

Massive Random Access With Trellis-Based Codes and Random Signatures

A. Korhan Tanc¹, *Member, IEEE*, and Tolga M. Duman², *Fellow, IEEE*

Abstract—We investigate unsourced and grant-free massive random access in which all the users employ the same codebook, and the basestation is only interested in decoding the distinct messages sent. To resolve the colliding user packets, a novel approach relying on user transmissions with random-like amplitudes selected from a large number of possible signatures is proposed. The scheme is combined with convolutional coding for error correction. The receiver operates by first identifying the signatures used by the transmitting nodes employing a sparsity-based detection algorithm, and then utilizing a trellis-based decoding algorithm. Despite its simplicity, the proposed solution offers excellent performance.

Index Terms—Convolutional codes, machine-type communications, multiple access, random signatures.

I. INTRODUCTION

A DISTINCTIVE feature of new generation wireless communication systems is that they support a huge number of nodes communicating with the same basestation [1]. The concept is closely related with wireless sensor networks with potential applications in environmental monitoring, logistic management, medical care, smart cities, and so on [2], [3]. In this concept, however, the traffic between the nodes and the basestation is unusual in the sense that the nodes are only sporadically active and channel access is not coordinated. Furthermore, the basestation is only interested in the transmitted messages as opposed to the information on who transmitted what.

With the aim of approaching the information theoretic limits of massive random access [4], different coding/signaling schemes have been proposed. In a pioneering work [5], a practical scheme based on serially concatenated coding and list decoding is introduced. In this scheme, while the inner code converts the access channel to a binary adder channel, a linear outer code provides error correction. Using the dispersion of the resulting equivalent channel, the required signal to noise ratio for a given packet error probability is computed. A more sophisticated scheme based on low density parity check (LDPC) codes and compressed sensing (CS) is presented in [6]. In this CS based study, along with an identical LDPC code, the users employ interleavers determined by some part of the message bits, making it possible to decode colliding

user transmissions through an iterative decoding procedure. Furthermore, the user packets are transmitted more than once as in irregular repetition slotted ALOHA (IRSA) [7], and successive interference cancellation is employed. Another CS based solution is proposed in [8] where a divide-and-conquer approach is utilized in decoding. Utilization of Tanner graphs with CS approach is also introduced [9]. Approximate message passing (AMP), which is an effective tool for inverse problems, is employed for massive random access in [10]. In addition, IRSA and chirp signaling based frameworks are developed in [11] and [12], respectively.

In this letter, our main motivation is to develop a simple, practical and well-performing coding/signaling scheme for massive random access. Specifically, we propose to use a set of random signatures (whose lengths match the codeword lengths), which effectively convert the unfavorable equal gain multiple access channels into “random-like” changing channels, introducing “diversity.” When transmitting, each user selects a signature at random from this set, and transmits its codeword by scaling the coded/modulated bits by the corresponding elements of the signature sequence. While different coding techniques can be used with this approach, due to its simplicity, we adopt convolutional codes. In other words, all the users employ the same convolutional code as dictated by the unsourced nature of transmissions, however, the receiver is helped by the fact that with high probability they will employ different signatures. We show that it is possible for the receiver to identify the signature sequences employed during a particular transmission using an all-ones pilot and sparsity based recovery. Once the utilized random amplitude sequences are determined, we execute a super-trellis based decoding for multi-user detection to recover the distinct messages. We also discuss optimization of the distribution from which the random amplitudes can be selected. We combine this scheme with the simple ALOHA based transmission (similar to [5]), and demonstrate via an extensive set of simulations that our proposed solution, despite employing very simple convolutional codes, outperforms the benchmark [5] in terms of the required energy per bit.

The rest of the letter is organized as follows. In Section II, we describe the proposed system model. In Section III, we present our decoding approach and formulate optimization of the distribution from which the random signatures are selected. We present an extensive set of numerical examples and demonstrate the superiority of the newly proposed simple solution in Section IV. Finally, we conclude the letter in Section V.

II. PROPOSED TRANSMISSION SCHEME

We consider a similar set up as in [5]. We assume that K_a active users are interested in sending their packets to a common basestation without any coordination. The communication

Manuscript received November 12, 2020; revised December 4, 2020; accepted December 20, 2020. Date of publication January 5, 2021; date of current version May 6, 2021. This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under the grant 119E589. The associate editor coordinating the review of this letter and approving it for publication was X. Jiang. (Corresponding author: A. Korhan Tanc.)

A. Korhan Tanc is with Vitrociset Belgium, 2201 Noordwijk, The Netherlands (e-mail: a.tanc@vitrocisetbelgium.com).

Tolga M. Duman is with the Department of Electrical and Electronics Engineering, Bilkent University, 06800 Ankara, Turkey (e-mail: duman@ee.bilkent.edu.tr).

Digital Object Identifier 10.1109/LCOMM.2021.3049195

1558-2558 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

is frame oriented, and each frame is divided into slots. Each user selects one slot per frame uniformly randomly and transmits its codeword. Hence, the number of user allocations per slot follows a binomial distribution. The activity of a user is only known to itself, not to the other users or the basestation. We represent frame and slot lengths as n and \tilde{n} bits, respectively. We also denote the number of slots per frame as $N = \lfloor n/\tilde{n} \rfloor$.

We construct a set of signature sequences, i.e., codeword-long sequences of amplitude levels. We denote the selected signature by the i th user by s_i , and the entire set of signatures by the columns of a matrix $\mathbf{S} \in \mathbb{R}^{\tilde{n} \times S_a}$, S_a being the number of signatures used. While different constructions are possible, we assume that these signatures are picked randomly (following a certain probability distribution for each element), and the set of signatures are made available to each transmitter and to the receiver. Each transmitting user selects one of the S_a signatures based on the first $k' = \log_2 S_a$ bits of its message, and the remaining $k - k'$ message bits are transmitted after encoding by a channel code.

With $k - k'$ bits to be transmitted in each codeword, all the users employ the same convolutional code with rate $(k - k')/(\tilde{n} - p)$, and the coded bits are appended with a length- p all-one sequence used as a common pilot. The coded bits are then converted into binary phase-shift keying (BPSK) symbols and scaled with the user-specific sequence of random amplitudes (based on the selected signature sequences). The transmitted signal vector for user i is given by $2s_i \circ c_i - s_i$, where s_i and c_i represent the random amplitude sequence and the coded bits appended with pilot bits, respectively, and the operation “ \circ ” refers to an elementwise product of two vectors of the same length. Assuming unit channel gains for all the users, the received signal at the j th slot is given by

$$\mathbf{y}_j = \sum_{i=1}^{K_a} \mathbf{w}_{ij} \circ (2s_i \circ c_i - s_i) + \mathbf{z}_j, \quad j = 1, 2, \dots, N. \quad (1)$$

Here $\mathbf{w}_{ij} \in \mathbb{B}^{\tilde{n}}$ (where $\mathbb{B} = \{0, 1\}$) and $\mathbf{z}_j \in \mathbb{R}^{\tilde{n}}$ are the indicator and additive white Gaussian noise (AWGN) vectors, respectively. If user i transmits in slot j , \mathbf{w}_{ij} is the all-one vector. Otherwise, it is the all-zero vector.

The receiver operates on a slot by slot basis. If there is an excessive number of transmissions in a given slot (beyond the error correction capabilities of the coding scheme), the transmissions in this slot would be non-decodable, hence the slot is abandoned. If the number of transmissions is below a certain (predetermined) number, the decoder attempts the decoding process. The decoder estimates the signature sequences employed in a given slot by using the all-one pilot sequences. Once the active signatures are determined, a combined trellis, determined from the convolutional code with labels obtained from the signature sequences, is employed for decoding. Details are given in Section III.

Along with possibly excessive number of collisions in a given slot, another source of error is the possible collisions among the signatures used. When the users select the same signature, it is not possible to decode their packets successfully. If L users are transmitting in a particular slot, the signature collision probability is simply $1 - S_a^{-L} S_a! / (S_a - L)!$, which can be made small with a sufficiently large S_a . However, it

should also be clear that increasing S_a will increase the pilot overhead and complexity, hence there is a trade-off.

III. RECEIVER STRUCTURE AND SIGNATURE OPTIMIZATION

A. Detection of Signatures Used

In each slot, the receiver first determines the specific signatures from the set \mathbf{S} by employing the received signal corresponding to the pilot bits. That is, for slot j , the received signal for the last p symbols is written as

$$\tilde{\mathbf{y}}_j = \tilde{\mathbf{S}}\mathbf{b}_j + \tilde{\mathbf{z}}_j, \quad j = 1, 2, \dots, N, \quad (2)$$

where $\tilde{\mathbf{S}}$ is the last p rows of the \mathbf{S} matrix, and $\mathbf{b}_j \in \mathbb{B}^{S_a}$ is a binary vector. The locations of non-zero elements of \mathbf{b}_j indicate the active columns of \mathbf{S} . Hence, successful recovery of \mathbf{b}_j leads to an accurate estimation of the non-zero $\mathbf{w}_{ij} \circ s_i$ vectors. Let L_j be the number of active users in slot j . When S_a is large enough, i.e., when $S_a \gg \max_j L_j$, \mathbf{b}_j will be sparse, and the recovery problem is formulated as

$$\hat{\mathbf{b}}_j = \underset{\mathbf{b}_j}{\operatorname{argmin}} \|\tilde{\mathbf{y}}_j - \tilde{\mathbf{S}}\mathbf{b}_j\|_2^2 + \lambda \|\mathbf{b}_j\|_0, \quad \mathbf{b}_j \in \mathbb{B}^{S_a}. \quad (3)$$

Here $\|\cdot\|_2$ and $\|\cdot\|_0$ are the Euclidean and L_0 norms, respectively. In order to recover \mathbf{b}_j , we resort to the computationally efficient basis pursuit denoising (BPD) algorithm [13]. Once \mathbf{b}_j is estimated using BPD, we apply an instance of thresholding such that L_j largest elements of $\hat{\mathbf{b}}_j$ are assigned to one, and other elements of $\hat{\mathbf{b}}_j$ are assigned to zero, respectively.

B. Multi-User Decoding

Once the non-zero $\mathbf{w}_{ij} \circ s_i$ vectors are estimated, a joint multi-user decoding takes place. In the presence of equal channel gains and AWGN, the decoder minimizes the cost function

$$f_j = \left\| \mathbf{y}_j - \sum_{i=1}^{K_a} \mathbf{w}_{ij} \circ (2s_i \circ \hat{c}_i - s_i) \right\|_2. \quad (4)$$

In order to implement this decoder efficiently, we utilize a Viterbi Algorithm with $2^{(K-1)L_j}$ states where K is the constraint length of the user-common binary convolutional encoder.

We note that it is possible to simplify the joint decoder using standard techniques. For instance, instead of Viterbi Algorithm, one can employ sequential decoding approaches, such as Fano or stack decoding, or other reduced complexity search algorithms such as M -algorithm that have been widely applied for trellis based codes [14]. In addition, as another approach, instead of employing a super-trellis structure with $2^{(K-1)L_j}$ states, L_j component decoders can be employed where each component decoder has $2^{(K-1)}$ states and component decoders are connected via a state node.

C. Signature Set Optimization

In principle, it is possible to conduct a performance analysis for convolutional codes over a multiple access channel as it is feasible to consider a product trellis for pairs of codewords, and determine a weight distribution [15], to apply the union

bound. With this approach it could be possible to perform code optimization. On the other hand, this computation is highly complex, hence we take an alternate approach and conduct a simple performance analysis based on the first-term approximation of the union bound for high signal-to-noise ratios (SNRs). For $L = 1$ and random amplitude $\theta > 0$, the bit error probability for the coded case can be bounded as

$$P_b \leq \int_0^\infty \alpha Q\left(\frac{\beta\theta}{\sqrt{2N_0}}\right) p(\theta) d\theta. \quad (5)$$

Here Q is the usual right tail probability of standard Gaussian, and $p(\theta)$ is the probability distribution function (PDF) of θ , respectively. We note that α and β are positive constants whose values are determined according to the specific channel code used [14]. For $L > 1$, the received signal at the joint decoder follows a pulse amplitude modulation (PAM) constellation with 2^L elements. Under the assumption of independent and identically distributed (i.i.d.) random signatures, we generalize the result in (5) to the multi-user case as

$$P_b \leq \int_0^\infty \int_0^\infty \dots \int_0^\infty \tilde{\alpha} Q\left(\frac{\tilde{\beta} g(\theta_1, \theta_2, \dots, \theta_L)}{\sqrt{2N_0}}\right) \times p(\theta_1) p(\theta_2) \dots p(\theta_L) d\theta_1 d\theta_2 \dots d\theta_L. \quad (6)$$

Here, $g(\theta_1, \theta_2, \dots, \theta_L)$ is one half of the minimum distance for the uncoded case, and $\tilde{\alpha}$ and $\tilde{\beta}$ values are determined according to the code structure. Since calculation of (6) is still intractable, we resort to maximization of the expectation

$$J = \int_0^\infty \dots \int_0^\infty g(\theta_1, \dots, \theta_L) p(\theta_1) \dots p(\theta_L) d\theta_1 \dots d\theta_L.$$

For example in the case of $L = 2$ with $\theta_1 \geq \theta_2$, the effective constellation points are $\theta_1 + \theta_2$, $\theta_1 - \theta_2$, $-(\theta_1 + \theta_2)$, $-(\theta_1 - \theta_2)$, similar to 4-ary PAM. In this case, half of the minimum distance is $g(\theta_1, \theta_2) = \min\{\theta_2, \theta_1 - \theta_2\}$. In fact, maximizing J is equivalent to the minimization of the bit-error probability for 2^L -ary PAM, and the above can be considered as constellation design for a multi-user communication scenario.

In order to select a suitable PDF, we experiment with the general beta distribution defined on the interval $\theta \in (0, 1)$,

$$p(\theta) = \frac{\Gamma(2a)}{\Gamma^2(a)} \theta^{a-1} (1 - \theta)^{a-1}, \quad (7)$$

where a is the shaping parameter, and Γ is the Gamma function, respectively [16]. We note that, for the special case of $a = 1$, uniform distribution is obtained. We shift the distribution to the right by θ_{min} and scale it so that $E\{x^2\} = 1$. For various values of θ_{min} , a and L , the resulting J values obtained via numerical evaluations are shown in Table I. As shown in the table, increasing θ_{min} improves the expectation J for $L = 1$ towards the limiting value of 1. However, for $L = 3$, the best value of J is obtained with $\theta_{min} = 0.1$. In addition, θ_{min} values of 0.1 and 0.2 result in the best J values for $L = 5$. Note that it may also be advantageous to employ signatures with corresponding J values smaller than the maximum due to possibly reduced pilot overhead requirements. In order to take such signatures into the account properly, the design can be carried out through FER simulations.

TABLE I
J VALUES FOR VARIOUS VALUES OF θ_{min} , a AND L

θ_{min}	a	$L = 1$	$L = 3$	$L = 5$
0	0.5	0.816	0.079	0.012
	1	0.866	0.108	0.016
	2	0.913	0.116	0.017
0.1	0.5	0.862	0.110	0.015
	1	0.901	0.118	0.017
	2	0.937	0.113	0.017
0.2	0.5	0.892	0.110	0.015
	1	0.924	0.113	0.016
	2	0.953	0.108	0.016
0.5	0.5	0.943	0.095	0.014
	1	0.961	0.103	0.016
	2	0.976	0.095	0.015

TABLE II
EXPERIMENTED TRELLIS BASED CODES

Code	Rate	Polynomial	Puncturing	\tilde{n}	N
C ₁	1/3	(17, 15, 13)	No	346	86
C ₂	1/2	(17, 15)	No	252	119
C ₃	2/3	(17, 15)	Yes	205	146

D. Other Coding Schemes

We highlight that while we are only considering trellis based codes in this letter, the idea of using different signature sequences can be adopted with any other coding scheme as well. For instance, LDPC codes, turbo codes or polar codes can also be used. With the conversion of the channel into a more favorable *i.i.d.* fading scenario, these more sophisticated coding schemes combined with IRS is expected to improve the system performance significantly.

IV. SIMULATION RESULTS

We now present several examples using the setup in [5] with $k = 100$, $n = 30000$ and frame error rate (FER) per user of 0.05. We select $S_a = 512$ and $p = 64$. We experiment with off-the-shelf 8-state convolutional codes given in Table II. In this table, \tilde{n} is calculated as follows: each user selects a signature using the first 9 bits of its message. Hence, 91 bits are encoded using the convolutional code. In addition, three all-zero bits are appended to the uncoded bits for trellis termination, and length- p all-one pilot bits are appended to the coded bits for signature recovery, respectively. For the generation of signatures, we choose $\theta_{min} = 0.1$ and $a = 1$ as Table I suggests. In addition, we set $\lambda = \sqrt{N_0/50}$, and we run the BPD algorithm with an augmented Lagrangian parameter of 0.1 and 100 iterations [13].

We select a threshold T , and discard the received signal for a slot if the number of colliding packets is more than this threshold value. If the number of colliding codewords $L_j \leq T$, then the detection takes place. Regarding our codes, we select $T = 3$ for $K_a = 50$, and $T = 5$ for $K_a = \{100, 150, 200, 250\}$, respectively. In Fig. 1, the average achieved E_b/N_0 values at FER = 0.05 are provided in comparison with the benchmark schemes. Fig 1 shows that, for all the values of K_a we have used, the proposed scheme performs better than the scheme in [5]. Despite its simplicity, the E_b/N_0 gain attained with the proposed approach is as much as 5 dB

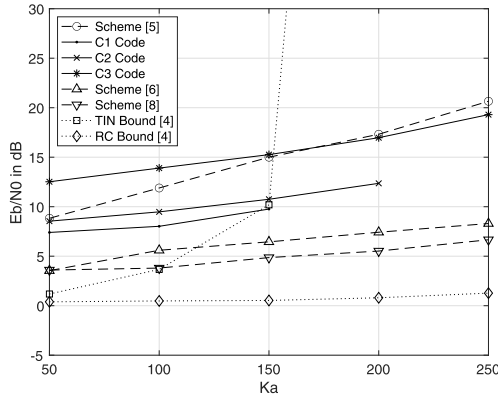
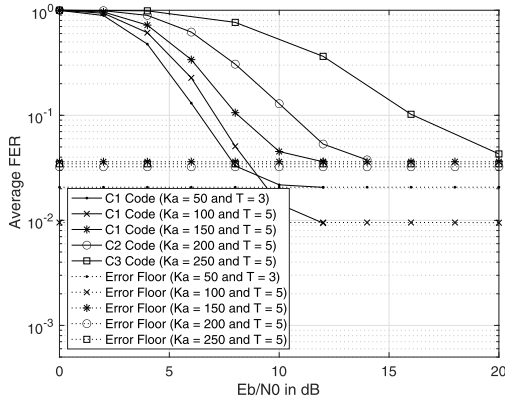
Fig. 1. Achieved E_b/N_0 values at FER = 0.05.

Fig. 2. Average FER results.

for $K_a \in \{150, 200\}$. For $K_a \in \{150, 200, 250\}$, the proposed scheme also overcomes the TIN approach. It is also noteworthy that our results are based on explicit simulations with low complexity, while those in [5] are based on finite length information theoretic bounds.

Although the performance of our scheme is worse than those of schemes [6] and [8], our scheme has both latency and complexity advantages. While the latency of our scheme is in the order of slot size, the latencies of schemes [6] and [8] are both in the order of frame size. In addition, schemes [6] and [8] both utilize successive interference cancellation (SIC) and iterative decoding which increase the decoding complexity considerably. In Fig. 2 we also show the detailed performance of our scheme along with the error floors accounting for the signature collisions. We observe that convergence is achieved which implies the success of signature estimation.

In our investigations, we have also experimented with the use of amplitude levels that remain identical for the entire codeword which was explored before in the context of IRSA and polar coding [17]. This approach converts the equal-gain multiple access channel into a *quasi-static* like fading multiple access channel, and it is possible to resolve the colliding packets, however, the energy per bit performance with convolutional coding is not competitive. We note that as a future research direction, generalization of the proposed random signature approach for the case of unknown (and unequal) channel gains can be considered. In this case, these gains can be estimated using clustering techniques [18], and by a proper modification of (4). Employing more sophisticated

channel codes with random signatures and IRSA is also a promising research direction.

V. CONCLUSION

We propose the use of random-like signatures along with simple convolutional codes for massive random access, and demonstrate that this novel approach is both very simple and implementable, and it is superior to the benchmark scheme in terms of the required E_b/N_0 . We note that the proposed solution of employing random-like signatures can also be used with the other previously reported approaches such as those employing more sophisticated channel coding with IRSA as well.

REFERENCES

- [1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [2] P. Sarigiannidis, T. Zygiridis, A. Sarigiannidis, T. D. Lagkas, M. Obaidat, and N. Kantartzis, "Connectivity and coverage in machine-type communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.
- [3] S. K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 426–471, 1st Quart., 2020.
- [4] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2523–2527.
- [5] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access Gaussian channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2528–2532.
- [6] A. Vem, K. R. Narayanan, J.-F. Chamberland, and J. Cheng, "A user-independent successive interference cancellation based coding scheme for the unsourced random access Gaussian channel," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8258–8272, Dec. 2019.
- [7] 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.
- [8] 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.
- [9] A. Pradhan, V. Amalladinne, A. Vem, K. R. Narayanan, and J.-F. Chamberland, "A joint graph based coding scheme for the unsourced random access Gaussian channel," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
- [10] A. Fengler, P. Jung, and G. Caire, "SPARCs and AMP for unsourced random access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 2843–2847.
- [11] E. Marshakov, G. Balitskiy, K. Andreev, and A. Frolov, "A polar code based unsourced random access for the Gaussian MAC," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Honolulu, HI, USA, Sep. 2019, pp. 1–5.
- [12] R. Calderbank and A. Thompson, "CHIRUP: A practical algorithm for unsourced multiple access," *Inf. Inference, J. IMA*, vol. 9, no. 4, pp. 875–897, Dec. 2020.
- [13] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.
- [14] C. B. Schlegel and L. C. Perez, *Trellis and Turbo Coding: Iterative and Graph-Based Error Control Coding*. Hoboken, NJ, USA: Wiley, 2015.
- [15] M. Dabirnia, S. Sharifi, A. K. Tanc, and T. M. Duman, "Short block length trellis-based codes for interference channels," *IET Commun.*, vol. 12, no. 18, pp. 2282–2289, Nov. 2018.
- [16] A. L. Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*. Upper Saddle River, NJ, USA: Pearson, 2008.
- [17] 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. (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–6.
- [18] K. Andreev, E. Marshakov, and A. Frolov, "A polar code based TIN-SIC scheme for the unsourced random access in the quasi-static fading MAC," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Los Angeles, CA, USA, Jun. 2020, pp. 3019–3024.