Resource Allocation in Low Density Spreading Uplink NOMA via Asymptotic Analysis

Hossein Asgharimoghaddam and Antti Tölli, *Senior Member, IEEE*
Centre for wireless communications (CWC), University of Oulu, Oulu, Finland
{hossein.asgharimoghaddam, antti.tolli}@oulu.fi

Abstract—Low density spreading non-orthogonal multiple-access (LDS-NOMA) is considered where K single-antenna user equipments (UEs) communicate with a base station (BS) over F fading sub-carriers. Each UE k spreads its data symbol over $d_k < F$ sub-carriers. Given $d_k, \forall k$ as design parameters, we characterize the resource allocation solutions that closely maximize the ergodic mutual information (EMI) in a scenario where the BS assigns resources solely based on the UEs' pathlosses. Conducting analysis in asymptotic limit where F, K, and $d_k, \forall k$ converge to $+\infty$ at the same rate, we present EMI in terms of a deterministic equivalent plus a residual term. The deterministic equivalent is given in terms of pathloss values and LDS-codes, and the small residual term scales as $\mathcal{O}(\frac{1}{d^2})$ where $d = \min\{d_k, \forall k\}$. We formulate an optimization problem to get the set $\bar{\mathcal{C}}^*$ of all spreading codes, irrespective of sparsity constraints, which maximize the deterministic equivalent of EMI. The spreading codes in $\bar{\mathcal{C}}^*$ with desired sparsity are obtained via a simple and efficient algorithmic solution. In the finite regime, the residual term is shown to be a small incremental gain for the sparse solutions in $\bar{\mathcal{C}}^*$, which is dictated mainly by $d_k, \forall k$ values. Accordingly, we show that the solutions in $\bar{\mathcal{C}}^*$ with desired sparsity yield close to optimum values of EMI in the finite regime. Numerical simulation validates the attainable spectral efficiency enhancement as compared to regular, and random spreading.

I. INTRODUCTION

In low density spreading non-orthogonal multiple-access (LDS-NOMA), an LDS code comprising a small number of d_k non-zero elements is employed for linearly modulating user k's symbol over a number of F shared radio resources. The sparse mapping between K user equipments (UEs) and F resources in LDS-NOMA can be either regular, where each UE occupies $d_k = d, \forall k$ resources, and each resource is used by $\frac{K}{F}d$ UEs; or irregular otherwise [1]. The objective here is to investigate spectral efficiency (SE) limits of LDS-NOMA.

The SE limits of LDS-NOMA with spreading in time domain have been studied in [1]–[6] under symmetric AWGN channel model. The irregular schemes with d_k being randomly Poissonian distributed with fixed mean [4], [5], and randomly uniformly distributed [3] are studied using replica method [7] and random matrix framework developed in [8], respectively. The regular scheme is considered in [1], [2] where in [2] a closed-form approximation is given for the optimum SE. These analyses indicate that the regular codes, in symmetric AWGN channel, yield superior SE limits as compared to irregular and dense spreading (the case with $d_k = F, \forall k$). The aforementioned works rely on the analysis of random

matrices in large system regime [7], [9] where F grows large with a fixed ratio of K/F. Such analysis yields rather accurate approximations in the finite regime that become arbitrarily tight as F grows large. Since the mathematical literature studying the limiting behaviour of sparse random matrices is distinctly smaller than that for non-sparse random matrices [10], the analysis in more generic settings is rather limited. The SE analysis in presence of fading is considered in [11] in a special setting with $d_k = 1, \forall k$.

In this paper, a multi-carrier LDS-NOMA scheme with spreading in frequency domain [12], [13] is considered. Different from the aforementioned works, the UEs are allowed to have distinct pathloss values, and fading is imposed on sub-carriers. Also, instead of assuming a particular sparse mapping, we consider $d_k, \forall k$ as design parameters, and identify the LDS-code allocation policies that closely attain the maximum of the ergodic mutual information (EMI). A key feature of these policies is that they assign the codes only based on pathloss values. Conducting analysis in large system limits where F, K, and $d_k, \forall k$ converge to $+\infty$ at the same rate, we present EMI in terms of a deterministic equivalent plus a residual term. The deterministic equivalent is given in terms of pathlosses and LDS-codes, and the small residual term is shown to quickly vanish inversely proportional to d^2 where $d = \min\{d_k, \forall k\}$. We first formulate an optimization problem to get the set $\bar{\mathcal{C}}^*$ of all spreading codes, irrespective of sparsity constraints, which maximize the deterministic equivalent of EMI. The desired sparse spreading codes in $\bar{\mathcal{C}}^*$ are obtained via a simple and efficient algorithmic solution. Conducting numerical analysis in the finite regime with a moderate number of sub-channels F, we observe that the residual term appears as a small incremental gain in EMI for the sparse solutions in $\bar{\mathcal{C}}^*$, which is dictated mainly by $d_k, \forall k$ values. Thus, interpreting the residual term as sparsity gain[1] and with a line of argument, we show that the solutions in $\bar{\mathcal{C}}^*$ with desired sparsity attain close to the optimum values of EMI in the finite regime. It is observed that regular spreading is asymptotically optimal for the symmetric scenarios with the same pathlosses for all UEs. However, in the generic asymmetric scenarios, an irregular structure might arise. Numerical simulation validates the attainable SE enhancement as compared to random and regular spreading.

---

[1]We refer to the residual term as sparsity gain in contrast to the loss associated with spreading in the considered system model. See [14] for spreading-coding trade-off.

II. SYSTEM MODEL AND PROBLEM STATEMENT

Consider an uplink multi-carrier system where single-antenna UEs $\mathcal{K} = \{1, ..., K\}$ communicate with a base station (BS) on a common frequency band. Exploiting the OFDM technique, the total frequency band is divided into a set of narrow band sub-channels $\mathcal{F} = \{1, ..., F\}$. Then, each UE k spreads its data symbol in frequency direction using an LDS code $\mathbf{w}_k \in \mathbb{R}^F$. The code is a sparse vector consisting of F chips with d_k, a small number, of non-zero values. The UEs' codes are not restricted to be orthogonal. Thus, in an overloaded scenario, the signals of UEs transmitting on the same sub-channel are superimposed. The vector of received signal within F sub-channels is

$$\mathbf{r} = \sum_{k \in \mathcal{K}} \frac{\rho}{\sqrt{d_k}} \text{diag}\{\mathbf{w}_k\}(a_k \mathbf{g}_k) s_k + \mathbf{n} \tag{1}$$

where normalized noise vector is given by $\mathbf{n} \sim \mathcal{CN}(0, \mathbf{I}_F)$, unit variance symbol of kth UE is denoted by s_k, and signal-to-noise-ratio (SNR) at the transmitters is represented by ρ. The channel vector for UE k is denoted by $a_k \mathbf{g}_k$ where a_k^2 includes the pathloss due to large scale fading. The entries of $\mathbf{G} = [\mathbf{g}_1, ... \mathbf{g}_K] \in \mathbb{C}^{F \times K}$ are independent complex Gaussian random variables. Each entry has zero-mean independent real and imaginary parts with variance of $\frac{1}{2}$. The assumptions are based on uncorrelated fading channel model utilized in the context of multi-carrier systems [15, Chapter 1] where the fadings on adjacent data symbols after inverse OFDM and de-interleaving are considered as uncorrelated [15]. This holds when, for example, a frequency and time interleaver with sufficient depth is applied in a frequency-selective fading channel. Hereafter, $\mathbf{h}_k = \frac{a_k}{\sqrt{d_k}} \text{diag}\{\mathbf{w}_k\}\mathbf{g}_k$ and $\mathbf{H} = [\mathbf{h}_1, ..., \mathbf{h}_K]$ denote the equivalent channel vectors and matrix, respectively.

A. Ergodic Capacity of the Channel

Let $\mathbf{W} = [\mathbf{w}_1, ..., \mathbf{w}_K] \in \mathbb{R}^{F \times K}$ denotes the spreading matrix. We define the set of power constrained sparse spreading matrices as $\mathcal{C}_1 = \{\mathbf{W} | w_{f,k} \in \mathbb{R}, \frac{1}{d_k} \sum_f w_{f,k}^2 \le 1, \|\mathbf{w}_k\|_0 = d_k, \forall k \in \mathcal{K}\}$ where $\|\mathbf{w}_k\|_0 = d_k$ restricts the number of non-zero elements in $\mathbf{w}_k$ to be equal to d_k. The interest is in the scenario where the BS assigns the spreading codes based on pathlosses. Implicit in this model is the assumption that the channel statistics vary much more slowly than the small-scale fading coefficients [16]. Given perfect knowledge of fading coefficients at the BS side, ergodic mutual information (EMI) between the transmitters and the receiver, for a spreading matrix $\mathbf{W} \in \mathcal{C}_1$ known to the BS and the UEs, is

$$J(\mathbf{W}, \rho) = \frac{1}{F} \mathbb{E} \log \left| \mathbf{I}_F + \sum_{k \in \mathcal{K}} \frac{\rho a_k^2}{d_k} \text{diag}\{\mathbf{w}_k\} \mathbf{g}_k \mathbf{g}_k^{\text{H}} \text{diag}\{\mathbf{w}_k\} \right|$$

where the expectation is with respect to the ergodic random realizations of the small-scale fading. Then, the ergodic capacity of the channel is equal to the maximum of $J(\mathbf{W}, \rho)$ over the set of all sparse spreading matrices in $\mathcal{C}_1$, i.e.,

$$C_{\text{E}} = \max_{\mathbf{W} \in \mathcal{C}_1} J(\mathbf{W}, \rho). \tag{2}$$

The corresponding conventional problem without sparsity constraints is considered in [17], [18] for additive white Gaussian

noise (AWGN) channel. While AWGN channel capacity depends on the spreading codes through their cross-correlations, the transmission of the signal over a Rayleigh fading channel destroys the orthogonality of the spreading codes [19]. This together with the assumption of obtaining the spreading codes based on pathlosses limit the search space to the set of vectors $\mathbf{w}_k \in \mathbb{R}^+$ with positive values, and independent of the small-scale fading gains.[2] Even though the cross-correlation properties of the spreading codes are not determining parameters in (2), the optimal pairing of UEs and the power loading on each sub-channel need to be studied. The binary sparsity constraints in (2) makes the problem non-convex. Also, the expectation in (2) needs to be evaluated in a concise form.

III. AN OPTIMIZATION APPROACH BASED ON ASYMPTOTIC ANALYSIS

In the following, we use theory of large random matrices [9] to characterize the EMI in terms of spreading codes and pathloss values. The large system analysis of the problem is carried out in the asymptotic regime where $F \to \infty$ with $KF^{-1} \in (0, \infty)$ and $d_k F^{-1} \in (0, 1], \forall k$. We use subscript F to denote the dependency of the entities on the system size.

Theorem 1. *Consider* $\mathbf{H}_F = [\mathbf{h}_1, ..., \mathbf{h}_K]$ *with* $\mathbf{h}_k = \frac{a_k}{\sqrt{d_k}} \text{diag}\{\mathbf{w}_k\}\mathbf{g}_k, \forall k \in \mathcal{K}$. *The entries of* $\mathbf{G}_F = [\mathbf{g}_1, ..., \mathbf{g}_K]$ *are i.i.d standard complex Gaussian random variables. The deterministic vectors* $\mathbf{w}_k, \forall k \in \mathcal{K}$ *are the columns of* $\mathbf{W}_F \in \mathcal{C}_1$, *each with* d_k *non-zero values. The scalars* $\{a_k\}_{1 \le k \le K}$ *are bounded real values. Then, as* $F \to \infty$ *with* $KF^{-1} \in (0, \infty)$ *and* $d_k F^{-1} \in (0, 1], \forall k$, *we have*

$$J_F(\mathbf{W}_F, \rho) = \bar{J}_F(\mathbf{W}_F, \rho) + \epsilon_F(\rho) \tag{3}$$

where $\epsilon_F(\rho) \le \frac{K}{F} \frac{C}{d^2}$ *with* C *being a generic constant independent of system size,* $d = \min\{d_k\}_{1 \le k \le K}$, *and*

$$\bar{J}_F(\mathbf{W}_F, \rho) = \frac{1}{F} \sum_{k \in \mathcal{K}} \log(1 + \frac{\rho a_k^2}{d_k} \sum_{f \in \mathcal{F}} w_{f,k}^2 r_f)$$

$$+ \frac{1}{F} \sum_{f \in \mathcal{F}} \log(1 + \rho \sum_{k \in \mathcal{K}} \frac{1}{d_k} w_{f,k}^2 a_k^2 \tilde{r}_k) - \frac{\rho}{F} \sum_{f \in \mathcal{F}} \sum_{k \in \mathcal{K}} \frac{w_{f,k}^2}{d_k} a_k^2 r_f \tilde{r}_k \tag{4}$$

where $r_f(\mathbf{W}_F, \rho)$ *and* $\tilde{r}_k(\mathbf{W}_F, \rho)$ *are the solutions of*

$$r_f = (1 + \rho \sum_{k \in \mathcal{K}} \frac{1}{d_k} w_{f,k}^2 a_k^2 \tilde{r}_k)^{-1}, \forall f \in \mathcal{F},$$

$$\tilde{r}_k = (1 + \frac{\rho a_k^2}{d_k} \sum_{f \in \mathcal{F}} w_{f,k}^2 r_f)^{-1}, \forall k \in \mathcal{K}. \tag{5}$$

Proof. The convergence $J_F(\mathbf{W}_F, \rho) - \bar{J}_F(\mathbf{W}_F, \rho) \to 0$ can be claimed relying on Girko's law [21, Section 3.2.3] [9, Theorem 6.10]. An alternative proof based on Replica method is also given in [22]. The convergence rate can be obtained using Nash-Poincaré inequality [23] as in the extended version of this work in [24], or alternatively, using the results in [25] by properly scaling the channel entries' variances while ensuring the assumptions therein remain valid. □

---

[2]This directly follows from the invariance of standard complex Gaussian random vector in distribution under unitary transformation [20].

According to Theorem 1, the EMI $\mathrm{J}_F(\mathbf{W}_F, \rho)$ converges asymptotically to the deterministic equivalent $\bar{\mathrm{J}}_F(\mathbf{W}_F, \rho)$ with a convergence rate of $\mathcal{O}(\frac{1}{d^2})$. In the finite scenarios of interest with a moderate number of sub-channels, d_k values can be small relative to F.[3] In such finite cases, the analysis in Sec. IV-B shows that the residual term $\epsilon_F(\rho)$ appears as a small incremental gain for sparse spreading, which is dictated mainly by the number of non-zero elements in the codes. Keeping this in mind, in Sec. IV-A, we first formulate an optimization problem to get the set $\bar{\mathcal{C}}^*$ of all spreading codes, irrespective of sparsity constraints, which maximize $\bar{\mathrm{J}}_F(\mathbf{W}_F, \rho)$. Then, in Sec. IV-B, we show that the solutions in $\bar{\mathcal{C}}^*$ with desired sparsity harness the incremental gain in $\epsilon_F(\rho)$ and attain close to optimal values of EMI $\mathrm{J}_F(\mathbf{W}_F, \rho)$ in the finite regime. Finally, We identify the desired subset of spreading matrices in $\bar{\mathcal{C}}^*$ that satisfy the sparsity constraints via a simple partitioning algorithm.

IV. MAXIMIZING THE ERGODIC MUTUAL INFORMATION

In the sequel, we omit the subscript F denoting the dependency on system size. Also, observe that $\bar{\mathrm{J}}(\mathbf{W}, \rho)$ in (4) depends only on squares of $w_{f,k}$ values. Therefore, with a change of variable as $v_{f,k} = \frac{1}{d_k}w_{f,k}^2$, hereafter, we express the EMI and the related entities as a function of matrix $\mathbf{V} = [v_{f,k}]_{f \in \mathcal{F}, k \in \mathcal{K}}$. Given a matrix $\mathbf{V}$, the corresponding set of spreading vectors $\{\mathbf{w}_k\}$ can be obtained up to an uncertainty in the sign of the entries of spreading vectors. As mentioned in Sec. II-A, the objective function under the considered i.i.d channel model is indifferent to the signs of spreading code entries. Thus, hereafter, we refer to $\mathbf{V}$ and $\mathbf{W}$ interchangeably as the spreading matrix.

A. The optimal spreading codes in asymptotic regime

Let us first neglect the sparsity constraints, and define $\mathcal{C}_2 \triangleq \{\mathbf{V} | v_{f,k} \in \mathbb{R}^+, \sum_{f \in \mathcal{F}} v_{f,k} \leq 1, \forall k \in \mathcal{K}\}$ to be the set of all spreading codes that satisfy the power constraints. Then, the problem of maximizing $\bar{\mathrm{J}}(\mathbf{V}, \rho)$ can be formulated as follows

$$C_E = \max_{\mathbf{V} \in \mathcal{C}_2} \bar{\mathrm{J}}(\mathbf{V}, \rho). \tag{6}$$

The Karush-Kuhn-Tucker (KKT) conditions [26] are necessary conditions for a matrix $\mathbf{V}$ to be a local optimal solution of the problem in (6). The sufficiency and the global-optimality are discussed later. The Lagrangian associated with (6) is

$$L(v_{f,k}, \lambda_{f,k}, \delta_k) = -\bar{\mathrm{J}}(\mathbf{V}, \rho) - \sum_{i \in \mathcal{F}, j \in \mathcal{K}} \lambda_{i,j} v_{i,j} + \sum_{j \in \mathcal{K}} \delta_j \Big(\sum_{i \in \mathcal{F}} v_{i,j} - 1\Big)$$

where δ_k and $\lambda_{f,k}$ are associated Lagrangian variables. The gradient of the Lagrangian is given as

$$\nabla_{f,k} L(v_{f,k}, \lambda_{f,k}, \delta_k) = -\frac{\partial \bar{\mathrm{J}}(\mathbf{V}, \rho)}{\partial v_{f,k}} - \lambda_{f,k} + \delta_k, \forall f \in \mathcal{F}, k \in \mathcal{K}.$$

Note that $\bar{\mathrm{J}}(\mathbf{V}, \rho)$ depends on the entries of $\mathbf{V}$ via $r_f(\mathbf{V}, \rho)$ and $\tilde{r}_k(\mathbf{V}, \rho)$ as in (4) and (5). Since they are the solutions to the saddle point equations, the partial derivatives $\frac{\partial \bar{\mathrm{J}}(\mathbf{V}, \rho)}{\partial r_f}$ and $\frac{\partial \bar{\mathrm{J}}(\mathbf{V}, \rho)}{\partial \tilde{r}_k}$ are zero at any point given by

$(\mathbf{V}, r_f(\mathbf{V}, \rho), \tilde{r}_k(\mathbf{V}, \rho))$. Thus, the chain rules of derivatives [27] yield $\frac{\partial \bar{\mathrm{J}}(\mathbf{V}, \rho)}{v_{f,k}} = \frac{\rho}{F} a_k^2 \tilde{r}_k r_f$. As a result, the KKT conditions can be evaluated as

$$\lambda_{f,k}^* \geq 0, \quad \lambda_{f,k}^* v_{f,k} = 0, \quad \delta_k^* \geq 0, \quad \delta_k^*\Big(\sum_{i \in \mathcal{F}} v_{i,k} - 1\Big) = 0,$$

$$-\frac{\rho}{F} a_k^2 \tilde{r}_k r_f - \lambda_{f,k}^* + \delta_k^* = 0, \quad \forall f \in \mathcal{F}, k \in \mathcal{K}$$

where $\lambda_{f,k}^*$ and δ_k^* denote the optimal values of the Lagrangian variables. Since $\lambda_{f,k}$ can be solved from the last equation, the KKT conditions can be simplified as

$$\delta_k^*\Big(\sum_{i \in \mathcal{F}} v_{i,k} - 1\Big) = 0, \quad \forall k \in \mathcal{K}, \tag{7a}$$

$$\Big(\delta_k^* - \frac{\rho}{F} a_k^2 \tilde{r}_k r_f\Big) v_{f,k} = 0, \quad \forall f \in \mathcal{F}, k \in \mathcal{K}, \tag{7b}$$

$$\delta_k^* \geq \frac{\rho}{F} a_k^2 \tilde{r}_k r_f, \quad \forall f \in \mathcal{F}, k \in \mathcal{K}. \tag{7c}$$

The KKT condition (7c) implies that $\delta_k \geq \frac{\rho}{F} a_k^2 \tilde{r}_k r_f > 0$, and thus, (7a) gives $\sum_{i \in \mathcal{F}} v_{i,k} = 1, \forall k \in \mathcal{K}$, i.e., all UEs are active, and transmit with full power. Also, at the optimum, we have $r_f = r^*, \forall f \in \mathcal{F}$ where

$$r^* = \Big(1 + \frac{\rho}{F} \sum_{k \in \mathcal{K}} \frac{a_k^2}{1 + \rho a_k^2 r^*}\Big)^{-1}. \tag{8}$$

This is justified because (7b) and (7c) imply that a given UE k transmits only on the sub-channels with the largest r_f value. Since, all other UEs also have the same preference, the condition in (7b) and (7c) are satisfied only if UEs assign their powers such that $r_f = r, \forall f$. It is shown in [24] that r is unique, i.e., $r = r^*$, for any solution satisfying (7). Plunging $r_f = r^*, \forall f \in \mathcal{F}$ into (5) equivalently yields $\tilde{r}_k^* = \frac{1}{1 + \rho a_k^2 r^*}, \forall k \in \mathcal{K}$. Observe that r^* and $\tilde{r}_k^*$ are the same for all the solutions satisfying KKT conditions. Plugging the r^* and $\tilde{r}_k^*$ values into (4) yields

$$\bar{\mathrm{J}}(\mathbf{V}, \tilde{r}_k^*, r^*, \rho) = -\frac{1}{F}\sum_{k \in \mathcal{K}} \log(\tilde{r}_k^*) - \log(r^*) - \frac{\rho}{F} r^* \sum_{k \in \mathcal{K}} a_k^2 \tilde{r}_k^*.$$

Note that the resultant objective function is given in terms of $\tilde{r}_k^*$ and r^* values, and irrespective of $v_{f,k}$ values. Thus, one can conclude that any solution satisfying KKT condition attains the maximum of $\bar{\mathrm{J}}(\mathbf{V}, \rho)$. This set of solutions, denoted hereafter by $\bar{\mathcal{C}}^*$, can be evaluated equivalently as the positive solutions of the following indeterminate system of equations

$$\sum_{f \in \mathcal{F}} v_{f,k} = 1, \forall k \in \mathcal{K}, \tag{9a}$$

$$\sum_{k \in \mathcal{K}} \beta_k v_{f,k} = \frac{1}{r^*} - 1, \forall f \in \mathcal{F}. \tag{9b}$$

where $\beta_k = \frac{\rho a_k^2}{1 + \rho a_k^2 r^*}$. In these equations, r^* is a fixed scalar, which is evaluated from (8). The equalities in (9b) are obtained by setting $r_f = r^*, \forall f$ in (5). These equalities follow since any power constrained spreading matrix $\mathbf{V}$ that gives $r_f = r^*, \forall f$, equivalently, satisfies the KKT conditions in (7) as well. In the special setting with the same pathlosses for all UEs, it can be verified that the regular LDS codes are among the solutions of (9). However, in general, this condition does not hold due to distinct pathlosses. Also, observe that dense spreading with $v_{f,k} = 1/F, \forall k, f$ is always a solution to (9).

---

[3]Note that while the limiting results are obtained in the asymptotic regime, those can be applied as approximations for the finite scenarios with dimensions as small as 8 and even 4 or 2 [9, Section 2.2.1].

B. On the optimality of the asymptotic sparse spreading codes

Based on the analysis in Sec. IV-A, we know that the spreading codes in $\mathcal{C}_2$ that maximize the deterministic equivalent $\bar{J}(V,\rho)$ in (6) are given as the solutions of the system of equations in (9). This set of solutions is denoted by $\bar{\mathcal{C}}^*$. Hereafter, we use $\bar{V}^*$ to refer to a member of the set $\bar{\mathcal{C}}^*$. The analysis in Theorem 1 shows that the residual term is bounded as $\epsilon(\rho) \leq \frac{K}{F}\frac{C}{d^2}$ with C being a generic constant, independent of system size. Therefore, $\epsilon(\rho)$ term vanishes asymptotically, and thus, the solutions in $\bar{\mathcal{C}}^*$ attain the maximum of EMI $J(V,\rho)$ in the asymptotic regime. However, in the finite regime, $\epsilon(\rho)$ term appears as a small incremental gain in the EMI formulation, which needs to be considered. Let V_d^* to be the unknown optimal spreading matrix that maximizes $J(V,\rho)$ subject to sparsity constraints. Also, let $J_d^* \triangleq J(V_d^*,\rho)$ to be the maximum of $J(V,\rho)$ attained by V_d^*. Now, the penalty when using any spreading matrix $\bar{V}^* \in \bar{\mathcal{C}}^*$, given as a solution of the system of equations in (9), instead of the optimal one V_d^* can be written as

$$\Delta_d = \underbrace{J(V_d^*,\rho)}_{J_d^*} - J(\bar{V}^*,\rho) \tag{10}$$

Writing $J(V_d^*,\rho)$ and $J(\bar{V}^*,\rho)$ in terms of the deterministic equivalents and the residual terms, we get

$$J(V_d^*,\rho) - \bar{J}(V_d^*,\rho) = \epsilon_1(\rho) \tag{11a}$$
$$J(\bar{V}^*,\rho) - \bar{J}(\bar{V}^*,\rho) = \epsilon_2(\rho) \tag{11b}$$

where the additional index i in $\epsilon_i(\rho)$ is added to distinguish the above differences. Subtracting the sides of above equalities, and rearranging the terms, we get

$$\underbrace{(J(V_d^*,\rho) - J(\bar{V}^*,\rho))}_{\Delta_d} + (\bar{J}(\bar{V}^*,\rho) - \bar{J}(V_d^*,\rho)) = \Delta_\epsilon$$

where $\Delta_\epsilon = \epsilon_1(\rho) - \epsilon_2(\rho)$. Since the subtraction $(\bar{J}(\bar{V}^*,\rho) - \bar{J}(V_d^*,\rho))$ in the left-hand is positive[4], it can be claimed that the gap to the optimum Δ_d is bounded as

$$0 \leq \Delta_d \leq \Delta_\epsilon^+ \tag{12}$$

where Δ_ϵ^+ denotes the positive values of Δ_ϵ. The lower bound by zero is imposed since Δ_d attains negative values when $J_d^* < J(\bar{V}^*,\rho)$. However, this can happen only when $\bar{V}^*$ violates the sparsity constraints, which is not the case of interest.

Fig.1 illustrates the variations in the residual term $\epsilon(\rho)$ in a numerical study based on the simulation assumptions in Sec. V with $F = 50$, $K = 100$, and $d_K = d, \forall k$. In deriving the results, a randomly selected drop of UEs is taken, and the mean and variance of the residual term $\epsilon(\rho)$ are evaluated over 1000 randomly selected spreading matrices in $\mathcal{C}_1$. Fig. 1 shows the mean and 20 times magnified variance of $\epsilon(\rho)$ values along with $\bar{J}(V,\rho)$ versus the number of non-zero elements in the codes. The first observation is that the residual term is small as compared to $\bar{J}(V,\rho)$. Moreover, the small variance of the residual term indicates that the values of $\epsilon(\rho)$ do not vary abruptly among the spreading matrices in $\mathcal{C}_2$ with the same number of non-zero elements. Thus, the values of the residual term are dictated mainly by the number of non-zero
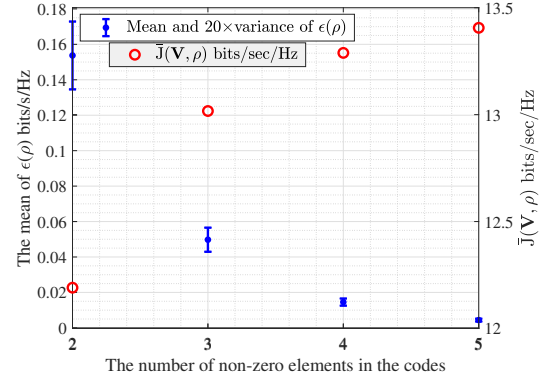
[4]Note that $\bar{J}(\bar{V}^*,\rho)$ is the maximum of the objective function in (6).

Fig. 1: Mean and 20 times magnified variance of $\epsilon(\rho)$ (left axis), and $\bar{J}(V,\rho)$ (right axis) vs. d, random spreading.

elements in spreading codes, which implies a decline in $\epsilon(\rho)$ as d increases. Generally, the residual term is larger for the sparse spreading matrices, while it almost disappears in the dense spreading with a moderate number of sub-channels F. This motivates the small incremental gain in $\epsilon(\rho)$ for sparse spreading to be interpreted as sparsity gain, in contrast to the loss associated with the spreading in the considered system model [14] Now, let us recall the Δ_d formulation in (12), which gives the gap to the optimum J_d^* when using a solution $\bar{V}^* \in \bar{\mathcal{C}}^*$ in the finite regime. Relying on the above analysis, we expect a solution $\bar{V}^*$ that satisfies the sparsity constraints to harness an incremental $\epsilon_2(\rho)$ value close to the $\epsilon_1(\rho)$ one, and thus, yields a close to zero performance gap of Δ_d. Next, we find the sparse solutions in $\bar{\mathcal{C}}^*$ via an algorithmic solution.

C. Algorithms for allocation of low density spreading codes

The system of equations in (9) unveils a simple rule for the allocation of spreading codes. This allows the desired sparse codes to be obtained using efficient algorithmic solutions. Let the elements $v_{f,k}$ of spreading codes to be taken from $\{0, \frac{1}{d_k}\}$. Based on (9b), the problem is to allocate $\frac{1}{d_k}\beta_k$ values to sub-channels such that the sums of $\frac{1}{d_k}\beta_k$ values on each sub-channel become the same, i.e., equal to $(\frac{1}{r^*} - 1)$, across all sub-channels. This problem falls within a class of partitioning problems that arise in number theory and computer science [28]. Although the partitioning problem is NP-complete, there are heuristics that solve the problem in many instances, either optimally or approximately [29]. One

---

Algorithm 1 Partitioning solution

1: Divide the total power of each UE k into d_k power fragments.
2: Set $v_{f,k} = 0$, $\forall f,k$, and $\mathcal{J} = \{1, ..., K\}$.
3: **while** $\mathcal{J}$ is non-empty **do**
4:     Set $k = \arg\max_{j \in \mathcal{J}} \frac{1}{d_j}\beta_j$.
5:     Set $f = \arg\min_{i \in \mathcal{F}} \eta_i$ with $\eta_i = \sum_{j \in \mathcal{K}} \beta_j v_{i,j}$.
6:     Set $v_{f,k} = 1/d_k$
7:     **if** $\sum_{i \in \mathcal{F}} v_{i,k} = 1$, i.e., UE k satisfies (9a). **then**
8:         Remove index k from $\mathcal{J}$.
9:     **end if**
10: **end while**

---

such approach is the greedy algorithm, which iterates through the $\frac{1}{d_k}\beta_k$ values in descending order, assigning each of them to whichever sub-channel has the smallest sum [30]. These steps are summarized in Alg. 1. Here, we try to make the sum terms η_f across the sub-channels as equal as possible. Let $\eta_{\max}^*$ denotes the maximum of $\eta_f, \forall f \in \mathcal{F}$ in an optimal partitioning solution. Alg. 1 yields η_f values such that $\frac{\max(\eta_f)}{\eta_{\max}^*} \leq \frac{4}{3} - \frac{1}{3F}$ [30]. Alg. 1 has a running time of $\mathcal{O}(2^F(d_{\max}K)^2)$ [30].

V. NUMERICAL RESULTS

The simulation results are generated in a scenario where a single-antenna BS serves the UEs in the uplink. Transmit power of each UE is 1 Watt, and the noise power is set to -120dB. The pathlosses are taken randomly and uniformly from the range of -150dB to -60dB to account for diverse received SNRs. The final channel gains are given by the product of pathlosses and fast fading effects as in (1).

Fig. 2 shows the SE enhancement for the coordinated allocation of sparse codes based on pathloss values in Alg.1, as compared to the uncoordinated scheme that allocates randomly generated sparse codes to UEs. The rates are averaged over 1000 random UE drops and in each drop the expectation in $J(\mathbf{V}, \rho)$ is evaluated over 1000 random realizations of small-scale fading. In the analysis, we use Alg.1 in the non-spreading case with $d = 1$ as well. Note that Alg.1 allocates sparse codes to UEs such that the deterministic equivalent of EMI is maximized. The motivation therein is that the residual term is small relative to the deterministic equivalent part, and the small gain in $\epsilon(\rho)$ is harnessed inherently due to the sparsity of the allocated codes. While we expect $\epsilon(\rho)$ to be relatively small for the cases with $d > 1$, due to the fast convergence rate of $\mathcal{O}(\frac{1}{d^2})$, the analysis in the non-spreading case with $d = 1$ might be considered as a heuristic attempt. Interestingly, the difference between $J(\mathbf{V}, \rho)$ and $\bar{J}(\mathbf{V}, \rho)$ is relatively small even in such a case, and the coordinated allocation of resources gives about 20% to 35% enhancement in spectral efficiency at 100% to 300% system load. The system load is defined as the ratio of $\frac{K}{F}$ in percentage. In the case with $d = 2$, the gain in the coordinated assignment of spreading codes is about 6.5% and 11% at 100% to 300% load, respectively, which is less than in the non-spreading case ($d = 1$).
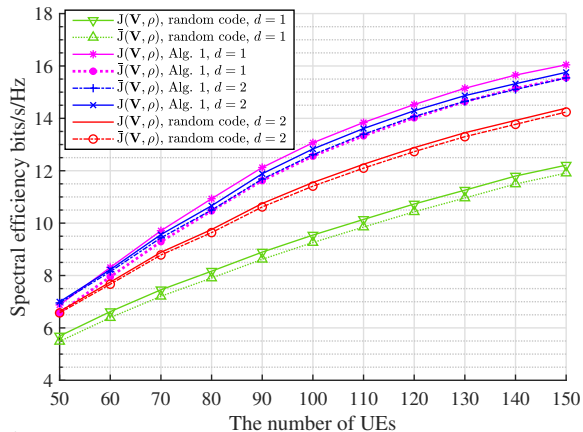
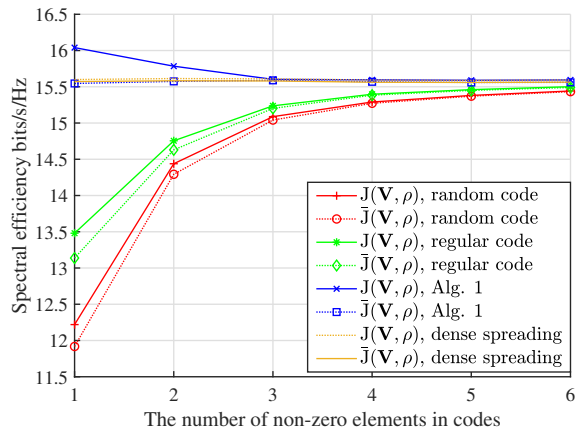Fig. 2: Spectral efficiency $J(\mathbf{V}, \rho)$ vs. K, $d = 1, 2$, $F = 50$.

Fig. 3: efficiency $J(\mathbf{V}, \rho)$ vs. d, $F = 50$, $\frac{K}{F} = 3$.

In Fig. 3, the attainable SEs are presented versus the number of non-zero elements in the codes d, for a system load of 300%. It can be seen that the performance of random spreading method improves as d grow larger. Spreading on more sub-channels allows UEs to attain interference diversity. This, in general, reduces the loss imposed by uncoordinated resource allocation. Note that, even though one can enhance the spectral efficiency of the uncoordinated method by spreading on further sub-channels, the number of UEs overlapping on the same sub-channel increases correspondingly. In a system with 300% load, the average number of overlapping UEs in the case with $d = 2$ and $d = 6$ is equal to 6 and 18, respectively. Thus, the detection complexity is greatly increased with larger d. In Fig. 3, the spectral efficiency of dense spreading is also depicted. As mentioned in Sec. IV-A, dense spreading is a solution of the optimization problem in (6), and thus, the values of deterministic $\bar{J}(\mathbf{V}, \rho)$ for both of dense spreading and the sparse spreading in Alg. 1 are the same. However, the values of $J(\mathbf{V}, \rho)$ for sparse spreading in Alg. 1 are better of dense spreading by an amount of $\epsilon(\rho)$. This can be seen from Fig. 3 where the curves of $\bar{J}(\mathbf{V}, \rho)$ and $J(\mathbf{V}, \rho)$ are almost overlapping for dense spreading while the curve of $J(\mathbf{V}, \rho)$ for Alg 1 at $d = 1$ is nearly 0.5 bits/s/Hz higher than the deterministic curve. This additional gain in $J(\mathbf{V}, \rho)$ for sparse spreading as compared to the dense spreading was referred to as sparsity gain in Sec. IV-B. Observe that the sparsity gain decreases as the number of non-zero elements in the codes increases. Finally, we observe that, in contrast to the symmetric model, the regular spreading is inferior to Alg.1 and dense spreading in the considered asymmetric scenario.

VI. CONCLUSIONS

A simple and efficient rule for close to the optimal allocation of sparse spreading codes was derived based on rigorous analysis. The analysis reduced EMI maximization dilemma to a partitioning problem, which was solved via an efficient algorithmic solution. The algorithm allocates the spreading codes based on the system load, sparsity constraints, and pathloss values. The simulation results showed that the performance of the proposed method with minimal coordination is superior to random, regular, and dense spreading schemes.

REFERENCES

[1] O. Shental, B. M. Zaidel, and S. S. Shitz, "Low-density code-domain NOMA: Better be regular," in *IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 2628–2632.

[2] B. M. Zaidel, O. Shental, and S. S. Shitz, "Sparse NOMA: A closed-form characterization," in *IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 1106–1110.

[3] G. C. Ferrante and M. D. Benedetto, "Spectral efficiency of random time-hopping CDMA," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6643–6662, Dec 2015.

[4] M. Yoshida and T. Tanaka, "Analysis of sparsely-spread CDMA via statistical mechanics," in *IEEE International Symposium on Information Theory*, July 2006, pp. 2378–2382.

[5] A. Montanari and D. Tse, "Analysis of belief propagation for non-linear problems: The example of CDMA (or: How to prove tanaka's formula)," in *IEEE Information Theory Workshop - ITW '06 Punta del Este*, March 2006, pp. 160–164.

[6] M. T. P. Le, G. C. Ferrante, G. Caso, L. De Nardis, and M. Di Benedetto, "On information-theoretic limits of code-domain NOMA for 5G," *IET Communications*, vol. 12, no. 15, pp. 1864–1871, 2018.

[7] R. R. Müller, "Random matrices, free probability and the replica method," in *12th European Signal Processing Conference*, Sep. 2004, pp. 189–196.

[8] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information theory*, vol. 45, no. 2, pp. 622–640, 1999.

[9] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge University Press, 2011.

[10] P. M. Wood, "Universality and the circular law for sparse random matrices," *Ann. Appl. Probab.*, vol. 22, no. 3, pp. 1266–1300, 06 2012.

[11] M. T. P. Le, G. C. Ferrante, T. Q. S. Quek, and M. Di Benedetto, "Fundamental limits of low-density spreading NOMA with fading," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4648–4659, July 2018.

[12] R. Hoshyar, R. Razavi, and M. Al-Imari, "LDS-OFDM an efficient multiple access technique," in *Vehicular Technology Conference*, May 2010, pp. 1–5.

[13] M. Al-Imari, M. Imran, and R. Tafazolli, "Low density spreading multiple access," *Information Technology Software Engineering*, vol. 2, 09 2012.

[14] V. V. Veeravalli and A. Mantravadi, "The coding-spreading tradeoff in CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 2, pp. 396–408, Feb 2002, doi=10.1109/49.983362.

[15] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX, Second Edition*. John Wiley & Sons, 2008, vol. 2.

[16] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," in *Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 974–978.

[17] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 40, no. 4, pp. 1261–1266, July 1994.

[18] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1984–1991, Sept 1999.

[19] M. Reinhardt and J. Lindner, "Transformation of a Rayleigh fading channel into a set of parallel AWGN channels and its advantage for coded transmission," *Electronics Letters*, vol. 31, no. 25, pp. 2154–2155, Dec 1995.

[20] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.4460100604

[21] R. R. Müller, G. Alfano, B. M. Zaidel, and R. de Miguel, "Applications of large random matrices in communications engineering," *ArXiv*, vol. abs/1310.5479, 2013. [Online]. Available: https://arxiv.org/abs/1310.5479

[22] R. R. Müller, *THE REPLICA METHOD IN MULTIUSER COMMUNICATIONS*, pp. 139–165.

[23] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans Inf Theory*, vol. 54, 01 2007.

[24] H. Asgharimoghaddam, A. Tölli, and J. Kaleva, "Capacity approaching low density spreading in uplink NOMA via asymptotic analysis," 2019. [Online]. Available: https://arxiv.org/abs/1911.11540

[25] W. Hachem, P. Loubaton, and J. Najim, "A CLT for information-theoretic statistics of Gram random matrices with a given variance profile," *Ann. Appl. Probab.*, vol. 18, no. 6, pp. 2071–2130, 12 2008, doi:10.1214/08-AAP515.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[27] H. Anton, *Calculus with Analytic Geometry*. John Wiley Sons Inc., 1995, vol. 10.

[28] R. E. Korf, "A complete anytime algorithm for number partitioning," *Artificial Intelligence*, vol. 106, no. 2, pp. 181 – 203, 1998.

[29] B. Hayes, "Computing science: The easiest hard problem," *American Scientist*, vol. 90, no. 2, pp. 113–117, 2002.

[30] R. L. Graham, "Bounds on multiprocessing timing anomalies," *SIAM Journal of Applied Mathematics*, vol. 17, pp. 416–429, 1969.