This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598

# Gaussian Process Kernels for Efficient Sequential Sampling of Electromagnetic Radiation on Spherical Surfaces

Yens Lindemans, *Graduate Student Member, IEEE*, Tim Claeys, *Member, IEEE*, Davy Pissoort, *Senior Member, IEEE*, Ivo Couckuyt, *Member, IEEE*, and Tom Dhaene, *Senior Member, IEEE*

***Abstract*—Traditional far-field characterization of wireless devices relies on dense, fixed-grid measurements over the entire measurement surface, leading to high time and cost overheads. This paper develops a data-efficient sequential sampling strategy for far-field radiation measurements using Bayesian optimization (BO) with Gaussian Process (GP) surrogate modeling on spherical domains. A central contribution is the adaptation and evaluation of GP kernels to capture the continuity and smoothness of spherical radiation patterns, enabling accurate surrogate models for BO. To demonstrate practicality, we compare kernel choices within the proposed framework and introduce a motion-aware acquisition function that reduces travel time between sampling points. The methodology is validated on both simple and more directive antenna scenarios, showing that the proposed approach reliably identifies key radiation features with significantly fewer samples than dense grid-based methods. These results highlight the potential of BO-driven sequential sampling for efficient and reliable radiation characterization on spherical domains.**

***Index Terms*—Bayesian Optimization, Gaussian Process, Manifold Gaussian Process, Sequential Sampling, Spherical Kernel Function**

## I. INTRODUCTION

As wireless devices evolve across applications ranging from radio frequency identification (RFID) to broadband communication systems, accurate antenna measurement has become increasingly important. Each device exhibits distinct radiation characteristics in both intended and spurious emissions, and compliance with regional regulations is mandatory. Within the European Union, for example, the radio equipment directive (RED) defines essential requirements that must be met by all radio devices[1]. Many associated standards impose limits on the effective radiated power (ERP) or effective isotropic radiated power (EIRP)[2], requiring the maximum radiated field to be identified and measured in spherical coordinates around the device under test (DUT). Similar requirements apply to out-of-band (OOB) and spurious emissions. These far-field measurements are particularly challenging because they require dense sampling, are time-consuming, and are typically performed late in the design cycle, when redesigning the DUT in case of non-compliance becomes costly [1].

To resolve data-gathering challenges, sequential sampling techniques have been widely studied in other contexts. Several surrogate-based techniques, including Bayesian optimization (BO), leave-one-out cross-validation, and local linear approximation, have been designed to reduce the number of required samples [2]–[6]. These methods have proven effective for improving sampling efficiency in complex design spaces, but they have been primarily developed for Euclidean (planar) domains. Their direct extension to spherical radiation measurements has not been systematically explored in the literature.

In this work, we develop a sequential sampling strategy tailored to spherical far-field characterization. We employ BO, which has proven highly effective for optimizing expensive black-box functions [4], [7]–[9], relying on a Gaussian process (GP) surrogate to balance exploration and exploitation when selecting new sampling points. However, directly applying standard GPs to spherical data poses challenges since these models fail to respect the geometry of the sphere. To address this, we implement and evaluate *manifold GPs* defined intrinsically on spherical domains and compare them against an extrinsic Cartesian model. Furthermore, we introduce a motion-aware acquisition function that accounts for probe travel times, thereby improving the practicality of sequential sampling in constrained measurement setups.

The proposed approach is validated using simulated far-field data from both simple and directive antenna scenarios. Results demonstrate that BO reliably identifies key radiation features with significantly fewer samples than dense grid-based methods, while the motion-aware strategy reduces travel overhead. To the best of our knowledge, this is the first work to systematically develop and apply BO-driven sequential sampling for spherical electromagnetic (EM) radiation measurements.

Building on the challenges outlined above, the remainder of this paper is organized as follows. Section II provides a brief summary of GPs, a theoretical overview of challenges in spherical modeling, and a discussion of this work's primary

---

Manuscript received xxx, 2024. This work is supported by Flemish Research Foundation (FWO-Vlaanderen G095224N) and the Flanders AI Research Program.

Yens Lindemans, Ivo Couckuyt, and Tom Dhaene are with the Department of Information Technology, Ghent University-imec, Technologiepark-Zwijnaarde 126, 9052 Ghent, Belgium (email: yens.lindemans@ugent.be; ivo.couckuyt@ugent.be; tom.dhaene@ugent.be)

Tim Claeys and Davy Pissoort are with the Waves: Core Research and Engineering (WaveCore), KULeuven, Spoorwegstraat 12, 8200 Sint-Michiels, Belgium (email: tim.claeys@kuleuven.be; davy.pissoort@kuleuven.be)

Davy Pissoort is also with Flanders Make, KULeuven, Gaston Geenslaan 8, 3001, Leuven, Belgium

[1] Available at https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32014L0053

[2] Accessed on 2023-08-14. Available at https://www.etsi.org/deliver/etsi_en/300200_300299/30022001/03.01.01_60/en_30022001v030101p.pdf

goals. Section III presents the methodology, detailing specific adaptations for spherical surfaces. In Section IV, the fundamentals of BO and its application to sequential sampling are presented. Section V introduces the evaluation metrics used to assess model performance. Section VI examines the efficacy of manifold methods in modeling and sequential sampling, including comparisons to standard GP models. Finally, Section VII concludes the paper and outlines directions for future research.

## II. THEORETICAL BACKGROUND

### A. Gaussian Process Model

GP regression works by sampling from a distribution of functions, where the mean of this distribution serves as the prediction and the variance represents the uncertainty of the prediction. The initial distribution, or *prior*, is a multivariate normal distribution with mean $m$ and covariance matrix $\Sigma$. This distribution is updated with data by making use of Bayes' rule

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{normalization constant}}\,, \tag{1}$$

to form a posterior distribution with a new mean and new covariance matrix. Here, the likelihood describes the probability of the observed data given the model parameters, and the posterior ultimately forms the basis for inference. For simplicity, the prior mean in GP regression is often set to zero. This is justified since the data can always be standardized. In contrast, the covariance matrix must be derived from a meticulously selected kernel function $k(\boldsymbol{x}, \boldsymbol{x}')$. For a dataset $D = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=0}^{N}$, the covariance matrix is defined as

$$[\Sigma]_{i,j} = k(\boldsymbol{x}_i, \boldsymbol{x}_j), \quad \boldsymbol{x}_i,\, \boldsymbol{x}_j \in D, \tag{2}$$

such that the covariance between outputs depends on the relationship between input points. The kernel thus enforces that points closer in the input space exhibit higher covariance, and therefore more similar function values, than those farther apart.

The notion of *closeness* is defined by a distance metric within the kernel function. Different kernels offer distinct properties, allowing flexibility in modeling diverse types of data. The most common kernels are the *Matérn* and *squared exponential* (SE) kernels, given by

$$k_\nu(\boldsymbol{x},\, \boldsymbol{x}') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\,\delta\right)^{\nu} K_\nu\left(\sqrt{2\nu}\,\delta\right) \quad \text{and} \tag{3}$$

$$k_{\text{SE}}(\boldsymbol{x},\, \boldsymbol{x}') = \sigma^2 \exp\left(-\frac{\delta^2}{2}\right), \tag{4}$$

with

$$\delta = \sqrt{\sum_{d=1}^{D} \frac{(x_d - x'_d)^2}{l_d^2}}\,. \tag{5}$$

Here, $\Gamma(\cdot)$ represents the Gamma function, $K_\nu(\cdot)$ is the modified Bessel function of the second kind, $\sigma^2$ is a positive scaling factor, $D$ is the input space dimensionality, $\nu$ is the *smoothness parameter* of the Matérn kernel, and $l_d$ denote characteristic length scales. The smoothness parameter, typically a half-integer starting from $1/2$, determines both the continuity of the kernel and, consequently, the continuity of model predictions. For half-integer values of $\nu$, the Matérn kernel can be simplified and written as a product of an exponential term and a polynomial of degree $\lceil \nu \rceil - 1$, e.g.

$$k_{1/2} = \sigma^2 \exp\left(-\delta\right), \tag{6}$$

$$k_{3/2} = \sigma^2 \left(1 + \sqrt{3}\,\delta\right) \exp\left(-\sqrt{3}\,\delta\right), \quad \text{and} \tag{7}$$

$$k_{5/2} = \sigma^2 \left(1 + \sqrt{5}\,\delta + \frac{5}{3}\delta^2\right) \exp\left(-\sqrt{5}\,\delta\right). \tag{8}$$

A GP model using a Matérn kernel is differentiable up to $\lceil \nu \rceil - 1$ times, illustrating the relationship between the Matérn and SE kernels. Specifically, the SE kernel is a limiting case of the Matérn kernel as $\nu \to \infty$, making it infinitely differentiable [10]. For most engineering applications, the Matérn $5/2$ or SE kernel are typical choices, as physical systems are generally represented by smooth functions governed by differential equations.

In GP models, the length scales of a kernel serve as hyperparameters optimized by maximizing the marginal log-likelihood of the posterior. After training the hyperparameters, the mean and covariance of the posterior are calculated as

$$\tilde{m}(\boldsymbol{x}^*) = K_{\boldsymbol{x}^*\boldsymbol{x}} \left[\Sigma_{\boldsymbol{x}\boldsymbol{x}}\right]^{-1} \cdot \boldsymbol{y} \tag{9}$$

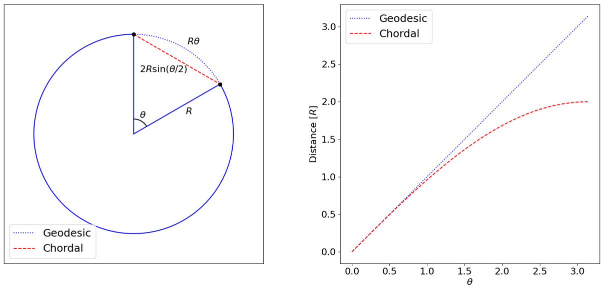
$$\tilde{\Sigma}_{\boldsymbol{x}^*\boldsymbol{x}^*} = K_{\boldsymbol{x}^*\boldsymbol{x}^*} - K_{\boldsymbol{x}^*\boldsymbol{x}} \left[\Sigma_{\boldsymbol{x}\boldsymbol{x}}\right]^{-1} K_{\boldsymbol{x}\boldsymbol{x}^*} \tag{10}$$

where $\boldsymbol{x}^*$ represents a test point, and $\boldsymbol{x}$ and $\boldsymbol{y}$ represent the training data. The matrices $K$ are constructed using the kernel function $k$, similarly to (2). This flexible framework, driven by an adaptable choice of kernels and equipped with uncertainty quantification, makes GP regression a powerful tool for BO, and consequently, sequential sampling.

### B. Problem Statement – Far-Field Data

The input of our GP model consists of coordinates on a 2D spherical surface surrounding the DUT. An intuitive approach for representing these coordinates is to use a spherical coordinate system $(r,\, \theta,\, \phi)$, with the radius $r$ kept constant, $\theta \in [0,\, \pi]$ the polar angle, and $\phi \in [0,\, 2\pi]$ the azimuthal angle. This simplifies the input for the GP model to $\boldsymbol{x} = (\theta,\, \phi)$. The problem with this approach is that the kernels in (3) and (4) assume a Euclidean space to formulate their distance metric. However, a sphere is not isometric to any Euclidean space, meaning that there exists no distance-preserving map from points on the sphere to points in a Euclidean plane [11], [12]. Consequently, inserting spherical coordinates directly into a GP model is problematic, as the Euclidean distance metric (5), or $L_2$ norm, is ill-defined on this coordinate system. For instance, there is no continuity between points located at $\phi = 0$ and $\phi = 2\pi$. Similarly, at the poles of the sphere the *nearness* of points is incorrectly defined since e.g. $\boldsymbol{x} = [\theta = 0, \phi = 0]$ and $\boldsymbol{x}' = [\theta = 0, \phi = \pi]$ are physically the same point but do not result in $\delta = 0$.

An intuitive solution would be to replace the $L_2$ norm (5) inside the kernel function with a geodesic distance defined on the sphere (see Fig. 1). However, this approach is not straightforward, as prior research has shown that the commonly used SE kernel, when paired with a geodesic distance



(a) (b)

Fig. 1. Geodesic and chordal distance on a circle [17].

on a non-Euclidean manifold, does not guarantee positive semi-definiteness across all length scales [13], [14]. While the geodesic SE kernel may work in specific cases (such as in [15]), it is not universally applicable for every dataset and can restrict the range of trainable length scales, potentially leading to detrimental problems when training the GP model [16]. Conversely, the authors in [13] demonstrate that the Matérn $1/2$ kernel avoids this issue and can reliably incorporate a geodesic distance on a spherical surface. However, this restriction in kernel choice may limit its applicability. To address this, a comparison of different kernel variants is included in this work to identify those best suited for sequential sampling of EM radiation on spherical domains.

### C. Goal Statement

The primary aim of this work is to develop and evaluate a data-efficient sequential sampling strategy for far-field radiation measurements around a DUT, using GP modeling and BO. Given the challenges of using Euclidean-based kernels on spherical data, central to this strategy is the selection of suitable GP kernels for spherical input domains, as kernels are the beating heart of GP regression and critically influence both model accuracy and optimization performance.

We define the domain of possible sampling points on the spherical surface around the DUT as $X \subset \mathcal{S}^2$. The objective is twofold: (1) to accurately model the radiated electric field around the DUT as a function $f : \boldsymbol{x} \mapsto f(\boldsymbol{x})$, and (2) to efficiently identify the maximum field value using BO. Notably, the proposed approach is not tied to a specific observable (e.g., emitted power, electric field strength, or magnetic field strength), as the general properties of the data remain consistent in the far-field regime.

This methodology emphasizes the importance of selecting a kernel that enables both accurate GP modeling of the radiated EM field (or power density) and efficient optimization on a non-Euclidean domain. Formally, the optimization problem consists of collecting a set of positions $P = \{\boldsymbol{p}_n = [\theta_n, \phi_n],\, \boldsymbol{p}_n \in X\}_{n=1}^{N}$ and corresponding field values $E = \{e_n = f(\boldsymbol{p}_n),\, \boldsymbol{p}_n \in P\}_{n=1}^{N}$ to locate the maximum field value. The task is to minimize the number of required samples $N$ while maximizing the information gained about the global maximum.

## III. Kernels for Spherical Gaussian Processes

We begin the discussion by giving an overview of different kernel methods suitable for GP regression on a spherical surface. The following sections introduce these kernels, exploring their mathematical formulations, theoretical advantages, and potential limitations when applied to spherical data.

### A. 2D Euclidean Kernel with Spherical Coordinates - 2DL2

One of the most straightforward approaches is to ignore the nonlinear structure of the input space and apply $\boldsymbol{x} = [\theta, \phi]$ directly in the Matérn or SE kernels, using the $L_2$ norm as defined in (3), (4). As discussed in Section II-B, this method disregards any symmetry of the input space and leads to discontinuities at the *prime meridian* of the sphere. Nonetheless, it will serve as a benchmark strategy to evaluate the performance of alternative kernel methods. We will refer to this approach as the *2DL2* method in subsequent sections, and the kernel functions will be written as $k_{\mathrm{2DL2},\nu}$ and $k_{\mathrm{2DL2},\infty}$ for the Matérn and SE kernel, respectively.

### B. Geodesic Kernel

Building on the discussion in Section II-B, a natural alternative to the $L_2$ norm is to define a geodesic distance, which respects the spherical structure of the input space. On the sphere, the shortest distance between two points $\boldsymbol{x}$ and $\boldsymbol{x}'$ is defined as

$$\delta_{\mathrm{g}}(\boldsymbol{x},\boldsymbol{x}') = r\arccos\left(\cos\theta\cos\theta' + \sin\theta\sin\theta'\cos\Delta\phi\right), \tag{11}$$

where $r$ is the radius of the sphere and $\Delta\phi = (\phi - \phi')$. An example of this distance on a circle is shown in Fig. 1. As mentioned previously, this geodesic distance only leads to provable positive semi-definite kernels when used within the Matérn $1/2$ kernel [13], [14], [16]. Therefore, the geodesic kernel used in this work is

$$k_{2\mathrm{D},\delta_{\mathrm{g}}}(\boldsymbol{x},\boldsymbol{x}') = \sigma^2 \exp\left(-\frac{\delta_{\mathrm{g}}(\boldsymbol{x},\,\boldsymbol{x}')}{l}\right), \tag{12}$$

with $\sigma^2$ a positive scaling factor, resulting in analogous properties as the original Matérn $1/2$ kernel but with a continuous distance metric on the sphere. It is important to note that this geodesic kernel does not support multiple length scales, unlike the kernels in (3) and (4), which contain a length scale for each dimension of $\boldsymbol{x}$. Consequently, if the objective function shows a significantly different behavior across its coordinates, this limitation may lead to a reduction of modeling accuracy.

### C. Geometric Kernel

Recent research has focused on defining alternative kernel functions on spheres [17] and, more generally, on any non-Euclidean manifold [18], [19]. In particular, the entire family of Matérn kernels (and by extension the SE kernel) can be constructed intrinsically on a manifold by considering the eigenfunctions of the Laplace-Beltrami operator specific to that manifold. The general idea behind these *geometric kernels*

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598

is that they can be expressed through a manifold-based Fourier feature expansion:

$$k_{\text{geo}}(\boldsymbol{x},\, \boldsymbol{x}') = \sigma^2 \sum_{n=0}^{\infty} \Psi(\lambda_n) f_n(\boldsymbol{x}) f_n(\boldsymbol{x}')\,, \tag{13}$$

where $\sigma^2$ is a positive scaling factor, $(\lambda_i, f_i)$ are eigenpairs of the Laplace-Beltrami operator, and $\Psi(\lambda_i)$ denotes the generalized spectral measure. The generalized spectral measure is fundamental in kernel theory, providing the necessary and sufficient condition for a function to qualify as a valid kernel by *Bochner's Theorem* [10]. In essence, the spectral measure acts as a generalized form of the Fourier transform of a kernel function. For instance, if the spectral measure can be expressed as $d\Psi(\boldsymbol{s}) = S(\boldsymbol{s})d\boldsymbol{s}$, then $S(\boldsymbol{s})$ corresponds to the Fourier transform of the kernel, satisfying

$$k(\boldsymbol{x},\, \boldsymbol{x}') = \int S(\boldsymbol{s}) e^{2\pi i \boldsymbol{s}\cdot(\boldsymbol{x}-\boldsymbol{x}')} d\boldsymbol{s}\,, \tag{14}$$

if and only if the kernel can be expressed as

$$k(\boldsymbol{x},\, \boldsymbol{x}') = k(\boldsymbol{x} - \boldsymbol{x}')\,. \tag{15}$$

For a spherical surface, the eigenvalues of the Laplace-Beltrami operator are given by $\lambda_n = n(n+1)$, with the corresponding eigenfunctions being spherical harmonics of degree $n$. This structure allows the Matérn or SE geometric kernel functions to be formulated analytically as an infinite sum

$$k_{\text{geo},\nu}(\boldsymbol{x}, \boldsymbol{x}') = \sum_{n=0}^{\infty} \rho_\nu(n) c_n \mathcal{C}_n^{1/2} \left(\cos\left(\delta_{\text{g}}(\boldsymbol{x}, \boldsymbol{x}')\right)\right)\,, \tag{16}$$

$$k_{\text{geo},\infty}(\boldsymbol{x}, \boldsymbol{x}') = \sum_{n=0}^{\infty} \rho_\infty(n) c_n \mathcal{C}_n^{1/2} \left(\cos\left(\delta_{\text{g}}(\boldsymbol{x}, \boldsymbol{x}')\right)\right)\,, \tag{17}$$

$$\rho_\nu(n) = \frac{\sigma^2}{C_\nu} \left(\frac{2\nu}{l^2} + n(n+1)\right)^{-(\nu+1)}\,, \tag{18}$$

$$\rho_\infty(n) = \frac{\sigma^2}{C_\infty} e^{-\frac{l^2}{2} n(n+1)}\,, \tag{19}$$

where $\delta_{\text{g}}(\boldsymbol{x}, \boldsymbol{x}')$ represents the geodesic distance (11), $l$ denotes a length scale, $C_\nu$ and $C_\infty$ are normalization constants, and $\mathcal{C}_n^{1/2}$ are *Gegenbauer polynomials*. The coefficients $c_n$ are defined as

$$c_n = (2n+1) \frac{\Gamma(n+1)}{\Gamma(2)\Gamma(n+1)} \frac{\Gamma(3/2)}{(2\pi)^{3/2} \mathcal{C}_n^{1/2}(1)}\,. \tag{20}$$

A complete analysis of the derivations can be found in Appendix B of [18]. It is important to note that no closed-form expression for these kernels is known to the authors and that the series is truncated for use in practical applications. The truncated implementation of geometric kernels is available in the open-source library on GitHub [20]. These kernels share the same limitation as the geodesic kernel in that they only allow for a single length scale, not multiple, potentially restricting the modeling accuracy.

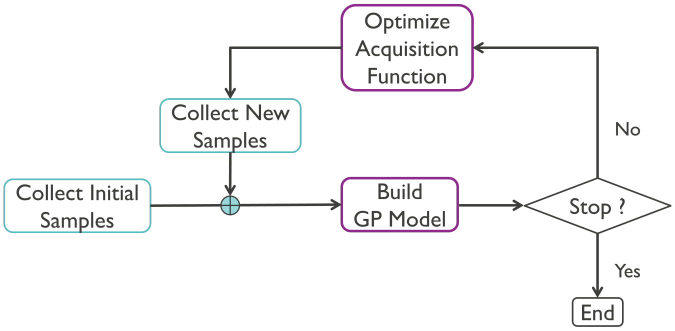


Fig. 2. Bayesian optimization workflow.

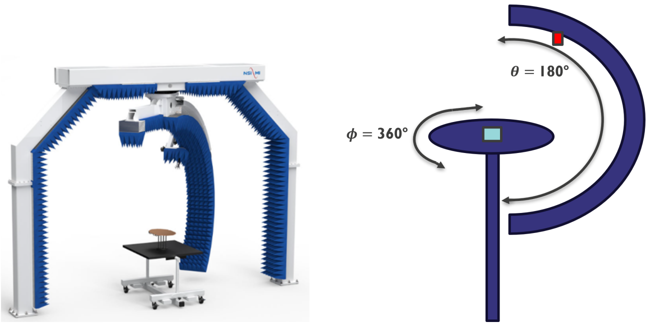
### D. 3D Euclidean Kernel with Cartesian Coordinates - 3DL2

The final kernel method we discuss uses Cartesian coordinates $\boldsymbol{x} = (x,\, y,\, z)$ within the standard Matérn (3) and SE (4) Euclidean kernels. Unlike previous methods, this approach falls into a different category, as it disregards the underlying spherical geometry and instead models directly in the ambient 3D Euclidean space that embeds the sphere. Therefore, it must be constrained a posteriori to the sphere, and while this does not pose any issues for pure modeling tasks, it introduces additional complexity when applied to BO for sequential sampling. For this reason, using 2D $(\theta,\, \phi)$-coordinates directly is often more intuitive. Similarly to the 2DL2 kernels, this method allows to define multiple length scales inside the distance function which can lead to performance improvements.

The main mathematical difference between this method (referred to as the *3DL2* method) and the previous two lies in the use of a chordal distance rather than a geodesic distance between points. This metric avoids discontinuities (unlike the 2DL2 method) but does not provide exact distances on the sphere itself. As shown in Fig. 1, the chordal distance diverges significantly from the geodesic distance for central angles greater than $\pi/2$. Although the two distances differ, they both increase monotonically with the central angle. Consequently, points that are farther apart on the spherical surface–and thus less correlated–are also less correlated in the 3D ambient Euclidean space. However, this relationship is specific for spherical surfaces and may not extend to other types of manifolds. An example of a manifold where these results fail is provided in Appendix A.

## IV. SEQUENTIAL SAMPLING

Once an appropriate kernel is selected, the next step is to use the GP model for sequential data collection. This section outlines the proposed sampling strategy, which combines BO with a motion-aware acquisition function to enable time- and sample-efficient exploration of the DUT's radiation pattern. This approach represents a novel application of BO to spherical far-field measurement setups, where physical movement incurs non-negligible time costs and traditional dense-grid methods are often slow.



(a) (b)

Fig. 3. Spherical measurement system: (a) picture [23] and (b) conceptual drawing. The DUT (green rectangle) is placed on a turntable capable of turning 360° while a measurement probe (red rectangle) can move along a semicircle.

### A. Bayesian Optimization

The typical workflow of BO is shown in Fig. 2. After building a GP using some initial data, a cheap-to-evaluate acquisition function is optimized to find new sampling positions. Commonly used acquisition functions include: probability of improvement (PoI), expected improvement (EI), and upper confidence bound (UCB) [21]. Here, the EI is selected since it balances exploration and exploitation without the need for extra hyperparameters, unlike UCB. Additionally, EI typically outperforms PoI in terms of sampling efficiency, as shown in [22]. The EI takes into account the magnitude of improvement and is defined as

$$EI(\boldsymbol{x}) = \mathbb{E}\left[\max(GP(\boldsymbol{x}) - f(\hat{\boldsymbol{x}}), 0)\right], \tag{21}$$

where $GP(\boldsymbol{x})$ is the prediction of the GP model at $\boldsymbol{x}$ and $f(\hat{\boldsymbol{x}})$ represents the current best objective value. Optimization of the EI yields the next sampling point, and the process repeats until a stopping criterion, such as a sampling budget, time constraint, or modeling error threshold, is reached.

### B. Motion-Aware Bayesian Optimization

Standard acquisition functions solely rely on the GP posterior to balance exploration and exploitation. However, they do not account for the travel time associated with collecting samples at different positions on the sphere. Since time-efficiency is an important objective of this research, a motion-aware acquisition function is employed, called motion-aware expected improvement (MEI) [24]. It is defined as

$$MEI(\boldsymbol{x}, \boldsymbol{p}_i) = \frac{EI(\boldsymbol{x})}{1 + C(\boldsymbol{x}, \boldsymbol{p}_i)}, \tag{22}$$

where $C(\boldsymbol{x}, \boldsymbol{p}_i)$ is the travel time required to reach point $\boldsymbol{x}$, starting from the current position $\boldsymbol{p}_i = (\theta_i, \phi_i)$ of the measurement probe.

The travel time is calculated based on the mechanical characteristics of the measurement setup. Specifically, for a rotational speed of $s_\theta$ along the polar angle and $s_\phi$ along the azimuthal angle, the travel time is given by

$$C(\boldsymbol{x}, \boldsymbol{p}_i) = \max\{\Delta t_\theta, \, \Delta t_\phi\}, \tag{23}$$

where

$$\Delta t_\theta = \frac{|\theta - \theta_i|}{s_\theta}, \qquad \text{and} \tag{24}$$

$$\Delta t_\phi = \frac{\min\bigl\{|\phi - \phi_i|, \, 2\pi - |\phi - \phi_i|\bigr\}}{s_\phi}. \tag{25}$$

Note that a minimum is taken in (25) to select the shortest path between clockwise and counterclockwise rotations.

These definitions of $\Delta t_\theta$ and $\Delta t_\phi$ are specifically for a measurement setup where the device is placed on a turntable that can rotate in two directions, along with a probe that can move up and down along the arc of a semicircle. An example of such a system is shown in Fig. 3. If a different type of system is used, (24) and (25) can be adapted as needed without affecting the general methodology. Furthermore, (23) assumes the possibility of simultaneously moving about both axes of rotation. If the motion of the system is limited to only one revolution at a time, the time cost (23) must be modified to

$$C(\boldsymbol{x}, \boldsymbol{p}_i) = \Delta t_\theta + \Delta t_\phi. \tag{26}$$

Finally, the new sampling point is determined by solving

$$\boldsymbol{p}_{i+1} = \arg\max_{\boldsymbol{x} \in X} MEI(\boldsymbol{x}) \tag{27}$$

which constitutes the final step of the loop in Fig. 2. This comprises the core of the proposed sampling strategy. By utilizing both a kernel, tailored towards spherical surfaces, and the MEI acquisition function (22), this strategy aims to balance the trade-off between locating the maximum field and minimizing the time cost to reach it. The optimization loop continues until a stop condition is met. Since the two main factors of interest in this study are the number of collected samples and the total sampling duration, one of these parameters will often serve as the stop condition.

## V. EVALUATION METRICS

To assess the performance and accuracy of the proposed modeling and sampling strategies, a set of metrics is defined on a large set of test samples at positions $X_s$. A similar sequential modeling and sampling strategy has been applied in [24] for cylindrical measurement setups. Therefore, similar metrics can be employed to ensure consistent comparisons across different studies.

First, the four different kernel strategies are evaluated by comparing the root mean squared error (RMSE) of the GP models against a simulated radiation pattern. This metric provides a direct measure of the model's accuracy in representing the field values on the spherical surface. However, since sampling is less likely in regions with low electric fields, and since such low field values often correspond to background noise, a constrained RMSE metric is introduced to more accurately evaluate model accuracy. This constrained RMSE considers only field values above a defined threshold:

$$X_f = \bigl\{\boldsymbol{x} \in X_s \, : \, f(\boldsymbol{x}) > \gamma \max_{\boldsymbol{x} \in X_s} f(\boldsymbol{x})\bigr\}, \tag{28}$$

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598

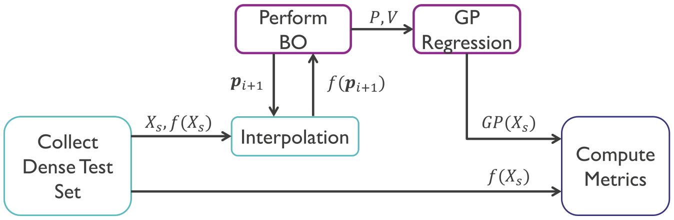
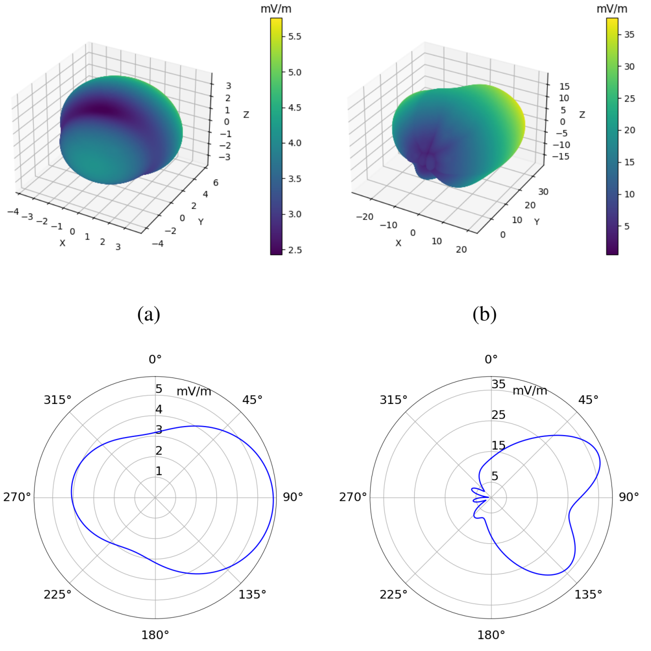


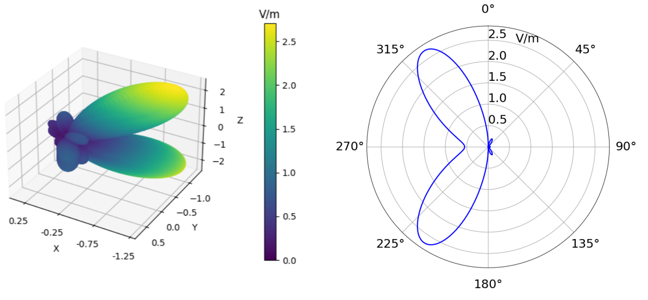
Fig. 4. Simulation method for evaluating different BO strategies.



(a) (b)

(c) (d)

Fig. 5. Open-ended microstrip line: Magnitude of the electric field, simulated on a spherical surface at 1 m distance, at (a) 645 MHz and (b) 4642 MHz. Figures (c) and (d) show corresponding two-dimensional cuts of the field in the $xy$-plane for (a) and (b), respectively.



(a) (b)

Fig. 6. Antenna array: Magnitude of the electric field at 5 GHz, simulated on a spherical surface at 1 m distance. (a) Full 3D electric field pattern. (b) Two-dimensional cut of the field magnitude at $\phi = 50°$, corresponding to the azimuthal direction of strongest radiation.

where $\gamma \in [0,\, 1)$. Using this set of test points, the constrained RMSE is computed as

$$\mathrm{RMSE}_\gamma = \sqrt{\sum_{\boldsymbol{x} \in X_f} \frac{\left(GP(\boldsymbol{x}) - f(\boldsymbol{x})\right)^2}{|X_f|}}\,. \tag{29}$$

Second, as the sampling approach is designed to locate the maximum of the electric field, its effectiveness is further evaluated using the *regret* $R$

$$R = \frac{\max_{\boldsymbol{x} \in X_s} f(\boldsymbol{x}) - \max_{\boldsymbol{p} \in P_i} f(\boldsymbol{p})}{\max_{\boldsymbol{x} \in X_s} f(\boldsymbol{x})}\,, \tag{30}$$

where $P_i$ represents the set of sampled positions. The regret quantifies the relative error of how closely the sampling approach comes to sampling the exact maximum and should therefore be as small as possible.

One drawback of this metric is its tendency to drop significantly when approaching local optima. Therefore, the prediction accuracy of the GP models is assessed simultaneously, as a low regret $R$ combined with a low $\mathrm{RMSE}_\gamma$ generally indicates optimal results.

## VI. RESULTS

In typical sequential sampling scenarios aimed at identifying maximum field intensity, measurements are acquired directly at positions selected by the acquisition function. However, in real-world applications, environmental noise or other time-dependent EM fluctuations can complicate direct comparisons of GP models and BO strategies when performed sequentially. To mitigate these effects, a simulation-based approach is adopted: the EM field is first simulated at a fixed, dense set of positions, $X_s$, producing a reference test set, $\{X_s, f(X_s)\}$. Sequential sampling (BO) is then performed virtually, with field values at each position $\boldsymbol{p}_i$ obtained by interpolation from the test set. This process ensures consistency in modeling, sampling behavior, and metric evaluation, with the test set serving as ground truth. Fig. 4 illustrates a schematic overview of this evaluation approach.

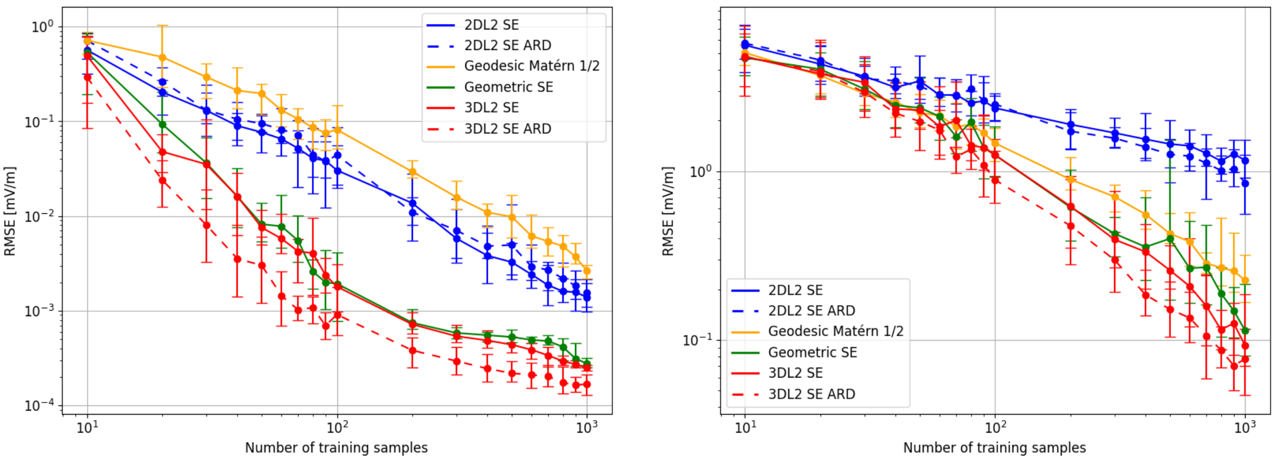
Two datasets are used to evaluate the proposed methodology. The first is based on simulated EM fields generated by an open-ended microstrip line, modeled using Keysight PathWave EM Design (EMPro) and sampled on a spherical grid with a radius of 1 m around the DUT. The microstrip is 30 mm in length and 2 mm in width, and is placed on a 100 mm × 100 mm PCB. This dataset is used to compare different kernel models and their impact on BO performance. Fig.5 shows the magnitude of the total electric field at frequencies of 645 MHz and 4642 MHz, highlighting the spatial complexity of the fields and the modeling challenges they present.

The second dataset involves simulated far-field electric field values generated by an antenna array at 5 GHz. This example is used to further demonstrate the sequential sampling strategy, specifically the practical benefits of motion-aware BO. While no kernel comparison is performed on this dataset, it provides additional validation of the proposed sampling method under a different, more complex radiation scenario. Fig. 6 shows the electric field magnitude for the antenna array example.

### A. Gaussian Process Models: Comparing Different Kernels

The four kernel methods discussed in Section III are employed to train GP models on two microstrip datasets (at

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598



(a) (b)

Fig. 7. Open-ended microstrip line: RMSE of GP model as a function of sample size for different kernels at (a) 645 MHz and (b) 4642 MHz. The solid lines represent kernels that use only 1 length scale, while the dashed lines represent kernels that use *automatic relevance detection* (ARD). ARD defines one length scale to each dimension of the input.

645 MHz and 4642 MHz), using a varying number of data samples. For each sample size, the RMSE is calculated and averaged over ten independent runs to ensure statistical robustness. In this procedure, the samples are generated randomly according to

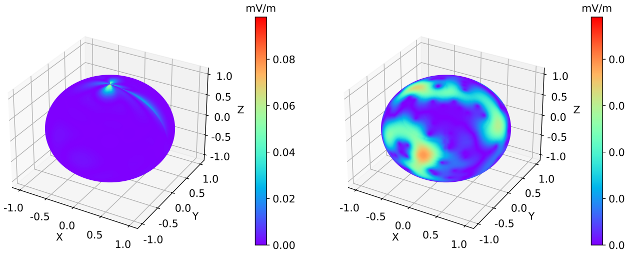
$$\theta = \arccos(2u - 1) \tag{31}$$

$$\phi = 2\pi v \tag{32}$$

where $u$ and $v$ are uniformly distributed in $[0, 1]$. This transformation prevents oversampling at the poles of the sphere by ensuring a uniform surface area distribution.

*1) Kernel Smoothness:* As stated above, the geodesic kernel is restricted to the Matérn $1/2$ implementation, whereas the other methods can employ any member of the Matérn family, including the SE kernel. Therefore, we first compare the performance of the Matérn $1/2$ kernel with that of the SE kernel for the three methods capable of implementing the latter, i.e., 2DL2, geometric, and 3DL2. In this section, only the results are discussed, but a detailed comparison between the Matérn $1/2$ and SE kernel is provided in Appendix B. Results indicate a significant performance gain when using the SE kernel across these methods. This improvement can be attributed to the fact that far-field radiation–being governed by Maxwell's equations–is generally continuously differentiable. Since the Matérn $1/2$ kernel models non-continuously differentiable functions, the SE kernel, which assumes infinite smoothness, provides a better fit for this application. Therefore, in the following analysis, we limit our discussion to the SE variant of the 2DL2, geometric, and 3DL2 methods, along with the geodesic Matérn $1/2$ kernel for further exploration.

*2) Number of Length Scales:* Next, we evaluate the impact of using multiple length scales (one for each input dimension) for the 2DL2 SE and 3DL2 SE kernels. This approach, referred to as *automatic relevance determination* (ARD) in Gaussian process literature, enhances modeling power, with a significant improvement observed for the 3DL2 SE kernel. For the 2DL2 SE kernel, the improvement is relatively minor, which can likely be attributed to inaccuracies introduced by the ill-defined $L_2$ norm in spherical coordinates.



(a) (b)

Fig. 8. Open-ended microstrip line: absolute error made by training GP models trained on 200 samples at 645 MHz using (a) the 2DL2 SE kernel and (b) the geodesic Matérn $1/2$ kernel, evaluated on a dense test set.

*3) General Comparison:* The comparative performance of the 2DL2 SE, geodesic Matérn $1/2$, geometric SE, and 3DL2 SE kernels is presented in Fig. 7, which shows the averaged RMSE values as a function of sample size. For the 2DL2 SE and 3DL2 SE kernels, results are shown with and without ARD. Notably, the 2DL2 SE and geodesic Matérn $1/2$ kernels exhibit the poorest performance.

This is expected for the 2DL2 SE kernel due to the discontinuity in its $(\theta, \phi)$-based $L_2$ norm. Fig. 8(a) further illustrates this by showing the absolute error of the model with 2DL2 SE kernel on a test set of the electric field at 645 MHz. The error is significantly higher near the poles and along the prime meridian compared to other regions, and hence the model performs better when inputs are outside these problematic regions.

In contrast, the absolute error of the geodesic Matérn $1/2$ model, shown in Fig. 8(b), exhibits a more random error distribution, albeit in the same order of magnitude as the 2DL2 SE model. The reason for this poor performance compared to the geometric SE or 3DL2 SE kernel stems from the previously

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598

mentioned limited differentiability of the Matérn $1/2$ kernel.

Since the electric fields in the $4642$ MHz dataset are less smooth, the RMSE difference between Matérn $1/2$ and SE kernels is smaller, but they remain present (see Fig. 7). This leads to the conclusion that both the 2DL2 SE and geodesic Matérn $1/2$ kernels are flawed and less suited for modeling far-field radiation on spherical surfaces.

Fig. 7 also shows that the performance difference between the geometric SE and 3DL2 SE kernels without ARD is negligible. However, including ARD provides a modest improvement by enabling finer modeling of the electric field's spatial variations. The geometric kernel only has one length scale, therefore defining a more restricted behavior in both $\theta$- and $\phi$-direction.

Additionally, the work in [25] demonstrates that using geometric kernels on $n$-dimensional non-Euclidean manifolds (such as a sphere) results in similar posterior contraction rates as regular kernels on the $(n+1)$-dimensional ambient Euclidean space in which the manifold is embedded. This implies that both methods converge similarly towards the true field values as the sample size approaches infinity. This finding aligns with our results, as the 3DL2 SE method, with or without ARD, shows comparable RMSE reduction with increased training points.

In conclusion, we can state that there are three important kernel features when modeling EM fields with GPs: smoothness, the number of length scales, and a correct distance metric. The geodesic method is therefore less suited in this problem setting due to its restriction to the Matérn $1/2$ kernel. The 2DL2 method ignores the spherical symmetries altogether and employs an ill-defined distance metric, resulting in poor performance at the prime meridian and poles of the sphere. Finally, the 3DL2 and geometric method share the benefits of a smooth kernel implementation and a valid distance metric. However, the 3DL2 method offers an additional benefit through its use of multiple length scales. This advantage may diminish when modeling on different types of manifolds, as discussed in Section III-D and further elaborated in Appendix A. An overview of these results is given in Table I.

### B. The Sequential Sampling Method

The objective is to identify the maximum field value in a data- and time-efficient manner. To this end, the best-performing configuration of each kernel method is selected to perform sequential sampling on the open-ended microstrip dataset at $645$ MHz and $4642$ MHz. Here, we assess whether the insights gained from modeling extend to the sequential sampling task and evaluate whether the motion-aware expected improvement (MEI) acquisition function outperforms the standard expected improvement (EI).

In each experiment, sequential sampling is conducted over $100$ iterations, starting from the same initial measurement position $(\theta, \phi) = (\pi/2,\, \pi)$. Each configuration is repeated for $10$ independent runs to capture the variability of the sampling process. In each run, the constrained RMSE with $\gamma = 0.80$ and regret $R$ are computed as functions of sample count and travel time. Travel times are calculated using (23) with antenna movement speeds of $s_\theta = 0.10$ rad/s and $s_\phi = 0.21$ rad/s. Summarized results are presented in Table II for the $645$ MHz dataset and in Table III for the $4642$ MHz dataset. Performance is evaluated based on how efficiently each method reduces the $80\%$-RMSE and regret below predefined thresholds, both in terms of the number of samples required and the total measurement time.

TABLE I
FEATURE COMPARISON FOR DIFFERENT KERNEL METHODS

| | 2DL2 | Geodesic | Geometric | 3DL2 |
|---|---|---|---|---|
| Smooth kernel possible | ✓ | X | ✓ | ✓ |
| Geodesic distance metric | X | ✓ | ✓ | X |
| ARD possible | ✓ | X | X | ✓ |

In addition to the radiating microstrip scenarios, a third experiment is conducted on a more directive dataset: the far-field radiation of an antenna array at $5$ GHz. The array consists of $16$ simple dipole antennas, each $15$ mm long and spaced $30$ mm apart (center-to-center), tuned in phase with each other. The radiation pattern is simulated using the NEC2 software. This case serves as a practical test of the sequential sampling strategy in a more intricate EM environment. As with the previous cases, MEI and EI are compared, and results are obtained from $10$ repeated runs of $100$ samples, summarized in Table IV. This final example highlights the applicability of the proposed methodology beyond simple radiation patterns and demonstrates the potential of motion-aware BO in more realistic measurement scenarios.

For the two microstrip datasets at $645$ MHz and $4642$ MHz, the geometric SE and 3DL2 SE kernels demonstrate the best overall modeling performance. They require the fewest samples to reach the predefined $80\%$-RMSE threshold and consistently achieve the $1\%$ regret target with fewer samples compared to the other kernels. In contrast, the 2DL2 SE kernel converges more slowly, while the geodesic Matérn $1/2$ kernel fails to reach the RMSE threshold within $100$ samples.

Travel-time results for these datasets follow a similar trend. At $645$ MHz, the 3DL2 SE kernel achieves the $80\%$-RMSE threshold in the shortest amount of time, while the geometric SE kernel performs best at $4642$ MHz. Both methods also exhibit the lowest travel times to reach the $1\%$ regret target across frequencies. Although the geodesic Matérn $1/2$ kernel occasionally reaches low regret values quickly, its inconsistent threshold attainment makes it a less reliable choice. Overall, the geometric SE and 3DL2 SE kernels provide the most robust and efficient performance across both metrics in this experiment.

An analysis of the acquisition functions reveals that the poorer-performing kernels, namely the 2DL2 SE and geodesic Matérn $1/2$, tend to prioritize exploitation over exploration. This behavior likely stems from limited predictive accuracy, which leads to higher sample counts and increased travel time when attempting to reach the $80\%$-RMSE threshold. These kernels often direct sampling efforts toward known high-field regions rather than exploring new areas, reducing their effectiveness in global field modeling.

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598

Comparing the two acquisition functions, EI and MEI, shows that MEI significantly reduces the average time required to meet both thresholds. While MEI does not necessarily reduce the number of samples required, it was not designed to do so; instead, it effectively minimizes travel time. Additionally, MEI often results in lower minimal regret values. This is likely because EI, over time, may direct samples toward uncertain regions that are far from the maximum field, whereas MEI continues focusing locally due to the implicit penalization of excessive movement.

In the additional experiment on the antenna array at 5 GHz, new insights are gained. The 2DL2 SE kernel performs much more competitively in this scenario, achieving the $1\%$ regret threshold with the fewest samples when paired with the MEI acquisition function. Moreover, it consistently reaches the $80\%$-RMSE target, whereas the geometric SE kernel does not. This suggests that for highly directional radiation patterns, where the field is concentrated in localized lobes, the limitations of the 2DL2 SE kernel become less critical. However, the discontinuity in its ($\theta$, $\phi$)-based $L_2$ norm may still lead to inconsistent results if the high-amplitude lobes are located near this discontinuity.

Although the geometric SE kernel performed well in the monopole cases, its modeling performance in the antenna array experiment degrades slightly. However, when combined with MEI, it still achieves the shortest average time to reach $1\%$ regret, highlighting its effectiveness in minimizing travel duration in scenarios where model accuracy is acceptable.

Across all experiments, the 3DL2 SE kernel remains the most consistent performer, offering both strong modeling accuracy and efficient convergence in sequential sampling. The 2DL2 SE kernel proves to be a valuable alternative, particularly in directional field scenarios. Meanwhile, the MEI acquisition function consistently improves time efficiency without sacrificing modeling quality, reinforcing its practical relevance in constrained measurement setups.

To contextualize the time savings offered by the proposed method, we estimate the total duration required to locate the maximum field value using a uniform dense sampling strategy. For a $20 \times 40$ spherical grid, with travel speed parameters $s_\theta = 0.10$ rad/s and $s_\phi = 0.21$ rad/s, the full grid sampling would take approximately $613.5$ s. In contrast, the best-performing BO configurations reach the $1\%$ regret threshold in significantly less time. The observed reductions range from $86.1\%$ to $96.7\%$, demonstrating the substantial measurement time efficiency enabled by the BO-based sequential sampling framework.

## VII. CONCLUSION

This work presents a novel strategy for data-efficient electromagnetic field characterization through motion-aware sequential sampling on spherical surfaces using Bayesian optimization (BO). Unlike conventional grid-based measurements, the proposed approach leverages Gaussian Process (GP) modeling to iteratively select informative sampling points, thereby reducing both the number of required measurements and the total acquisition time. The method is validated on simulated far-field electric fields of two antenna types: an open-ended microstrip line and a directional antenna array.

A detailed comparison of four GP kernel constructions identifies the 3DL2 SE kernel as the most reliable across different scenarios, providing accurate field modeling with fewer samples and shorter acquisition times. The geometric SE kernel offers competitive performance but is limited by its single length scale formulation, while the 2DL2 SE kernel shows case-dependent accuracy. The latter performs well in directional field distributions with prominent lobes but is less reliable for more isotropic or smooth patterns. In contrast, the geodesic Matérn $1/2$ kernel proves insufficient for practical use.

To further improve applicability, a motion-aware expected improvement (MEI) acquisition function is derived, which significantly reduces total measurement time compared to standard EI. This confirms its value in real-world measurement setups where time efficiency is critical.

In summary, BO-driven sequential sampling with GP models is developed as an effective alternative to dense spherical sampling for electromagnetic measurements. By identifying the most suitable kernel choices and incorporating motion-aware acquisition, the proposed method achieves an estimated $86.1\%$ to $96.7\%$ reduction in measurement time for locating the maximum field value compared to a uniform grid sampling strategy. Future research will explore multi-length scale geometric kernels and enhanced acquisition strategies to extend this framework to broader spherical and non-Euclidean domains.

## APPENDIX A
## EXAMPLE MANIFOLD WHERE 3D EXTRINSIC KERNEL FAILS

As stated in Section III-D, the extrinsic three-dimensional Euclidean distance (3DL2) kernel method works effectively in our problem setting due to a similar distance correlation compared to the geodesic distance on a spherical surface. Specifically, as the $L_2$ distance increases, so does the geodesic distance, maintaining a positive correlation. However, as noted by the authors in [25], intrinsic geometric kernels can achieve higher modeling performance than extrinsic methods on manifolds where this correlation does not hold. One such example is the one-dimensional dumbbell manifold shown in Fig. 9. The colormap illustrates values of the Matérn $5/2$ kernel (i.e., the correlation) relative to a reference point, emphasizing the differences between intrinsic geometric and extrinsic $L_2$ norm kernels. For instance, when a sample is taken near the black dot on the figure, the intrinsic geometric Gaussian process (GP) correlates nearby points along the manifold, while the $L_2$ norm GP correlates points that are distant on the manifold but spatially close in 2D Euclidean space. Consequently, geometric kernels may offer improved modeling on complex shapes where geodesic and Euclidean distances lack a positive correlation.

TABLE II
OPEN-ENDED MICROSTRIP LINE AT 645 MHZ - PERFORMANCE COMPARISON OF SEQUENTIAL SAMPLING STRATEGIES

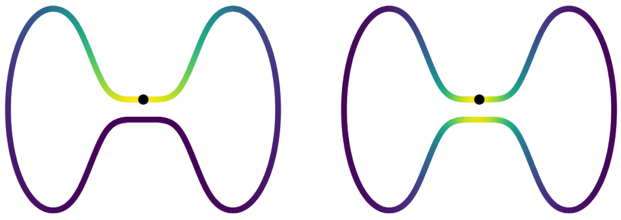
| 645 MHz | 2DL2 SE ARD | | Geodesic Matérn 1/2 | | Geometric SE | | 3DL2 SE ARD | |
|---|---|---|---|---|---|---|---|---|
| | EI | MEI | EI | MEI | EI | MEI | EI | MEI |
| Av. samples for 80%-RMSE = 1e-2 mV/m | 41 | 46 | X | X | 29 | 21 | **18** | 27 |
| Av. samples for regret = 1% | 11 | 9 | 42 | 15 | 10 | 8 | 10 | **7** |
| Av. time [s] for 80%-RMSE = 1e-2 mV/m | 477.28 | 368.00 | X | X | 242.98 | 108.38 | 155.80 | **88.55** |
| Av. time [s] for regret = 1% | 92.83 | 54.96 | 146.32 | 27.57 | 58.00 | **20.36** | 58.73 | 38.88 |

TABLE III
OPEN-ENDED MICROSTRIP LINE AT 4642 MHZ - PERFORMANCE COMPARISON OF SEQUENTIAL SAMPLING STRATEGIES

| 4642 MHz | 2DL2 SE ARD | | Geodesic Matérn 1/2 | | Geometric SE | | 3DL2 SE ARD | |
|---|---|---|---|---|---|---|---|---|
| | EI | MEI | EI | MEI | EI | MEI | EI | MEI |
| Av. samples for 80%-RMSE = 1 mV/m | 79 | 46 | X | X | 51 | **25** | 45 | 28 |
| Av. samples for regret = 1% | 48 | 36 | 32 | X | 60 | **19** | 38 | **19** |
| Av. time [s] for 80%-RMSE = 1 mV/m | 902.60 | 418.11 | X | X | 447.69 | **146.93** | 361.07 | 233.53 |
| Av. time [s] for regret = 1% | 549.26 | 281.97 | 94.80 | X | 567.19 | 112.23 | 361.07 | **85.39** |

TABLE IV
ANTENNA ARRAY AT 5 GHZ - PERFORMANCE COMPARISON OF SEQUENTIAL SAMPLING STRATEGIES

| 5 GHz | 2DL2 SE ARD | | Geodesic Matérn 1/2 | | Geometric SE | | 3DL2 SE ARD | |
|---|---|---|---|---|---|---|---|---|
| | EI | MEI | EI | MEI | EI | MEI | EI | MEI |
| Av. samples for 80%-RMSE = 1e-1 V/m | 58 | 75 | X | X | X | 47 | **24** | 45 |
| Av. samples for regret = 1% | 59 | **23** | 82 | X | 68 | 31 | 37 | 46 |
| Av. time [s] for 80%-RMSE = 1e-1 V/m | 539.97 | 295.74 | X | X | X | 198.58 | 178.04 | **166.11** |
| Av. time [s] for regret = 1% | 514.04 | 82.03 | 456.31 | X | 358.00 | **65.09** | 326.13 | 122.93 |



(a) (b)

Fig. 9. Matérn kernel evaluated on a dumbell-shaped line (a) intinsically and (b) extrinsically. The color implicitly represents the distance from the reference point (black dot) with yellow being closer and purple being further away from it [25].

## APPENDIX B
## COMPARISON BETWEEN MATÉRN 1/2 AND SQUARED EXPONENTIAL KERNEL

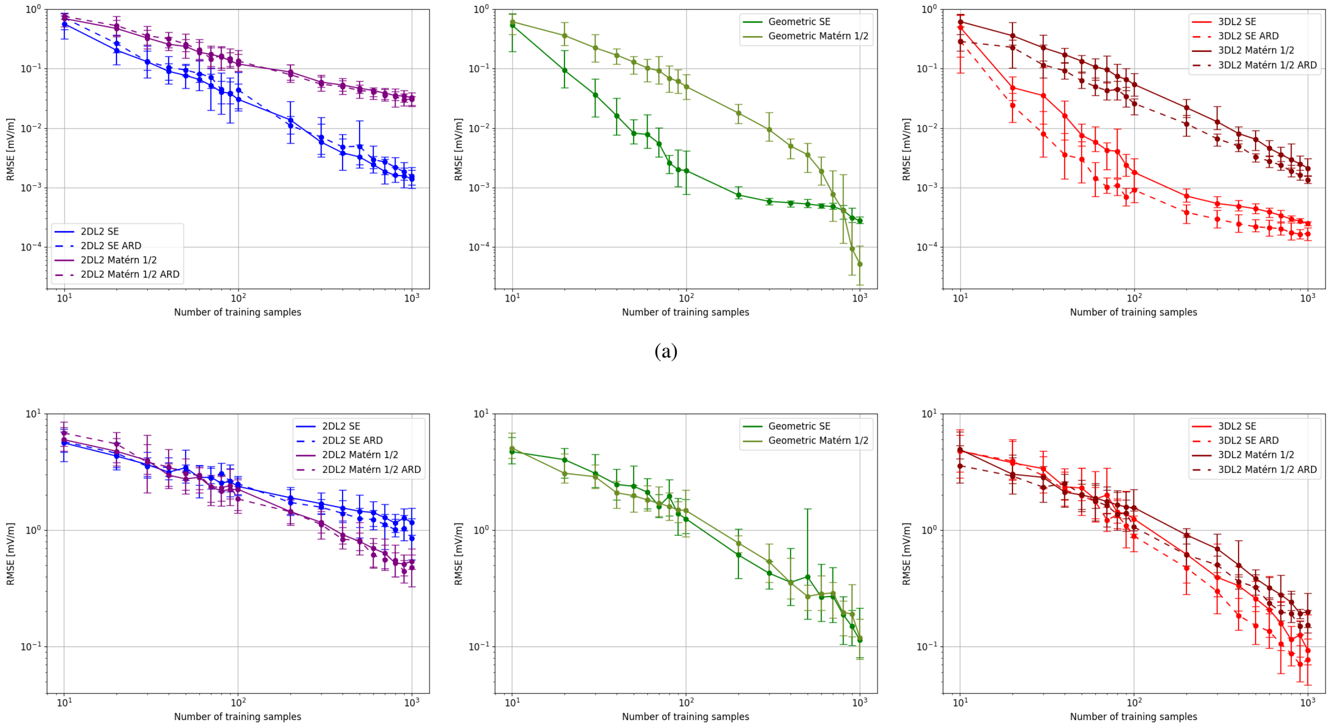
The smoothness of a Gaussian process (GP) kernel is an important feature to consider when fitting a model to a dataset. The Matérn kernels have a half-integer smoothness parameter, $\nu$, which determines the level of differentiability of the GP model's predictions, making them $\lceil \nu \rceil - 1$ times continuously differentiable. The squared exponential (SE) kernel is a limiting case of the Matérn kernel as $\nu \to \infty$, making it infinitely continuously differentiable. In practice, $\nu$ is rarely set above $5/2$ for Matérn kernels, as higher values produce results nearly indistinguishable from the SE kernel.

In this paper, we work with far-field electromagnetic (EM) fields, specifically electric fields, which are inherently continuously differentiable. This motivates the use of a smooth kernel. Since the geodesic kernel method is limited to the Matérn $1/2$ kernel to assure positive semi-definiteness, we perform tests to assess the performance differences between the non-smooth Matérn $1/2$ and the infinitely smooth SE kernel for far-field data. These tests are conducted as described in Section VI and cover the 2DL2, geometric, and 3DL2 methods explained in Section II-A.

The results, shown in Fig. 10, compare the 645 MHz and 4642 MHz datasets of the total electric field magnitude on a sphere at 1 m from an open-ended microstrip line. Across all methods, the SE kernel generally achieves lower RMSE values with fewer training points than the Matérn $1/2$ kernel. This suggests that the SE kernel is better suited for far-field EM data, as it captures the smooth spatial behavior and continuity typical for far-field radiation. However, the performance difference is frequency dependent, with a smaller RMSE gap between both kernels at 4642 MHz than at 645 MHz. This can be a result of the increased number of lobes in the field pattern at higher frequencies, which reduces the smoothness of the modeling landscape.

In summary, the SE kernel is theoretically and practically a

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598



(a)

(b)

Fig. 10. Open-ended microstrip line: comparison of Matérn $1/2$ and squared exponential kernel at two frequencies: (a) 645 MHz and (b) 4642 MHz. The dashed lines represent kernels that use only one length scale, while the solid lines represent kernels that use *automatic relevance detection* (ARD). ARD defines one length scale for each dimension of the input.

better fit for modeling far-field EM radiation patterns due to its smoothness assumptions, which align well with the continuity of EM fields in the far-field regime.

## REFERENCES

[1] K. Armstrong, *The physical basis of EMC*. Nutwood UK Ltd., 2010, pp. 61–77.

[2] N. Mutonkole and D. de Villiers, “Multivariate adaptive sampling of parameterized antenna responses,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 3, pp. 1073–1080, 2017.

[3] J. N. Fuhg, A. Fau, and U. Nackenhorst, “State-of-the-art and comparative review of adaptive sampling methods for kriging,” *Archives of Computational Methods in Engineering*, vol. 28, pp. 2689–2747, 2021.

[4] Z. Zhou, Z. Wei, J. Ren, *et al.*, “Bayesian-inspired sampling for efficient machine-learning-assisted microwave component design,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 72, no. 2, pp. 996–1007, 2024.

[5] P. Lagouanelle, F. Freschi, and L. Pichon, “Adaptive sampling for fast and accurate metamodel-based sensitivity analysis of complex electromagnetic problems,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 65, no. 6, pp. 1820–1828, 2023.

[6] C. Yang, T. Monopoli, S. Götschel, X. Wu, F. Grassi, and C. Schuster, “Adaptive on-the-fly scan method for fast and efficient planar near-field acquisition,” *IEEE Antennas and Wireless Propagation Letters*, vol. 24, no. 8, pp. 2143–2147, 2025.

[7] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, “Taking the human out of the loop: A review of Bayesian optimization,” *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016.

[8] F. Hong, P. Wei, J. Fu, and M. Beer, “A sequential sampling-based Bayesian numerical method for reliability-based design optimization,” *Reliability Engineering & System Safety*, vol. 244, p. 109 939, 2024, ISSN: 0951-8320.

[9] S. Greenhill, S. Rana, S. Gupta, P. Vellanki, and S. Venkatesh, “Bayesian optimization for adaptive experimental design: A review,” *IEEE Access*, vol. 8, pp. 13 937–13 948, 2020.

[10] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. The MIT Press, 2006, vol. 2.

[11] P. Robinson, “The sphere is not flat,” *The American Mathematical Monthly*, vol. 113, no. 2, pp. 171–173, 2006.

[12] F. Pearson, *Map Projections: Theory and Applications*. CRC press, 1990.

[13] A. Feragen, F. Lauze, and S. Hauberg, “Geodesic exponential kernels: When curvature and linearity conflict,”

This article has been accepted for publication in IEEE Transactions on Antennas and Propagation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAP.2025.3645598

in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015.

[14] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi, “Kernel methods on Riemannian manifolds with Gaussian RBF kernels,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 12, pp. 2464–2477, 2015.

[15] M. Kumru and E. Özkan, “3D extended object tracking using recursive Gaussian processes,” in *2018 21st International Conference on Information Fusion (FUSION)*, IEEE, 2018, pp. 1–8.

[16] A. Feragen and S. Hauberg, “Open problem: Kernel methods on manifolds and metric spaces. what is the probability of a positive definite geodesic exponential kernel?” In *Conference on Learning Theory*, PMLR, 2016, pp. 1647–1650.

[17] J. Jeong, M. Jun, and M. G. Genton, “Spherical process models for global spatial statistics,” *Statistical science: a review journal of the Institute of Mathematical Statistics*, vol. 32, no. 4, p. 501, 2017.

[18] V. Borovitskiy, A. Terenin, P. Mostowsky, and M. Deisenroth, “Matérn Gaussian processes on Riemannian manifolds,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 426–12 437, 2020.

[19] I. Azangulov, A. Smolensky, A. Terenin, and V. Borovitskiy, “Stationary kernels and Gaussian processes on Lie groups and their homogeneous spaces I: The compact case,” *Journal of Machine Learning Research*, vol. 25, no. 280, pp. 1–52, 2024.

[20] P. Mostowsky, V. Dutordoir, I. Azangulov, *et al.*, “The geometrickernels package: Heat and Matérn kernels for geometric learning on manifolds, meshes, and graphs,” *arXiv preprint arXiv:2407.08086*, 2024.

[21] J. Wilson, F. Hutter, and M. Deisenroth, “Maximizing acquisition functions for Bayesian optimization,” in *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31, Curran Associates, Inc., 2018.

[22] W. Gan, Z. Ji, and Y. Liang, “Acquisition functions in Bayesian optimization,” in *2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*, 2021, pp. 129–135.

[23] NSI-MI Technologies, “Spherical near-field vertical arch measurement system” NSI-MI.com Accessed: Sept. 25, 2024. [Online.] Available: https://www.nsi-mi.com/system-solutions/near-field-systems/spherical-near-field-scanner-systems/spherical-nf-vertical-arch-measurement-system..

[24] F. Garbuglia, T. Claeys, I. Couckuyt, D. Deschrijver, and T. Dhaene, “Bayesian active learning for radiation pattern sampling over cylindrical surfaces,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 64, no. 5, pp. 1391–1398, 2022.

[25] P. Rosa, V. Borovitskiy, A. Terenin, and J. Rousseau, “Posterior contraction rates for Matérn Gaussian processes on Riemannian manifolds,” in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., vol. 36, Curran Associates, Inc., 2023, pp. 34 087–34 121.