

An Architecture for 3D UAV Positioning over 5G mmWave Networks

Bismark C. Teixeira, Diego A. Sousa, Daniel C. Araújo

Abstract—This paper presents an architecture for three-dimensional outdoor positioning of Unmanned Aerial Vehicles (UAVs) over a 5G Millimeter Wave (mmWave) network at 28 GHz, evaluated through a ray-tracing simulation of the University of Brasília School of Engineering Sciences and Technology (UnB-FCTE) campus generated with Sionna RT. Each Access Point (AP) performs local Direction of Arrival (DoA) estimation from a single Physical Resource Block (PRB) pilot snapshot and forwards only the B -bit quantized angles over the backhaul. The central processing unit then estimates the UAV position via the Extra Trees Regressor (ETR), trained offline on the ray-traced database. The architecture is benchmarked against two classical positioning baselines, angular Multilateration (MT) and an Extended Kalman Filter (EKF) with a 3D Gauss–Markov process model, on 400 Monte Carlo trajectories. Experiments sweep the four DoA algorithms: Delay and Sum (Bartlett), Minimum Variance Distortionless Response (Capon), Multiple Signal Classification (MUSIC) and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT). The evaluation also sweeps the number of active APs from 2 to 6 (each equipped with a Uniform Planar Array (UPA)), as well as the angular quantizer bit-depth from 4 to 16 bits. The proposed regressor reaches competitive accuracy at $N_{AP}^{act} \geq 4$ with lower per-slot complexity than the EKF baseline, while 8-bit quantization recovers the unquantized performance and a single PRB suffices for reliable pilot-based DoA estimation.

Keywords—UAV positioning, 5G mmWave, DoA estimation, Extra Trees, ray-tracing, 3GPP

I. INTRODUCTION

Three-dimensional (3D) localization of UAVs underpins Unmanned Aircraft System Traffic Management (UTM), collision avoidance and beyond-visual-line-of-sight operations in cellular-connected aerial networks [1], [2]. Global Navigation Satellite Systems (GNSSs) degrade in dense urban canyons, where multipath, blockage and Non-Line-of-Sight (NLOS) propagation corrupt the position fix. 5G infrastructure is a natural complement in this regime [1]: the Frequency Range 2 (FR2) and mmWave bands offer wide bandwidth and large antenna arrays at each AP [3], translating directly into sub-meter ranging and fine angular resolution. Single-anchor localization is feasible by jointly exploiting DoA, Angle of Departure (AoD) and Time-of-Arrival (ToA), including NLOS and obstructed-Line-of-Sight (LOS) regimes [4], [5].

B. C. Teixeira and D. C. Araújo are with the Telecommunications Laboratory (LabTelecom), University of Brasília (UnB), Brasília, DF 72444-240, Brazil (e-mail: bismark.teixeira@aluno.unb.br; daniel.araujo@unb.br).

D. A. Sousa is with the Federal Institute of Education, Science and Technology of Ceará (IFCE), Campus Paracuru, Paracuru, CE 62680-000, Brazil (e-mail: diego.sousa@ifce.edu.br).

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

DoA-based positioning in particular avoids the inter-AP time synchronization that ToA/Time-Difference-of-Arrival (TDoA) schemes require [6], which motivates this work.

DoA-based positioning at scale is backhaul-constrained: forwarding raw in-phase/quadrature (IQ) samples or full covariances to a central unit is prohibitively expensive, so each AP must estimate and quantize its own angle locally. This shapes the downstream positioning method: the three families that dominate DoA-based localization each break down under it differently. (i) Geometric methods (angular MT) solve the bearing-line intersection in closed form [7], but collapse when few APs leave the geometry ill- or under-determined. (ii) Bayesian sequential estimators (the EKF) couple a kinematic process model to nonlinear angular observations [8], but need enough visible APs to stay well-conditioned. Below that, they degrade to open-loop drift. (iii) Data-driven methods learn the angle-to-position mapping directly, predominantly via deep neural networks: Mousa and Al-Rubaye [9] report a Mean Absolute Error (MAE) of 1.3–1.7 m refining 5G ToA estimates in LOS/NLOS scenarios, with comparable gains for Channel State Information (CSI)-fingerprinting [10], but such models demand large datasets, costly inference and limited interpretability, conflicting with embedded UAV avionics.

None of these families, moreover, has been evaluated against a common front-end quantifying how many angle bits and active APs are needed before its assumptions break down, compounded by the fact that classical DoA estimators (Bartlett [11], Capon [12], MUSIC [13], ESPRIT [14]) already trade off resolution, cost and robustness before any quantization or AP-count constraint applies [15]–[17]. This paper addresses this joint constraint with a ray-traced 5G mmWave dataset of an outdoor university campus, contributing to:

- An architecture in which each AP estimates a local DoA from a single PRB pilot and forwards only B -bit quantized angles to a central unit running an ETR, keeping the backhaul payload at two scalars per slot.
- A benchmark of the proposed regressor against MT [7] and EKF [8] on a Sionna RT pipeline aligned with 3GPP TR 38.901, sweeping the four DoA estimators, N_{AP}^{act} and B , which identifies 8-bit quantization with $N_{AP}^{act} \geq 4$ as the recommended operating point.

II. SYSTEM MODEL AND PROBLEM STATEMENT

We consider an outdoor cellular-connected UAV positioning scenario where N_{AP} ground APs at known pose $\{\mathbf{p}_i, \mathbf{R}_i\}_{i=1}^{N_{AP}} \in \mathbb{R}^3 \times \text{SO}(3)$ cooperate to localize a single UAV transmitter from its 5G uplink pilots, as illustrated in Fig. 1. Each AP carries an $M_x \times M_y$ UPA of half-wavelength spacing $d = \lambda/2$ and

cross polarization. The UAV has a single vertically polarized isotropic element, which decouples orientation from positioning. Propagation is modeled by ray tracing (Sionna RT [18]) following 3GPP TR 38.901 [19].

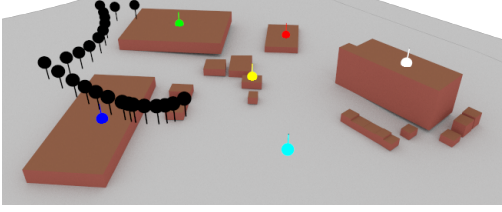


Fig. 1: Sionna RT scene of the UnB-FCTE campus with six APs (colored markers) and an example of UAV trajectory.

Throughout the paper: a single UAV transmits the positioning pilot per slot, the array response $\mathbf{a}_i(\mathbf{u}^g)$ is taken frequency-independent across the K subcarriers of the allocated PRB, path gains, delays and directions remain constant over the L pilot snapshots of the slot. The anchor poses $\{\mathbf{p}_i, \mathbf{R}_i\}$ are perfectly known at the central processing unit, the receiver noise is spatially white, $\mathbf{n}_{k,i,\ell} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$, independent across snapshots, APs, slots and each propagation path is a plane wave at the array, an assumption valid for array apertures much smaller than the UAV-AP distance.

Under the above assumptions, the uplink channel between the UAV at $\mathbf{p}_k$ and AP i at the ℓ -th pilot resource element of carrier frequency f_ℓ is the sum of $P_{k,i}$ paths returned by the ray tracer, $\mathbf{h}_{k,i,\ell} = \sum_{p=1}^{P_{k,i}} \beta_{k,i,p} \mathbf{a}_i(\mathbf{u}_{k,i,p}^g) e^{-j2\pi f_\ell \tau_{k,i,p}}$, where $\beta_{k,i,p} \in \mathbb{C}$ is the complex path gain, $\tau_{k,i,p}$ the propagation delay, $\mathbf{u}_{k,i,p}^g \in \mathbb{S}^2$ the unit propagation direction in the global frame and $\mathbf{a}_i(\cdot)$ the UPA steering vector of AP i .

Array-level angles are expressed in the AP's local frame via $\mathbf{R}_i$. With the global DoA defined by Zenith Angle of Arrival (ZoA)/Azimuth Angle of Arrival (AoA) (θ^g, ϕ^g) of a given path, the global direction reads

$$\mathbf{u}^g = [\sin \theta^g \cos \phi^g, \sin \theta^g \sin \phi^g, \cos \theta^g]^T, \quad (1)$$

and the corresponding local direction is $\mathbf{u} = \mathbf{R}_i^T \mathbf{u}^g$, from which the local angles (θ, ϕ) seen by the array processor follow as

$$\theta = \arccos([\mathbf{u}]_3), \quad \phi = \arctan 2([\mathbf{u}]_2, [\mathbf{u}]_1). \quad (2)$$

Letting $\mathbf{r}_{m,n} = d[m, n, 0]^T$, $m=0, \dots, M_x-1$, $n=0, \dots, M_y-1$, denote the UPA element positions in the local frame, the steering vector for a global direction $\mathbf{u}^g$ is

$$[\mathbf{a}_i(\mathbf{u}^g)]_{m,n} = \exp\left(j \frac{2\pi}{\lambda} (\mathbf{R}_i^T \mathbf{u}^g)^T \mathbf{r}_{m,n}\right), \quad (3)$$

written compactly as $a(\theta, \phi)$ in the local angles (2) used throughout Table I.

The uplink follows a generic 5G FR2 Orthogonal Frequency Division Multiplexing (OFDM) numerology with carrier frequency f_c , subcarrier spacing Δf (numerology index μ) and a single PRB of K subcarriers, giving pilot band occupation $B_{\text{pos}} = K \Delta f$, a small fraction of a typical FR2 carrier. Within each 14-symbol slot, N_p fixed-position OFDM symbols carry the pilot, while the remaining $14 - N_p$ are available for payload, giving overhead $N_p/14$. The pilots follow a Kronecker

pattern with $L = K N_p$ unimodular QPSK symbols s_ℓ on the K subcarriers of the N_p pilot symbols; since LS pilot removal divides by these known entries, the noise statistics match those of a Zadoff-Chu sequence. Since $L < 2M_x M_y$, $\hat{\mathbf{R}}_{k,i}$ is rank-deficient and a diagonal loading $10^{-5} \mathbf{I}$ is added before inversion. After pilot dewhitening, the received snapshot at AP i is

$$\mathbf{y}_{k,i,\ell} = \mathbf{h}_{k,i,\ell} + \mathbf{n}_{k,i,\ell}, \quad (4)$$

which is the spatial-array sufficient statistic processed by the DoA estimators of Section III.

The trajectory follows the 3D Gauss-Markov mobility model of IEEE 802.16m [20], with state $\mathbf{x}_k = [\mathbf{p}_k^T, \mathbf{v}_k^T]^T \in \mathbb{R}^6$ stacking position and velocity,

$$\mathbf{v}_{k+1} = \alpha \mathbf{v}_k + (1-\alpha) \bar{\mathbf{v}} + \sqrt{1-\alpha^2} \sigma_v \mathbf{w}_k, \quad (5)$$

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{v}_k \Delta t, \quad (6)$$

with driving noise $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_3)$, correlation coefficient $\alpha \in (0, 1)$, mean velocity $\bar{\mathbf{v}}$, velocity standard deviation σ_v and time step Δt .

A. Problem Definition

To emulate the finite-resolution DoA report that an embedded AP would forward to a positioning unit, every estimated angle is quantized in its sine/cosine representation, which is bounded and immune to 2π -wraparound artifacts. With B bits and step $\Delta = 2/(2^B - 1)$, the quantizer $Q_B(\cdot)$ maps

$$Q_B(\hat{\theta}) = \arctan 2(s_q, c_q), \quad \begin{aligned} s_q &= \text{round}\left(\frac{\sin \hat{\theta} + 1}{\Delta}\right) \Delta - 1, \\ c_q &= \text{round}\left(\frac{\cos \hat{\theta} + 1}{\Delta}\right) \Delta - 1. \end{aligned} \quad (7)$$

The same map is applied to both $\hat{\theta}$ (ZoA) and $\hat{\phi}$ (AoA) and to all four DoA estimators, ensuring a fair comparison for any chosen bit-depth B . Given the resulting set of B -bit quantized estimates $\{Q_B(\hat{\theta}_{k,i}), Q_B(\hat{\phi}_{k,i})\}_{i \in \mathcal{S}_k}$ from a subset $\mathcal{S}_k$ of $N_{\text{AP}}^{\text{act}} \leq N_{\text{AP}}$ active APs, equivalently the per-AP sample covariances $\{\hat{\mathbf{R}}_{k,i}\}$, the goal is to produce a 3D position estimate $\hat{\mathbf{p}}_k \in \mathbb{R}^3$ of the UAV along the trajectory (5)–(6). Performance is measured by the 3D MAE

$$\text{MAE} = \frac{1}{N_{\text{traj}}^{\text{test}} T_{\text{steps}}} \sum_{n=1}^{N_{\text{traj}}^{\text{test}}} \sum_{k=1}^{T_{\text{steps}}} \|\hat{\mathbf{p}}_k^{(n)} - \mathbf{p}_k^{(n)}\|_2, \quad (8)$$

evaluated over the $N_{\text{traj}}^{\text{test}}$ held-out test trajectories, with trajectory length T_{steps} , both defined in Section IV. The architecture proposed in Section III solves this problem with an ensemble regression-tree estimator at the central processing unit and is benchmarked in Section IV against two classical baselines operating on the same quantized angles: angular MT and an EKF with a 3D Gauss-Markov process model.

III. PROPOSED ARCHITECTURE

We propose the two-stage architecture sketched in Fig. 2.

Each AP forms the spatial sample covariance from the L pilot snapshots of Section II,

$$\hat{\mathbf{R}}_{k,i} = \frac{1}{L} \sum_{\ell=1}^L \mathbf{y}_{k,i,\ell} \mathbf{y}_{k,i,\ell}^H, \quad (9)$$

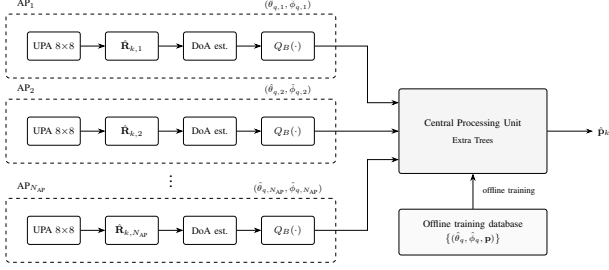


Fig. 2: Proposed architecture for estimating the 3D UAV position.

runs a DoA estimator on $\hat{\mathbf{R}}_{k,i}$ to obtain a ZoA/AoA pair $(\hat{\theta}_{k,i}, \hat{\phi}_{k,i})$, applies the inverse rotation $\mathbf{R}_i$ to express the angles globally and forwards their B -bit quantized counterparts $Q_B(\hat{\theta}_{k,i}), Q_B(\hat{\phi}_{k,i})$ over the backhaul. Shipping only the B -bit sine/cosine codes of two angles per AP per slot reduces the backhaul load from $\mathcal{O}(M_x M_y L)$, required by raw IQ samples or full covariance matrices, to $\mathcal{O}(B)$ bits per slot. The central processing unit hosts the ETR [21], trained offline on a labeled angle-position database.

We evaluate four classical DoA estimators on $\hat{\mathbf{R}}_{k,i}$: Bartlett [11], Capon [12], MUSIC [13] and ESPRIT [14], with the spectral and subspace forms of Bartlett, Capon, MUSIC and ESPRIT reported in Table I. In the table $\mathbf{a} \equiv \mathbf{a}(\theta, \phi)$ is the steering vector (3), $\mathbf{E}_S/\mathbf{E}_N$ are the signal and noise-subspace eigenvectors of $\mathbf{R}$ and $\mathbf{J}_1, \mathbf{J}_2$ are the selection matrices of two displaced UPA sub-arrays for each dimension ($\mathbf{J}_{1,x}, \mathbf{J}_{2,x}, \mathbf{J}_{1,y}, \mathbf{J}_{2,y}$). For Bartlett, Capon and MUSIC the strongest peak of the 2D (θ, ϕ) pseudospectrum is taken as the path-1 estimates, while ESPRIT returns the leading eigenpair directly. The pair is then quantized via (7) and pushed onto the backhaul.

TABLE I: Direction-of-arrival estimators evaluated at each AP.

Algorithm	Spatial estimator
Bartlett [11]	$P_B(\theta, \phi) = \mathbf{a}^H \hat{\mathbf{R}} \mathbf{a}$
Capon [12]	$P_C(\theta, \phi) = [\mathbf{a}^H \hat{\mathbf{R}}^{-1} \mathbf{a}]^{-1}$
MUSIC [13]	$P_M(\theta, \phi) = [\mathbf{a}^H \mathbf{E}_N \mathbf{E}_N^H \mathbf{a}]^{-1}$
ESPRIT [14]	$\Psi = (\mathbf{J}_1 \mathbf{E}_S)^\dagger (\mathbf{J}_2 \mathbf{E}_S),$ $\lambda_{x,k} = e^{j2\pi \frac{d}{\lambda} \sin \theta_k \cos \phi_k}$ $\lambda_{y,k} = e^{j2\pi \frac{d}{\lambda} \sin \theta_k \sin \phi_k}$

At the central processing unit, the global-frame angles received from the active APs feed the ETR [21], which produces the 3D UAV position directly from the quantized angular features. For each Cartesian axis $a \in \{x, y, z\}$, an ensemble of N_{tree} regression trees $T_{t,a}(\cdot)$ is trained on the full training set and averaged,

$$f_a(\boldsymbol{\xi}) = \frac{1}{N_{\text{tree}}} \sum_{t=1}^{N_{\text{tree}}} T_{t,a}(\boldsymbol{\xi}), \quad (10)$$

and the three axis-wise outputs are stacked into the position estimate $\hat{\mathbf{p}}_k = [f_x(\boldsymbol{\xi}_k), f_y(\boldsymbol{\xi}_k), f_z(\boldsymbol{\xi}_k)]^T$. Each tree is a piecewise-constant function partitioning the feature space into

J_t disjoint leaves $\{R_{t,j}\}_{j=1}^{J_t}$,

$$T_{t,a}(\boldsymbol{\xi}) = \sum_{j=1}^{J_t} c_{t,a,j} \mathbf{1}[\boldsymbol{\xi} \in R_{t,j}], \quad c_{t,a,j} = \frac{1}{|\mathcal{D}_{t,j}|} \sum_{(\boldsymbol{\xi}, \mathbf{p}) \in \mathcal{D}_{t,j}} [\mathbf{p}]_a, \quad (11)$$

where $\mathcal{D}_{t,j}$ collects the training samples falling in leaf j of tree t . The partition is grown by recursive axis-aligned splits $R \rightarrow \{R \cap \{\xi_d \leq \tau\}\} \cup \{R \cap \{\xi_d > \tau\}\}$, with τ drawn uniformly at random in $[\min \xi_d, \max \xi_d]$ and the split minimizing the local squared loss retained. The Base feature vector concatenates, for each active AP i , the quantized angles $(Q_B(\hat{\theta}_{k,i}), Q_B(\hat{\phi}_{k,i}))$ and the known position $\mathbf{p}_i$. The Base+Signal-to-Noise Ratio (SNR) augmentation appends one scalar per AP, the post-beamforming SNR computed on the noiseless signal. For each $N_{\text{AP}}^{\text{act}}$, the active APs are randomly drawn and a dedicated regressor is trained. Section IV evaluates the two feature configurations against two classical baselines that consume the same global-frame angles, angular MT [7] and the EKF [8].

A. Training Protocol of the Data-Driven Branch

The dataset is partitioned at the trajectory level, with all T_{steps} snapshots of a trajectory in a single split, eliminating temporal leakage. For every combination of DoA estimator, quantization depth B , number of active APs $N_{\text{AP}}^{\text{act}}$ and feature set, the per-slot tuples $(\boldsymbol{\xi}_k, \mathbf{p}_k)$ from the training split form the supervised dataset $\mathcal{D}_{\text{train}} = \{(\boldsymbol{\xi}_k^{(n)}, \mathbf{p}_k^{(n)})\}$. An independent triplet (f_x, f_y, f_z) of axis-wise regressors is then fit by minimizing the squared loss

$$f_a = \arg \min_f \frac{1}{|\mathcal{D}_{\text{train}}|} \sum_{(\boldsymbol{\xi}, \mathbf{p}) \in \mathcal{D}_{\text{train}}} ([\mathbf{p}]_a - f(\boldsymbol{\xi}))^2, \quad a \in \{x, y, z\}, \quad (12)$$

the native objective in scikit-learn. At test time, the frozen regressors process held-out feature vectors $\boldsymbol{\xi}_k$ slot-by-slot and emit $\hat{\mathbf{p}}_k = [f_x(\boldsymbol{\xi}_k), f_y(\boldsymbol{\xi}_k), f_z(\boldsymbol{\xi}_k)]^T$. The same data partition is reused by MT and EKF for a fair head-to-head comparison. The 3D MAE (8) on the test split is the figure of merit throughout Section IV.

B. Computational Complexity

Let $M = M_x M_y$ be the array size, $K_\theta K_\phi$ the angular search grid, L the number of pilot snapshots, $N_{\text{AP}}^{\text{act}}$ the number of active anchors, N_{tree} the ensemble size and D_t the average tree depth. At each AP, building the covariance $\hat{\mathbf{R}}_{k,i}$ costs $\mathcal{O}(M^2 L)$ regardless of the DoA method. Bartlett then performs an angular spectrum search at cost $\mathcal{O}(M^2 K_\theta K_\phi)$ without further matrix algebra. Capon adds the $\mathcal{O}(M^3)$ inversion $\hat{\mathbf{R}}^{-1}$. MUSIC requires an $\mathcal{O}(M^3)$ eigendecomposition followed by the same $\mathcal{O}(M^2 K_\theta K_\phi)$ scan over the noise-subspace pseudospectrum. ESPRIT keeps the $\mathcal{O}(M^3)$ eigendecomposition but bypasses the angular search, returning the angles directly from the rotational-invariance eigenvalues at $\mathcal{O}(P^3)$ where P is the model order. Thus Bartlett is the cheapest at the array, Capon and MUSIC have comparable cost dominated by the $K_\theta K_\phi$ scan, and ESPRIT is the most efficient when $K_\theta K_\phi \gg M$.

At the central unit, angular MT solves a fixed 3×3 linear system per slot at $\mathcal{O}(N_{\text{AP}}^{\text{act}})$ for the accumulator and $\mathcal{O}(1)$ for the inversion, with no training phase. The EKF updates a 6×6 state covariance and inverts a $2N_{\text{AP}}^{\text{act}} \times 2N_{\text{AP}}^{\text{act}}$ innovation matrix per slot at $\mathcal{O}((N_{\text{AP}}^{\text{act}})^3)$, plus $\mathcal{O}(N_{\text{AP}}^{\text{act}})$ for the analytical Jacobian. Per inference, the proposed regressors traverse each of the N_{tree} trees in $\mathcal{O}(D_t)$ for a per-axis $\mathcal{O}(N_{\text{tree}}D_t)$ that is independent of $N_{\text{AP}}^{\text{act}}$ and dominated by the chosen tree depth. Training is offline and pays a one-off $\mathcal{O}(N_{\text{tree}}F|\mathcal{D}_{\text{train}}|\log|\mathcal{D}_{\text{train}}|)$ cost, where $F \propto N_{\text{AP}}^{\text{act}}$ is the feature-vector dimension, amortized across deployments. At the recommended operating point of Section IV, this places ETR below the EKF in per-slot cost while consuming the same backhaul as MT.

IV. RESULTS

The scene is the UnB-FCTE campus with $N_{\text{AP}}=6$ APs, each carrying an $M_x \times M_y=8 \times 8$ UPA. The carrier is set to $f_c=28$ GHz with FR2 numerology $\mu=2$, giving subcarrier spacing $\Delta f=60$ kHz. A single PRB of $K=12$ subcarriers is reserved for positioning, so the band occupation $B_{\text{pos}}=K\Delta f$ evaluates to 720 kHz, less than 0.8% of a typical 100 MHz FR2 carrier. Each 14-symbol slot carries $N_p=2$ pilot OFDM symbols (positions 2 and 7), yielding overhead of $N_p/14 \approx 14\%$ and $L=KN_p=24$ snapshots per (UAV, AP) link. The UAV transmit power is $P_t=15$ dBm and the AP noise figure $\text{NF}=7$ dB, which with $T=303$ K set $\sigma_n^2=kTB_{\text{pos}}\text{NF}$ in (4).

The 3D Gauss–Markov model (5)–(6) uses $\alpha=0.8$, $\bar{v}=10$ m/s, $\sigma_v=3.0$ m/s and $\Delta t=1$ s. Trajectories have length $T_{\text{steps}}=25$ steps with altitude constrained to $[45, 100]$ m, representative of stable drone operations. The dataset comprises $N_{\text{traj}}^{\text{total}}=400$ Monte Carlo trajectories, split into $N_{\text{traj}}^{\text{train}}=280$ training and $N_{\text{traj}}^{\text{test}}=120$ test trajectories ($N_{\text{traj}}^{\text{train}}T_{\text{steps}}=7000$ and $N_{\text{traj}}^{\text{test}}T_{\text{steps}}=3000$ tuples, respectively). We jointly vary $N_{\text{AP}}^{\text{act}} \in \{2, 3, 4, 5, 6\}$, $B \in \{4, 8, 12, 16\}$ and the DoA estimator across Bartlett, Capon, MUSIC and ESPRIT. At $B=8$ each sine/cosine code fits one byte and Q_B reduces to a 2×256 lookup table, taken as the default operating choice. All reported MAE values use (8) on the held-out test split. Both baselines consume the same quantized angles: the geometry-based MT [7] and the recursive EKF [8], whose process model is the zero-mean-velocity Cartesian form of (5)–(6), with the angular noise variance set from the training-set DoA error and the correction gated by ≥ 3 valid APs. Each trajectory is warm-started at the true initial position, which makes the EKF an optimistic model-based reference.

The $B=12$ case is omitted because, for MT and ETR at $N_{\text{AP}}^{\text{act}} \geq 4$, it lies within ≤ 0.06 m of $B=16$. Table II reports the 3D MAE for ETR, MT and EKF using Bartlett, MUSIC and ESPRIT as the representative estimators, since, for ETR, Bartlett and Capon stay within centimeters of MUSIC and ESPRIT is at most 0.74 m worse at $N_{\text{AP}}^{\text{act}}=2$, matching MUSIC from $N_{\text{AP}}^{\text{act}}=4$; MT is more sensitive to the estimator choice. Both MT and ETR decrease with $N_{\text{AP}}^{\text{act}}$, up to minor fluctuations at $B=4$. At $N_{\text{AP}}^{\text{act}}=2$, MT is critically determined and the EKF runs in prediction-only mode, producing the order-of-magnitude inflation visible in the EKF column. ETR already

matches MT at $N_{\text{AP}}^{\text{act}}=4$ without a geometric model and at lower per-slot complexity than the EKF. Across bit-depths, $B=4$ injects biases that the baselines cannot filter and that ETR partially absorbs, whereas $B=8$, $B=12$ (omitted) and $B=16$ are indistinguishable for ETR and MT, confirming 8 bits as the operating point.

TABLE II: 3D MAE (m) across $N_{\text{AP}}^{\text{act}}$, B , DoA estimator and central-unit algorithm, Base feature vector. Capon agrees with Bartlett within ≤ 0.18 m for ETR and is omitted.

$N_{\text{AP}}^{\text{act}}$	DoA	3D MAE [m]								
		$B=4$			$B=8$			$B=16$		
		ETR	MT	EKF	ETR	MT	EKF	ETR	MT	EKF
2	Bartlett	7.35	7.94	63.6	3.31	3.83	63.6	3.25	3.80	63.6
	MUSIC	7.11	8.21	63.6	3.13	3.81	63.6	3.13	3.76	63.6
	ESPRIT	7.63	9.42	63.6	3.87	5.71	63.6	3.75	5.69	63.6
4	Bartlett	3.69	4.33	7.90	1.63	1.88	19.84	1.59	1.86	15.27
	MUSIC	3.65	4.41	8.79	1.69	1.76	12.26	1.63	1.72	14.34
	ESPRIT	3.77	4.89	7.22	1.68	2.66	9.57	1.66	2.64	9.51
6	Bartlett	3.12	3.59	7.83	1.44	1.32	1.27	1.39	1.30	1.26
	MUSIC	3.12	3.66	2.88	1.42	1.26	1.17	1.35	1.23	1.14
	ESPRIT	3.20	3.96	3.19	1.51	1.86	2.37	1.44	1.84	3.13

Table III complements this with the 10th, 50th and 90th percentiles of the 3D error for ETR with ESPRIT. The upper tail compresses consistently as $N_{\text{AP}}^{\text{act}}$ grows: at $N_{\text{AP}}^{\text{act}}=6$, the p_{90} error drops from 5.39 m at $B=4$ to 2.76 m at $B=8$, while p_{50} more than doubles when quantization drops to 4 bits. The EKF’s non-monotonic MAE, occasionally exceeding the 90th-percentile error, suggests trajectory-level divergence under ill-conditioned geometries that coarser quantization masks.

TABLE III: 10th, 50th and 90th percentiles of the 3D positioning error (m) with the ESPRIT estimator, Base feature vector, comparing ETR, MT and EKF.

$N_{\text{AP}}^{\text{act}}$	Alg.	3D positioning error [m]								
		$B=4$			$B=8$			$B=16$		
		p_{10}	p_{50}	p_{90}	p_{10}	p_{50}	p_{90}	p_{10}	p_{50}	p_{90}
2	ETR	2.91	6.75	13.31	1.03	2.79	7.98	1.00	2.60	7.66
	MT	3.35	7.30	14.88	0.51	2.36	12.68	0.44	2.30	12.83
	EKF	10.42	59.16	122.2	10.42	59.16	122.2	10.42	59.16	122.2
4	ETR	1.49	3.32	6.53	0.47	1.24	3.25	0.42	1.18	3.31
	MT	2.00	4.12	7.76	0.30	1.24	6.23	0.25	1.19	6.35
	EKF	1.38	2.86	5.79	0.27	0.78	3.46	0.24	0.77	3.58
6	ETR	1.31	2.81	5.39	0.48	1.13	2.76	0.44	1.05	2.59
	MT	1.67	3.38	6.20	0.26	0.89	4.23	0.20	0.86	4.24
	EKF	1.15	2.39	4.80	0.23	0.67	2.82	0.20	0.66	2.83

Table IV isolates the marginal value of appending per-AP SNR to the Base vector, exclusive to ETR since MT and EKF consume only global-frame angles (marked “–”). The gain is largest at low anchor count and bit-depth (34% at (2, 4), 23% at (2, 8)), narrowing to $<5\%$ at $N_{\text{AP}}^{\text{act}}=6$ for $B \geq 8$ as geometric redundancy removes most SNR-dependent residuals. The table also exposes ETR robustness to coarse quantization and few anchors: from $B=8$ to $B=4$ at $N_{\text{AP}}^{\text{act}}=6$, both MT (1.86 $\rightarrow$ 3.96 m) and ETR (1.51 $\rightarrow$ 3.20 m) grow by $2.1\times$ and +SNR drives the latter to 2.47 m, 37% below MT. At $N_{\text{AP}}^{\text{act}}=2$, where

the EKF loses its update step and inflates to 63.6 m, ETR stays at 3.87 m (2.99 m with +SNR) for $B=8$ and 7.63 m (5.06 m with +SNR) for $B=4$, dominating MT (5.71 and 9.42 m), making it the most resilient choice under aggressive backhaul compression and limited anchor visibility.

TABLE IV: 3D MAE (m) with the ESPRIT estimator under Base vs. Base+SNR, swept over N_{AP}^{act} and B . MT and EKF do not consume the SNR feature, marked “–”.

N_{AP}^{act}	Alg.	3D MAE [m]					
		$B=4$		$B=8$		$B=16$	
		Base	+SNR	Base	+SNR	Base	+SNR
2	ETR	7.63	5.06	3.87	2.99	3.75	2.87
	MT	9.42	–	5.71	–	5.69	–
	EKF	63.6	–	63.6	–	63.6	–
4	ETR	3.77	2.67	1.68	1.61	1.66	1.59
	MT	4.89	–	2.66	–	2.64	–
	EKF	7.22	–	9.57	–	9.51	–
6	ETR	3.20	2.47	1.51	1.47	1.44	1.38
	MT	3.96	–	1.86	–	1.84	–
	EKF	3.19	–	2.37	–	3.13	–

V. CONCLUSION

This paper addressed the backhaul constraint limiting DoA-based UAV positioning: each AP estimates and quantizes its own angle locally. A central unit running an ensemble regression-tree estimator (ETR) recovers the 3D position from the resulting B -bit angles. On the UnB-FCFE ray-traced dataset, this design directly answers the failure mode of each baseline: unlike MT, whose accuracy degrades under ill-conditioned or underdetermined geometry, ETR reaches an MAE approaching 1.5 m at $N_{AP}^{act}=6$ and matches MT already at $N_{AP}^{act}=4$. Unlike the EKF, which degrades to unbounded open-loop drift with too few visible APs, ETR remains stable at $N_{AP}^{act}=2$, precisely where the EKF loses its update step. Unlike deep neural networks, the proposed ETR is trained offline on a modest dataset and runs at lower per-slot complexity. This resilience extends to the front-end: under 4-bit quantization and as few as two active APs, ETR still outperforms both baselines. Appending the per-AP SNR to the feature vector recovers a further 23–34% of the resulting error.

$B=8$ recovers the unquantized accuracy of $B \in \{12, 16\}$, enabling a 2×256 lookup-table. With ETR, the four DoA estimators agree to within centimeters once $N_{AP}^{act} \geq 4$, the recommended operating point, with the SNR-augmented feature set as a fallback under fewer APs or coarser quantization.

Future work will address multi-UAV scenarios, hybrid ToA/DoA fusion and bit-depth-aware training.

REFERENCES

- [1] A. Fotouhi, H. Qiang, M. Ding, *et al.*, “Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges,” *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3417–3442, 2019. DOI: 10.1109/COMST.2019.2906228.
- [2] B. Li, Z. Fei, and Y. Zhang, “UAV communications for 5g and beyond: Recent advances and future trends,” *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2241–2263, 2019. DOI: 10.1109/JIOT.2018.2887086.

- [3] H. Wymeersch, D. Shrestha, N. Garcia, and G. Seco-Granados, “5g mmWave positioning for vehicular networks,” *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, 2017. DOI: 10.1109/MWC.2017.1600374WC.
- [4] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, “Position and orientation estimation through millimeter-wave MIMO in 5G systems,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, 2018. DOI: 10.1109/TWC.2017.2785788.
- [5] A. Shahmansoori, B. Uguen, G. Destino, G. Seco-Granados, and H. Wymeersch, “Tracking position and orientation through millimeter wave lens MIMO in 5G systems,” *IEEE Signal Process. Lett.*, vol. 26, no. 8, pp. 1222–1226, 2019. DOI: 10.1109/LSP.2019.2925969.
- [6] Y. Wang, Y. Zeng, P. H. Tan, S. Sun, and Y. Ma, “Drone controller localization based on TDoA,” in *2025 IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2025, pp. 1–6. DOI: 10.1109/WCNC61545.2025.10978550.
- [7] D. Torrieri, “Statistical theory of passive location systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-20, no. 2, pp. 183–198, 1984. DOI: 10.1109/TAES.1984.310439.
- [8] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, “State estimation for non-linear dynamic systems,” in *Estimation with Applications to Tracking and Navigation*. Wiley, 2001, ch. 10, pp. 371–420. DOI: 10.1002/0471221279.ch10.
- [9] M. Mousa and S. Al-Rubaye, “Intelligent 5g-aided UAV positioning in high-density environments using neural networks for NLOS mitigation,” *Aerospace*, vol. 12, no. 6, 2025, ISSN: 2226-4310. DOI: 10.3390/aerospace12060543.
- [10] X. Wang, L. Gao, and S. Mao, “CSI phase fingerprinting for indoor localization with a deep learning approach,” *IEEE Internet Things J.*, vol. 3, no. 6, pp. 1113–1123, 2016. DOI: 10.1109/JIOT.2016.2558659.
- [11] M. S. Bartlett, “Periodogram analysis and continuous spectra,” *Biometrika*, vol. 37, no. 1/2, pp. 1–16, 1950. DOI: 10.1093/biomet/37.1-2.1.
- [12] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969. DOI: 10.1109/PROC.1969.7278.
- [13] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986. DOI: 10.1109/TAP.1986.1143830.
- [14] R. Roy and T. Kailath, “ESPRIT-estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989. DOI: 10.1109/29.32276.
- [15] T. Rice, D. Pandey, D. Ramirez, and E. Knightly, “Experimental evaluation of AoA estimation for UAV to massive MIMO,” in *MILCOM 2023 - 2023 IEEE Mil. Commun. Conf. (MILCOM)*, 2023, pp. 839–844. DOI: 10.1109/MILCOM58377.2023.10356267.
- [16] S. R. and S. Jacob, “Comparative study of DOA estimation algorithms,” in *2022 IEEE 19th India Council Int. Conf. (INDICON)*, 2022, pp. 1–4. DOI: 10.1109/INDICON56171.2022.10040076.
- [17] M. Espandar and H. R. Bakhshi, “AoA and AoD estimation for rectangular antenna array in non-selective and slow fading in MIMO channels,” in *2009 Int. Conf. Future Comput. Commun. (ICFCC)*, 2009, pp. 390–394. DOI: 10.1109/ICFCC.2009.88.
- [18] J. Hoydis, F. A. Aoudia, S. Cammerer, *et al.*, “Sionna RT: Differentiable ray tracing for radio propagation modeling,” in *2023 IEEE Globecom Workshops (GC Wkshps)*, 2023, pp. 317–321. DOI: 10.1109/GCWkshps58843.2023.10465179.
- [19] 3GPP, “Study on channel model for frequencies from 0.5 to 100 GHz,” 3rd Generation Partnership Project, Tech. Rep. TR 38.901, 2022.
- [20] IEEE 802.16 Broadband Wireless Access Working Group, “IEEE 802.16m evaluation methodology document (EMD),” Institute of Electrical and Electronics Engineers, Tech. Rep. IEEE 802.16m-08/004r5, 2009.
- [21] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely randomized trees,” *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006. DOI: 10.1007/s10994-006-6226-1.