CITYDEPLOY-BENCH: BENCHMARKING PHYSICS-GROUNDED SPATIAL SET PLANNING FOR MULTI-TRANSMITTER NETWORK DEPLOYMENT

Anonymous authors
Paper under double-blind review

ABSTRACT

ABSTRACT

Automating city-scale wireless deployment remains challenging under complex urban propagation and network-wide interference. We introduce **CityDeploy-Bench**, a benchmark that reframes multi-transmitter deployment as *physics-grounded spatial set planning* under a unified ray-tracing verifier. The benchmark separates utility representation from planning dynamics, enabling controlled comparison between direct scalar rewards, relational models, and higher-order interaction structures across diverse planners. Our experiments reveal a clear transition in planning behavior as physical coupling grows. Deployment quality becomes increasingly dependent on whether the learned utility captures collective transmitter interactions, whereas stronger search alone cannot compensate for missing relational structure. This establishes multi-transmitter deployment as a coordination problem over physically interacting sets rather than a collection of independent spatial decisions. We release CityDeploy-Data and the benchmark framework as a reproducible testbed for research linking decision learning with physically grounded wireless network design.

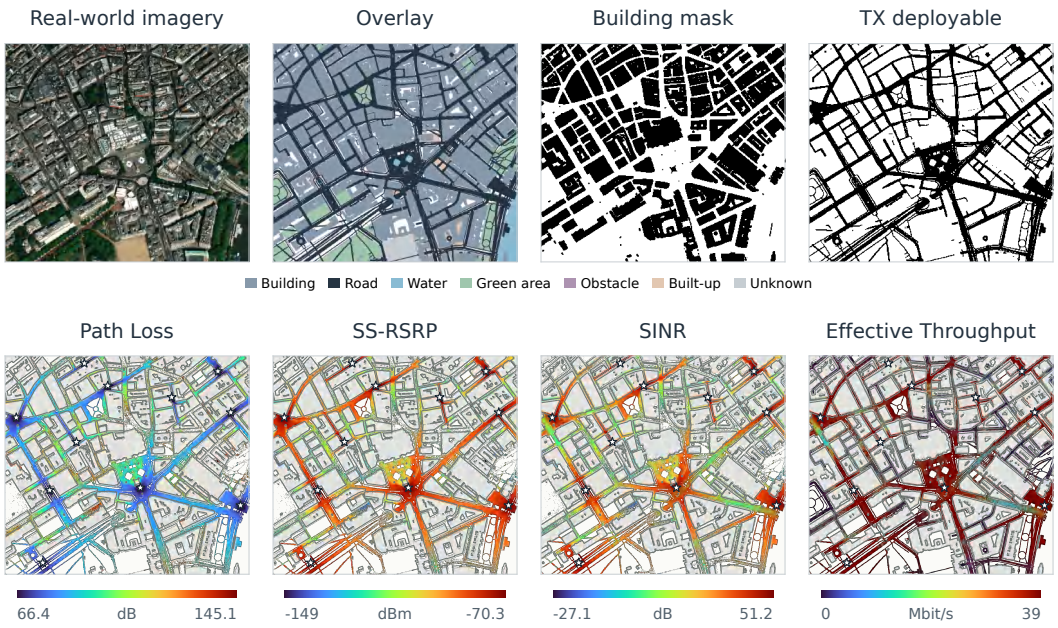


Figure 1: Trafalgar Square illustrates the central challenge of CityDeploy-Bench: effective multi-transmitter deployment requires coordinated decisions across the complete transmitter set under geometry-induced propagation and interference coupling.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

1 INTRODUCTION

Wireless service in cities is governed by a strongly structured physical environment. Buildings, roads, materials, visibility, and multipath propagation jointly determine where transmitters (TXs) can provide useful coverage and how simultaneously active TXs interfere with one another (Levie et al., 2021; Tian et al., 2021). Designing a multi-transmitter (multi-TX) deployment therefore requires more than placing individually strong sites: the complete TX set must be coordinated with the surrounding urban geometry and with the service requirements imposed on the network. In practice, such deployments are commonly designed through expert-driven iteration and repeatedly validated with ray-tracing tools, resulting in substantial engineering and computational cost (Li et al., 2022). Recent advances in machine learning and open-source differentiable or programmable ray tracing make automatic wireless planning increasingly feasible (Hoydis et al., 2023; Khan et al., 2022).

We view this problem through the lens of decision learning and formulate city-scale multi-TX deployment as *physics-grounded spatial set planning*. Given an urban scene, a planner must generate an unordered set of TX locations whose joint operation satisfies prescribed service requirements after physical verification. This perspective differs from independent site selection. The utility of one TX depends on the remaining deployment through serving-cell reassignment, coverage redundancy, and co-channel interference. As deployment density grows, planning therefore changes from locating individually useful sites into coordinating a physically coupled set. This set-level structure makes urban deployment a natural testbed for studying how learned utility representations and search algorithms interact under a high-fidelity physical evaluator.

A growing body of work has explored automatic wireless planning using optimization, learning, and generative reasoning. Black-box and population-based optimizers search deployment configurations directly (Eriksson et al., 2019; Bonyadi & Michalewicz, 2017; Deb & Jain, 2014; Qiu et al., 2024); Monte Carlo methods construct reward-induced target distributions and explore them through importance weighting, sequential sampling, or Hamiltonian dynamics (Piché et al., 2019; Naesseth et al., 2014; Lioutas et al., 2023; Hoffman & Gelman, 2014; Cardoso et al., 2022); and diffusion- or score-based planners progressively refine noisy candidate configurations under learned rewards or constraints (Ho et al., 2020; Janner et al., 2022; Dong et al., 2024; Pan et al., 2024; Akhound-Sadegh et al., 2024; **?**). Large language models have also recently been applied to wireless planning and engineering workflows under specific task formulations (Quan et al., 2025; Li et al., 2026; Modesto et al., 2026; Deng et al., 2025; Yuan et al., 2026). However, these studies differ substantially in environment construction, infrastructure assumptions, physical verification, objective definitions, and evaluation metrics. Consequently, it remains difficult to determine whether performance gains arise from the planner itself, the learned reward, the physical setting, or the evaluation protocol.

Two challenges are particularly important for a standardized benchmark. First, *spatial search is non-smooth and scene dependent*. Urban geometry creates abrupt changes in visibility and propagation, so nearby TX coordinates can induce markedly different service responses (Yuan & Cheng, 2026). Multi-TX planning must therefore search a high-dimensional configuration space containing fragmented feasible regions and multiple locally attractive solutions. Second, *deployment utility is relational*. A TX may improve path-loss or received-power coverage while simultaneously increasing interference or duplicating service already provided by other TXs. The value of a deployment is thus a property of the complete set rather than the sum of independent site scores. This raises a second question beyond search: how should the utility of an interacting TX set be represented? Pairwise graph models provide one form of relational inductive bias (Battaglia et al., 2016; Kipf et al., 2018), while higher-order hypergraph formulations offer a natural extension when useful or harmful interactions involve multiple TXs jointly (Feng et al., 2019; Battiston et al., 2021; Wang et al., 2023).

In this work, we introduce **CityDeploy-Bench**, a benchmark for physics-grounded spatial set planning in city-scale wireless networks. CityDeploy-Bench couples physically reconstructed urban scenes, ray-tracing-derived service responses, learned set-level utility models, and diverse planning algorithms under a common executable deployment interface. Rather than evaluating search methods in isolation, the benchmark separates *utility representation* from *planning dynamics*, enabling controlled study of how relational structure and inference interact. Our main contributions are:

1 **A unified formulation of city-scale multi-TX deployment.** We formulate wireless deployment as physics-grounded spatial set planning over unordered TX configurations, with a common ray-tracing-based service utility and executable road-constrained action space.

This establishes a standardized bridge between scenario-specific wireless engineering and learning-based decision making.

2 **CityDeploy-Data and structured utility representations.** We construct a geographically partitioned dataset of urban scenes and multi-TX deployments with dense physical responses and deployment-level service labels. On top of this interface, we benchmark pooled, pairwise, latent-relational, and higher-order hypergraph utility models, allowing the effect of reward structure to be separated from the effect of the planner.

3 **A broad benchmark of search and sampling dynamics.** We evaluate nine planners spanning annealed score-based search, Monte Carlo inference, and derivative-free optimization under the same learned utility interfaces and physical verification protocol. This design supports controlled comparison of representation–inference compatibility rather than treating each reward–planner combination as an unrelated method.

4 **Mechanistic findings on multi-TX planning.** Our experiments reveal that deployment difficulty is governed by physical coupling rather than TX count alone: densification can move a scene from a coverage-limited to an interference-limited regime; the useful relational order grows with coupling but is not universally maximal; reward representation and search dynamics exhibit strong compatibility effects; and high peak coverage does not necessarily imply reliable target attainment or infrastructure efficiency.

5 **Multi-faceted evaluation with physical verification.** We evaluate simulator-verified joint and marginal service quality, planning efficiency, infrastructure efficiency, seed stability, and geographic transfer/adaptation under consistent deployment constraints and radio assumptions. This provides a reproducible basis for studying learned planning beyond a single scene, objective, or optimization paradigm.

2 Problem Formulation

2.1 Multi-TX Deployment and Service Utility

Deployment. Given an urban scene S, we seek an unordered set of N transmitter (TX) locations $P = \{p_i\}_{i=1}^{N}$ on the deployable road domain $\mathcal{D}_S$. Feasible deployments satisfy

$$\mathcal{F}_N(S) = \{P \subseteq \mathcal{D}_S : |P| = N,\ \|p_i - p_j\|_2 \geq d_{\min}\}\,, \tag{1}$$

with $d_{\min} = 20\,\mathrm{m}$. Scene construction, coordinate conventions, and continuous-to-road mapping are detailed in Appendices A and E.

Physical response and service utility. For a feasible deployment, the ray-tracing operator

$$Y = F_{\mathrm{RT}}(S, P), \qquad Y(x) = \big(\mathrm{PL}, \mathrm{SS\text{-}RSRP}, \mathrm{SINR}, T_{\mathrm{eff}}\big)(x) \tag{2}$$

combines strongest-power serving association with full-frequency-reuse interference. Over the road evaluation cells Ω_S, we define joint-metric (JM) coverage as

$$U(S, P) = \frac{1}{|\Omega_S|} \sum_{x \in \Omega_S} \mathbf{1}[Y(x) \text{ satisfies } \boldsymbol{\tau}]\,, \tag{3}$$

where $\boldsymbol{\tau}$ denotes the fixed service thresholds. For a prescribed TX cardinality N, the planning objective is

$$P_N^{\star}(S) \in \arg\max_{P \in \mathcal{F}_N(S)} U(S, P). \tag{4}$$

Because both serving association and interference depend on the complete TX set, U is inherently relational and need not improve monotonically with TX cardinality. The radio profile, service definitions, thresholds, and metric dependencies are given in Appendix B.

2.2 Learned Utility Representations

Direct evaluation of Eq. equation 4 is expensive because each utility query requires ray tracing. We therefore learn a permutation-invariant surrogate

$$\widehat{U}_\theta(X_S, \mathbf{u}) \approx U(S, P), \tag{5}$$

where X_S is the rasterized scene representation and $\mathbf{u} = \{u_i\}_{i=1}^{N}$ denotes normalized TX coordinates. The predictor is invariant to any reordering of the TX set.

We benchmark four utility representations with increasing relational structure. The *Regressor* pools independently encoded TX features without explicit relations; the Interaction Network (*IN*) (Battaglia et al., 2016) introduces pairwise message passing; the *NRI* predictor (Kipf et al., 2018) learns latent relation types; and the Hypergraph Potential Energy Model (*HPEM*) represents higher-order TX interactions through subset potentials.

For HPEM, the predicted utility is parameterized by a learned energy,

$$\widehat{U}_\theta = -E_\theta, \qquad E_\theta = E_{\text{global}} + \sum_{k \in \mathcal{K}_N} w_k \operatorname*{Mean}_{\substack{e \subseteq \mathbf{u} \\ |e|=k}} E_{\theta,k}(X_S, e), \tag{6}$$

where $\mathcal{K}_N = \{k \in \mathcal{K} : k \le N\}$ and $\mathcal{K} \subseteq \{1, 2, 3, 4\}$. The learned potentials provide a structured predictive decomposition rather than an assumed one-to-one correspondence with physical interaction energies. Architectures, symmetry constraints, and training objectives are detailed in Appendix D.

2.3 Reward-Guided Planning

All planners consume the same learned utility through the penalized score

$$r_S(\mathbf{u}) = \widehat{U}_\theta(X_S, \mathbf{u}) - \lambda_{\text{road}}\phi_{\text{road}}(\mathbf{u}) - \lambda_{\text{sep}}\phi_{\text{sep}}(\mathbf{u}), \tag{7}$$

where the penalties encourage road feasibility and the minimum TX separation. A planner $\mathcal{A}$ produces continuous candidate deployments, which are converted to an executable TX set by a common selection-and-repair operator,

$$\mathcal{C} = \mathcal{A}[r_S], \qquad \widehat{P} = \mathcal{R}_S(\mathcal{C}) \in \mathcal{F}_N(S). \tag{8}$$

The selected deployment is then evaluated by F_{RT}; simulator outputs are used only for final physical verification and never to rerank candidates during search.

We consider annealed score-based, Monte Carlo, and derivative-free planners, all operating through this common scalar interface. Their objectives, constraints, and implementations are given in Appendix E.

3 Benchmarks

CityDeploy-Bench couples scene-conditioned deployment data with three planning regimes under the common physical model defined in Section 2.1. All reported deployments satisfy the same road constraints and are evaluated by the same ray-tracing verifier.

3.1 CityDeploy-Data

Urban scenes and deployments. CityDeploy-Data is constructed from geographic urban regions spanning multiple cities and spatial scales. Each region is converted into an aligned simulation scene containing urban geometry, semantic structure, radio-material assignments, deployable road locations, and receiver evaluation grids. For each scene, we sample feasible unordered multi-TX configurations across multiple cardinalities and evaluate them using the same ray-tracing pipeline used throughout the benchmark. Scene provenance, reconstruction rules, deployment generation, and radio-material assignments are detailed in Appendix A.

Scene-conditioned supervision. Each deployment is paired with dense path-loss, SS-RSRP, SINR, and effective-throughput maps, together with their marginal and joint service labels. Serving association, interference, and service utility are recomputed for the complete active TX set, so the labels preserve the physical interactions specific to each deployment. This supervision directly matches the learned utility interface $\widehat{U}_\theta(X_S, \mathbf{u})$ introduced in Section 2.2. Dataset statistics, deployment-family construction, and storage conventions are reported in Appendix A.

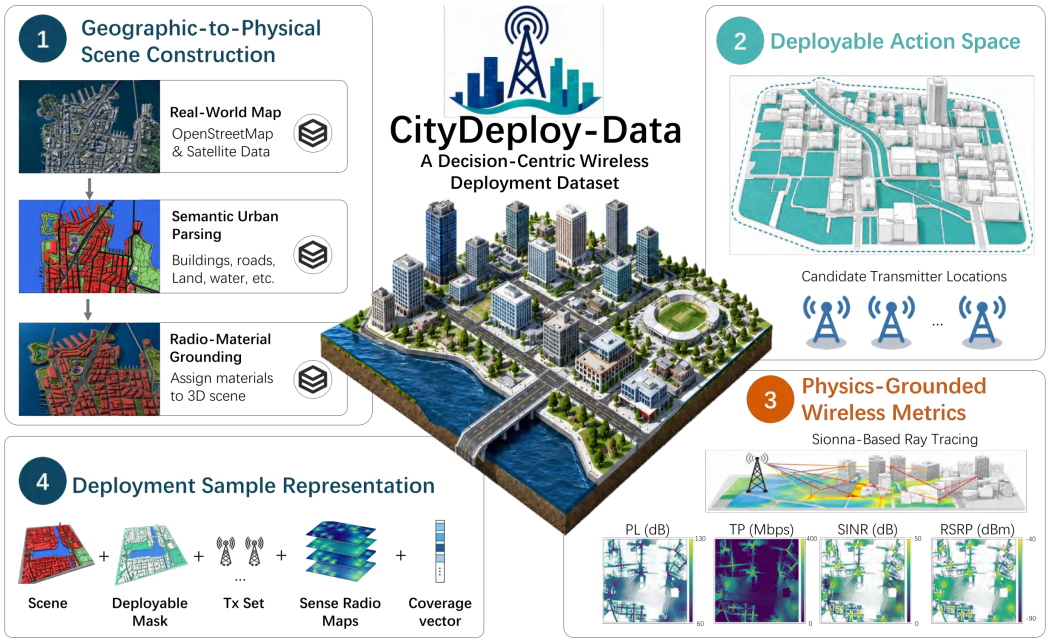


Figure 2: **Overview of CityDeploy-Data.** (1) Geographic urban regions are converted into simulation-ready physical scenes through semantic parsing, three-dimensional reconstruction, and radio-material grounding. (2) Urban semantics determine the scene-specific deployable action space and the feasible transmitter locations. (3) Multi-transmitter configurations are evaluated using Sionna RT to obtain dense path-loss (PL), reference signal received power (RSRP), signal-to-interference-plus-noise ratio (SINR), and throughput (TP) maps. (4) Each benchmark sample associates the physical scene, deployable mask, unordered transmitter set, dense wireless responses, and deployment-level service labels.

Geographic partitioning. To reduce spatial leakage, geographically overlapping scenes are grouped before partitioning, and each connected geographic group is assigned wholly to one split. The benchmark therefore evaluates transfer to held-out geographic regions rather than random held-out deployments from overlapping areas. The base partition is not interpreted as unseen-city generalization unless an entire target city is excluded from training. Complete split composition and integrity checks are provided in Appendix A.4.

3.2 Benchmark Tasks

Fixed-cardinality planning. Given a scene S and prescribed TX cardinality N, planners optimize the fixed-cardinality objective in Eq. equation 4 and return one executable deployment for physical verification. We report simulator-verified JM coverage as the primary planning-quality metric, together with the marginal service coverages and planning cost. This task measures how effectively a planner uses a fixed infrastructure budget.

Target-coverage planning. Given a target JM coverage γ_{tar}, the ideal objective is

$$N^{\star}_{\gamma}(S) = \min\{N : \exists P \in \mathcal{F}_N(S),\ U(S,P) \geq \gamma_{\mathrm{tar}}\}. \tag{9}$$

In practice, each cardinality is searched independently in ascending order, and the procedure stops at the first simulator-verified success. Because each inner search is approximate, the returned cardinality is the first successful tested count rather than a claimed global minimum. We report success rate, first successful TX count, best verified JM coverage, and cumulative planning cost. This task measures reliable service attainment and infrastructure efficiency under finite search.

Transfer and adaptation. We evaluate whether learned utility models remain useful when planning shifts to new geographic scenes. *Zero-shot* evaluation applies the base checkpoint directly, while

few-shot and *full-pool* adaptation update the utility model using increasing amounts of target-scene supervision. Adaptation and evaluation deployments remain disjoint. Planning performance is measured using the same fixed-cardinality or target-coverage protocols above. City-level generalization is claimed only when the complete target city is excluded from base training. Adaptation pools, provenance, stopping rules, and computational accounting are specified in Appendices D.6 and F.

4 EXPERIMENTS

We evaluate CityDeploy-Bench along the three planning regimes defined in Section 3. Rather than treating the utility model and the planner as a single method, our experiments separate two complementary design axes: *how deployment utility is represented* and *how the resulting reward landscape is searched*. This allows us to study not only which combinations achieve high simulator-verified service, but also how physical coupling, relational structure, search dynamics, and geographic shift determine planning behavior. All reported service metrics are computed on the final road-feasible deployment using the physical simulator rather than the learned utility model. Detailed training configurations, planner hyperparameters, additional scenes, target levels, and ablations are provided in the appendices.

4.1 EXPERIMENTAL SETUP

Utility representations. We evaluate four permutation-invariant utility representations introduced in Section 2.2: the pooled *Regressor*, the pairwise Interaction Network (*IN*), the latent-relation *NRI* model, and the higher-order Hypergraph Potential Energy Model (*HPEM*). These models expose the same scalar utility interface but differ in the relational structure available to represent interactions among simultaneously active TXs.

Planning algorithms. We benchmark nine search algorithms spanning three planning families. The annealed score-based planner uses reward gradients under a decreasing noise schedule. Monte Carlo methods include Gaussian and Langevin weighted sampling (*WS-G* and *WS-L*), NUTS (Hoffman & Gelman, 2014), and the BR-SNIS adaptation (Cardoso et al., 2022). Derivative-free methods include Greedy search, particle swarm optimization (*PSO*), a genetic algorithm (*GA*), and TuRBO (Eriksson et al., 2019). Each compatible planner can consume any of the learned utility representations, making utility representation and search strategy separable experimental factors.

Evaluation protocol. For fixed-cardinality planning, we report simulator-verified JM coverage and the associated marginal service coverages. For target-coverage planning, we additionally report success rate, the first successful tested TX count, best verified JM coverage, and cumulative planning cost. Planning time, model-evaluation counts, and random-seed variability are reported separately because heterogeneous planner iterations are not treated as equivalent computational units. Unless explicitly stated otherwise, results aggregate independent random seeds; full per-scene results and implementation settings are given in the appendix.

4.2 FINDINGS

We summarize six findings that characterize how physical coupling, utility representation, and search interact in multi-TX planning.

Finding 1: More coverage does not necessarily mean better service. Multi-TX deployment has a non-monotone marginal structure. To isolate the physical effect of adding one transmitter, we compare nested deployments S and $S \cup \{p\}$ from the same scene and deployment family, while keeping all existing TX locations unchanged.

Figure 3 shows that an added TX can increase path-loss coverage while simultaneously reducing JM coverage. Across 9,320 nested deployment pairs from the held-out scenes, 3,955 (42.44%) exhibit lower joint coverage after the additional TX is activated. This statistic is descriptive of the available nested pairs rather than an estimated probability over arbitrary deployments.

The mechanism follows directly from the shared-spectrum physical model. An added TX can strengthen the serving signal or cover previously weak regions, while also introducing co-channel

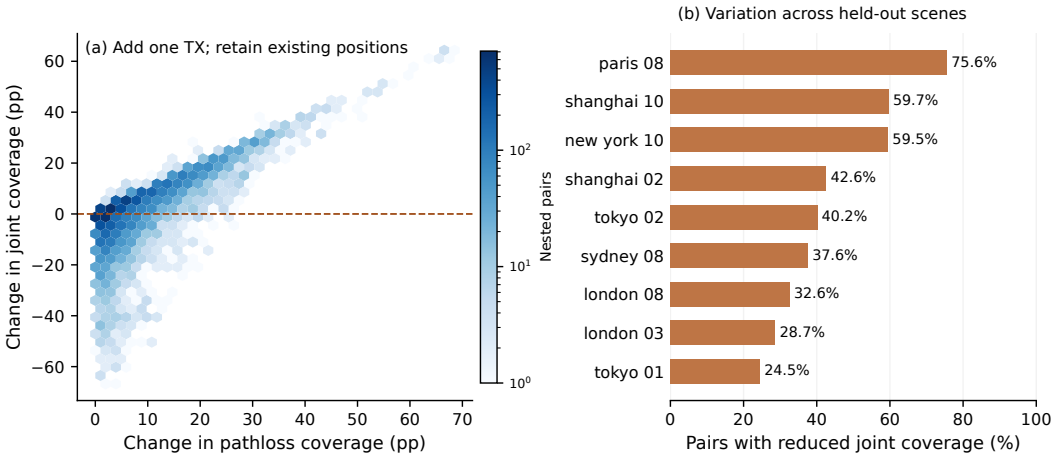


Figure 3: **Marginal effect of adding one transmitter.** We compare nested deployments S and $S \cup \{p\}$ while retaining all existing TX positions. **Left:** change in path-loss coverage versus change in JM coverage across matched deployment pairs; density below the horizontal zero line shows cases where serving-link coverage improves while joint service deteriorates. **Right:** fraction of nested pairs with reduced JM coverage in each held-out scene. Across all 9,320 pairs, 42.44% exhibit a decrease in JM coverage after the additional TX is activated.

interference and changing serving association. Its marginal value therefore depends on the existing deployment:

$$U(S \cup \{p\}) - U(S)$$

can be positive or negative even when serving-link coverage improves. Multi-TX planning must consequently evaluate transmitters as a coordinated set rather than rank locations by individual coverage quality.

Aggregate cardinality-level trends exhibit the same coverage–interference transition across deployment scales (Figure 17 in Appendix M.1).

Finding 2: Relational complexity should match physical coupling. No single utility representation is uniformly preferable across deployment regimes. At low TX cardinality, pooled or pairwise representations can match or exceed higher-order HPEM guidance for some planners; as the active set grows, replacing HPEM produces increasingly systematic degradation (Figure 18 in Appendix M.2). The relevant representation complexity is therefore determined by the interaction structure induced by geometry, cardinality, and interference rather than by cardinality alone.

The controlled ablations sharpen this observation. Increasing generic model capacity does not monotonically reduce utility-prediction error, whereas increasing the represented interaction order improves verified planning quality at matched inference budgets (Appendix L). Higher order is therefore useful when the deployment requires it, rather than as a generic consequence of using a larger model.

Finding 3: Representation and search provide complementary gains. Utility representation and search strength are separable experimental factors, but their effects are not interchangeable. The reward-sensitivity analysis shows that changing the utility model can alter different planners by substantially different amounts (Figure 18). This indicates that search behavior depends on the geometry of the learned reward landscape rather than only on its scalar prediction accuracy.

The inference-budget ablation provides a complementary result. More particles and additional refinement steps consistently improve planning within each tested interaction order, while higher-order representations retain an advantage under matched inference budgets (Appendix L.2). Search effort can therefore extract more value from a given reward landscape, but it does not eliminate the benefit of representing the underlying TX interactions. Scaling inference and improving relational structure address different limitations and are most effective when used together.

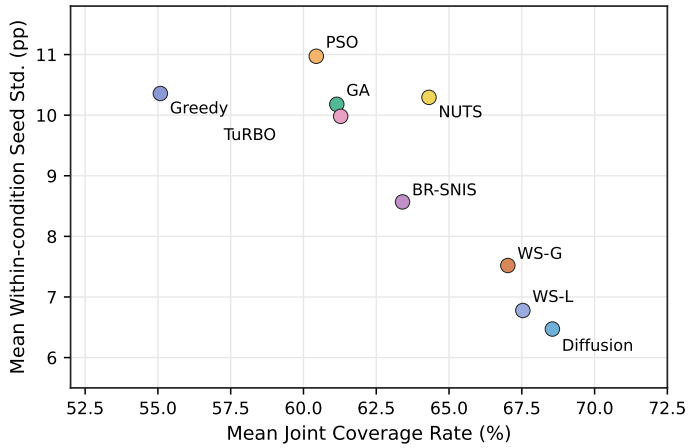


Figure 4: **Planning quality and random-seed stability.** Mean simulator-verified JM coverage is plotted against the mean within-condition standard deviation over five random seeds. Each condition is defined by city, deployment scale, TX cardinality, and utility representation. Several stochastic planners occupy the high-quality, low-variance region, showing that stochastic search does not inherently imply unstable planning.

Finding 4: Planning performance is multi-dimensional. Search methods occupy different operating regions in service quality, computational cost, and seed sensitivity. The quality–efficiency frontier changes with deployment scale (Figure 19 in Appendix M.3), so no single search strategy defines the preferred operating point across all regimes.

Stochasticity also does not imply poor repeatability. As shown in Figure 4, several sampling-based planners combine high simulator-verified JM coverage with relatively low within-condition variation across random seeds. Planning methods should therefore be compared jointly in terms of service quality, cost, and repeatability rather than by a single best-coverage value.

Finding 5: Peak quality does not imply reliable target attainment. Target-coverage planning exposes a distinction that is largely hidden by fixed-cardinality evaluation. A planner must not only discover high-quality deployments, but do so reliably, with few TXs and without excessive cumulative search cost. These criteria need not agree.

At the 70% JM target in the representative Paris scene, Diffusion succeeds in all 15 trials with a first-successful count of 2.20 ± 0.45 TXs. TuRBO reaches a comparable best verified JM coverage ($76.39 \pm 5.92\%$ versus $76.82 \pm 3.42\%$ for Diffusion), but succeeds in only 66.67% of trials and requires 3.00 ± 1.00 TXs among successful searches. WS-G reaches similarly high peak coverage with only 2.40 ± 0.55 TXs on successful trials, yet its success rate is 73.33%. Thus, methods that occasionally find strong deployments can differ substantially in how reliably they satisfy a prescribed service requirement.

The distinction becomes sharper when computational cost is included. WS-L attains a 93.33% success rate with the lowest cumulative planning time among the high-success methods, whereas Greedy is both less reliable and requires substantially more infrastructure on successful trials. Target-coverage evaluation separates four properties that are conflated by a single best-utility number: *attainability*, *infrastructure efficiency*, *peak service quality*, and *search cost*. For deployment planning, these quantities define different operating regimes rather than a single scalar notion of performance. Additional target levels and scenes are reported in Appendix I.

Finding 6: Interaction structure transfers more readily than its calibration. Zero-shot utility models retain useful planning behavior on target scenes, but target-scene adaptation consistently improves simulator-verified deployment quality. Table 2 shows this effect in a challenging 12-TX Paris setting: all three utility representations improve from zero-shot to few-shot and full-pool adaptation.

Table 1: **Target-coverage planning at 70% JM coverage in the representative Paris scene using HPEM guidance.** We report the first successful tested TX count (Req. TXs), success rate (SR), best simulator-verified JM coverage, marginal service coverages, and cumulative planning time. Values are mean±std over 15 independent trials. Required-TX statistics are computed over successful trials; failures remain in the success-rate denominator.

Method	Req. TXs↓	SR (%)↑	Best JM (%)↑	PL (%)↑	SS-RSRP (%)↑	SINR (%)↑	Eff. TP (%)↑	RT (s)↓
Diffusion	2.20±0.45	100.00	76.82±3.42	87.35±5.33	84.80±6.49	79.49±3.15	78.51±2.49	66.09±49.36
BR-SNIS	3.60±1.14	86.67	70.92±3.46	84.74±6.52	79.25±2.95	73.72±3.55	83.88±3.89	31.70±13.98
NUTS	3.40±0.89	86.67	74.96±3.96	81.92±6.39	82.00±6.38	77.87±3.73	81.60±4.03	36.09±14.33
TuRBO	3.00±1.00	66.67	76.39±5.92	83.68±9.22	83.79±9.20	78.98±5.50	82.84±5.69	30.21±14.64
PSO	4.40±2.61	80.00	72.78±3.66	79.19±2.95	87.42±5.29	75.14±4.43	81.69±3.35	41.48±32.41
GA	4.40±2.70	73.33	72.39±3.20	78.81±7.63	78.90±7.60	75.96±1.59	80.01±3.41	42.32±33.47
Greedy	6.20±3.83	40.00	65.91±6.89	75.84±5.06	75.96±5.02	70.14±6.95	76.29±3.68	100.92±186.00
WS-G	2.40±0.55	73.33	75.80±4.25	83.88±3.44	83.94±3.42	77.91±3.98	82.52±2.83	18.72±7.03
WS-L	3.60±3.05	93.33	72.97±1.95	78.94±4.03	79.01±4.05	75.45±1.06	78.47±1.46	18.07±6.93

Table 2: **Target-scene adaptation in a challenging Paris deployment with 12 TXs.** We compare zero-shot transfer, few-shot adaptation, and full-pool adaptation for utility models with increasing relational structure. Results are reported as mean±std. Adaptation consistently improves simulator-verified service, while higher-order relational modeling retains an advantage across all three regimes.

Utility	Setting	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
Regressor	Zero-shot	64.82±1.06	34.76±3.31	43.28±4.48	55.62±4.37	41.54±3.72	47.31±4.16
	Few-shot	66.31±1.42	38.21±2.76	46.78±3.71	59.08±3.56	45.02±3.08	50.82±3.43
	Full-pool	66.76±1.15	39.47±2.28	48.06±3.12	60.31±2.98	46.34±2.58	52.11±2.87
Pairwise	Zero-shot	70.18±1.11	37.64±2.91	46.21±4.02	58.46±3.89	44.37±3.31	50.18±3.72
	Few-shot	71.82±1.53	41.03±2.38	49.66±3.31	61.84±3.17	47.86±2.74	53.64±3.03
	Full-pool	72.31±1.22	42.36±1.94	50.94±2.76	63.08±2.63	49.21±2.28	54.91±2.52
4th-Order	Zero-shot	80.73±1.43	39.42±2.68	48.03±3.74	60.31±3.57	46.16±3.04	52.06±3.43
	Few-shot	82.64±1.81	43.18±2.14	51.84±3.02	64.02±2.87	49.94±2.48	55.82±2.78
	Full-pool	83.17±1.47	44.72±1.73	53.26±2.48	65.34±2.32	51.41±2.03	57.18±2.26

The relational advantage also persists across adaptation regimes. For example, the 4th-order model improves JM coverage from $39.42 \pm 2.68\%$ in the zero-shot setting to $43.18 \pm 2.14\%$ after few-shot adaptation and $44.72 \pm 1.73\%$ with full-pool supervision, while remaining above the pooled and pairwise alternatives at each stage. This suggests that part of the interaction structure transfers across scenes, whereas its quantitative calibration remains sensitive to local geometry and propagation.

Updating the utility model therefore provides a direct mechanism for adapting planning without relearning the search procedure. These experiments measure target-scene or geographic transfer unless the complete target city is excluded from base training; we do not interpret scene-level separation as unseen-city generalization. Additional adaptation results are reported in Appendix D.6.

5 Conclusion

We introduced CityDeploy-Bench, a physics-grounded benchmark for spatial set planning in multi-transmitter wireless deployment. By separating utility representation from search under a common ray-tracing verifier, the benchmark exposes how physical coupling changes both the structure of the planning objective and the requirements placed on learned rewards. Our results show that densification can shift deployment from coverage expansion toward interference-aware coordination, that useful relational complexity grows with physical coupling, and that relational structure and inference budget provide complementary improvements rather than interchangeable ones. Target-coverage and transfer experiments further show that peak deployment quality, reliable service attainment, and cross-scene calibration are distinct dimensions of planning performance. CityDeploy-Bench provides a reproducible interface for studying these questions at the intersection of decision learning and physically grounded wireless network design.

AI USE STATEMENT

In this work, we used ChatGPT to improve the readability and clarity of the manuscript and to assist with organizing the structure of the paper. We did not use generative AI tools for any tasks requiring mandatory disclosure under the ICLR 2027 AI Policy for Authors. All AI-assisted edits were reviewed and, where necessary, revised by the authors. We take responsibility for the final content of this work, including text produced or modified with the aid of generative AI.

REFERENCES

Tara Akhound-Sadegh, Jarrid Rector-Brooks, Joey Bose, Sarthak Mittal, Pablo Lemos, Cheng-Hao Liu, Marcin Sendera, Siamak Ravanbakhsh, Gauthier Gidel, Yoshua Bengio, Nikolay Malkin, and Alexander Tong. Iterated denoising energy matching for sampling from boltzmann densities. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 760–786, 2024.

Peter W. Battaglia, Razvan Pascanu, Matthew Lai, Danilo Jimenez Rezende, and Koray Kavukcuoglu. Interaction networks for learning about objects, relations and physics. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Federico Battiston, Enrico Amico, Alain Barrat, Ginestra Bianconi, Guilherme Ferraz de Arruda, Benedetta Franceschiello, Iacopo Iacopini, Sonia Kéfi, Vito Latora, Yamir Moreno, Micah M. Murray, Tiago P. Peixoto, Francesco Vaccarino, and Giovanni Petri. The physics of higher-order interactions in complex systems. *Nature Physics*, 17(10):1093–1098, October 2021. ISSN 1745-2481. doi: 10.1038/s41567-021-01371-4. URL http://dx.doi.org/10.1038/s41567-021-01371-4.

Mohammad Reza Bonyadi and Zbigniew Michalewicz. Particle swarm optimization for single objective continuous space problems: A review. *Evolutionary Computation*, 25(1):1–54, 2017. doi: 10.1162/EVCO_r_00180.

Olivier Cappé, Simon J Godsill, and Eric Moulines. An overview of existing methods and recent advances in sequential monte carlo. *Proceedings of the IEEE*, 95(5):899–924, 2007.

Gabriel Cardoso, Sergey Samsonov, Achille Thin, Eric Moulines, and Jimmy Olsson. Br-snis: bias reduced self-normalized importance sampling. *Advances in Neural Information Processing Systems*, 35:716–729, 2022.

Kalyanmoy Deb and Himanshu Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, 18(4):577–601, 2014. doi: 10.1109/TEVC.2013.2281535.

Zongyuan Deng, Yujie Cai, Qing Liu, Shiyao Mu, Bin Lyu, and Zhen Yang. Teleplannet: An ai-driven framework for efficient telecom network planning. In *2025 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–6, 2025. doi: 10.1109/ICCC65529.2025.11148630.

Zibin Dong, Jianye Hao, Yifu Yuan, Fei Ni, Yitian Wang, Pengyi Li, and Yan Zheng. Diffuserlite: Towards real-time diffusion planning. *Advances in Neural Information Processing Systems*, 37: 122556–122583, 2024.

David Eriksson, Michael Pearce, Jacob R. Gardner, Ryan Turner, and Matthias Poloczek. Scalable global optimization via local bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Yifan Feng, Haoxuan You, Zizhao Zhang, Rongrong Ji, and Yue Gao. Hypergraph neural networks, 2019. URL https://arxiv.org/abs/1809.09401.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Matthew D. Hoffman and Andrew Gelman. The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(47):1593–1623, 2014.

Jakob Hoydis, Fayçal Aït Aoudia, Sebastian Cammerer, Merlin Nimier-David, Nikolaus Binder, Guillermo Marcus, and Alexander Keller. Sionna rt: Differentiable ray tracing for radio propagation modeling. In *2023 IEEE Globecom Workshops (GC Wkshps)*, pp. 317–321, 2023. doi: 10.1109/GCWkshps58843.2023.10465179.

Michael Janner, Yilun Du, Joshua B. Tenenbaum, and Sergey Levine. Planning with diffusion for flexible behavior synthesis. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 9902–9915, 2022.

Latif U. Khan, Zhu Han, Walid Saad, Ekram Hossain, Mohsen Guizani, and Choong Seon Hong. Digital twin of wireless systems: Overview, taxonomy, challenges, and opportunities. *IEEE Communications Surveys & Tutorials*, 24(4):2230–2254, 2022. doi: 10.1109/COMST.2022.3198273.

Thomas Kipf, Ethan Fetaya, Kuan-Chieh Wang, Max Welling, and Richard Zemel. Neural relational inference for interacting systems. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2688–2697, 2018.

Ron Levie, Çağkan Yapar, Gitta Kutyniok, and Giuseppe Caire. Radiounet: Fast radio map estimation with convolutional neural networks. *IEEE Transactions on Wireless Communications*, 20(6):4001–4015, 2021. doi: 10.1109/TWC.2021.3054977.

Xiaomeng Li, Yuru Zhang, Qiang Liu, Mehmet Can Vuran, Nathan Huynh, Li Zhao, Mizan Rahman, and Eren Erman Ozguven. Automatic network planning with digital radio twin. In *IEEE International Conference on Communications (ICC)*, pp. 1–6, 2026. doi: 10.1109/ICC59461.2026.11587682.

Ying Li, Djordje Tujkovic, and Po-Han Huang. Analytics and machine learning powered wireless network optimization and planning, 2022. URL https://arxiv.org/abs/2209.06352.

Vasileios Lioutas, Jonathan Wilder Lavington, Justice Sefas, Matthew Niedoba, Yunpeng Liu, Berend Zwartsenberg, Setareh Dabiri, Frank Wood, and Adam Scibior. Critic sequential monte carlo. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=ObtGcyKmwna.

Cláudio Modesto, Lucas Mozart, Glauco Gonçalves, Cleverson Nahum, Bruno Castro, and Aldebaro Klautau. Lorawan gateway placement for network planning using ray tracing-based channel models. *IEEE Open Journal of the Communications Society*, 7:6612–6624, 2026. doi: 10.1109/OJCOMS.2026.3704345.

Christian A Naesseth, Fredrik Lindsten, and Thomas B Schön. Sequential monte carlo for graphical models. *Advances in neural information processing systems*, 27, 2014.

Chaoyi Pan, Zeji Yi, Guanya Shi, and Guannan Qu. Model-based diffusion for trajectory optimization. *Advances in Neural Information Processing Systems*, 37:57914–57943, 2024.

Alexandre Piché, Valentin Thomas, Cyril Ibrahim, Yoshua Bengio, and Chris Pal. Probabilistic planning with sequential monte carlo methods. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=ByetGn0cYX.

Kehai Qiu, Stefanos Bakirtzis, Ian Wassell, Hui Song, Jie Zhang, and Kezhi Wang. Large language model-based wireless network design. *IEEE Wireless Communications Letters*, 13(12):3340–3344, 2024. doi: 10.1109/LWC.2024.3462556.

Hongye Quan, Wanli Ni, Tong Zhang, Xiangyu Ye, Ziyi Xie, Shuai Wang, Yuanwei Liu, and Hui Song. Large language model agents for radio map generation and wireless network planning. *IEEE Networking Letters*, 7(3):166–170, 2025. doi: 10.1109/LNET.2025.3539829.

Yu Tian, Shuai Yuan, Weisheng Chen, and Naijin Liu. Radionet: Transformer based radio map prediction model for dense urban environments, 2021. URL https://arxiv.org/abs/2105.07158.

Toshihide Ubukata, Jialong Li, and Kenji Tei. Diffusion model for planning: A systematic literature review. *arXiv preprint arXiv:2408.10266*, 2024.

Yuxin Wang, Quan Gan, Xipeng Qiu, Xuanjing Huang, and David Wipf. From hypergraph energy functions to hypergraph neural networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 35605–35623, 2023.

Chenyang Yuan and Xiaoyuan Cheng. Planning with language and generative models: Toward general reward-guided wireless network design, 2026. URL https://arxiv.org/abs/2602.00357.

Chenyang Yuan, Jinbo Hou, Gang Yu, Kehai Qiu, Kezhi Wang, Haonan Hu, and Jie Zhang. Llm-assisted optimisation of multi-ris placement and beamforming in smart warehouses. *IEEE Transactions on Mobile Computing*, 25(8):12722–12734, 2026. doi: 10.1109/TMC.2026.3674609.

A CITYDEPLOY-DATA CONSTRUCTION

CityDeploy-Data pairs urban scenes and feasible TX sets with spatial radio maps and service labels. This appendix describes scene construction, deployment sampling, and data organization; Appendix B specifies the radio calculations.

A.1 GEOGRAPHIC SOURCES AND SCENE CONSTRUCTION

Geographic inputs. The dataset contains ten scenes from each of London, New York, Paris, Shanghai, Sydney, and Tokyo. OpenStreetMap provides building footprints, roads, and land-use features within each scene's geographic bounding box. We project these features to a local metric frame to construct three-dimensional geometry and aligned semantic rasters, retaining the source snapshot, query record, geographic extent, and coordinate reference system.

Geometry and material priors. Building footprints are extruded using `building:height` or `height` attributes, followed by level counts at $3.2\,\mathrm{m}$ per level and a deterministic building-type prior when heights are missing. The benchmark construction uses a flat ground surface. Semantic classes are mapped to Sionna radio-material priors: concrete for buildings and roads, medium-dry ground for ground and green areas, and wet ground for water. These assignments specify the simulation model rather than measured material composition. Scene manifests record the height policy, material profile, software versions, and hashes of the source and generated assets.

A.2 SPATIAL COORDINATES AND DEPLOYABLE REGIONS

One coordinate convention. A scene of width L_x and height L_y uses a centered horizontal frame in meters. The normalized coordinates supplied to the learning and planning modules are

$$u_i = \left(\frac{p_{i,x}}{L_x} + \frac{1}{2}, \frac{p_{i,y}}{L_y} + \frac{1}{2} \right), \qquad p_i = \left((u_{i,x} - \tfrac{1}{2})L_x, (u_{i,y} - \tfrac{1}{2})L_y \right). \tag{10}$$

Thus $u_i \in [0,1]^2$, while separation constraints are always evaluated in meters. The three scene sizes are 128×128, 256×256, and $512 \times 512\,\mathrm{m}^2$.

Deployment and evaluation domains. The road raster is thresholded at 0.5 on a $1\,\mathrm{m}$ grid. The centers of its active cells form $\mathcal{D}_S$; the corresponding receiver cells form Ω_S. For an $H \times W$ raster, column c and row r have coordinates

$$p_{x,c} = -\frac{L_x}{2} + (c + \tfrac{1}{2})\frac{L_x}{W}, \qquad p_{y,r} = -\frac{L_y}{2} + (r + \tfrac{1}{2})\frac{L_y}{H}. \tag{11}$$

The dataset explicitly uses the road mask for service evaluation. Other scene masks, including a general non-water evaluation mask, do not define the denominator in Eq. equation 3.

A.3 DEPLOYMENT SAMPLING AND SUBSET FAMILIES

Master deployments. We first sample a feasible master set from $\mathcal{D}_S$. A fixed schedule allocates 50% of masters to uniform sampling, 30% to maximin sampling, and 20% to clustered sampling. Uniform sampling traverses a random candidate order; maximin sampling favors candidates farthest from the selected set; clustered sampling traverses candidates by distance to a randomly chosen center. Each strategy enforces the $20\,\mathrm{m}$ separation constraint during selection. The maximin search considers at most 4,096 candidates at each addition. An incomplete maximin selection falls back to uniform sampling; failure to complete a feasible set raises a generation error.

Related subsets. Each master defines a *family* of deployments at prescribed cardinalities (Table 3). When all combinations are requested, they are enumerated. Otherwise, subsets are sampled as prefixes of random permutations of the master, retaining unique subsets and cross-cardinality containment relations. All subsets inherit spatial feasibility from the master. Family links expose controlled changes in TX membership without assuming additive or monotone utility.

Table 3: Deployment family construction. Counts describe the completed dataset. Memberships count appearances in families before deduplication. The final generation uses three blocks of the base master schedule.

Scene side	Scenes	Master size	Masters/scene	Subsets per master
128 m	12	2	180	$N = 1 : 2;\ N = 2 : 1$
256 m	18	4	90	$N = 2 : 6;\ N = 3 : 4;\ N = 4 : 1$
512 m	30	9	60	$N = 5, 6, 7, 8 : 4$ each; $N = 9 : 1$
Total	60		5,580 families	54,900 memberships

Propagation reuse and label recomputation. The simulator computes a separate propagation-gain map for every TX in a master. For a subset, we select the corresponding gain maps and recompute serving association, interference, service maps, and coverage labels using only the active TXs. Channel reuse preserves per-TX propagation; subset labels still require a joint service calculation.

A.4 Dataset Composition and Geographic Splits

Unique deployments. The completed dataset contains 54,193 unique scene–deployment records and 54,900 family memberships. A deployment is identified by its scene, dataset contract, and sorted candidate indices. When the same deployment occurs in multiple families, one physical record is retained and the membership table preserves its family links. The counts by $N = 1, \ldots, 9$ are respectively 3,655; 11,838; 6,480; 1,620; 7,200; 7,200; 7,200; 7,200; and 1,800.

Geographic separation. Scenes with overlapping geographic bounding boxes are connected, and the transitive connected components define 54 geographic groups. Each group belongs to a single split, preventing overlapping scene extents from crossing the training, validation, and test boundaries. The fixed assignments yield the composition in Table 4. All six cities occur in the dataset splits; this partition measures transfer to held-out geographic areas within the included cities. An evaluation that holds out an entire city requires a separate partition.

Table 4: Dataset partition. Size columns report numbers of scenes; deployment counts refer to unique records after deduplication.

Split	Groups	Scenes	128 m	256 m	512 m	Deployments
Training	36	42	8	12	22	38,162
Validation	9	9	2	3	4	8,019
Test	9	9	2	3	4	8,012
Total	54	60	12	18	30	54,193

A.5 Sample Schema and Data Integrity

Stored objects. Scene features are stored once per scene; family memberships are stored separately from unique deployment records. Each record contains identifiers, candidate indices, meter-valued and normalized coordinates, simulation provenance, cell counts, and coverage labels. Associated files store the four service maps, serving-TX indices, evaluation and coverage masks, and per-TX gains.

Frozen construction contract. The dataset contract fixes the radio profile, coordinate convention, deployment and evaluation policies, and minimum separation. The scene registry hashes the manifest, geometry, scene inputs, semantic masks, and candidates; the split manifest records geographic assignments. Deduplication retains the first stored evaluation and its seed, so repeated family memberships do not represent independent Monte Carlo replications.

B WIRELESS SIMULATION AND SERVICE UTILITY

This appendix specifies the forward operator in Eq. equation 2. Its calculation follows a single chain: per-TX propagation gains, received powers, serving association and interference, service maps, and cell-wise coverage labels. The same definitions apply to dataset construction and the simulator evaluation of planned deployments.

B.1 RADIO AND PROPAGATION CONFIGURATION

We use Sionna RT (Hoydis et al., 2023) with a fixed radio profile (Table 5). Each TX and RX uses a single-element isotropic antenna with vertical polarization and zero azimuth orientation. TXs are placed at a fixed height of $10\,\mathrm{m}$, and receivers lie on a horizontal plane at $1.5\,\mathrm{m}$. Ray tracing uses 5×10^6 samples per TX, maximum depth 5, and random seed 42. All active TXs use the same carrier and bandwidth. The configured propagation mechanisms are line of sight, specular and diffuse reflection, refraction, diffraction, and edge diffraction. Material scattering coefficients are set to 0.30 for concrete, 0.10 for medium-dry ground, and 0.05 for wet ground. The completed dataset records Sionna version 1.2.1 and the CUDA polarized ray-tracing backend.

Table 5: Fixed radio profile used to construct service labels. Transmit power is per TX; the receiver noise figure is included in the thermal-noise calculation.

Quantity	Value
Carrier frequency; system bandwidth W	$3.65\,\mathrm{GHz}$; $20\,\mathrm{MHz}$
TX power $P_{\mathrm{tx,dBm}}$; TX/RX heights	$24\,\mathrm{dBm}$; $10/1.5\,\mathrm{m}$
TX/RX antenna	Single element, isotropic, vertical polarization
Receiver noise figure F_{NF}	$7\,\mathrm{dB}$
Subcarrier spacing; resource blocks	$30\,\mathrm{kHz}$; 51
Subcarriers per block; SSB power offset Δ_{SSB}	12; $0\,\mathrm{dB}$
Receiver-cell size; maximum ray depth	$1 \times 1\,\mathrm{m}^2$; 5
Samples per TX	5×10^6
Throughput efficiency η; resource share ρ	0.65; 0.50
Spectral-efficiency cap $c_{\max}$	$6\,\mathrm{bit/s/Hz}$

The declared system bandwidth $W = 20\,\mathrm{MHz}$ is used in both the thermal-noise and throughput calculations. The $30\,\mathrm{kHz}$ subcarrier spacing and $51 \times 12 = 612$ active subcarriers are recorded for the reference-signal power conversion only; we do not replace W by the corresponding occupied bandwidth of $18.36\,\mathrm{MHz}$ in Eqs. equation 16 or equation 17.

B.2 RECEIVED POWER AND SERVING ASSOCIATION

Let $g_i(x) \geq 0$ be the linear propagation gain from TX i to receiver cell x. The received power in watts and serving index are

$$q_i(x) = P_{\mathrm{tx,W}} g_i(x), \qquad P_{\mathrm{tx,W}} = 10^{(P_{\mathrm{tx,dBm}}-30)/10}, \qquad s(x) = \arg\max_i q_i(x). \tag{12}$$

Ties use the first maximum in the stored TX order; equal-power ties do not change the scalar service metrics. For a positive serving gain, path loss and wideband received power are

$$\mathrm{PL}(x) = -10\log_{10} g_{s(x)}(x), \qquad P_{\mathrm{rx,dBm}}(x) = 10\log_{10} q_{s(x)}(x) + 30. \tag{13}$$

The reference-signal power proxy distributes wideband power over the configured subcarriers:

$$\text{SS-RSRP}(x) = P_{\mathrm{rx,dBm}}(x) - 10\log_{10}(N_{\mathrm{RB}} N_{\mathrm{SC/RB}}) + \Delta_{\mathrm{SSB}}, \tag{14}$$

where $N_{\mathrm{RB}} N_{\mathrm{SC/RB}} = 51 \times 12 = 612$. This calculation defines a reproducible reference-signal power label; it does not simulate an explicit synchronization-signal waveform.

B.3 INTERFERENCE AND SERVICE METRICS

SINR. Every non-serving active TX contributes co-channel interference. Received powers are added without coherent combining across TXs. With W expressed in hertz, receiver noise and linear

SINR are

$$P_{\mathrm{n,dBm}} = -174 + 10\log_{10} W + F_{\mathrm{NF}}, \qquad P_{\mathrm{n,W}} = 10^{(P_{\mathrm{n,dBm}}-30)/10}, \tag{15}$$

$$I(x) = \sum_{i \neq s(x)} q_i(x), \qquad \gamma(x) = \frac{q_{s(x)}(x)}{I(x) + P_{\mathrm{n,W}}}. \tag{16}$$

The SINR service map is $\mathrm{SINR}(x) = 10\log_{10}\gamma(x)$ in dB. Strongest-power association and interference are recomputed whenever the active TX set changes.

Effective throughput. We convert SINR to a capped Shannon-based throughput proxy,

$$T_{\mathrm{eff}}(x) = \frac{W\eta\rho}{10^6}\min\{\log_2(1+\gamma(x)), c_{\max}\} \quad [\mathrm{Mbps}]. \tag{17}$$

The fixed efficiency η and resource share ρ summarize implementation losses and resource allocation. With the configured values, the maximum is $39\,\mathrm{Mbps}$. The proxy measures service under a common allocation assumption; it does not model traffic-dependent scheduling or user load.

B.4 Coverage Labels and Joint Service Utility

Cell-level decisions. For metric j, let $b_j(x)$ indicate that its value is finite and satisfies the corresponding condition in Table 6. Marginal coverage and JM coverage are

$$C_j(S,P) = \frac{1}{|\Omega_S|}\sum_{x\in\Omega_S} b_j(x), \qquad U(S,P) = \frac{1}{|\Omega_S|}\sum_{x\in\Omega_S}\prod_{j=1}^{4} b_j(x). \tag{18}$$

Thus JM coverage is the spatial fraction of the intersection of the four service masks, rather than a product or average of the four marginal fractions. All road cells remain in the denominator, including cells without a valid serving link. Before metric computation, non-finite and non-positive propagation gains are replaced by zero. When no positive serving power exists, the logarithmic metrics are marked invalid and effective throughput is zero, so the cell fails the joint criterion.

Table 6: Service conditions defining $\boldsymbol{\tau}$. These are per-cell radio requirements; a target JM coverage is a separate deployment-level requirement.

Service map	Condition	Interpretation
Path loss	$\mathrm{PL} \leq 106\,\mathrm{dB}$	Serving-link attenuation
SS-RSRP	$\text{SS-RSRP} \geq -110\,\mathrm{dBm}$	Reference-signal power proxy
SINR	$\mathrm{SINR} \geq 5\,\mathrm{dB}$	Interference and noise tolerance
Effective throughput	$T_{\mathrm{eff}} \geq 10\,\mathrm{Mbps}$	Rate under the fixed resource share

Dependence between service conditions. The four service maps represent distinct radio quantities, but the corresponding binary threshold tests are not independent under the fixed benchmark profile.

First, because transmit power and the reference-signal conversion are fixed,

$$\text{SS-RSRP} = 24 - \mathrm{PL} - 10\log_{10}(612) = -\mathrm{PL} - 3.87 \quad [\mathrm{dBm}]. \tag{19}$$

Hence

$$\mathrm{PL} \leq 106\,\mathrm{dB} \quad \Longrightarrow \quad \text{SS-RSRP} \geq -109.87\,\mathrm{dBm},$$

which is slightly stronger than the benchmark requirement of $-110\,\mathrm{dBm}$.

Second, T_{eff} is a deterministic monotone function of SINR before the spectral-efficiency cap. At the benchmark SINR threshold,

$$\mathrm{SINR} = 5\,\mathrm{dB} \quad \Longrightarrow \quad T_{\mathrm{eff}} \approx 13.37\,\mathrm{Mbps},$$

which exceeds the $10\,\mathrm{Mbps}$ throughput requirement.

Therefore, for valid cells under the present radio profile and thresholds,

$$\mathbf{1}_{\mathrm{JM}}(x) = \mathbf{1}[\mathrm{PL}(x) \leq 106]\,\mathbf{1}[\mathrm{SINR}(x) \geq 5]. \tag{20}$$

This equivalence is specific to the chosen profile and thresholds; it is not part of the general definition of JM coverage. We retain all four service maps and marginal labels because they expose distinct physical/service quantities and allow the same benchmark interface to support alternative radio profiles or threshold settings.

C Related Work

C.1 Optimization-Based Models

Optimization-based methods directly search the deployment space using network utility as black-box feedback. Bayesian optimization improves evaluation efficiency by fitting a probabilistic surrogate to the objective, while TuRBO restricts modeling and candidate selection to adaptive local trust regions for scalable high-dimensional search (Eriksson et al., 2019). Heuristic methods instead explore candidate configurations through local or population-based updates, including particle-swarm and evolutionary optimization (Bonyadi & Michalewicz, 2017; Deb & Jain, 2014; Qiu et al., 2024). These approaches provide strong training-free planning baselines, but their effectiveness increasingly depends on whether repeated objective evaluations can resolve fragmented and highly coupled multi-TX configuration spaces.

C.2 Monte Carlo Sampling

Monte Carlo methods reinterpret planning as stochastic inference over a target distribution whose probability increases with deployment utility. Sequential Monte Carlo represents this distribution with weighted particles that are iteratively propagated, reweighted, and resampled (Piché et al., 2019; Naesseth et al., 2014; Lioutas et al., 2023). Importance-sampling methods approximate the same target from weighted proposal samples, with BR-SNIS reducing the bias induced by self-normalization (Cardoso et al., 2022). Gradient-informed MCMC methods such as NUTS instead exploit Hamiltonian dynamics to explore continuous high-probability regions (Hoffman & Gelman, 2014). These methods provide complementary mechanisms for exploring multi-modal deployment distributions, but their inference remains governed by the utility distribution defined by the underlying reward.

C.3 Diffusion Sampling

Diffusion-based planners generate solutions through iterative stochastic refinement, providing a flexible mechanism for modeling multi-modal planning distributions (Ho et al., 2020; Janner et al., 2022; Ubukata et al., 2024; Dong et al., 2024). Reward- and energy-guided formulations further steer this reverse process toward task objectives or physical constraints (Akhound-Sadegh et al., 2024; Pan et al., 2024; **?**). For wireless deployment, recent work formulates AP placement as sampling from a reward-induced Gibbs distribution and shows that deployment quality can be strongly limited by reward estimation accuracy (Yuan & Cheng, 2026). This makes wireless planning a particularly relevant setting for examining whether stronger generative inference can compensate for imperfect representations of physical deployment utility.

C.4 Relational Graph and Structured Reward Modeling

Relational models represent multi-entity systems through objects and their interactions rather than independent entity features. Interaction Networks explicitly decompose object-centric and relation-centric computation, while Neural Relational Inference learns latent interaction graphs directly from observations (Battaglia et al., 2016; Kipf et al., 2018). These formulations provide a natural basis for multi-TX planning, where the utility of one transmitter depends on the remaining configuration through coverage complementarity, redundancy, and interference.

Standard graph formulations primarily encode pairwise relations. However, many physical systems exhibit collective effects that arise only from groups of interacting entities (Battiston et al., 2021). Hypergraph neural networks represent such dependencies through hyperedges connecting multiple entities (Feng et al., 2019), while hypergraph energy formulations provide a principled way to express these higher-order relations as structured potentials (Wang et al., 2023). For wireless deployment, this view extends the reward from a scalar assessment of a complete TX set to an interaction-aware utility representation over individual, pairwise, and higher-order TX relations. The open question is therefore not simply whether relational modeling improves planning, but which interaction structure provides the appropriate reward representation under different urban environments and deployment scales.

D UTILITY MODELS AND TRAINING DETAILS

This appendix specifies the predictors in Section 2.2, using the saved base-model architectures and training configurations.

D.1 SCENE AND TX-SET INPUTS

Scene rasters. The input contains 11 channels: buildings, roads, green areas, water, obstacles, built-up areas, explicitly open areas, unknown features, background, TX feasibility, and the stored evaluation mask. Rasters retain their native 128^2, 256^2, or 512^2 resolution; batches share one scene and therefore one spatial shape. The stored evaluation-mask feature can represent non-water areas; labels use the road-only domain Ω_S (Appendix B.4).

Spatial encoding. A convolutional encoder produces a dense feature field D_S and a global scene embedding g_S. At each normalized TX location u_i, bilinear interpolation extracts a local feature $\ell_i = D_S(u_i)$ using border padding and `align_corners=False`. Physical dimensions (L_x, L_y) accompany the coordinates, and an active-TX mask excludes padding from mean/max reductions. All base configurations use the cardinality feature $c = N/9$. For Reward, IN, and NRI, the map-size feature is $s = \frac{1}{2}\log_2((L_x, L_y)/128)$, applied elementwise. HPEM instead uses

$$s_H = \left[\log \frac{(L_x, L_y)}{512},\ \frac{(L_x, L_y)}{\sqrt{L_x^2 + L_y^2}}\right]. \tag{21}$$

Table 7: Saved base-model architectures. Parameter counts include auxiliary heads. IN and NRI share the scene encoder structure but have different relation modules.

Component	Reward	IN	NRI	HPEM
Convolution channels	32/64/96	32/64/48	32/64/48	32/64/48
Kernel sizes	5/3/3	5/3/3	5/3/3	5/3/3
Strides	2/2/2	2/2/1	2/2/1	2/2/1
GroupNorm groups	4/8/8	8/8/8	8/8/8	8/8/8
Local / global scene width	96/96	48/96	48/96	48/128
TX embedding width	128	128	128	192
Relation / hyperedge width	–	128	128	160
Output-head hidden widths	256/128/64	192/192/192	192/192/192	192/192
Output nonlinearity	Sigmoid	Sigmoid	Sigmoid	None
Output-head dropout	0.2	0.2	0.2	0
Auxiliary scalar heads	0	0	0	5
Trainable parameters	244,641	521,649	833,909	1,270,778

D.2 SET-POOLED AND RELATIONAL PREDICTORS

Reward Predictor. A two-layer TX encoder maps $[u_i, \ell_i]$ to $h_i \in \mathbb{R}^{128}$. The output head receives the masked mean and maximum of $\{h_i\}$ together with g_S, s, and c. Its sigmoid output lies in $[0, 1]$. Convolutions use GroupNorm and SiLU; the head uses SiLU and dropout.

Interaction Network. The IN node encoder receives $[u_i, \ell_i, g_S, s, c]$. For receiver i and sender $j \neq i$, it constructs

$$a_{ij} = [h_i, h_j, \delta_{ij}, |\delta_{ij}|, \|\delta_{ij}\|_2], \qquad \delta_{ij} = u_j - u_i. \tag{22}$$

A shared relation MLP gives $v_{ij} = f_{\text{rel}}(a_{ij})$; node updates use mean incoming messages:

$$h_i' = f_{\text{node}}\left(h_i, \frac{1}{N-1}\sum_{j\neq i} v_{ij}, g_S, s, c\right), \qquad N > 1. \tag{23}$$

The readout combines mean/max-pooled updated nodes, the mean of all valid directed relation features, and the scene context. There is one message-passing block and no self-loop; for $N = 1$, the relation aggregate is zero.

NRI-style predictor. NRI encodes a_{ij} as four logits and averages opposite directions before computing type probabilities. Type zero contributes no message; other types use separate MLPs:

$$v_{ij} = \sum_{k=1}^{3} q_{ij,k}\, f_k(a_{ij}). \tag{24}$$

Training uses soft Gumbel–Softmax samples at temperature 0.5. Evaluation uses deterministic softmax probabilities without Gumbel noise, yielding stable predictions and coordinate gradients. Although the logits are symmetrized, directed message features can differ between $i \to j$ and $j \to i$. The IN aggregation pattern provides the readout; edge types have no observed physical labels.

D.3 Hypergraph Potential Energy Model

TX embeddings. HPEM augments positions with Fourier features

$$\psi(u_i) = \left[u_i, \{\sin(2\pi f u_i), \cos(2\pi f u_i)\}_{f\in\{1,2,4\}}\right]. \tag{25}$$

A shared MLP encodes $[\psi(u_i), \ell_i, g_S, s_H, c]$ into $h_i \in \mathbb{R}^{192}$. MLPs in this model have two SiLU hidden layers followed by a linear output layer. The singleton energy is obtained from a scalar head on each node, followed by a masked mean.

Symmetric subset features. For each enabled order $k \geq 2$, all active k-TX combinations are enumerated. Order-specific encoders and mean/max pooling produce symmetric member features. This is concatenated with an eight-dimensional geometric descriptor: normalized centroid (2), coordinate standard deviations (2), and mean, standard deviation, minimum, and maximum of within-subset pair distances (4). Distances are computed in meters and divided by the map diagonal; standard deviations use the population convention. The centroid's scene feature, g_S, s_H, and c complete the input to an order-specific feature MLP and energy head.

Order aggregation and global readout. Let $\mathcal{H}_k(P)$ contain all active subsets of size k. The model averages their energies,

$$\overline{E}_k = \begin{cases} |\mathcal{H}_k(P)|^{-1} \sum_{e\in\mathcal{H}_k(P)} E_{\theta,k}(X_S, e), & \mathcal{H}_k(P) \neq \varnothing, \\ 0, & \mathcal{H}_k(P) = \varnothing, \end{cases} \tag{26}$$

and uses zero for disabled orders. With $\overline{h}$ denoting the masked mean of node embeddings, the global and final energies are

$$E_{\text{global}} = f_g([\overline{h}, g_S, s_H, c, \overline{E}_1, \ldots, \overline{E}_4]), \qquad E_\theta = E_{\text{global}} + \sum_{k\in\mathcal{K}} w_k \overline{E}_k. \tag{27}$$

The base model enables orders 1–4 with $w_k = 1$. Its prediction $\widehat{U}_\theta = -E_\theta$ has no sigmoid or inference-time clipping. Order means control subset-count scaling, and the global branch couples their contributions. Five auxiliary heads read the same global input through a separate shared trunk and predict JM, path-loss, SS-RSRP, SINR, and effective-throughput coverage.

Permutation invariance. Shared encoders, exhaustive active pairs/subsets, and symmetric reductions preserve predictions when coordinates and masks are permuted together, up to floating-point error. This holds in deterministic evaluation mode; independent NRI training-time Gumbel draws need not agree across forward passes. Coordinate gradients exist almost everywhere.

D.4 Training Objectives and Sampling

Supervision. Let $y_b = U(S_b, P_b)$ be the JM label for deployment b. Reward and IN minimize $\mathcal{L}_{\text{MSE}} = \text{Mean}_b(\widehat{U}_b - y_b)^2$. NRI adds a regularizer toward a uniform four-type distribution:

$$\mathcal{L}_{\text{NRI}} = \mathcal{L}_{\text{MSE}} + 10^{-4}\mathcal{L}_{\text{edge}}, \qquad \mathcal{L}_{\text{edge}} = \text{Mean}_b\, \text{Mean}_{i\neq j \text{ active}} \sum_{k=0}^{3} q_{bij,k} \log \frac{q_{bij,k}}{1/4}. \tag{28}$$

The regularizer uses the sampled soft probabilities during training and is zero for a deployment without valid TX pairs.

HPEM objectives. The energy model combines utility regression, ranking, and auxiliary supervision:

$$\mathcal{L}_{\mathrm{HPEM}} = \mathrm{MSE}(-E_\theta, y) + 0.25\mathcal{L}_{\mathrm{rank}} + 0.25\mathcal{L}_{\mathrm{aux}}, \tag{29}$$

$$\mathcal{L}_{\mathrm{rank}} = \frac{1}{|\mathcal{Q}|} \sum_{(a,b)\in\mathcal{Q}} [0.02 + \mathrm{sign}(y_a - y_b)(E_a - E_b)]_+ . \tag{30}$$

Here $\mathcal{Q}$ contains within-batch pairs with $a < b$ and $|y_a - y_b| > 10^{-4}$. The ranking loss is zero if $\mathcal{Q}$ is empty. The auxiliary loss averages the five coverage-head MSEs. Planning uses the main prediction $-E_\theta$.

Batch and sample units. Batches contain one scene and one TX cardinality, so ranking compares deployments at a common scene and size. Pairs need not belong to the same master family, and the ranking loss does not compare different cardinalities. TX order is randomly permuted during training. The HPEM loader expands family memberships, while Reward, IN, and NRI use unique physical deployments. Repeated membership therefore changes some HPEM sample weights. The geographic split is shared, while sampling weights, capacity, and auxiliary supervision differ.

D.5 Base-Model Configuration

All base models use AdamW with $(\beta_1, \beta_2) = (0.9, 0.999)$ and $\epsilon = 10^{-8}$. The saved runs use seed 42, batch size at most 32, gradient-norm clipping at 5, and a maximum of 100 epochs. HPEM uses cosine learning-rate decay over the configured training duration and selects the checkpoint with the lowest total validation objective. The remaining predictors use validation-MSE plateau scheduling and early stopping. Table 8 gives the remaining settings.

Table 8: Training settings from the saved base runs. HPEM uses a different selection loss and sample weighting from the other predictors.

Setting	HPEM	Reward / IN / NRI
Learning rate	3×10^{-4}	10^{-4}
Weight decay	10^{-4}	10^{-5}
Learning-rate schedule	Cosine, period 100, minimum 0	Plateau, factor 0.5, patience 5, minimum 10^{-6}
Early-stopping patience	Disabled	15 epochs
Checkpoint criterion	Strictly lower total validation loss	Validation MSE improvement $> 10^{-5}$
Sampling unit	Family membership	Unique deployment
Auxiliary objectives	Ranking and five coverage heads	None / none / edge KL

D.6 Target-Scene Adaptation

Adaptation updates all parameters with the base losses. From a pool of 990 unique deployments, separate from 200 evaluation examples, 297 examples are used for 30 epochs, followed by all 990 for 70 additional epochs. AdamW is initialized once and retained between stages: learning rates are 3×10^{-5} for HPEM and 10^{-5} otherwise, with weight decay 10^{-5}, batch size 32, clipping at 5, and seed 42. There is no scheduler or early stopping; end-of-stage checkpoints are saved without evaluation-based selection. The retained figures call full-pool fine-tuning *In-distribution*; target-scene and target-city provenance remain separate properties.

E Planning Algorithms and Implementation

All planners share a learned utility interface and a final mapping to road deployments. They differ in how they propose and update continuous TX coordinates. This appendix specifies the search and selection procedures.

E.1 Shared Planning Interface and Constraints

Utility and search score. For fixed S and N, write $\mathbf{u} = (u_1, \ldots, u_N) \in [0,1]^{N\times 2}$. The model receives the scene features X_S, normalized coordinates, and a mask identifying active TXs, and returns $\widehat{U}_\theta(X_S, \mathbf{u})$. The scalar prediction approximates the simulator utility $U(S,P)$. During continuous search, all planners evaluate

$$r_S(\mathbf{u}) = \widehat{U}_\theta(X_S, \mathbf{u}) - \lambda_{\text{road}}\phi_{\text{road}}(\mathbf{u}) - \lambda_{\text{sep}}\phi_{\text{sep}}(\mathbf{u}). \tag{31}$$

Let $m_S(u)$ be the bilinear interpolation of the binary road feasibility mask and $[a]_+ = \max(a, 0)$. The penalties are

$$\phi_{\text{road}}(\mathbf{u}) = \frac{1}{N}\sum_i [1 - m_S(u_i)]_+, \tag{32}$$

$$\phi_{\text{sep}}(\mathbf{u}) = \binom{N}{2}^{-1} \sum_{i<j} \left[1 - \frac{\|p_i - p_j\|_2}{d_{\min}}\right]_+^2, \quad N \geq 2, \tag{33}$$

with $\phi_{\text{sep}} = 0$ for $N = 1$ and p_i recovered using Eq. equation 10. Mask interpolation uses zero padding and cell-centered coordinates (`align_corners=False`). The score r_S is a search objective, not a physical coverage fraction. Soft penalties encourage feasibility during search; the discrete checks in Appendix E.5 enforce it on the returned deployment.

Bounded and unconstrained coordinates. Reflection acts elementwise as $R(a) = 1 - |1 - (a \bmod 2)|$. The diffusion planner uses it to keep updates in the unit cube. For Monte Carlo methods using unconstrained coordinates, let $\mathbf{u} = \text{sigmoid}(z)$. Their unnormalized log target is

$$\ell_\beta(z) = \beta r_S(\text{sigmoid}(z)) + \sum_{k=1}^{2N} \log\big(u_k(1-u_k)\big), \tag{34}$$

where $\beta > 0$ controls concentration around high scores. The Jacobian term preserves the intended density $\pi_\beta(\mathbf{u}) \propto \exp(\beta r_S(\mathbf{u}))$ with respect to uniform volume on the cube. Numerical logit conversion clips coordinates to $[10^{-5}, 1 - 10^{-5}]$.

E.2 Diffusion-Based Planning

Reward-induced score. The diffusion planner constructs a score from the learned utility and its coordinate gradients; it does not require a separately trained denoiser. For Gaussian perturbations $\epsilon \sim \mathcal{N}(0, I)$, define the smoothed positive function

$$h_\sigma(\mathbf{u}) = \mathbb{E}_\epsilon\left[\exp(\beta r_S(R(\mathbf{u} + \sigma\epsilon)))\right]. \tag{35}$$

At each step, M perturbations $v_m = \mathbf{u} + \sigma\epsilon_m$ estimate its log gradient:

$$w_m = \frac{\exp(\beta r_S(R(v_m)))}{\sum_{a=1}^M \exp(\beta r_S(R(v_a)))}, \qquad \widehat{s}_\sigma(\mathbf{u}) = \beta \sum_{m=1}^M w_m \nabla_{v_m} r_S(R(v_m)). \tag{36}$$

Gradients pass through reflection, and antithetic perturbations pair ϵ with $-\epsilon$. Equation equation 36 is a self-normalized Monte Carlo approximation, with finite-sample error.

Annealed updates. Noise scales decrease geometrically from $\sigma_{\max}$ to $\sigma_{\min}$ over T steps. Starting from $R(\frac{1}{2} + \sigma_{\max}\epsilon)$, the planner applies a variance-exploding reverse-diffusion-style update,

$$\mathbf{u}_{t+1} = R\Big(\mathbf{u}_t + \Delta_t \widehat{s}_{\sigma_t}(\mathbf{u}_t) + \sqrt{\Delta_t}\,\xi_t\Big), \qquad \Delta_t = \sigma_t^2 - \sigma_{t+1}^2, \tag{37}$$

where $\xi_t \sim \mathcal{N}(0, I)$ and the last stochastic term is disabled by default. The implementation retains the best visited state of each trajectory. This finite-step search does not establish exact sampling from π_β.

E.3 Monte Carlo Planning

Annealed SMC. Sequential Monte Carlo evolves L particles through temperatures $\beta_t = \beta(t/T)^a$, starting from uniform samples on the unit cube (Cappé et al., 2007). Before mutation, particle log weights are incremented by

$$\log \widetilde{w}_t^{(l)} = \log w_{t-1}^{(l)} + (\beta_t - \beta_{t-1}) r_S(\mathbf{u}_{t-1}^{(l)}). \tag{38}$$

After normalization, systematic resampling is triggered when $\mathrm{ESS} = 1/\sum_l (w_t^{(l)})^2$ falls below κL; resampled weights are reset to uniform. Mutation operates in logit coordinates with target ℓ_{β_t}. *SMC-Gaussian* uses a Gaussian random walk and a Metropolis–Hastings acceptance test. *SMC-Langevin* proposes $z' = z + \frac{1}{2}s_t^2 \nabla \ell_{\beta_t}(z) + s_t \epsilon$ and includes the forward and reverse proposal densities in the acceptance ratio. The proposal scale s_t decreases geometrically. A score-ranked archive retains high-scoring particle states across temperatures.

NUTS. The No-U-Turn sampler (Hoffman & Gelman, 2014) also uses ℓ_β in logit coordinates. It samples Gaussian momenta with an identity mass matrix and recursively extends a leapfrog trajectory until a no-U-turn condition, a divergence test, or the maximum tree depth stops the expansion. Each chain adapts its step size by dual averaging during warmup. The planner archives only post-warmup states, and final deployment selection uses their search scores.

BR-SNIS adaptation. The implementation uses the iterated sampling–importance resampling mechanism underlying BR-SNIS (Cardoso et al., 2022). Each chain combines its retained deployment with M new uniform proposals on the cube. It selects the next retained state from these $M + 1$ candidates with probabilities proportional to $\exp(\beta r_S)$. All candidates are eligible for the score-ranked archive. This adapts a sampling mechanism to deployment search: the returned plan is a concrete candidate, and no weighted average of TX coordinates is used.

E.4 Derivative-Free Optimization

Greedy construction. A scrambled Sobol sequence provides a finite pool of normalized TX locations. Beam search grows deployments from one TX to N TXs, scoring all single-location extensions and retaining the highest-scoring beam at each cardinality. Coordinate-replacement rounds then refine the complete sets using the same pool. The method therefore queries the utility model on partial as well as complete deployments. It is a search heuristic for the learned score; no submodularity assumption is imposed.

Particle swarm optimization. PSO maintains a population of deployments, their personal best states, and a global or ring-neighborhood best state. Velocity updates combine inertia with random cognitive and social attraction terms. The inertia weight decreases linearly, velocities are clipped, and positions are clamped to the unit cube. An out-of-bounds velocity component is reversed and halved. The archive retains the best visited deployments.

Genetic algorithm. The GA applies tournament selection, simulated binary crossover, and polynomial mutation to flattened $2N$-dimensional coordinates. Elitism carries high-scoring deployments into the next generation, and offspring remain within $[0, 1]^{2N}$. When no positive mutation probability is specified, the per-coordinate probability is $1/(2N)$.

TuRBO. The TuRBO-1 adaptation (Eriksson et al., 2019) fits a Gaussian process with an ARD Matérn-$5/2$ kernel to standardized search scores. A trust region is centered on the current best point, with side lengths scaled by the learned length scales. Thompson samples over Sobol candidates select the next evaluation batch. Consecutive improvements expand the trust region; consecutive failures shrink it, and a region below the minimum size triggers a new Sobol initialization. All restarts share one explicit score-evaluation budget, and the final archive pools evaluated candidates across restarts.

E.5 Candidate Selection and Feasibility Repair

Selection order. Each planner returns a collection of candidate deployments and their continuous search scores. The shared procedure sorts these scores in descending order and maps every TX to its

nearest road candidate in meters. A mapped set is accepted only if its candidate locations are distinct and all pairwise distances satisfy $d_{\min}$, up to a 10^{-6} m numerical tolerance. The first valid mapped set is returned.

Repair fallback. When a mapped set fails, the procedure attempts a constrained assignment of its TXs to road candidates. It tests cyclic rotations of both the original and reversed TX orders. For each order, each TX greedily takes its nearest unused candidate that is sufficiently separated from already assigned TXs. Among successful orders, the assignment with the smallest total squared displacement is retained. The first successful repair in score order is saved as a fallback while the scan continues for a directly valid nearest-candidate mapping. The fallback is used only when the entire scan yields no such mapping; if neither is available, the procedure reports failure. This finite assignment heuristic need not find the globally nearest feasible deployment.

Final evaluation. The selected coordinates define the executable deployment $P \in \mathcal{F}_N(S)$. Its unpenalized model prediction is recomputed, and its service maps and $U(S, P)$ are then obtained from the simulator. The selection order is based on scores before mapping; candidates are not reranked by post-mapping predictions or simulator outcomes.

E.6 Algorithm Configurations

Table 9 records implementation defaults; run configurations specify overrides, population size, and seeds. These reference settings do not imply equal compute budgets across methods. The single-run interface defaults to $\lambda_{\text{road}} = \lambda_{\text{sep}} = 10$. The TX cardinality N, particle count L, and number of score evaluations are separate quantities.

Table 9: Reference implementation defaults. $d = 2N$ is the coordinate dimension. SMC-Gaussian and SMC-Langevin share the annealing settings; only their mutation kernels differ.

Planner	Default configuration
Diffusion	100 steps; $M = 10$; $\beta = 1$; $\sigma : 0.35 \to 0.005$; antithetic noise; reflection; deterministic final step
SMC	100 temperatures; $\beta = 1$; $a = 1$; $\kappa = 0.5$; one mutation per temperature; $s : 0.45 \to 0.08$
NUTS	50 warmup and 50 retained iterations per chain; $\beta = 1$; initial step size 0.15; target acceptance 0.8; maximum depth 7; divergence threshold 1,000
BR-SNIS	100 rounds; 32 new uniform proposals per chain and round; $\beta = 1$
Greedy	256 candidate locations; one coordinate-replacement round; scoring batches of at most 4,096 deployments
PSO	100 steps; inertia $0.9 \to 0.4$; cognitive/social coefficients 1.49618; velocity limit 0.2; global topology
GA	100 generations; tournament size 3; elite fraction 0.1; crossover probability 0.9; mutation probability $1/d$; crossover/mutation indices 15/20
TuRBO	200 score evaluations; $\max(2d, 8)$ initial points; batch size 4; 256 candidates; 25 GP fitting steps; trust-region length 0.8, bounds $[2^{-7}, 1.6]$; 3 successes to expand; $\max(4, \lceil d/4 \rceil)$ failures to shrink

Recorded outputs. The evaluation interface stores selected coordinates, post-mapping predictions, simulated service maps and utility, configurations, and search traces. Score evaluations, gradients, and simulator calls are distinct operations: diffusion uses multiple perturbed model evaluations per update, while NUTS trajectory lengths vary. Iteration counts alone therefore do not specify equal computation.

F Evaluation Protocols and Metrics

This appendix consolidates the task metrics used in Section 3.2. Physical service labels are defined once in Appendix B; learned search scores and final selection are defined in Appendix E.

F.1 Fixed-Cardinality Evaluation

For every scene, cardinality, method, and seed, the planner returns one deployment after road mapping and separation checks. The simulator produces $U(S, P)$ and the four marginal coverage fractions

$C_j(S, P)$. Reported percentages are $100U$ and $100C_j$. Higher coverage is preferred; lower planning time is preferred. Comparisons use the same scene, active TX count, deployment constraints, radio profile, and service thresholds.

F.2 Target Coverage and Failure Handling

For a trial r, let $\mathcal{N}_r$ be the consecutively tested TX counts and $U_{r,N}$ the verified utility of its selected deployment at count N. The success indicator and first successful count are

$$s_r = \mathbf{1}[\exists N \in \mathcal{N}_r : U_{r,N} \geq \gamma_{\mathrm{tar}}], \qquad \widehat{N}_r = \min\{N \in \mathcal{N}_r : U_{r,N} \geq \gamma_{\mathrm{tar}}\}. \tag{39}$$

The value $\widehat{N}_r$ is defined only for successful trials. Success rate is $100R^{-1}\sum_{r=1}^{R} s_r$, and Best JM is $100 \max_{N \in \mathcal{N}_r} U_{r,N}$ for a trial with verified deployments. The search stops at its first success or exhausts the configured count range. A missing smaller-count result does not establish minimality across that gap. When reporting required TX counts over successful seeds, their conditional nature is explicit and success rate is reported alongside them.

F.3 Transfer and Adaptation Provenance

Geographic-group test scenes, external scenes, and held-out cities have different relationships to the training data. The six-city split in Appendix A.4 provides geographic-group separation. For target adaptation, the saved scene IDs, base-checkpoint split, and disjoint adaptation/evaluation record lists determine the transfer setting. The retained adaptation configurations use `paris_11`, `new_york_12`, and `tokyo_01`; the first two are outside the base dataset, while `tokyo_01` belongs to its test split. All three cities occur in base training. The labels Zero-shot, Few-shot, and In-distribution in the retained figures refer to adaptation stages; their precise training procedure is given in Appendix D.6.

F.4 Computation and Aggregation

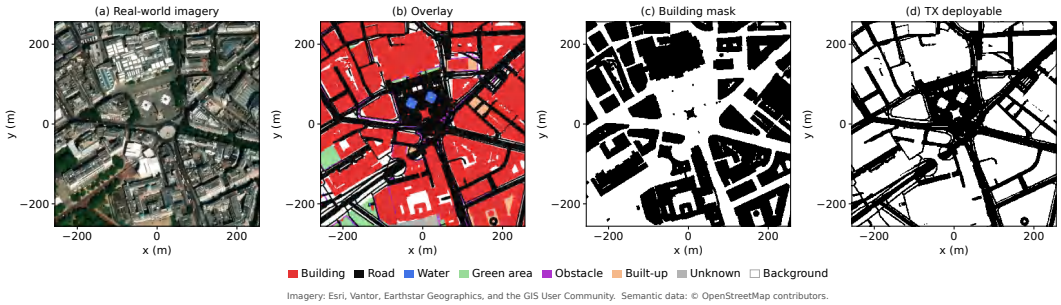
The single-deployment evaluator measures GPU-synchronized planner execution separately from road mapping and ray tracing. For target coverage, cumulative planning time adds the completed planner calls across tested TX counts; ray-tracing time and end-to-end wall time are separate fields. Model evaluations count candidate-level calls, including the perturbed candidates and gradient-bearing calls used by the search method. Iteration counts, population sizes, model evaluations, and physical simulations are therefore reported as distinct budget quantities.

Repeated trials are summarized with means and standard deviations where available, using matched scene, cardinality, and radio conditions. Target-coverage failures contribute to success rate rather than disappearing from the comparison. Numerical results and their existing captions are retained in the experimental sections; this appendix defines the quantities needed to interpret them.

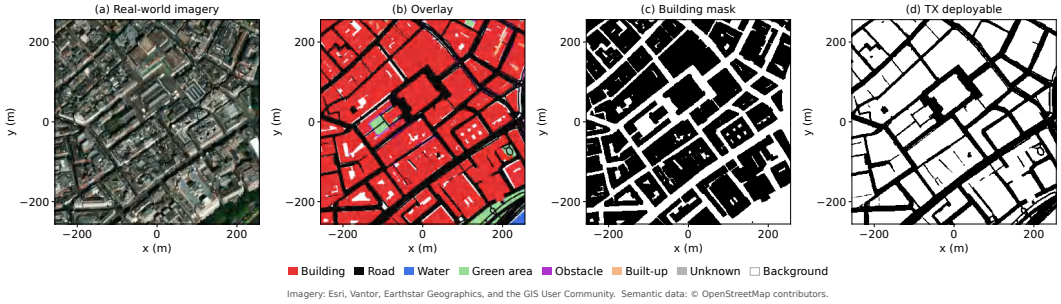
G Additional Results and Visualizations

G.1 Urban Scene Visualizations across Six Cities

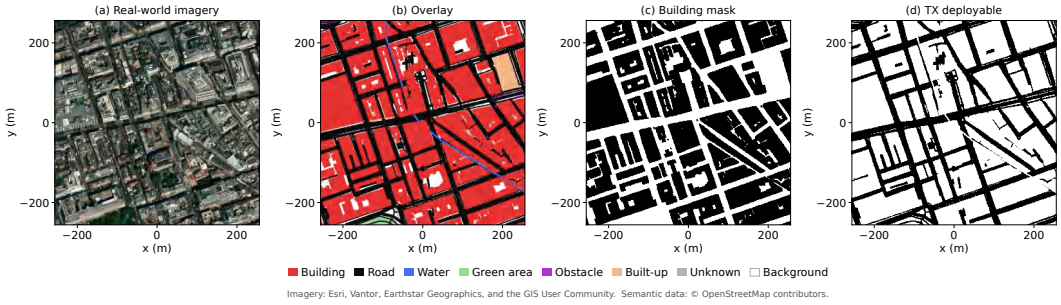
We present three landmark-centered examples from each of the six cities. Each example covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. From left to right, each row shows geographically aligned real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask. Coordinates are expressed in meters in the local scene frame. The examples illustrate different urban layouts, including public squares, dense commercial districts, circular intersections, and waterfront areas. Imagery acquisition dates may differ from those of the OpenStreetMap snapshots; the original masks are displayed without modification. Imagery and map-data credits are retained in each row.



(i) Trafalgar Square (`london_01`).

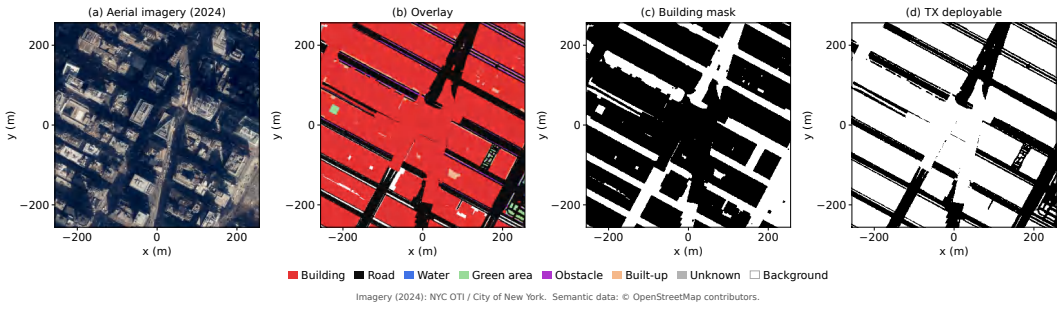


(ii) Covent Garden Market / Royal Opera House (`london_02`).

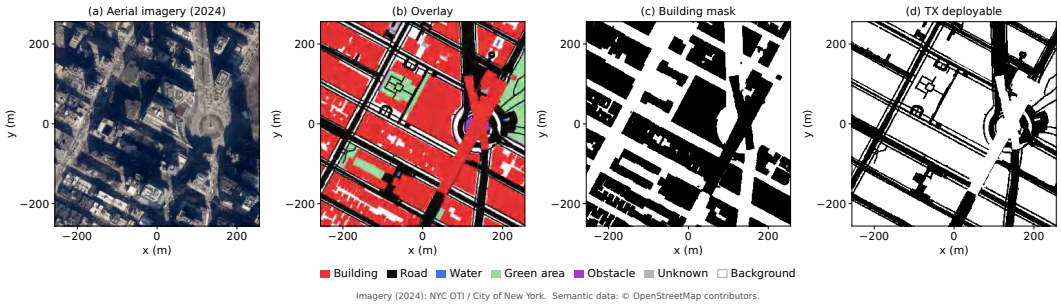


(iii) Oxford Street / Bond Street (`london_03`).

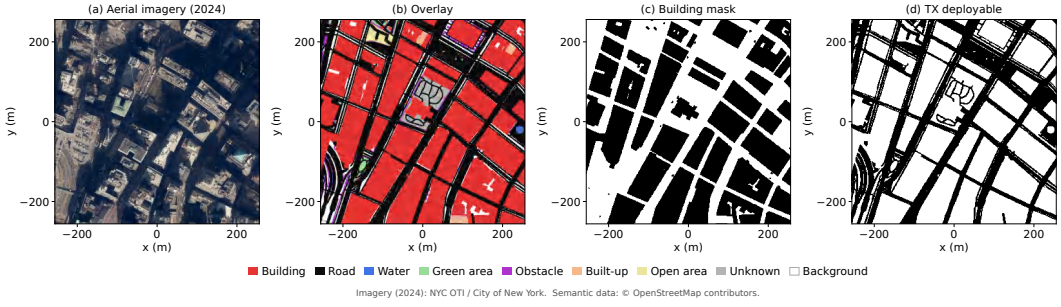
Figure 5: Urban scene examples in London. Each row covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region and shows real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask, from left to right.



(i) Times Square south / 42nd Street (`new_york_02`).

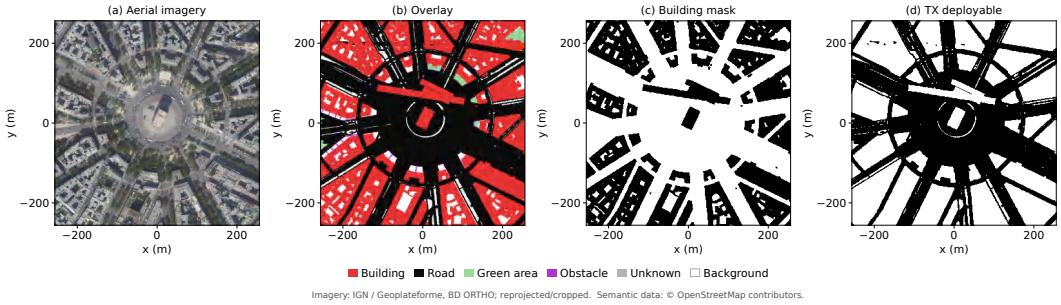


(ii) Columbus Circle (`new_york_03`).

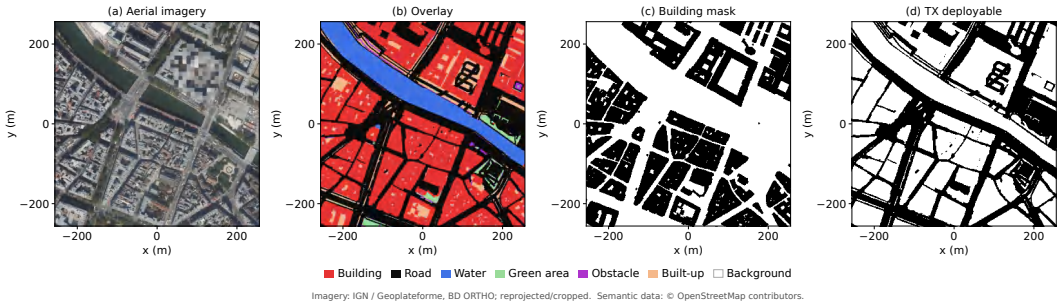


(iii) Wall Street / New York Stock Exchange (`new_york_04`).

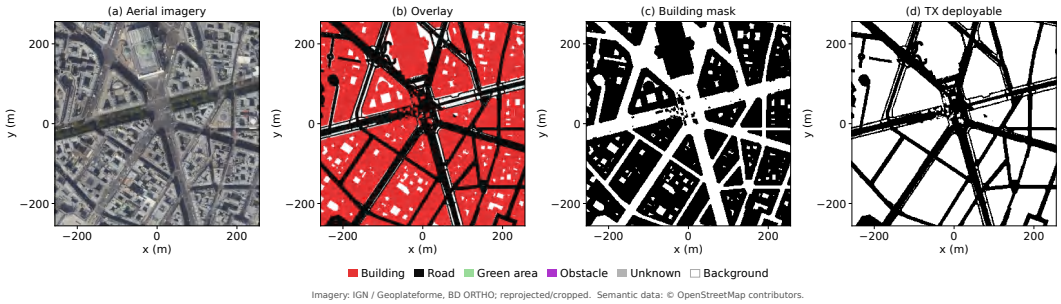
Figure 6: Urban scene examples in New York. Each row covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region and shows real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask, from left to right. The Times Square example shows its southern portion.



(i) Arc de Triomphe / Place Charles-de-Gaulle (`paris_01`).

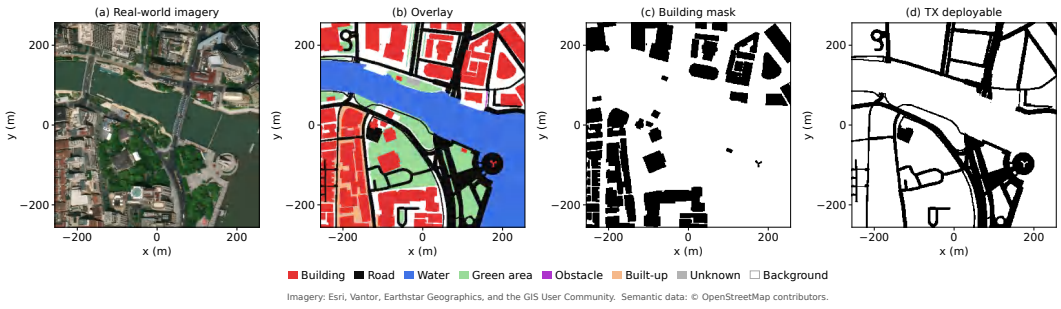


(ii) Place Saint-Michel / Sainte-Chapelle surroundings (`paris_03`).

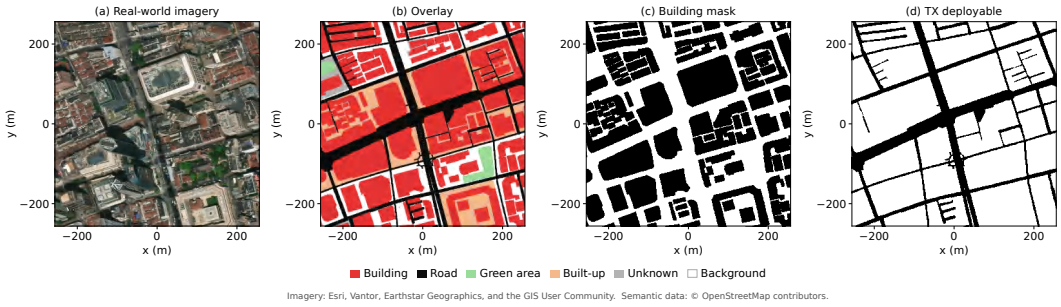


(iii) Palais Garnier / Place de l'Opéra (`paris_05`).

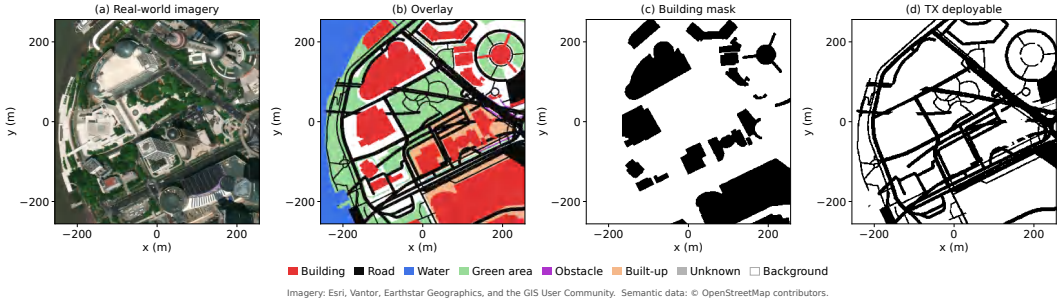
Figure 7: Urban scene examples in Paris. Each row covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region and shows real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask, from left to right. The pixelated patch in the second row is present in the source imagery.



(i) Waibaidu Bridge / Huangpu Park (`shanghai_01`).

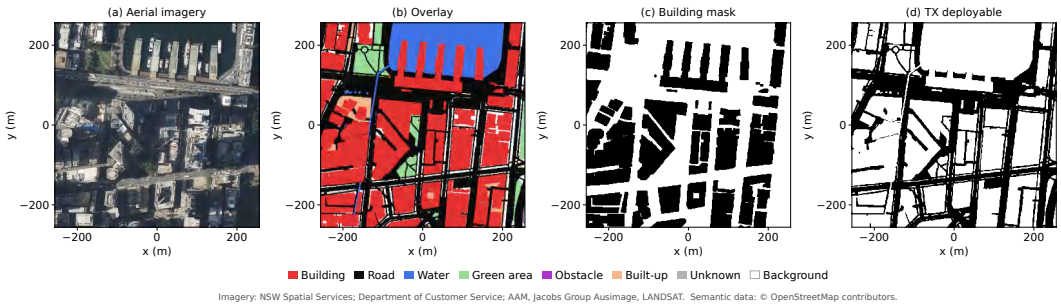


(ii) East Nanjing Road pedestrian street (`shanghai_02`).

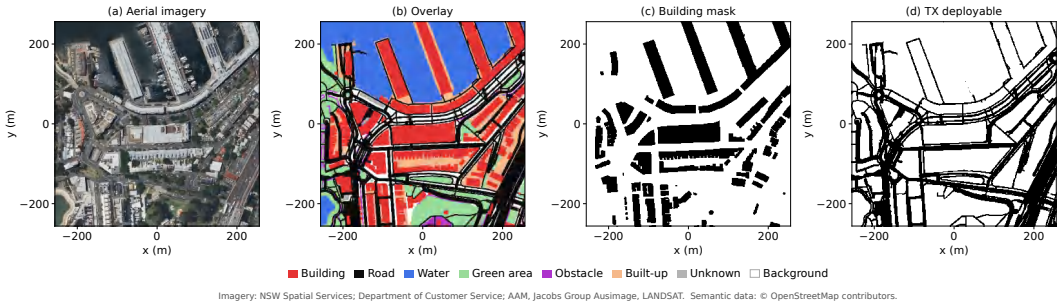


(iii) Oriental Pearl Tower / Lujiazui waterfront (`shanghai_05`).

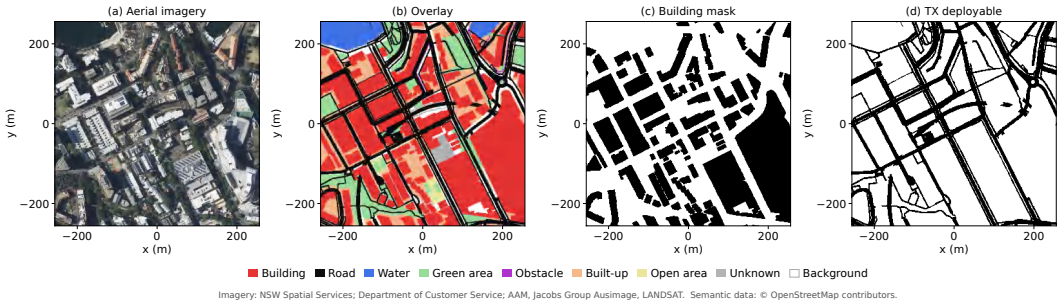
Figure 8: Urban scene examples in Shanghai. Each row covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region and shows real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask, from left to right.



(i) Circular Quay / Customs House (`sydney_01`).

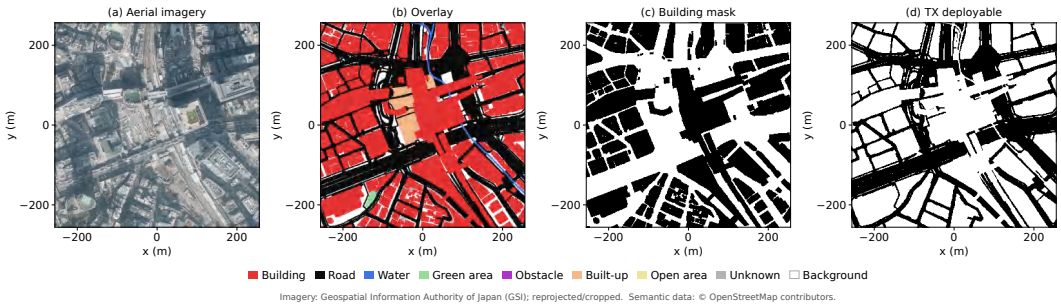


(ii) Observatory Hill / The Rocks surroundings (`sydney_02`).

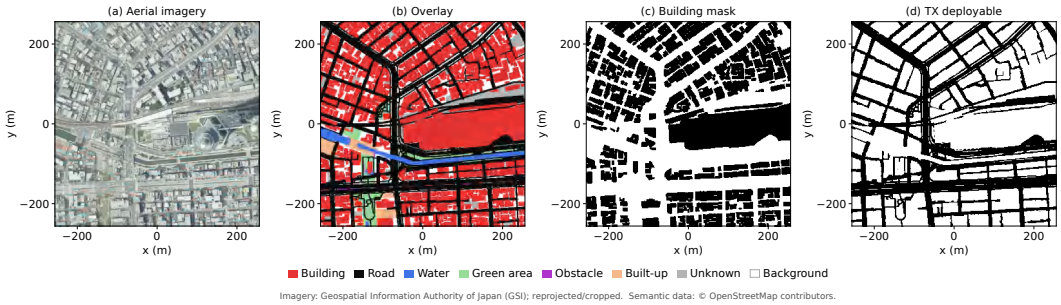


(iii) The Star / Darling Island, Pyrmont (`sydney_03`).

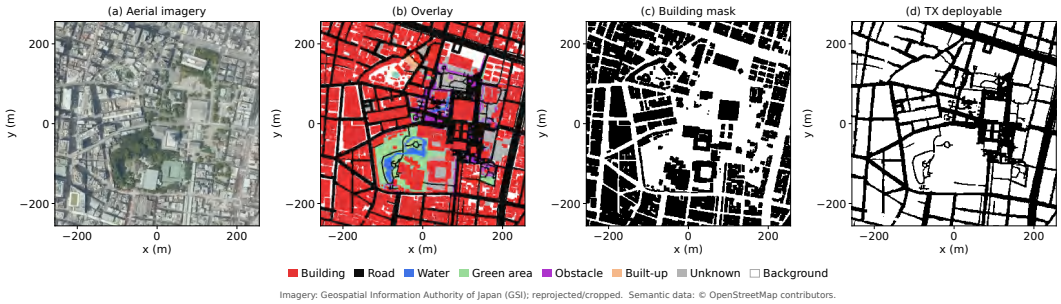
Figure 9: Urban scene examples in Sydney. Each row covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region and shows real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask, from left to right.



(i) Shibuya Scramble Crossing (`tokyo_02`).



(ii) Tokyo Skytree / Tokyo Solamachi (`tokyo_03`).



(iii) Sensō-ji / Nakamise (`tokyo_04`).

Figure 10: Urban scene examples in Tokyo. Each row covers a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region and shows real-world imagery, the semantic overlay, the building mask, and the TX-deployable mask, from left to right.

G.2 Qualitative Comparison of Deployment Layouts and Radio Maps

We visualize deployment layouts and Sionna-evaluated radio maps for nine sampling methods across six urban scenes. All methods use HPEM as the reward model and deploy $N_{\mathrm{TX}} = 9$ transmitters within a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region.

From top to bottom, the rows correspond to Diffusion, WS-L, WS-G, NUTS, BR-SNIS, PSO, GA, Greedy, and TuRBO. From left to right, the columns show joint coverage, path loss, SS-RSRP, SINR, and effective throughput. Stars indicate TX locations, and each panel title reports the corresponding coverage rate. In the joint coverage maps, red and blue denote road locations that satisfy and fail the joint coverage criterion, respectively.

Each figure uses a common seed across all nine methods and shows individual deployments rather than seed-averaged results. Seeds were selected based on availability across methods, not deployment performance. Continuous metric maps retain their original per-panel color scales; their numerical colorbar values should therefore be consulted when comparing panels.

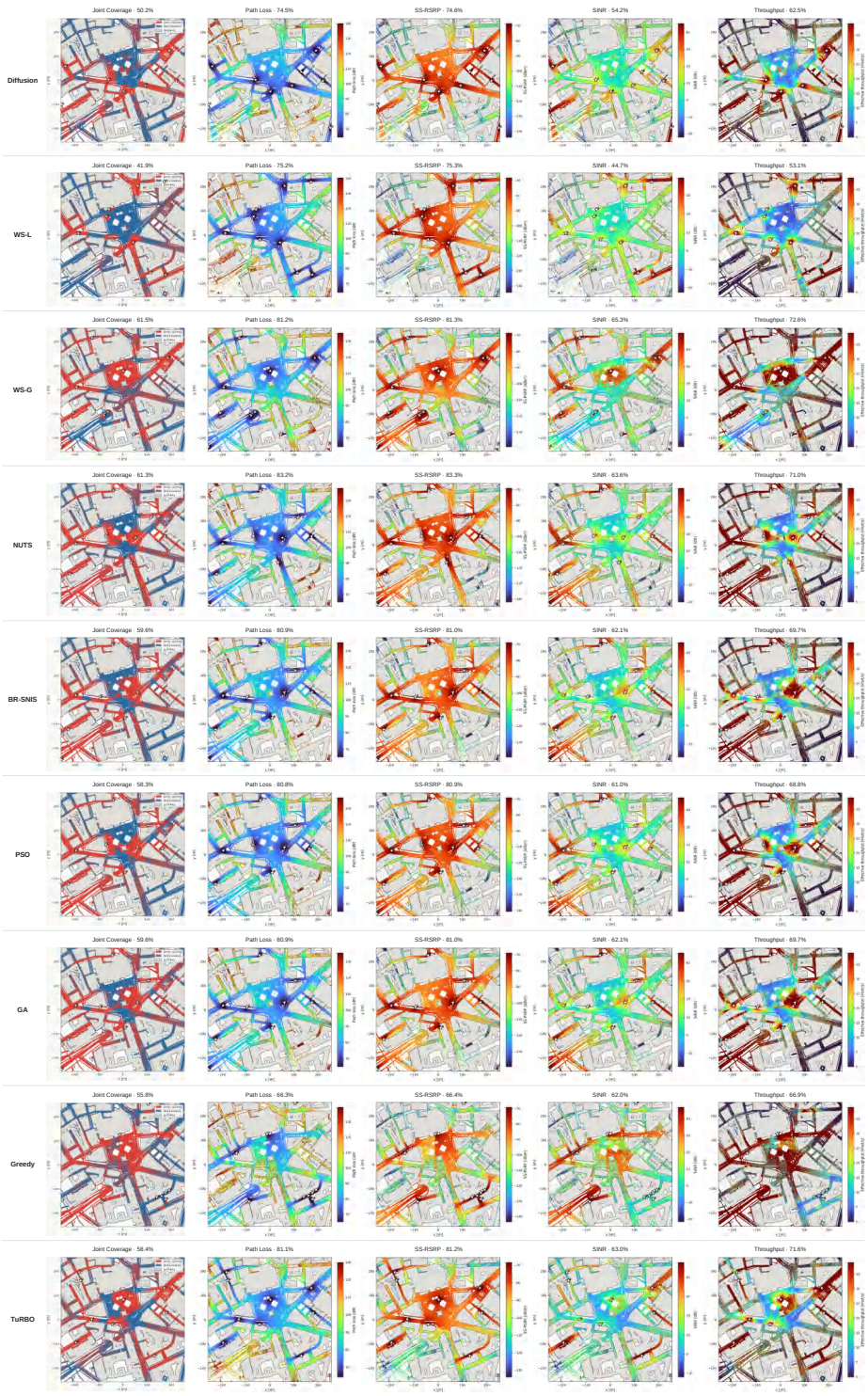


Figure 11: Deployment layouts and radio maps for nine HPEM-guided sampling methods in London. The scene is `london_01`, with $N_{\mathrm{TX}} = 9$ over a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. All rows use random seed 62. Results are individual Sionna-evaluated deployments.

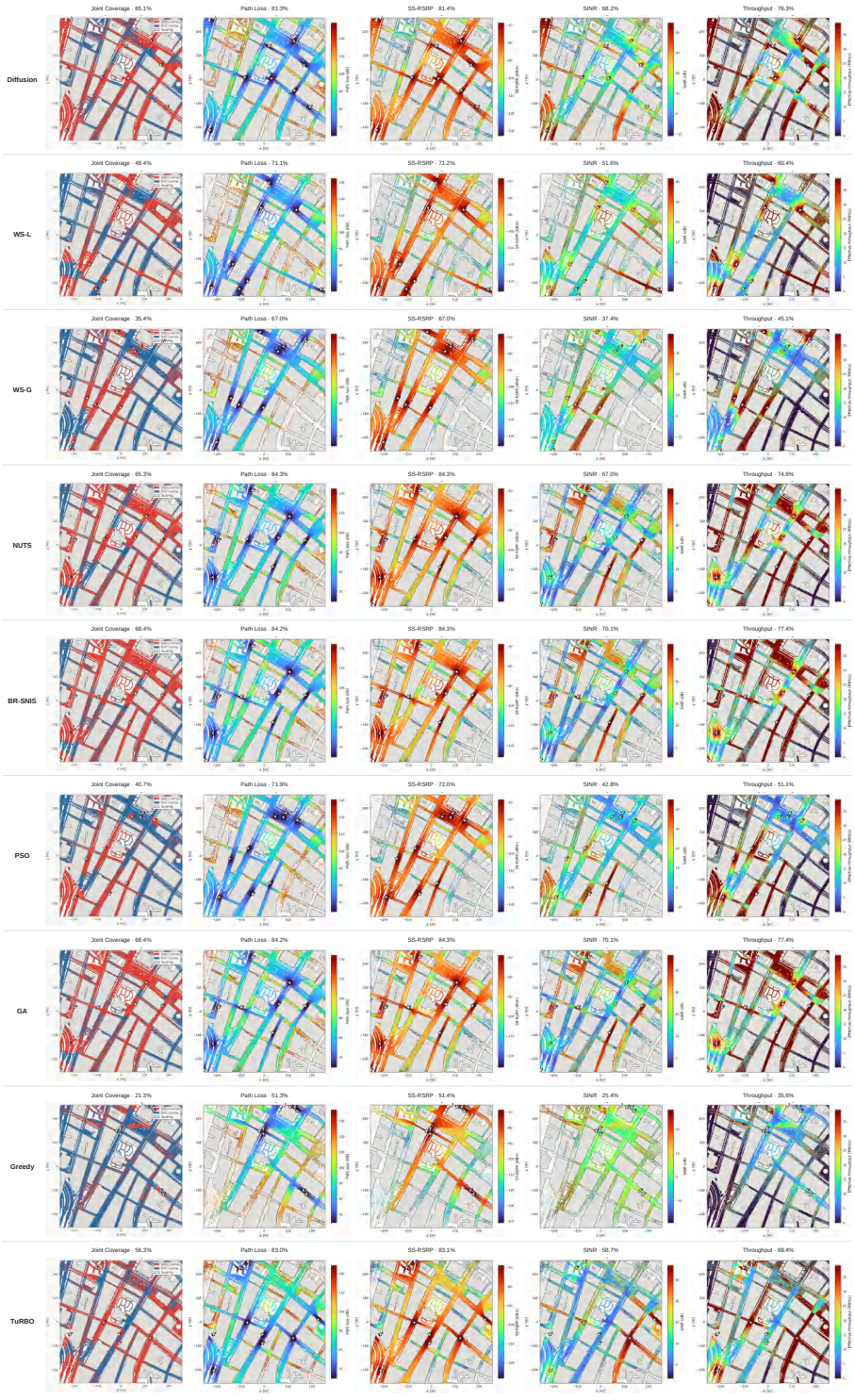


Figure 12: Deployment layouts and radio maps for nine HPEM-guided sampling methods in New York. The scene is `new_york_04`, with $N_{\mathrm{TX}} = 9$ over a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. All rows use random seed 62. Results are individual Sionna-evaluated deployments.

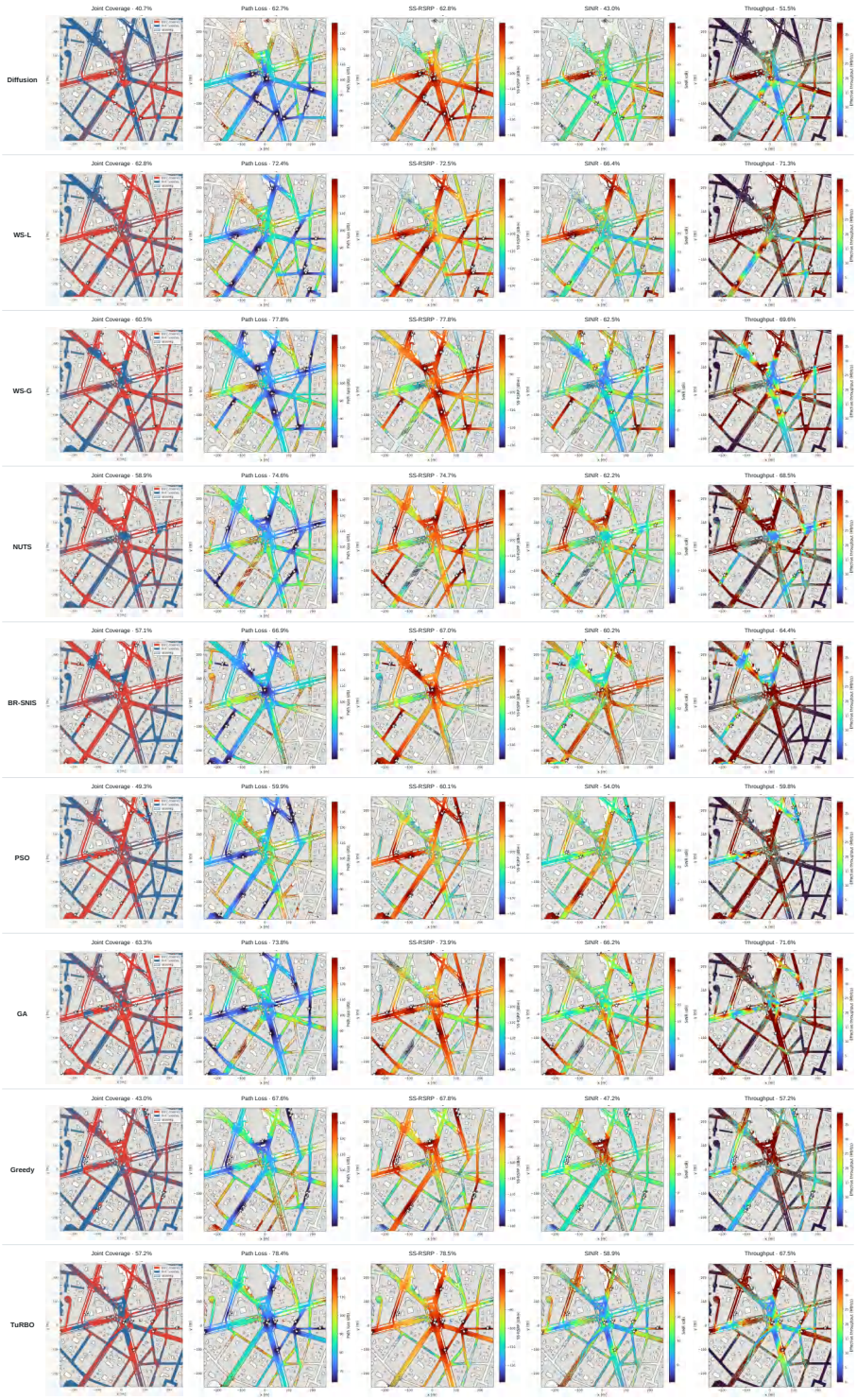


Figure 13: Deployment layouts and radio maps for nine HPEM-guided sampling methods in Paris. The scene is `paris_05`, with $N_{\mathrm{TX}} = 9$ over a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. All rows use random seed 62. Results are individual Sionna-evaluated deployments.

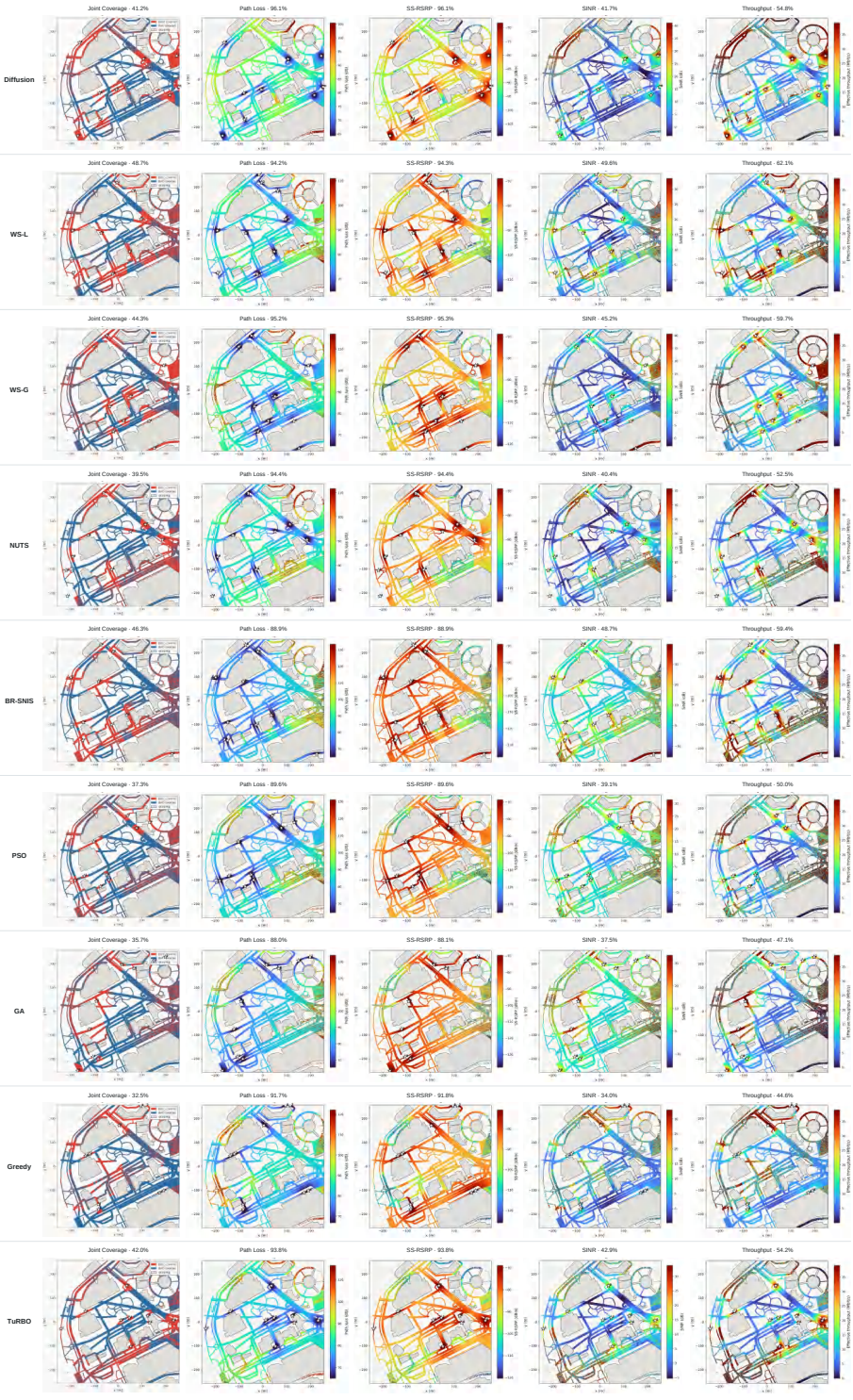


Figure 14: Deployment layouts and radio maps for nine HPEM-guided sampling methods in Shanghai. The scene is `shanghai_05`, with $N_{\mathrm{TX}} = 9$ over a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. All rows use random seed 62. Results are individual Sionna-evaluated deployments.

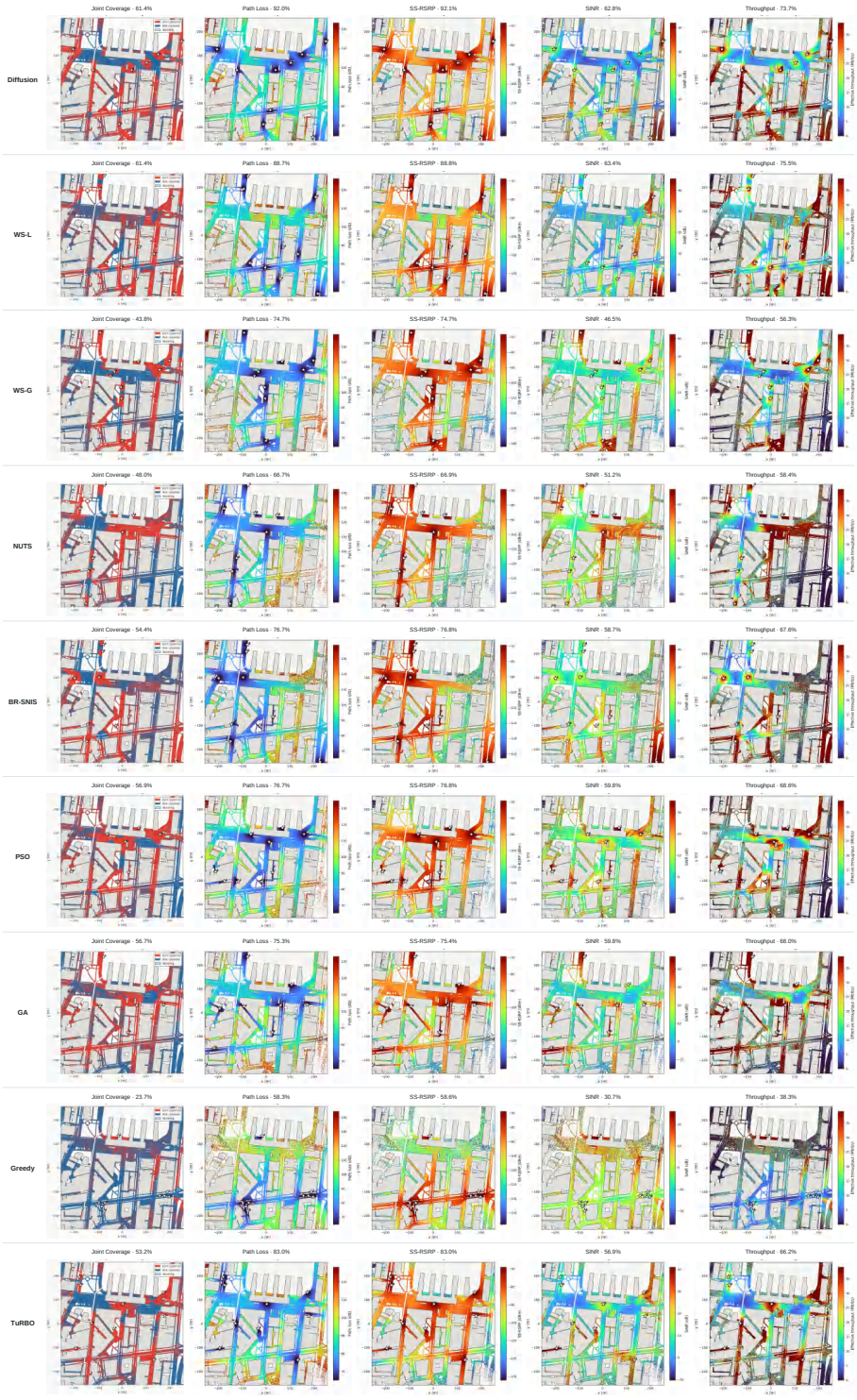


Figure 15: Deployment layouts and radio maps for nine HPEM-guided sampling methods in Sydney. The scene is `sydney_01`, with $N_{\mathrm{TX}} = 9$ over a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. All rows use random seed 62. Results are individual Sionna-evaluated deployments.

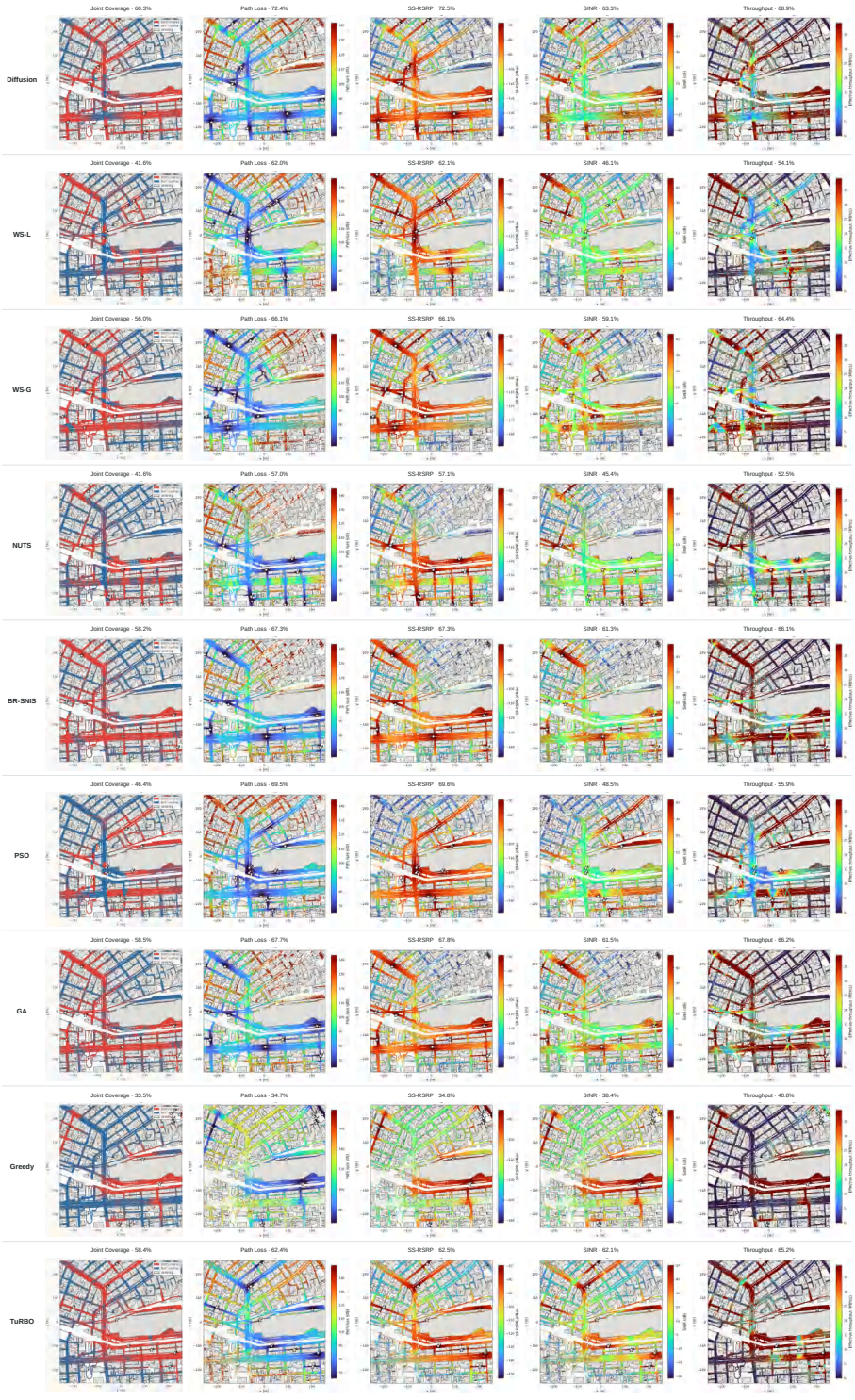


Figure 16: Deployment layouts and radio maps for nine HPEM-guided sampling methods in Tokyo. The scene is `tokyo_03`, with $N_{\mathrm{TX}} = 9$ over a $512\,\mathrm{m} \times 512\,\mathrm{m}$ region. All rows use random seed 42. Results are individual Sionna-evaluated deployments.

H SIMULATION RESULTS WITH FIXED TX NUMBER

Table 10: **Performance comparison across different TX deployment scales in the** $128\,\mathrm{m} \times 128\,\mathrm{m}$ **Sydney scene.** The simulation results are reported as mean±std. For each BS scale and evaluation metric, the *top three* methods are highlighted in (1st), (2nd), and (3rd), respectively, while the worst result in each column is underlined.

TXs	Method	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
1	Diffusion	4.82±0.21	91.34±0.84	94.62±0.72	96.48±0.58	93.11±0.91	95.08±0.69
	WS-G	5.46±0.28	90.42±0.96	93.74±0.87	95.48±0.72	92.18±1.00	94.16±0.84
	WS-L	6.91±0.37	90.08±1.01	93.41±0.91	95.16±0.78	91.87±1.04	93.89±0.88
	NUTS	8.14±0.52	90.91±0.93	94.20±0.81	96.11±0.66	92.72±0.97	94.73±0.78
	BR-SNIS	5.13±0.24	88.96±1.18	92.36±1.08	94.22±0.93	90.71±1.21	92.65±1.04
	Greedy	1.79±0.13	86.72±0.69	90.18±0.63	92.04±0.58	88.27±0.71	90.56±0.61
	PSO	13.72±0.88	89.84±1.03	93.05±0.94	94.91±0.82	91.43±1.07	93.46±0.90
	GA	15.18±1.03	87.63±1.16	91.08±1.05	92.91±0.91	89.21±1.19	91.44±1.01
	TuRBO	10.96±0.71	91.70±0.79	94.91±0.69	95.96±0.63	93.46±0.86	95.31±0.71
2	Diffusion	14.62±0.47	95.08±0.62	97.01±0.54	98.21±0.42	96.37±0.67	97.48±0.51
	WS-G	16.18±0.61	94.21±0.78	96.42±0.70	97.61±0.57	95.42±0.82	96.71±0.66
	WS-L	20.06±0.79	93.88±0.84	96.11±0.75	97.34±0.63	95.08±0.88	96.39±0.71
	NUTS	23.42±0.88	94.62±0.71	96.72±0.63	97.94±0.49	96.01±0.73	97.22±0.58
	BR-SNIS	15.37±0.54	92.94±0.91	95.19±0.82	96.56±0.69	94.04±0.95	95.51±0.79
	Greedy	5.61±0.29	90.84±0.66	93.21±0.61	94.86±0.54	91.93±0.69	93.67±0.59
	PSO	48.76±1.91	93.76±0.80	95.94±0.72	97.18±0.60	94.91±0.84	96.22±0.68
	GA	55.42±2.34	91.92±0.98	94.24±0.88	95.61±0.76	93.02±1.02	94.48±0.85
	TuRBO	34.91±1.21	94.91±0.66	97.16±0.57	97.82±0.52	95.82±0.76	97.61±0.55

Table 11: **Performance comparison across different TX deployment scales in the $256\,\mathrm{m} \times 256\,\mathrm{m}$ Paris scene.** The simulation results are reported as mean±std. For each BS scale and evaluation metric, the *top three* methods are highlighted in (1st), (2nd), and (3rd), respectively, while the worst result in each column is underlined.

TXs	Method	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
2	Diffusion	18.74±0.62	76.82±1.12	84.91±1.06	89.12±0.91	80.63±1.21	86.47±1.02
	WS-G	21.06±0.78	75.31±1.34	83.42±1.26	87.61±1.08	79.02±1.42	84.93±1.21
	WS-L	24.18±0.91	74.96±1.41	83.08±1.32	87.28±1.14	78.67±1.48	84.57±1.27
	NUTS	29.73±1.06	76.31±1.19	84.47±1.11	88.71±0.96	80.21±1.26	86.02±1.07
	BR-SNIS	20.11±0.71	73.18±1.47	81.27±1.39	85.42±1.21	76.89±1.53	82.74±1.34
	Greedy	6.87±0.34	68.42±0.93	76.21±0.88	80.53±0.79	71.94±1.01	77.82±0.86
	PSO	61.34±2.18	74.13±1.39	82.17±1.31	86.36±1.13	77.86±1.46	83.68±1.25
	GA	69.82±2.64	70.91±1.56	78.86±1.47	83.11±1.29	74.48±1.62	80.39±1.42
	TuRBO	44.36±1.52	77.03±1.07	85.18±1.01	88.46±0.98	80.94±1.18	86.73±0.99
3	Diffusion	28.61±0.88	83.06±0.94	89.48±0.89	92.63±0.76	86.39±1.03	90.74±0.86
	WS-G	32.14±1.07	81.56±1.16	88.02±1.08	91.17±0.92	84.86±1.22	89.27±1.04
	WS-L	36.72±1.24	81.21±1.21	87.68±1.13	90.84±0.97	84.51±1.28	88.91±1.09
	NUTS	44.53±1.41	82.47±1.01	89.03±0.95	92.21±0.81	85.98±1.08	90.28±0.91
	BR-SNIS	30.47±0.96	79.36±1.28	85.79±1.21	88.96±1.05	82.67±1.35	87.14±1.17
	Greedy	10.41±0.46	74.21±0.84	80.46±0.81	83.72±0.73	77.43±0.91	81.82±0.78
	PSO	91.58±2.91	80.31±1.21	86.72±1.14	89.91±0.98	83.61±1.28	88.08±1.09
	GA	103.46±3.42	76.63±1.41	82.94±1.34	86.17±1.16	79.84±1.48	84.27±1.28
	TuRBO	65.82±2.06	82.84±0.97	89.71±0.91	91.94±0.84	85.72±1.10	90.96±0.88
4	Diffusion	39.46±1.14	87.08±0.82	92.31±0.79	94.82±0.67	89.94±0.91	93.24±0.76
	WS-G	44.17±1.32	85.63±1.03	90.89±0.97	93.38±0.82	88.46±1.09	91.82±0.93
	WS-L	49.84±1.51	85.27±1.08	90.54±1.01	93.04±0.87	88.11±1.14	91.47±0.98
	NUTS	59.68±1.72	86.53±0.88	91.82±0.84	94.41±0.71	89.53±0.96	92.81±0.81
	BR-SNIS	41.92±1.24	83.47±1.16	88.67±1.10	91.24±0.94	86.29±1.22	89.64±1.06
	Greedy	14.63±0.58	78.62±0.76	83.71±0.73	86.34±0.65	81.47±0.83	84.82±0.71
	PSO	124.37±3.64	84.39±1.08	89.56±1.03	92.11±0.89	87.18±1.14	90.57±0.99
	GA	140.18±4.21	80.91±1.28	86.04±1.22	88.67±1.04	83.76±1.35	87.12±1.17
	TuRBO	88.75±2.47	86.92±0.85	92.57±0.81	94.13±0.74	89.31±0.98	93.47±0.78

Table 12: **Performance comparison across different TX deployment scales in the $512\,\mathrm{m} \times 512\,\mathrm{m}$ Tokyo scene.** The simulation results are reported as mean±std. For each BS scale and evaluation metric, the *top three* methods are highlighted in (1st), (2nd), and (3rd), respectively, while the worst result in each column is underlined.

TXs	Method	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
6	Diffusion	58.40±1.62	59.62±1.34	70.84±1.41	77.26±1.28	65.18±1.56	72.41±1.39
	WS-G	64.75±1.88	58.14±1.56	69.32±1.62	75.71±1.43	63.46±1.71	70.83±1.55
	WS-L	72.63±2.11	57.78±1.64	68.96±1.69	75.28±1.51	63.01±1.78	70.42±1.62
	NUTS	88.91±2.54	59.08±1.42	70.36±1.48	76.82±1.31	64.72±1.61	71.96±1.44
	BR-SNIS	62.10±1.77	55.96±1.72	67.28±1.78	73.64±1.60	61.11±1.86	68.57±1.69
	Greedy	19.84±0.83	51.62±1.21	62.48±1.30	68.77±1.18	56.72±1.34	63.54±1.25
	PSO	176.35±4.96	56.91±1.61	68.13±1.67	74.56±1.49	62.12±1.74	69.46±1.58
	GA	198.42±5.81	53.47±1.79	64.61±1.84	70.96±1.66	58.73±1.92	65.87±1.76
	TuRBO	131.28±3.72	60.03±1.29	71.19±1.36	76.41±1.34	65.63±1.50	72.86±1.33
7	Diffusion	71.86±1.91	66.08±1.20	76.72±1.29	82.84±1.14	71.53±1.42	78.34±1.27
	WS-G	79.42±2.15	64.47±1.43	75.08±1.51	81.06±1.31	69.77±1.57	76.62±1.45
	WS-L	88.36±2.47	64.05±1.51	74.67±1.58	80.65±1.38	69.34±1.63	76.18±1.51
	NUTS	107.64±2.93	65.51±1.29	76.21±1.36	82.36±1.19	71.08±1.47	77.91±1.32
	BR-SNIS	76.21±2.04	62.28±1.57	72.93±1.64	78.96±1.45	67.45±1.70	74.52±1.57
	Greedy	23.96±0.96	57.44±1.14	67.82±1.22	73.92±1.08	62.43±1.27	69.24±1.17
	PSO	211.67±5.62	63.18±1.48	73.72±1.56	79.72±1.37	68.21±1.61	75.31±1.49
	GA	239.84±6.38	59.72±1.66	70.19±1.72	76.26±1.52	64.67±1.79	71.55±1.64
	TuRBO	156.41±4.14	65.84±1.24	77.03±1.31	82.09±1.22	70.82±1.45	78.61±1.29
8	Diffusion	86.74±2.18	70.12±1.08	80.13±1.18	85.82±1.02	75.42±1.31	81.93±1.16
	WS-G	95.68±2.46	68.64±1.31	78.62±1.38	84.26±1.19	73.81±1.44	80.36±1.33
	WS-L	105.47±2.73	68.22±1.37	78.21±1.44	83.87±1.25	73.39±1.50	79.94±1.39
	NUTS	127.93±3.32	69.54±1.16	79.72±1.24	85.41±1.07	75.03±1.35	81.51±1.21
	BR-SNIS	91.66±2.31	66.53±1.42	76.61±1.49	82.21±1.30	71.58±1.55	78.31±1.44
	Greedy	28.31±1.08	61.37±1.06	71.12±1.14	76.84±1.01	66.29±1.19	72.97±1.10
	PSO	251.43±6.11	67.37±1.36	77.34±1.42	82.98±1.24	72.41±1.49	79.08±1.37
	GA	284.76±7.04	63.72±1.54	73.61±1.61	79.25±1.41	68.74±1.67	75.31±1.53
	TuRBO	184.52±4.63	69.91±1.12	80.42±1.20	85.07±1.11	74.81±1.37	82.17±1.18
9	Diffusion	102.18±2.47	73.04±0.98	82.56±1.09	87.71±0.94	78.31±1.19	84.22±1.07
	WS-G	112.53±2.78	71.62±1.21	81.08±1.28	86.17±1.10	76.72±1.32	82.71±1.24
	WS-L	123.42±3.06	71.21±1.27	80.67±1.34	85.79±1.16	76.31±1.38	82.29±1.29
	NUTS	149.88±3.74	72.52±1.07	82.13±1.15	87.32±0.99	77.92±1.23	83.84±1.11
	BR-SNIS	108.46±2.61	69.48±1.33	78.91±1.40	84.02±1.22	74.55±1.45	80.63±1.35
	Greedy	33.17±1.21	64.42±0.99	73.68±1.08	78.81±0.95	69.32±1.13	75.41±1.04
	PSO	294.64±6.83	70.37±1.26	79.81±1.33	84.91±1.15	75.43±1.37	81.51±1.27
	GA	332.18±7.92	66.71±1.45	76.03±1.52	81.12±1.33	71.83±1.58	77.69±1.45
	TuRBO	215.73±5.14	72.83±1.03	82.83±1.11	87.01±1.03	77.73±1.25	84.48±1.09

I SIMULATION RESULTS WITH TARGET JM COVERAGE

Table 13: **Performance comparison at a 85% target joint coverage in the $128\,\mathrm{m} \times 128\,\mathrm{m}$ Shanghai scene under the HPEM reward.** We report the required transmitter (TX) count, success rate (SR), best achieved joint coverage (Best JM), per-metric coverages (PL, SS-RSRP, SINR, and Eff. TP), and runtime (RT). Values are mean±std over 15 independent trials. Per-metric coverages are evaluated at the TX count corresponding to Best JM. For each column, the three best results are highlighted as 1st, 2nd, and 3rd, with ties sharing the same rank; the worst result is underlined.

Method	Req. TXs↓	SR (%)↑	Best JM (%)↑	PL (%)↑	SS-RSRP (%)↑	SINR (%)↑	Eff. TP (%)↑	RT (s)↓
Diffusion	2.20±0.41	100.00	90.38±2.24	95.21±2.71	94.08±2.94	92.44±2.18	93.02±2.42	57.82±18.94
WS-G	2.73±0.59	86.67	88.74±2.67	94.86±3.06	93.61±3.19	91.36±2.54	92.37±2.76	31.92±9.31
WS-L	2.87±0.64	80.00	88.21±2.82	94.31±3.17	94.53±3.23	90.78±2.68	93.36±2.81	37.44±11.14
NUTS	2.47±0.52	93.33	89.91±2.46	94.68±2.89	93.88±3.05	91.81±2.36	92.61±2.57	49.16±14.53
BR-SNIS	3.27±0.70	73.33	86.84±3.06	91.86±3.58	90.76±3.81	88.17±3.01	89.54±3.28	27.61±8.04
Greedy	4.47±1.02	53.33	83.52±3.57	88.32±4.17	91.22±4.05	84.81±3.55	86.18±3.79	14.74±4.81
PSO	3.07±0.80	80.00	87.96±2.91	93.13±3.35	95.18±3.12	89.72±2.84	93.91±3.04	81.24±28.41
GA	3.73±0.88	46.67	82.91±3.47	89.81±3.88	91.07±4.02	85.31±3.41	85.74±3.66	92.11±31.22
TuRBO	2.33±0.49	93.33	90.72±2.31	95.63±2.69	94.21±2.84	91.47±2.21	93.24±2.45	44.07±13.08

Table 14: **Performance comparison at a 75% target joint coverage in the $256\,\mathrm{m} \times 256\,\mathrm{m}$ New York scene under the HPEM reward.** We report the required transmitter (TX) count, success rate (SR), best achieved joint coverage (Best JM), per-metric coverages (PL, SS-RSRP, SINR, and Eff. TP), and runtime (RT). Values are mean±std over 15 independent trials. Per-metric coverages are evaluated at the TX count corresponding to Best JM. For each column, the three best results are highlighted as 1st, 2nd, and 3rd, with ties sharing the same rank; the worst result is underlined.

Method	Req. TXs↓	SR (%)↑	Best JM (%)↑	PL (%)↑	SS-RSRP (%)↑	SINR (%)↑	Eff. TP (%)↑	RT (s)↓
Diffusion	3.13±0.64	100.00	82.66±2.58	90.21±3.06	88.57±3.39	84.71±2.61	86.91±2.82	75.91±23.21
WS-G	3.47±0.67	86.67	82.14±2.83	89.46±3.28	88.96±3.47	84.36±2.92	84.62±3.06	41.57±11.54
WS-L	3.80±0.75	80.00	80.31±3.05	87.86±3.51	87.54±3.63	83.81±2.98	85.77±3.13	48.04±13.82
NUTS	3.07±0.59	93.33	81.62±2.79	89.73±3.19	88.16±3.48	85.08±2.71	85.41±2.98	63.88±18.31
BR-SNIS	4.33±0.90	66.67	77.96±3.35	85.34±3.88	84.26±4.08	79.91±3.28	81.61±3.55	36.04±10.19
Greedy	5.87±1.24	40.00	73.36±3.88	79.81±4.47	80.29±4.51	75.26±3.82	77.18±4.08	19.18±5.71
PSO	4.07±0.89	73.33	78.91±3.21	86.38±3.74	89.42±3.71	81.48±3.15	86.19±3.27	109.26±35.12
GA	5.00±1.07	33.33	72.74±3.68	81.66±4.15	79.92±4.35	76.81±3.61	76.56±3.85	123.31±39.18
TuRBO	3.27±0.59	93.33	81.87±2.69	89.51±3.11	88.41±3.32	84.17±2.64	85.68±2.88	57.92±17.16

Table 15: **Performance comparison at a 55% target joint coverage in the $512\,\mathrm{m} \times 512\,\mathrm{m}$ London scene under the HPEM reward.** We report the required transmitter (TX) count, success rate (SR), best achieved joint coverage (Best JM), per-metric coverages (PL, SS-RSRP, SINR, and Eff. TP), and runtime (RT). Values are mean±std over 15 independent trials. Per-metric coverages are evaluated at the TX count corresponding to Best JM. For each column, the three best results are highlighted as 1st, 2nd, and 3rd, with ties sharing the same rank; the worst result is underlined.

Method	Req. TXs↓	SR (%)↑	Best JM (%)↑	PL (%)↑	SS-RSRP (%)↑	SINR (%)↑	Eff. TP (%)↑	RT (s)↓
Diffusion	5.47±0.82	86.67	61.76±3.11	73.16±3.82	70.38±4.13	65.82±3.01	68.24±3.34	127.84±36.27
WS-G	5.93±0.87	80.00	60.36±3.36	73.82±4.04	71.14±4.21	64.61±3.28	66.51±3.58	67.04±17.38
WS-L	6.20±0.94	73.33	59.94±3.44	71.91±4.18	69.47±4.45	63.94±3.36	69.02±3.69	78.11±20.18
NUTS	5.33±0.72	93.33	62.41±3.19	72.91±3.93	70.06±4.24	65.21±3.14	66.92±3.48	109.12±29.41
BR-SNIS	6.87±1.06	60.00	57.31±3.81	68.26±4.55	66.38±4.67	60.44±3.74	63.01±4.05	57.86±15.48
Greedy	8.60±1.43	33.33	53.22±4.25	62.81±5.13	63.62±5.12	56.27±4.18	58.42±4.49	31.38±8.21
PSO	6.47±1.16	66.67	58.81±3.58	69.91±4.28	71.68±4.39	62.26±3.49	68.11±3.79	175.91±52.11
GA	7.53±1.31	26.67	52.57±4.11	64.82±4.83	62.91±5.04	57.29±4.03	57.94±4.34	198.16±59.27
TuRBO	5.07±0.70	86.67	62.08±3.27	72.98±3.86	70.81±4.05	64.88±3.08	67.82±3.39	95.84±26.35

J ENERGY MODEL PERFORMANCE WITH A FIXED NUMBER OF TXS

Table 16: **Cross-city planning performance under different adaptation regimes in Sydney scene with 4 TXs.** We compare zero-shot transfer, few-shot target-city adaptation, and in-distribution evaluation for different relational energy models. Results are reported as mean±std. Few-shot adaptation substantially narrows the cross-city generalization gap, while in-distribution evaluation provides the strongest overall performance. Higher-order relational modeling consistently improves joint and per-metric coverage at the cost of moderately increased inference time.

Method	Setting	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
Diffusion	Zero-shot	21.84±0.46	54.82±2.71	63.48±3.84	72.16±3.62	60.74±3.05	66.31±3.51
	Few-shot	22.67±0.71	58.06±2.18	66.82±3.13	75.41±2.94	64.11±2.56	69.72±2.88
	In-distribution	22.91±0.53	59.42±1.82	68.15±2.67	76.58±2.46	65.47±2.17	71.03±2.39
Pairwise	Zero-shot	24.16±0.51	57.31±2.32	65.92±3.31	74.86±3.08	63.41±2.73	68.87±3.02
	Few-shot	24.88±0.76	60.68±1.86	69.28±2.72	78.12±2.53	66.82±2.21	72.16±2.48
	In-distribution	25.14±0.57	61.96±1.54	70.46±2.28	79.26±2.08	68.13±1.88	73.42±2.05
4th-Order	Zero-shot	28.73±0.64	59.18±2.14	67.81±3.02	76.92±2.78	65.27±2.48	70.84±2.77
	Few-shot	29.61±0.88	62.76±1.69	71.36±2.43	80.31±2.20	68.94±1.98	74.38±2.21
	In-distribution	29.96±0.69	64.08±1.38	72.61±2.06	81.42±1.84	70.26±1.69	75.62±1.87

Table 17: **Cross-city planning performance under different adaptation regimes in Tokyo scene with 6 TXs.** We compare zero-shot transfer, few-shot target-city adaptation, and in-distribution evaluation for different relational energy models. Results are reported as mean±std. As scene complexity increases, absolute joint coverage decreases and the cross-city transfer gap becomes more pronounced. Few-shot adaptation recovers a substantial portion of this gap, while higher-order relational modeling provides consistent gains across adaptation regimes.

Method	Setting	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
Diffusion	Zero-shot	32.46±0.68	46.24±3.08	56.31±4.22	65.42±4.01	52.18±3.47	58.76±3.91
	Few-shot	33.52±0.94	49.76±2.51	59.87±3.48	68.91±3.27	55.72±2.91	62.28±3.21
	In-distribution	33.86±0.72	51.04±2.08	61.16±2.94	70.13±2.76	57.08±2.46	63.61±2.72
Pairwise	Zero-shot	35.74±0.73	49.38±2.67	59.42±3.76	68.71±3.51	55.26±3.08	61.84±3.42
	Few-shot	36.81±1.02	52.63±2.16	62.73±3.07	71.86±2.84	58.61±2.54	65.16±2.86
	In-distribution	37.18±0.78	53.84±1.79	63.96±2.58	72.98±2.37	59.87±2.13	66.42±2.39
4th-Order	Zero-shot	42.61±0.91	51.02±2.49	61.17±3.52	70.54±3.24	56.94±2.86	63.52±3.18
	Few-shot	43.82±1.16	54.71±1.98	64.86±2.84	74.12±2.58	60.73±2.31	67.21±2.63
	In-distribution	44.23±0.87	56.03±1.61	66.11±2.36	75.28±2.12	62.08±1.94	68.47±2.17

Table 18: **Cross-city planning performance under different adaptation regimes in Paris scene with 12 TXs.** We compare zero-shot transfer, few-shot target-city adaptation, and in-distribution evaluation for different relational energy models. Results are reported as mean±std. The high-complexity setting exhibits the largest transfer gap and the lowest absolute joint coverage. Few-shot adaptation nevertheless recovers most of the performance lost under zero-shot transfer, and higher-order relational modeling remains beneficial across all three adaptation regimes.

Method	Setting	RT (s)↓	JM (%)↑	PL (%)↑	TP (%)↑	SINR (%)↑	RSRP (%)↑
Diffusion	Zero-shot	64.82±1.06	34.76±3.31	43.28±4.48	55.62±4.37	41.54±3.72	47.31±4.16
	Few-shot	66.31±1.42	38.21±2.76	46.78±3.71	59.08±3.56	45.02±3.08	50.82±3.43
	In-distribution	66.76±1.15	39.47±2.28	48.06±3.12	60.31±2.98	46.34±2.58	52.11±2.87
Pairwise	Zero-shot	70.18±1.11	37.64±2.91	46.21±4.02	58.46±3.89	44.37±3.31	50.18±3.72
	Few-shot	71.82±1.53	41.03±2.38	49.66±3.31	61.84±3.17	47.86±2.74	53.64±3.03
	In-distribution	72.31±1.22	42.36±1.94	50.94±2.76	63.08±2.63	49.21±2.28	54.91±2.52
4th-Order	Zero-shot	80.73±1.43	39.42±2.68	48.03±3.74	60.31±3.57	46.16±3.04	52.06±3.43
	Few-shot	82.64±1.81	43.18±2.14	51.84±3.02	64.02±2.87	49.94±2.48	55.82±2.78
	In-distribution	83.17±1.47	44.72±1.73	53.26±2.48	65.34±2.32	51.41±2.03	57.18±2.26

K ENERGY MODEL PERFORMANCE WITH TARGET JM COVERAGE

Table 19: **Cross-city target-coverage planning under different adaptation regimes in New York scene.** We compare zero-shot transfer, few-shot target-city adaptation, and in-distribution evaluation for increasingly expressive relational energy models. Results are reported as mean±std over 12 independent trials. Higher target joint coverage requires more TXs and longer planning time. Few-shot adaptation substantially reduces the cross-city transfer gap, while in-distribution evaluation achieves the strongest overall performance.

Target JM (%)	Method	Setting	Req. TXs↓	SR (%)↑	Best JM (%)↑	RT (s)↓
45	Diffusion	Zero-shot	6.42±0.79	83.33	46.58±1.72	61.84±7.48
		Few-shot	5.58±0.67	91.67	48.31±1.36	63.12±7.86
		In-distribution	5.17±0.58	100.00	49.26±1.12	63.76±7.31
	Pairwise	Zero-shot	5.83±0.72	91.67	48.02±1.49	68.31±8.12
		Few-shot	5.00±0.60	100.00	49.86±1.21	69.74±8.46
		In-distribution	4.67±0.49	100.00	50.78±0.98	70.28±7.92
	4th-Order	Zero-shot	5.17±0.58	91.67	49.31±1.33	80.42±9.51
		Few-shot	4.42±0.51	100.00	51.24±1.06	82.05±9.86
		In-distribution	4.08±0.29	100.00	52.16±0.89	82.73±9.18
65	Diffusion	Zero-shot	9.08±0.90	58.33	63.72±2.31	84.61±10.76
		Few-shot	7.92±0.79	83.33	66.08±1.87	86.28±11.14
		In-distribution	7.33±0.65	91.67	67.21±1.52	87.06±10.42
	Pairwise	Zero-shot	8.33±0.78	66.67	65.12±2.04	92.47±11.35
		Few-shot	7.08±0.67	91.67	67.43±1.58	94.18±11.72
		In-distribution	6.58±0.51	100.00	68.61±1.27	94.93±10.96
	4th-Order	Zero-shot	7.42±0.67	83.33	67.08±1.71	108.36±13.42
		Few-shot	6.25±0.62	100.00	69.62±1.29	110.58±13.91
		In-distribution	5.83±0.39	100.00	70.74±1.03	111.42±12.84
85	Diffusion	Zero-shot	12.17±1.11	33.33	81.62±3.28	108.74±14.63
		Few-shot	10.75±0.97	58.33	84.18±2.36	111.06±15.14
		In-distribution	10.08±0.79	75.00	85.48±1.86	112.14±14.28
	Pairwise	Zero-shot	11.08±1.00	50.00	83.54±2.57	118.83±15.71
		Few-shot	9.67±0.89	75.00	85.72±1.82	121.27±16.22
		In-distribution	9.00±0.74	83.33	86.83±1.47	122.36±15.08
	4th-Order	Zero-shot	9.75±0.87	75.00	85.96±1.91	139.24±18.36
		Few-shot	8.42±0.67	91.67	88.14±1.43	142.17±18.92
		In-distribution	7.83±0.58	100.00	89.31±1.12	143.45±17.61

Table 20: **Cross-city target-coverage planning under different adaptation regimes in London scene.** We compare zero-shot transfer, few-shot target-city adaptation, and in-distribution evaluation for increasingly expressive relational energy models. Results are reported as mean±std over 12 independent trials. Compared with the easy scene, higher deployment complexity leads to lower success rates and a larger number of required TXs, particularly at aggressive coverage targets.

Target JM (%)	Method	Setting	Req. TXs↓	SR (%)↑	Best JM (%)↑	RT (s)↓
35	Diffusion	Zero-shot	7.17±0.72	75.00	35.92±1.69	72.16±8.84
		Few-shot	6.25±0.62	91.67	37.68±1.38	73.72±9.18
		In-distribution	5.83±0.58	100.00	38.62±1.15	74.35±8.57
	Pairwise	Zero-shot	6.50±0.67	83.33	37.21±1.51	79.64±9.36
		Few-shot	5.58±0.51	100.00	39.17±1.22	81.26±9.74
		In-distribution	5.17±0.39	100.00	40.08±1.01	81.91±9.02
	4th-Order	Zero-shot	5.83±0.58	91.67	38.84±1.38	93.58±11.42
		Few-shot	5.00±0.43	100.00	40.96±1.08	95.47±11.86
		In-distribution	4.67±0.49	100.00	41.93±0.91	96.15±10.93
55	Diffusion	Zero-shot	9.83±0.94	41.67	52.31±2.64	96.24±12.17
		Few-shot	8.50±0.80	66.67	55.47±2.02	98.36±12.72
		In-distribution	7.92±0.67	83.33	56.82±1.61	99.27±11.86
	Pairwise	Zero-shot	9.00±0.85	58.33	54.64±2.18	105.38±13.42
		Few-shot	7.75±0.75	83.33	57.18±1.66	107.76±13.91
		In-distribution	7.17±0.58	91.67	58.43±1.34	108.82±12.97
	4th-Order	Zero-shot	8.00±0.74	75.00	56.91±1.84	124.73±15.68
		Few-shot	6.92±0.67	91.67	59.62±1.39	127.56±16.11
		In-distribution	6.42±0.51	100.00	60.91±1.12	128.74±15.06
75	Diffusion	Zero-shot	13.58±1.16	16.67	71.48±3.04	122.61±16.08
		Few-shot	11.92±1.00	41.67	73.92±2.32	125.47±16.72
		In-distribution	11.08±0.90	58.33	75.31±1.82	126.61±15.54
	Pairwise	Zero-shot	12.42±1.08	25.00	73.64±2.43	134.26±17.42
		Few-shot	10.83±0.94	58.33	75.96±1.83	137.48±18.06
		In-distribution	10.08±0.79	66.67	77.18±1.47	138.75±16.84
	4th-Order	Zero-shot	10.92±0.90	58.33	77.82±1.68	157.94±20.31
		Few-shot	9.42±0.67	83.33	79.86±1.28	161.72±20.94
		In-distribution	8.75±0.62	91.67	81.06±1.04	163.18±19.47

Table 21: **Cross-city target-coverage planning under different adaptation regimes in Shanghai scene.** We compare zero-shot transfer, few-shot target-city adaptation, and in-distribution evaluation for increasingly expressive relational energy models. Results are reported as mean±std over 12 independent trials. The high-complexity scene exhibits the largest transfer gap and the strongest degradation in success rate at demanding target coverage levels. Few-shot adaptation recovers a substantial fraction of the lost performance, while higher-order relational models remain consistently advantageous.

Target JM (%)	Method	Setting	Req. TXs↓	SR (%)↑	Best JM (%)↑	RT (s)↓
25	Diffusion	Zero-shot	7.92±0.79	66.67	25.84±1.78	89.73±10.86
		Few-shot	6.92±0.67	83.33	27.62±1.43	91.64±11.24
		In-distribution	6.42±0.51	91.67	28.51±1.19	92.48±10.52
	Pairwise	Zero-shot	7.25±0.75	75.00	27.14±1.59	98.16±11.73
		Few-shot	6.25±0.62	91.67	29.04±1.23	100.27±12.14
		In-distribution	5.83±0.58	100.00	30.02±1.01	101.14±11.32
	4th-Order	Zero-shot	6.50±0.67	83.33	28.93±1.41	115.62±13.98
		Few-shot	5.58±0.51	100.00	31.05±1.09	118.16±14.46
		In-distribution	5.17±0.39	100.00	32.11±0.90	119.23±13.41
40	Diffusion	Zero-shot	10.58±0.90	25.00	37.12±2.51	112.48±14.13
		Few-shot	9.25±0.87	50.00	39.74±2.02	114.93±14.68
		In-distribution	8.67±0.78	66.67	41.03±1.66	115.96±13.72
	Pairwise	Zero-shot	9.67±0.89	41.67	39.16±2.16	122.36±15.21
		Few-shot	8.42±0.67	66.67	41.87±1.69	125.18±15.74
		In-distribution	7.83±0.58	83.33	43.12±1.37	126.31±14.63
	4th-Order	Zero-shot	8.58±0.79	58.33	41.86±1.76	143.85±17.62
		Few-shot	7.42±0.67	83.33	44.48±1.34	147.16±18.13
		In-distribution	6.83±0.58	91.67	45.72±1.09	148.52±16.88
55	Diffusion	Zero-shot	14.08±1.11	8.33	49.62±3.37	137.94±17.84
		Few-shot	12.42±1.08	25.00	52.13±2.72	141.06±18.46
		In-distribution	11.58±0.90	33.33	53.72±2.21	142.41±17.12
	Pairwise	Zero-shot	13.00±1.04	16.67	51.28±2.92	149.63±19.24
		Few-shot	11.42±0.90	41.67	54.42±2.13	153.08±19.86
		In-distribution	10.58±0.79	50.00	55.81±1.73	154.46±18.41
	4th-Order	Zero-shot	11.42±0.90	33.33	54.91±2.15	175.42±21.73
		Few-shot	9.92±0.79	58.33	57.43±1.61	179.61±22.46
		In-distribution	9.17±0.72	75.00	58.72±1.31	181.18±20.82

L Ablation Studies

L.1 Potential Model Capacity

We examine whether increasing the capacity of the learned potential systematically improves utility prediction. We vary the hidden and hypernetwork dimensions while keeping the remaining training configuration fixed.

Table 22: **Potential-model capacity ablation.** Coverage prediction errors under different hidden and hypernetwork dimensions. Values are reported as mean±standard deviation; lower prediction error is better.

Hidden	Hyper	Runtime (s)	Joint (%)	Pathloss (%)	Throughput (%)	SINR (%)	RSRP (%)
64	128	1.08 ± 0.13	20.58 ± 3.48	20.91 ± 3.32	34.83 ± 5.00	33.61 ± 5.20	29.70 ± 4.39
128	128	1.08 ± 0.14	19.75 ± 3.81	20.03 ± 3.79	33.36 ± 4.27	32.21 ± 4.27	28.59 ± 3.88
256	64	1.09 ± 0.11	$\mathbf{17.52 \pm 1.69}$	$\mathbf{18.58 \pm 2.28}$	$\mathbf{27.70 \pm 1.75}$	$\mathbf{25.88 \pm 1.84}$	$\mathbf{24.22 \pm 1.77}$
256	128	1.09 ± 0.10	18.55 ± 2.92	19.51 ± 2.55	29.76 ± 4.28	28.14 ± 5.02	25.80 ± 3.52
256	256	1.05 ± 0.09	19.05 ± 2.83	19.58 ± 2.26	31.07 ± 3.83	29.70 ± 4.60	26.48 ± 3.29

Increasing model capacity does not yield monotonic gains. The $256/64$ configuration produces the lowest error across all five service targets, while further enlarging the hypernetwork degrades prediction despite similar runtime. This suggests that utility quality depends on how capacity is allocated rather than parameter count alone.

L.2 Inference Budget and Interaction Order

We next vary relational order and inference budget while keeping the remaining planning procedure fixed.

Across the tested budget grid, higher interaction order produces a consistent advantage, and the gap becomes more pronounced as the inference budget grows. Increasing particles or sampling steps also improves coverage within each interaction order, indicating that representation capacity and search effort provide complementary gains in this setting. Importantly, the improvement from higher-order structure persists at matched inference budgets, so it cannot be attributed solely to stronger search.

M Aggregate Planning Diagnostics

M.1 Physical Scaling and Service Coupling

Figure 17 summarizes how the service metrics evolve with TX cardinality across representative deployment scales. Increasing the number of TXs generally improves serving-link coverage, as reflected by path-loss and SS-RSRP coverage. The behavior of SINR and JM coverage is more strongly scene dependent: in compact deployments, additional co-channel interference can offset the gain in serving-link coverage, whereas larger regions can remain predominantly coverage limited over the evaluated TX range. The resulting non-monotonicity motivates treating multi-TX deployment as a coupled set-planning problem rather than reasoning about TX locations independently.

M.2 Sensitivity to Utility Representation

Figure 18 isolates the effect of utility representation while holding the planner fixed. The preferred representation changes across deployment regimes: pairwise or latent-relational models can be competitive at low TX cardinality, while higher-order structure becomes increasingly useful in more strongly coupled settings. The variation across planner rows further shows that representation quality and search dynamics are not independent factors; the same reward replacement can help one planner while providing little gain to another.

Table 23: **Interaction order and inference-budget ablation.** Simulator-verified JM coverage (%) under different maximum interaction orders, particle populations, and sampling steps. All settings use the same planning protocol on the New York scene with $N = 3$ TXs.

Inference Budget		Maximum Interaction Order			
Particles	Steps	Unary	Pairwise	3rd-Order	4th-Order
10	20	38.90	40.15	41.47	**42.82**
	40	39.62	40.91	42.28	**43.67**
	60	40.18	41.52	42.94	**44.40**
	80	40.52	41.89	43.35	**44.84**
	100	40.68	42.08	43.55	**45.06**
15	20	39.58	40.91	42.31	**43.74**
	40	40.31	41.69	43.14	**44.63**
	60	40.88	42.31	43.82	**45.36**
	80	41.23	42.69	44.23	**45.81**
	100	41.40	42.88	44.44	**46.04**
20	20	40.19	41.58	43.04	**44.55**
	40	40.93	42.37	43.89	**45.45**
	60	41.51	43.01	44.58	**46.19**
	80	41.87	43.40	45.01	**46.65**
	100	42.05	43.60	45.22	**46.88**
25	20	40.72	42.16	43.69	**45.27**
	40	41.47	42.96	44.55	**46.18**
	60	42.06	43.61	45.25	**46.93**
	80	42.43	44.01	45.69	**47.39**
	100	42.61	44.21	45.90	**47.63**

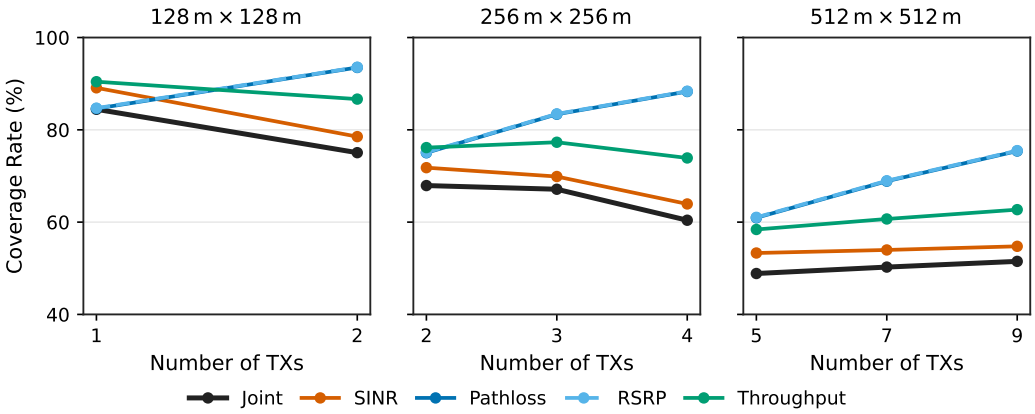


Figure 17: **Service trade-offs under increasing TX cardinality.** Each point aggregates six cities, nine planners, four utility representations, and five random seeds. Additional TXs improve serving-link coverage, while the resulting SINR and JM behavior depends on the balance between coverage gain and co-channel interference at each deployment scale.

M.3 Quality, Efficiency, and Stability

Planning quality alone does not determine the usefulness of a search method. Figure 19 compares simulator-verified JM coverage against planning time and shows that the empirical Pareto frontier changes with deployment scale. Fast local procedures can remain competitive in simpler settings, while more structured stochastic exploration becomes increasingly valuable as the deployment problem becomes more strongly coupled.

Figure 4 provides a complementary view of repeatability. Stochastic search does not necessarily imply unstable outcomes: several sampling methods combine high average JM coverage with relatively

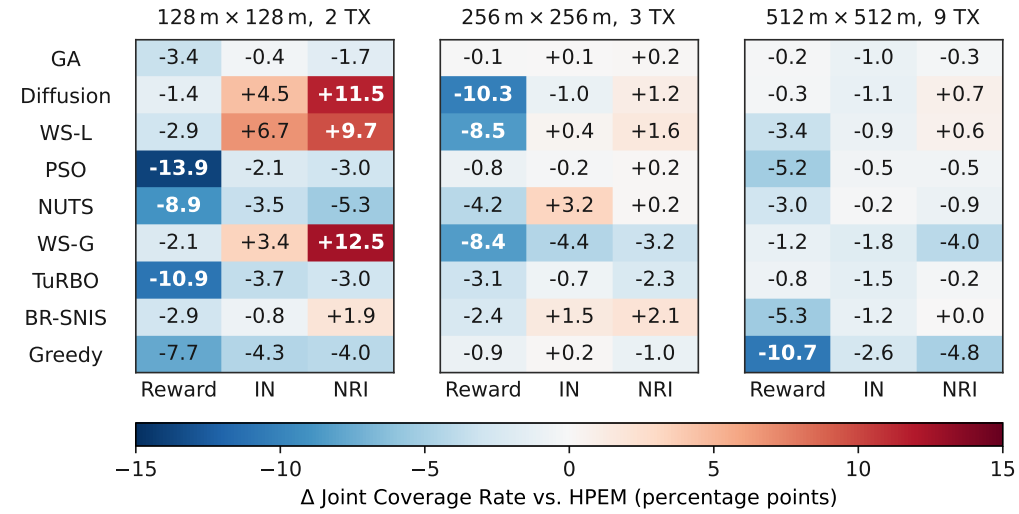


Figure 18: **Sensitivity of planning quality to utility representation.** Each cell reports the mean change in simulator-verified JM coverage (percentage points) when HPEM is replaced by the pooled Regressor, IN, or NRI under the same planner. Results are averaged over five random seeds. Positive values favor the alternative representation; negative values favor HPEM.

small variation across random seeds. Together, these results motivate evaluating planning methods jointly in terms of quality, cost, and stability rather than using a single performance axis.

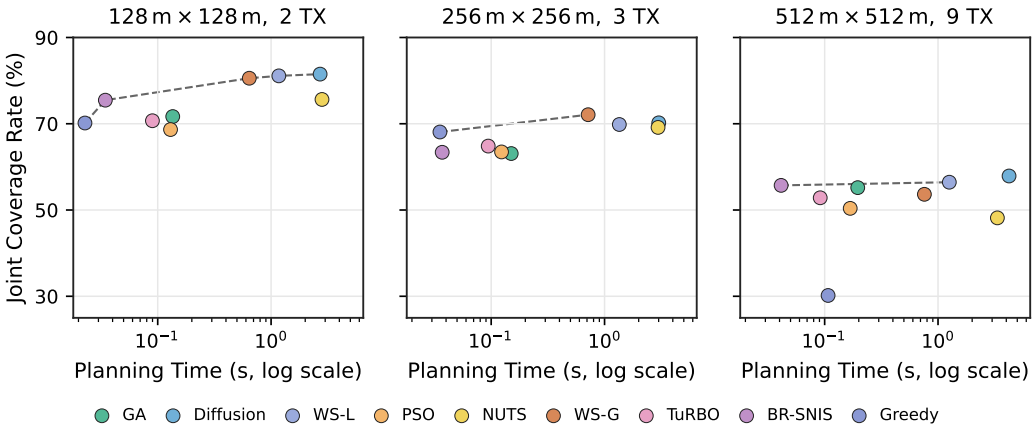


Figure 19: **Quality–efficiency trade-off across planning methods.** The horizontal axis reports median planning time on a logarithmic scale and the vertical axis reports mean simulator-verified JM coverage. Results aggregate six cities, four utility representations, and five random seeds for representative deployment regimes. Dashed curves connect the empirical Pareto-optimal methods.