Compressive Sensing Based User Clustering for Downlink NOMA Systems with Decoding Power

Zhaohui Yang, Cunhua Pan, Wei Xu, *Senior Member, IEEE*, and Ming Chen

Abstract—This letter investigates joint power control and user clustering for downlink non-orthogonal multiple access systems. Our aim is to minimize the total power consumption by taking into account not only the conventional transmission power but also the decoding power of the users. To solve this optimization problem, it is firstly transformed into an equivalent problem with tractable constraints. Then, an efficient algorithm is proposed to tackle the equivalent problem by using the techniques of reweighted ℓ_1-norm minimization and majorization-minimization. Numerical results validate the superiority of the proposed algorithm over the conventional algorithms including the popular matching-based algorithm.

Index Terms—Non-orthogonal multiple access, power control, user clustering, majorization-minimization.

I. INTRODUCTION

Non-orthogonal multiple access (NOMA) has been deemed as a promising technology for future fifth generation systems [1]–[3]. By applying superposition coding at the transmitter and employing successive interference cancellation (SIC) at the users, NOMA serves multiple users with the same time-frequency resource, which makes NOMA more spectral efficient than orthogonal multiple access (OMA). In [4], it was shown that NOMA can achieve superior performance in terms of ergodic sum rate compared with OMA. The power control problem of maximizing sum rate was investigated in [5] for downlink NOMA systems. Besides, the impact of user pairing in NOMA with fixed power allocation was investigated in [6], which showed that the paired users should have distinctive channel gains to obtain large sum rate. Moreover, the authors in [7] studied joint power control and user pairing, where the user pairing subproblem was solved by using matching theory.

In addition to spectral efficiency, power minimization for NOMA has been attracting research attention lately. To minimize total transmission power, a distributed power control algorithm was proposed by using the game theoretic approach [8] for uplink NOMA. In [9], the standard interference function was applied to solve the downlink sum power minimization problem for a two-cell NOMA system. It is proven in [10] that the downlink power minimization problem for NOMA with multiple subcarriers is NP-hard, and a relax-then-adjust algorithm was accordingly proposed. However, the

above works [8]–[10] all ignored the decoding power in the power consumption model, even though the decoding power is comparable to the transmission power [11].

In this letter, we investigate the downlink power minimization problem for NOMA. There are two main contributions in this letter. One contribution is that we consider the decoding power consumption, which is important but ignored in many existing works. The other contribution is to tackle this nonconvex power minimization problem. In particular, we first successfully transform the original problem with nonlinear rate constraints into an equivalent problem with linear rate constraints, and then adopt the penalty method and compressive sensing method to solve a sequence of tractable convex problems.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

Consider a downlink NOMA system with one single-antenna base station (BS) and M single-antenna users. In this system, there are N subcarriers. The channel gain between the BS and the m-th user on the n-th subcarrier is denoted by h_{mn}. Without loss of generality, the channels are sorted as $|h_{1n}|^2 \leq |h_{2n}|^2 \leq \cdots \leq |h_{Mn}|^2$, for all $n = 1, \cdots, N$. To ensure the sort of channel gains, user m is denoted as the m_n-th user on the n-th subcarrier.

According to the NOMA principle, the BS simultaneously transmits signal to all the users. The transmitted signal x_n on the n-th subcarrier can be expressed as

$$x_n = \sum_{m=1}^{M} \sqrt{p_{mn}} x_{mn}, \quad (1)$$

where x_{mn} and p_{mn} are the message and allocated power for the m-th user on the n-th subcarrier, respectively. The observation at the m-th user on the n-th subcarrier is

$$y_{mn} = h_{mn} \sum_{l=1}^{M} \sqrt{p_{ln}} x_{ln} + z_{mn}, \quad (2)$$

where z_{mn} represents the additive zero-mean Gaussian noise with variance σ^2.

For downlink NOMA, SIC is carried out at the users. Assume that the bandwidth for each subcarrier is B. According to [4], the achievable rate of the m-th user on the n-th subcarrier is

$$r_{mn} = B \log_2 \left(1 + \frac{H_{mn} p_{mn}}{H_{mn} \sum_{l=m+1}^{M} p_{ln} + \sigma^2} \right) \quad (3)$$

for $1 \leq m \leq M - 1$, and

$$r_{Mn} = B \log_2 \left(1 + \frac{H_{Mn} p_{Mn}}{\sigma^2} \right), \quad (4)$$

where $H_{mn} = |h_{mn}|^2$, $\forall m = 1, \cdots, M$.

Since one user can be allocated with multiple subcarriers, the achievable rate of user m is given by

$$r_m = \sum_{n=1}^{N} r_{m_n n}. \tag{5}$$

B. Power Consumption Model

The total power consumption of the system consists of two parts: the transmission power of the BS and the decoding power of the users. By summing the transmission power for all users on all subcarriers, the total transmission power of the BS can be calculated as

$$\sum_{m=1}^{M} \sum_{n=1}^{N} p_{mn}. \tag{6}$$

According to [12], the decoding complexity of each user grows linearly with the decoded rate of each user. A linear function relating to the rate and the power consumed by the decoder at user m can be given by [12]–[14]

$$P_m^{\text{dec}}(\bar{r}_m) = \lambda_m \bar{r}_m, \tag{7}$$

where λ_m is the decoder efficiency of user m, and $\bar{r}_m$ is the decoded rate of user m. Since user m should decode the signals of weak users before decoding its own message [4], the decoded rate of user m on the n-th subcarrier is $\sum_{s=1}^{m_n} r_{sn}$ if user m occupies the n-th subcarrier, i.e., $p_{m_n n} > 0$. According to (7), the decoding power of user m is

$$P_m^{\text{dec}}(\bar{r}_m) = \sum_{n=1}^{N} \sum_{s=1}^{m_n} \lambda_m r_{sn} \|p_{m_n n}\|_0, \tag{8}$$

where $\|\cdot\|_0$ is the ℓ_0-norm.

The total power consumption, denoted as P_{total}, is given by

$$P_{\text{total}} = \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn} + \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{s=1}^{m_n} \lambda_m r_{sn} \|p_{m_n n}\|_0. \tag{9}$$

From (9), the total power consumption contains two parts that conflict each other. On one hand, the transmission power part decreases with the number of users occupying one subcarrier (according to [10]). This is due to the fact that users can occupy more subcarriers and SIC is helpful in mitigating inter-user interference for larger number of users occupying one subcarrier. On the other hand, the decoding power part increases with the number of users occupying one subcarrier for higher decoded rate of the users.

C. Problem Formulation

According to (3), (4), (5) and (9), the total power optimization problem can be formulated as:

$$\min_{\boldsymbol{p} \geq \boldsymbol{0}} \quad \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn} + \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{s=1}^{m_n} \lambda_m r_{sn} \|p_{m_n n}\|_0 \tag{10a}$$

$$\text{s.t.} \quad \sum_{n=1}^{N} r_{m_n n} = R_m, \quad \forall m \tag{10b}$$

$$\|\boldsymbol{p}_n\|_0 \leq L, \quad \forall n, \tag{10c}$$

where $\boldsymbol{p} = [p_{11}, \cdots, p_{M1}, \cdots, p_{MN}]^T$, r_{sn} is the achievable rate of the s-th user on the n-th subcarrier defined in (3), R_m is the rate demand of user m, L is the maximum number of multiplexed users on each subcarrier, and $\boldsymbol{p}_n = [p_{1n}, \cdots, p_{Mn}]^T$. Due to the practical limitations of the receiver's design complexity and the signal processing time for SIC, L is a parameter with $L \leq M$.

III. Joint Power Allocation and User Clustering

A. Equivalent Transformation

Theorem 1: The original total power minimization Problem (10) can be equivalently transformed to the following problem as:

$$\min_{\boldsymbol{r} \geq \boldsymbol{0}} \quad \sum_{n=1}^{N} \sum_{m=1}^{M} \left(\frac{\sigma^2}{H_{mn}} - \frac{\sigma^2}{H_{(m+1)n}} \right) 2^{\sum_{s=1}^{m} \frac{r_{sn}}{B}}$$

$$+ \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{s=1}^{m_n} \lambda_m r_{sn} \|r_{m_n n}\|_0 \tag{11a}$$

$$\text{s.t.} \quad \sum_{n=1}^{N} r_{m_n n} = R_m, \quad \forall m \tag{11b}$$

$$\|\boldsymbol{r}_n\|_0 \leq L, \quad \forall n, \tag{11c}$$

where $\boldsymbol{r} = [r_{11}, \cdots, r_{M1}, \cdots, r_{MN}]^T$, $\boldsymbol{r}_n = [r_{1n}, \cdots, r_{Mn}]^T$, and we set $\frac{\sigma^2}{H_{(M+1)n}} = 0$ for all n.

Proof: Please refer to Appendix A. ∎

From Theorem 1, nonlinear constraints (10b) in Problem (10) are converted to linear constraints (11b) in Problem (11). To deal with constraints (11c) with non-smooth ℓ_0-norm, we adopt the penalty method [15]. Specially, Problem (11) can be reformulated as:

$$\min_{\boldsymbol{r} \geq \boldsymbol{0}} \quad \sum_{n=1}^{N} \sum_{m=1}^{M} \left(\frac{\sigma^2}{H_{mn}} - \frac{\sigma^2}{H_{(m+1)n}} \right) 2^{\sum_{s=1}^{m} \frac{r_{sn}}{B}}$$

$$+ \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{s=1}^{m_n} \lambda_m r_{sn} \|r_{m_n n}\|_0 + \sum_{n=1}^{N} \frac{\|\boldsymbol{r}_n\|_0^K}{(L+0.5)^K} \tag{12a}$$

$$\text{s.t.} \quad \sum_{n=1}^{N} r_{m_n n} = R_m, \quad \forall m, \tag{12b}$$

where K is a large positive constant. When K becomes infinity, Problem (11) is equivalent to Problem (12) with fewer constraints. The reason is that $\lim_{K \to +\infty} \left(\frac{\|\boldsymbol{r}_n\|_0}{L+0.5} \right)^K = 0$ for $\|\boldsymbol{r}_n\|_0 \leq L$, and $\lim_{K \to +\infty} \left(\frac{\|\boldsymbol{r}_n\|_0}{L+0.5} \right)^K = +\infty$ for $\|\boldsymbol{r}_n\|_0 > L$.

B. Compressive Sensing Based Algorithm

The difficulty to solve Problem (12) is the non-smooth ℓ_0-norm in the objective function (12a). Since $\|\boldsymbol{r}_n\| \leq L$ and L is usually small in practical situations, $\boldsymbol{r}_n$ can be viewed as a sparse vector in the compressive sensing method. To deal with (12a), non-smooth ℓ_0-norm minimization can be approximately solved via a sequence of weighted ℓ_1-norm minimizations in compressive sensing according to [16] and [17]. Specifically, non-smooth ℓ_0-norm is approximated by

$$\|x\|_0 = \lim_{\tau \to 0} \frac{\ln(1 + x\tau^{-1})}{\ln(1 + \tau^{-1})}, \quad \forall x \geq 0. \tag{13}$$

Replacing non-smooth ℓ_0-norm in Problem (12) with the logarithmic function according to (13), we have

$$\min_{\boldsymbol{r}\geq 0} \sum_{n=1}^{N}\sum_{m=1}^{M}\left(\frac{\sigma^2}{H_{mn}}-\frac{\sigma^2}{H_{(m+1)n}}\right)2^{\sum_{s=1}^{m}\frac{r_{sn}}{B}}+\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{s=1}^{m_n}$$

$$\lambda_m r_{sn}\frac{\ln(1+r_{m_n n}\tau^{-1})}{\ln(1+\tau^{-1})}+\sum_{n=1}^{N}\frac{\left(\sum_{m=1}^{M}\ln(1+r_{mn}\tau^{-1})\right)^K}{(L+0.5)^K\ln^K(1+\tau^{-1})}$$

$$(14a)$$

$$\text{s.t.}\ \sum_{n=1}^{N}r_{m_n n}=R_m,\quad\forall m. \tag{14b}$$

Using (13) and the first-order approximation, we approximate the ℓ_0-norm in the objective function (12a) as

$$\|\boldsymbol{r}_n\|_0 \approx \sum_{m=1}^{M}(w_{mn}^{(t)}r_{mn}+\alpha_{mn}^{(t)}), \tag{15}$$

with $w_{mn}^{(t)}$ and $\alpha_{mn}^{(t)}$ iteratively updated according to

$$w_{mn}^{(t)}=\frac{1}{(r_{mn}^{(t)}+\tau)\ln(1+\tau^{-1})}, \tag{16}$$

and

$$\alpha_{mn}^{(t)}=\frac{(r_{mn}^{(t)}+\tau)\ln(1+\tau^{-1}r_{mn}^{(t)})-r_{mn}^{(t)}}{(r_{mn}^{(t)}+\tau)\ln(1+\tau^{-1})}, \tag{17}$$

where $r_{mn}^{(t)}$ is value of r_{mn} in the t-th iteration, and τ is a constant regularization factor. According to (16) and (17), we can obtain

$$r_{sn}\|r_{m_n n}\|_0 \approx r_{sn}(w_{m_n n}^{(t)}r_{m_n n}+\alpha_{m_n n}^{(t)})$$
$$=0.25w_{m_n n}^{(t)}[(r_{sn}+r_{m_n n})^2-(r_{sn}-r_{m_n n})^2]+\alpha_{m_n n}^{(t)}r_{sn}$$
$$\leq 0.25w_{m_n n}^{(t)}(r_{sn}+r_{m_n n})^2-0.25w_{m_n n}^{(t)}[(r_{sn}^{(t)}-r_{m_n n}^{(t)})^2$$
$$+2(r_{sn}^{(t)}-r_{m_n n}^{(t)})(r_{sn}-r_{sn}^{(t)}-r_{m_n n}+r_{m_n n}^{(t)})]+\alpha_{m_n n}^{(t)}r_{sn}$$
$$\triangleq f_{sm_n n}^{(t)}(r_{sn},r_{m_n n}), \tag{18}$$

where the inequality follows from the fact that the quadratic term $(r_{sn}-r_{m_n n})^2$ is a convex function which is lower bounded by the first order Taylor series.

Based on (16), (17) and (18), the optimization Problem (14) after approximation is formulated as:

$$\min_{\boldsymbol{r}\geq 0}\ \sum_{n=1}^{N}\sum_{m=1}^{M}\left(\frac{\sigma^2}{H_{mn}}-\frac{\sigma^2}{H_{(m+1)n}}\right)2^{\sum_{s=1}^{m}\frac{r_{sn}}{B}}+\sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{s=1}^{m_n}$$

$$\lambda_m f_{sm_n n}^{(t)}(r_{sn},r_{m_n n})+\sum_{n=1}^{N}\frac{\left(\sum_{m=1}^{M}(w_{mn}^{(t)}r_{mn}+\alpha_{mn}^{(t)})\right)^K}{(L+0.5)^K}$$

$$(19a)$$

$$\text{s.t.}\ \sum_{n=1}^{N}r_{m_n n}=R_m,\quad\forall m. \tag{19b}$$

Since the objective function (19a) is convex and the constraints (19b) are linear, Problem (19) is a convex problem, which can be solved by using the popular interior-point method [15]. We now summarize the proposed joint power control and user clustering (JPCUC) algorithm for solving total power minimization Problem (14) in Algorithm 1.

3

Algorithm 1 JPCUC

1: Initialize a feasible $\boldsymbol{r}^{(0)}$ of Problem (14) and the iteration number $t=0$. Obtain the values of $w_{mn}^{(0)}$ and $\alpha_{mn}^{(0)}$ according to (16) and (17), respectively.
2: **repeat**:
3: Obtain the optimal $\boldsymbol{r}^{(t+1)}$ by solving Problem (19).
4: Set $t=t+1$, and update the values of $w_{mn}^{(t)}$ and $\alpha_{mn}^{(t)}$ according to (16) and (17), respectively.
5: **until** convergence

C. Convergence and Complexity Analysis

Theorem 2: Starting with any feasible point, the sequence $\{\boldsymbol{r}^{(t)}\}_{t=1}^{t=\infty}$ generated by Algorithm 1 is guaranteed to converge to a stationary point of Problem (14).

Proof: Please refer to Appendix B. ∎

For Algorithm 1, the major complexity in each iteration lies in solving the convex optimization (19). Considering that the dimension of the variables in Problem (19) is MN, the complexity of solving Problem (19) by using the standard interior point method is $\mathcal{O}(M^3N^3)$ [15, Page 487, 569]. Hence, the total complexity of the Algorithm 1 is $\mathcal{O}(TM^3N^3)$, where T denotes the total number of iterations.

IV. NUMERICAL RESULTS

There are $M=10$ users uniformly distributed in a square area of size 300 m × 300 m with the BS in the center. We set $B=1$ MHz, $\sigma^2=-174$ dBm/Hz, $\alpha_m=0.01$ Joule/Mbits, $\forall m$, and $K=10$. For propagation model, the large-scale path loss is $L(d)=128.1+37.6\log(d)$ [17], where the unit of d is kilometer, and the standard deviation of shadow fading is 4 dB. We consider equal minimum rate demand, i.e., $R_1=\cdots=R_M$. We compare the proposed JPCUC algorithm for NOMA systems (labeled as 'JPCUC-NOMA') with the relax-then-adjust algorithm for NOMA systems in [10] (labeled as 'RTA-NOMA'), the suboptimal matching for subcarrier assignment algorithm for NOMA systems in [7] (labeled as 'SOMSA-NOMA'), the dynamic programming algorithm for OMA systems in [18] (labeled as 'DP-OMA'), and the suboptimal matching for subcarrier assignment algorithm [7] with single-user detection (labeled as 'SOMSA-WSD').

According to Fig. 1, NOMA with $L=2$ achieves significant power savings over OMA, especially when the rate demand is high. This is because the transmission power part dominates the decoding power part which increases with number of users occupying one subcarrier. It is also observed that the performance of the proposed JPCUC-NOMA with $L=1$ is worse than OMA, since optimal subcarrier assignment is obtained in DP-OMA. Compared with RTA-NOMA, the proposed JPCUC-NOMA consumes lower total power. This is due to that JPCUC-NOMA does not restrict each subcarrier to be multiplexed by $L=2$ users. From Fig. 2, the JPCUC-NOMA outperforms the other two algorithms for NOMA, which shows the superiority of the proposed algorithm with ℓ_1 minimization and MM. The SOMSA-WSD yields the worst performance since SIC is conducted in NOMA to effectively mitigate inter-user interference. It is shown in Fig. 2 that sum power can be reduced when more users can be multiplexed on
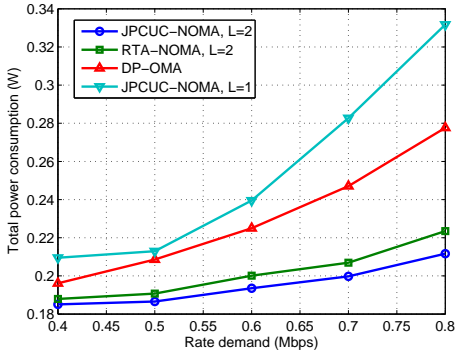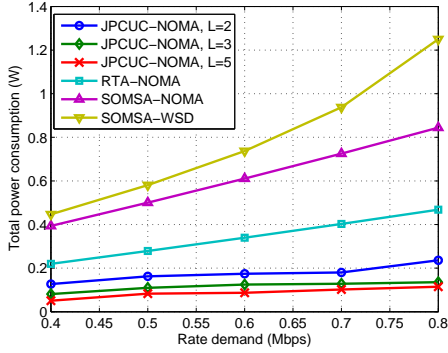
Fig. 1. Total power consumption with $N = 10$.

Fig. 2. Total power consumption with respect to L with $N = 5$.

each subcarrier, and the energy saving is marginal for large L. This is due to the tradeoff of the transmission power of the BS and the decoding power of the users according to (9).

V. CONCLUSIONS

We have formulated the power optimization in NOMA via joint power control and user clustering as an ℓ_0-norm form. An efficient algorithm with polynomial complexity is proposed and numerical results show the performance gain of the proposed algorithm with especially the popular matching-based user pairing algorithm.

APPENDIX A

The proof of Theorem 1 is established on showing that power vector $\boldsymbol{p}$ can be replaced by rate vector $\boldsymbol{r}$ without loss of optimality. According to (3), we can obtain

$$2^{\frac{r_{mn}}{B}} H_{mn} \sum_{l=m+1}^{M} p_{ln} + \sigma^2 \left(2^{\frac{r_{mn}}{B}} - 1\right) = H_{mn} \sum_{l=m}^{M} p_{ln}, \forall m, n. \tag{20}$$

To solve those MN equations, we first define

$$a_{mn} = \sum_{l=m}^{M} p_{ln}, \quad \forall m, n. \tag{21}$$

Substituting (21) into (20) yields

$$a_{mn} = 2^{\frac{r_{mn}}{B}} a_{(m+1)n} + \frac{\sigma^2}{H_{mn}} \left(2^{\frac{r_{mn}}{B}} - 1\right), \quad \forall m, n. \tag{22}$$

To solve (22), we denote $\boldsymbol{a}_n = [a_{1n}, \cdots, a_{Mn}]^T$,

$$\boldsymbol{b}_n = \left[\frac{\sigma^2}{H_{1n}} \left(2^{\frac{r_{1n}}{B}} - 1\right), \cdots, \frac{\sigma^2}{H_{Mn}} \left(2^{\frac{r_{Mn}}{B}} - 1\right)\right]^T, \tag{23}$$

and

$$\boldsymbol{W}_n = \begin{bmatrix} 0 & 2^{\frac{r_{1n}}{B}} & & & \\ & 0 & 2^{\frac{r_{2n}}{B}} & & \\ & & \ddots & \ddots & \\ & & & 0 & 2^{\frac{r_{(M-1)n}}{B}} \\ & & & & 0 \end{bmatrix}. \tag{24}$$

Equations in (22) can be rewritten as

$$(\boldsymbol{I} - \boldsymbol{W}_n)\boldsymbol{a}_n = \boldsymbol{b}_n, \tag{25}$$

where $\boldsymbol{I}$ is an identity matrix of size M. From (25), we have

$$\boldsymbol{a}_n = (\boldsymbol{I} - \boldsymbol{W}_n)^{-1}\boldsymbol{b}_n. \tag{26}$$

From the special structure of $\boldsymbol{W}_n$ in (24), we can obtain manipulation $\boldsymbol{W}_n^l$ for $l \geq 1$ by using the recursion method, and specifically $\boldsymbol{W}_n^M = \boldsymbol{0}$. Since

$$(\boldsymbol{I} - \boldsymbol{W}_n)\left(\boldsymbol{I} + \sum_{l=1}^{M-1} \boldsymbol{W}_n^l\right) = \boldsymbol{I} - \boldsymbol{W}_n^M = \boldsymbol{I}, \tag{27}$$

we have

$$(\boldsymbol{I} - \boldsymbol{W}_n)^{-1} = \boldsymbol{I} + \sum_{l=1}^{M-1} \boldsymbol{W}_n^l. \tag{28}$$

Substituting (28) into (26) yields

$$a_{mn} = \sum_{l=m}^{M} \frac{\sigma^2}{H_{ln}} \left(2^{\frac{r_{ln}}{B}} - 1\right) 2^{\sum_{s=m}^{l-1} \frac{r_{sn}}{B}}, \quad \forall m, n, \tag{29}$$

where we define $2^{\sum_{s=m}^{m-1} \frac{r_{sn}}{B}} = 2^0$. From (21) and (29), we have

$$\sum_{m=1}^{M} p_{mn} = a_{1n} = \sum_{m=1}^{M} \frac{\sigma^2}{H_{mn}} 2^{\sum_{s=1}^{m} \frac{r_{sn}}{B}} - \sum_{m=1}^{M} \frac{\sigma^2}{H_{mn}} 2^{\sum_{s=1}^{m-1} \frac{r_{sn}}{B}}$$

$$= \sum_{m=1}^{M} \left(\frac{\sigma^2}{H_{mn}} - \frac{\sigma^2}{H_{(m+1)n}}\right) 2^{\sum_{s=1}^{m} \frac{r_{sn}}{B}} - \frac{\sigma^2}{H_{1n}}. \tag{30}$$

Equation (30) implies that power vector $\boldsymbol{p}$ can be expressed by rate vector $\boldsymbol{r}$. From (3), we can observe that $p_{mn} = 0$ if and only if $r_{mn} = 0$, and $p_{mn} > 0$ if and only if $r_{mn} > 0$. Thus, we can obtain

$$\|\boldsymbol{p}_n\|_0 = \|\boldsymbol{r}_n\|_0, \quad \forall n. \tag{31}$$

Substituting (30) and (31) into Problem (10), we can find that Problem (11) is equivalent to Problem (10).

APPENDIX B

Since $\ln(1 + x)$ is concave, the inequality $\ln(1 + x)^\beta \leq ((1 + x_0)^{-1}x + \ln(1 + x_0) - (1 + x_0)^{-1}x_0)^\beta$ holds for any $x \geq 0$, $x_0 \geq 0$ and $\beta > 0$, and achieves equality if and only if $x = x_0$. Therefore, we can obtain (14a) $\leq$ (19a) with equality hold if and only if $r_{mn} = r_{mn}^{(t)}$, $\forall m, n$.

Algorithm 1 is equivalent to an majorization-minimization (MM) algorithm, which can be proved to converge to a stationary point of the original problem if the approximate objective function satisfies three conditions. An MM algorithm is guaranteed to converge to a stationary point of the original problem if the approximate objective function satisfies the following conditions according to [17, Appendix A]: 1) it is continuous, 2) it is a tight upper bound of the original objective function and 3) it has the same first-order derivative as the original objective function at the point where the upper bound is tight. Obviously, function in (19a) satisfies all these sufficient conditions. Thus, Algorithm 1 must converge.

REFERENCES

[1] L. Dai, B. Wang, Y. Yuan, and S. Han, "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.

[2] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[3] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Annu. Symp. Personal, Indoor and Mobile Radio Commun.*, London, U.K., Sep. 2013, pp. 611–615.

[4] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Jul. 2014.

[5] Z. Yang, W. Xu, C. Pan, Y. Pan, and M. Chen, "On the optimality of power allocation for noma downlinks with individual qos constraints," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1649–1652, Jul. 2017.

[6] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[7] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3722–3732, Sep. 2016.

[8] C. W. Sung and Y. Fu, "A game-theoretic analysis of uplink power control for a non-orthogonal multiple access system with two interfering cells," in *IEEE Veh. Technol. Conf.*, Nanjing, China, May 2016, pp. 1–5.

[9] Y. Fu, Y. Chen, and C. W. Sung, "Distributed downlink power control for the non-orthogonal multiple access system with two interfering cells," in *IEEE Int. Conf. Commun.*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[10] L. Lei, D. Yuan, and P. Värbrand, "On power minimization for non-orthogonal multiple access (NOMA)," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2458–2461, Dec. 2016.

[11] C. Xiong, G. Y. Li, Y. Liu, and S. Xu, "When and how should decoding power be considered for achieving high energy efficiency?" in *Proc. IEEE Annu. Symp. Personal, Indoor and Mobile Radio Commun.*, Sydney, NSW, Australia, Sep. 2012, pp. 2427–2431.

[12] J. Rubio and A. Pascual-Iserte, "Energy-aware broadcast multiuser-MIMO precoder design with imperfect channel and battery knowledge," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3137–3152, Jun. 2014.

[13] T. M. Nguyen, A. Yadav, W. Ajib, and C. Assi, "Energy efficiency with adaptive decoding power and wireless backhaul small cell selection," in *Proc. IEEE Global Commun. Conf.*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[14] C. G. Blake and F. R. Kschischang, "Energy consumption of VLSI decoders," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3185–3198, Jun. 2015.

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

[16] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, 2008.

[17] B. Dai and W. Yu, "Energy efficiency of downlink transmission strategies for cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 1037–1050, Apr. 2016.

[18] D. Yuan, J. Joung, C. K. Ho, and S. Sun, "On tractability aspects of optimal resource allocation in OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 863–873, Feb. 2013.