

A Joint PHY and MAC Layer Design for Coded Random Access with Massive MIMO

Lorenzo Valentini, Marco Chiani, Enrico Paolini

CNIT/WiLab, DEI, University of Bologna, Italy

Email: {lorenzo.valentini13, marco.chiani, e.paolini}@unibo.it

Abstract—Grant-free access schemes are candidates to support future massive multiple access applications owing to their capability to reduce control signaling and latency. As a promising class of grant-free schemes, coded random access schemes can achieve high reliabilities also with uncoordinated transmissions and therefore in presence packet collisions. In this paper, an analysis tool for coded random access, based on density evolution, is proposed and exploited for system design and optimization. In sharp contrast with the existing literature, where such tools have been developed under simplified channel assumptions, the proposed tool captures not only MAC layer features, but also the physical wireless fading channel and a realistic physical layer signal processing based on multiple antennas and randomly-chosen orthogonal pilots. Theoretical results are validated by comparison with symbol-level Monte Carlo simulations.

I. INTRODUCTION

The ever-growing spread and pervasiveness of machine-type communications and of the Internet of Things has recently boosted the attention towards massive multiple access (MMA) problems, where a massive number of devices, each with a sporadic but unpredictable activity, transmit short packets to a common base station (BS) [1]–[3]. A main target for MMA protocols is to achieve a very high scalability, defined as the number of simultaneously active users that the system can support, in presence of latency and reliability constraints [4]–[6]. In this context, grant-free multiple access schemes (as opposed to grant-based ones) can drastically reduce control signalling for connection establishment, bringing benefits in terms of both scalability and latency.

A class of grant-free medium access control (MAC) schemes that has attracted an increasing interest in the past few years, both in satellite and terrestrial applications, is that of coded random access (CRA) schemes [7], [8]. Based on a very simple uncoordinated access mechanism on the device side and on successive interference cancellation (SIC) across different slots on the BS side, they are intrinsically related to iterative decoding of codes on sparse graphs. Instances of CRA schemes include contention resolution diversity slotted ALOHA (CRDSA) [9], irregular repetition slotted ALOHA (IRSA) [10], and coded slotted ALOHA (CSA) [11].

While several works on CRA have addressed the collision channel model owing to its amenability to analysis, more recently the focus has been shifted to wireless channels models with realistic physical (PHY) layer processing. In this respect, an important direction of investigation focuses on CRA in contexts where the receiver is a BS with a massive number of

antennas (e.g., [3], [12]–[14]). This interest is justified by the possibility to exploit massive multiple input multiple output (MIMO) to implement forms of multi-packet reception (MPR), meaning that multiple packets can be correctly decoded in a single slot with a consequent substantial boost in scalability.

To achieve a high scalability under tightening reliability constraints, CRA schemes necessitate of a careful design. For example, with reference to the IRSA scheme, this typically means finding an optimum probability distribution for the packet repetition rate employed by the generic active user. The solution to this problem benefits from the bridge with codes on sparse graphs, so that it is usually tackled by tools developed in coding theory, like density evolution [15]. So far, the design of CRA schemes is well-established over channel models, such as the collision channel [10], [11] or its MPR extension [16], where the communication process over the PHY channel as well as signal processing at PHY layer are modeled in a very simple, often idealized, way.

In this paper we aim at addressing this problem by extending the density evolution analysis of IRSA to a more realistic setting. More specifically, we consider IRSA over a Rayleigh block fading channel with a massive MIMO BS and power control and we show that it is possible to tailor density evolution analysis, and consequent system optimization, to this channel model and to a specific PHY signal processing based on randomly-chosen orthogonal pilots for channel estimation and maximal ratio combining (MRC). Notably, the results are easily extendable to the more general CSA setting. The main contributions of the paper can be summarized as follows:

- We extend the simple collision channel to the case where orthogonal resources are available within each slot.
- We generalize CRA density evolution analysis to capture both PHY processing and Rayleigh fading.
- We show that IRSA distributions that are optimum over the collision channel may turn sub-optimum when a realistic setting is considered.

The paper is organized as follows. Section II introduces preliminary concepts, the system model, and some background material. Section III addresses density evolution generalizations. Numerical results are provided in Section IV. Finally, conclusions are drawn in Section V.

II. PRELIMINARIES AND BACKGROUND

We assume that the time is organized in MAC frames, each composed by N_s slots. In a frame time, K_a active users

(with K_a random and unknown to the receiver) contend for transmission of one information packet each. Considering an IRSA access protocol, the generic user sends r replicas of its packet (or “bursts”) in r randomly-chosen slots. For all users, r is generated randomly according to the same probability distribution. As remarked later, although we stick to IRSA, the developed results have a broader applicability to CRA schemes.

A. Density Evolution over the Collision Channel

This section reviews density evolution equations for IRSA over the collision channel [10]. The system can be represented by a bipartite graph where K_a user nodes, also known as burst nodes (BNs), are connected with N_s slot nodes, also referred to as sum nodes (SNs). A BN corresponding to an active user sending r packet replicas has r edges towards the r SNs associated with the chosen slots. The repetition degree r is a random variable with probability generating function (PGF) $\Lambda(x) = \sum_r \Lambda_r x^r$, independently drawn by each active user. Then, the probability that an edge is connected to a degree- r BN, λ_r , is given by

$$\lambda_r = \frac{\Lambda_r r}{\Lambda'(1)}. \quad (1)$$

On the other hand, a SN has c edges representing the number of users that have selected the corresponding slot to transmit a packet replica. We denote by ρ_c the probability that an edge is connected to a SN of degree c ; this is given by

$$\rho_c = \frac{\Psi_c c}{\sum_h \Psi_h h} \quad (2)$$

where Ψ_c is probability that c users performed a transmission in the considered slot.

The PGF $\Lambda(x)$, or BN degree distribution (or simply IRSA distribution), is under the control of the system designer. This is indeed not the case for the SN degree distribution $\Psi(x) = \sum_c \Psi_c x^c$, that is fully defined by the system load G and by the average burst repetition rate, $\sum_r r \Lambda_r = \Lambda'(1)$. In particular, for a large users' population size K , it is licit to assume that the number of transmissions in a slot follows a Poisson distribution. Specifically, we can write

$$\Psi_c = \frac{(G \Lambda'(1))^c}{c!} \exp(-G \Lambda'(1)) \quad (3)$$

which also yields

$$\rho_c = \frac{(G \Lambda'(1))^{c-1}}{(c-1)!} \exp(-G \Lambda'(1)). \quad (4)$$

The collision channel assumptions can be summarized as:
Assumption 1: If the number of arrivals in a slot is larger than one, then the receiver is unable to successfully decode any of these bursts.
Assumption 2: If there is only one arrival in a slot, then the burst is successfully decoded with zero error probability and perfectly subtracted from the slot.
Assumption 3: Whenever a burst is successfully decoded in a slot, the interference generated by each of its replicas can be perfectly subtracted from the corresponding slots.

Under these assumptions, let $q_\ell^{(r)}$ be the probability that an edge, connected to a degree- r BN, is unknown at the end of SIC iteration ℓ . Let us also define $p_\ell^{(c)}$ as the probability that an edge, connected to a degree- c SN, is unknown at the end of SIC iteration ℓ . We can exploit the edge-oriented probabilities λ_r and ρ_c to define the average probabilities q_ℓ and p_ℓ as

$$q_\ell = \sum_r \lambda_r q_\ell^{(r)} \quad \text{and} \quad p_\ell = \sum_c \rho_c p_\ell^{(c)}. \quad (5)$$

Next, consider a degree- r BN. An edge is revealed whenever any of the other edges connected to the same BN has been revealed. Thus, the probability that a burst has not yet been retrieved by the “MAC layer repetition code” is

$$q_\ell^{(r)} = p_{\ell-1}^{r-1}. \quad (6)$$

Similarly, consider a degree- c SN. An edge is revealed whenever all the other edges have been revealed due to the collision channel assumptions. Hence, the probability that a burst in a slot has not yet been cancelled by SIC is

$$p_\ell^{(c)} = 1 - (1 - q_\ell)^{c-1}. \quad (7)$$

Combining equations from (4) to (7) leads to the density evolution recursion over the collision channel in the form

$$p_\ell = 1 - \exp\left(-G \sum_r r \Lambda_r p_{\ell-1}^{r-1}\right) \quad (8)$$

with initial condition $p_0 = 1$ (there are no revealed edges at the beginning of the process). The asymptotic packet loss rate at SIC iteration ℓ is given by

$$Q_\ell = \sum_r \Lambda_r p_\ell^r. \quad (9)$$

Finally, the asymptotic load threshold of an IRSA distribution $\Lambda(x)$ is defined as

$$G^* = \sup\{G > 0 : Q_\ell \rightarrow 0 \text{ as } \ell \rightarrow \infty\} \quad (10)$$

where we note that $Q_\ell \rightarrow 0$ if and only if $p_\ell \rightarrow 0$. Here, “asymptotic” refers to the fact that density evolution recursion assumes statistical independence of messages along the edges of the graph, a condition that is met in the limit as $K \rightarrow \infty$, $N_s \rightarrow \infty$, and the ratio K/N_s being constant.

Remark. We highlight that, in density evolution recursion of CRA protocols, q_ℓ depends on the MAC layer, while p_ℓ on the channel model and PHY layer processing. For this reason, the results developed in this paper, aimed at evaluating p_ℓ over a realistic channel model and PHY processing, are amenable to extension to other MAC CRA protocols such as CSA [11].

B. Scenario Definition

The availability of a BS with a massive number of antennas is a key feature to enable MPR (i.e., decoding of multiple packets in the same slot even without SIC across slots) at the receiver. A simple approach to achieve MPR is represented by use of orthogonal pilots in combination with massive MIMO processing [3], [13], [14]. Since in MMA applications the number of users is typically extremely large compared with

the number of available pilots, N_P , it is however not possible to pre-assign one specific pilot to each user. Rather, each active user may pick one pilot randomly from the set of N_P orthogonal ones, without any coordination with the other users (hence, possibly yielding pilot collisions in a slot). In combination with IRSA, this approach was dubbed “coded pilot random access” in [3].

Next, we review the PHY layer model and processing adopted in this paper, which is the one also considered in [13], [14]. The model may be summarized as follows:

- The channel is a Rayleigh block fading channel with power control and coherence time equal to the slot time.
- The receiver has M antennas, each with independent fading coefficient per user.
- An active user picks, for each burst, a pilot uniformly at random from a set of N_P orthogonal ones for channel estimation purposes.
- Quadrature phase-shift keying (QPSK) modulation with Gray mapping and hard decision is employed.
- The payload is composed by N_D symbols and protected using a Bose–Chaudhuri–Hocquenghem (BCH) code able to correct up to t errors per codeword.
- The receiver performs a “channel hardening-based” SIC processing, as described in [3], [13], [14].

Then, the received signal in a slot may be expressed as $[\mathbf{P}, \mathbf{Y}] \in \mathbb{C}^{M \times (N_P + N_D)}$ where

$$\mathbf{P} = \sum_{k \in \mathcal{A}} \mathbf{h}_k \mathbf{s}(k) + \mathbf{Z}_p \quad \text{and} \quad \mathbf{Y} = \sum_{k \in \mathcal{A}} \mathbf{h}_k \mathbf{x}(k) + \mathbf{Z}. \quad (11)$$

In (11), \mathcal{A} is the set of users transmitting a burst in the considered slot; $\mathbf{h}_k = (h_{k,1}, \dots, h_{k,M})^T \in \mathbb{C}^{M \times 1}$ is the k -th user channel coefficient vector, whose elements are independent and identically distributed (i.i.d.) random variables with distribution $\mathcal{CN}(0, \sigma_h^2)$ for all $k \in \mathcal{A}$ due to power control. Without losing generality it is assumed $\sigma_h^2 = 1$. Moreover, $\mathbf{s}(k) \in \mathbb{C}^{1 \times N_P}$ and $\mathbf{x}(k) \in \mathbb{C}^{1 \times N_D}$ are the pilot sequence and payload transmitted by user k . Finally, $\mathbf{Z}_p \in \mathbb{C}^{M \times N_P}$ and $\mathbf{Z} \in \mathbb{C}^{M \times N_D}$ are matrices of Gaussian noise samples.

The processing is split into two phases. In phase 1, all slots are processed in order. In each slot, the BS first attempts channel estimation for all possible pilots by computing $\phi_j \in \mathbb{C}^{M \times 1}$ and consequently attempting payload estimation, for all $j \in \{1, \dots, N_P\}$, as

$$\phi_j = \frac{\mathbf{P} \mathbf{s}_j^H}{\|\mathbf{s}_j\|^2} = \sum_{k \in \mathcal{A}^j} \mathbf{h}_k + \mathbf{z}_j \quad \text{and} \quad \hat{\mathbf{x}}_j = \frac{\phi_j^H \mathbf{Y}}{\|\phi_j\|^2}. \quad (12)$$

Here, \mathcal{A}^j is the set of active users employing pilot j in the current slot, $\mathbf{s}_j \in \mathbb{C}^{1 \times N_P}$ is the j -th pilot sequence, and $\mathbf{z}_j \in \mathbb{C}^{M \times 1}$ is a noise vector.

If a generic user ℓ is the only one picking pilot j in the currently processed slot ($\mathcal{A}^j = \{\ell\}$) and if the number of antennas M is large enough, then $\hat{\mathbf{x}}_j$ approximates the user’s payload $\mathbf{x}(\ell)$. Upon successful channel decoding performed on $\hat{\mathbf{x}}_j$, the packet symbols are stored in a buffer waiting for the SIC phase.

In phase 2, SIC is performed across slots relying on information about replica number and positions along with the employed pilots for each of them retrieved in the decoded payload. This can be implemented, for example, by letting the number of replicas, their positions, and their pilot indexes be functions of the random message. By subtracting interference, it is possible that new users can be found re-attempting channel decoding procedure. Then, whenever a new user is successfully decoded its information is stored and phase 2 iterates until the buffer contain no more users. More details on the SIC and channel model, here omitted for space reasons, are available in [3], [13], [14].

III. DENSITY EVOLUTION WITH REALISTIC PHY LAYER

In this section we develop a density evolution analysis for IRSA over the realistic channel model and PHY processing that was reviewed in Section II-B. This analysis is presented in Section III-B. Before that, in Section III-A, we extend the above-reviewed density evolution analysis over the collision channel to the case of a collision channel featuring N_P orthogonal resources per slot. To keep a clean and compact notation, we denote the probability that a random variable A takes the value a , $\Pr(A = a)$, as $P(a)$. Similarly, we write $P(a, b|c)$ to indicate the probability $\Pr(A = a, B = b | C = c)$, and $P(\mathcal{E})$ to denote the probability that an event \mathcal{E} occurs.

A. Collision Channel with Orthogonal Resources

Consider a simple extension of the collision channel in which N_P orthogonal resources are available in each slot of the frame. In IRSA over this channel, an active user generates its repetition rate r according to the distribution $\Lambda(x)$, chooses r slots uniformly at random (without replacement), then chooses one resource per slot uniformly at random, and finally transmits r replicas of its packet into these slot-resource pairs. For this channel, Assumptions 1-3 listed in Section II-A remain valid, upon simply replacing “slot” with “slot-resource pair”. Note that the collision channel corresponds to $N_P = 1$.

To extend IRSA density evolution equations to this channel, we initially observe that ρ_c in (4), q_ℓ in (5), and $q_\ell^{(r)}$ in (6) remain unchanged as they only depend on the IRSA multiple access protocol. To update the expression of $p_\ell^{(c)}$, consider a slot with c burst arrivals. Consider any such burst and let H , $0 \leq H \leq c - 1$, be a random variable representing the number of interferers not yet recovered in the slot at the current iteration ℓ . Defining the failure event $\mathcal{F}_\ell = \{\text{The replica of an edge is not yet decoded at iteration } \ell\}$, we have

$$\begin{aligned} p_\ell^{(c)} &= P(\mathcal{F}_\ell | c) = \sum_{h=0}^{c-1} P(h) P(\mathcal{F}_\ell | c, h) \\ &= \sum_{h=0}^{c-1} \binom{c-1}{h} q_\ell^h (1 - q_\ell)^{c-1-h} \left[1 - \left(1 - \frac{1}{N_P} \right)^h \right] \\ &= 1 - \left(1 - \frac{q_\ell}{N_P} \right)^{c-1}. \end{aligned} \quad (13)$$

where the factor $[1 - (1 - 1/N_P)^h]$ is the probability that at least one of the h interferers collides with the considered

burst in the same resource. Repeating the steps reviewed for the collision channel, we obtain density evolution recursion

$$p_\ell = 1 - \exp\left(-\frac{G}{N_P} \sum_r r \Lambda_r p_{\ell-1}^{r-1}\right). \quad (14)$$

From (14) and (8) we deduce that, over the collision channel with N_P orthogonal resources per slot: (i) the asymptotic threshold G^* of a distribution $\Lambda(x)$ is N_P times larger than the threshold over the collision channel for the same distribution; (ii) for a given average number of replicas $\Lambda'(1)$, the distribution that maximizes the threshold G^* over the collision channel remains optimum also on the collision channel with N_P orthogonal resources per slot.

B. MIMO Block Fading Channel with Realistic Processing

In this section we extend density evolution equations to the scenario described in Section II-B. Referring to the collision channel assumptions, pointed out in Section II-A and generalized in Section III-A to $N_P > 1$, we note that Assumption 1 still holds due to the hypothesis of power control. On the other hand, Assumptions 2 and 3 do not hold any more because of the realistic payload and channel estimations. As the PHY processing and interference cancellation in a slot are captured by (7), next we focus on $p_\ell^{(c)}$ under realistic assumptions.

In density evolution, the probability $p_\ell^{(c)}$ is associated with the occurrence of the event that a packet replica, arriving in a slot where c users have transmitted (i.e., with $c-1$ interfering bursts), is not successfully decoded in that slot at iteration ℓ . From a user's viewpoint, we can therefore define the random variable l , representing the number of other interfering bursts that have picked the same pilot (pilot-colliding bursts). The random variable l is considered to account for the behaviour explicitly shown in [14] in which the probability to fail the decoding of a user packet depends on the initial number of user choosing the same pilot. We can write the probability that a user has exactly i pilot-colliders, given c total arrivals in the slot and N_P available pilots, as

$$P(i|c) = \binom{c-1}{i} \left(\frac{1}{N_P}\right)^i \left(1 - \frac{1}{N_P}\right)^{c-1-i}. \quad (15)$$

Since to successfully decode a burst within a slot it is necessary that it is the only one with a given pilot, we define the random variable S as the number of pilot-colliding bursts that have been subtracted so far. Then, considering that from previous interference cancellations bursts can be subtracted with probability $1 - q_\ell$, the probability that exactly s subtractions have been performed, given i pilot-colliding users, is

$$P(s|i) = \binom{i}{s} (1 - q_\ell)^s q_\ell^{i-s}. \quad (16)$$

Noting that $P(s|i, c) = P(s|i)$, we can write

$$p_\ell^{(c)} = \sum_i \sum_s P(\mathcal{F}_\ell|i, s, c) P(s|i) P(i|c). \quad (17)$$

Since it is not possible to successfully decode a pilot-collided replica due to the power control hypothesis, we have $P(\mathcal{F}_\ell|i, s, c) = 1$, for all $s \neq i$. On the other hand, when $s = i$ and a realistic channel is considered, there is a non-zero probability to successfully decode the burst. Using the approximation analytically derived in [14], we can write

$$P(\mathcal{F}_\ell|i, s, c) = \begin{cases} P_{\text{fail}}((i+1)c-1), & s = i \\ 1, & s \neq i \end{cases} \quad (18)$$

where

$$P_{\text{fail}}(n) = 1 - \sum_{d=0}^t \binom{N_D}{d} P_e^d(n) (1 - P_e(n))^{N_D-d} \quad (19)$$

and

$$P_e(n) = \text{erfc}\left(\sqrt{\frac{M}{2n}}\right) - \frac{1}{4} \text{erfc}^2\left(\sqrt{\frac{M}{2n}}\right). \quad (20)$$

It is possible to note that P_{fail} depends on the number of payload symbols N_D , the number of antennas M , the error correction capability t of the PHY channel code, and the number of interfering users. Finally, we can write

$$p_\ell^{(c)} = 1 - \left(1 - \frac{q_\ell}{N_P}\right)^{c-1} + \sum_{i=0}^{c-1} (1 - q_\ell)^i P_{\text{fail}}((i+1)c-1) P(i|c). \quad (21)$$

Using $p_\ell^{(c)}$ from (21) instead of (7) in the density evolution recursion, allows us evaluating the asymptotic threshold G^* of an IRSA distribution $\Lambda(x)$ over the realistic wireless channel and signal processing.

Remark. Equation (18) accounts for the fact that, in realistic settings, the above collision channel Assumptions 2 and 3 do not hold. In this respect we note that, letting $P(\mathcal{F}_\ell|i, s, c) = 1 - \delta_{i,s}$ where $\delta_{i,s}$ is the discrete delta function, it is easy to retrieve (13). This observation confirms that the procedure derived in this section can be seen as a “generalization” of the density evolution over collision channel.

IV. PERFORMANCE EVALUATION

As main outcomes of our analysis: (i) we use the developed analysis tool with differential evolution optimization [17] to derive optimal $\Lambda(x)$ distributions; (ii) we verify the consistency between the optimization results and Monte Carlo simulations (with both PHY and MAC layers); (iii) we show that an over-simplification of the PHY channel model may lead to a wrong design when adopted in a realistic system. Before presenting the results, we briefly review differential evolution and the PHY layer aware system adopted to verify the theoretical results, respectively in Section IV-A and Section IV-B.

A. Differential Evolution Optimization

We search for the optimal $\Lambda(x)$ distribution which maximizes G^* in (10). The optimization problem is constrained by imposing the average number of transmitted packets $\Lambda'(1)$,

related to the average energy spent per user packet. We solved the optimization problem using differential evolution [17], an evolutionary, population-based metaheuristic search algorithm used to find the global minimum of a real-valued function of a vector of continuous parameters. The main steps are similar to those of evolutionary optimization algorithms [18]. An initial population of vectors, each representing a $\Lambda(x)$ distribution, is first generated. A competitor for each population element is then constructed by mutation and crossover over the current vector population. Next, each population element is compared with its own competitor and only one is selected, resulting in an evolved population. The mutation, crossover and selection steps are iterated until a certain stopping criterion is fulfilled. Differential evolution was proposed for the optimization of low-density parity-check (LDPC) codes degree profile in [19] and for IRSA and CSA optimization in [10] and [11].

B. Simulation Setup

With reference to Section II-B, we consider a system where users transmit payloads encoded with an $(n = 511, k = 421, t = 10)$ binary BCH code. Part of the information bits are used to validate the decoded packets, for example by cyclic redundancy check (CRC). After zero padding the BCH codeword with a final bit, we map the encoded bits onto a QPSK constellation with Gray mapping, obtaining $N_D = 256$ symbols per codeword. Simulations have been carried out with symbol rate $B_s = 1$ Msps and $M = 256$ BS antennas. Imposing a maximum latency constraint of $\Omega = 50$ ms, we compute the number of slots per frame N_s as [13]

$$N_s = \left\lceil \frac{\Omega B_s}{2(N_P + N_D)} \right\rceil \quad (22)$$

where the number of orthogonal pilot symbols, N_P , equals the total number of available pilot sequences. These sequences are constructed using Hadamard matrices. In particular, we assume $N_P = 64$, resulting in a number of slots $N_s = 78$.

C. Numerical Results

We start by presenting numerical results that illustrate the accuracy of the proposed threshold analysis. To this aim, we ran Monte Carlo simulations for some IRSA distributions $\Lambda(x)$ over both the collision channel with orthogonal resources and the MIMO block fading channel with actual signal processing, and performed threshold analysis for the same distributions over these channels. In practice, we declared a value of G as achievable (i.e., $G < G^*$) when density evolution recursion yielded $Q_\ell < 10^{-4}$ after a sufficiently large number of iterations.

In Fig. 1 and Fig. 2 we report simulation results (in terms of packet loss rate versus the number of active users over the frame) in solid lines, while thresholds are marked by dashed vertical lines. In these figures, the “thresholds” are defined as $K_a^* = N_s G^*$, which represents an approximation of the number of simultaneously active users the scheme can support. Fig. 1 shows that the IRSA distribution with average packet repetition rate $\Lambda'(1) = 3$ and maximum repetition

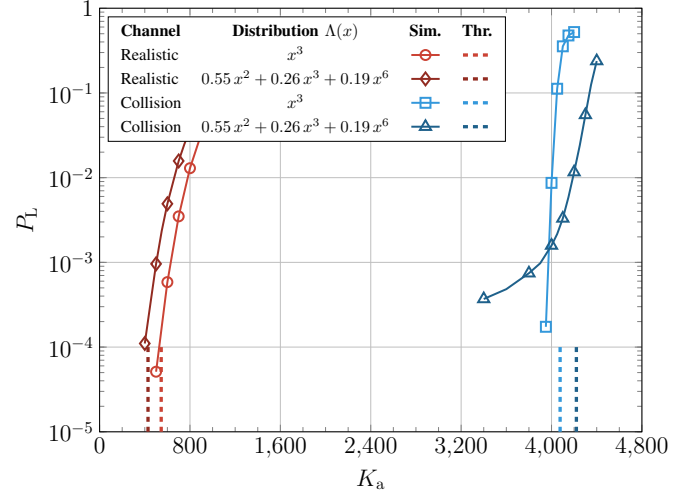


Fig. 1. Packet loss rate comparison between the IRSA distribution with $\Lambda'(1) = 3$ and maximum repetition degree 6 being optimal over the collision channel (with or without orthogonal resources) and the concentrated distribution $\Lambda(x) = x^3$. Channels: Collision channel with N_P orthogonal resources and MIMO block fading channel with realistic signal processing. Parameters: $N_P = 64$, $N_s = 78$, $M = 256$. Dashed lines: Values of $G^* N_s$.

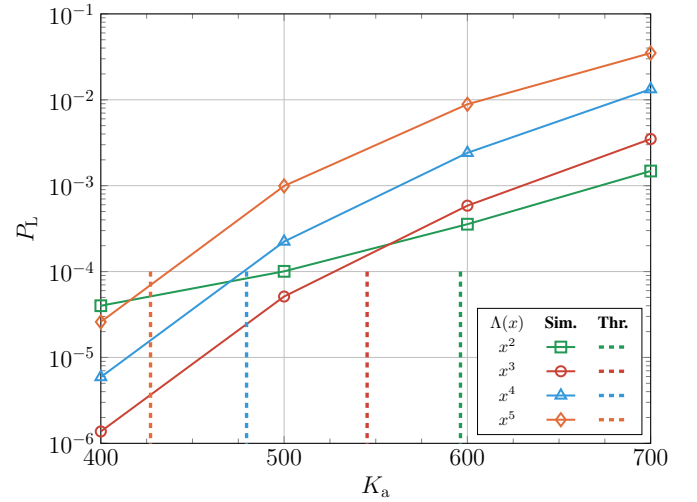


Fig. 2. Packet loss rate comparison between concentrated distributions characterized by different repetition rates $r \in \{2, 3, 4, 5\}$ over realistic channel. Solid: Monte Carlo simulations. Parameters: $N_P = 64$, $N_s = 78$, and $M = 256$. Dashed lines: $N_s G^*$.

rate 6, having the largest threshold over the collision channel model [11], becomes sub-optimal when the realistic channel and signal processing is considered. In fact, its threshold is outperformed by that of the distribution with a constant repetition rate $\Lambda(x) = x^3$ (that is, CRDSA with repetition rate 3). Very remarkably, as predicted by our threshold analysis, this result is in perfect agreement with the Monte Carlo simulation. Fig. 2 shows similar results for other distributions, which again reveal the effectiveness and reliability of the proposed analysis over massive MIMO block fading channels and realistic PHY layer processing. Note that all concentrated (CRDSA) distributions considered in Fig. 2 exhibit the best threshold constrained

TABLE I
ASYMPTOTIC THRESHOLDS OBTAINED THROUGH DENSITY EVOLUTION
UNDER REALISTIC CHANNEL ASSUMPTIONS.

IRSA Distribution	$\Lambda'(1)$	G^*
$\Lambda(x) = x^2$	2	7.64
$\Lambda(x) = x^3$	3	6.99
$\Lambda(x) = x^4$	4	6.15
$\Lambda(x) = x^5$	5	5.48
$\Lambda(x) = 0.55x^2 + 0.26x^3 + 0.19x^6$	3	5.49
$\Lambda(x) = 0.50x^2 + 0.50x^3$	2.5	6.64
$\Lambda(x) = 0.51x^2 + 0.27x^3 + 0.22x^8$	3.6	4.63
$\Lambda(x) = 0.55x^2 + 0.16x^3 + 0.29x^6$	3.3	4.97

to the corresponding integer $\Lambda'(1)$. Looking again at Fig. 2 we can see that, as expected, the proposed density evolution analysis is unable to capture error floor phenomena such as the one affecting the distribution $\Lambda(x) = x^2$. In Table I we list the thresholds estimated through density evolution for some $\Lambda(x)$ distributions with the previous choice of the system parameters.

Lastly, we show in Table II the results of another analysis we carried out using the proposed tool. For a constrained average repetition rate $\Lambda'(1) = 3$, we let the number of BS antennas M vary, searching for the optimum distribution (in terms of G^*) for each considered M . For all values of M , differential evolution optimization returned the same distribution $\Lambda(x) = x^3$. It is interesting to observe that, while the asymptotic threshold G^* increases monotonically with M , the ratio G^*/M (which represents a sort of efficiency per antenna) is not monotonically increasing but exhibits a maximum value. We attribute the decrease of G^*/M for large M to the constant number of orthogonal pilots $N_P = 64$.

V. CONCLUSIONS

The increasing interest on grant-free protocols for massive multiple access in next generation wireless networks requires reliable design tools in order to optimize and compare different solutions. We propose a joint PHY and MAC layer design targeting a massive MIMO system over Rayleigh fading channel. The design is presented step-by-step to facilitate extending the analysis to other PHY processing strategies and channel assumptions. As a main outcome of our analysis we show that IRSA distributions, optimum over the simple collision channel model, may turn suboptimum under a realistic setting.

ACKNOWLEDGEMENTS

This work has been carried out in the framework of the CNIT National Laboratory WiLab and the WiLab-Huawei Joint Innovation Center.

REFERENCES

- [1] L. Liu and W. Yu, "Massive connectivity with massive MIMO—Part I: Device activity detection and channel estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2933–2946, Mar. 2018.
- [2] L. Sanguinetti, A. A. D'Amico, M. Morelli, and M. Debbah, "Random access in massive MIMO by exploiting timing offsets and excess antennas," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6081–6095, 2018.

TABLE II
OPTIMUM ASYMPTOTIC THRESHOLDS CONSTRAINED TO $\Lambda'(1) = 3.0$
VERSUS THE NUMBER OF ANTENNAS M . OPTIMUM DISTRIBUTION
 $\Lambda(x) = x^3$ IN ALL CASES, $N_P = 64$, $t = 10$, $N_D = 256$.

M	G^*	G^*/M
8	0.1356	0.0169
16	0.4409	0.0276
32	1.0562	0.0330
64	2.0778	0.0325
128	3.8167	0.0298
256	6.9909	0.0273

- [3] J. H. Sørensen, E. De Carvalho, Č. Stefanović, and P. Popovski, "Coded pilot random access for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8035–8046, Dec. 2018.
- [4] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, Jun. 2013.
- [5] L. Liu, E. G. Larsson, W. Yu, P. Popovski, Č. Stefanović, and E. De Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, Sep. 2018.
- [6] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.
- [7] E. Paolini, Č. Stefanović, G. Liva, and P. Popovski, "Coded random access: Applying codes on graphs to design random access protocols," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 144–150, Jun. 2015.
- [8] M. Berlioli, G. Cocco, G. Liva, and A. Munari, "Modern random access protocols," *Foundations and Trends in Networking*, vol. 10, 2016.
- [9] E. Casini, R. De 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, no. 4, pp. 1408–1419, Apr. 2007.
- [10] 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.
- [11] 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.
- [12] E. De Carvalho, E. Björnson, J. H. Sørensen, P. Popovski, and E. G. Larsson, "Random access protocols for massive MIMO," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 216–222, May 2017.
- [13] L. Valentini, A. Faedi, M. Chiani, and E. Paolini, "Coded random access for 6G: Intra-frame spatial coupling with ACKs," in *Proc. 2021 IEEE Global Commun. Conf. Workshops*, Madrid, Spain, Dec. 2021.
- [14] —, "Impact of interference subtraction on grant-free multiple access with massive MIMO," in *Proc. 2022 IEEE Int. Conf. Commun.*, Seoul, South Korea, May 2022.
- [15] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.
- [16] Č. Stefanović, E. Paolini, and G. Liva, "Asymptotic performance of coded slotted ALOHA with multipacket reception," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 105–108, Jan. 2018.
- [17] R. Storn and K. Price, "Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [18] T. Bäck, D. B. Fogel, and Z. Michalewicz, "Handbook of evolutionary computation," *Release*, vol. 97, no. 1, p. B1, 1997.
- [19] A. Shokrollahi and R. Storn, "Design of efficient erasure codes with differential evolution," in *Proc. 2000 IEEE Int. Symp. Inf. Theory*, 2000.