Two-Step Random Access in 5G New Radio: Channel Structure Design and Performance

Elena Peralta*‡, Toni Levanen*‡, Frank Frederiksen†, Mikko Valkama‡

*Nokia Bell Labs, Finland, †Nokia Bell Labs, Denmark, ‡Department of Electrical Engineering, Tampere University, Finland,

Email: elena.peralta_calvo@nokia.com

Abstract—A common design of the random access procedure on the physical random access channel (PRACH) is required for the diverse usage scenarios in the fifth generation new radio (5G NR) mobile networks. Based on the latest 3GPP specifications and evaluation assumptions agreed for Release 16, the 2 step-RACH (2SR) enhancement, composed of the denoted MsgA and MsgB, not only reduces the latency but also the control-signalling overhead due to the reduced number of messages transmitted. The channel structure of MsgA comprises RACH preamble and data in the physical uplink shared channel (PUSCH) while MsgB combines the random access response and the contention resolution. This procedure should operate in local area (LA), medium range (MR) and wide area (WA) cells despite the lack of time alignment (TA) in the PUSCH part of MsgA. The demodulation performance degradation observed without time offset compensation at the base station (gNB), specially for MR or WA cells, highlight that practical gNB implementations relying in MAC control element-based TA command for PUSCH time alignment are not conceivable for 2SR. Furthermore, in the case that all preambles from multiple users (UEs) trying to perform the initial access are mapped to the same PUSCH physical resources, the associated data parts overlap and may result in unsuccessful decoding. There is therefore a trade-off between the collision probability of the PUSCH part of MsgA and the resource overhead for 2SR. This paper addresses the channel structure design of this procedure for the preamble and data parts of MsgA together with the receiver processing framework. The performance results suggest that using lower payload sizes provide higher resource utilization and allow more UEs to be multiplexed within the same PUSCH occasion. In addition, using different DMRS ports for UEs sharing same physical resources decrease the probability of failure in the decoding of the data part of MsgA while reduces the resource overhead for 2SR.

Keywords—5G New Radio (NR), random access procedures, physical layer design, advance receivers, detection performance

I. INTRODUCTION

The random access (RA) procedure on the physical random access channel (PRACH) in the fifth generation new radio (5G NR) mobile networks should meet the low latency requirements of the diverse usage scenarios and applications [1]. In particular, the RACH procedure can be designed to support multiple use cases, e.g, enhanced mobile broadband, ultra-reliable low latency communications and massive machine type communications, following a common design for both licensed and unlicensed spectrum. In that aspect, the 2-step RACH (2SR) enhancement has been discussed in 3GPP standards motivated by the reduced latency. There is a work item in 5G NR standardization, see [2], aiming to specify the 2SR as a general MAC procedure assuming mainly contention based RACH procedures, where contention free RA in 2SR is only supported for handover [3]. As depicted in [4] and illustrated in Fig. 1, the messages in time order in the 4-step RACH (4SR) procedure are named as Msg1, Msg2, Msg3, Msg4. While in the 2SR, the messages are named as MsgA and MsgB.

More specifically, the channel structure of MsgA comprises preamble (Msg1) and data part (Msg3) in the physical uplink shared channel (PUSCH) and, MsgB combines the random access response (Msg2) and the contention resolution (Msg4). Consequently, there is only one round-trip cycle between the user equipment (UE) and the base station (gNB) to complete the 2SR procedure instead of the two round-trip cycles required in 4SR [1]. This enhanced procedure not only reduces the latency but also the control-signalling overhead. Besides, the decreased number of messages transmitted reduces the number of listen before talk attempts in NR unlicensed spectrum. Although this reduction may imply increased overhead and potential increased complexity in the receiver processing. In this context, the fall-back from 2SR to 4SR will be supported [4], where the fallback after MsgA transmission is feasible only if the detection of the UE without decoding of the data is possible.

The channel structure of MsgA and related physical layer design are currently under discussion in 3GPP NR Release 16. The study item [2] aim is to have a common design for the several use cases of 5G NR and to operate in any cell size despite the lack of time alignment (TA). Based on 3GPP 5G NR specifications in Release 15, the gNB estimates the initial TA from the preamble sent by the UE and relies in the MAC control element (CE) TA command for the time synchronization of the data part. However, without TA adjustment in the 2SR procedure, the time offset (TO) can be twice the propagation delay for the PUSCH part of MsgA compared to connected UEs. On one hand, the demodulation of the data part may not be impacted considering local area (LA) gNBs, with TOs values within the cyclic prefix (CP). Although in this case the percentage of CP available for the several number of fading taps in a tapped-delay-line (TDL) fading channel model [3] may not be enough. On the other hand, without implementing PUSCH time adjustment in the receiver, the demodulation performance may be degraded for larger inter-site distances (ISD) in medium range (MR) and wide area (WA) cells. In this context, 3GPP Release 15 PUSCH performance tests are not taking into account timing offset errors and, therefore, practical gNB implementations relying in MAC CE-based TA command for PUSCH time alignment are not conceivable for 2SR procedure [5]. Under these premises, the demodulation performance requirements for the 2SR procedure in 3GPP NR Release 16 should consider a time offset estimation (TOE) based on the demodulation reference signal (DM-RS) or, joint preamble detection and data decoding at the gNB. In this paper, the demodulation performance for different cell sizes are analyzed, considering LA, MR and WA gNBs, respectively.

The intra-cell interference from multiple UEs trying to perform the initial access has a significant impact on the final uplink (UL) performance. Moreover, the interference of the data part of MsgA will be even stronger under the assumptions of

equal transmitted power from multiple UEs in the serving cell. A typical interference-aware receiver is the linear minimum-mean-square-error (MMSE) receiver, where the covariance of the interference is considered. The interference-aware MMSE-IRC (MMSE-interference rejection combining) receiver can suppress inter-cell as well as intra-cell interference and it is expected to provide better performance than a common MMSE receiver [6]. Based on 3GPP agreements for the 2SR study item [7], MMSE-IRC receiver shall be considered as baseline to suppress the interference. Other advanced receivers with interference cancellation capabilities are not precluded and left to gNB implementation. The MMSE-IRC receiver requires channel and covariance matrix estimation [8]. Therefore, this paper investigates the gain in performance considering two estimation schemes based on the transmitted demodulation reference signals. At first, it is assumed that the knowledge of the effective channel of other interfering links is unknown at the receiver. In a second approach, the interference covariance matrix is estimated based on the effective channels of interfering links in addition to the desired signal.

The 2SR procedure should be applied for different radio resource control (RRC) states, e.g inactive, connected or idle. In this context, this paper presents a comprehensive overview of the 2SR procedure, focusing on describing and evaluating a flexible channel structure design for MsgA. On one hand, the PRACH occasion (RO) for the transmission of the preamble and, on the other hand, the PUSCH ocassion (PO) for the transmission of the data part, which consist of a set of 2SR PUSCH resource units (PRU). The link budget is analyzed based on the SNR requirements to achieve a target miss-detection probability given a false alarm rate for one long and one short preamble format. In addition, several payload sizes are evaluated for different number of physical resource blocks (PRBs) and transport block sizes (TBS) for single and multiple UEs sharing the same physical resources within the PO. Finally, the maximum coupling losses (MCL) of preamble and data parts of MsgA are evaluated and compared. Larger MCL differences between MsgA PRACH and PUSCH parts are observed as increasing the payload size. In addition to higher resource overhead requirements and diminished capacity of the number of UEs that can be multiplexed within the same PUSCH occasion. Therefore, the 2SR configuration may be different per cell size or between RRC states. Besides, the results indicate that using a MMSE-IRC receiver where the interference covariance matrix is estimated based on the effective channels of multiple interfering links provides significant gain in performance. In particular, we compare the cell coverage in a multiple-to-one mapping scenario when using different DMRS ports or DMRS sequences for several UEs sharing the same physical resources. It can be observed a performance gain for the benefit of using different DMRS ports.

The rest of the paper is organized as follows: Section II describes the two types of random access procedures: 2-step and 4-step RACH. In Section III, the mapping between preamble and data part of MsgA is described. Then, in Section IV, the channel structure design principles of MsgA are analyzed. In section V, the DMRS based time offset estimation algorithm, an enhanced MMSE-IRC receiver and related performance metrics are described. Then, Section VI presents the demodulation performance for the preamble and data part of MsgA. Finally, in Section VII, the conclusions of this study are drawn.
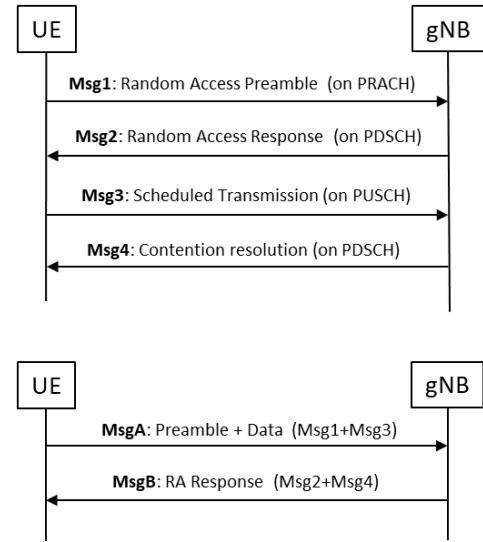

Fig. 1: Random access procedures operation in four-step RACH (top) and two-step RACH (bottom)

II. RANDOM ACCESS PROCEDURES

There are two types of RA procedures supported for NR to perform initial access and achieve UL synchronization between the UE and the gNB [9], contention-based and contention free random accesses. The 4SR procedure follows the steps illustrated on the top part of Fig. 1 for the contention-based random access. In this case, one PRACH preamble will be randomly selected by the UE among available preambles and transmitted to the gNB on the shared PRACH. Each preamble transmission is associated with the random access radio network temporary identifier (RA-RNTI). Upon reception of Msg1, the gNB will calculate the TA information and send random access response (RAR) messages including temporary cell-RNTI (C-RNTI) and initial UL grant resources to one or more UEs trying to perform initial access. The UE will then transmit the RRC connection request on PUSCH using the initial UL grant resources. Finally, the contention resolution is performed in case of successful decoding of the data transmitted from a single UE at the gNB side and a permanent C-RNTI is sent. Otherwise, nothing will be sent back from the gNB due to collision between more than one UE selecting same PRACH preamble, RA-RNTI and C-RNTI. Failure in the RACH procedure due to collision will be solved after a back-off period when the UE retries a new RA procedure.

A 2SR procedure may be applied for the case that an UE is in RRC-connected active mode, in handover or when is transitioning from RRC-connected inactive mode to RRC-connected active mode [2]. The principle behind the 2SR procedure is to combine the PRACH preamble corresponding to Msg1 and data information corresponding to Msg3 in the 4SR into one message, for instance, MsgA. Then, the gNB will respond upon successful reception of MsgA with MsgB, which combines the RAR corresponding to Msg2 and the contention resolution corresponding to Msg4 in the 4SR. The steps followed by the 2SR procedure are illustrated in the bottom part of Fig. 1. Being MsgA the first message transmitted to the gNB, this procedure does not support TA alignment and proper methodologies should be defined to reserve PUSCH resources of MsgA. On the other hand, due to

the diminished transmission of messages, the 2SR procedure can reduce latency and network overhead as increasing the number of POs associated with each RO. The UE considers the 2SR procedure as successfully completed upon reception of MsgB including a success RAR, which will be sent after successful decoding of the PUSCH part of MsgA. Otherwise, in case of unsuccessful detection of the preamble or failure in the decoding of the data, the UE will re-transmit MsgA or fall-back to 4SR procedure. Therefore, MsgB can contain a back-off indication for next RACH transmission or a fallback RAR, to fall back to 4SR with the transmission of Msg3.

III. Mapping between preamble and data of MsgA

Preamble and data parts of MsgA are expected to be time division multiplexed as described in [2]. Based on 3GPP discussions, there are several possible configurations for the PO of MsgA, e.g, POs can be separately configured from ROs or the relative time and frequency location of the PO can be configured with respect to the associated RO. The preambles of the 2SR procedure are separated from the 4SR ones and transmitted in the corresponding 2SR ROs, which can be the same or different RO as for the 4SR procedure. In addition, the PUSCH part of MsgA is defined by the time and frequency resources of the 2SR PRU and the corresponding DMRS port or DMRS sequence. Multiple PRUs can be contained in each PO, each of them associated with specific DMRS configuration. These PRUs are configured consecutively in frequency and time domain for the assumed numerology. This methodology is illustrated in Fig.2, where one PO can occupy multiple PRBs in frequency domain and assuming two PRUs per PO as example, e.g PRU:{DMRS$_{conf1}$, DMRS$_{conf2}$} associated with corresponding DMRS port or DMRS sequence. In this case, aiming for an efficient use of the physical resources, the network can configure multiple 2SR PUSCH resource groups with different configurations according to the resource allocation and payload size [2], [10].

The discussion in 3GPP focused on whether to support one-to-one or multiple-to-one mapping between the preambles in the RO and associated PUSCH data. In the case that all preambles are mapped to the same physical resources, and more than one preamble is successfully detected, the associated PUSCH parts overlap and may result in unsuccessful decoding. On the other hand, if each preamble is mapped to separate physical resources in the PO, the probability of failure in the decoding of the data decreases while the overhead of the RACH procedure increases significantly. There is therefore a trade-off between the collision probability of the PUSCH part of MsgA and the resource overhead for 2SR. In this context, the mapping between PRACH and PUSCH parts of MsgA can be found in [9]. The UE will try to access the network using the PRACH resource that corresponds to the detected SS/PBCH block index (Synchronization / Physical Broadcast Channel) [11]. For a specific preamble transmission, SS/PBCH block indexes are mapped to valid ROs in the following order [9]: (1) in increasing order of preamble indexes within a single RO, (2) in increasing order of frequency resource indexes for frequency multiplexed RO, (3) in increasing order of time resource indexes for time multiplexed RO within a PRACH slot, (4) in increasing order of indexes for PRACH slots. Afterwards, the mapping of one or multiple preambles of a PRACH slot to a PO with a DMRS resource is in the following order [9]: (1) in increasing order of frequency resource indexes for
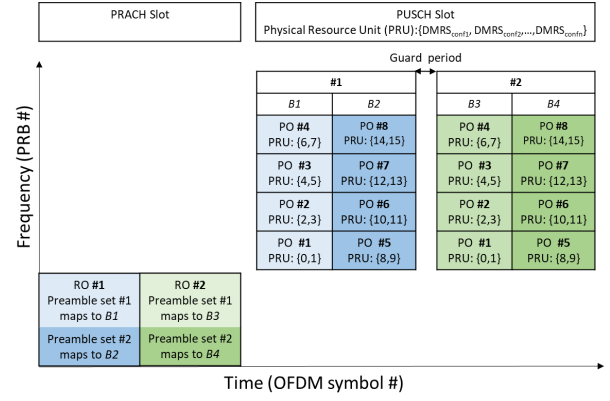

Fig. 2: Example of the mapping between the preamble in the RACH occasion (RO) and the physical resource unit (PRU) in the PUSCH occasion (PO) of MsgA in two-step RACH procedure

frequency multiplexed POs, (2) in increasing order of DMRS indexes within a PO, where a DMRS index is determined first in an ascending order of a DMRS port index and second in an ascending order of a DMRS sequence index [12], (3) in increasing order of time resource indexes for time multiplexed PO within a PUSCH slot, (4) in increasing order of indexes for PUSCH slots.

IV. Channel Structure

A. Preamble configuration

Each UE transmits a RACH preamble to the gNB, which are randomly selected from a maximum of 64 available preambles [12]. The random access preambles are generated from CAZAC codes known as Zadoff Chu (ZC) sequences. Therefore, they have a constant amplitude and their auto-correlation with a cyclically shifted version of itself is zero. A ZC sequence that has not been shifted is known as a root sequence. Hence, different preambles are generated from one or several root sequences by applying cyclic shifts. The detailed base sequence generation algorithm is summarized in [12]. In this study, the number of samples of the cyclic shift N_{cs} is 13. There are multiple preamble formats comprised of one or more PRACH OFDM symbols, different cyclic prefix and guard time. In this study, one short and one long preamble formats are chosen to compare the cell coverage, more specifically, formats A1 and 0.

B. Data resource configuration

The PRU is defined as the PO and DMRS configuration used for a MsgA payload transmission aiming to increase the overall capacity of the 2SR procedure. However, multiple configurations of 2SR PUSCH resources may be required depending on the coverage situation. For instance, smaller allocations may be used for cell-edge UEs due to coverage restrictions while relatively large payload size can be used for cell-center UEs with good channel conditions. Besides, low payload sizes can be assumed for RRC idle or inactive state, while larger payload size can be used when the UE is in RRC connected state. In general, 56 or 72 bits is the most generic operation mode and it is assumed as baseline in this study [2]. Additional simulations are carried out for varying values of payload size and number of PRBs to verify the PUSCH decoding performance.
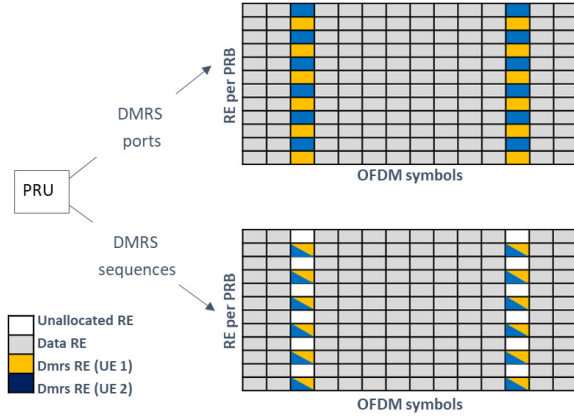
Fig. 3: Slot configuration example of a MsgA payload transmission from 2UEs within the same PO. PRU: {$DMRS_{conf1}$, $DMRS_{conf2}$}, associated with two different DMRS ports (top) or two different DMRS sequences (bottom).

The capacity of the data part of MsgA is limited due to the maximum supported number of UEs sharing the same PUSCH resources for a multiple-to-one mapping between the preamble and data parts. Aiming to improve the resource utilization of the 2SR procedure, we investigate the demodulation performance impact when multiple UEs transmit the MsgA PUSCH in the same resources using either different DMRS sequences and same port or, using different DMRS antenna ports but same sequence. In this case, each UE will randomly select a scrambling ID out of a set of available DMRS sequences or transmit on multiple antennas ports differentiated by either the cyclic shift or orthogonal cover code as described in [12]. These slot configurations are illustrated in Fig. 3 assuming two simultaneously transmitting UEs.

V. METHODS FOR 2SR AND PERFORMANCE METRICS

A. Time offset estimation and compensation

There are at least two approaches for the MsgA PUSCH time adjustment at the gNB side in the 2SR procedure. On one hand, a TOE algorithm based on the PRACH part of MsgA increases the gNB processing complexity and may not be suitable for 2SR procedure due to the increased latency. On the other hand, the accuracy of the DMRS based TOE depends on the number of PRBs allocated, which tend to be rather small in 2SR. The latter has been considered in this study, where the basic processing approach is to coherently combine the two received DMRS symbols and estimate the delay d as follows:

$$d = \arg\max_{n}\{c(n)\} \qquad (1)$$

where $\mathbf{c}$ vector is the channel input response of n samples:

$$\mathbf{c} = ||\sqrt{L} \times F^{-1}(\hat{\mathbf{h}}_{raw})||^2 \qquad (2)$$

and $\hat{\mathbf{h}}_{raw}$ is the raw channel estimation vector of length L with $F^{-1}()$ denoting the inverse Fourier transform and $||\cdot||$ the L2-norm of a vector. Based on the actual DMRS pilots transmitted, the received pilots $\mathbf{y}$ can be defined as:

$$y(f,k) = h(f,k)s(f,k) + n(f,k) \qquad (3)$$

where $\mathbf{h}$ is the unknown complex channel coefficient vector, $\mathbf{s}$ is the transmitted DMRS pilot vector and $\mathbf{n}$ is the noise

vector for the set of received reference subcarriers indexes f and OFDM symbol k. Therefore, the raw channel estimates can be defined as:

$$\hat{h}_{raw}(f,k) = y(f,k) \cdot s^*(f,k) \qquad (4)$$

In this study, a practical wiener filter [13] is used to estimate the channel response. At first, the channel response is estimated in the frequency direction for the OFDM symbols carrying the DMRS. Then, interpolation in the time direction is performed to obtain the channel estimate for the rest of the subframe. In this case, the time offset is compensated (TOC) in the wiener interpolation of the channel in frequency direction by a phase shift to the fourier transform of the delay power spectrum, the autocorrelation function $\mathcal{R}(f)$, for all received data symbols:

$$\mathcal{R}(f) = \mathcal{R}(f)e^{-j2\pi fd} \qquad (5)$$

B. Enhanced 2SR receiver

Intra-cell interference in 2SR procedure should be properly handled when multiple UEs are transmitting simultaneously using same physical resources. Linear interference aware receivers in a multiple antenna receiver can suppress part of the intra-cell interference together with a proper scheduling of the resources. Based on 3GPP assumptions [2], the MMSE-IRC receiver is assumed as baseline for this study, while serial interference aware receivers are not precluded and can be implemented in time or frequency domain. In this case, the covariance matrix of the received signal should be properly estimated, as the more accurate the estimate of the covariance matrix the better the MMSE-IRC receiver will perform. The interference plus noise covariance matrix can be estimated using transmitted reference signals. At first, it can be assumed that the knowledge of the effective channel of other interfering links is unknown at the receiver. Therefore, the estimation will rely on the knowledge obtained from the own reference signal of the target UE. The so-called differential covariance estimation methodology is used in this case. It relies on the assumption that the channel estimation of two consecutive subcarriers is approximately the same. However, under the assumptions of equal transmitted power from multiple UEs in the serving cell, the interference cannot be suppressed properly and the MMSE-IRC receiver may not perform good enough. An enhanced MMSE-IRC receiver could be used, where we can assume we have the knowledge of the interference signals at the gNB node, i.e reference signals used in overlapping transmissions by other UEs, in addition to the desired signal. Consequently, we can also estimate the effective channel of interfering links and the interference covariance matrix can be estimated from the effective channels of multiple interfering links. Based on the fore mentioned assumptions, the received signal $\mathbf{y}$ can be defined as:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} + \mathbf{i} = \mathbf{H}\mathbf{s} + \mathbf{r} \qquad (6)$$

where $\mathbf{H}$ is the channel matrix and, $\mathbf{s}$, $\mathbf{n}$ and $\mathbf{i}$ are the transmitted signal, noise and transmitted interfering signal vectors, respectively. Defining $\mathbf{r}$ as the total interference plus noise vector, the interference covariance matrix of the received signal $\mathbf{y}$ fed to the MMSE-IRC receiver for demodulation can be defined as:

$$C_y = E[\mathbf{y}\mathbf{y}^H] = \sigma_s^2\mathbf{H}\mathbf{H}^H + \mathbf{R}_r \qquad (7)$$

where $(\cdot)^H$ corresponds to the conjugate-transpose and σ_s^2 is the power of the signal $\mathbf{s}$. The following covariance estimation methods can be specified:

1) The differential covariance estimation method: relies on the knowledge obtained from the target UE reference signal and, therefore, the raw interference plus noise covariance estimate can be defined as:

$$\hat{r}(f,k) = y(f,k) - s(f,k)s^*(f+6,k)y(f+6,k) \quad (8)$$

where interference plus noise samples are created differentially from the received samples at neighboring pilot subcarriers, assuming the channel remains approximately unchanged for two contiguous pilot subcarriers, i.e; $h(f,k) \approx h(f+6,k)$. Consequently, the interference plus noise covariance estimate for all subcarriers is defined as:

$$\hat{\mathbf{R}}_r(f,k) = \hat{\mathbf{r}}(f,k)\hat{\mathbf{r}}(f,k)^H \quad (9)$$

2) The enhanced covariance estimation method: relies on the knowledge obtained from the reference signals used in neighboring UEs. Therefore, the effective channel of interfering links can be calculated. The generic form for the interference covariance matrix is defined as:

$$\hat{\mathbf{R}}_r = \sum_{i=1}^{N} \sigma_i^2 \mathbf{H}_i \mathbf{H}_i^H + \sigma_{N_0}^2 \mathbf{I} \quad (10)$$

where N is the number of interferers and σ_{N_0} is the noise power.

C. Performance metrics

The main purpose of the upcoming evaluations is to analyze the coverage of preamble and data parts of MsgA. To this end, the following metrics are used in the evaluations:

Missed detection probability is defined as the probability of not detecting a transmitted preamble with the correct timing estimate. Target value is 1%.

False alarm probability is defined by the ratio of detecting a no transmitted preamble (e.g incorrect preamble or wrong timing estimation) and the total number of possible detection occurrences. Target value is 0.1%.

The link budget calculates the total gain and loss in a system and, therefore, it is an important tool to estimate the coverage areas of 5G cells [14]. The exact link budget analysis is given as follows:

    (1) Tx Power (dBm)
    (2) Thermal noise density (dBm/Hz)
    (3) eNB receiver noise figure (dB)
    (4) Interference margin (dB)
    (5) Occupied channel bandwidth (Hz)
    (6) Effective noise power = (2)+(3)+(4)+10log(5) (dBm)
    (7) Required SINR (dB)
    (8) Receiver sensitivity = (6)+(7) (dBm)
    (9) Receiver processing gain
    (10) MCL = (1)-(8)+(9) (dB)

VI. PERFORMANCE EVALUATIONS

In this section, performance evaluations are provided focusing on the coverage area of 5G cells for the preamble and the data parts of MsgA in 2SR. All evaluations are performed using a 3GPP 5G NR standardization compliant radio link

TABLE I: Physical layer parameterization

Parameter	Value
Carrier frequency [GHz]	4
Sub-carrier spacing [kHz]	30
ISD [m]	[180, 270, 450, 690, 1140]
Time offset [μs]	[0.6, 0.9, 1.5, 2.3, 3.8]
Frequency offset	0.05ppm at Tx, and 0.1 ppm at Rx
Slot duration [ms]	0.5
Channel model	TDL-C DS = 300ns
Antenna configuration	1 Tx $\times$ 2 Rx
Waveform	CP-OFDM
UE velocity [kph]	3
PRACH Preamble Format	A1 and 0
DMRS configuration	Type 1
DMRS allocation density	2 per slot
Payload size [bits]	72, 198, 408, 1032
Number of UEs	[1,3]
Channel estimation	Practical
Receiver	MMSE-IRC

simulator based on the agreed simulation assumptions for the Release-16 study item [2]. Table I summarizes the exact link level simulation assumptions used in the upcoming evaluations.

A. Time adjustment

The demodulation performance of the PUSCH part of MsgA is compared with and without applying DMRS based TOE and TOC algorithms at the receiver side for several TO values as depicted in Fig. 4. In particular, for time delays of [0.6, 0.9, 1.5, 2.3, 3.8] μs, which correspond to distances from the UE to the gNB of [90, 135, 225, 345, 570] meters, respectively. It can be observed a performance degradation of up to 0.2, 0.4 and 2.3 dBs for distances of 90, 135 and 225 meters, respectively. While the performance degradation is significantly increasing for higher distances. Based on 3GPP Release 15 performance tests, gNBs using state-of-the-art channel estimators, e.g the practical wiener filter [13] presented in section V, may assume perfect time alignment for the data part of MsgA. Hence, they will fail in decoding the data based on the presented evaluations. More specifically, the time offset should be estimated and compensated at the gNB in the 2SR procedure to be able to demodulate the data part of MsgA in at least MR and WA cells. In this case, assuming asynchronous UEs with TOs within the CP boundaries, no inter carrier interference is expected as shown in Fig. 4, where no significant performance degradation is observed assuming time adjustment at the receiver side but for TO values above the CP length, i.e, 3.8 μs (CP of 2.34 μs for 30kHz SCS). In this context, TOs within the CP length corresponds to a subcarrier phase rotation of the received signal, which can be estimated and compensated. However, for TOs above the CP length, the orthogonality of the subcarriers may not be preserved impacting the demodulation performance.

Based on the agreed simulation assumptions in Release 16 [5], the 2SR procedure should operate in any cell size despite the lack of TA. Hence, TO estimation and compensation is assumed when specifying gNB demodulation requirements for 2SR in WA and MR cells. In this case, high and medium level TO sets are defined in [5] based on a cycling TO as follows: (X, Δt, Y) μs. For 30kHz SCS, medium and high level TO cycling values correspond to [0,0.2,1] μs and [0,0.1,3.8] μs, respectively. Consequently, for high level TO cycling values,
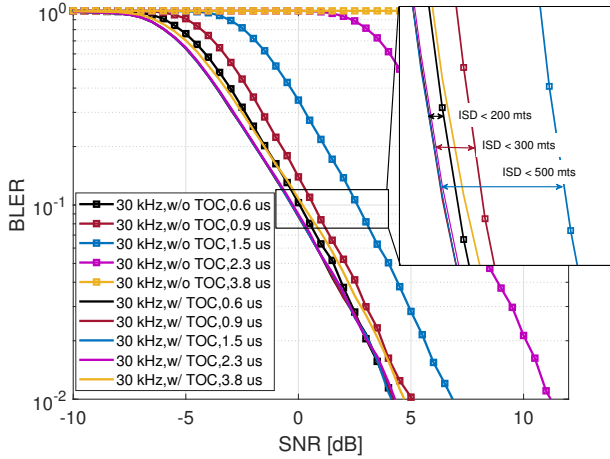
Fig. 4: BLER performance of the PUSCH part of MsgA with (w/) and without (w/o) time offset estimation and compensation at the gNB for several TOs.

TABLE II: Link budget calculation of RACH preamble and the data part of MsgA for single UE or 2 and 3 colliding UEs using different DMRS antenna ports or DMRS sequences and payload sizes of 72, 198, 408 and 1032. The listed values represent the MCL values required to meet the BLER target value of 1%, where SNR results are shown in brackets. (-) BLER target has not been reached for these channel conditions

	RACH configuration			
	Format 0		Format A1	
	131.2 (-5.5)		138.8 (-7.1)	
PUSCH configuration				
Payload	1 UE	2 UEs		3 UEs	
(bits, PRBs)		Ports	Sequences	Ports	Sequences
72, 1PRB	133.8 (2.6)	130.6 (5.8)	129.5 (6.9)	127.6 (8.8)	-
72, 2PRB	134.8 (-1.4)	133.3 (0.1)	132.8 (0.6)	131.9 (1.5)	129.8 (3.6)
72, 3PRB	134.6 (-3.0)	133.7 (-2.1)	133.1 (-1.4)	131.9 (-0.3)	130.16 (1.5)
198, 3PRB	130.8 (0.8)	129.9 (2.7)	128.3 (3.3)	125.9 (5.7)	-
198, 6PRB	131.8 (-3.1)	131.4 (-2.7)	131.1 (-2.4)	129.7 (-1.0)	128.9 (-0.2)
408, 6PRB	134.8 (0.3)	127.9 (0.7)	127.6 (1.1)	125.6 (3.0)	124.5 (4.1)
408, 12PRB	134.6 (-4.6)	129.6 (-4.0)	129.5 (-3.9)	128.3 (-2.7)	127.6 (-2)
1032, 12PRB	136.5 (-0.9)	125.5 (0.1)	125.1 (0.5)	123.0 (2.6)	121.5 (4.1)

38% of TO values within the range are above the CP. Based on evaluations, the SNR requirements to meet the BLER target value of 1% without TOC at the receiver correspond to 4.6 and 15 dB for medium and high level TO sets, respectively. While a similar demodulation performance of 4.2 dB is observed for both TO sets assuming TOC at the receiver but not shown due to space limitations. Further evaluations were performed for a higher TO level to verify that there is no significant performance degradation unless more that approximately 60% of the TO values within the range are above the CP. Based on these results, 3GPP Release 16 gNBs assuming a DMRS based TOE algorithm can support the 2SR procedure and operate in any cell size. Therefore, time adjustment at the receiver side is assumed in all the upcoming evaluations.

B. Preamble detection

The SNR requirements to achieve a target miss-detection probability of 1%, for a given false alarm target of 0.1%, are provided in Table II for the evaluated preambles and single user transmission. Additional simulations were carried out to verify that increasing the number of transmitting UEs for a fixed false alarm rate leads to an increase of the miss-detection rate of preamble detection. Therefore, the final capacity of the 2SR procedure in terms of maximum number of UEs supported is limited due to the increased collision probability between RACH preambles. Based on these results, the detection performance relies on the exact preamble format and number of simultaneously transmitting UEs for a given false alarm target. Besides, it should be noted that the cost of a miss-detection for 2SR is higher than that of 4SR. Mainly due to the fact that a 2SR re-transmission involves re-transmitting the MsgA PRACH and MsgA PUSCH parts, respectively, leading to higher energy transmission requirements from the UE and, therefore, higher interference. For that reason, to optimize the performance of 2SR, the preamble power control parameters could be set in a different manner than those of 4SR but it is out of the scope of this study.

C. Data resource configuration

Assuming that the preamble has been successfully detected at the gNB, the SNR requirements to achieve a target BLER performance of 1% for the PUSCH part of MsgA are shown in Table II for a single UE transmission. Several payload sizes are used in the evaluations according to [2], where the number of PRBs is assumed to be 1, 2, 3, 6 or 12. In general, higher SNR requirements as well as increased number of PRBs are required to support larger payload sizes. The results show that lower MCL differences between MsgA PRACH and PUSCH parts are observed for short preamble format, e.g format A1, than for long preamble format, e.g format 0. In addition, larger MCL differences between MsgA PRACH and PUSCH parts are observed using a larger payload size of 1032 bits compared to a payload of 72 bits. Therefore, it can be expected that higher number of repetitions will be required for larger payload size transmissions, resulting in larger latency. Besides, the resource overhead in terms of number of PRBs required to achieve the target BLER of 1% is considerably larger as increasing the payload size. For instance, payload sizes of 408 or 1032 bits would require at least 6 and 12 PRBs respectively to ensure a BLER of 1% for a single UE allocation within each resource. These observations suggest that a payload size of 72 bits should be prioritized for the 2SR procedure, where the exact configuration may be different per cell size or between RRC states. For instance, lower payload size for RRC-inactive state and larger payload size for RRC-connected state and UEs located in the cell center as the network can control the load and potential PUSCH allocation.

There are multiple configured POs containing several PRUs in the 2SR procedure. In order to increase the resource utilization and decrease the collision probability between UEs, either different DMRS ports or DMRS sequences can be used together with advanced receivers at the gNB. In this context, differential and enhanced covariance estimation methods using a MMSE-IRC receiver are compared in terms of BLER performance in Fig. 5. It should be noted that the MMSE-IRC gain increases with the number of reception antennas, and therefore, the number of Rx antennas was increased up to four. In this case, performance results are shown for the case of 2UEs allocated within the same PRU and using different DMRS antenna ports. These results indicate that the so-called enhanced MMSE-IRC receiver provides significant gain in
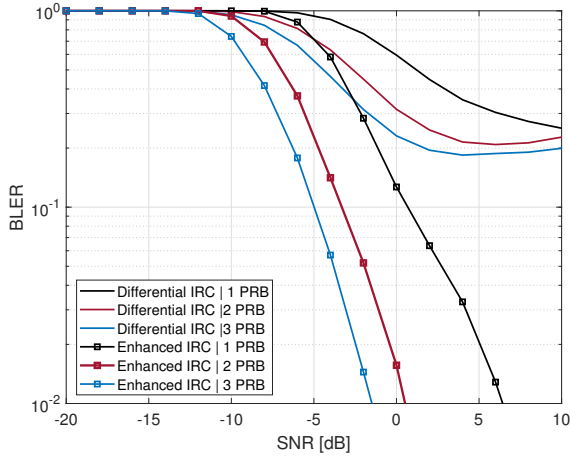
Fig. 5: BLER performance of the PUSCH part of MsgA using 72 bits payload size and different number of PRBs when 2UEs are sharing same PRU resources and, comparing differential and enhanced covariance estimation methods using a MMSE-IRC receiver.

performance as the intra-cell interference cannot be properly suppressed without relying on the knowledge obtained from the reference signals used from other UEs transmitting in the same PRU. Therefore, all the evaluations analyzed in this paper for multiple colliding UEs sharing same physical resources are carried out employing an enhanced MMSE-IRC receiver.

Performance results are presented in Table II for the case of two or three colliding UEs sharing the same physical resources and using different DMRS ports or DMRS sequences for all payload sizes evaluated, respectively. Although one-to-one mapping between preamble and data parts of MsgA can provide the best performance at cost of resource overhead, we observe that increasing the number of DMRS ports or DMRS sequences can reduce the collision probability within the PRU and allow successful decoding of the data. Comparing the MCL obtained when different DMRS ports or DMRS sequences are used, we can observe a difference of 0.5 up to 1 dB for the benefit of using different DMRS ports in the case of a 72 bits payload size. While similar observations can be extracted for other payload sizes. Results adopting a multiple-to-one mapping between preamble and data parts of MsgA show that increasing the number of users multiplexed in a PRU from two to three degrades the BLER performance. In addition, when the payload size increases from 72 to 1032 bits, the performance degradation is more significant and the number of PRBs should be increased to achieve higher MCL. Therefore, more UEs can be multiplexed within the same PO using smaller payload size while achieving higher resource utilization. Consequently, several DMRS ports or DMRS sequences can be used to decrease the collision probability within the PRUs, although the number of UEs trying to perform the 2SR procedure simultaneously on the same PO should be rather small to avoid high collision probabilities. In this case, the signal power of the received SS block could be used to differentiate UEs located at different distances within the cell. Those UEs with similar received signal power, and therefore, similar propagation delays, could be allocated within the same PRU. Based on the results presented in this study, up to three different DMRS ports or sequences provide the best trade-off between resource overhead and achievable performance.

VII. CONCLUSIONS

The 2SR procedure, composed of the denoted MsgA and MsgB, is described and analyzed in this paper based on 3GPP 5G NR Release 16 agreements. The channel structure of MsgA is evaluated for the preamble and data parts, which are allocated in the RACH and PUSCH occasions, respectively. The capacity of the data part of MsgA is limited for a multiple-to-one mapping between the preamble and data parts. In particular, there is a compromise between the collision probability of the PUSCH part of MsgA and the resource overhead for 2SR. The performance results suggest that to suppress the intra-cell interference in the 2SR procedure, a MMSE-IRC receiver where the covariance estimation relies on the knowledge obtained from the reference signals used from other UEs transmitting in the same PRU, in addition to the desired signal, is required. In addition, this procedure should operate in any cell size despite the the lack of TA in the PUSCH part of MsgA. The demodulation performance degradation observed without time adjustment at the receiver side highlight that gNBs assuming a DMRS based TOE algorithm can support the 2SR procedure and operate in any cell size, while practical Release 15 gNB implementations relying in MAC CE-based TA command for PUSCH time alignment are not conceivable. The results indicate that lower payload sizes provide higher resource utilization and allow more UEs to be multiplexed within the same PUSCH occasion. In particular, the cell coverage is higher when using different DMRS ports for UEs sharing same physical resources than when using different DMRS sequences. In this context, up to three different DMRS ports or sequences provide the best trade-off between PRU resource overhead and performance in a multiple-to-one mapping scenario.

REFERENCES

[1] M. Enescu et al., "5G New Radio A Beam-Based Air Interface," *Wiley*, 2020.
[2] "RP-182894 New work item: 2-step RACH for NR," Dec 2018.
[3] "3GPP TR 38900 V15.0.0, "Study on channel model for frequency spectrum above 6 GHz", Tech. Spec. Group Radio Access Network, Rel. 15," June 2018.
[4] "3GPP TR 38889 V16.0.0, "Study on NR-based access to unlicensed spectrum", Tech. Spec. Group Radio Access Network, Rel. 16," Dec 2019.
[5] "R4-2012705, Way forward on BS demodulation requirements for 2-step RACH," 2020, 3GPP TSG-RAN WG1 Meeting #96-e.
[6] "3GPP TR 36884 V13.1.0, "Performance requirements of MMSE-IRC receiver for LTE BS", Tech. Spec. Group Radio Access Network, Rel. 13," Sep 2016.
[7] "R1-1903834 Email discussion on potential link-level simulation assumption for 2-step RACH , Tech. Spec. Group Radio Access Network, Rel. 15," Mar 2019.
[8] Y. Ohwatari et al., "Performance of advanced receiver employing interference rejection combining to suppress inter-cell interference in lte-advanced downlink," *IEEE VTC Fall*, pp. 1–7, 2011.
[9] "3GPP TR 38213 V16.2.0, "Physical layer procedures control", Tech. Spec. Group Radio Access Network, Rel. 16," Jul 2020.
[10] "R1-1906746 On 2-step RACH Channel Structure , Tech. Spec. Group Radio Access Network, Rel. 15," May 2019.
[11] E. Dahlman et al., "5G NR: The next generation wireless access technology," *Academic Press*, 2018.
[12] "3GPP TR 38211 V16.2.0, "Physical channels and modulation", Tech. Spec. Group Radio Access Network, Rel. 16," Jul 2020.
[13] W. C. Jakes et al., "Microwave mobile communications," *Wiley-IEEE Press*, 1994.
[14] H. Holma et al., "5G Technology: 3GPP New Radio," *John Wiley and Sons*, 2020.