NETAGENT: AGENTIC WIRELESS BASE-STATION PLANNING WITH DIFFERENTIABLE PROPAGATION OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Reliable wireless service and quiet-zone protection depend on base-station placement. This is a constrained inverse-design problem in which sparse transmitter decisions are judged by the dense propagation fields they induce. Locations, power allocation, and station count interact: strengthening one region can illuminate another that should remain quiet. Electromagnetic simulation captures these effects but makes configuration search costly, while fast channel models miss local building geometry. A radio-map generator offers another route when transmitter locations are not prescribed: gradients of regional objectives shape a candidate field whose dominant peak identifies a deployable site. NetAgent trains a RadioDiff-based architecture without its original condition encoder and freezes the resulting radio-map prior. Short deterministic denoising allows joint optimization of spatial latents and source power weights; station coordinates are read from the optimized single-source fields. One shared multi-source optimization produces weight-ranked nested deployments, from which a constraint-first rule selects station count without separate searches. An agentic interface converts drawn or natural-language requests into validated regional constraints. On thirty OpenStreetMap-derived scenarios scored by an independent electromagnetic simulator, NetAgent Auto-K reaches 0.5885 mean target attainment in 8.21 s with 1.27 stations on average over the three frozen seeds, compared with 0.4661 for AutoBS and 0.6431 for a 101.82 s exhaustive ray-tracing reference.

1 INTRODUCTION

Reliable wireless access depends on deployments that serve designated regions while keeping received power low where signals should remain quiet. As networks become denser, station coordinates, power allocation, and station count affect the same propagation field: improving one region can degrade another, as Figure 1 illustrates. This is also a constrained inverse-design problem in machine learning. A small set of physical decisions is evaluated through a dense, geometry-dependent output field, so the planning representation has to connect field-level objectives back to deployable sites.

Obtaining that connection from conventional propagation models is difficult. Buildings reflect, diffract, and scatter radio waves, creating local patterns that simple path-loss models cannot capture (3GPP, 2020). Site-specific electromagnetic simulation resolves these patterns but evaluates fixed placements, and the number of multi-station configurations grows combinatorially (Sato & Suto, 2024). Learned radio-map models reduce the cost of forward prediction, yet models conditioned on an already chosen transmitter location still start from a site and predict its field (Levie et al., 2021; Wang et al., 2025). A generator trained without prescribed transmitter locations offers a different representation: its dominant field peak identifies a candidate source, while the rest of the field describes how that source illuminates and leaks across the map. Regional objectives can shape this field through continuous generator inputs, with the discrete site read out after optimization.

NetAgent instantiates this field-space inverse design with a RadioDiff-based architecture (Wang et al., 2025) retrained without its original condition encoder and then frozen. A short deterministic denoising trajectory (Song et al., 2021) renders one field per source; gradients of service, suppression, margin, and whole-map objectives update the spatial latents and source power weights. Each source's deployable coordinate is extracted from the peak of its optimized field inside its allowed region. A shared optimization over $K_{\max}$ sources also produces a weight ordering and nested Top-K

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.

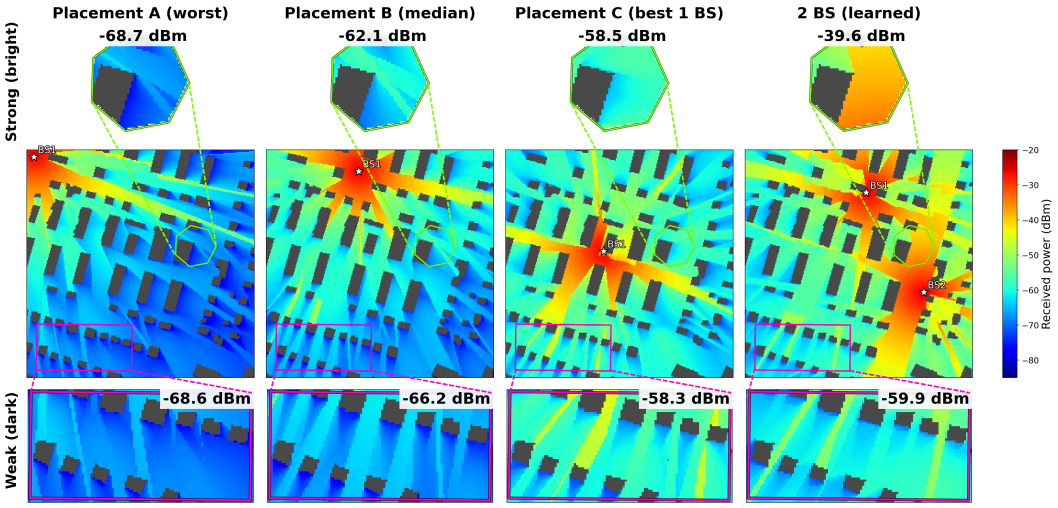


Figure 1: Electromagnetic simulations on one map show the core difficulty. The first three columns place one station at three positions, each showing the strong region above the full received-power map and the quiet region below, and every position that strengthens the strong region also lifts the quiet one, so a single station cannot serve and stay quiet at once. The fourth column deploys two stations with optimized positions and weights, which brightens the strong region while the quiet region stays dark, so joint multi-source planning resolves the conflict.

deployments, allowing a constraint-first rule to select a count from the same run. NetAgent uses no electromagnetic-simulator feedback during this search. An agentic front-end translates drawings or natural-language requests into the same validated regional task; the language model does not predict radio maps or decide the deployment.

We evaluate whether decisions produced in the learned field space transfer to an independent propagation scorer, with methods sharing the same scenarios and radio protocol. This tests the value of a generative model as a planning representation rather than only a forward predictor. The main contributions of this paper are summarized as follows.

- **Field-space inverse planning.** We optimize regional objectives through a frozen diffusion radio-map prior whose transmitter locations are not prescribed, then decode deployable sites from the optimized single-source fields without enumerating coordinate configurations during NetAgent's search.
- **Shared station-count planning.** One multi-source optimization jointly shapes source fields and power weights, yielding weight-ranked nested deployments. A constraint-first rule selects the smallest surrogate-feasible prefix, or the least-violating prefix when none is feasible, without an independent optimization for each count.
- **Independent deployment evaluation.** On thirty OpenStreetMap-derived scenarios scored by WinProp, NetAgent Auto-K achieves 0.5885 mean target attainment in 8.21 s with 1.27 stations on average over the three frozen seeds, compared with 0.4661 in 3.55 s for AutoBS and 0.6431 in 101.82 s for the exhaustive ray-tracing reference.

2 Related Work

Base-station planning. Recent planners reduce expensive simulator calls. Sato and Suto jointly optimize placement and transmit power with Bayesian optimization (Sato & Suto, 2024). Related lines cover neural placement surrogates benchmarked on 167k city-map scenarios (Çağkan Yapar, 2026), resilient mmWave IAB planning (Zhang et al., 2025), reinforcement-learning deployment on digital twins (Lee & Molisch, 2025), and LLM-in-the-loop site selection (Deng et al., 2025; Quan et al., 2025).

Learned propagation models. RadioUNet and PMNet predict radio maps from environmental and transmitter inputs, both built on the U-Net architecture (Levie et al., 2021; Lee & Molisch, 2024; Ronneberger et al., 2015). RadioDiff uses conditional diffusion for RM construction (Wang

et al., 2025), and follow-up work extends it in several directions (Jia et al., 2025; Fang et al., 2025; O'Shea & Brennan, 2026; Liao et al., 2026; Quang et al., 2026; Saeizadeh et al., 2026; Luo et al., 2025; Sun et al., 2025; Wang et al., 2026). Differentiable ray tracing pushes the same idea into the simulator itself (Hoydis et al., 2023; Aoudia et al., 2026; Pegurri et al., 2025).

Frozen diffusion priors for inverse problems. Optimizing the latent input of a frozen diffusion model is itself an established inverse-problem technique rather than a contribution of this work. The diffusion image prior shows that early stopping inside a pretrained model's reconstruction process supplies a strong prior for image restoration (Chihaoui & Favaro, 2025), and diffusion generative inverse design samples designs from a denoising model guided by a simulator-derived cost (Vlastelica et al., 2023). This work therefore claims no novelty for latent optimization as such and instead claims the domain-specific content built on it, namely regional-constraint planning on a generated field and the nested Top-K construction under a fixed power budget.

Language-model agents. Wang et al. study prompt- and agent-based base-station siting (Wang et al., 2024). NetAgent instead gives the agent a narrower role, in which it validates the spatial task while the differentiable planner selects the deployment.

3 Problem Formulation

Let $B \in [0,1]^{H\times W}$ be an OSM-derived binary building map and $\Omega = \{u : B(u) < 0.5\}$ its free-space pixels.[1] A task contains K source-specific deployment regions $\mathcal{D}_k$, strong regions $\{\mathcal{S}_i\}$, and weak regions $\{\mathcal{W}_j\}$, given as exact circle, ellipse, rotated-rectangle, and polygon masks. We seek coordinates $x_k \in \mathcal{D}_k \cap \Omega$ and normalized weights $w_k \geq 0$ with $\sum_k w_k = 1$.

Let $P_k(u; x_k)$ be the single-source received power predicted by the frozen propagation evaluator. Multi-source power is fused in linear units as

$$P(u) = 10\log_{10}\left(\sum_{k=1}^{K} K w_k 10^{P_k(u;x_k)/10}\right). \tag{1}$$

The factor K keeps the total budget equal to K nominal base stations, so uniform weights give one unit to every source. The total transmit budget therefore scales with the station count, since each count is allocated its nominal per-station power, so a fixed budget is fixed within a count rather than across counts, and the station-count results are read under that setting.

Every regional threshold τ_i^s, τ_j^w, and τ^g, every reported per-region value, and the Strong–Weak margin M below are RSS statistics on this dBm scale.

Each strong region $\mathcal{S}_i$ and weak region $\mathcal{W}_j$ carries its own threshold τ_i^s, τ_j^w and target ratio ρ_i^s, ρ_j^w. The common metrics are

$$C_i^s = |\mathcal{S}_i|^{-1}\sum_{u\in\mathcal{S}_i}\mathbf{1}[P(u)\geq\tau_i^s], \qquad S_j^w = |\mathcal{W}_j|^{-1}\sum_{u\in\mathcal{W}_j}\mathbf{1}[P(u)\leq\tau_j^w],$$
$$C_g = |\Omega|^{-1}\sum_{u\in\Omega}\mathbf{1}[P(u)\geq\tau^g], \qquad M = \min_i \overline{P}_i^s - \max_j \overline{P}_j^w, \tag{2}$$

where $\overline{P}_i^s$ and $\overline{P}_j^w$ are mean dBm values over the corresponding free-space regions, and $C_s = |\{i\}|^{-1}\sum_i C_i^s$ and $S_w = |\{j\}|^{-1}\sum_j S_j^w$ are the aggregate Strong-coverage and Weak-suppression fractions. A scenario is feasible only when every $C_i^s \geq \rho_i^s$, every $S_j^w \geq \rho_j^w$, and $M \geq M_{\min}$. A requirement is met in full when the measured value reaches the required level, and some combinations of threshold and required fraction are not attainable under the fixed power budget, so such a target is approached as closely as the budget allows rather than declared a failure. Section 5.1 defines the protocol score, which ranks fixed-K placement quality, and the attainment, which measures how fully a deployment reaches the goals of its own map; runtime is reported separately and affects neither. WinProp is the source of the final metrics and feasibility decision (Altair Engineering, 2024).

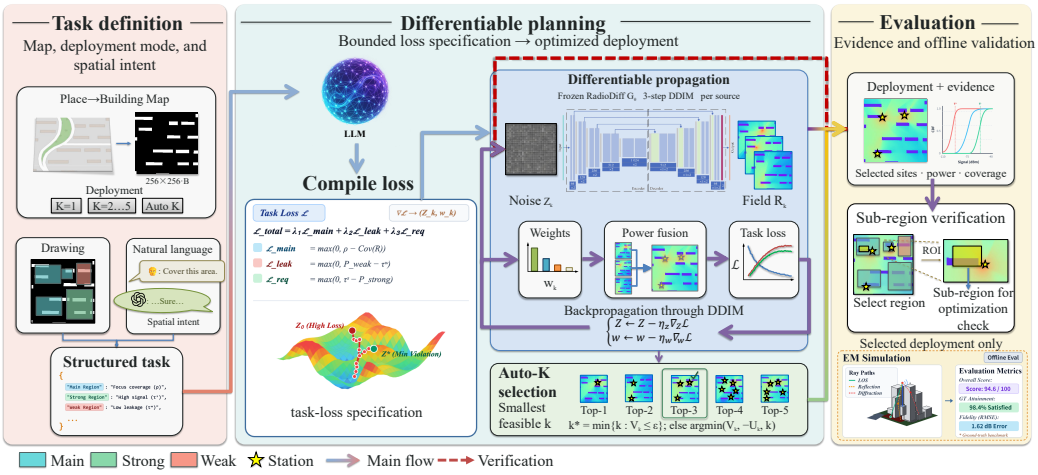


Figure 2: The NetAgent pipeline from map and region inputs to the simulator-scored deployment.

4 NetAgent

4.1 Constraint Specification by Drawing or Natural Language

Both input modes yield the identical structured task of one Main region per source plus Strong and Weak regions with per-region thresholds. Drawing mode sketches these regions and validates geometry, map bounds, and free space, while natural-language mode parses a prompt such as BS1 in the top-left with strong coverage on the left and weak coverage in the top-right into the same constraint set; the LLM never executes code or touches the optimizer.

4.2 From Constraints to the Loss Objective

Binary coverage tests are replaced by sigmoid approximations $\widetilde{C}_i^s$ and $\widetilde{S}_j^w$ with violations $v_i^s = [\rho_s - \widetilde{C}_i^s]_+$ and $v_j^w = [\rho_w - \widetilde{S}_j^w]_+$; and source-integrity constraints keep each source's dominant peak inside its own deployment region. With v_r and q_r a violation and its priority, the constraint loss is the augmented Lagrangian (Deng et al., 2026)

$$\mathcal{L}_{con} = \frac{\sum_r q_r(\lambda_r v_r + \frac{\mu}{2} v_r^2)}{\sum_r q_r} + \gamma(\max_r v_r)^2. \tag{3}$$

The best iteration is selected by source integrity, then task violation, then utility.

4.3 Differentiable Propagation Optimization

Let G_θ be the frozen RadioDiff prior (Wang et al., 2025), a denoising diffusion model (Ho et al., 2020). For each source, NetAgent initializes a spatial diffusion input Z_k and generates

$$R_k = G_\theta(Z_k; B). \tag{4}$$

The coordinate is then

$$x_k^* = \arg\max_{u \in \mathcal{D}_k \cap \Omega} R_k(u). \tag{5}$$

For $K > 1$, source logits a_k produce

$$w_k = w_{\min} + (1 - K w_{\min}) \frac{e^{a_k}}{\sum_\ell e^{a_\ell}}, \qquad w_{\min} = 0.05, \tag{6}$$

and the inputs and logits are optimized jointly.

The planning medium is the field that the frozen prior renders rather than a learned scorer, and it is this field that the task gradients steer. The prior concentrates energy near a source, carries it along the open street, and lets it decay with occlusion rather than with distance alone, so the regional objectives are differentiable in the field and their gradients move every source location and power weight. The brightest region of that field is then the extracted source of Equation 5, so the peak is

[1]Map data from OpenStreetMap contributors under ODbL at `https://www.openstreetmap.org/copyright`.

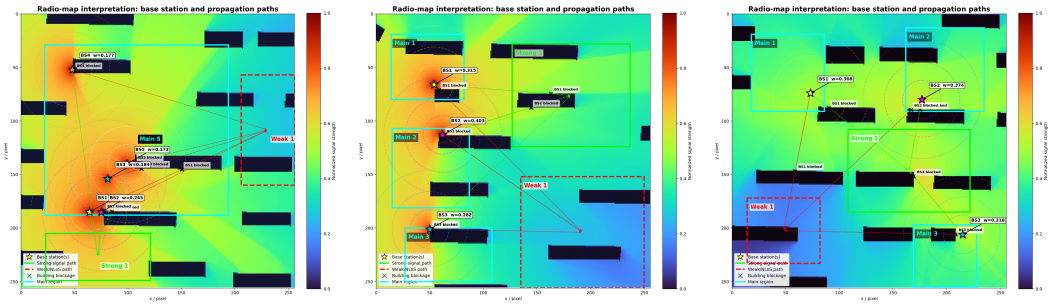


Figure 3: Fields rendered by NetAgent itself, one per scenario, where the brightest region is the extracted base-station position. The normalized signal strength concentrates near that source and along the open street and is attenuated where a facade blocks the direct path.

read out of the same field that the gradients shaped. Figure 3 shows three rendered fields with the regional masks and the extracted sources annotated.

5 Experimental Setup

All methods run on thirty variable-K scenarios from one generation pipeline. Ten are large OSM-derived maps with per-map thresholds, target ratios, and margins frozen in an earlier study, and twenty are maps from the same task library, stratified by building density, which share one uniform configuration with $\tau^s = -50\,\text{dBm}$ at 90%, $\tau^w = -60\,\text{dBm}$ at 90%, and a $15\,\text{dB}$ margin. All thresholds, region geometry, and seeds were frozen before any method ran, and the seeds come from operating-system entropy. NetAgent's frozen RM diffusion prior was trained on RadioMapSeer building maps, while the evaluation maps come from OpenStreetMap, so no map file is shared between the two sets, and both use the same 256×256 raster at $1\,\text{m}$ per pixel under the fixed radio protocol below. The thirty evaluation scenarios were drawn and frozen independently of the training pipeline.

5.1 Evaluation Metrics of Score and Attainment

Strong coverage S is the mean fraction of free-space pixels in Strong regions meeting the map's Strong threshold τ^s between -40 and $-55\,\text{dBm}$, and weak suppression W is the analogous fraction of Weak-region pixels at or below τ^w between -60 and $-70\,\text{dBm}$. The margin M is the Strong–Weak gap in dB over region means, and whole-map coverage C is the fraction of free-space pixels meeting τ^s.

Score. The protocol score is a fixed weighting of the four raw quantities,

$$\text{score} = 0.30\,\frac{S}{100} + 0.30\,\frac{W}{100} + 0.30\,m(M) + 0.10\,\frac{C}{100}, \qquad m(M) = \text{clip}\Big(\tfrac{M}{M^\star},\,-1,\,1\Big), \quad (7)$$

Here M is the measured Strong–Weak margin in dB, and $M^\star$ is the required minimum margin of that scenario, which is read from the scenario configuration and ranges from 10 to $25\,\text{dB}$ across the thirty scenarios, so the margin term is a single bounded ratio of the measured margin to the required one and follows the scenario's own scale rather than one global constant. The tabulated score is consequently not recoverable from the tabulated component means. The lead over the strongest deployable baseline averages 10.7% and ranges from 4.9% to 16.1% across the five counts.

Attainment. The attainment measures goal completion against each map's own targets,

$$\text{attainment} = \frac{1}{4}\left[\text{clip}_{[0,1]}\Big(\frac{S}{\rho^s}\Big) + \text{clip}_{[0,1]}\Big(\frac{W}{\rho^w}\Big) + \text{clip}_{[0,1]}\Big(\frac{M}{M_{\min}}\Big) + \text{clip}_{[0,1]}\Big(\frac{C}{90}\Big)\right], \quad (8)$$

where $\rho^s, \rho^w \in [80\%, 95\%]$ and $M_{\min} \in [10, 25]\,\text{dB}$ are per-map targets and the coverage target is fixed at 90%. Every term is capped at 1, so over-achieving one target cannot compensate a deficit on another, and it is the variable-K selection criterion. Since ρ^s, ρ^w and $M_{\min}$ are per-map targets that are not tabulated, the tabulated attainment is likewise not recoverable from the tabulated component values. The reported metrics therefore measure how closely each deployment approaches the goals of its map, and the regional shortfall of Appendix C quantifies the remaining distance.

5.2 Methods and Search Protocols

NetAgent Fixed-K. NetAgent runs an independent 10-epoch optimization for each $K \in \{1, \ldots, 5\}$ with learned per-source softmax weights, which is the prescribed fixed-K protocol rather than a post-processing of $K = 5$.

NetAgent Auto-K. One shared $K_{\max} = 5$, 10-epoch optimization of $\{(Z_k, w_k)\}_{k=1}^{K_{\max}}$ with $\sum_k w_k = 1$ covers all counts. Sorting sources by weight, $\pi = \mathrm{argsort}(w)$ descending, yields the five nested prefixes

$$\mathcal{C}_k = \{\pi_1, \ldots, \pi_k\}, \qquad \hat{w}_j = w_{\pi_j} \Big/ \textstyle\sum_{i \le k} w_{\pi_i}, \qquad k = 1, \ldots, K_{\max}, \tag{9}$$

fused via Eq. equation 1 and scored by the differentiable task loss into utility–violation pairs (U_k, V_k). Constraint-first K selection scans these five prefixes with a feasibility tolerance τ and takes

$$k^\star = \min\big\{k : V_k \le \tau\big\} \quad \text{if that set is nonempty, otherwise} \quad k^\star = \arg\min_k \big(V_k,\, -U_k,\, k\big), \tag{10}$$

with $\tau = 10^{-6}$ on the normalized violation. The rule inspects only these five nested prefixes rather than every count and configuration, it ranks a prefix by its constraint violation before its utility, and it therefore reports the smallest feasible prefix when one exists and a minimum-violation fallback otherwise. Auto-K is run with three frozen entropy seeds over the thirty scenarios, the reported deployment is the first of them, and the seeds are drawn from operating-system entropy and frozen before any run.

Sionna Exhaustive. Sionna ray tracing runs on the same 32×32 candidate grid on a GPU, where each source draws its candidate positions from the free grid cells inside its own deployment region, one ray-traced map is computed per candidate, and the five counts are searched one after another. For each count the multi-source deployment is chosen by enumerating the full product of the per-source candidate sets and scoring every configuration exactly on the GPU in chunks, and the five counts together cost 101.82 s. Because it scores every configuration of that product exactly, it attains the best mean that this candidate space allows, and its candidate sets are the free grid cells inside each source's own deployment region, so it is the strongest reference under the protocol rather than a bound over all possible placements. Every method is therefore compared as a quality-versus-time trade-off rather than partitioned into deployable and non-deployable groups, since NetAgent Auto-K decides in 8.21 s at 0.5885, Sionna Exhaustive reaches 0.6431 at 101.8 s, the TR 38.901 Proxy reaches 0.4396 at 181.48 s, and AutoBS reaches 0.4661 at 3.55 s.

TR 38.901 Proxy Exhaustive. A fast analytical baseline replaces propagation learning with the 3GPP TR 38.901 UMi Street Canyon path-loss formulas and a deterministic building-mask occlusion test with a frozen 12 dB penetration penalty and no shadowing. It enumerates the same candidates as Sionna Exhaustive, selects per K by the proxy score, and only its selections are simulated.

AutoBS. AutoBS is the sequential PPO baseline (Schulman et al., 2017; Lee & Molisch, 2025) under the same regional evaluator with three seeds.

Ablation. The single ablation contrasts the shared Auto-K optimization with NetAgent Full, which runs an independent optimization per K under the same ten-epoch schedule and the same learning rates, latent 2×10^{-3} and weight logits 4×10^{-2}, so the difference is not attributable to a learning-rate difference. The ablation supports the implementation as a whole, but it does not isolate whether the gain comes from the shared optimization, the weight ordering, or the candidate construction. NetAgent learns its power weights while the baselines use uniform weights, so the main comparison favours NetAgent in that respect, and the weight ablation itself is limited to the ten-scenario subset.

5.3 Radio Protocol and Fusion

The carrier is 1.8 GHz with 10 MHz bandwidth, the nominal transmit power is 43 dBm/BS, the noise is -104 dBm, and transmitters sit at 20 m with receivers at 1.5 m and omni antennas. NetAgent uses a frozen RadioDiff (Wang et al., 2025) checkpoint trained without preset base-station inputs, and all experiments run on an NVIDIA RTX PRO 6000 GPU dedicated to one job at a time. All times are end-to-end decision times on that GPU, where fixed-K methods report per-K times and their five-run sums and variable-K one 10-epoch optimization.

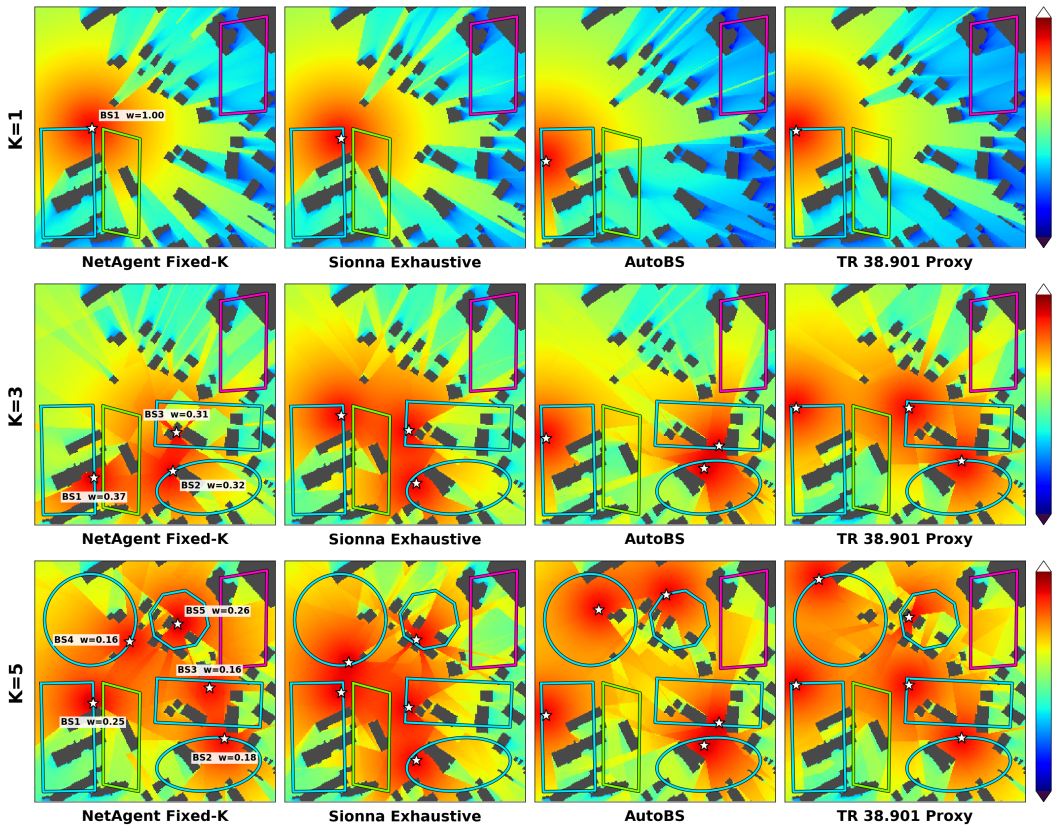


Figure 4: One map at $K = 1$, 3 and 5, one method per column, with the Strong and Weak regions overlaid.

6 RESULTS

All main-table values come from the retained WinProp records under the metrics of Section 5.1, averaged over the thirty scenarios, except the budget-curve comparison rows of Appendix B, which use four comparison rows on the same fifteen scenarios at the same four cuts under their own selection rules, while the Auto-K row is measured on all thirty scenarios, so that margin is indicative. Appendix A reports the weight ablation, Appendix C the regional shortfalls, and Appendix D the implementation details.

6.1 FIXED-K COMPARISON

Table 1 reports the protocol score of Eq. 7 per K for every method, the five-K average, and the five-run decision-time sum, and Figure 4 shows one map at three counts.

Three patterns stand out. Sionna Exhaustive attains the highest score at every count and averages 0.4869, leading NetAgent by 0.0407 at $K = 1$. NetAgent Fixed-K is the strongest deployable method overall with 0.4268 against 0.3855 for the Proxy and 0.3837 for AutoBS, a 10.7% average lead that holds at all five counts and ranges from $+4.9\%$ to $+16.1\%$.

Table 1: Fixed-K score and decision time. The Rel. row reports the score gain over the strongest deployable baseline at each count and the decision-time reduction over the Sionna Exhaustive reference.

Method	$K = 1 \uparrow$	$K = 2 \uparrow$	$K = 3 \uparrow$	$K = 4 \uparrow$	$K = 5 \uparrow$	ALL $\uparrow$	total time for the five runs [s] $\downarrow$
NetAgent Fixed-K	**0.4518** $\pm$0.2091	**0.4321** $\pm$0.1111	**0.4142** $\pm$0.0836	**0.4128** $\pm$0.0993	**0.4233** $\pm$0.0684	**0.4268** $\pm$0.1237	**29.84**
Sionna Exhaustive	0.4925 $\pm$0.2404	0.4936 $\pm$0.1239	0.4771 $\pm$0.0981	0.4882 $\pm$0.0496	0.4832 $\pm$0.0446	0.4869 $\pm$0.1304	101.82
TR 38.901 Proxy	0.3754 $\pm$0.2029	0.3723 $\pm$0.1217	0.3926 $\pm$0.0983	0.3937 $\pm$0.0949	0.3936 $\pm$0.0848	0.3855 $\pm$0.1266	181.48
AutoBS	0.4049 $\pm$0.2224	0.3651 $\pm$0.1444	0.3727 $\pm$0.1292	0.3902 $\pm$0.1008	0.3857 $\pm$0.0936	0.3837 $\pm$0.1456	3.55
Rel. vs. best baseline	**+11.6%**	**+16.1%**	**+5.5%**	**+4.9%**	**+7.5%**	**+10.7%**	**-70.7%**

NetAgent Fixed-K is the strongest deployable method on Strong coverage at every count and leads the average by 8.2%. Weak suppression collapses for every method from 32 to 42% at $K = 1$ to below 1% at $K = 5$, a structural consequence of linear-power fusion, and NetAgent trails the reference

there at every count, by 17.8% to 51.6%. On the strong–weak margin NetAgent leads the deployable baselines at every K, on average by 33.3%. Whole-map coverage, reported in Appendix C, is the least discriminative component, since all methods average between 65.0% and 66.1% and NetAgent stays within 0.7% of the strongest baseline.

Table 2: Fixed-K strong-region coverage in percent.

Method	$K=1\uparrow$	$K=2\uparrow$	$K=3\uparrow$	$K=4\uparrow$	$K=5\uparrow$	ALL↑
NetAgent Fixed-K	**72.01** ±30.70	**86.19** ±23.54	**91.30** ±18.39	**91.83** ±20.32	**94.24** ±13.49	**87.11** ±23.18
Sionna Exhaustive	73.06 ±31.44	92.09 ±19.78	93.69 ±18.73	97.18 ±8.54	97.43 ±5.33	90.69 ±21.14
TR 38.901 Proxy	58.04 ±35.72	77.90 ±30.60	87.47 ±22.16	88.96 ±22.09	90.12 ±21.24	80.50 ±29.24
AutoBS	60.79 ±33.91	78.26 ±31.57	83.17 ±29.32	89.02 ±23.77	89.58 ±23.43	80.16 ±30.46
Rel.	**+18.5%**	**+10.1%**	**+4.4%**	**+3.2%**	**+4.6%**	**+8.2%**

Table 3: Fixed-K weak-region suppression in percent.

Method	$K=1\uparrow$	$K=2\uparrow$	$K=3\uparrow$	$K=4\uparrow$	$K=5\uparrow$	ALL↑
NetAgent Fixed-K	**31.83** ±35.07	**6.16** ±11.73	**1.77** ±3.22	**1.16** ±2.18	**0.52** ±1.27	**8.29** ±20.32
Sionna Exhaustive	38.71 ±37.62	9.17 ±15.91	3.66 ±6.84	1.58 ±2.86	0.72 ±1.38	10.77 ±23.26
TR 38.901 Proxy	33.44 ±36.75	6.63 ±13.11	2.80 ±5.71	1.58 ±3.60	0.52 ±1.20	8.99 ±21.45
AutoBS	42.08 ±38.79	6.58 ±13.02	2.72 ±6.86	1.26 ±1.78	0.37 ±0.91	10.60 ±24.41
Rel.	**-24.4%**	**-7.1%**	**-36.8%**	**-26.6%**	**+0.0%**	**-21.8%**

Table 4: Fixed-K strong–weak margin in dB.

Method	$K=1\uparrow$	$K=2\uparrow$	$K=3\uparrow$	$K=4\uparrow$	$K=5\uparrow$	ALL↑
NetAgent Fixed-K	**7.91** ±11.86	**8.50** ±8.55	**5.23** ±7.58	**5.04** ±6.52	**5.17** ±5.20	**6.37** ±8.28
Sionna Exhaustive	11.30 ±1.49	13.00 ±7.99	11.49 ±6.10	11.57 ±3.94	10.75 ±3.50	11.62 ±7.16
TR 38.901 Proxy	5.78 ±8.49	5.39 ±5.97	4.75 ±5.94	4.23 ±4.60	3.75 ±3.33	4.78 ±5.89
AutoBS	5.86 ±10.02	4.93 ±7.69	4.16 ±7.06	3.74 ±4.96	2.77 ±3.84	4.29 ±7.10
Rel.	**+35.0%**	**+57.7%**	**+10.1%**	**+19.1%**	**+37.9%**	**+33.3%**

6.2 Variable-K and Auto-K Analysis

Table 5 summarizes the variable-K comparison. Sionna Exhaustive attains 0.6431 with five stations everywhere at 101.8 s per decision, since it ray-traces every candidate on the grid for each count. Among the planners that decide in seconds NetAgent Auto-K attains the highest mean at 0.5885, 26.3% above AutoBS at 0.4661, while selecting on average only 1.27 stations, which is the mean over the three frozen seeds, against 3.27 for AutoBS, with the seeds spanning 1.133 to 1.400, and it is the only adaptive method that needs no per-K sweep, deciding in 8.21 s.

Table 5: Variable-K selection.

Method	Mean K	Attainment↑	Time [s]↓
NetAgent Auto-K	**1.27**	**0.5885** ±0.1093	**8.21**
AutoBS	3.27	0.4661 ±0.0909	3.55
TR 38.901 Proxy	3.07	0.4396 ±0.1139	181.48
Sionna Exhaustive	5.00	0.6431 ±0.0454	101.82
Rel.	**-61.2%**	**+26.3%**	–

Table 6: Auto-K versus Full. Win rate counts strict wins, two scenarios tied.

Variant	Attainment↑	Time [s]↓	Win rate↑
Auto-K	**0.5885** ±0.1093	**8.21**	**20/30**
Full	0.5393 ±0.1348	29.84	8/30
Rel.	**+9.1%**	**-72.5%**	–

Table 6 contrasts the shared Auto-K optimization with NetAgent Full, which optimizes every K independently under the same settings and selection rule. The shared optimization is not a compromise, attaining 0.5885 versus 0.5393 for Full, a paired gain of +0.0492 that wins on 20 of 30 scenarios with $p < 0.05$, in 8.21 s against 29.8 s for Full. The exhaustive row shows that a majority of per-scenario wins can coexist with a lower mean, since a few scenarios dominate the shortfall and the mean responds to them most. Sionna Exhaustive scores every configuration of the per-source product exactly, so it is the strongest reference under the protocol rather than a bound over all possible placements, and the comparison against it concerns cost and per-scenario reliability rather than the mean, and Auto-K decides in 8.21 s against 101.82 s.

6.3 Scenario and Seed Variation

Table 7 reports variation across the thirty scenarios, the three frozen Auto-K seeds, and the paired comparisons. The reported seed selects 24 scenarios at $K = 1$, 5 at $K = 2$ and 1 at $K = 4$, the second seed selects 26 at $K = 1$ and 4 at $K = 2$, and the third selects 20 at $K = 1$, 8 at $K = 2$ and 2 at $K = 3$, and the seeds select the same count in 16 of 30 scenarios, where the mean pairwise displacement of the selected source positions is 16.3 m, with a median of 14.8 m and a maximum of 43.1 m. Across the three frozen seeds the selected count is the quantity that moves, since the runs deploy 38, 34 and 42 stations at means of 1.267, 1.133 and 1.400 for a three-seed average of 1.27, while the achieved quality moves far less, since attainment spans 0.0337 in a scoring pass that

evaluated the three deployments together and its seed-level spread of about 0.017 stays far below the per-scenario spreads of the attainment rows in the main tables, which lie between 0.0909 and 0.1348.

The count moves because the constraint-first rule compares the five nested prefixes on constraint violation first, so a prefix that is only marginally feasible can be accepted under one initialization and rejected under another, which makes the deployed count a near-tie outcome rather than a tightly identified property of the map. AutoBS produces identical deployments across its three seeds. Auto-K is ahead of NetAgent Fixed-K, AutoBS, and the TR 38.901 Proxy at $p < 0.05$, with paired differences of $+0.0492$, $+0.1224$, and $+0.1489$; it wins 20 of the 30 scenarios against NetAgent Fixed-K and 25 against AutoBS, the TR 38.901 Proxy, and Sionna Exhaustive. Against Sionna Exhaustive the paired difference is -0.0546 and is also significant at $p < 0.05$, so the ray-tracing reference retains the higher mean attainment while Auto-K wins the per-scenario majority.

Table 7: Paired comparisons.

Comparison	ΔAttainment	Auto-K wins	p	Judged by
Auto-K vs. NetAgent Fixed-K	$+0.0492$	$20/30$	< 0.05	Auto-K on the mean
Auto-K vs. AutoBS	$+0.1224$	$25/30$	< 0.05	Auto-K on the mean
Auto-K vs. TR 38.901 Proxy	$+0.1489$	$25/30$	< 0.05	Auto-K on the mean
Auto-K vs. Sionna Exhaustive	-0.0546	$25/30$	< 0.05	Auto-K on wins and cost

7 LIMITATIONS

Limits. The benchmark uses OSM-derived digital twins and a frozen RadioDiff prior, so missing heights, materials, and antenna details can affect the reported deployments, although the transfer from the generated field to the scored deployment is not verified here, since no measured field validation was performed. Strict feasibility is not the yardstick here, since some threshold and required-fraction combinations are unreachable under the fixed power budget, so the regional shortfall of Appendix C reports how closely each deployment approaches its targets rather than whether it passed a test. The released records hold the planner's own per-scenario evaluation and flags rather than the scorer's per-scenario scored metrics.

8 CONCLUSION

We presented NetAgent, which plans base-station deployments by differentiating regional constraints through a frozen RM diffusion prior instead of enumerating candidate configurations, and one shared optimization yields nested Top-1 through Top-5 deployments that the constraint-first K selection scans. Judged solely by WinProp on thirty OSM-derived scenarios, the shared 10-epoch Auto-K optimization attains the highest mean goal completion among the planners that decide in seconds at 0.5885, selecting on average 1.27 stations and deciding in 8.21 s. Against NetAgent Fixed-K, AutoBS, and the TR 38.901 Proxy the paired differences are $+0.0492$, $+0.1224$, and $+0.1489$, all at $p < 0.05$, and Auto-K wins 20 of the 30 scenarios against NetAgent Fixed-K and 25 against AutoBS and the TR 38.901 Proxy. Against Sionna Exhaustive the paired difference is -0.0546 at $p < 0.05$, so the ray-tracing reference keeps the higher mean while Auto-K wins the per-scenario majority. The ablation isolates the gain, since independently optimized per-K deployments cost 0.0492 in attainment at roughly four times the decision cost, and measured RMs with physical deployments remain important future work.

AI USE STATEMENT

This statement is required by ICLR 2027 and does not count toward the page limit. In this work, generative AI tools were used for language editing of the manuscript, LaTeX table formatting, literature retrieval and discovery, and assistance with experiment-visualization code and figures. Generative AI tools were not used for any task that requires disclosure in the research itself, including generating synthetic datasets, developing theoretical models or conceptual frameworks, formulating or proving mathematical claims, proposing or refining hypotheses, designing the research methodology or experiments, implementing methods, cleaning or reformatting datasets, translating text, or analyzing and interpreting results; all methods, experiments, and conclusions were conceived, executed, and verified by the authors. All AI-assisted text and code were reviewed by the authors, who take full responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

Reproducibility statement

All thirty scenario definitions, comprising OSM-derived maps, Main, Strong, and Weak region masks, per-region thresholds, target ratios, minimum margins, and per-cell random seeds drawn from operating-system entropy, are frozen as metadata with checksums before any method runs in Section 5. The fixed radio protocol, the linear-domain fusion rule, and the exact score and attainment formulas are given in Eqs. 7 and 8. Every method shares the same candidate grid, power budget, and electromagnetic-simulation evaluation procedure in Section 5.1, and the retained records contain every coordinate, weight, timing field, and the planner's own per-scenario regional metrics, constraint evaluation, and feasibility flags, while the scorer's per-scenario scored metrics are not part of those records. The scenario configurations and the evaluation pipeline are available in the anonymous repository at `https://anonymous.4open.science/r/netagent-release-F36D/`, and the retained records will be released upon publication.

References

3GPP. TR 38.901: Study on channel model for frequencies from 0.5 to 100 GHz. Technical report, 3rd Generation Partnership Project, 2020. URL `https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3173`.

Altair Engineering. Altair Feko WinProp: Propagation modeling and network planning. User Guide, Release 2024, 2024. URL `https://help.altair.com/feko/`.

Fayçal Aït Aoudia, Jakob Hoydis, Merlin Nimier-David, Baptiste Nicolet, Sebastian Cammerer, and Alexander Keller. Sionna rt: Technical report, 2026. URL `https://arxiv.org/abs/2504.21719`.

Hamadi Chihaoui and Paolo Favaro. Diffusion image prior. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 24636–24644, 2025.

Kangkang Deng, Rui Wang, Zhenyuan Zhu, Junyu Zhang, and Zaiwen Wen. The augmented lagrangian methods: Overview and recent advances. *SIAM Review*, 68(3):471–576, 2026. doi: 10.1137/24M1706153. URL `https://doi.org/10.1137/24M1706153`.

Zongyuan Deng, Yujie Cai, Qing Liu, Shiyao Mu, Bin Lyu, and Zhen Yang. Teleplannet: An ai-driven framework for efficient telecom network planning, 2025. URL `https://arxiv.org/abs/2505.13831`.

Zheng Fang, Kangjun Liu, Ke Chen, Qingyu Liu, Jianguo Zhang, Lingyang Song, and Yaowei Wang. Radioformer: A multiple-granularity radio map estimation transformer with 1‰spatial sampling, 2025. URL `https://arxiv.org/abs/2504.19161`.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851. Curran Associates, Inc., 2020. URL `https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf`.

Jakob Hoydis, Faycal Ait 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.

Haozhe Jia, Wenshuo Chen, Zhihui Huang, Lei Wang, Hongru Xiao, Nanqian Jia, Keming Wu, Songning Lai, Bowen Tian, and Yutao Yue. Physics-informed representation alignment for sparse radio-map reconstruction. In *Proceedings of the 33rd ACM International Conference on Multimedia*, MM '25, pp. 12352–12360, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400720352. doi: 10.1145/3746027.3758161. URL `https://doi.org/10.1145/3746027.3758161`.

Ju-Hyung Lee and Andreas F. Molisch. A scalable and generalizable pathloss map prediction. *IEEE Transactions on Wireless Communications*, 23(11):17793–17806, 2024. doi: 10.1109/TWC.2024.3457431. URL https://doi.org/10.1109/TWC.2024.3457431.

Ju-Hyung Lee and Andreas F. Molisch. Autobs: Autonomous base station deployment with reinforcement learning and digital network twins, 2025. URL https://arxiv.org/abs/2502.19647.

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.

Cunyi Liao, Yi Zheng, Ji Wang, and Shouyin Liu. Visual transformer based unified framework for radio map estimation and optimized site selection. *IEICE Transactions on Communications*, E109-B(1):50–64, 2026. doi: 10.23919/transcom.2025EBP3055.

Xuanhao Luo, Zhizhen Li, Zhiyuan Peng, Mingzhe Chen, and Yuchen Liu. Denoising diffusion probabilistic model for radio map estimation in generative wireless networks. *IEEE Transactions on Cognitive Communications and Networking*, 11(2):751–763, 2025. doi: 10.1109/TCCN.2025.3529879.

Jonathan O'Shea and Conor Brennan. Ea-rmenet – path loss prediction in urban environments using deep learning, 2026. URL https://arxiv.org/abs/2607.16449.

Roberto Pegurri, Francesco Linsalata, Eugenio Moro, Jakob Hoydis, and Umberto Spagnolini. Toward digital network twins: Integrating sionna rt in ns-3 for 6g multi-rat networks simulations. In *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, 2025. doi: 10.1109/INFOCOMWKSHPS65812.2025.11152946.

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, 2025. URL https://arxiv.org/abs/2501.11283.

Nguyen Duc Minh Quang, Chang Liu, Huy-Trung Nguyen, Shuangyang Li, Derrick Wing Kwan Ng, and Wei Xiang. 3d dynamic radio map prediction using vision transformers for low-altitude wireless networks, 2026. URL https://arxiv.org/abs/2511.19019.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi (eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241, Cham, 2015. Springer International Publishing. ISBN 978-3-319-24574-4.

Ali Saeizadeh, Miead Tehrani-Moayyed, Davide Villa, J. Gordon Beattie, Pedram Johari, Stefano Basagni, and Tommaso Melodia. Airmap: Ai-generated radio maps for wireless digital twins. *IEEE Transactions on Wireless Communications*, 25:18871–18884, 2026. doi: 10.1109/TWC.2026.3704457.

Koya Sato and Katsuya Suto. Bayesian optimization framework for channel simulation-based base station placement and transmission power design. *IEEE Networking Letters*, 6(4):217–221, 2024. doi: 10.1109/LNET.2024.3469175.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms, 2017. URL https://arxiv.org/abs/1707.06347.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=St1giarCHLP.

Zeyao Sun, Bohao Fan, Qingyu Liu, Shuhang Zhang, and Lingyang Song. Radiopit: Radio map generation with pixel transformer driven by ultra-sparse real-world data. In *2025 30th Asia-Pacific Conference on Communications (APCC)*, pp. 1–6, 2025. doi: 10.23919/APCC64555.2025.11279807.

Marin Vlastelica, Tatiana López-Guevara, Kelsey Allen, Peter Battaglia, Arnaud Doucet, and Kimberley Stachenfeld. Diffusion generative inverse design. In *ICML Workshop on Structured Probabilistic Inference and Generative Modeling*, 2023.

Xiucheng Wang, Keda Tao, Nan Cheng, Zhisheng Yin, Zan Li, Yuan Zhang, and Xuemin Shen. Radiodiff: An effective generative diffusion model for sampling-free dynamic radio map construction. *IEEE Transactions on Cognitive Communications and Networking*, 11(2):738–750, 2025. doi: 10.1109/TCCN.2024.3504489.

Xiucheng Wang, Yuhao Pan, Nan Cheng, Çağkan Yapar, Ruijin Sun, Zhisheng Yin, Conghao Zhou, Wenchao Xu, Yuxiang Zhang, Jianhua Zhang, Shuguang Cui, and Xuemin Shen. A tutorial on learning-based radio map construction: Data, paradigms, and physics-awareness, 2026. URL https://arxiv.org/abs/2603.17499.

Yanhu Wang, Muhammad Muzammil Afzal, Zhengyang Li, Jie Zhou, Chenyuan Feng, Shuaishuai Guo, and Tony Q. S. Quek. Large language model as a catalyst: A paradigm shift in base station siting optimization, 2024. URL https://arxiv.org/abs/2408.03631.

Jie Zhang, Mostafa Rahmani Ghourtani, Swarna Bindu Chetty, Paul Daniel Mitchell, and Hamed Ahmadi. Digital twin-assisted resilient planning for mmWave IAB networks via graph attention networks. *arXiv preprint arXiv:2509.12499*, 2025. URL https://arxiv.org/abs/2509.12499.

Çağkan Yapar. Learning coverage- and power-optimal transmitter placement from city maps: A comparative study of direct and indirect neural approaches, 2026. URL https://arxiv.org/abs/2604.22056.

A Weight Ablation Details

Table 8 ablates the learned per-source weights against frozen uniform weights under the same fixed-K protocol with ten epochs on the ten-scenario subset, so these values are not directly comparable with the thirty-scenario tables of the main text. The two arms coincide at $K = 1$ because the single source carries the whole budget. At $K \geq 2$ the learned weights raise the score at all four non-trivial counts, most at $K = 2$ with 0.4380 against 0.4186, and overall with 0.4006 against 0.3933, at comparable runtime. Tables 9 and 10 report the remaining per-arm metrics, including protocol score, strong–weak margin, strong-region coverage, weak-region suppression, whole-map coverage, and runtime per K with ten scenarios as mean $\pm$ std.

Table 8: Weight ablation, score by K.

Arm	K=1 ↑	K=2 ↑	K=3 ↑	K=4 ↑	K=5 ↑	ALL↑
Learned	**0.4784**	**0.4380**	**0.3614**	**0.3683**	**0.3568**	**0.4006**
Uniform	0.4784	0.4186	0.3557	0.3609	0.3528	0.3933
Rel.	**+0.0%**	**+4.6%**	**+1.6%**	**+2.1%**	**+1.1%**	**+1.9%**

B Wall-Clock Budget Curves

Table 11 scores the deployment that each method can deliver within four wall-clock budgets of 30, 60, 120, and 240 s. The four comparison rows are scored on the same fifteen scenarios at the same four cuts, and each cell is what that method can deliver at that moment under its own selection rule, so a column is a progress snapshot rather than a run of fixed length. A method that has already finished returns the same deployment at every later cut, and no cell is empty, since every method completes at least one run before 30 s. The Auto-K row is the reported 10-epoch decision, whose single decision costs 8.21 s and therefore finishes before the smallest budget, so it is flat at 0.5885 and is measured on all thirty scenarios rather than on the fifteen-scenario subset, which is why it carries the dagger. NetAgent Auto-K leads the four comparison rows at every budget by 27.1% to 32.9% over the strongest of them, the Proxy, but that margin is indicative rather than a like-for-like comparison, because the Auto-K row covers a different and larger scenario set. Sionna Exhaustive stays flat at 0.489 because its first completed sweep already fixes its quality within these budgets, and the Proxy and AutoBS are flat from 60 s on at 0.443 and 0.440.

Table 9: Weight ablation, learned arm.

K	Score ↑	Margin [dB] ↑	S% ↑	W% ↑	Cov% ↑	Time [s]
1	0.4784 ±0.1495	+13.92 ±5.35	53.9 ±25.7	46.9 ±37.7	36.87 ±20.89	4.75 ±0.89
2	0.4380 ±0.1321	+12.10 ±7.21	74.7 ±18.8	11.9 ±18.5	57.14 ±20.49	6.11 ±1.77
3	0.3614 ±0.1264	+5.98 ±5.45	73.4 ±31.5	3.6 ±5.3	65.65 ±20.20	4.80 ±0.71
4	0.3683 ±0.1128	+5.20 ±5.88	77.8 ±22.9	1.2 ±1.7	74.32 ±19.38	5.55 ±1.30
5	0.3568 ±0.1492	+4.17 ±5.01	77.3 ±33.4	0.5 ±0.8	78.01 ±19.62	5.87 ±1.10
All	**0.4006** ±0.1382	**+8.27** ±6.87	**71.4** ±27.4	**12.8** ±25.3	**62.40** ±24.32	**5.41** ±1.29

Table 10: Weight ablation, uniform arm.

K	Score ↑	Margin [dB] ↑	S% ↑	W% ↑	Cov% ↑	Time [s]
1	0.4784 ±0.1495	+13.92 ±5.35	53.9 ±25.7	46.9 ±37.7	36.87 ±20.89	4.49 ±1.11
2	0.4186 ±0.1400	+11.09 ±7.42	70.9 ±22.2	11.4 ±18.6	56.55 ±20.06	5.53 ±0.96
3	0.3557 ±0.1218	+5.20 ±5.82	72.7 ±31.2	4.3 ±5.6	65.66 ±20.39	5.24 ±0.98
4	0.3609 ±0.1061	+4.56 ±5.44	78.1 ±21.9	1.0 ±1.6	74.59 ±19.22	5.93 ±1.59
5	0.3528 ±0.1450	+3.71 ±4.59	77.6 ±33.0	0.4 ±0.7	78.39 ±19.81	6.24 ±1.43
All	**0.3933** ±0.1371	**+7.69** ±6.90	**70.6** ±27.5	**12.8** ±25.3	**62.41** ±24.39	**5.49** ±1.34

Table 11: Attainment at fixed wall-clock budgets. †Reported Auto-K deployment, measured on all thirty scenarios.

Method	30 s	60 s	120 s	240 s	ALL
NetAgent Auto-K	**0.5885**†	**0.5885**†	**0.5885**†	**0.5885**†	**0.5885**† ±0.1093
Sionna Exhaustive	0.4890	0.4890	0.4890	0.4893	0.4891 ±0.0396
TR 38.901 Proxy	0.4629	0.4429	0.4429	0.4429	0.4479 ±0.0910
AutoBS	0.4397	0.4397	0.4397	0.4397	0.4397 ±0.0771
Rel. vs. best baseline	**+27.1%**	**+32.9%**	**+32.9%**	**+32.9%**	**+31.4%**

C Regional Shortfall Details

The regional shortfall measures how far a deployment falls short of the requirements it is scored against. For strong coverage it is the percentage-point deficit against the required fraction, for weak suppression the same deficit on the weak fraction, and for the margin the relative deficit against the required minimum margin. Every deficit is clipped at zero, so a satisfied requirement contributes nothing. Requirements differ across the thirty scenarios, and the values below are computed against the thirty-scenario mean requirement of $\rho^s = 89.33\%$, $\rho^w = 88.83\%$, and $M_{\min} = 15.83\,\mathrm{dB}$, so they are aggregate indicators rather than per-scenario means. Only these aggregate deficits are reported here, because a per-scenario spread would require the scorer's per-scenario output and is not reproduced. Because the clipping at zero is convex, averaging the per-scenario deficit can only raise these values, and the numbers reported here therefore bound the shortfall from below. Because some of those combinations are unreachable under the fixed power budget, these deficits read how closely a deployment approaches its targets rather than whether it passed. The margin shortfall can exceed one hundred percent, because a deployment whose Strong–Weak gap is negative misses the required margin by more than that margin. As in the main tables, the underlined value is the strongest deployable baseline of that column and lower values are better. Each Rel. row reports NetAgent against that underlined baseline, so a negative entry means a smaller shortfall than the baseline and is therefore the better outcome. The strong-coverage entry at $K = 5$ is undefined because both values are zero and is marked accordingly.

D Implementation Details

Table 16 lists the full hyperparameter configuration, which is shared by every method and frozen before evaluation.

D.1 Search Protocol

The exhaustive references share one candidate space. A scenario is a 256×256 array at 1 m per pixel and the candidate lattice is a regular 32×32 grid whose nodes are the centers of the 8×8 pixel blocks, so consecutive candidates sit 8 m apart. A node is admissible for a source when its pixel is free space and lies inside that source's own deployment region, hence a source holds at most 32 admissible nodes in the thirty scenarios and fewer wherever its region covers fewer free

Table 12: Strong-coverage shortfall in percentage points.

Method	$K=1\downarrow$	$K=2\downarrow$	$K=3\downarrow$	$K=4\downarrow$	$K=5\downarrow$	ALL$\downarrow$
NetAgent Fixed-K	**17.32**	**3.14**	**0.00**	**0.00**	**0.00**	**4.09**
Sionna Exhaustive	16.27	0.00	0.00	0.00	0.00	3.25
TR 38.901 Proxy	31.29	11.43	1.86	0.37	0.00	8.99
AutoBS	28.54	11.07	6.16	0.31	0.00	9.22
Rel. vs. best baseline	**-39.3%**	**-71.6%**	**-100.0%**	**-100.0%**	**–**	**-54.5%**

Table 13: Weak-suppression shortfall in percentage points.

Method	$K=1\downarrow$	$K=2\downarrow$	$K=3\downarrow$	$K=4\downarrow$	$K=5\downarrow$	ALL$\downarrow$
NetAgent Fixed-K	**57.00**	**82.67**	**87.06**	**87.67**	**88.31**	**80.54**
Sionna Exhaustive	50.12	79.66	85.17	87.25	88.11	78.07
TR 38.901 Proxy	55.39	82.20	86.03	87.25	88.31	79.84
AutoBS	46.75	82.25	86.11	87.57	88.46	78.23
Rel. vs. best baseline	**+21.9%**	**+0.6%**	**+1.2%**	**+0.5%**	**+0.0%**	**+3.0%**

Table 14: Margin shortfall in percent of the required margin.

Method	$K=1\downarrow$	$K=2\downarrow$	$K=3\downarrow$	$K=4\downarrow$	$K=5\downarrow$	ALL$\downarrow$
NetAgent Fixed-K	**50.1**	**46.3**	**66.9**	**68.2**	**67.3**	**59.8**
Sionna Exhaustive	28.6	17.9	27.4	26.9	32.1	26.6
TR 38.901 Proxy	63.5	66.0	70.0	73.3	76.3	69.8
AutoBS	63.0	68.9	73.7	76.4	82.5	72.9
Rel. vs. best baseline	**-20.6%**	**-29.8%**	**-4.4%**	**-7.0%**	**-11.7%**	**-14.4%**

Table 15: Fixed-K whole-map coverage in percent.

Method	$K=1\uparrow$	$K=2\uparrow$	$K=3\uparrow$	$K=4\uparrow$	$K=5\uparrow$	ALL$\uparrow$
NetAgent Fixed-K	**39.52** $\pm$17.42	**59.50** $\pm$17.90	**70.00** $\pm$16.74	**77.35** $\pm$15.22	**81.76** $\pm$13.20	**65.63** $\pm$21.99
TR 38.901 Proxy	38.00 $\pm$31.76	60.86 $\pm$17.57	70.74 $\pm$16.27	78.09 $\pm$14.93	82.73 $\pm$13.74	66.09 $\pm$22.48
Sionna Exhaustive	39.34 $\pm$17.47	58.84 $\pm$17.74	69.41 $\pm$15.31	76.29 $\pm$14.35	81.26 $\pm$12.88	65.03 $\pm$21.49
AutoBS	36.78 $\pm$16.15	58.32 $\pm$16.69	68.49 $\pm$15.82	78.01 $\pm$14.51	84.02 $\pm$12.79	65.12 $\pm$22.54
Rel.	**+4.0%**	**-2.2%**	**-1.0%**	**-0.9%**	**-2.7%**	**-0.7%**

blocks. A configuration at count K selects one admissible node for each of the K sources, and the search enumerates the full product of the per-source sets rather than a pruned subset. At $K = 5$ with every source at the lattice maximum that product contains $32^5 = 33{,}554{,}432$ configurations, which is the largest product over the study, and a product above the configured ceiling raises an error instead of being subsampled, so every reported selection is an exact optimum over the product. Each per-source candidate is ray traced once, and the ray-traced search then evaluates every configuration of the product from those cached fields directly on the GPU in chunks of 4096 configurations, where a chunk reduces to its best scalar and index before the next chunk is materialized, so the working set stays bounded while the count grows. The five counts are searched one after another and independently, so a selection at count K is not a prefix of the selection at count $K + 1$.

D.2 Differentiable Optimization and Coordinate Extraction

Each iteration decodes the latent with three deterministic DDIM steps at $\eta = 0$, so the generated map is a deterministic differentiable function of the latent, no noise is injected along the decode, and the gradient that reaches a latent is the gradient of the task loss through a fixed field. The optimizer holds two parameter groups, the latents and the weight logits, and the prior is never one of them, since it is loaded in evaluation mode and its weights receive no update, so only the $\{Z_k\}$ and the logits behind Eq. equation 6 move during a run. The latent learning rate is 2×10^{-3} and the weight logits use 4×10^{-2} in every ten-epoch run, which covers NetAgent Fixed-K, the Full arm of the ablation, both arms of the weight ablation, and the shared Auto-K arm, so the per-count and the shared arms share one learning-rate setting; the legacy two-hundred-epoch configuration keeps the library default of 2×10^{-3} for both.

The constraint term of Eq. equation 3 is rescheduled once per epoch rather than once per iteration. With $\mu = 5$ fixed, a multiplier rises after every epoch as $\lambda_r \leftarrow \min(25,\ [\lambda_r + \mu v_r]_+)$, so the penalty on a requirement that is still violated is compounded across epochs until it reaches its ceiling of 25. The update is monotone and never returns a multiplier to zero, so a requirement that was violated early keeps a residual penalty afterwards, and the term therefore behaves close to a plain penalty at

Table 16: Hyperparameters and implementation details.

Component	Setting
Diffusion prior	RadioDiff (Wang et al., 2025), frozen; U-Net backbone
Sampling	Deterministic three-step DDIM (Song et al., 2021), $\eta = 0$
Optimization	10 epochs; Adam over two parameter groups, the latents and the weight logits
Station count	$K \in \{1, \ldots, 5\}$; Auto-K trains once at $K_{\max} = 5$
Weight parameterization	Softmax with floor $w_{\min} = 0.05$
Constraint loss	Augmented Lagrangian (Deng et al., 2026), $\mu = 5$, $\lambda_r \leq 25$, dual update every epoch
Coverage softening	Sigmoid temperature 3 dB; CVaR level $\alpha = 0.9$
Radio protocol	1.8 GHz, 10 MHz, 43 dBm/BS, noise -104 dBm (3GPP, 2020)
Scenarios	Thirty OSM-derived maps, 256×256 at 1 m per pixel
Baseline search	Sionna Exhaustive over a 32×32 candidate grid
Evaluation	WinProp, fixed radio protocol of Section 5
Hardware	NVIDIA RTX PRO 6000 GPU
Auto-K selection	Five nested prefixes of the single $K_{\max} = 5$ run; feasibility tolerance 10^{-6} on the normalized violation; fallback ranks prefixes by violation, then utility, then count
Utility weighting	Strong 0.30, weak 0.30, margin 0.30, whole-map 0.10; quality $\epsilon = 0.20$
Region priorities	Strong, weak, margin, location 1.0; worst-tail 0.75
Tail risk	CVaR $\alpha = 0.90$; strong and weak tail tolerance 5 dB
Region loss weights	Main 2.0, Strong 1.0, Weak 1.0
Source integrity	$\lambda_{\text{source}} = 2.0$; inside 2.0, outside 1.0; peak margin 12 dB; hotspot margin 10 dB over 25 pixels; peak guard 12 px
Component overlap	Overlap and diversity weight 0.02
Spatial regularization	Total variation 3×10^{-4}, decayed by $(1 - \text{progress})^2$
Learning rates	Latents 2×10^{-3} and weight logits 4×10^{-2} in every ten-epoch run, covering the per-count arms, the weight ablation, and the shared Auto-K arm, so all reported ten-epoch results share one learning-rate setting; the legacy two-hundred-epoch configuration keeps the library default of 2×10^{-3} for both
Diffusion steps	250 prior sampling steps; three deterministic DDIM steps per optimization iteration
Task thresholds	Strong -50 dBm, weak -60 dBm, whole map -50 dBm; required fraction 0.90; margin 15 dB; location fraction 0.85
Softening	Region temperature 3 dB; extreme temperature 1 dB; location softmax $\beta = 12$
Path-loss mapping	Floor -147.0 dB, ceiling -47.84 dB, transmit power 43 dBm

the start of a run and close to a hard constraint at the end on whichever requirements remain binding under the ten-epoch budget.

The coordinate of a source is read from the retained field of the best iteration once the optimization stops, as the lattice pixel of largest value over the free pixels of that source's own deployment region, which is Eq. equation 5 evaluated directly on the 1 m lattice. The readout contains no local-maximum search, no sub-pixel interpolation and no separation rule, so a coordinate is an integer pixel that is reproducible from the retained field alone. Separation between sources is encouraged only through the peak-guard term of the source-integrity loss, and because a source is confined to its own region in the readout as well as in the loss, two sources report distinct coordinates whenever their deployment regions are disjoint.

D.3 Power Fusion in Linear Units

Eq. equation 1 combines the sources in linear units, which is the physical rule. A single-source field is reported in dBm, so the corresponding power is $p_k = 10^{P_k/10}$ in that same reference unit, and fields that overlap add as powers rather than as levels, hence $P = 10 \log_{10} \sum_k K w_k p_k$ with every weight scaled by K so that uniform weights give one nominal station per source and the sum equals K nominal stations. Combining the same fields by an arithmetic mean of their dBm values instead returns the level of the geometric mean of the p_k, because a dB value is proportional to the logarithm of a power, whereas superposition returns the level of a weighted arithmetic mean of the same powers. At uniform weights and equal per-source levels the fused level exceeds the averaged level by $10 \log_{10} K$ dB, which is 3.0 dB at $K = 2$ and 7.0 dB at $K = 5$, since the fused level is the level of K times the arithmetic mean of the per-source powers while the averaged level is the level of their geometric mean, which never exceeds the arithmetic mean. The gap widens as soon as the per-source levels differ, so averaging on the dB scale would understate the coverage fractions

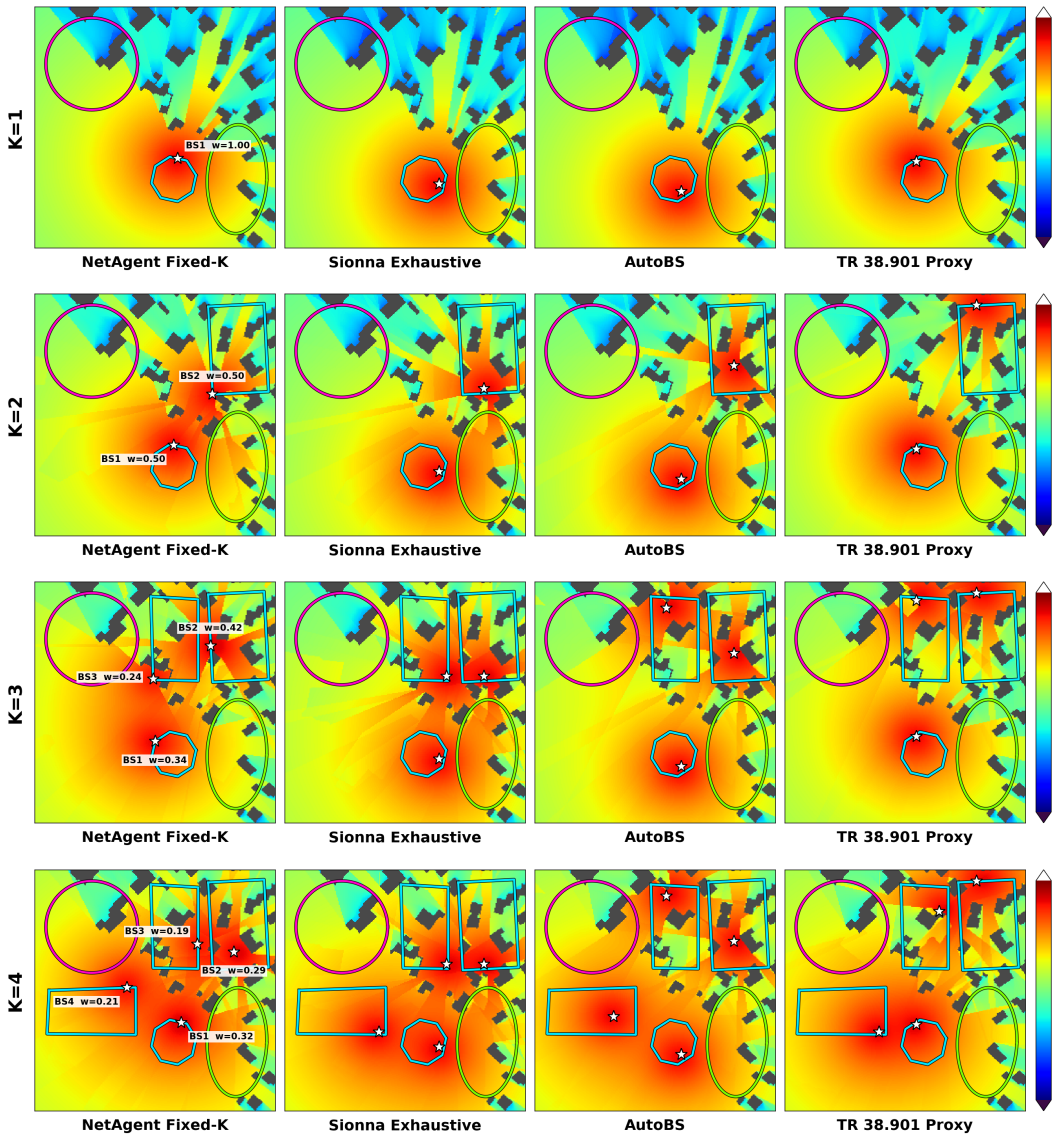
and the Strong–Weak margin by discarding the constructive addition of every extra transmitter. The fusion rule is therefore applied in linear units at the optimization stage and at the evaluation stage alike, and no stage of the pipeline averages per-source levels on the dB scale.

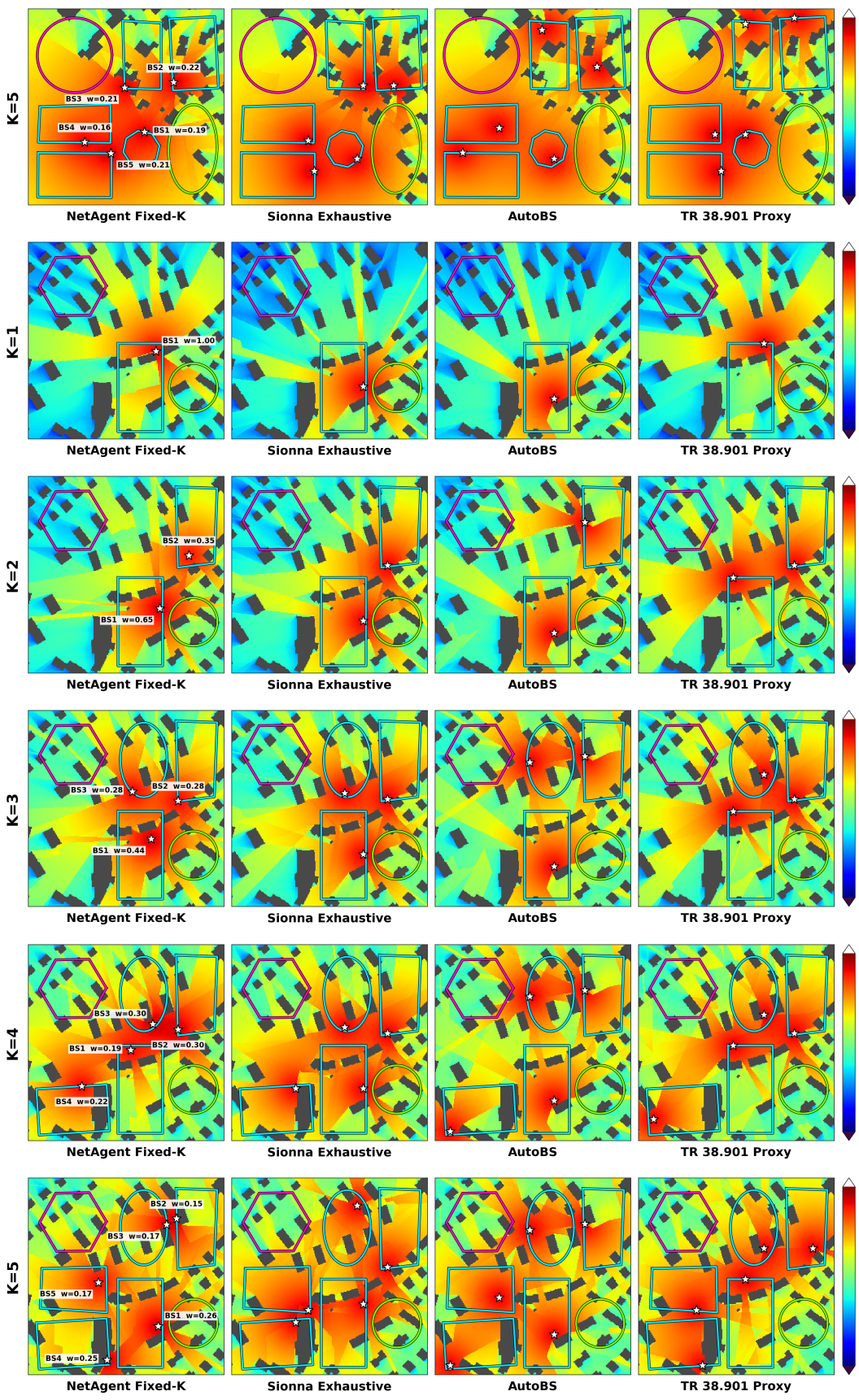
E Qualitative Simulator Comparisons

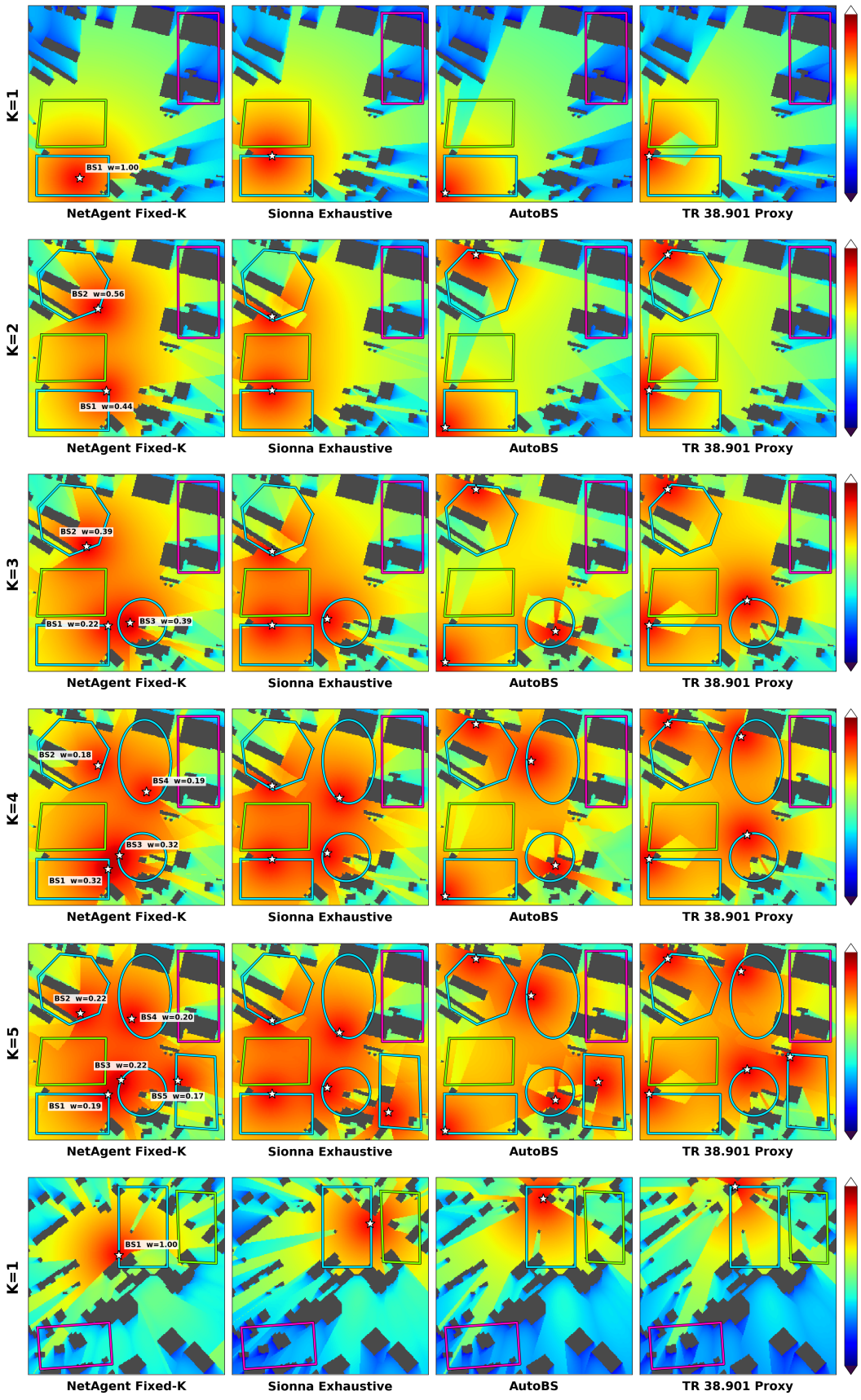
Every deployment behind the tables is rendered by the same scorer that produces the reported metrics. The Main region is cyan, the Strong region green and the Weak region magenta, one colour bar sits at the far right of each row, and the NetAgent panels label every source with its learned weight. For space each experiment shows a subset of its rows, grouped per map and ranked by how far NetAgent beats the strongest deployable baseline on that map and count, and the remaining rows and cells are omitted while the full set is released with the code.

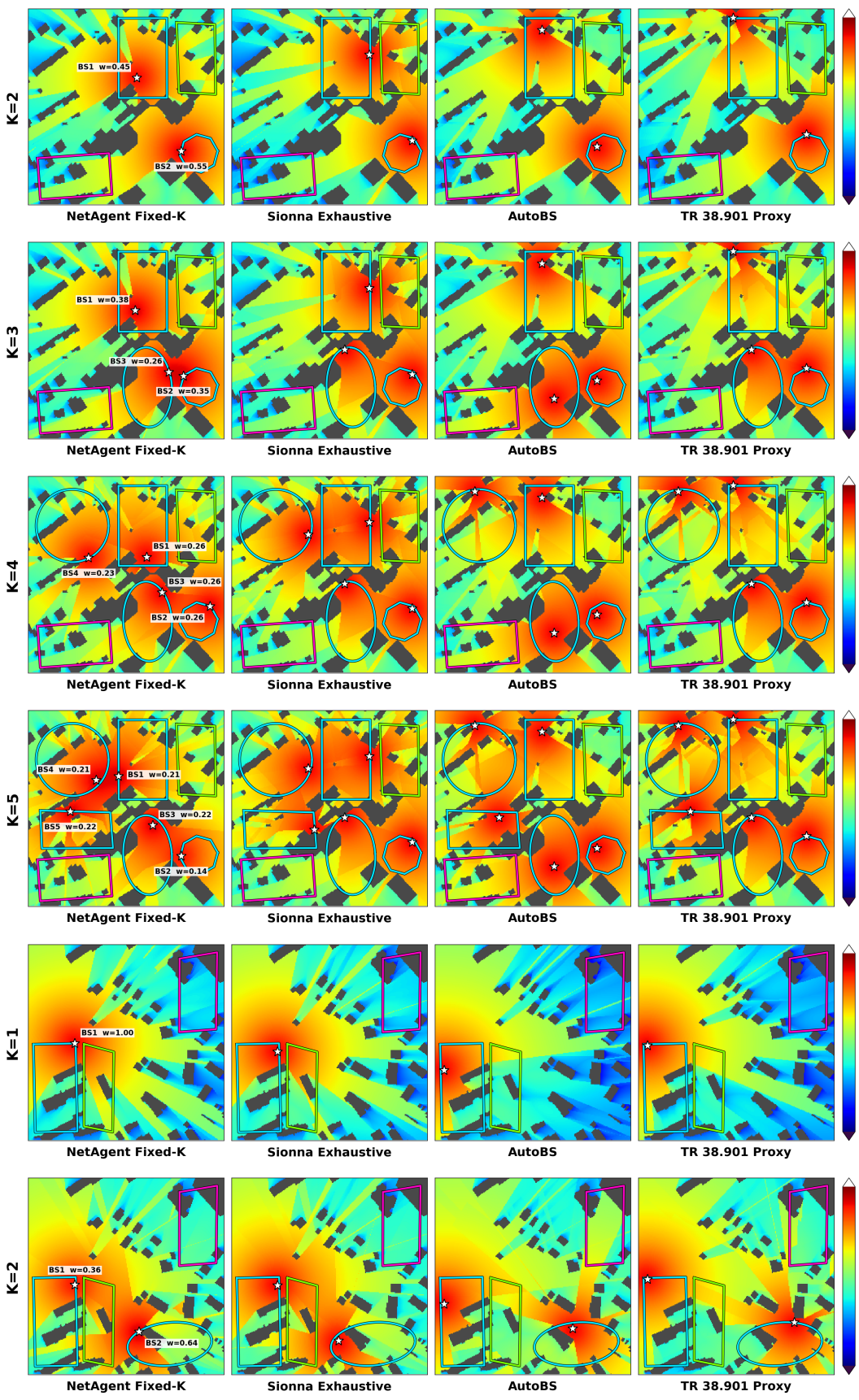
E.1 Fixed-K deployments

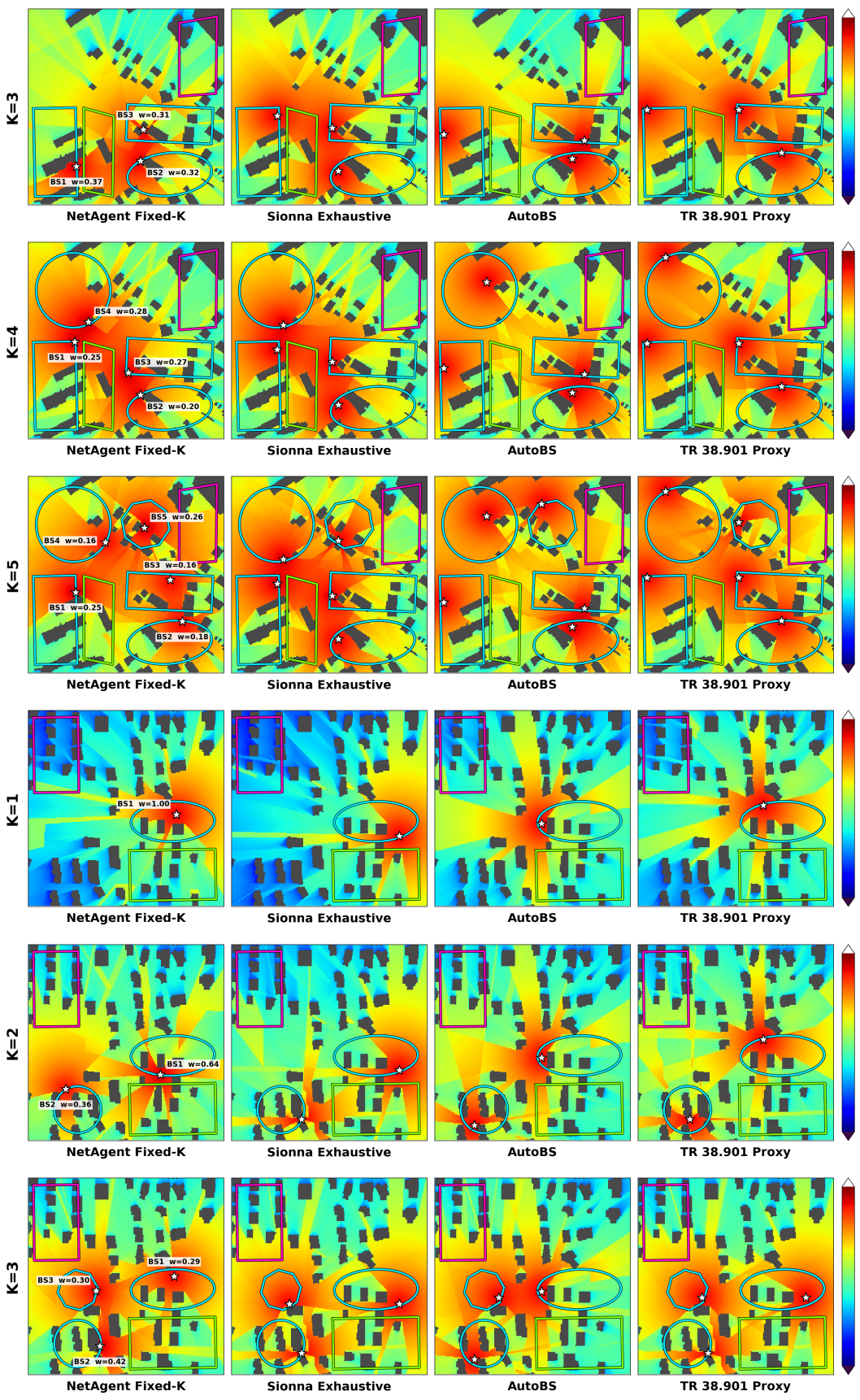
NetAgent Fixed-K | Sionna Exhaustive | AutoBS | TR 38.901 Proxy, one row per map and count. Six maps are shown in full, each at all five counts in ascending order, chosen because they are the maps where NetAgent leads the strongest deployable baseline most on their best cell, and the remaining maps are omitted while the full set is released with the code.

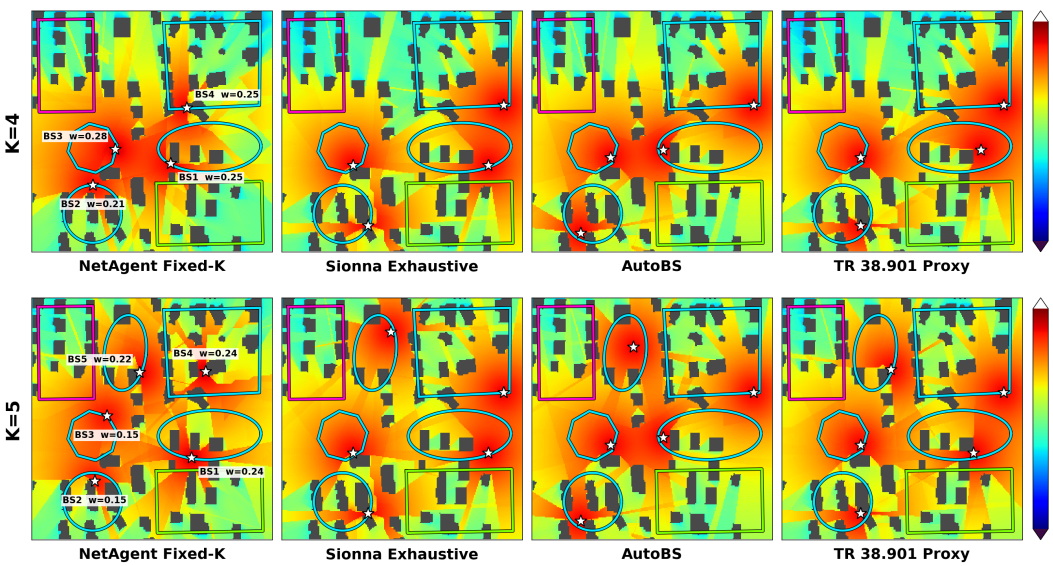



K=5
BS2 w=0.22
BS3 w=0.21
BS4 w=0.16
BS1 w=0.19
BS5 w=0.21
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=1
BS1 w=1.00
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=2
BS2 w=0.35
BS1 w=0.65
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=3
BS3 w=0.28
BS2 w=0.28
BS1 w=0.44
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=4
BS3 w=0.30
BS1 w=0.19
BS2 w=0.30
BS4 w=0.22
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=5
BS2 w=0.15
BS3 w=0.17
BS5 w=0.17
BS1 w=0.26
BS4 w=0.25
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy


K=1
BS1 w=1.00
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=2
BS2 w=0.56
BS1 w=0.44
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=3
BS2 w=0.39
BS1 w=0.22
BS3 w=0.39
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=4
BS2 w=0.18
BS4 w=0.19
BS3 w=0.32
BS1 w=0.32
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=5
BS2 w=0.22
BS4 w=0.20
BS3 w=0.22
BS5 w=0.17
BS1 w=0.19
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=1
BS1 w=1.00
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy

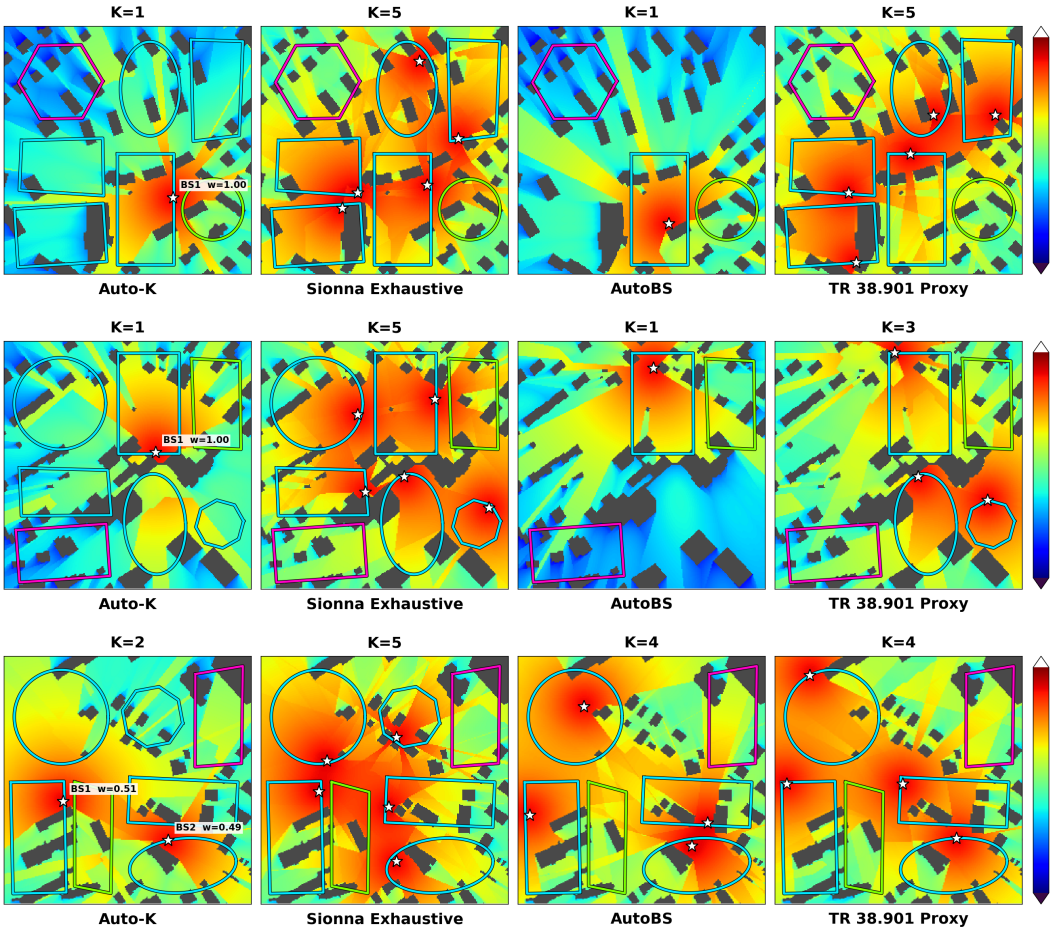

K=2
BS1 w=0.45
BS2 w=0.55
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=3
BS1 w=0.38
BS3 w=0.26
BS2 w=0.35
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=4
BS1 w=0.26
BS4 w=0.23
BS3 w=0.26
BS2 w=0.26
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=5
BS4 w=0.21
BS1 w=0.21
BS3 w=0.22
BS5 w=0.22
BS2 w=0.14
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=1
BS1 w=1.00
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=2
BS1 w=0.36
BS2 w=0.64
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy

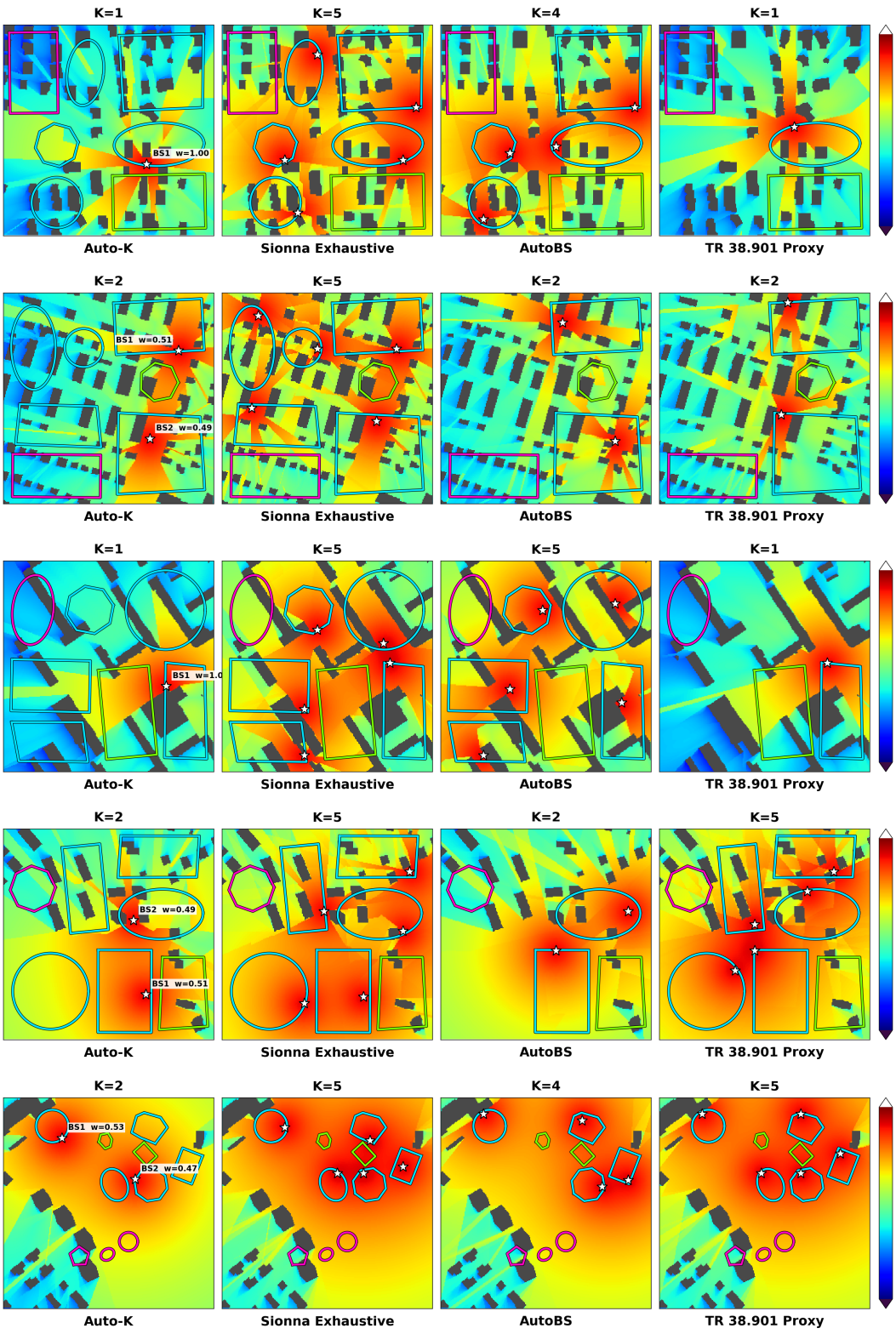

K=3
BS3 w=0.31
BS1 w=0.37
BS2 w=0.32
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=4
BS4 w=0.28
BS1 w=0.25
BS3 w=0.27
BS2 w=0.20
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=5
BS5 w=0.26
BS4 w=0.16
BS3 w=0.16
BS1 w=0.25
BS2 w=0.18
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=1
BS1 w=1.00
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=2
BS1 w=0.64
BS2 w=0.36
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy
K=3
BS1 w=0.29
BS3 w=0.30
BS2 w=0.42
NetAgent Fixed-K
Sionna Exhaustive
AutoBS
TR 38.901 Proxy



E.2 VARIABLE-K COMPARISON

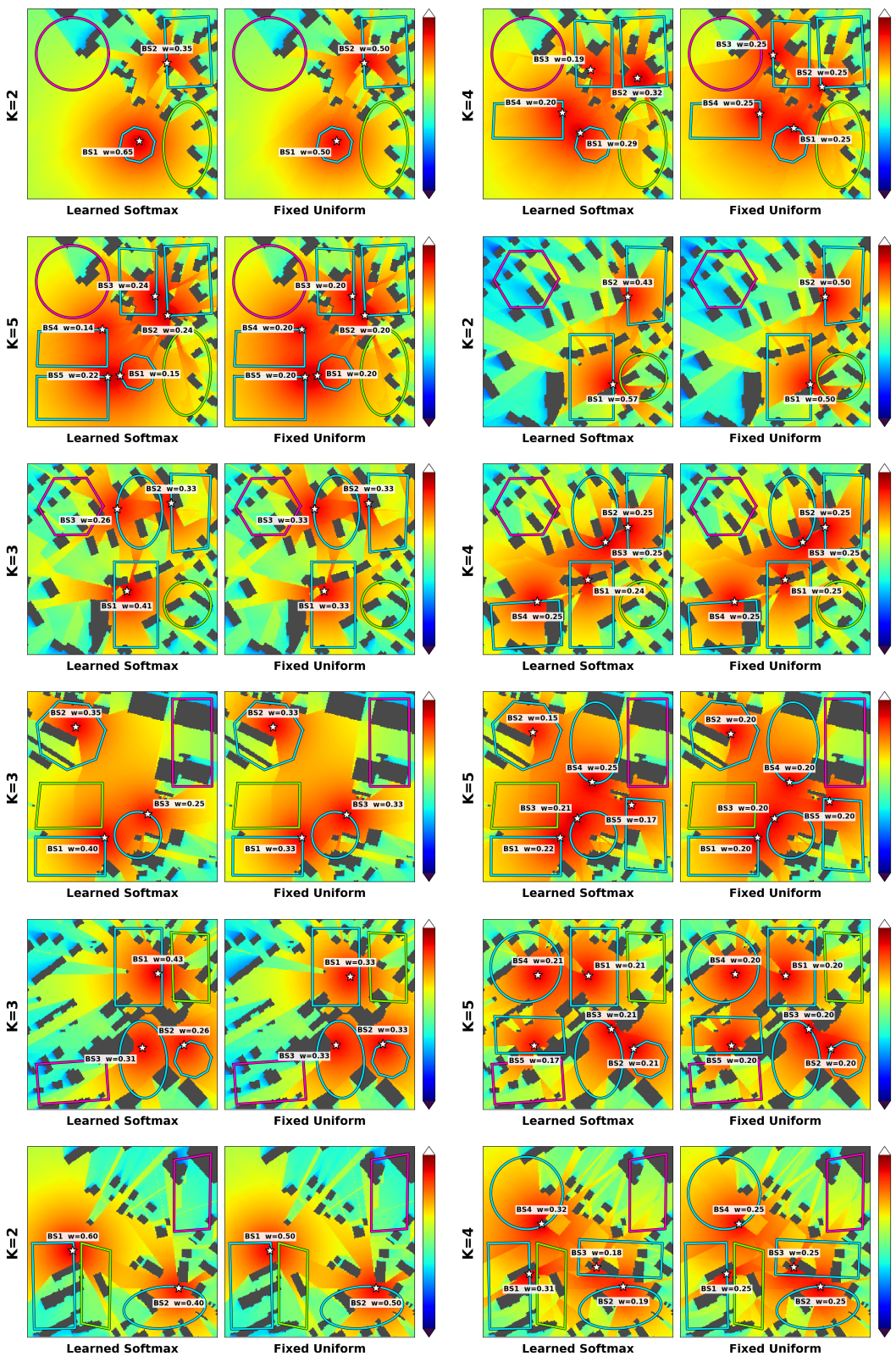
Auto-K | Sionna Exhaustive | AutoBS | TR 38.901 Proxy, each method at the station count it fixes for itself; the header above each panel states that count, and the first eight rows are shown, which is the better half of the evaluated deployments under that ranking.





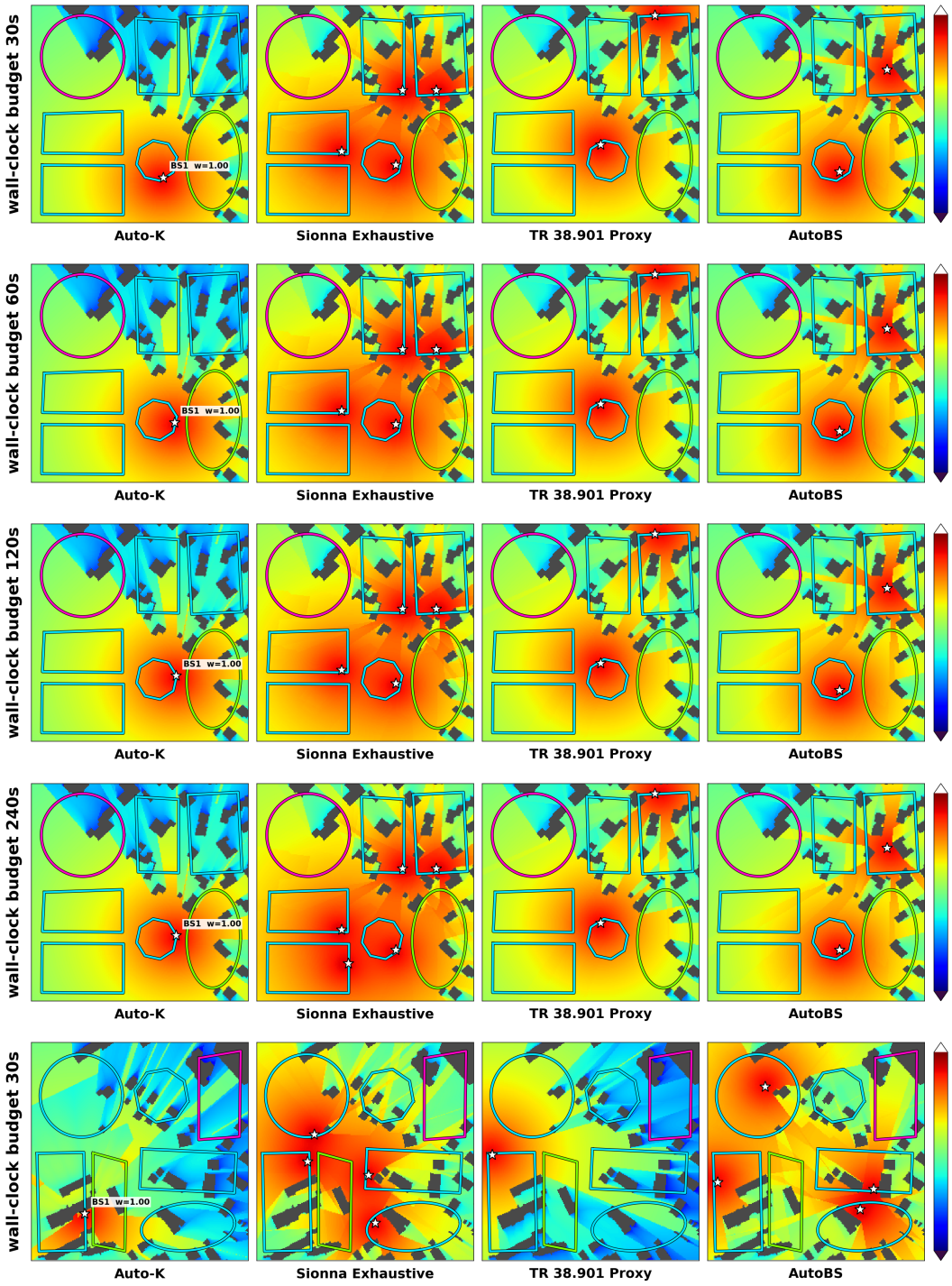
E.3 Weight ablation

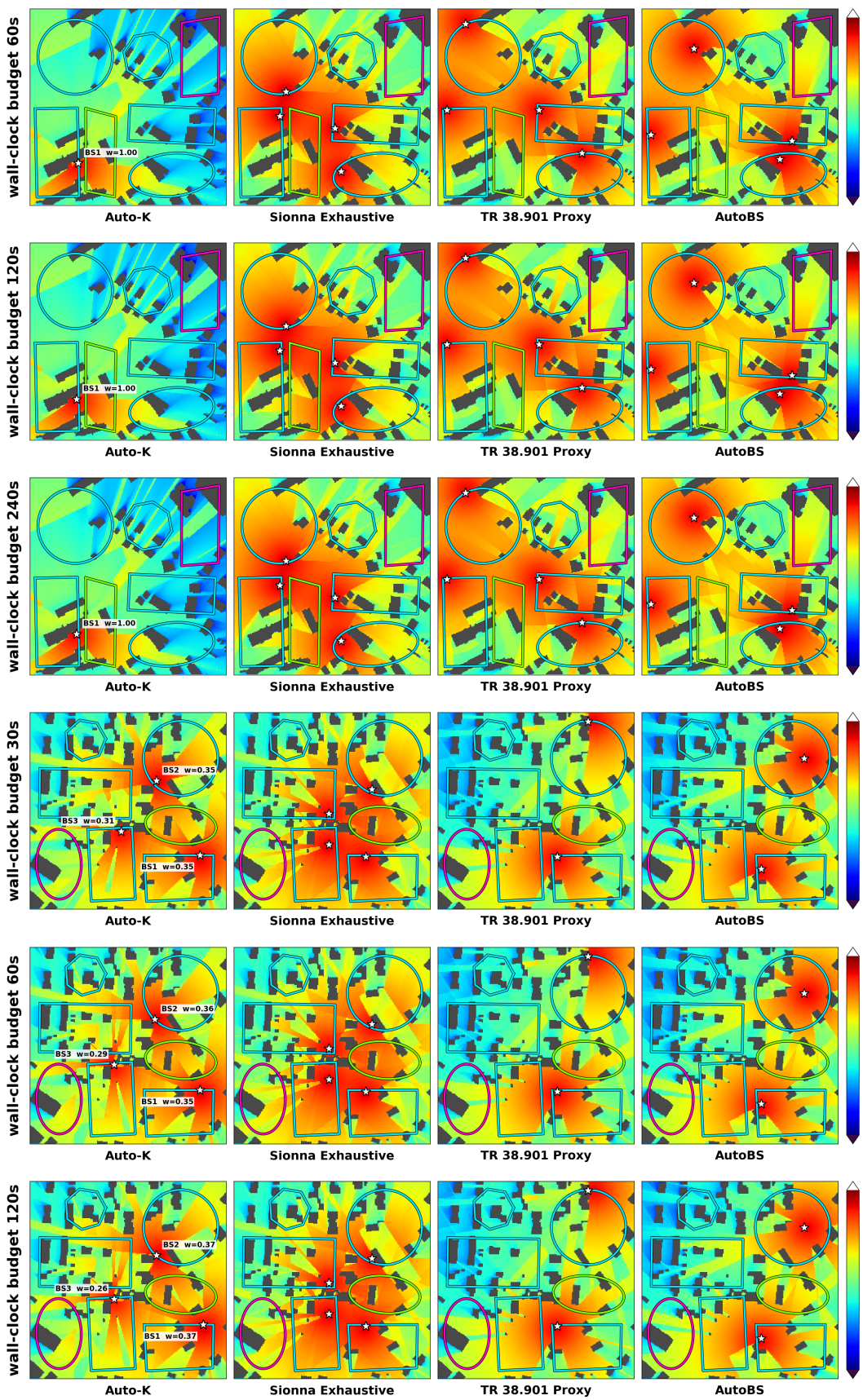
Learned Softmax | Fixed Uniform on the ten-scenario subset, of which the first twelve rows are shown as six pairs, which is the better half of the evaluated deployments under that ranking.

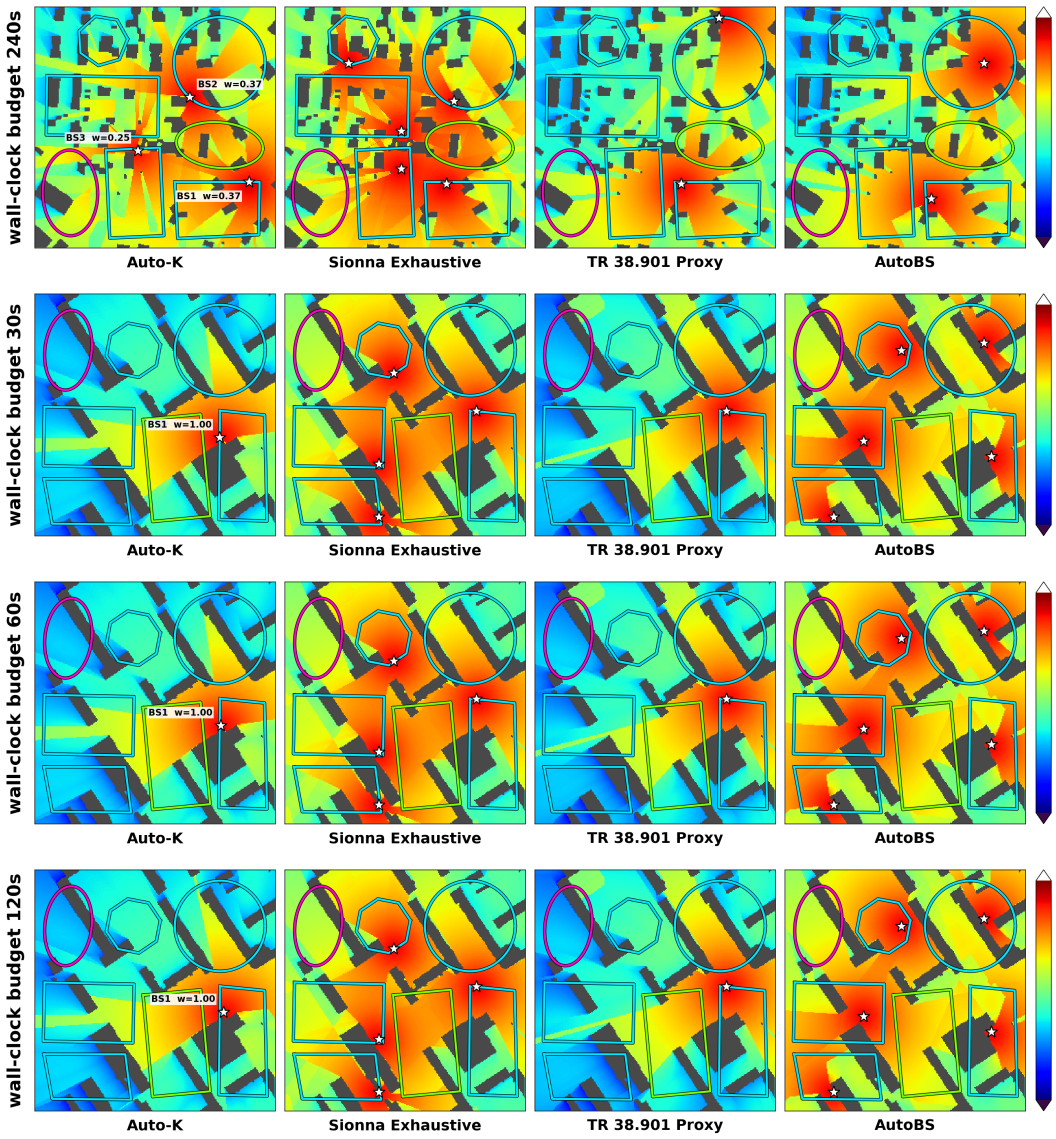

K=2
BS2 w=0.35
BS1 w=0.65
BS2 w=0.50
BS1 w=0.50
Learned Softmax
Fixed Uniform
K=4
BS3 w=0.19
BS2 w=0.32
BS4 w=0.20
BS1 w=0.29
BS3 w=0.25
BS2 w=0.25
BS4 w=0.25
BS1 w=0.25
Learned Softmax
Fixed Uniform
K=5
BS3 w=0.24
BS4 w=0.14
BS2 w=0.24
BS5 w=0.22
BS1 w=0.15
BS3 w=0.20
BS4 w=0.20
BS2 w=0.20
BS5 w=0.20
BS1 w=0.20
Learned Softmax
Fixed Uniform
K=2
BS2 w=0.43
BS1 w=0.57
BS2 w=0.50
BS1 w=0.50
Learned Softmax
Fixed Uniform
K=3
BS2 w=0.33
BS3 w=0.26
BS1 w=0.41
BS2 w=0.33
BS3 w=0.33
BS1 w=0.33
Learned Softmax
Fixed Uniform
K=4
BS2 w=0.25
BS3 w=0.25
BS1 w=0.24
BS4 w=0.25
BS2 w=0.25
BS3 w=0.25
BS1 w=0.25
BS4 w=0.25
Learned Softmax
Fixed Uniform
K=3
BS2 w=0.35
BS3 w=0.25
BS1 w=0.40
BS2 w=0.33
BS3 w=0.33
BS1 w=0.33
Learned Softmax
Fixed Uniform
K=5
BS2 w=0.15
BS4 w=0.25
BS3 w=0.21
BS5 w=0.17
BS1 w=0.22
BS2 w=0.20
BS4 w=0.20
BS3 w=0.20
BS5 w=0.20
BS1 w=0.20
Learned Softmax
Fixed Uniform
K=3
BS1 w=0.43
BS2 w=0.26
BS3 w=0.31
BS1 w=0.33
BS2 w=0.33
BS3 w=0.33
Learned Softmax
Fixed Uniform
K=5
BS4 w=0.21
BS1 w=0.21
BS3 w=0.21
BS5 w=0.17
BS2 w=0.21
BS4 w=0.20
BS1 w=0.20
BS3 w=0.20
BS5 w=0.20
BS2 w=0.20
Learned Softmax
Fixed Uniform
K=2
BS1 w=0.60
BS2 w=0.40
BS1 w=0.50
BS2 w=0.50
Learned Softmax
Fixed Uniform
K=4
BS4 w=0.32
BS3 w=0.18
BS1 w=0.31
BS2 w=0.19
BS4 w=0.25
BS3 w=0.25
BS1 w=0.25
BS2 w=0.25
Learned Softmax
Fixed Uniform

E.4 WALL-CLOCK BUDGET CURVES

Auto-K | Sionna Exhaustive | TR 38.901 Proxy | AutoBS at the four wall-clock cuts of the fifteen-scenario subset, of which the first fifteen rows are shown, one per map with the cuts in ascending order of budget, which is the better half of the evaluated deployments under that ranking.




wall-clock budget 60s
wall-clock budget 120s
wall-clock budget 240s
wall-clock budget 30s
wall-clock budget 60s
wall-clock budget 120s
BS1 w=1.00
BS2 w=0.35
BS3 w=0.31
BS1 w=0.35
BS2 w=0.36
BS3 w=0.29
BS1 w=0.35
BS2 w=0.37
BS3 w=0.26
BS1 w=0.37
Auto-K
Sionna Exhaustive
TR 38.901 Proxy
AutoBS



F PER-SCENARIO RESULTS

Tables 17 and 18 list every deployment behind the fixed-K and weight-ablation tables, read from the same scored records as those aggregates. The Auto-K comparison is not repeated here. Scenario identifiers name the map, and the count column gives the station count of that deployment.

Table 17: Fixed-K per-scenario results: thirty maps $\times$ $K = 1, \ldots, 5$.

Scenario	K	score↑	attain.↑	margin [dB]↑	S%↑	W%↑	Cov%↑
group_01	1	0.4686	0.6236	9.01	77.78	26.25	66.44
group_01	2	0.4728	0.6353	8.17	94.22	9.41	80.31
group_01	3	0.3708	0.5155	1.46	87.98	1.69	87.20
group_01	4	0.4092	0.5628	3.58	93.09	1.12	90.77
group_01	5	0.3958	0.5398	2.33	93.80	0.34	90.05
group_02	1	0.3205	0.4355	12.08	44.71	12.98	26.61
group_02	2	0.3325	0.4750	15.86	39.54	6.91	34.54
group_02	3	0.2942	0.4363	9.73	48.15	0.98	49.53
group_02	4	0.1215	0.2323	1.62	15.40	0.29	58.21
group_02	5	0.2651	0.4135	5.16	47.80	0.00	70.09

Scenario	K	score↑	attain.↑	margin [dB]↑	S%↑	W%↑	Cov%↑
group_03	1	0.6858	0.7508	24.77	99.96	23.97	66.34
group_03	2	0.5454	0.7374	15.81	100.00	0.93	84.54
group_03	3	0.5564	0.7477	16.60	100.00	0.65	88.53
group_03	4	0.5368	0.7500	14.39	100.00	0.00	92.91
group_03	5	0.5224	0.7500	12.65	100.00	0.00	95.86
group_04	1	0.5095	0.5751	9.07	52.40	75.19	36.07
group_04	2	0.3536	0.4817	5.34	65.28	13.86	62.82
group_04	3	0.3602	0.4949	5.86	66.76	12.01	65.32
group_04	4	0.4272	0.5687	7.57	85.78	6.59	74.45
group_04	5	0.3096	0.4663	1.75	69.60	0.76	80.95
group_05	1	0.7033	0.7620	18.95	61.27	97.91	36.29
group_05	2	0.5387	0.6416	19.55	85.56	14.16	43.97
group_05	3	0.4971	0.6291	15.87	93.31	5.59	41.80
group_05	4	0.5404	0.6837	17.76	96.58	3.71	61.96
group_05	5	0.2405	0.3554	-1.96	54.68	0.19	75.91
group_06	1	0.4771	0.5690	10.27	50.58	61.13	39.29
group_06	2	0.4134	0.5391	5.70	89.73	4.54	73.61
group_06	3	0.3938	0.5271	4.06	88.10	2.92	80.07
group_06	4	0.3610	0.4995	-0.92	90.26	0.20	89.59
group_06	5	0.4048	0.5185	1.48	98.32	0.00	95.01
group_07	1	0.3076	0.3484	7.88	65.81	0.06	31.21
group_07	2	0.4498	0.5255	12.20	90.78	0.06	55.20
group_07	3	0.4405	0.5390	9.26	92.41	0.00	70.72
group_07	4	0.3830	0.4876	4.35	89.80	0.00	70.08
group_07	5	0.3686	0.4780	1.09	93.19	0.00	78.17
group_08	1	0.5335	0.5700	17.86	16.04	99.85	7.18
group_08	2	0.4503	0.5705	16.60	36.19	53.92	14.02
group_08	3	0.1785	0.3201	7.12	19.21	10.76	17.45
group_08	4	0.1580	0.2646	3.92	27.27	4.13	24.62
group_08	5	0.4544	0.6160	13.35	90.74	3.65	37.67
group_09	1	0.5260	0.5596	10.28	80.26	51.65	27.51
group_09	2	0.4456	0.5374	8.31	98.17	0.94	65.11
group_09	3	0.4191	0.5149	4.28	99.95	0.15	75.97
group_09	4	0.4575	0.5823	7.08	99.60	0.04	87.72
group_09	5	0.4888	0.6264	10.11	99.01	0.00	90.66
group_10	1	0.4458	0.4951	18.76	73.63	0.14	36.92
group_10	2	0.4900	0.5681	17.87	85.55	0.00	54.62
group_10	3	0.3379	0.4453	5.73	70.45	0.00	69.24
group_10	4	0.2957	0.4279	0.46	69.36	0.00	83.01
group_10	5	0.3354	0.4698	0.97	80.02	0.00	85.62
group_11	1	0.4927	0.5587	3.72	100.00	33.33	55.47
group_11	2	0.4669	0.5869	3.61	100.00	15.61	84.02
group_11	3	0.4334	0.5513	2.83	100.00	6.81	84.72
group_11	4	0.4608	0.5986	5.48	100.00	6.69	85.93
group_11	5	0.4540	0.5948	5.59	100.00	3.76	86.81
group_12	1	0.6051	0.6356	25.88	99.17	0.00	48.81
group_12	2	0.5326	0.6922	16.34	100.00	0.00	69.18
group_12	3	0.4916	0.6618	9.71	100.00	0.00	94.57
group_12	4	0.4512	0.5930	5.58	100.00	0.00	95.36
group_12	5	0.4838	0.6472	8.83	100.00	0.00	95.47
group_13	1	0.6535	0.7682	11.24	100.00	61.02	58.13
group_13	2	0.4883	0.6296	8.70	100.00	8.44	76.04
group_13	3	0.4966	0.6515	10.15	100.00	5.73	77.90
group_13	4	0.5190	0.6904	12.15	100.00	5.90	79.71
group_13	5	0.5019	0.6772	11.91	100.00	0.24	82.08
group_14	1	0.3720	0.4501	-3.23	100.00	0.00	72.03
group_14	2	0.4219	0.5447	3.94	100.00	0.00	82.45
group_14	3	0.4713	0.6294	8.68	100.00	0.00	84.47
group_14	4	0.4351	0.5715	4.84	100.00	0.00	86.72
group_14	5	0.4050	0.5237	1.62	100.00	0.00	88.82
group_15	1	0.5019	0.5831	3.72	100.00	33.55	64.04
group_15	2	0.4376	0.5681	5.77	100.00	0.00	79.90
group_15	3	0.4338	0.5700	4.20	100.00	0.00	91.82

Scenario	K	score↑	attain.↑	margin [dB]↑	S%↑	W%↑	Cov%↑
group_15	4	0.4679	0.6255	7.53	100.00	0.00	92.58
group_15	5	0.4590	0.6078	6.47	100.00	0.00	94.30
group_16	1	0.4714	0.5870	11.41	100.00	2.22	50.67
group_16	2	0.3974	0.4964	2.10	100.00	0.15	75.96
group_16	3	0.3863	0.4897	-4.32	100.00	0.00	86.30
group_16	4	0.3895	0.4985	-2.65	100.00	0.00	89.45
group_16	5	0.4480	0.5948	5.69	100.00	0.00	91.07
group_17	1	0.6931	0.7569	19.74	100.00	51.61	40.85
group_17	2	0.5800	0.6633	22.12	100.00	0.00	58.80
group_17	3	0.4789	0.6349	10.10	100.00	0.00	77.99
group_17	4	0.5093	0.6964	12.16	100.00	0.00	87.77
group_17	5	0.4944	0.6658	9.95	100.00	0.00	94.88
group_18	1	0.2281	0.3184	-16.27	56.76	0.00	57.85
group_18	2	0.3373	0.4440	-5.97	88.74	0.00	71.10
group_18	3	0.3822	0.4783	-1.05	100.00	0.00	82.19
group_18	4	0.4049	0.5219	1.75	100.00	0.00	87.36
group_18	5	0.4808	0.6471	9.47	100.00	0.00	86.15
group_19	1	0.8617	0.8502	25.50	100.00	90.23	36.07
group_19	2	0.5212	0.6716	15.95	100.00	0.00	61.76
group_19	3	0.5497	0.6964	17.91	100.00	0.00	70.68
group_19	4	0.4442	0.5887	5.58	100.00	0.00	88.50
group_19	5	0.4342	0.5655	3.93	100.00	0.00	94.92
group_20	1	0.1054	0.1465	-11.48	24.65	1.70	26.39
group_20	2	0.4542	0.5441	6.68	99.77	11.14	54.66
group_20	3	0.4296	0.5321	5.76	100.00	2.47	64.52
group_20	4	0.3739	0.4571	-0.96	99.77	0.00	74.55
group_20	5	0.4396	0.5597	6.78	100.00	0.46	70.34
group_21	1	0.0815	0.1244	-9.66	8.33	10.01	26.44
group_21	2	0.1082	0.1978	-11.09	15.76	2.75	52.69
group_21	3	0.2855	0.3555	-8.97	77.78	0.99	49.20
group_21	4	0.3507	0.4766	0.21	88.21	0.92	81.19
group_21	5	0.3851	0.4831	-2.39	100.00	0.61	83.32
group_22	1	0.6007	0.6572	9.61	93.99	66.83	22.09
group_22	2	0.5972	0.6733	16.65	100.00	34.16	28.22
group_22	3	0.4108	0.4908	5.12	100.00	1.80	54.14
group_22	4	0.4409	0.5386	6.88	100.00	4.72	57.92
group_22	5	0.4198	0.5018	4.75	100.00	5.05	57.13
group_23	1	0.8110	0.7978	22.56	100.00	89.13	18.08
group_23	2	0.4292	0.5099	8.92	100.00	0.00	40.04
group_23	3	0.4558	0.5853	8.77	100.00	0.00	68.09
group_23	4	0.5392	0.6934	16.96	100.00	0.00	69.60
group_23	5	0.4665	0.6162	8.68	100.00	0.00	79.76
group_24	1	0.3263	0.4077	5.74	78.28	0.00	34.03
group_24	2	0.2606	0.3391	-7.05	69.24	0.00	52.83
group_24	3	0.3640	0.4474	-11.95	97.66	0.00	71.05
group_24	4	0.3806	0.4740	-7.90	100.00	0.00	80.65
group_24	5	0.3887	0.4965	-5.77	100.00	0.00	88.72
group_25	1	0.4946	0.5496	5.72	75.83	61.08	26.65
group_25	2	0.4207	0.5470	10.62	90.42	0.00	43.21
group_25	3	0.4266	0.5379	6.54	98.92	0.00	64.43
group_25	4	0.3749	0.4591	-1.88	99.86	0.00	75.28
group_25	5	0.3769	0.4637	-6.57	100.00	0.00	76.93
group_26	1	0.0206	0.0571	-16.48	0.00	0.00	20.57
group_26	2	0.5051	0.6194	14.17	100.00	7.70	40.28
group_26	3	0.3792	0.4423	2.49	100.00	0.00	54.28
group_26	4	0.4456	0.5558	8.63	100.00	0.52	57.80
group_26	5	0.4616	0.5899	9.56	100.00	0.52	64.50
group_27	1	0.2121	0.2658	3.65	50.00	0.90	22.87
group_27	2	0.4165	0.5113	5.60	100.00	0.00	60.45
group_27	3	0.3666	0.4475	-3.71	98.51	0.00	71.11
group_27	4	0.3694	0.4427	-0.41	100.00	0.00	69.39
group_27	5	0.4162	0.5339	3.51	100.00	0.00	81.15
group_28	1	0.1930	0.2175	-7.77	52.97	4.35	20.97

Scenario	K	score↑	attain.↑	margin [dB]↑	S%↑	W%↑	Cov%↑
group_28	2	0.1541	0.2210	-7.59	37.09	0.18	42.31
group_28	3	0.3556	0.4045	-7.99	99.77	0.35	55.27
group_28	4	0.3584	0.4142	-3.61	99.77	0.00	59.12
group_28	5	0.4482	0.5675	8.44	100.00	0.06	63.63
group_29	1	0.3679	0.4386	-0.18	100.00	0.00	67.89
group_29	2	0.4057	0.5083	3.18	100.00	0.00	73.88
group_29	3	0.4085	0.5181	2.99	100.00	0.00	78.61
group_29	4	0.4081	0.5217	2.58	100.00	0.00	82.34
group_29	5	0.4465	0.5860	6.38	100.00	0.00	82.70
group_30	1	0.4841	0.6143	14.87	97.82	0.00	41.94
group_30	2	0.5371	0.6907	16.94	99.67	0.00	68.65
group_30	3	0.5707	0.7022	19.79	100.00	0.00	72.79
group_30	4	0.5707	0.7386	18.48	100.00	0.00	85.91
group_30	5	0.5034	0.6887	11.32	100.00	0.00	90.18

Table 18: Weight-ablation per-scenario results: ten maps $\times$ $K = 1, \ldots, 5$, both arms.

Arm	Scenario	K	score↑	attain.↑	margin [dB]↑	S%↑	W%↑	Cov%↑
Learned	group_01	1	0.4686	0.6236	9.01	77.78	26.25	66.44
Learned	group_01	2	0.4626	0.6447	9.47	87.01	9.29	78.98
Learned	group_01	3	0.4511	0.6193	7.41	93.34	4.23	84.26
Learned	group_01	4	0.3996	0.5489	2.73	92.66	1.22	90.67
Learned	group_01	5	0.4023	0.5328	1.90	96.78	0.41	91.71
Learned	group_02	1	0.4525	0.5398	17.50	46.72	39.83	17.77
Learned	group_02	2	0.3010	0.4394	13.26	40.28	4.61	33.72
Learned	group_02	3	0.1893	0.2703	0.38	45.19	0.05	49.79
Learned	group_02	4	0.2208	0.3438	2.45	43.63	0.00	65.42
Learned	group_02	5	0.1699	0.2734	-0.86	35.75	0.00	62.67
Learned	group_03	1	0.6543	0.7346	24.51	100.00	12.34	72.12
Learned	group_03	2	0.5903	0.7407	19.90	100.00	2.33	84.32
Learned	group_03	3	0.5263	0.7503	13.36	100.00	0.09	92.44
Learned	group_03	4	0.5546	0.7500	16.17	100.00	0.00	92.86
Learned	group_03	5	0.5228	0.7500	12.67	100.00	0.00	96.11
Learned	group_04	1	0.6540	0.7271	17.30	58.35	91.42	31.67
Learned	group_04	2	0.3056	0.4302	1.59	59.63	14.69	66.73
Learned	group_04	3	0.4842	0.6148	12.11	86.25	13.41	64.14
Learned	group_04	4	0.2847	0.4303	-2.62	65.56	3.34	77.99
Learned	group_04	5	0.3133	0.4702	0.90	73.44	0.66	81.98
Learned	group_05	1	0.4964	0.6228	15.20	33.30	75.67	17.53
Learned	group_05	2	0.3622	0.5196	13.25	52.98	7.17	49.32
Learned	group_05	3	0.3863	0.4891	6.16	87.38	3.34	52.58
Learned	group_05	4	0.4418	0.5863	9.23	91.68	2.99	65.48
Learned	group_05	5	0.4646	0.6070	8.66	99.20	2.04	74.30
Learned	group_06	1	0.3661	0.4603	7.46	42.70	39.49	45.04
Learned	group_06	2	0.2916	0.4091	0.30	67.86	3.70	73.91
Learned	group_06	3	0.2415	0.3720	-4.90	53.73	0.07	80.12
Learned	group_06	4	0.3053	0.4436	-2.36	72.69	0.13	86.86
Learned	group_06	5	0.3855	0.5085	0.68	95.03	0.00	93.63
Learned	group_07	1	0.3456	0.3866	9.99	69.54	1.70	31.96
Learned	group_07	2	0.4272	0.5028	10.83	88.41	0.03	53.57
Learned	group_07	3	0.4214	0.5140	7.89	91.97	0.00	66.63
Learned	group_07	4	0.4511	0.5364	7.86	99.25	0.00	74.80
Learned	group_07	5	0.3819	0.4992	4.44	86.89	0.00	76.83
Learned	group_08	1	0.4165	0.5548	12.37	10.81	84.26	7.58
Learned	group_08	2	0.6979	0.7452	23.44	86.19	62.22	18.27
Learned	group_08	3	0.1521	0.3261	8.82	0.69	13.14	22.46
Learned	group_08	4	0.2298	0.3934	7.45	38.68	4.36	26.16
Learned	group_08	5	0.0357	0.0901	-2.76	0.00	1.73	30.47
Learned	group_09	1	0.6929	0.7263	17.04	65.45	97.79	32.78
Learned	group_09	2	0.4749	0.5833	12.76	80.09	15.19	61.51

Arm	Scenario	K	score↑	attain.↑	margin [dB]↑	S%↑	W%↑	Cov%↑
Learned	group_09	3	0.4011	0.4961	2.54	98.47	1.58	75.59
Learned	group_09	4	0.4779	0.6028	9.45	99.63	0.00	84.48
Learned	group_09	5	0.4748	0.6131	9.05	97.80	0.00	90.89
Learned	group_10	1	0.2369	0.3107	8.79	34.38	0.00	45.82
Learned	group_10	2	0.4669	0.5389	16.21	84.58	0.00	51.06
Learned	group_10	3	0.3606	0.4652	6.00	77.37	0.00	68.48
Learned	group_10	4	0.3172	0.4403	1.61	74.19	0.00	78.52
Learned	group_10	5	0.4174	0.5426	7.02	88.57	0.00	81.47
Uniform	group_01	1	0.4686	0.6236	9.01	77.78	26.25	66.44
Uniform	group_01	2	0.4504	0.6264	8.95	86.52	8.02	77.27
Uniform	group_01	3	0.4549	0.6273	7.64	93.17	4.49	85.47
Uniform	group_01	4	0.3862	0.5209	1.26	94.53	0.00	89.96
Uniform	group_01	5	0.3982	0.5286	1.66	96.63	0.35	90.66
Uniform	group_02	1	0.4525	0.5398	17.50	46.72	39.83	17.77
Uniform	group_02	2	0.2929	0.4282	12.69	39.80	4.21	34.00
Uniform	group_02	3	0.1849	0.2624	-0.67	45.19	0.00	49.28
Uniform	group_02	4	0.2218	0.3455	2.54	43.63	0.00	65.50
Uniform	group_02	5	0.1799	0.2831	-0.77	38.97	0.00	62.95
Uniform	group_03	1	0.6543	0.7346	24.51	100.00	12.34	72.12
Uniform	group_03	2	0.5455	0.7377	15.75	100.00	1.21	84.37
Uniform	group_03	3	0.5218	0.7503	12.85	100.00	0.09	93.06
Uniform	group_03	4	0.5546	0.7500	16.17	100.00	0.00	92.86
Uniform	group_03	5	0.5121	0.7500	11.59	100.00	0.00	96.14
Uniform	group_04	1	0.6540	0.7271	17.30	58.35	91.42	31.67
Uniform	group_04	2	0.3222	0.4507	2.65	62.53	13.70	67.01
Uniform	group_04	3	0.4466	0.5910	10.43	78.74	14.03	63.92
Uniform	group_04	4	0.2892	0.4349	-1.18	66.98	3.41	77.98
Uniform	group_04	5	0.3059	0.4619	0.39	72.90	0.36	82.16
Uniform	group_05	1	0.4964	0.6228	15.20	33.30	75.67	17.53
Uniform	group_05	2	0.3404	0.4888	12.27	52.24	4.32	47.99
Uniform	group_05	3	0.3840	0.4850	5.80	87.70	3.28	53.03
Uniform	group_05	4	0.4213	0.5595	7.98	89.06	2.73	66.07
Uniform	group_05	5	0.4502	0.5864	7.17	99.17	1.51	76.38
Uniform	group_06	1	0.3661	0.4603	7.46	42.70	39.49	45.04
Uniform	group_06	2	0.1834	0.3001	-4.18	35.06	2.81	69.83
Uniform	group_06	3	0.2413	0.3709	-5.92	53.38	0.55	79.52
Uniform	group_06	4	0.2958	0.4353	-2.24	69.42	0.13	87.14
Uniform	group_06	5	0.3731	0.5012	0.09	92.85	0.00	93.64
Uniform	group_07	1	0.3456	0.3866	9.99	69.54	1.70	31.96
Uniform	group_07	2	0.4321	0.5091	11.26	88.34	0.06	54.37
Uniform	group_07	3	0.4215	0.5143	7.94	91.83	0.00	66.58
Uniform	group_07	4	0.4442	0.5305	7.35	98.71	0.00	74.52
Uniform	group_07	5	0.3862	0.5038	4.84	86.92	0.00	77.02
Uniform	group_08	1	0.4165	0.5548	12.37	10.81	84.26	7.58
Uniform	group_08	2	0.6765	0.7423	22.85	81.73	61.63	17.89
Uniform	group_08	3	0.1583	0.3355	9.06	1.37	13.80	22.16
Uniform	group_08	4	0.2389	0.3948	6.80	44.25	3.75	26.91
Uniform	group_08	5	0.0353	0.0889	-2.58	0.00	1.75	30.03
Uniform	group_09	1	0.6929	0.7263	17.04	65.45	97.79	32.78
Uniform	group_09	2	0.4874	0.5978	13.63	78.64	17.89	61.49
Uniform	group_09	3	0.3935	0.4737	-0.45	99.74	6.86	73.65
Uniform	group_09	4	0.4431	0.5612	5.67	99.97	0.00	86.50
Uniform	group_09	5	0.4696	0.5957	7.66	99.88	0.00	93.44
Uniform	group_10	1	0.2369	0.3107	8.79	34.38	0.00	45.82
Uniform	group_10	2	0.4556	0.5278	15.06	84.58	0.00	51.23
Uniform	group_10	3	0.3502	0.4576	5.27	75.85	0.00	69.93
Uniform	group_10	4	0.3141	0.4371	1.22	74.48	0.00	78.50
Uniform	group_10	5	0.4174	0.5426	7.02	88.57	0.00	81.47