IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 8 February 2024, accepted 3 March 2024, date of publication 11 March 2024, date of current version 15 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3375769

RESEARCH ARTICLE

Power Efficient and Ultra Dense Open-RAN Vehicular Networks With Non-Linear Processing

GEORGE N. KATSAROS AND KONSTANTINOS NIKITOPOULOS, (Senior Member, IEEE)
Wireless Systems Laboratory, 5G and 6G Innovation Centre, Institute for Communication Systems, University of Surrey, GU2 7XH Guildford, U.K.

Corresponding author: Konstantinos Nikitopoulos (k.nikitopoulos@surrey.ac.uk)

This work was supported in part by the "Flex-5G" Project, a Winner of the U.K.'s Department for Science, Innovation, and Technology (DSIT), Future RAN Competition; and in part by the HiPer-RAN Project, a Winner of U.K.'s DSIT Open Networks Ecosystem Competition.

ABSTRACT Transitioning to more intelligent, autonomous transportation systems necessitates network infrastructure capable of accommodating both substantial uplink traffic and massive vehicle connectivity. Current approaches addressing these throughput and connectivity requirements rely on the utilization of the multiple input, multiple output (MIMO) technology. However, When traditional linear detection/precoding processing methods are adopted, they require the deployment of an extensive number of co-located, access-point antennas to support a comparatively much smaller number of data streams. Such a setup significantly increases the power consumption on the radio side, raising substantial concerns about the operational costs and sustainability of such deployments, particularly in densely deployed scenarios, across extensive road networks. Addressing these concerns, this work proposes an Open Radio Access Network (Open-RAN) deployment that incorporates a Massively Parallelizable, Nonlinear (MPNL) MIMO processing framework and assesses, for the first time, its impact on the power consumption and vehicular connectivity in various Vehicle-to-Infrastructure (V2I) and Network (V2N) scenarios. We show that flexible, Open-RAN physical layer deployments, incorporating MPNL, emerge as a critical power efficiency enabler, especially when flexibly activating/deactivating employed RF elements. Our field-programmable gate array (FPGA) based evaluation of MPNL, reveals that it can lead to significant power savings on the radio side, by eliminating the need for a "massive" number of base station antennas and radio frequency (RF) chains. Additionally, our findings show substantial connectivity gains, exceeding 400%, in terms of concurrently transmitting vehicles compared to traditional processing approaches, without significantly affecting the access point power consumption budgets, thereby catalyzing the evolution towards more intelligent, fully autonomous, and sustainable transportation systems.

INDEX TERMS Power efficiency, C-V2X, open-RAN, massive MIMO, non-linear processing.

I. INTRODUCTION

The standardization of 5G New Radio (5G-NR) based Cellular Vehicle-to-Everything (C-V2X) [1] technology marks a significant milestone in the transformation of our transportation systems. It promises to improve road safety, traffic efficiency, and user convenience by enabling seamless communication between vehicles, infrastructure, pedestrians, and network services, laying the foundation for autonomous and cooperative driving.

However, the transition to even more intelligent transportation systems requires that vehicles, infrastructure, and pedestrians continuously transmit substantial volumes of real-time sensor data with high reliability and low latency. For instance, it is projected that a single autonomous vehicle will generate an astounding 1.4 to 40TB of data per hour [2], [3]. When comparing this with the average smartphone data usage in 2021 [4], the projected data output for autonomous vehicles

The associate editor coordinating the review of this manuscript and approving it for publication was Luyu Zhao.

exceeds the average user consumption by over 7000 times. Adding to this challenge is the simultaneous requirement for massive device connectivity. As an example, consider that a busy intersection (e.g., Times Square) could host more than 400 concurrently transmitting vehicles and more than four times that number for pedestrians and passengers [5]. This underscores the need for highly spectral efficient solutions capable of concurrently transmitting and receiving large data volumes.

In the latest mobile generations, advancements in Multiple-Input Multiple-Output (MIMO) and Massive MIMO (mMIMO) technologies have been the primary drivers behind device connectivity and throughput gains, particularly in the sub-6 GHz frequency range (FR1), where the multipath propagation environment enables substantial spatial multiplexing gains [6]. However, the high computational complexity associated with MIMO detection/precoding algorithms has led the community to resort to less computationally demanding but suboptimal linear processing approaches based on the Zero Forcing (ZF) and Minimum Mean Square Error (MMSE) principles. Despite the practical computational complexity benefits, linear MIMO processing typically comes with the requirement for a disproportionately large number of antennas and RF chains employed on the base station than the number of supported MIMO streams. This leads to a substantial increase in the power consumption budgets required on the radio unit side [7], [8], [9]. The power consumption of the radio units becomes particularly pressing within the context of vehicular networks. On top of the high operational costs and concerns about the carbon footprint of radio infrastructures, a large portion of existing road networks even lay outside the reach of the electrical power grid [10]. This creates the requirement for a novel, ultra-power efficient solutions capable of supporting even potential solar-based roadside unit deployments [11].

Open RAN [12], [13], [14] emerges as a potential core enabler of novel processing approaches in the radio access network (RAN). With standardized open interfaces and multiple disaggregation options, Open-RAN promises, in addition to a highly diversified RAN ecosystem, high flexibility with programmatic control, essentially creating the conditions for a boost in innovation [15]. Open RAN can allow for a simplification of the radio unit, moving computationally intensive parts of the physical layer (PHY) processing chain (or even the entire PHY) away from the radio side to more centralized locations through the deployment of different functional splits [16]. Although functional splits do not directly address the high power consumption from the use of multiple power-hungry RF-Chains the imposed flexibility allows for the implementation of more advanced, computationally demanding MIMO processing techniques, such as nonlinear MIMO processing, which, as we show in this work, have the potential to substantially improve system-level energy efficiency and vehicular connectivity capabilities. However, traditional, existing nonlinear processing approaches (e.g., based on the Sphere Decoder [17], [18], [19]) are highly computationally demanding, with exponentially increasing complexity with the number of MIMO streams [20], [21], to the point of impracticality for real-time MIMO deployments. Recent proposals, such as massively parallelizable nonlinear (MPNL) processing [19], [22], [22], [23], [24], attempt to address this issue by aggressively parallelizing the corresponding functionalities. This allows to efficently leverage any modern parallel processing architectures, significantly reducing the detection and precoding processing latency, making them compatible with the real-time processing requirements of 5G-NR. Still, despite the potential of MPNL processing, a practical realization of MPNL has not been discussed in the context of Open RAN. Consequently, its actual system-level gains in terms of connectivity and power consumption are yet to be fully understood, as well as the impact of such deployment and the new trade-offs it unlocks within a practical and highly dynamic V2X environment.

It is worth noting that while the algorithmic principles of MPNL are applicable across any wireless technology that supports MIMO, in this work, we focus on cellular-based Open-RAN deployments. Compared to WiFi approaches (typically based on IEEE 802.11p and 802.11bd) that mainly emphasize vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication, cellular radio access technologies can explore a broader scope of vehicular applications, including high throughput vehicle-to-network (V2N) applications. In addition, C-V2X has shown higher resilience to interference, enhanced Non-Line of Sight (NLOS) performance compared to 802.11p, as well as better future-proofing aspects such as the ability to be directly integrated with existing cellular infrastructures [25], [26], and strong regulatory support [27], [28].

In this work, we propose and evaluate MPNL MIMO processing as a means of substantially improving power efficiency and vehicular connectivity in the context of Open-RAN-based vehicular network deployments. Examining different MIMO dimensions in both urban and highway mobility settings, we find that an Open-RAN, MPNL-based deployments can unlock substantial power consumption gains, requiring less than half the number of base station antennas compared to a conventional MMSE-based system to achieve the same packet error rate (PER) performance, and for the same system configuration. Furthermore, by examining indicative uplink-intensive V2I and V2N use cases [29], such as advanced driving, teleoperated / remote driving, and with respect to their estimated data rate requirements, we quantified the connectivity gains of MPNL in terms of concurrently supported vehicles per slot by a radio unit. Our findings reveal a substantial increase, in cases of more than 400%, in the number of concurrently connected vehicles, and without any performance loss, compared to an equivalent MMSE-based system. Moreover, by developing a Field Programmable Gate Array (FPGA)-Based architecture of

MPNL capable of supporting up to 16 MIMO streams in real-time, we quantified the system-level power gains of a practical MPNL deployment. Considering the computational power overhead of MPNL, we find that with an MPNL FPGA unit of 30 Watts of computational power, more than 400 Watts can be passively saved from the radio unit side. Finally, we propose a highly flexible method for delivering MPNL power savings in future practical deployments. Taking advantage of Open-RAN's RF Channel Reconfiguration that offers flexible a activation/deactivation of antenna elements based on traffic, we discuss a dynamic power state mechanism based on the intrinsic properties of MPNL to elastically trade PER performance for computational power by activating more parallel processing units.

II. BACKGROUND

In this Section, we discuss the basics of our adopted massively parallelizable non-linear MIMO detection approach as well as its integration to an Open RAN system, discussing the corresponding trade-offs of each deployment option.

A. MASSIVELY PARALLELIZABLE NON-LINEAR PROCESSING

Traditional, linear MIMO algorithms use linear transformations to convert the MIMO channel into separate SISO channels. This greatly simplifies the processing requirements and allows well-known SISO techniques to be applied to MIMO systems. However, this transformation to equivalent SISO channels with degraded characteristics leads to substantial underutilization of the available MIMO capacity. This, in practice, translates to limited effective throughput and the number of streams that can be concurrently transmitted. To overcome those limitations, linear transceivers rely on disproportionality, increasing the number of employed antennas and, inevitably, the corresponding power consumption.

An alternative approach involves reexamining the fundamentals of MIMO transceivers and developing strategies to minimize the probability of transmission errors in MIMO channels. However, addressing this challenge directly is highly complex and computationally intensive due to its inherently NP-hard nature. This complexity has led to the development of various non-linear algorithms aimed at reducing computational demands, though often at the expense of achievable throughput. On top of that, traditional non-linear processing solutions also often have impractical characteristics, including exponential computational complexity and limited capabilities for parallel processing. For example, a traditional non-linear soft detection of 24 information streams at a 17 dB Tx SNR such as [30] would necessitate a detection time of 15 minutes per sample, even by assuming a processing clock frequency of 1 GHz.

In response to those challenges, the Massively Parallelizable Non-Linear (MPNL) Processing framework was recently introduced [18], [19], [22], [23], [31]. The main principle of MPNL is, before the transmission and based only on the characteristics of the MIMO channel, to identify a subset of the possible vector solutions for the MIMO detection problem. This subset, identified in terms of relative positions from any possible transmitted vector, is selected to have the highest probability of containing the solution to the non-linear transmission problem (pre-processing). After the transmission begins, the set of relative position vectors (RPVs) is demapped into actual symbols and then processed in parallel to extract the required hard or soft information (post-processing). Due to MPNL's ability to identify and focus the available processing power only on the most promising solutions, MPNL can substantially decrease the required computational complexity, enabling efficient non-linear MIMO detection, scalable for large MIMO systems. It is worth noting that while MPNL can consistently improve performance by increasing the subset of examined RPVs, in practice, and as we discuss later in detail, a small number of RPVs (often single digit), flexibly allocated, is typically required to achieve the standard 10% PER. Consider that a traditional tree search approach would require examining Q^N potential vector solutions, where N is the number of MIMO layers, each layer modulated at an order Q.

When compared to recently proposed Deep Learning (DL) based MIMO detectors, such as GEPNet, RE-MIMO, DetNET, and MMNet, it was shown [32] that MPNL is significantly more efficient both in terms of computational complexity and error rate performance. Moreover, MPNL, unlike DL approaches, does not require training overhead and exhibits greater adaptability to the transmission conditions, which is particularly important in vehicular environments.

Lastly, it is significant to notice that while MPNL implementations are being discussed in [23], existing solutions exhibit a lack of flexibility, supporting only a limited number of MIMO configurations and RPVs, and are incapable of providing soft information essential for 5G-NR decoding. Moreover, this study is the first to systematically examine the gains and trade-offs of Open-RAN-enabled MPNL regarding power consumption and connectivity, particularly within the highly dynamic environment of vehicular communications.

B. FUNCTIONAL SPLIT DEPLOYMENTS FOR OPEN-RAN-ENABLED MPNL

The introduction of multiple functional splits [16], [33] in the Radio Access Network (RAN) architecture, between the Radio Unit (RU) and the Distributed Unit (DU), expands deployment flexibility, allowing for a range of MPNL processing solutions, each tailored for specific vehicular communication scenario. Specifically, for intra-PHY split options (i.e., option 7), their practical application lies in simplifying the radio unit by moving some of the most computationally intensive PHY functionalities from the RU to the DU. This has immediate results not only on the power consumption of the RUs but also on the manufacturing process, allowing for more compact and lighter RU deployments [34]. At the same time, it enables higher degrees of centralization for the baseband processing, allowing for

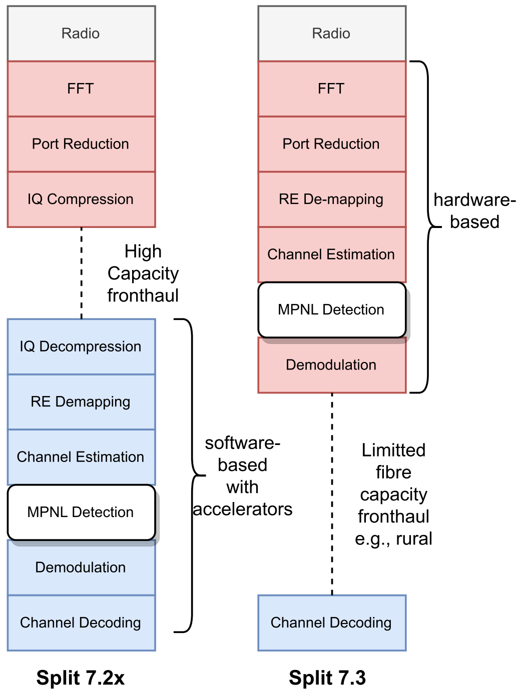


FIGURE 1. Functional split deployment options for MPNL. Red blocks indicate the functionalities realized in the radio unit (RU), while the blue ones indicate the functionalities realized in the distributed unit (DU).

potential joint processing, inter-cell coordination, and control optimization [35].

The first Open-RAN split option we explore is 7.2x, as shown in Figure 1 (left), where the primary PHY processing functions located in the Radio Unit (RU) are the Fourier transforms, with the remainder of the PHY processing tasks carried out on the Distributed Unit (DU). This setup is particularly suitable for urban areas with denser RU deployments. It strikes a balance between RU complexity and the need for high-capacity fronthaul bandwidth, making it ideal for urban settings, with better access to high-capacity fiber connections. Moreover, since the 7.2x split simplifies RU architecture, it effectively reduces the corresponding radio power consumption. The associated PHY processing, including the MIMO detector in this case, can be largely software-driven, potentially virtualized, and centralized. This creates the opportunity for a single DU to serve multiple RUs, and it could even take advantage of existing cloud service infrastructures. As we discuss later in Section IV-C2, this deployment method also presents substantial opportunities for leveraging the Massively Parallelizable Non-Linear (MPNL) Processing's capability to flexibly balance power between RU and DU according to traffic demands and available power budgets. In addition, it is important to mention that specifically split 7.2x has been selected by the ORAN alliance as the standard functional split for any current and ongoing ORAN deployment, counting over 25 active field trials and deployments announced by Mobile Network Operators.[1]

[1]https://map.o-ran.org/

The second scenario we consider in this work corresponds to a rural/highway scenario utilizing split 7.3 Fig.1 (right). In cases where access to fiber may be limited, limiting the fronthaul capacity requirement may be necessary. In such cases, the PHY processing chain, up to demodulation and including the MIMO detector, is implemented on the radio unit, reducing the fronthaul bandwidth requirement drastically compared to higher split options, making it proportional to the utilized modulation order. As shown in [35], the 7.3 split in downlink scenarios can reduce the frontal bandwidth requirements by up to 16 times compared to the 7.2x when employing QPSK modulation and can be four times more bandwidth-efficient with 256QAM. For uplink scenarios, the required fronthaul bandwidth can be reduced for up to four times in a 7.3 split, with four bits assigned to each soft information bit. The drawback is a more complicated RU design compared to splits such as 7.2x; therefore, low-power hardware-based designs might be required to keep RU within the required power budgets.

1) FUTURE EXTENSIONS OF 7.2X

It is worth noting that new potential extensions of 7.2x to additional intermediate splits are still part of an active discussion within the Open-RAN community. In particular, there's an active discussion within the ORAN alliance regarding extending the 7.2x split options to improve performance in uplink massive MIMO (mMIMO) systems for systems that employ more than 16 antennas at the radio unit. The current 7.2x split, which places the MIMO detector in the DU, has raised performance concerns in uplink massive MIMO [36], [37]. Particularly a performance gap observed between the 7.2x and 7.3 splits, that is often attributed to issues like channel aging, polluting the combining matrix required to be applied before the data is being transferred to the DU [38]. These concerns have opened the floor for further proposals under the term 'Uplink Pefromance Improvement' (ULPI). ULPI aims to enhance uplink performance for mMIMO by relocating parts of the MIMO detector and channel estimation to the RU. There are two main split classes that are currently under investigation by ORAN alliance. The first suggests placing DMRS processing in the RU, including the equalization function, while the second suggests having DMRS processing in both DU and RU. While the investigation of ULPI proposals on aspects such as the effect on RU complexity, on the fronthaul bandwidth requirements, and the backward compatibility of the proposals is still ongoing, a recent survey [39] indicates that more than a third of Operators currently realizing open ran solutions are waiting for ULPI to proceed with mMIMO deployments.

III. ADOPTED METHODOLOGY

In this Section, we discuss the methodology applied in this work. We present the procedures for simulating and evaluating MPNL in different vehicular connectivity scenarios (i.e., rural and urban). This allows, within the context of practical Open-RAN-based vehicular communication systems, to assess MPNL gains in terms of connectivity and power efficiency compared to the currently established linear detection algorithms.

A. SIMULATED SYSTEM

1) CHANNEL GENERATION

In assessing MPNL processing in the context of vehicular communications, we generated, for every examined MIMO dimension, a unique set of 100 independent cluster delay (CDL) MIMO channels. Specifically, we examined scenarios with 2, 4, 8, 12, and 16 MIMO streams corresponding to unique single-antenna vehicles, as well as RUs equipped with antennas ranging from 1 to 64. A comprehensive outline of our simulation parameters for channel generation and detector evaluation can be found in Table 1. Our evaluations, thoroughly reported and discussed in Section IV-C2, explore two distinct mobility scenarios. The first scenario represents urban mobility at a vehicle speed of 30 km/h, while the second represents highway mobility at a speed of 100 km/h. To account for the potential obstructions to line of sight (LOS) – such as buildings and other structures – between the transmitter and receiver in an urban setting, we assumed a CDL-B channel model, which is a non-line-of-sight outdoor channel model. In contrast, for highway scenarios, we used the CDL-D channel model, which considers both LOS and diffuse multipath components in the environment, presuming fewer obstructions between RU and the transmitting vehicle. The channels were generated utilizing MATLAB's 5G Toolbox *nrCDLChannel* with parameters presented in Table 1. Specifically, we assumed that vehicles are randomly distributed, with corresponding arrival angles varying between 0 and 60 degrees. We also assumed that the relative vehicle movement is either towards or away from the radio unit and that the communication is established at the 3.5GHz carrier frequency.

TABLE 1. Simulated system parameters.

Simulation Parameters	Values
Channel Models	CDL-B and CDL-D
Uncorrelated Channels per Scenario	100
MIMO Detection Algorithm	MMSE, MPNL
Modulation Coding Scheme (MCS)	9, 28
Supported MIMO Streams (N)	2 to 16
Base Station Antennas (M)	Up to 64
Vehicular Speed	30 km/h and 100 km/h
Packet Error Rate (PER) Target	10%
Relative Position Vectors (RPVs) (N_{RPV})	Up to 64
Signal Noise Ratio (SNR)	15, 25 dB
Carrier Frequency	3.5 GHz
Utilized Bandwidth (BW)	100 MHz
Subcarrier Spacing (SCS)	30 KHz
Number of Data Symbols (N^s_{data})	12
Number of DMRS Symbols (N^s_{dmrs})	2
Angles of Arrival (AoA)	0° to 60°
Maximum LDPC Iterations (I)	30

2) MODULATION CODING SCHEME SELECTION

To make comparative assessments, we conducted simulations on both the MPNL and the de facto standard Minimum Mean Square Error (MMSE) detectors. For each mobility scenario, we examined one low-rate and one high-rate transmission case, with the indicative corresponding modulation coding schemes (MCS) of 9 at 15dB SNR and 28 at 25dB SNR. We selected the indicative MCS 9 and 28 specifically because they correspond to the highest spectral efficiency for the constellations QPSK and 64-QAM, respectively. In 5G-NR, the Modulation and Coding Scheme (MCS) values are selected from three tables defined in 3GPP 38.214. The choice depends on UE capabilities and RAN configuration. This work focuses on 38.214-Table 5.1.3.1-1 due to its wide applicability in various channel conditions and its role as a common benchmark. Moreover, 38.214-Table 5.1.3.1-1 contains 28 MCS values, each corresponding to a different modulation order and coding rate, ordered in ascending order based on their associated targetted spectral efficiency. In practice, the UE measures channel conditions and reports them to the network, selecting an MCS value from the table that best matches these conditions. A higher MCS value is chosen under better conditions to maximize throughput, while a lower value is selected in worse conditions to maintain reliability. Different MCS configurations modify the number of transmission resources and, consequently, the transmission bandwidth. In this context, practical rate adaptation and user scheduling are open and challenging topics, particularly in vehicular environments, due to the highly fluctuating channel conditions, and require advanced Radio Resource Management (RRM) approaches beyond this paper's scope. In the sphere of Open RAN, these challenges are driving the ongoing development of different xAPPs and rAPPs. With the substantial data volumes that can be collected from the DUs and RUs, valuable data features and models can be identified and extracted to enhance intelligent management and control within the RAN, This information can then be combined with additional network-wide policies and external vehicular information [40] (i.e., vehicle localization and planned path) to drive fine-grained and real-time radio resource management. In this context, Machine Learning (ML)-based approaches are expected to become increasingly more important due to their ability to solve parameter-rich and computationally complex problems with no analytical solutions. Still, multiple additional open challenges appear in collecting the datasets and testing ML-based control at scale, as well as mitigating conflicts between RAN control decisions among different xAPPs/rAPPs [13], [41], [42].

3) SIMULATION METHODOLOGY

Our methodology involves identifying, for each of the examined mobility scenarios, MCS and number of supported MIMO streams, the minimum number of base station antennas required to achieve a packet error rate performance of 10 percent. For MPNL specifically, we also identified the minimum number of relative vectors (N_{rpv}) required to reach the PER target. The generated channels (Sec. III-A1) were grouped based on the number of supported MIMO streams N (i.e., 2, 4, 8, 12, 16). Each group comprised of 100 channels for each number of employed base station antennas M, where $M \in [N, 64]$. For each examined number of MIMO streams, starting from the smallest number of employed antennas ($M = N$), we evaluated over all the corresponding channels the average PER performance of each detector (i.e., MMSE and MPNL). If the resulted PER didn't satisfy our 10% maximum threshold, we increased the number of base station antennas by one and repeated the experiment either until the performance reached the acceptable threshold or until the number of antennas reached the maximum examined in our simulations, i.e., 64; in such case, the examined scenario is considered not supported. For the MPNL detector, an additional simulation loop was required since its PER performance is also a function of the number of examined solutions (N_{RPV}). Starting from a single examined solution, we gradually incremented N_{RPV} until the average BLER dropped below the acceptable threshold or until the number of solutions exceeded our maximum supported value of 64. In the cases where the targeted PER was not achievable, the experiment was repeated after increasing the number of antennas.

Furthermore, we assumed a transmission bandwidth of 100MHz within the 5G-NR's Frequency Range 1 (FR1), with a subcarrier spacing of 30kHz, corresponding to a slot duration of half a millisecond based on 5G-NR numerology. Each slot is presumed to consist of 12 data symbols and two symbols reserved for the demodulation reference signal (DMRS). We focussed on uplink heavy slots for two primary reasons: first, from a computational standpoint, uplink transmission is more computationally demanding than downlink (e.g., additional matrix inversions, more computationally demanding Low-Density Parity-Check (LDPC) decoder compared to the encoder). Hence, by assuming uplink transmission, we can estimate the upper bounds in terms of computational power consumption and processing latency in the case of Time Division Duplex (TDD) transmission. The second reason we emphasize uplink is that in the context of vehicular communications, and unlike traditional communications, V2X traffic is envisioned to be heavily uplink-based [43] since large data volume from multiple cameras and on-board sensors will need to be transmitted by the vehicles to the network infrastructure.

IV. RESULTS AND DISCUSSION

This Section discusses our findings regarding the power consumption reductions and the connectivity gains that MPNL can contribute to an Open-RAN C-V2X setting. Figure 2 presents our simulation results for the examined urban mobility scenario. Specifically, we show the number of base-station antennas required for supporting various high and low-rate MIMO streams (9,28 MCS) at a 10 percent PER target. This comparison is made between traditional MMSE

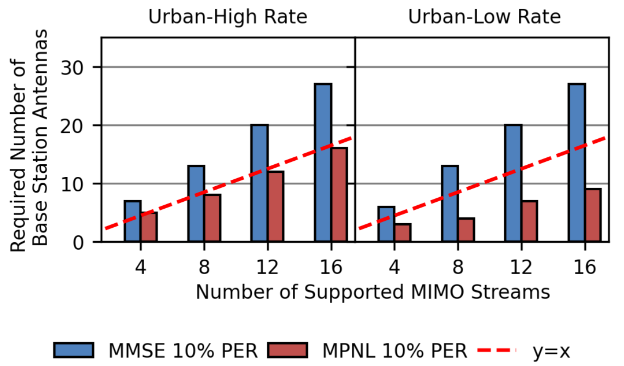


FIGURE 2. Comparison of the required base-station antennas to support urban mobility vehicles (30 km/h) using linear (MMSE) and non-linear (MPNL) MIMO processing. The analysis is presented for 10% Packet Error Rates (PER), considering both high-rate (MCS 28) and low-rate (MCS 9) Modulation and Coding Schemes.

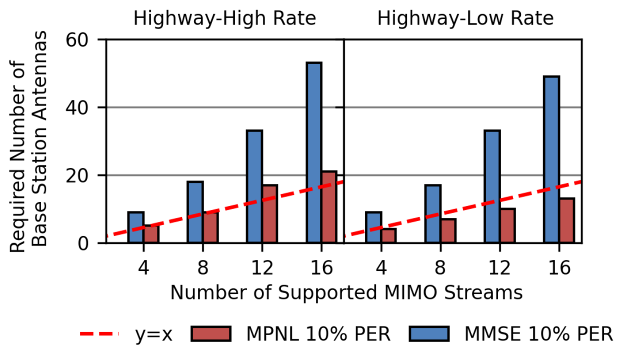


FIGURE 3. Comparison of the required base-station antennas to support highway mobility vehicles (100 km/h) using linear (MMSE) and non-linear (MPNL) MIMO processing. The analysis is presented for 10% Packet Error Rates (PER), considering both high-rate (MCS 28) and low-rate (MCS 9) Modulation and Coding Schemes.

processing and our proposed MPNL method. Similarly, in Figure 3, we provide the equivalent results corresponding to the examined highway mobility scenario. As shown, regardless of the mobility scenario, MPNL's gains in reducing the required number of antennas without PER loss compared to MMSE are substantial. In specific cases, the number of required antennas is cut by more than 3.7 times (i.e., for high-mobility and 16 low-rate MIMO streams). Moreover, we report an average reduction of 2.1 times in the required antennas for urban mobility scenarios at 10 PER, while an average of 2.5 times fewer antennas are needed for supporting highway mobility vehicles.

Furthermore, in the majority of the examined scenarios, MPNL can consistently support at least the same number of users as base station antennas. A noteworthy exception is observed in low-rate scenarios, especially urban mobility cases. In this case, MPNL can achieve the required PER performance with fewer base station antennas than the number of supported MIMO streams. This capability of MPNL to overload the employed RF chains is highly advantageous for use cases requiring massive connectivity with

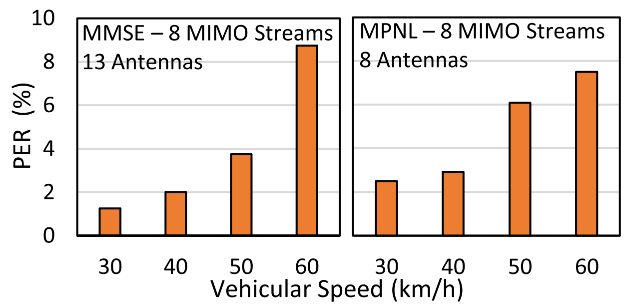


FIGURE 4. The PER for different vehicular speed values in an indicative urban mobility scenario of 8 concurrently transmitting vehicles.

lower data rate requirements without necessitating advanced Non-orthogonal multiple acces (NOMA) schemes [44].

Finally, Fig. 4 shows the PER for different vehicle speeds (i.e., 30 to 60km/h) in an urban mobility scenario for an indicative case of 8 concurrently transmitting vehicles. The base station antennas were fixed at the values obtained from the simulations in Fig. 2. That is 13 antennas for supporting 8 MIMO streams with MMSE detection and 8 antennas for the MPNL detector. While the vehicular speed within the examined range of 30-60km/h presents an observable increase in the PER, it does not recommend employing additional antennas, as it does not surpass the targetted 10% PER within the examined speed range. The rest of this Section is divided into two parts, aiming to translate the aforementioned gains into practical connectivity gains and system-level power consumption reductions.

A. MPNL FOR MASSIVE VEHICULAR CONNECTIVITY

As a measure of connectivity, we refer to the number of concurrently transmitted vehicles that can be scheduled within the duration of a slot to support the targeted data rate of the examined use case. To comprehensively understand the connectivity gains offered by non-linear processing, we've examined four use cases, each presenting a different indicative uplink data rate requirement [29], [45]. Our first examined use-case focuses on Advanced Driving, an encompassing term for features including autonomous and cooperative driving. We assumed a per-vehicle uplink throughput of 300Mbps for this application. Second, we consider remote or tele-operated driving, with an estimated uplink throughput of 50Mbps. Third, we examine Basic Safety features, which comprise collision avoidance, pedestrian safety, road condition alerts, and emergency vehicle alerts, estimated to necessitate an uplink data rate of 30Mbps. Our final examined use case is in-vehicle entertainment, which we assume a throughput requirement of up to 250Mbps.

Furthermore, for our analysis, we assumed 85% utilization of the 100MHz bandwidth (FR1) designated for uplink transmission in an urban mobility scenario involving single-antenna vehicles. We separately examined three base stations performing MMSE and MPNL MIMO detection with a fixed number of antennas, that is, 8, 16, 32 antennas, and a maximum of 16 MIMO Streams. Based on the required data rate per vehicle for the given use cases and the estimated spectral efficiency for different MCS values ranging from 0 to 28, we deduced the maximum number of vehicles that can be accommodated in a single slot for each examined use case. We assumed the scheduling unit of 1 resource block. The results of our analysis are shown in Table 2. The maximum number of supported vehicles per base station (NV_i^{max}) for a given use-case (i) can be approximated by Equation (1).

$$\mathrm{NV}_i^{\max} = \left\lfloor \frac{NRB}{\left\lceil \frac{R_\mathrm{i}}{\mathrm{SE_{MCS}} \times \mathrm{SCS} \times \mathrm{SC_{RB}}} \right\rceil} \right\rfloor \times \mathrm{N} \tag{1}$$

where R_i is the uplink throughput requirement corresponding to the use case i, $\mathrm{SE_{MCS}}$ is the spectral efficiency for a given MCS value, SCS is the subcarrier spacing, $\mathrm{SC_{RB}}$ is the number of subcarrier per resource block, NRB is the total number of resource blocks in a symbol and N is the number of supported MIMO streams by our base station. Moreover, in cases where the throughput requirement of the application is sufficiently low compared to the achievable spectral efficiency of the vehicles, that is when $\frac{R_\mathrm{i}}{\mathrm{SE_{MCS}} \times \mathrm{SCS} \times \mathrm{SC_{RB}}} < 1$ then the scheduling unit dictates the maximum supported vehicles in the slots $\mathrm{NV}_i^{\max} = NRB \times N$. Given the relationship between the number of required base station antennas and supported MIMO streams that we presented in Figures 2 and 3, we can reason about the range of vehicles that a fixed number of antennas can support under a specific target PER for each given use case. While it's important to acknowledge that signaling overhead can pose a practical constraint, it pertains to the optimization of the Physical Downlink Control Channel (PDCCH), a topic that exceeds the scope of this work. However, it is noteworthy to mention strategies such as Semi-Persistent Scheduling (SPS), introduced in 5G-NR, which could alleviate such overhead. Instead of scheduling each individual transmission separately—which would necessitate a considerable amount of signaling—SPS enables the network to prearrange a series of transmissions at fixed intervals.

Moreover, by assuming an indicative uplink power consumption per RF chain power at 15.6 W [46], Figure 5 shows the number of supported vehicles as a function of the radio unit power consumption under MMSE and MPNL detection. For this analysis, we focus on the urban mobility scenarios: basic safety, remote driving. We consider a maximum of 16 MIMO streams and a varying number of base station antennas. It's important to note that in a practical, real-world setting, and as we discuss in Sec. III-A2, vehicles may not all be transmitting at the same MCS. Thus, the figure presents a spectrum of values bounded by the maximum and the minimum MCS, where the intermediate values correspond to a mixture of vehicles transmitting at rates within the indicative range of MCS 9 to 28. As shown in Figure 5, the use of MPNL processing facilitates support for a maximum number of vehicles under 16 MIMO streams of MCS 28, equivalent to a power reduction of over 250 Watts

TABLE 2. The simulated gains of MPNL in terms of supported single-antenna vehicles within a slot compared to traditional MMSE for different V2I/V2N use-cases.

Use Case	Throughput (Mbps)	Max single antenna vehicles†	MPNL gains (in vehicles) for a maximum number of 16 MIMO streams ‡		
			8-Antennas RU	16-Antennas RU	32-Antennas RU
Advanced Driving	300	29	+11 (466%)	+15 (214%)	+7 (131%)
In-Vehicle Entertainment	250	35	+13 (425%)	+18 (225%)	+9 (134%)
Remote Driving	50	176	+66 (400%)	+88 (200%)	+44 (133%)
Basic Safety	30	293	+110 (405%)	+147 (201%)	+73 (133%)

†The maximum number of supported single antenna vehicles, assuming 16 MIMO streams, 100MHz of bandwidth utilized for uplink data transmission.
‡Compared to the equivalent MMSE-based system

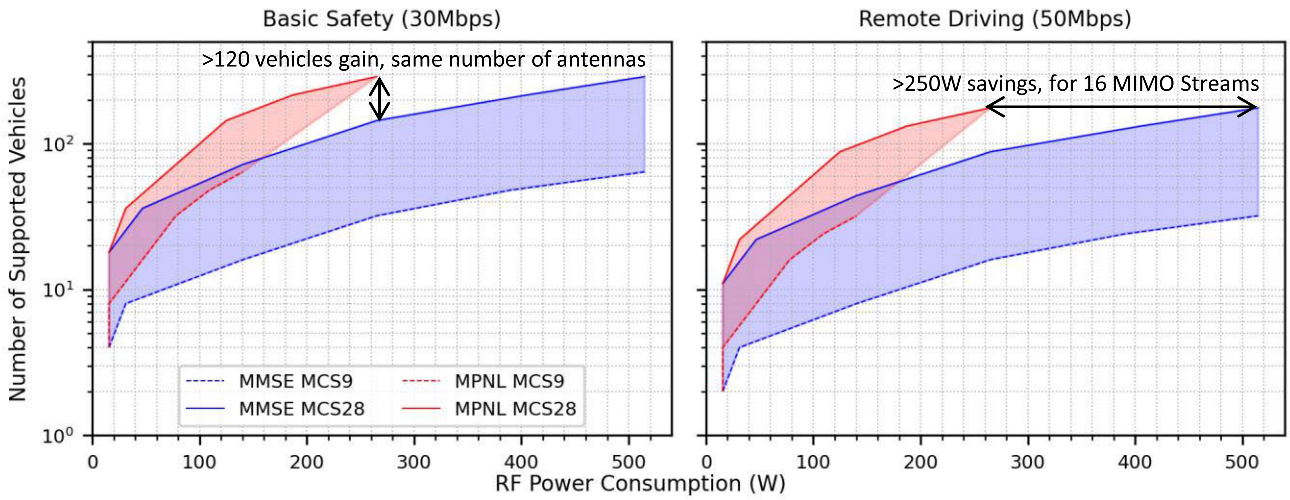


FIGURE 5. The number of supported vehicles as a function of the RF power consumption, given maximum support of 16 concurrently transmitted MIMO streams at 100MHz bandwidth. For different V2I/V2N use cases with different data rate requirements, comparing our proposed MPNL approach with traditional MMSE MIMO detection.

compared to an MMSE-based system. Further, a fixed power of 250 Watts on the RU, which is roughly equivalent to 16 RF chains, can support nearly double the quantity of vehicles by simply transitioning to MPNL detection.

B. MPNL FOR HIGH POWER EFFICIENCY

While in the previous Section, we discussed the gains of MPNL in terms of reducing the number of RF chains and enhancing device connectivity, in this Section, we will discuss the power consumption cost of MPNL and, therefore, provide the complete image of the overall gains of our proposed approach in the power domain. To provide this holistic assessment of our solution's power gains, we evaluated MPNL as an FPGA accelerator.

1) FPGA DESIGN

Our architecture is similar to Flexcore [23], with additional capabilities for the processing of soft information as described in [19]. The fundamental building block of our design is the Processing Element (PE), which corresponds to the fully instantiated logic required to process an RPV per clock cycle in a pipeline manner. Moreover, PEs can be spatially multiplied, allowing multiple RPVs and subcarriers to be processed in parallel, flexibly reducing the corresponding latency. For example, a single Processing Element (PE) can be instantiated to handle all the required RPV (N_{RPV}) in a pipeline manner, thus detecting a subcarrier for every N_{RPV} clock cycles. In a massively parallel scenario, a total of $N_{sbc} * N_{RPV}$ instances can be used to maximize the detection throughput to one entire data symbol per clock cycle. It is important to note that we distinguish between the number of RPVs (N_{RPV}) and the number of instantiated PEs (N_{PE}), considering that each PE is capable of serving one or more RPVs. This allows for a flexible trade-off between computational power (or design area), computational latency, and PER performance. For the evaluations presented in this Section, we restricted the maximum instantiated logic to 75% of the total utilization of our FPGA (Xilinx VCU118). This constraint ensures that performance is preserved by avoiding routing congestion [47]. The design operates at a clock frequency of 350MHz, and power consumption was estimated at the post-routing stage using the Vivado Power Estimator under worst-case static power conditions [48].

2) MPNL POWER EVALUTATION

Figure 6 presents the power consumption of our FPGA-based solution as a function of the number of parallel PEs instantiated and for supporting different numbers of MIMO streams. Our evaluations reveal a range between 2 and 45 watts to support up to 16 MIMO streams and up to

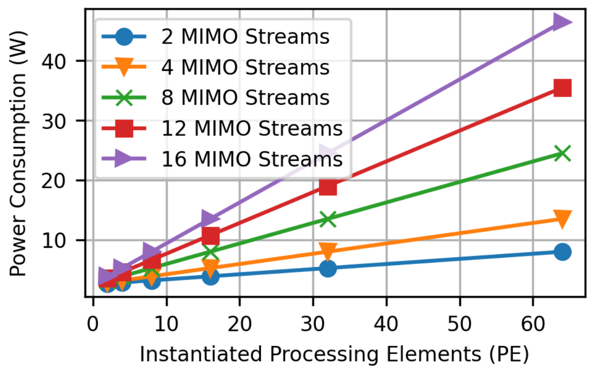


FIGURE 6. The dynamic FPGA power consumption for supporting MPNL detection for different number of concurrently transmitted MIMO Streams and different numbers of instantiated processing elements.

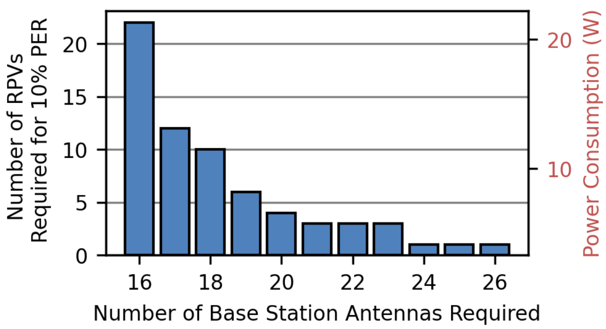


FIGURE 7. On the left axis, the number of required MPNL Paths to achive 10% PER in urban mobility scenario of supporting concurrently 16 high-rate (MCS28) MIMO streams. On the right axis the overall FPGA power consumption assuming a fully data-parallel implementation where each MPNL path is allocated to an individual parallel PE.

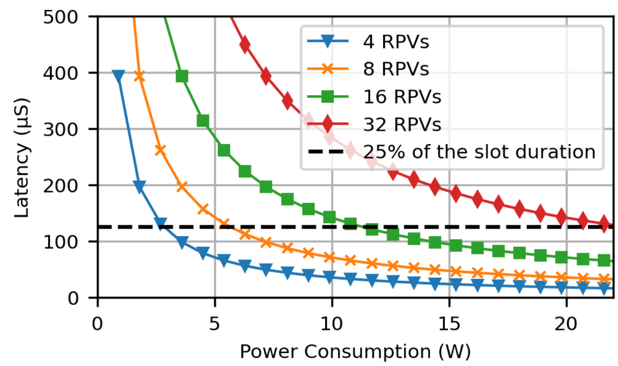


FIGURE 8. The FPGA processing latency versus the corresponding power consumption for supporting 16 MIMO streams at 100MHZ for different number of examined RPVs. The 25-percent of the slot duration is our latency target that dictates the number of employed parallel instances in our overall system evaluations.

64 parallel PEs. As already discussed, the PER performance of MPNL is a function of the number of examined RPVs. The number of employed RPVs, with respect to the processing latency targets and computational power budgets (or area budgets), determines the number of PEs to be instantiated in the design. Although MPNL can process any number of RPVs in parallel by spatially multiplying the instantiated PEs, this might not always be necessary. As such, a single parallel PE may sequentially process multiple RPVs, trading computational latency for power consumption.

To evaluate this trade-off, first, in Figure 7, we capture the number of RPVs required for achieving a maximum PER 10 %. The examined scenario corresponds to high-rate urban mobility vehicles, and the results show the required number of MPNL RPVs for supporting 16 MIMO streams with different numbers of employed antennas, from 16 to 26 antennas. On the right axis, we have the maximum power consumption requirement, assuming that each PE handles a single RPV. As shown, the required number of RPVs decreases rapidly with increasing the number of antennas. It is worth recalling here that in the same scenario, in order for MMSE to achieve a PER of 10 percent, it requires over 28 antennas (Figure 2). Moreover, as seen, in all examined cases the power gain from reducing the number of antennas by one is consistently larger than the power overhead related to the additional RPVs required to be examined. Even in the most demanding case examined, reducing the active number of antennas from 17 to 16, the power overhead of supporting the additional 12 RPVs is still 30 percent less than the assumed power-per-antenna value (i.e., 15.6 Watts [46]). Note that this is still under the assumption that a single RPV is being allocated to a single PE, which is the most power-consuming configuration that corresponds to minimum processing latency.

Furthermore, in Figure 8, we present the total processing latency of our FPGA MPNL accelerator supporting 16 MIMO streams as a function of its dynamic power consumption and for different numbers of MPNL RPVs. Assuming a latency budget equal to 25% of the total slot duration for 100MHz of FR1 with 30kHz SCS and 14 data symbols, we find that the power consumption of an RF chain is approximately equivalent to supporting 22 MPNL RPVs. It is worth mentioning here that 25% of the total slot duration is selected as the processing latency target, given that the real-time requirement of the entire PHY is to finish the corresponding processing within the slot duration. In practice, this real-time constraint can be further relaxed, allowing the PHY processing time to exceed the slot duration and, consequently, more time for the MIMO detection process. That is because 5G-NR allows scheduling a Hybrid ARQ(Automatic Repeat reQuest) response within a 4-slot timeframe.

In addition, we find that for the 100MHz bandwidth, approximately 0.81 PEs per RPV are sufficient to ensure an overall processing time of 25% of the slot length. In the case of utilizing a narrower continuous bandwidth, for instance, 50MHz, and for maintaining the same latency budget, a single PE instance can handle almost twice as many RPVs in real-time. Finally, given our previous discussion in Figure 9, we present the overall power savings of a base station running MPNL compared to MMSE. We assumed 10 percent PER and

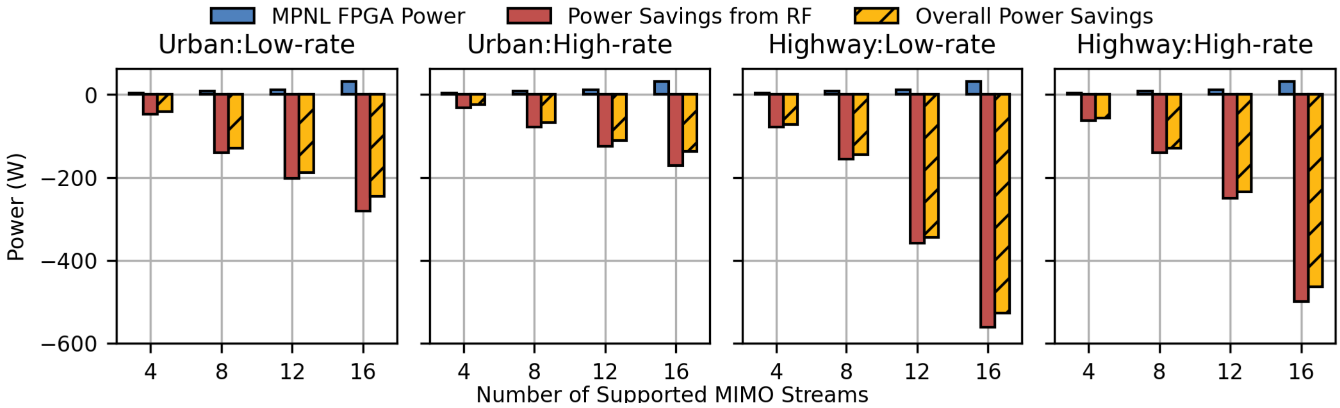


FIGURE 9. **The overall power gains for different vehicles with different mobility and rate, for supporting FPGA-based MPNL detector and different MIMO Streams.**

an overall processing latency for the detector equal to 25% of the overall processing slot (0.5ms). As seen, the overall power savings can exceed 200 Watts in urban mobility cases for the largest MIMO dimensions, while over 400 Watts can be saved from highway deployments.

C. IMPLICATIONS AND FUTURE DEVELOPMENTS

1) TOWARDS SUSTAINABLE, CENTRALIZED, AND SOFTWARE-BASED RAN

The implications of the demonstrated gains are substantial, particularly considering that the additional power overhead required by MPNL can be centralized (i.e., Open RAN split 7.2x) and potentially realized to cloud-based infrastructures. These facilities can employ advanced resource allocation methods and centralized "smart" cooling systems to enhance their power output and energy efficiency. Moreover, the fact that the power gains of MPNL originate from the radio side is particularly significant. By reducing the power consumption of the RF-Chains, often by more than half, more flexible and potentially denser deployments of radio/roadside units can be realized, with the potential to support even solar-powered MIMO roadside infrastructure deployments. Additionally, Open-RAN realizations are expected to rely heavily on software solutions, that is due to software's inherent flexibility, expedited development and ability to integrate of new features faster. However, it's worth noting that the power efficiency of software solutions tends to be orders of magnitude lower than that of hardware-based alternatives [49]. Therefore, the power gains provided by MPNL deployment can enable a larger portion of the system to transition towards more flexible software-based platforms without sacrificing system-level power efficiency. For example, a power reduction of 200 Watts is approximately equivalent to deploying a modern multicore server for performing real-time MIMO PHY processing. Still, the efficient softwarization of MIMO PHY processing remains an open challenge due to the very stringent processing latency constraints of 5G-NR PHY. A very first live demo of a software-based MPNL in real-time and over-the-air has been presented at [50]. Still exploring higher bandwidths and larger MIMO dimensions would require employing more aggressive parallelization schemes harnessing multiple processing cores as well as the extensive vectorization of the individual processing functions. Exploring the gains of MPNL in a software-driven PHY framework represents a promising research avenue, indicating a significant area for future research.

2) ADVANCED OPEN RAN RF-CHANNEL RECONFIGURATION

The power gains presented by the use of MPNL can work additively with existing base station power-saving strategies, such as sleep states [51]. Importantly, recent Open RAN proposals such as the RF Channel Reconfiguration [52] allows for embedding even greater flexibility and intelligence in the base station power management. With RF-Channel Reconfiguration, in periods of low load, when the expected traffic volume or the number of connected users is below a configured threshold, power consumption can be reduced by switching off RF elements, potentially with a symbol-level granularity [53]. This process can be managed either in the Non-Real Time RAN Intelligent Controller (Non-RT RIC) or in the Near-Real Time RIC, where KPIs and power measurements collected from the O-RUs can be used by xApps/rApps to determine reconfiguration recommendations. With MPNL, the gains of RF-Channel Reconfiguration are not only passively amplified, approximately by a factor of 2, but it also allows for a much more flexible power-saving trade-off. The intelligent controller can manage the allocation of active processing elements, increasing computational power while keeping RF power below a specified threshold or vice versa. For instance, an increase in traffic can now be managed either by increasing the active number of antennas on the radio side or by allocating more MPNL PEs on the baseband processing side. This flexibility fosters a dynamic power-saving strategy, which can adapt effectively to various demands (i.e., power budgets, QoS) and conditions. Essentially, it transforms the concept of distinct sleep stages model into a much more dense spectrum of power modes, which can span from activating and deactivating antennas to launching additional PE or reallocating the existing MPNL PEs from one RU to another in a centralized location.

V. RELATED WORK

Numerous PHY processing solutions have been proposed in the literature. In the sphere of software-based Open-RAN solutions, prominent open-source 3GPP-based projects such as OpenAirInterface (OAI) [54] and srsRAN [55] offer complete and software-driven base station implementations. Specifically, srsRAN is capable of supporting all the FR1 bands in real-time and to the best of our knowledge is the only openly available solution that currently includes a complete implementation of the 3GPP Sidelink PHY [56]. Still, it offers very limited MIMO support. Moreover, OpenAirInterface (OAI) (and solutions based on OAI [57], [58]), despite perhaps being the most advanced open platform that supports 5G-NR and provides software implementations for both RAN and User Equipment (UE), it is still not capable of supporting MU-MIMO processing with a large number of concurrently transmitted information streams. Earlier notable MIMO PHY realizations, including BigStation [59] and Agora [60] explore larger MIMO dimensions, exploiting aggressive pipelining and data parallelization techniques. However, both projects rely on linear processing techniques (MMSE) without accounting for the consequential practical power implications both on the radio side or the processor side. In this work, our targets partially align with those studies, as we target real-time large MIMO processing capable of meeting substantial traffic and uplink-throughput demands. However, we differentiate from them by departing from conventional linear MIMO approaches. We examine and propose massively parallelizable non-linear processing, heavily investigating its gains and trade-offs in the power consumption and connectivity domains.

Moreover, our research relates to previous works evaluating massively parallelizable non-linear MIMO detectors. Previous studies have found limited integration into practical systems. This can be attributed to the substantial computational efforts required [18], [19], and to the absence of essential 5G-NR features such as the computation of soft information [23]. It's also important to note that this is the first time the power and connectivity benefits associated with Massively Parallel Non-Linear (MPNL) are being thoroughly analyzed on a system level, within the context of practical Open RAN V2X deployment. Our research also relates to existing studies on power-saving mechanisms for base stations, specifically those concerning base station sleep states [51], and RF chain activation/deactivation [52], [53], [61], [62]. In this work, we go a step further; apart from passively amplifying the power savings discussed in these studies, our work discusses a more flexible and dynamic power-saving mechanism built on the foundation of Open RAN's RF Channel Reconfiguration [52] and the properties of the discussed MPNL algorithm as discussed in Sec.IV-C2.

VI. CONCLUSION AND FUTURE WORK

In this study, we discuss and evaluate the potential of MPNL MIMO processing in substantially improving the system-level power efficiency and connectivity capabilities in Open RAN vehicular networks, assessing its impact in various V2I and V2N scenarios. Our FPGA-based evaluation underscores MPNL's substantial power efficiency gains, in cases exceeding 400W reduction in the RU power consumption by reducing the need for numerous base station antennas and RF chains. Additionally, we quantified MPNL's connectivity gains, showing gains of over 400% in terms of concurrently transmitting vehicles, compared to traditional processing approaches and without a significant increase in the power consumption targets.

This work has been introductory to the HiPer-RAN project,[2] examining the feasibility and highlighting the potential of Open-RAN-based MPNL processing in advancing toward more intelligent, autonomous, and sustainable transportation systems. In future works, we will consider holistic aspects of RAN, including the rate adaptation challenge and the development of xAPPs and rAPPs for intelligent and energy-efficient Radio Resource Management (RRM) that pose significant challenges, particularly in vehicular environments with highly fluctuating channel conditions. Furthermore, future work will involve the creation of prototypes capable of fully leveraging the parallelization potential inherent in MPNL approaches, as well as the joint optimization of both radio and computational resources in order to deliver adaptively the discussed gains in energy efficiency.

REFERENCES

[1] *NR and NG-RAN Overall Description: Stage-2*, Standard TS 38.300, 3GPP, Release 16, 2021.
[2] J. Zhang and K. B. Letaief, "Mobile edge intelligence and computing for the Internet of Vehicles," *Proc. IEEE*, vol. 108, no. 2, pp. 246–261, Feb. 2020.
[3] A. Y. Alhilal, T. Braud, and P. Hui, "Distributed vehicular computing at the dawn of 5G: A survey," 2001, *arXiv:2001.07077*.
[4] OFCOM. (2022). *Communications Market Report*. [Online]. Available: https://www.ofcom.org.uk/
[5] T. Karunathilake and A. Förster, "A survey on mobile road side units in VANETs," *Vehicles*, vol. 4, no. 2, pp. 482–500, May 2022.
[6] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.
[7] H. Holtkamp, G. Auer, S. Bazzi, and H. Haas, "Minimizing base station power consumption," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 297–306, Feb. 2014.
[8] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, C. Desset, and O. Blume, "Cellular energy efficiency evaluation framework," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–6.

[2] HiPer-RAN Project https://hiper-ran.com/

[9] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

[10] R. Atallah, M. Khabbaz, and C. Assi, "Energy harvesting in vehicular networks: A contemporary survey," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 70–77, Apr. 2016.

[11] Y.-J. Ku and S. Dey, "Sustainable vehicular edge computing using local and solar-powered roadside unit resources," in *Proc. IEEE 90th Veh. Technol. Conf.*, Sep. 2019, pp. 1–7.

[12] M. Yang, Y. Li, D. Jin, L. Su, S. Ma, and L. Zeng, "OpenRAN: A software-defined ran architecture via virtualization," in *Proc. ACM Conf. SIGCOMM*, Aug. 2013, pp. 549–550.

[13] W. Azariah, F. A. Bimo, C.-W. Lin, R.-G. Cheng, N. Nikaein, and R. Jana, "A survey on open radio access networks: Challenges, research directions, and open source approaches," *Sensors*, vol. 24, no. 3, p. 1038, Feb. 2024.

[14] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Commun. Surveys Tuts.*, pp. 1376–1411, 4th Quart., 2023, doi: 10.1109/COMST.2023.3239220.

[15] OFCOM. (Jan. 2022). *Open RAN and the Link Between Competition and Innovation*. [Online]. Available: https://www.ofcom.org.uk/research-and-data/economics-discussion-papers/open-ran-competition-innovation

[16] O-RAN Alliance, "O-RAN architecture description 8.0," Tech. Rep. OAD-R003-V08.00, Mar. 2023.

[17] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.

[18] K. Nikitopoulos, J. Zhou, B. Congdon, and K. Jamieson, "Geosphere: Consistently turning MIMO capacity into throughput," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, pp. 631–642, Aug. 2014.

[19] C. Jayawardena and K. Nikitopoulos, "G-MultiSphere: Generalizing massively parallel detection for non-orthogonal signal transmissions," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 1227–1239, Feb. 2020.

[20] J. Jalden and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[21] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[22] K. Nikitopoulos, "Massively parallel, nonlinear processing for 6G: Potential gains and further research challenges," *IEEE Commun. Mag.*, vol. 60, no. 1, pp. 81–87, Jan. 2022.

[23] C. C. M. Husmann, G. Georgis, K. Nikitopoulos, and K. Jamieson, "FlexCore: Massively parallel and flexible processing for large MIMO access points," in *Proc. 14th USENIX Symp. Netw. Syst. Des. Implement. (NSDI)*, 2017, pp. 197–211.

[24] K. Nikitopoulos and R. Tafazolli, "Parallel processing of sphere decoders and other vector finding approaches using tree search," WIPO Patent 2 016 198 845 A1, Aug. 6, 2019.

[25] S. Gyawali, S. Xu, Y. Qian, and R. Q. Hu, "Challenges and solutions for cellular based V2X communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 222–255, 1st Quart., 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9217500

[26] G. Naik, B. Choudhury, and J.-M. Park, "IEEE 802.11BD & 5G NR V2X: Evolution of radio access technologies for V2X communications," *IEEE Access*, vol. 7, pp. 70169–70184, 2019.

[27] Federal Commun. Commission (FCC). (2020). *FCC Modernizes 5.9 GHz Band to Improve Wi-Fi and Automotive Safety*. Accessed: Jan. 25, 2024. [Online]. Available: https://www.fcc.gov/document/fcc-modernizes-59-ghz-band-improve-wi-fi-and-automotive-safety

[28] Federal Commun. Commission (FCC). (2023). *Request for Waiver of 5.9 GHz Band Rules To Permit Initial Deployment of Cellular Vehicle-toEverything Technology*. Accessed: Jan. 25, 2024. [Online]. Available: https://docs.fcc.gov/public/attachments/DA-23-343A1.pdf

[29] A. Bishen. (2021). *Vehicular Connectivity: C-V2X and 5G*. [Online]. Available: https://www.5gamericas.org/vehicular-connectivity-c-v2x-and-5g/

[30] C. Studer, A. Burg, and H. Bolcskei, "Soft-output sphere decoding: Algorithms and VLSI implementation," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 290–300, Feb. 2008.

[31] K. Nikitopoulos, *Massively Parallel and Flexible Processing for MIMO Systems*. Hoboken, NJ, USA: Wiley, 2019, pp. 1–21.

[32] J. C. De Luna Ducoing, C. Jayawardena, and K. Nikitopoulos, "An assessment of deep learning versus massively parallel, non-linear methods for highly-efficient MIMO detection," *IEEE Access*, vol. 11, pp. 97493–97502, 2023.

[33] *Study on New Radio Access Technology: Radio Access Architecture and Interfaces*, Standard TR 38.80, 3rd Gener. Partnership Project (3GPP), 2017.

[34] O-RAN Work. Group 4. (2023). *Control, User and Synchronization Plane Specification*. Accessed: May 5, 2023.

[35] V. Q. Rodriguez, F. Guillemin, A. Ferrieux, and L. Thomas, "Cloud-RAN functional split for an efficient fronthaul network," *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, pp. 245–250, 2020.

[36] C. Gabriel. (2023). *New O-RAN Interface May Address Massive MIMO Challenge at the Expense of Supply-chain Diversity*. Accessed: Jan. 25, 2024. [Online]. Available: https://www.analysysmason.com

[37] 5G Americas. (2023). *Open RAN Update*. Accessed: Jan. 25, 2024. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2023/11/Open-RAN-Update-Id.pdf

[38] K. Muralidhar, L. Vijayan, J. Abdoli, and Y.-H. Nam, "Linear channel prediction-based receiver for split 7.2 uplink massive MIMO," in *Proc. IEEE 33rd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*. Kyoto, Japan: IEEE, Sep. 2022, pp. 1093–1098.

[39] *Operator Survey: Open RAN Splits for 5G Massive MIMO: Heavy Reading*, 2023.

[40] F. Linsalata, E. Moro, M. Magarini, U. Spagnolini, and A. Capone, "Open RAN-empowered V2X architecture: Challenges, opportunities, and research directions," 2023, *arXiv:2303.06938*.

[41] C. Adamczyk and A. Kliks, "Conflict mitigation framework and conflict detection in O-RAN near-RT RIC," *IEEE Commun. Mag.*, vol. 61, no. 12, pp. 1–7, Dec. 2023.

[42] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "ColO-RAN: Developing machine learning-based XAPPS for open RAN closed-loop control on programmable experimental platforms," 2021, *arXiv:2112.09559*.

[43] H. Yang and E. G. Larsson, "Can massive MIMO support uplink intensive applications?" in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.

[44] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[45] M. H. C. Garcia, A. Molina-Galan, M. Boban, J. Gozalvez, B. Coll-Perales, T. Sahin, and A. Kousaridas, "A tutorial on 5G NR V2X communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1972–2026, 3rd Quart., 2021.

[46] Y. Gong, L. Zhang, R. Liu, K. Yu, and G. Srivastava, "Nonlinear MIMO for industrial Internet of Things in cyber–physical systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5533–5541, Aug. 2021.

[47] D. J. Smith and M. Dini, *FPGA-Based Prototyping Methodology Manual: Best Practices in Design-For-Prototyping*. Amsterdam, The Netherlands: Elsevier, 2011.

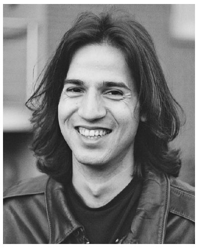
[48] Xilinx. (2022). *Vivado Design Suite User Guide: Power Analysis and Optimization*.

[49] H. Lee, K. Oh, M. Cho, Y. Jang, and J. Kim, "Efficient low-latency implementation of CORDIC-based sorted QR decomposition for multi-gbps MIMO systems," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 65, no. 10, pp. 1375–1379, Oct. 2018.

[50] K. Nikitopoulos, M. Filo, G. N. Katsaros, C. Jayawardena, and R. Tafazolli, *MU-MIMO, Open-RAN PHY With Linear and Massively Parallelizable Non-Linear Processing*. New York, NY, USA: Association for Computing Machinery, 2023.

[51] J. Wu, Y. Zhang, M. Zukerman, and E. K. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 803–826, 2nd Quart., 2015.

[52] O-RAN Allianc, ‘‘O-RAN network energy savings use cases technical report 1.0,’’ Tech. Rep. R003, Mar. 2023.
[53] Deutsche Telekom, Orange, Telefónica, TIM, and Vodafone. *Open RAN Technical Priorities-Focus on Energy Efficiency*. [Online]. Available: https://www.vodafone.com/about-vodafone/what-we-do/technology/open-ran/research-reports
[54] N. Nikaein, M. K. Marina, S. Manickam, A. Dawson, R. Knopp, and C. Bonnet, ‘‘OpenAirInterface: A flexible platform for 5G research,’’ *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 5, pp. 33–38, Oct. 2014.
[55] SRSRAN. *Software Radio Systems*. [Online]. Available: https://www.srslte.com/
[56] SRSRAN. (2023). *C-V2X Application Note*. [Online]. Available: https://docs.srsran.com/
[57] G. Georgis, M. Filo, A. Thanos, C. Husmann, J. C. De Luna Ducoing, R. Tafazolli, and K. Nikitopoulos, ‘‘SWORD: Towards a soft and open radio design for rapid development, profiling, validation and testing,’’ *IEEE Access*, vol. 7, pp. 186017–186040, 2019.
[58] M. Filo, Y. Xia, and K. Nikitopoulos, ‘‘SACCESS: Towards a software acceleration framework for 5G radio access networks,’’ in *Proc. IEEE Int. Medit. Conf. Commun. Netw. (MeditCom)*, Sep. 2021, pp. 318–323.
[59] Q. Yang, X. Li, H. Yao, J. Fang, K. Tan, W. Hu, J. Zhang, and Y. Zhang, ‘‘BigStation: Enabling scalable real-time signal processingin large mu-MIMO systems,’’ *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 399–410, Aug. 2013, doi: 10.1145/2534169.2486016.
[60] J. Ding, R. Doost-Mohammady, A. Kalia, and L. Zhong, ‘‘Agora: Real-time massive MIMO base band processing in software,’’ in *Proc. 16th Int. Conf. Emerg. Netw. Exp. Technol.*, Spain: ACM, Nov. 2020, pp. 232–244.
[61] M. Feng, S. Mao, and T. Jiang, ‘‘Dynamic base station sleep control and RF chain activation for energy-efficient millimeter-wave cellular systems,’’ *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9911–9921, Oct. 2018.
[62] *Study on Network Energy Savings for NR*, Standard TR 38.864, V18.1.0, 3rd Gener. Partnership Project (3GPP), 2023.



GEORGE N. KATSAROS received the B.Sc. degree in physics and the M.Sc. degree in control and computing from the University of Athens, Greece, in 2018 and 2020, respectively. He is currently pursuing the Ph.D. degree with the 5G/6G Innovation Center (5G/6GIC), University of Surrey, U.K. His research interests include architectural and algorithmic design of real-time and power-efficient physical layer architectures, with a particular application to future wireless and open-RAN communication systems.

KONSTANTINOS NIKITOPOULOS (Senior Member, IEEE) is currently an Associate Professor (Reader) with the Institute for Communication Systems, University of Surrey, U.K., and the Director of the Wireless Systems Laboratory. He is also an Active Academic Member of the 5G/6G Innovation Centre (5G/6GIC), where he leads the ‘‘Theory and Practice of Advanced Concepts in Wireless Communications’’ research area. His research interests include the physical layer (PHY) aspects of pragmatic, energy and latency-efficient wireless communication systems that ‘‘work in practice’’ and, in particular, on the intersection of advanced, highly efficient, physical layer (PHY) processing design, advanced computing architectures, and system level design and demonstration. He is also the Technical Lead of the Open-RAN Implementation of the University of Surrey. He was a recipient of the Prestigious First Grant of the U.K.'s Engineering and Physical Sciences Research Council.