This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2772268, IEEE Transactions on Communications
The final version of record is available at https://dx.doi.org/10.1109/TCOMM.2017.2772268

Relay Selection for Multi-Channel Cooperative Multicast: Lexicographic Max-Min Optimization

Yitu Wang, *Student Member*, Wei Wang, *Senior Member, IEEE*, Lin Chen, *Member, IEEE*, Pan Zhou, *Member, IEEE*, Zhaoyang Zhang, *Member, IEEE*

***Abstract*—Cooperative multicast has been demonstrated to achieve significant performance gain over the classic source-destination transmission paradigm by exploiting spatial diversity through the participation of multiple relay nodes. As a major technical challenge, the selection of relays for a multicast session has significant impact on the multicast performance. The challenge is even more pronounced when the number of channels are limited as the relay selection is in this context coupled with channel allocation. The goal of this paper is to design a fair multicast relay selection scheme with limited channel resources. Specifically, we establish an analytical framework for this joint relay selection and channel allocation problem and develop a lexicographic max-min multicast relay selection scheme. Our design consists of two technical steps. 1) We consider the maximization of the minimal data rate. By decoupling relay selection and channel allocation, the problem is transformed to a max-min-max problem, which is difficult to solve. To make this problem tractable, we reformulate it as a convex optimization problem via relaxation and smoothing, and prove the asymptotic equivalence from a geometrical perspective. 2) We propose an adjustment algorithm based on the initial max-min solution, and prove that the proposed scheme achieves lexicographic optimality. Finally, our proposed algorithm is evaluated by simulation to show its superiority over the conventional schemes.**

I. Introduction

Nowadays emerging wireless multimedia applications such as mobile TV are more and more exigent on data rate over limited spectrum resources. Multicast is a spectrum-efficient paradigm for one-to-many transmissions over wireless channels by enabling service providers to send multimedia data to multiple users simultaneously [2], [3], [4]. To further proliferate multimedia applications over wireless networks and enable quality of service (QoS)-aware wireless video streaming, Scalable video coding (SVC) stands out with its graceful rate adaptation capabilities to cope with bandwidth scarcity and network variation, and is often employed in multicast so as to improve radio resource utilization and provide differentiated QoS [5]. In a nutshell, SVC divides a multimedia stream into one base layer and multiple enhancement layers, where the base layer provides a minimum quality of the multimedia while the enhancement layers gradually increase the quality. QoS-aware streaming can be achieved by including different number of the enhancement layers during transmission to conform to network variation, hardware heterogeneity or users requirement.

As a recently emerged technique, cooperative multicast takes the merit of spatial diversity to combat the influence of path loss and channel fading in order to further improve the multicast capacity [6], [7], [8], [9]. In the canonic two-hop cooperative multicast system, the source node first transmits data to relay nodes, then the users requesting the same data can be logically grouped as multicast groups and served by designated relay nodes respectively. Although cooperative multicast has the potential to increase the capacity, an improper relay selection scheme will result in an even lower data rate than that in the non-relay-assisted transmission paradigm. Therefore, the relay selection scheme should be carefully designed so as to fully exploit the performance gain brought by cooperative relay.

A. Motivation

Due to its importance, a large body of works have considered the relay selection problem in cooperative multicast. For one channel case, maximal ratio combining (MRC) is usually adopted, which combines the signals received from different transmitters to increase the signal-to-noise ratio [10]. However, the MRC technique is generally not compatible with the SVC paradigm due to the implementation issue. Different transmitters may transmit different layers of the broadcast content, which cannot be combined by simply applying MRC. With perfect synchronization and coordination, it is possible to incorporate MRC with SVC, however, imperfect synchronization results in significant performance loss and the associated communication overhead is fairly large, and thus makes MRC costly to implement for the scenario with multiple relay nodes and multiple destination nodes[1].

To address the above problem, a natural approach is to use multiple orthogonal channels for different transmitters. In this research strand, most existing works assume that the

Part of this work has been presented at IEEE WCNC 2017, San Francisco, U.S.A. [1].

This work is supported in part by National Natural Science Foundation of China (Nos. 61571396, 61725104, 61401169), Zhejiang Provincial Natural Science Foundation of China (No. LR17F010001), Young Elite Scientist Sponsorship Program by CAST (No. 2016QNRC001), Talent Project of ZJAST (No. 2017YCGC011), and Fundamental Research Funds for the Central Universities (No. 2015XZZX001-03). The corresponding author is Wei Wang (Email: wangw@zju.edu.cn).

Yitu Wang, Wei Wang and Zhaoyang Zhang are with College of Information Science and Electronic Engineering, Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Zhejiang University, Hangzhou 310027, China.

Lin Chen is with Laboratoire de Recherche en Informatique (LRI), University of Paris-Sud and CNRS, Orsay 91405, France.

Pan Zhou is with School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China.

[1]In Section IV.E, we discuss incorporating MRC to further improve the performance.

number of orthogonal channels are sufficient to avoid co-channel interference among relays [11], [12], [13]. However, in many practical networks, e.g., IEEE 802.11 [14], the number of available channels is rather limited, thus rendering this assumption invalid. Under this context with limited number of channels, it is clear that a subset of relay nodes need to be deactivated to avoid interference, which brings a new dimension on the problem of relay selection.

B. Main Contribution and Results

Motivated by the above analysis, in this paper we address the problem of relay selection in cooperative multicast with a limited number of channels. Our goal is to achieve a lexicographic max-min optimal solution. Lexicographic optimization is a well-recognized fair optimility criterion [15], [16], [17] for multi-objective optimization problems. Theoretically, it is proven that the lexicographically optimal vector is uniquely optimal over any given convex and compact set, and such a solution is always Pareto optimal [18].

The main technical challenge in our problem comes from the limitation of the channel number, which creates complex interdependence among relays and thus makes the problem of relay selection an involved optimization problem. To address this challenge, we decompose the problem and proceed by two steps. 1) We consider the problem of maximizing the minimal data rate. By decoupling relay selection and channel allocation, we transform the problem to a max-min-max problem. To make the transformed problem tractable, we reformulate it as a convex optimization problem via relaxation and smoothing, and establish the asymptotic equivalence from a geometrical perspective. 2) With the solution in the first step, we further develop an adjustment procedure, which is proved to produce a lexicographically optimal solution.

The rest of the paper is organized as follows. Section II discusses the related works. Section III presents the system model. Section IV proposes the details in designing the relay selection algorithm. The performance of the proposed algorithm is evaluated by simulation in Section V. Finally, this paper is concluded in Section VI.

II. Related Works

This paper develops an analytical framework for the problem of relay selection in cooperative multicast with a limited number of channels. In this section, we briefly review the existing works on the receiver design and the fairness considerations.

A. Receiver Design

Multicast relay selection has attracted increasing attention recently for its spectral efficient nature. It is more efficient and realistic compared with the conventional model in [11], which restricts each relay node to be assigned to only one destination node for cooperative multicast. Multicast relaying provides an extra degree of freedom to achieve higher capacity, while brings additional constraint during the scheme design.

For one channel case, MRC is usually adopted and nodes transmit and receive data at the same channel. Most existing works on the cooperative multicast networks using MRC in SVC scenario investigate the case that a source node broadcasts data to multiple relay nodes, and then the selected relay nodes broadcast data to one receiver [19], [20], where the PHY layer behavior is analysed, e.g., the outage probability. In [21], the capacity and SNR performance is analysed in a multi-receiver case, but the source node broadcasts the basic layer and one enhancement layer, then the relay nodes transmit the basic layer only. There are also a few publications investigating cooperative multicast networks using MRC without SVC [22], [23]. In this case, the relay nodes transmit the same data to the destination nodes, where the user diversity is not fully exploited. The MRC technique is generally not compatible with the SVC paradigm because different transmitters may transmit different number of the enhancement layers of the broadcast content, which cannot be combined by simply applying MRC.

To address the above problem, a natural approach is to use multiple orthogonal channels for different transmitters. For multi-channel case, in [12] and [13], a group of intended users receive the same data from the source, where the former assumes that the selected multicast relays transmit in orthogonal channels, and investigates the optimal relay scheduling and power allocation strategies to minimize the total power consumption, while the latter proposes a distributed energy efficient multicast relay selection scheme. In [24], the authors propose an optimal multicast relay selection scheme achieving maximal capacity in deterministic relay networks without cross-channel interference. The authors in [25] consider cooperative multicast of videos over wireless networks, where the relays use TDMA to forward packets, and adopt layered video coding to provide different video qualities to the users according to their channel conditions, which does not provide any discussion on the algorithm optimality. Most existing publications assume that enough orthogonal channels are available for cooperative multicast without interference among relays. In many practical networks, e.g., IEEE 802.11 [14], the number of available channels is rather limited, thus rendering this assumption invalid. However, few publication discusses multicast relay selection scheme considering a limited number of channels.

B. Fairness Consideration

One of the mostly used criteria for resource allocation is maximizing the total throughput of the system known as max-sum criterion. The weak point of this criterion lies in the fact that the users achieve less resources due to their poor link qualities. To achieve the fairness, more resources should be allocated to the users with the poorest channel condition, which is known as the max-min optimization. The main problem of the max-min optimization is that the optimal resource allocation is not necessarily Pareto optimal. In other words, starting from the max-min optimal resource allocation, one can increase the utility of one individual without decreasing the utilities of the others, which is clearly not a desirable property of an efficient resource allocation algorithm. Moving one step ahead, the solution of the lexicographic max-min

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2772268, IEEE Transactions on Communications
The final version of record is available at https://dx.doi.org/10.1109/TCOMM.2017.2772268

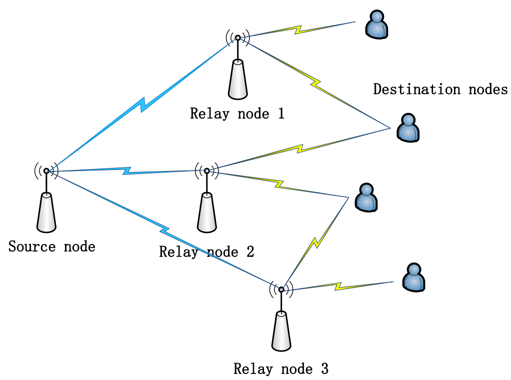


Fig. 1. Cooperative Multicast System

optimization is a refinement of the standard max-min concept. This ordering takes both fairness and efficiency into account, which is addressed in few previous works [15], [16], [17] in cooperative networks. In [15], the cooperative transmission in cognitive radio networks is considered to provide reliable communication for secondary users by lexicographically optimizing the received data rates. In [16], a lexicographic resource allocation scheme is proposed by allowing the relays to transmit their own data frames while performing cooperative transmission. In [17], the joint optimization of subcarrier-relay assignment and power allocation is investigated to obtain a lexicographically optimal solution in an OFDMA system, without considering the coupling of channel allocation and relay selection, such that channel allocation and relay selection can be optimized separately, which significantly simplifies the problem. Different to [17], we limit the number of channels in the system, and address the coupling of the relay selection and the channel allocation head-on to derive a lexicographically optimal solution.

III. System Model

A. Cooperative Multicast Model

Consider a wireless cooperative multicast network, consisting of a source node $\mathcal{S} = \{s\}$, M relay nodes $\mathcal{R} = \{r_1, r_2, \cdots, r_M\}$ and N destination nodes $\mathcal{D} = \{d_1, d_2, \cdots, d_N\}$. Relay nodes multicast with SVC to improve the wireless resource utilization and provide differentiated QoS according to the weakest channel conditions in their corresponding multicast group. The time is slotted and the duration of each time slot is assumed to be a unit of time without loss of generality. Due to the long distance or the shielding effect caused by some barrier between the source node and the destination nodes, the destination nodes are not within the communication range of the source node, so that all signals received at the destination nodes need to be forwarded by the assisting relay nodes[2]. A two-stage cooperative relay in multicast transmission paradigm is adopted in this work, which is commonly used to improve the performance of the users [22], [23], [27]. It takes two time slots to accomplish the cooperative multicast [3]. In the first time slot, as the blue links shown in Fig. 1, the source node s broadcasts the data to the relay nodes according to the weakest source-relay channel condition γ. To focus on the relay selection problem, γ is assumed to be large enough to support the transmission of at least the basic layer [28]. In the second time slot, the relay nodes multicast the received data to the destination nodes simultaneously, where the multicast data rates are determined according to the weakest channel conditions in their corresponding multicast groups, as the yellow links in Fig. 1. As in [8], we assume that K orthogonal channels are available in the network (e.g., using OFDMA), denoted as $\mathcal{C} = \{c_1, c_2, \cdots, c_K\}$, which are flat and remain constant over time slot [29], [30].

Let $\mathcal{G} = (V, E)$ denote the conflict graph of the relay nodes, where each relay node $r_i \in \mathcal{R}$ is one of the vertices in V of the conflict graph, and $(r_i, r_j) \in E$ implies that r_i and r_j cannot transmit on the same channel simultaneously since their transmissions interfere with each other. If any destination node can receive the signal from both the relay nodes r_i and r_j, then $(r_i, r_j) \in E$. Taking Fig. 1 as an example, we obtain that the relay nodes r_1 and r_2 conflict with each other, r_2 and r_3 conflict while r_1 and r_3 do not conflict.

Both the source node and the relay nodes transmit the signal with a unit power. For the destination node d_j, its received signal from the relay node r_i can be written as

$$y_{ij} = \sqrt{D_{ij}^{-\alpha}} h_{ij} x + n_{ij}, \quad (1)$$

where h_{ij} is the fading coefficient between destination node d_j and relay node r_i, α is the path loss exponent depending on the propagation environment, D_{ij} is the distance between destination node d_j and relay node r_i, and n_{ij} is additive white Gaussian noise with variance N_0. Therefore, when the relay node r_i transmits, the recieved signal-to-noise ratio of d_j is

$$\gamma_{ij} = \frac{|h_{ij}|^2 D_{ij}^{-\alpha}}{N_0}. \quad (2)$$

When γ_{ij} is greater than a certain threshold, the signal can be received and decoded successfully. To this end, the enhancement layers are selected according to the channel capacity, i.e., the total size of the basic layer and the selected enhancement layers should not be larger than the channel capacity, while the number of the enhancement layers should be chosen as large as possible to provide the best possible QoS.

Decode-and-forward transmission mode is adopted in the cooperative multicast systems to be compatible with SVC. The relay node r_i decodes the signal received from the source node s and then transmits the decoded data to the destination node d_j. The capacity from source node s to destination node d_j with the assistance of relay node r_i is

$$C_{ij} = \frac{W}{2} \min\{\log_2(1+\gamma), \log_2(1+\gamma_{ij})\}. \quad (3)$$

When the signals from multiple relay nodes reach a destination node, the selection combiner is adopted because these

[2]Such a two-hop relaying model is practical and has been widely adopted to extend the communication coverage, e.g., the Type-II relaying in IEEE and 3GPP standards [26].

[3]By embracing multicast, more relay nodes are able to receive the content, and thus more destination nodes are able to receive the content from a nearby relay node, such that better QoS can be potentially achieved.

signals are transmitted over different channels for avoiding the interference. The destination node receives the signal from only one relay node which has the largest number of the enhancement layers so as to achieve the best QoS. Following [31], it is assumed that the channel state information between the destination node and the relay nodes are known by this destination node from the report of the relay nodes. Specifically, relay nodes report a received-signal-strength information via control messages, and the destination nodes derive the channel state information according to the messages before choosing a relay node.

B. Problem Formulation

Denote the relay selection scheme as $\mu : \mathcal{D} \to \mathcal{R}$, where $\mu(d_j) = r_i$ indicates the relay node r_i is selected to help the transmission to d_j via cooperative multicast. Note that one relay node can help the transmission of multiple destination nodes by multicast, i.e., it is possible that $\mu(d_i) = \mu(d_j)$ for $i \neq j$, which is different from the models in [8], [11], where a relay node can be assigned to assist only one destination node. The channel allocation matrix of the relay nodes is denoted as $\boldsymbol{\tau} = \{\tau_{ik}\}_{M\times K}$, where $\tau_{ik} = 1$ represents that relay node r_i is activated using channel c_k.

Considering the multicast nature of wireless communication systems, a relay node r_i multicasts data to the destinations in $\mathcal{D}_i = \{d_j | \mu(d_j) = r_i, \forall d_j \in \mathcal{D}\}$ with a maximal rate of

$$R_i = \sum_{c_k \in \mathcal{C}} \tau_{ik} \min_{d_j \in \mathcal{D}_i} \{C_{ij}\}, \quad (4)$$

such that all the destination nodes in $\mathcal{D}_i$ can successfully decode the data.

To provide a unique fair solution that outperforms all possible classic max-min solutions [18], our goal is to design a lexicographic max-min relay selection scheme, in which the lexicographically optimal rate vector is no lexicographically less than that of any other scheme. We define the lexicographic optimality formally as follows:

Definition 1 (Lexicographic Optimality)**.** *Let* $\mathbf{R} = (\nu_1, \nu_2, ..., \nu_N)$ *be an achievable rate vector which is sorted in non-descending order, where* ν_i *represents the* i*-th smallest data rate. Two vectors* $\mathbf{R}$ *and* $\mathbf{R}'$ *have the following relationships:*

- *If* $\nu_i = \nu_i'$, $\forall i = 1, 2, \cdots, N$, *then* $\mathbf{R}$ *is lexicographically equal to* $\mathbf{R}'$.
- *If there exist a prefix* $(\nu_1, \nu_2, ...\nu_i)$ *of* $\mathbf{R}$ *and a prefix* $(\nu_1', \nu_2', ...\nu_i')$ *of* $\mathbf{R}'$ *such that* $\nu_i > \nu_i'$, *and* $\nu_j = \nu_j'$ *for* $1 \le j \le i-1$, *then* $\mathbf{R}$ *is lexicographically greater than* $\mathbf{R}'$.

A rate vector $\mathbf{R}$ *is lexicographically optimal if it is no lexicographically less than all other feasible rate vectors.* ■

Remark 1. *The idea for finding the lexicographic max-min solution is to sequentially identify all the max-min solutions and to sort the vectors in weakly decreasing order to identify the lexicographically optimal one. Specifically, lexicographic max-min optimization takes into account maximizing the second smallest item, maximizing the third smallest one and further to be hierarchically maximized.*

The received data rates for destination nodes are lexicographically optimized by determining which relay nodes should be activated and which destination nodes that these relay nodes forward data to. Based on the definition of lexicographic optimality in Definition 1, we can formulate the lexicographic max-min problem as follows.

$$\begin{aligned} \operatorname{lex}\max_{\boldsymbol{\tau},\mu}\ & \mathbf{R} \\ s.t.\ & \sum_{c_k \in \mathcal{C}} \tau_{ik} \le 1 \\ & \tau_{ik} + \tau_{lk} \le 1, \forall (r_i, r_l) \in E \\ & \tau_{ik} \in \{0, 1\}, \end{aligned} \quad (5)$$

where $\operatorname{lex}\max$ indicates the operation of lexicographic maximization. The first constraint indicates that a relay node can use at most one channel which depends on the fact that usually only one radio interface is deployed in a device. The second constraint represents that if two relay nodes r_i and r_l interfere with each other, they must engage different channels to avoid cross-channel interference. The optimization problem in Eq. (5) lexicographically maximizes the rate vector by determining the relay selection scheme μ and channel allocation scheme $\boldsymbol{\tau}$.

IV. Algorithm Design

In this section, we propose an analytical framework for lexicographic max-min multicast relay selection scheme for cooperative multicast with a limited number of channels. As we discussed before, the major technical challenge is induced by the complicated coupling between relay selection and channel allocation. It is difficult to decouple them in the lexicographic optimization problem. To overcome this challenge, instead of solving the lexicographic optimization directly, we first consider the optimization problem to maximize the minimal data rate, where it is possible to decouple relay selection and channel allocation. Technically, we address the problem in two steps:

- Consider only the minimal data rate and solve the max-min problem to obtain an initial relay selection solution.
- Optimize the rates of other nodes by further adjusting the relay selection to achieve lexicographic optimality.

A. Decoupling in Max-Min Subproblem

To design the lexicographic max-min scheme, we consider only the minimal data rate first. Maximizing the minimal data rate among the destination nodes is equivalent to maximizing the minimal multicast rate among the relay nodes according to Eq. (4), we can formulate the max-min subproblem from Eq. (5) as follows:

$$\begin{aligned} \max_{\boldsymbol{\tau},\mu} \min_{i}\ & R_i \\ s.t.\ & \sum_{c_k \in \mathcal{C}} \tau_{ik} \le 1 \\ & \tau_{ik} + \tau_{lk} \le 1, \forall (r_i, r_l) \in E \\ & \tau_{ik} \in \{0, 1\}. \end{aligned} \quad (6)$$

This max-min problem involves both the relay selection scheme μ and the channel allocation scheme $\boldsymbol{\tau}$, which are coupled with each other due to the limited number of channels. To decouple these two aspects, we exploit the property of the max-min problem and suggest a capacity-based relay selection scheme which is independent to channel allocation in the following lemma.

Lemma 1 (Capacity-Based Relay Selection)**.** *The max-min solution of the cooperative multicast system is achievable only by adjusting channel allocation if each destination node d_j joins the multicast group of the relay r_i which has the largest channel capacity C_{ij}, i.e.,*

$$\max_{\boldsymbol{\tau},\mu} \min_{i} \; R_i = \max_{\boldsymbol{\tau}} \min_{i} \; R_i(\hat{\mu}), \tag{7}$$

where

$$\hat{\mu}(d_j) = \arg\max_{r_i \in \mathcal{R}} \{C_{ij}\}. \tag{8}$$

Proof: The weakest link in a cooperative multicast system is the link of the destination node d_k satisfying

$$d_k = \arg\min_{d_j \in \mathcal{D}} \{\max_{r_i \in \mathcal{R}} \{C_{ij}\}\}. \tag{9}$$

To maximize the minimal data rate of the system, we can maximize the received data rate for the destination node d_k instead. Under the relay selection scheme $\hat{\mu}$ in Lemma 1, even though not all destination nodes receive a data rate as much as the channel link capacity according to Eq. (4), the destination node k indeed receives data at a rate of $\max_{r_i \in \mathcal{R}}\{C_{ij}\}$, which is maximized among all possible choices. ∎

It is noted that such a scheme $\hat{\mu}$ in Lemma 1 does not achieve lexicographic optimality, since the data rates of other destination nodes are not taken into consideration. We will later propose a relay selection scheme μ^* that achieves lexicographic optimality based on $\hat{\mu}$.

To decouple relay selection and channel allocation in Eq. (6), we adopt the relay selection scheme $\hat{\mu}$ in Lemma 1 and transform the max-min problem in Eq. (6) to a max-min-max problem which is a channel allocation problem only.

$$\begin{aligned} \max_{\boldsymbol{\tau}} \;\; & \Phi(\boldsymbol{\tau}) = \min_{j} \max_{i} \sum_{c_k \in \mathcal{C}} \tau_{ik} C_{ij} \\ s.t. \;\; & \sum_{c_k \in \mathcal{C}} \tau_{ik} \le 1 \\ & \tau_{ik} + \tau_{lk} \le 1, \forall (r_i, r_l) \in E \\ & \tau_{ik} \in \{0,1\}. \end{aligned} \tag{10}$$

B. Tractable Reformulation for Channel Allocation

Due to the non-smooth structure of max-min-max function, the max-min-max problem in Eq. (10) is very difficult to solve both in theoretical analysis and in numerical calculation [32]. To solve the max-min-max problem directly by numerical calculation, the result comes close to the exhausting method, which definitely faces the curse of dimensionality, i.e., the complexity exponentially grows with the size of the problem, and is not applicable to a real system [33]. To solve the problem, we transform the max-min-max problem into a convex one by relaxation and smoothing techniques, and further prove that the relaxation and smoothing are tight.

The max-min-max problem is a combinational optimization problem, which is not differentiable due to the 0-1 integral constraint in Eq. (10), so that traditional optimization techniques cannot approach the problem. To obtain the optimal solution, we relax the 0-1 integral constraint in Eq. (10) to a box constraint as

$$\tau_{ik} \in [0,1], \tag{11}$$

so that the objective in Eq. (10) is a continuous function of $\boldsymbol{\tau}$.

With the above relaxation, the max-min-max problem is continuous but still non-differentiable, we further adopt a smoothing technique to approximate the original max-min-max optimization problem, and the transformed approximation problem is differentiable about $\boldsymbol{\tau}$.

Adopting the smoothing technique in [34], the objective $\Phi(\boldsymbol{\tau})$ of the max-min-max problem in Eq. (10) can be approximated by

$$\Phi_\epsilon(\boldsymbol{\tau}) = \frac{1}{\epsilon} \ln \left(\sum_{i=1}^{M} \frac{1}{\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}} \right) + \frac{\ln M}{\epsilon}, \tag{12}$$

where ϵ is the approximation parameter. For given ϵ, the objective function in Eq. (12) can be safely transformed into

$$\ln \left(\sum_{i=1}^{M} \frac{1}{\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}} \right). \tag{13}$$

Consider the exponential of the objective function in Eq. (13), the optimality is preserved according to the monotone property of the exponential function [36]. The optimization problem in Eq. (10) can be transformed to

$$\begin{aligned} \min_{\boldsymbol{\tau}} \;\; & \sum_{i=1}^{M} \frac{1}{\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}} \\ s.t. \;\; & \sum_{k=1}^{K} \tau_{ik} \le 1 \\ & \tau_{ik} + \tau_{jk} \le 1, \forall (r_i, r_j) \in E \\ & 0 \le \tau_{ik} \le 1. \end{aligned} \tag{14}$$

We prove the convexity of the problem in Eq. (14) in the following Theorem

Theorem 1 (Convexity of Problem (14))**.** *The optimization problem in Eq. (14) is convex.*

Proof: We analyze the Hessian of the objective function in Eq. (14), which is proved to be positive semi-definite and hence it is convex. Since all the constraints in Eq. (14) are linear, we prove that the problem in Eq. (14) is a convex optimization problem. More details of the proof are provided in Appendix A. ∎

C. Geometrical Analysis

To obtain some critical insights of the above convex problem, we consider the relay selection optimization problem in Eq. (14) as a geometrical problem which investigates the relationship of positions of a line and multiple points in two-dimensional space. This method can decrease the computational complexity significantly. More importantly, the relaxation is proved to be tight by adopting geometrical analysis.

We first consider the optimality condition of the problem in Eq. (14). According to Lagrangian duality [36], the problem can be transformed to

$$\begin{aligned} \min_{\boldsymbol{\tau}} \quad P = & \sum_{i=1}^{M} \frac{1}{\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}} + \sum_{i=1}^{M} \lambda_i \left(\sum_{k=1}^{K} \tau_{ik} - 1 \right) \\ & + \sum_{k=1}^{K} \sum_{i=1}^{M} \sum_{j,(r_i,r_j)\in E} \beta_{ijk}(\tau_{ik} + \tau_{jk} - 1) \\ s.t. \quad & \lambda_i \left(\sum_{k=1}^{K} \tau_{ik} - 1 \right) = 0 \\ & \beta_{ijk}(\tau_{ik} + \tau_{jk} - 1) = 0, \end{aligned} \tag{15}$$

where λ_i is the Lagrangian multiplier for the first constraint in Eq. (14) and β_{ijk} is the Lagrangian multiplier for the second constraint in Eq. (14).

The derivative of P with respect to $\boldsymbol{\tau}$ is

$$\frac{\partial P}{\partial \tau_{in}} = -\frac{1}{\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}} \sum_{j=1}^{N} \epsilon C_{ij} + \lambda_i + \sum_{j,(r_i,r_j)\in E} \beta_{ijn}. \tag{16}$$

Based on different cases of optimal $\boldsymbol{\tau}^*$, the optimality conditions are provided as

$$\begin{cases} \frac{\partial P}{\partial \tau_{in}} \mid_{\tau_{in}=1} \leq 0 \longrightarrow \tau_{in}^* = 1 \\ \frac{\partial P}{\partial \tau_{in}} \mid_{\tau_{in}=\tau_{in}^*} = 0 \longrightarrow 0 < \tau_{in}^* < 1 \\ \frac{\partial P}{\partial \tau_{in}} \mid_{\tau_{in}=0} \geq 0 \longrightarrow \tau_{in}^* = 0. \end{cases} \tag{17}$$

A relay node r_i is activated to transmit using the channel c_n if $\tau_{in} > 0$. When $0 < \tau_{in}^* < 1$, the optimal solution is achieved if $\frac{\partial P}{\partial \tau_{in}} \mid_{\tau_{in}=\tau_{in}^*} = 0$. When $\tau_{in}^* = 1$, because the Lagrangian multiplier λ is configurable for each relay node, we can let $\frac{\partial P}{\partial \tau_{in}} = 0$ by adjusting λ. Therefore, the optimal solution of a relay node transmitting with a channel is $\frac{\partial P}{\partial \tau_{in}} \mid_{\tau_{in}=\tau_{in}^*} = 0$ for $\tau_{in} > 0$. Therefore, it is necessary to analyze the relationship between Eq. (16) and zero. We rewrite Eq. (16) by multiplying $\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}$ as

$$-\sum_{j=1}^{N} \epsilon C_{ij} + \left(\lambda_i + \sum_{j,(r_i,r_j)\in E} \beta_{ijn} \right) \sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}}. \tag{18}$$

Since $\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}} > 0$, the multiplication does not affect the relationship between Eq. (16) and zero.

To analyze Eq. (18), we rewrite the optimality condition for the problem in Eq. (18) as an expression of a line in two-dimensional space [38],

$$y_{in} = A_i x_{in}, \tag{19}$$

where

$$\begin{aligned} x_{in} &= \lambda_i + \sum_{j,(r_i,r_j)\in E} \beta_{ijn} \\ A_i &= \sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{ik} C_{ij}} \\ y_{in} &= \sum_{j=1}^{N} \epsilon C_{ij}. \end{aligned}$$

From a geometrical perspective, each relay r_i using the channel c_n has a corresponding point $S_{in} = (x_{in}, y_{in})$ in the two dimensional space. Define $\mathcal{S}_i$ as the set of all points for r_i, i.e., $\mathcal{S}_i = \{S_{in}, \forall n\}$. For given λ, β and the relay node r_i, the coordinates x_{in} and y_{in} are determinate. The problem is to determine A_i, which is the slope of the line $Y_i = A_i X_i$ of the relay node r_i. In such a case, the problem is transformed to a geometrical problem on the position relationship of the points $S_{in} \in \mathcal{S}_i$ and the line $Y_i = A_i X_i$. The geometrical problem is to find a line $Y_i = A_i X_i$ by adjusting $\boldsymbol{\tau}$ to let some of the points $S_{in} \in \mathcal{S}_i$ on the line and all the other points under the line.

If only one relay node r_i transmits using the channel c_n, i.e., the line $Y_i = A_i X_i$ goes through $S_{in} = (x_{in}, y_{in})$ and all the other points $S_{ik} = (x_{ik}, y_{ik}), k \neq n$ are under the line, the optimal conditions are

$$y_{in} = A_i x_{in} \tag{20}$$

$$y_{ik} < A_i x_{ik}, \forall k \neq n. \tag{21}$$

In this way, we obtain the optimal conditions for the reformulated convex optimization problem in Eq. (14). Besides providing the optimal conditions, the geometrical analysis also gives important insights on the relationship between the reformulated convex optimization problem and the original max-min-max problem in Eq. (10). From the geometrical perspective, we prove that the relaxation and smoothing in the tractable reformulation step are asymptotically tight in the following Theorem

Theorem 2 (Asymptotic Equivalence)**.** *With sufficiently large smoothing parameter ϵ, the solution $\boldsymbol{\tau}^*_{approx}$ of the convex optimization problem in Eq.* (14) *is asymptotically optimal for the original max-min-max problem in Eq.* (10)*, where the approximation error satisfies*

$$\Phi(\boldsymbol{\tau}^*) - \Phi(\boldsymbol{\tau}^*_{approx}) = o\left(\frac{1}{\epsilon}\right), \tag{22}$$

where $\boldsymbol{\tau}^*$ represents the optimal solution solving the original max-min-max problem in Eq. (10)*.*

Proof: Since each relay node can use at most one channel to transmit, and the line of Eq. (20) in a two-dimensional space can be found with probability 0 to go through 3 points, we prove that the relaxation of $\boldsymbol{\tau}$ in Eq. (11) is tight through reasoning by contradiction. Then, the approximation error is analyzed to prove the asymptotically equivalent property. More details of the proof are provided in Appendix B. ■

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2772268, IEEE Transactions on Communications
The final version of record is available at https://dx.doi.org/10.1109/TCOMM.2017.2772268

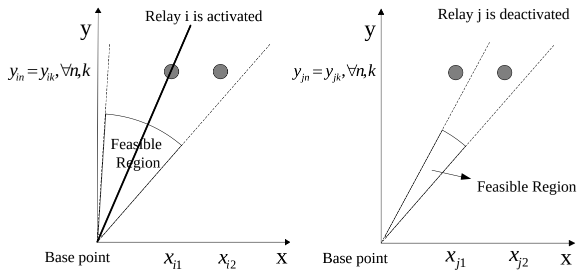


Fig. 2. Relationship between lines and points

D. Proposed Algorithm

Solving the geometrical problem in Eq. (20) is equivalent to finding the slope of a line $Y_i = A_i X_i$ for each relay node r_i, such that some of the points $S_{in} = (x_{in}, y_{in})$ on the line and all the other points $S_{ik} = (x_{ik}, y_{ik}), k \neq n$ under the line, because the choice of $\boldsymbol{\tau}$ only poses influence on the slope of each line. Accordingly, the following rule can be determined

$$\begin{aligned} \tau_{ik} &= 1, \exists A_i = \frac{y_{ik}}{x_{ik}} \\ \tau_{ik} &= 0, \forall A_i > \frac{y_{ik}}{x_{ik}}, \end{aligned} \tag{23}$$

where $k = \arg\min_j x_{ij}$. To determine the slope of the line $Y_i = A_i X_i$, we only need to compare A_i with the slope of a line, containing the leftmost point (x_{ik}, y_{ik}) and the base point $(0, 0)$, for each relay node r_i. The feasible region of the slope A_i is determined according to (20) by varying τ_{ik} from 0 to 1. As the example in Fig. 2, there exists a slope A_i that the line passes through the left-most point (x_{i1}, y_{i1}). If so, relay node r_i is activated to transmit with channel 1. While there does not exist a slope A_j that the line passes through the left-most point (x_{j1}, y_{j1}). As a result, relay node r_j is deactivated.

Towards the lexicographic optimality, we propose a relay selection scheme μ^* based on the initial relay selection scheme $\hat{\mu}$ in Lemma 1 and the channel allocation $\boldsymbol{\tau}$. Even though $\hat{\mu}$ achieves a max-min data rate of the system, it is not lexicographically optimal. A destination node may actually receives a higher data rate from another relay node than that from the relay node which holds the largest channel capacity between them, because the received data rate is the multicast rate which depends on the channel capacities of all the destination nodes in the multicast group.

We propose a lexicographically optimal relay selection scheme μ^* to improve the performance. We first divide the destination nodes into two categories, including the node with the weakest channel quality in each multicast group and other destination nodes. Denote the set of the weakest nodes in all multicast groups as $\mathcal{J}$, i.e.,

$$\mathcal{J} = \{d(r_i) | d(r_i) = \arg\min_{d_j \in \mathcal{D}_i} \{C_{ij}\}, \forall r_i \in \mathcal{R}\}, \tag{24}$$

where $\mathcal{D}_i = \{d_j | \hat{\mu}(d_j) = r_i, \forall d_j \in \mathcal{D}\}$ and $d(r_i)$ is the destination node which has the weakest channel quality in the multicast group of the relay node r_i.

For the destination node $d_j \in \mathcal{J}$,

$$\mu^*(d_j) = \hat{\mu}(d_j) = \arg\max_{r_i \in \mathcal{R}} \{C_{ij}\}. \tag{25}$$

For the destination node $d_j \in \mathcal{D}/\mathcal{J}$,

$$\mu^*(d_j) = \arg\max_{r_i \in \mathcal{R}} \{\min\{C_{ij}, R_i\}\}, \tag{26}$$

where R_i is the multicast rate of the relay node r_i with the initial relay selection scheme $\hat{\mu}$ and the channel allocation $\boldsymbol{\tau}$ obtained in the first step.

Remark 2 (Interpretation of the Relay Selection Scheme)**.** *The destination nodes in $\mathcal{J}$ have the worst channel conditions of their multicast group, and hence limit the multicast rate of each relay node. Their relay selection schemes in μ^* are the same as those in $\hat{\mu}$. Each of the other destination nodes selects the relay which provides the maximal possible data rate for this destination node.*

Theorem 3 (Lexicographic Optimality)**.** *The proposed relay selection scheme μ^* is lexicographically optimal.*

Proof: Reasoning by contradiction, suppose there exists a relay selection scheme μ' that is lexicographically greater than μ^*. Despite the cases whether the weakest links in each multicast group are changed or not, we can prove that μ^* is not lexicographically less than μ'. Therefore, the proposed relay selection scheme μ^* achieves the lexicographic optimality. More details of the proof are provided in Appendix C. ■

Algorithm 1 Lexicographic Max-Min Relay Selection

1: **loop**
2: Each destination node selects a relay node according to Eq. (8).
3: Adjust λ and β by augmented Lagrange method.
4: **for** each user i and channel k **do**
5: **if** $\exists A_i = y_{ik}/x_{ik}$, where $k = \arg\min_n x_{in}$, and $\tau_{lk} = 0, \forall (r_i, r_l) \in E$. **then**
6: Set $\tau_{ik} = 1$.
7: **end if**
8: **end for**
9: **end loop** until the approximation gap is within a given threshold by updating the approximation parameter ϵ.
10: Each destination node selects a relay node according to Eq. (25) and Eq. (26).
11: The multicast rate of each relay node is determined according to Eq. (4).

We provide the pseudocodes of the proposed algorithm in Algorithm 1, which is launched at the beginning of each time slot. In the pseudo-codes, Line 2 presents the max-min relay selection scheme, Lines 3-9 determine the channel allocation, Line 10 deploys the lexicographically optimal relay selection scheme, and Line 11 determines the scheduled transmission rate for each activated relay nodes. Any strictly increasing updating rule of ϵ can be used [34].

Remark 3 (Computational Complexity)**.** *The complexity is mainly brought by channel allocation whose complexity is*

$O(K^2M)$. Therefore, the complexity of the proposed scheme is $O(cK^2M)$, where c represents the iteration rounds.

Remark 4 (Compatibility with Dynamic Scenarios)**.** *In dynamic scenarios, the communication overhead could be large if we execute the proposed algorithm at the beginning of each time slot. However, if the resource allocation is updated less frequently, the performance will slightly deviate from the optimality. Therefore, there is an inherent trade-off between the communication overhead and the performance.*

In slow fading channel, the channel quality can be considered to be unchanged during a whole signal block, which is commonly considered in cooperative networks [35]. In such a case, we execute the proposed algorithm for lexicographically optimal resource allocation at the beginning of each signal block instead of each time slot, and the performance will not be degraded.

Moreover, in dynamic scenarios, we may not able to estimate the channel quality accurately due to the delay of estimation process. To address this issue, we budget some allowance to enlarge the average decoding success rate. The less number of the enhancement layers is transmitted, the larger average decoding success rate we have.

E. Discussions on Further Improving the Performance

In this section, we discuss two possible approaches to further improve the performance, i.e., optimizing the time intervals of the two cooperative transmission stages and incorporating MRC by allocating channels to the deactivated relay nodes.

1). **Optimizing the time intervals of the two stages:** In this paper, we adopt a two-stage cooperative relay in multicast transmission paradigm, where each time slot for multicast services T is divided into two time intervals T_1 and T_2. To determine the lexicographically optimal rate vector, we need to find $\mathbf{R}$ for all possible T_1 and T_2 satisfying $T_1 + T_2 = T$. We demonstrate that with the proposed algorithm, we can easily find the optimal T_1 and T_2. For given T_1 and T_2, we can calculate the associated lexicographically optimal rate vector using the proposed algorithm. Then adopting exhaustive search, we can find the optimal T_1 and T_2 as well as the associated lexicographically optimal rate vector.

2). **Incorporating MRC:** Consider an example that we have allocated the channel c_k to the relay node r_i and its neighbor relay node r_j is deactivated due to the limitation of the number of channels. Observing that the neighboring relay nodes r_i and r_j share a part of the destination nodes in their multicast groups, who are able to receive the signal from both the relay nodes, it is possible for those destination nodes to receive a larger number of enhancement layers by incorporating MRC. To this end, we allocate the channel c_k to the relay node r_j who transmits the same number of the enhancement layers as the relay node r_i, where the exact number of the enhancement layers can be determined through further coordination between the two relay nodes. In such a case, it is equivalent to merging the two relay nodes r_i, r_j into a new mega-node $r_{(i,j)}$, where $\forall r_l \in \mathcal{R}$, any $(r_l, r_i) \in E$ or $(r_l, r_j) \in E$, it is satisfied that $(r_l, r_{(i,j)}) \in E$. By doing so, we form a new conflict graph, where the proposed algorithm can be applied to obtain the lexicographically optimal resource allocation. However, the problem of finding the optimal MRC combination is NP-hard, because the MRC weights vary according to different combinations [40]. It is an interesting problem to derive a low-complexity algorithm which we are willing to investigate in our future work.

V. SIMULATION

In this section, we evaluate the performance of the proposed algorithm through simulation. In the simulation, the source node is located at the center, the destination nodes are randomly distributed in a circular area with a radius of 500 meters and the relay nodes are randomly distributed in a circular area with a radius of 100 meters. Following the simulation parameter settings in [11], we set the bandwidth as 22 MHz for all channels. The transmission power is the same for each node, i.e., $P_s = P_{r_i} = 1$ Watt for the source node s and the relay nodes $r_i \in \mathcal{R}$. For the transmission model, we assume that the path loss exponent $\alpha = 4$ and the abient noise is 10^{-10}.

For performance comparison with the proposed lexicographic max-min scheme μ^*, we adopt three baseline schemes as follows:

- **Random**: The relay nodes are activated randomly and the destination nodes choose the relay node with the best channel quality.
- **Throughput-based scheme** [41]: The channels are allocated to the multicast relay nodes to maximize total data rate.
- **Max-Min**: The max-min scheme $\hat{\mu}$ proposed in Lemma 1.

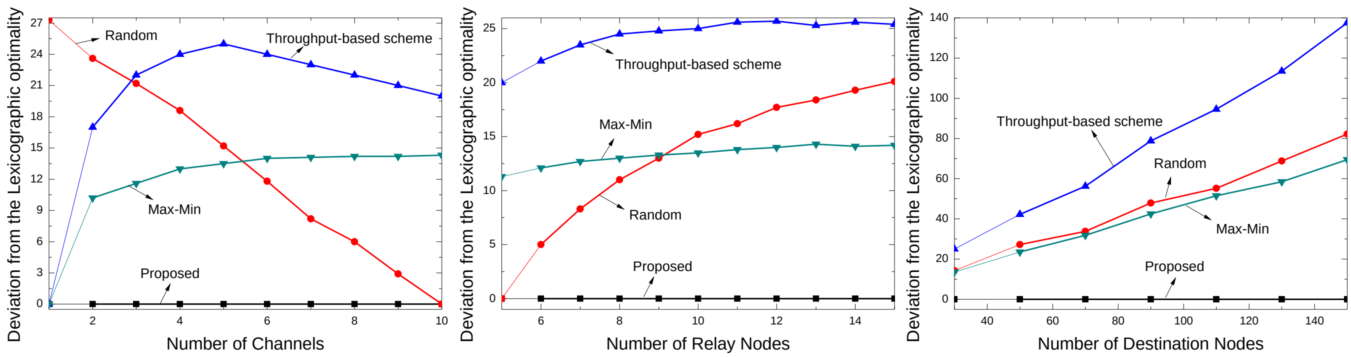
For performance comparison, the deviation from the lexicographic optimality is defined as $N - i$, where i is given such that the prefix $(\nu_1^*, \nu_2^*, ...\nu_i^*)$ of the lexicographically optimal rate vector $\mathbf{R}^*$ and the prefix $(\nu_1', \nu_2', ...\nu_i')$ of $\mathbf{R}'$ under a baseline scheme satisfy that $\nu_i^* > \nu_i'$, and $\nu_j^* = \nu_j'$ for $1 \le j \le i-1$.

We first analyze the deviation from the lexicographic optimality of the four schemes, and then further discuss the data rates achieved by all destination nodes. We evaluate the performance with different number of available channels, different number of the relay nodes and different number of the destination nodes. For each case, 10 instances are generated for obtaining the average performance.

Fig. 3(a) demonstrates the deviation from the lexicographic optimality with different numbers of available channels, where 10 relay nodes and 30 destination nodes are deployed. Even though the max-min scheme achieves the same minimal capacity as the proposed scheme, the deviation from the lexicographic optimality is quite large, which verifies that the solution of the lexicographic max-min optimization is a refinement of the standard max-min concept. The performance of the proposed scheme is better than those of the other two baselines. The throughput-based scheme deviates from the lexicographic optimality the most, because it sacrifices the performance of the worst user.

TABLE I
SORTED RATE VECTOR $\mathbf{R}$

Proposed	12.7496	12.7496	12.7496	14.7067	14.7067	15.8894	15.8894	15.8894	17.2415	17.2415
Random	12.2972	12.2972	12.2972	14.2155	14.2155	14.2155	15.0306	15.0306	18.2415	18.2415
EPSA	11.5488	11.5488	11.5488	13.4623	18.0838	18.0838	18.0838	20.2415	20.2415	20.2415
Max-Min	12.7496	12.7496	12.7496	12.7496	14.7067	14.7067	14.7067	15.8894	15.8894	17.2415



(a) Varying the number of channels (b) Varying the number of relay nodes (c) Varying the number of destination nodes

Fig. 3. The deviation from the lexicographic optimality

Fig. 3(b) provides the deviation from the lexicographic optimality with different numbers of relay nodes when there are 5 channels and 30 destination nodes. The proposed scheme outperforms the three baseline schemes. The number of the relay nodes has small influence to the deviation from the lexicographic optimality, because the number of channels is small and thus limits the choice of relay selection.

Fig. 3(c) provides the deviation from the lexicographic optimality with different numbers of destination nodes when there are 10 relay nodes and 5 channels. As the number of the destination nodes is large, the performance gain of the proposed algorithm is significant.

Besides the deviation from the lexicographic optimality, we further analyze the data rates of all destination nodes for the lexicographic optimality. Table I provides the rate vector $\mathbf{R}$ for all four schemes when there are 10 destination nodes and 10 relay nodes with 3 available channels. It can be found that the rate vector of the proposed scheme is lexicographically greater than those of three baseline schemes. It is worth mentioning that the proposed scheme lexicographically dominates the max-min one, even though the minimal capacities of the two schemes are the same. As a result, we can safely draw the conclusion that the proposed scheme achieves superior max-min fairness than the three baseline schemes and provides relatively homogeneous service quality to all the destination nodes in the cooperative multicast system.

VI. CONCLUSION

In this paper, we construct an analytical framework for lexicographic max-min multicast relay selection for cooperative multicast with a limited number of channels. Specifically, we design the algorithm in two steps. 1) We consider the maximization of the minimal data rate. By decoupling relay selection and channel allocation, the problem is transformed to a max-min-max problem, which is difficult to solve. To make this problem tractable, we reformulate it via relaxation and smoothing, and prove the asymptotic equivalence from a geometrical perspective. 2) We propose an adjustment algorithm based on the initial max-min solution, and prove that the proposed scheme achieves lexicographic optimality. The simulation results show that the proposed scheme achieves the data rate lexicographically greater than those of the conventional schemes.

APPENDIX A
PROOF OF THEOREM 1

Denote the objective function in Eq. (14) as $f(\boldsymbol{\tau})$. To analyze the convexity of $f(\boldsymbol{\tau})$, we take the first-order derivative of $f(\boldsymbol{\tau})$ with respect to τ_{ln} as

$$\frac{\partial f(\boldsymbol{\tau})}{\partial \tau_{ln}} = -\frac{1}{\sum_{j=1}^{N} e^{\epsilon \sum_{k=1}^{K} \tau_{lk} C_{lj}}} \sum_{j=1}^{N} \epsilon C_{lj} < 0, \tag{27}$$

which is a strictly decreasing function of $\boldsymbol{\tau}$.

Then we consider the second-order derivative of $f(\boldsymbol{\tau})$

$$\begin{cases} \dfrac{\partial^2 f(\boldsymbol{\tau})}{\partial^2 \tau_{ln}} = \left(\dfrac{\sum_{j=1}^{N} \epsilon C_{lj}}{\sum_{j=1}^{N} e^{\epsilon (\sum_{k=1}^{K} \tau_{lk} C_{lj})}} \right)^2 > 0 \\ \dfrac{\partial^2 f(\boldsymbol{\tau})}{\partial \tau_{ln} \partial \tau_{mp}} = \dfrac{\partial^2 f(\boldsymbol{\tau})}{\partial \tau_{ln} \partial \tau_{mn}} = 0 \\ \dfrac{\partial^2 f(\boldsymbol{\tau})}{\partial \tau_{ln} \partial \tau_{lp}} = \left(\dfrac{\sum_{j=1}^{N} \epsilon C_{lj}}{\sum_{j=1}^{N} e^{\epsilon (\sum_{k=1}^{K} \tau_{lk} C_{lj})}} \right)^2 = \dfrac{\partial^2 f(\boldsymbol{\tau})}{\partial^2 \tau_{ln}}. \end{cases} \tag{28}$$

We derive the Hessian matrix of $f(\boldsymbol{\tau})$ as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2772268, IEEE Transactions on Communications
The final version of record is available at https://dx.doi.org/10.1109/TCOMM.2017.2772268

$$\begin{pmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & A_M \end{pmatrix}, \tag{29}$$

where

$$A_l = \begin{pmatrix} a_l & \dots & a_l \\ \vdots & \ddots & \vdots \\ a_l & \dots & a_l \end{pmatrix}, \tag{30}$$

and $a_l = \left(\frac{1}{\sum_{j=1}^{N} e^{\epsilon(\sum_{k=1}^{K} \tau_{lk} C_{lj})}}\right)^2$. Note that the Hessian of $f(\boldsymbol{\tau})$ is a symmetric matrix. Each sub-block A_l is also a symmetric matrix. It can be derived according to [37] that

$$\mathbf{E} = (N, 0, \cdots, 0), \tag{31}$$

where $\mathbf{E}$ is the eigenvector of the matrix in Eq. (30).

Therefore, the eigenvalue of the Hessian of $f(\boldsymbol{\tau})$ is non-negative. According to [36], [37], the Hessian is positive semi-definite, and $f(\boldsymbol{\tau})$ is convex with respect to $\boldsymbol{\tau}$. Since all the constraints in Eq. (14) are linear, we prove that the problem in Eq. (14) is a convex optimization problem.

Appendix B
Proof of Theorem 2

For the optimality, each relay node can use at most one channel to transmit, which makes the relaxation tight.

Reasoning by contradiction, suppose there exists more than one channels used by a relay node. Assume there are two channels c_n and c_m used by the relay node r_i, i.e., $\tau_{in} > 0$ and $\tau_{im} > 0$. Base on Eq. (11), we have $1 \geq \tau_{in} \geq 0, 1 \geq \tau_{im} \geq 0$. For the optimality, the following equations need to be satisfied according to the optimality conditions,

$$\begin{aligned} A_i x_{in} - y_{in} &= 0 \\ A_i x_{im} - y_{im} &= 0. \end{aligned} \tag{32}$$

Thus, it follows from Eq. (32) that

$$\frac{x_{in}}{y_{in}} = A_i = \frac{x_{im}}{y_{im}}. \tag{33}$$

According to Eq. (20),

$$y_{in} = y_{im}. \tag{34}$$

Thus, it comes from Eq. (33) that

$$x_{in} = x_{im}. \tag{35}$$

Observing the intercept of the line in Eq. (20) is zero, i.e., the line must go through the base point. To allocate the channels to the relay nodes efficiently, we assume that a relay node always wants to use a channel with smaller index as in [39], such that the lagrange multipliers are adjusted differently for different channels. Channels c_n and c_m are different, and a line in a two-dimensional space can be found with probability 0 to go through all these 3 points, which is a contradiction. Therefore, the relaxation of $\boldsymbol{\tau}$ in Eq. (11) is tight.

According to [34], the approximation error satisfies

$$\frac{2\ln K}{\epsilon} \leq \Phi_\epsilon(\boldsymbol{\tau}) - \Phi(\boldsymbol{\tau}) \leq \frac{\ln(MK)}{\epsilon}. \tag{36}$$

With sufficiently large smoothing parameter ϵ, the value of the approximating function converges to a stationary point with the approximation error $\Phi_\epsilon(\boldsymbol{\tau}) - \Phi(\boldsymbol{\tau}) = o\big(1/\epsilon\big)$, such that the smoothing in Eq. (12) is tight. Then, we analyse the approximation error. According to Eq. (36), we have

$$\frac{2\ln K}{\epsilon} \leq \Phi_\epsilon(\boldsymbol{\tau}^*_{\text{approx}}) - \Phi(\boldsymbol{\tau}^*_{\text{approx}}) \leq \frac{\ln(MK)}{\epsilon}, \tag{37}$$

$$\frac{2\ln K}{\epsilon} \leq \Phi_\epsilon(\boldsymbol{\tau}^*) - \Phi(\boldsymbol{\tau}^*) \leq \frac{\ln(MK)}{\epsilon}. \tag{38}$$

Let Eq. (37) minus Eq. (38),

$$\begin{aligned} \frac{\ln K - \ln M}{\epsilon} &\leq \Phi(\boldsymbol{\tau}^*) - \Phi(\boldsymbol{\tau}^*_{\text{approx}}) \\ &+ \Phi_\epsilon(\boldsymbol{\tau}^*_{\text{approx}}) - \Phi_\epsilon(\boldsymbol{\tau}^*) \leq \frac{\ln M - \ln K}{\epsilon}. \end{aligned} \tag{39}$$

Observing that $\Phi_\epsilon(\boldsymbol{\tau}^*_{\text{approx}}) - \Phi_\epsilon(\boldsymbol{\tau}^*) \geq 0$ and $\Phi(\boldsymbol{\tau}^*) - \Phi(\boldsymbol{\tau}^*_{\text{approx}}) \geq 0$, we have

$$\Phi(\boldsymbol{\tau}^*) - \Phi(\boldsymbol{\tau}^*_{\text{approx}}) \leq \frac{\ln M - \ln K}{\epsilon} = o\big(\frac{1}{\epsilon}\big). \tag{40}$$

Since the relaxation of $\boldsymbol{\tau}(t)$ in Eq. (11) and the smoothing in Eq. (12) are tight, the optimization problem in Eq. (12) and the one in Eq. (10) are asymptotically equivalent with the approximation error $o\big(1/\epsilon\big)$.

Appendix C
Proof of Theorem 3

Reasoning by contradiction, suppose there exists a relay selection scheme μ' that is lexicographically greater than the optimal μ^* at the position k, i.e., $\nu'_k > \nu^*_k$ and $\nu'_i = \nu^*_i, \forall i < k$.

To prove this theorem, we respectively consider the cases whether the weakest links in each multicast group are the same under μ' and μ^*, i.e., $\mu'(d_j) = \mu^*(d_j), \forall d_j \in \mathcal{J}$.

1) **At least one of the weakest links in each multicast group is not the same.** First, let us consider the case where one of the weakest links is not the same under μ' and μ^*. Suppose that one destination node $d_j \in \mathcal{J}$ belongs to the multicast group of r_l under μ^*, and belongs to the multicast group of r_n under μ'. The received data rates $a^*_j = R^*_l$ under μ^* and $a'_j = R'_n$ under μ' should satisfy $a^*_j > a'_j$ due to the optimality of μ^*.

Since the weakest link is removed from the multicast group of r_l under μ^*, the multicast rate of r_l under μ' is not less than that under μ^*, i.e., $R'_l \geq R^*_l$. As for any destination node $d_q \in \mathcal{D}$, the received data rates under μ' and μ^* have the following relationship:

$$\begin{aligned} &a'_q \geq a^*_q, \text{if } R^*_l \leq a^*_q < R'_l, \\ &a'_q \leq a^*_q, \text{if } a^*_q = R^*_n, \\ &a'_q = a^*_q, \text{if } a^*_q < R^*_n, \text{or } R^*_n < a^*_q < R^*_l, \text{or } a^*_q \geq R'_l. \end{aligned} \tag{41}$$

For the first equation, the destination nodes in the multicast group of r_l under μ^* (i.e., $a^*_q = R^*_l$) receive higher data rates under μ' than those under μ^*, and the destination nodes whose data rates satisfy $R^*_l < a^*_q < R'_l$ under μ^* may reselect r_l as their relay node under μ' for achieving higher data rates. In this case, their position numbers in the rate vector sorted in non-descending order under μ' are larger than those of the destination nodes in the multicast group of r_l under μ^*. For the second equation, the destination nodes in the multicast group of r_n under μ^* (i.e., $a^*_q = R^*_n$) receive lower data rates under μ' than those under μ^* due to the reduction of the data rate of the multicast group of r_n. Note that some destination nodes in the multicast group of r_n under μ^* may reselect their relay nodes under μ', but their data rates under μ' are still lower than R^*_n due to the optimality of μ^*. In this case, their position numbers in the rate vector sorted in non-descending order under μ' are smaller than those of the destination nodes in the multicast group of r_l under μ^*. Apart from the above two cases, the received data rates of the destination nodes under μ' and μ^* are the same, which is described in the third equation.

According to the definition of lexicographic optimality in Definition 1, we focus on the first position (i.e., the smallest position number) that results in different data rates of the rate vectors sorted in non-descending order under μ' and μ^*, which is caused by the difference between $\mu'(d_j)$ and $\mu^*(d_j)$. Next, we discuss the difference between μ' and μ^* from the perspective of the position numbers of the destination nodes in the multicast group of $r_n = \mu'(d_j)$ under μ' and μ^* in the rate vectors.

a). The position numbers of the destination nodes in the multicast group of r_n under μ' are the same as those under μ^*, which is achieved if the multicast rate of r_n under μ' and μ^* are the same despite of d_j, or the reduction of the multicast rate of r_n under μ' is small enough. To explicitly describe such circumstances, we consider two sub-cases as follows:

a-1). The multicast rate of r_n under μ' is the same as that under μ^*. Denote the largest position number of the destination nodes in the multicast group of r_n under μ^* as p, i.e., $\nu^*_p \le \nu^*_{p+1}$. Here, the largest position number of the destination nodes in the multicast group of r_n under μ' becomes $p+1$ due to the joining of d_j, i.e., $\nu'_p = \nu'_{p+1} \le \nu'_{p+2}$. Since the multicast rate of r_n under μ' is the same as that under μ^*, we have $\nu^*_p = \nu'_p$. Since μ' is assumed to be lexicographically greater than μ^* at the position k, and the first position that results in different data rates under μ' and μ^* is $p+1$, we have $k = p+1$. Therefore, we have $\nu'_k = \nu'_{p+1} = \nu'_p = \nu^*_p \le \nu^*_{p+1} = \nu^*_k$, hence μ^* is lexicographically greater than or equal to μ', which contradicts the premise that μ' is lexicographically greater than μ^*.

a-2). The multicast rate of r_n under μ' is smaller than that under μ^*. Denote the smallest position number of the destination nodes in the multicast group of r_n under μ^* as g, i.e., $\nu^*_g \ge \nu^*_{g-1}$. Here, the multicast rate of r_n under μ' is smaller than that under μ^*, i.e., $\nu'_g < \nu^*_g$. Since the position numbers of the destination nodes in the multicast group of r_n under μ' are the same as those under μ^*, we have $\nu^*_{g-1} \le a'_j = \nu'_g < \nu^*_g$. Since the first position that results in different data rates under μ' and μ^* is $k = g$, and it is satisfied that $\nu'_k = \nu'_g < \nu^*_g = \nu^*_k$, hence μ^* is lexicographically greater than μ', which contradicts the premise that μ' is lexicographically greater than μ^*.

b). The position numbers of the destination nodes in the multicast group of r_n under μ' are smaller than those under μ^*, hence it is satisfied that $a'_j < \nu^*_{g-1} \le \nu^*_g$. According to (42), for all destination nodes d_q whose data rate under μ^* satisfies $a^*_q < R^*_n = \nu^*_g$, their data rates are the same under μ' and μ^*, i.e., $a'_q = a^*_q$. Since the rate vector under μ' is sorted in non-decreasing order, the smallest position number of the destination nodes in the multicast group of r_n under μ' is determined as h that satisfies $\nu'_h > \nu^*_{h-1}$ and $\nu'_h \le \nu^*_h$. In this context, the first position that results in different data rates under μ' and μ^* is $k = h$. Therefore, we have $\nu'_k = \nu'_h \le \nu^*_h = \nu^*_k$, hence μ^* is lexicographically greater than μ', which contradicts the premise that μ' is lexicographically greater than μ^*.

Then, let us consider the case where multiple weakest links are not the same under μ' and μ^*. Denote the destination nodes in $\mathcal{J}$ under μ^* that belong to different multicast groups under μ' and μ^* as $\mathcal{W}$, and denote their selected relay nodes under μ' as $\mathcal{Q}$ and that under μ^* as $\mathcal{T}$. Due to the optimality of μ^*, it is satisfied that $a'_j < a^*_j, \forall d_j \in \mathcal{W}$. Since the weakest links are removed from the multicast group of $r_i \in \mathcal{T}$ under μ^*, their multicast rates under μ' in not less than those under μ^*, i.e., $R'_i > R^*_i, \forall r_i \in \mathcal{T}$. As for any destination node $d_q \in \mathcal{D}$, the received data rate under μ' and μ^* have the following relationship:

$$
\begin{aligned}
&a'_q \ge a^*_q, \text{if } \min_{r_i\in\mathcal{T}}\{R^*_i\} \le a^*_q < \max_{r_i\in\mathcal{T}}\{R'_i\},\\
&a'_q \le a^*_q, \text{if } a^*_q = R^*_i, \exists r_i \in \mathcal{Q},\\
&a'_q = a^*_q, \text{if } a^*_q < \min_{r_i\in\mathcal{Q}}\{R^*_i\} \text{ or } \min_{r_i\in\mathcal{Q}}\{R^*_i\} < a^*_q < \min_{r_i\in\mathcal{T}}\{R^*_i\}\\
&\qquad\qquad \& a^*_q \ne R^*_i, \forall r_i \in \mathcal{Q}, \text{or } a^*_q \ge \max_{r_i\in\mathcal{T}}\{R'_i\}.
\end{aligned}
$$

According to the definition of lexicographic optimality in Definition 1, we focus on the first position that results in different data rates of the rate vectors sorted in non-descending order under μ' and μ^*, which are caused by the difference between $\mu'(d_j)$ and $\mu^*(d_j)$, where $d_j = \arg\min_{d_i\in\mathcal{W}}\{a'_i\}$. Here, we only have to consider the influence brought by d_j, while the influence brought by the rest of the destination nodes in $\mathcal{W}$ will not affect the result of the analysis between μ' and μ^* in view of lexicographic optimality, because any destination node $d_z \in \mathcal{W}/\{d_j\}$ has a higher data rate under μ' than that of d_j under μ', i.e., $a'_z \ge a'_j$, and thus their position numbers under μ' will be larger than that of d_j under μ', hence they do not account for the first position that results in the different data rates of the rate vectors under μ' and μ^*. In this context, the case where multiple weakest links are not the same is reduced to the case where one weakest link is not the same. Following the same analysis when discussing the difference between μ' and μ^* from the perspective of the position numbers of the destination nodes in the multicast group of $\mu'(d_j)$ under μ' and μ^* in the rate vectors, it is proved that under all circumstances, we have $\nu^*_k \ge \nu'_k$, and thus μ^* is lexicographically greater than μ', which contradicts the premise that μ' is lexicographically greater than μ^*.

2) **The weakest links in each multicast group are the same under μ' and μ^*.** If $\forall d_j \in \mathcal{J}$ it is satisfied that $\mu'(d_j) = \mu^*(d_j)$, then we have $\forall d_j \in \mathcal{D}$ that $a'_j \le a^*_j$. Because the proposed relay selection scheme μ^* selects the relay node which provides the maximal possible data rate for the rest of

the destination nodes. Thus, when $\mu'(d_j) = \mu^*(d_j), \forall d_j \in \mathcal{J}$, it is not possible that $\exists d_j \in \mathcal{D}/\mathcal{J}$ such that $a'_j > a^*_j$.

Therefore, there does not exist a relay selection scheme μ' that is lexicographically greater than μ^*, which means that the proposed relay selection scheme μ^* achieves the lexicographic optimality.

References

[1] Y. Wang, W. Wang, L. Chen, P. Zhou, and Z. Zhang, "Lexicographic relay selection and channel allocation for multichannel cooperative multicast," *Proc. of IEEE WCNC 2017*, Mar. 2017.

[2] L. Militano, D. Niyato, M. Condoluci, G. Araniti, A. Iera, and G. M. Bisci, "Radio resource management for group-oriented services in LTE-A," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3725-3739, Sept. 2015.

[3] B. Hu, H. V. Zhao, and H. Jiang, "Wireless multicast using relays: Incentive mechanism and analysis," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2204-2219, Dec. 2012.

[4] L. Lv, J. Chen, Q. Ni, and Z. Ding, "Design of cooperative non-orthogonal multicast cognitive multiple access for 5G systems: User scheduling and performance analysis," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2641-2656, Mar. 2017.

[5] P. Li, H. Zhang, and B. Zhao, "Scalable video multicast with adaptive modulation and coding in broadband wireless data systems," *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 57-68, Aug. 2012.

[6] X. Shi, M. Medard, and D. E. Lucani, "Whether and where to code in the wireless packet erasure relay channel," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1379-1389, Aug. 2013.

[7] H. Khamfroush, P. Pahlevani, D. E. Lucani, M. Hundeboll, and F. H. Fitzek, "On the coded packet relay network in the presence of neighbors: Benefits of speaking in a crowded room," *Proc. of IEEE ICC 2014*, pp. 1928-1933, Jun. 2014.

[8] Y. Shi, S. Sharma, and Y. T. Hou, "Optimal relay assignment for cooperative communications," *Proc. of ACM ISMNC 2008*, pp. 3-12, Feb. 2008.

[9] Y. Lu, W. Wang, L. Chen, Z. Zhang, and A. Huang, "Distance-based energy-efficient opportunistic broadcast forwarding in mobile delay tolerant networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5512 - 5524, Jul. 2016.

[10] R. Senanayake, P. J. Smith, P. A. Martin, and J. S. Evans, "Performance analysis of reconfigurable antenna arrays," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2726-2739, Mar. 2017.

[11] D. Yang, X. Fang, and G. Xue, "OPRA: Optimal relay assignment for capacity maximization in cooperative networks," *Proc. of IEEE ICC 2011*, Jun. 2011.

[12] I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1080-1088, Aug. 2004.

[13] B. Mergen and A. Scaglione, "On the power efficiency of cooperative broadcast in dense wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 497-507, Feb. 2007.

[14] P. Kyasanur and N. H. Vaidya, "Capacity of multi-channel wireless networks: impact of number of channels and interfaces," *Proc. of IEEE ICMCN 2005*, pp. 43-57, Aug. 2005.

[15] A. Homayounzadeh and M. Mahdavi, "Performance analysis of cooperative cognitive radio networks with imperfect sensing," *Proc. of IEEE ICCSPA 2015*, Feb. 2015.

[16] S. Narayanan, M. Di Renzo, F. Graziosi, and H. Haas, "Distributed spatial modulation: A cooperative diversity protocol for half-duplex relay-aided wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 2947-2964, May 2016.

[17] A. Attarkashani and W. Hamouda, "Joint power allocation and subcarrier-relay assignment for OFDM-based decode-and-forward relay networks," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2312-2315, Aug. 2016.

[18] S. Yang and J. A. McCann, "Distributed optimal lexicographic max-min rate allocation in solar-powered wireless sensor networks," *ACM Trans. Sensor Netw.*, vol. 11, no. 1, pp. 1-35, Nov. 2014.

[19] Y. S. Wu, C. H. Yeh, W. Y. Tseng, and M. K. Chang, "Scalable video streaming transmission over cooperative communication networks based on frame significance analysis," *Proc. of IEEE ICSPCC 2012*, pp. 274-279, Aug. 2012.

[20] T. V. Nguyen, P. C. Cosman, and L. B. Milstein, "Double-layer video transmission over decode-and-forward wireless relay networks using hierarchical modulation," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1791-1804, Apr. 2014.

[21] D. Hwang, P. Chau, J. Shin, and T. J. Lee, "Two cooperative multicast schemes of scalable video in relay-based cellular networks," *IET Commun.*, vol. 9, no. 7, pp. 982-989. May 2015.

[22] Y. Zhou, H. Liu, Z. Pan, L. Tian, and J. Shi, "Spectral-and energy-efficient two-stage cooperative multicast for LTE-advanced and beyond," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 34-41, Apr. 2014.

[23] Y. Zhou, H. Liu, Z. Pan, L. Tian, J. Shi, and G. Yang, "Two-stage cooperative multicast transmission with optimized power consumption and guaranteed coverage," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 274-284, Feb. 2014.

[24] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Netw.*, vol. 14, pp. 2425-2432, Oct. 2006.

[25] O. Alay, T. Korakis, and Y. Wang, "Layered wireless video multicast using omni-directional relays," *Proc. of IEEE APSP 2008*, pp.2149-2152, Mar. 2008.

[26] K. Xiong, P. Fan, C. Zhang, and K. B. Letaief, "Wireless information and energy transfer for two-hop non-regenerative MIMO-OFDM relay networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1595-1611, Aug. 2015.

[27] B. Niu, J. Hai, and H. Zhao, "A cooperative multicast strategy in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 3136-3143, Jul. 2010.

[28] D. Jiang, Z. Xu, and Z. Lv, "A multicast delivery approach with minimum energy consumption for wireless multi-hop networks," *Telecommun. Syst.*, vol. 62, no. 4, pp. 771-782, Jul. 2016.

[29] S. Ren and M. V. Schaar, "Pricing and distributed power control in wireless relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2913-2926, Jun. 2011.

[30] B. Medepally and N. B. Mehta, "Voluntary energy harvesting relays and selection in cooperative wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3543-3553, Nov. 2010.

[31] F. E. Lapiccirella, X. Liu and Z. Ding, "Distributed control of multiple cognitive radio overlay for primary queue stability," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 112-122, Jul. 2013.

[32] D. Cao, M. Chen, and H. Wang, "Interval method for global solutions of a class of min-max-min problems," *Appl. Math. Comput.*, vol. 196, no. 2, pp. 594-602, Oct. 2008.

[33] E. Polak and J. O. Royset, "Algorithms for finite and semi-infinite min-max-min problems using adaptive smoothing techniques," *J. Optimiz. Theory App.*, vol. 119, no. 3, pp. 421-457, Dec. 2003.

[34] A .Tsoukalas, P. Parpas, and B.Rustem, "A smoothing algorithm for finite min-max-min problems," *Optim. Lett.*, vol. 3, no. 1, pp. 49-62, Jul. 2009.

[35] T. Li, P. Fan, and K. B. Letaief, "Outage probability of energy harvesting relay-aided cooperative networks over Rayleigh fading channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 972-978, Feb. 2016.

[36] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge university press, 2004.

[37] B. Noble and J. W. Daniel, *Applied Linear Algebra*, New Jersey: Prentice-Hall, 1988.

[38] W. Wang, W. Wang, Q. Lu, K. G. Shin, and T. Peng, "Geometry-based optimal power control of fading multiple access channels for maximum sum-rate in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 1843-1848, Jun. 2010.

[39] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. I. Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47 no. 3, pp. 1083-1102, Mar. 2001.

[40] H. Kellerer, U. Pferschy, and D. Pisinger, "Introduction to NP-Completeness of knapsack problems," *Knapsack Problems*, Springer Berlin Heidelberg, pp. 483-493, 2004.

[41] J. Liu, W. Chen, Z. Cao, and K. Letaief, "Dynamic power and subcarrier allocation for OFDMA-based wireless multicast systems," *Proc. of IEEE ICC 2008*, May 2008.

Yitu Wang (S'16) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2013. From August to November 2014, he was a visiting student with the University of Paris-Sud, Orsay, France. He is currently a Ph.D candidate at Zhejiang University, Hangzhou, China. His research interests mainly focus on stochastic optimization for cross-layer resource allocation in wireless networks and cache-enabled networks.

Wei Wang (S'08-M'10-SM'15) received the B.S. and Ph.D. degrees from the Beijing University of Posts and Telecommunications, China, in 2004 and 2009, respectively. From 2007 to 2008, he was a Visiting Student with the University of Michigan, Ann Arbor, USA. From 2013 to 2015, he was a Hong Kong Scholar with the Hong Kong University of Science and Technology, Hong Kong. He is currently an Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University, China. His research interests mainly focus on stochastic optimization for cross-layer resource allocation in wireless networks, and caching and computing in wireless networks. He is the Editor of the book entitled Cognitive Radio Systems, and serves as an Editor of the IEEE Access, Transactions on Emerging Telecommunications Technologies, and KSII Transactions on Internet and Information Systems.

Lin Chen (S'07-M'10) received his B.E. degree in Radio Engineering from Southeast University, China in 2002 and the Engineer Diploma from Telecom ParisTech, Paris in 2005. He also holds a M.S. degree of Networking from the University of Paris 6. He currently works as associate professor in the department of computer science of the University of Paris-Sud. He serves as Chair of IEEE Special Interest Group on Green and Sustainable Networking and Computing with Cognition and Cooperation, IEEE Technical Committee on Green Communications and Computing. His main research interests include modeling and control for wireless networks, distributed algorithm design and game theory.

Pan Zhou (S'07-M'14) is currently an associate professor with School of Electronic Information and Communications, Wuhan, P.R. China. He received his Ph.D. in the School of Electrical and Computer Engineering at the Georgia Institute of Technology (Georgia Tech) in 2011, Atlanta, USA. He received his B.S. degree in the Advanced Class of HUST, and a M.S. degree in the Department of Electronics and Information Engineering from HUST, Wuhan, China, in 2006 and 2008, respectively. He held honorary degree in his bachelor and merit research award of HUST in his master study. He was a senior technical member at Oracle Inc, America during 2011 to 2013, and worked on Hadoop and distributed storage system for big data analytics at Oralce could Platform. His current research interest includes: big data analytics and machine learning, security and privacy, and information networks.

Zhaoyang Zhang (M'10) received the Ph.D. degree in communication and information systems from Zhejiang University, Hangzhou, China, in 1998. He is currently a Full Professor with the College of Information Science and Electronic Engineering, Zhejiang University. He has coauthored more than 150 refereed international journal and conference papers, as well as two books in his areas of interest. His research interests are mainly focused on information theory and coding theory, signal processing techniques, and their applications in wireless communications and networking.

Dr. Zhang coreceived three conference Best Paper Awards/Best Student Paper Award. He is currently serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, IET Communications, and several other international journals. He has served as a Technical Program Committee Cochair or a Symposium Cochair for many international conferences, such as the 2013 International Conference on Wireless Communications and Signal Processing and the 2014 IEEE Global Communications Conference Wireless Communications Symposium.