

AI-assisted Indoor Wireless Network Planning with Data-Driven Propagation Models

Stefanos Bakirtzis, *Member, IEEE*, Ian Wassell,
Marco Fiore, *Senior Member, IEEE* and Jie Zhang, *Senior Member, IEEE*

Abstract—Propelled by rapid advances in artificial intelligence (AI), the design and operation of 5G and beyond networks are anticipated to be radically different from those of legacy communication systems. Indeed, AI can be exploited to automate and optimize various essential functionalities of the wireless ecosystem, such as resource allocation, channel modeling, or network planning. This article explores how AI-driven propagation models can be leveraged for the automated and expedient deployment of small cells in indoor environments. To this end, we couple a generalizable data-driven propagation model with an AI-based optimizer to determine the optimal network topology with respect to a target key performance indicator. Our approach reduces the computational time of indoor wireless network design by two to three orders of magnitude, thus enabling accurate planning that would be extremely expensive to conduct using conventional indoor propagation tools and yielding significant gains in the resulting indoor planning quality and performance.

Index Terms—Ultra-dense networks, artificial intelligence, deep learning, network planning, propagation modeling, 5G.

I. INTRODUCTION

FIFTH-generation (5G) and beyond (B5G) mobile networks are expected to support a plethora of applications with diverse requirements in terms of data rate, quality of service, and reliability, along with an unprecedented increase in the users' wireless traffic demands, rendering network planning and optimization of 5G/B5G networks a complicated and multifarious procedure. The problem of robust and efficient network planning is especially critical in ultra-dense networks (UDNs), which can effectively improve the coverage, the spectral, and the energy efficiency of 5G/B5G systems, by deploying a large number of small cells within the existing radio access network infrastructure **Cell-Free MIMO**, [1].

Ultra-dense deployments of small cells will primarily target indoor environments where the signal coverage of the outdoor base stations (BSs) deteriorates significantly. These are highly relevant scenarios as the largest part of the wireless communications traffic is generated by indoor users [2]. Thus, it is necessary to have reliable tools that can assist the effective deployment of new UDN components by creating realistic digital replicas of the expanding wireless ecosystem. Specifically, such tools will provide the means to proactively configure the wireless network through high-fidelity digital twins that can be exploited to assess network performance in advance of the real network operation and ensure the optimal orchestration of its services [3].

From a planning perspective, UDNs are characterized by very large degrees of freedom in the reciprocal locations of a large number of small cells, and a diversity of possible target performance indicators to be optimized. Thus,

existing strategies for UDN planning based on ray tracing techniques [4], [5], [6] are unfit for the task, as they can only explore a limited number of network topologies, leading to deployments that are largely sub-optimal with respect to what could be achieved in practice. This calls for optimization frameworks that can efficiently and automatically explore the vast space of feasible network topologies. Formally, near-optimal UDN planning requires enhancing one or more of three components: (i) a radio propagation modeling tool; (ii) a suitable target key performance indicator (KPI) metric that quantitatively represents the performance of the deployment; (iii) an optimization problem formulation that can return a high-performing deployment of dense femto-base stations (fBSs) [1].

Notably, the selection of the propagation solver for the generation of radio maps assumes a crucial role in the efficiency and quality of the optimization process [7]. On one hand, deployment frameworks leveraging empirical propagation tools, e.g. log-distance model, can generate radio maps rapidly, yet the fidelity of these maps is questionable and can jeopardize the effectiveness of the final deployment [8], [9]. On the other hand, deterministic propagation solvers generate realistic radio maps via solutions that abide by the fundamental laws of physics, but they are computationally demanding. Consequently, this severely impacts the optimization computational cost and compromises the search space process since the solvers are invoked a large number of times by iterative optimization algorithms to gradually adapt the network topology towards meeting the target KPI.

This is precisely why computationally expensive full-wave or asymptotic solvers, e.g., ray tracing or numerical solvers of Maxwell's equations, which require minutes or even hours for a single simulation cannot support thorough explorations of the deployment space [4], [5], [6], [10]. The problem can be severe to the point that UDN planning with deterministic propagation solvers becomes computationally unfeasible using common optimization techniques [11] even in relatively small-scale scenarios with a limited number of fBSs. For instance, the use of a ray-tracer or a full-wave solver led to prolonged optimization times in [6] and [10], respectively, while in [4], [5], the capacity, spectrum, and energy efficiency of a UDN was only evaluated for a small set of predetermined deployments owing to the use of ray tracing-generated radio maps.

In this paper, we propose an original approach for indoor UDN planning, which leverages recent advances in machine learning (ML)-based propagation modeling [12]. These models hinge upon a set of descriptive indoor environment features that allow an artificial intelligence (AI) module agnostic to

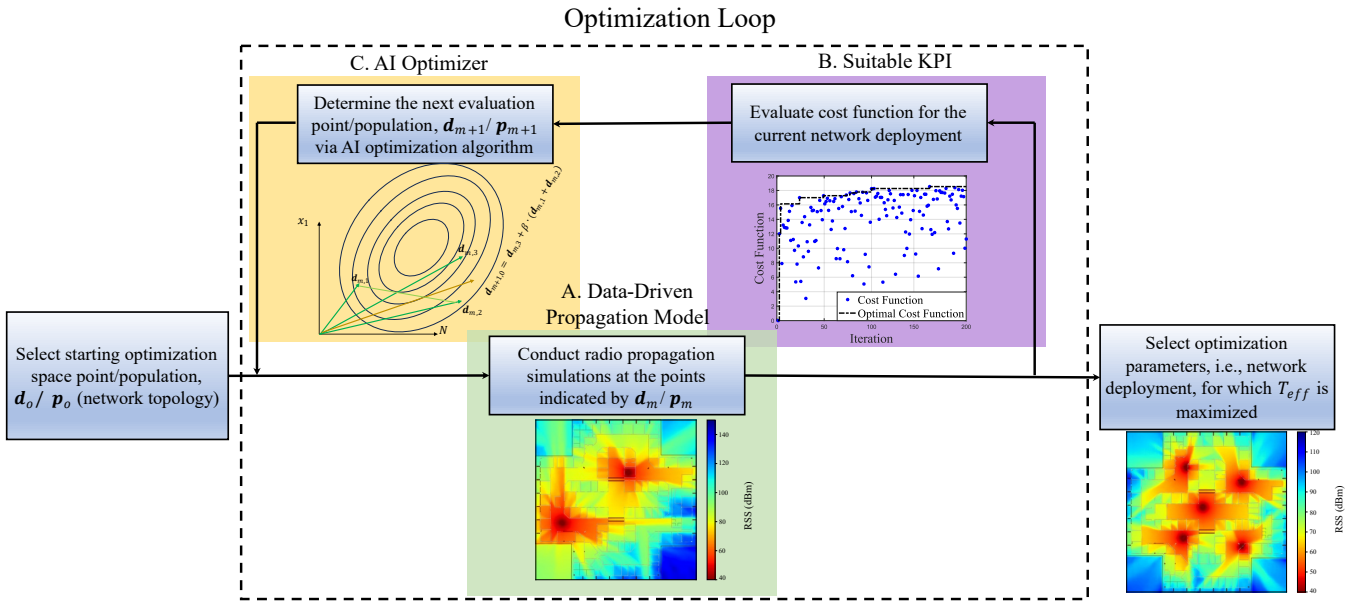


Fig. 1: Complete framework for AI-assisted indoor wireless network planning with a data-driven propagation model, including (A) an expedient data-driven propagation model leveraged to readily emulate wireless signal propagation (green box), (B) evaluation of the metric describing the performance of the network (purple box), and (C) a *black-box* computational intelligence optimizer that guides the optimization space point selection by observing the harvested cost function for each point of the optimization space, i.e., possible deployment (orange box). Starting from a randomly selected deployment (left), the network performance is measured iteratively for deployments indicated by the computational intelligence algorithm, ultimately leading to a nearly-optimal network design (right).

the laws of electromagnetism to develop an understanding of how such features affect radio wave propagation. Commonly, ML algorithms are used to process this information and find a mapping that transforms it to a physical quantity of interest, e.g., path loss (PL). The estimates of the target quantity can be real-world measurements or synthetic data generated by a high-performance propagation solver. In the first case, the aim is to make data-driven solvers more realistic than legacy ones, whilst in the second, the aim is to overcome the computational barrier imposed by conventional deterministic models. Due to the scarcity of measured data and the subsequent challenges in training, most existing data-driven propagation models use synthetic data; once trained offline, they can be used as standalone propagation models that replicate the radio maps of high-performance solvers in a dramatically reduced computational time. Consequently, they are ideal tools for network planning-related tasks, as they can narrow the efficiency-accuracy gap exhibited by conventional radio propagation modeling tools [7].

However, while prior works have discussed how outdoor or indoor data-driven propagation models can be created [12], their practical advantages on wireless network-related applications have not yet been thoroughly addressed. This article explores the integration of fully-fledged high-fidelity data-driven propagation models in automated network design, factually introducing a new class of planners based on the coupling of data-driven propagation solvers and AI-based optimizers [11]. We show that our proposed strategy can overcome the severe computational impediments of legacy ray-tracers, reducing planning time by two to three orders of magnitude compared to current state-of-the-art approaches. In turn, this allows exploring deployment options much more rapidly, hence supporting

computational intelligence optimizers with higher performance as well as scaling to larger scenarios. By speeding up the planning process, and providing higher flexibility in terms of deployment choices, one can identify low-cost deployments that can increase network spectral efficiency, ensuring higher-performance network operation. Ultimately, our study paves the way for automated AI-powered network design solutions by demonstrating how legacy radio propagation solvers can be eventually replaced by data-driven ones for expedient and high-quality UDN planning.

II. UDN PLANNING WITH DATA-DRIVEN MODELS

The framework for the proposed AI-assisted indoor UDN planning is illustrated in Fig. 1, comprising three sub-components: (A) a data-driven propagation model used to swiftly evaluate signal distribution throughout the building (green sub-box), (B) an appropriate KPI that quantifies the performance of the network topology (purple sub-box), and (C) an AI-based optimizer used to explore potential network topologies (orange box). Note that the proposed strategy is general and can accommodate any implementation of subcomponents (A)-(C). Yet, our work advances the state-of-the-art in network planning by exploiting a generalizable data-driven propagation model, which enables conducting accurate and unprecedentedly fast simulations of wireless signal levels.

A. Data-driven propagation solver

We employ a data-driven propagation model, called Electromagnetic DeepRay (EM DeepRay) [13], which uses four input features as shown in Fig. 2. Each feature is represented as a distinct tensor channel depicting: (i) the relative permittivity, ϵ_r , (ii) the conductivity, σ , of the building materials at each

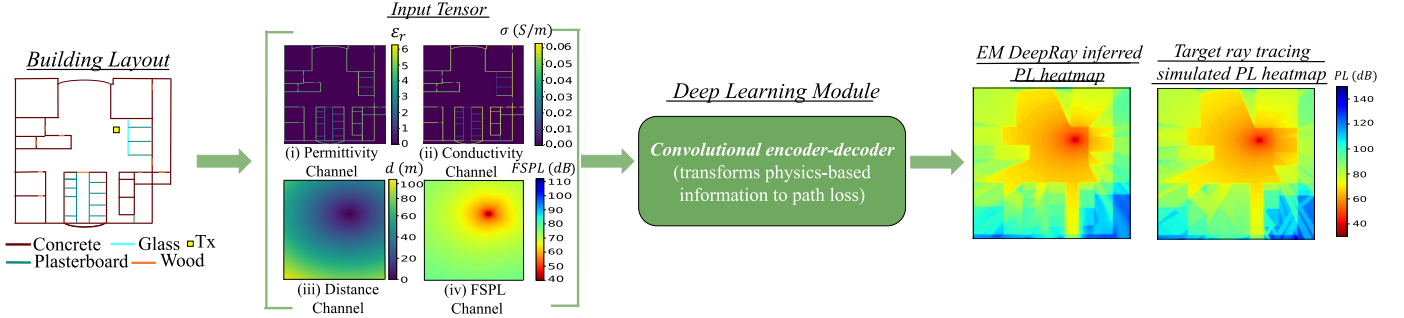


Fig. 2: Schematic of a physics-based data-driven propagation model inferring signal attenuation. The model receives at its input a tensor conveying some characteristic features of the propagation scenario (wall permittivity & conductivity, distance, and FSPL), which is encoded through consecutive convolutions and decoded as a signal attenuation heatmap for the target fBS.

point of the simulated grid, *(iii)* the distance, d , between the transmitter and every point within the simulated grid, *(iv)* and the free space path loss (FSPL). The propagation environment description, i.e., channels *(i)* and *(ii)*, is a common type of information used by propagation solvers, that is typically acquired through dedicated floor plan design software. This software digitizes blueprints or takes LIDAR measurements to generate computer-aided design floor plans. In general, the more detailed the description, the higher the fidelity of the propagation solver.

The input channels are concatenated and passed to a convolutional encoder-decoder which applies consecutive non-linear learnable convolutional operations to the physics-based input tensor and transforms it into a PL heatmap. After exploring various ML modules, we choose to use a U-Net with stacked dilated convolutions as our backbone encoder-decoder, and we refer the reader to [13] for its implementation details. During the training phase, the weights of the convolutional operations are computed so that the error between the predicted PL heatmap of EM DeepRay and that produced by a ray-tracer is minimized. Then, at the inference stage, with the weights properly tuned, the same transformations can be applied to reliably estimate the PL distribution for any target indoor environment and frequency.

The model used in this work assumes the same architecture as in [13], but it is trained and tested over a significantly larger volume of data. In particular, EM DeepRay is trained with approximately 100,000 samples derived by conducting ray tracing simulations at 100 diverse indoor environments in various frequency bands, and its accuracy is tested over 25,000 samples coming from 30 indoor geometries not seen during the training phase. For the unknown indoor environment layouts and frequencies, its predictions yielded a 3.09 dB mean absolute error with respect to the ground truth ray tracing simulation results. However, unlike the time-consuming ray tracing approach which requires a few minutes for a single simulation, the PL inference via EM DeepRay requires only a few milliseconds, including the time required to extract the physics-based input tensor. A comparison between the simulated PL via ray tracing and that predicted with EM DeepRay for an unknown geometry is shown in Fig. 2, demonstrating close correspondence between the two.

The one-time monetary cost of training EM DeepRay includes *(i)* the cost of simulating the ground-truth propagation

data, and *(ii)* the time needed to train the model. Through an automated pipeline, the ground-truth data generation process lasted approximately two weeks in a workstation equipped with an Intel Core i9-9960X X-series Processor with 16 central processing units and two 24 GB Nvidia Titan RTX graphics processing units. The offline training process itself took approximately 48 hours. It is paramount to note that these are non-recurring costs: once trained, EM DeepRay can be leveraged directly as a standalone, accurate, and extremely fast propagation solver for unknown propagation scenarios not seen during training, i.e., it does not require additional data generation or training to make predictions for new scenarios.

B. Indoor Wireless Network Performance Metric

A key component of UDN planning is the identification of the relevant KPI that the femtocell deployment shall optimize. For our analysis, we rely on one of a set of metrics introduced in [14] to evaluate indoor wireless network performance (IWNP). These metrics represent compound measures that can be used by telecommunication practitioners and wireless engineers to assess the effectiveness of a network topology and guide network deployment. Specifically, the cost function used in this work is the throughput effectiveness metric, T_{eff} , measured in Mbps, defined over all locations ω in the area Ω of an indoor environment:

$$T_{eff} = \frac{1}{\Omega} \int_{\Omega} \Xi \left(\frac{S(\omega)}{\sum_{i \in O} I_i'(\omega) + \sum_{j \in A} I_j(\omega) + N_{ch}} \right) w(\omega) d\omega, \quad (1)$$

where Ξ is a mapping function, $S(\omega)$ represents the signal from the serving cell, $I_i'(\omega)$ indicates co-channel interfering received power from i -th macro BS from the set O of the outdoor macro BSs, $I_j(\omega)$ corresponds to the co-channel interfering received power from the j -th fBSs from the set A of the indoor BSs, N_{ch} is the channel noise and $w(\omega)$ is a weighting function. To compute the throughput for a given 5G channel configuration, we use the 3GPP Technical Specification (TS) 38.306 [14]. Specifically, the function Ξ accounts for the channel conditions and maps the simulated signal-to-interference-plus-noise ratio onto the channel quality indicator. This designates the attainable spectral efficiency of the system, achieved by using the best possible modulation order and coding rate. The weighting function is used to control the contribution of certain building areas toward the

compound metric. The calculation of (1) requires knowledge of the signal distribution over space, which can be determined via a propagation model. A data-driven propagation model, such as EM DeepRay, can bridge the efficiency-accuracy gap that conventional propagation models exhibit [7], providing reliable PL estimates in a computationally affordable manner.

C. Optimization

To determine the UDN deployment that optimizes the KPI relying on ML-based radio signal estimates, we need to explore the solution space, i.e., the number of fBS and their placement, in a smart way. Therefore, the last component of our proof-of-concept UDN planning framework using data-driven propagation models, is an optimizer. We explore different optimization tools, which are employed in cases where the optimization objective function, f , does not have a known special structure, but it is an unconstrained *black-box*. In particular, we consider the following state-of-the-art AI optimization algorithms: (i) Bayesian optimization (BO), (ii) genetic algorithms (GA), (iii) particle swarm optimization (PSO), and (iv) differential evolution (DE) [11]. BO is a non-population-based optimizer, i.e., at each iteration, the cost function is evaluated at a single point, \mathbf{d} , of the multi-dimension optimization space. The other algorithms are population-based, retrieving at each iteration K different solutions by evaluating the cost function at a set of K distinct optimization space points, referred to as the population, \mathbf{p} . The goal of each optimizer is to find a global optimum point, and as shown in Fig. 1, a similar procedure is followed irrespective of the optimization algorithm.

In the subscript (m, k) let m indicate the iteration of the optimization algorithm, whilst for population-based methods the subscript k designates a specific optimization point from the population. Initially, a starting point, \mathbf{d}_o , comprising the optimization parameters, or a population of K starting points, $\mathbf{p}_o = [\mathbf{d}_{o,1}, \mathbf{d}_{o,2}, \dots, \mathbf{d}_{o,K}]$, is randomly selected and the cost function f is evaluated at it. Then, based on the computed values of f , the optimization algorithm is employed to determine the point or set of points from the optimization space, \mathbf{d}_{m+1} or \mathbf{p}_{m+1} , at which f will be evaluated in the next iteration. This process is repeated iteratively until a convergence criterion is met or a certain maximum number of iterations is reached. The difference between the various optimization algorithms lies in the way \mathbf{d}_{m+1} or \mathbf{p}_{m+1} are chosen [11]. Ultimately, rather than conducting brute-force simulations, and scanning the entire optimization space, the AI optimization algorithms allow the points at which f will be evaluated to be chosen in a smart manner.

III. EVALUATION IN PRACTICAL USE CASES

To showcase our approach, we emulate the digital twin of a city and the UDN deployed in it. Specifically, we generate the urban environment shown in Fig. 3, and we consider the digital representation of two indoor environments in which fBSs will be installed. Then, an outdoor-to-indoor ray tracing simulation is conducted to derive the signal heatmaps within the indoor environments owing to the 4 outdoor macro BSs shown as

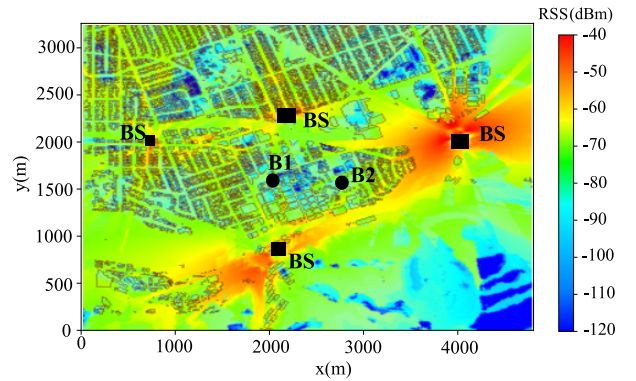


Fig. 3: Digital twin of an urban environment used to conduct an outdoor-to-indoor ray tracing simulation, providing the signal heatmaps of four macro BSs (black squares) within the digital replicas of two buildings, B1 and B2 (black dots), where fBSs will be deployed.

black squares. The buildings are shown as black dots and are annotated as B1 and B2 in Fig. 3. They are constructed using various materials, and neither of them has been used during the training phase of EM DeepRay.

To simulate the outdoor-to-indoor received signal strength (RSS), the Ranplan Wireless ray tracing simulator is used; we remark that this reference is not affected by the UDN deployment, hence it is only calculated once, and computing it via an expensive ray tracing approach is affordable. The macro BSs use directional antennas with 18 dBi gain, and a half power beam width equal to 57° and 15° in the horizontal and the vertical planes, respectively. The fBSs are equipped with antennas having an omnidirectional beam pattern and a 3 dBi antenna gain. The transmission power for the macro BSs and the fBSs is constant and equal to 49.4 and 20 dBm, respectively. Both the macro BSs and the fBSs operate at 2 GHz over a 5 MHz bandwidth channel. The sub-carrier spacing is 15 kHz, and the number of resource blocks is 25. For each point within the simulated domain, the fBS with the strongest signal is selected as the serving one. We assume that there is no interference between the fBSs due to the use of a distributed antenna system or interference mitigation techniques, such as coordinated multipoint, and that interference is only introduced by the outdoor macro BS to the serving fBS. To calculate the channel noise, we assume a constant noise power spectral density equal to -174 dBm/Hz and a noise figure at the receiving signal chain equal to 5 dB.

The optimization returns the number N of fBSs to be deployed, along with their position within the building, i.e., $\mathbf{d} = [N, x^{(1)}, y^{(1)}, \dots, x^{(N)}, y^{(N)}]$, where $x^{(n)}$ and $y^{(n)}$ are the Cartesian coordinates of the n -th fBS. For the illustrative experiments carried out in this paper, we bound the maximum number of fBSs to 5, as this is already sufficient to demonstrate the practical computational advantage of the proposed approach. As shown in Fig. 1, at the m -th iteration, the optimizer designates the number, N_m , of fBSs and their positions. Then, N_m radio propagation simulations are conducted to evaluate the received signal throughout the building, one for each one of the N_m fBSs. For population-based methods, for each optimization point $\mathbf{d}_{m,k}$ in the population \mathbf{p}_m , we conduct

TABLE I: OPTIMIZATION PERFORMANCE STATISTICS.

| | Metric (Mbps) & Cost (sec) | B1 | | | B2 | | |
|-----|-------------------------------|----------------|---------------------|------------------|-----------------|---------------------|------------------|
| | | Ray Tracing | ML-based model | Hybrid | Ray Tracing | ML-based model | Hybrid |
| BO | T_{eff} (Mbps) | 15.04 | 16.16 \pm 0.74 | 16.02 \pm 0.73 | 17.59 | 19.95 \pm 0.55 | 17.04 \pm 0.75 |
| | CE | 149.00 | 108.63 \pm 31.73 | - | 87.00 | 98.81 \pm 33.50 | - |
| | CT (sec) | 23,040 (6.4 h) | 134.78 \pm 21.61 | - | 62,280 (17.2 h) | 143.27 \pm 67.64 | - |
| GA | T_{eff} (Mbps) | - | 15.18 \pm 1.92 | 14.30 \pm 2.43 | - | 20.70 \pm 0.34 | 17.02 \pm 0.63 |
| | CE | - | 980.78 \pm 139.61 | - | - | 977.45 \pm 20.18 | - |
| | CT (sec) | - | 352.62 \pm 99.75 | - | - | 678.44 \pm 107.59 | - |
| PSO | T_{eff} (Mbps) | - | 15.61 \pm 0.90 | 15.01 \pm 1.21 | - | 20.67 \pm 0.48 | 17.62 \pm 0.65 |
| | CE | - | 475.09 \pm 134.60 | - | - | 553.81 \pm 47.14 | - |
| | CT (sec) | - | 257.03 \pm 74.68 | - | - | 353.96 \pm 49.13 | - |
| DE | T_{eff} (Mbps) | - | 16.49 \pm 0.81 | 16.25 \pm 1.03 | - | 20.78 \pm 0.32 | 17.96 \pm 0.75 |
| | CE | - | 236.27 \pm 75.01 | - | - | 183.54 \pm 60.98 | - |
| | CT (sec) | - | 182.47 \pm 63.70 | - | - | 110.33 \pm 40.18 | - |

$N_{m,k}$ simulations, where $N_{m,k}$ indicates the number of fBSs for the k -th optimization point of the population at the m -th iteration. Consequently, the value of T_{eff} is computed for the current network topology, or the K different network topologies in population-based methods, and it is employed to determine the optimization point/population for the next iteration according to the optimization scheme used.

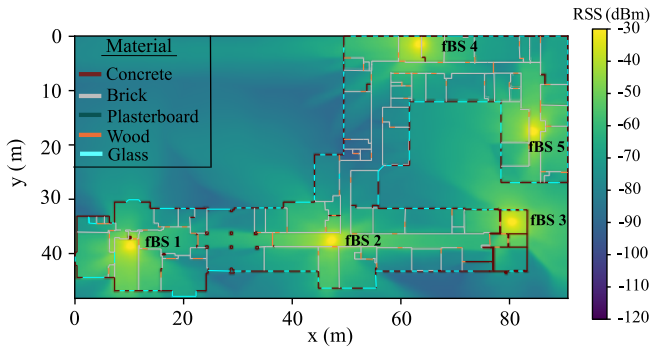
The optimization setup for each algorithm is as follows; for BO we use the expected improvement acquisition function and we allow up to 200 iterations. For GA we allow 42 iterations, whilst at each iteration the population comprises 8 distinct points of the optimization space, out of which 4 are chosen as parents to determine the points of the next population, using a steady-state selection scheme. A scattered crossover operation, with a crossover probability 0.8, is applied to the parents, and a random mutation scheme with a mutation probability equal to 0.3 is used. In the PSO case, we use the global-best particle swarm optimization algorithm with the cognitive parameters c_1 and c_2 set to 0.3 and 0.5, respectively, whilst the inertia parameter is set to 0.7. The swarms consist of 10 particles and we allow 60 iterations. Finally, for DE we allow 25 iterations employing the best1bin evolution strategy, with a population size equal to 6, a differential weight β drawn at each iteration from $U(0.5, 1)$, and a recombination probability equal to 0.7. For a detailed discussion on the optimization algorithm parameters, we refer the reader to [11].

The overall optimization results for the two buildings are illustrated in Table I, presenting the values of the cost function, the convergence epoch (CE), and the convergence time (CT). The CE and CT measured from the start of the optimization process until the epoch and time, respectively, at which the optimal solution occurs, i.e., the iteration and time at which we identify the network deployment with the highest T_{eff} . The optimization results of the ML-based propagation model are benchmarked against those of ray tracing. Additionally, we consider a *Hybrid* optimization approach, conducting ray tracing simulations for the solution space optimal point determined during the optimization process with the ML-based model. This approach lets us corroborate whether the topologies indicated by the data-driven propagation model optimization also yield a nearly-optimal point when signal propagation is approximated via a high-performance propagation solver. Since the proposed framework leverages metaheuristic algo-

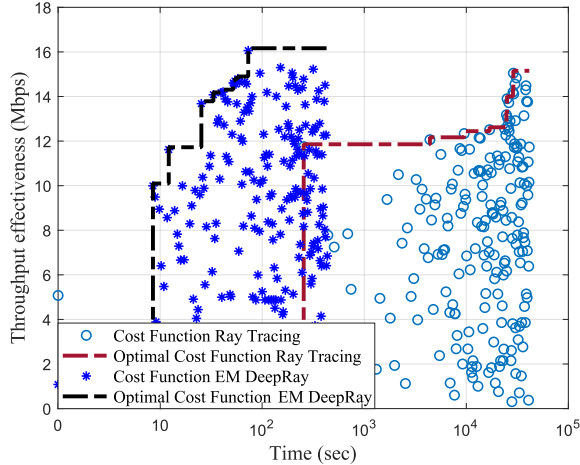
ritms to retrieve a nearly-optimal solution, to quantify the uncertainty entwined with the intrinsic stochasticity of these algorithms, the optimization process with each algorithm is repeated 10 times assuming different seeds, and in Table I we report the statistics of this process. As we will discuss later, such an approach is not feasible when using the ray-tracer due to the unbearable computational cost.

The first conclusion to be drawn from Table I is that for each building all optimization algorithms converge approximately to the same optimal T_{eff} value. Indeed, the mean values attained by each algorithm are very close, while the standard deviation over the 10 optimization runs per algorithm remains small, yielding tight confidence intervals around the indicated nearly-optimal solution. For both buildings, the optimization based on ray tracing returns a topology with the maximum number of 5 fBSs, whilst EM DeepRay yields equivalently good topologies that in some cases reduce that number to 4 fBSs, i.e., do not need to leverage all allowed fBSs to ensure optimal wireless performance. Unsurprisingly, multiple optimization runs show that there can be multiple network deployments that attain nearly-optimal throughput effectiveness. The potential of the data-driven model to conduct on-the-fly optimization of network performance enables the identification of several of these deployments, which translates into higher flexibility for MNOs to select the deployment that, e.g., implies the lowest realization cost in terms of cabling. Overall, DE attains the highest sheer performance, achieving the best trade-off between T_{eff} and CT.

Another key takeout from Table I is the equivalency between the network topologies indicated by the optimization with the data-driven propagation model and these determined via the ray-tracer-based optimization. That is clearly highlighted when comparing the cost function values attained via the ray tracing optimization and these of the *Hybrid* optimization scheme (i.e., when T_{eff} is computed via ray tracing but at the optimal points indicated via the data-driven propagation model optimization process). Indeed, the *Hybrid* optimization yields similar or even higher throughput effectiveness compared to the ray-tracer-based optimization for all the buildings and optimization algorithms. That is to say, *that if one conducts the optimization using a ray-tracer or the data-driven propagation model they should expect to find a topology that has approximately the same performance and is equally good*. Indicatively,



(a) RSS for the indicated nearly-optimal deployment.



(b) Throughput effectiveness.

Fig. 4: Example of identified nearly-optimal network deployment for B1, and evolution of T_{eff} during the optimization, indicating the cost function value achieved at each iteration, as well as its best value up until that iteration, while using ray tracing and EM DeepRay to emulate signal propagation

the RSS heatmap for the nearly-optimal network deployment for B1 at the end of the optimization procedure with BO is depicted in Fig. 4(a), which also shows the building blueprint. We observe that the optimization places the fBSs in diverse areas of the building, to maximize the coverage and capacity of the indoor network.

The distinct advantage of conducting the optimization process with the data-driven propagation model is the sharp decrease in the time required to identify the nearly-optimal network topology. A ray tracing simulation for a single fBS would take a few minutes. On the other hand, via the data-driven model, one can compute T_{eff} in less than a second. When plugged into the UDN planning framework, this yields a dramatic drop in the optimization time, as the optimization with a data-driven model requires a few minutes, whereas with a ray-tracer it takes hours. We note that this is the reason why we only conducted ray tracing optimizations with BO, since it is not population-based and it requires the lowest number of epochs to converge; otherwise, the optimization process could require close to a week. Moreover, the results indicate that a ray-tracing-based optimization for larger scenarios consid-

ering a higher number of fBSs is prohibitive, as the time will grow exponentially. Finally, it is worth noting that the result obtained with the ray-tracer does not carry standard deviation information, as executing multiple optimization runs is too expensive; this means that the flexibility granted by a data-driven approach (with multiple proposed deployments of similar quality to choose among, as discussed earlier) is lost with a solution relying on ray tracing.

To better illustrate the computational advantage of our proposed data-driven approach, we present in Fig. 4(b) the evolution of the cost function overtime on a \log -scale. The values of T_{eff} associated with each iteration of the optimization process with EM DeepRay and the ray-tracer are depicted as asterisks and circles, respectively. Over time, i.e., moving along the x-axis, the optimizers identify increasingly better deployments with higher T_{eff} , indicated by the dashed black and dark red lines representing the optimal value of T_{eff} attained with time for the data-driven model and ray tracing, respectively. Evidently, our proposed strategy finds a UDN deployment in a time that is more than two orders of magnitude lower, i.e., around 3 minutes versus more than 6 hours for the benchmark.

IV. DISCUSSION AND FUTURE CHALLENGES

The drastic decrease in the optimization time renders frameworks leveraging data-driven propagation models a key enabler for the timely installation of indoor cellular and industrial private networks, which are expected to be extensively deployed in 5G/B5G systems. Future research can explore multi-objective optimization problems where a set of conflicting target criteria ought to be simultaneously met, e.g., having KPIs related to latency or meeting energy and budget constraints associated with the fBS transmit power and their number. Moreover, the performance of advanced optimization algorithms, such as the simultaneous perturbation stochastic approximation or zeroth-order optimization, is worth considering instead of meta-heuristics. In addition to the initial deployment stage, such frameworks can be employed to redesign indoor wireless networks, which are much more agile and likely to be subject to changes compared to legacy outdoor networks.

Future research on data-driven propagation models can further empower their role in network planning and optimization. Specifically, most of the existing generalizable data-driven propagation models are trained to predict physical quantities related to the signal level. On the contrary, data-driven models aiming at inferring other wireless channel characteristics, e.g., the angle of arrival/departure or the delay spread, currently exhibit limited scalability. However, these characteristics are important for many applications of wireless communication systems, such as dynamic beamforming and beam steering, that can enable real-time network optimization to accommodate user requirements. Likewise, embedding functionalities related to controllable secondary radiation sources, e.g., re-configurable intelligent surfaces, will make data-driven models stand out from the time-consuming legacy solvers, and assist in real-time network optimization.

Finally, generative AI, probing the distribution of real propagation data, can be employed to transfuse real-world accuracy to data-driven solvers developed with synthetic data and eventually make them more realistic than legacy models. That can further compensate for limited knowledge in the representation of the propagation environment or inaccuracies related to the selected material coefficients, which are both necessary for ray-tracing and any data-driven model. Hence, one can narrow the gap between radio propagation solvers and real-world propagation conditions.

V. CONCLUSION

The use of AI yields unprecedented potential to revolutionize the foundations of radio wave propagation inference, and consequently establish a new generation of generalizable ML-based propagation models. In particular, these models can be leveraged to enable rigorous and fully automated AI-based network planning, through promptly generated digital twins, and according to a KPI of interest. To this end, we coupled a pre-trained data-driven indoor propagation model with AI optimizers to proactively assess multiple network topologies, and eventually identify the one that maximizes the IWNP. Our analysis highlights how the once cumbersome procedure of finding the optimal network topology now takes only a few minutes. More importantly, it paves the way for the shift from conventional and obsolete network planning techniques, that rely on human experience or simplistic mathematical models, to a fully automated AI-powered network design.

ACKNOWLEDGMENT

This work was supported by European Commission through the Horizon 2020 Framework Programme, H2020-MSCA-ITN-2019, Grant No. 860239, BANYAN. The work of Stefanos Bakirtzis is supported by the Onassis Foundation and the Foundation for Education and European Culture.

REFERENCES

- [1] X. Ge et al., "5G ultra-dense cellular networks," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 72–79, Feb. 2016.
- [2] "Cisco vision: 5G-thriving indoors," Cisco, San Jose, CA, USA, White Paper, 2019. [Online]. Available: <https://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/ultra-services-platform/5g-ran-indoor.pdf>.
- [3] P. Almasan et al., "Network digital twin: Context, enabling technologies, and opportunities," *IEEE Com. Mag.*, vol. 60, no. 11, pp. 22–27, 2022.
- [4] S. F. Yunas et al., "Spectral and energy efficiency of ultra-dense networks under different deployment strategies," *IEEE Com. Mag.*, vol. 53, no. 1, pp. 90–100, Jan. 2015.
- [5] S. F. Yunas et al., "Deployment strategies and performance analysis of macrocell and femtocell networks in suburban environment with modern buildings," in *Proc. IEEE 39th Conf. Local Comput. Netw.*, Sep. 2014, pp. 643–651.
- [6] A. W. Reza et al., "A novel integrated mathematical approach of ray-tracing and genetic algorithm for optimizing indoor wireless coverage," *Prog. Electromagn. Res.*, vol. 110, pp. 147–162, 2010.
- [7] T. K. Sarkar et al., "A survey of various propagation models for mobile communication," *IEEE Antennas Propag. Magazine*, vol. 45, no. 3, pp. 51–82, Jun. 2003.

- [8] Y.-C. Wang et al., "Small-cell planning in LTE HetNet to improve energy efficiency," *Int. J. Commun.*, vol. 31, no. 5, e3492, 2018.
- [9] M. J. Nassiri et al., "A novel approach on femtocell placement in the commercial buildings using genetic algorithm," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 9, e4285, 2021.
- [10] A. Oustry et al., "Optimal deployment of indoor wireless local area networks," *Networks*, vol. 81, no. 1, pp. 23–50, 2023.
- [11] A. P. Engelbrecht, *Computational intelligence: an introduction*. Hoboken, NJ, USA: John Wiley & Sons, 2007.
- [12] A. Seretis et al., "An overview of machine learning techniques for radiowave propagation modeling," *IEEE Trans. Antennas and Propag.*, vol. 70, no. 6, pp. 3970–3985, Jun. 2022.
- [13] S. Bakirtzis et al., "EM DeepRay: An expedient, generalizable, and realistic data-driven indoor propagation model," *IEEE Trans. Antennas and Propag.*, vol. 70, no. 6, pp. 4140–4154, Jun. 2022.
- [14] S. Bakirtzis et al., "Stochastic evaluation of indoor wireless network performance with data-driven propagation models," in *Proc. IEEE Global Commun. Conf.*, Dec. 2022, pp. 3587–3592.