

Traffic-Aware Network Average SNR Maximization for Generalized Pinching-Antenna Systems

Yanqing Xu[†], Xiu Yin Zhang^{§¶}, Trung Q. Duong^{*||}, and Tsung-Hui Chang[‡]

[†]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China

[§] School of Microelectronics, South China University of Technology, Guangzhou, China

[¶]Pazhou Laboratory, Guangzhou, China

^{*}Faculty of Engineering and Applied Science, Memorial University, St. John's, Canada

^{||}School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K.

[‡]School of Artificial Intelligence, The Chinese University of Hong Kong, Shenzhen, China

Email: xuyanqing@cuhk.edu.cn, eexyz@scut.edu.cn, tduong@mun.ca, tsunghui.chang@ieee.org

Abstract—This paper studies traffic-aware deployment for generalized pinching-antenna systems from a network-level perspective. Different from existing works that optimize antenna positions for specific instantaneous users, we aim to adapt generalized pinching antenna deployment to long-term spatial traffic distributions. To this end, the user traffic over the service region is modeled by a Gaussian-mixture hotspot distribution, and a traffic-weighted network-average signal-to-noise ratio (SNR) metric is adopted. Under a grid-based formulation, we derive the average-SNR expression and show that the resulting non-convex optimization problem decomposes into independent scalar subproblems across waveguides. Furthermore, by analyzing the hotspot-induced structure of the objective, we show that each component is unimodal and use this property to develop a candidate-based global optimization algorithm with low complexity. Numerical results verify that the proposed method matches exhaustive search while requiring much lower computational cost, and consistently outperforms benchmark schemes based on projected gradient, hotspot-center placement, and fixed antennas.

Index Terms—Generalized pinching antenna, network-average SNR maximization, traffic-aware deployment

I. INTRODUCTION

Generalized pinching-antenna systems have recently attracted growing attention as a promising flexible-antenna architecture for future wireless networks [1], [2]. By enabling the radiation point to move along a guided medium, such as a dielectric waveguide, a leaky coaxial cable, or other waveguiding structures, generalized pinching-antenna systems offer a new spatial degree of freedom beyond conventional fixed-aperture arrays, thereby facilitating dynamic adaptation of the radiating position to the communication environment [3], [4].

This spatial reconfigurability has stimulated a rapidly growing literature on generalized pinching-antenna systems. In the context of dielectric-waveguide implementations, several works have examined the role of in-waveguide attenuation and shown, for both LoS-dominant channels and probabilistic blockage models, that its impact is often modest in practical operating regimes [5], [6]. To combat the in-waveguide attenuation issue in long waveguide-based system, a segmented waveguide-enabled pinching-antenna system was proposed in [7]. On the communication-design side, early studies considered rate-oriented transmission for single-waveguide downlink systems

[8], while subsequent works expanded the scope to multiuser and multi-waveguide architectures, where jointly designing pinching-point locations and transmit strategies was shown to improve spectral and/or energy efficiency [9]–[12]. User multiplexing with pinching antennas has also been explored through NOMA-based schemes [13]–[15]. Beyond conventional communication objectives, pinching antennas have been incorporated into emerging frameworks such as integrated sensing and communication [16], [17]. More recently, their ability to reshape line-of-sight (LoS) connectivity has also motivated environment division multiple access, which further illustrates the promise of pinching antennas for geometry-aware multiuser transmission [18].

Despite this progress, the literature on generalized pinching antennas remains largely focused on link-oriented optimization, with antenna positions designed for specific user locations or instantaneous channel conditions. While appropriate for quasi-static settings, this approach is less suitable for practical networks with dynamic user distributions, especially in wide-area or hotspot-dominated environments. In such cases, the objective shifts from optimizing individual links to enhancing long-term spatial service quality [19]. This calls for a traffic-aware regional design framework for generalized pinching-antenna systems, in which antenna positions are adapted to the spatial traffic distribution across the service region rather than to individual users. This is especially relevant when user activity is uneven and concentrated in dominant hotspots, such as offices, meeting areas, production lines, or access corridors.

Motivated by the above observations, this paper investigates traffic-aware deployment optimization for dielectric-waveguide-based pinching-antenna systems, an important realization of the generalized pinching-antenna concept. The spatial user distribution is modeled by a Gaussian-mixture hotspot model, and the traffic-weighted network-average SNR is adopted to emphasize high-demand regions. The problem is challenging because the objective is non-concave and depends nonlinearly on antenna positions under random LoS/NLoS propagation. To address this challenge, we derive a closed-form expression for the traffic-weighted average SNR under a grid-based approximation and show that the objective is

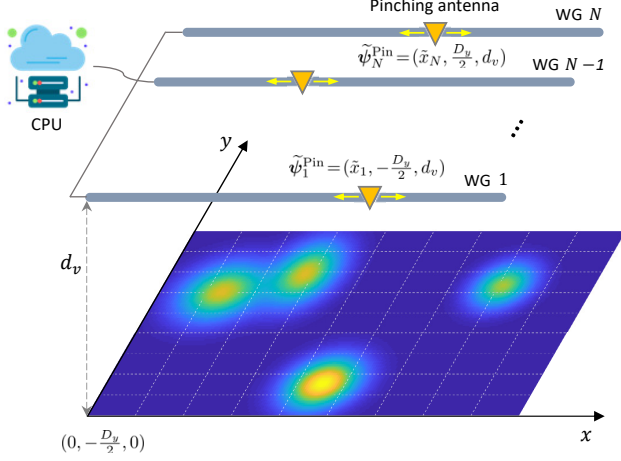


Fig. 1: Illustration of the considered geometry-aware pinching-antenna systems with multiple hotspots and waveguides.

separable across waveguides, reducing the problem to independent one-dimensional subproblems. Exploiting the Gaussian-mixture structure, we further establish the unimodality of each hotspot-induced component, which leads to a candidate-based low-complexity global optimization method. Numerical results show that the proposed method closely approaches exhaustive search with far lower complexity and significantly outperforms projected-gradient, hotspot-center, and fixed-antenna baselines.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a downlink traffic-aware transmission setting based on dielectric-waveguide pinching antennas. As illustrated in Fig. 1, the service area is a rectangle of size $D_x \times D_y$, above which N waveguides are installed at height d_v . These waveguides extend along the x -direction and are evenly distributed along the y -axis. Let $\mathcal{N} = \{1, \dots, N\}$ denote the waveguide set. The spacing between neighboring waveguides is chosen as $d_h = D_y/(N-1)$, so that the lateral position of waveguide n is $\tilde{y}_n = (n-1)d_h - \frac{D_y}{2}$, $n \in \mathcal{N}$. Each waveguide equips one pinching antenna. The location of the pinching antenna on waveguide n is written as $\psi_n^{\text{pin}} = [\tilde{x}_n, \tilde{y}_n, d_v]^T$, where $\tilde{x}_n \in [0, D_x]$ is the controllable horizontal coordinate. The deployment variables of all pinching antennas are collected in $\tilde{\mathbf{x}} = [\tilde{x}_1, \dots, \tilde{x}_N]^T$. All waveguides are coordinated by a common processor, which allows coherent transmission across antennas.

For a specific user located at $\psi = [x, y, 0]^T$ with $0 \leq x \leq D_x$ and $-\frac{D_y}{2} \leq y \leq \frac{D_y}{2}$. Given antenna deployment $\tilde{\mathbf{x}}$, the channel vector from the N pinching antennas to the user can be written as

$$\mathbf{h}(\psi, \tilde{\mathbf{x}}) = [h_1(\psi, \tilde{\mathbf{x}}), \dots, h_N(\psi, \tilde{\mathbf{x}})]^T. \quad (1)$$

The maximum-ratio transmission scheme is applied, the instantaneous received SNR is given by

$$\Gamma(\psi, \tilde{\mathbf{x}}) = \frac{P}{\sigma^2} \|\mathbf{h}(\psi, \tilde{\mathbf{x}})\|_2^2 \triangleq \rho \|\mathbf{h}(\psi, \tilde{\mathbf{x}})\|_2^2, \quad (2)$$

where P is the transmit power, σ^2 is the noise power, and $\rho \triangleq P/\sigma^2$.

A. Channel and Traffic Models

In this work, we model the channel from the n th pinching antenna to the user as the sum of a distance-dependent LoS term and an NLoS term with random LoS availability, namely,

$$h_n(\psi, \tilde{\mathbf{x}}) = \xi_n h_n^{\text{LoS}}(\psi, \tilde{\mathbf{x}}) + h_n^{\text{NLoS}}(\psi, \tilde{\mathbf{x}}), \quad (3)$$

where $\xi_n \in \{0, 1\}$ is a Bernoulli random variable indicating whether the LoS path exists. For a user at $\psi = [x, y, 0]^T$, define $r_n(\psi, \tilde{\mathbf{x}}) \triangleq \sqrt{(x - \tilde{x}_n)^2 + (y - \tilde{y}_n)^2 + d_v^2}$. Then, the LoS channel is modeled as

$$h_n^{\text{LoS}}(\psi, \tilde{\mathbf{x}}) = \frac{\sqrt{\eta} e^{-j\left(\frac{2\pi}{\lambda} r_n(\psi, \tilde{\mathbf{x}}) + \frac{2\pi}{\lambda_g} \tilde{x}_n\right)}}{r_n(\psi, \tilde{\mathbf{x}})}, \quad (4)$$

where λ and λ_g denote the free-space wavelength and guided wavelength, respectively, and $\eta = \frac{\lambda^2}{16\pi^2 f_c^2}$ is a free-space path loss constant with f_c denoting the working frequency.

The LoS existence probability is assumed to depend on the propagation distance and is modeled as

$$\Pr[\xi_n = 1] = p_{\text{LoS}}(r_n(\psi, \tilde{\mathbf{x}})) = e^{-\beta r_n^2(\psi, \tilde{\mathbf{x}})}, \quad (5)$$

where $\beta \geq 0$ is an environment-dependent blockage parameter. The NLoS component is modeled as

$$h_n^{\text{NLoS}}(\psi, \tilde{\mathbf{x}}) \sim \mathcal{CN}\left(0, \frac{\mu_n^2}{r_n^2(\psi, \tilde{\mathbf{x}})}\right), \quad (6)$$

where μ_n^2 denotes the average NLoS power coefficient.

To capture long-term spatial traffic heterogeneity, we model the location of a typical scheduled user by a Gaussian-mixture hotspot distribution [20]. Let $\mathbf{U} = [X, Y]^T$ denote the random horizontal position of the scheduled user. Its probability density function is given by

$$f_{\mathbf{U}}(x, y) = \sum_{\ell=1}^L \alpha_{\ell} \mathcal{N}([x, y]^T; \boldsymbol{\mu}_{\ell}, \boldsymbol{\Sigma}_{\ell}), \quad (7)$$

where L is the number of hotspots, $\alpha_{\ell} \geq 0$ is the traffic weight of hotspot ℓ with $\sum_{\ell=1}^L \alpha_{\ell} = 1$, and $\boldsymbol{\mu}_{\ell}$ and $\boldsymbol{\Sigma}_{\ell}$ are the mean vector and covariance matrix of hotspot ℓ , respectively.

B. Traffic-Weighted Network-Average SNR

Given a deployment $\tilde{\mathbf{x}}$, we first define the local average SNR at location $(x, y, 0)$ as

$$\bar{\Gamma}(x, y; \tilde{\mathbf{x}}) \triangleq \mathbb{E}[\Gamma(\psi, \tilde{\mathbf{x}}) | \psi = [x, y, 0]^T], \quad (8)$$

where the expectation is taken over the random LoS/NLoS channel realization.

We adopt the traffic-weighted network-average SNR as the network-level performance metric, i.e.,

$$\bar{\Gamma}_{\text{net}}(\tilde{\mathbf{x}}) \triangleq \iint \bar{\Gamma}(x, y; \tilde{\mathbf{x}}) f_{\mathbf{U}}(x, y) dx dy. \quad (9)$$

This metric characterizes the long-term average SNR experienced by a typical active user whose position follows the traffic map in (7). Hence, it emphasizes regions with high traffic density and is well suited for traffic-aware network-level pinching antenna deployment.

C. Grid-Based Approximation and Problem Formulation

To enable tractable optimization, the service region is discretized into N_h grids along the x -direction and N_v grids along the y -direction. Let $\Delta_x = \frac{D_x}{N_h}$, $\Delta_y = \frac{D_y}{N_v}$ denote the grid sizes. The center of grid (u, v) is given by

$$x_u = \left(u - \frac{1}{2}\right) \Delta_x, \quad y_v = -\frac{D_y}{2} + \left(v - \frac{1}{2}\right) \Delta_y, \quad (10)$$

for $u = 1, \dots, N_h$ and $v = 1, \dots, N_v$. The corresponding representative position is denoted by $\psi_{u,v} = [x_u, y_v, 0]^T$. The traffic probability of grid (u, v) is approximated by

$$p_{u,v} \approx f_U(x_u, y_v) \Delta_x \Delta_y, \quad (11)$$

followed by normalization such that $\sum_{u=1}^{N_h} \sum_{v=1}^{N_v} p_{u,v} = 1$.

For grid (u, v) , define the average SNR as $\bar{\Gamma}_{u,v}(\tilde{\mathbf{x}}) \triangleq \mathbb{E}[\Gamma(\psi_{u,v}, \tilde{\mathbf{x}})]$. Then, the continuous metric in (9) is approximated by

$$\bar{\Gamma}_{\text{net}}(\tilde{\mathbf{x}}) \approx \sum_{u=1}^{N_h} \sum_{v=1}^{N_v} p_{u,v} \bar{\Gamma}_{u,v}(\tilde{\mathbf{x}}). \quad (12)$$

Accordingly, the traffic-aware network-level design problem is formulated as

$$\max_{\tilde{\mathbf{x}}} \sum_{u=1}^{N_h} \sum_{v=1}^{N_v} p_{u,v} \bar{\Gamma}_{u,v}(\tilde{\mathbf{x}}) \quad (13a)$$

$$\text{s.t.} \quad 0 \leq \tilde{x}_n \leq D_x, \quad \forall n \in \mathcal{N}. \quad (13b)$$

Problem (13a) is non-convex due to the coupled distance-dependent channel terms and the traffic-weighted summation over all grids. In the next section, we show that by exploiting the structure of (12), the problem can be decomposed into independent one-dimensional subproblems, which further enables a candidate-based low-complexity global optimization method.

III. STRUCTURE ANALYSIS AND CANDIDATE-BASED GLOBAL OPTIMIZATION

In this section, we first derive an explicit expression for the traffic-weighted network-average SNR and reveal its separable structure across waveguides. Then, by exploiting the Gaussian-mixture traffic model, we further characterize the shape of each hotspot-induced component and develop a candidate-based low-complexity global optimization method.

A. Closed-Form Objective and Problem Decomposition

Recall that the instantaneous SNR at grid (u, v) is given by

$$\Gamma_{u,v}(\tilde{\mathbf{x}}) = \rho \sum_{n=1}^N |h_n(\psi_{u,v}, \tilde{\mathbf{x}})|^2. \quad (14)$$

By averaging over the random LoS/NLoS channel realization, we obtain the following result.

Lemma 1. For a given antenna deployment $\tilde{\mathbf{x}}$, the average SNR at grid (u, v) is

$$\bar{\Gamma}_{u,v}(\tilde{\mathbf{x}}) = \rho \sum_{n=1}^N \frac{\eta e^{-\beta r_{n,u,v}^2(\tilde{x}_n)} + \mu_n^2}{r_{n,u,v}^2(\tilde{x}_n)}, \quad (15)$$

where $r_{n,u,v}(\tilde{x}_n) \triangleq \sqrt{(x_u - \tilde{x}_n)^2 + (y_v - \tilde{y}_n)^2 + d_v^2}$. Accordingly, the traffic-weighted network-average SNR can be written as

$$\bar{\Gamma}_{\text{net}}(\tilde{\mathbf{x}}) = \rho \sum_{u=1}^{N_h} \sum_{v=1}^{N_v} p_{u,v} \sum_{n=1}^N \frac{\eta e^{-\beta r_{n,u,v}^2(\tilde{x}_n)} + \mu_n^2}{r_{n,u,v}^2(\tilde{x}_n)}. \quad (16)$$

Proof: The result follows by substituting the LoS/NLoS channel model into the instantaneous SNR expression and taking expectation over the Bernoulli LoS indicator and the Gaussian NLoS fading. The detailed proof is presented in the full version of this paper [21]. ■

Based on (16), the traffic-aware pinching-antenna deployment problem can be formulated as

$$\max_{\tilde{\mathbf{x}}} \bar{\Gamma}_{\text{net}}(\tilde{\mathbf{x}}) \quad (17a)$$

$$\text{s.t.} \quad 0 \leq \tilde{x}_n \leq D_x, \quad \forall n \in \mathcal{N}. \quad (17b)$$

Although problem (17a) is generally non-convex, its objective admits a favorable separable structure, as shown below.

Lemma 2. The objective in (17a) can be decomposed as

$$\bar{\Gamma}_{\text{net}}(\tilde{\mathbf{x}}) = \sum_{n=1}^N f_n(\tilde{x}_n), \quad (18)$$

where

$$f_n(\tilde{x}_n) \triangleq \rho \sum_{u=1}^{N_h} \sum_{v=1}^{N_v} p_{u,v} \frac{\eta e^{-\beta r_{n,u,v}^2(\tilde{x}_n)} + \mu_n^2}{r_{n,u,v}^2(\tilde{x}_n)}. \quad (19)$$

Hence, problem (17a) is equivalent to solving N independent scalar subproblems

$$\max_{0 \leq \tilde{x}_n \leq D_x} f_n(\tilde{x}_n), \quad \forall n \in \mathcal{N}. \quad (20)$$

Proof: The result follows directly by interchanging the order of summation in (16). ■

Lemma 2 shows that the original N -dimensional deployment problem reduces to N scalar maximization problems. However, each $f_n(\tilde{x}_n)$ is still non-concave in general, since it is a weighted superposition of grid-dependent terms centered at different traffic locations. Next, we exploit the hotspot structure of the traffic model to further reveal the geometry of $f_n(\tilde{x}_n)$.

B. Hotspot-Induced Structure

To expose the structure more clearly, we temporarily use the continuous traffic model. For waveguide n , define

$$\tilde{f}_n(\tilde{x}_n) \triangleq \iint_{\mathcal{D}} \bar{\Gamma}_n(\xi, \zeta; \tilde{x}_n) f_U(\xi, \zeta) d\xi d\zeta, \quad (21)$$

where $\mathcal{D} = [0, D_x] \times [-D_y/2, D_y/2]$ and

$$\bar{\Gamma}_n(\xi, \zeta; \tilde{x}_n) \triangleq \rho \frac{\eta e^{-\beta r_n^2(\xi, \zeta; \tilde{x}_n)} + \mu_n^2}{r_n^2(\xi, \zeta; \tilde{x}_n)}, \quad (22)$$

with $r_n(\xi, \zeta; \tilde{x}_n) \triangleq \sqrt{(\xi - \tilde{x}_n)^2 + (\zeta - \tilde{y}_n)^2 + d_v^2}$. Substituting (7) into (21) yields

$$\tilde{f}_n(\tilde{x}_n) = \sum_{\ell=1}^L \alpha_\ell F_{n,\ell}(\tilde{x}_n), \quad (23)$$

where

$$F_{n,\ell}(\tilde{x}_n) \triangleq \iint_{\mathcal{D}} \bar{\Gamma}_n(\xi, \zeta; \tilde{x}_n) \mathcal{N}([\xi, \zeta]^T; \boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell) d\xi d\zeta. \quad (24)$$

Hence, the per-waveguide objective is a traffic-weighted superposition of hotspot-induced component functions.

Lemma 3. *Assume that each hotspot is sufficiently away from the service-region boundary so that truncation effects are negligible. Then, for any waveguide n and hotspot ℓ , the function $F_{n,\ell}(\tilde{x}_n)$ is smooth and has a unique peak on $[0, D_x]$, attained at $\tilde{x}_n = \mu_{\ell,x}$, where $\mu_{\ell,x}$ is the horizontal center of hotspot ℓ .*

Proof: Since $\bar{\Gamma}_n(\xi, \zeta; \tilde{x}_n)$ depends on $\tilde{x}_n$ only through $(\xi - \tilde{x}_n)^2$, averaging it over the Gaussian hotspot centered at μ_ℓ, x yields a function symmetric with respect to $\tilde{x}_n = \mu_{\ell,x}$. In addition, the average SNR decreases with the squared distance, which implies that $F_{n,\ell}(\tilde{x}_n)$ is unimodal. The detailed proof is omitted for brevity. ■

Lemma 3 implies that each hotspot pulls the optimal pinching antenna position toward its own horizontal center. Therefore, multimodality of $f_n(\tilde{x}_n)$ arises not from each component individually, but from the superposition of multiple hotspot-induced unimodal components.

C. Candidate-Based Global Optimization

We first study the two-hotspot case, which already captures the essential interaction among hotspot-induced components and directly motivates the candidate construction used in the proposed algorithm.

1) *Two-Hotspot Case:* For $L = 2$, the per-waveguide objective can be written as

$$\tilde{f}_n(x) = \alpha_1 F_{n,1}(x) + \alpha_2 F_{n,2}(x), \quad (25)$$

where, without loss of generality, we assume that $\mu_{1,x} < \mu_{2,x}$. By Lemma 3, each component $F_{n,\ell}(x)$ is smooth and unimodal, and attains its unique maximum at $x = \mu_{\ell,x}$. Then, we have

$$\tilde{f}'_n(\mu_{1,x}) > 0, \quad \tilde{f}'_n(\mu_{2,x}) < 0. \quad (26)$$

Therefore, there exists at least one stationary point in the interval $(\mu_{1,x}, \mu_{2,x})$. The possible local profiles of $\tilde{f}_n(x)$ in $(\mu_{1,x}, \mu_{2,x})$ are characterized as follows.

Proposition 1. *Let $\mu_{1,x} < \mu_{2,x}$. Then, $\tilde{f}_n(x)$ must admit at least one stationary point in the open interval $(\mu_{1,x}, \mu_{2,x})$. Furthermore, the profile of $\tilde{f}_n(x)$ over this interval falls into one of the following patterns:*

- 1) *Merged case:* if there is exactly one stationary point in $(\mu_{1,x}, \mu_{2,x})$, then it is a local maximum, so the two hotspot contributions combine into a single peak between the two hotspot centers.
- 2) *Split case:* if there are three stationary points in $(\mu_{1,x}, \mu_{2,x})$, written as $x_L < x_M < x_R$, then x_L and x_R are local maxima and x_M is a local minimum; hence the objective exhibits two separated peaks.

Proof: Please see Appendix A. ■

TABLE I: Default simulation parameters.

Parameter	Value
Carrier frequency (f_c)	28 GHz
Communication-region size ($D_x \times D_y$)	$60 \times 200 \text{ m}^2$
Waveguide height (d_v)	10 m
Transmit power (P)	40 dBm
Noise power (σ^2)	-70 dBm
NLoS power (μ^2)	-60 dBm
Number of PAs / waveguides (N)	6
Number of grids ($N_h \times N_v$)	400×120
LoS blockage parameter (β)	0.01
Number of hotspots (L)	3

Proposition 1 indicates that, instead of exhaustively searching over $[0, D_x]$, it suffices to identify a small set of stationary candidates and compare their objective values.

2) *General Multi-Hotspot Case:* For $L > 2$, the per-waveguide objective remains

$$\tilde{f}_n(x) = \sum_{\ell=1}^L \alpha_\ell F_{n,\ell}(x), \quad (27)$$

where each $F_{n,\ell}(x)$ is unimodal with its maximum at $x = \mu_{\ell,x}$. Let $\mu_{(1),x} \leq \dots \leq \mu_{(L),x}$ denote the sorted hotspot centers. Then, the stationary-point search is restricted to the adjacent intervals $(\mu_{(k),x}, \mu_{(k+1),x})$, $k = 1, \dots, L-1$.

In implementation, we apply this idea to the discrete surrogate $f_n(x)$ in (19). In particular, its derivative is

$$f'_n(x) = 2\rho \sum_{u=1}^{N_h} \sum_{v=1}^{N_v} p_{u,v}(x_u - x) \times \frac{\eta e^{-\beta r_{n,u,v}^2(x)} (\beta r_{n,u,v}^2(x) + 1) + \mu_n^2}{r_{n,u,v}^4(x)}. \quad (28)$$

For each adjacent interval, coarse sampling is first used to detect sign changes of $f'_n(x)$, and each bracketed root is then refined by bisection. The stationary points corresponding to a positive-to-negative derivative change, together with the boundary points 0 and D_x , form the candidate set. The global maximizer is finally obtained by evaluating $f_n(x)$ over this set.

Compared with dense exhaustive search over $[0, D_x]$, the proposed method only evaluates a small set of structurally meaningful candidates. It therefore significantly reduces the computational cost while preserving global optimality with respect to the one-dimensional decomposed subproblems.

IV. SIMULATION RESULTS

This section evaluates the proposed traffic-aware pinching antenna deployment method. Unless otherwise specified, the simulation setup follows Table I. For each Monte Carlo trial, the hotspot centers are independently generated within the service area, the hotspot weights are randomly drawn and normalized, and all hotspots share the covariance matrix $\boldsymbol{\Sigma} = \text{diag}((0.15D_x)^2, (0.2D_y)^2)$. All results are averaged over 100 independent traffic realizations.

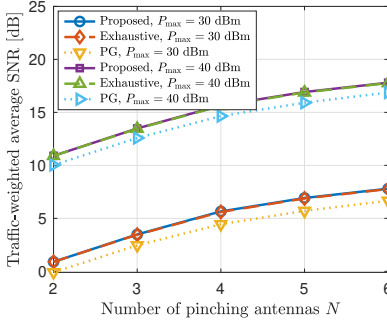


Fig. 2: Traffic-weighted network average-SNR performance comparison among the proposed algorithm with two baselines.

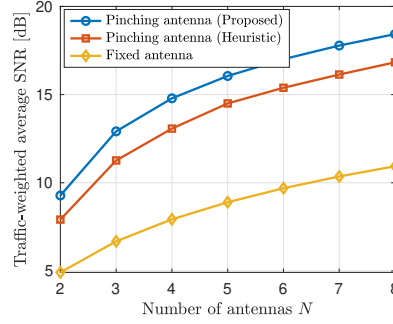


Fig. 3: Traffic-aware network average SNR comparisons between different schemes versus the number of antennas N .

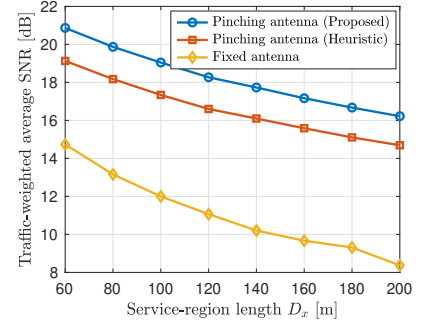


Fig. 4: Traffic-aware network average-SNR comparisons between different schemes versus the communication-region length D_x .

A. Optimality and Complexity of the Proposed Method

We first compare the proposed candidate-based method with two references in Fig. 2: an exhaustive one-dimensional search and a projected-gradient (PG) method. Owing to the separable structure of the objective, all methods solve N independent scalar subproblems. In the exhaustive benchmark, the interval $[0, D_x]$ is densely scanned with step size $\Delta_x = \lambda/5$. In the proposed method, the interval is first coarsely partitioned to identify sign changes of the derivative, and then each bracketed stationary point is refined by bisection. For the PG scheme, each scalar variable is iteratively updated using projected gradient steps under a prescribed iteration budget. Fig. 2 shows that the proposed method matches exhaustive search in traffic-weighted network-average SNR over the tested range of N for both $P = 30$ dBm and $P = 40$ dBm, confirming that it finds the global optimum of the decomposed one-dimensional subproblems in these experiments. By contrast, the PG method consistently attains lower values, as the non-convex problem may cause gradient-based iterations to converge to suboptimal stationary points. Moreover, higher transmit power shifts all curves upward, while more waveguides improve the network-average SNR with diminishing gains.

Table II reports the average runtime of the considered schemes. Several observations can be made. First, exhaustive search has by far the highest runtime, since it evaluates the objective over a dense set of search points and becomes even slower when the search resolution is refined from $\Delta_x = \lambda/5$ to $\Delta_x = \lambda/10$. Second, the proposed algorithm is consistently much faster than exhaustive search, because it only refines a small number of structurally meaningful candidates rather than scanning the entire interval. Third, PG has a runtime comparable to the proposed method for a small iteration budget, but its cost increases more noticeably when the maximum number of iterations is enlarged. Overall, the results verify that the proposed method achieves exhaustive-search-level performance with a much lower computational burden.

B. Performance Gain over Representative Baselines

Next, we evaluate the network-level performance improvement enabled by traffic-aware pinching-antenna deployment. In Fig. 3, the proposed scheme is benchmarked against two ref-

TABLE II: Execution time (in seconds) of the considered algorithms.

	$\Delta x = \lambda/5, J = 10, K_{\max} = 30$				
	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
Exhaustive	7.83	11.71	15.6	19.56	23.43
Proposed	0.034	0.048	0.059	0.073	0.089
PG	0.030	0.044	0.060	0.070	0.085
	$\Delta x = \lambda/10, J = 20, K_{\max} = 50$				
	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$
Exhaustive	15.61	23.35	30.88	38.66	46.61
Proposed	0.033	0.049	0.064	0.077	0.098
PG	0.044	0.072	0.096	0.130	0.159

erence designs. The first is a hotspot-center rule, which places each pinching antenna at the center of a hotspot and serves as a simple traffic-informed placement strategy. The second is a fixed-array baseline, where transmission is performed by a conventional N -element half-wavelength-spaced array located at $x = D_x/2$, corresponding to a non-reconfigurable and traffic-agnostic architecture.

From Fig. 3, the proposed design consistently provides the largest traffic-weighted network-average SNR for all tested values of N . This demonstrates the benefit of directly optimizing the antenna positions with respect to the network-level objective rather than relying on heuristic placement. The hotspot-center rule performs better than the fixed benchmark, indicating that incorporating traffic information is already beneficial even without solving the optimization problem explicitly. However, it still falls short of the proposed method because it is not derived by optimizing the full traffic-weighted objective. The fixed-array scheme performs worst, as its radiation locations remain fixed and cannot adjust to the hotspot pattern.

Fig. 4 examines the effect of the service-region length D_x . For all considered schemes, the network-average SNR drops as D_x increases, since the propagation distances become larger on average. Nevertheless, the proposed method remains the best-performing design throughout the entire range of D_x , which shows the robustness of traffic-aware pinching-antenna optimization under different coverage sizes. The hotspot-center baseline again outperforms the fixed-antenna benchmark, while still remaining inferior to the proposed design. Moreover, the

advantage of reconfigurable pinching-antenna deployment over the fixed benchmark becomes more pronounced for larger D_x , suggesting that position flexibility is particularly valuable in extended service regions.

V. CONCLUSION

This paper studied traffic-aware network-level deployment for dielectric-waveguide-based generalized pinching-antenna systems. By adopting a traffic-weighted network-average SNR metric, we optimized the pinching-antenna positions according to long-term spatial traffic statistics. We showed that the resulting problem is separable across waveguides, which reduces the original multi-variable optimization to independent one-dimensional subproblems. By further exploiting the Gaussian-mixture hotspot structure, we revealed the unimodality of each hotspot-induced component and developed a candidate-based low-complexity global optimization method. Simulation results showed that the proposed design matches exhaustive-search performance with much lower complexity and outperforms projected-gradient, hotspot-center, and fixed-antenna baselines. These results highlight the potential of traffic-aware pinching-antenna deployment for improving long-term network-level performance in spatially heterogeneous environments.

APPENDIX A PROOF OF PROPOSITION 1

Recall from Lemma 3 that each component $F_{n,\ell}(x)$ is differentiable and has a unique peak at $x = \mu_{\ell,x}$. Hence, for the two-hotspot case,

$$F'_{n,1}(\mu_{1,x}) = 0, F'_{n,2}(\mu_{2,x}) = 0, \quad (29)$$

and, under $\mu_{1,x} < \mu_{2,x}$, the monotonicity of the two unimodal functions further gives

$$F'_{n,2}(\mu_{1,x}) > 0, F'_{n,1}(\mu_{2,x}) < 0. \quad (30)$$

Since $\tilde{f}_n(x) = \alpha_1 F_{n,1}(x) + \alpha_2 F_{n,2}(x)$, its derivative at the two hotspot centers satisfies

$$\tilde{f}'_n(\mu_{1,x}) = \alpha_1 F'_{n,1}(\mu_{1,x}) + \alpha_2 F'_{n,2}(\mu_{1,x}) = \alpha_2 F'_{n,2}(\mu_{1,x}) > 0, \quad (31)$$

$$\tilde{f}'_n(\mu_{2,x}) = \alpha_1 F'_{n,1}(\mu_{2,x}) + \alpha_2 F'_{n,2}(\mu_{2,x}) = \alpha_1 F'_{n,1}(\mu_{2,x}) < 0. \quad (32)$$

Because $\tilde{f}_n(x)$ is continuous, there must exist at least one $x_b \in (\mu_{1,x}, \mu_{2,x})$ such that $\tilde{f}'_n(x_b) = 0$, which proves the existence of an interior stationary point. We next distinguish the possible local profiles.

If $\tilde{f}_n(x)$ has only one stationary point in $(\mu_{1,x}, \mu_{2,x})$, then the derivative is positive near $\mu_{1,x}$ and negative near $\mu_{2,x}$. Therefore, this stationary point must correspond to a change of $\tilde{f}'_n(x)$ from positive to negative, and thus it is a local maximizer by the first-derivative criterion. This is exactly the merged case.

If, instead, there are three stationary points in $(\mu_{1,x}, \mu_{2,x})$, denoted by $x_L < x_M < x_R$, then continuity implies that the derivative sign alternates across consecutive roots. Combined with $\tilde{f}'_n(\mu_{1,x}) > 0$ and $\tilde{f}'_n(\mu_{2,x}) < 0$, the sign pattern must be

$$+ \rightarrow - \text{ at } x_L, \quad - \rightarrow + \text{ at } x_M, \quad + \rightarrow - \text{ at } x_R.$$

Hence, x_L and x_R are local maximizers, while x_M is a local minimizer. This corresponds to the split case. The proposition is therefore proved. ■

REFERENCES

- [1] Y. Xu, J. Cui, Y. Zhu, Z. Ding, T.-H. Chang, R. Schober, V. W. Wong, O. A. Dobre, G. K. Karagiannidis, H. V. Poor, and X. You, "Generalized pinching-antenna systems: A tutorial on principles, design strategies, and future directions," *IEEE Commun. Surv. Tutorials*, 2026.
- [2] K. Wang, Z. Ding, and L. Hanzo, "Generalized pinching-antenna systems: A leaky-coaxial-cable perspective," 2025, Available online: arXiv:2512.04979.
- [3] Z. Ding, R. Schober, and H. V. Poor, "Flexible-antenna systems: A pinching-antenna perspective," *IEEE Trans. Commun.*, vol. 73, no. 10, pp. 9236–9253, 2025.
- [4] Z. Yang, N. Wang, Y. Sun, Z. Ding, R. Schober, G. K. Karagiannidis, V. W. Wong, and O. A. Dobre, "Pinching antennas: Principles, applications and challenges," 2025, arXiv preprint arXiv:2501.10753.
- [5] Y. Xu, Z. Ding, R. Schober, and T.-H. Chang, "Pinching-antenna systems with in-waveguide attenuation: Performance analysis and algorithm design," *IEEE Trans. Wireless Commun.*, early access, 2026.
- [6] Y. Xu, Z. Ding, O. A. Dobre, and T.-H. Chang, "Pinching-antenna system design with LoS blockage: Does in-waveguide attenuation matter?" *IEEE Trans. Wireless Commun.* (submitted), 2025, available online: arXiv:2508.07131.
- [7] C. Ouyang, H. Jiang, Z. Wang, Y. Liu, and Z. Ding, "Uplink and downlink communications in segmented waveguide-enabled pinching-antenna systems (SWANs)," *IEEE Trans. Commun.*, vol. 74, pp. 3688–3703, 2026.
- [8] Y. Xu, Z. Ding, and G. K. Karagiannidis, "Rate maximization for downlink pinching-antenna systems," *IEEE Wireless Commun. Lett.*, vol. 14, no. 5, pp. 1431–1435, May 2025.
- [9] Z. Ding and H. V. Poor, "Analytical optimization for antenna placement in pinching-antenna systems," submitted to *IEEE Trans. Wireless Commun.*, 2025, available online: arXiv:2507.13307.
- [10] A. Bereyhi, C. Ouyang, S. Asaad, Z. Ding, and H. V. Poor, "MIMO-PASS: Uplink and downlink transmission via MIMO pinching-antenna systems," *IEEE Trans. Commun.*, pp. 1–1, 2026.
- [11] L. Zhang, X. Mu, A. Liu, and Y. Liu, "Two-timescale joint transmit and pinching beamforming for pinching-antenna systems," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2025.
- [12] M. Qian, X. Mu, L. You, and M. Matthaiou, "Pinching-antenna-based communications: Spectral efficiency analysis and deployment strategies," *IEEE Trans. Commun.*, vol. 74, pp. 4758–4771, 2026.
- [13] P. Liu, Z. Fei, M. Hua, G. Chen, X. Wang, and R. Liu, "Wireless powered mec systems via discrete pinching antennas: Tdma versus noma," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 12 034–12 049, 2026.
- [14] Y. Xu, Z. Ding, D. Cai, and V. W. Wong, "QoS-aware NOMA design for downlink pinching-antenna systems," *IEEE Trans. Commun.*, vol. 73, no. 12, pp. 13 611–13 625, 2025.
- [15] B. Zhuo, J. Gu, W. Duan, M. Wen, L. Peng, and P.-H. Ho, "P-noma for pinching-antenna systems (PASS)," *IEEE Wireless Commun. Lett.*, vol. 15, pp. 131–135, 2026.
- [16] Z. Wang, C. Ouyang, Y. Liu, and A. Nallanathan, "Wireless sensing via pinching-antenna systems," *IEEE Wireless Commun. Lett.*, vol. 14, no. 11, pp. 3475–3479, 2025.
- [17] Y. Hu, T. Zhang, X. Xu, and Y. Liu, "Pinching antenna-enabled integrated sensing and communication for low-altitude UAV," *IEEE Trans. Veh. Technol.*, pp. 1–6, 2025.
- [18] Z. Ding, R. Schober, and H. V. Poor, "Environment division multiple access (EDMA): A feasibility study via pinching antennas," 2025, available online: arXiv:2511.03820.
- [19] L. Li, Y. Xu, Y. Xue, F. Yin, C. Shen, R. Zhang, and T.-H. Chang, "PEMNet: Towards autonomous and enhanced environment-aware mobile networks," *IEEE Commun. Mag.* (submitted), 2026, available online: arXiv:2601.11025.
- [20] Q. Zhang, W. Saad, M. Bennis, X. Lu, M. Debbah, and W. Zuo, "Predictive deployment of UAV base stations in wireless networks: Machine learning meets contract theory," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 637–652, Jan. 2021.
- [21] Y. Xu, Z. Ding, X. Y. Zhang, T. Q. Duong, and T.-H. Chang, "Environment-aware network-level design of generalized pinching-antenna systems—part I: Traffic-aware case," *arXiv preprint arXiv:2602.17023*, 2026.