Evolution of the 5G New Radio Two-Step Random Access towards 6G Unsourced MAC

Patrick Agostini[†], Jean-Francois Chamberland[‡], Federico Clazzer[§], Johannes Dommel[†], Gianluigi Liva[§], Andrea Munari[§], Krishna Narayanan[‡], Yury Polyanskiy[⋆], Slawomir Stanczak[†] and Zoran Utkovski[†]

[†]Fraunhofer Heinrich-Hertz-Institute
[‡]Department of Electrical and Computer Engineering, Texas A&M University
[⋆]Laboratory for Information and Decision Systems, Massachusetts Institute of Technology
[§]German Aerospace Center, DLR

Version 1.0

April 18, 2024

This report summarizes some considerations on possible evolutions of grant-free random access in the next generation of the 3GPP wireless cellular standard. The analysis is carried out by mapping the problem to the recently-introduced unsourced multiple access channel (UMAC) setup. By doing so, the performance of existing solutions can be benchmarked with information-theoretic bounds, assessing the potential gains that can be achieved over legacy 3GPP schemes. The study focuses on the two-step random access (2SRA) protocol introduced by Release 16 of the 5G New Radio standard, investigating its applicability to support large MTC / IoT terminal populations in a grant-free fashion. The analysis shows that the existing 2SRA scheme may not succeed in providing energy-efficient support to large user populations. Modifications to the protocol are proposed that enable remarkable gains in both energy and spectral efficiency while retaining a strong resemblance to the legacy protocol.

Early versions of this work have been presented at the *Algorithmic Structures for Uncoordinated Communications and Statistical Inference in Exceedingly Large Spaces* workshop held in Banff (Canada), March 10 - 15, 2024. P. Agostini, F. Clazzer, J. Dommel, G. Liva, A. Munari, S. Stanczak and Z. Utkovski acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the program of "Souverän. Digital. Vernetzt." Joint project 6G-RIC, project identification numbers: 16KISK022, 16KISK020K and 16KISK030.

Contents

1 Introduction

This report explores efficient access paradigms and algorithmic frameworks for machine-driven data transfers. The impetus for this initiative is the increasing heterogeneity of wireless traffic, mainly driven by unattended devices. The wireless community has recognized that traditional infrastructures designed for human-centric Internet interactions face challenges in accommodating the fundamentally different ways in which machines and Internet of Things (IoT) devices transfer information. Unlike humans who typically establish sustained connections and send large packets, machines transmit short sporadic payloads. Furthermore, the density of machine is expected to far exceed that of humans in the near future. Thus, the current architectures, based on the *acquisition-estimation-scheduling* paradigm, are ill-suited for this emerging digital landscape.

To address this challenge, modern wireless standards include modalities based on efficient random access (RA) protocols. Such mechanisms are attuned to the unpredictable nature of machine-type communication, yet their efficiency can be enhanced. This report explores novel access framework and algorithmic enhancements tailored for small payload communication and sporadic traffic, in the context of future 3GPP cellular standard releases. A departure from more traditional paradigms is key to eliminate the need for individualized feedback and making high-throughput connections more cost-effective for machine-type data transfers.

In a recent release [1], the 5G New Radio (5GNR) standard included a modificaton of the legacy four-step random access (4SRA) protocol adopted by earlier relesases of 5GNR as well as by the Long Term Evolution (LTE) and the Narrowband IoT (NB-IoT) specifications [2, 3]. The modification, aiming at reducing the latency entailed by the four-step handshake of the 4SRA protocol, reduces the RA procedure to two steps, consisting of the transmission of an uplink message, followed by a feedback acknowledgment. The method, referred to as two-step random access (2SRA), allows resuming the legacy 4SRA procedure whenever a transmission fails (details on the protocol behavior are discussed in Section 3). The first part of the 2SRA protocol, though, provides a first form of grant-free transmission mechanism. It is hence natural to ask whether the existing 2SRA procedure can be used as a blueprint to develop a grant-free scheme for massive machine-type communication (MTC)/IoT scenarios. This document aims at providing some first answers to this question. To enable a fair assessment on the efficiency of the 5GNR 2SRA protocol in massive grant-free scenarios, we use the information theoretic model of [4], referred to as unsourced multiple access (UMAC) model in the following. Monte Carlo simulation results over the Gaussian multiple access (MAC) and over the quasi-static fading MAC reveal that the current specification of the 2SRA protocol — adapted to the grant-free setting of [4]— is substantially sub-optimal in terms of both energy and spectral efficiency. Large performance gaps are identified with respect to the finite-length performance benchmarks of [4], as well as with respect to the performance achieved by recently-introduced UMAC schemes [5–9]. The analysis of the 2SRA protocol reveals that the scheme may be improved in two main areas: by introducing an enlarged family of physical random access channel (PRACH) preambles, and by enriching the set of patterns that users can adopt, when accessing the channel. Based on these observations, we formulate a modification of 2SRA protocol that fully embraces the use of

preambles to signal a rich set of access patterns. The scheme can be viewed as a combination of two UMAC schemes that attracted some attention in the research community, namely, the sparse interleaver division multiple access (IDMA) construction of [6] and the contention resolution diversity slotted Aloha (CRDSA) / coded slotted Aloha (CSA) protocols of [10–12]. As for sparse IDMA, the proposed protocol uses preambles to identify interleaving patterns that users will employ to transmit coded bits. As for CRDSA/CSA, transmissions by a user are organized in *bursts* (or *segments*). The receiver makes use of successive interference cancellation (SIC) to mitigate multiuser interference. Its behavior is assumed to be similar to the one introduced in [6], with decoding of the user data that exploits the combined observations of the segments transmitted by the user. Noting that this option is made possible by the detection of the user preambles, and that the original CRDSA/CSA construction relies on the decoding of the individual segments, we recognize that the proposed protocol is closer in spirit to sparse IDMA. Leaning on this observation, and to the bursty nature of the construction, we referred to the new protocol as sparse block interleaver division multiple access (SB-IDMA).

The rest of the report is organized as follows.

- Section 2 outlines the reference notation, as well as the channel models used to analyze the performance of the various schemes.
- Section 3 provides a high-level description of the 2SRA protocol. Standard-specific details that are omitted/disregarded to put emphasis on the 2SRA basic structure and performance drivers. A performance analysis over the simple models of Section 2 is provided, focusing on configurations that are close enough to the ones used in the recent UMAC literature.
- Section 4 focuses on possible improvements of the 2SRA protocol, providing a description of SB-IDMA. Numerical results are provided for the settings already used in the 2SRA simulations of Section 3.
- Section 5 contains the main conclusions of the study, suggesting directions that may enable the development of a high-efficiency grant-free random access option based on 2SRA for grant-free in future releases (6G) of the 3GPP wireless cellular standard.

To facilitate the reading of the report, a list of the used acronyms and of the main notation elements is included at the end of the document.

2 Preliminaries

2.1 Communication Models

In the following, two channel models will be used when comparing the performance of various random access schemes, namely the Gaussian MAC model, and a quasi-static fading MAC model. We denote by $\mathcal{X} \subseteq \mathbb{C}$ the alphabet used for transmission, i.e., the set of symbols that can be used by the transmitter. We assume that the user population comprises K user terminals (UTs), out of which $K_a \ll K$ are active. Each active user attempts the transmission of k information bits. We consider transmission over n channel uses, and we refer to n as the UMAC frame length. In the Gaussian MAC case, the channel output (for a single channel use) is modeled as

$$Y = \sum_{i=1}^{K_a} X^{(i)} + Z$$

where $X^{(i)} \in \mathcal{X}$ is the symbol transmitted by the ith active UT, and Z is the complex circularly symmetric additive white Gaussian noise (AWGN) term, i.e., $Z \sim \mathcal{CN}(0, \sigma^2)$. The sequence of symbols transmitted by the ith user is

$$\boldsymbol{X}^{(i)} = \left(X_1^{(i)}, X_2^{(i)}, \ldots, X_n^{(i)}\right)$$

and we enforce the power constraint

$$||\boldsymbol{X}^{(i)}||_2^2 \leq nP.$$

The per-user signal-to-noise ratio (SNR) is

$$\frac{E_b}{N_0} = \frac{nP}{k\sigma^2}$$

where E_b is the energy per information bit, and N_0 is the single-sided noise power spectral density. In the quasi-static fading MAC case, we assume each UT transmission to be affected by Rayleigh block fading, with fading coefficients that are constant over the whole UMAC frame, and that are independent across UTs. Therefore,

$$Y = \sum_{i=1}^{K_a} H_i X^{(i)} + Z$$

where $H_i \sim \mathcal{CN}(0, 1)$. The average per-user SNR is

$$\frac{\bar{E}_b}{N_0} = \frac{nP}{k\sigma^2}$$

where $\bar{E}_b$ is the average energy per information bit.

Following the UMAC model [4], all users use the same code $\mathcal{C}$ with blocklength n, where the number of codewords is $|\mathcal{C}| = 2^k$, i.e., each user can transmit k information bits. In the idealized case where the receiver is able to obtain an accurate estimate of the number of active users, the

Table 1. 5G NR numerology Release 17

numerology	SCS	OFDM sym. len.	CP len.	slots per subframe	slot len.	PRB wid.
$\mu = 0$	15 kHz	66.67 μs	4.8 μs	1	1000 μs	180 kHz
$\mu = 1$	30 kHz	33.33 μs	2.4 μs	2	500 μs	360 kHz
$\mu = 2$	60 kHz	16.67 μs	1.2 μs	4	250 μs	720 kHz
$\mu = 3$	120 kHz	8.33 μs	0.6 μs	8	125 μs	1440 kHz
$\mu = 4$	240 kHz	4.17 μs	0.3 μs	16	62.5 μs	2880 kHz
$\mu = 5$	480 kHz	2.09 μs	0.15 μs	32	31.25 μs	5760 kHz
$\mu = 6$	960 kHz	1.04 μs	0.075 μs	64	15.63 μs	11520 kHz

decoder outputs an unordered list $\mathcal{L}$ of K_a codewords. The per-user probability of error is then given by

$$\mathsf{PUPE} = \frac{1}{K_a} \sum_{i=1}^{K_a} \mathsf{P}\Big(\boldsymbol{X}^{(i)} \notin \mathcal{L}\Big) . \tag{1}$$

Box 1: On the simulation setup. *In the simulation results provided in this report, we make use of the following assumptions:*

[A1] *Unless otherwise stated, pilot sequences are sampled at random, with entries that are i.i.d. complex Gaussian with zero mean and variance P.*

[A2] *Similarly, if no specific statement is provided, preambles are generated at random, with entries that are i.i.d. complex Gaussian with zero mean and variance P. For the special case of the 5GNR two-step random access protocol, preambles are generated according to the standard, i.e., from sets of Zadoff-Chu sequences.*

[A3] *In the definition of the signal-to-noise ratio, we will neglect the energy overhead entailed by the orthogonal frequency-division modulation (OFDM) cyclic prefix (CP).*

2.2 Basics of 5G NR Numerology

5G NR defines a flexible numerology to accommodate various scenarios and use cases. The system is based on OFDM and each configuration is characterized by a specific sub-carrier spacing (SCS) and a dedicated CP duration. The numerology $\mu \in \{0, 1, \ldots, 6\}$ identifies the SCS according to $15 \cdot 2^{\mu}$ kHz. In the time domain, 14 OFDM symbols fit into one slot, $1/2^{\mu}$ slots compose one subframe whose duration is 1 ms, finally 10 subframes constitute one frame. In the frequency domain, 12 consecutive subcarriers form one physical resource block (PRB) so that its bandwidth occupation corresponds to $180 \cdot 2^{\mu}$. Note that the minimum resource allocation for data tranmission is 1 PRB. We list in Tab. 1, relevant parameters for each numerology.

In 5G NR the frequency range up to 52.6 GHz is considered. It is subdivided in two ranges, FR1

Table 2. 5G NR numerology to FR mapping for Data and Synchronization

numerology	SCS	data FR1	data FR2	sync. FR1	sync. FR2
$\mu = 0$	15 kHz	✓		✓	
$\mu = 1$	30 kHz	✓		✓	
$\mu = 2$	60 kHz	✓	✓		
$\mu = 3$	120 kHz		✓		
$\mu = 4$	240 kHz				✓
$\mu = 5$	480 kHz		✓		✓
$\mu = 6$	960 kHz		✓		✓

Table 3. Number of PRBs as function of transmission bandwidth and SCS

SCS	5 MHz	10 MHz	15 MHz	20 MHz	25 MHz	30 MHz	35 MHz	40 MHz
15 kHz	25	52	79	106	133	160	188	216
30 kHz	11	24	38	51	65	78	92	106
60 kHz	–	11	18	24	31	38	44	51
SCS	45 MHz	50 MHz	60 MHz	70 MHz	80 MHz	90 MHz	100 MHz	
15 kHz	242	270	–	–	–	–	–	
30 kHz	119	133	162	189	217	245	273	
60 kHz	58	65	79	93	107	121	135	

and FR2. The former includes the sub-6 GHz frequencies while the latter includes millimiter wave ranging between 24.25 GHz and 52.6 GHz. Since our focus is on massive IoT communications, FR1 appears to be the best suited frequency range to be considered as it provides long-range capabilities. Not all available numerologies of Tab. 1 can be used in FR1 or FR2, as can be seen from Tab. 2. To be noted that in the field *data FR1* of Tab. 2 we consider both physical uplink shared channel (PUSCH) and physical downlink shared channel (PDSCH) as possible physical data channels while, for synchronization – *sync. FR1* in Tab. 2 – we include primary synchronization signal (PSS), secondary synchronization signal (SSS) and physical broadcast channel (PBCH).

Depending on the transmission bandwidth selected and the SCS, different number of PRBs are available for data transfer. The corresponding PRBs as a function of transmission bandwidth and SCS are listed in Tab. 3.

The data transmission is refered as to PUSCH occasion. Once the modulation, code rate and few additional aspects related to the 5G NR configuration are fixed, one can compute the transport block size (TBS) and consequently the dimension of one data transmission in terms of PRBs. More detail can be found in [13, 14].

Prior to data transmission, a preamble is sent in order to notify the basestation of the presence of a novel UT wanting to access the network and send data. The preamble is sent over the PRACH resources which are independent of the PUSCH ones. The SCS in the PRACH can differ from

Table 4. Preamble length as a function of PRACH SCS

SCS	139	571	1151
15 kHz	✓	–	✓
30 kHz	✓	✓	–
60 kHz	✓	–	–
120 kHz	✓	✓	✓
240 kHz	–	–	–
480 kHz	✓	–	✓
960 kHz	✓	–	–

Table 5. Short preamble formats 5G NR (durations normalized to symbol duration)

format	cyclic prefix	symbols	repetitions	guard period
A1	0.141	2	2	0
A2	0.281	4	4	0
A3	0.422	6	6	0
B1	0.105	2	2	0.035
B2	0.176	4	4	0.105
B3	0.246	6	6	0.176
B4	0.457	12	12	0.387
C0	0.605	1	1	0.535
C2	1	4	4	1.422

the SCS for the PUSCH and a mapping between the two can be derived as per [15]. There are two types of preambles, long preambles composed by 839 symbols inherited from LTE, which also feature a smaller carrier spacing of either 1.5 kHz or 5 kHz and can be used only in FR1, and short preambles. The latter are more relevant to us and feature an SCS$\in \{15, 30, 60, 120, 480, 960\}$ kHz and a duration of 139, 571 or 1151 symbols depending on the configuration. The preamble formats are listed in table 5 and differ in terms of CP length, number of repetitions and the guard duration [16]. Note that all durations in the table are in terms of PRACH symbols.

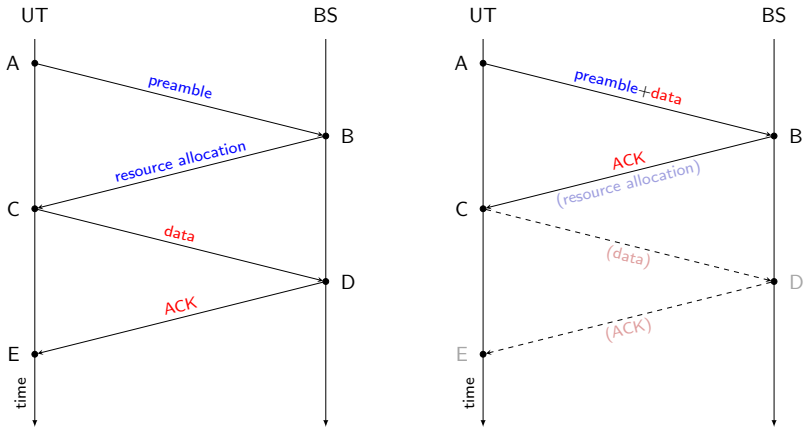
- Subcarrier spacing options in FR1 and FR2
- PRB block definition and bandwidth
- Frame, subframe, number of slots, slot is 14 OFDM symbols.
- CP lenght
- Tx bandwidth, subcarrier spacing and number of PRBs
- PRACH details, format, CP, num. symbols

FC: Here comment on how at MAC layer, the physical resources are mapped to the PRACH

preambles transmissions and the PUSCH occasions if wanted.

3 5GNR Two-Step Random Access

The 5GNR random access channel inherits a four-step handshake mechanism from the LTE standard. From a high-level point of view, the 4SRA protocol works as follows. An active user terminal picks a preamble at random for a set of (up to) 64 Zadoff-Chu sequences. The preamble is sent over the shared PRACH. At the base station, orthogonal resources—over a PUSCH—are granted to each detected user preamble, with the allocation sent back to the UT. In the subsequent phase, the UTs transmit their data units over the allocated resources. Upon decoding the transmitted packets, the receiver acknowledges the success to the transmitters. The 4SRA protocol handshake is summarized in Figure 1a.



(a) Four-step RA (b) Two-step RA

Figure 1. Random access procedures employed by LTE/5GNR standards. (a) Four-step random access: the UTs transmits a preamble (A). Upon detecting the transmitted preambles, the base station (BS) provides a resource allocation to each detected user (B). UTs transmit their data packets in the allocated resources (C). The BS acknowledges the correctly decoded packets (D). The procedure ends when the UT receives the acknowledgement (E). (b) Two-step random access (Release 16 of the 5GNR standard): A UT transmits a preamble, that directly points to the resource that will be used to transmit the data packet. The data packet transmission follows w/o waiting for a resource allocation (A). At the BS, preambles are detected and decoding is attempted in the resources pointed by the preambles. For detected UT transmissions, and acknowledgment is sent to the UTs that are successfully decoded (B). For detected UT transmissions that do not yield to successful decoding, orthogonal resources are allocated for the retransmission of the data packet, resuming the four-step random access procedure.

With Release 16 of the 5GNR standard [1], a grant-free access mechanism is introduced through the so-called 2SRA protocol. In 2SRA, each active UT picks a random preamble from a set of (up to) 64 Zadoff-Chu sequences, and transmits it over the PRACH. Each preamble points to a resource, in the form of a PUSCH occasion (PO), which is used by the UT to transmit its data unit. The mapping can be one-to-one (OTO) or many-to-one (MTO) [17, 18]. In the OTO mapping case, each preamble points to a distinct PO. In the MTO mapping case, several preambles can point to the same PO, possibly employing different pilot sequences to enable channel estimation even in the presence of collisions. At the BS, the receiver attempts demodulation/decoding at each PO pointed by the detected preambles. For decoded packets, BS acknowledges reception through a feedback message. If a preamble is detected, but the

decoding of the associated packet transmission fails, the legacy 4SRA procedure is resumed: the BS signals back an orthogonal resource allocation to the corresponding UT, which proceeds with the retransmission of its packet in the granted resource unit. The procedure defined by the 2SRA protocol is outlined in Figure 1b.

Box 2: Mapping the 5GNR 2SRA protocol on the UMAC setting. *With the perspective of enabling massive grant-free connectivity, we next focus on the first phase of the 2SRA, i.e., we study its performance in isolation, excluding the possibility of exploiting feedback to resume the 4SRA grant-based procedure for unsuccessful packet transmission. We do so by removing ancillary aspects of the protocol, such as the specific mapping of the PRACH and of POs in the 5GNR framing structure. More specifically, we consider our channel model as a sequence of channel uses — The reader should bear in mind that the channel uses refer to specific time/frequency resources in the 5GNR OFDM grid. With reference to Figure 2, we denote by*

- n_{PRE} *the preamble length in channel uses (c.u.s);*
- N *the number of available POs;*
- n_{PO} *the length in c.u.s of each PO.*

It results that the frame length is given by

$$n = n_{\mathsf{PRE}} + N \times n_{\mathsf{PO}}.$$

We will consider parameters sets that may not fully be aligned with the 5GNR numerology, when seeking the need to compare 2SRA with alternative solutions. When doing so, we will nevertheless consider configurations that are sufficiently close to configurations that can be obtained using the 5GNR numerology/framing—hence, obtaining realistic estimates on the performance that can be achieved by 2SRA.

3.1 Configurations

In the following, we will make use of the notion of *access pattern*, defined below.

Definition 1 (Access Pattern) *With reference to Figure 2, and assuming an order of the POs, the access pattern of a user is defined as a binary N-tuple*

$$\boldsymbol{a} = (a_0, a_1, \ldots, a_{N-1})$$

where $a_i = 1$ if the user transmits in the ith PO, and it is zero otherwise.

Access patterns define how users can access the set of available POs. Obviously, each preamble is associated with a specific access pattern. Figure 2 provides a simplified description of the first phase of the 2SRA protocol. In the example, user 1 selects a preamble that is associated with the access pattern with $a_1 = 1$ and $a_i = 0$ for $i \neq 1$. Upon detecting the preambles of the active users, the BS attempts decoding at the POs identified by the corresponding access patterns. In case of two or more users transmitting at the same PO, correct decoding may be hindered by the mutual interference caused by the colliding users. Two configurations will be considered:

- A OTO configuration (Table 6), where each preamble points to a different PO;
- A MTO configuration (Table 7) where several preambles point to the same PO.

The MTO configuration considered in the example represents an extreme case of a MTO allocation where all preambles are mapped to a single PO. In this case, a few random access channels are bundled together, to obtain a number of resources (channel uses) that is comparable with the OTO configuration. Note that the parameters in the MTO setting are provided only for the fading MAC case, since the scheme is obviously outperformed by its OTO counterpart over the Gaussian MAC.

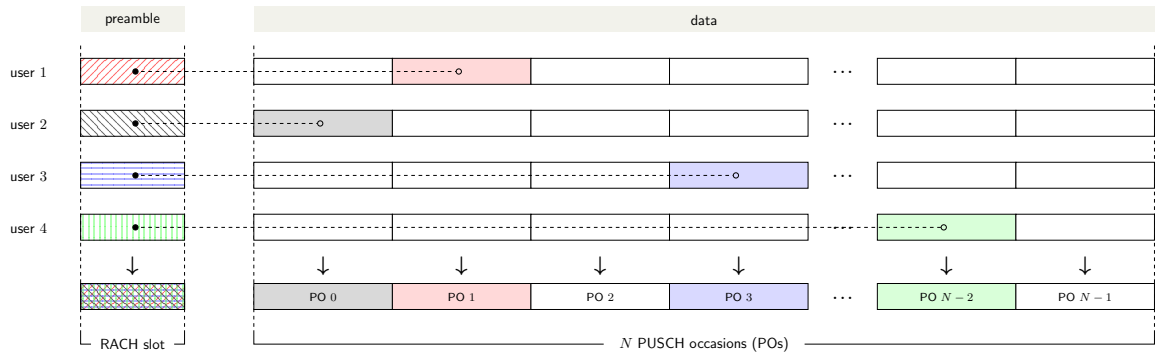


Figure 2. Illustration of the two-step random access protocol of the 5GNR standard (first transmission only). The number of PUSCH occasions is N.

Table 6. Parameters used for 2SRA simulations – OTO allocation.

Parameter	Gaussian MAC	Fading MAC	Unit
Information bits per user	100	100	bits
Frame length (n)	16278	19478	c.c.u.
# POs (N)	64	64	-
Preamble length	139	139	c.c.u.
Preamble repetition rate	2	2	-
Channel code	5GNR LDPC (BG2)	5GNR LDPC (BG2)	-
Block length (n_c)	500	500	bits
Modulation	QPSK	QPSK	-
# pilot symbols per PO	-	50	-

Table 7. Parameters used for 2SRA simulations – MTO allocation.

Parameter	Gaussian MAC	Fading MAC	Unit
Information bits per user	-	100	bits
Frame length (n)	-	20230	c.c.u.
# POs (N)	-	1 ($\times$35)	-
Preamble length	-	139	c.c.u.
Preamble repetition rate	-	2	-
Channel code	-	5GNR LDPC (BG2)	-
Block length (n_c)	-	500	bits
Modulation	-	QPSK	-
# pilot symbols per PO	-	50	-

3.2 Receiver Algorithms

The following box provides a high-level description of the receiver algorithms used for the simulations of 2SRA (the same algorithms will be used for the simulations of the modifications of 2SRA of Section 4).

Box 3: On Detection and Decoding. *Simulations for 2SRA and for the variations introduced in Section 4 rely on the following detection / decoding algorithms.*

- *Preamble detection: the orthogonal matching pursuit (OMP) algorithm [19, 20] is used to detect preambles. The algorithm is set to provide a list of L preambles, where L is chosen to be larger than the number of active users K_a. A typical value of L is $\approx 1.5K_a$. This choice allows to reduce the number of preamble misdetections. Preambles that have not been transmitted trigger a decoding stage, which is responsible of filtering out transmissions that did not take place. For each preamble identified by the OMP algorithm, the steps below are performed.*

- *Channel estimation: on the Gaussian MAC, the (unitary) channel coefficient of each user is assumed to be known at the receiver. On the quasi-static fading MAC, a pilot field is included in each PO where a user is attempting transmission. The pilot field is uniquely determined by the preamble chosen by the user. Denote by* $\boldsymbol{x}_{\mathsf{P}}^{(i)}$ *the pilot field used by the* i*th user active in a given PO, and by* $\boldsymbol{y}_{\mathsf{P}}$ *the corresponding observation (affected by fading, AWGN and multiuser interference). The estimate of the channel for* i*th user is obtained as*

$$\hat{h}_i = \frac{\langle \boldsymbol{y}_{\mathsf{P}}, \boldsymbol{x}_{\mathsf{P}}^{(i)} \rangle}{\left\| \boldsymbol{x}_{\mathsf{P}}^{(i)} \right\|_2^2}. \tag{2}$$

- *Interference-plus-Noise Power: Within each PO, the number of active transmissions is unknown to the receiver. Hence, the power of noise plus interference needs to be estimated on a per-PO basis. We use a blind estimator that evaluates the interference-plus-noise power as*

$$\mathsf{NI} = \frac{1}{n_{\mathsf{PO}}} \|\boldsymbol{y}\|_2^2 \tag{3}$$

  where $\boldsymbol{y}$ *is the observation associated with the PO.*

- *Computation of log-likelihood ratios (LLRs): LLRs are computed using the channel estimate provided by* (2) *and the interference-plus-noise power estimate from* (3)*. Note that, although the quasi-static fading MAC maintains the channel coefficients of the active users constant across a frame, channel estimation will be performed on a per-PO basis. Hence, the LLRs computed within a PO will be based only on the PO-specific channel estimate. Moreover, we shall observe that the interference-plus-noise power estimate from* (3) *does not depurate the estimate from the useful signal power. Therefore, it tends to provide a pessimistic quantification of the interference-plus-noise power.*

- *Decoding: When packets are encoded with low-density parity-check (LDPC) codes, we assume belief propagation (BP) decoding with* 50 *iterations. When packets are encoded with polar codes, we assume successive cancellation list (SCL) decoding with adaptive list size, with maximum list size set to* 128 *[21, 22].*

- *Packet Validation: For LDPC codes, error detection relies on the output of the BP decoder: an error flag is raised in the decoded word does not fulfill the code parity-check equations. For the polar code case, a specific error detection method will be discussed in some detail in Section 4.*

A receiver that operates according to the steps described above is referred to as treat-interference-as-noise (TIN) receiver. We will consider also the improvement obtained by applying SIC, where

- When a packet is correctly decoded, the corresponding modulated codeword is used, together with the pilot field, to re-estimate the channel coefficient using an approach that is analogous to the one of (2). Using the improved channel estimate, the pilot field and codeword contributions are removed (cancelled) from the corresponding PO.

- The contribution of the corresponding preamble is also subtracted from the PRACH. The

channel estimate used for the preamble cancellation is the one delivered by the OMP algorithm.

- After interference cancellation, the whole procedure described in Box 3 is performed again, i.e., preamble detection followed by the decoding attempts over the residual signal. The iterative procedure is repeated until no more valid packets are recovered.

Note that the estimation / detection procedures outlined above can be improved, for example, by developing more sophisticated interference-plus-noise power estimators, by performing joint decoding and channel estimation, by adopting more powerful preamble detection techniques, and (in case of non-orthogonal pilot sequences) by replacing (2) with a minimum mean square error (MMSE) channel estimator. For the results provided in this document, it was decided to use simpler — yet, suboptimal — algorithms which may well reflect the behavior of existing implementations.

3.3 Two-Step Random Access: Gaussian MAC Performance

To enable comparisons with recently-proposed UMAC schemes, the protocol parameters have been customized to approximate a classical scenario addressed in the UMAC literature, which relies on a MAC frame composed by 15000 complex channel uses (c.c.u.s) — equivalent to 30000 real c.u.s. The parameters used for the simulation refer to the OTO configuration, which achieves the best performance over the Gaussian MAC, and are summarized in Table 6. In a nutshell, the configuration relies on

- A frame length of 16278 c.c.u.s, equivalent to 32556 real c.u.s;
- 64 POs, where each PO consists of 250 c.c.u.s;
- Within a PO, transmission takes place by means of the rate-1/5 5GNR $(500, 100)$ LDPC code (derived from the so-called *base graph* 2 [23]);
- quadrature phase shift keying (QPSK) modulation;
- Preambles chosen from the Zadoff-Chu dictionary with length 139, with two-fold preamble repetition according to configuration A1 [24, Chapter 16].

Figure 3 reports the performance of 2SRA over the Gaussian MAC. The performance is provided in terms of minimum SNR required to achieve a target per-user probability of error (PUPE) of 5×10^{-2}. On the same chart the performance of several UMAC schemes from the literature is shown, together with the UMAC achievability random coding union (RCU) bound of [4] (——). The bound as well as the competing schemes assume 30000 real c.u.s. The schemes for which the performance is provided are T-fold irregular repetition slotted Aloha (IRSA) [5] (—×—), sparse IDMA [6] (—■—), coded compressed sensing (CCS) [7] (—◀—), the CCS sparse regression code (SPARC)-based construction of [8] (—+—), and the spread-spectrum scheme with data-dependent spreading codes from [9] (—●—).

The performance for the tested OTO 2SRA configuration shows a remarkable gap from the Gaussian MAC achievability bound, as well as from the reference UMAC scheme from literature. The gap is visible under both TIN (-⋆-) and TIN-SIC (-○-) decoding, with a saturation of the number of supported users at high SNR that is around $K_a = 15$ in the former case, and $K_a = 35$ in the latter case.

3.4 Two-Step Random Access: Quasi-Static Fading MAC Performance

Over the quasi-static fading MAC, the performance is provided only for the case of TIN-SIC decoding. Figure 4 depicts the performance in terms of minimum SNR required to achieve a target PUPE of 10^{-1}. The results are given for a frame length in the order of 20000 c.c.u.s. As theoretical reference, a non-rigorous estimation on the performance of the optimal decoder applied to a Gaussian codebook (via replica method) [25] is provided (——). The estimation is developed in the asymptotic regime $n \to \infty$, and adapted to the finite frame length as suggested in [25].

Both the OTO (-○-) and the MTO (-○-) configuration yield a performance that is far from the reference. The MTO allocation shows a largely improved performance over the OTO allocation. The reason for the gain stems from the improved multi-packet reception (MPR) provided by the MTO mapping: while the average number of users colliding in a PO is larger than in the OTO case (due to the smaller number N of available POs, see Tables 6 and 7), most of the times colliding users employ a different preamble and, hence, a different pilot sequence, allowing for a sufficiently-accurate estimate of the corresponding channel coefficients. On the contrary, in the OTO allocation, two users colliding in the same PO use the same preamble, hence, the same pilot sequence — hindering the possibility of accurate channel estimation when two users have similar channel gains.

3.5 Two-Step Random Access: Some Considerations

The performance of the tested 2SRA configurations points to the following observations:

- MTO configurations allow to leverage from an improved MPR capability, with respect to the OTO case.
- This comes at the expense on the number of available POs: since more PRACHs have to be allocated, a smaller number of POs can be allocated for a fixed total number of channel uses.
- The tension between the two points above may be solved by increasing the preamble set size, enabling the use of OTO mappings with a single PRACH, and the larger number of POs. This point will be at the core of the solutions investigated in Section 4.
- There are other two main limitations of the existing 2SRA design, which may not be

immediately recognized from the simulation results in Figure 3 and in Figure 4: the limited capability of the 5GNR LDPC codes to resolve multiple collisions, and the inherent lack of diversity with respect to interference. The first point is due to the lowest rate achievable by the 5GNR LDPC codes, which is 1/5. Lower rates may allow for a stronger MPR capability. This issue is exacerbated, at small block length, by the tangible suboptimality of short LDPC codes. The second point is due to the limited number of access patterns entailed by the 5GNR 2SRA scheme: the number of access patterns is given by N, i.e., by the number of POs. The combined effect of these two points is to further limit the performance achievable by the 5GNR 2SRA protocol.

- 2SRA implements a variation of slotted Aloha [26,27], where the POs selected by the users are announced by a preamble transmitted over the PRACH. One may observe that slotted Aloha does not require additional preamble transmission: a classical slotted Aloha receiver would simply attempt decoding at every available PO. By omitting the transmission of the preamble, a significant energy savings may be achieved, shifting the 2SRA curves of Figures 3 and 4 left by approx.

$$10 \log_{10} \frac{n_{\mathsf{PRE}} + n_{\mathsf{PO}}}{n_{\mathsf{PO}}} = 2.84 \, [\mathrm{dB}].$$

A natural question would be then if 2SRA could be operated without the initial preamble transmission. System-level considerations lead us to exclude this possibility, since the initial preamble transmission serves additional purposes. As discussed earlier, it allows to resume the 4SRA procedure in case of decoding failure. Moreover, preambles allow measuring the delay of user transmissions, providing essential information to implement time advance (TA) [24, Chapter 15]. Therefore, the possibility of sparing the preamble transmission should be carefully addressed taking into account its impact at the system level.

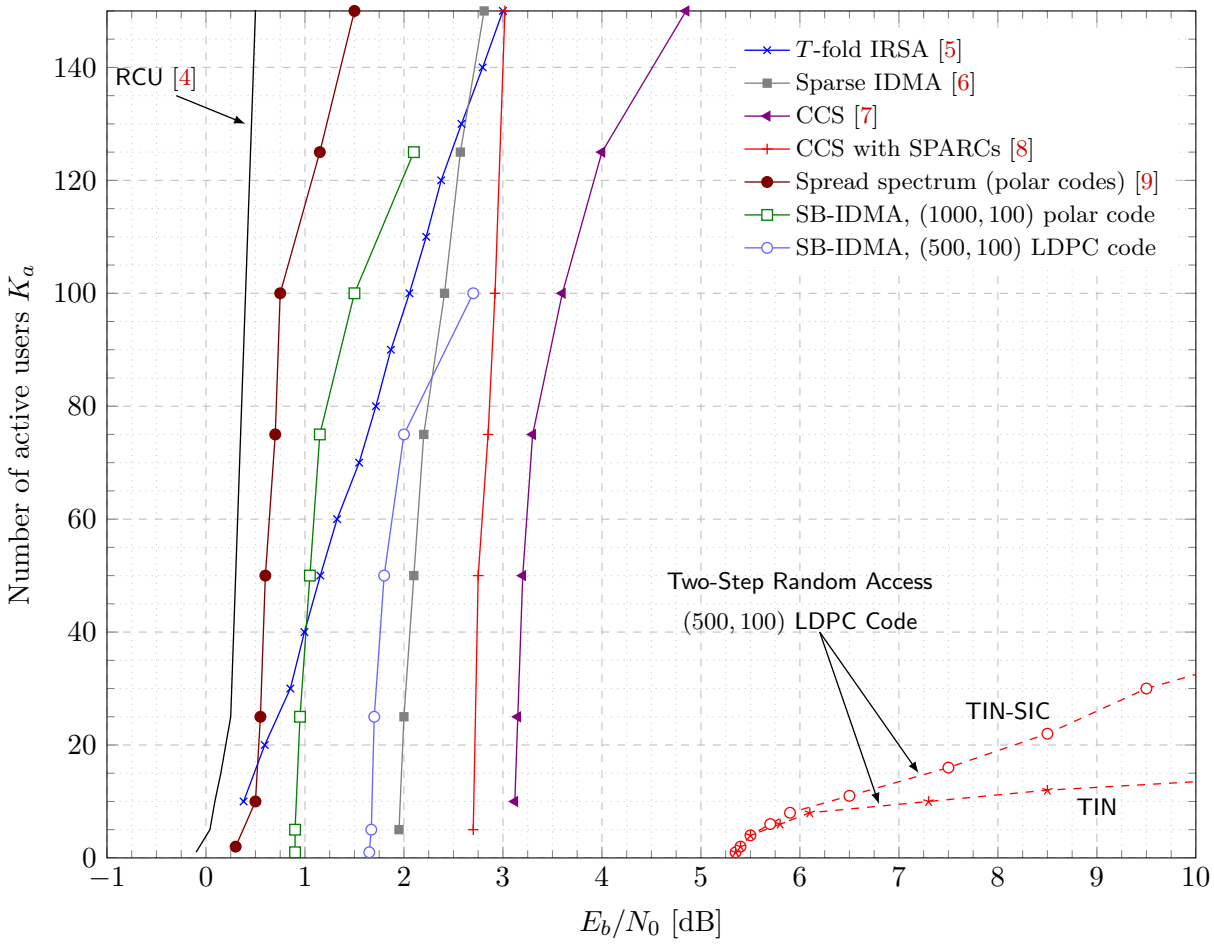


Figure 3. Number of supported active users vs. SNR for a target PUPE $= 5 \times 10^{-2}$. AWGN channel, $n \approx 15000$ c.c.u.s.

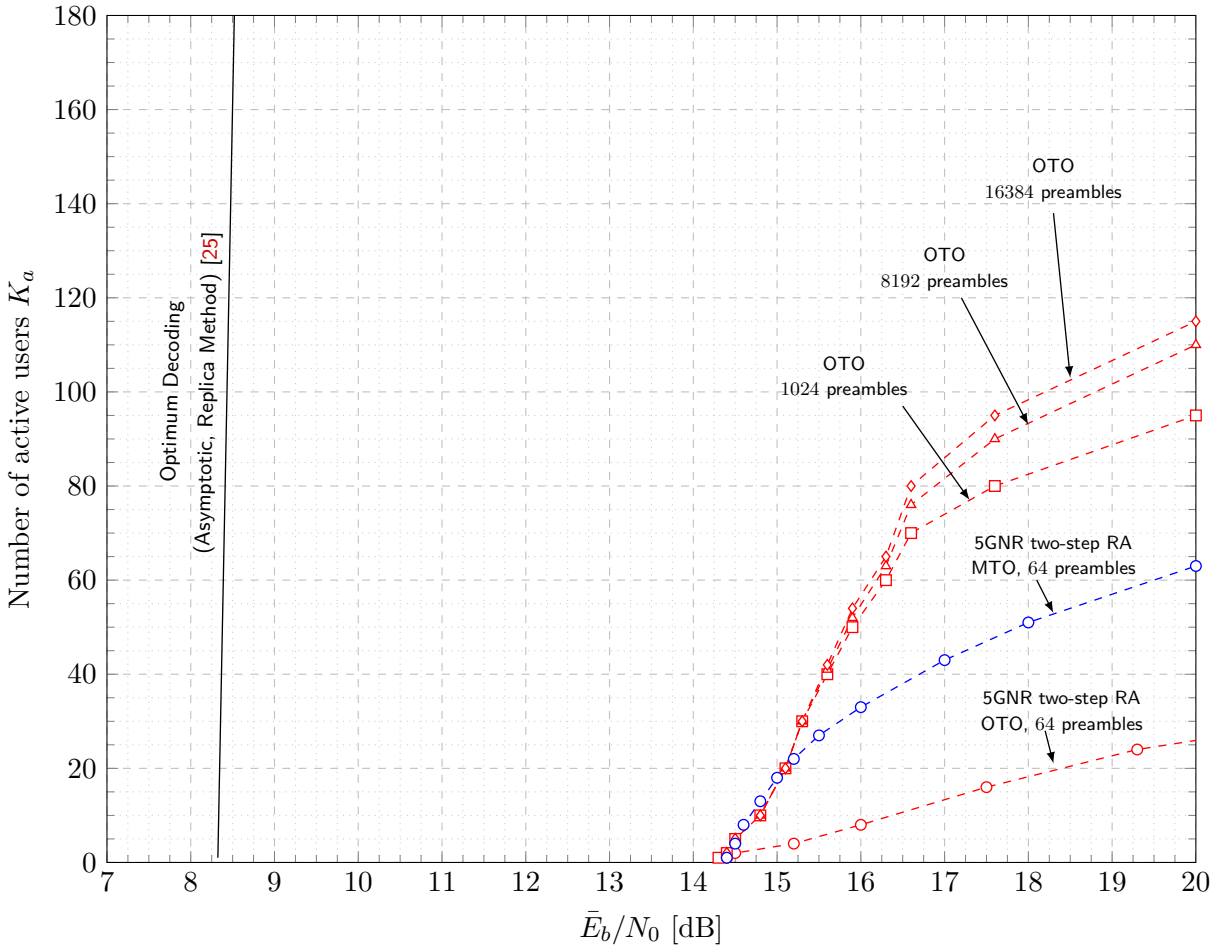


Figure 4. Number of supported active users vs. average SNR for a target PUPE $= 10^{-1}$. Quasi-static Rayleigh fading channel, $n \approx 20000$ channel uses. Single antenna at the base station. The 2SRA performance is provided for the OTO configuration (Table 6).

4 Improving the 5G NR Two-Step Random Access

This section investigates possible improvements of the 5GNR 2SRA protocol. The modifications that will be outlined in the following subsections aim at addressing some of the issues identified in Section 3.5, namely:

- We study the impact of using larger preamble sets (Section 4.1).
- Inspired by the sparse IDMA scheme of [6], we introduce a richer access pattern family, that adapts sparse IDMA to the framing structure of 5GNR. We refer to the proposed scheme as SB-IDMA (Section 4.2).
- For SB-IDMA, we analyze the performance gain that can be achieved, in the short block length regime, by replacing the 5GNR LDPC codes [23] with the 5GNR polar codes [28].

4.1 Extended Preamble Set

A first improvement over the 5GNR 2SRA protocol may be obtained by using preamble sets with cardinality larger than the one of the standardized Zadoff-Chu sequence family. The rationale behind this choice is to remove the tension between the limited MPR capability of OTO configurations (due to the use of identical pilot sequences for colliding users) and the preamble transmission overhead incurred by a repeated use of MTO configurations (see Section 3.5).

Figure 4 reports the performance achievable over the quasi-static fading MAC by enlarging the preamble set to 1024 (-□-), 8192 (-△-), and 16384 (-◇-) preambles. In all three cases, non-orthogonal preambles with symbols drawn independently from a complex Gaussian distribution were used. The preamble length has been fixed to the one used in the 2SRA simulations (see Section 3.1), i.e., $n_{\mathsf{PRE}} = 2 \times 139$ c.c.u.s. Each preamble points one of the $N = 64$ POs, realizing a MTO configuration, and to a unique pilot sequence. The rest of the configuration parameters is unmodified with respect to Table 6. The result with 1024 preambles shows a remarkable gain over the OTO mapping configuration with 64 preambles: the number of active users supported by the system is multiplied by a factor four. The improvement over the MTO configuration with 64 preambles is remarkable, too, although limited to a 50% gain in the number of supported users.

The performance for the case of 8192 and 16384 assumes ideal preamble detection. The reason for this is that, when performing preamble detection via OMP, a large preamble misdetection rate yields a visible performance degradation, calling for the use of stronger preamble detection algorithms, or for an optimized preamble set design. Nevertheless, the results provide still an important insight, that is, enlarging the preamble set beyond a certain limit gives diminishing returns. In fact, the performance achieved with 8192 under ideal preamble detection only marginally improves the performance obtained by the configuration employing 1024 preambles (and actual, OMP-based, preamble detection).

Even for 16384 preamble with ideal detection, the performance shows signs of saturation at about $100 \div 120$ supported active users. The reason of this lies mainly in the limited MPR capability provided by the error correction scheme, and by the limited number of access patterns inherent to the slotted Aloha nature of 2SRA (see Section 3.5). These points are addressed in the following subsection.

4.2 Sparse-Block Interleaver Diversity Multiple Access

The modification of the 2SRA proposed in this section is inspired by the sparse IDMA scheme of [7]. In a nutshell, sparse IDMA works as follows:

- Each user splits its message in two parts: a first part, that is used to selected a preamble, as well as an interleaving pattern, and a second part that is encoded via a binary linear block code, genereting a codeword $\boldsymbol{c}$.
- The codeword $\boldsymbol{c}$ is repeated d times. Zero-padding follows up, resulting in a vector $\boldsymbol{v}$ of a prescribed vector length.
- The preamble is transmitted over the channel, followed by an interleaved version of $\boldsymbol{v}$, with interleaver determined by the first message part (and, hence, by the preamble).
- At the receiver side, upon detecting the preambles of the active users, joint decoding of the users transmission is performed. In particular, in [7] the channel code used to encode the second message part is a suitably-designed LDPC code. Joint decoding is performed by exploiting the knowledge of the users interleaving patterns obtained from the detected preambles: BP decoding is performed over the joint Tanner graph [29] of the detected users.

The SB-IDMA approach closely resembles these steps, with some important differences. First, the entire user message is encoded via an (n_c, k) binary linear block code , no splitting of the user message is performed. Still, the user message is used to select the preamble by hashing the message. Second, while the interleaving used in sparse IDMA applies an unconstrained permutation at the symbol level, in SB-IDMA the interleaver is restricted to permute blocks of symbols, where each block (segment, in the following) is allocated to specific PO. In other words, the interleaver used in SB-IDMA defines the user access pattern (see Definition 1). This choice stems from the need of adapting sparse IDMA to the 5GNR framing, and to transmission over fading channels. The key intuition is that, by clustering symbols in segments, channel estimation can be performed on a per-PO basis by appending a pilot field to each segment. The use of sufficiently-large segments allows to reduce the pilot field overhead. Finally, at the receiver side, we do not assume joint multi-user decoding. We rather consider the use of single-user detection and decoding, enhanced by SIC (TIN-SIC).

4.2.1 Transmission Chain

The transmission diagram of SB-IDMA is given in Figure 5, with steps that are detailed in Figure 6. Recalling that the access frame is composed by N POs, the scheme works as follows:

1. The k-bits message $\boldsymbol{u}$ is hashed, generating an index $\phi(\boldsymbol{u})$ which uniquely identifies (a) a preamble within the preamble dictionary, (b) a set of n_s pilot sequences, and a set of n_s distinct indexes in $[1, 2, \ldots, N]$.

2. The message $\boldsymbol{u}$ is encoded via a binary linear block code $\mathcal{C}$. The resulting codeword is QPSK-modulated, resulting in the n_c-symbols vector $\boldsymbol{c}$ (Figure 6(a)).

3. The modulated codeword $\boldsymbol{c}$ is repeated d times (Figure 6(b)).

4. The resulting vector, composed by dn_c QPSK symbols, is split into n_s segments of equal length (Figure 6(c)).

5. Each of the n_s pilot fields identified in step 1 is appended to the respective segment, i.e., the first pilot field is appended to the first segment, the second pilot field is appended to the second segment, etc. (Figure 6(d)).

6. The preamble selected in Step 1 is appended to the sequence of segments (Figure 6(e)).

7. The preamble is sent over the PRACH. The interleaving pattern selected in Step 1 determines the POs where the individual segments must be transmitted (Figure 6(f)).

4.2.2 Receiver Chain

The receiver behavior is largely based on the procedures described in Box 3, and already adopted for the 2SRA simulations. Iterative TIN-SIC decoding is assumed, as visualized in Figure 7. Some important details about the receiver behavior follow:

- In each iteration of TIN-SIC, as set of L preambles is obtained via OMP, applied to the observation of the PRACH.

- For each detected preamble, the sequence of POs used to transmit the user segments is determined, as well as the sequence of pilot fields.

- In each PO, channel estimation and interference-plus-noise power estimation are performed (see Box 3). LLRs for the codeword bits are computed accordingly. The d LLRs associated with the repetition of each codeword bit are combined (summed) and passed to the decoder of $\mathcal{C}$.

- Assuming an incomplete decoding algorithm, the outcome of decoding can be either a detected error or a valid codeword decision. In the former case, no further action is needed: the decoder output is simply discarded. In the latter case, an additional error

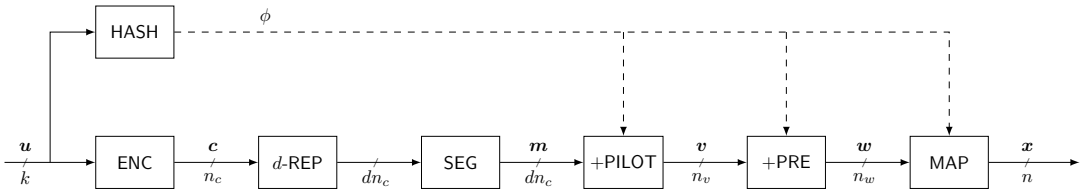


Figure 5. SB-IDMA transmission chain.

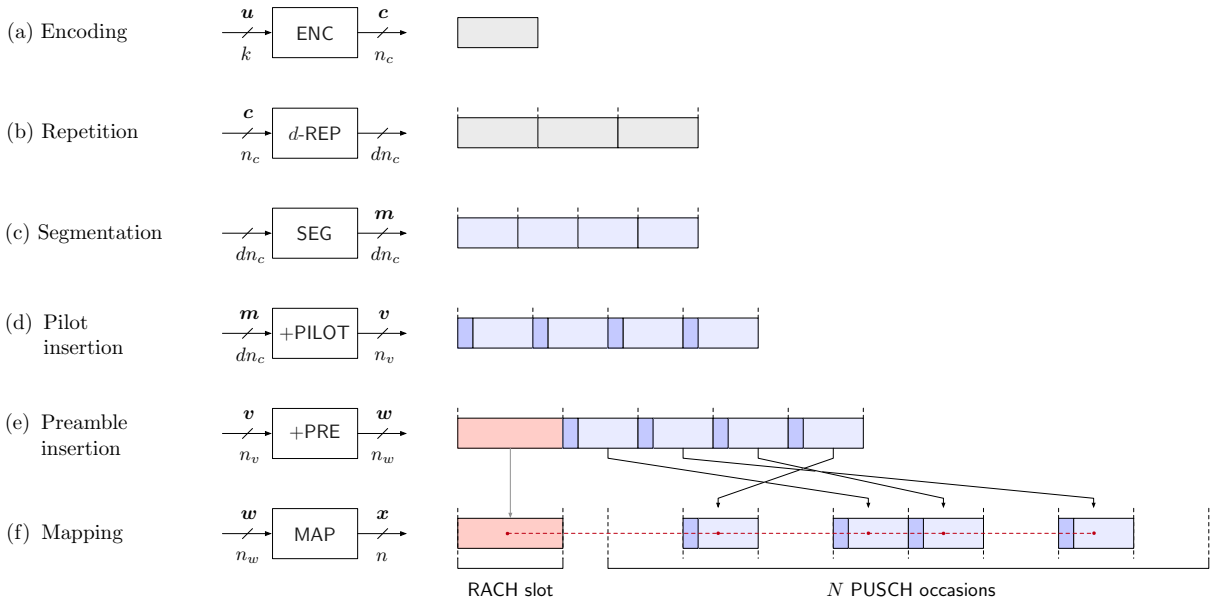


Figure 6. Detailed description of the transmission steps of Figure 5.

detection step is performed by (a) computing the hash of the decoded message and (b) comparing it with the preamble index associated with the decoding attempt: if the two indexes are different, the decoder output is discarded. Otherwise, the decoded message is deemed to be correct, and it is passed at the receiver output.

- For decoded messages that are considered correct, SIC is performed. In particular, as for the discussion that follows Box 3, the channel coefficient is re-estimated, for each segment, using the decoded data as an extended pilot field. The interference contribution of each segment is then removed from the respective PO. Similarly, the interference contribution of the corresponding preamble is cancelled from the PRACH observation, with channel coefficient provided by the OMP algorithm.

4.2.3 SB-IDMA: Performance

In the tested SB-IDMA configurations, two types of error correcting codes are considered: the 5GNR LDPC codes, and the 5GNR polar codes. In the former case, BP decoding is assumed with 50 decoding iterations. In the latter case, adaptive SCL decoding is used [22], with a maximum list size set to 128.

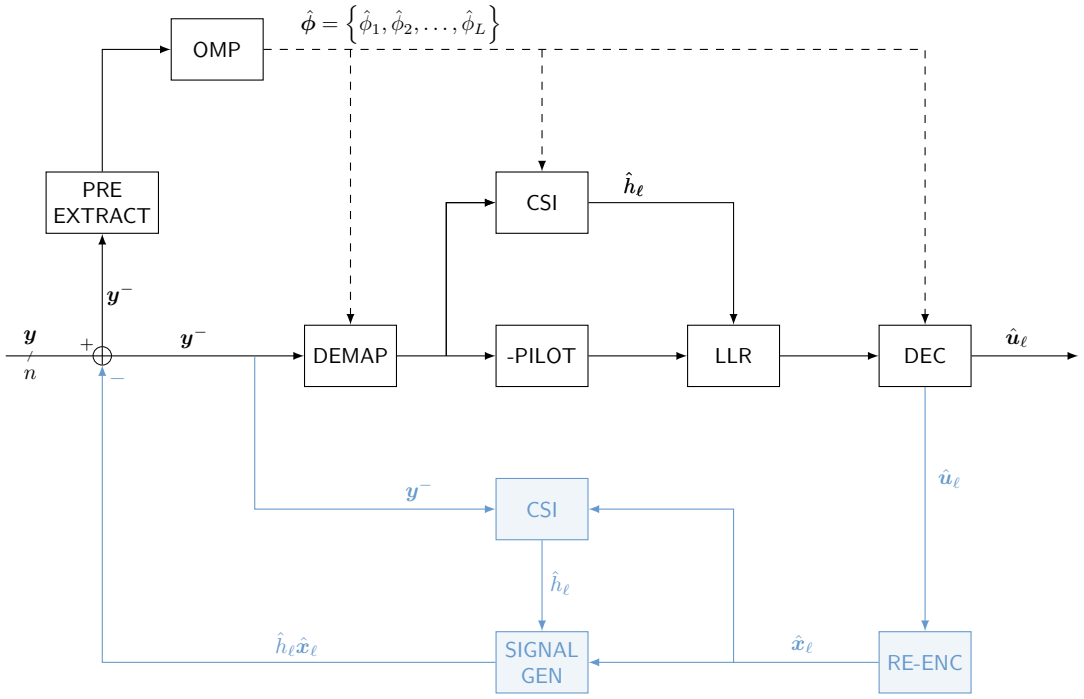


Figure 7. Schematic description of the SB-IDMA receiver chain.

Gaussian MAC Performance

A first set of results is provided for the Gaussian MAC in Figure 3. The configurations have been tailored for $n = 15000$ c.c.u.s (30000 real c.u.s), hence are directly comparable with the schemes whose performance is provided on the same chart. The details of the SB-IDMA configurations are summarized in Table 8. The configurations are somehow "extreme", in the sense that they employ POs of relatively small size (25 complex channel uses each). As a result, each transmission yields a relatively large number of segments ($n_s = 80$). Preambles and pilots have been drawn randomly, with symbols that are independently distributed according to a circularly-symmetric Gaussian distribution (see Box 1).

The performance of SB-IDMA with the $(500, 100)$ LDPC code (—○—) from the NR standard is with 1.8 dB from the achievability bound (——) up to a moderate-large number of active users, competing with some of the best schemes from literature. The gap from the bound can be dissected in two main components. A 0.6 dB loss is caused by the energy overhead introduced by the preamble. The remaining 1.2 dB loss can be attributed to the LDPC code: that is the gap displayed by the $(500, 100)$ 5GNR LDPC code, at a block error rate of 5×10^{-2}, with respect to the RCU bound of [30] over the single-user AWGN channel. The configuration that employs the 5GNR polar code (—□—) exploits the fact that the polar code rate can be further lowered, with respect to the LDPC code. A rate-1/10 polar code is hence used. The performance over the overall scheme results in a gain of about 0.8 dB over its LDPC-based counterpart, operating within 1 dB from the RCU up to 80 active users.

Quasi-Static Fading MAC Performance

Table 8. Parameters used for SB-IDMA simulations – Gaussian MAC.

Parameter	LDPC-based	Polar-based	Unit
Information bits per user	100	100	bits
Frame length (n)	15000	15000	c.c.u.
# POs (N)	589	589	-
PO size (n_{PO})	25	25	c.c.u.
Preamble length	275	275	c.c.u.
# Preambles	2048	2048	-
Channel code	5GNR LDPC (BG2)	5GNR Polar (CRC-11)	-
Block length (n_c)	500	1000	bits
Modulation	QPSK	QPSK	-
Repetition rate (d)	8	4	-
# Segments (n_s)	80	80	-

A second set of results was obtained over the quasi-static fading MAC (Figure 8). The parameters used for the simulations are given in Table 9, and are closely aligned to the 2SRA OTO configuration of Table 6. The performance of the LDPC-based (—△—) and of the polar-based (—△—) scheme shows a remarkable gain over the various 2SRA configurations (including the non-standard extension using 1024 preambles). A saturation of the number of supported users, that happens around $160 \div 200$ users in the polar code case, can still be observed, with a two-fold improvement over the 2SRA OTO mapping with 1024 preambles (- □ -). The saturation is here caused by the limited number of preambles, and can be largely improved by enabling a even larger preamble set. However, as observed for the 2SRA case, the number of channel uses allocated to the PRACH (278) limits the number of preambles that can be supported under OMP detection. To enlarge the preamble set, longer preambles can be used. On the same chart, the performance of an SB-IDMA configuration that trades the number of POs ($N = 59$) for longer preambles (12×139 c.c.u.s) is shown (- △ -). To limit the energy overhead caused by the longer preamble, a power back-off of 10 dB is applied to the preambles. The resulting performance allows to support more than 260 active users, without displaying signs of saturation.

A final set of results, aiming at assessing the effect of multiple antennas at the base station, is provided in Figure 9. Here, the number of receiver antennas is moderate, and set to $N_{\text{RX}} = 2$ and to $N_{\text{RX}} = 4$. The configurations used for the simulations are the ones described in Table 6 and in Table 9. The results confirm all the insights gathered in the previous simulation setups.

Table 9. Parameters used for SB-IDMA simulations – Quasi-static fading MAC.

Parameter	LDPC-based	Polar-based	Unit
Information bits per user	100	100	bits
Frame length (n)	19478	19478	c.c.u.
# POs (N)	64	64	-
PO size (n_{PO})	300	300	c.c.u.
Preamble length	278	278	c.c.u.
# Preambles	1024	1024	-
Channel code	5GNR LDPC (BG2)	5GNR Polar (CRC-11)	-
Block length (n_c)	500	500	bits
Modulation	QPSK	QPSK	-
Repetition rate (d)	3	3	-
# Segments (n_s)	3	3	-
# pilot symbols per PO	50	50	-

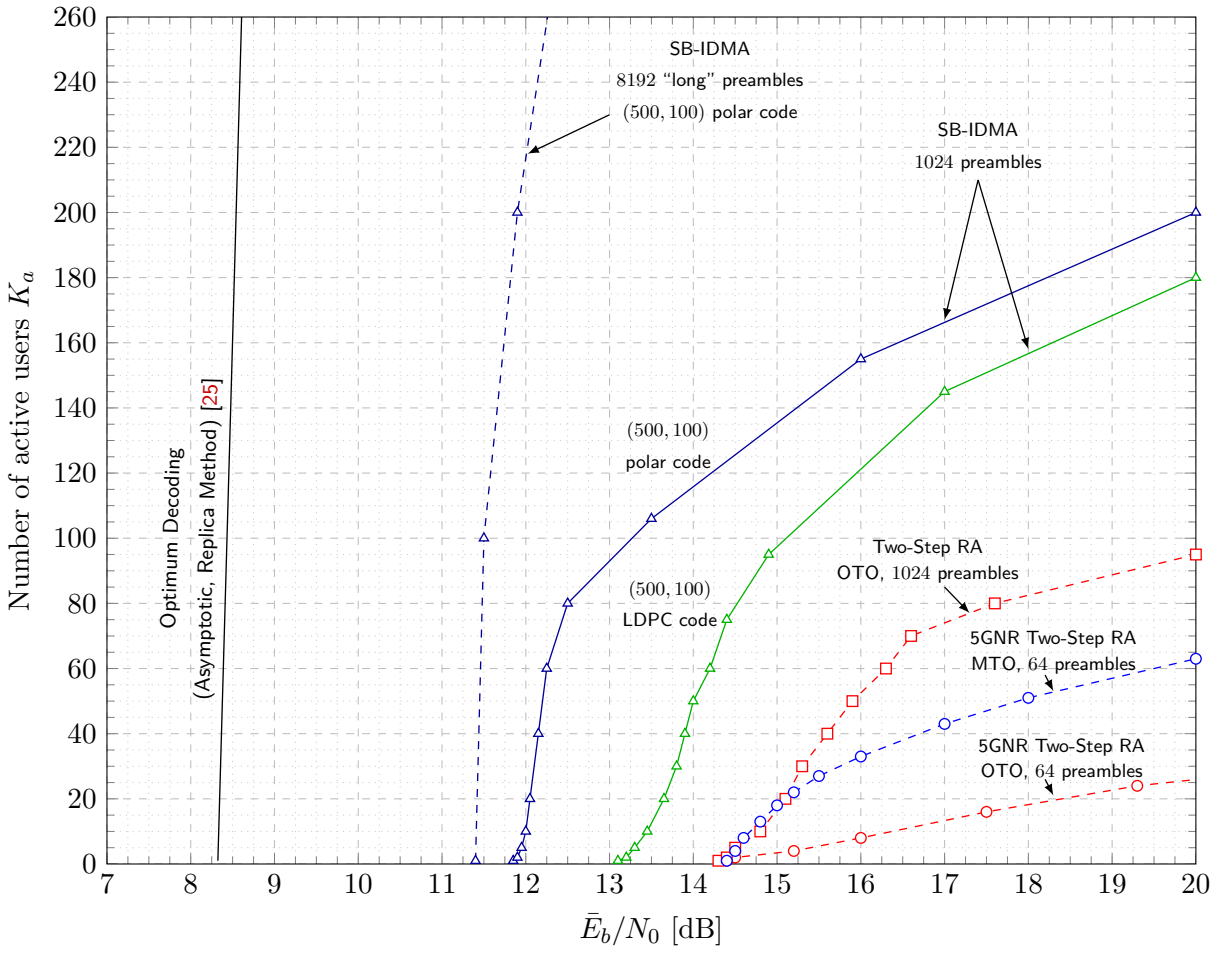


Figure 8. Number of supported active users vs. average SNR for a target PUPE $= 10^{-1}$. Quasi-static Rayleigh fading channel, $n \approx 20000$ channel uses. Single antenna at the base station.

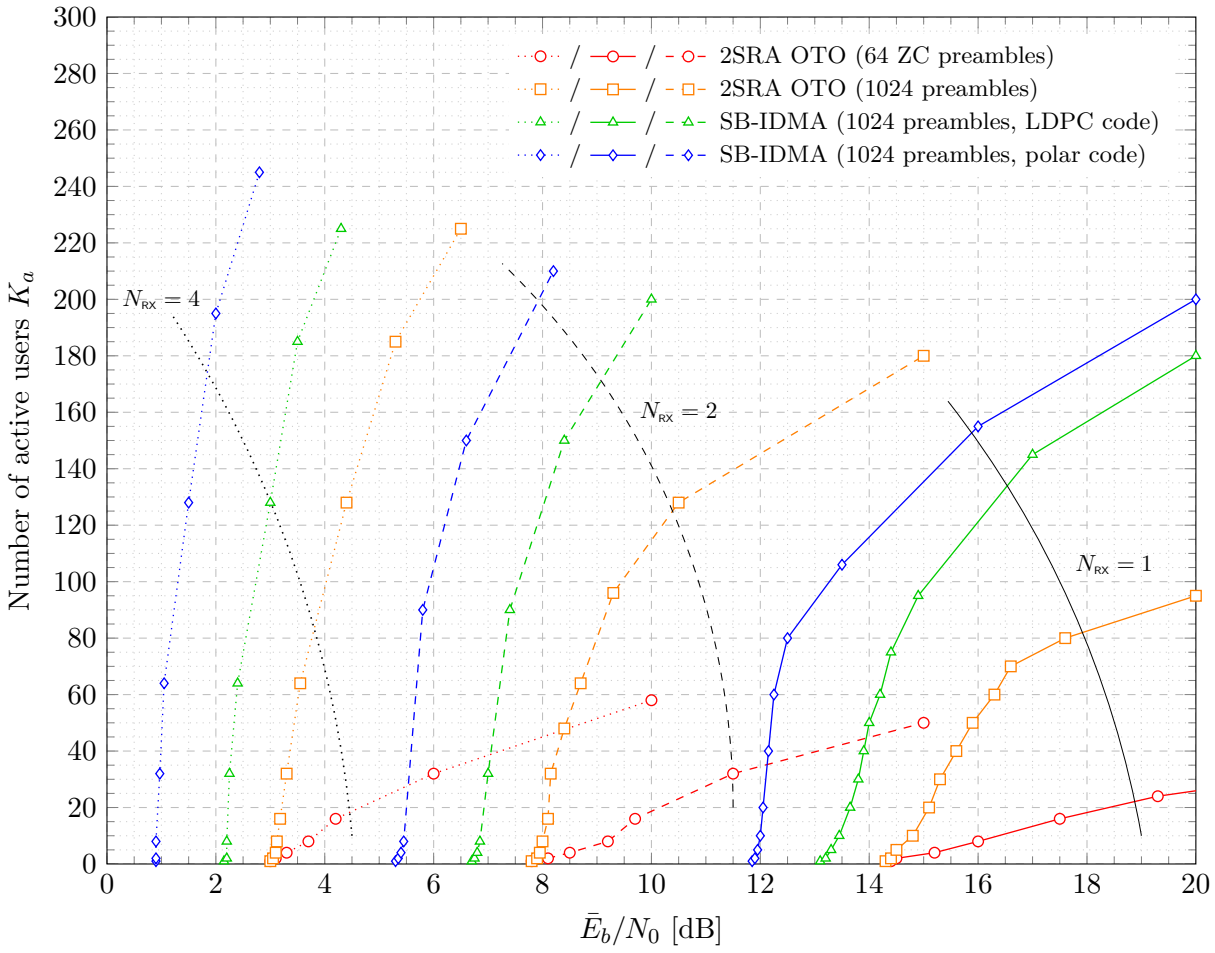


Figure 9. Number of supported active users vs. average SNR for a target $\mathsf{PUPE} = 10^{-1}$. Quasi-static Rayleigh fading channel, $n \approx 20000$ channel uses. Multiple antennas at the base station.

5 Conclusions

This report provided a first investigation on possible improvements of the random access channel in future evolutions of the 3GPP standards. The analysis addressed the design of massive random access schemes for grant-free access, taking as reference the initial component of the 2SRA protocol included in the 5GNR specification. The study identified three directions that may yield substantial improvements to the existing protocol, namely

1. The extension of the set of preambles to be used over the PRACH, together with the introduction of preamble-dependent pilot fields;

2. The definition of a more flexible use of POs, e.g., by allowing transmission to take place over several POs.

3. The adoption of polar codes, when short data packets have to be transmitted.

The identified improvement points have been embodied in a new scheme – SB-IDMA – which requires relatively minor updates of the specification, while delivering sizable gains over both Gaussian and fading MACs.

Acknowledgments

The authors would like to thank William Powell, Thomas Nitsche, Basavana Rudra, Christian Mensing and Reiner Stuhlfauth from Rohde und Schwarz for the useful discussions on the 5GNR 2SRA.

References

[1] *5GNR: Medium Access Control (MAC) protocol specification*, 3rd Generation Partnership Project (3GPP) Std. TS 138.321, Rev. 16.1.0, Jul. 2020.

[2] *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E- UTRA); Multiplexing and channel coding (Release 9)*, 3rd Generation Partnership Project (3GPP) Std. TS 136.212, Rev. 8.3.0, May 2008.

[3] *LTE: Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification*, 3rd Generation Partnership Project (3GPP) Std. TS 136.300, Rev. 13.2.0, Aug. 2016.

[4] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory*, 2017.

[5] E. Marshakov, G. Balitskiy, K. Andreev, and A. Frolov, "A Polar Code Based Unsourced Random Access for the Gaussian MAC," in *Proc. Vehicular Technology Conference*, Honolulu, HI, USA, Sep. 2019.

[6] A. K. Pradhan, V. K. Amalladinne, A. Vem, K. R. Narayanan, and J.-F. Chamberland, "Sparse IDMA: A Joint Graph-Based Coding Scheme for Unsourced Random Access," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7124–7133, Nov. 2022.

[7] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, "A coded compressed sensing scheme for unsourced multiple access," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6509–6533, Oct. 2020.

[8] A. Fengler, P. Jung, and G. Caire, "SPARCs for unsourced random access," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6894–6915, Oct. 2021.

[9] A. K. Pradhan, V. K. Amalladinne, K. R. Narayanan, and J.-F. Chamberland, "Polar coding and random spreading for unsourced multiple access," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2020.

[10] E. Casini, R. D. Gaudenzi, and O. del Rio Herrero, "Contention Resolution Diversity Slotted ALOHA (CRDSA): An Enhanced Random Access Scheme for Satellite Access Packet Networks." *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1408–1419, Apr. 2007.

[11] G. Liva, "Graph-Based Analysis and Optimization of Contention Resolution Diversity Slotted ALOHA," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[12] E. Paolini, G. Liva, and M. Chiani, "Coded slotted aloha: A graph-based method for uncoordinated multiple access," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6815 – 6832, Dec. 2015.

[13] *5GNR: Physical layer procedures for data*, 3rd Generation Partnership Project (3GPP) Std.

TS 138.214, Rev. 17.6.0, 2023.

[14] aa., "Mcs / tbs / code rate in a nutshell," https://www.sharetechnote.com/html/5G/5G_MCS_TBS_CodeRate.html, 2024, accessed: 2024-05-06.

[15] *5GNR: Physical channels and modulation*, 3rd Generation Partnership Project (3GPP) Std. TS 138.211, Rev. 17.7.0, Apr. 2024.

[16] A. Chakrapani, "On the design details of ss/pbch, signal generation and prach in 5g-nr," *IEEE Access*, vol. 8, pp. 136 617–136 637, 2020.

[17] J. Kim, G. Lee, S. Kim, T. Taleb, S. Choi, and S. Bahk, "Two-Step Random Access for 5G System: Latest Trends and Challenges," *IEEE Netw.*, vol. 35, no. 1, pp. 273–279, Jan./Feb. 2021.

[18] E. Peralta, T. Levanen, F. Frederiksen, and M. Valkama, "Two-Step Random Access in 5G New Radio: Channel Structure Design and Performance," in *Proc. Vehicular Technology Conference*, Apr. 2021.

[19] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, US, Nov. 1993.

[20] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[21] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.

[22] B. Li, H. Shen, and D. Tse, "An Adaptive Successive Cancellation List Decoder for Polar Codes with Cyclic Redundancy Check," *IEEE Commun. Lett.*, vol. 16, no. 12, p. 2044–2047, Dec. 2012.

[23] T. Richardson and S. Kudekar, "Design of low-density parity check codes for 5G new radio," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 28–34, 2018.

[24] E. Dahlman, S. Parkvall, and J. Sköld, *5GNR The Next Generation Wireless Technology*. Academic Press, 2018.

[25] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient coded random access for the wireless uplink," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4694–4708, Aug. 2020.

[26] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications," in *Proc. 1970 Fall Joint Computer Conference*, vol. 37, 1970, pp. 281–285.

[27] L. G. Roberts, "ALOHA packet system with and without slots and capture," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 5, no. 2, p. 28–42, Apr. 1975.

[28] V. Bioglio, C. Condo, and I. Land, "Design of polar codes in 5G New Radio," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 29–40, 2020.

[29] M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533–547, Sep. 1981.

[30] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

Acronyms and Notation

List of Acronyms

2SRA	two-step random access
4SRA	four-step random access
5GNR	5G New Radio
AWGN	additive white Gaussian noise
BP	belief propagation
BS	base station
CCS	coded compressed sensing
c.c.u.	complex channel use
CP	cyclic prefix
CRDSA	contention resolution diversity slotted Aloha
CSA	coded slotted Aloha
c.u.	channel use
IDMA	interleaver division multiple access
IoT	Internet of Things
IRSA	irregular repetition slotted Aloha
LDPC	low-density parity-check
LLR	log-likelihood ratio
LTE	Long Term Evolution
MAC	multiple access
MMSE	minimum mean square error
MPR	multi-packet reception
MTC	machine-type communication
MTO	many-to-one
NB-IoT	Narrowband IoT
NR	new radio
OFDM	orthogonal frequency-division modulation
OMP	orthogonal matching pursuit
OTO	one-to-one
PBCH	physical broadcast channel
PDSCH	physical downlink shared channel
PO	PUSCH occasion
PRACH	physical random access channel
PRB	physical resource block
PSS	primary synchronization signal
PUPE	per-user probability of error
PUSCH	physical uplink shared channel
QPSK	quadrature phase shift keying
RA	random access
RCU	random coding union

SB-IDMA	sparse block interleaver division multiple access
SCL	successive cancellation list
SCS	sub-carrier spacing
SIC	successive interference cancellation
SNR	signal-to-noise ratio
SPARC	sparse regression code
SSS	secondary synchronization signal
TA	time advance
TBS	transport block size
TIN	treat-interference-as-noise
UT	user terminal
UMAC	unsourced multiple access

Notation Summary

$\mathcal{C}$	Error correcting code
E_b	Energy per information bit
k	Number of information bits per active user
K	Total number of users
K_a	Number of active users
n	Frame length
n_s	Number of segments (SB-IDMA)
N	Number of PUSCH occasions
N_0	Single-sided noise power spectral density
n_c	Block length of the error correcting code
n_{PO}	number of channel uses in a PO
n_{PRE}	Preamble length in channel uses
N_{RX}	Number of antennas at the base station
PUPE	Per-user probability of error