Internet of Vehicles via Rate-Splitting Multiple Access With Antenna and RIS Partitioning

Maurice Nduwayezu, Kofi Amoako Ofori-Amanfo, Derek Kwaku Pobi Asiedu, Ji-Hoon Yun

Vehicular Communication via Rate-Splitting Multiple Access With Antenna and RIS Partitioning

Maurice Nduwayezu, Kofi Amoako Ofori-Amanfo, Derek Kwaku Pobi Asiedu, *Member, IEEE*, and Ji-Hoon Yun, *Senior Member, IEEE*

***Abstract*—Integrating rate-splitting multiple access (RSMA) and reconfigurable intelligent surfaces (RIS) into vehicular communication systems offers significant potential for enhancing quality of service but also introduces challenges in radio resource allocation due to increased system complexity and the heterogeneous mobility patterns of vehicles. In this paper, we propose a vehicle-to-infrastructure (V2I) communication system that leverages a partitioned antenna array and RIS elements for reduced computational complexity, with RSMA serving multiple vehicles and users therein. To mitigate inter-vehicular interference, we employ a zero-forcing technique that decouples precoding at the BS and decoding at the user equipment (UE) into a zero-forcing factor and a precoding factor. For sum-rate maximization, we develop a block coordinate descent (BCD) algorithm, incorporating a dynamic partitioning scheme that divides the BS antenna array and RIS elements into subplanes. Each subplane is dedicated to a single vehicle for transmitting private streams to UEs therein, along with common streams for all UEs across all vehicles. The passive beamforming for each RIS subplane is optimized using the Riemannian conjugate gradient (RCG) method, while precoding at the BS subarray is optimized using a weighted minimum mean square error (WMMSE) approach. Simulation results and comparisons with benchmarks demonstrate that the proposed partitioning scheme achieves robust performance in terms of system sum rate while significantly reducing computational complexity.**

***Index Terms*—vehicular communication, reconfigurable intelligent surface, BS array and RIS elements partitioning, rate splitting multiple Access, mmWave, computational complexity.**

I. Introduction

The Internet of Vehicles (IoV) is a rapidly evolving paradigm that enables seamless data exchange between vehicles, network infrastructure, and other connected devices, supporting applications such as traffic management, autonomous driving, and enhanced road safety [1]. Reliable and high-spectral-efficiency vehicular communication systems are essential for the success of IoV, as they offer real-time data transmission among vehicles. Achieving this requires leveraging wide bandwidth and ensuring strong channel gains. Furthermore, vehicular communication systems must implement efficient multiple access schemes to support concurrent user streams while maintaining low-complexity radio resource allocation algorithms. This is particularly challenging in the presence of heterogeneous vehicle velocities, which result in diverse channel conditions and introduce channel estimation errors.

Millimeter-wave (mmWave) communication systems have emerged as a promising solution to address the exponential growth in mobile data traffic, leveraging the abundant bandwidth available in the mmWave spectrum. The short wavelength of mmWave signals enables the integration of a large number of antenna elements within a compact physical space, making it well-suited for high-density antenna arrays at the base station (BS). This integration realizes highly directional beamforming to mitigate severe signal attenuation while simultaneously supporting spatial multiplexing [2] [3] [4] [5]. In the context of vehicular communication, mmWave technology holds significant potential for delivering high-speed data services to vehicles. However, the challenging nature of vehicular propagation environments—characterized by vehicle mobility and encountering obstacles such as buildings and other vehicles—poses significant difficulties in maintaining reliable mmWave links between BSs and vehicles or users therein [2] [6] [7]. These conditions may result in signal blockages and amplify the inherent high signal attenuation of the mmWave spectrum, making it difficult to ensure consistent and reliable connectivity in mmWave vehicular networks.

Emerging reconfigurable intelligent surfaces (RIS), consisting of arrays of passive reflecting elements capable of dynamically adjusting their reflection coefficients, provide practical solutions to mitigate blockages in mmWave vehicular communication [8] [9]. However, deploying RIS with a fixed number of reflecting elements to assist specific vehicles may not align with the highly dynamic nature of vehicular channel environments. Furthermore, the computational complexity associated with large-scale RIS configurations can make the system less cost-efficient and less responsive to real-time environmental changes. Many works in vehicular communication adopted large scale of unit elements at antenna array and/or RIS technologies in order to improve the system's throughput [2] [8] [10] [11]. On the one hand, larger units elements will cause larger array dimension and hence high computational complexity for the optimization of the transmit beamforming at BS and passive beamforming at RIS.

This work was supported in part by the Institute of Information Communications Technology Planning & Evaluation (IITP) Grant funded by the Korean Government (MSIT) under Grant 2022-0-00214 and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant RS-2019-NR040071.

(Corresponding author: Ji-Hoon Yun.)

Maurice Nduwayezu, Kofi A. Ofori-Amanfo, and Ji-Hoon Yun are with the Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Seoul 01811, Republic of Korea (Emails: ndumaurice@naver.com, kofiaoforiamanfo@seoultech.ac.kr, and jhyun@seoultech.ac.kr).

Derek Kwaku Pobi Asiedu is with the Department of Mathematics and Electrical Engineering, Institut Mines-Telecom (IMT) Atlantique, 29238 Brest, France (e-mail: derek.asiedu@imt-atlantique.fr).

To enable concurrent multi-user streams with high spectral efficiency in advanced vehicular communication systems, a promising approach is the adoption of rate-splitting multiple access (RSMA). RSMA provides the flexibility to split each user's rate into multiple message streams with varying proportions, effectively managing interference by optimally encoding messages into common and private streams [12] [13]. Integrating RSMA with RIS in vehicular communication introduces significant challenges in radio resource allocation due to the system's inherent complexity. The optimization process involves numerous coupled variables, including beamforming, RIS reflection coefficients, and rate-splitting strategies, all of which are highly interdependent. This makes real-time optimization computationally intensive, particularly in dynamic vehicular environments. Additionally, RSMA's sensitivity to interference requires careful management of inter-stream interference, ensuring that it is either mitigated or effectively treated as noise through optimal message encoding and decoding. The primary challenge lies in achieving high spectral efficiency while minimizing computational complexity, making the deployment of RIS and RSMA in vehicular communication networks a demanding yet promising endeavor.

In this paper, we propose a novel V2I communication system, named PARIS-RSMA-V2I (Partitioned Array and RIS with RSMA-based V2I system), tailored to support downlink communication for UEs inside moving vehicles. Additionally, we present associated low-complexity optimization algorithms aimed at maximizing the sum-rate. The PARIS-RSMA-V2I system strategically partitions the BS antenna array elements and RIS elements, ensuring that each pair is dedicated to transmitting data streams to UEs within a specific vehicle. This partitioning and the resulting subarray–subsurface–vehicle mapping enable low-complexity radio resource allocation by effectively addressing the heterogeneous channel characteristics and estimation errors across vehicles, while exploiting the similar channel conditions among UEs within the same vehicle. Furthermore, this approach, combined with per-vehicle encoding and decoding, enables efficient inter-vehicle interference suppression using zero forcing, allowing the system to concurrently support multiple user streams under RSMA.

The proposed radio resource allocation algorithms for the PARIS-RSMA-V2I system are designed to balance the trade-off between maximizing the overall sum rate and minimizing computational complexity. These algorithms optimize key aspects, including the partitioning of BS subarrays and RIS subsurfaces, precoding and beamforming at the BS, rate-splitting strategies, and passive beamforming at the RIS. To achieve this, we develop a block coordinate descent (BCD) algorithm that incorporates a dynamic partitioning scheme, dividing the BS antenna array and RIS elements into subplanes. The passive beamforming for each RIS subplane is optimized using the Riemannian conjugate gradient (RCG) method, while the precoding at the BS subarray is refined through a weighted minimum mean square error (WMMSE) approach.

Simulation results and comparisons with benchmarks demonstrate that the proposed partitioning scheme achieves robust performance in terms of system sum rate while significantly reducing computational complexity.

The main contributions of our work are summarized as follows:

- *Design of the PARIS-RSMA-V2I system architecture*: We propose a novel system in which BS antennas and RIS elements are partitioned, enabling a one-to-one mapping between BS antenna subarrays, RIS element subsurfaces, and per-vehicle RSMA stream groups. This design incorporates per-vehicle encoding and decoding, enabling efficient suppression of inter-vehicular interference using zero forcing, under speed-dependent heterogeneous channel characteristics and estimation errors.
- *Development of a dynamic partitioning algorithm*: We introduce a partitioning algorithm that adapts to the varying conditions of vehicles, dynamically dividing the available BS antenna arrays and RIS elements into subarrays and subsurfaces. Additionally, we provide a computational complexity analysis that demonstrates the effectiveness of the proposed algorithm in reducing complexity.
- *Development of hybrid precoding and beamforming optimization algorithms*: We formulate optimization algorithms where precoding and beamforming at the BS are addressed using the WMMSE and matched filter (MF) methods, respectively. Meanwhile, passive beamforming for the RIS subsurfaces is optimized using the RCG method.
- *Comprehensive simulation and performance evaluation*: We present an extensive simulation study that highlights the performance gains of the proposed algorithm compared to various benchmarks. The results demonstrate improvements in computational complexity and sum-rate under diverse network conditions and operational strategies.

The remainder of this article is organized as follows. Section II reviews and discusses recent works. Section III presents the system and channel models. The solution architecture and problem formulation are described in Section IV, while the associated solution algorithms are developed in Section V. Section VI provides the performance evaluation, including simulation environment settings, results, and discussions. Finally, concluding remarks are presented in Section VII.

Notations: $\mathbb{C}^{M\times N}$ denotes the space of $M \times N$ complex-valued matrices. For a square matrix $\mathbf{F}$, $\mathrm{Tr}(\mathbf{F})$ represents its trace. For a complex-valued vector $\mathbf{x}$, $|\mathbf{x}|$ denotes its Euclidean norm. Given a complex-valued scalar x, $\mathfrak{Re}(x)$ and $\mathfrak{Im}(x)$ denote its real and imaginary parts, respectively, while x^* represents its complex conjugate. The matrix $\mathbf{I}_N$ denotes an $N \times N$ identity matrix. The operator $\odot$ represents the Hadamard product. For an independent and identically distributed (i.i.d.) random variable (RV) s, $s \sim \mathcal{CN}(0, \sigma)$ indicates that s follows a complex Gaussian distribution with zero mean and variance σ. The frequently used notations are listed in Table II.

II. Related Work

Several studies have explored the interplay between RSMA and RIS. In [25], B. Clerckx et al. and in [26], A. S. de Sena et al. presented a general model for RIS-aided RSMA

TABLE I
SUMMARY OF RELATED WORKS ON RSMA+V2I, ANTENNA PARTITIONING, RIS PARTITIONING, AND RIS+V2I

Reference	Objective	V2X scenario	Multiple access scheme	RIS	RIS Partitioning	Antenna Partitioning	Heterogeneous vehicular speeds
Cai *et al.* [14]	Channel capacity maximization by optimizing transmit covariance and RIS reflection phase in a MIMO system		NOMA	Yes	Yes	No	No
Pala *et al.* [15]	Sum-rate maximization by optimizing precoder at BS, block-length of common and private symbols of each user, and passive beamforming at RIS in a URLLC system		RSMA	Yes	No	No	No
Duan *et al.,* [16]	Time-averaged maximization by optimizing active beamforming at the VS and passive beamforming at RIS based on statistical CSI feedback		OMA	Yes	No	No	Yes
Wu *et al.* [17]	Sum-rate maximization by optimizing analog and digital beamforming at BS, and passive beamforming at RIS using a transformer in a Terahertz MIMO system		RSMA	Yes	No	No	Yes
Khaleel *et al.* [18]	Spectrum efficiency enhancement by improving the ergodic rate and maximizing user fairness via the optimization of RIS partitioning in an environment where the BS and RIS serve users in different clusters		NOMA	Yes	Yes	No	No
Wang *et al.* [11]	Sum-rate maximization by optimizing the splitting of RIS elements into sub-surfaces and then optimizing the passive beamforming in each sub-surface in a large scale RIS elements based multi-user wireless network environment		NOMA	Yes	Yes	No	No
Kim *et al.* [19]	Transmit power and computational complexity minimization by optimizing the splitting of RIS elements into sub-surfaces in RIS assisted NOMA system		NOMA	Yes	Yes	No	No
Jiang *et al.* [20]	Aim to achieving the balance between the modeling accuracy and complexity by partitioning the RIS elements for the investigation of the small-scale fading characteristics of the RIS-assisted V2V channel	Yes	OMA	Yes	Yes	No	No
Wang *et al.,* [5]	Received signal power maximization by jointly optimizing the BS transmit precoding vector and RIS reflection coefficient matrix in a RIS assisted mmWave communication system		OMA	Yes	No	No	No
Gao *et al.,* [4]	Energy efficiency maximization by optimizing partitioning of the antenna array and SIC precoding in mmWave MIMO systems		OMA	No	No	Yes	No
Chen *et al.,* [7]	Link capacity maximization by jointly optimizing the BS transmit power, RIS reflection coefficient matrix and spectrum allocation in a RIS assisted V2I communication	Yes	OMA	Yes	No	No	No
Pala *et al.,* [21]	Sum-rate maximization by jointly optimizing the BS transmit power and RIS reflection coefficient matrix under ICSI with DRL in a RIS-aided Full-Duplex-RSMA V2I Communications	Yes	RSMA	Yes	No	No	Yes
Houran *et al.,* [22]	overall network sum-rate maximization by assigning either a single or multiple RIS to the BS-user pair in a RIS aided RAN V2I	Yes	OMA	Yes	No	No	No
Saikia *et al.,* [23]	Achievable sum-rate maximization by optimizing RIS elements phase shifts with DRL in a 2 RIS based FD 6G V2I communication	Yes	OMA	Yes	No	No	No
Cui *et al.,* [6]	channel capacity maximization by jointly optimizing active beamforming at the BS and passive beamforming at the RIS with a 2 phase transformer network architecture in a RIS assisted Internet of Vehicles	Yes	OMA	Yes	No	No	No
Zhang *et al.,* [24]	Energy efficiency maximization by jointly optimizing the RIS dynamic element partitioning and beamforming with DRL for green symbiotic communications in IoT		OMA	Yes	Yes	No	No
Wang *et al.,* [11]	sum-rate maximization by optimizing the RIS partitioning, passive beamforming of each beamforming and active beamforming at BS in Segmented RIS Assisted Multi-User Wireless Networks		OMA	Yes	Yes	No	No
This work	Sum-rate maximization by optimizing BS antenna and RIS partitioning, beamforming, and RIS phase shifts in a RSMA-RIS based V2I system	Yes	RSMA	Yes	Yes	Yes	Yes

systems and illustrated the rate regions under both perfect and imperfect channel conditions. Using simulation case studies, [25] showed that RIS can enable RSMA to ensure a private message is sent exclusively to the intended UE. Sama Wahb et al. in [27] further integrated RSMA with cooperative relaying and RIS to enhance the spectral efficiency of multi-user communications. They proposed a selective approach where certain users decode the common stream using successive interference cancellation (SIC) while treating private streams as noise, whereas other users treat the common stream as noise without decoding it. The authors employed the alternating optimization (AO) technique to jointly optimize time slot allocation, the precoding matrix, common rate allocation, and RIS phase shifts for both direct and cooperative transmission phases iteratively.

There have been some studies on the partitioning of RIS elements for OMA and NOMA, but not for RSMA. In line with the concept of separating messages intended for specific users, as discussed in [27] and [25], the work in [14] explored the partitioning of RIS into subsurfaces. This approach optimizes the phase shifts at the subsurface level to direct the reflected signal toward a desired direction, with the subsurface sizes adaptively adjusted based on channel conditions. The paper introduces a scalable solution to the joint active and passive beamforming design problem, adopting a linear phase variation structure for RIS partitioning to achieve anomalous reflection. Yonghwi Kim et al. in [19] proposed a closed-form partition-based multiple access (MA) scheme that simultaneously allocates RIS elements and determines the decoding order. The proposed method efficiently covers only the essential spatial area by forming multiple beams simultaneously, aligning suitable channels for each user to leverage NOMA. By exploiting the benefits of RIS-NOMA, the partition-based MA approach achieves low complexity and reduced transmit power. Aymen Khaleel in [18] proposed a RIS partitioning scheme in which the BS and RIS are dedicated to serving different clusters of users, with each subsurface utilized to serve a specific user. The authors formulated an RIS partitioning optimization problem to allocate RIS elements among users, aiming to maximize user fairness. Similarly, Xiaoqing Wang *et al.* in [11] tackled the sum-rate maximization problem under a RIS segmentation structure. Their framework adaptively adjusts the distribution and size of each segmentation to balance optimality and complexity in RIS-assisted multi-user systems. The system divides RIS arrays into subsurfaces and optimizes passive beamforming by calculating a reflection coefficient shared by all elements within a subsurface. In the work of Hao Jiang [20], the RIS is partitioned to achieve a planar wavefront assumption for the partitioned RIS subarrays. This approach balances modeling accuracy and computational complexity in analyzing the small-scale fading characteristics of RIS-assisted V2V channels.

Some studies have worked on the partitioning of BS antenna arrays. Sung Hoo Chae *et al.* in [28] proposed a general rate-splitting-enabled hybrid beamforming framework applicable to both fully connected and subarray hybrid beamforming structures. This framework enables flexible assignment of streams to each user by dividing streams into an arbitrary number of common and private substreams Private substreams are recoverable only by their intended user, while common substreams can be decoded by all users. The authors developed a joint optimization of analog and digital beamforming to transmit common and private substreams, maximizing both the sum rate and the minimum rate across multiple users. In the work presented by Zhiqiang Wei *et al.* in [3], where a beam-splitting technique dynamically partitions the entire antenna array associated with an RF chain into multiple subarrays, forming disjoint beams. This method creates multiple analog beams for each NOMA group of UEs, facilitating downlink NOMA transmission by leveraging channel sparsity and the large-scale antenna array at the BS.

The literature reviewed in this section highlights the promising advantages of integrating RIS and RSMA, and also shows that significant sum-rate performance can be achieved while reducing computational complexity by partitioning RIS and BS antenna elements. However, no prior work has explored the simultaneous partitioning of both BS antennas and RIS elements for the transmission of multiple user streams under RSMA. Furthermore, existing studies have not considered partitioning techniques tailored to V2I communication while accounting for the unique channel characteristics of UEs within vehicles. To the best of our knowledge, this work is the first to partition both BS antennas and RIS elements while forming beams using paired subarrays and subsurfaces, dynamically adapting to the heterogeneous conditions of vehicles. A detailed comparison of our work with closely related studies is provided in Table I.

III. System Model

The system's general model and communication architecture are illustrated in Fig. 1. We consider a downlink communication scenario where an M-antenna BS, with an antenna set denoted as $\mathcal{M} = \{1, \cdots, M\}$, serves K UEs, each equipped with L antennas, and placed in vehicles $\mathcal{V} = \{1, \cdots, V\}$. Each vehicle v contains a set of UEs denoted as $\mathcal{K}_v = \{1, \cdots, K_v\}$, such that the total number of UEs satisfies $\sum_{v \in \mathcal{V}} K_v = K$. We consider a scenario where the line-of-sight (LOS) paths between the BS and the vehicles are obstructed by obstacles such as buildings or other vehicles. To overcome this, we deploy an N-element RIS on a roadside unit (RSU) to establish virtual end-to-end (E2E) LOS links. Both the BS and RIS are equipped with uniform planar array (UPA) structures, with the RIS reflecting element set defined as $\mathcal{N} = \{1, \cdots, N\}$. To ensure that all vehicles can be served by the BS, we assume $M \geq V$ and $N \geq V$.

We define sets $\mathcal{A}$ and $\mathcal{R}$ to represent all subarrays and subsurfaces as $\mathcal{A} = \{\mathcal{A}_1, \cdots, \mathcal{A}_A\}$ and $\mathcal{R} = \{\mathcal{R}_1, \cdots, \mathcal{R}_S\}$, respectively, where the number of all subarrays is A and the number of all subsurfaces is S. The subsets $\mathcal{A}_a$ and $\mathcal{R}_s$ represent the elements within subarray a and subsurface s, respectively. Specifically, $\mathcal{A}_a = \{1, \cdots, M_a\}$ denotes the set of elements in subarray s, where m_a refers to antenna element m within subarray a. Similarly, $\mathcal{R}_s = \{1, \cdots, N_s\}$ represents the set of RIS elements within subsurface s, where n_s denotes RIS element n of subsurface s.

TABLE II
LIST OF NOTATIONS

Notation	Definition
$\mathcal{M}$	The set of transmit antennas at BS
M	Total number of transmit antennas at BS
$\mathcal{N}$	The set of reflecting elements at RIS
N	The total number of reflecting elements at RIS
K	Total number of UEs
L	Receive antennas at UE
$\mathcal{V}$	The set of all vehicles
V	Total number of vehicles
$\mathcal{K}_v$	The set of UEs within vehicle v
K_v	Number of UEs within vehicle v
$\mathcal{A}$	The set of subarrays
$\mathcal{R}$	The set of subsurfaces
A_s	Total number of subarrays
A_s	Total number of subsurfaces
$\mathcal{A}_s$	The set of antennas in subarray s
$\mathcal{R}_s$	The set of RIS elements in subsurface s
m_s	Antenna element in subarray set $\mathcal{A}_s$
n_s	RIS element in subsurface set $\mathcal{R}_s$
M_s	Total number of antennas in $\mathcal{A}_s$
N_s	Total number of RIS elements in subsurface $\mathcal{R}_s$
$\mathbf{H}_{b,s}$	Channels between subarray b and subsurface s
$\mathbf{H}_{s,k}$	Channels between subsurface s and UE k
$\eta_{b,s}$	The complex path gain of the link b and s
α_b	BS subarray response vector
α_s	RIS subsurface response vector
ϑ_b^D	Azimuth angle of departure at the BS subarray
φ_b^D	Elevation angle of departure at the BS subarray
ϑ_s^A	Azimuth angle of arrival at the RIS subsurface
φ_s^A	Elevation angle of arrival at the RIS subsurface
ϑ_k^A	Azimuth angle of arrival at the UE k
φ_k^A	Elevation angle of arrival at UE k
$\mathcal{B}_{b,s}$	Reference path gain at a distance of 1 meter
$d_{b,s}$	Distance between subarray b and subsurface s
f_D	Doppler frequency
v^k	Speed of UE k
λ	Wavelength of the system carrier frequency
$\mathbf{\Theta}$	Vector of reflection coefficients for all subsurfaces
$\mathbf{\Theta}_s$	Diagonal reflection coefficients of subsurface s
$\kappa_{s,n}$	Amplitude of the reflection coefficient of element n_s
$\theta_{s,n}$	Phase of the reflection coefficient of element n_s
$\boldsymbol{\delta}_s$	Reflection coefficients associated with subsurface s
$\delta_{s,n}$	Reflection coefficient associated with element n_s

A. Channel Models

We define the channel from BS subarray $\mathcal{A}_a$ to RIS subsurface $\mathcal{R}_s$ as $\mathbf{H}_{a,s} \in \mathbb{C}^{M_a \times N_s}$ and the channel from RIS subsurface $\mathcal{R}_s$ to UE k as $\mathbf{H}_{s,k} \in \mathbb{C}^{N_s \times L}$. These channels are modeled using the Saleh-Valenzuela mmWave channel model [29], [30], where each cluster (scatterer) is assumed to have a single propagation path. Since the RIS establishes a virtual line-of-sight (LOS) link, we assume this virtual LOS component dominates.

The channel between a transmitting point t and a receiving point r (from a BS antenna to a RIS element or from a RIS element to a UE antenna) is given by

$$h_{t,r} = \sqrt{M_t N_r}\frac{\sqrt{\mathcal{B}_{t,r}}}{d_{t,r}}\alpha_t(\vartheta_t^D, \varphi_t^D)\alpha_r(\vartheta_r^A, \varphi_r^A), \quad (1)$$

where M_t and N_r denote the number of transmitting and receiving antennas (or elements), respectively, $\mathcal{B}_{t,r}$ is the reference path gain at a distance of 1 meter (m), and $d_{t,r}$ represents the propagation distance between the transmitter and the receiver. Additionally, $\alpha_t(\vartheta_t^D, \varphi_t^D)$ is the transmitter response vector, where $(\vartheta_t^D, \varphi_t^D)$ denote the azimuth and elevation angles of departure (AOD) at the transmitter, while $\alpha_r(\vartheta_r^A, \varphi_r^A)$ is the receiver response vector, where $(\vartheta_r^A, \varphi_r^A)$ denote the azimuth and elevation angles of arrival (AOA) at the receiver.

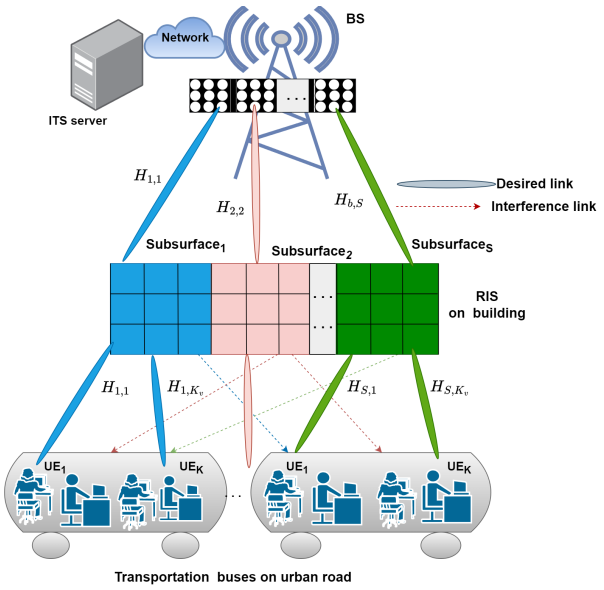


Fig. 1. System model

The RIS operates by reflecting incident signals using a reflection coefficient matrix, denoted as

$$\mathbf{\Theta} = [\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_s, \ldots, \mathbf{\Theta}_S], \quad (2)$$

where $\mathbf{\Theta}_s \in \mathbb{C}^{N_s \times N_s}$ represents the reflection coefficient matrix for subsurface s and is defined as

$$\mathbf{\Theta}_s = \text{diag}(\kappa_{s,n} e^{j\theta_{s,n}}), \quad \forall n_s \in \mathcal{R}_s, \quad (3)$$

where $\kappa_{s,n}$ and $\theta_{s,n}$ denote the reflection amplitude and phase, respectively, of element n_s. The vector $\boldsymbol{\delta}_s$ contains all the reflection coefficients for subsurface s and is expressed as

$$\boldsymbol{\delta}_s = [\delta_{s,1}, \cdots, \delta_{s,n}, \cdots, \delta_{s,N_s}], \quad \text{where} \quad \delta_{s,n} = \kappa_{s,n} e^{j\theta_{s,n}}. \quad (4)$$

We assume that the reflecting elements are ideally passive, meaning that $\kappa_{s,n} = 1$.

We define the cascaded channel $\mathbf{H}_{a,s,k} \in \mathbb{C}^{M_a \times L}$, which represents the channel from the subarray a of the BS to UE k through the subsurface s of the RIS as follows:

$$\mathbf{H}_{a,s,k}^H = \mathbf{H}_{s,k}^H \mathbf{\Theta}_s \mathbf{H}_{a,s}^H. \quad (5)$$

B. Imperfect Channel Estimation

We consider imperfect channel state information (CSI) due to vehicle mobility and delays in CSI reporting from the UE to the BS. Consequently, we express the channel in Eq. (5) at time t as

$$\mathbf{H}_{a,s,k}(t) = \epsilon \mathbf{H}_{a,s,k}(t-T) + \sqrt{1-\epsilon^2}\mathbf{e}_{a,s,k}(t), \quad (6)$$

where T is the time elapsed since the latest channel estimation, ϵ represents the time correlation coefficient, and $\mathbf{e}_{a,s,k}(t) \in \mathbb{C}^{M_a \times L}$ denotes the channel estimation error,

whose elements are independently and identically distributed (i.i.d.) as $\mathcal{CN}(0,1)$. The time correlation coefficient ϵ follows Jake's model, given by

$$\epsilon = J_0(2\pi f_D T), \tag{7}$$

where J_0 is the zeroth-order Bessel function and f_D represents the Doppler frequency, expressed as

$$f_D = \frac{-\upsilon_k}{\lambda} \sin\vartheta_k^A \sin\varphi_k^A, \tag{8}$$

where λ is the wavelength of the carrier frequency, and υ_k denotes the speed of the vehicle in which UE k is located.

IV. Solution Architecture and Problem Formulation

A. Architecture

Our proposed PARIS-RSMA-V2I system is designed to concurrently transmit multiple user streams to UEs within vehicles using RSMA, while ensuring both efficiency and low complexity. To achieve this, our solution incorporates the following key features:

- **Per-vehicle precoding:** Since UEs within the same vehicle share similar channel characteristics, including channel estimation errors, due to their common velocity, treating them as a group with both per-vehicle and per-UE precoders enhances processing efficiency and simplifies inter-vehicle interference suppression, as discussed next.
- **Inter-vehicle interference suppression:** Transmitting multiple user streams using non-orthogonal multiple access introduces inter-UE interference across vehicles. To mitigate this, the proposed solution applies per-vehicle precoding alongside zero-forcing precoders and decoders to effectively suppress inter-vehicle interference. Meanwhile, inter-UE interference within the same vehicle is managed using successive interference cancellation (SIC), following the standard RSMA decoding process.
- **Subarray–subsurface–vehicle mapping:** To maintain low complexity, our solution establishes a one-to-one mapping between BS antenna subarrays, RIS element subsurfaces, and vehicles. Each vehicle is assigned a dedicated subarray and subsurface, leading to a partitioning of M BS antennas and N RIS elements into S subplanes. Consequently, we set $A = S = V$ and use a unified notation s to indicate the paired subarray a and subsurface s. This leads to the notation $\mathbf{H}_{s,s,k}$ for $\mathbf{H}_{a,s,k}$ and allows the mapped vehicle index v and s to be used interchangeably, depending on the context and convenience.

The structure of PARIS-RSMA-V2I is illustrated in Fig. 2. In the RSMA framework, the BS multiplexes the messages intended for all UEs, where each message $x_{v,k}$ for UE k in vehicle v is split into two components: a common part $x_{v,k}^c$ and a private part $x_{v,k}^p$. The private messages are then encoded into private data streams, while the common parts from all UEs are aggregated and encoded into a single common message x^c. The common message stream is intended to be decoded by all UEs across all vehicles in the network, whereas each private part is exclusively decoded by its corresponding UE. All private parts intended for the UEs, along with the encoded common stream containing messages from all UEs, are superposed. To enable this process, a dedicated subarray at the BS and a corresponding subsurface of the RIS are assigned to each vehicle, ensuring efficient signal transmission and superposition.

Before transmission, a precoder $\mathbf{F}$ is applied, which is defined as

$$\mathbf{F} = [\mathbf{f}_1, \cdots, \mathbf{f}_v, \cdots, \mathbf{f}_V], \tag{9}$$

where $\mathbf{f}_v \in \mathbb{C}^{1+K_v}$ represents the precoding vector for vehicle v. The message streams x_v^c and $x_{v,k}^p$, for $k \in \mathcal{K}_v$, intended for vehicle v, are individually precoded using the vector

$$\mathbf{f}_v = [f_v^c, f_{v,1}^p, \cdots, f_{v,k}^p, \cdots, f_{v,K_v}^p], \tag{10}$$

where $f_v^c \in \mathbb{C}$ and $f_{v,k}^p \in \mathbb{C}$ denote the precoding coefficients for the common and private streams, respectively, of the message intended for UE k in vehicle v. After precoding, the streams are further processed using beamforming, represented by $\mathbf{G}$:

$$\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_v, \cdots, \mathbf{g}_V], \tag{11}$$

where $\mathbf{g}_v \in \mathbb{C}^{M_s \times 1}$ is the beamforming vector for the subarray s assigned to serve vehicle v. At the receiver, the signal is processed using a decoder vector $\mathbf{\Psi}_k \in \mathbb{C}^{L\times 1}$ for UE k.

B. Inter-Vehicle Interference Suppression

We employ a complete zero-forcing (CZF) precoding and decoding scheme [31] at the BS and UE, respectively, to eliminate inter-vehicle interference. To achieve this, we decompose the BS active precoders for both common and private streams, f_v^c and $f_{v,k}^p$ in Eq. (10), as well as the UE active decoder Ψ_k, into two components: one dedicated to eliminating the interference from other vehicles (denoted with a superscript z) and the other responsible for common and private stream beamforming (denoted with a superscript d). Accordingly, we define

$$f_v^c = f_v^{c-z} f_v^{c-d},\ f_{v,k}^p = f_{v,k}^{p-z} f_{v,k}^{p-d},\ \mathbf{\Psi}_k = \mathbf{\Psi}_k^z * \mathbf{\Psi}_k^d, \tag{12}$$

where $f_v^{c-z} \in \mathbb{C}$ and $f_v^{c-d} \in \mathbb{C}$ represent the ZF precoder and the data stream precoder for the common part of the message, respectively, while $f_{v,k}^{p-z} \in \mathbb{C}$ and $f_{v,k}^{p-d} \in \mathbb{C}$ correspond to the CZF precoder and the data stream precoder for the private part of the message, respectively, for UE k. Similarly, $\mathbf{\Psi}_k^z \in \mathbb{C}^{L\times 1}$ and $\mathbf{\Psi}_k^d \in \mathbb{C}^{L\times 1}$ denote the decoders for ZF and stream beamforming, respectively, and $*$ is an element-wise multiplication operator. Since the common and private streams intended for the same UE k experience the same inter-vehicular interference, both the corresponding ZF precoders f_v^{c-z} and $f_{v,k}^{p-z}$ take the same value, thus we denote $f_v^{c-z} = f_{v,k}^{p-z} = f_{v,k}^z$.

To null-out inter-vehicle interference, the following two conditions must be satisfied:

$$\begin{aligned} &\mathbf{\Psi}_k^z \mathbf{H}_{a,s,k} f_{v,k}^z \neq 0, \forall k \in \mathcal{K}_v, \\ &\mathbf{\Psi}_{k'}^z \mathbf{H}_{a,s,k'} f_{v,k}^z = 0, \forall k' \in \mathcal{K} - \mathcal{K}_v. \end{aligned} \tag{13}$$

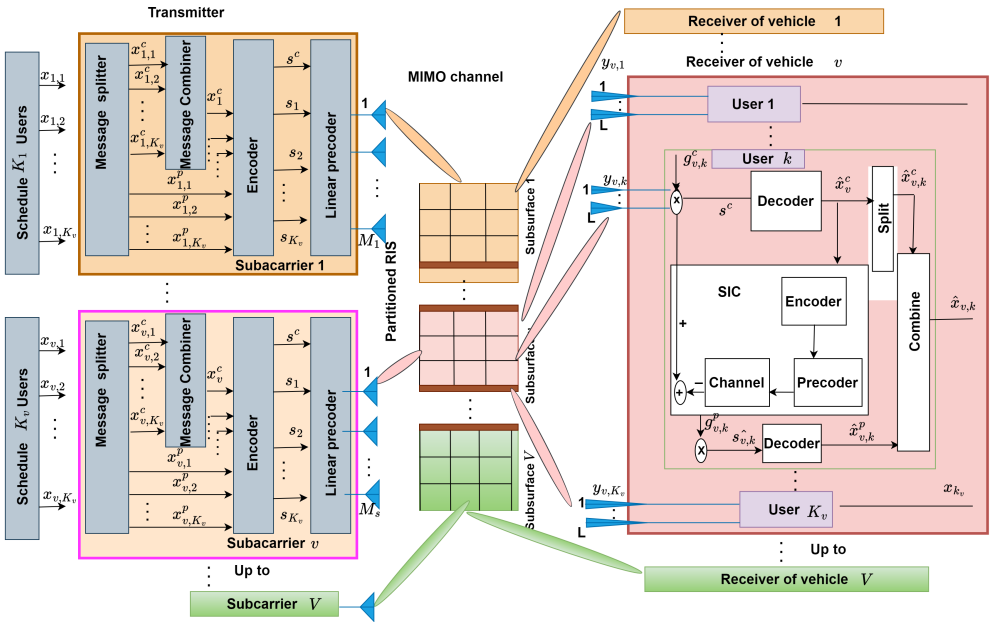


Fig. 2. RSMA encoding and decoding model

A sufficient condition for satisfying the second criterion in Eq. (13) is given by

$$\mathbf{H}_{a,s'} f^z_{v,k} = 0 \quad \text{and} \quad \mathbf{\Psi}^z_{k'} \mathbf{H}_{s,k'} = 0, \tag{14}$$

where the first condition indicates that the interference from a BS subarray to unintended subsurfaces (paired with other vehicles) is zero, and the second condition indicates that the interference from a subsurface to other-vehicle users is zero.

The feasible regions for $\mathbf{\Psi}^z_k$ and $f^z_{v,k}$ can be determined using the horizontally concatenated matrices $\mathbf{P}_{s,k} \in \mathbb{C}^{M_a \times N_s(S-1)}$ and $\mathbf{D}_{s,k} \in \mathbb{C}^{N_s \times L(K-1)}$, which aggregate the channels as

$$\begin{aligned} \mathbf{P}_{s,k} &= [\mathbf{H}_{a,\mathcal{R}_1}, \cdots, \mathbf{H}_{a,s'}, \cdots, \mathbf{H}_{a,\mathcal{R}_S}], s' \neq s, \\ \mathbf{D}_{s,k} &= [\mathbf{H}_{s,1}, \cdots, \mathbf{H}_{s,k'}, \cdots, \mathbf{H}_{s,K}], k' \neq k. \end{aligned} \tag{15}$$

According to the singular value decomposition (SVD) entropy theorem in [32], for any real or complex matrix $\mathbf{H} \in \mathbb{C}^{m \times n}$, there exist unitary matrices $\mathbf{U} \in \mathbb{C}^{m \times m}$ and $\mathbf{V} \in \mathbb{C}^{n \times n}$, which are also orthogonal, such that

$$\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H, \tag{16}$$

where $\mathbf{\Sigma} \in \mathbb{C}^{m \times n}$ is a diagonal matrix containing the singular values of $\mathbf{H}$. Therefore, the SVD of $\mathbf{P}_{s,k}$ and $\mathbf{D}_{s,k}$ can be expressed as

$$\mathbf{P}_{s,k} = \mathbf{U}_P \mathbf{\Sigma}_P \mathbf{V}_P^H = \begin{bmatrix} \mathbf{U}_P^1 & \mathbf{U}_P^0 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_P^1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_P^1 \\ \mathbf{V}_P^0 \end{bmatrix},$$

$$\mathbf{D}_{s,k} = \mathbf{U}_D \mathbf{\Sigma}_D \mathbf{V}_D^H = \begin{bmatrix} \mathbf{U}_D^1 & \mathbf{U}_D^0 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_D^1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_D^1 \\ \mathbf{V}_D^0 \end{bmatrix}. \tag{17}$$

Based on the interference elimination criteria using channel decomposition via SVD, as applied in [31], inter-vehicle interference is suppressed by assigning

$$f^z_{v,k} = V_p^0, \quad \mathbf{\Psi}^z_k = \mathbf{U}_D^0, \quad \forall v \in \mathcal{V}, k \in \mathcal{K} \tag{18}$$

where V_p^0 is the first element of $\mathbf{V}_p^0$.

After computing these per-vehicle precoders and decoders, the remaining precoders and decoders including the per-UE ones, namely f^{c-d}_v, $f^{p-d}_{v,k}$, and $\mathbf{\Psi}^d_k$, are formulated as part of the main optimization problem and solved subsequently.

C. Received Signal and SINR

The received signal at UE k in vehicle v is given by

$$\begin{aligned} y_k = & \underbrace{\mathbf{\Psi}_k \mathbf{H}^H_{a,s,k} \mathbf{g}_v f^c_v x^c}_{\text{Common signal for this vehicle}} + \underbrace{\mathbf{\Psi}_k \mathbf{H}^H_{a,s,k} \sum_{l \in \mathcal{K}_v} \mathbf{g}_v f^p_{v,l} x^p_{v,l}}_{\text{Private signals for this vehicle}} + \\ & \underbrace{\sum_{m \in \mathcal{M} \setminus a} \sum_{n \in \mathcal{N} \setminus s} \mathbf{\Psi}_k \mathbf{H}^H_{m,n,k} \sum_{u \in \mathcal{V} \setminus v} \left(\mathbf{g}_u f^c_u x^c + \sum_{l \in \mathcal{K}_u} \mathbf{g}_u f^p_{u,l} x^p_{u,l} \right)}_{\text{Signals for other vehicles}} \\ & + n_k \end{aligned} \tag{19}$$

where $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ denotes complex Gaussian noise at UE k with power σ_k^2. The signal-to-interference-plus-noise ratio (SINR) for decoding the common stream at UE k is given by

$$\gamma^c_k = \frac{|\mathbf{\Psi}_k \mathbf{H}^H_{a,s,k} \mathbf{g}_v f^c_v|^2}{\sum_{l \in \mathcal{K}_v} |\mathbf{\Psi}_k \mathbf{H}^H_{a,s,l} \mathbf{g}_v f^p_{v,l}|^2 + I^p_{v,k} + \sigma^2_k} \tag{20}$$

where $I^p_{v,k}$ is the interference power from private signal streams of vehicles other than v, given by

$$I^p_{v,k} = \sum_{u \in \mathcal{V} \setminus v} \left(\sum_{l \in \mathcal{K}_u} |\mathbf{\Psi}_k \mathbf{H}^H_{a,s,k} \mathbf{g}_u f^p_{u,l}|^2 \right). \tag{21}$$

Similarly, the SINR for decoding the private stream at UE k is expressed as

$$\gamma^p_k = \frac{|\mathbf{\Psi}_k \mathbf{H}^H_{a,s,k} \mathbf{g}_v f^p_{v,k}|^2}{\sum_{l \neq k} |\mathbf{\Psi}_k \mathbf{H}^H_{a,s,k} \mathbf{g}_v f^p_{v,l}|^2 + I^p_{v,k} + \sigma^2_k}. \tag{22}$$

We recall that the channel considered in the received signal equations above consists of both the true part $\epsilon \mathbf{H}_{a,s,k}(t-T)$ and the error part $\sqrt{(1-\epsilon^2)}\mathbf{e}_{a,s,k}(t)$. However, when analyzing the channel components separately, the expressions for the SINR are modified as follows. The SINR for decoding the common stream at UE k becomes

$$\gamma_k^c = \frac{|\boldsymbol{\Psi}_k \epsilon \mathbf{H}_{a,s,k}^H \mathbf{g}_v f_v^c|^2 + |\sqrt{(1-\epsilon^2)}\boldsymbol{\Psi}_k \mathbf{e}_{a,s,k}^H \mathbf{g}_v f_v^c|^2}{\sum_{l\in\mathcal{K}_v}(|\epsilon \boldsymbol{\Psi}_k \mathbf{H}_{a,s,k}^H \mathbf{g}_v f_{v,l}^p|^2 + |\sqrt{(1-\epsilon^2)}\boldsymbol{\Psi}_k \mathbf{e}_{a,s,k}^H \mathbf{g}_v f_{v,l}^p|^2) + I_{v,k}^p + \sigma_k^2}. \tag{23}$$

Similarly, the SINR for decoding the private stream at UE k is given by

$$\gamma_k^p = \frac{|\epsilon \boldsymbol{\Psi}_k \mathbf{H}_{a,s,k}^H \mathbf{g}_v f_{v,k}^p|^2 + |\sqrt{(1-\epsilon^2)}\boldsymbol{\Psi}_k \mathbf{e}_{a,s,k}^H \mathbf{g}_v f_{v,k}^p|^2}{\sum_{l\neq k}(|\epsilon \boldsymbol{\Psi}_k \mathbf{H}_{a,s,k}^H \mathbf{g}_v f_{v,l}^p|^2 + |\sqrt{(1-\epsilon^2)}\boldsymbol{\Psi}_k \mathbf{e}_{a,s,k}^H \mathbf{g}_v f_{v,l}^p|^2) + I_{v,k}^p + \sigma_k^2}. \tag{24}$$

The interference $I_{v,k}^p$ resulting from signal streams of vehicles other than v under CZF is expressed as

$$I_v = \sum_{u\in\mathcal{V}\setminus v}\left(\sum_{l\in\mathcal{K}_u} |\sqrt{(1-\epsilon^2)}\boldsymbol{\Psi}_k \mathbf{e}_{c,t,k}^H \mathbf{g}_u f_{u,l}^p|^2\right). \tag{25}$$

The updated expressions in Eqs. (23), (24), and (25) result from decomposing the overall channel in Eq. (6) as

$$\begin{aligned}&|\boldsymbol{\Psi}_k \mathbf{H}_{a,s,k}^H \mathbf{g}_v f_{v,k}^p|^2 \approx \\ &|\epsilon \boldsymbol{\Psi}_k \mathbf{H}_{a,s,k}^H \mathbf{g}_v f_{v,k}^p|^2 + |\sqrt{(1-\epsilon^2)}\boldsymbol{\Psi}_k \mathbf{e}_{a,s,k}^H \mathbf{g}_v f_{v,k}^p|^2,\end{aligned} \tag{26}$$

where the approximation holds because the product terms involving both the true and error components are neglected due to the fact that as $0 \leq \epsilon < 1$, the product of itself and $\sqrt{(1-\epsilon^2)}$ gets even smaller, thus can be neglected [33]. A similar approximation applies to the common stream component. However, when deriving the expressions involving precoders and decoders from other vehicles, the true channel component vanishes due to CZF, leaving only the error component. This follows from the CZF design, which eliminates interference from other vehicles by canceling the true channel component while the residual error remains [33]. Consequently, the interference term $I_{v,k}^p$ contains only the error terms, as the BS can track and remove the true components using CZF.

For brevity, we will refer to the two components of the channel $\epsilon \mathbf{H}_{a,s,k}^H$ and $\sqrt{(1-\epsilon^2)}\mathbf{e}_{a,s,k}^H$ simply as $\mathbf{H}_{a,s,k}^H$ and $\mathbf{e}_{a,s,k}^H$, respectively, omitting the explicit notation of the scaling factors ϵ and $\sqrt{(1-\epsilon)}$. But, they will be implicitly included in all terms involving their respective channel components.

The rates of the common and private streams of UE k are obtained as

$$r_k^c = w\log_2(1+\gamma_k^c), \quad r_k^p = w\log_2(1+\gamma_k^p). \tag{27}$$

To ensure successful decoding of the common stream x_v^c by all UEs in $\mathcal{V}$, its encoding rate is set as the minimum rate among all UEs:

$$R^c = \min_{k\in\mathcal{V}} r_k^c. \tag{28}$$

Since the common message transmitted by the BS contains parts of messages from all UEs, the rate allocation of R^c among users is proportional to the contribution of each UE's message to the common stream. Letting C_k denote the portion of R^c allocated to UE k for its common message, the total common rate allocated across all UEs is

$$R^c = \sum_{k\in\mathcal{V}} C_k. \tag{29}$$

The total rate for UE k is the sum of its allocated common rate and its private rate:

$$R_k = C_k + r_k^p. \tag{30}$$

Defining $\bar{k} \triangleq \arg\min_{k\in\mathcal{K}}(r_k^c)$, and assuming equal distribution of R^c among all UEs in all vehicles, we set $C_k = \frac{1}{K}\min_{k\in\mathcal{V}} r_k^c$, where $K = \sum_v K_v$. Thus, the total rate for UE k is

$$R_k = \frac{1}{K} r_{\bar{k}}^c + r_k^p. \tag{31}$$

The total system capacity is the sum of the private rates of all UEs and the minimum common stream rate:

$$C = \sum_{v\in\mathcal{V}} C_v \tag{32}$$

where $C_v = \sum_{k\in\mathcal{K}_v}(r_k^p + \frac{1}{K} r_{\bar{k}}^c)$.

D. Problem Formulation

We aim to maximize the system capacity C for V2I communication under BS antennas and RIS partitioning and vehicle mapping. To achieve this, we optimize subarray partitioning $\mathcal{A}$, subsurface partitioning $\mathcal{R}$, the RIS reflection coefficient matrix (passive beamforming) $\boldsymbol{\Theta}$, the precoder $\mathbf{F}$ and decoder matrix $\boldsymbol{\Psi}_k$, as well as the beamforming $\mathbf{G}$. The optimization problem is formulated as follows:

$$\begin{aligned}
\mathcal{P}: \max_{\mathcal{A},\mathcal{R},\boldsymbol{\Theta},\boldsymbol{\Psi}_k^d,\mathbf{F},\mathbf{G}} \quad & C && \text{(33a)}\\
\text{s.t. } & |\delta_{s,n}| = 1 && \text{(33b)}\\
& 0 \leq \theta_{s,n} \leq \pi && \text{(33c)}\\
& \|[g_v]_{j,r}\|_F^2 = 1, j = 1,\dots M_s, && \text{(33d)}\\
& \sum_v \text{tr}(\mathbf{W}_v \mathbf{W}_v^H) \leq P_t && \text{(33e)}\\
& R^c = \min_{k\in\mathcal{V}} r_k^c, && \text{(33f)}\\
& \sum_{s\in\mathcal{S}} M_s = M, \ \sum_{s\in\mathcal{S}} N_s = N, && \text{(33g)}
\end{aligned}$$

where $\mathbf{W}_v := \mathbf{g}_v(\mathbf{f}_{v,k}^{p-d})^H$. In Eq. (33), constraint (33b) ensures that the reflection coefficients of the RIS elements have a unit modulus. Constraint (33c) enforces the reflection phases of the RIS elements to be within the range $[0,\pi]$. Constraint (33d) ensures that the beamformer applies constant modulus phase shifts to the antenna beams, maintaining their normalization to one. Here, r represents the r-th RF chain. Constraint (33e) limits the total active beamforming power of the precoder so that it does not exceed the allocated power for a BS subarray. Constraint (33f) guarantees that the common stream rate is equal to the minimum rate among all UEs in all vehicles, ensuring that every UE can successfully decode the common stream. Constraint (33g) enforces that the sum of all subarray antenna elements matches the total BS antenna

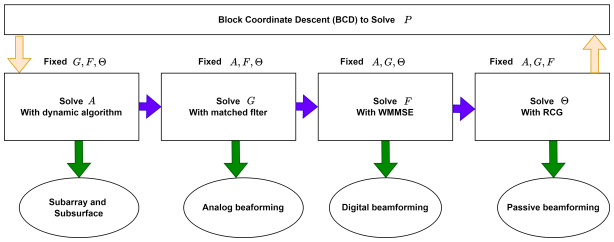


Fig. 3. BCD diagram to solve the formulated problem

budget, and the total number of RIS elements is sufficient to cover all RIS subsurface elements. This ensures proper resource allocation and system configuration.

V. Solution Algorithms

The problem (33) is inherently non-convex due to its non-convex objective function and constraints, making it difficult to solve directly. To tackle this challenging optimization problem, we decompose it into two subproblems, $\mathcal{P}1$ and $\mathcal{P}2$. In $\mathcal{P}1$, we focus on optimizing $\mathcal{A}$, $\mathcal{R}$, and the beamformer $\mathbf{G}$, while keeping the RIS phase shift $\mathbf{\Theta}$ and the precoder and decoder $\mathbf{F}$ and $\mathbf{\Psi}$ fixed. In $\mathcal{P}2$, we concentrate on optimizing the passive and active beamforming at the level of a single subarray-subsurface pair. This involves determining the optimal values for $\mathbf{\Theta}s$ and $\mathbf{\Psi}_k^d$ along with $\mathbf{f}_{v,k}^d$, as introduced in Eq. (12), for each subarray-subsurface pair.

The BCD diagram to solve the whole problem is shown in Fig.3.

A. Subarray and Subsurface Partitioning Algorithm

The subarray and subsurface partitioning scheme developed in this subsection dynamically allocates the available BS antenna elements and RIS elements to different vehicles in the V2I network at each transmission interval T_c. To determine this allocation, the algorithm evaluates the system capacity C, as defined in Eq. (32), and assigns an antenna element m and its corresponding RIS element n to a vehicle based on the resulting capacity gain.

However, assigning more antennas to a particular vehicle results in a higher SINR and, consequently, greater capacity compared to other vehicles. Additionally, the first allocated antenna contributes the largest SINR increment, leading to a greater initial capacity gain. As a result, maximizing system capacity alone may violate Jain's fairness criterion [34]. To balance user fairness with the objective of capacity maximization, we address this problem by considering the maximum capacity increment Δ_C [9].

Our algorithm sets the number of UEs in each vehicle as the threshold for the minimum number of antenna elements and RIS elements that must be allocated to the corresponding vehicle, subarray, and subsurface. This initial allocation serves as the first condition for partitioning fairness, as the number of UEs determines the required number of antennas in any MIMO communication system [9]. We denote this threshold as A_{th}^s for subarrays and R_{th}^s for subsurfaces.

The partitioning algorithm is presented in Algorithm 1. Its objective is to optimize the allocation of additional antenna and RIS elements to each vehicle to maximize the objective function. We define τ_a and τ_s as a step size vector representing the additional antennas and RIS elements assigned to each subarray and subsurface, respectively. The problem is formulated as follows.

$$\mathcal{P}1: \max_{\tau_a,\tau_s,\mathbf{G}} \Delta_C \tag{34a}$$

$$\text{s.t. } (33d), (33e), (33g) \tag{34b}$$

$$\mathcal{A} = \{A_{th}^1 + \tau_1, \ldots, A_{th}^a + \tau_a, \ldots, A_{th}^A + \tau_A\} \tag{35}$$

$$\mathcal{R} = \{R_{th}^1 + \tau_1, \ldots, R_{th}^s + \tau_s, \ldots, R_{th}^S + \tau_S\} \tag{36}$$

$$\Delta_C = C(A_{th}^a + \tau_s, R_{th}^s + \tau_s) - C(A_{th}^a, R_{th}^s) \tag{37}$$

We address the problem (34) by optimizing the vector $\boldsymbol{\tau} = [\tau_1, \cdots, \tau_a, \cdots, \tau_A]$ and $\boldsymbol{\tau} = [\tau_1, \cdots, \tau_s, \cdots, \tau_S]$ to maximize Δ_C, which is achieved by iteratively increasing the step size τ_a and τ_s for each subarray and subsurface, respectively.

By allocating each antenna and RIS elements to an RF chain based on the maximum increment of Δ_C, we ensure a balance between user (vehicle) fairness and the maximization of the system's total capacity [9].

Upon incrementing τ_a and τ_s for each vehicle, we compute the system's capacity increment and iteratively identify the vehicle (subarray and subsurface) that achieve the highest increase in Δ_C. The subarray and subsurface associated with the greatest capacity gain is assigned the additional antenna and RIS elements, as it provides the most significant improvement in Δ_C compared to other subarrays and subsurfaces. The algorithm continues until all available antennas and RIS elements are allocated to their respective subarrays and subsurface, at which point the optimized allocation vector is denoted as $\boldsymbol{\tau}_a^*$ and $\boldsymbol{\tau}_s^s$.

To track the remaining unpartitioned antenna elements, we define the variable α, which is initialized as $\max\{\alpha_a = M - \sum_{v\in\mathcal{V}} A_{th}^V, \alpha_s = N - \sum_{v\in\mathcal{V}} R_{th}^V\}$.

The precoding and decoding values are fixed using zero-forcing techniques, while the passive beamforming values are set as constants determined by the BS.

1) Beamforming: To optimize beamforming, we adopt the matched filter (MF) approach, considering the practical limitation that the phase shifts are quantized into B bits. To maximize the channel gain between the subarray elements and the subsurface reconfigurable elements, the optimal set of beamformers is formulated as follows:

$$\begin{aligned} &g_v = e^{-j\frac{2\pi n_0}{B}}, \\ &n_0 = \arg\min_{n\in\{0,1,\cdots,2^B-1\}} |\angle h_{m_s,n_s} - \tfrac{2\pi n}{B}|, \\ &\forall m_s \in \mathcal{A}_s, \forall n_s \in \mathcal{R}_s \end{aligned} \tag{38}$$

where h_{m_s,n_s} represents the channel between the subarray antenna m_s and the subsurface IRS element n_s. Further details on this formulation can be found in [17].

Algorithm 1 Dynamic subarray and Subsurface partitioning

Initialize: M and N
Initialize $\mathcal{K}_v$
Initialize A as $A_0 = \{A_{th}^1, \ldots, A_{th}^A\}$
Initialize R as $R_0 = \{R_{th}^1, \ldots, R_{th}^S\}$
Initialize optimal τ_a as $\tau_a^{Opt} = \{0, \ldots, 0\}$
Initialize optimal τ_s as $\tau_s^{Opt} = \{0, \ldots, 0\}$
Initialize Δ_C by using Eq. 37 with τ_{Opt}
for $i = 1 : \alpha$ **do**
  for $v = 1 : V$ **do**
    Increment τ_a^{Opt} and τ_s^{Opt} by 1
    Update C_v by using the incremented τ_a^{Opt} and τ_s^{Opt}
    $\Delta_C \leftarrow \Delta_{C_v}$
    record argmax Δ_C as v_{best}
    Increment τ_a^{Opt} of v_{best} by 1
    Increment τ_s^{Opt} of v_{best} by 1
  end for
end for

In the unpartitioned antenna scenario, all vehicles are served by a single beam formed using all available antennas. In this case, the beamforming is given by

$$g = \frac{1}{\sqrt{M}}[1, \cdots, e^{j(M)\frac{2\pi}{\lambda} d \sin \vartheta_b^D}]. \tag{39}$$

Now that we have solved the subarray antenna and subsurface element assignment, as well as the beamforming, we proceed to solve for the active and passive precoding along with receiver decoding. By utilizing the relationship in Eq. (3), we can replace $\mathbf{\Theta}_s$ with $\boldsymbol{\delta}_s$. Consequently, the optimization problem is reformulated as

$$\mathcal{P}^2 : \max_{\boldsymbol{\delta}_s, f_{v,k}^d, \Psi_k^d} C_v(\boldsymbol{\delta}_s, f_{v,k}^{c-d}, f_{v,k}^{p-d}, \Psi_k^d) \tag{40a}$$

$$\text{s.t. (33b), (33c), (33e)} \tag{40b}$$

To solve this problem, we first optimize $\boldsymbol{\delta}_s$ while keeping $f_{v,k}^{c-d}, f_{v,k}^{p-d}$, and Ψ_k^d fixed.

B. Optimization of RIS Phase Shifts

We address the optimization problem by first solving for the RIS subsurface reflection coefficient $\boldsymbol{\delta}_s$, while keeping the remaining variables $f_{v,k}^d$ and Ψ_k^d fixed. This reformulates the problem as follows:

$$\mathcal{P}^2 : \max_{\boldsymbol{\delta}_s} C_v(\boldsymbol{\delta}_s, f_{v,k}^{c-d}, f_{v,k}^{p-d}, \Psi_k^d) \tag{41a}$$

$$\text{s.t. (33b), (33c)}$$

where

$$C_v = \sum_k (\log_2(1+\frac{|\Psi_k^d \mathbf{H}_{s,k}^{eff} g_v f_{v,k}^{p-d}|^2 + |\Psi_k^d e_{s,k}^{eff} g_v f_{v,k}^{p-d}|^2}{\sum_{i\in\mathcal{K}, i\neq k}(|\Psi_i^d \mathbf{H}_{s,k}^{eff} g_v f_{v,i}^{p-i}|^2 + |\Psi_i^d e_{s,k}^{eff} g_v f_{v,i}^{p-i}|^2) + I_v + \sigma_k})) + (\log_2(1+\frac{|\Psi_k^d \mathbf{H}_{s,\bar{k}}^{eff} g_v f_{v,\bar{k}}^{c-d}|^2 + |\Psi_k^d e_{s,\bar{k}}^{eff} g_v f_{v,\bar{k}}^{c-d}|^2}{\sum_{i\in\mathcal{K}}(|\Psi_k^d \mathbf{H}_{s,\bar{k}}^{eff} g_v f_{v,k}^{p-i}|^2 + |\Psi_k^d e_{s,\bar{k}}^{eff} g_v f_{v,k}^{p-i}|^2) + I_v + \sigma_k})) \tag{42}$$

where the effective channel and channel estimation error are defined as follows:

$$\mathbf{H}_{s,k}^{eff} \triangleq \Psi_k^z \mathbf{H}_{s,k} \text{diag}(\boldsymbol{\delta}_s) \mathbf{H}_{b,s,k} g_v f_{v,k}^z, \forall k \in \mathcal{K} \tag{43}$$

$$e_{s,k}^{eff} \triangleq \Psi_k^z e_{s,k} \text{diag}(\boldsymbol{\delta}_s) e_{b,s,k} g_v f_{v,k}^z, \forall k \in \mathcal{K}. \tag{44}$$

To further simplify the problem, we introduce the following quantities:

$$\begin{aligned} \upsilon_k^p &\triangleq (\Psi_k^d \text{diag}(\Psi_k^z \mathbf{H}_{s,k}^H) \mathbf{H}_{b,s,k} f_{v,k}^z g_v f_{v,k}^{p-d}) \\ \upsilon_{ke}^p &\triangleq (\Psi_k^d \text{diag}(\Psi_k^z e_{s,k}^H) e_{b,s,k} f_{v,k}^z g_v f_{v,k}^{p-d}) \\ \upsilon_k^c &\triangleq (\Psi_k^d \text{diag}(\Psi_k^z \mathbf{H}_{s,k}^H) \mathbf{H}_{b,s,k} f_{v,k}^z g_v f_{v\bar{k}}^{c-d}) \\ \upsilon_{ke}^c &\triangleq (\Psi_k^d \text{diag}(\Psi_k^z e_{s,k}^H) e_{b,s,k} f_{v,k}^z g_v f_{v\bar{k}}^{c-d}). \end{aligned} \tag{45}$$

With these definitions, problem (41) can be rewritten as

$$\mathcal{P}^2 : \max_{\boldsymbol{\delta}_s} C_v(\boldsymbol{\delta}_s, f_{v,k}^{c-d}, f_{v,k}^{p-d}, \Psi_k^d) \tag{46a}$$

$$\text{s.t. (33b), (33c)} \tag{46b}$$

where the total sum rate associated with a single vehicle is expressed as

$$C_v = \sum_k (\log_2\left(1 + \frac{|\upsilon_k^p \boldsymbol{\delta}_s|^2 + |\upsilon_{ke}^p \boldsymbol{\delta}_s|^2}{I_v^p + I_{v'} + \sigma_k}\right) + \log_2\left(1 + \frac{|\upsilon_k^c \boldsymbol{\delta}_s|^2 + |\upsilon_{ke}^c \boldsymbol{\delta}_s|^2}{I_v^c + I_v + \sigma_k}\right)) \tag{47}$$

where the interference terms are given by

$$I_{v'} = \sum_{v'\in\mathcal{V}, v'\neq v} \sum_{k\in\mathcal{K}_{v'}} |\upsilon_{v',ke}^p \delta_{v'}|^2, \tag{48}$$

$$I_v^p = \sum_{i\in\mathcal{K}, i\neq k} (|\upsilon_i^p \boldsymbol{\delta}_s|^2 + \upsilon_{ie}^p \boldsymbol{\delta}_s|^2), \tag{49}$$

$$I_v^c = \sum_{i\in\mathcal{K}} (|\upsilon_i^c \boldsymbol{\delta}_s|^2 + |\upsilon_{ie}^c \boldsymbol{\delta}_s|^2). \tag{50}$$

Since the problem (46) is continuously differentiable with respect to $\boldsymbol{\delta}_s$, and the constraint (33b) forms a complex cyclic manifold [35], we adopt the Riemannian Conjugate Gradient (RCG) method to determine the optimal phase shifts. Similar to [35], the constraints in (46) can be interpreted as a Riemannian manifold. The rich geometric structure of Riemannian manifolds allows for the definition of gradients for cost functions, enabling efficient optimization. The RCG method consists of three key steps, which are developed as follows. The complete algorithm for computing the phase shift matrix is detailed in Algorithm 2.

a) Riemannian Gradient: The Euclidean gradient of C_k is evaluated as follows:

$$\Delta C_v(l+1) = 2(R_{\bar{k}c} + \sum_{k\in\mathcal{K}} R_{kp})$$

where

$$R_{\bar{k}c} = \left(\frac{\left((v_k^c)^H (v_k^c)^H \boldsymbol{\delta}_s\right) + \left((v_{ke}^c)^H (v_{ke}^c)^H \boldsymbol{\delta}_s\right) + \frac{\Delta I_v^c}{2} + \frac{\Delta I_{v'}}{2}}{|(v_k^c)^H \boldsymbol{\delta}_s|^2 + |(v_{ke}^c)^H \boldsymbol{\delta}_s|^2 + I_v^c + I_{v'} + \sigma_{\bar{k}}} \right) - \left(\frac{\frac{\Delta I_v^c}{2} + \frac{\Delta I_{v'}}{2}}{I_v^c + I_{v'} + \sigma_{\bar{k}}} \right) \tag{51}$$

and

$$R_{kp} = \left(\frac{\left((v_k^p)^H (v_k^p)^H \boldsymbol{\delta}_s\right) + \left((v_{ke}^p)^H (v_{ke}^p)^H \boldsymbol{\delta}_s\right) + \frac{\Delta I_v^p}{2} + \frac{\Delta I_{v'}}{2}}{|(v_k^p)^H \boldsymbol{\delta}_s|^2 + |(v_{ke}^p)^H \boldsymbol{\delta}_s|^2 + I_v^p + I_{v'} + \sigma_k} \right) - \left(\frac{\frac{\Delta I_v^p}{2} + \frac{\Delta I_{v'}}{2}}{I_v^p + I_{v'} + \sigma_k} \right).$$

where l represents the iteration step.

Next, $\Delta C_v(l+1)$ is projected orthogonally onto the complex circle to obtain the Riemannian gradient:

$$\begin{aligned} &\text{grad}C_v = \text{Project}_{\boldsymbol{\delta}_s}(\Delta C_v(l+1)) \\ &= \Delta C_v(l+1) - \mathcal{R}[\Delta C_v(l+1) \odot \boldsymbol{\delta}_s^*(l+1)] \odot \boldsymbol{\delta}_s(l+1) \end{aligned} \tag{52}$$

where $\boldsymbol{\delta}_s^*$ denotes the complex conjugate of $\boldsymbol{\delta}_s$.

b) Search direction: The search direction is defined as the tangent vector conjugate to the Riemannian gradient, which is evaluated as

$$d(l+1) = -gradC_v(l+1) + \mathcal{W}_1(l+1)\mathcal{J}(l+1) \tag{53}$$

where $\mathcal{W}_1$ represents the Polak-Ribiére parameter, calculated as given in [36]:

$$\mathcal{W}_1(l+1) = \max\left\{ \frac{(\Delta C_v(l+1))^H [\Delta C_v(l+1) - \Delta C_v(l)]}{(\Delta C_v(l+1))^H \Delta C_v(l)}, 0 \right\}. \tag{54}$$

The term $\mathcal{J}$ denotes the transport function, which is obtained as

$$\mathcal{J}(l+1) = d(l) - \mathcal{R}[d(l) \odot \boldsymbol{\delta}_s^*(l+1)] \odot \boldsymbol{\delta}_s(l+1). \tag{55}$$

c) Retraction: To update the optimization variable, the tangent vector, known as a transport vector, is mapped from one tangent space to another. In other words, this transport vector is projected back onto the complex circle manifold, and the new $\boldsymbol{\delta}_s$ is updated as

$$\boldsymbol{\delta}_s(l+1) \leftarrow \frac{\boldsymbol{\delta}_s(l) + \mathcal{W}_2(l)d(l)}{|\boldsymbol{\delta}_s(l) + \mathcal{W}_2(l)d(l)|} \tag{56}$$

where $\mathcal{W}_2$ represents the Armijo backtracking line search step size, which is computed as

$$\mathcal{W}_2 = \varepsilon\beta \tag{57}$$

with ε initialized to 1 and β being a given parameter such that $\beta \in (0,1)$, as referenced in [37] and [38].

Algorithm 2 RCG for optimal phase shift matrix

1: Initialize $\boldsymbol{\delta}_s$
2: Initialize $l = 0$
3: Initialize $d(0) = -gradC_v(0)$ using 52
4: Repeat
5: Evaluate $\mathcal{W}_2(l)$ using 57
6: Evaluate $\boldsymbol{\delta}_s(l+1)$ using 56
7: Evaluate $R_{\bar{k}c}$ and $R_{\bar{k}p}$
8: Evaluate $gradC_v(l+1)$ and $R_{\bar{k}p}$ using 52
9: Evaluate $\mathcal{W}_1(l+1)$ using 54
10: Evaluate $\mathcal{J}(l+1)$ using 55
11: Evaluate $d(l+1)$ using 53
12: $l \leftarrow l+1$
13: Until convergence is reached

C. Optimization of BS Beamforming

After determining the preliminary active beamforming in Subsection III-B and optimizing the passive beamforming in Subsection V-B, we now compute the remaining active beamforming factors, $f_{v,d}^{c-d}$ and $f_{v,k}^{p-d}$, at the BS, as well as the decoder Ψ_k^d at the UE. The optimization problem $\mathcal{P}^2$ in Eq. (40) can then be reformulated as

$$\mathcal{P}^2 : \max_{f_{v,k}^d, \Psi_k^d} C_v(\boldsymbol{\delta}_s, f_{v,k}^{c-d}, f_{v,k}^{p-d}, \Psi_k^d) \tag{58a}$$

$$\text{s.t. (33e).}$$

We begin by determining the receiver decoder Ψ_k^d. By fixing the precoder $f_{v,k}^d$ to the same value as in the previous problem (46), we can then solve for Ψ_k^d. Consequently, the problem can be reformulated without constraints as

$$\mathcal{P}^2 : \max_{\Psi_k^d} C_v(\boldsymbol{\delta}_s, f_{v,k}^{c-d}, f_{v,k}^{p-d}, \Psi_k^d). \tag{59a}$$

By incorporating the optimized passive beamforming into the effective channel, $\mathbf{H}_{b,s,k}^{eff}$ is expressed as follows (noting that at this stage, we do not consider the error channel):

$$\mathbf{H}_{b,s,k}^{eff} = \Psi_{v,k,a}^H (\mathbf{H}_{s,b,k} \text{diag}(\boldsymbol{\delta}_s^*) \mathbf{H}_{s,u,k}) f_{v,k}^z, \forall k \in \mathcal{K}. \tag{60}$$

For brevity, we denote $\mathbf{H}_{b,s,k}^{eff}$ as $\hat{\mathbf{H}_{b,k}}$. According to [31], the left singular vector of the SVD of $\mathbf{H}s^{eff}$ in Eq. (60) is a matrix containing the first $\mathbf{H}_{s,u,k}$ left singular vectors. Thus,

$$\begin{aligned} &\Psi_{v,k,a}^H (\mathbf{H}_{s,b,k} \text{diag}(\boldsymbol{\delta}_s^*) \mathbf{H}_{s,u,k}) f_{v,k}^z \\ &= \begin{bmatrix} U_b^1 & U_b^0 \end{bmatrix} \begin{bmatrix} \Sigma_b^1 & 0 \\ 0 & \Sigma_b^0 \end{bmatrix} \begin{bmatrix} V_b^1 \\ V_b^0 \end{bmatrix} \end{aligned} \tag{61}$$

where $\Psi_k^d = U_b^1$, considering that in (18), we selected the first column of the first singular vectors for Ψ_k^z.

Next, we substitute the value of Ψ_k^d into problem $\mathcal{P}^2$ and proceed to solve for the remaining precoder $f_{v,k}^d$, which consists of private and common active beamforming. The optimization problem $\mathcal{P}^2$ is now reformulated as

$$\mathcal{P}^2 : \max_{f_{v,k}^d} C_v(\boldsymbol{\delta}_s, f_{v,k}^{c-d}, f_{v,k}^{p-d}, \Psi_k^d) \tag{62a}$$

$$\text{s.t. (33e).}$$

To solve (62), we employ the WMMSE algorithm. By defining $\hat{s}_{c,\bar{k}} = e_{c,\bar{k}} y_{\bar{k}}$ and $\hat{s}_{p,k} = e_{p,k}(y_k - \hat{\mathbf{H}_{b,k}}^H f_{v,d}^c s_c)$ as the estimated common and private streams at UE $\bar{k}$ and k, respectively, where $e_{c,\bar{k}}$ and $e_{p,k}$ are their corresponding equalizers, the mean squared errors (MSEs) for decoding the common and private data streams are formulated as

$$\begin{aligned} \epsilon_{c,\bar{k}} &= E\{|\hat{s}_{c,\bar{k}} - s_{c,\bar{k}}|^2\} \\ &\approx |e_{c,\bar{k}}|^2 T_{c,\bar{k}} - 2\mathcal{R}\{e_{c,\bar{k}} \hat{\mathbf{H}}_{b,\bar{k}} f_{v,\bar{k}}^{c-d}\} + 1, \\ \epsilon_{p,k} &= E\{|\hat{s}_{p,k} - s_{p,k}|^2\} \\ &\approx |e_{p,k}|^2 T_{p,k} - 2\mathcal{R}\{e_{p,k} \hat{\mathbf{H}}_{b,k} f_{v,k}^{p-d}\} + 1 \end{aligned} \tag{63}$$

where

$$\begin{aligned} T_{c,\bar{k}} &= I_v^c + I_{v'} + \sigma\bar{k}, \\ T_{p,k} &= I_v^p + I_{v'} + \sigma k \end{aligned} \tag{64}$$

represent the average received power of the common signal and the private signal (after removing the common signal), respectively.

By setting $\frac{\partial \epsilon_{c,\bar{k}}}{\partial e_{c,\bar{k}}} = 0$ and $\frac{\partial \epsilon_{p,k}}{\partial e_{p,k}} = 0$, the minimum mean squared error (MMSE) equalizers are derived as

$$e_{c,\bar{k}}^* = \frac{\hat{\mathbf{H}}_{b\bar{k}} f_{v,\bar{k}}^{c-d}}{T_{c,\bar{k}}}, \quad e_{p,k}^* = \frac{\hat{\mathbf{H}}_{bk} f_{v,k}^{p-d}}{T_{p,k}}. \tag{65}$$

Substituting Eq.(65) into Eq.(63), the MMSE expressions are obtained as

$$\epsilon_{c,\bar{k}}^* = 1 - \frac{|\hat{\mathbf{H}}_{b\bar{k}} f_{v,\bar{k}}^{c-d}|^2}{T_{c,\bar{k}}}, \quad \epsilon_{p,k}^* = 1 - \frac{|\hat{\mathbf{H}}_{bk} f_{v,k}^{p-d}|^2}{T_{p,k}}. \tag{66}$$

The SINRs of the common and private data streams can then be rewritten as $\gamma_{\bar{k}}^c = \frac{1}{\epsilon_{c,\bar{k}}^*} - 1$ and $\gamma_k^p = \frac{1}{\epsilon_{p,k}^*} - 1$. Accordingly, the achievable rates for the common and private streams are expressed as $R_{\bar{k}}^c = -\log_2 \gamma_{\bar{k}}^c$ and $R_k^p = -\log_2 \gamma_k^p$.

An augmented weighted mean squared error (AWMSE) is introduced to establish a feasible optimization framework, expressed as

$$\Upsilon_{c,\bar{k}} = \mu_c \epsilon_{c,\bar{k}} - \log_2 \mu_c, \quad \Upsilon_{p,k} = \mu_p \epsilon_{c,\bar{k}} - \log_2 \mu_p \tag{67}$$

where μ_c and μ_p are the weights associated with the MSEs for the $\bar{k}$-th and k-th UEs, respectively. By setting $\frac{\partial \Upsilon_{c,\bar{k}}}{\partial \mu_{c,\bar{k}}} = 0$ and $\frac{\partial \Upsilon_{p,k}}{\partial \mu_{p,k}} = 0$, the optimal weights are derived as

$$\mu_{c,\bar{k}}^* = (\epsilon_{c,\bar{k}}^*)^{-1}, \quad \mu_{p,k}^* = (\epsilon_{p,k}^*)^{-1}. \tag{68}$$

Substituting Eqs. (66) and (68) into Eq. (67), the rate-AWMSE relationship is given by

$$\Upsilon_{c,\bar{k}}^* = 1 + \log_2 \epsilon_{c,\bar{k}}^*, \qquad \Upsilon_{p,k}^* = 1 + \log_2 \epsilon_{p,k}^*. \tag{69}$$

Using this relationship, the optimization problem $\mathcal{P}^1$ can be formulated as

$$\begin{aligned} \max_{f_{v\bar{k}}^{c-d}, f_{v,k}^{p-d}, e, \mu} \quad & \Upsilon^w = \sum_k \left(\frac{1}{K} \Upsilon_{c,\bar{k}} + \Upsilon_{p,k}\right) \\ \text{s.t.} \quad & \text{(33e)} \end{aligned} \tag{70}$$

where $\mathbf{e} = [e_{c,1}, \cdots, e_{c,\bar{K}}, e_1, e_2, \cdots, e_K]^H$ and $\boldsymbol{\mu} = [\mu_{c,1}, \cdots, \mu_{c,\bar{K}}, \mu_1, \mu_2, \cdots, \mu_K]^H$ represent the equalizer and weight vectors for both the common and private streams, respectively.

Algorithm 3 WMMSE based RSMA baseband precoding

1: Initialize $f_{v,\bar{k}}^{c-d}$ and $f_{v,k}^{p-d}$
2: **for** $i = 1 : I$ **do**
3: Update e and μ using (65) and (68) with fixed $f_{v,\bar{k}}^{c-d}$ and $f_{v,k}^{p-d}$
4: Update $f_{v,\bar{k}}^{c-d}$ and $f_{v,k}^{p-d}$ by solving (70) with fixed e and μ
5: **end for**

By adopting Lemma 1 from [39], we derive a closed-form solution for the optimization problem (70), as it is subject to only one constraint. The lemma states that the optimal solution $f_{v,k}^*$ of the original problem (70) and the optimal solution $f_{v,k}^{**}$ of its unconstrained transformation satisfy the following relationship:

$$f_{v,k}^* = \eta(f_{v,k}^{**}), \tag{71}$$

where $\eta = \frac{\sqrt{P_t}}{(\sum_k Tr(f_{v,k}^{**}(f_{v,k}^{**})^H))^{\frac{1}{2}}}$ is a scalar factor that ensures the power constraint is satisfied. To transform problem (70) into an unconstrained form, the expressions for (64) are rewritten as follows [39], [40]

$$\begin{aligned} T_{c,\bar{k}} &= I_v^c + I_{v'} + \sigma_{\bar{k}} \frac{\sum_k (f_{v\bar{k}}^{c-d})(f_{v\bar{k}}^{c-d})^H}{P_t}, \\ T_{p,k} &= I_v^p + I_{v'} + \sigma_k \frac{\sum_k (f_{v,k}^{p-d})(f_{v,k}^{p-d})^H}{P_t}. \end{aligned} \tag{72}$$

Then, the closed-form expressions for $f_{v,\bar{k}}^{c,d}$ and $f_{v,k}^{p-d}$ are given by

$$\begin{aligned} f_{v,\bar{k}}^{c-d} &= \frac{1}{K} \left(b_{c,\bar{k}} + \sum_{k=1} \alpha_{c,\bar{k}} \hat{\mathbf{h}}_{s,k} (\hat{\mathbf{h}}_{s,k})^H \right)^{-1} \left(\alpha_{c,\bar{k}} \hat{\mathbf{h}}_{s,k} \right), \\ f_{v,k}^{p-d} &= \left(b_{p,k} + \sum_{k=1} \alpha_{p,k} \hat{\mathbf{h}}_{s,k} (\hat{\mathbf{h}}_{s,k})^H \right)^{-1} (\alpha_{p,k} \hat{\mathbf{h}}_{s,k}) \end{aligned} \tag{73}$$

where $\alpha_{c,\bar{k}} = e_{c,\bar{k}} \mu_{c,\bar{k}}$, $\alpha_k = e_k \mu_k$, and

$$\begin{aligned} b_{c,\bar{k}} &= \frac{\sigma_k}{p_c} \sum_{k=1} \left(\mu_{c,\bar{k}} |e_{c,\bar{k}}|^2 + \mu_k |e_k|^2 \right) \\ b_{p,k} &= b_{c,\bar{k}} + \sigma_k^{CSI} \sum_{k' \in \mathcal{K}_v \setminus k} \mu_{c,\bar{k}} |e_{c,\bar{k}}|^2. \end{aligned} \tag{74}$$

The algorithm for updating $f_{v,\bar{k}}^{c-d}$ and $f_{v,k}^{p-d}$, along with the equalizer vector e and the weight vector μ, is presented in Algorithm 3.

Next, we formulate the Block Coordinate Descent (BCD) method for the joint optimization of both active and passive beamforming at the subarray level. A detailed description can be found in Section 4.

D. Computational Complexity Analysis

In this subsection, we analyze the computational complexity of the proposed solution, highlighting the impact of partitioning on complexity reduction. Table III summarizes the complexity of each component of the solution, both with and without partitioning. Additionally, we provide computation

Algorithm 4 BCD-based joint precoding and beamforming phase shift matrix optimization

1: Input: Antenna and RIS elements partitioning vectors $\mathcal{A}, \mathcal{R}$ from Algorithm 1, Maximum power P_t, optimal beamforming $G, \mathbf{H_{b,s}}, \mathbf{H_{s,u}}$, set ϱ to a small value
2: Initialize the convergence checker ϱ to a small value.
3: Initialize the precoder f_v by zero forcing precoder and phase shift matrix as $\boldsymbol{\delta}_s$ to fixed phase by the BS, set the iteration interval $t = 0$
4: **for** All vehicles **do**
5: Repeat
6: Compute the subcarrier optimal beamforming using Eq. (38)
7: Run Algorithm 2 to compute the optimal phase shift matrix $\boldsymbol{\delta}_s$
8: Run Algorithm 3 to compute the optimal precoder $f_{\bar{k}}^c f_k^p$
9: Realize $\bar{k} = \arg\min_{k \in \mathcal{K}} r_k^c$ and evaluate the problem C_v
10: Until $t > 2$ and $C_v(t) - C_v(t-1) \leq \varrho$
11: **end for**

time measurements in the simulation section, as shown in Tables V and VI.

For Algorithm 1, which handles the partitioning of BS antennas and RIS elements by allocating a total of M BS antennas and N RIS elements to V vehicles, the computational complexity is given by:

$$\mathcal{O}(V(M+N)). \tag{75}$$

For Algorithm 2, which computes the RIS reflection coefficients, the computational complexity is primarily dominated by the gradient computation and retraction steps (Steps 6 and 8 in Algorithm 2). It can be expressed as

$$\mathcal{O}(I_{rcg}(VK_vN_s^2)) + \mathcal{O}(I_{rcg}(VK_v^2N_s^2)), \tag{76}$$

where I_{rcg} represents the total number of iterations required for convergence.

For Algorithm 3, which employs the WMMSE algorithm to compute private and common stream precoding, the computational complexity is given by

$$\mathcal{O}(I_{wmmse}VK_v^2), \tag{77}$$

where I_{wmmse} is the total number of iterations required for convergence.

In the non-partitioned scheme, there is no complexity associated with the partitioning algorithm. However, the complexity increases due to beamforming processing at the BS and RIS, which must handle all elements for each vehicle. The computational complexity associated with RCG in this case is

$$\mathcal{O}(I_{rcg}(VK_vN^2)) + \mathcal{O}(I_{rcg}(VK_v^2N^2)), \tag{78}$$

while the complexity of WMMSE remains

$$\mathcal{O}(I_{wmmse}VK_v^2), \tag{79}$$

but with an increased overhead due to the larger channel matrix size at each UE.

TABLE III
COMPUTATION COMPLEXITY COMPARISON

Scheme	RCG alg	WMMSE Alg	Partitioning Alg
Main scheme	$\mathcal{O}(I_{rcg}(VK_vN_s^2)) + \mathcal{O}(I_{rcg}(VK_v^2N_s^2))$	$\mathcal{O}(I_{dig}VK_V^2)$	$\mathcal{O}(V(M+N))$
No partitioning	$\mathcal{O}(I_{rcg}(VK_vN^2)) + \mathcal{O}(I_{rcg}(VK_v^2N^2))$	$\mathcal{O}(I_{dig}VK_V^2)$	N/A

TABLE IV
SIMULATION PARAMETERS (DEFAULT VALUES ARE UNDERLINED)

Parameter	Value (s)
Road width	750/2 m
Road length	1299/2 m
Lane width	3.5 m
BS position on the urban grid	(750, 1299/2) m
RIS position on the urban grid	(750/2, 1299) m
Number of vehicles ($\lvert\mathcal{V}\rvert$)	[3, **4**, 5, 6]
Number of vehicle random drops	10
Number of UEs per vehicle (K)	[1, 2, **3**, 4]
Number of antennas (M)	[16, 32, **64**, 80]
Number of RIS elements (N)	[80, 100, **128**, 148]
BS transmit power budget (P_t)	[25, 30, **35**, 40]
UE noise power (σ_k)	-96 dBm
Maximum speed of vehicles (v_{max})	[0, **10**, 20, 30] m/s
Carrier frequency (f)	30 GHz
Channel bandwidth (B)	60 MHz
Antennas and RIS elements separation distance (d)	$\frac{\lambda}{2}$
Azimuth angle (ϑ)	90
Elevation angle (φ)	60
Communication time (T_c)	1 ms

VI. PERFORMANCE EVALUATION

In this section, we present simulation results and evaluate the performance of our proposed approach in comparison with various benchmark schemes.

A. Simulation Setup

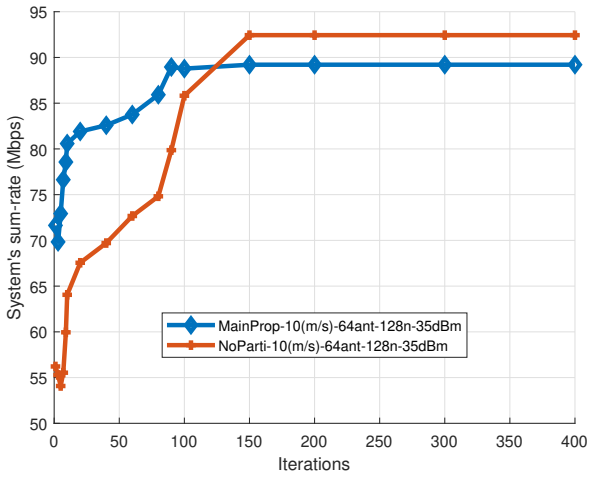
We adopt the simulation setup for an urban road grid scenario as specified in 3GPP TR 36.885 [41]. Vehicles are randomly placed on the grid and move in one of four possible road directions based on predefined probabilities. The specifications for road grid dimensions, drop model, speed range, and movement directions are detailed in Table IV. The positions of the BS and RIS are also provided in Table IV, where they remain fixed during each period T_c and are updated at the start of each new period. The table also underscores the values of the key variables and parameters used in the simulation.

For assigning the movement speeds of individual vehicles, we set their maximum speed (v_{max}) and randomly select a speed value between $\max 0, v_{max} - 9$ and v_{max}.

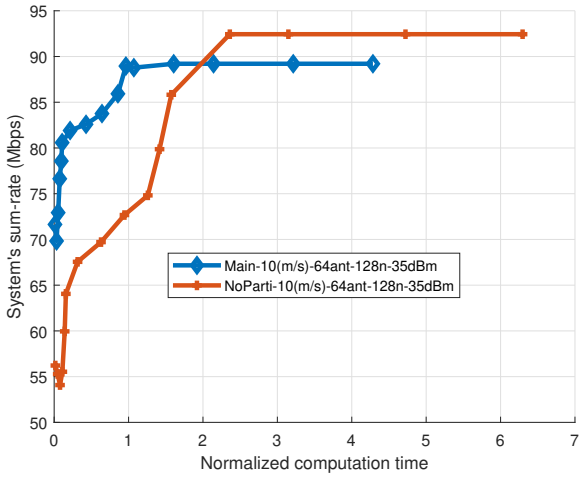
B. Benchmarks

In this subsection, we define the benchmark schemes used for comparison. Our proposed scheme is referred to as Prop-Main.

- *No partitioning (NoParti)*: In this scheme, the large-scale antenna array at the BS and the RIS elements remain unpartitioned, meaning all antennas and RIS elements



(a) Convergence iterations



(b) Convergence over normalized time complexity

Fig. 4. Convergence analysis over BCD iterations

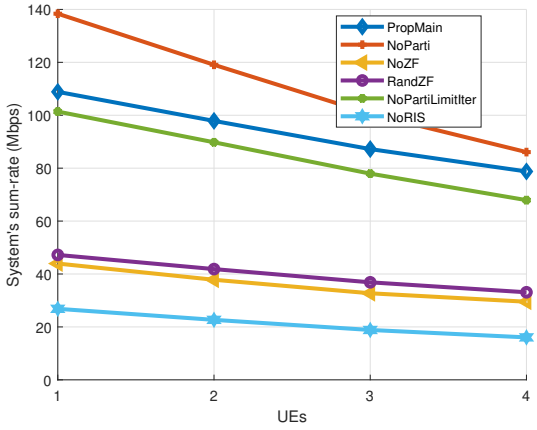


Fig. 5. Total system's capacity over UEs

jointly serve all vehicles without subarray-subsurface mapping.

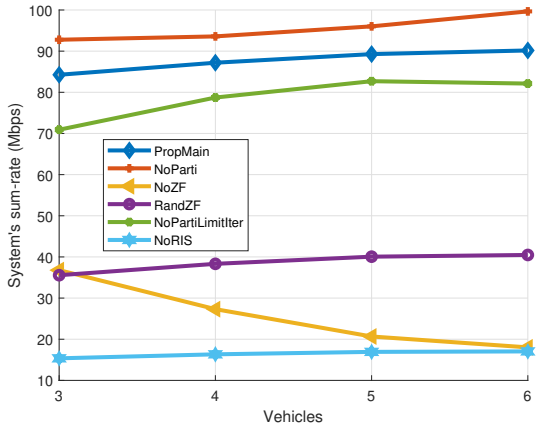


Fig. 6. Total system's capacity over deployed vehicles

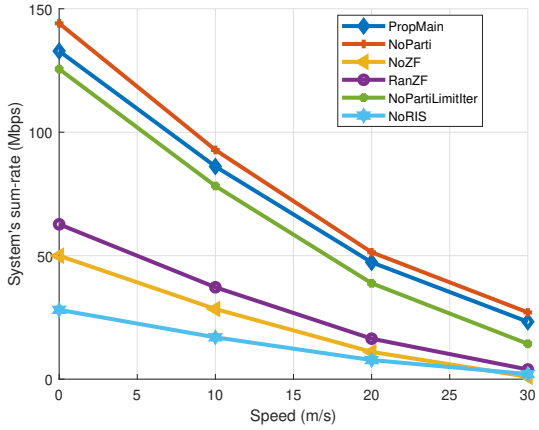


Fig. 7. Total system's capacity over speed

- *NoPartiLimitIter*: This benchmark is similar to NoParti, but the simulation is conducted over a limited number of iterations to analyze the impact of computational constraints.
- *No Zero Forcing scheme (NoZF)*: This scheme evaluates the performance of the main scheme when CZF is not applied, allowing inter-vehicular interference to persist.
- *Random Reflection Coefficient with CZF (RanZF)*: This scheme assesses the system performance when the RIS reflection coefficients are randomly generated instead of being optimized, while still applying CZF.
- *Non-RIS (NoRIS)*: This scheme evaluates the performance of PropMain without the RIS, assuming a direct communication link between the BS and vehicles, bypassing RIS-assisted signal enhancement.
- *Random UE grouping (RandGroupedUEs)*: In this scheme, UEs are randomly grouped across vehicles, whereas PropMain groups UEs within the same vehicle. The number of UEs per group remains consistent with PropMain.
- *NOMA*: This scheme evaluates the performance of the proposed PropMain, but replaces RSMA with NOMA for multiple access.

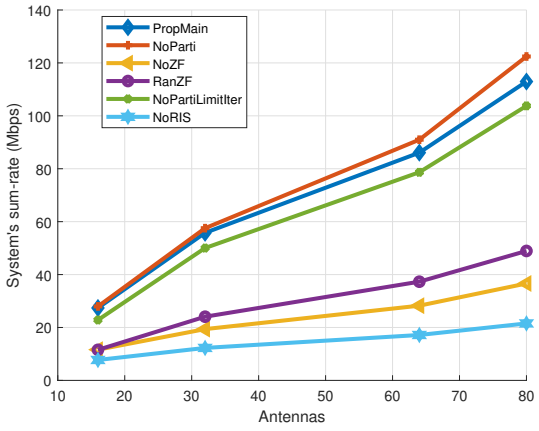


Fig. 8. Total system's capacity over antennas budget at the BS

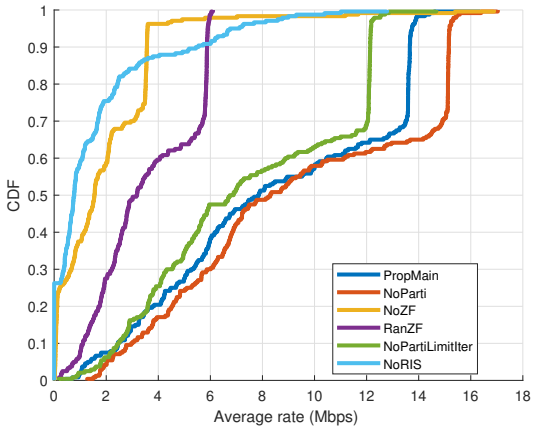


Fig. 11. CDF over speed of vehicles

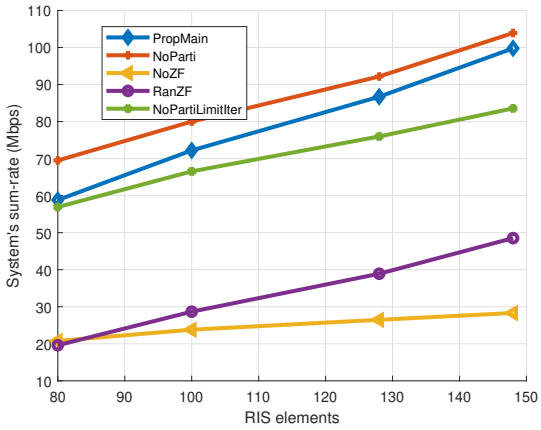


Fig. 9. Total system's capacity over RIS elements

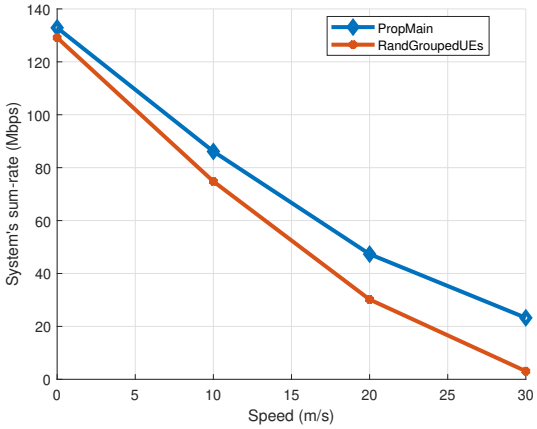


Fig. 12. Comparing total system's capacity over speed by UEs grouping method

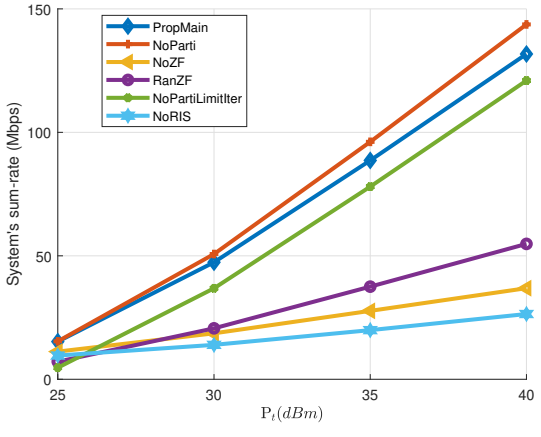


Fig. 10. Total system's capacity over transmit powers, P_t

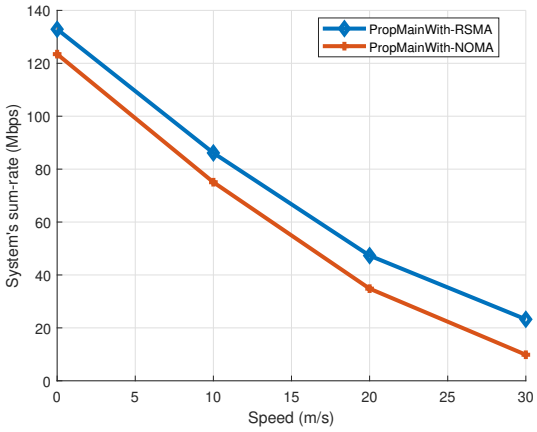


Fig. 13. Comparison of the proposed scheme under RSMA and with the proposed scheme under NOMA

C. Performance Results

1) Convergency and complexity: The convergence behavior of our proposed partitioning-based scheme (PropMain) and the NoParti scheme, both in terms of iterations and computation time, is illustrated in Fig. 4, with numerical values provided in Tables V and VI in the Appendix. As shown in Fig. 4(a), PropMain converges in approximately 100 iterations, whereas NoParti requires 150 iterations, demonstrating a faster convergence rate. When comparing the total computation time required for convergence, the gap between the two schemes

becomes even more significant. Fig. 4(b) shows that PropMain requires substantially less computation time per iteration, ultimately leading to a 50% reduction in total computation time compared to NoParti. This indicates that when computational resources are constrained and NoParti is unable to reach full convergence within the available budget, PropMain may achieve a higher sum rate. As defined earlier, NoPartiLimitIter represents a scenario where the computational budget is restricted to match the convergence time of PropMain. As shown in Fig. 4(b), while PropMain achieves 89 Mbps after convergence, NoPartiLimitIter achieves only 73 Mbps, which is an 18% lower performance than PropMain.

From the tables in the Appendix, it is evident that partitioning significantly reduces computational time but comes at a trade-off in sum rate performance. However, the discrepancy between the sum rate gap and the reduction in total computational time highlights the overall benefits of our dynamic partitioning approach. In Table V, we observe that the BCD algorithm's computation time is primarily dominated by the partitioning, WMMSE, and RCG algorithms, while the computational costs of other components in our solution are relatively minor.

2) Impact of the number of UEs per vehicle: The impact of the number of UEs per vehicle on the system sum rate is illustrated in Fig. 5. As the number of UEs increases, the sum rate decreases due to inter-user interference, which CZF alone cannot fully mitigate.

Among all schemes, NoParti achieves the highest sum rate across different numbers of UEs. However, when its iterations are limited (NoPartiLimitIter), its performance drops below that of the proposed scheme. This demonstrates that under a similar computational budget, the proposed scheme achieves the highest sum rate among all benchmarks.

The other three schemes show significantly lower sum rates, highlighting the importance of optimal resource allocation. Specifically, NoRIS emphasizes the crucial role of RIS in improving system performance, while RandZF illustrates the necessity of optimizing RIS configurations, as its performance remains suboptimal despite the presence of RIS. Lastly, NoZF confirms that suppressing inter-vehicle interference is an essential component of the proposed architecture for maintaining high sum-rate performance.

3) Impact of the number of vehicles: Fig. 6 illustrates the system sum rate as the number of vehicles in the network varies. Due to the CZF technique, the sum rate slightly increases with the number of vehicles, as interference suppression allows for more efficient resource allocation.

The proposed scheme achieves the highest sum rate among all benchmarks except NoParti. However, PropMain consistently outperforms NoPartiLimitIter, regardless of the number of vehicles, demonstrating the efficiency of partitioning in optimizing performance under a constrained computational budget.

The other three schemes show significantly lower sum rates, underscoring the importance of integrating RIS deployment, optimal RIS configuration, and CZF within our solution. In particular, NoZF shows a declining sum rate as the number of vehicles increases, due to the growing impact of inter-vehicle interference. This further confirms that CZF-based interference suppression is crucial for maximizing the system sum rate in the proposed architecture.

4) Impact of vehicular speed: The impact of vehicle speed (and consequently, UE mobility) on the system sum rate is illustrated in Fig. 7. The sum rates of all schemes are highly influenced by vehicle speed, with performance decreasing by nearly 50% when vehicles move at 10 m/s, compared to a stationary scenario. This decline is primarily due to channel degradation and increased estimation errors caused by the Doppler frequency shift resulting from vehicular motion [42]. Despite these challenges, our proposed solution effectively mitigates the impact of increasing vehicular speed.

The relative performance gap between NoParti, PropMain, and NoPartiLimitIter remains nearly constant across different speeds. Similarly, the other three schemes show significantly lower sum rates, further highlighting the importance of partitioning, RIS configuration, and CZF in maintaining system performance.

In Fig. 12, we compare the sum rate of our main scheme, which groups UEs per vehicle, with a scheme where UEs are randomly grouped across all vehicles. For instance, a UE group in the random scheme could consist of the 1st UE from Vehicle 1, 4th UE from Vehicle 2, 3rd UE from Vehicle 3, and 1st UE from Vehicle 4. The results show that when UEs are randomly grouped across vehicles, performance deteriorates more significantly as speed increases, compared to our main scheme, which groups UEs within the same vehicle. This highlights the benefits of per-vehicle UE grouping in mitigating the adverse effects of mobility on system performance.

5) Impact of the number of antennas and RIS elements: The impact of the number of BS antenna elements and RIS elements on the system sum rate is illustrated in Figs. 8 and 9, respectively. The sum rate steadily increases as the number of antennas grows across all schemes.

Among the evaluated schemes, NoParti achieves the highest sum rate. However, its advantage over PropMain is not substantial, particularly when the total number of antenna elements is below 64, where both schemes show nearly identical performance.

A notable performance gap is observed between the first three schemes (NoParti, PropMain, and NoPartiLimitIter) and the last three schemes (RandZF, NoZF, and NoRIS). Furthermore, the first three schemes demonstrate a better ability to leverage an increasing number of BS antennas for sum rate enhancement compared to the latter three schemes. A similar trend is observed when increasing the number of RIS elements, reinforcing the importance of optimal resource allocation and interference suppression in achieving higher system performance.

6) Impact of BS transmit power: Fig. 10 illustrates the impact of the BS transmit power P_t on the system sum rate. As P_t increases, all schemes show a corresponding increase in sum rate. NoParti outperforms PropMain; however, at P_t = 25 dBm, both schemes achieve nearly identical sum rates.

For higher P_t, the performance gap between PropMain and the other three schemes (RandZF, NoZF, and NoRIS) widens. However, at P_t = 25 dBm, the relative gain of PropMain over

these schemes diminishes significantly, and NoPartiLimitIter even falls below them in performance.

These trends indicate that, at higher transmit power levels, the benefit of RIS-assisted signal enhancement becomes less pronounced, as the schemes rely less on the path gain provided by RIS.

7) CDF: Fig. 11 compares the cumulative distribution function (CDF) of user rates for the proposed scheme and the benchmark schemes. PropMain consistently outperforms NoPartiLimitIter across most samples, demonstrating its efficiency in optimizing user rates. The CDF curve of PropMain closely follows that of NoParti up to 60% of the samples. However, for the remaining 40% of samples with higher user rates, the gap between the two schemes widens. This indicates that the performance gain of NoParti primarily comes from higher user rates for UEs experiencing good channel conditions, rather than an overall system-wide improvement. NoZF and NoRIS show significantly lower user rates, with 20% of samples having zero rate, highlighting the importance of interference suppression and RIS deployment in maintaining user performance. RandZF, while benefiting from RIS assistance, achieves much higher and more stable user rates than NoZF and NoRIS. However, it still falls significantly short of PropMain, emphasizing the critical role of optimal RIS configuration in maximizing user rates.

8) Impact of UE grouping: Fig. 12 shows the impact of UE grouping on the system sum rate for varying vehicular speed by comparing two different schemes: PropMain and RandGroupedUEs. As the speed increases, both schemes experience a significant decline in sum rate, which is expected due to the Doppler effect, channel variations, and increased estimation errors at higher speeds. However, PropMain consistently outperforms RandGroupedUEs across all speed levels, demonstrating the advantage of per-vehicle UE grouping in maintaining higher sum rates. At low speeds (0–10 m/s), the gap between the two schemes is relatively small, suggesting that even random grouping can achieve reasonable performance under stable channel conditions. However, as the speed increases beyond 10 m/s, the performance gap widens, with RandGroupedUEs showing a much steeper decline in sum rate. This indicates that random UE grouping exacerbates the effects of mobility-induced channel degradation, whereas PropMain's structured UE grouping better preserves communication efficiency in high-mobility scenarios. At 30 m/s, the sum rate for RandGroupedUEs is nearly zero, whereas PropMain still maintains a non-negligible sum rate, further highlighting the robustness of the proposed method in handling high-mobility environments.

9) Impact of multiple access scheme: Fig. 13 compares the sum rate of PropMain with RSMA and PropMain with NOMA across different vehicular speeds. As expected, the sum rate decreases for both schemes as speed increases, due to channel weakening and estimation errors caused by Doppler shifts. However, PropMain with RSMA consistently outperforms PropMain with NOMA across all speeds, demonstrating the superior interference management and spectral efficiency of RSMA. The performance gap between the two schemes remains relatively stable across speed levels, indicating that RSMA is more resilient to mobility-induced impairments. At 30 m/s, the sum rate of PropMain with RSMA is significantly higher than PropMain with NOMA, highlighting RSMA's ability to better handle high-mobility environments. This suggests that RSMA is a more suitable multiple access scheme for high-speed vehicular communication systems compared to NOMA.

VII. Conclusion

In this paper, we investigated the trade-off between sum rate maximization and computational complexity in a mmWave V2I communication system using RSMA with BS antenna and RIS element partitioning. We developed a dynamic partitioning algorithm that efficiently creates subarrays and subsurfaces, enabling the transmission of private RSMA streams to their corresponding UEs within a vehicle. This approach significantly reduces computational time while maintaining a high system sum rate. To mitigate inter-vehicular interference, we incorporated CZF. Additionally, we designed two key optimization algorithms: RCG for passive beamforming at RIS subsurfaces and WMMSE for precoding at BS subarrays. Through extensive simulations, we compared our proposed scheme against multiple benchmarks. The results demonstrate that our scheme achieves a competitive sum rate while significantly reducing computational cost.

For future work, we plan to enhance our approach by integrating machine learning models to optimize beamforming at each subarray, further improving the partitioning strategy to match or even exceed the sum rate performance of the no-partitioning scheme.

Appendix

TABLE V
Computational Complexity Results of Main scheme along with Sum Rate (Time is sec and Sum rate in Mbps)

Iterations for Main	BCD Time	BCD Sum Rate	RCG Time	WMMSE Time	Partitioning AlgTime
Iterations 1	28.92	71.64	23.436	2.8564	1.3756
Iterations 5	144.6	72.928	117.18	14.282	6.8779
Iterations 10	289.2	80.585	234.36	28.564	13.756
Iterations 60	1735.2	83.76	1406.2	171.38	82.535
Iterations 100	2892	88.78	2343.6	285.64	137.56
Iterations 150 (Compare to NoParti)	4338.1	89.207	3515.5	428.46	206.34
Iterations 200	5784.1	89.207	4687.3	571.28	275.12
Iterations 300	8676.1	89.207	7030.9	856.92	412.67

References

[1] K. N. Qureshi, S. Din, G. Jeon, and F. Piccialli, "Internet of vehicles: Key technologies, network model, solutions and challenges with future aspects," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1777–1786, Mar, 2021.

[2] F. S. Woon and C. Y. Leow, "Intelligent reflecting surfaces aided millimetre wave blockage prediction for vehicular communication," in *Proc. IEEE International Symposium on Telecommunication Technologies (ISTT)*, 2022, pp. 11–15.

[3] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, "Multi-beam NOMA for hybrid mmwave systems," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1705–1719, Feb 2019.

TABLE VI
COMPUTATIONAL COMPLEXITY RESULTS OF NOPARTI SCHEME ALONG WITH SUM RATE (TIME IS IN SEC AND SUM RATE IN MBPS)

Iterations for NoParti	BCD Time	BCD Sum Rate	RCG Time	WMMSE Time	Partitioning AlgTime
Iterations 1	42.498	56.216	34.346	3.1636	NaN
Iterations 5	212.49	54.08	171.73	15.818	NaN
Iterations 10	424.98	64.05	343.46	31.636	NaN
Iterations 60	2549.9	72.668	2060.7	189.82	NaN
Iterations 100 (Run Limited)	4249.8	85.817	3434.6	316.36	NaN
Iterations 150 (Compare to Main)	6374.7	92.437	5151.8	474.54	NaN
Iterations 200	8499.6	92.437	6869.1	632.72	NaN
Iterations 300	12749	92.437	10304	949.09	NaN

[4] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, Apr 2016.

[5] P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14 960–14 973, Dec 2020.

[6] Y. Cui, G. Wang, D. Wu, P. He, R. Wang, and Y. Liu, "RIS-assisted unsupervised beamforming in internet of vehicles," *IEEE Transactions on Vehicular Technology*, pp. 1–14, Sept 2024.

[7] Y. Chen, Y. Wang, J. Zhang, and Z. Li, "Resource allocation for intelligent reflecting surface aided vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12 321–12 326, Oct 2020.

[8] X. Gu, G. Zhang, W. Duan, J. Choi, M. Wen, and P.-H. Ho, "RIS-V2V communications: The way to mitigate doppler shifts and multipath spread," *IEEE Wireless Communications Letters*, vol. 13, no. 3, pp. 746–750, 2024.

[9] A. Khaleel and E. Basar, "A novel NOMA solution with RIS partitioning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 70–81, Jan 2022.

[10] S. Zhang and R. Zhang, "Intelligent reflecting surface aided multi-user communication: Capacity region and deployment strategy," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5790–5806, 2021.

[11] X. Wang, R. Zheng, F. Du, X. Zhao, Y. Zhang, Y. Xu, S. Geng, and P. Qin, "Joint beamforming and reflecting elements optimization for segmented RIS assisted multi-user wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 3, pp. 3820–3831, March 2024.

[12] B. Clerckx, Y. Mao, E. A. Jorswieck, J. Yuan, D. J. Love, E. Erkip, and D. Niyato, "A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1265–1308, 2023.

[13] B. C. Y. Mao and V. O. K. Li, "Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming SDMA and NOMA," *Journal of Wireless Communication Network,*, may 2018.

[14] C. Cai, X. Yuan, and Y.-J. A. Zhang, "RIS partitioning based scalable beamforming design for large-scale MIMO: Asymptotic analysis and optimization," *IEEE Transactions on Wireless Communications*, vol. 22, no. 9, pp. 6061–6077, January 2023.

[15] S. Pala, M. Katwe, K. Singh, B. Clerckx, and C.-P. Li, "Spectral-efficient RIS-aided RSMA URLLC: Toward mobile broadband reliable low latency communication (mBRLLC) system," *IEEE Transactions on Wireless Communications*, vol. 23, no. 4, pp. 3507–3524, Apr 2024.

[16] W. Duan, X. Gu, G. Zhang, M. Wen, Z. Ding, and P.-H. Ho, "Sum-rate maximization for RIS-IoV: From instantaneous to statistical CSI," *IEEE Transactions on Wireless Communications*, vol. 23, no. 7, pp. 8071–8084, Jul 2024.

[17] M. Wu, Z. Gao, Y. Huang, Z. Xiao, D. W. K. Ng, and Z. Zhang, "Deep learning-based rate-splitting multiple access for reconfigurable intelligent surface-aided tera-Hertz massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1431–1451, 2023.

[18] A. Khaleel and E. Basar, "A novel NOMA solution with RIS partitioning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 1, pp. 70–81, Jan 2022.

[19] Y. Kim, J.-H. Kim, and C.-B. Chae, "Partition-based RIS-assisted multiple access: NOMA decoding order perspective," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 9083–9088, Aug 2022.

[20] H. Jiang, B. Xiong, H. Zhang, and E. Basar, "Hybrid far- and near-field modeling for reconfigurable intelligent surface assisted V2V channels: A sub-array partition based approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 8290–8303, Nov 2023.

[21] S. Pala, M. Katwe, K. Singh, T. A. Tsiftsis, and C.-P. Li, "Robust transmission design for RIS-aided full-duplex-RSMA V2X communications via multi-agent DRL," *IEEE Transactions on Vehicular Technology*, pp. 1–15, Sept 2024.

[22] M. A. Houran, M. Asad, G. Srivastava, J. Mirza, A. Ranjha, M. A. Javed, and X. Yang, "Intelligent reflecting surfaces assisted cellular V2X based open RAN communications," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 7, pp. 9226–9233, Jul 2024.

[23] P. Saikia, S. Pala, K. Singh, S. K. Singh, and W.-J. Huang, "Proximal policy optimization for RIS-assisted full duplex 6G-V2X communications," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 7, pp. 5134–5149, May 2024.

[24] T. Zhang, P. Ren, D. Xu, and Z. Ren, "RIS subarray optimization with reinforcement learning for green symbiotic communications in internet of things," *IEEE Internet of Things Journal*, vol. 10, no. 22, pp. 19 454–19 465, Nov 2023.

[25] H. Li, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for 6g—part III: Interplay with reconfigurable intelligent surfaces," *IEEE Communications Letters*, vol. 26, no. 10, pp. 2242–2246, Oct 2022.

[26] A. S. de Sena, P. H. J. Nardelli, D. B. da Costa, P. Popovski, and C. B. Papadias, "Rate-splitting multiple access and its interplay with intelligent reflecting surfaces," *IEEE Communications Magazine*, vol. 60, no. 7, pp. 52–57, 2022.

[27] S. Wahb, A. El-Mahdy, and F. Dressler, "Selective multi-user cooperative rate splitting assisted by reconfigurable intelligent surface," in *Proc. International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, May 2024, pp. 1–10.

[28] S. H. Chae, H. W. Kim, H. J. Park, and S.-W. Jeon, "Rate splitting-based hybrid beamforming for multi-user downlink cellular networks," *IEEE Transactions on Communications*, vol. 72, no. 6, pp. 3338–3351, Jun 2024.

[29] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter wave communication: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, 2018.

[30] Y. Lee, J.-H. Lee, and Y.-C. Ko, "Beamforming optimization for IRS-assisted mmwave V2I communication systems via reinforcement learning," *IEEE Access*, vol. 10, pp. 60 521–60 533, 2022.

[31] B. Ning, P. Wang, L. Li, Z. Chen, and J. Fang, "Multi-IRS-aided multi-user MIMO in mmWave/THz communications: A space-orthogonal scheme," *IEEE Transactions on Communications*, vol. 70, no. 12, pp. 8138–8152, 2022.

[32] Q. Li, B. Bu, and J. Zhao, "A novel hierarchical situation awareness model for CBTC using SVD entropy and GRU with PRD algorithms," *IEEE Access*, vol. 9, pp. 132 290–132 300, Sept, 2021.

[33] O. Dizdar, Y. Mao, and B. Clerckx, "Rate-splitting multiple access to mitigate the curse of mobility in (massive) MIMO networks," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6765–6780, October 2021.

[34] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[35] Y. Xiu, Y. Zhao, Y. Liu, J. Zhao, O. Yagan, and N. Wei, "IRS-assisted millimeter wave communications: Joint power allocation and beamforming design," in *Proc. IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2021, pp. 1–6.

[36] S. K. Singh, K. Agrawal, K. Singh, B. Clerckx, and C.-P. Li, "RSMA for hybrid RIS-UAV-aided full-duplex communications with finite block-length codes under imperfect SIC," *IEEE Transactions on Wireless Communications*, vol. 22, no. 9, pp. 5957–5975, 2023.

[37] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, 2016.

[38] A. R. M. Zargham and A. Jadbabaie, "A distributed line search for network optimization," in *Proc. 2012 American Control Conference (ACC), Montreal, QC, Canada, 2012*, Oct 2012, pp. 472–477.

[39] H. Jiang, M. Cui, D. W. K. Ng, and L. Dai, "Accurate channel prediction based on transformer: Making mobility negligible," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2717–2732, 2022.

[40] Q. Hu, Y. Cai, Q. Shi, K. Xu, G. Yu, and Z. Ding, "Iterative algorithm induced deep-unfolding neural networks: Precoding design for multiuser MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1394–1410, 2021.

[41] T. S. G. R. A. Network;, "Study LTE-based V2X services; (release 14),document 3GPP TR 36.885," *3rd Generation Partnership Project*, vol. 14.0.0, Jun 2016.

[42] Y. Wang, G. Wang, R. He, B. Ai, and C. Tellambura, "Doppler shift and channel estimation for intelligent transparent surface assisted communication systems on high-speed railways," *IEEE Transactions on Communications*, vol. 71, no. 7, pp. 4204–4215, 2023.