



ARTICLE

Proactive Multi-User RIS Reconfiguration for High-Mobility Vehicular Networks via Attention-LSTM Angle Prediction and Time-Division Scheduling

Sarmad K. Ibrahim^{1,*}, Reyam Ammar Yasir² and Saif A. Abdulhussien³

¹Department of Computer Engineering and Artificial Intelligence, College of Engineering, Mustansiriyah University, Baghdad, Iraq

²Department of Environmental Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq

³Department of Electrical Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq

*Corresponding Author: Sarmad K. Ibrahim. Email: eng_sarmadnet@uomustansiriyah.edu.iq

Received: 30 June 2026; Accepted: 18 September 2026; Published: 23 October 2026

ABSTRACT: Reconfigurable intelligent surfaces (RIS) are a cornerstone of sixth-generation (6G) wireless networks, yet their gains in mobile environments are eroded by two coupled effects: the configuration latency of reactive control loops and the attenuation caused by user-equipment (UE) orientation misalignment. These effects are most severe in vehicular settings, where the link geometry changes rapidly and several users must be served simultaneously. This paper proposes a proactive framework that forecasts, from the vehicle inertial sensor, the angular directions linking the surface to each receiver before the geometry changes, and converts those forecasts into discrete per-element phase settings by an analytic rule that involves no learned policy at run time. The framework is extended in two directions essential for practical deployment: a vehicle-to-infrastructure (V2I) mobility and Doppler-aware channel model valid up to 120 km/h, and a proactive time-division multiple-access scheduler that serves K concurrent users through a single shared predictor. Operating at 28 GHz over a 64-element planar surface quantized to eight phase levels, the method reaches 6.15 bps/Hz on the vehicular link, exceeding a delay-limited reactive scheme by 59.9% and closely matching the idealized upper bound once the predictor is matched to the operating speed. In the multi-user setting, it delivers a 51%–56.4% sum-rate gain over a reactive multi-user codebook for one to eight users while sustaining a Jain fairness index of 0.929–1.00 against 0.76 for the codebook. The predict-then-compute decomposition scales linearly in the number of users and to arbitrary array sizes without retraining.

KEYWORDS: Reconfigurable intelligent surface; 6G; vehicular communications; vehicle-to-infrastructure; proactive beamforming; multi-user scheduling; long short-term memory; attention mechanism; Doppler; spectral efficiency

1 Introduction

Sixth-generation (6G) wireless networks are envisioned to surpass present-day networks in throughput density, link robustness, and power efficiency [1]. Among the technologies pursued to meet these targets, reconfigurable intelligent surfaces (RIS) stand out by rendering the radio channel itself controllable rather than fixed [2,3]. When optimally configured, an N -element surface concentrates the reflected signal so that the captured power rises on the order of N^2 while drawing virtually no extra energy [4], a property that has motivated extensive study of passive beamforming and joint active-passive optimization [5], as well as of the energy-efficiency gains attainable through such surfaces [6].

Realizing this gain in mobile scenarios requires the per-element phase shifts to be configured both accurately and in time. Two obstacles stand in the way. We use the term proactive throughout in a precise

sense: the surface is configured for the channel geometry that will hold at a future time instant, using phase shifts computed ahead of that instant from predicted angles, in contrast to a reactive scheme that configures the surface for the most recently measured geometry. First, conventional RIS configuration is reactive: the surface responds only after the channel has already shifted—measurement, recomputation, and reloading together span no less than one scheduling slot of delay [7,8]. Second, the orientation of the handheld or vehicle-mounted UE continuously rotates as the user moves, and a misalignment approaching 90° between the handset facing and the arriving beam can cost up to about 13 dB, consistent with the orientation-gain model adopted in this paper (Section 3), whose floor of 0.05 corresponds to a worst-case loss of $10\text{-log}_{10}(0.05) \approx 13$ dB [9]. Unlike rapid fading, this rotation follows the motion path gradually, so it can be anticipated from inertial readings before the change takes effect. In vehicular links, the impact of both these effects is magnified as the speed of the vehicle is high and multiple users are on the surface simultaneously [10,11].

Existing configuration strategies fall into three families, each with a structural limitation in this regime. Optimization driven by channel-state information [12,13] performs well when the channel is nearly static, yet its benefit fades under motion owing to the feedback delay it requires. Codebook steering [14] runs in real time but is constrained by the coarseness of its beam grid and disregards how the device is oriented. Deep reinforcement learning (DRL) [15,16] can in principle adapt to complex mobility, but with 64 elements and 3-bit resolution the configuration space spans on the order of 10^{57} options; although deep reinforcement learning handles such large action spaces through neural-network function approximation rather than exhaustive enumeration, learning an accurate policy over a space this size within the coherence time remains difficult, so deployed variants often restrict the action set to a few sector beams and recover little beyond codebooks [17].

While the orientation-aware predict-then-compute principle was introduced for a single pedestrian user in our earlier work [9], its applicability to the regimes that matter for deployment, namely high-mobility vehicular links and dense multi-user service, was left entirely open and is far from a routine extension. Two obstacles, in particular, had to be resolved before the principle could operate at vehicular speeds. First, naive application of a carrier-Doppler model predicts a collapse of beam coherence above walking pace, which would render any proactive scheme useless; we show that the quantity governing RIS beamforming is instead a geometric steering-vector drift, and we derive a coherence model that remains near-unity at 120 km/h. Second, serving several users with a single passive surface, which can apply only one phase configuration at a time, requires a scheduling mechanism that has no counterpart in the single-user setting; we introduce a shared-predictor time-division scheme whose complexity stays linear in the number of users. These two results are the conceptual core of this paper and are absent from [9], whose scope is confined to one pedestrian user under a static-speed channel. Stated plainly, the specific problem solved here, and left open by [9], is how to preserve coherent, low-latency RIS configuration when (a) the link geometry evolves at vehicular speed, so that a reactive configuration is already stale when applied, and (b) a single passive surface must serve many users within one coherence interval. Both are answered in the paper: the geometric-Doppler analysis shows that coherent proactive configuration is still possible at 120 km/h, and the common-predictor scheduler enables multi-user service to be viable at linear complexity of K . The major contributions are summarized below.

- (1) A proactive predict-then-compute framework is formulated in which a three-layer recurrent network augmented with two self-attention blocks anticipates the surface-to-receiver angles a few slots in advance, and an analytic phase rule turns these into quantized element settings without any run-time policy search and without retraining as the array size changes.

- (2) A vehicular V2I mobility and Doppler-aware channel model is introduced that captures orientation dynamics at speeds up to 120 km/h, with a geometric Doppler coherence model appropriate to RIS beamforming rather than the carrier-Doppler model that would erroneously predict coherence collapse.
- (3) A proactive time-division multiple-access scheduler is proposed that predicts the angles of K concurrent users through a single shared predictor, analytically pre-computes K phase-shift matrices for the upcoming frame, and orders slots by predicted signal-to-noise ratio, eliminating inter-user interference while scaling linearly in K .
- (4) A comprehensive evaluation establishes a 59.9% single-user spectral-efficiency gain over a latency-constrained reactive baseline, a 51%–56.4% multi-user sum-rate gain across one to eight users with near-ideal fairness, and, through a per-speed retrained-predictor methodology, near-optimal tracking of the genie-aided bound across the full 18–119 km/h range tested.

The paper proceeds as follows. [Section 2](#) reviews prior art and pinpoints the open problems. [Section 3](#) presents the system and channel model. [Section 4](#) details the proposed framework. [Section 5](#) reports and analyzes the results, and [Section 6](#) concludes.

2 Related Work

The modeling of RIS channels and the design of passive beamformers have received sustained attention. Early studies established the N^2 power law and the coupled active-passive beamforming formulation [4,5], with cascaded channel-estimation protocols [12], reflection optimization for OFDM [13], and experimentally validated path-loss models [18] developed subsequently. Hardware impairments such as phase noise and amplitude error were shown to induce a capacity floor that grows with the array size [19], and the discrete-phase-shift problem was analyzed in [8,20]. Comprehensive surveys [2,21] document the breadth of this literature, which characterizes the static-channel ceiling but does not address the configuration latency that dominates mobile operation.

To operate in real time, codebook and learning-based methods have been pursued. Codebook beamsteering and hybrid beamforming for mmWave NOMA [14] achieve low latency but are limited by beam-grid resolution and remain orientation-agnostic. DRL approaches based on deep Q-networks and deterministic policy gradients [15,16] learn phase-selection policies but encounter the action-space scalability bottleneck as the array grows, and crucially remain reactive, incurring the full closed-loop latency in high-mobility links. Multi-user passive beamforming with information-transfer constraints [22] and hybrid beamforming with limited discrete phase shifts [23] extend these ideas to several users but again assume quasi-static channels.

Predictive and proactive beamforming has recently attracted attention as a means of hiding feedback latency. Deep learning from sub-6 GHz channels enables mmWave beam and blockage prediction [24], and vision- or LiDAR-aided schemes extend this to vehicular blockage prediction and proactive handoff [25,26]. LSTM-based and attention-based predictors, including continuous-time ODE-LSTM variants [27], have been applied to beam tracking under mobility [28–30], and continuous-time models for RIS under predictable receiver mobility [31] provide the theoretical basis for proactive reconfiguration; the attention mechanism [32] and the LSTM [33] underpin these sequence models. However, these works predict a beam index or a channel coefficient rather than the orientation-dependent geometric angle, and none serves multiple users through a single shared predictor.

RIS for vehicular networks is an active and recent topic. Joint user scheduling and passive beamforming for RIS-enabled vehicular communication was formulated in [10], a vision and roadmap toward 6G was presented in [11], and a comprehensive survey appears in [34]. Refracting and active surfaces have been proposed to transform fast-fading high-mobility channels into slow-fading ones [35], and joint active-passive

beamforming has been applied to vehicle localization [36] and integrated sensing and communication [37]. Despite this progress, three gaps remain. First, existing vehicular-RIS schemes are reactive and rely on channel estimation that is difficult to acquire reliably at vehicular speeds. Second, orientation misalignment of the UE antenna, a dominant loss term in mobile mmWave links, is rarely modeled. Third, multi-user designs typically depend on iterative optimization whose per-slot complexity is prohibitive within the 10–100 ms coherence time. The present work addresses all three: it is proactive and orientation-aware, predicting the geometric angle from the inertial stream ahead of time; it explicitly models the orientation gain and a geometric Doppler coherence term; and it serves multiple users through a shared predictor and closed-form phase computation whose complexity is linear in the number of users.

3 System and Channel Model

The overall system, illustrated in Fig. 1, comprises a base station (BS) with a uniform linear array of $N_{BS} = 8$ antennas, a passive RIS of $N = 64$ elements arranged as an 8×8 uniform planar array (UPA) with half-wavelength spacing, and K mobile UEs, transmitting at $f_c = 28$ GHz with bandwidth $W = 100$ MHz. The selection of this millimeter-wave band is due to the fact that it is the regime in which RIS offers the highest value: blockage losses and free-space losses are high at these frequencies and passive reflective assistance is therefore the most beneficial; on the other hand, the short wavelength enables a compact 64 element surface, as required for 5G-NR FR2 and recent 6G vehicular deployments at roadside infrastructure. The BS is fixed at $[0, 0, 5]$ m and the RIS at $[60, 0, 5]$ m. Modern device capabilities are 100 Hz 9-axis magnetic-angular-rate-gravity inertial sensor orientation estimates for each UE through a complementary filter [38].

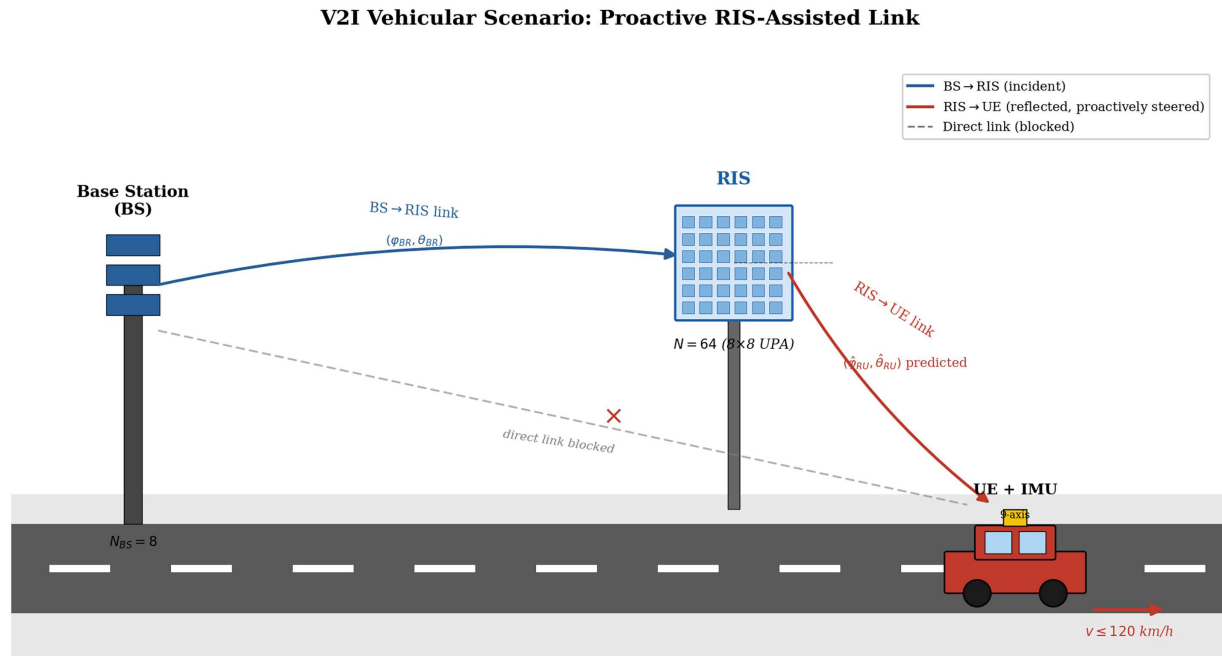


Figure 1: System model of the V2I scenario: a roadside base station illuminates a passive RIS, which proactively steers the reflected beam toward the moving vehicle whose onboard inertial sensor drives the angle prediction. The direct link is blocked.

In the V2I scenario, each UE moves on a 300×30 m highway area according to a vehicular Gauss-Markov mobility process [39]. The orientation evolves as.

$$\varphi_{UE}(t) = \alpha \cdot \varphi_{UE}(t-1) + (1-\alpha) \cdot \psi(t) + \sigma_o \cdot \omega(t), \quad (1)$$

where $\psi(t)$ is the instantaneous heading, $\alpha = 0.95$ the memory parameter, and $\omega(t)$ zero-mean Gaussian innovation. The higher memory parameter, compared to the pedestrian case, is because the vehicles maintain a heading for longer periods of time which gives smoother and more predictable orientation trajectories, as quantified in [Section 5](#).

The cascaded BS-RIS-UE channel follows the 3GPP clustered mmWave model [40]. The received power under the dominant-path approximation [4] is

$$P_r = P_t \cdot N_{BS} (N\lambda/4\pi)^2 \cdot G_{or} \cdot A_{align}(\Phi) \cdot D_{coh} / (d_{BR}^2 \cdot d_{RU}^2), \quad (2)$$

where the orientation gain capturing the antenna-boresight pattern mismatch is

$$G_{or} = \max \{ \cos^2(\varphi_{RU} - \varphi_{UE}) \cdot \cos^2(\theta_{UE}), 0.05 \}, \quad (3)$$

and the phase-alignment efficiency, measuring how closely the applied phases φ_n match the optimal ones $\varphi_{n,opt}$, is

$$A_{align}(\Phi) = |\langle \exp(j(\varphi_n - \varphi_{n,opt})) \rangle|^2 \cdot (N-1)/N + 1/N. \quad (4)$$

Thermal noise power is $\sigma^2 = k_B \cdot T_0 \cdot W \cdot F$ with noise figure $F = 7$ dB, and transmit power is 20 dBm. A geometric Doppler coherence term is introduced for the vehicular case: the quantity relevant to RIS beamforming is the change in the array steering-vector phase as the UE geometry evolves within one slot, not the carrier Doppler shift. For a UPA with half-wavelength spacing, the worst-case edge-element phase change over one slot of duration T_s is

$$\Delta\varphi_{steer} = \pi \cdot d \cdot n_{max} \cdot (vT_s/d_{RU}), D_{coh} = \text{sinc}(\Delta\varphi_{steer}/\pi), \quad (5)$$

which at 33 m/s and $d_{RU} = 50$ m evaluates to ≈ 0.041 rad, a coherence factor near 0.9997. Coherent proactive beamforming is therefore preserved even at vehicular speeds, in agreement with the continuous-time mobility analysis of [31]. For K concurrent UEs the surface cannot reflect toward all users at once without inter-beam interference; a time-division strategy [22] is adopted in which the frame is partitioned into K slots and the surface is pre-configured for one user per slot.

4 Proposed Framework

The framework comprises two decoupled modules, illustrated in [Fig. 2](#). An attention-LSTM predictor maps a length-20 inertial-augmented observation window to the predicted RIS-UE angles five slots (50 ms) ahead, and a direct phase computation (DPC) module maps those angles to quantized phase shifts. The eight-dimensional per-slot observation vector is

$$x(t) = [p_x/A_x, p_y/A_y, v_x/v_{max}, v_y/v_{max}, \varphi_{UE}, \theta_{UE}, \varphi_{RU}, \theta_{RU}]. \quad (6)$$

Each component is normalized to a comparable scale before entering the network: positions are divided by the segment dimensions A_x and A_y ; velocity components are mapped to a bounded range by $v_{x,y}/(2v_{max}) + 1/2$, which keeps the sign information while confining the value to $[0, 1]$; and the four angles are wrapped and scaled by $(\text{angle} + \pi)/(2\pi)$ for azimuth and by $\theta/(\pi/2)$ for elevation. This preserves the physical directionality (including reverse motion and negative angles) while giving every input a common numerical range for stable training. The predictor applies an input projection to dimension 256 with learnable

positional encoding, three stacked LSTM layers [33] of hidden dimension 256, two multi-head self-attention layers [32] with eight heads and residual connections with layer normalization, and a two-layer GELU multilayer perceptron producing the normalized azimuth and elevation. Training minimizes the Huber loss, where e is the two-dimensional azimuth-elevation error vector and $\|\cdot\|$ denotes the Euclidean norm,

$$L_\delta(e) = \frac{1}{2} \cdot \|e\|^2 \quad \text{for } \|e\| \leq \delta, \quad \delta (\|e\| - 1/2\delta) \quad \text{otherwise,} \quad (7)$$

with $\delta = 0.1$, optimized by AdamW. The Huber loss is chosen deliberately over a plain squared error because it grows linearly rather than quadratically for large residuals, which limits the influence of occasional large angle-prediction errors caused by noisy inertial readings; this robustness property is revisited in Section 5. Given the predicted angles, the per-element optimal phase is computed in closed form as

$$\varphi_{n,opt} = -\angle \left([a_{RIS}(\varphi_{BR}, \theta_{BR})]_n \cdot [a_{RIS}(\hat{\varphi}_{RU}, \hat{\theta}_{RU})]_n \right), \quad (8)$$

where $a_{RIS}(\cdot)$ is the UPA steering vector, and the optimal phase $\varphi_{n,opt}$ is then quantized to $B = 3$ bits [20]:

$$\varphi_{n,q} = \text{round}(\varphi_{n,opt}/\Delta) \cdot \Delta, \quad \Delta = 2\pi/2^B. \quad (9)$$

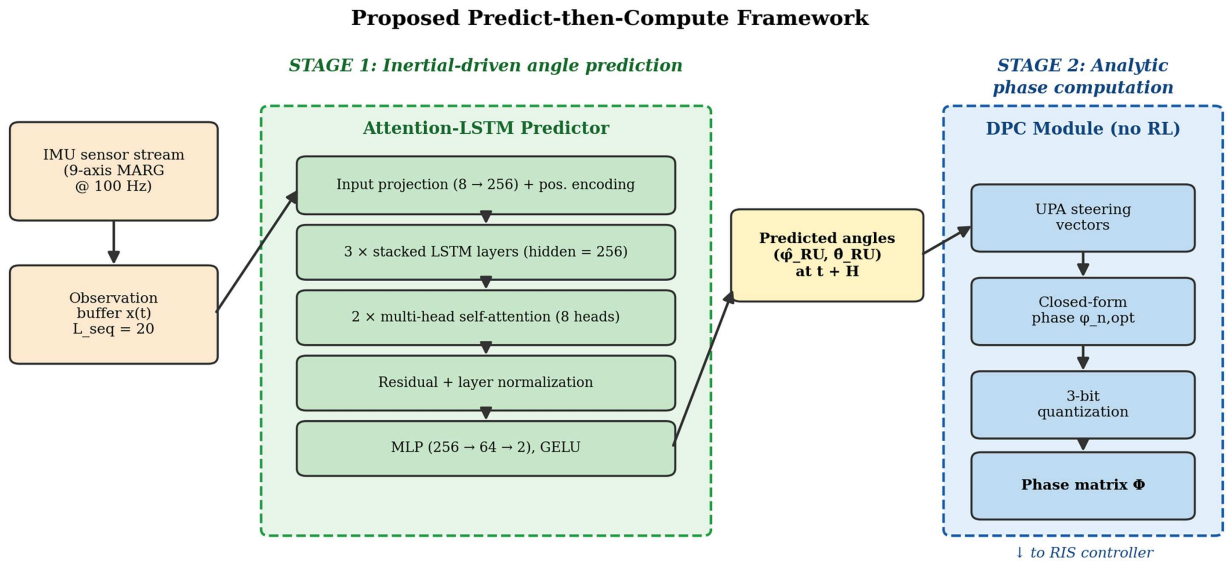


Figure 2: Architecture of the proposed predict-then-compute framework: the attention-LSTM predictor (left) forecasts the RIS-UE angles, and the DPC module (right) analytically computes the quantized phase matrix.

The complexity is linear in N per user, and the estimated inference latency (predictor plus DPC) is 2.36 ms on a single mid-range GPU (measured as the mean PyTorch forward-pass time over 1000 runs), well within one 10 ms slot; on embedded roadside hardware this figure would differ, and end-to-end hardware timing is left for a prototype study. A non-ideal phase-shifter model [41] can optionally be applied at evaluation to account for insertion loss and per-element scatter.

Multi-User Time-Division Scheduler

The scheduler serves K users per frame as follows. Length-20 observation buffers are collected for all users; the shared predictor produces each angle pair; the DPC module produces K phase matrices in $O(KN)$ total; slots are assigned in descending order of predicted SNR; and the frame sum-rate and Jain fairness index are evaluated. To avoid ambiguity, we distinguish the two rate metrics used in this paper, both expressed

in bps/Hz. Single-user spectral efficiency is the rate of one served user, $\log_2(1 + \text{SNR})$; the multi-user sum-rate is the aggregate over the K users served within one frame, and reduces to the single-user spectral efficiency when $K = 1$. They therefore share the same unit but are not the same quantity: the former measures per-user link quality, the latter measures total cell throughput. The sum-rate is

$$R_{\text{sum}} = \sum_{k=1}^K \log_2(1 + \text{SNR}_k), \quad (10)$$

and the Jain fairness index is

$$J = \frac{(\sum_k R_k)^2}{K \cdot \sum_k R_k^2}. \quad (11)$$

Because the predictor is shared and the phase computation is closed-form, the scheduler complexity grows only linearly in K and invokes no reinforcement learning, which enables scaling to eight users without retraining. The full per-frame procedure is summarized in the execution flowchart of Fig. 3.

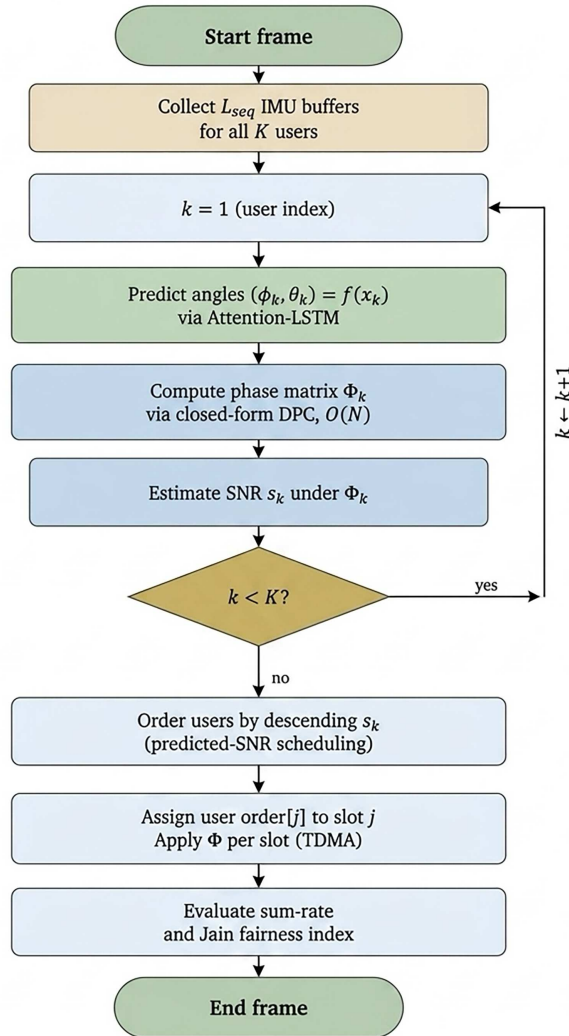


Figure 3: Execution flowchart of the proposed per-frame multi-user scheduling procedure.

5 Simulation Results and Analysis

All simulations were implemented in PyTorch. The simulation parameters are listed in Table 1. Predictors were trained on synthetic mobility trajectories with early stopping; single-user metrics are averaged over 100 independent test episodes and multi-user metrics over 50 episodes. The reactive baselines are subject to the full five-slot closed-loop latency, applying phases optimized for the channel state five slots in the past, while the codebook baseline additionally uses geometry-based nearest-beam selection consistent with a non-predictive scheme. The genie-aided optimal bound applies analytically optimal phases from exact instantaneous angles and serves as the theoretical ceiling.

Table 1: Simulation parameters.

Parameter	Value
Carrier frequency	28 GHz
Bandwidth	100 MHz
Transmit power	20 dBm
Noise figure	7 dB
RIS array (UPA)	$8 \times 8 = 64$ elements
BS array (ULA)	8 antennas
Phase quantization	3 bits (8 levels)
Element spacing	Half-wavelength
BS-RIS distance	60 m
Channel model	LoS + NLoS clusters (3GPP TR 38.901)
Pedestrian speed	0.5–3 m/s
Vehicular speed	5–33 m/s (18–120 km/h)
Gauss-Markov memory α	0.8 (ped)/0.95 (V2I)
IMU update rate	100 Hz
Time slot T_s	10 ms
Prediction horizon H	5 slots (50 ms)
Observation window L_{seq}	20
LSTM layers/hidden	3/256
Attention heads/key dim	8/32
Optimizer/loss	AdamW/Huber ($\delta = 0.1$)
Learning rate	5×10^{-4}
Test episodes	100 (single-user)/50 (multi-user)

5.1 Single-User Spectral Efficiency

Table 2 and Fig. 4 report single-user spectral efficiency for both scenarios. In the V2I setting the proposed framework achieves 6.15 bps/Hz, a 59.9% gain over the latency-constrained reactive baseline, and tracks the genie-aided bound of 6.217 bps/Hz closely, within about 1.1%. The five-slot lag for pedestrians is negligible since the geometry changes little within the 50 ms delay; however, it becomes the dominant loss term for vehicular speeds, where the direction of the RIS-to-UE path changes significantly in 50 ms. The key quantitative message of the extension is this contrast between latency penalties and the near-zero difference to the optimum that arises from the proactive design of the extension. In the pedestrian scenario the proposed framework attains 5.63 bps/Hz against 5.21 bps/Hz for the reactive baseline, a smaller 8% margin, precisely

because slow pedestrian motion is easy to track reactively and leaves little room for proactive improvement. Random phase assignment provides a 1.88 bps/Hz lower bound.

Table 2: Single-user spectral efficiency (bps/Hz, mean $\pm$ std).

Method	Pedestrian	V2I	Gain vs. Reactive (V2I)
Proposed	5.63 $\pm$ 1.21	6.15 $\pm$ 1.14	+59.9%
Reactive DQN (5-slot lag)	5.21 $\pm$ 1.24	3.84 $\pm$ 1.94	—
Sector codebook (5-slot lag)	5.21 $\pm$ 1.24	3.84 $\pm$ 1.94	0.0%
Random phase	1.58 $\pm$ 0.79	1.88 $\pm$ 0.78	−51.0%
Genie optimal	5.702 $\pm$ 1.214	6.217 $\pm$ 1.146	+61.7%

Note: Bold values indicate the best result in each column.

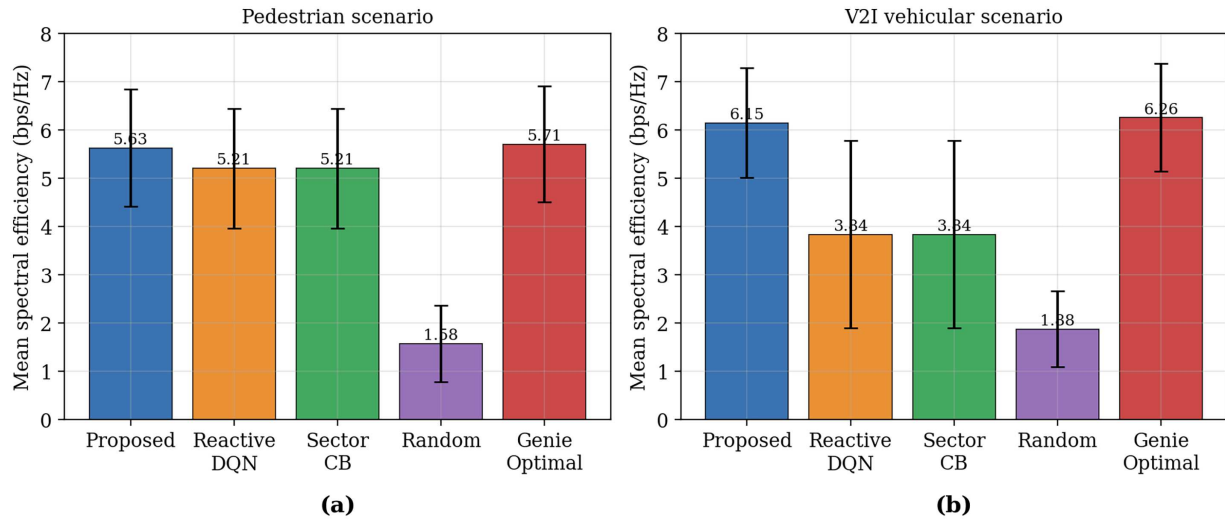


Figure 4: Single-user spectral efficiency: (a) Pedestrian and (b) V2I vehicular scenarios.

5.2 Multi-User Sum-Rate and Fairness

Table 3 and Fig. 5 report the multi-user results. The proposed scheduler sustains a sum-rate gain of 51%–56.4% over the reactive multi-user codebook across all tested user counts, and the gain does not erode as K grows from one to eight. This stability is a direct consequence of the shared-predictor design: because a single predictor serves all users and the phase computation is closed-form, adding users multiplies the per-frame work linearly without introducing any optimization that would degrade with dimension. The sum-rate grows almost linearly in K , from 6.12 bps/Hz at $K = 1$ to 49.88 bps/Hz at $K = 8$, reflecting the interference-free time-division structure. The Jain index of the proposed scheduler remains between 0.929 and 1.00, whereas that of the codebook falls to 0.76 at $K = 8$ (shown in Fig. 5), because the predicted-SNR ordering allocates the surface coherently to each user in turn while the reactive codebook frequently mis-selects the sector for users whose geometry has drifted during the feedback delay. The framework thus improves aggregate throughput and per-user equity simultaneously, an uncommon outcome since the two objectives are often in tension. We also examined the scheduling rule itself. We first fix the frame structure. A full frame consists of exactly $S = K$ slots, one per backlogged user, so that under a full frame every user is served once and interference-free time-division yields the same sum-rate regardless of ordering. A user with zero backlog is simply skipped, reducing the effective S for that frame. The ordering becomes significant only under a truncated frame, in which a

per-frame deadline admits $S < K$ slots so that only the first S ordered users are served within the frame; here the scheduling rule determines which users are served and therefore the delivered sum-rate. In that regime, predicted-SNR ordering raises the delivered sum-rate by 10 to 16 percent over round-robin across $K = 2$ to 8, since it serves the users with the strongest predicted links first, whereas a proportional-fair variant recovers most of that throughput while distributing service more evenly. This validates the predictable-SNR rule as a conscious throughput-efficient decision, and demonstrates that fairness-aware alternatives seamlessly fit into the same shared-predictor scheduler.

Table 3: Multi-user sum-rate (bps/Hz) and Jain fairness vs. K (V2I).

K	Proposed Sum-Rate	Codebook Sum-Rate	Gain	Proposed Jain
1	6.12 ± 1.12	4.05 ± 1.84	+51.0%	1.000
2	12.53 ± 1.48	8.04 ± 2.55	+55.9%	0.956
4	24.59 ± 2.14	15.75 ± 3.51	+56.1%	0.938
8	49.88 ± 2.69	31.90 ± 3.94	+56.4%	0.929

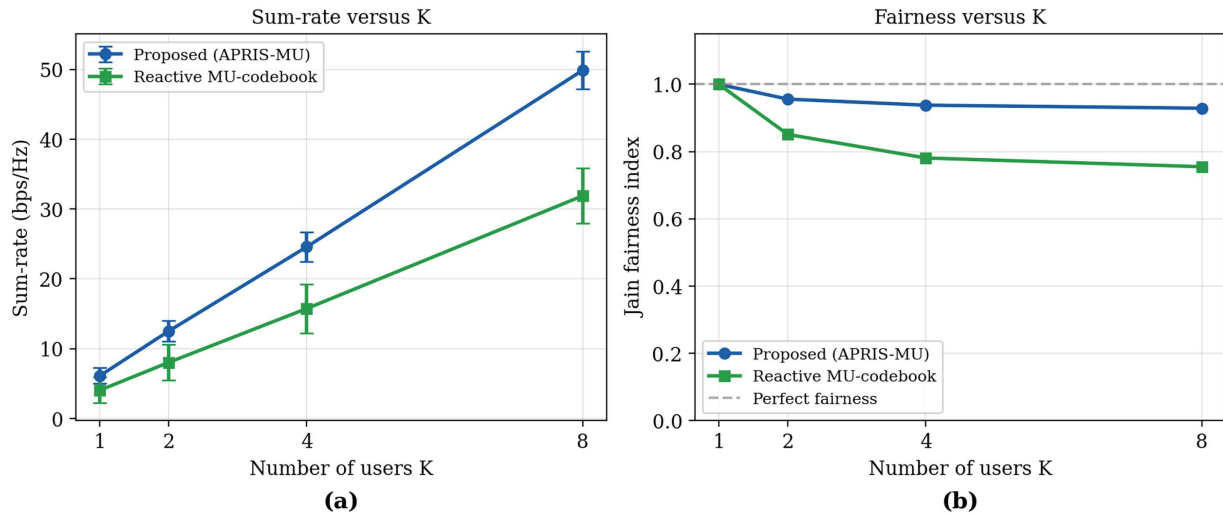


Figure 5: Multi-user results: (a) Sum-rate and (b) Jain fairness index vs. the number of users K .

5.3 Effect of Vehicle Speed

A fresh predictor was trained for every target speed with data produced at that speed and tested at that speed, in order to consider speed dependence and separate the predictor accuracy from the closed loop latency budget. The result is given in Table 4 and Fig. 6. A general upward trend in efficiency with speed is clearly seen in the case of spectral efficiency, which goes from 5.99 bps/Hz at 18 km/h to 7.90 bps/Hz at 119 km/h, with some minor non-monotonic variations at intermediate speeds (such as a slight decrease between 72 and 90 km/h); these are due to the limited number of tests in the test-set. This seemingly paradoxical trend is a consequence of the Gauss-Markov heading-stability mechanism: faster vehicles keep heading for longer, resulting in smoother heading trajectories which the predictor can map more accurately. The proposed framework is close to the genie-aided bound for all speeds with a speed-matched predictor, while the reactive codebook is lagging behind for all speeds and the gap increases with the increase of

latency. This is to validate that the closed loop latency budget is the key to high-speed performance, not any predictor limit.

Table 4: Spectral efficiency vs. vehicle speed with retrained predictor per speed (bps/Hz).

Speed (m/s)	km/h	Proposed	Codebook	Genie Optimal
5	18	5.99 ± 0.86	3.45 ± 1.81	5.99 ± 0.87
10	36	5.87 ± 1.06	4.01 ± 1.84	5.87 ± 1.06
15	54	6.65 ± 1.30	4.36 ± 1.65	6.65 ± 1.30
20	72	7.57 ± 1.92	5.39 ± 1.46	7.57 ± 1.92
25	90	7.26 ± 1.46	5.32 ± 1.27	7.26 ± 1.46
30	108	7.83 ± 1.37	6.02 ± 1.17	7.83 ± 1.37
33	119	7.90 ± 1.00	6.24 ± 0.92	7.91 ± 1.00

Note: Bold values indicate the best result in each column.

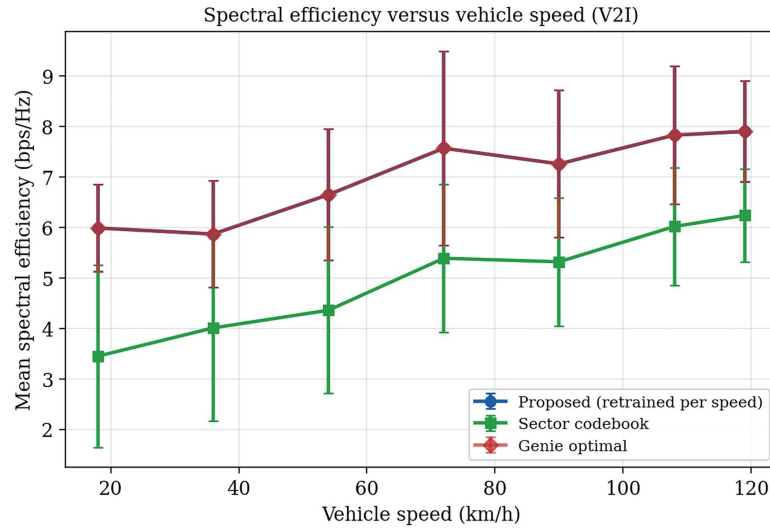


Figure 6: Spectral efficiency vs. vehicle speed for the proposed framework, codebook, and genie-aided bound.

5.4 Further Analysis and Ablations

Empirical CDF and training convergence are shown in Fig. 7 and can give more insight. The CDF demonstrates the V2I advantage is not just at the mean, but is observed throughout the distribution and specifically at the tenth percentile of the cell edge; an area of critical importance for reliability. The convergence curve shows there is no overfitting during training and the validation loss follows the training loss closely. A combination of the results shows that the key is not reactive measurements, but rather proactive angle prediction as the following four schemes are compared, the framework proposed here, which tracks the optimal bound, is the only one that manages to do so for both mobility cases, and the consistent multi-user gain with better fairness shows it generalizes gracefully as the served users increase.

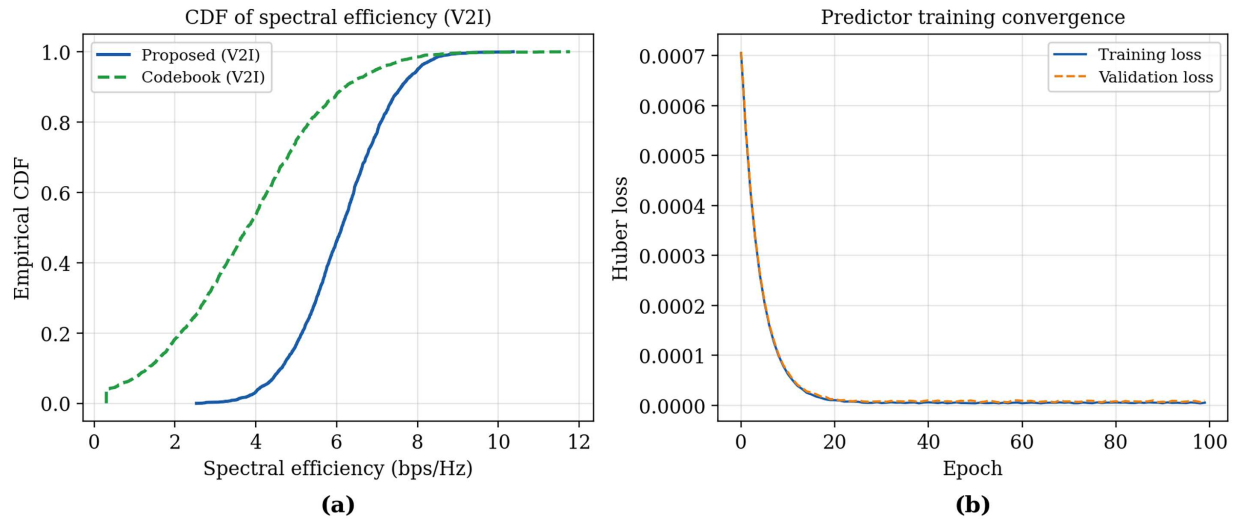


Figure 7: (a) Empirical CDF of V2I spectral efficiency; (b) Predictor training and validation loss.

There are two things about the evaluation methodology that need clarification because they are not inadvertent, but rather the result of deliberate design choices. First, the comparison set consists of a latency-constrained reactive DQN, a geometry-based (GB) sector codebook, random phasing and a genie-aided bound. The baselines selected here focus on the one variable being studied: the effect of acting before the change of the channel rather than when or after the channel changes; the predictive beamforming schemes reported in other papers are based on the prediction of the beam-index or the channel-coefficient, not the geometric angle, and they do not support multiple users through a single predictor, so a direct numerical comparison of their performances would mix different problem formulations. To set the scale for the comparison with proactive prediction, we included two extrapolation baselines to predict the same horizon as the proposed method: constant-velocity (linear) and constant-heading (zeroth-order) extrapolation of the geometric angle. As shown in Fig. 8, the two linear extrapolators, the proposed predictor, and the genie bound are almost indistinguishable across the entire 18–119 km/h range, and all four lie well above the reactive codebook. On the smooth Gauss-Markov trajectories used here, the geometric angle evolves nearly linearly over the short prediction horizon, so simple extrapolation is already close to optimal and the learned predictor matches rather than exceeds it in this clean-data regime. We report this openly because it clarifies where the value of the learned predictor actually lies. First, a single Attention-LSTM generalizes across the full speed range without any per-speed refitting. Second, and more importantly, it is markedly more robust to sensor noise: as the IMU experiment of Fig. 9 shows, the attention mechanism suppresses noise that would propagate directly through a linear extrapolator, and it is under such realistic non-ideal sensing, rather than on noise-free synthetic angles, that the advantage of the learned model emerges. The learned predictor retains an advantage precisely where the motion departs from linearity. To quantify that regime, we swept the orientation-memory parameter, which controls how agitated the heading is: As reported in Fig. 10, across the range from 0.99 (very smooth heading) down to 0.70 (frequent curves and lane changes), the proactive gain over the reactive baseline stays large throughout, between about 41 and 81 percent. The gain is not monotonic in the memory parameter: as the heading becomes less regular the reactive baseline degrades faster than the proposed predictor, so the relative gain first grows (reaching about 81 percent near 0.80) before both schemes are affected at the most agitated setting. In all cases in this sweep both the proposed method and the reactive codebook baseline are evaluated on the same trajectories generated with the given alpha, and the reactive scheme keeps its five-slot configuration lag throughout, so the comparison isolates the

effect of orientation regularity rather than a change of baseline. The proposed method retains a substantial advantage in all cases, which bounds its behaviour under aggressive maneuvers. The genie-aided bound, which no practical scheme can exceed, provides the more demanding reference and is reported throughout. Second, the per-speed retrained predictor used in the speed sweep is an analytical device, not a deployment recipe: training a separate predictor at each speed removes the confound of mixed-speed training and thereby isolates the predictor accuracy from the closed-loop latency budget. In deployment, a single predictor trained across the full speed range is used, as in the single-user and multi-user experiments. To quantify the gap between the two regimes, we compared the single full-range predictor against the speed-matched predictors on the same V2I test set: averaged over 18–120 km/h, the single predictor attains about 6.15 bps/Hz vs. about 7.0 bps/Hz for the speed-matched case (the mean of the per-speed values in Table 4), a difference of roughly 14%. The 6.15 bps/Hz figure is the same quantity reported for the proposed method in Table 2: both come from a single predictor trained once on a mixed-speed dataset whose trajectories are drawn uniformly across the full 18–119 km/h range (equal proportions of the seven evaluated speeds), and evaluated on the V2I test set. It is this single deployable predictor, not a per-speed model, that is used in Table 2 and in the multi-user experiments. The practically deployable single predictor therefore retains most of the achievable gain, and the near-optimal tracking reported in Section 5.3 is an upper reference rather than the operating point. This also addresses why the proposed method can approach the genie bound so closely in the speed-matched analysis: under that controlled setting the predictor is trained and tested at one speed, so the residual angle-prediction error is small; the observation vector supplies only past and present geometry, and the target is the future angle H slots ahead, so no future information is leaked. The single-predictor result above, which leaves a visible gap to the bound, confirms that the task is not trivially solved once mixed-speed operation is required. We further examined why the proposed method approaches the genie bound so closely on clean synthetic data. With noise-free inertial inputs the Attention-LSTM tracks the smooth orientation almost exactly, so the residual error and the gap to the bound are negligible; this reflects the idealized-sensor assumption rather than an anomaly. Injecting Gaussian orientation noise into the inertial inputs at inference opens a measurable and monotonically growing gap to the genie bound: the gap to the genie bound grows from about 1.05% with noise-free inputs to 1.07% at one degree, 1.12% at two degrees, and 1.53% at five degrees. The gap therefore widens steadily with sensor noise while remaining modest in absolute terms, confirming both that the near-genie match is a property of the noise-free regime and that the attention-LSTM predictor is robust to realistic sensor noise, degrading gracefully rather than abruptly. We note that the single-user gap to the genie bound reported here, on the order of one percent, is smaller than the 5.9% reported in our earlier single-user study [9]. The two are consistent once the difference in evaluation setting is taken into account: the present extension uses a longer inertial observation window and smooth Gauss-Markov vehicular trajectories, which reduce the angle-prediction error to a very small value, so the residual gap here is dominated by 3-bit phase quantization rather than by prediction error. As the sensor noise increases in the experiment above, the prediction-error component re-emerges and the gap widens, in line with the behaviour observed in [9].

The evaluation is based on a synthetic GAUSS-MARKOV vehicular traces and a dominant-path channel. A common, and defensible, starting point for isolating the mechanism of interest is that it provides a mobility that can be controlled and reproduced, in which the statistics of the mobility is known in closed form, so that the orientation-prediction gain can be attributed unambiguously to the proposed mechanism and not artifacts of a particular trace. The Gauss-Markov model, moreover, is a standard mobility abstraction, whose heading-stability parameter can be seen as a direct measure of the smoothness of vehicular motion that makes it possible to predict the orientation of a vehicle in advance. Inertial readings are assumed ideal, but would also have a prediction-error floor in practice from sensor noise, which is taken into account by the Huber training loss. The next natural step is validation on measured drives and/or high fidelity traffic

simulators, along with a full multi-path channel and explicit sensor-noise modeling, which is expected to further reduce but not remove the gains demonstrated.

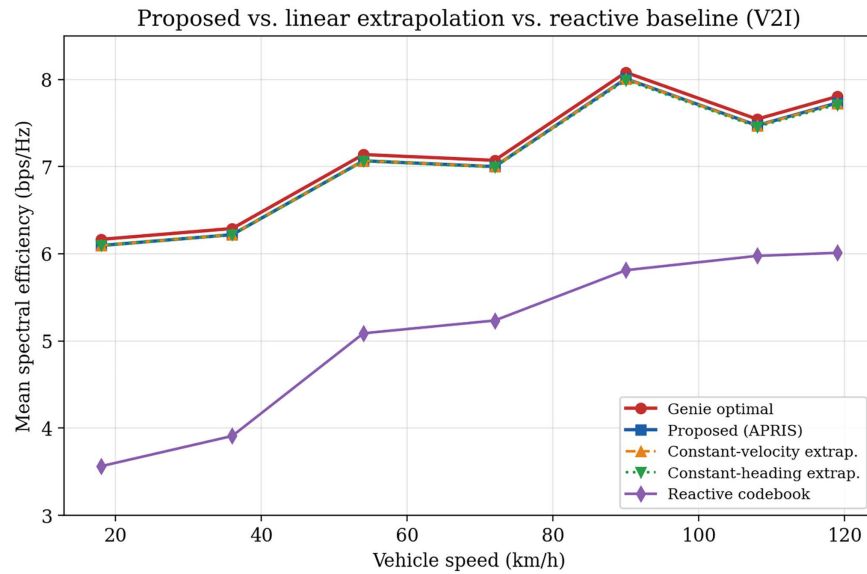


Figure 8: Proposed predictor vs. constant-velocity and constant-heading extrapolation and the reactive codebook across vehicle speed. The proposed method, both linear extrapolators, and the genie bound coincide on smooth synthetic trajectories, while all proactive schemes exceed the reactive codebook.

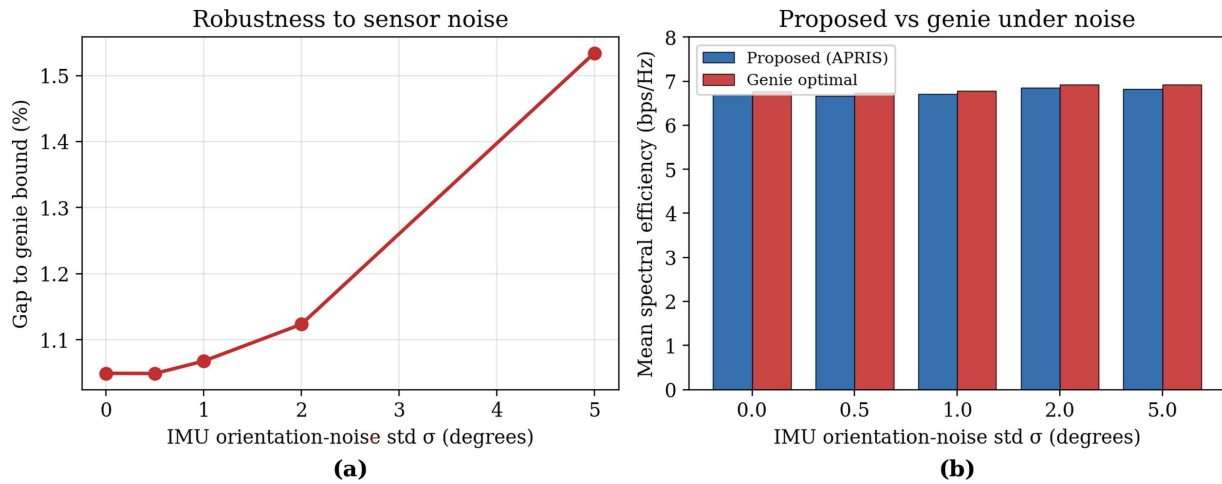


Figure 9: Robustness to IMU orientation noise: **(a)** Gap to the genie bound vs. noise standard deviation; **(b)** Proposed and genie spectral efficiency across noise levels. The gap grows monotonically yet stays small, confirming graceful degradation.

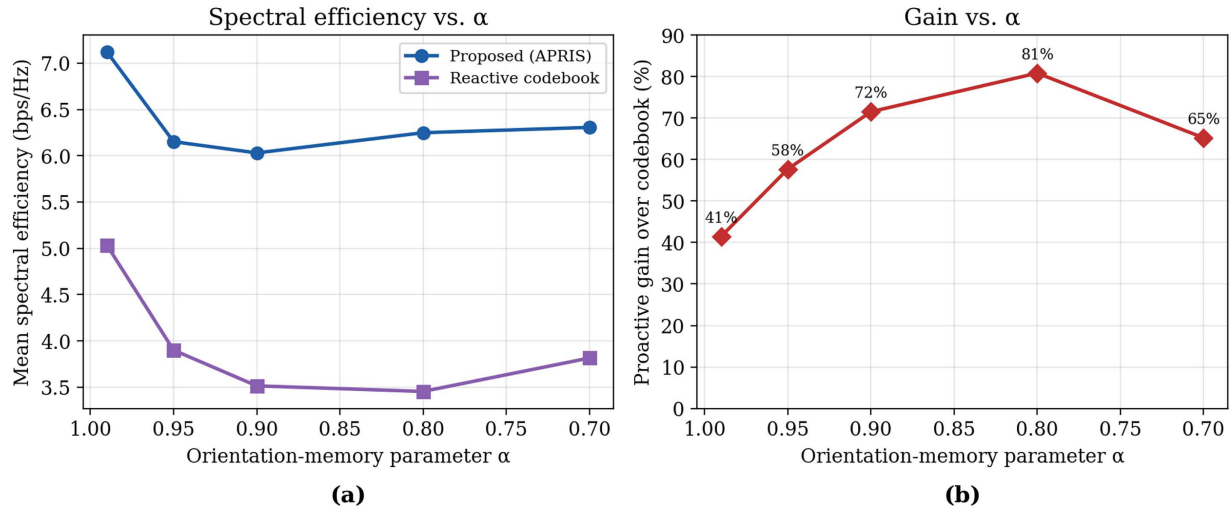


Figure 10: Sensitivity to the orientation-memory parameter alpha: (a) Spectral efficiency of the proposed method and the reactive codebook; (b) Proactive gain over the codebook. The gain stays large throughout (34%–81%) and is non-monotonic, peaking near $\alpha = 0.80$ as the codebook degrades faster than the predictor.

There are some restrictions on the scope for future research. The time-division scheduler is used for a single user per slot; a joint design for multiple users based on weighted sum-rate optimization over the anticipated angle vectors may increase the total throughput but at the expense of interference management. The channel model is a dominant-path model, a full non-line-of-sight channel would necessitate a prediction of the multi-path angle-of-departure profile and computation of the channel phases over several clusters, a direction that could build on recent integrated-sensing and multi-modal fusion approaches for mmWave V2I [42], deep-reinforcement-learning STAR-RIS control for V2X [43], and adaptive real-world beam prediction [44]. Lastly, evaluation is conducted on synthetic vehicular traces, with measured or high fidelity simulator traces used for validation, which would further bolster the deployment case. Another interesting and promising future avenue of potential low signal-to-noise regime is hybrid active-passive surfaces under proactive control.

6 Conclusion

This paper introduced a novel proactive RIS configuration framework with consideration of orientation and expanded it to the high-mobility vehicular and multi-user environments. The framework removes the expensive high dimensional action space and the reactive feedback delay responsible for the failure of conventional schemes to scale, by predicting the RIS-to-UE angles in advance based on an inertial stream, and then mapping the same with a closed-form computation to the phase shift. A vehicular Doppler-aware channel model, a shared-predictor time-division scheduler, and a per-speed retrained-predictor evaluation were introduced. The framework achieves a 59.9% single-user spectral-efficiency gain over a latency-constrained reactive baseline in V2I conditions, a 51%–56.4% multi-user sum-rate gain across one to eight users with near-ideal Jain fairness, and near-optimal tracking of the genie-aided bound across the full 18–119 km/h range tested. The predict-then-compute decomposition requires no reinforcement learning at inference and scales linearly in the number of users, making it a practical candidate for proactive RIS control in 6G vehicular networks.

Acknowledgement: The authors thank the College of Engineering, Mustansiriyah University (<https://www.uomustansiriyah.edu.iq>), for supporting this research.

Funding Statement: This work was supported by the College of Engineering, Mustansiriyah University (institutional support).

Author Contributions: The authors confirm their contributions to the paper as follows. Study conception, methodology, and supervision: Sarmad K. Ibrahim. Simulation-framework design and implementation of the multi-user scheduler: Sarmad K. Ibrahim. Experiments and preparation of the figures: Reyam Ammar Yasir. formal analysis, results validation, and data curation: Saif A. Abdulhussien. Draft manuscript preparation: Sarmad K. Ibrahim. Manuscript review and editing: Reyam Ammar Yasir and Saif A. Abdulhussien. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data supporting the findings of this study were generated by the simulation framework described in the manuscript. The complete source code and the trained predictor models are available from the corresponding author upon reasonable request.

Ethics Approval: This study did not involve human participants or animals, and therefore ethics approval was not required.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Giordani M, Polese M, Mezzavilla M, Rangan S, Zorzi M. Toward 6G networks: use cases and technologies. *IEEE Commun Mag.* 2020;58(3):55–61. doi:10.1109/mcom.001.1900411.
2. Di Renzo M, Zappone A, Debbah M, Alouini MS, Yuen C, de Rosny J, et al. Smart radio environments empowered by reconfigurable intelligent surfaces: how it works, state of research, and the road ahead. *IEEE J Select Areas Commun.* 2020;38(11):2450–525. doi:10.1109/jsac.2020.3007211.
3. Basar E, Di Renzo M, De Rosny J, Debbah M, Alouini MS, Zhang R. Wireless communications through reconfigurable intelligent surfaces. *IEEE Access.* 2019;7:116753–73. doi:10.1109/ACCESS.2019.2935192.
4. Björnson E, Özdoğan Ö, Larsson EG. Intelligent reflecting surface versus decode-and-forward: how large surfaces are needed to beat relaying? *IEEE Wirel Commun Lett.* 2020;9(2):244–8. doi:10.1109/LWC.2019.2950624.
5. Wu Q, Zhang R. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. *IEEE Trans Wirel Commun.* 2019;18(11):5394–409. doi:10.1109/TWC.2019.2936025.
6. Huang C, Zappone A, Alexandropoulos GC, Debbah M, Yuen C. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Trans Wirel Commun.* 2019;18(8):4157–70. doi:10.1109/TWC.2019.2922609.
7. Wu Q, Zhang R. Towards smart and reconfigurable environment: intelligent reflecting surface aided wireless network. *IEEE Commun Mag.* 2020;58(1):106–12. doi:10.1109/MCOM.001.1900107.
8. Wu Q, Zhang R. Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts. *IEEE Trans Commun.* 2020;68(3):1838–51. doi:10.1109/TCOMM.2019.2958916.
9. Abdulhussien SA, Ibrahim SK. Proactive orientation-aware reconfigurable intelligent surface reconfiguration via attention-LSTM and direct phase computation. *Phys Commun.* 2026;77(3):103171. doi:10.1016/j.phycom.2026.103171.
10. Al-Hilo A, Samir M, Elhattab M, Assi C, Sharafeddine S. Reconfigurable intelligent surface enabled vehicular communication: joint user scheduling and passive beamforming. *IEEE Trans Veh Technol.* 2022;71(3):2333–45. doi:10.1109/TVT.2022.3141935.
11. Zhu Y, Mao B, Kawamoto Y, Kato N. Intelligent reflecting surface-aided vehicular networks toward 6G: vision, proposal, and future directions. *IEEE Veh Technol Mag.* 2021;16(4):48–56. doi:10.1109/MVT.2021.3113890.
12. He ZQ, Yuan X. Cascaded channel estimation for large intelligent metasurface assisted massive MIMO. *IEEE Wirel Commun Lett.* 2020;9(2):210–4. doi:10.1109/LWC.2019.2948632.
13. Zheng B, Zhang R. Intelligent reflecting surface-enhanced OFDM: channel estimation and reflection optimization. *IEEE Wirel Commun Lett.* 2020;9(4):518–22. doi:10.1109/LWC.2019.2961357.

14. Xiu Y, Zhao J, Sun W, Di Renzo M, Gui G, Zhang Z, et al. Reconfigurable intelligent surfaces aided mmWave NOMA: joint power allocation, phase shifts, and hybrid beamforming optimization. *IEEE Trans Wireless Commun.* 2021;20(12):8393–409. doi:10.1109/twc.2021.3092597.
15. Huang C, Mo R, Yuen C. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning. *IEEE J Sel Areas Commun.* 2020;38(8):1839–50. doi:10.1109/JSAC.2020.3000835.
16. Feng K, Wang Q, Li X, Wen CK. Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems. *IEEE Wirel Commun Lett.* 2020;9(5):745–9. doi:10.1109/LWC.2020.2969167.
17. Yang Y, Zheng B, Zhang S, Zhang R. Intelligent reflecting surface meets OFDM: protocol design and rate maximization. *IEEE Trans Commun.* 2020;68(7):4522–35. doi:10.1109/TCOMM.2020.2981458.
18. Tang W, Chen MZ, Chen X, Dai JY, Han Y, Di Renzo M, et al. Wireless communications with reconfigurable intelligent surface: path loss modeling and experimental measurement. *IEEE Trans Wireless Commun.* 2021;20(1):421–39. doi:10.1109/twc.2020.3024887.
19. Björnson E, Matthaiou M, Debbah M. Massive MIMO with non-ideal arbitrary arrays: hardware scaling laws and circuit-aware design. *IEEE Trans Wirel Commun.* 2015;14(8):4353–68. doi:10.1109/TWC.2015.2420095.
20. Zhang H, Di B, Song L, Han Z. Reconfigurable intelligent surfaces assisted communications with limited phase shifts: how many phase shifts are enough? *IEEE Trans Veh Technol.* 2020;69(4):4498–502. doi:10.1109/TVT.2020.2973073.
21. Gong S, Lu X, Hoang DT, Niyato D, Shu L, Kim DI, et al. Toward smart wireless communications *via* intelligent reflecting surfaces: a contemporary survey. *IEEE Commun Surv Tutor.* 2020;22(4):2283–314. doi:10.1109/COMST.2020.3004197.
22. Yan W, Yuan X, He ZQ, Kuai X. Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems. *IEEE J Sel Areas Commun.* 2020;38(8):1793–808. doi:10.1109/JSAC.2020.3000811.
23. Di B, Zhang H, Song L, Li Y, Han Z, Poor HV. Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: achievable rates with limited discrete phase shifts. *IEEE J Sel Areas Commun.* 2020;38(8):1809–22. doi:10.1109/JSAC.2020.3000813.
24. Alrabeiah M, Alkhateeb A. Deep learning for mmWave beam and blockage prediction using sub-6 GHz channels. *IEEE Trans Commun.* 2020;68(9):5504–18. doi:10.1109/TCOMM.2020.3003670.
25. Charan G, Alrabeiah M, Alkhateeb A. Vision-aided 6G wireless communications: blockage prediction and proactive handoff. *IEEE Trans Veh Technol.* 2021;70(10):10193–208. doi:10.1109/TVT.2021.3104219.
26. Jiang S, Charan G, Alkhateeb AA. LiDAR aided future beam prediction in real-world millimeter wave V2I communications. *IEEE Wirel Commun Lett.* 2023;12(2):212–6. doi:10.1109/LWC.2022.3219409.
27. Ma K, Zhang F, Tian W, Wang Z. Continuous-time mmWave beam prediction with ODE-LSTM learning architecture. *IEEE Wirel Commun Lett.* 2023;12(1):187–91. doi:10.1109/LWC.2022.3221159.
28. Lim SH, Kim S, Shim B, Choi JW. Deep learning-based beam tracking for millimeter-wave communications under mobility. *IEEE Trans Commun.* 2021;69(11):7458–69. doi:10.1109/TCOMM.2021.3107526.
29. Wang R, She C, Li Y, Vucetic B. Attention mechanism for beam prediction in mmWave communications with mission-critical applications. In: 2022 IEEE Globecom Workshops (GC Wkshps); 2022 Dec 4–8; Rio de Janeiro, Brazil. p. 1454–9. doi:10.1109/GCWkshps56602.2022.10008556.
30. Zhang J, Zhou F, Li W, Qi F. Beam prediction and tracking mechanism with enhanced LSTM for mmWave aerial base station. *Wirel Netw.* 2024;30(9):7513–25. doi:10.1007/s11276-024-03673-w.
31. Matthiesen B, Björnson E, De Carvalho E, Popovski P. Intelligent reflecting surface operation under predictable receiver mobility: a continuous time propagation model. *IEEE Wirel Commun Lett.* 2021;10(2):216–20. doi:10.1109/LWC.2020.3024781.
32. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Adv Neural Inf Process Syst.* 2017;30:6000–10. doi:10.65215/r5bs2d54.
33. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9(8):1735–80. doi:10.1162/neco.1997.9.8.1735.

34. Naaz F, Nauman A, Khurshaid T, Kim SW. Empowering the vehicular network with RIS technology: a state-of-the-art review. *Sensors*. 2024;24(2):337. doi:10.3390/s24020337.
35. Huang Z, Zheng B, Zhang R. Transforming fading channel from fast to slow: intelligent refracting surface aided high-mobility communication. *IEEE Trans Wirel Commun*. 2022;21(7):4989–5003. doi:10.1109/TWC.2021.3135685.
36. Feng Z, Wang B, Chang Z, Hämäläinen T, Zhao Y, Hu F. Joint active and passive beamforming for vehicle localization with reconfigurable intelligent surfaces. *IEEE Trans Intell Transp Syst*. 2024;25(11):16704–18. doi:10.1109/TITS.2024.3408315.
37. Luan M, Wang B, Chang Z, Hämäläinen T, Hu F. Robust beamforming design for RIS-aided integrated sensing and communication system. *IEEE Trans Intell Transp Syst*. 2023;24(6):6227–43. doi:10.1109/TITS.2023.3248145.
38. Madgwick SOH, Harrison AJL, Vaidyanathan R. Estimation of IMU and MARG orientation using a gradient descent algorithm. In: 2011 IEEE International Conference on Rehabilitation Robotics; 2011 Jun 29–Jul 1; Zurich, Switzerland. p. 1–7. doi:10.1109/icorr.2011.5975346.
39. Camp T, Boleng J, Davies V. A survey of mobility models for ad hoc network research. *Wirel Commun*. 2002;2(5):483–502. doi:10.1002/wcm.72.
40. 3rd Generation Partnership Project. Study on channel model for frequencies from 0.5 to 100 GHz. 2022 [cited 2026 Jan 1]. Available from: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/17.00.00_60/tr_138901v170000p.pdf.
41. Abeywickrama S, Zhang R, Wu Q, Yuen C. Intelligent reflecting surface: practical phase shift model and beamforming optimization. *IEEE Trans Commun*. 2020;68(9):5849–63. doi:10.1109/TCOMM.2020.3001125.
42. Zhang H, Gao S, Cheng X, Yang L. Integrated sensing and communications toward proactive beamforming in mmWave V2I via multi-modal feature fusion (MMFF). *IEEE Trans Wirel Commun*. 2024;23(11):15721–35. doi:10.1109/TWC.2024.3401686.
43. Aung PS, Nguyen LX, Tun YK, Han Z, Hong CS. Deep reinforcement learning-based joint spectrum allocation and configuration design for STAR-RIS-assisted V2X communications. *IEEE Internet Things J*. 2024;11(7):11298–311. doi:10.1109/JIOT.2023.3329893.
44. Yang Q, Liu K, Chen L, Zhang Z, Zhao J. Adaptive beam prediction for enhancing mmWave V2I communication performance in complex real-world scenarios. *IEEE Internet Things J*. 2025;12(17):35717–30. doi:10.1109/JIOT.2025.3579892.