

Bayesian Optimization Framework for Channel Simulation-Based Base Station Placement and Transmission Power Design

Koya Sato and Katsuya Suto *Members, IEEE*

Abstract—This study proposes an adaptive experimental design framework for a channel-simulation-based base station (BS) design that supports the joint optimization of transmission power and placement. We consider a system in which multiple transmitters provide wireless services over a shared frequency band. Our objective is to maximize the average throughput within an area of interest. System operators can design the system configurations prior to deployment by iterating them through channel simulations and updating the parameters. However, accurate channel simulations are computationally expensive; therefore, it is preferable to configure the system using a limited number of simulation iterations. We develop a solver for the problem based on Bayesian optimization (BO), a black-box optimization method. The numerical results demonstrate that our proposed framework can achieve 18-22% higher throughput performance than conventional placement and power optimization strategies.

Index Terms—Channel simulation, adaptive experimental design, Bayesian optimization, log-normal shadowing

I. INTRODUCTION

Designing wireless systems before deployment is a typical, but essential task for efficient spectrum utilization. For example, in a sixth-generation (6G) network [1], operators must configure the parameters for dense base stations (BSs), such as BS placement and transmission power values, to provide stable wireless service in an area of interest under limited frequency resources. Furthermore, public Wi-Fi [2] and smart stadiums [3] are examples of such systems.

If the wireless channel is influenced solely by path loss, then BS placement can be optimized using lightweight channel simulations based on empirical path loss models. However, in practice, shadowing effects often degrade the system performance. The shadowing effect, which is a stochastic process, can only be quantified after determining the positions of the transmitters through channel simulations or actual measurements, rendering white-box optimization frameworks impractical. To proactively incorporate shadowing effects into the system design, it is necessary to iterate through parameter selection, channel simulation, and performance evaluation. However, high-precision channel simulations (e.g., ray tracing)

are computationally intensive and often require several hours to days per evaluation, thereby limiting the feasibility of conducting numerous simulations. This challenge herein motivates us to explore *how to find the optimal BS configurations with as few channel simulations as possible*.

Recently, adaptive experimental design has gained traction across various fields, including robotics, neuroscience, and materials discovery [4]. This data-driven framework designs experiments by iteratively sampling parameters, evaluating the performance, and predicting the performance for unknown inputs based on data-driven models. By employing a Gaussian process (GP)-based black-box optimization method known as Bayesian optimization (BO), optimal parameters can be identified with a limited number of simulations, greatly reducing design costs; BO will enable high-performance BS configurations with reduced simulation costs.

In this letter, we propose an adaptive experimental design framework for a simulation-based BS design that supports the joint optimization of placement and transmission power. Some recent studies have shown the benefits of BO in wireless systems (e.g., task offloading in mobile edge computing [5] and wireless resource allocation [6]). These studies suggest that BO-based optimizers can enhance system performance with fixed BS placements over time-varying wireless channels, even when performance metrics are treated as black boxes. In contrast to these studies, our approach focuses on exploring BS placements and transmission power settings to achieve high throughput over spatially-varying channels *before system deployment*. The objective is to maximize the average throughput in an area of interest, modeled as a black-box function due to shadowing. We propose an optimizer based on a nested BO architecture, and numerical simulations demonstrate that our framework achieves superior average throughput per unit area with fewer channel simulations.

II. SYSTEM MODEL

We consider a scenario in which multiple BSs provide wireless communication services in a two-dimensional area. Their operator tunes the BS placements and transmission power values before system deployment through several channel simulation trials to ensure that the average throughput within the area of interest is maximized, as shown in Fig. 1.

We define the set of BSs as $\mathcal{B} = \{\mathcal{B}_i \mid i = 1, 2, \dots, N_{\text{Tx}}\}$, where \mathcal{B}_i denotes the i -th BS and N_{Tx} is the number of BSs. We also express the region of interest as $\mathcal{A} = \bigcup_{i=1}^{N_{\text{Tx}}} \mathcal{A}_i$, where

This work was supported in part by JST ACT-X Grant Number JPMJAX21AA, JST PRESTO Grant Number JPMJPR23P3, and JST ASPIRE Grant Number JPMJAP2346, Japan. Koya Sato is with the Artificial Intelligence eXploration Research Center, The University of Electro-Communications, 182-8585, Tokyo, Japan (e-mail: k_sato@ieee.org). Katsuya Suto is with the Graduate School of Informatics and Engineering, The University of Electro-Communications, Chofu 182-8585, Japan (e-mail: k.suto@uec.ac.jp). Corresponding author: *Koya Sato*

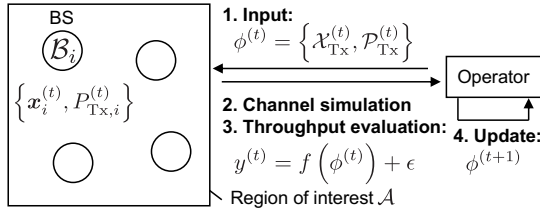


Fig. 1. System model.

$\mathcal{A}_i (\in \mathbb{R}^2)$ is the possible coordinate of the i -th BS. Assuming that the channel follows path loss and shadowing, the received signal power from the i -th coordinate $\mathbf{x}_{Tx,i} (\in \mathcal{A}_i)$ to the coordinate $\mathbf{x} (\in \mathcal{A})$ can be modeled as

$$P_{Rx,i}(\mathbf{x}) = P_{Tx,i} d_i^{-\eta} w(\mathbf{x}, \mathbf{x}_{Tx,i}) \quad [\text{mW}], \quad (1)$$

where $d_i = \|\mathbf{x}_{Tx,i} - \mathbf{x}\|$ [m] is the communication distance ($\|\cdot\|$ is the Euclidean distance), η is the path loss index, and $P_{Tx,i}$ [mW] is the transmission power of the i -th BS. Furthermore, $w(\mathbf{x}, \mathbf{x}_{Tx,i})$ represents the shadowing, indicating a spatial correlation between peer-to-peer wireless links [7].

When all BSs transmit signals simultaneously over a shared frequency band, the channel capacity of the i -th BS at \mathbf{x} is given by

$$C_i(\mathbf{x}) = B \log_2 \left(1 + \frac{P_{Rx,i}(\mathbf{x})}{\sum_{\mathcal{B} \setminus \mathcal{B}_i} P_{Rx,j}(\mathbf{x}) + BN_0} \right) \quad [\text{bps}] \quad (2)$$

where B [Hz] is the bandwidth and N_0 [mW/Hz] is the additive white Gaussian noise (AWGN). Herein, we define a set of input parameters as $\phi = \{X_{Tx}, P_{Tx}\}$ where

$$X_{Tx} = \{\mathbf{x}_{Tx,i} \mid i = 1, 2, \dots, N_{Tx}\}, \quad (3)$$

$$P_{Tx} = \{P_{Tx,i} \mid i = 1, 2, \dots, N_{Tx}\}. \quad (4)$$

The objective is to maximize the average throughput per unit area [bps/m²] in the region of interest \mathcal{A} ; i.e.,

$$\underset{\phi}{\text{maximize}} \quad \left[f(\phi) := \frac{1}{|\mathcal{A}|} \sum_{i=1}^{N_{Tx}} \int_{\mathcal{A}} C_i(\mathbf{x}) d\mathbf{x} \right] \quad (5a)$$

$$\text{subject to} \quad 0 \leq P_{Tx,i} \leq P_{\max}, \quad \forall i, \quad (5b)$$

$$\mathbf{x}_{Tx,i} \in \mathcal{A}_i, \quad \forall i, \quad (5c)$$

where P_{\max} [mW] is the maximum transmission power.

Note that if the system can utilize multiple channels, the objective function can be extended to $f'(\phi) = \sum_{l=1}^{n_f} f_l(\phi)$ because of the orthogonality between channels, where n_f denotes the number of channels and f_l is the average throughput per unit area in the l -th channel. Based on the additive decomposition of the objective function [8], the proposed method can support this by optimizing the placement and transmission powers for each f_l by using the proposed framework. Since this extension is straightforward, this study focuses on improving the efficiency of a single channel.

Furthermore, the proposed method can be extended to a 3D environment without modification by redefining the 2D vector $\mathbf{x}_{Tx,i}$ as a 3D vector. However, this extension increases the number of parameters that must be tuned, which may slightly degrade the convergence performance with respect to the number of simulations.

III. ADAPTIVE EXPERIMENTAL DESIGN VIA BO

Before introducing the proposed framework, this section summarizes the BO. To present the BO separately from problem (5), we denote the set of input parameters to be optimized and the objective function as ψ and $h(\psi)$, respectively. The objective in the BO can be given by

$$\psi_{\text{opt}} = \underset{\psi \in \Psi_{\text{FR}}}{\text{argmax}} \quad h(\psi), \quad (6)$$

where Ψ_{FR} denotes the feasible region for the input parameters. The method can be divided into three steps: (i) initialization, (ii) Gaussian process regression (GPR), and (iii) parameter update based on acquisition function. We detail these steps as follows. This algorithm is summarized in Alg. 1.

1) *Initialization*: Initially, several simulation trials are performed to develop an initial dataset

$$\mathcal{D}^{(0)} = \{[\psi^{(t)}, y^{(t)}] \mid t = 1, 2, \dots, N_{\text{init}}\}, \quad (7)$$

where N_{init} indicates the number of initial trials and $\psi^{(t)}$ indicates the t -th initial input randomly selected from the feasible region. In addition, $y^{(t)} = h(\psi^{(t)}) + \epsilon$ is the observation value and ϵ is the observation noise.

2) *GPR*: This step estimates the distribution of the observation $p(y \mid \mathcal{D}^{(t)})$ based on GPR, a kernel-based non-parametric regression method that assumes a GP for the target function. To this end, we first tune a set of hyperparameters $\theta = \{\omega, \sigma_\epsilon\}$, where ω represents the hyperparameter vector for a kernel function k . Note that k measures the similarity between data points in a multi-dimensional space (e.g., the radial basis function (RBF)). The tuning of θ can be achieved by maximizing the following log-marginal likelihood:

$$\log p(\mathbf{y} \mid \mathbf{X}, \theta) = -\frac{1}{2}(\mathbf{y} - \mathbf{m})^T (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} (\mathbf{y} - \mathbf{m}) - \frac{1}{2} \log \det(\mathbf{K} + \sigma_\epsilon^2 \mathbf{I}) - \frac{|\mathcal{D}|}{2} \log 2\pi, \quad (8)$$

where \mathbf{I} is the $|\mathcal{D}| \times |\mathcal{D}|$ identity matrix and $\mathbf{K} \in \mathbb{R}^{|\mathcal{D}| \times |\mathcal{D}|}$ is the kernel matrix, where its element is $K_{ij} = k(\psi^{(i)}, \psi^{(j)})$. Further, \mathbf{m} is a vector with $(t + N_{\text{init}})$ elements, where its i -th element $m(\psi^{(i)})$ represents the prior mean at $\psi^{(i)}$; it can be estimated as $m(\psi^{(i)}) = \frac{1}{|\mathcal{D}|} \sum_{t=1}^{|\mathcal{D}|} y^{(t)}, \forall i$.

After the maximum likelihood estimation (MLE), the GPR predicts the distribution of the output for unknown inputs $\Psi_* = \{\psi_{*,i} \mid i = 1, 2, \dots, N_{\text{test}}\}$, where N_{test} indicates the number of unknown inputs that must be interpolated. For an unknown input ψ_* , the output mean and variance can be expressed by the following equations:

$$\mu(\psi_*) = m(\psi_*) + \mathbf{k}^T(\Psi, \psi_*) (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} (\mathbf{y} - \mathbf{m}) \quad (9)$$

$$\sigma^2(\psi_*) = k(\psi_*, \psi_*) - \mathbf{k}^T(\Psi, \psi_*) (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{k}(\Psi, \psi_*), \quad (10)$$

where $\Psi = \{\psi_i \mid i = 1, 2, \dots, t + N_{\text{init}}\}$, and

$$\mathbf{k}(\Psi, \psi_*) = [k(\psi_1, \psi_*), \dots, k(t + N_{\text{init}}, \psi_*)]^T. \quad (11)$$

Finally, we can predict the observation distribution at ψ_* as the normal distribution with mean $\mu(\psi_*)$ and variance $\sigma^2(\psi_*)$.

Algorithm 1 BO for adaptive experimental design

Require: $\mathcal{D}^{(0)} = \{[\psi^{(t)}, y^{(t)}] \mid t = 1, 2, \dots, N_{\text{init}}\}$

- 1: **for** $t = 1, 2, \dots, T$ **do**
- 2: Kernel tuning with MLE and $\mathcal{D}^{(t)}$
- 3: Randomly sample a set Ψ_*
- 4: Calculate μ and σ^2 for $\psi_{*,i} \in \Psi_*$ by Eqs. (9)(10)
- 5: Calculate α for $\psi_{*,i} \in \Psi_*$ by Eq. (12)
- 6: Select $\psi^{(t+1)}$ by Eq. (13).
- 7: Observe $y^{(t+1)} = h(\psi^{(t+1)}) + \epsilon$
- 8: Update the dataset as $\mathcal{D}^{(t+1)} = \mathcal{D}^{(t)} \cup \{[\psi^{(t+1)}, y^{(t+1)}]\}$
- 9: **end for**
- 10: **return** ψ_{opt} calculated by Eq. (14)

3) *Parameter Update:* Subsequently, this step estimates the acquisition function, which is a criterion for evaluating the goodness of the input parameters. We employed expected improvement (EI), which quantifies the anticipated performance gain from sampling a specific point by considering both the potential for discovering a better solution and the uncertainty in the model's predictions; i.e.,

$$\alpha(\psi) = \mathbb{E} [\max(y - y^+, 0)], \quad (12)$$

where y^+ denotes the maximum value of the dataset. Finally, the next input can be determined by

$$\psi^{(t+1)} = \underset{\psi \in \Psi_*}{\operatorname{argmax}} \alpha(\psi). \quad (13)$$

Steps 2 and 3 are repeated T times; the optimized input ψ_{opt} is $\psi^{(t)} \in \mathcal{D}^{(T)}$ which maximizes $y^{(t)}$; that is,

$$\psi_{\text{opt}} = \psi^{(t_{\text{best}})}, \quad t_{\text{best}} = \underset{t \in \{1, 2, \dots, |\mathcal{D}^{(T)}|\}}{\operatorname{argmax}} y^{(t)}. \quad (14)$$

IV. APPLYING BO TO WIRELESS SYSTEM DESIGN

BO models the set of observation values as a GP and interpolates the values at unobserved regions to identify better inputs. In practice, $P_{\text{Rx},i}$ follows the GP over the dBm domain spatially [7], suggesting that C_i can be approximated as GP at high signal-to-interference-plus-noise ratio (SINR). Considering this background, we apply the BO algorithm to solve problem (5).

Although Alg. 1 can simultaneously search for both \mathcal{X}_{Tx} and \mathcal{P}_{Tx} , this straightforward implementation would suffer from the curse of dimensionality as N_{Tx} increases. To avoid this problem, our framework introduces a nested BO architecture that separates the input parameters into two sets (Fig. 2). This approach focuses on the fact that the transmission power values \mathcal{P}_{Tx} can be optimized using a one-channel simulation for a given \mathcal{X}_{Tx} , which can improve the throughput performance in the BO without increasing the number of channel simulations.

As with pure BO, our framework is divided into three steps, as outlined below. This algorithm is summarized in Alg. 2.

1) *Initialization:* At the beginning of the t -th step, this method randomly samples a set of transmitter placements $\mathcal{X}_{\text{Tx}}^{(t)}$ and evaluates throughput performance. It then optimizes the transmission power values such that

$$\mathcal{P}_{\text{Tx,opt}}^{(t)} = \underset{\mathcal{P}_{\text{Tx}}}{\operatorname{argmax}} f(\mathcal{P}_{\text{Tx}} \mid \mathcal{X}_{\text{Tx}}^{(t)}). \quad (15)$$

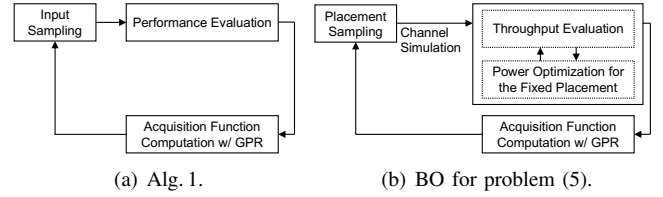


Fig. 2. Workflows in the pure BO and proposed framework.

This result can be obtained by Alg. 1 assuming that $\mathcal{X}_{\text{Tx}}^{(t)}$ is fixed.

In this framework, throughput performance $y^{(t)}$ can be obtained from a channel simulation. Assuming that the channel simulation result for the i -th transmitter can express the average channel gain at a pair of transmitter and receiver coordinates, its logarithmic form can be expressed as a function

$$g_i^{(t)}(\mathbf{x}) = -10\eta \log_{10} \left\| \mathbf{x}_{\text{Tx},i}^{(t)} - \mathbf{x} \right\| + 10 \log_{10} W(\mathbf{x}, \mathbf{x}_{\text{Tx},i}^{(t)}) + \epsilon_{\text{M}} \quad (16)$$

where ϵ_{M} is the simulation error and $\mathbf{x} \in \mathcal{A}$; after the simulation is performed for all possible transmitters. Accordingly, a set of channel simulation results can be constructed as

$$\mathcal{G}^{(t)} = \left\{ g_i^{(t)}(\mathbf{x}) \mid i = 1, 2, \dots, N_{\text{Tx}} \right\}. \quad (17)$$

Then, we obtain throughput $y^{(t)}$ using $\mathcal{P}_{\text{Tx,opt}}^{(t)}$ and $\mathcal{G}^{(t)}$, and Eq. (5a). Once $y^{(t)}$ is observed, the dataset is updated as

$$\mathcal{D}^{(t)} = \mathcal{D}^{(t-1)} \cup \{\phi^{(t)}, y^{(t)}\}, \quad (18)$$

where $\phi^{(t)} = \{\mathcal{X}_{\text{Tx}}^{(t)}, \mathcal{P}_{\text{Tx,opt}}^{(t)}\}$. These steps are repeated N_{init} times for the initial dataset.

2) *GPR:* To tune the set of placements, this step performs (a) MLE with $\mathcal{D}^{(t)}$ as with simultaneous BO, and (b) GPR over the set of placements \mathcal{X}_{Tx} . The GPR is performed for a set of unknown inputs $\mathcal{X}_* = \{\mathcal{X}_{*,i} \mid i = 1, 2, \dots, N_{\text{test}}\}$, and $p(y \mid \mathcal{D}^{(t)})$ is modeled for $\mathcal{X}_{\text{Tx}} \in \mathcal{X}_*$. As with Eqs. (9)(10), the mean and variance of the output at $\mathcal{X}_{*,i}$ can be given by

$$\mu(\mathcal{X}_{*,i}) = m(\mathcal{X}_{*,i}) + \mathbf{k}_{*,i}^{\text{T}} \left(\mathbf{K} + \sigma_{\epsilon}^2 \mathbf{I} \right)^{-1} (\mathbf{y} - \mathbf{m}) \quad (19)$$

$$\sigma^2(\mathcal{X}_{*,i}) = k(\mathcal{X}_{*,i}, \mathcal{X}_{*,i}) - \mathbf{k}_{*,i}^{\text{T}} \left(\mathbf{K} + \sigma_{\epsilon}^2 \mathbf{I} \right)^{-1} \mathbf{k}_{*,i}, \quad (20)$$

where $\mathbf{k}_{*,i} = k(\xi, \mathcal{X}_{*,i})$, and

$$\xi = \left\{ \mathcal{X}_{\text{Tx}}^{(i)} \mid i = 1, 2, \dots, t + N_{\text{init}} \right\}. \quad (21)$$

3) *Placement Update and Power Optimization:* This step calculates the acquisition function $\alpha(\mathcal{X}_{\text{Tx}})$ for $\mathcal{X}_{\text{Tx}} \in \mathcal{X}_*$ based on EI. The next set of placements can be then selected by

$$\mathcal{X}_{\text{Tx}}^{(t+1)} = \underset{\mathcal{X}_{\text{Tx}} \in \mathcal{X}_*}{\operatorname{argmax}} \alpha(\mathcal{X}_{\text{Tx}}). \quad (22)$$

After the channel information $\mathcal{G}^{(t+1)}$ is simulated, based on Alg. 1, we optimize the transmission power values such that

$$\mathcal{P}_{\text{Tx,opt}}^{(t+1)} = \underset{\mathcal{P}_{\text{Tx}}}{\operatorname{argmax}} f\left(\mathcal{P}_{\text{Tx}} \mid \mathcal{X}_{\text{Tx}}^{(t+1)}, \mathcal{G}^{(t+1)}\right). \quad (23)$$

Algorithm 2 BO for channel simulation-based BS design

Require: $\mathcal{D}^{(0)} = \emptyset$

- 1: **for** $t = 1, 2, \dots, N_{\text{init}}$ **do**
- 2: Randomly select the set of placements $\mathcal{X}_{\text{Tx}}^{(t)}$
- 3: Construct $\mathcal{G}^{(t)}$ by channel simulation.
- 4: Find $\mathcal{P}_{\text{Tx,opt}}^{(t)}$ by Eq. (15) and Alg. 1
- 5: Set $\phi^{(t)} = \left\{ \mathcal{X}_{\text{Tx}}^{(t)}, \mathcal{P}_{\text{Tx,opt}}^{(t)} \right\}$
- 6: Observe throughput $y^{(t)}$ by $\mathcal{P}_{\text{Tx,opt}}^{(t)}$ and $\mathcal{G}^{(t)}$.
- 7: Update $\mathcal{D}^{(t-1)}$ to $\mathcal{D}^{(t)}$ by Eq. (18)
- 8: **end for**
- 9: **for** $t = 1, 2, \dots, T$ **do**
- 10: Kernel tuning with MLE and $\mathcal{D}^{(t)}$
- 11: Sample $\mathcal{X}_* = \{ \mathcal{X}_{*,i} \mid i = 1, 2, \dots, N_{\text{test}} \}$
- 12: Calculate μ and σ^2 for $\mathcal{X}_{*,i} \in \mathcal{X}_*$ by Eqs. (19)(20)
- 13: Calculate α for $\mathcal{X}_{*,i} \in \mathcal{X}_*$ by Eq. (12)
- 14: Select $\mathcal{X}_{\text{Tx}}^{(t+1)}$ by Eq. (22)
- 15: Simulate $\mathcal{G}^{(t+1)}$ for $\mathcal{X}_{\text{Tx}}^{(t+1)}$
- 16: Find $\mathcal{P}_{\text{Tx,opt}}^{(t+1)}$ by Eq. (23) and Alg. 1
- 17: Observe throughput $y^{(t+1)}$ by $\mathcal{P}_{\text{Tx,opt}}^{(t+1)}$ and $\mathcal{G}^{(t+1)}$.
- 18: Update $\mathcal{D}^{(t)}$ to $\mathcal{D}^{(t+1)}$ by Eq. (24)
- 19: **end for**
- 20: **return** ϕ_{opt} calculated by Eq. (25)

Then, the dataset is updated to

$$\mathcal{D}^{(t+1)} = \mathcal{D}^{(t)} \cup \{ \phi^{(t+1)}, y^{(t+1)} \}. \quad (24)$$

Steps 2 and 3 are repeated T times. Finally, the optimized input ϕ_{opt} is $\phi^{(t)} \in \mathcal{D}^{(T)}$ that maximizes $y^{(t)}$: this operation can be given by,

$$\phi_{\text{opt}} = \phi^{(t_{\text{best}})}, t_{\text{best}} = \underset{t \in \{1, 2, \dots, |\mathcal{D}^{(T)}\}}{\text{argmax}} y^{(t)}. \quad (25)$$

V. PERFORMANCE EVALUATION

The proposed framework was implemented using Optuna 3.2.0 [9], BoTorch 0.9.2 [10], and Python 3.10.6. Furthermore, the performance of the proposed framework (Alg.2) was compared with that of the following three baseline methods.

- **Placement Optimization:** Assuming the maximum transmission power, this method optimizes \mathcal{X}_{Tx} via Alg. 1.
- **Regular Placement with Power Optimization:** This method determines \mathcal{X}_{Tx} as a regular hexagonal grid. It then optimizes \mathcal{P}_{Tx} based on Alg. 1.
- **Pure BO:** This method simultaneously searches for \mathcal{X}_{Tx} and \mathcal{P}_{Tx} based on Alg. 1 only.
- **Random Search:** This method samples the parameters in ϕ based on the uniform distribution.

The simulation parameters are listed in Table I. We simulated wireless channels based on Eq. (1). Shadowing exhibits spatial correlations that depend on the extent of change in the coordinates of the transmitter and receiver [7]. To consider this effect, we assume that shadowing has a spatial correlation with the standard deviation σ_S [dB]. Note that although it can be simulated by a multivariate log-normal distribution, generating it using a pure multivariate log-normal distribution

TABLE I
SIMULATION PARAMETERS.

Parameter	Detail
Kernel function k	RBF kernel without scaling factor
Optimizer for MLE	L-BFGS-B [11]
N_{init} and N_{test}	8 and 256
Size of \mathcal{A}	1000 [m] \times 1000 [m]
Range of \mathcal{A}_i	$[0, 1000]$ (x axis), $\left[\frac{1000}{N_{\text{Tx}}}i, \frac{1000}{N_{\text{Tx}}}(i+1) \right]$ (y axis)
Grid size	20 [m] \times 20 [m]
η	4.0
d_{cor}	200 [m]
σ_S	6.0 [dB]
P_{max}	10 [mW]
N_0	-174 [dBm/Hz]
B	20 [MHz]

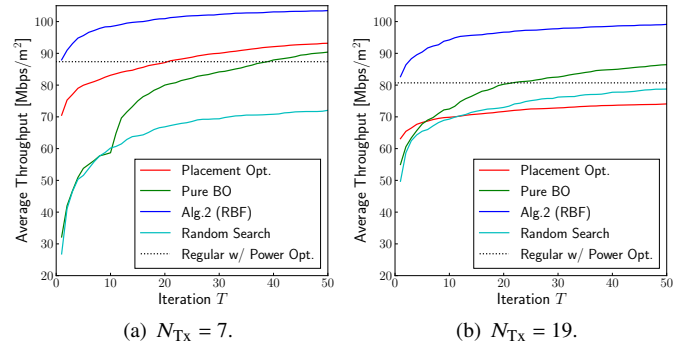


Fig. 3. Effects of the number of iterations.

in our simulation environment is computationally expensive. Thus, we generated the shadowing based on the following: (i) allocating N_s points on \mathcal{A} based on the Poisson point process (PPP), (ii) assigning i.i.d. log-normal random variables to the allocated points, and (iii) interpolating shadowing values on the missing coordinates based on the nearest neighborhood search. The point density in PPP is designed such that the nearest neighbor distance r satisfies $\mathbb{E}[r] = \mathbb{E}[d_{\text{cor}}]$. Note that we define the distance between a pair of wireless links as $d_{ij} \triangleq \sqrt{\| \mathbf{x}_{\text{Tx},i} - \mathbf{x}_{\text{Tx},j} \|^2 + \| \mathbf{x}_{\text{Rx},i} - \mathbf{x}_{\text{Rx},j} \|^2}$.

Fig. 3 shows the effects of T on the average throughput performance. Assuming the errorless channel simulation, we iterated the independent simulations 100 times and calculated the average throughput for each condition. The number of transmitters was set to $N_{\text{Tx}} = 7$ (Fig.3(a)) and $N_{\text{Tx}} = 19$ (Fig. 3(b)). For $N_{\text{Tx}} = 7$, the average throughput in the regular placement was 87.4 Mbps/m². The other methods improved the throughput values as the number of iterations increased. At $T = 50$, each method achieves the following throughput values: 93.2 Mbps/m² (placement optimization), 90.4 Mbps/m² (pure BO), 103.5 Mbps/m² (proposed method), and 72.1 Mbps/m² (random search). Furthermore, placement optimization and pure BO outperformed the regular placement method at $T \geq 21$ and $T \geq 39$, respectively. The proposed framework resulted in a higher throughput than these methods. It outperformed regular placement, even with fewer iterations, and an 18.4% improvement was confirmed at $T = 50$. Moreover, the random search method is inferior to the other methods.

Similar to the case of $N_{\text{Tx}} = 7$, an improvement in throughput with an increasing number of iterations was also

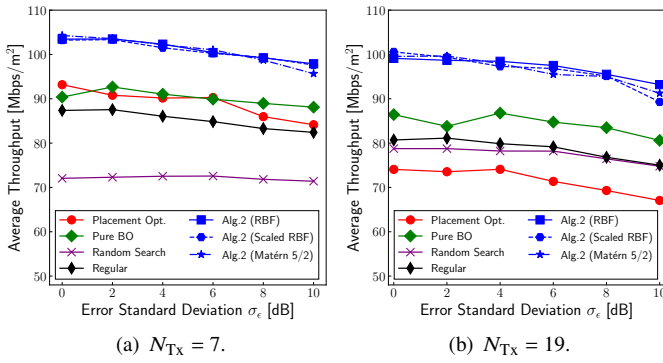


Fig. 4. Effects of channel simulation error at $T = 50$.

observed for $N_{Tx} = 19$. In contrast to the case of $N_{Tx} = 7$, the performance of the placement optimization was inferior to that of the regular placement method and random search for $N_{Tx} = 19$. For $N_{Tx} = 7$, the number of transmitters was small relative to the area size; a higher transmission power tends to improve the throughput performance in this case. However, as N_{Tx} increases, the interference between the transmitters becomes more significant, and the throughput deteriorates. Pure BO and the proposed framework outperformed the other methods. Specifically, our framework improved throughput by 22.8% compared to the regular placement method at $T = 50$.

Overall, for both $N_{Tx} = 7$ and 19, the proposed methods achieved a higher throughput performance with a limited number of simulations compared with the baseline methods.

We then evaluated the effects of the observation error on the average throughput performance to discuss more practical performance¹. We model the error factor ϵ_M in Eq. (16) as a random variable following the normal distribution $\mathcal{N}(0, \sigma_\epsilon^2)$, where σ_ϵ [dB] is the standard deviation. Fig. 4 shows the effects of the standard deviation σ_ϵ on the average throughput performance where $N_{Tx} = 7$ (Fig. 4(a)) and 19 (Fig. 4(b)). We changed σ_ϵ from zero (noiseless) to 10 and evaluated the average throughput at $T = 50$. Furthermore, to confirm effects of kernel function, we also evaluated the proposed method with Matérn 5/2 kernel and scaled RBF kernel.

The throughput of all methods decreases as the noise standard deviation increases. For the proposed method (with the RBF kernel), the throughput decreased by 5.4% at $N_{Tx} = 7$ and by 6.0% at $N_{Tx} = 19$ compared to the noiseless case when $\sigma_\epsilon = 10$. However, the proposed framework achieves better throughput than the baseline methods, as in the noiseless case. For example, for $N_{Tx} = 7$, the throughput was approximately 18.8% higher than that of the regular placement at $\sigma_\epsilon = 10$. By modeling the observation with noise (i.e., $y^{(t)} = f(\phi^{(t)}) + \epsilon$) and performing black-box optimization while considering the uncertainties, BO facilitates error-robust transmission power and placement designs.

Focusing on the impact of the kernel in the proposed method, similar trends were observed for all types of kernels.

¹For example, based on [12], the RMSE of the ray-tracing simulation in sub-6GHz outdoor scenarios was approximately 4-13 dB. Furthermore, it is necessary to average the samples that are within a half-wavelength radius around a certain reception point to mitigate the effects of multipath fading; however, this can cause several errors in shadowing estimation.

In particular, at $\sigma_\epsilon = 10$, the RBF without scaling exhibited the best performance, achieving 103.5 Mbps/m² at $N_{Tx} = 7$ and 99.1 Mbps/m² at $N_{Tx} = 19$. This kernel has fewer trainable parameters compared to other kernel functions. When the relationship between samples can be sufficiently modeled with this kernel alone, it tends to exhibit better convergence properties in comparison to more complex kernels. Therefore, the RBF without scaling is a reasonable choice for this problem.

VI. CONCLUSION

We proposed a BO framework for channel simulation-based joint placement and transmission power design. The results demonstrate that our proposed framework can achieve 18-22% higher throughput performance compared to conventional strategies at $T = 50$; further, it works even though the channel simulation includes the computation error. The proposed approach can help system operators design system parameters before deployment, for example, cell planning in 6G networks or access point deployment in public Wi-Fi.

Note that the proposed framework utilizes the spatial correlation of throughput performances to search for optimal BS settings. It depends on the spatial correlation of shadowing; thus, its performance improvement is limited under short correlation distance environments, including indoor scenarios with super-high-frequency (SHF) systems. Furthermore, this study assumed isotropic antennas and static transmission positions for the BSs. Introducing other adjustable parameters, such as beamformer and path planning in unmanned aerial vehicle (UAV)-based BSs, will allow further throughput; however, it increases the input dimensionality in the BO. The BO design for more parameters remains a topic for future work.

REFERENCES

- [1] Y. L. Lee et al., "6G massive radio access networks: Key applications, requirements and challenges," *IEEE Open J. Veh. Technol.*, vol. 2, pp. 54-66, 2021.
- [2] H. Yu et al., "Public Wi-Fi monetization via advertising," *IEEE/ACM Trans. Netw.*, vol. 25, no. 4, pp. 2110-2121, Aug. 2017.
- [3] C.-W. Wu et al., "Enhancing fan engagement in a 5G stadium with AI-based technologies and live streaming," *IEEE Syst. J.*, vol. 16, no. 4, pp. 6590-6601, Dec. 2022.
- [4] S. Greenhill et al., "Bayesian optimization for adaptive experimental design: A review," *IEEE Access*, vol. 8, pp. 13937-13948, 2020.
- [5] J. Yan et al., "Bayesian optimization for online management in dynamic mobile edge computing," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3425-3436, Apr. 2024.
- [6] Y. Zhang et al., "Bayesian and multi-armed contextual meta-optimization for efficient wireless radio resource management," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 5, pp. 1282-1295, Oct. 2023.
- [7] Z. Wang et al., "Joint shadowing process in urban peer-to-peer radio channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 52-64, Jan. 2008.
- [8] P. Rolland et al., "High-dimensional Bayesian optimization via additive models with overlapping groups," in *Proc. AISTATS 2018*, Lanzarote, Spain, Apr. 2018, pp. 298-307.
- [9] T. Akiba et al., "Optuna: A next-generation hyperparameter optimization framework," in *Proc. ACM SIGKDD'19*, Anchorage, AK, USA, 2019, pp. 2623-2631.
- [10] M. Balandat et al., "BoTorch: A framework for efficient Monte-Carlo Bayesian optimization," in *Proc. NeurIPS'20*, Vancouver, BC, Canada, 2020.
- [11] C. Zhu et al., "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, no. 4, pp. 550-560, Dec. 1997.
- [12] A. W. Mbugua et al., "Review on ray tracing channel simulation accuracy in sub-6GHz outdoor deployment scenarios," *IEEE Open J. Antennas Propag.*, vol. 2, pp. 22-37, 2021.