

Ray-Tracing-Based Feature-Aware Bayesian Optimization for Joint Base Station Selection and Configuration

Koki Kanzaki and Koya Sato

Artificial Intelligence eXploration Research Center,
The University of Electro-Communications, Chofu, Tokyo, Japan
Email: k-kanzaki@uec.ac.jp, k_sato@ieee.org

Abstract—This paper proposes an efficient framework for the joint base-station selection and configuration optimization that maximizes the communication coverage over a target area. Base stations must be carefully designed prior to the system deployment; however, the objective function for practical wireless channels cannot be expressed in a closed form, requiring a closed-box optimization method that iterates parameter search and channel simulation. This problem is difficult to solve efficiently because the channel simulation, such as ray tracing, is computationally expensive, and the optimization follows a mixed combinatorial/continuous problem with high-dimensional inputs. To address this, the proposed method consists of i) ray-tracing-based feature extraction and ii) Bayesian optimization (BO). The method first extracts propagation fingerprints from path-gain distributions obtained for representative transmit directions. After combining them with candidate-position information and reducing their dimensionality, standard BO searches for the optimal base-station configuration in this feature space. Performance evaluation for a mmWave communication system in two urban scenarios demonstrates that the proposed method achieves approximately 3.4 – 6.6 percent point higher required-throughput achievement ratios than baseline methods such as the genetic algorithm.

Index Terms—Bayesian optimization, Gaussian process regression, closed-box optimization

I. INTRODUCTION

The amount of data delivered over the Internet has been increasing over the last decade, and content that requires real-time responsiveness has also rapidly become widespread. In addition, these services are increasingly accessed not only through high-speed fixed-line networks but also through wireless communication devices such as smartphones and autonomous vehicles. However, wireless signals are strongly affected by site-specific propagation effects such as building geometry, blockage, and reflection. To provide high-speed and responsive services over communication area of interest under limited wireless resources, it is important to carefully design base-station deployment locations and antenna configurations in advance of the system deployment [1].

The base-station design faces two practical challenges. First, coverage evaluation requires modeling site-specific

propagation effects. While ray tracing can accurately capture these effects, its high computational cost limits large-scale candidate evaluation. Second, the design problem involves both selecting base station locations from discretized candidate points and optimizing their antenna parameters. This feature leads to an optimization problem combining discrete selection and conditional continuous variables.

Bayesian optimization (BO) is a method for efficiently optimizing an expensive closed-box function with a small number of trials; it has also been applied to base-station location and transmit-power design based on channel simulation [2]. However, typical BO mainly focuses on continuous optimization problem; i.e., its naive implementation cannot directly handle discrete structures such as the base-station candidate selection. Some BO methods have been proposed for discrete (or combinatorial) structures, including Bayesian optimization of combinatorial structures [3] and combinatorial Bayesian optimization using the graph Cartesian product [4]. However, these methods mainly target combinatorial spaces and are not directly applicable to the conditional mixed-variable space considered in this study, where continuous antenna parameters must be optimized jointly with base-station selection.

In this paper, we propose feature-aware BO for base-station selection and continuous-valued parameter configuration by leveraging propagation features obtained from ray tracing. In the proposed method, in addition to the geometric positions of the base-station candidate points, path-gain information derived from ray tracing for representative transmit directions is extracted as features. Using these features can map combinations of selected base stations and their antenna configurations into a unified continuous feature space. As a result, candidate points that are similar in terms of location or propagation characteristics are also treated as similar inputs in the BO surrogate model, enabling efficient search with a small number of ray-tracing evaluations.

The main contributions of this paper are as follows. (i) We formulate the base-station design problem in millimeter-wave communication environments as a conditional mixed-variable closed-box optimization problem that jointly handles base-station candidate selection and antenna configuration. (ii) We propose feature-aware BO that maps discrete base-station selection and continuous-valued parameters into a unified fea-

This work was supported by JST PRESTO under Grant Number JPMJPR23P3, CRONOS under Grant Number JPMJCS24N1, ASPIRE under Grant JPMJAP2346, and BOOST under Grant Number JPMJBS2415.

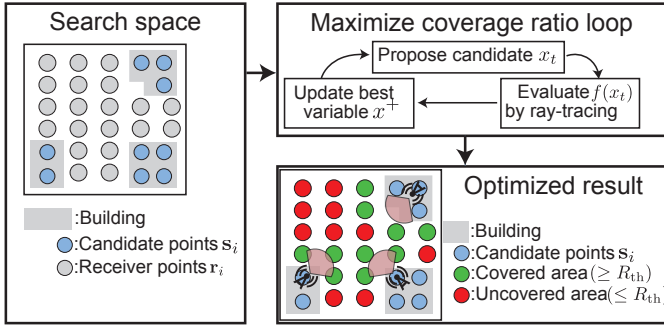


Fig. 1. System model.

ture space by using the location information of the base-station candidate points and ray-tracing-derived propagation features. (iii) Through evaluations in urban environments using Sionna RT [5], we show that the proposed method achieves a higher required-throughput achievement ratio than other methods.

II. SYSTEM MODEL

Fig. 1 shows the system model. We consider a base-station configuration problem in a three-dimensional target region $\mathcal{A} \subseteq \mathbb{R}^3$. Our objective is to jointly find the deployment locations of multiple base stations and their continuous-valued configuration parameters (the antenna angle and transmission power) so that the communication coverage is maximized. The set of candidate points indices is defined as $\mathcal{C} \triangleq \{1, 2, \dots, N\}$, where N is the number of base-station candidate points. Further, let the position of each candidate point be denoted by $s_i \in \mathbb{R}^3$. We assume that $M (\leq N)$ base stations are deployed at candidate points in $\mathcal{C}$. When a base station is deployed at the candidate point i , its continuous-valued parameter vector is defined as

$$\psi_i = [\alpha_i, \beta_i, p_i]^\top, \quad (1)$$

where α_i is the azimuth angle, β_i is the tilt angle, and p_i is the transmission power. All parameters to be optimized is modeled as

$$\mathbf{x} \triangleq [b_1, b_2, \dots, b_N, \psi_1^\top, \psi_2^\top, \dots, \psi_N^\top]^\top, \quad (2)$$

where b_i is a binary parameter representing the selection status of base station i : i.e.,

$$b_i = \begin{cases} 1 & \text{if the base station } i \text{ is selected,} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Given a design variable $\mathbf{x}$, let $\text{SINR}_i(\mathbf{r}, \mathbf{x})$ denote the signal-to-interference-plus-noise ratio from base station i at an arbitrary receiver position $\mathbf{r} \in \mathcal{A}$. When $b_i = 0$, we set $\text{SINR}_i(\mathbf{r}, \mathbf{x}) = 0$. Because each receiver connects to the base station providing the highest SINR, the effective $\text{SINR}(\mathbf{r}, \mathbf{x})$ at receiver position $\mathbf{r}$ is defined as

$$\text{SINR}(\mathbf{r}, \mathbf{x}) = \max_{i \in \mathcal{C}} \text{SINR}_i(\mathbf{r}, \mathbf{x}). \quad (4)$$

Algorithm 1 Standard BO for continuous problem.

Require: Continuous feasible search space Ξ , initial sample size t_0 , and evaluation budget T .

Ensure: Best continuous variable ξ^+ and best objective value q^+ .

- 1: Generate initial feasible points $\{\xi_j\}_{j=1}^{t_0} \subset \Xi$ at random.
- 2: **for** $j = 1, 2, \dots, t_0$ **do**
- 3: Evaluate $y_j = q(\xi_j)$.
- 4: **end for**
- 5: Set $\mathcal{T}_{t_0} \leftarrow \{(\xi_j, y_j)\}_{j=1}^{t_0}$.
- 6: Set (ξ^+, q^+) to the best observed pair in $\mathcal{T}_{t_0}$.
- 7: **for** $t = t_0, t_0 + 1, \dots, T - 1$ **do**
- 8: Construct the GP surrogate model using Eqs. (10)–(13).
- 9: Select the next point ξ_{t+1} by Eq. (15).
- 10: Evaluate $y_{t+1} = q(\xi_{t+1})$.
- 11: Update $\mathcal{T}_{t+1} \leftarrow \mathcal{T}_t \cup \{(\xi_{t+1}, y_{t+1})\}$.
- 12: Update (ξ^+, q^+) if y_{t+1} improves the current best value.
- 13: **end for**
- 14: **return** ξ^+, q^+ .

Then, the throughput at $\mathbf{r}$ can be derived as,

$$C(\mathbf{r}, \mathbf{x}) = B \log_2 (1 + \text{SINR}(\mathbf{r}, \mathbf{x})), \quad (5)$$

where B is the bandwidth. The performance metric is defined as the fraction of the target region in which throughput exceeds the target value R_{th} . Receiver height is fixed at h_{rx} , and the horizontal plane at height h_{rx} of the target region $\mathcal{A}$ is divided into U grid cells. Let $\mathbf{r}_u \in \mathcal{A}$ denote the receiver position corresponding to each receiver point $u \in \{1, 2, \dots, U\}$. The objective function $f(\mathbf{x})$ for the design variable $\mathbf{x}$ can then be defined as

$$f(\mathbf{x}) \triangleq \frac{1}{U} \sum_{u=1}^U \mathbf{1}[C(\mathbf{r}_u, \mathbf{x}) \geq R_{th}], \quad (6)$$

where $\mathbf{1}[\cdot]$ is an indicator function that returns 1 if the condition is true and 0 otherwise. Thus, $f(\mathbf{x}) \in [0, 1]$ represents the coverage ratio. The objective value $f(\mathbf{x})$ is computed by ray-tracing simulation. By iterating searching over design parameters and performing ray tracing, we solve the following closed-box optimization problem;

$$\mathbf{x}_* = \arg\max_{\mathbf{x}} f(\mathbf{x}). \quad (7)$$

III. STANDARD BO FOR CONTINUOUS OPTIMIZATION

Evaluating the objective function in Eq. (6) by ray tracing is computationally expensive, and the objective has no closed-form expression. Therefore, this study develops a BO-based solver for the problem in Eq. (7). Before presenting the proposed method, we review standard BO for the following continuous closed-box optimization problem:

$$\xi_* = \arg \max_{\xi \in \Xi} q(\xi), \quad (8)$$

where $\xi \in \mathbb{R}^{d_\xi}$ is a continuous variable, $\Xi \subseteq \mathbb{R}^{d_\xi}$ is the feasible search space, and $q(\cdot)$ is an expensive closed-box objective function. This auxiliary problem is used only to explain standard BO and differs from the mixed combinatorial/continuous base-station design problem in Eq. (7). Sec. IV extends this framework to the proposed feature-aware formulation.

BO efficiently optimizes an expensive closed-box objective with a limited number of evaluations, without requiring its closed-form expression or gradient information [6]. It consists of a surrogate model and an acquisition function. Given the observation dataset $\mathcal{T}_t = \{(\xi_j, y_j)\}_{j=1}^t$, we assume

$$y_j = q(\xi_j) + \epsilon_j, \quad \epsilon_j \sim \mathcal{N}(0, \sigma_n^2), \quad (9)$$

where ϵ_j is Gaussian observation noise with standard deviation σ_n . The surrogate model approximates the objective probabilistically, while the acquisition function selects the next point by balancing exploration and exploitation.

We use a Gaussian process (GP) as the surrogate model. Let $\ell(\xi)$ be a latent function approximating $q(\xi)$. Assuming a zero-mean GP prior,

$$\ell(\xi) \sim \text{GP}(0, k(\xi, \xi')), \quad (10)$$

the predictive distribution at an arbitrary point $\xi^\dagger$ is given by

$$\ell(\xi^\dagger) | \mathcal{T}_t \sim \mathcal{N}(\mu_t(\xi^\dagger), \sigma_t^2(\xi^\dagger)). \quad (11)$$

The posterior mean and variance are

$$\mu_t(\xi^\dagger) = \mathbf{k}_t^\top(\xi^\dagger)(\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}_t, \quad (12)$$

$$\sigma_t^2(\xi^\dagger) = k(\xi^\dagger, \xi^\dagger) - \mathbf{k}_t^\top(\xi^\dagger)(\mathbf{K}_t + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_t(\xi^\dagger), \quad (13)$$

where $\mathbf{K}_t$ is the kernel matrix of the observed points, $\mathbf{k}_t(\xi^\dagger)$ is the kernel vector between the observed points and $\xi^\dagger$, and $\mathbf{y}_t$ is the observation vector.

As the acquisition function, we use expected improvement (EI), defined as

$$\text{EI}(\xi^\dagger) = \mathbb{E} \left[\max \left(\ell(\xi^\dagger) - q_t^+, 0 \right) \right], \quad (14)$$

where q_t^+ is the best observed value up to iteration t . The next point is selected by maximizing EI:

$$\xi_{t+1} = \arg \max_{\xi \in \Xi} \text{EI}(\xi). \quad (15)$$

Using these components, BO can be performed as follows. First, t_0 initial evaluation points are collected by random sampling, and the observation dataset $\mathcal{T}_{t_0}$ is constructed. Next, a GP surrogate model is constructed from the observation data, and the next evaluation point is selected by the acquisition function. The selected point is then evaluated and added to the observation dataset. This procedure is repeated until the evaluation budget is reached, and the final best observed value and the corresponding ξ are taken as the optimization result. This procedure is summarized in Algorithm 1.

Algorithm 2 Ray-tracing-based feature-aware BO.

Require: Feasible search space $\mathcal{X}$, candidate positions $\{\mathbf{s}_i\}_{i=1}^N$, representative directions $\mathcal{D}$, path-gain samples, PCA dimension p , initial sample size t_0 , and evaluation budget T .

Ensure: Best design variable $\mathbf{x}^+$ and best objective value f^+ .

- 1: Define the representative direction set by Eq. (??).
 - 2: **for** $i = 1, 2, \dots, N$ **do**
 - 3: Compute the ray-tracing-based propagation feature $\mathbf{h}_i$ by Eq. (17).
 - 4: **end for**
 - 5: Construct the propagation feature matrix $\mathbf{H}$ by Eq. (18).
 - 6: Standardize $\mathbf{H}$ and apply PCA to obtain the low-dimensional propagation feature matrix $\mathbf{E}$.
 - 7: Construct $\mathbf{S}$ using Eq. (20).
 - 8: Construct and standardize the candidate feature matrix using Eq. (21).
 - 9: Extract the standardized candidate feature vector ϕ_i for each candidate point.
 - 10: Generate initial feasible design variables $\{\mathbf{x}_j\}_{j=1}^{t_0} \subset \mathcal{X}$ at random.
 - 11: **for** $j = 1, 2, \dots, t_0$ **do**
 - 12: Compute the feature vector $z(\mathbf{x}_j)$ by Eq. (22).
 - 13: Evaluate y_j using Eq. (6).
 - 14: **end for**
 - 15: Set $\mathcal{T}_{t_0}^z \leftarrow \{(z(\mathbf{x}_j), y_j)\}_{j=1}^{t_0}$.
 - 16: Set $(\mathbf{x}^+, f^+)$ to the best observed pair among the initial evaluations.
 - 17: **for** $t = t_0, t_0 + 1, \dots, T - 1$ **do**
 - 18: Construct the GP surrogate model on the feature space using Eq. (24).
 - 19: Compute the posterior distribution as in Eqs. (12) and (13), with ξ replaced by $z(\mathbf{x})$.
 - 20: Select the next design variable $\mathbf{x}_{t+1}$ by Eq. (26).
 - 21: Compute the feature vector $z(\mathbf{x}_{t+1})$ by Eq. (22).
 - 22: Evaluate y_{t+1} using Eq. (6).
 - 23: Update $\mathcal{T}_{t+1}^z \leftarrow \mathcal{T}_t^z \cup \{(z(\mathbf{x}_{t+1}), y_{t+1})\}$.
 - 24: Update $(\mathbf{x}^+, f^+)$ if y_{t+1} improves the current best value.
 - 25: **end for**
 - 26: **return** $\mathbf{x}^+, f^+$.
-

IV. RAY-TRACING-BASED FEATURE-AWARE BO

The standard BO in Sec. III is effective for expensive closed-box optimization, but its direct application to problem (7) is difficult. This is because the problem has the following two structural characteristics: (i) the search space contains a combinatorial component corresponding to the selection of M base stations from N candidate points, and (ii) the antenna parameters are conditional continuous variables that are meaningful only when the corresponding candidate point is selected, and the continuous values associated with unselected candidate points do not affect the objective function. These structures make it difficult for a GP surrogate model to

capture similarity between design variables when the original design variable $\mathbf{x}$ is used directly as input. The key idea of the proposed method is to keep the BO procedure largely unchanged while changing the input representation used by the GP surrogate model. Specifically, each feasible design variable $\mathbf{x} \in \mathcal{X}$ is mapped to a continuous feature vector $z(\mathbf{x})$ constructed from the selected candidate points, their antenna parameters, and ray-tracing-derived propagation features. The GP surrogate model and acquisition function are evaluated in this feature space, while acquisition maximization is still performed over the original feasible search space $\mathcal{X}$. In other words, the proposed method selects the next design variable $\mathbf{x}_{t+1}$ in the original mixed-variable space, but it models and compares candidates through their feature vector $z(\mathbf{x})$. Thus, standard BO can handle the conditional mixed-variable structure through a low-dimensional continuous representation.

To construct the feature vector $z(\mathbf{x})$, we first define features for each base-station candidate point. The candidate-point features consist of geometric position and ray-tracing-derived propagation information. These features place geographically or propagation-wise similar candidate points close to each other in the feature space. Let

$$\mathcal{D} \triangleq \{(\bar{\alpha}_d, \bar{\beta}_d)\}_{d=1}^D \quad (16)$$

be the set of representative transmit directions, where D is the number of representative transmit directions, and $(\bar{\alpha}_d, \bar{\beta}_d)$ denotes the azimuth and tilt angles of the d -th direction. Let $\gamma_{i,d,v}$ denote the path gain at receiver position $\mathbf{r}_v$ for candidate point i and direction $(\bar{\alpha}_d, \bar{\beta}_d)$. The propagation feature of candidate point i is

$$\mathbf{h}_i \triangleq [\gamma_{i,1,1}, \dots, \gamma_{i,1,V}, \gamma_{i,2,1}, \dots, \gamma_{i,D,V}]^T \in \mathbb{R}^{DV}, \quad (17)$$

where V is the number of points at which the path gain is evaluated. The propagation feature is computed for each candidate point, and the matrix $\mathbf{H}$ that collects the propagation features of all candidate points is defined as

$$\mathbf{H} \triangleq [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N]^T \in \mathbb{R}^{N \times DV}. \quad (18)$$

To obtain compact features that capture the dominant variations in the propagation characteristics, $\mathbf{H}$ is first standardized column-wise, and principal component analysis (PCA) is then applied to the standardized feature matrix. The resulting p -dimensional PCA features of all candidate points are collected into the low-dimensional propagation feature matrix

$$\mathbf{E} \triangleq [\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_N]^T \in \mathbb{R}^{N \times p}, \quad (19)$$

where $\boldsymbol{\eta}_i \in \mathbb{R}^p$ denotes the PCA feature vector of candidate point i . The matrix obtained by arranging the position vectors of all candidate points is denoted by $\mathbf{S}$ and is defined as

$$\mathbf{S} \triangleq [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N]^T \in \mathbb{R}^{N \times 3}. \quad (20)$$

The candidate feature matrix Φ is obtained by concatenating the low-dimensional propagation features and the position information, is defined as

$$\Phi \triangleq [\mathbf{E}, \mathbf{S}] \in \mathbb{R}^{N \times (p+3)}. \quad (21)$$

Let $\tilde{\Phi}$ denote the standardized version of Φ , and let ϕ_i be the feature vector corresponding to candidate point i extracted from $\tilde{\Phi}$. Using this vector, the GP input feature vector $z(\mathbf{x})$ associated with design variable $\mathbf{x}$ is defined as

$$z(\mathbf{x}) \triangleq \sum_{i=1}^N b_i \phi_i \otimes \mathbf{v}_i, \quad \mathbf{v}_i = \begin{bmatrix} 1, \cos \alpha_i, \sin \alpha_i, \tilde{\beta}_i, \tilde{p}_i \end{bmatrix}^T. \quad (22)$$

Here, $\tilde{\beta}_i$ and $\tilde{p}_i$ are normalized versions of β_i and p_i , respectively. This feature mapping combines the features of the selected candidate points with their azimuth angle, tilt angle, and transmission power into a single continuous feature vector. Moreover, because each term is multiplied by b_i , the continuous-valued parameters of unselected candidate points do not affect $z(\mathbf{x})$. Therefore, this feature mapping makes it possible to treat the conditional mixed-variable space, which includes both discrete base-station selection and continuous-valued parameters that are meaningful only for selected base stations, in a low-dimensional continuous form.

BO is then performed using $z(\mathbf{x})$ defined above. In the standard BO described in Sec. III, the design variable $\mathbf{x}$ is assumed to be directly modeled by a GP, whereas in this study, we assume that the latent function that approximates the objective function over the feature space follows a GP. That is, for the input obtained by the feature mapping $z(\mathbf{x})$, we introduce an unknown function

$$g : \mathbb{R}^{5(p+3)} \rightarrow \mathbb{R}, \quad (23)$$

and assume

$$g \sim \text{GP}(0, k(\cdot, \cdot)), \quad (24)$$

where the covariance between two designs $\mathbf{x}$ and $\mathbf{x}'$ is given by $k(z(\mathbf{x}), z(\mathbf{x}'))$. The objective function $f(\mathbf{x})$ is modeled using this latent function as

$$f(\mathbf{x}) \approx g(z(\mathbf{x})). \quad (25)$$

This mapping reduces the original $4N$ -dimensional representation to $5(p+3)$ dimensions. Based on this model, given the feature-space observation data $\mathcal{T}_t^z = \{(z(\mathbf{x}_j), y_j)\}_{j=1}^t$, the posterior predictive distribution is obtained in the same manner as in Sec. III. Using the surrogate model on this feature space, the proposed method evaluates EI through the feature mapping $z(\mathbf{x})$. Therefore, the design variable to be evaluated next is selected as

$$\mathbf{x}_{t+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \text{EI}(z(\mathbf{x})). \quad (26)$$

Importantly, Eq. (26) still searches over the original feasible design space $\mathcal{X}$, and the selected design variable $\mathbf{x}_{t+1}$ is evaluated by ray tracing using Eq. (6). Therefore, the proposed method preserves the standard BO loop while the GP surrogate model operates on the feature representation $z(\mathbf{x})$ instead of the original design variable $\mathbf{x}$. The overall procedure is summarized in Algorithm 2.

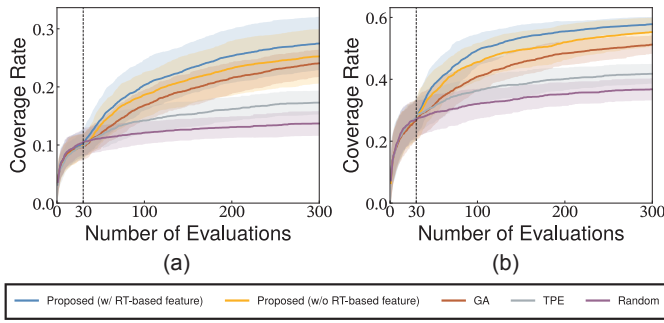


Fig. 2. Evaluation results. (a) Munich, (b) Etoile.

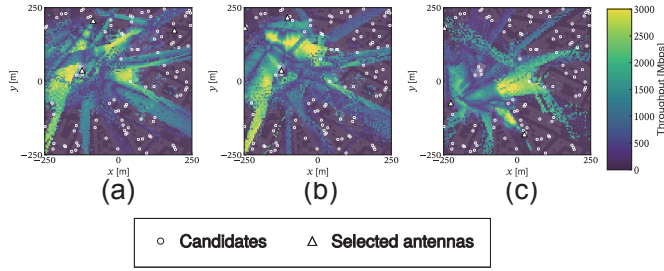


Fig. 3. Example of optimization. (a) Proposed (w/ RT-based feature), (b) Proposed (w/o RT-based feature), (c) GA.

V. PERFORMANCE EVALUATION

This section evaluates the effectiveness of the proposed method by maximizing the objective function. The objective function was evaluated using ray-tracing results obtained with NVIDIA Sionna RT [5]. The evaluation environments were the two models included in Sionna RT, namely Munich and Etoile. For each model, $N = 100$ candidate points were randomly generated at locations 5 m above rooftops, and the problem was to select $M = 3$ points from among them. The search range of the azimuth angle was set to $[0, 2\pi)$, and that of the tilt angle was set to $[0, \pi/3]$. The transmit antenna was modeled according to 3GPP TR 38.901 [7] and arranged as an 8×8 array, the carrier frequency was set to 28 GHz, and the bandwidth was set to 800 MHz. The transmission power was searched over the range in which the EIRP of each base station was $[21.9, 48.0]$ dBm. The required throughput was set to $R_{th} = 100 \times 10^6$ [bps]. The number of initial samples was set to $t_0 = 30$, the evaluation budget was set to $T = 270$, the compressed PCA dimension was set to $p = 4$, and the Matérn 3/2 kernel was used as the kernel function.

The results are shown in Fig. 2. The experiments were independently repeated 100 times, the solid lines show the mean values, and the shaded regions represent the range of $\pm\sigma$, where σ is the standard deviation of each method's performance. An example optimization result is shown in Fig. 3.

In the graphs, Proposed (w/ RT-based feature) denotes the method proposed in Sec. IV, and Proposed (w/o RT-based feature) denotes a version of the proposed method that removes the ray-tracing-derived features from the features used by the proposed method and uses only

the position information¹. GA denotes the genetic algorithm [8], TPE denotes the method using the Tree-structured Parzen Estimator [9], and Random denotes random search. The results show that, in Munich, Proposed (w/ RT-based feature) outperformed GA, which achieved the highest coverage rate among the baseline methods, by 3.4 percentage point, and it also outperformed Proposed (w/o RT-based feature) by 1.2 percentage point. Similarly, in Etoile, it outperformed GA by 6.6 percentage point and Proposed (w/o RT-based feature) by 4.0 percentage point. These results indicate that the proposed method, which uses radio propagation characteristics as features, outperforms the other methods, thereby demonstrating its effectiveness. The results also show that Proposed (w/o RT-based feature) outperforms GA, indicating that incorporating only the geometric arrangement of base stations can also improve the accuracy to some extent.

VI. CONCLUSION

This study proposed a method for efficient base-station configuration that maximizes the communication service area based on BO. The proposed method utilizes ray-tracing-based feature acquisition and PCA-based dimension reduction to support mixed combinatorial/continuous and high-dimensional inputs under the standard BO framework. Numerical results demonstrated that the proposed method achieves higher performance than conventional methods including GA and TPE. Our study assumes the beams of the base stations are static. In future, we will extend the method to more dynamic environment, including multiple-input multiple-output systems considering practical pedestrian flow.

REFERENCES

- [1] J. Li, Y. Niu, H. Wu, B. Ai, S. Chen, Z. Feng, Z. Zhong, and N. Wang, "Mobility Support for Millimeter Wave Communications: Opportunities and Challenges," *IEEE Commun. Surv. Tutor.*, vol. 24, no. 3, pp. 1816–1842, 2022.
- [2] K. Sato and K. Suto, "Bayesian Optimization Framework for Channel Simulation-Based Base Station Placement and Transmission Power Design," *IEEE Netw. Lett.*, vol. 6, no. 4, pp. 217–221, Feb. 2024.
- [3] R. Baptista and M. Poloczek, "Bayesian Optimization of Combinatorial Structures," in *In Proc. 35th Int. Conf. Mach. Learn.* PMLR, Jul. 2018, pp. 462–471.
- [4] C. Oh, J. Tomczak, E. Gavves, and M. Welling, "Combinatorial Bayesian Optimization using the Graph Cartesian Product," in *Adv. Neural Inf. Process. Syst.*, vol. 32. Curran Associates, Inc., 2019.
- [5] J. Hoydis, S. Cammerer, F. Ait Aoudia, M. Nimier-David, L. Maggi, G. Marcus, A. Vem, and A. Keller, "Sionna," 2022.
- [6] J. Snoek, H. Larochelle, and R. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," in *Adv. Neural Inf. Process. Syst.*, vol. 25. Curran Associates, Inc., 2012.
- [7] 3rd Generation Partnership Project (3GPP), "Study on channel model for frequencies from 0.5 to 100 ghz," Technical Report TR 38.901, 2017.
- [8] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
- [9] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for Hyper-Parameter Optimization," in *Adv. Neural Inf. Process. Syst.*, vol. 24. Curran Associates, Inc., 2011.

¹The results do not include costs in RT-based feature calculation. However, the calculation cost adjusted by tuning ray-tracing parameters such as the number of rays. Therefore, it is possible to reduce the total execution time enough to be ignored.