IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 4, 2021, accepted April 21, 2021, date of publication May 12, 2021, date of current version May 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3078721

Non-Linear Base-Station Processing Within a 3GPP Compliant Framework

KONSTANTINOS NIKITOPOULOS, (Senior Member, IEEE), **MARCIN FILO**,
CHATHURA JAYAWARDENA, (Member, IEEE), **J. C. DE LUNA DUCOING**, (Member, IEEE),
AND RAHIM TAFAZOLLI, (Senior Member, IEEE)
Wireless Systems Laboratory, 5G and 6G Innovation Centre, Institute for Communication Systems, University of Surrey, Guildford GU2 7XH, U.K.
Corresponding author: Konstantinos Nikitopoulos (k.nikitopoulos@surrey.ac.uk)

ABSTRACT MIMO mobile systems, with a large number of antennas at the base-station side, enable the concurrent transmission of multiple, spatially separated information streams, and therefore, enable improved network throughput and connectivity both in uplink and downlink transmissions. Traditionally, such MIMO transmissions adopt *linear* base-station processing, that translates the MIMO channel into several single-antenna channels. While such approaches are relatively easy to implement, they can leave on the table a significant amount of unexploited MIMO capacity and connectivity capabilities. Recently-proposed *non-linear* base-station processing methods claim this unexplored capacity and promise substantially increased network throughput and connectivity capabilities. Still, to the best of the authors' knowledge, non-linear base-station processing methods not only have not yet been adopted by actual systems, but have not even been evaluated in a standard-compliant framework, involving all the necessary algorithmic modules required by a practical system. In this work, for the first time, we incorporate and evaluate non-linear base-station processing in a 3GPP standard environment. We outline the required research platform modifications and we verify that significant throughput gains can be achieved, both in indoor and outdoor settings, even when the number of base-station antennas is much larger than the number of transmitted information streams. Then, we identify missing algorithmic components that need to be developed to make non-linear base-station practical, and discuss future research directions towards potentially transformative next-generation mobile systems and base-stations (i.e., 6G) that explore currently unexploited non-linear processing gains.

INDEX TERMS Multi-user (MU)-MIMO, signal detection, precoding, non-linear signal processing.

LIST OF ACRONYMS

3GPP	3rd Generation Partnership Project
AMC	adaptive modulation and coding
BS	base-station
CSI	channel state information
DCI	downlink control information
DM-RS	demodulation reference signal
HARQ	hybrid automatic repeat request
LO	local oscillator
LOS	line-of-sight
MCS	modulation and coding scheme
MIMO	multiple-input multiple-output
MMSE	minimum-mean-square-error
MPNL	massively parallel non-linear
MU	multi-user
NL	non-linear
NR	New Radio
OAI	OpenAirInterface
OFDM	orthogonal frequency-division multiplexing
OTA	over-the-air
PDSCH	physical downlink shared channel
PE	processing element
PUSCH	physical uplink shared channel
RRC	radio resource control
RT	real-time
SDR	software-defined radio
SNR	signal-to-noise ratio
SRS	sounding reference signal
SWORD	SoftWare Open Radio Design
TDD	time-division duplexing
TPC	transmit power control

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

UE	user equipment
ULA	uniform linear array
USRP	Universal Software Radio Peripheral
VP	vector perturbation
ZF	zero-forcing

I. INTRODUCTION

Much of the current communication systems research focuses on finding new, breakthrough ways to increase the achievable throughput (both at a user and system-level) and user connectivity capabilities, while meeting very tight latency requirements. In this direction, a plethora of ideas have been proposed. Still, very few of these ideas, and perhaps the simplest in terms of practical realization, have finally been adopted by actual wireless systems and standards. In the natural question "why is this happening?" someone can give several answers. In many of the works published, the proposed ideas are only evaluated via simulations and, therefore, the results may be heavily *assumption-dependent*. Namely, the showed gains can be a strong function of the simulated environment that can sufficiently differ from the actual transmission environment. To facilitate more realistic evaluations, many researchers use "proof-of-concept" systems. However, this approach comes with its own challenges and practical limitations. Such a challenge is the availability of appropriate research platforms able to realize and validate proposed novel ideas. Additionally, in many cases and especially in physical layer research, the proposed ideas are not evaluated in a complete, standard-compliant environment. As a result, *additional algorithmic components* may be required to make a new idea adoptable by a practical system or communication standards. As we will discuss later in detail, such components are often related to limitations imposed from the existing system design. These limitations are often related to conventional signaling procedures or other mechanisms required for transmission optimization, such as transmission rate selection. As a result, implementing and evaluating new physical layer concepts and ideas in a standard-compliant environment can be of substantial importance. This is not only to verify the potential gains in more realistic transmission scenarios, but also to identify any need for new "building blocks" or required modifications in the standards (e.g., the signaling, pilots, or access methods). In other words, testing and verifying new physical layer ideas in a research-grade, standard-compliant environment can be an important step towards future systems design and evaluation. This can highly increase our confidence in newly-proposed approaches and help identify further requirements and missing components that will enable the adoption of novel ideas in actual systems.

In this work, we focus on recently-proposed ideas to improve base-station (BS) processing in multiple-input multiple-output (MIMO) spatially-multiplexed systems. The use of a large number of antennas at the BS side has been shown to be a very efficient way to increase the achievable throughput and the user connectivity capabilities of mobile systems, both in uplink and downlink transmissions, by enabling several concurrently-transmitting, spatially-separated users (i.e., multi-user (MU)-MIMO) [1]–[3]. Traditionally, in such systems, *linear* precoding (in the downlink) and detection (in the uplink) approaches are employed at the BS, based on the zero-forcing (ZF) or the minimum-mean-square-error (MMSE) principles. Such linear approaches have two major practical benefits. Their implementation is relatively simple, and since they practically translate the mutually interfering information streams into traditional, non-interfering ones, they can be easily adopted by standards with minimum modifications. Still, their main drawback is that in order for these approaches to be efficient in terms of achievable throughput, the number of BS antennas needs to be much larger than the number of concurrently-transmitted information streams, and, therefore, the number of served users [1], [2]. However, since by increasing the number of antennas, the capacity of the MIMO channel generally increases [4], such an approach leaves on the table a significant amount of unexploited capacity [5]. Equivalently, it unnecessarily increases the number of BS antennas for a certain number of users, significantly increasing the BS cost and reducing power efficiency. In contrast, *non-linear* BS processing approaches, like "hard" and "soft" *maximum-likelihood* detection, in the uplink [5], [6], and vector perturbation (VP) in the downlink [7], promise substantially increased achievable throughput and user connectivity. Still, to the best of our knowledge, such approaches have not yet been adopted by practical systems and have not even been validated in a standard-compliant scenario. Additionally, and perhaps as a consequence, it is not obvious what further system changes are required to deliver the promised gains in practice.

Contributions of This Paper: This work focuses on the integration of non-linear processing into a 3GPP framework. In particular, in this work, we:

- for the first time, incorporate and evaluate non-linear base-station processing in a 3GPP standard-compliant environment;
- outline the corresponding necessary improvements to available research platforms;
- verify that significant throughput gains can be achieved, both in indoor and outdoor settings, and even when the number of base-station antennas is much larger than the number of transmitted information streams;
- identify the missing algorithmic components that are further required to make non-linear processing practical for future base-station processing; and
- discuss some future research directions towards potentially-transformative next-generation mobile systems and base-stations (i.e., 6G) that explore currently unexploited non-linear processing gains.

II. OUTLINE OF NON-LINEAR PROCESSING TECHNIQUES FOR BASE-STATION PROCESSING

As discussed, current MIMO deployments mostly employ linear BS processing, both for uplink and downlink transmissions, but such approaches may leave a significant amount of capacity unexploited. Instead, non-linear BS processing approaches promise consistent gains compared to the linear ones, both in terms of achievable throughput and user connectivity. For uplink transmissions, "hard" *maximum-likelihood* detection methods have been proposed, both exact and approximate, and with most of them being realized in terms of sphere decoding [5], [8]–[12]. However, most of these approaches have been evaluated using simulations and assuming *Rayleigh fading* or other mathematically-modeled MIMO channels. While such mathematical models are convenient for the theoretical analysis of such systems, they do not necessarily capture the spatial multiplexing capabilities of the actual MIMO channels. In addition, the provided performance of these methods is often presented in terms of (uncoded) bit-error-rate, which is not adequate for evaluating system-level throughput gains. The sphere decoding approaches of [5], [9], [13] are evaluated in actual transmission channel environments and in terms of achievable throughput. Still, their evaluations are based on a very limited number of transmission rates (i.e., combinations of QAM constellation size and coding rates) that are selected based on their average performance. Furthermore, since the processing takes place off-line, the reported achievable rates do not include the impact of the higher layers of the protocol stack (e.g., the impact of the hybrid automatic repeat request (HARQ) mechanism). Moreover, "hard" detection cannot be used jointly with state-of-the-art "soft" channel encoding and decoding schemes (e.g., LDPC, Turbo) adopted in current standards, and therefore, are of limited practical interest.

For use with soft channel decoding schemes, soft-output sphere decoders have been proposed to reduce the complexity of optimal soft detection [6], [14]–[16]. Again, most soft sphere decoding approaches, including the sequential sphere decoder of [6] and the soft fixed complexity sphere decoder of [15], are evaluated by assuming mathematically-modeled MIMO channels. The massively parallel hard and soft detectors of [17]–[19] that enable practical, low complexity, and low latency non-linear detection, are evaluated both in mathematically-modeled and measured channels. Still, these evaluations are based on a limited number of transmission rates. The performance is reported for rates that are chosen by an exhaustive search to maximize the average throughput, across all positions, rather than optimizing the rate per packet transmission. Additionally, similarly to the hard detection approaches, they do not capture the impact of the higher layers of the protocol stack.

In the downlink, non-linear, theoretical precoding approaches exist which claim the MIMO channel capacity that is currently unexploited by linear precoders [7], [20]–[24]. These approaches are based on Dirty Paper Coding principles which can achieve the capacity of the Gaussian broadcast channel [25]. In this direction, the non-linear Tomlinson-Harashima precoding [26] can substantially improve on the throughput achievable by traditional linear precoding. Furthermore, improving on the Tomlinson-Harashima precoding, VP precoding [7] can further bridge the gap to the MIMO channel capacity limit. VP perturbs the transmitted constellation symbols so that the corresponding perturbation effect can be efficiently compensated at the receiver side. Again, most evaluations of vector perturbation precoding [7], [20] are limited to simulations employing mathematically-modeled channels. The massively-parallel vector perturbation precoder of [27], which promises practical non-linear precoding, is evaluated by over-the-air experiments, but also with off-line processing, inheriting the corresponding evaluation drawbacks.

As mentioned above, massively parallel non-linear (MPNL) processing [13], [27] enables the use of non-linear detection and precoding in practical applications. In general, the MIMO detection and precoding objective function can be expressed in terms of one (e.g., hard detection, vector-perturbation precoding) or several (e.g., soft demapping) integer least-squares problem of the type

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}^K} \|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2, \tag{1}$$

where $\mathbf{y}$ is a vector of observables, $\mathbf{A}$ is a matrix that depends on the MIMO transmission channel $\mathbf{H}$, and $\mathbf{x}$ is a potentially transmitted or perturbation vector, with dimensions $K \times 1$, where K is the number of concurrently-transmitted information streams. Each element of $\mathbf{x}$ is selected from the set $\mathcal{X}$. The noise variance is denoted as σ^2. Equation (1) can be optimally solved using sequential sphere decoders or encoders [5], [8], by converting the problem into a tree search with height K. However, these algorithms suffer from a high complexity and processing latency due, in part, to their sequential nature. In contrast, the MPNL detection and decoding techniques are able to search the tree in parallel using a high number of processing elements (PEs).

A block diagram that compares the MPNL processing of [17]–[19], [27] to conventional linear processing is shown in Fig. 1. At the pre-processing stage, linear processing performs a matrix inversion, followed by a precoding stage for the downlink, or a detection/demodulation stage for uplink communications. However, MPNL pre-processing involves a matrix triangularization (QR decomposition) that allows the detection/precoding optimization problem to be translated into an equivalent tree search. This is followed by the key part of the MPNL approach, which is identifying the most promising tree paths where the solution to (1) is most likely to be found. The identified most promising tree paths are then processed in parallel by the PEs. At the final stage, the tree path with the shortest Euclidean distance, encountered during the searches by all the PEs, is selected as the output of the algorithm. The MPNL algorithms exhibit a processing latency comparable to that of linear processing, and

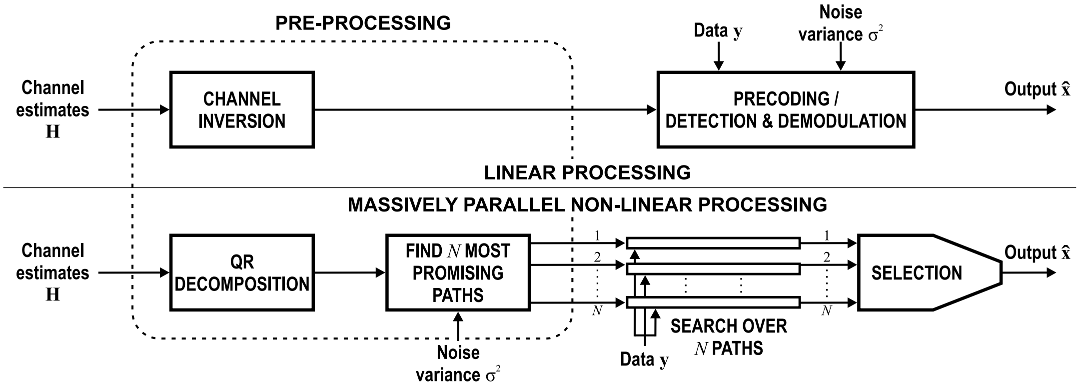

FIGURE 1. **Block diagram of conventional linear processing (top) and massively parallel non-linear processing (bottom).**

computational complexity of at least an order of magnitude less compared to other parallel approaches, e.g., [15].

As discussed, none of the above non-linear approaches has been evaluated in a standard-compliant framework to the best of our knowledge, neither a corresponding attempt has been reported that would identify missing algorithmic components and further challenges that need to be resolved.

III. CHALLENGES, ADOPTED APPROACHES & LESSONS LEARNED

Here we describe our experience by trying to incorporate and validate the performance of non-linear processing approaches in a 3GPP compliant environment. We outline some of the main challenges we have faced, as well as our adopted solutions together with their corresponding limitations, and the related lessons we have learned. As we will discuss in detail, such an attempt came with numerous challenges, ranging from finding (and extending) an appropriate software and hardware platform to perform our evaluations, to challenges related to missing components and practical aspects of the algorithms that, to the best of our knowledge, have not been highlighted/identified before.

A. SEEKING FOR THE APPROPRIATE SOFTWARE PLATFORM

There are several software platforms that aim at providing a 3rd Generation Partnership Project (3GPP)-compliant protocol stack, capable of running on general-purpose processors. They can be broadly classified as commercial and open-source. The commercial solutions include, among others, the LTE and NR Network Software Suit by Amarisoft [28], the National Instruments LTE Application Framework for LabVIEW Communications System Design Suite [29], and Intel's FlexRAN [30]. The most complete is perhaps the solution provided by Amarisoft, which in contrast to other options, provides a full protocol stack implementation on the BS side and user equipment (UE) side. Although it supports many features and transmission modes, the Amarisoft solution cannot be openly used for physical layer research due to its closed-source nature. In contrast to commercial platforms, open-source solutions which include srsLTE [31], openLTE [32] and OpenAirInterface (OAI) [33] are freely available to the public. Among those, it seems that the most advanced platform is OAI, the open-source solution with the largest developer community actively working towards adding new features into the existing code base (e.g., support for 5G New Radio (NR)).

1) OUR ADOPTED SOLUTION

For our evaluations, we extended our recently proposed SoftWare Open Radio Design (SWORD) platform [34], which overcomes the missing support for large/massive MIMO setups, as well as the inherent inability of existing approaches to investigate non-linear processing without prohibitive software and hardware optimization necessary. To support downlink and uplink MU-MIMO transmission schemes, which were in our main interest for testing non-linear processing approaches, SWORD significantly extends the OAI code base [35] and introduces an entirely new mode of operation, which we call *pseudo-real-time (RT)*. As described in [34] in detail, this new mode combines the properties and builds upon two existing modes of operations already supported by OAI, which permit RT over-the-air (OTA) transmission and emulation of an entire radio access network without the use of software-defined radio (SDR) modules. Compared to the generally adopted method of *offline processing* in which a received signal is stored in a raw format on the receiver side and then processed, the *pseudo-RT* can be effectively used to evaluate the impact of advanced physical layer approaches on the overall system performance.

In contrast to *offline processing*, the pseudo-RT mode makes use of a pause period between each transmission to facilitate signal processing on both sides. As a result, it preserves the dependence between consecutive events, allowing for a more realistic setting in which the full-protocol stack is executed. To enable pseudo-RT processing, several changes and extensions to the OAI code were required to ensure proper synchronization of subframe processing, and enable appropriate handling of multiple UEs and multiple BS radio chains. Details can be found in [34].

2) FURTHER CHALLENGES AND LESSONS LEARNED

The effective investigation of advanced physical layer approaches requires supporting large/massive MIMO setups and pseudo-RT mode of operation, which are not yet available in existing platforms. While the SWORD platform provides these features, the current implementation of the pseudo-RT mode of operation mandates that processing for all UEs and BS is performed by a single process, executed on a single workstation [35]. Although beneficial during the development and debugging of new features, we found that this architecture does not scale well for a higher number of UEs and BSs due to limited computational power. In the next iteration of our software platform, we intend to adopt a new architecture that permits UE processing to be executed in a separate process (and a separate machine) to allow for better flexibility in the allocation of resources for processing. Note that this is also a key enabler in providing more flexibility in interconnecting SDR modules, as the current software architecture mandates that all radio modules are connected to the same workstation. As a result, in order to conduct measurements under various channel conditions, long, low-attenuation cables are required which interconnect UE antennas with SDR modules on the UE side. We noticed that these cables, due to their limited length, can significantly restrict the set of scenarios that can be investigated. To address this, we foresee to rework the *subframe processing synchronization mechanism* that constitutes the core of the pseudo-RT mode, thus eliminating the need for such cables. Given the new architecture, the reworked mechanism would allow for the flexibility in interconnecting SDR modules used on the UE side with any workstation dedicated to UE processing.

B. SEEKING FOR THE APPROPRIATE HARDWARE PLATFORM

There are several hardware platforms capable of supporting MIMO setups that aim to be open to everyone for experimentation and can be potentially used for the evaluation of new physical layer solutions. One example of such a hardware platform is COSMOS [36], a city-scale testbed deployed in New York City aimed at providing means for real-world experimentation on next-generation wireless technologies and applications. Another example is POWDER [37], another city-scale testbed run by the University of Utah. Contrary to COSMOS, POWDER provides hardware components specifically dedicated for large/massive MIMO experimentation, with up to 64 antennas per site/sector. Interestingly, both COSMOS and POWDER allow for the use of various open-source software platforms such as OAI, srsLTE, or openLTE. Yet another example of a hardware platform is LuMaMi [38] of Lund University. LuMaMi is much smaller in scale compared to COSMOS and POWDER. However, in contrast to the other two testbeds, it is specifically dedicated for conducting large/massive MIMO-related research and supports up to 128 radio chains.

Although all three setups have a broad range of capabilities, they come with certain limitations that make them non-appropriate for meeting our objectives, at least at their current design stage. For instance, in the case of COSMOS, the capabilities of SDR modules used in the deployed nodes are limited to a maximum of four radio chains per site/sector. This can be potentially circumvented by considering a distributed MIMO setup; however, due to additional challenges, this type of setup currently is not our main focus. The situation is slightly different in the case of POWDER. In this case, the limitation resides on the UE side, as only two SDR modules in POWDER's massive MIMO setup seem to be currently dedicated to run as UEs. This means that the non-linear processing gains would be difficult to demonstrate since they target supporting numbers of users that are similar to the numbers of BS antennas [5], [13], [27]. The main limitation of LuMaMi is that, contrary to the other testbeds, it heavily relies on proprietary hardware and software solutions from National Instruments [39]. This means that any experiments would have to be based on National Instruments' software. Note that LuMaMi was not designed to be used for evaluation of physical layer approaches as part of a full 3GPP compliant protocol stack. It is not clear if LuMaMi would support the National Instruments software extensions, which could potentially bridge this gap. In addition to the above, in all three cases, the lack of physical access to nodes dedicated for experimentation on the UE side restricts investigation to a limited set of scenarios.

1) OUR ADOPTED SOLUTION

The identified limitations of the existing publicly available hardware platforms convinced us to invest in the development of our own hardware platform that can be easily moved around and permits investigation of scenarios with a different number of BS antennas, and different number of UEs. The main hardware component of our SWORD hardware platform is a multi-core x86_64 workstation with a large number of PCIe lanes used to interface with SDR modules of our choice. The SDR module selected is the Universal Software Radio Peripheral (USRP) X series with UBX daughterboard. USRP X series hosts two independent radio chains and is one of the SDR modules recommended by Ettus for applications that require phase alignment [40]. To synchronize and maintain phase alignment across multiple SDR modules, we use the Ettus Research Octoclock-G CDA-2990 [41], a highly accurate external clock reference and pulse distribution module. Circulators are used to connect the TX and RX paths of each radio chain to an antenna port. A more detailed description with a rationale behind using specific building blocks can be found in [34].

2) FURTHER CHALLENGES AND LESSONS LEARNED

In order to investigate and demonstrate the benefits of non-linear processing, a movable hardware platform that can run multiple UEs and support large/massive MIMO setup is needed. While our SWORD hardware solution meets these requirements, maintaining phase alignment through reference clock sharing across multiple USRPs proved difficult. It required frequent execution of time-division duplexing (TDD) reciprocity calibration to compensate for any drifts. We observed that such drifts had a significant negative impact on system performance, particularly when the number of UEs in a setup approached the number of BS antennas. To improve this, we plan to achieve phase alignment through the local oscillator (LO) sharing, rather than reference clock sharing. As highlighted by Ettus in [42], LO sharing can significantly reduce short-term and long-term phase drift. Note that the USRP N32X series would be required for this purpose.

Furthermore, our existing hardware setup is currently based on the use of circulators which, due to the limited output power of USRP X series, significantly limits the range of scenarios that can be investigated. To overcome this, we intend to replace circulators with external power amplifiers in the next iteration of our platform. This current version of the SWORD platform supports up to 8 BS antennas and 4 UEs at present. Therefore, our current evaluations are limited by these numbers. Nevertheless, future versions of the testbed will include support for more BS antennas and UEs.

C. REMAINING SYSTEM CHALLENGES AND TWEAKS AROUND THEM

While trying to evaluate the non-linear approaches, we came across several practical issues that needed to be resolved or bypassed. These are:

1) ENABLING NON-LINEAR PROCESSING

As many non-linear decoding approaches are designed for orthogonal frequency-division multiplexing (OFDM) transmission, in order to test non-linear processing in the uplink, we modified the processing of the physical uplink shared channel (PUSCH) in our LTE-based platform by making transform precoding optional (see LTE PUSCH processing in [43]). To inform UEs about the use of transform precoding, we extended radio resource control (RRC) signaling in line with the 5G-NR specification (note that transform precoding in 5G-NR is optional and can be dynamically enabled or disabled using RRC signaling). We faced similar issues with non-linear precoding approaches in the downlink, which adopt vector perturbation and thus require a modulo operation to be applied at the transmitter side [7], [23]. To revert this operation on the receiver side, we modified UE processing of the physical downlink shared channel (PDSCH) accordingly. Furthermore, we extended RRC signaling to inform UEs about the use of vector perturbation. Note that to enable more dynamic switching existing set of downlink control information (DCI) in LTE and 5G-NR used for scheduling transmission opportunities could be extended to include information on whether the incoming transmission underwent vector perturbation.

2) TRANSMISSION RATE SELECTION

adaptive modulation and coding (AMC) is an important aspect of 3GPP systems that enables the efficient utilization of the available spectrum resources. However, AMC for non-linear is still an open problem. As discussed before, to evaluate the performance of non-linear algorithms, the research community usually conducts an exhaustive search by running experiments for a small number of rates (i.e., QAM constellations and channel coding rates) and shows the average performance per rate. Although useful, the number of rates is in general very limited. In order to better evaluate the performance of non-linear approaches, and in the absence of AMC, we have applied an ‘‘adaptive’’ rate adaptation algorithm that selects the employed modulation and coding scheme (MCS) based on the reported ACK/NACK information. More specifically, the employed algorithm tracks the erroneous and correctly received transmissions in both uplink and downlink. Based on this information, and given a maximum and minimum MCS, the algorithm attempts to adjust the MCS value after a predefined number of consecutive ACKs, or NACKs is received (resetting an ACK counter, when a NACK is received, and NACK counter, when an ACK is received). To prevent excessive MCS switching, the proposed MCS selection approach implements a simple ‘‘cool-off period mechanism’’ that prevents any MCS changes for a specific number of frames after the last MCS change. Still, while our adopted approach can provide an improved throughput evaluation compared to traditionally used approaches that use a limited number of rates (and they depict the rate that maximizes the average performance across channels), it is far from being realistic. It can only be used to reliably evaluate the performance in a static environment where the channel remains relatively unchanged over multiple radio frames.

3) DOWNLINK CHANNEL ESTIMATION

As indicated in [7], non-linear precoding approaches, such as vector perturbation, results in an extended symbol constellation. As we have here identified, this makes non-linear approaches more sensitive to the channel estimation errors than linear approaches. In order to evaluate the performance of non-linear precoding techniques adopting vector-perturbation, we compensated for the impact of the channel estimation errors by boosting the transmit power of UE-specific demodulation reference signal (DMRS) used in LTE and 5G-NR for channel estimation. We note that DMRS is only used for detection at the UE side and not for beamforming, which is based on the sounding reference signal (SRS). We also note that LTE and 5G-NR already support power boosting for UE-specific DMRS; however, only a predefined set of power boosting values can be used for this purpose. In LTE, a 3dB power boosting is used when more than two layers are transmitted. In the case of 5G-NR, 3dB or 4.77dB power boosting can be applied, depending on the DMRS configuration used [44]. To inform UEs about the non-standard compliant values, we also extended RRC

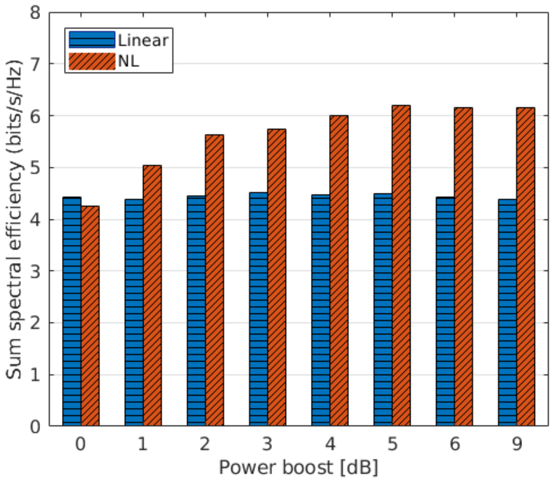

FIGURE 2. **The impact of UE specific DMRS power boosting on system performance.**

signaling. Figure 2 presents data for a single indoor measurement location and depicts the impact of UE-specific DMRS power boosting on downlink sum spectral efficiency for a 4 × 4 MU-MIMO configuration. As seen, power boosting of UE-specific DMRS can lead to significant performance improvements when non-linear precoding is used.

4) CHANNEL STATE INFORMATION (CSI) ESTIMATION

Another issue that we came across, independent of processing type (i.e., linear or non-linear), is related to the estimation of channel state information (CSI). For CSI estimation 5G-NR and LTE employ a special signal transmitted in uplink termed SRS. The existing implementation of the SRS transmit power control (TPC) mechanism may result in a partial loss of CSI, which in turn can limit the performance of a precoder. In particular, the signal amplitude difference between multiple UEs in a cell is lost. The fundamental objective of the TPC mechanism is to assure that signals transmitted by multiple UEs arrive at BS with approximately the same strength, which in turn results in the loss of amplitude information. To circumvent this, as a first attempt solution, we set the SRS transmit power to a constant value. To achieve it and at the same time retain the benefits of TPC for uplink, we separated the TPC for SRS and other uplink signals so that they are not conducted jointly. Note that in 5G-NR, a separate TPC for SRS and other uplink signals is already part of the standard. Separate TPC for a new variant of SRS (termed ‘‘additional’’ SRS) has also been recently introduced in LTE release 16.

5) FURTHER CHALLENGES AND LESSONS LEARNED

Adapting non-linear processing in a real system requires a number of changes in 3GPP standards, which include changes in PUSCH and PDSCH processing. Additional changes are also required in the signaling procedures. These primarily include extensions of RRC signaling, which is used to inform UEs about the non-linear processing settings (e.g., additional power boosting for UE-specific DMRS), but can also affect DCIs, e.g., to allow for a ‘‘per transmission’’ parameter selection. Additionally, AMC for non-linear systems is a critical missing component that, as we also discuss later in more detail, can determine the system performance. In this context, its absence may be one of the main reasons preventing the adoption of non-linear approaches to actual systems.

IV. EVALUATION RESULTS

This section presents results obtained by the OTA measurements that validate our design and provide some indicative performance evaluation of advanced non-linear (NL) processing against linear (i.e., ZF) approaches that serve as the baseline approach for linear processing. Without a loss of generality, we employed the soft, near-optimal, non-linear detection algorithm discussed in [17] and vector perturbation–based, non-linear precoder introduced in [27], since they are the most promising in terms of processing latency and complexity. The number of processing elements assumed is 40 and 32 for uplink and downlink, respectively, which have been observed to provide a good trade-off between error performance and computational complexity. The measurements were conducted using the developed hardware and software SWORD platform for an MU-MIMO setup with 4 and 8-antenna BS setup and 4 single-antenna UEs. While the examined MIMO dimensions are small, as we will discuss later in detail, they have been sufficient to verify the gains of non-linear processing. To demonstrate the level of potential improvements in a way that is transparent to the adopted physical layer numerology, we present our results in the form of relative gains.

It is significant to note that the aim of our evaluations is not to determine the gains of non-linear processing over linear. This has already been done extensively in the literature (please see Section II). Instead, our aim is to identify all changes that are required to adopt non-linear approaches in a 3GPP context (i.e., actual base-stations), to validate our proposed approach, and to showcase that the promised in the literature gains can still be achieved after performing all practical algorithmic tasks of an actual BS (e.g., channel estimation and calibration, antenna synchronization), and with all the dependencies of the full protocol stack involved.

6) MEASUREMENT SETUP

Several indicative locations were selected for measurements, including four indoor locations (for both uplink and downlink measurements) and three outdoor locations (for uplink measurements only).

Figure 3 shows pictures of the measurement setups for all four indoor locations. The selected indoor locations provided multiple reflective, obstructive, and scattering surfaces. Note that our platform does not currently integrate external power amplifiers. Therefore, due to the limited output power, the UEs were placed approximately 10m from the BS in the indoor locations, with a line-of-sight (LOS) or partially-obstructed LOS (on indoor measurement location 3)

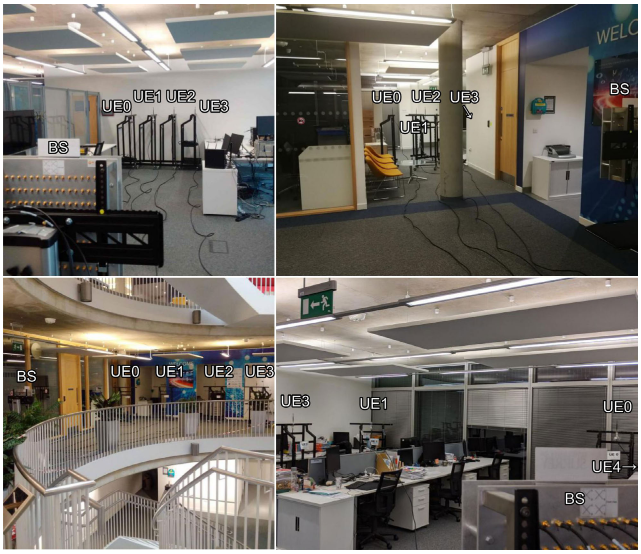

FIGURE 3. Indoor measurement locations. Clockwise from top left are indoor locations 1, 3, 4, and 2.

path between the BS and the UEs. The signal-to-noise ratios (SNRs) were in the range of 15dB or below.

In the case of outdoor measurements, three distinct setups have been examined. In the first measurement location, the UEs have been placed about 3m apart. In the second, they have been arranged about 1.5m apart, and in the third, about 4.5m apart. In all three cases, the UEs were placed approximately 17m from the base-station. Figure 4 shows the pictures of the setup for the three measurement locations.

The platform was set in TDD mode, with 5 MHz channel bandwidth and an operating frequency of 3.55 GHz. The LTE downlink/uplink slot configuration number 3 was used for the measurements, which includes 6 downlink slots, 3 uplink slots, and 1 special slot [43]. A subset of antenna array elements with the same polarisation and equivalent to a uniform linear array (ULA) was used. Furthermore, the scheduler for MU-MIMO was set to always schedule transmission to all UEs with the same number of resource blocks. CSI estimates at the transmitter were obtained using SRS transmitted in every frame, with a moving average filter applied to reduce the effect of thermal noise. To compensate for any thermal phase drift, each measurement instance was preceded by the TDD reciprocity calibration procedure.

7) INDOOR RESULTS

As shown in Figure 5 for our indicative indoor uplink evaluations and in Figure 6 for indoor downlink evaluations, the use of NL processing results in a consistent increase in overall system performance compared to linear processing. In the case of the uplink tests, the average gains of NL approach 57% and 9% for 4-antenna and 8-antenna setup, respectively. Furthermore, for the downlink tests, NL offered average relative gains of approx. 50% and 2% in the case of 4×4 and 8×4 MIMO configurations, respectively. While the downlink NL gains compared to linear processing are consistent, they are less prominent compared to the uplink. This is due to reasons like channel aging, as well as the imperfections of the SRS-based channel estimation and TDD calibration, which can be further improved.

FIGURE 4. Outdoor measurement locations. From top to bottom are outdoor locations 1, 2, and 3.

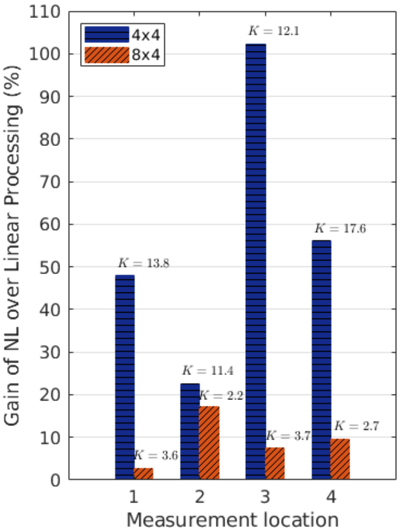

FIGURE 5. Relative gain of NL over linear processing in uplink for 4×4 and 8×4 MU-MIMO configuration with average channel condition number K for indoor measurement locations.

8) OUTDOOR RESULTS

The uplink outdoor scenario results are presented in Figure 7. Similarly to the indoor cases, an increase in system

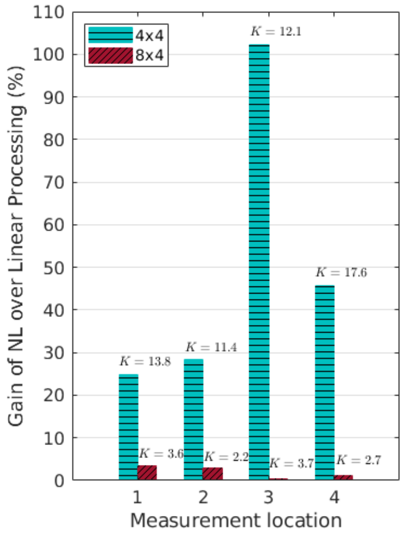

FIGURE 6. **Relative gain of NL over linear processing in downlink for 4 × 4 and 8 × 4 MU-MIMO configuration with average channel condition number *K* for indoor measurement locations.**

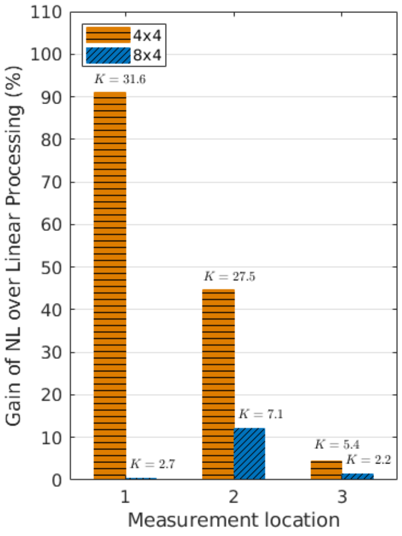

FIGURE 7. **Relative gain of NL over Linear processing in uplink for 4 × 4 and 8 × 4 MU-MIMO configuration with average channel condition number *K* for outdoor measurement locations.**

performance through the use of NL detection was obtained in all measured locations. Relative average gains of approx. 47% and 5% were achieved for the 4-antenna and 8-antenna cases, respectively. From the results above, it is interesting to notice the performance difference between the outdoor locations 1 and 2. In location 1, where the UEs are not very close to each other, the channel condition can be substantially improved by increasing the number of base-station antennas from 4 to 8. This results in an increased beam directionality that can diminish the gains of NL processing. Still, this is not the case for location 2, where UEs are closer together. NL processing can then provide significant gains, and much more BS antennas would be required to make linear approaches equally efficient to the non-linear ones.

The indoor and outdoor measurement results presented above show that NL offers a reduced gain in the 8-antenna cases. This is expected, since increasing the number of BS antennas while maintaining the same number of UEs allows simplifying the signal detection processing, at the cost of highly under-utilizing the MIMO channel [17]. Still, in contrast to what has been expected, the gains of the NL processing in the 8 × 4 MU-MIMO cases are non-negligible.

It is also worth noting that the average channel condition number K shown in all the figures is a good indicator of the gains that can be achieved when using NL approaches. This can be seen in particular for the outdoor measurements. Still, the average channel condition number does not fully describe the full MIMO environment; hence additional metrics are needed to fully predict the NL gains.

V. REMAINING CHALLENGES AND WAY FORWARD

Here we discuss some of the remaining challenges that need to be addressed in order to develop future BSs that can benefit from the non-linear processing approaches.

A. TRANSMISSION RATE ADAPTATION

As it has become evident, one of the essential missing components needed to make non-linear processing both practical and effective, is efficient rate adaptation. In this direction, two approaches can be potentially examined. The first one is to try to develop *non linear–specific* AMC methods, and the other is to adopt *rateless* (or *fountain*) channel coding.

The direction towards developing non linear–specific AMC methods is particularly challenging in the uplink. In this case, the per-user SNR would typically differ, and therefore, each user should use its own transmission rate. This issue could be partially handled by retaining the transmit power control mechanism of SRS signals (see discussion in Section III-C). Still, the maximum achievable transmission rate is a function of the MIMO channel, and the adopted detection method makes the AMC problem even more complicated. A promising direction towards non linear–specific AMC would be to consider the mathematical framework used for identifying the "most promising" vector solutions in the massively parallel methods of [13], [17], since the corresponding *metrics-of-promise* are related to the achievable error-rate probability. In the downlink, predicting the modulation order and the coding rate that maximizes the throughput could be easier, since typical non-linear precoding approaches result in the same SNR per user. Still, if a per-user *power-loading* approach is adopted (that by itself is an interesting research direction), the problem becomes similar to the uplink case. Then the duality between uplink and downlink transmissions could perhaps be explored.

Alternatively, rateless codes can be used, which would avoid choosing an MCS mode [45]–[47]. This is achieved by initially transmitting high-rate information, followed by parity information. This approach decreases the effective information rate until the transmitted information is correctly decoded. Still, such approaches would require revisiting the way ACK/NACK signaling is transmitted. It is noted that akin to the rate adoption problem (which is still open) and perhaps needs to be considered jointly with AMC, are the *user-selection* schemes that allocate users to MIMO transmissions (or MIMO antennas) in order to maximize system performance.

B. SCALING TO LARGE NUMBERS OF USERS

Channel estimation is an essential aspect of every MIMO system. To allow for effective channel estimation in 3GPP systems, each data stream is assigned with a DMRS, which is orthogonal with respect to DMRS allocated to other streams. As 3GPP systems have not been specifically designed for non-linear processing (which enables supporting very large numbers of users), the number of orthogonal DMRS allocations in 3GPP is limited to 8 in LTE [43] and 12 in 5G-NR [48]. While these limits seem reasonable when systems are based on linear processing, such limits may become a bottleneck in the case of systems with non-linear processing (in particular, when the number of concurrently supported UEs is larger than the number of BS antennas, as will be discussed in the following section). Note also that capability of SRS to support multiple users is highly dependent on the channel delay spread, and there is only a limited number of cyclic shifts that can be used in practice. Because of this, and given that the periodicity of SRS needs to reflect changes in the channel coherence time, the structure of SRS may not be sufficient to maintain the CSI reliability.

VI. CONCLUSION & FUTURE RESEARCH

We have verified, for the first time, in a 3GPP compliant framework that non-linear processing is a promising approach for increasing the achievable throughput and user connectivity of mobile systems. Still, further research remains to be done before such approaches are adopted by actual wireless systems and standards. Among them, two of the most significant open questions are how to perform rate adaption and how to redesign the corresponding wireless systems in order to be able to support a much larger number of users. Especially since, as has already been shown in the literature, the gains of non-linear processing increase when increasing the number of concurrently-served users. It is significant to notice that since, at the time of this research, no 5G-NR base system has been available, this work has been based on an LTE realization. Still, NL processing is inherently transparent to the generations of mobile systems. Consequently, the developed algorithmic blocks, the achievable throughput gains (relative to linear processing), the proposed modifications, and the identified missing blocks still apply to 5G systems and beyond.

Despite the verified gains, the most interesting capabilities that non-linear BS processing can offer, and perhaps revolutionize future wireless systems, have not yet been explored. In this direction, we can identify two promising research pathways: (a) non-linear processing for supporting numbers of transmitted information streams that are larger than the number of BS antennas and (b), practical, non-linear, and iterative BS processing for further bridging the gap between the theoretical capacity and the achievable throughput of systems with large connectivity.

A. TRANSMITTING MORE STREAMS THAN BASE-STATION ANTENNAS

In a ''fully connected'' wireless ecosystem, future communication systems will need to support a very large number of users. Traditional wireless system designs with linear processing are not capable of supporting more information streams than the number of BS antennas and, in practice, can efficiently support only a much smaller number of information streams. Non-linear processing approaches can negate this limitation, at least theoretically, and promise to support a much larger number of information streams than the number of BS antennas [10], [49], even without the need for specifically designated *Non-Orthogonal Multiple Access* (NOMA) techniques [50]–[52]. Still, as already discussed, for developing such systems, we will need to revisit the signaling procedures, as well as the way channel estimation is performed.

B. NON-LINEAR ITERATIVE BASE-STATION PROCESSING

Iterative systems that exchange ''soft'' information between a non-linear detector and a ''soft'' channel decoder promise substantial gains [53]–[55]. Still, such approaches are not scalable to a large number of information streams due to their exponential increase in complexity and latency requirements. For example, the approximate non-linear approach of [55] would require 10^{14} multiplications for a 12 × 12 MIMO system. On the other hand, currently proposed massively parallel, soft-output approaches that can substantially reduce the corresponding complexity and processing latency requirements [17], [49] are not applicable to the iterative case. This is because such approaches are heavily based on the geometrical properties of the transmitted signal constellation, which is destroyed by the existence of prior information (from previous iterations). Furthermore, existing iterative schemes cannot currently support a larger number of information streams than the BS antennas. The development of non-linear, massively parallel, iterative detection and decoding techniques, able to support more users than BS antennas, could give a significant connectivity boost and allow access to unexploited capacity resources.

REFERENCES

[1] 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,'' in *Proc. ACM SIGCOMM*, 2013, vol. 43, no. 4, pp. 399–410.

[2] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, ''Argos: Practical many-antenna base stations,'' in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2012, pp. 53–64.

[3] K. Tan, H. Liu, J. Fang, W. Wang, J. Zhang, M. Chen, and G. M. Voelker, ''SAM: Enabling practical spatial multiple access in wireless LAN,'' in *Proc. 15th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2009, pp. 49–60.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[5] K. Nikitopoulos, J. Zhou, B. Congdon, and K. Jamieson, "Geosphere: Consistently turning MIMO capacity into throughput," in *Proc. ACM SIGCOMM*, vol. 44, 2014, pp. 631–642.

[6] 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.

[7] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[8] 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.

[9] L. Barbero and J. Thompson, "Fixing the complexity of the sphere decoder for MIMO detection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2131–2142, Jun. 2008.

[10] T. Cui and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," in *Proc. IEEE 60th Veh. Technol. Conf. (VTC-Fall)*, vol. 5, Sep. 2004, pp. 3689–3693.

[11] M. Shabany, K. Su, and P. Glenn Gulak, "A pipelined scalable high-throughput implementation of a near-ML K-best complex lattice decoder," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 3173–3176.

[12] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

[13] 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. Design Implement. (NSDI)*, 2017, pp. 197–211.

[14] B. Mennenga and G. Fettweis, "Search sequence determination for tree search based detection algorithms," in *Proc. IEEE Sarnoff Symp.*, Mar. 2009, pp. 1–6.

[15] L. G. Barbero, T. Ratnarajah, and C. Cowan, "A low-complexity soft-MIMO detector based on the fixed-complexity sphere decoder," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 2669–2672.

[16] J. Jalden and B. Ottersten, "Parallel implementation of a soft output sphere decoder," in *Proc. Conf. Rec. 39th Asilomar Conf. Signals, Syst. Comput.*, 2005, pp. 581–585.

[17] K. Nikitopoulos, G. Georgis, C. Jayawardena, D. Chatzipanagiotis, and R. Tafazolli, "Massively parallel tree search for high-dimensional sphere decoders," *IEEE Trans. Parallel Distrib. Syst.*, vol. 30, no. 10, pp. 2309–2325, Oct. 2019.

[18] K. Nikitopoulos, "Massively parallel and flexible processing for MIMO systems," in *Wiley 5G Ref: The Essential 5G Reference Online*. Hoboken, NJ, USA: Wiley, 2019, pp. 1–21.

[19] K. Nikitopoulos and R. Tafazolli, "Parallel processing of sphere decoders and other vector finding approaches using tree search," U.S. Patent 10 374 841, Aug. 6, 2019.

[20] M. Mohaisen and K. Chang, "Fixed-complexity sphere encoder for multi-user MIMO systems," *J. Commun. Netw.*, vol. 13, no. 1, pp. 63–69, Feb. 2011.

[21] D. A. Schmidt, M. Joham, and W. Utschick, "Minimum mean square error vector precoding," *Eur. Trans. Telecommun.*, vol. 19, no. 3, pp. 219–231, 2008.

[22] C. Yuen and B. M. Hochwald, "WLCp1-13: How to gain 1.5 dB in vector precoding," in *Proc. IEEE Globecom*, Nov. 2006, pp. 1–5.

[23] H.-S. Han, S.-H. Park, and I. Lee, "Improved vector perturbation with modulo loss reduction for multiuser downlink systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2009, pp. 1–5.

[24] W. Yao, S. Chen, and L. Hanzo, "Generalized MBER-based vector precoding design for multiuser transmission," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 739–745, Feb. 2011.

[25] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[26] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 774–780, Aug. 1972.

[27] C. Husmann and K. Nikitopoulos, "ViPer MIMO: Increasing large MIMO efficiency via practical vector-perturbation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018.

[28] Amarisoft. (2019). *LTE and NR Network Software Suit*. [Online]. Available: https://www.amarisoft.com/products/custom-projects/

[29] National Instruments. (2019). *LabVIEW Communications LTE Application Framework*. [Online]. Available: http://www.ni.com/en-gb/shop/select/labview-communications-lte-application-framework

[30] Intel. (2018). *FlexRAN SW Wirelss Access Solutions for 4G and 5G VNFS*. [Online]. Available: https://software.intel.com/en-us/videos/an-overview-of-flexran-sw-wireless-access-solutions

[31] I. Gomez-Miguelez, A. Garcia-Saavedra, P. D. Sutton, P. Serrano, C. Cano, and D. J. Leith, "srsLTE: An open-source platform for LTE evolution and experimentation," in *Proc. 10th ACM Int. Workshop Wireless Netw. Testbeds, Exp. Eval., Characterization*, 2016, pp. 25–32.

[32] openLTE. (2019). *OpenLTE; An Open Source Implementation of the 3GPP LTE Specifications*. [Online]. Available: https://sourceforge.net/projects/openlte/

[33] F. Kaltenberger, A. P. Silva, A. Gosain, L. Wang, and T.-T. Nguyen, "OpenAirInterface: Democratizing innovation in the 5G era," *Comput. Netw.*, vol. 176, Jul. 2020, Art. no. 107284.

[34] 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.

[35] G. Georgis, A. Thanos, M. Filo, and K. Nikitopoulos, "A DSP acceleration framework for software-defined radios on X86 64," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 1648–1652.

[36] COSMOS. (2019). *Cloud Enhanced Open Software Defined Mobile Wireless Testbed for City-Scale Deployment*. [Online]. Available: http://www.cosmos-lab.org/

[37] POWDER. (2019). *Platform for Open Wireless Data-driven Experimental Research*. [Online]. Available: https://powderwireless.net/

[38] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, "The world's first real-time testbed for massive MIMO: Design, implementation, and validation," *IEEE Access*, vol. 5, pp. 9073–9088, 2017.

[39] *5G Massive MIMO Testbed: From Theory to Reality*, National Instruments, Austin, TX, USA, 2014.

[40] Ettus. (2019). *UBX–Knowledge Base*. [Online]. Available: https://kb.ettus.com/UBX

[41] (Jan. 2019). *Octoclock Specification Sheet*. [Online]. Available: https://www.ettus.com/wp-content/uploads/2019/01/Octoclock_Spec_Sheet.pdf

[42] D. Baker and B. Avenell. (2019). *RF System Synchronization–LO's*. [Online]. Available: https://www.gnuradio.org/grcon/grcon19/presentations/phase_synchronizat ion_techniques/Dan%20Baker%20-%20Phase%20Synchronization%20Techniques.pdf

[43] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, 3rd Generation Partnership Project (3GPP), Technical Specification document (TS) 36.211, Dec. 2019, version 16.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2425

[44] *NR; Physical Layer Procedures for Data*, 3rd Generation Partnership Project (3GPP), Technical Specification document (TS) 38.214, Dec. 2019, version 16.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3216

[45] O. Etesami and A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2033–2051, May 2006.

[46] M. Luby, A. Shokrollahi, M. Watson, and T. Stockhammer, *Raptor Forward Error Correction Scheme for Object Delivery*, document RFC 5053, Oct. 2007.

[47] J. Perry, P. A. Iannucci, K. E. Fleming, H. Balakrishnan, and D. Shah, "Spinal codes," in *Proc. ACM SIGCOMM Conf. Appl., Technol., Archit., Protocols Comput. Commun.*, 2012, pp. 49–60.

[48] *NR; NR and NG-RAN Overall Description; Stage 2*, 3rd Generation Partnership Project (3GPP), Technical Specification document (TS) 38.300, Dec. 2019, version 16.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191

[49] 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.

[50] 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.

[51] R. Hoshyar, R. Razavi, and M. Al-Imari, "LDS-OFDM an efficient multiple access technique," in *Proc. IEEE 71st Veh. Technol. Conf.*, 2010, pp. 1–5.

[52] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 332–336.

[53] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[54] C. Studer and H. Bolcskei, "Soft–input soft–output single tree-search sphere decoding," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4827–4842, Oct. 2010.

[55] K. Nikitopoulos and G. Ascheid, "Approximate MIMO iterative processing with adjustable complexity requirements," *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 639–650, Feb. 2012.

KONSTANTINOS NIKITOPOULOS (Senior Member, IEEE) is currently a Reader in signal processing for communication systems with the Electrical and Electronic Engineering Department, Institute for Communication Systems, University of Surrey, Guildford, U.K. He is a member of the 5G and 6G Innovation Centre, where he leads the "Theory and Practice of Advanced Concepts in Wireless Communications" work area, and is the Director of the Wireless Systems Laboratory of the Institute for Communication Systems. He has held research positions with RWTH Aachen University, Germany, the University of California at Irvine, USA, and University College London, U.K. He was a recipient of the prestigious First Grant of the U.K.'s Engineering and Physical Sciences Research Council and the Principal Investigator of several research projects, including the AutoAir I and II projects, of the U.K.'s 5G Testbeds and Trials Programme.

MARCIN FILO received the Ph.D. degree from the University of Surrey, Guildford, U.K. He is currently a Research Fellow with the University of Surrey. Prior to joining the University of Surrey, he worked as a Researcher at WRC EIT+, Wroclaw, Poland, and Create-Net, Trento, Italy, where he was involved in a number of FP7 ICT projects. His main research interests include ultra dense networks, interference management for cellular networks, coexistence of heterogeneous wireless systems, and dynamic spectrum access.

CHATHURA JAYAWARDENA (Member, IEEE) received the Ph.D. degree from the University of Surrey, Guildford, U.K. He is currently a Research Fellow with the University of Surrey. His research interest includes signal processing for communications, with an emphasis on detection methods for non-orthogonal transmission schemes.

J. C. DE LUNA DUCOING (Member, IEEE) received the B.Sc. degree in electronic engineering from the Instituto Tecnólogico de Nuevo Laredo, Mexico, in 1995, and the M.Sc. and Ph.D. degrees in wireless communications from the University of Surrey, U.K., in 2013 and 2018, respectively. He has considerable experience in the telecommunications industry, particularly in international voice and data communication networks. In 2018, he joined the Institute for Communication Systems, as a Research Fellow. His research interests include, but not limited to, algorithm design and signal processing for large-scale MIMO wireless communications.

RAHIM TAFAZOLLI (Senior Member, IEEE) has been a Professor of mobile and satellite communications, since April 2000, the Director of ICS, since January 2010, and the Founder and the Director of the 5G Innovation Centre, University of Surrey, U.K. In 2011, he was appointed as a Fellow of the Wireless World Research Forum (WWRF) in recognition of his personal contributions to the wireless world and heading one of Europe leading research groups. He was an Advisor to the Mayor of London in regard to the London Infrastructure Investment 2050 Plan, from May 2014 to June 2014. He has over 25 years of experience in digital communications research and teaching. He has authored and coauthored more than 500 research publications. He is a co-inventor of more than 30 granted patents, all in the field of digital communications. He is regularly invited by many governments for advice on 5G technologies.

• • •