matrix. Inducing inputs on a grid and kernel interpolation, however, enable efficient computations via Krylov subspace methods. The key to efficient computation is not to construct the matrices involved in (9a) and (9b) explicitly but to keep them in terms of a factorized format of 3 smaller matrices, one for each input dimension. Based on [20], we use preconditioned conjugate gradient to find the solution  $\alpha$  to the linear system given by

$$\mathbf{A} \alpha = \text{vec}(\mathbf{Y}^\top). \quad (11)$$

While in D-SKI fast matrix-vector-multiplications (MVMs) are computed via FFT, we compute them like in [17] via Kronecker MVMs as described in [33]. In this way, the predictive mean of a 3D map is computed with a computational complexity of

$$\mathcal{O} \left( J \left( 3N + M_{\text{ind}} \sum_{d=1}^3 M_{\text{ind}}^{(d)} \right) \right) = \mathcal{O}_{\text{ind}}, \quad (12)$$

where  $J$  is the number of iterations in the conjugate gradient. To compute the predictive variance of the magnetic field map, we build on the Lanczos Variance Estimates approach by [21], which is based on the Lanczos tridiagonalization algorithm. We use the Lanczos Variance Estimates within the D-SKI framework to find a low-rank approximation for  $\mathbf{A}$  given by

$$\mathbf{A} \approx \mathbf{Q}_T \mathbf{T}_T \mathbf{Q}_T^\top, \quad (13)$$

where  $\mathbf{Q}_T \in \mathbb{R}^{3N \times T}$  contains  $T$  orthonormal vectors corresponding to the first  $T$  leading eigenvalues and  $\mathbf{T}_T \in \mathbb{R}^{T \times T}$  has a tridiagonal structure [34]. Once an approximation of  $\mathbf{A}$  is found,  $\mathbf{C}$  from (10b) can be computed as

$$\begin{aligned} \mathbf{C} &= \mathbf{K}_{\mathbf{u}, \mathbf{u}} (\partial \mathbf{W}_f)^\top \mathbf{A}^{-1} (\partial \mathbf{W}_f) \mathbf{K}_{\mathbf{u}, \mathbf{u}} \\ &\approx \mathbf{K}_{\mathbf{u}, \mathbf{u}} (\partial \mathbf{W}_f)^\top \mathbf{Q}_T \mathbf{T}_T^{-1} \mathbf{Q}_T^\top (\partial \mathbf{W}_f) \mathbf{K}_{\mathbf{u}, \mathbf{u}}. \end{aligned} \quad (14)$$

As described in [21], we again exploit Kronecker algebra to compute the MVMs in (14) efficiently. The computation of  $\mathbf{C}$  has an associated computational complexity of

$\mathcal{O}(2T(3N + M_{\text{ind}}(M_{\text{ind}}^{(1)} + M_{\text{ind}}^{(2)} + M_{\text{ind}}^{(3)})))$  for magnetic field modeling [21]. In numerical implementations, the complexity is higher due to the full reorthogonalization required for the Lanczos tridiagonalization algorithm, scaling linearly with the number of training points  $N$  and Lanczos iterations  $T$ .

Alternatively, the predictive variance could also be computed using conjugate gradient. However, computing  $\mathbf{C}$  by solving a linear system to find  $\mathbf{A}^{-1} (\partial \mathbf{W}_f) \mathbf{K}_{\mathbf{u}, \mathbf{u}}$  needs to be done for each column of  $(\partial \mathbf{W}_f) \mathbf{K}_{\mathbf{u}, \mathbf{u}}$  sequentially, which is not particularly efficient. In [35], the variance is stochastically estimated by drawing samples from the predictive distributions [36]. While this approach can reduce the computational complexity associated with the computation of the predictive variances, it introduces significant accuracy losses.

Once  $\alpha$  and  $\mathbf{C}$  are computed, predictions in new locations can be computed with (9a) and (9b) again by exploiting Kronecker algebra and the structure of the interpolation matrices.

## V. EXPERIMENTS

In our experiments, we first compare our method to existing ones in simulations with synthetic data, then show our method's scalability with large-scale magnetometer data collected with a motion capture suit. All computations are done on a 2016 HP ZBook Studio G3 laptop (Intel Core i7 @ 2.60 GHz, 8GB RAM).

### A. Accuracy analysis for growing mapping area

In the first simulation, we compare the accuracy of magnetic field maps computed with our and existing methods. For this, we create a synthetic data set of 6000 data points, representing a magnetic field map. Each input is a random 3D vector lying in a box confined by  $[-20, 20] \times [-20, 20] \times [0.01, 0.01]$  and the corresponding output is sampled from a GP prior with a curl-free kernel with hyperparameters, length scale, signal variance, and noise variance,  $[\ell, \sigma_f^2, \sigma_y^2] = [2, 1, 0.01]$  which is equivalent to drawing samples from model (1) [27].

We compute two maps with different sizes: area 1 of size  $20 \times 20$  and area 2 of size  $40 \times 40$ . For area 1, a subset of the 6000 data points is used, laying in the white square shown in Fig. 2, and for area 2, all data is used. We divide the data points in each area into 80% training data and 20% testing data. With the testing data, we compute root mean square errors (RMSEs) as a metric for accuracy.

We compare our method to a GP where we downsample the data, as well as to the basis function approach by [16], showing how the area size impacts the accuracy of the map for different  $M_{\text{ind}}$ ,  $N_{\text{down}}$  and  $M_{\text{bf}}$  in the respective methods. Since the domain on which basis functions are computed needs to be a bit bigger than the mapping area [16], we add twice the length scale in each dimension.

To be able to compare all methods, we compute the computational complexity of our method  $\mathcal{O}_{\text{ind}}$  as defined in (12), and impose this complexity as the computational budget for the other methods. The complexity of the approach of downsampling the data is equal to the cubic complexity of full GP. The complexity of the basis function approach is linear

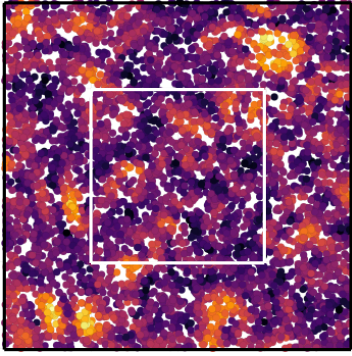


Fig. 2: Synthetic data with white and black squares denoting area 1 of size  $20 \times 20$  ( $N \approx 1526$ ) and area 2 of size  $40 \times 40$  ( $N = 6000$ ). The scattered dots are data points and the color corresponds to the magnitude of the magnetic field.

in the number of data points and quadratic in the number of basis functions. By equating the complexities

$$\mathcal{O}_{\text{ind}} = \mathcal{O}(M_{\text{bf}}^2, 3N) = \mathcal{O}(N_{\text{dwn}}^3), \quad (15)$$

and solving for  $M_{\text{bf}}$  and  $N_{\text{dwn}}$ , those numbers can be used in the corresponding methods.

For each of the two area sizes, we have six different settings, where we vary the number of inducing inputs, i.e.,  $M_{\text{ind}}^{(1)} = M_{\text{ind}}^{(2)} = [10, 20, 40, 80, 100, 200]$  and  $M_{\text{ind}}^{(3)} = 5$ . The number of iterations in the conjugate gradient,  $J$ , is based on the tolerance for accuracy. Table I summarizes all settings used in the simulations. The first two rows are the total number of inducing inputs  $M_{\text{ind}}$  and the other rows are the number of basis functions and downsampled data points that result from the imposed computational budget  $\mathcal{O}_{\text{ind}}$ . We run the simulation 100 times, where each time new data is sampled. An example of the data is illustrated in Fig. 2.

Fig. 3 shows the RMSEs on the testing data computed with the three methods for area 1 (solid line) and area 2 (dash-dotted line). The mean and standard deviation of the RMSE from 100 runs are plotted. The horizontal axis denotes the 6 different settings as described in Table I. The mean RMSE of the full GP is given for area 2 as a reference (dashed black line), for area 1 the RMSE is very similar and therefore not shown. The figure shows that for each method, the RMSEs are larger for the  $40 \times 40$  area than for the  $20 \times 20$  area. This implies that a larger  $M_{\text{ind}}$ ,  $N_{\text{dwn}}$ , and  $M_{\text{bf}}$  are required for a larger area to achieve low RMSE. Also, all methods converge to the RMSE of the full GP in the limit. Comparing our methods to the other two, for both areas our method has lower RMSEs. In addition, the difference in RMSEs between our methods and the other methods is more significant for the larger area. The main takeaway of this simulation is that our method has better accuracy than the other methods when the computational budget  $\mathcal{O}_{\text{ind}}$  is imposed for all methods. It follows that the computational cost to compute a map of a specified accuracy

is lower for our approach compared to the others. Especially when computing large-scale maps, this becomes important: Since the number of basis functions needed to approximate the kernel function sufficiently is known to scale with the domain size [16], a trade-off between accuracy and computational cost needs to be made. A similar trade-off is necessary for downsampling the data, because more data points are required for larger areas.

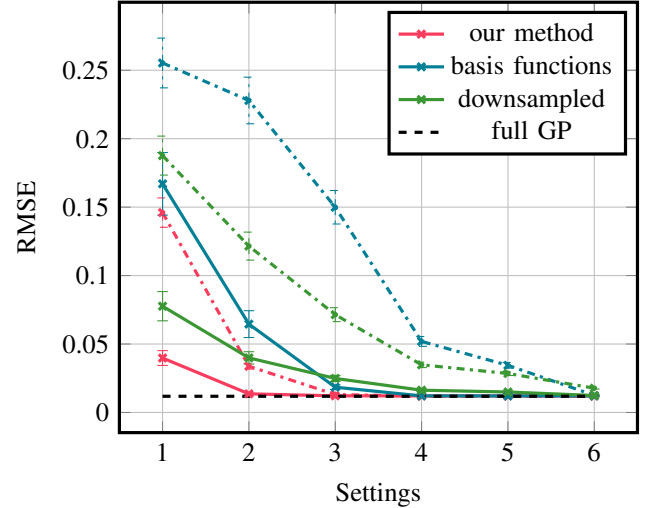


Fig. 3: RMSE for area 1 (solid line) and area 2 (dash-dotted line). For reference, the RMSE of full GP is given for area 2 only, because the other value is very similar. The horizontal axis are the six different settings described in Table I. Mean and standard deviation are plotted for 100 runs of the simulation.

### B. Analysis of maps with divided mapping area

As mentioned in the previous section, the basis function approach requires a large number of basis functions for growing mapping areas. To lower the computational complexity, an alternative approach for computing large-scale maps with basis functions is dividing the area into smaller areas for each of which a GP approximation with basis functions is made [18]. To train each smaller map, not only training data from the mapping area is used, but also training data in close proximity to that area. In a second simulation, we show that this strategy may result in inconsistencies at boundaries. We use synthetic data sampled from a GP prior with a curl-free kernel and with hyperparameters  $[\ell, \sigma_{\mathbf{f}}^2, \sigma_{\mathbf{y}}^2] = [5, 1, 0.01]$ , divide the data into four regions and downsample the data by different factors in every region, as shown in Fig. 4 (a). The reason for it is to analyze how the amount of data points in a neighboring area influences the inconsistencies at the boundaries. For training each map, we use data points in each area as well as data from an overlap of size  $0\ell, 0.1\ell, 0.3\ell$  to the neighboring areas. The domain on which we compute basis functions is then the size of the mapping area plus the overlap plus twice the length scale. The result of the reconstruction is shown in Fig. 4 (c)-(e). The figure shows that for no overlap, there are visible

	Setting	1	2	3	4	5	6
$M_{\text{ind}}$	Area 1	500	2K	8K	32K	50K	200K
	Area 2	500	2K	8K	32K	50K	200K
$M_{\text{bf}}$	Area 1	35 - 36	97 - 98	262 - 266	726 - 738	1012 - 1027	2843 - 2887
	Area 2	15 - 16	57 - 58	152 - 153	420 - 422	584 - 587	1635 - 1646
$N_{\text{dwn}}$	Area 1	55 - 56	107 - 108	208 - 210	411 - 415	512 - 518	1020 - 1030
	Area 2	49 - 50	120 - 121	231 - 232	455 - 457	567 - 569	1126 - 1131

TABLE I:  $M_{\text{ind}}$ ,  $M_{\text{bf}}$  and  $N_{\text{dwn}}$  used in the 6 settings of the simulation. The first and second rows for basis function approach and downsampled data are the values for the area 1 and 2, respectively. Value ranges for 100 simulations.

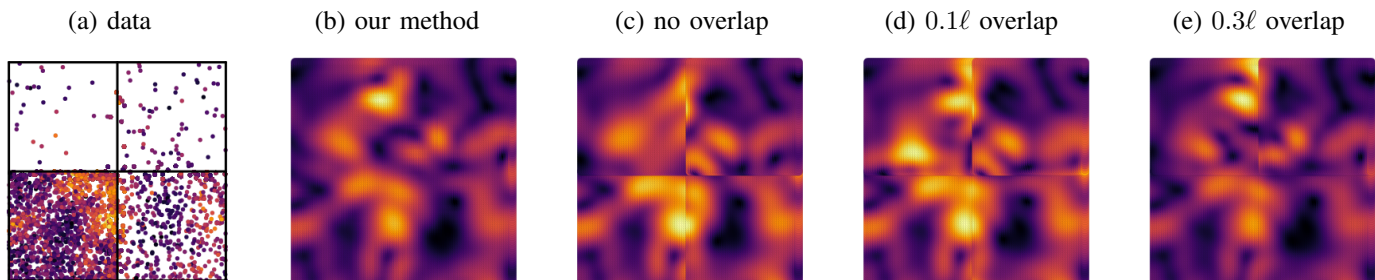


Fig. 4: Magnetic field data (a). Predictions with our method for the whole area (b). Predictions with basis functions in smaller areas separately using an overlap of 0,  $0.1\ell$ , and  $0.3\ell$  for the training data, respectively (c)-(e).

inconsistencies in the mean on the magnetic field prediction. For  $0.1\ell$  and  $0.3\ell$ , the inconsistencies are smaller but still present. In addition, the inconsistencies are more visible at the boundaries of areas with fewer data. As a comparison, a map reconstructed with our method is shown in Fig. 4 (b), where the mapping area is not divided.

### C. Large-scale map in university building

For our experiments, we use data collected with a motion capture suit (Xsens MVN Link [37]). The suit contains 17 inertial measurement units (IMUs) equipped with magnetometers tightly attached to segments all over the body. The IMUs provide accelerometer, gyroscope, and magnetometer data with position and orientation data in the navigation frame at a maximum sample rate of 240 Hz. The magnetometers have been calibrated using software available with the suit, such that after calibration the norm of the undisturbed Earth's magnetic field is 1 [38]. In a pre-processing step, the data is first rotated to the global frame defined by the magnetic North pole, and second, the mean is subtracted from the  $x$ -,  $y$ -, and  $z$ -component, since we only model the anomalies of the magnetic field. We use the data collected from one IMU that is located at the pelvis. In the first large-scale experiment, the magnetic field of one of the university hallway wings at the TU Delft is computed based on  $N = 21\,931$  magnetic field measurements. The area of interest for this experiment is a rectangle bounded by  $[-34\text{ m}, 34\text{ m}] \times [-5.25\text{ m}, 5.25\text{ m}]$  located at a height of approximately 1 m. The magnetic field is estimated with our method using an inducing point grid of size  $400 \times 40 \times 4$ , for a total of 64 000 inducing points. The

number of inducing points per dimension is chosen such that several inducing points are present per characteristic length scale. The hyperparameters are trained on a subset of the data by minimizing the log marginal likelihood with the curl-free kernel, resulting in  $\ell = 0.5\text{ m}$ ,  $\sigma_f = 0.2$  and  $\sigma_y = 0.01$ . Fig. 1 shows the magnitude of the magnetic field predictions, computed from the three components. Darker regions in the figure correspond to a higher magnitude of the magnetic field. Since there are metallic lockers located in the hallway, we expected a strong magnetic anomaly, which is visible in the figure. Fig. 5 shows a smaller section of the magnetic field map, in terms of its magnitude, as well as its  $x$ -,  $y$ - and  $z$ -component. The magnetic disturbance caused by the lockers is mostly visible in the  $z$ -components, as shown in the upper right part of Fig. 5 (d). The transparency in the figure indicates the certainty of the map.

Regarding computational time, computing the map with 21 931 measurements took approximately 1 min for training and 18 s for testing. In a second experiment with 41 383 data points, the training took approximately 97 s for training and 18 s for testing. While with full GP the map would not be feasible to compute on our laptop, our method scales very well: The computing time approximately only doubling when doubling the data points, thus is approximately linear in  $N$ .

## VI. CONCLUSION

In this paper, we described an algorithm to efficiently compute large-scale magnetic field maps using approximate Gaussian process regression. We used inducing inputs on a grid and structured kernel interpolation with derivative to

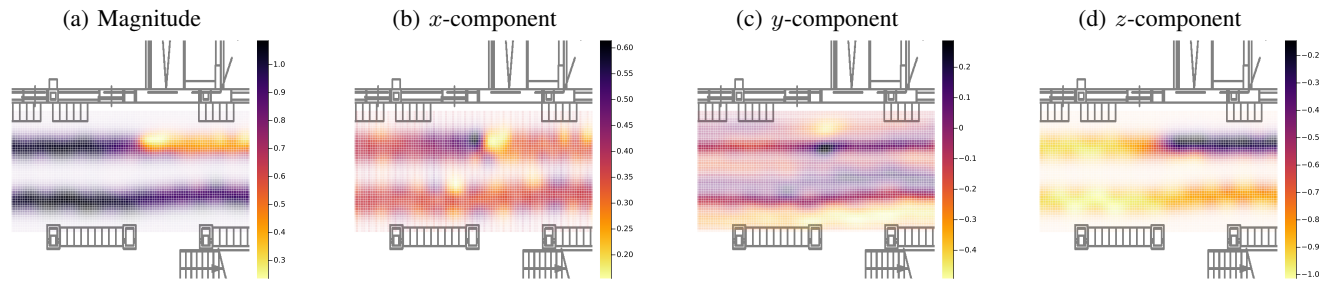


Fig. 5: Magnetic field maps for a smaller region of map in Fig. 1, located on the left part of the hallway. In the  $z$ -components, the lockers are strongly visible, indicating that the metal in the lockers mainly disturbs the magnetic field in that direction. Transparency indicates the certainty of the prediction.

compute the predictive mean and an algorithm based on Lanczos tridiagonalization to compute variance estimates. We compared our method to existing methods in simulations and showed its scalability in large-scale experiments. There are multiple directions for future work. The kernel in the scalar potential model can be complemented by a linear kernel as in [19], [22] in order to also model the underlying Earth’s magnetic field. The linear kernel, however, is not a product kernel and can thus not be decomposed as a Kronecker product. When using the equivalent curl-free model instead, the kernel consists of a constant term and a curl-free kernel [27], where the constant term can again be decomposed as a Kronecker product. In this way, a model considering both the Earth’s magnetic field and the spatial anomalies can be used in the D-SKI framework. In addition, the presented method can be extended to online mapping to enable its use in e.g. simultaneous localization and mapping (SLAM) algorithms. The SKI framework for online GPs has been described in [39] and can be adapted to magnetic field mapping.

#### ACKNOWLEDGMENT

This publication is part of the project “Sensor Fusion For Indoor Localisation Using The Magnetic Field” with project number 18213 of the research program Veni which is (partly) financed by the Dutch Research Council (NWO).

We would like to thank Fabian Girrbaach from Movella Technologies for post-processing data collected with the motion capture suit.

#### REFERENCES

- [1] F. Zafari, A. Gkelias, and K. K. Leung, “A survey of indoor localization systems and technologies,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.
- [2] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, “Recent advances in indoor localization: A survey on theoretical approaches and applications,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2016.
- [3] S. He and S.-H. G. Chan, “Wi-fi fingerprint-based indoor positioning: Recent advances and comparisons,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 466–490, 2015.
- [4] A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafessah, S. Al-Hadhrani, M. A. Al-Ammar, and H. S. Al-Khalifa, “Ultra wideband indoor positioning technologies: Analysis and recent advances,” *Sensors*, vol. 16, no. 5, p. 707, 2016.
- [5] 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.
- [6] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Röning, “Simultaneous localization and mapping using ambient magnetic field,” in *2010 IEEE Conference on Multisensor Fusion and Integration*. IEEE, 2010, pp. 14–19.
- [7] 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*, 2011, pp. 141–154.
- [8] P. Robertson, M. Frassl, M. Angermann, M. Doniec, B. J. Julian, M. G. 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 *International conference on indoor positioning and indoor navigation*. IEEE, 2013, pp. 1–10.
- [9] G. Berkovich, D. Churikov, J. Georgy, and C. Goodall, “Coursa venue: Indoor navigation platform using fusion of inertial sensors with magnetic and radio fingerprinting,” in *2019 22th International Conference on Information Fusion (FUSION)*. IEEE, 2019, pp. 1–6.
- [10] S.-E. Kim, Y. Kim, J. Yoon, and E. S. Kim, “Indoor positioning system using geomagnetic anomalies for smartphones,” in *2012 International conference on indoor positioning and indoor navigation (IPIN)*. IEEE, 2012, pp. 1–5.
- [11] J. Torres-Sospedra, D. Rambla, R. Montoliu, O. Belmonte, and J. Huerta, “Ujiindoorloc-mag: A new database for magnetic field-based localization problems,” in *2015 International conference on indoor positioning and indoor navigation (IPIN)*. IEEE, 2015, pp. 1–10.
- [12] F. Viset, R. Helmons, and M. Kok, “An extended Kalman filter for magnetic field SLAM using Gaussian process regression,” *Sensors*, vol. 22, no. 8, p. 2833, 2022.
- [13] A. Solin, S. Särkkä, J. Kannala, and E. Rahtu, “Terrain navigation in the magnetic landscape: Particle filtering for indoor positioning,” in *2016 European Navigation Conference (ENC)*. IEEE, 2016, pp. 1–9.
- [14] C. E. Rasmussen and H. Nickisch, “Gaussian processes for machine learning (GPML) toolbox,” *The Journal of Machine Learning Research*, vol. 11, pp. 3011–3015, 2010.
- [15] D. Nguyen-Tuong, J. Peters, and M. Seeger, “Local Gaussian process regression for real-time online model learning,” *Advances in neural information processing systems*, vol. 21, 2008.
- [16] A. Solin and S. Särkkä, “Hilbert space methods for reduced-rank Gaussian process regression,” *Statistics and Computing*, vol. 30, no. 2, pp. 419–446, 2020.
- [17] A. Wilson and H. Nickisch, “Kernel interpolation for scalable structured Gaussian processes (KISS-GP),” in *Proceedings of the International conference on machine learning*. PMLR, 2015, pp. 1775–1784.
- [18] M. Kok and A. Solin, “Scalable magnetic field SLAM in 3D using Gaussian process maps,” in *2018 21st international conference on information fusion (FUSION)*. IEEE, 2018, pp. 1353–1360.
- [19] A. Solin, M. Kok, N. Wahlström, T. B. Schön, and S. Särkkä, “Modeling and interpolation of the ambient magnetic field by Gaussian processes,” *IEEE Transactions on robotics*, vol. 34, no. 4, pp. 1112–1127, 2018.
- [20] D. Eriksson, K. Dong, E. Lee, D. Bindel, and A. G. Wilson, “Scaling Gaussian process regression with derivatives,” *Advances in neural information processing systems*, vol. 31, 2018.



- [21] G. Pleiss, J. Gardner, K. Weinberger, and A. G. Wilson, "Constant-time predictive distributions for Gaussian processes," in *International Conference on Machine Learning*. PMLR, 2018, pp. 4114–4123.
- [22] N. Wahlström, M. Kok, T. B. Schön, and F. Gustafsson, "Modeling magnetic fields using Gaussian processes," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 3522–3526.
- [23] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Röning, "Magnetic field-based SLAM method for solving the localization problem in mobile robot floor-cleaning task," in *2011 15th international conference on advanced robotics (ICAR)*. IEEE, 2011, pp. 198–203.
- [24] N. Akai and K. Ozaki, "Gaussian processes for magnetic map-based localization in large-scale indoor environments," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 4459–4464.
- [25] Y. Li, Z. He, J. Nielsen, and G. Lachapelle, "Using Wi-Fi/magnetometers for indoor location and personal navigation," in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2015, pp. 1–7.
- [26] T. N. Lee and A. J. Canciani, "MagSLAM: Aerial simultaneous localization and mapping using Earth's magnetic anomaly field," *Navigation*, vol. 67, no. 1, pp. 95–107, 2020.
- [27] N. Wahlström, "Modeling of magnetic fields and extended objects for localization applications," Ph.D. dissertation, Linköping University, 2015.
- [28] S. Särkkä, "Linear operators and stochastic partial differential equations in Gaussian process regression," in *International Conference on Artificial Neural Networks*. Springer, 2011, pp. 151–158.
- [29] C. Williams and M. Seeger, "Using the Nyström method to speed up kernel machines," *Advances in neural information processing systems*, vol. 13, 2000.
- [30] J. Quinero-Candela and C. E. Rasmussen, "A unifying view of sparse approximate Gaussian process regression," *The Journal of Machine Learning Research*, vol. 6, pp. 1939–1959, 2005.
- [31] Y. Saatçi, "Scalable inference for structured Gaussian process models," Ph.D. dissertation, University of Cambridge, 2012.
- [32] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE transactions on acoustics, speech, and signal processing*, vol. 29, no. 6, pp. 1153–1160, 1981.
- [33] A. G. Wilson, "Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes," Ph.D. dissertation, University of Cambridge, 2014.
- [34] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU press, 2013.
- [35] A. G. Wilson, C. Dann, and H. Nickisch, "Thoughts on massively scalable Gaussian processes," *arXiv preprint arXiv:1511.01870*, 2015.
- [36] G. Papandreou and A. L. Yuille, "Efficient variational inference in large-scale Bayesian compressed sensing," in *Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011, pp. 1332–1339.
- [37] Movella Inc., "MVN Link," [Online; accessed March, 2023], <https://www.movella.com/products/motion-capture/xsens-mvn-link>.
- [38] Xsens, "Interpreting magnetic field data represented as an arbitrary unit (a.u.)," *Xsens knowledge base*, Aug 2022, [Online; accessed November 14, 2022], <https://base.xsens.com/s/article/Interpreting-magnetic-field-data-represented-as-an-arbitrary-unit-a-u>.
- [39] S. Stanton, W. Maddox, I. Delbridge, and A. G. Wilson, "Kernel interpolation for scalable online Gaussian processes," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021, pp. 3133–3141.