Bayesian Active Learning for Radiation Pattern Sampling Over Cylindrical Surfaces

Federico Garbuglia ⓘ, Tim Claeys ⓘ, *Member, IEEE*, Ivo Couckuyt ⓘ, *Member, IEEE*,
Dirk Deschrijver ⓘ, *Senior Member, IEEE*, and Tom Dhaene ⓘ, *Senior Member, IEEE*

Abstract—In this article, a new motion-aware sampling strategy (MASS) is presented to speed up the measurement of radiation patterns around cylindrical surfaces. Differently from preexisting sampling techniques, the MASS directly chooses positions that reduce the overall travel time of the field antenna, rather than minimizing the total number of samples. The proposed strategy employs a Gaussian process model that is adapted to the field over a cylindrical surface. Moreover, a new acquisition function for Bayesian active learning is developed in order to efficiently search the peaks of the measured field and predict their values. Next, the proposed strategy is tested on the experimental data from a radiation pattern of a comb generator. Finally, the results are compared to standard grid sampling and Bayesian optimization strategies.

Index Terms—Active learning, Bayesian optimization (BO), design of experiments (DoE), electromagnetic (EM) compatibility, radiation pattern.

I. INTRODUCTION

NOWADAYS, electronic devices with wireless capabilities are ubiquitous and new ones are being produced at an accelerating pace. Despite the availability of simulators, physical measurements are essential to check the device compliance with radiated emission limits. In particular, the main peaks of the emitted field need to be identified by measuring the emitted electromagnetic (EM) waves on a cylindrical surface in a semianechoic chamber (SAC) using a receiving antenna. In order to accelerate this operation, the device under test (DUT) is usually moved with software-controlled motors. Thereby, the consecutive positions of the receiving antenna can be selected via a sampling strategy.

In recent years, several adaptive sampling strategies have been presented to reduce the time cost of near-field measurements on planar surfaces [1]–[3]. These techniques aim at minimizing the total number of samples by selecting the positions that carry the most relevant information about the field. Their efficacy lies in the assumption that the main contribution to the total time cost is the measurement time introduced by the instrumentation, rather than the travel time of the antenna. Unfortunately, the travel time is substantially high when the antenna speed is low, relative to the scanning area. This is often the case when sampling 3-D cylindrical surfaces around the DUT.

The problem of reducing the travel time has been addressed by line-based sampling strategies such as [4]–[6]. In particular, an active learning strategy built on Bayesian optimization (BO) has been proposed in [5]. This approach updates a Gaussian process (GP) model of the emitted field in a sequential manner and selects new positions that are most likely to present a field maximum. Then, additional positions are sampled along the way of the antenna, thus increasing the number of samples without adding travel delay. This contribution presents a new Bayesian active learning policy denoted as motion-aware sampling strategy (MASS). Firstly, this strategy employs a GP model that is adapted to cylindrical sampling surfaces. Secondly, time efficiency is improved over [5] by selecting positions that present the best tradeoff between travel cost and the probability of finding the field maximum. As a result, the new strategy minimizes the total sampling duration while focusing on peaks of the measured field.

The rest of the article is organized as follows. Section II states the main goal of the sampling and provides an overview on the different types of strategies. Subsequently, Section III analyzes the learning model adopted in the new motion-aware strategy. Section IV describes the MASS procedure and its main functional blocks. Then, suitable sampling metrics are defined in Section V, while Section VI demonstrates the effectiveness of the MASS compared to other sampling strategies. Finally, Section VII concludes the article.

II. GOAL STATEMENT

Let $X \subset \mathbb{R}^2$ be the domain of all possible sampling positions. Let the vector $\boldsymbol{x} = [h, \phi] \in X$ define the position of the antenna on a cylindrical surface around the DUT, where $h, \phi \in \mathbb{R}$ are the height and the azimuth of the antenna, respectively. The radiated field can be described by a function f over X: $f = f(\boldsymbol{x})$. Then, measuring the field of interest means collecting a set of positions in cylindrical coordinates $P = \{\boldsymbol{p}_n = [h_n, \phi_n], \boldsymbol{p}_n \in X\}_{n=1}^N$ and a set of measured field

values $V = \{v_n = f(\boldsymbol{p}_n), \boldsymbol{p}_n \in P\}$, where n is the sample index and N is the total number of samples. Additionally, let $T = \{t_n : \boldsymbol{x}(t_n) = \boldsymbol{p}_n, t_n \in \mathbb{R}_+, \boldsymbol{x} \in X\}_{n=1}^N$ be the set of sampling instants, where t_n represents the time at which the antenna reaches the position $\boldsymbol{p}_n$. In the considered measurement setting, the main goal of the sampling strategy is to choose P that maximizes the information about the absolute maximum and the local maxima of f, while limiting the total duration of the sampling t_N.

A. Related Work

The definition of the sampling strategy is in essence a design of experiments (DoE) problem [7]. In fact, a sampling strategy has to identify which positions are most relevant for the characterization of the function under test. The first task of DoE is to gather some information over the entire measurement space, known as *exploration* in literature. Exploration focuses on filling the function domain as evenly as possible for a given amount of sampling points.

The most immediate way to perform exploration is using a one-shot strategy, such as simple grid search or Latin hypercube design [8]. In one-shot strategies, a predefined number of samples is spread across the domain, without considering any observation of f. On one hand, this approach allows to compute all the positions upfront, thus saving computation time. On the other hand, the number of samples has to be estimated in advance: Too many data samples lead to excessive measurement time (oversampling), while too few data samples may be insufficient to characterize the function (undersampling). Thus, one-shot strategies lack flexibility and may require multiple reiterations.

The second task of DoE is *exploitation*, i.e., selecting sampling positions where the most desirable outcome of f is predicted. This task can be performed by training a surrogate model to predict the outcome of f from the available samples. Indeed, surrogate models allow to define a sequential sampling strategy, also known as *active learning:* The surrogate model identifies the next sampling point according to an optimum criterion; next, the new sample is used to update the model. In this way, samples are collected one-by-one at consecutive iterations.

Depending on the goal of the sampling, exploration and exploitation can be balanced by selecting a proper surrogate model. If the function maximum is requested, a *local* surrogate model is sufficient to estimate the maximum position, thus ignoring exploration. Conversely, if the goal is to obtain an accurate approximation of f over the whole domain, a *global* surrogate model can be constructed from points where the prediction is most uncertain. Such points are typically spread across the domain, leading to a good exploration. Sequential sampling strategies like local linear approximation (LOLA)-Voronoi [9], [10] are based on the latter approach. In fact, the LOLA-Voronoi algorithm selects new sampling locations that deviate most from the linear approximation of the function, thus focusing where the estimated model error is the largest.

However, radiation pattern measurements exhibit an additional challenge for defining a proper sampling strategy. In fact, the measurement process is inherently sequential and the cost of sampling at one position is not constant, but it depends on the travel time from the previous position. For this reason, *line-based* strategies have been proposed to identify the optimal trajectory through the sampling points. Available line-based strategies are based on one-shot approaches, such as [4], [11], or sequential approaches such as [1], [6], [12]. As alternative, line-based samplings like [6], [12] follow a hybrid approach: They compute sampling points in different batches rather than singularly, thus generating a new trajectory for each iteration.

BO has already been successfully employed in DoE for EM devices design [13]–[15]. In fact, BO relies on the statistical uncertainty of a GP model in order to automatically balance exploration and exploitation. Thus, BO constitutes a good candidate for DoE in the context of radiation patterns sampling. The MASS strategy proposed in this work is fully sequential and line-based. Additionally, its Bayesian approach matches the purpose of sampling: The surrogate model combines global and local accuracy to best locate multiple maxima of the function, that are spread over the domain.

III. GP MODELING OVER CYLINDRICAL SURFACES

A. GP Regression

As discussed in the previous section, a surrogate model can drive the sampling process by predicting which points in the domain space are most informative for the purpose of the experiment. For this purpose, several machine learning methods have been used as surrogate models in DoE for EM devices, such as artificial neural networks [17], support vector machines [18], and GPs [19]. In this study, a GP is selected to model the measured field: $f \sim GP(X)$. Compared to other models, the GP can be trained on a low amount of data, thanks to its self-regularizing capability. At the limit, it can be updated with one sample at a time, allowing for the definition of a fully sequential strategy. Moreover, unlike classical interpolation techniques, the GP model estimates the uncertainty on the predicted values, which is crucial to define an efficient sampling strategy. The GP represents each available data point $\boldsymbol{x}$ as a random Gaussian variable with mean m and covariance k with respect to any other point $\boldsymbol{x}' \in X$

$$m(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})] \tag{1a}$$

$$k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - m(\boldsymbol{x}))(f(\boldsymbol{x}') - m(\boldsymbol{x}'))] \tag{1b}$$

where $\mathbb{E}$ is the expectation operator. While the mean can initially be set at zero, the covariance function has to be carefully chosen. Indeed, a well-designed covariance function, which is generically called *kernel*, is crucial for the GP model to achieve high accuracy. Section III-B describes how to construct a suitable kernel for the considered experimental setting.

Once the kernel is specified, Gaussian process regression (GPR) [19], also known as Kriging, can be executed. The GPR exploits Bayesian inference to compute the posterior distribution, i.e., the probability distribution of values conditioned to the available data points. At any test point $\boldsymbol{x}^*$, the posterior is Gaussian and it is fully determined by the mean μ and variance σ^2

$$\mu(\boldsymbol{x}^*) = K_{\boldsymbol{x}\boldsymbol{x}^*}^T (K_{\boldsymbol{x}\boldsymbol{x}} + \sigma_r^2 I)^{-1} f(\boldsymbol{x}) \tag{2a}$$

$$\sigma^2(\boldsymbol{x}^*) = K_{\boldsymbol{x}^*\boldsymbol{x}^*} - K_{\boldsymbol{x}\boldsymbol{x}^*}^T(K_{\boldsymbol{x}\boldsymbol{x}} + \sigma_r^2 I)^{-1}K_{\boldsymbol{x}\boldsymbol{x}^*} \tag{2b}$$

where the matrix K contains the covariance function computed between pairs of points specified by the subscript. Additionally, the tunable variance σ_r^2 accounts for an additive Gaussian noise on the measured data points. On one hand, the mean μ represents the prediction of the GP for one particular test point. On the other hand, the variance σ^2 of the posterior indicates the uncertainty of the model on the function value. Thus, in the setting of radiation pattern measurements, the field value at any position can be predicted by applying the GPR on the positions that have been sampled so far.

B. Cylindrical Kernel

The predictive accuracy of the GP depends on a suitable kernel k. The assumptions made about the function f are reflected in the choice of the kernel [19]. In fact, the kernel defines how strong the correlation is between a pair of points in the domain. Many kernels have been designed for GPR and other machine learning applications [22]. A popular choice is the squared exponential kernel

$$k_{SE}(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\sum_{d=1}^{2}\frac{(\boldsymbol{x}_d - \boldsymbol{x}'_d)^2}{2\ell_d^2}\right) \tag{3}$$

where d is the dimension, while ℓ is the tunable length-scale parameter. Since this kernel is stationary, it enforces a correlation among the f values that depends on the absolute distance among two sampling positions. In other words, the GP model will assign similar probability of values to positions that are close together. Additionally, being infinitely differentiable, the squared exponential function assumes a high smoothness of f. These properties are appropriate for a radiated EM field at sufficiently low frequencies.

However, this kernel is not suitable for a function defined on a cylindrical surface. In fact, the right-hand side of (3) depends only on the Euclidean distance between two points in X, such that the periodicity of f along the azimuthal coordinate is not reproduced. Thus, independently from the trajectory of the measuring antenna, the squared exponential would lead to a discontinuity of the GP model between the values at $\phi = 0$ and at $\phi = 2\pi$ in the cylindrical reference frame, even though they correspond to the same positions in the physical space. To overcome this issue, a new cylindrical kernel is constructed as follows. First, a 1-D periodic kernel is chosen for ϕ, i.e., the azimuthal component of $\boldsymbol{x}$

$$k_\phi(\phi, \phi') = \exp\left(\frac{2\sin^2\left(|\phi - \phi'| \cdot \pi/p\right)}{\ell_\phi^2}\right) \tag{4}$$

where p is the period length parameter and ℓ_ϕ is the length scale parameter over ϕ. If $p = 2\pi$ is assigned, k_ϕ acquires the same periodicity of f over ϕ. Second, the 1-D squared exponential kernel is chosen for h, i.e., the vertical component of $\boldsymbol{x}$

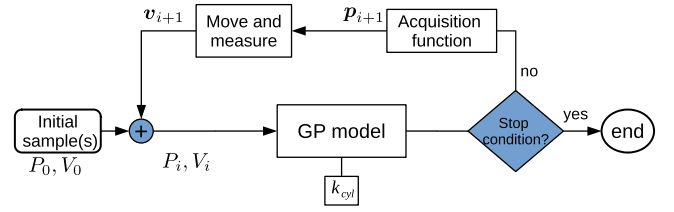$$k_h(h, h') = \exp\left(-\frac{(h - h')^2}{2\ell_h^2}\right) \tag{5}$$

Fig. 1. Flowchart of the MASS.

where ℓ_h is the length scale parameter over h. Then, by executing the tensor product between kernels [20], k_ϕ and k_h can be combined into a single 2-D kernel k_{cyl}, defined over both the dimensions of $\boldsymbol{x}$

$$k_{\text{cyl}}(\boldsymbol{x}, \boldsymbol{x}') = k_h(h, h')k_\phi(\phi, \phi'). \tag{6}$$

The resulting kernel preserves the stationarity and the smoothness properties of the squared exponential, while being periodical over ϕ. Thus, it is suitable for the GP modeling of functions defined over cylindrical surfaces.

IV. MASS

After constructing the adequate GP model, the MASS adopts the scheme in Fig. 1. The first step is to collect the initial position for the antenna $P_0 = \{\boldsymbol{p}_0\}$ and measure the corresponding field value $V_0 = \{v_0\}$. These initial sets are extended iteratively with one sample at a time, such that at each iteration i the sets become $P_i = \{\boldsymbol{p}_n\}_{n=0}^{i}$ and $V_i = \{v_n\}_{n=0}^{i}$. The second step is building the GP model on the collected samples P_i, V_i using the cylindrical kernel defined in Section III-B. At this point, it is possible to compute the posterior for the field values at any position, by executing the GPR as described in Section III-A. Next, the stop condition is checked: If the condition is met, the sampling process is ended. Otherwise, the *acquisition function* is computed on the GP model posterior to determine the next sampling position $\boldsymbol{p}_{i+1}$. Then, a new field value v_{i+1} is acquired by moving the antenna and measuring at the next position. Finally, $\boldsymbol{p}_{i+1}$ and v_{i+1} are added to the corresponding sets P_i and V_i, respectively. These steps are repeated until the stop condition is met.

A. Motion-Aware Expected Improvement

The task of identifying the next position for the antenna is delegated to the acquisition function. Thus, the acquisition function has to be designed in accordance with the goals of the sampling strategy, balancing the exploration and exploitation tasks. If the only objective is to find the absolute maximum of the modeled function f, the proposed scheme corresponds to a BO. In BO, the acquisition function α estimates the probability of finding the maximum of f for each point in the domain X. Thus, the next point to sample is chosen by solving an optimization problem

$$\boldsymbol{p}_{i+1} = \underset{\boldsymbol{x} \in X}{\arg\max}\, \alpha(\boldsymbol{x}). \tag{7}$$

Standard acquisition functions for BO, such as the expected improvement (EI) [21], rely on the GP posterior to achieve a good balance between exploration and exploitation. However, they do not take into account the time cost required to sample at different positions. Consequently, in this problem setting, choosing a standard acquisition function would produce a sampling strategy that is not optimized for time efficiency. In order to minimize the total sampling duration, a new acquisition function called motion-aware expected improvement (MEI) is constructed. First, let $EI(\boldsymbol{x})$ be the standard EI [21] over the sampling domain X

$$EI(\boldsymbol{x}) = \mathbb{E}[\max(GP(\boldsymbol{x}) - f(\hat{\boldsymbol{x}}), 0)] \tag{8}$$

where $\hat{\boldsymbol{x}}$ is the position corresponding to the highest value of f sampled so far

$$\hat{\boldsymbol{x}} = \operatorname*{argmax}_{\boldsymbol{x} \in P_i} f(\boldsymbol{x}). \tag{9}$$

The maximum of $EI(\boldsymbol{x})$ corresponds to the highest expected increase over the current maximum $f(\boldsymbol{x}^+)$. Based on the idea of cost-aware modeling in [24], the MEI is obtained from the normalization of the $EI(\boldsymbol{x})$ by a specific cost function C

$$MEI(\boldsymbol{x}) = \frac{EI(\boldsymbol{x})}{1 + C}. \tag{10}$$

The purpose of the normalization is to penalize the points that require higher time to be reached by the antenna. Therefore, the cost function is based on the minimum travel time between any position $\boldsymbol{x}$ and the current position of the antenna $\boldsymbol{p}_i$

$$C = \Delta t_{\min}(\boldsymbol{x}, \boldsymbol{p}_i). \tag{11}$$

In general, the travel time can be derived from the mechanical characteristics of the measurement setup. For instance, consider the motion of the antenna on the cylindrical surface as a superposition of a motion along the azimuth and a motion along the vertical axis. Then, assuming that the antenna can switch between zero and constant speed instantaneously, moving across each direction takes the following time intervals:

$$\Delta t_{phi} = \frac{\min[|\phi - \phi_i|, 2\pi - |\phi - \phi_i|]}{s_\phi} \tag{12a}$$

$$\Delta t_h = \frac{h - h_i}{s_h} \tag{12b}$$

where $\boldsymbol{x} = [\phi, h]$ and $\boldsymbol{p}_i = [\phi_i, h_i]$, while s_ϕ and s_h are the maximum angular speed along ϕ and the maximum vertical speed along h, respectively. Note that the $\min$ operation in (12a) selects the shortest path between a clockwise and a counterclockwise revolution on the cylindrical surface. However, in some settings, the motion of the measuring antenna is limited to only one revolution around the device such that, in order to reach the next position only one azimuthal direction is allowed. In this case, (12a) can be modified in

$$\Delta t_{phi} = \frac{|\phi - \phi_i|}{s_\phi}. \tag{13}$$

Furthermore, two different situations may occur. In case the antenna can move along the two dimensions simultaneously, the minimum moving time is approximately

$$\Delta t_{\min}^{(1)}(\boldsymbol{x}, \boldsymbol{p}_i) \approx \max[\Delta t_{phi}, \Delta t_h]. \tag{14}$$

Indeed, in this condition, it is necessary to wait for the slowest of the motions to complete. Conversely, if antenna can only move along one direction at a time

$$\Delta t_{\min}^{(2)}(\boldsymbol{x}, \boldsymbol{p}_i) \approx \Delta t_{phi} + \Delta t_h \tag{15}$$

as the two motions are executed consecutively. Note that, given s_ϕ and s_h, $\Delta t_{\min}^{(1)} \le \Delta t_{\min}^{(2)}, \forall(\boldsymbol{x}, \boldsymbol{p}_i)$. (14) and (15) provide simple approximations for a wide variety of setups. Nonetheless, the physical derivation of $\Delta t_{\min}$ should be adjusted to the setup that is actually in use. Finally, having obtained the cost function C, the next position in the proposed sampling scheme is located by optimizing the MEI acquisition function

$$\boldsymbol{p}_{i+1} = \operatorname*{argmax}_{\boldsymbol{x} \in X} MEI(\boldsymbol{x}). \tag{16}$$

By employing both the cylindrical kernel and the MEI acquisition function, the resulting strategy aims at the positions that present a good tradeoff between the chance of finding the maximum and the time cost required to reach them.

Following the flowchart in Fig. 1, the collection of sampling points continues until a stop condition is met. The most straightforward implementation of the stop condition is fixing a maximum amount of samples or a maximum duration for the sampling. Since collecting different samples requires different amounts of time, the latter criterion is chosen for the MASS.

V. Evaluation Metrics

In order to evaluate the performance of the MASS, it is necessary to measure a large set of samples at test positions X_s. Thereby, sampling metrics can be defined by comparing the GP model prediction with the actual field values. Since the first goal of the sampling is to search for the field maximum, the primary sampling metric is the *regret R*

$$R = \max_{\boldsymbol{x} \in X_s} f(\boldsymbol{x}) - \max_{\boldsymbol{x} \in P_i} f(\boldsymbol{x}). \tag{17}$$

The regret denotes the discrepancy between the measured maximum of f and the highest value sampled so far; thus, a low regret suggests that the antenna has come near the main peak of the emitted field. Unfortunately, the regret may drop significantly when sampling near a local maxima, even if the absolute maximum is far from the collected samples. For this circumstance, it is useful to evaluate the overall predictive accuracy of the surrogate model by computing the root-mean-square error (rmse) over X_s. Although, in radiation patterns, lower field values are often close to the background noise. Hence, as suggested in [1], it is more relevant to consider only the positions where the field is high

$$X_f = \{\boldsymbol{x}_j \in X_s : f^*(\boldsymbol{x}_j) > \gamma \cdot \max[f^*(\boldsymbol{x})]\} \tag{18}$$

where γ is a chosen threshold and f^* is the field function scaled to the range [0, 1]. Then, a constrained rmse indicator can be computed over the test positions

$$\gamma\text{-}rmse = \sqrt{\sum_{\boldsymbol{x} \in X_f} \frac{(GP(\boldsymbol{x}) - f(\boldsymbol{x}))^2}{|X_f|}}. \tag{19}$$
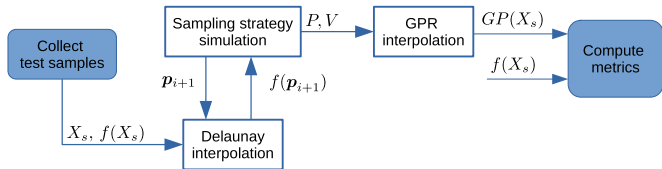
Fig. 2. Procedure for the evaluation of sampling strategies on a common test set X_s.
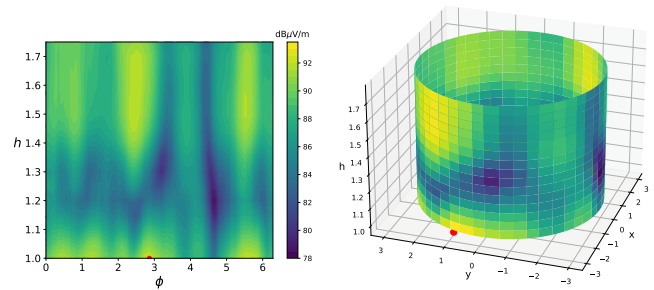

Fig. 3. Electrical field magnitude for the vertical polarization, @740 MHz at 3 m from the DUT. On the left: Plane view of the field in cylindrical coordinates; on the right: 3D view of the field in Euclidean coordinates; the red dot indicates the position of the maximum recorded value (92.21 dB μV/m).

The most immediate way to test a sampling strategy is to execute the measurement at the indicated positions and compute the metrics of interest. However, noise or other time-dependant fluctuations could alter the emitted field, compared to the test data. In this condition, the sampling metrics would return inconsistent values. This problem is exacerbated if several strategies are tested consecutively. Therefore, rather than performing multiple sets of measurements, sampling strategies are *simulated* by using the test set as ground truth for the function $f(\boldsymbol{x})$. Then, the sampling metrics can be computed for each strategy by following the procedure of Fig. 2. First, test values are collected through physical measurement at the position X_s. Second, the strategy is simulated by computing the next sampling position $\boldsymbol{p}_{i+1}$. For each new position, the corresponding value $f(\boldsymbol{p}_{i+1})$ is obtained via a Delaunay interpolation over the test samples. Once the simulation is concluded, a final interpolation of the sampled values of X_s is performed. By definition, GP already constitutes an interpolation model for the MASS and other strategies based on BO. Similarly, the samples from one-shot strategies can be interpolated with an additional GPR. Finally, the GP model prediction and the test values can be used to compute the validation metrics over the same positions X_s. This procedure allows to directly compare the metrics from the MASS with any other sampling strategy.

Furthermore, in order to evaluate the metrics in a function of time, the sampling instants t_i can be estimated as

$$t_i = \sum_{n=0}^{i-1} \left[\Delta t_{\min}(\boldsymbol{p}_n, \boldsymbol{p}_{n+1}) + t_{\text{meas}} + t_{\text{comp}} \right] \quad (20)$$

where $\Delta t_{\min}$ is approximated as described in Section IV-A and the term t_{meas} models the delay introduced by the measuring instrumentation. Additionally, t_{comp} is the delay caused by the software control of the antenna. For one-shot strategies, t_{comp} is negligible since all the sampling positions are computed beforehand. On the contrary, for sequential strategies, t_{comp} accounts for the computation of $\boldsymbol{p}_{i+1}$ at each iteration. In particular, for the proposed motion-aware sampling, both the GPR and the acquisition function optimization contribute to the computational time.

VI. APPLICATION EXAMPLE: COMB GENERATOR RADIATION PATTERN

A. Experimental Setup

The proposed MASS is applied to a DUT, which is a comb generator fed by a resonator and powered by a regular lab dc power supply. The electric field emitted by the device is measured in a SAC with an R&S broadband antenna (HL562) and electromagnetic interference (EMI) receiver (ELS6). The antenna scans a cylindrical surface around the device as a result of two simultaneous motions: 1) the translation of the antenna along the vertical axis and 2) the rotation of the device around its own vertical axis. The vertical position of the antenna is restricted to the range $[1\,\text{m}, 1.75\,\text{m}]$ from the ground, while the DUT can perform a full rotation, such that its angular position covers the range $[0, 2\pi]$. The maximum speeds along the axes are determined by the stepper motors employed: The maximum vertical speed is $s_h = 0.10$ m/s, while the angular speed is $s_\phi = 0.16$ rad/s. The instrumentation introduced a measurement delay of $t_{\text{meas}} \sim 1$ s. Note that, in this setup, the antenna takes almost 20 s to shift to the opposite side of the device. Since the motions over the two axes are simultaneous, $\Delta t_{\min}$ is approximated with 14 for both the MEI acquisition function and the estimation sampling instants.

The antenna is placed at 3 m distance from the device, which is sufficient to measure in far field conditions, while its orientation is adjusted to the vertical polarization of the electrical field. Then, the magnitude of electrical field is measured at a frequency of 740 MHz on a uniform 91x16 grid on the cylindrical surface. The measured samples are collected in a test set X_s for the evaluation of the metrics. The contour plot in Fig. 3 represents the measured values over the two cylindrical coordinates. Note that the maximum value of field of 92.21 dB μV/m is recorded in proximity of the lower boundary of h. In addition, the field presents several peaks scattered across the space.

B. Results and Discussion

Following the procedure described in Section V, the different sampling strategies are simulated on the test set. The simulation is performed by assuming no limit on the number of revolutions around the DUT. Thus, the trajectories are allowed to cross the 2π boundary and the azimuthal time intervals are approximated with (13). Note that the test values in X_s are converted to a linear scale (mV/m) for the execution of the
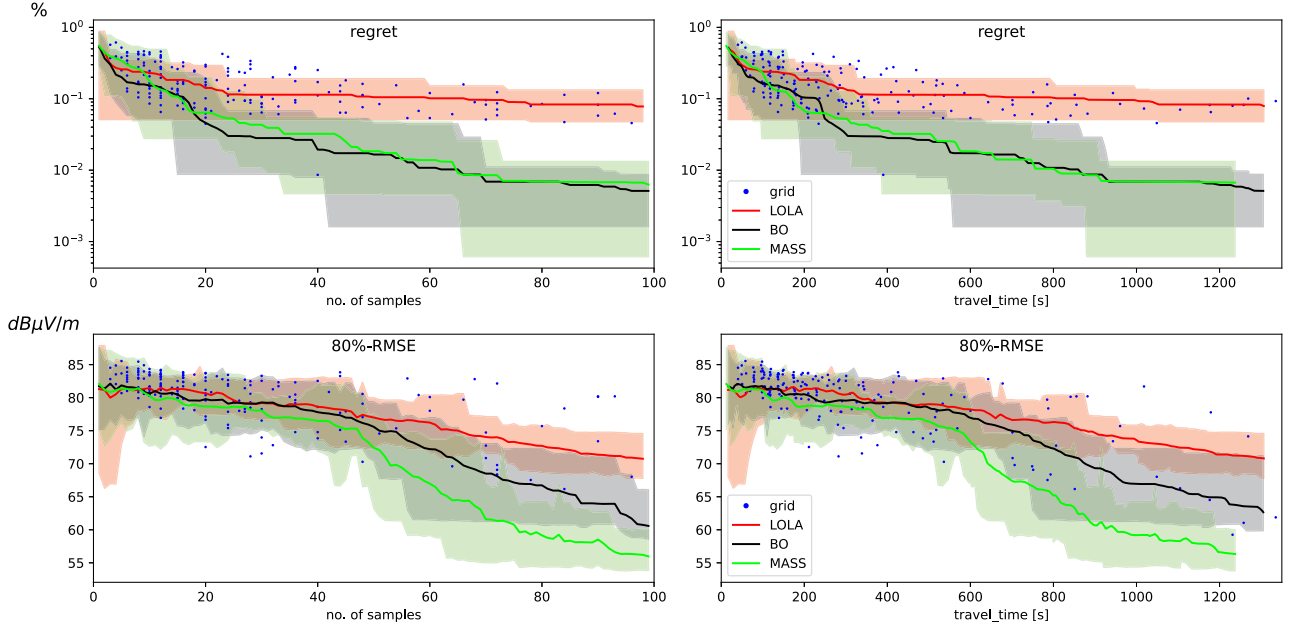
Fig. 4. Evolution of sampling metrics among different strategies. The solid lines show the metric values for sequential strategies, averaged across ten runs. The colored bands represent the deviation between the minimum and the maximum recorded metric among ten runs. Each blue dot represents a complete run of grid sampling for random step sizes in the ranges: $\Delta\phi \in [0.1, 2.0]$ and $\Delta h \in [0.03, 0.50]$. Starting position is chosen randomly on the cylindrical surface.
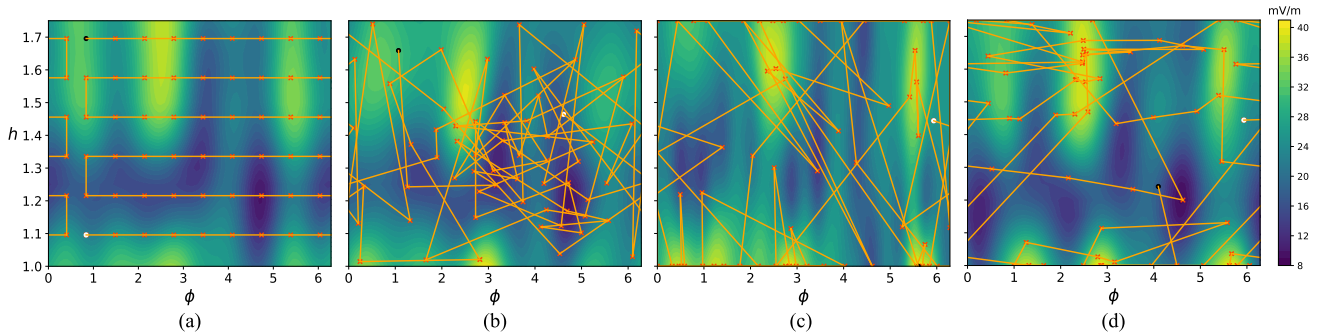

Fig. 5. Example of GP model prediction and traveled trajectory for different strategies after 60 samples. Red crosses indicate the sampling positions, while the white and black dots are the initial and the last sample, respectively. (a) Grid. (b) LOLA-Vornoi. (c) BO. (d) MASS.

TABLE I
SIMULATION SETTINGS FOR EACH SAMPLING STRATEGY

Strategy	Type	GP Interp. kernel	t_{comp}
Grid	One-shot	cylindrical	~ 0 s
LOLA-V	sequential	cylindrical	~ 0.1 s
Bayesian Opt. (BO)	sequential	squared-exponential	~ 1 s
BO + k_{cyl}	sequential	cylindrical	~ 1 s
BO + MEI	sequential	squared-exponential	~ 1 s
MASS (BO + k_{cyl} + MEI)	sequential	cylindrical	~ 1 s

GPR and the computation of the sampling metrics. Table I summarizes the settings employed for each simulated strategy. A random position is chosen to initialize each sampling strategy. The sampling instants are obtained by recording the computation time for a computer with an Intel(R) Core(TM) i7-9850H CPU

and 16 GB of RAM. Specifically for the grid strategy, the grid spacing across the azimuthal and vertical directions is drawn from uniform distributions with ranges: $\Delta\phi \in [0.1, 2.0]$ and $\Delta h \in [0.03, 0.50]$, respectively. Then, the samples are sorted according to a horizontal scanning trajectory [Fig. 5(a)], also known as Boustrophedon [11].

The evolution of the performance metrics over the iterations is shown in Fig. 4, which compares MASS with preexisting sampling strategies. Here, each blue dot represents a complete run of the grid sampling for specific step sizes. The bottom half of the figure reports the $rmse$ at 80% of the field maximum. It is apparent from the plot that the MASS has the lowest 80%-rmse over time, while its regret is comparable with the standard BO. In other words, the GP interpolator of the MASS reaches a high accuracy for both the global and the local maxima of the field. Note that the performance of grid sampling across multiple

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

GARBUGLIA *et al.*: BAYESIAN ACTIVE LEARNING FOR RADIATION PATTERN SAMPLING OVER CYLINDRICAL SURFACES 7
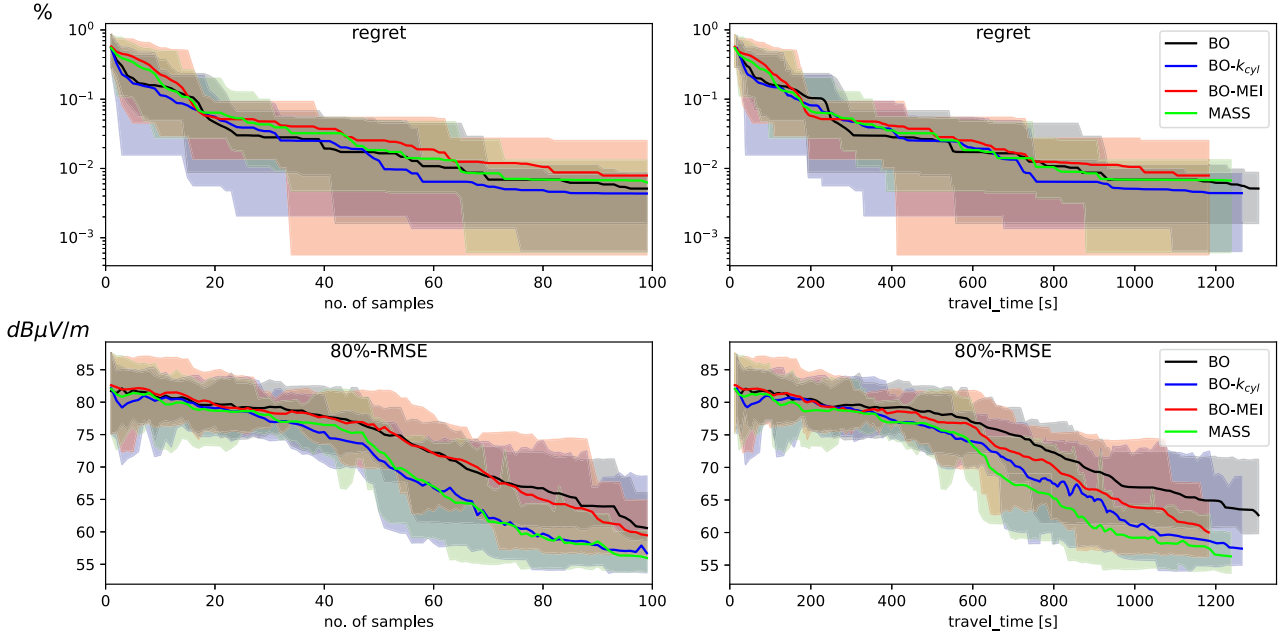
Fig. 6. Comparison of performance metrics for strategies with and without cylindrical kernel and MEI.

TABLE II
SAMPLE- AND TIME-EFFICIENCY FOR THE SIMULATED SEQUENTIAL STRATEGIES, FOR A FIXED THRESHOLD ON THE EVALUATION METRICS

	LOLA-Voronoi	BO	BO+k_{cyl}	BO+MEI	**MASS** (BO+MEI+k_{cyl})
Avg. samples @ regret=5%	78	**24**	**24**	26	27
Avg. samples @ 80%-RMSE=74 dBμV/m	61	56	**37**	53	39
Avg. time @ regret=5%	17 min 33 s	5 min 16 s	5 min 15 s	**4 min 53 s**	5 min 3 s
Avg. time @ 80%-RMSE=74 dBμV/m	13 min 43 s	12 min 16 s	8 min 9 s	10 min 23 s	**7 min 44 s**

runs is inconsistent and strongly depends on the chosen step sizes.

The differences among sampling approaches are evident from Fig. 5, which provide an example of the travelled trajectory for 60 samples. Since grid strategy is purely explorative, the sampling positions look evenly scattered in the domain. Conversely, BO and the MASS sampling focus near few regions of interest, near the maxima of the function.

Furthermore, Fig. 6 analyzes separately the effect of the cylindrical kernel and the MEI, when employed in the standard BO strategy. Interestingly, the regret presents no substantial decrease with either the new kernel or the new acquisition function. On the contrary, the two modifications contribute to the reduction 80%-rmse across time, thus making the MASS more efficient in identifying the local maxima compared to the standard BO. Finally, Table II compares the sample efficiency and the time efficiency of the sequential strategies for a fixed threshold on the recorded metrics. What stands out is that the MASS, which uses both the cylindrical kernel and the MEI acquisition function, reaches the same level of 80%-rmse much faster than standard BO, while it locates the field maximum in a similar amount of time. In fact, 8 min, equivalent to 39 samples, are sufficient for the MASS to decrease the error to 74 dB μV/m around the local maxima of the field.

VII. CONCLUSION

In this work, the new MASS has been presented for the measurement of radiated patterns on a cylindrical surface around the device under test. Based on active learning, the new strategy sequentially collects data and trains a GP model of the emitted field which is highly accurate near the maxima. Moreover, the MASS identifies the sampling positions with the best trade-off between the probability of finding the maximum and the required travel time. In cases where the antenna movement is slow relative to the sampling area, this approach shows better time efficiency than minimizing the number of samples. The experiment indicates that MASS effectively reduces the total sampling time, compared to preexisting explorative sampling strategies and standard BO.

Further studies are needed to extend the presented approach to the field measurement over multiple frequencies. An additional challenge for the research is the design of a stop condition that automatically terminates the sampling based on the surrogate model, without the need of a prefixed time budget.

REFERENCES

[1] S. Serpaud, A. Boyer, S. Ben-Dhia, and F. Coccetti, "Fast and accurate near-field measurement method using sequential spatial adaptive sampling (SSAS) algorithm," *IEEE Trans. Electromagn. Compat.*, vol. 63, no. 3, pp. 858–869, Oct. 2021.

[2] S. Tao, H. Zhao, and Z. D. Chen, "An adaptive sampling strategy based on region growing for near-field-based imaging of radiation sources," *IEEE Access*, vol. 9, pp. 9550–9556, 2021.

[3] D. Deschrijver, F. Vanhee, D. Pissoort, and T. Dhaene, "Automated near-field scanning algorithm for the EMC analysis of electronic devices," *IEEE Trans. Electromagn. Compat.*, vol. 54, no. 3, pp. 502–510, Sep. 2012.

[4] T. Van Steenkiste, J. van der Herten, D. Deschrijver, and T. Dhaene, "ALBATROS: Adaptive line-based sampling trajectories for sequential measurements," *Eng. Comput.*, vol. 35, pp. 537–550, 2018.

[5] R. Delanghe, T. Van Steenkiste, I. Couckuyt, D. Deschrijver, and T. Dhaene, "A Bayesian optimisation procedure for estimating optimal trajectories in electromagnetic compliance testing," *Eng. Proc.*, vol. 3, no. 1, 2020.

[6] M. Rahimi, M. Hansen, W. J. Kaiser, G. S. Sukhatme, and D. Estrin, "Adaptive sampling for environmental field estimation using robotic sensors," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 3692–3698.

[7] D.C Montgomery, *Design and Analysis of Experiments*. Hoboken, NJ, USA: Wiley, 1952.

[8] F. A. Viana, G. Venter, and V. Balabanov, "An algorithm for fast optimal Latin hypercube design of experiments," *Int. J. Numer. Methods Eng.*, vol. 82, no. 2, pp. 135–156, 2010.

[9] K. Crombecq, D. Gorissen, D. Deschrijver, and T. Dhaene, "A novel hybrid sequential design strategy for global surrogate modeling of computer experiments," *SIAM J. Sci. Comput.*, vol. 33, no. 4, pp. 1948–1974, 2011.

[10] J. van der Herten, I. Couckuyt, D. Deschrijver, and T. Dhaene, "A fuzzy hybrid sequential design strategy for global surrogate modeling of high-dimensional computer experiments," *SIAM J. Sci. Comput.*, vol. 37, no. 2, pp. A1020–A1039, 2015.

[11] H. Choset, "Coverage of known spaces: The boustrophedon cellular decomposition," *Auton. Robots*, vol. 9, pp. 247–253, 2000.

[12] M. A. Qureshi, C. H. Schmidt, and T. F. Eibert, "Adaptive sampling in spherical and cylindrical near-field antenna measurements," *IEEE Antennas Propag. Mag.*, vol. 55, no. 1, pp. 243–249, 2013.

[13] J. E. Rayas-Sánchez, S. Koziel, and J. W. Bandler, "Advanced RF and microwave design optimization: A journey and a vision of future trends," *IEEE J. Microw.*, vol. 1, no. 1, pp. 481–493, 2021.

[14] H. M. Torun, M. Larbi, and M. Swaminathan, "A Bayesian framework for optimizing interconnects in high-speed channels," in *Proc. IEEE MTT-S Int. Conf. Numer. Electromagn. Multiphys. Model. Optim.*, 2018, pp. 1–4.

[15] J. Qing, N. Knudde, I. Couckuyt, D. Spina, and T. Dhaene, "Bayesian active learning for electromagnetic structure design," in *Proc. 14th Eur. Conf. Antennas Propag.*, 2020, pp. 1–5.

[16] F. Garbuglia *et al.*, "Bayesian active learning for multi-objective feasible region identification in microwave devices," *Electron. Lett.*, vol. 57, no. 10, pp. 400–403, 2021.

[17] Qi-Jun Zhang, K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for RF and microwave design—From theory to practice," *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 4, pp. 1339–1350, Apr. 2003.

[18] G. Angiulli, M. Cacciola, and M. Versaci, "Microwave devices and antennas modelling by support vector regression machines," *IEEE Trans. Magn.*, vol. 43, no. 4, pp. 1589–1592, Mar. 2007.

[19] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning, 1st Edition*. Cambridge, MA, USA: MIT Press, 2008.

[20] A. Wilson and H. Nickisch, "Kernel interpolation for scalable structured Gaussian processes (KISS-GP)," in *Proc. 32th Int. Conf. Mach. Learn.*, vol. 37, 2015, pp. 1775–1784.

[21] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Glob. Optim.*, vol. 13, pp. 455–492, 1998.

[22] A. Wilson and R. Adams, "Gaussian process kernels for pattern discovery and extrapolation," in *Proc. 30th Int. Conf. Int. Conf. Mach. Learn.*, 2013, pp. 1067–1075.

[23] D. Zhan and H. Xing, "Expected improvement for expensive optimization: A review," *J. Glob. Optim.*, vol. 78, pp. 507–544, Jul. 2020.

[24] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," 2012, *arXiv:1206.2944*.

Federico Garbuglia received the M.Sc. degree in electronic engineering from the Polytechnic University of Marche, Anchona, Italy, in 2019. Since 2019, he has been working toward the Ph.D. degree in electrical engineering from the IDLab, Ghent University, Ghent, Belgium, working in the Surrogate Modelling Lab (SUMO Lab).

His current research interests include machine learning for electronics design automation and optimization, and adaptive system identification.

Tim Claeys (Member, IEEE) received the Ph.D. degree in electrical engineering from KU Leuven, Leuven, Belgium, in 2018.

Since 2018, he has been a Postdoctoral Researcher with the M-Group Research Group, KU Leuven. His research interests include near-field scanning, characterization methods for shielding materials and gaskets, electromagnetic interference resilience of wireless protocols, and global reliability of electronic systems.
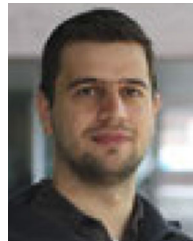
Ivo Couckuyt (Member, IEEE) received the M.S. degree in computer science from the University of Antwerp (UA), Antwerp, Belgium, in 2007, and the Ph.D. degree in engineering from the IDLab, Ghent University, Ghent, Belgium, in 2013.

In October 2007, he was with the research group Computer Modeling and Simulation (COMS). Since 2021, he has been an Associate Professor with the IDLab Research Group, Ghent University, working on automation in machine learning, data analytics, data-efficient machine learning, and surrogate modeling.

Dirk Deschrijver (Senior Member, IEEE) received the Ph.D. degree in computer science from the Department of Mathematics and Computer Science, in 2007, and received a second Ph.D. degree in engineering from Ghent University, Ghent, Belgium, in 2012.

From 2008 to 2014, he was an FWO Postdoctoral Research Fellow with the IBCN Research Group in IDLab, Ghent University. Since 2016, he has been an Associate Professor with the IDLab Research Group, Ghent University, working on data analytics, machine learning, and surrogate modeling.

Tom Dhaene (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from Ghent University, Ghent, Belgium, in 1989 and 1993, respectively.

Since 2000, he has been a Professor with the Computer Modeling and Simulation Research Group, University of Antwerp, Antwerp, Belgium. Since 2007, he has been a Full Professor with the IDLab, Ghent University, researching on distributed scientific computing, machine learning, bioinformatics, signal integrity, electromagnetic compatibility, optimal design, surrogate modeling, and circuit and EM modeling.