IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 1 June 2022, accepted 4 July 2022, date of publication 14 July 2022, date of current version 20 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3190491

THEORY

Energy and Rate Allocation for Massive Multiple Access With Interference Cancelation

FRANCESC MOLINA, (Member, IEEE), JOSEP SALA-ALVAREZ, (Senior Member, IEEE), JAVIER VILLARES, (Senior Member, IEEE), AND FRANCESC REY, (Member, IEEE)

Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

Corresponding author: Francesc Molina (francesc.molina@upc.edu)

This work was supported in part by the Spanish Ministry of Science and Innovation through project RODIN (PID2019-105717RB-C22/AEI/10.13039/501100011033) and in part by Grant 2017 SGR 578.

ABSTRACT This article addresses the problem of energy and code allocation to many users accessing, under spreading-based nonorthogonal multiple access, a wireless node set up with a successive interference cancellation architecture aided by redundancy-check error control. As an application, we consider the asynchronous access of a delay-tolerant satellite system, where users employ finite-length channel codes and are subject to a known power unbalance induced by the known distribution of the channel's attenuation. The article develops, as a mathematically tractable approximation to massively populated systems, a unified framework to compute the best energy and code allocation rules that maximize the spectral efficiency of a network that handles asymptotically many users. Concretely, the presented approach circumvents the exponential complexity in the number of users when modeling the propagation of packet decoding failures through the receiver's decoding scheme. It also enables a deterministic analysis of the more complex features affecting the receiver, making the related performance optimization problem amenable to systematic tools from differential and variational calculus. The derived expressions evidence the most favorable three-way unbalance between energy, rate, and reliability for receiver performance. Low-level system simulations are carried out for validation.

INDEX TERMS Massive multiple access, successive interference cancellation, satellite, energy and code allocation, asymptotic analysis.

I. INTRODUCTION

Large-scale heterogeneous networks of fully-connected devices with human-less interaction constitute a driver for the next generation wireless systems. In this context, the data traffic generated by multiple devices is characterized by many sporadic transmissions of short bit-payload, demanding massive connectivity and high throughput, among others [1]. Satellite communication networks are re-emerging as promising candidates to manage this addition to global network traffic. At the same time, they solve the problem of ubiquitous access in Earth areas where the current terrestrial infrastructure does not operate [2]–[4].

A problematic aspect of the satellite radio link is its large propagation delay, which causes the inefficient operation of, for instance, the demand assignment protocol adopted in the Digital Video Broadcasting standard [5], the sensing-based techniques in [6], and the orthogonal schemes active in terrestrial networks [7]–[9]. To date, the best (performance-wise) techniques that circumvent the above drawback have released the strict coordination between transmitters and the receiver in terms of the time/frequency resources used to assess the channel [10]. The successful idea has been to enable nonorthogonal multiple access (NOMA) and adopt a multiuser receiver to resolve destructive packet collisions. This scheme enables the autonomous operation of multiple devices in exchange for increasing the complexity of the receiver, which must adopt an algorithm for detecting many users [11], [12]. Such is the case for the slotted [13] or

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

the asynchronous [14] versions of the contention resolution diversity ALOHA protocol, which transmit many copies of the same packet and rely on an interference cancellation (IC) based demodulator. Another even more promising candidate to solve the massive multiple access problem is Enhanced Spread Spectrum ALOHA (E-SSA) [15], which adopts spreading-based NOMA and a bank of matched filters in front of the successive IC (SIC) receiver; a feasible solution in terms of computational complexity in the number of users, rather than the optimal detector [16], [17]. Remarkably, both protocols have in common the adoption of IC schemes whose decoding performance is improved substantially under packet-power unbalance. They have been duly adopted in numerous settings due to their high performance and complexity-constrained operation [18]–[21]. Thus, the investigation of the best power unbalance required by the iterative receiver deserves special attention. For the sake of clarity, we will use the term *unbalance* to refer to the variation of a specific magnitude among different users.

Precisely, this work gathers the best allocation designs to mitigate the impact of short-packet collisions on the system performance of massive access future networks into a unified theory. The objective is to drive performance optimization of networks operating under the E-SSA umbrella, extending the optimal unbalance to energy, rate, and reliability.

A. RELATED WORK

Information Theory constitutes the guideline for addressing the above goal [22]. The ultimate data rate that a SIC receiver can reliably achieve imposes a specific unbalance on the energy and rate transmitted by different users. Novel research findings on massive access future networks invalidate the straightforward applicability of previous results since future wireless networks will operate with short packets at moderate/low error rates [23]–[25]. The interest of the present work stems precisely from the reason above: an evident need to devise adequate allocation designs that, accounting for the receiver implementation specificities, bring the achievable performance closer to the theoretical bounds.

The adoption of a powerful SIC receiver to decode short packets has extended the mentioned unbalance to four magnitudes, totaling a four-way interplay between energy, rate, reliability, and blocklength (decoding latency). The majority of works have studied the interplay over two of them and for specific implementations highly sensitive to system performance [26]–[32]. The authors in [26] discuss the latency-reliability trade-off. In [27] and [28], users are subject to heterogeneous reliability constraints and energy expenditure is minimized by varying the blocklength and energy. Therein, a rough approximation for users' decoding success probabilities along SIC receiver stages is adopted. The same simplified model is adopted in [29]. Another relevant unbalance is between users' energies and rates at the receiver input. It has been analyzed mainly for latency-tolerant settings: generally, satellite applications operating under slotted ALOHA or E-SSA. In this respect, the best allocation design for an asymptotic number of nonorthogonal spread spectrum users demands equal coding rates and exponentially distributed symbol energies at the receiver input [30]. The analysis assumes an ideal SIC receiver and capacity-achieving codes, but it can be extended to coding schemes with abrupt packet error rate versus signal-to-interference-plus-noise ratio (SINR) curves. Anyhow, its interest for the satellite context is that disparate values in the best energy distribution can be more easily obtained by taking advantage of the large unbalance in the distribution of channel attenuation between the satellite and users [31], [32]. A practical implementation of the above result is that adopted in E-SSA [32], [33]. Therein, users employ the same coding scheme and transmission protocol, and the receiver adopts a powerful SIC receiver and a sliding window approach to deal with packet collisions. This approach aids the SIC receiver with redundancy-check error control and performs cancellation only when successful decoding occurs, which differs from the hard cancellation in [34], optimal if users are allocated to low packet error rates.

As a matter of fact, the rigorous validation of the allocation designs in [30] is put in doubt since the adopted receiver incorporates practical (realistic) features not contemplated therein. Then, pioneering allocation designs are reformulated to adapt them to more practical cases. The authors in [32] and [33] design, under some heuristics, a very competitive user-energy distribution for a coded modulation system common to all users and a per-user quality of service. The optimal energy distribution resulted roughly exponential. Other works [35], [36] consider simplified models for the demodulator adopted in E-SSA, assessing the benefits of a quasi-exponential distribution at the receiver input when users employ the same physical layer.

B. TECHNICAL CONTRIBUTIONS

This work develops a unified framework for designing the energy and rate allocation in a massive access wireless network where many users are decoded by a powerful SIC receiver. The work contributes to increasing the understanding of a complex SIC receiver by providing accurate system models of its behavior, and derives the best allocation rules for many users simultaneously accessing the common receiver. As a scenario of practical interest, we consider a satellite equipped with a single antenna serving static/low mobility terminals with directive antennas, and an open-loop control system broadcasting a known pilot from which users simply: estimate their individual channel gains accurately, and use them to aid the allocation function designed in this work. Research findings can be also applied to multi-beam satellites with a low inter-beam interference pattern.

The main contributions are listed below:

- We analyze a SIC receiver aided by redundancy-check error control; a relevant approach also adopted in the E-SSA protocol [32], whose randomness in decoding many

users challenges the computation of system performance: the exact contribution to performance of all possible combinations of packet decoding error/success events during SIC has exponential complexity in the number of users.

- We study the asymptotic regime as the number of users grows with the spreading gain. For this case, the receiver adopts the form of a dynamical system that we characterize through deterministic expressions; thus, circumventing the aforementioned exponential computational complexity.
- Taking advantage of the asymptotic analysis and tools from the calculus of variations, we investigate the best allocation designs as smooth functions of the user indexing, the per-user channel power gains, and the performance of employed encoders. This approach has revealed groundbreaking results to maximize network spectral efficiency. In contrast to the semi-analytic optimization in [33] or some discretizations in [30], this approach obtains explicit forms or implicit expressions that can be solved in linear time complexity through the proposed algorithms.
- For a variety of cases, we prove the existence of piecewise smooth allocation functions satisfying the stationary point equations of the spectral efficiency maximization problem. We fill the void in the existing literature [27]–[30], [33], [35]–[37] by considering several encoder sets with coding schemes achieving optimal first- and second-order coding rates, and by exploring the best unbalance in terms of users' energy, rate, and reliability. Remarkably, we extend previous works by studying in-depth the three-way unbalance between the above magnitudes. Research findings show that when the per-user channel power gains are balanced and users are constrained by fair reliability, they are allocated to the same SINR or remain silent. The fraction of silent users only depends on how second-order coding rates approach capacity. For other cases, the optimal allocation achieves such energy, rate and reliability unbalance that the asymptotic SINR cancellation factor (Φ in Section II-D) does not increase at each stage.
- Lastly, we assess the performance of the above study through a low-level implementation of a SIC receiver, achieving competitive results and affordable computation times for networks above a hundred users.

C. PAPER ORGANIZATION AND NOTATION

The rest of this paper is structured as follows: Section II describes the system model of the wireless scenario and the receiver operation at both the signal and SINR level; Section III performs system optimization, complemented by simulations in Section IV; and Section V states conclusions. The Appendices contain the more intricate derivations.

Regarding paper notation: $\mathbb{E}[\cdot]$ is the expectation operator. ∇_x denotes the gradient under variable x, and partial derivatives of $G(x, y)$ are indicated by G_x and G_y. Table 1 summarizes the notation for the main variables.

TABLE 1. Notation table for the main variables.

Variable	Definition
t	Time variable (non-italic font)
n, n_o, n_e	Packet, preamble and payload lengths (in symbols)
K	Number of users
N	Spreading gain
k	User index ($k \in \{1, \ldots, K\}$)
s_k	Packet transmitted by user k
$c_{k,m}(\mathrm{t})$	Spreading waveform for user k and symbol m
$E_{\mathrm{x}}[k]$	k-th user transmitted symbol energy
$E_{\mathrm{r}}[k]$	k-th user received symbol energy
$R[k]$	k-th user coding rate
$h_l[k]$	k-th user channel power gain from the l-th path
η	Average LOS power gains relative to multipath
$N_{\mathrm{t}}[k]$	Noise plus interference term at stage k of SIC
N_0	Power spectral density of thermal noise
$\xi_{\mathrm{prv}}[k]$	Interference from users ($i < k$) already processed
$\xi_{\mathrm{rem}}[k]$	Interference from users ($i > k$) yet to be processed
ξ_{m}	Random multipath
$\Gamma[k]$	k-th user SINR after symbol despreading
$\mathrm{PER}[\Gamma, R]$	Packet error rate versus SINR curve associated with the R-rate encoder
$\varepsilon[\Gamma, R]$	Residual energy versus SINR curve associated with the R-rate encoder

II. SYSTEM MODEL

We consider a large population of K low-throughput users, equipped with low/medium complexity terminals, accessing a common multiple access (MA) node.[1] Users transmit n-symbol packets that contain n_o preamble and n_e coded symbols. $s_k[0 \leq m < n]$ denotes the packet for user k, in which data has been first encapsulated next to a cyclic redundancy check (CRC) for error control and encoded thereafter. For encoding, users may avail themselves of an encoder from $\mathcal{M} \triangleq \{\mathcal{M}_1, \ldots, \mathcal{M}_p\}$; $\mathcal{M}_i$ maps b_i bits to n_e symbols, so the coding rate is $R_i = b_i/n_e$ bits/symbol. We denote the k-th user's coding rate and transmitted symbol energy $R[k]$ (bits per symbol) and $E_{\mathrm{x}}[k]$ (Joules per symbol). Although we use this general definition for user data packets, typical user data rates may be found in [32] and [33].

To comply with power-efficient transmission subject to limited peak-to-average power ratios and enable practically uncoordinated and reliable multiple access to many users, we deem the adoption of the same baseline in E-SSA [33] appropriate. Specifically, users expand the signal space by adopting nonorthogonal spreading waveforms with a large time-bandwidth product [16] and the receiver relies on SIC to separate users in the signal space. The baseband signal transmitted by user k is

$$x_k(\mathrm{t}) = \sqrt{E_{\mathrm{x}}[k]} \sum_{m=0}^{n-1} s_k[m] c_{k,m}(\mathrm{t} - mT). \tag{1}$$

[1] We assume a common configuration for an Internet of Things setting comprising many users that generate low data traffic because they transmit short bit-payload packets sporadically.

$c_{k,m}(\mathrm{t})$ is the unit-energy spreading waveform for user k and symbol m, built from a long pseudonoise sequence of period far exceeding the number of chips N in the symbol interval T (long code model) and a unit-energy chip pulse [16, eq. (2.18)]. The average modulus-squared cross-correlation over a symbol interval between spreading waveforms is inversely proportional to the spreading factor N as $1/N$ [17].

We investigate this system for a channel where the signal transmitted by each user is received from L paths. Accordingly, let $h[k]$ be the channel power gain between the k-th user and the MA node, comprising: (i) the line-of-sight (LOS) power gain $h_0[k]$, obtained from the attenuation at the k-th user's location with respect to the receiver's antenna broadside direction; and (ii) the multipath power gains $h_l[k]$ for $l = 1, \ldots, L-1$, independent random variables with expectations $\bar{h}_l[k] \triangleq \mathbb{E}[h_l[k]]$. Then, the received baseband signal at the MA node is

$$y(\mathrm{t}) = \sum_{k=1}^{K} \sum_{l=0}^{L-1} A_l[k] \sum_{m=0}^{n-1} s_k[m] c_{k,m}(\mathrm{t} - mT - \tau_l[k]) + w(\mathrm{t}), \tag{2}$$

with $A_l[k] \triangleq (E_{\mathrm{x}}[k] h_l[k])^{1/2} e^{j\varphi_l[k]}$ the k-th user received complex amplitude from the l-th path, and $\varphi_l[k]$ the corresponding random phase. $w(\mathrm{t})$ is white Gaussian noise.

Although the formulation presented herein is for a generic wireless channel with additive noise, we are mostly thinking of a satellite return link covering a large Earth area, where the channel's autocorrelation holds practically constant for some seconds [38]. This usually happens in scenarios with sufficiently directive antennas and high elevation angles. In this regard, we assume that the spreading gain and bandwidth are adjusted so that the packet duration does not exceed the channel coherence time.

A. PROBLEM STATEMENT

In the analyzed setting, multiple access interference constitutes the relevant factor contributing to performance degradation, as the system is dominated by the high interference level from the simultaneous transmission of many users. This paper studies the adoption of a practical SIC receiver that, aided by CRC error control, provides very reliable decisions of correct decoding of received packets; and to the benefit of receiver performance, we enable users to vary the transmitted symbol energies and coding rates

$$E_{\mathrm{x}}[1], \ldots, E_{\mathrm{x}}[K] \tag{3a}$$

$$R[1], \ldots, R[K] \tag{3b}$$

so as to maximize network spectral efficiency. In awareness of energy-efficient communications, we address the design of the best energy and code allocation strategy subject to the average energy constraint over all users:

$$\frac{1}{K} \sum_{k=1}^{K} E_{\mathrm{x}}[k] = \bar{E}. \tag{4}$$

B. RECEIVER OPERATION

The SIC receiver operates under a user decoding order set by the known distribution of the LOS channel power gains. When a direct LOS exists between users and the common MA node and users are equipped with sufficiently directive antennas, the respective LOS components strongly dominate the channel power gains $h[1], \ldots, h[K]$. That is, multipath has a minor relevance in front of the LOS power gain $h_0[k]$; we measure such relevance through the magnitude

$$\eta \triangleq \left(\sum_{k=1}^{K} h_0[k] \right) \left(\sum_{k=1}^{K} \sum_{l=1}^{L-1} \bar{h}_l[k] \right)^{-1}. \tag{5}$$

Typical values for η may be extracted from [38]. Fig. 1 will exemplify later on related results. Based on the superiority of the LOS channel gains in front of multipath, the receiver proceeds to successive decoding in non-increasing order of the LOS channel power gains, as

$$h_0[1] \geq \cdots \geq h_0[k] \geq \cdots \geq h_0[K]. \tag{6}$$

In practice, the above ordering can be implemented either if the receiver perfectly knows the channel power gain affecting each user, or if the receiver accurately estimates and orders the received symbol energies from all users. The latter approach is sustained by [35] and by simulations in Section IV, where it is shown that the received symbol energy distribution after our energy allocation design retains the order set by the LOS channel gains. Therefore, we adopt a SIC strategy in which users are sequentially decoded while treating the random multipath components as additive uncorrelated noise. This scheme, although suboptimal (because LOS-dominated multipath is treated as noise), simplifies the ordering procedure and justifies performance-wise not incorporating a more complex receiver. Optimality is approached as η increases.

In the literature, a number of scenarios exploiting the SIC receiver have been investigated from an abstract perspective. Specifically, information-theoretic analyses and wireless applications under SIC receivers operating ideally [30]. Generally, those are of interest for gaining insight into SIC operation, but of moderate practical usefulness owing to more complex effects affecting the receiver. In our case, we consider of significant interest the analysis of a practical SIC receiver that performs one decoding attempt per user, and error control under CRC enables reliable packet error detection. The system model for such a receiver is detailed in the sequel. Specifically, the receiver proceeds through K stages and, at stage $1 \leq k \leq K$, performs:

(i) initial estimation of the received amplitude and delay $A_0[k]$, $\tau_0[k]$ from the known preamble: $\hat{A}_0[k]$, $\hat{\tau}_0[k]$.
(ii) demodulation and channel decoding. Specifically, the receiver operates with the received waveform $y(\mathrm{t})$ minus the contribution of users already processed $\hat{r}_i(\mathrm{t})$ in

previous ($i < k$) stages

$$y_k(\mathrm{t}) = y(\mathrm{t}) - \sum_{i=1}^{k-1} \hat{r}_i(\mathrm{t}); \tag{7}$$

and, after despreading $y_k(\mathrm{t})$ as

$$y_k[m] = \int y_k(\mathrm{t}) c^*_{k,m}(\mathrm{t}-mT - \hat{\tau}_0[k])\mathrm{dt}, \tag{8}$$

recovers the useful packet data and its associated CRC.

(iii) If the CRC does not check out, no cancellation occurs: $\hat{r}_k(\mathrm{t}) = 0$; whereas if the CRC checks out, the initial preamble-based estimations $\hat{A}_0[k]$, $\hat{\tau}_0[k]$ are improved using the successfully recovered packet $s_k[0 \leq m < n]$, and waveform cancellation is performed. The k-th user reconstructed waveform for cancellation is

$$\hat{r}_k(\mathrm{t}) = \hat{A}_0[k] \sum_{m=0}^{n-1} s_k[m] c_{k,m}(\mathrm{t}-mT - \hat{\tau}_0[k]). \tag{9}$$

This cancellation approach is sustained by the operation described in highly relevant works [32], [33].

Herein, we seek an accurate system model that captures the main features of the SIC receiver and that allows for further optimization. Following [35]–[37], the model is developed for a finite number of users in Section II-C, and extended to the user infinite case in Section II-D. In Section IV, we consider low-level simulation details to assess its accuracy.

C. THE MODEL FOR K USERS

Unlike the ALOHA protocol, we assume that the traffic delivered to the central node fluctuates at a low magnitude and consider both the number of users K and the traffic as deterministic magnitudes [8, Sec. 2.2.3]. Firstly, some considerations about the statistics of interference are stated before entering into more specific details of the SIC system model. As explained in Section II, we adopt long spreading codes for each user, which entails that interfering signals resemble Gaussian signals after despreading [17]. Consequently, we treat them as equivalent Gaussian additive noise terms, for which the relevant magnitudes are their variances. This motivates the computation of an SINR-dependent model for user k based on the ratio between the symbol energy received from the strong LOS path

$$E_{\mathrm{r}}[k] = E_{\mathrm{x}}[k] h_0[k] \tag{10}$$

and the noise plus interference term $N_{\mathrm{t}}[k]$, as

$$\Gamma[k] = \frac{E_{\mathrm{r}}[k]}{N_{\mathrm{t}}[k]} = \frac{E_{\mathrm{r}}[k]}{N_0 + \xi_{\mathrm{prv}}[k] + \xi_{\mathrm{rem}}[k] + \xi_{\mathrm{m}}}. \tag{11}$$

$N_{\mathrm{t}}[k]$ is decomposed into four terms: N_0, the noise power spectral density (PSD) of thermal noise; $\xi_{\mathrm{prv}}[k]$, the aggregate interference from previously ($i < k$) processed users; $\xi_{\mathrm{rem}}[k]$, the remaining ($i > k$) interference from yet unprocessed users; and ξ_{m}, the interference from multipath, uncorrelated with the main LOS component.

For a SIC receiver that knows the user decoding order, the main effects characterizing its behavior are nonideal channel decoding and imperfect cancellation. Similarly to [36] and [35], we opt for a model where both features are modeled using their known performance versus the SINR Γ characteristics. Specifically, for the former effect, the packet error rate (PER) curve of each encoder-decoder pair PER[Γ, $R_{1 \leq i \leq p}$] is used to characterize the performance of channel decoding at each stage, in combination with statistical independence when decoding different users; for the latter effect, the average energy fraction that remains uncanceled after each user cancellation is incorporated using the residual energy (RE) function $\varepsilon[\Gamma, R_i]$. The relevance of such a model is its high accuracy and the fact that both functions are easy to compute and to be incorporated into our model later on.

Then, with obvious identification of terms, the denominator of (11) yields

$$N_{\mathrm{t}}[k] = N_0 + \frac{\theta}{N} \sum_{i<k} \epsilon[i] E_{\mathrm{r}}[i] + \frac{\theta}{N} \sum_{i>k} E_{\mathrm{r}}[i] + \frac{\theta}{N} \sum_{i,l \geq 1} E_{\mathrm{x}}[i] h_l[i], \tag{12}$$

with θ/N the decorrelation between user signatures in the long-code model [17], and $0 \leq \theta \leq 1$ associated with time misalignments. $\epsilon[i]$ is the binary random variable associated with the decoding-cancellation operation for user i. We adopt a powerful CRC with vanishing undetected-error probability, such that $\epsilon[i]$ equals: 1 when the CRC fails, with probability PER[$\Gamma[i]$, $R[i]$], or the RE $\varepsilon[\Gamma[i], R[i]]$ when the CRC checks out, with complementary probability PSR[$\Gamma[i]$, $R[i]$] $\triangleq$ 1 − PER[$\Gamma[i]$, $R[i]$]. PSR reads as packet success rate.

The problematic aspect of this model is its intractable nature: the randomness associated with successful or failed user cancellation is propagated through the SIC stages. We need the 2^K combinations to compute system performance; an approach difficult to afford, especially when K is large. For these cases, a very competitive approach to deal with randomness is by resorting to the user-asymptotic regime; an approach that deserves an individual study, as follows.

D. THE MODEL FOR INFINITE USERS

We investigate the regime in which the number of users K grows with the spreading gain N at $\alpha \triangleq K/N$. We follow previous works [35]–[37] and analyze the user-limit behavior of the expressions in Section II-C by defining, with a little abuse of notation, the asymptotic indexing

$$t \triangleq \lim_{K \to \infty} k/K, \tag{13}$$

which, relying on some context-awareness, must not be confused with the time variable t (in non-italic font) introduced at the beginning of Section II. Consequently, $1 \leq k \leq K$, now defined by (13), becomes $0 \leq t \leq 1$ in the user-limit. The first user $k = 1$ corresponds to $t = 0$ and

TABLE 2. Notation for the main variables in the model for infinite users.

Definition	Previous	New
User index (according to (13))	k	t
Traffic load	K/N	α
LOS channel power gain	$h_0[k]$	$h_0(t)$
Coding rate	$R[k]$	$R(t)$
Transmitted symbol energy	$E_{\rm x}[k]$	$E_{\rm x}(t)$
Received symbol energy	$E_{\rm r}[k]$	$E_{\rm r}(t)$
Noise plus interference level	$N_{\rm t}[k]$	$N_{\rm t}(t)$
SINR	$\Gamma[k]$	$\Gamma(t)$

the last user $k = K$ to $t = 1$. Hence, the user ordering now corresponds to the ordered LOS channel function

$$h_0(t) \triangleq \lim_{K\to\infty} h_0[tK]. \tag{14}$$

Analogously: $\Gamma[k], N_{\rm t}[k], E_{\rm r}[k], E_{\rm x}[k], R[k]$ are turned into $\Gamma(t), N_{\rm t}(t), E_{\rm r}(t), E_{\rm x}(t), R(t)$, now denoted *profiles* defined over [0, 1]. The main changes of notation are summarized in Table 2. The summations in (12) are substituted by integrals under the differential $\mathrm{d}t \triangleq \lim_{K\to\infty} 1/K$. The received energy profile is $E_{\rm r}(t) = E_{\rm x}(t)h_0(t)$. The SINR profile reads

$$\Gamma(t) = \frac{E_{\rm r}(t)}{N_{\rm t}(t)} = \frac{E_{\rm r}(t)}{N_0 + \xi_{\rm prv}(t) + \xi_{\rm rem}(t) + \xi_{\rm m}}, \tag{15}$$

with the noise plus interference profile

$$N_{\rm t}(t) = N_0 + \xi_{\rm m} + \alpha\theta\int_0^t \epsilon[K\tau]E_{\rm r}(\tau)\mathrm{d}\tau + \alpha\theta\int_t^1 E_{\rm r}(\tau)\mathrm{d}\tau. \tag{16}$$

Remarkably, the interference terms that are random when K is finite ($\xi_{\rm prv}[k], \xi_{\rm m}$) become deterministic in the case of an infinite number of users, as proved in the sequel.

1) INTERFERENCE FROM PROCESSED USERS

The term $\xi_{\rm prv}(t)$ is obtained by multiplying and dividing $\xi_{\rm prv}[Kt]$ by K in (12): $\xi_{\rm prv}[Kt] = \frac{\alpha\theta}{K}\sum_{i<Kt}\epsilon[i]E_{\rm r}[i]$. Making use of Kolmogorov's strong law of large numbers [39], it is straightforward to show that $\lim_{K\to\infty}\xi_{\rm prv}[Kt] = \lim_{K\to\infty}\mathbb{E}[\xi_{\rm prv}[Kt]]$ provided that $E_{\rm r}^2[i] < \infty$ for $1 \le i \le K$ and $S \triangleq \sum_{i\le Kt} i^{-2}\mathrm{VAR}[\epsilon[i]E_{\rm r}[i]] < \infty$. Specifically: $\max_i E_{\rm r}[i] < \infty$ is subject to channel power gain profiles with non-infinite gain, and $S \le \sum_{i\le Kt} i^{-2} \le \pi^2/6 < \infty$ [39, Theorem 2.3.10]. Thus, each of the random variables $\epsilon[K\tau]$ in (16) can be substituted by their expectations

$$\mathbb{E}[\epsilon[K\tau]] \triangleq 1 - (1 - \varepsilon[\Gamma(\tau), R(\tau)])\mathrm{PSR}[\Gamma(\tau), R(\tau)]. \tag{17}$$

2) INTERFERENCE FROM MULTIPATH

Analogously, since the multipath components are statistically independent of the transmitted symbol energies, $\xi_{\rm m}$ tends to the deterministic quantity

$$\xi_{\rm m} = \alpha\theta\,\frac{\bar{h}_0}{\eta}\int_0^1 E_{\rm x}(\tau)\mathrm{d}\tau. \tag{18}$$

$\bar{h}_0 = \frac{1}{K}\sum_k h_0[k]$ is the average LOS power gain.

3) CONCLUDING REMARKS

The asymptotic large-user assumption provides a very comfortable result in that $N_{\rm t}(t)$ in (16), the denominator of $\Gamma(t)$, turns out to be deterministic rather than the sum of random variables in (12). The relevant finding is that the SIC receiver can be seen as a dynamical system, whose behavior can be characterized through deterministic expressions either in differential or integral form. Henceforth, we use the integral form:

$$N_{\rm t}(t) = N_{\rm t}(0)\exp\left(-\alpha\int_0^t \Phi\left[\frac{E_{\rm r}(\tau)}{N_{\rm t}(\tau)}, R(\tau)\right]\mathrm{d}\tau\right) \tag{19a}$$

$$N_{\rm t}(0) = N_0 + \alpha\theta\int_0^1 E_{\rm x}(\tau)\left(h_0(\tau) + \frac{\bar{h}_0}{\eta}\right)\mathrm{d}\tau. \tag{19b}$$

$\Phi[\Gamma, R] \triangleq \theta(1 - \varepsilon[\Gamma, R])\Gamma \cdot \mathrm{PSR}[\Gamma, R]$ is a known function that describes the decoding-cancellation system implementation. In particular, $\Phi[\Gamma, R] = a(\Gamma, R)\theta\Gamma$ with $0 \le a(\Gamma, R) \le 1$ quantifying the closeness of the practical SIC implementation to a genie-aided receiver. Although our formulation focuses on a SIC policy aided by redundancy-check error control, it can be easily extended to other SIC schemes by re-defining Φ as in [37].

III. SYSTEM OPTIMIZATION

Inspired by [30], we adopt the user-aggregate spectral efficiency as the key performance indicator for system optimization. We have adapted it for short-packet transmissions with non-zero error performances. Specifically, the spectral efficiency of the evaluated spread spectrum multiple access network is computed by aggregating all the user effective rates as (20), which, as the number of users tends to infinity, converges almost surely to the utility (21):

$$\mathrm{SE} = \frac{1}{N}\sum_{k=1}^{K} R[k]\cdot\mathrm{PSR}[\Gamma[k], R[k]], \tag{20}$$

$$\mathrm{ASE} = \alpha\int_0^1 R(t)\cdot\mathrm{PSR}[\Gamma(t), R(t)]\mathrm{d}t. \tag{21}$$

We denote the latter asymptotic spectral efficiency (ASE) to distinguish it from the case when K is finite. We investigate the maximum of the functional ASE when the average energy over the user population is limited to $\bar{E}$ Joules per symbol:

$$\frac{1}{K}\sum_{k=1}^{K} E_{\rm x}[k] \xrightarrow{\text{a.s. } K\to\infty} \int_0^1 E_{\rm x}(t)\mathrm{d}t = \bar{E}. \tag{22}$$

We derive the best user-energy and -rate allocation policies that, leveraging the perfectly-known strong LOS paths $h_0(t)$, provide the necessary unbalance between received energies and transmitted rates so that ASE is maximized. We seek an asymptotic allocation function independent of K defined on a function space. Our approach consists: firstly, in deriving analytical expressions for optimal energy and rate allocation using tools from the calculus of variations; and secondly, in addressing a numerical implementation for their resolution. This substantially differs from [30, Section III-B] where

the authors turn the optimization in a function space to a multivariate optimization by simple discretization; an approach that, to the best of our knowledge, can be done more accurately and with low computation times. More concretely, the approach presented in the sequel turns a potentially infinite-dimensional problem into a computationally much simpler resolution where a few scalars need to be found.

A. ASE WITH BALANCED CHANNELS

We first analyze the case in which the LOS channel power gains are not unbalanced: $h_0(t) = \bar{h}_0$. In practice, this situation corresponds to satellites serving ultra-dense networks, in which users are concentrated in such a limited geographic area on Earth that are affected by practically equal strong LOS paths within the same spot beam.

1) RESULTS IN THE INFINITE BLOCKLENGTH REGIME

Our first analysis considers: for users, an infinite set ($p \to \infty$) of capacity-achieving encoders $\mathcal{M}$ from which each user selects the coding rate as a function of its allocated SINR, as

$$r(\Gamma) = \log(1 + \Gamma); \tag{23}$$

and for the receiver, a genie-aided SIC approach with perfect cancellation and error-free decoding: $\Phi[\Gamma, R] = \Phi[\Gamma] = \theta\Gamma$.

Making use of the analytical findings reported in [30, Sec. 3.5], all users can be admitted in the system as long as: they regulate the transmitted symbol energies $E_{\mathrm{x}}(t)$ to arrive with an exponential energy distribution, which corresponds to a uniform SINR allocation $\Gamma(t) = \bar{\Gamma}$ in $0 \le t \le 1$ with

$$\bar{\Gamma} = \frac{1}{\alpha\theta} \ln\left(1 + \frac{\alpha\theta\bar{E}\bar{h}_0}{N_0 + \alpha\theta\bar{E}\bar{h}_0\eta^{-1}}\right), \tag{24}$$

and they choose the same coding rate $\log(1 + \bar{\Gamma})$, for which the optimal ASE reads

$$\mathrm{ASE} = \alpha \log(1 + \bar{\Gamma}). \tag{25}$$

2) RESULTS IN THE FINITE BLOCKLENGTH REGIME

We extend the previous analysis by allowing users to transmit n_e-symbol codewords rather than unboundedly large packets, and by adopting a *nonideal* SIC receiver with decoding success $\mathrm{PSR}[\Gamma, R] = q$ and RE factor ε. The receiver characteristic is $\Phi[\Gamma, R] \equiv \Phi[\Gamma] = \theta(1-\varepsilon)q\Gamma$. We consider that users employ second-order coding rates that achieve [23]

$$r(\Gamma) = \log(1 + \Gamma) - \sqrt{\frac{V(\Gamma)}{n_e}}\mathcal{Q}^{-1}(1 - q) \tag{26}$$

for a given reliability $0 < q < 1$. In the above expression: $V(\Gamma) \triangleq (1 - (1 + \Gamma)^{-2})\log^2 e$ is the channel dispersion and $\mathcal{Q}(\cdot)$ is the Q-function. Note that $q = 1$ enables reliable communication if $n_e \to \infty$; the case analyzed before.

For this analysis, we adopt a broader function space with two-piecewise continuously differentiable functions defined by the user index $0 < t_* \le 1$, above which null energy is allocated: $E_{\mathrm{x}}(t) = 0$ in $t > t_*$. The user index t_* is optimized as well, ensuring that ASE does not decay as α increases.

We address the following variational calculus problem under the adopted function space, which for convenience we formulate in terms of the received energy profile $E_{\mathrm{r}}(t)$ as

$$\max_{t_* \le 1} \ \max_{E_{\mathrm{r}}(t)} \ \alpha \int_0^{t_*} r\left(\frac{E_{\mathrm{r}}(t)}{N_{\mathrm{t}}(t)}\right) \mathrm{d}t \tag{27a}$$

$$\text{s.t. } \bar{E} = \frac{1}{\bar{h}_0}\int_0^{t_*} E_{\mathrm{r}}(t)\mathrm{d}t \tag{27b}$$

$$\text{s.t. } \frac{N_{\mathrm{t}}(t)}{N_{\mathrm{t}}(0)} = \exp\left(-\alpha\int_0^t \Phi\left[\frac{E_{\mathrm{r}}(\tau)}{N_{\mathrm{t}}(\tau)}\right]\mathrm{d}\tau\right) \tag{27c}$$

$$\text{s.t. } N_{\mathrm{t}}(0) = N_0 + \alpha\theta\bar{E}\bar{h}_0(1 + \eta^{-1}) \tag{27d}$$

Unlike the previous solution, the optimal allocation enables the transmission only to users able to attain a minimum SINR requirement, computed as the $\Gamma_* \ge 0$ solution to

$$r'(\Gamma_*)\Gamma_* = r(\Gamma_*). \tag{28}$$

Since the same reliability constraint $q < 1$ is superimposed to all users, the maximum ASE is achieved when *active* users arrive with exponentially distributed symbol energies and experience the same SINRs when they are decoded. The optimal SINR profile is

$$\Gamma(t) = \begin{cases} \max(\bar{\Gamma}, \Gamma_*) & \text{if } 0 \le t \le t_* \\ 0 & \text{if } t_* < t \le 1, \end{cases} \tag{29}$$

with $t_* = \min(\bar{\Gamma}/\Gamma_*, 1)$ and $\bar{\Gamma}$ computed by solving

$$\Phi[\bar{\Gamma}] = \frac{1}{\alpha}\ln\left(\frac{N_0 + \alpha\theta\bar{E}\bar{h}_0(1 + \eta^{-1})}{N_0 + \alpha\theta\bar{E}\bar{h}_0(\eta^{-1} + 1 - q + \varepsilon q)}\right). \tag{30}$$

The achieved ASE is

$$\mathrm{ASE} = \alpha t_* q \cdot r(\max(\bar{\Gamma}, \Gamma_*)). \tag{31}$$

B. ASE WITH UNBALANCED CHANNELS AND FAIR RELIABILITY

Our next study analyzes the impact of unbalanced LOS channel power gains $h_0(t)$ on ASE. A simple setting was analyzed in [30] and solved numerically with a lack of theoretical insights. In this sense, we provide a theoretical study into the maximization of the ASE of a broader problem that considers short packets and a nonideal receiver. Specifically, we consider users employing second-order coding rates constrained to a reliability q and a nonideal SIC receiver characterized by $\Phi[\Gamma, R] = \Phi[\Gamma] = \theta(1 - \varepsilon)q\Gamma$. With regard to system optimization, we consider two-piecewise continuously differentiable functions defined by the admission user index $0 < t_* \le 1$, above which null rate and energy are allocated. The following optimization problem determines the corresponding smooth $E_{\mathrm{r}}(t)$ in $0 \le t \le t_*$:

$$\max_{0 < t_* \le 1} \quad \max_{E_{\mathrm{r}}(t)} \ \alpha \int_0^{t_*} r\left(\frac{E_{\mathrm{r}}(t)}{N_{\mathrm{t}}(t)}\right) \mathrm{d}t \tag{32a}$$

$$\text{s.t. } E_{\mathrm{r}}(t_*) \ge 0 \quad \text{s.t. } \bar{E} = \int_0^{t_*} \frac{E_{\mathrm{r}}(t)}{h_0(t)}\mathrm{d}t \tag{32b}$$

$$\text{s.t. (27c)} - \text{(27d)} \tag{32c}$$

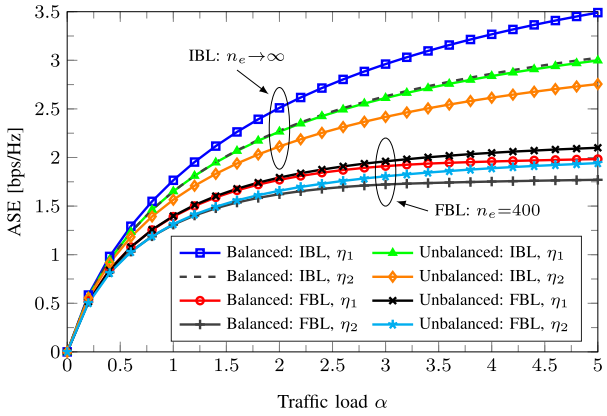


FIGURE 1. ASE versus traffic load α for a proof-of-concept scenario. The simulation parameters are: $\bar{E}/N_0 = 12$ dB, $q = 1 - 10^{-4}$, $\theta = 1$, $\varepsilon[\Gamma, R] = 0.05$ and log-normally distributed $h_0(t)$ with $\bar{h}_0 = 1$, $\eta_1 = 15$ dB and $\eta_2 = 10$ dB.

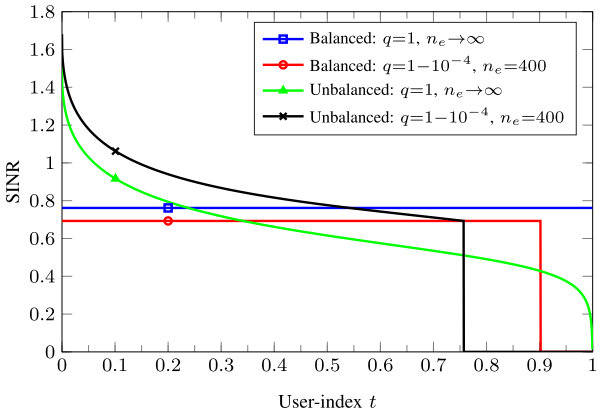


FIGURE 2. Optimal SINR profiles for the same scenario simulated in Figure 1. The traffic load is $\alpha = 4.00$.

This problem is subject to the average energy over the user population (22) formulated in terms of $E_r(t)$ in (32b), and (32c) contains the evolution of the noise plus interference level in the SIC receiver. The user index t_* is optimized in the outer problem guaranteeing the non-negativeness of $E_r(t)$.

As derived in Appendix I-D, the optimal $E_r(t)$ is obtained from the $\Gamma(t) = E_r(t)/N_t(t)$ that satisfies

$$r'(\Gamma(t)) = \lambda \left(\frac{N_t(t)}{h_0(t)} + c - \alpha\theta(1-\varepsilon)qI_x(t) \right) \tag{33}$$

in $0 \le t \le t_*$, with

$$I_x(t) = \int_t^{t_*} \frac{E_r(\tau)}{h_0(\tau)} d\tau, \quad I(t) = \int_t^{t_*} E_r(\tau) d\tau \tag{34a}$$

$$c = \alpha\theta\bar{E} \left(1 + \alpha\theta I(0) \cdot \frac{1-(1-\varepsilon)q}{N_0 + \xi_m} \right) \tag{34b}$$

and the coding rates adapted as $R(t) = r(\Gamma(t))$.

1) PROPERTIES OF THE SOLUTION

The optimum energy allocation results into the following important properties:

i. $\Gamma(t)$ and $R(t)$ are non increasing in t.
ii. $\Gamma(t)$ and $R(t)$ are no longer uniform in t except if the LOS channel power gains are balanced: $h_0(t) = \bar{h}_0$.
iii. the slope of $R(t)$ is close to $\nabla_t h_0(t)/h_0^2(t)$.
iv. $\Gamma(t) \ge \Gamma_*$ in $0 \le t \le t_*$, with Γ_* the solution to

$$r'(\Gamma_*)\Gamma_* = r(\Gamma_*) \tag{35}$$

and users $t > t_*$ unable to attain such a SINR remain silent. If $t_* < 1$, then $\Gamma(t_*) = \Gamma_*$.

2) SOLUTION PROCEDURE

We obtain the optimal t_* and $E_r(t)$ by solving (33) jointly with the integral (27c)–(27d) and a halting criterion. Our numerical implementation discretizes $t \in [0, 1]$ in r intervals of length $\Delta t \triangleq \frac{1}{r}$, resulting in the left endpoints $t_0, \ldots, t_{r-1}$. For temporarily known λ and $I(t_0)$, $N_t(t_0)$ is computed as (27d): $N_t(0) = N_0 + \alpha\theta I(t_0) + \alpha\theta\bar{E}\bar{h}_0\eta^{-1}$. Then, at each t_i, $\Gamma(t_i)$ is computed from (33), after which $I_x(t_{i+1}) = I_x(t_i) - E_x(t_i)\Delta t$ and $N_t(t_{i+1}) = N_t(t_i)\exp(-\alpha\Phi[\Gamma(t_i)]\Delta t)$ are updated for the next t_{i+1} computation. The validity of the computation for t_i is conditioned on $\Gamma(t_i) > \Gamma_*$. A halting policy stops the algorithm when $\Gamma(t_i) \le \Gamma_*$, and thus we set $t_* = t_{i-1}$. We determine λ and $I(0)$ by bisection searches since we found in all the simulated cases that the end points $I(t_*)$ and $I_x(t_*)$ are monotonic in λ and $I(0)$, respectively. The optimal $\Gamma(t)$ satisfies $I(t_*) = I_x(t_*) = 0$.

Fig. 1 and Fig. 2 exemplify the above findings relative to the ASE optimization with balanced channels.

C. ASE WITH UNBALANCED CHANNELS AND OPTIMAL RELIABILITY

In the next approach, we go one step forward and determine the optimum system performance releasing the fair reliability constraint. We use the normal approximation of the maximal channel coding rate [23] to express the PER function as a bivariate function of the variables Γ and R as

$$\mathrm{PER}[\Gamma, R] = Q\left(\sqrt{n_e} \cdot \frac{\log(1+\Gamma) - R}{\sqrt{V}} \right). \tag{36}$$

We design jointly the coding rate $R(t)$ and the transmitted energy $E_r(t)$ profiles to maximize ASE, adopting two-piecewise continuously differentiable functions defined by $0 < t_* \le 1$; t_* is optimized as well. We maximize ASE by formulating the following variational problem:

$$\max_{0<t_*\le 1} \max_{R(t),E_r(t)} \alpha \int_0^{t_*} R(t)\mathrm{PSR}\left[\frac{E_r(t)}{N_t(t)}, R(t) \right] dt \tag{37a}$$

$$\text{s.t. } \bar{E} = \int_0^{t_*} \frac{E_r(t)}{h_0(t)} dt \tag{37b}$$

$$\text{s.t. (19a)} - \text{(19b)} \tag{37c}$$

As shown in Appendices I-A and I-B, the solution of the inner problem is cast in terms of those $\Gamma(t)$ and $R(t)$ that jointly satisfy the following equations in $0 \le t \le t_*$

$$\lambda = \frac{R(t)\mathrm{PSR}_\Gamma[\Gamma(t), R(t)]}{N_t(t)\left(\frac{1}{h_0(t)} + \frac{\alpha\theta\bar{E}}{N_0+\xi_m}\right) - \rho(t)\Phi_\Gamma[\Gamma(t), R(t)]} \tag{38a}$$

$$\lambda = -\frac{\mathrm{PSR}[\Gamma(t), R(t)] + R(t)\mathrm{PSR}_R[\Gamma(t), R(t)]}{\rho(t)\Phi_R[\Gamma(t), R(t)]} \tag{38b}$$

with

$$\rho(t) \triangleq \alpha I_{\mathrm{x}}(t) + \alpha \frac{\alpha\theta\bar{E}}{N_0 + \xi_{\mathrm{m}}} I(t). \tag{39}$$

$\mathrm{PSR}_\Gamma, \mathrm{PSR}_R$ and Φ_Γ, Φ_R are the partial derivatives of $\mathrm{PSR}[\Gamma, R]$ and $\Phi[\Gamma, R]$ with respect to Γ and R.

The optimum allocation policy sets a minimum SINR and rate requirement to perform allocation, so that users unable to attain it remain silent. The optimum admission user index (determined in Appendix I-C) is: $t_* = 1$, or such $t_* < 1$ that allocates the last active user t_* to a critical SINR and rate $\Gamma(t_*) = \Gamma_*$ and $R(t_*) = R_*$ computed by solving

$$\begin{aligned} \Gamma_* \cdot \mathrm{PSR}_\Gamma[\Gamma_*, R_*] &= \mathrm{PSR}[\Gamma_*, R_*], \\ -R_* \cdot \mathrm{PSR}_R[\Gamma_*, R_*] &= \mathrm{PSR}[\Gamma_*, R_*]. \end{aligned} \tag{40}$$

A theoretical insight on the optimal allocation is provided next using the *elasticity* magnitude (commonly used in economics). Certainly, energy and rate are allocated guaranteeing those fractional variations in Γ, R entail fractional changes in PSR of a lesser magnitude (inelastic); in particular, the last user admitted in the system achieves the critical SINR and rate at which PSR varies in the same proportion.

1) SOLUTION PROCEDURE

Likewise in the previous problem, obtaining the solution needs bisection searches of $\lambda, I(0) > 0$. We follow similar rationales as those explained in Section III-B: for the temporarily known pair $\{\lambda, I(0)\}$, $\Gamma(t_i), R(t_i)$ are computed from (38a)–(38b) evaluated at $t = t_i$, after which $I_{\mathrm{x}}(t_{i+1}) = I_{\mathrm{x}}(t_i) - E_{\mathrm{x}}(t_i)\Delta t$, $I(t_{i+1}) = I(t_i) - E_{\mathrm{r}}(t_i)\Delta t$ and $N_{\mathrm{t}}(t_{i+1}) = N_{\mathrm{t}}(t_i)\exp(-\alpha\Phi[\Gamma(t_i), R(t_i)]\Delta t)$ are updated for the next t_{i+1} computation. A halting criterion stops the algorithm either when $\Gamma(t_i) < \Gamma_*$ or $R(t_i) < R_*$.

D. ASE WITH UNBALANCED CHANNELS AND A FINITE ENCODER SET

We now consider the practical case in which users have a finite set of p given encoders, instead of the many coding schemes adopted before. We consider coding schemes with non-crossing PER characteristics, as shown for the examples depicted in Fig. 3. Then, we seek the optimum symbol energy and coding rate $E_{\mathrm{r}}(t)$ and $R(t) \in \{R_{1\le i\le p}\}$ profiles. To that aim, we first consider a SIC policy in which users are decoded in subsets of users that share the same encoder. Further, we consider an arbitrary permutation (i out of $p!$) of the encoder set that we denote $\mathcal{M}^i \triangleq \{\mathcal{M}_1^i, \mathcal{M}_2^i, \ldots, \mathcal{M}_p^i\}$, and the partition of the user set

$$\tau^i \triangleq \{t_0^i = 0, t_1^i, \ldots, t_p^i = t_*^i \le 1\}, \tag{41}$$

wherein users in the interval $[t_{j-1}^i, t_j^i)$ use $\mathcal{M}_j^i$. In this case, we consider a more general form of piecewise continuously differentiable functions: those having at most $p+1$ pieces and defined within $0 \le t \le 1$. Then, we introduce the left and right limits at any point t_k^i where the sought profiles may have a discontinuity denoted *corner*: t_k^{i-} and t_k^{i+}. Then, for a given permutation $\mathcal{M}^i$ the following needs to be solved:

$$\max_{\{t_1^i,\ldots,t_p^i\}} \max_{E_{\mathrm{r}}(t)} \alpha \sum_{k=1}^{p} \int_{t_{k-1}^{i+}}^{t_k^{i-}} R_k \mathrm{PSR}\left[\Gamma(t), R_k\right] \mathrm{d}t \tag{42a}$$

$$\text{s.t. } \bar{E} = \sum_{k=1}^{p} \int_{t_{k-1}^{i+}}^{t_k^{i-}} \frac{E_{\mathrm{r}}(t)}{h_0(t)} \mathrm{d}t \tag{42b}$$

$$\text{s.t. (19a)} - \text{(19b)} \tag{42c}$$

The analytic solution of the inner subproblem is found to fulfill a discrete version of (38a). Specifically, we have $k = 1, \ldots, p$ equations of the type

$$\lambda = \frac{R_k \mathrm{PSR}_\Gamma[\Gamma(t), R_k]}{N_{\mathrm{t}}(t)(\frac{1}{h_0(t)} + \frac{\alpha\theta\bar{E}}{N_0+\xi_{\mathrm{m}}}) - \rho(t)\Phi_\Gamma[\Gamma(t), R_k]} \tag{43}$$

to be fulfilled in each $t_{k-1}^{i+} \le t < t_k^{i-}$. The optimum partition $t_1^i, \ldots, t_{p-1}^i$ (outer problem) is found by taking univariate differentiation over each t_k^i. This results, as proved in Appendix II, into the following condition at each point t_k^i ($t_k^i > t_{k-1}^i$) where $E_{\mathrm{r}}(t)$ has a corner:

$$\lambda = \frac{R_k \mathrm{PSR}[\Gamma(t_k^{i-}), R_k] - R_{k+1}\mathrm{PSR}[\Gamma(t_k^{i+}), R_{k+1}]}{F(t_k^{i-}, R_k) - F(t_k^{i+}, R_{k+1})}, \tag{44}$$

with $F(t, R) = E_{\mathrm{r}}(t)(\frac{1}{h_0(t)} + \frac{\alpha\theta\bar{E}}{N_0+\xi_{\mathrm{m}}}) - \rho(t)\Phi[\Gamma(t), R]$. With regard to the last user t_p^i, the corner determines the admission user-index apart from which null energy is allocated. Specifically, when $t_p^i < 1$, the last user SINR is mapped to $\Gamma(t_p^{i-}) = \Gamma_*$ such that

$$\Gamma_* \mathrm{PSR}_\Gamma[\Gamma_*, R_p] = \mathrm{PSR}[\Gamma_*, R_p]. \tag{45}$$

Summarizing results, we have one stationary point equation (43) for each k out of p intervals in order to find each of the pieces $E_{\mathrm{r}}(t_{k-1}^{i-} \le t \le t_k^{i+})$ plus $p-1$ equations (44) to find the corner points $t_1^i, t_2^i, \ldots, t_{p-1}^i$, and the additional equation (45) to determine t_p^i when $t_p^i < 1$.

Solution procedure: A fast computational method for solving (42a)–(42c) is proposed, where the encoder's set is jointly determined with the optimum profiles, with only exhaustive search over an initial encoder $\mathcal{M}_1$, and bisection searches of the parameters $\lambda, I(0)$. We discretize, as in Section III-B, the user-variable t in r intervals of length $\Delta t = r^{-1}$ and use the endpoint representation $t_0^+, \ldots, t_r^+$ for the right limits of every interval and $t_i^- = t_{i-1}^+$ for the lower limits. Firstly, assume an initial encoder (primary encoder) $\mathcal{M}_1$. Secondly, for every encoder $\mathcal{M}_{1\le i\le p}$, solve (43) at $t_k^+ \equiv t_k^+(\mathcal{M}_i)$ to obtain $\Gamma(t_k^+(\mathcal{M}_i))$. Thirdly, check if (44) is verified for the SINRs obtained by the primary encoder $\Gamma(t_k^-(\mathcal{M}_1))$ and the rest of encoders $\Gamma(t_k^+(\mathcal{M}_{i>1}))$. If (44) is verified for $\Gamma(t_k^+(\mathcal{M}_j))$, change the primary encoder to $\mathcal{M}_j$ for the next t_{i+1}^+ computation, otherwise continue with encoder $\mathcal{M}_i$. Finally, update, under finite differences, the variables $I_{\mathrm{x}}(t_{i+1}^+), I(t_{i+1}^+), N_{\mathrm{t}}(t_{i+1}^+)$ for next t_{i+1}^+ computation.

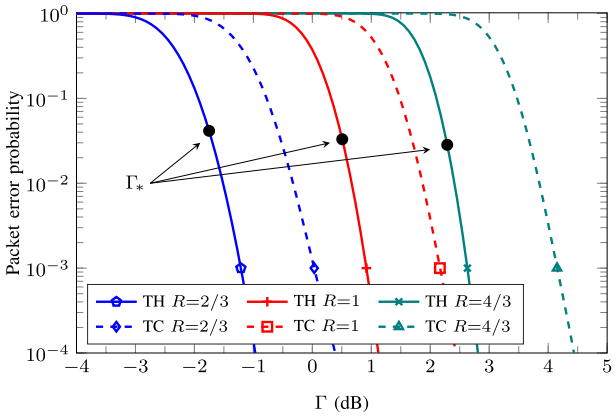


FIGURE 3. PER versus Γ curves. Theoretical encoders (TH) are obtained from second-order coding rates at blocklength n_e and rate R. Practical encoders (TC) correspond to 8-state turbo codes [40] with QPSK modulation and blocklength $n_e = 500$.

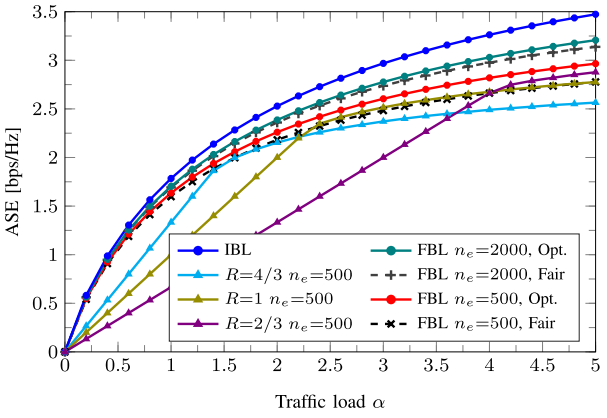


FIGURE 4. ASE versus traffic load α for several cases: (i) the infinite blocklength regime (IBL) $n_e \to \infty$; (ii) the finite blocklength regime (FBL) with fair ($1 - q = 10^{-3}$) and optimal reliability; and (iii) a single encoder of rate R and n_e symbols.

IV. SIMULATIONS AND NUMERICAL RESULTS

Numerical results are shown in the following two subsections where: firstly, we analyze a representative setting; and secondly, we simulate a low-level implementation of a SIC receiver to assess the performance of our study.

A. A REPRESENTATIVE CASE STUDY

We first address a theoretical study of a representative setting in which many users are subject to log-normal shadowing; a realistic approximation used for a Land Mobile Satellite system employing E-SSA in [32]. Specifically, we compute the LOS channel power gains as $h_0(t) = F^{-1}(1-t)$, with $F(h)$ the cumulative distribution of a log-normal random variable with mean 0 dB and deviation 3 dB. These parameters are extracted from [32]. To avoid unrealistic channel power gains, we neglect the subset $\{h, F_H(h) \geq 1 - 10^{-3}\}$. Moreover, we consider the average energy constraint $\bar{E}/N_0 = 12$ dB, the average LOS channel power gain $\bar{h}_0 = 1$, and the strong LOS path $\eta = 15$ dB. At the SIC receiver, the factor θ is set to 1 and the RE after cancellation is assumed Γ-independent [33], [36] with $\varepsilon[\Gamma, R] = 0.01$.

We then analyze ASE and the optimal energy and rate allocation computed by solving the stationary point equations under $r = 5000$ intervals.

1) ASYMPTOTIC SPECTRAL EFFICIENCY

We depict in Fig. 4 ASE (21) versus the traffic load α for a number of cases following the same structure in Section III.

Firstly, we consider $p \to \infty$ encoders and analyze the optimal ASE for several blocklengths n_e. The ASE in the IBL regime attains error-free performance. Severe performance degradation is experimented at $n_e \leq 2000$ when users transmit short packets and employ second-order coding rates for the Gaussian channel [23]. The gap between the IBL and FBL regimes increases with α. The cause is threefold: more users are left unsuccessfully decoded; the achievable coding rates are found to be far from the capacity of the Gaussian channel for moderate packet lengths and sufficiently low SNRs; and the minimum SINR and rate requirements given in (35) and (40) severely penalize ASE. Moreover, the optimal allocation constrained by fair reliability is shown to degrade substantially ASE, especially so at low blocklength.

Secondly, we simulate a more practical approach in which all users employ the same encoder ($p = 1$), as in [35]. We adopt theoretical encoders achieving optimal second-order coding rates at blocklength $n_e = 500$ and different rates. Their error performance versus SINR are shown in Fig. 3 under the labels "TH". Simulation results in Fig. 4 show that ASEs increase almost linearly in α with slopes close to $R_{1 \leq i \leq 3}$, up to the critical traffic loads at which energy allocation sets the last user SINR to $\Gamma(1) = \Gamma_*$. The respective SINR points Γ_* are also depicted in Fig. 3. At the critical traffic loads, a single encoder practically attains the maximum ASE computed with many encoders for the same blocklength $n_e = 500$. At higher traffic loads, the admission user index $t_1 < 1$ is enabled to silence users with poor channel power gains; specifically, users $t_1 < t \leq 1$ remain silent. This causes ASE to increase with a less steep slope.

The previous analysis is extended to the case of $p = 2$ or $p = 3$ encoders available to users. The ASE is drawn in Fig. 5. The optimum ASE in the FBL regime with many encoders is practically attained when users employ either 2 or 3 encoders, provided that they take advantage of the optimal allocation. Specifically, the optimization problem (42a)–(42c) is satisfied only for piecewise continuously differentiable symbol energy profiles $E_x(t)$ and $E_r(t)$. Moreover, the optimum permutation of the encoder set is $\mathcal{M}^* = \{\mathcal{M}_1^*, \mathcal{M}_2^*, \mathcal{M}_3^*\}$ such that $R_1 > R_2 > R_3$ (decreasing order of coding rates). Remarkably, strong users exploit high rate encoders whereas weak users employ more protective encoders. This result can be extrapolated to more encoders as long as their PER

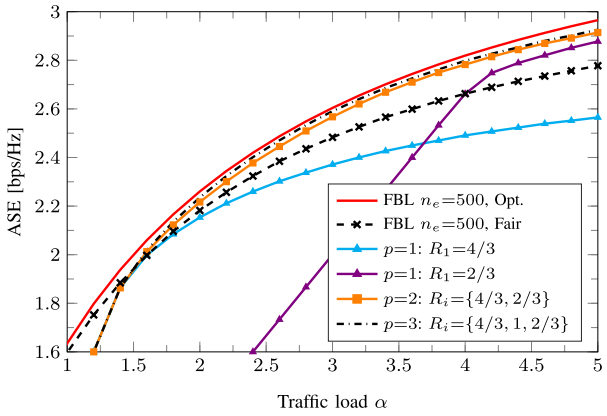


FIGURE 5. ASE versus traffic load α at blocklength $n_e = 500$ symbols and different encoder sets.

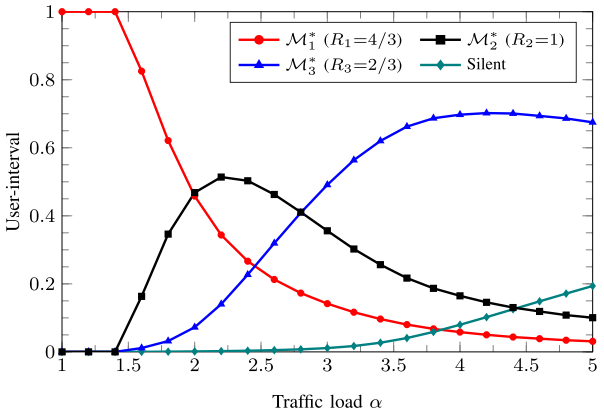


FIGURE 6. Fraction of users employing encoders $\mathcal{M}^*_{1\le i\le 3}$ versus traffic load α. The fraction of users remaining silent corresponds to $1 - t_3$.

characteristics do not intersect; a hypothesis that generally holds if the coding rates are sufficiently distant.

2) ENERGY AND RATE ALLOCATION

The fraction of users employing each encoder is depicted in Fig. 6. Only computations relative to the optimal reliability case are shown. At $\alpha \le 1.40$, enough energy is available to allocate a sufficiently high SINR to users when they employ the encoder with the highest rate $\mathcal{M}_1^*$. At $\alpha > 1.40$, $\bar{E}/N_0 = 12$ dB is not enough to provide successful performance to all users if they employ $\mathcal{M}_1^*$. Instead, more encoders are used, allowing thus the transmission of more users while improving ASE. In this regime, more users employ $\mathcal{M}_2^*$, $\mathcal{M}_3^*$ in detriment of the number of users using $\mathcal{M}_1^*$. Particularly, there is always a nonzero fraction of users using each encoder beyond $\alpha = 1.40$. At $\alpha \ge 2.80$, it is better to decrease the fraction of users employing both $\mathcal{M}_1^*$, $\mathcal{M}_2^*$ and enforce the transmission of users under the most protective encoder $\mathcal{M}_3^*$. Moreover, the user admission index $t_3 < 1$ is enabled so as to avoid the transmission of users with degraded channels. This allows, although penalizing the individual performance of those users, the improvement of network performance in terms of ASE. The idea is to adapt, as the level of MA interference increases, the use of more protective channel encoders to the detriment of the use of high rate encoders. Remarkably, users encode information using coding rates that decrease according to the magnitude of the LOS channel power gain profile $h_0(t)$. In the IBL or FBL regimes, $R(t)$ decreases softly from high to low coding rates and enables only some of the available coding rates when p is finite.

Energy allocation takes advantage of unbalanced channel power gains to maximize the transfer of energy without penalizing ASE. Examples of the optimal energy allocation rules are depicted in Fig. 7. This enforces stronger users to employ high coding rates and transmit high energies. Broadly speaking, the optimal allocation adjusts the transmitted symbol energies and rates to fix the individual error performances (the individual probability of error) in

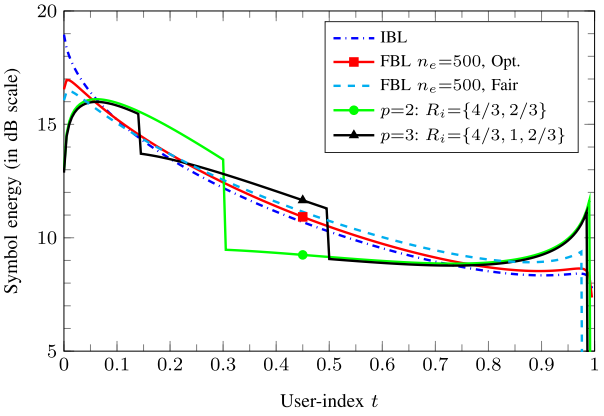


FIGURE 7. Optimal transmitted symbol energy profiles at $\alpha = 3.00$.

the interval $[10^{-4}, 10^{-1}]$. The optimal transmitted symbol energy profiles $E_r(t)$ are drawn in Fig. 7, exhibiting nonincreasing distributions regardless of the adopted set of encoders. This generates, in combination with $R(t)$, a SINR profile $\Gamma(t)$ such that the effects of non-ideal decoding and imperfect cancellation have a lesser impact on the initial SIC stages. This way, the impact of users unsuccessfully decoded and users imperfectly canceled propagated along SIC stages is minimized. One consequence of the previous rationale is that the associated SINR profile $\Gamma(t)$ is a nonincreasing function of the user indexing t.

B. A LOW-LEVEL IMPLEMENTATION OF THE RECEIVER

We assess the accuracy of our system model and performance optimization for a more realistic satellite return link and a low-level receiver implementation at the gateway station.

We have simulated a satellite equipped with a dish antenna and operating in the geostationary orbit with European coverage. We have computed the LOS channel power gains from the path-loss attenuation between the satellite's position and the users' Earth locations plus the antenna radiation diagram. The LOS path is $\eta = 15$ dB

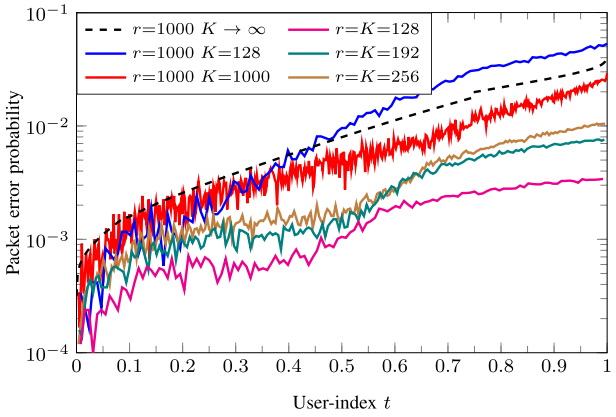


FIGURE 8. Asymptotic and empirical PER profiles PER[$\Gamma(t)$, $R(t)$] at $\alpha = 2.00$. Empirical results have been averaged under $5 \cdot 10^4$ independent Monte Carlo runs.

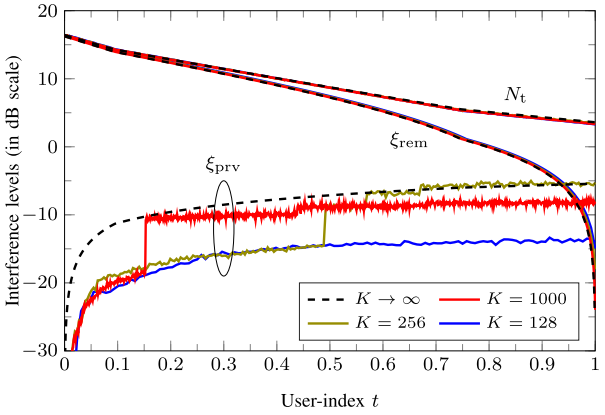


FIGURE 9. Asymptotic and empirical interference profiles at $\alpha = 2.00$. Computations are performed under $r = 1000$ points.

TABLE 3. Spectral efficiency (in bps/Hz) / Network PER ($\times 10^{-3}$) for several number of users K and discretization intervals r.

r	Theoretical	$K = 128$	$K = 256$	$K = 1000$
128	1.77 / 12.70	1.77 / 1.50	1.77 / 0.53	1.77 / 0.36
256	1.83 / 12.01	1.83 / 4.47	1.83 / 3.53	1.83 / 1.07
1000	1.88 / 11.34	1.87 / 16.42	1.87 / 14.11	1.88 / 7.33

superior to multipath. Clear sky conditions are evaluated and the turbulent behavior of the ionosphere over transmitted wavefronts is reasonably discarded [33]. The resulting LOS profile resembles a log-normal distribution. From the users' point of view, we have adopted three encoders from the 8-state parallel concatenated convolutional code [40] with different puncturing patterns to obtain the coding rates $R_1 = 4/3$, $R_2 = 1$ and $R_3 = 2/3$ bits per symbol. Packets contain $n_e = 500$ encoded symbols plus $n_o = 50$ preamble symbols. QPSK is adopted due to its low peak-to-average-power ratio, together with real-valued binary spreading codes and a square-root raised cosine pulse with a roll-off factor of 0.25. In all the simulations, the chip rate is fixed to 3.84 Mchps.

We adopt a SIC receiver operating at the waveform level as follows. It demodulates the strongest user at each stage and performs channel decoding under 20 max-log-map iterations. A perfect CRC is used for error control. For packet reconstruction, complex amplitude is estimated by correlating the successfully recovered packet with the despread signal. Cancellation is produced, and the receiver processes the next user received with the highest LOS power gain.

Recall that our system model and optimization are aided by the PER and RE versus SNR curves of each encoder. We compute them through separate simulations and use them later as lookup tables. For the PER curves, we have estimated the packet error probability in the range $[10^{-4}, 1]$ by averaging 200 packet error events over the total number of decoding trials at different SNRs, and by interpolating and smoothing the resulting trace. The results are shown in Fig. 3. The magnitude of the cancellation error $\varepsilon[\Gamma]$ is simulated jointly with the PER curves, rather than adopting the SINR-independent cancellation error in Section IV-A.

We compute the asymptotic ($K \to \infty$) profiles following the procedure explained in Section III-D, adopting the above PER and RE curves as lookup tables. Specifically, the optimal energy and coding rate profiles $E_{\mathrm{x}}(t), R(t)$ are computed under r points. To be applied for a practical scenario with K users, they are interpolated at $t = k/K$ and $k = 1, \ldots, K$. In this section, we assess the performance and accuracy of the system optimization for relevant designs that relate the number of users K with the number of intervals r used to solve the stationary point equations.

Fig. 8 depicts the packet error probabilities experienced by users. Table 3 shows a number of performance/complexity ratios in terms of spectral efficiency and network PER achieved by theoretical analyses and empirical validations. Theoretical computations result in spectral efficiencies increasing in r; network PER remains practically invariant. The result is clear: higher performance is achieved the more complexity is added to solve the stationary point equations. The spectral efficiencies achieved by empirical simulations are practically equal to those predicted theoretically. Contrarily, some mismatch between theoretical and empirical PERs is evidenced. Generally, empirical PERs are lower except when r is large and K sufficiently low. The reason is that the numerical resolution for the stationary point equations evaluates a conservative computation of the noise plus interference term, whose implementation is highly beneficial when r is low. This gap is remedied as K grows if $r > K$. Generally, a few hundred of users are needed to reach a network PER similar to that theoretically predicted.

This behavior is better highlighted in Fig. 9, where the interference and total noise plus interference terms are shown in dB for a single realization. Looking at ξ_{prv}, for finite K we observe the random appearance of packet errors as sporadic steep increments of $\xi_{\mathrm{prv}}[k]$ ($k = Kt$), of decreasing

magnitude in dB as t progresses. Note how in this process $\xi_{\text{prv}}[k]$ approaches the more pessimistic user-asymptotic prediction $\xi_{\text{prv}}(t)$, which assumes a continuous density of appearance of detrimental packet errors. For low K, this conservative computation is more relevant to the accuracy of the interference from previous users. As shown, this is remedied as K grows. Nonetheless, the theoretical model for the terms ξ_{rem} and N_{t} fits the empirical computations very nicely regardless of the number of users. In fact, ξ_{rem} displays a much higher level than ξ_{prv} for the first users (low t), while for the last users where ξ_{rem} is low the accuracy of ξ_{prv} has improved. In summary, the model for the total noise plus interference N_{t} fits empirical computations quite closely for the whole range of the user-index t.

V. CONCLUSION

The spectral efficiency of a massively populated spread spectrum multiple access network with a successive interference cancellation gateway demodulator has been optimized in terms of the energy/code allocation scheme for a population of users communicating under short packets and nonorthogonal spreading waveforms. We have shown that, in the user-asymptotic case, a number of optimum (according to different criteria) energy/code allocation policies can be obtained in a deterministic way as functions of the per-user channel coefficients and the packet error rate curves of the employed coding schemes. We have resorted to a systematic optimization tool over function spaces to derive the asymptotic allocation rules that maximize the user-aggregate spectral efficiency. Theoretical insight and computational advantages advocate for variational calculus versus typical user-finite multivariate optimization techniques in prior literature. Simulations together with the proposed system optimization have shown three relevant results. Firstly, the optimum allocation policy maximizes the energy transfer without penalizing the aggregate throughput of transmitters. Secondly, the use of a few encoders practically attains the performance achieved when using a large encoder set. Thirdly, the asymptotic performance analysis accurately predicts that of a realistic implementation of the iterative receiver.

APPENDIX I
OPTIMAL ALLOCATION UNDER INFINITELY MANY ENCODERS

The ASE maximization in (37a)–(37c) is undertaken. We divide it into two parts. Firstly, we determine for a fixed t_* the continuously differentiable $\Gamma(t)$ and $R(t)$ that maximize

$$\max_{R(t),\Gamma(t)} \alpha \int_0^{t_*} R(t)\text{PSR}\left[\Gamma(t), R(t)\right] \text{d}t \tag{46a}$$

$$\text{s.t. } \bar{E} = \int_0^{t_*} \frac{\Gamma(t) N_{\text{t}}(t)}{h_0(t)} \text{d}t \tag{46b}$$

where (37c) has been substituted into (37a) and (37b) in order to determine, instead of $E_{\text{r}}(t)$, the optimum SINR profile $\Gamma(t)$. Secondly, $0 < t_* \leq 1$ is optimized in Appendix I-C.

The Lagrangian is

$$\mathcal{L} \triangleq \int_0^{t_*} R(t)\text{PSR}[\Gamma(t), R(t)]\text{d}t - \lambda \left(\int_0^{t_*} \frac{\Gamma(t) N_{\text{t}}(t)}{h_0(t)} \text{d}t - \bar{E} \right) \tag{47}$$

From [41], we adopt the space of continuously differentiable functions $\mathcal{S} \triangleq \{g(t) \in \mathcal{C}^1[0, t_*]\}$ under the support $t \in [0, t_*]$, where differentiability is assumed under the norm $\mathcal{D}_0 \triangleq \max_{0 \leq t \leq t_*} |f(t)| + \max_{0 < t < t_*} |\nabla_t f(t)|$. Next, we consider the following variations (with amplitudes a, b and directions $\phi(t), \varphi(t)$) around any $\Gamma_0(t), R_0(t) \in S$:

$$\begin{aligned} \Gamma(t) &= \Gamma_0(t) + a \cdot \phi(t) \\ R(t) &= R_0(t) + b \cdot \varphi(t). \end{aligned} \tag{48}$$

We determine the optimum $\Gamma(t)$ for a given $R(t)$ in Appendix I-A, and the optimum $R(t)$ for a given $\Gamma(t)$ in Appendix I-B. Hence, the stationary point profiles $\Gamma_0(t)$ and $R_0(t)$ will be those that jointly satisfy the respective stationary point equations for the same λ. For simpler notation, hereinafter we drop the explicit arguments in $\text{PSR}[\Gamma, R]$ and $\Phi[\Gamma, R]$; they will be denoted PSR and Φ, respectively. The explicit argument t of all variables will also be implicit by context.

A. INNER PROBLEM: OPTIMUM SINR PROFILE

The stationary point SINR profile Γ_0 is such that the gradient $\nabla_{a \to 0} \mathcal{L}$ vanishes for any admissible ϕ [41]. The Lagrangian (47) contains three terms that depend on a through Γ, viz: PSR in the first integral and $\Gamma \cdot N_{\text{t}}$ in the second term. The gradient of the first term is straightforward whereas that of the second term needs the computation of $\nabla_{a \to 0}[\Gamma \cdot N_{\text{t}}]$, with N_{t} in (19a) depending on a through $N_{\text{t}}(0)$ and Φ. Thus,

$$\begin{aligned} \nabla_{a \to 0} \mathcal{L} = &\int_0^{t_*} R_0 \cdot \text{PSR}_\Gamma \phi \ \text{d}t - \lambda \int_0^{t_*} \frac{N_{\text{t}}}{h_0} \phi \ \text{d}t \\ &- \lambda \bar{E} \cdot \frac{\nabla_{a \to 0} N_{\text{t}}(0)}{N_{\text{t}}(0)} + \lambda \alpha \int_0^{t_*} E_{\text{x}} \int_0^t \Phi_\Gamma \phi \text{d}\tau \text{d}t. \end{aligned} \tag{49}$$

The term $N_t^{-1}(0) \nabla_{a \to 0} N_{\text{t}}(0)$ is computed first, by solving for $N_{\text{t}}(0)$ from $N_{\text{t}}(0) = N_0' + \alpha\theta \int_0^{t_*} E_{\text{r}} \text{d}t$ (19a), with $N_0' \triangleq N_0 + \alpha\theta \bar{E} \bar{h}_0 \eta^{-1}$ and $E_{\text{r}} = \Gamma_0 N_{\text{t}}(0) \exp(-\alpha \int_0^t \Phi \text{d}\tau)$:

$$N_{\text{t}}(0) = \frac{N_0'}{1 - \alpha\theta \int_0^{t_*} \Gamma_0 \exp\left(-\alpha \int_0^t \Phi \text{d}\tau\right) \text{d}t}, \tag{50}$$

and by differentiating later on:

$$\frac{\nabla_{a \to 0} N_{\text{t}}(0)}{N_{\text{t}}(0)} = \frac{\alpha\theta}{N_0'} \int_0^{t_*} \left(N_{\text{t}} \phi - \alpha E_{\text{r}} \int_0^t \Phi_\Gamma \phi \text{d}\tau \right) \text{d}t. \tag{51}$$

It can be simplified by integrating by parts the second term with $u \triangleq \int_0^t \Phi_\Gamma \phi \text{d}\tau$ and $\text{d}v \triangleq E_{\text{r}} \text{d}t$ as

$$\frac{\nabla_{a \to 0} N_{\text{t}}(0)}{N_{\text{t}}(0)} = \frac{\alpha\theta}{N_0'} \int_0^{t_*} \left(N_{\text{t}} - \alpha I \Phi_\Gamma \right) \phi \ \text{d}t, \tag{52}$$

with $I \equiv I(t) \triangleq \int_t^{t_*} E_{\text{r}}(\tau) \text{d}\tau$ the integrated $E_{\text{r}}(t)$ profile.

Similarly, the last term in (49) can still be simplified under the same integration by parts. This gives

$$\alpha \int_0^{t_*} E_{\mathrm{x}} \left(\int_0^t \Phi_\Gamma \phi \, \mathrm{d}\tau \right) \mathrm{d}t = \alpha \int_0^{t_*} I_{\mathrm{x}} \Phi_\Gamma \phi \, \mathrm{d}t. \tag{53}$$

Now, substituting (52) and (53) into (49), we finally get

$$\int_0^{t_*} \left[R_0 \mathrm{PSR}_\Gamma - \lambda \left(N_{\mathrm{t}} \left(\frac{1}{h_0} + \frac{\alpha\theta\bar{E}}{N_0'} \right) - \rho \Phi_\Gamma \right) \right] \phi \mathrm{d}t \tag{54}$$

with $\rho \triangleq \alpha I_{\mathrm{x}} + \alpha(\alpha\theta\bar{E}/N_0')I$. The previous expression is set to zero, and the stationary point $\Gamma(t)$ must verify it for every admissible ϕ. The Fundamental Lemma of the Calculus of Variations (FLCV) [41] requires that its integrand vanishes. Hence, we have the following in $0 \le t \le t_*$:

$$\lambda = \frac{R_0 \mathrm{PSR}_\Gamma[\Gamma_0, R_0]}{N_{\mathrm{t}}(\frac{1}{h_0} + \frac{\alpha\theta\bar{E}}{N_0'}) - \rho \Phi_\Gamma[\Gamma_0, R_0]}. \tag{55}$$

B. INNER PROBLEM: OPTIMUM RATE PROFILE

Analogously, we determine the stationary point $R(t)$, $R_0(t)$. The gradient $\nabla_{b\to 0}\mathcal{L}$ is

$$\nabla_{b\to 0}\mathcal{L} = \int_0^{t_*} (\mathrm{PSR} + R_0 \cdot \mathrm{PSR}_R)\varphi \, \mathrm{d}t - \lambda\bar{E} \frac{\nabla_{b\to 0} N_{\mathrm{t}}(0)}{N_{\mathrm{t}}(0)} + \lambda\alpha \int_0^{t_*} E_{\mathrm{x}} \int_0^t \Phi_R \varphi \mathrm{d}\tau \, \mathrm{d}t. \tag{56}$$

Following a rationale identical to that in Section V-A, the term $N_t^{-1}(0)\nabla_{b\to 0}N_{\mathrm{t}}(0)$ computed from (50) yields

$$\frac{\nabla_{b\to 0} N_{\mathrm{t}}(0)}{N_{\mathrm{t}}(0)} = -\alpha \frac{\alpha\theta}{N_0'} \int_0^{t_*} I \Phi_R \phi \, \mathrm{d}t. \tag{57}$$

Then, the gradient pursued reads

$$\nabla_{b\to 0}\mathcal{L} = \int_0^{t_*} [\mathrm{PSR} + R_0 \mathrm{PSR}_R + \lambda\rho\Phi_R] \, \varphi \, \mathrm{d}t, \tag{58}$$

which is set to zero for any admissible φ. Applying the FLCV [41] in $0 \le t \le t_*$, we obtain

$$\lambda = -\frac{\mathrm{PSR}[\Gamma_0, R_0] + R_0 \mathrm{PSR}_R[\Gamma_0, R_0]}{\rho \Phi_R[\Gamma_0, R_0]}. \tag{59}$$

C. OUTER PROBLEM: USER-ADMISSION OPTIMIZATION

This section solves the univariate optimization under t_*. When $t_* < 1$ is enabled, (55) and (59) must be satisfied at $t = t_*$. Let us denote $\Gamma_* = \Gamma(t_*)$ and $R_* = R(t_*)$. Hence

$$\lambda N_{\mathrm{t}}(t_*) \left(\frac{1}{h_0(t_*)} + \frac{\alpha\theta\bar{E}}{N_0'} \right) = R_* \mathrm{PSR}_\Gamma[\Gamma_*, R_*], \tag{60}$$

$$-\mathrm{PSR}[\Gamma_*, R_*] = R_* \mathrm{PSR}_R[\Gamma_*, R_*]. \tag{61}$$

The optimal t_* is determined from $\nabla_{t_*}\mathcal{L} = 0$, which yields

$$R_* \mathrm{PSR}[\Gamma_*, R_*] = \lambda \left(E_{\mathrm{x}}(t_*) + \frac{\alpha\theta\bar{E}}{N_0'} \cdot E_{\mathrm{r}}(t_*) \right). \tag{62}$$

Dividing by (60) and after some straightforward manipulations, we get the optimal values of Γ_* and R_*

$$\Gamma_* \mathrm{PSR}_\Gamma[\Gamma_*, R_*] = \mathrm{PSR}[\Gamma_*, R_*] \tag{63}$$

$$-R_* \mathrm{PSR}_R[\Gamma_*, R_*] = \mathrm{PSR}[\Gamma_*, R_*]. \tag{64}$$

D. PARTICULAR CASE: FAIR RELIABILITY

We analyze the particular case in which active users are constrained by the same reliability q. In this case, since decoding is error free, we may substitute $R(t)\mathrm{PSR}[\Gamma(t), R(t)]$ by $q \cdot r(\Gamma(t))$ and $\Phi[\Gamma(t), R(t)]$ by $\theta(1-\varepsilon)q\Gamma(t)$. Then, the stationary point equation (55) is simplified as

$$r'(\Gamma(t)) = \lambda \left(\frac{N_{\mathrm{t}}(t)}{h_0(t)} + c - \alpha\theta(1-\varepsilon) I_{\mathrm{x}}(t) \right), \tag{65}$$

with $c \triangleq \alpha\theta\bar{E}(N_0 + \xi_{\mathrm{m}} + \alpha\theta\varepsilon I(0))/(N_0 + \xi_{\mathrm{m}})$.

APPENDIX II
OPTIMAL ALLOCATION UNDER A FINITE ENCODER SET

This appendix solves the optimization problem (42a)–(42c). We consider piecewise smooth profiles, for which the Lagrangian can be expressed as

$$\mathcal{L} = \sum_{k=1}^{p} \int_{t_{k-1}^{i+}}^{t_k^{i-}} R_k \mathrm{PSR}\,[\Gamma(t), R_k]\,\mathrm{d}t \tag{66a}$$

$$- \lambda \left(\sum_{k=1}^{p} \int_{t_{k-1}^{i+}}^{t_k^{i-}} \frac{\Gamma(t) N_{\mathrm{t}}(t)}{h_0(t)} \mathrm{d}t - \bar{E} \right). \tag{66b}$$

The optimum set $t_1^i, t_2^i, \ldots, t_p^i$ is found by taking univariate optimizations over the former indices, resulting in the well-known Erdmann-Weierstrass conditions [41]. Hence, differentiating the first term with respect to t_k^i, we have

$$\nabla_{t_k^i} \left[\sum_{k=1}^{p} \int_{t_{k-1}^{i+}}^{t_k^{i-}} R_k \mathrm{PSR}\,[\Gamma(t), R_k]\,\mathrm{d}t \right] = R_k \mathrm{PSR}[\Gamma(t_k^{i-}), R_k] - R_{k+1}\mathrm{PSR}[\Gamma(t_k^{i+}), R_{k+1}]. \tag{67}$$

Analogously, the differentiation of the second term reads

$$\nabla_{t_k^i} \left[\sum_{k=1}^{p} \int_{t_{k-1}^{i+}}^{t_k^{i-}} \frac{E_{\mathrm{r}}(t) N_{\mathrm{t}}(t)}{h_0(t)} \mathrm{d}t \right] = E_{\mathrm{x}}(t_k^{i-}) \nabla_{t_k^{i-}} N_{\mathrm{t}}(t_k^{i-}) - E_{\mathrm{x}}(t_k^{i+}) \nabla_{t_k^{i+}} N_{\mathrm{t}}(t_k^{i+}), \tag{68}$$

Now, from an extensive analysis, it is easy to show that

$$E_{\mathrm{x}}(t_k^{i-}) \nabla_{t_k^{i-}} N_{\mathrm{t}}(t_k^{i-}) = F(t_k^{i-}, R_k) \tag{69}$$

$$E_{\mathrm{x}}(t_k^{i+}) \nabla_{t_k^{i+}} N_{\mathrm{t}}(t_k^{i+}) = F(t_k^{i+}, R_{k+1}) \tag{70}$$

with

$$F(t, R) \triangleq E_{\mathrm{r}}(t) \left(\frac{1}{h_0(t)} + \frac{\alpha\theta\bar{E}}{N_0'} \right) - \rho(t) \Phi[\Gamma(t), R]. \tag{71}$$

Therefore, we finally obtain the sought equation at every point t_k^i ($k = 1, \ldots, p-1$) where the function has a corner

$$\lambda = \frac{R_k \mathrm{PSR}[\Gamma(t_k^{i-}), R_k] - R_{k+1}\mathrm{PSR}[\Gamma(t_k^{i+}), R_{k+1}]}{F(t_k^{i-}, R_k) - F(t_k^{i+}, R_{k+1})}. \tag{72}$$

Notice that, if $t_p < 1$, then the previous equation remains still valid particularizing $R_{p+1} = 0$ and $\Gamma(t_p^{i+}) = 0$. This gives, together with (43), the SINR for the last user allocated with nonzero energy $\Gamma(t_p^{i-}) = \Gamma_*$, with

$$\Gamma_* \mathrm{PSR}_\Gamma[\Gamma_*, R_p] = \mathrm{PSR}[\Gamma_*, R_p]. \tag{73}$$

REFERENCES

[1] *Ericsson Mobility Report*, Ericsson, Stockholm, Sweden, Nov. 2020.

[2] S. Cioni, R. De Gaudenzi, O. Del Rio Herrero, and N. Girault, "On the satellite role in the era of 5G massive machine type communications," *IEEE Netw.*, vol. 32, no. 5, pp. 54–61, Sep. 2018.

[3] M. Bacco, P. Cassara, M. Colucci, and A. Gotta, "Modeling reliable M2M/IoT traffic over random access satellite links in non-saturated conditions," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 5, pp. 1042–1051, May 2018.

[4] S. Scalise, C. P. Niebla, R. De Gaudenzi, O. del Rio Herrero, D. Finocchiaro, and A. Arcidiacono, "S-MIM: A novel radio interface for efficient messaging services over satellite," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 119–125, Mar. 2013.

[5] M. Luglio, C. Roseti, and F. Zampognaro, "Performance evaluation of TCP-based applications over DVB-RCS DAMA schemes," *Int. J. Satell. Commun. Netw.*, vol. 27, no. 3, pp. 163–191, May 2009.

[6] *IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Standard 802.11-2007 (Rev. of IEEE Std 802.11-1999), IEEE, Jun. 2007, p. 1076.

[7] O. D. R. Herrero and R. De Gaudenzi, "High efficiency satellite multiple access scheme for machine-to-machine communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 2961–2989, Oct. 2012.

[8] R. De Gaudenzi, O. Del Rio Herrero, G. Gallinaro, S. Cioni, and P.-D. Arapoglou, "Random access schemes for satellite networks, from VSAT to M2M: A survey," *Int. J. Satell. Commun. Netw.*, vol. 36, no. 1, pp. 66–107, 2018.

[9] H. Peyravi, "Medium access control protocols performance in satellite communications," *IEEE Commun. Mag.*, vol. 37, no. 3, pp. 62–71, Mar. 1999.

[10] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[11] M. Berioli, G. Cocco, G. Liva, and A. Munari, *Modern Random Access Protocols*. Now Foundations and Trends, 2016.

[12] R. De Gaudenzi, O. Del Río Herrero, S. Cioni, and A. Mengali, *Random Access Versus Multiple Access*. Cham, Switzerland: Springer, 2019, pp. 535–584.

[13] 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.

[14] R. D. Gaudenzi, O. D. R. Herrero, G. Acar, and E. G. Barrabés, "Asynchronous contention resolution diversity ALOHA: Making CRDSA truly asynchronous," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6193–6206, Nov. 2014.

[15] M. Andrenacci, G. Mendola, F. Collard, D. Finocchiaro, and A. Recchia, "Enhanced spread spectrum Aloha demodulator implementation, laboratory tests and satellite validation," *Int. J. Satell. Commun. Netw.*, vol. 32, no. 6, pp. 521–533, Nov. 2014.

[16] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[17] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Feb. 1999.

[18] X. A. Davix and D. Judson, "Successive interference cancellation in asynchronous CC-CDMA systems under rician fading channels," *Telecommun. Syst.*, vol. 72, no. 2, pp. 261–271, Oct. 2019.

[19] R. M. Buehrer, "Equal BER performance in linear successive interference cancellation for CDMA systems," *IEEE Trans. Commun.*, vol. 49, no. 7, pp. 1250–1258, Jul. 2001.

[20] F. Ghanami, G. A. Hodtani, B. Vucetic, and M. Shirvanimoghaddam, "Performance analysis and optimization of NOMA with HARQ for short packet communications in massive IoT," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4736–4748, Mar. 2021.

[21] M. Krondorf, M. Goblirsch, R. de Gaudenzi, G. Cocco, N. Toptsidis, and G. Acar, "Towards the implementation of advanced random access schemes for satellite IoT," *Int. J. Satell. Commun. Netw.*, vol. 38, no. 2, pp. 177–199, Mar. 2020.

[22] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[23] 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.

[24] E. MolavianJazi and J. N. Laneman, "A second-order achievable rate region for Gaussian multi-access channels via a central limit theorem for functions," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6719–6733, Dec. 2015.

[25] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, Jun. 2014.

[26] K. Gatsis, H. Hassani, and G. J. Pappas, "Latency-reliability tradeoffs for state estimation," *IEEE Trans. Autom. Control*, vol. 66, no. 3, pp. 1009–1023, Mar. 2021.

[27] Y. Xu, C. Shen, T.-H. Chang, S.-C. Lin, Y. Zhao, and G. Zhu, "Transmission energy minimization for heterogeneous low-latency NOMA downlink," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1054–1069, Feb. 2020.

[28] Y. Xu, C. Shen, D. Cai, and G. Zhu, "Latency constrained non-orthogonal packets scheduling with finite blocklength codes," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 12312–12316, Oct. 2020.

[29] A. Agarwal, A. K. Jagannatham, and L. Hanzo, "Finite blocklength non-orthogonal cooperative communication relying on SWIPT-enabled energy harvesting relays," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3326–3341, Jun. 2020.

[30] D. V. Djonin and V. K. Bhargava, "Asymptotic analysis of the conventional decision feedback receiver in fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1066–1078, Sep. 2003.

[31] R. De Gaudenzi and O. del Rio Herrero, "Advances in random access protocols for satellite networks," in *Proc. Int. Workshop Satellite Space Commun.*, Sep. 2009, pp. 331–336.

[32] R. De Gaudenzi, O. del Rio Herrero, and G. Gallinaro, "Enhanced spread Aloha physical layer design and performance," *Int. J. Satellite Commun. Netw.*, vol. 32, no. 6, pp. 457–473, 2014.

[33] F. Collard and R. De Gaudenzi, "On the optimum packet power distribution for spread Aloha packet detectors with iterative successive interference cancelation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6783–6794, Dec. 2014.

[34] A. Agrawal, J. G. Andrews, J. M. Cioffi, and T. Meng, "Iterative power control for imperfect successive interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 878–884, May 2005.

[35] F. Molina, J. Sala-Álvarez, F. Rey, and J. Villares, "Channel-aware energy allocation for massive low-rate multiple access," in *Proc. IEEE Int. Conf. Commun.*, May 2019, pp. 1–6.

[36] J. Sala-Alvarez, F. Rey, J. Villares, and F. Molina, "Minimum PER user-energy profile for massive SIC receivers under an average energy constraint," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2017, pp. 1–6.

[37] F. Molina, J. Sala-Álvarez, J. Villares, and F. Rey, "Optimal power control law for equal-rate DS-CDMA networks governed by a successive soft interference cancellation scheme," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2018, pp. 1–5.

[38] J. E. Xu, J.-H. Jong, and C. Ravishankar, "Channel modeling for a land mobile satellite system," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2007, pp. 4596–4600.

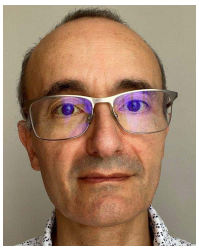
[39] P. K. Sen and J. M. Singer, *Large Sample Methods in Statistics: An Introduction With Applications*. Boca Raton, FL, USA: CRC Press, 2017.

[40] *Universal Mobile Telecommunications Systems (UMTS); Multiplexing and Channel Coding (FDD)*, Standard ETSI TS 125 212 v12.1.0 Release 12, TS ETSI, Jan. 2015.

[41] I. M. Gelfand and R. A. Silverman, *Calculus of Variations*. North Chelmsford, MA, USA: Courier Corporation, 2000.

FRANCESC MOLINA (Member, IEEE) was born in Barcelona, Spain, in 1993. He received the M.Sc. degree in telecommunications engineering and the Ph.D. degree in signal theory and communications from Universitat Politècnica de Catalunya, Barcelona, in 2017 and 2021, respectively.

In 2016, he joined as an Assistant Researcher at the Signal Processing and Communications (SPCOM) Research Group, Universitat Politècnica de Catalunya. From 2018 to 2021, he held the grant for the recruitment of new research staff FI-2018 from the Catalan Administration AGAUR to complete his Ph.D. degree. His research interests include the massive access, wireless communication, and information theory.



JOSEP SALA-ALVAREZ (Senior Member, IEEE) was born in Barcelona, Spain, in 1967. He received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Technical University of Catalonia (UPC), Barcelona, in 1991 and 1995, respectively.

In 1992, he worked at the European Space Operations Centre (ESOC) of the European Space Agency (ESA), Darmstadt, Germany, in the area of software engineering for telemetry processing. From 1993 to 1994, he held a Grant from the Generalitat de Catalunya in support of the Ph.D. degree at the Department of Signal Theory and Communications, UPC. In 1994 he joined UPC as Assistant Professor and was promoted to Associate Professor in 1997. He has participated in space-related communication projects for ESA and in wireless communications projects at the national/European level with industry and institutions. His current research interests include signal processing, communications, and information theory. He was a recipient of the IEEE Signal Processing Society Best (Senior) Paper Award 2003, the International Symposium on Turbo-Codes and Applications (ISTC '03) Best Poster Paper Award, and the Best Ph.D. Thesis in Telecommunications National Award, Spain, in 1995.

JAVIER VILLARES (Senior Member, IEEE) was born in Barcelona, Spain, in 1974. He received the M.S. and Ph.D. degrees in telecommunications engineering from the Technical University of Catalonia (UPC), Barcelona, in 1999 and 2005, respectively.

He has been an Associate Professor with the Department of Signal Theory and Communications, UPC, since 2003. His research interests include digital communications, statistical signal processing, and information theory.

FRANCESC REY (Member, IEEE) was born in Barcelona, Spain, in 1973. He received the M.Sc. and Ph.D. degrees in telecommunication engineering from the Technical University of Catalonia (UPC), Barcelona, in 1997 and 2006, respectively.

In 1998, he joined the Department of Signal Theory and Communications, UPC, where he has been an Associate Professor, since April 2001. From 1998 to 2001, he held a grant from the Generalitat de Catalunya in support of the Ph.D. thesis. His current research interests include signal processing and communications, wireless MIMO channels, and distributed signal processing. He has been involved in many research projects in the framework of the research programs of the European Union, the European Space Agency and the industries.