Decentralized Random Energy Allocation for Massive Non-Orthogonal Code-Division Multiple Access

Francesc Molina, *Student Member, IEEE,* Javier Villares, *Member, IEEE,*
Francesc Rey, *Member, IEEE,* and Josep Sala-Alvarez, *Senior Member, IEEE*

Abstract—This work studies the spectral efficiency achievable when a very large number of terminals are connected simultaneously to a central node (uplink) through independent and identically-distributed flat-fading channels. Assuming that terminals only have statistical channel state information (CSI), the optimum random transmitted-energy allocation is formulated considering a non-orthogonal direct-sequence code-division multiple access (DS-CDMA) where all users transmit using the same modulation and error correcting code and the receiver implements successive interference cancellation (SIC). Focusing on low-power terminals, optimization is carried out by imposing constraints on both the average and peak per-user transmitted energy. Simulations have revealed that a limited number of random energy levels, whose number is determined by the channel power gain variance, is sufficient to achieve approximately the maximum spectral efficiency that would be obtained under direct optimization of the received energy profile.

Index Terms—Non-orthogonal multiple access, DS-CDMA, successive interference cancellation, random energy allocation.

I. Introduction

IN the context of the Internet of Things and 5G applications, new multiple access techniques enabling the connection of a large number of low-rate devices need to be developed [1][2]. In this kind of massive access scenarios, the system throughput is fundamentally limited by multiple-access interference, and advanced interference management strategies are required to boost the system spectral efficiency.

A classical result in information theory states that the corners of the multiple access channel capacity region are achievable by means of non-orthogonal multiple access (NOMA) in conjunction with SIC at the receiver [3, Sec. 6.1]. However, to attain this bound, each terminal has to transmit in the uplink using a distinct capacity-achieving forward error correcting (FEC) code with a coding rate dependent on the instantaneous CSI. Unfortunately, if we focus on applications involving hundreds, or even thousands of users, it is not feasible to generate and distribute this large collection of low-rate near-capacity-achieving codes. A prevalent, practical alternative is to assign

the same (moderate coding rate) FEC code to all users and to distinguish users using different non-orthogonal spreading sequences, following a so-called *code-domain* NOMA [1][2]. Among all the existing code-domain NOMA techniques, in this letter we consider the classical (non-orthogonal) DS-CDMA presented in the seminal paper by Verdú et al. [4]. A prominent implementation of non-orthogonal DS-CDMA, which exploits the multipacket reception capability of SIC, has been developed and standardised in recent times [5][6]. For the interested reader, a comprehensive enumeration of other relevant *code-domain* NOMA techniques is given in [1][2][7].

In the case of non-orthogonal DS-CDMA, it is crucial to unbalance the users' received energy appropriately, as is done in *power-domain* NOMA [8], with the aim of enhancing SIC performance [9] and approaching the sum capacity of DS-CDMA [4]. If every transmitter is aware of its individual (instantaneous) CSI, the optimum user-asymptotic (deterministic) energy allocation was derived in [10] for an arbitrary modulation and coding scheme (MODCOD). In the case of power-domain NOMA without spreading and *two* colliding users having perfect CSI, Xu et al. investigate in [11] the random energy probability distribution that maximizes the SIC throughput when using capacity-achieving codes, concluding, as was already conjectured in [12], that a multilevel discrete energy distribution is the optimum distribution. In the same line as above, a simplified random discrete multilevel energy allocation method is proposed in [13] for the Gaussian channel that extends [11] to deal with an arbitrary number of users.

In this paper, we go one step forward and determine the random energy allocation strategy that maximizes the system spectral efficiency when a *massive* number of DS-CDMA users access the central SIC receiver sharing the same *practical* MODCOD and having only *statistical* CSI. Two main results are obtained in this case: 1) the aggregate spectral efficiency is maximized by transmitting a limited, discrete number of energy levels; 2) random energy allocation achieves roughly the spectral efficiency of deterministic energy allocation with perfect instantaneous CSI [10] as well as the maximum spectral efficiency that could be attained in case of a constant channel and direct energy profile design at reception [14].

II. Problem Statement

We consider K wireless terminals transmitting simultaneously to a central receiver using slotted random access. The received baseband signal in the slot of interest is given by

$$y(t) = \sum_{k=1}^{K} \sqrt{E_x[k]h[k]} \sum_i a_i[k]c_{k,i}(t - iT - \tau[k]) + w(t) \quad (1)$$

where $E_x[k]$ is the (average) energy per symbol transmitted by user k, $h[k]$ the channel (power) gain between user k and the

central receiver, $a_i[k]$ is the sequence of unit-variance coded symbols sent by user k, $c_{k,i}(t)$ the unit-energy waveform of effective duration T and limited bandwidth B assigned to user k during the transmission of symbol i, and $w(t)$ complex zero-mean white Gaussian noise of two-sided power spectral density N_0. The information of all users is encoded using the same practical modulation and coding scheme (MODCOD), which packs R_c information bits in every transmitted symbol $a_i[k]$. Thus, all users transmit at the same fixed rate: R_c information bits per channel use.

We consider a DS-CDMA system based on *long* pseudo-random spreading sequences spanning many symbol intervals [4]. The long spreading signal $c_k(t) \doteq \sum_i c_{k,i}(t - iT)$ assigned to user k fulfils $\int c_{k,i}(t)c_{k',i'}(t)dt = \frac{1}{SF}$ for $k \neq k'$ and all i, i' with $SF \doteq B \cdot T$ the spreading factor [4]. For the sake of clarity, we restrict ourselves to a quasi-synchronous scenario in which users are received with almost identical delay and the receiver is correctly synchronized, i.e. $\tau[k] \approx 0$.

At every time slot, each user selects randomly its transmitted symbol energy $E_x[k]$ from the ordered set $\{0, x_1, x_2, \ldots, x_M\}$, with probabilities $\{p_x[m]\}_{m=0,\ldots,M}$, where $p_x[0]$ is the backoff probability associated with the null energy level. Accordingly, there will be, on average, $K_0 \doteq (1 - p_x[0])K$ users that decide to transmit during the slot (active users) and $K - K_0 = p_x[0]K$ that decide to remain silent (idle users).

A frequency-flat slow block fading channel model is considered, in which the channel (power) gain $h[k]$ of user k is constant during the time slot. Moreover, the K channel gains are independent and identically distributed according to a known probability distribution function $F_h(h) \doteq \Pr(h[k] \leq h)$.

At every time slot, the K received packets are sorted according to the received symbol energy $E_y[k] \doteq E_x[k]h[k]$, and decoded sequentially from the strongest to the weakest user. If all users share the same MODCOD and thus transmit at the same rate, this decoding order is the one that maximizes the aggregate spectral efficiency. Moreover, to enable mathematical tractability, we consider that users are always decoded in the correct order despite errors in the estimation of the received symbol energies $\{E_y[k]\}_{k=1,\ldots,K}$. Finally, note that the order of individual users changes at every time slot due to the randomness of $E_y[k]$. With some abuse of notation, index k will be used hereinafter to reindex users at every time slot according to their instantaneous received energy $E_y[k]$ from $k = K$ (strongest user) to $k = 1$ (weakest user).

After ordering the K users, the following SIC policy is implemented; if packet k is decoded successfully, which is checked using a Cyclic Redundancy Check (CRC) code[1], the signal of user k is reconstructed and subtracted from the received signal $y(t)$. Due to reconstruction errors, cancellation will be imperfect in that case. Otherwise, if the packet is not correctly decoded, this user is not cancelled and the receiver proceeds with the next user. For this SIC architecture [14], the SINR of user k is given by the following expression:

$$\Gamma[k] = \frac{E_y[k]}{N_0 + \frac{1}{SF} \sum_{j=1}^{k-1} E_y[j] + \frac{1}{SF} \sum_{j=k+1}^{K} \epsilon[j]E_y[j]} \quad (2)$$

[1]To simplify the analysis, we will consider hereinafter that the CRC code is error-free, that is, that its misdetection probability is null.

where $\epsilon[k]$ stands for the fraction of residual interference energy from user k that remains in the received signal $y(t)$ after processing this user. To characterize $\epsilon[k]$, two positive decreasing functions of the SINR are required: RI(Γ), which models the fraction of residual interference energy after imperfect cancellation, and PSR(Γ) $\doteq 1 - $ PER(Γ), which is the Packet Success Rate (PSR) curve of the selected MODCOD (R_c information bits per symbol). Thus, in the adopted SIC policy, $\epsilon[k]$ is a Bernouilli random variable that can take two different values: if user k is correctly decoded (CRC ok), $\epsilon[k]$ is equal to RI($\Gamma[k]$) < 1 (imperfect cancellation) with probability PSR($\Gamma[k]$), whereas, if user k is incorrectly decoded, $\epsilon[k] = 1$ (no cancellation) with probability PER($\Gamma[k]$).

To facilitate the system analysis and optimization in Sec. III, the received symbol energies $\{E_y[k]\}_{k=1,\ldots,K}$ are discretized using $N + 1$ equally-spaced energy levels $\{n\Delta_y\}_{n=0,\ldots,N}$ with Δ_y a sufficiently small step size, and N a large integer that limits the probability of $E_y[k]$ exceeding the maximum energy level $N\Delta_y$. Thus, the probability of quantizing $E_y[k]$ with the nth quantization level $y_n \doteq n\Delta_y$ is given by

$$p_y[n] = \sum_{m=0}^{M} p_h(y_n|x_m)p_x[m] \quad (3)$$

with

$$p_h(y_n|x_m) \doteq \begin{cases} \delta(y_n) & x_m = 0 \\ F_h\left(\frac{y_n + \Delta_y/2}{x_m}\right) - F_h\left(\frac{y_n - \Delta_y/2}{x_m}\right) & x_m \neq 0 \end{cases} \quad (4)$$

If the number of users and the spreading factor go to infinity while their ratio $\alpha \doteq K/SF$ (load factor) remains constant, from the law of large numbers[2] and after some manipulations, we obtain the following asymptotic approximation to the SINR of all the users whose received energy $E_y[k]$ is quantized with the energy level $y_n \doteq n\Delta_y$:

$$\Gamma_n \doteq \frac{y_n}{N_0 + \alpha\bar{E}_y - \alpha\sum_{i=n+1}^{N}(1 - \text{RI}(\Gamma_i))\text{PSR}(\Gamma_i)y_ip_y[i]} \quad (5)$$

where $\bar{E}_y = \sum_{n=1}^{N} y_np_y[n]$ is the per-user long-term average received symbol energy. Note that the negative term in the denominator of (5) corresponds to the interference energy that the SIC has removed after processing all the stronger users ($i > n$) but none of the $Kp_y[n]$ users that are received with the same quantized energy level ($i = n$). Thus, the approximation in (5) is actually a lower bound to the true SINR, which is tight if Δ_y is sufficiently small.

After some straightforward manipulations, it can be shown that Γ_n can be computed iteratively, starting at $n = N - 1$ and running the following nonlinear difference equation:

$$\Gamma_n = \Gamma_{n+1}\frac{c_n}{1 - \alpha\Phi(\Gamma_{n+1})p_y[n+1]} \quad (6)$$

where $c_n \doteq n/(n+1)$, $\Phi(\Gamma) \doteq (1 - \text{RI}(\Gamma))\text{PSR}(\Gamma)\Gamma$ is an implementation-dependent function of the SINR, $p_y[n]$ is given in (3) and, the initial value of (6) is $\Gamma_N = \frac{N\Delta_y}{N_0 + \alpha\bar{E}_y}$.

[2]If the number of users K goes to infinity, the number of users that are received with energy level y_n converges to $Kp_y[n]$. In addition, the ratio K/SF converges to constant α (load factor). Note that α can exceed 1 in the studied NOMA setup while it is limited to 1 ($K \leq SF$) in the case of code-domain orthogonal multiple access (OMA) [2].

The objective of this letter is to find the optimum probability distribution $\{p_x[m]\}_{m=0,\ldots,M}$ and the associated energy levels $\{x_m\}_{m=1,\ldots,M}$ that maximize the asymptotic *aggregate* spectral efficiency (bps/Hz):

$$SE_\infty \doteq \lim_{K,SF\to\infty} \frac{R_c}{SF} \sum_{k=1}^{K} \mathrm{PSR}(\Gamma[k]) = \alpha R_c \sum_{n=1}^{N} \mathrm{PSR}(\Gamma_n)p_y[n] \tag{7}$$

subject to a *maximum* transmitted symbol energy constraint, $x_M \le E_{\max}$, and a per-user long-term *average* transmitted symbol energy constraint:

$$\sum_{m=1}^{M} x_m p_x[m] \le \bar{E}_x. \tag{8}$$

Note that, although the instantaneous spectral efficiency of individual users changes from slot to slot due to the randomness of the received symbol energy, in the studied setup, all users achieve eventually the same *long-term* average spectral efficiency when the average is carried out over many time slots. Thus, long-term fairness is ensured to all the network terminals. The long-term average spectral efficiency of any user is therefore a fraction $1/K$ of the aggregate spectral efficiency (7). Moreover, the normalized throughput, which is a traditional performance metric in the random access literature, is given by the product of (7) and the constant factor $1/(R_c\alpha)$.

III. PROBLEM SOLUTION

By introducing the following vector definitions

$$\mathbf{p}_x \doteq [p_x[1], p_x[2], \ldots, p_x[M]]^T \tag{9}$$
$$\mathbf{x} \doteq [x_1, x_2, \ldots, x_M]^T \tag{10}$$
$$\boldsymbol{\Gamma} \doteq [\Gamma_1, \Gamma_2, \ldots, \Gamma_N]^T, \tag{11}$$

the asymptotic spectral efficiency SE_∞ in (7) can be recast using vectorial notation as follows

$$SE_\infty = \mathbf{v}_1^T(\boldsymbol{\Gamma})\mathbf{H}_1(\mathbf{x})\mathbf{p}_x \tag{12}$$

where

$$[\mathbf{v}_1(\boldsymbol{\Gamma})]_n \doteq \alpha R_c \mathrm{PSR}(\Gamma_n) \tag{13}$$
$$[\mathbf{H}_1(\mathbf{x})]_{n,m} \doteq p_h(y_n|x_m) \tag{14}$$

are the nth element of vector $\mathbf{v}_1(\boldsymbol{\Gamma})$ and the element in the nth row and mth column of $\mathbf{H}_1(\mathbf{x})$, respectively.

Regarding the SINR vector $\boldsymbol{\Gamma}$ in (12), it follows from (6) that $\boldsymbol{\Gamma}$ is an *implicit* nonlinear function of $\mathbf{p}_x$ and $\mathbf{x}$, that is solved recursively by iterating (6). This equation constitutes a constraint on the set of admissible SINR vectors that, after rearranging terms and stacking below the initialization of Γ_N, can be compacted into the following vectorial equation:

$$\mathbf{f}(\boldsymbol{\Gamma}, \mathbf{p}_x, \mathbf{x}) \doteq \mathbf{v}_2(\boldsymbol{\Gamma}) - \alpha \begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{\Gamma})\mathbf{H}_2(\mathbf{x}) \\ \mathbf{e}_N^T\mathbf{H}_1(\mathbf{x}) \end{bmatrix} \mathbf{p}_x = \mathbf{0} \tag{15}$$

with the following definitions:

$$[\mathbf{v}_2(\boldsymbol{\Gamma})]_n \doteq \begin{cases} 1 - c_n\frac{\Gamma_{n+1}}{\Gamma_n} & 1 \le n < N \\ N/\Gamma_N - N_0/\Delta_y & n = N \end{cases} \tag{16}$$

$$[\boldsymbol{\Phi}(\boldsymbol{\Gamma})]_{n,n'} \doteq \begin{cases} \Phi(\Gamma_{n+1}) & n = n' \\ 0 & n \ne n' \end{cases} \tag{17}$$

$$[\mathbf{H}_2(\mathbf{x})]_{n,m} \doteq p_h(y_{n+1}|x_m) \tag{18}$$

Consequently, the optimization problem at hand reads

$$\max_{\mathbf{p}_x, \mathbf{x}} \mathbf{v}_1^T(\boldsymbol{\Gamma})\mathbf{H}_1(\mathbf{x})\mathbf{p}_x \tag{19}$$

subject to equation (15) and the following set of constraints:

$$\mathbf{x}^T\mathbf{p}_x \le \bar{E}_x \tag{20}$$
$$x_M \le E_{\max} \tag{21}$$
$$p_x[0] = 1 - \mathbf{1}^T\mathbf{p}_x \ge 0 \tag{22}$$
$$\mathbf{p}_x \ge \mathbf{0} \tag{23}$$
$$\mathbf{Dx} < \mathbf{0} \tag{24}$$

to which we have appended the following definitions:

$$\mathbf{e}_N \doteq [1, 2, \ldots, N]^T \tag{25}$$
$$\mathbf{1} \doteq [1, 1, \ldots, 1]^T \tag{26}$$
$$[\mathbf{D}]_{m,m'} = \begin{cases} -1 & m = m' \\ 1 & m = m' + 1 \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

Note that equation (20) is the vectorial counterpart of (8), which imposes the per-user long-term average transmitted energy. The maximum transmitted energy is limited by (21). The rest of equations are structural constraints imposing non-negative backoff probability (22), probability positiveness (23) and energy levels ordering, i.e. $0 < x_1 < \ldots < x_M$ (24).

Finally, omitting the argument of $\mathbf{H}_1(\mathbf{x})$ and $\mathbf{v}_1(\boldsymbol{\Gamma})$ for the sake of clarity, the gradient of (19) can be written as follows:

$$\frac{dSE_\infty}{d\mathbf{p}_x} = \left(\frac{d\mathbf{v}_1}{d\mathbf{p}_x^T}\right)^T \mathbf{H}_1\mathbf{p}_x + \mathbf{H}_1^T\mathbf{v}_1 \tag{28}$$

$$\frac{dSE_\infty}{d\mathbf{x}} = \left(\frac{d\mathbf{v}_1}{d\mathbf{x}^T}\right)^T \mathbf{H}_1\mathbf{p}_x + \mathbf{Dg}(\mathbf{p}_x)\dot{\mathbf{H}}_1^T\mathbf{v}_1 \tag{29}$$

where $\mathbf{Dg}(\mathbf{p}_x)$ is a diagonal matrix whose main diagonal is $\mathbf{p}_x$, and matrices $\frac{d\mathbf{v}_1}{d\mathbf{p}_x^T}$, $\frac{d\mathbf{v}_1}{d\mathbf{x}^T}$ and $\dot{\mathbf{H}}_1$ are derived in Appendix A.

IV. NUMERICAL RESULTS

In this Section, the transmitted-energy probability distribution is optimized in a typical wireless environment in which the signal suffers log-normal shadowing. In particular, we consider that the channel (power) gain of all users is independent and follows the log-normal distribution $F_h(h) = 1 - Q\left(\frac{\ln h - \mu}{0.1\ln(10)\sigma}\right)$ with $Q(\cdot)$ the Gaussian tail probability function. The parameter μ is set to $-\ln^2(10)\sigma^2/200$ in order to fix the average channel gain $\mathbb{E}\{h\}$ to 1 whereas the parameter σ (log-gain standard deviation *in dB*) is varied from 0 (constant channel) to 4dB in order to test different degrees of shadowing.

The average and maximum energy per symbol of all users is set to $\bar{E}_x = 0.25$ and $E_{\max} = 1$, respectively. The average E_s/N_0 ($\bar{E}_x/N_0$) is set to 20dB (interference-limited scenario). The received energy is discretized in the interval $[0, y_{\max}]$ using $N + 1 = 500 \cdot y_{\max}$ energy levels where $y_{\max} = h_{\max}E_{\max}$ and $h_{\max}$ is selected so that $F_h(h_{\max}) = 1 - 10^{-3}$.

As in [14][16][17], a simple imperfect cancellation model is adopted in which the fraction of residual interference energy is independent of the SINR. In particular, we consider $\mathrm{RI}(\Gamma) = 0.1$ because this is the simulated value in [14]. In that reference, the aggregate spectral efficiency is optimized for a
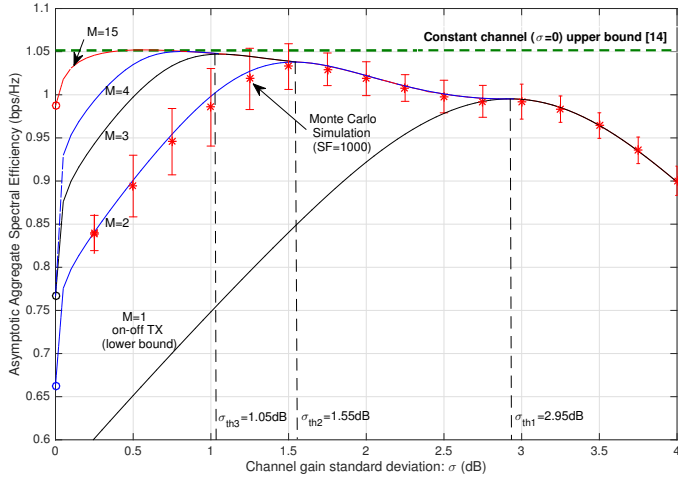
Fig. 1. Asymptotic aggregate spectral efficiency as a function of the shadowing standard deviation σ for different values of M and $\alpha = 1.5$. The asymptotic aggregate spectral efficiency derived in [14] for the constant channel ($\sigma = 0$) is also represented (green dashed upper line). For finite SF, a Monte Carlo simulation for $M = 2$ and $SF = 1000$ is superimposed using asterisks and vertical interval bars.

Fig. 2. Received energy probability distribution for several values of M and $\alpha = 1.5$. For each M, the channel standard deviation σ is set to σ_{th_M}, which is the value that maximizes the spectral efficiency in Fig. 1. The optimal energy distribution for $\sigma = 0$ (constant channel) in [14] is also depicted (green dashed curve). Note that the backoff probability $p_y[0] \approx p_x[0]$ is not plotted here but is given in Fig. 3.

constant deterministic channel (i.e., $h[k] = 1$ for all k) and it constitutes therefore an upper-bound to the spectral efficiency that can be achieved when dealing with independent random channel gains, as considered in this paper.

In order to exemplify the optimization in Sec. III, in this section we show numerical results for a typical MODCOD consisting of QPSK and a convolutional code of rate 1/2 and constraint length 7. While QPSK is the usual choice in low-rate reliable communications, convolutional coding is selected because its reduced decoding latency, which makes it suitable for the studied large-user setup. Note, however, that optimization in Sec. III can be applied to any other MODCOD provided that its PSR(Γ) curve is available.

The optimization problem in Section III is solved numerically with Matlab using a state-of-the-art Sequential Quadratic Programming (SQP) algorithm [15, Sec. 15.3]. To accelerate its execution, the solver is assisted with the closed-form expression of the cost function gradient given in (28) and (29).

In Fig. 1, the asymptotic aggregate spectral efficiency (7) is depicted as a function of the shadowing log-gain standard deviation σ for different values of M. For every value of M and σ, the optimum energy allocation is determined (Sec. III). It is found that on-off transmission (i.e., $M = 1$) is optimum for σ greater than a scenario-dependent threshold σ_{th_1}. For lower values of σ, the number of active energy levels depends on σ. In particular, $M > 1$ energy levels are optimum for σ in the interval $\sigma_{th_M} < \sigma < \sigma_{th_{M-1}}$ where σ_{th_M} stands for the value of σ that maximizes the spectral efficiency for M active transmitted levels. If the channel is constant ($\sigma = 0$), the *continuous* energy profile derived in [14] is optimum and is attained asymptotically as $M \to \infty$. Otherwise, for any random channel ($\sigma > 0$), we conclude that the optimum energy distribution is *discrete*, as already announced in [11][12]. Moreover, it is evidenced in Fig. 1 that some degree of channel randomness allows to attain practically the maximum spectral
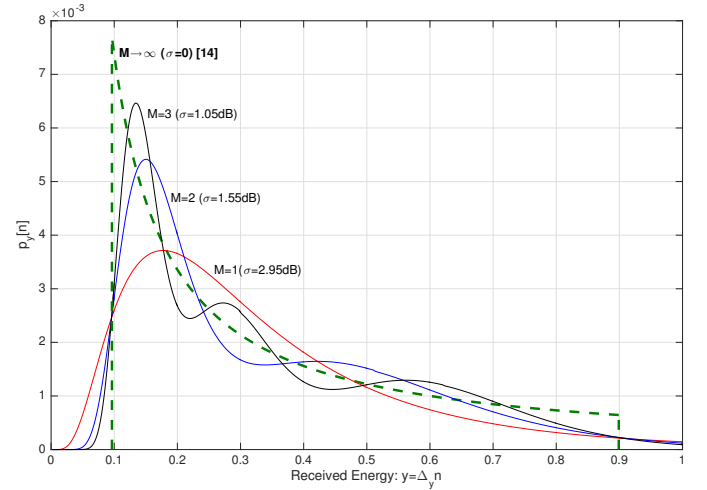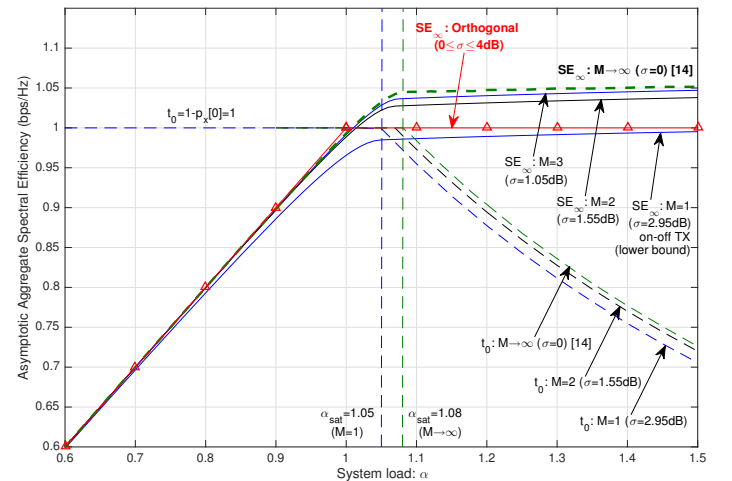

Fig. 3. Asymptotic aggregate spectral efficiency as a function of the load factor α for different values of M. For each M, the channel gain standard deviation σ is the one maximizing the spectral efficiency in Fig. 1. The maximum spectral efficiency for $\sigma = 0$ (constant channel) given in [14] is represented (green dashed upper curve). The transmission probability $t_0 \doteq 1 - p_x[0]$ is superimposed showing how t_0 decreases when the system enters into saturation ($\alpha > \alpha_{\text{sat}}$). Lastly, the aggregate spectral efficiency of *orthogonal* DS-CDMA is represented ($\triangle$ markers). In the orthogonal case, the denominator of the SINR (5) is interference-free for $\alpha \leq 1$, where we have found that on-off transmission ($M = 1$) is optimum for any value of σ. For $\alpha > 1$, orthogonal DS-CDMA does not admit more than SF users and it is surpassed by non-orthogonal DS-CDMA.

efficiency [14] with a reduced number of transmitted energy levels (e.g., $M = 3$), although the received energy distribution for small M is only a rough approximation to the optimum one, as illustrated in Fig. 2.

In order to validate the user-asymptotic analysis, we have computed, via Monte Carlo simulations, the aggregate spectral efficiency that is achieved by a large but finite number of users. In particular, we have simulated a system with $K = 1500$ users

($SF = 1000$, $\alpha = 1.5$) that transmit using $M = 2$ non-zero energy levels. In the finite-user case, the aggregate spectral efficiency changes in every slot randomly. For this reason, in Fig. 1, we have plotted the mean value (red asterisks) and $\pm$ the standard deviation (interval bars) of the aggregate spectral efficiency computed over 1000 time slots.

Finally, the asymptotic aggregate spectral efficiency is represented in Fig. 3 as a function of the load factor α for different values of M. For whichever value of M, the spectral efficiency curve exhibits an inflection point at $\alpha = \alpha_{\mathrm{sat}}$, as indicated in Fig. 3. When the system enters into saturation ($\alpha > \alpha_{\mathrm{sat}}$), the number of *active* terminals in a given slot, $K_0 = (1 - p_x[0])K$, is automatically limited to $K_{\mathrm{sat}} \doteq \alpha_{\mathrm{sat}} SF$ without the need of adding a new constraint to the optimization program. The remaining $K - K_0 = p_x[0]K$ terminals stay silent during the slot, saving energy to transmit in subsequent time slots. The saturation load α_{sat} is characteristic of the adopted MODCOD as well as other system parameters (e.g., E_s/N_0). However, it is quite insensitive to M. In the simulated scenario (Fig. 3), α_{sat} ranges from 1.05 ($M = 1$) to 1.08 ($M \to \infty$).

The probability of transmission $t_0 \doteq 1 - p_x[0]$ is also represented in Fig. 3 to visualize how users refrain from transmitting when the system is overloaded ($\alpha > \alpha_{\mathrm{sat}}$). Note that t_0 can be interpreted as the duty-cycle of the terminals, i.e., the fraction of time that users are active on average.

V. CONCLUSION

Discrete random energy allocation is shown to be optimum when a massive number of users access a central receiver in the presence of random frequency-flat fading. If users know the channel gain variance, they can select the number of transmitted energy levels in an optimum way to maximize the system spectral efficiency. For moderate channel gain variance, a reduced number of energy levels suffices to reach approximately the maximum spectral efficiency attainable in the case of a deterministic constant channel.

APPENDIX A
GRADIENT CALCULATION

Using the chain rule, we have that

$$\frac{d\mathbf{v}_1}{d\mathbf{p}_x^T} = \frac{d\mathbf{v}_1}{d\mathbf{\Gamma}^T} \frac{d\mathbf{\Gamma}}{d\mathbf{p}_x^T}. \tag{30}$$

The first factor in (30) is given by

$$\left[\frac{d\mathbf{v}_1}{d\mathbf{\Gamma}^T}\right]_{n,n'} = \begin{cases} \alpha R_c \, \mathrm{PSR}'(\Gamma_n) & n = n' \\ 0 & n \neq n' \end{cases} \tag{31}$$

where $\mathrm{PSR}'(\Gamma) \doteq \frac{d\mathrm{PSR}(\Gamma)}{d\Gamma}$. The second factor in (30) is obtained invoking the implicit function theorem:

$$\frac{d\mathbf{\Gamma}}{d\mathbf{p}_x^T} = -\left(\frac{\partial \mathbf{f}}{\partial \mathbf{\Gamma}^T}\right)^{-1} \frac{\partial \mathbf{f}}{\partial \mathbf{p}_x^T} = \alpha \left(\frac{\partial \mathbf{f}}{\partial \mathbf{\Gamma}^T}\right)^{-1} \begin{bmatrix} \mathbf{\Phi}(\mathbf{\Gamma})\mathbf{H}_2 \\ \mathbf{e}_N^T \mathbf{H}_1 \end{bmatrix} \tag{32}$$

where $\mathbf{f}(\mathbf{\Gamma}, \mathbf{p}_x, \mathbf{x})$ is the implicit function imposed in (15) and

$$\left[\frac{\partial \mathbf{f}}{\partial \mathbf{\Gamma}^T}\right]_{n,n'} = \begin{cases} -\frac{N}{\Gamma_N^2} & n' = n = N \\ \frac{c_n \Gamma_{n+1}}{\Gamma_n^2} & n' = n \neq N \\ -\frac{c_n}{\Gamma_n} - \alpha \Phi'(\Gamma_{n+1})\left[\mathbf{H}_2\mathbf{p}_x\right]_n & n' = n + 1 \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

where $c_n \doteq n/(n + 1)$ and $\Phi'(\Gamma) \doteq \frac{d\Phi(\Gamma)}{d\Gamma}$ is the derivative of the function $\Phi(\Gamma)$ introduced previously.

Regarding vector $\mathbf{x}$, we have that

$$\frac{d\mathbf{v}_1}{d\mathbf{x}^T} = \frac{d\mathbf{v}_1}{d\mathbf{\Gamma}^T} \frac{d\mathbf{\Gamma}}{d\mathbf{x}^T} = \alpha \frac{d\mathbf{v}_1}{d\mathbf{\Gamma}^T} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{\Gamma}^T}\right)^{-1} \begin{bmatrix} \mathbf{\Phi}(\mathbf{\Gamma})\dot{\mathbf{H}}_2 \\ \mathbf{e}_N^T \dot{\mathbf{H}}_1 \end{bmatrix} \mathbf{Dg}\left(\mathbf{p}_x\right) \tag{34}$$

and $\dot{\mathbf{H}}_1$ and $\dot{\mathbf{H}}_2$ are $N \times M$ and $N - 1 \times M$ matrices defined as

$$\left[\dot{\mathbf{H}}_1\right]_{n,m} \doteq \frac{d\left[\mathbf{H}_1\right]_{n,m}}{dx_m} = \frac{y_n - \Delta_y/2}{x_m^2} f_h\left(\frac{y_n - \Delta_y/2}{x_m}\right)$$
$$- \frac{y_n + \Delta_y/2}{x_m^2} f_h\left(\frac{y_n + \Delta_y/2}{x_m}\right) \tag{35}$$

$$\left[\dot{\mathbf{H}}_2\right]_{n,m} \doteq \frac{d\left[\mathbf{H}_2\right]_{n,m}}{dx_m} = \left[\dot{\mathbf{H}}_1\right]_{n+1,m} \tag{36}$$

with $f_h(h) \doteq \frac{dF_h(h)}{dh}$ the channel probability density function.

REFERENCES

[1] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan and L. Hanzo, "Nonorthogonal Multiple Access for 5G and Beyond," *Proceedings of the IEEE*, vol. 105, pp. 2347-2381, Dec. 2017.

[2] L. Dai, B. Wang, Y. Yuan, S. Han, C. I and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," in *IEEE Commun. Mag.*, vol. 53, pp. 74-81, Sep. 2015.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge, UK: Cambridge University Press, 2005.

[4] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. on Inf. Theory*, vol. 45, pp. 622-640, Mar. 1999.

[5] R. De Gaudenzi, O. del Rio Herrero, and G. Gallinaro. "Enhanced spread Aloha physical layer design and performance," *Int. Journal of Satellite Commun. and Networking*, vol. 32.6, pp. 457-473, 2014.

[6] ETSI TS, "Satellite Earth Stations and Systems; Air Interface for S-band Mobile Interactive Multimedia (S-MIM)," ETSI TS 102 721, Dec. 2011.

[7] Q. Wang, R. Zhang, L. Yang and L. Hanzo, "Non-Orthogonal Multiple Access: A Unified Perspective," in *IEEE Wireless Communications*, vol. 25, pp. 10-16, Apr. 2018.

[8] S. M. R. Islam, N. Avazov, O. A. Dobre and K. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 19, pp. 721-742, Secondquarter 2017.

[9] D. V. Djonin and V. K. Bhargava, "Asymptotic analysis of the conventional decision feedback receiver in fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, pp. 1066-1078, Sep. 2003.

[10] F. Molina, J. Sala-Alvarez, F. Rey and J. Villares, "Channel-aware energy allocation for throughput maximization in massive low-rate multiple access," in *IEEE 53rd Int. Commun. Conf. (ICC)*, Shanghai, May 2019.

[11] C. Xu, L. Ping, P. Wang, S. Chan and X. Lin, "Decentralized Power Control for Random Access with Successive Interference Cancellation," *IEEE J. on Sel. Areas in Commun.*, vol. 31, pp. 2387-2396, Nov. 2013.

[12] R. LaMaire, A. Krishna, and M. Zorzi, "On the randomization of transmitter power levels to increase throughput in multiple access radio systems," in *Wireless Networks*, vol. 4, pp. 263-277, March 1998.

[13] H. Lin, K. Ishibashi, W. Shin and T. Fujii, "A Simple Random Access Scheme With Multilevel Power Allocation," *IEEE Communications Letters*, vol. 19, pp. 2118-2121, Dec. 2015.

[14] 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 IEEE 18th Int. Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Sapporo, Jul. 2017.

[15] A. Antoniou and W.S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, New York, NY, USA: Springer, 2007.

[16] F. Collard and R. De Gaudenzi, "On the Optimum Packet Power Distribution for Spread Aloha Packet Detectors With Iterative Successive Interference Cancelation," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 6783-6794, Dec. 2014.

[17] J. Andrews, T. Meng, "Optimum power control for successive interference cancellation with imperfect channel estimation", in *IEEE Trans. on Wireless Communications*, vol. 2, pp. 375-383, Mar. 2003.