# Topology management and outage optimization for multicasting over slowly fading multiple access networks<sup>☆</sup>

Avi Zanko[1],[*], Amir Leshem[1], Ephraim Zehavi[1]

**Abstract**

This paper examines the problem of rate allocation for multicasting over slow Rayleigh fading channels using network coding. In the proposed model, the network is treated as a collection of Rayleigh fading multiple access channels. In this model, rate allocation scheme that is based solely on the statistics of the channels is presented. The rate allocation scheme is aimed at minimizing the outage probability. An upper bound is presented for the probability of outage in the fading multiple access channel. A suboptimal solution based on this bound is given. A distributed primal-dual gradient algorithm is derived to solve the rate allocation problem.

*Keywords:* Network coding for multicasting, wireless networks, outage capacity, Rayleigh fading, multiple access channels

## 1. Introduction

Network coding extends the functionality of intermediate nodes from storing/forwarding packets to performing algebraic operations on received data. If network coding is permitted, the multicast capacity of a network with a single source has been shown to be equal to the minimal min-cut between the source and each of its destinations [2]. In the past decade, the concept of combining

data by network coding has been extensively extended by e.g. [3–5] and it is well known that in order to achieve the multicast rate, a linear combination over a finite field suffices if the field size is larger than the number of destinations. Moreover, centralized linear network coding can be designed in polynomial time [6]. Decentralized linear network coding can be implemented using a random code approach [7]. A comprehensive survey of network coding can be found in e.g., [8, 9].

Many network resource allocation problems can be formulated as a constrained maximization of a certain utility function. The problem of network utility maximization has been explored extensively in the past few decades [10, 11]. We briefly introduce related work on topology management and rate allocation for network coding in multicast over wireless networks. The problem of finding a minimum-cost scheme (while maintaining a certain multicast rate) in coded networks was studied by Lun et al. [12, 13]. They showed that there is no loss of optimality when the problem is decoupled into: finding the optimal coding rate allocation vector (also known as subgraph selection) and designing the code that is applied over the optimal subgraph. Moreover, in many cases, optimal subgraphs can be found in polynomial time. If in addition the cost function is also convex and separable, the solution can be found in a decentralized manner, where message passing is required solely between directly connected nodes. This decentralized solution, if coupled with random network coding (e.g. [14, 15]) provides a fully distributed scheme for multicast in coded wireline networks. This has prompted many researchers to develop different algorithms that find minimum-cost rate allocation solutions distributively; e.g. [16–19].

When addressing the problem of rate allocation for multicast with network coding in wireless networks, Lun et al., [13, 20] tackled the problem through the so-called *wireless multicast advantage* phenomenon. This phenomenon simply comes down to the fact that when interference is avoided in the network (e.g., by avoiding simultaneous transmissions), communication between any two nodes is overheard by their nearby nodes due to the broadcast nature of the wireless medium. In [20], the wireless multicast advantage was used to reduce the transmission energy of the multicast scheme (since when two nodes communicate, some of their nearby nodes get the packet for "free"). Therefore, their wireline minimum-cost optimization problem was updated accordingly [see 20, eq.(1) and (40)]. In [19] interference is allowed but is assumed to be limited. Joint optimal power control, network coding and congestion control is presented for the case of very high SINR (signal to noise plus interference ratio). This interference assumption implies that there are some limitations on simultaneous transmissions and this is taken into account in the optimization problem. In [21] the problem

of joint power control, network coding and rate allocation was studied. They showed that the throughput maximization problem can be decomposed into two parts: subgraph selection at the network layer and power control at the physical layer. A primal dual algorithm was given that converges to the optimal solution provided that the capacity region is convex with respect to the power control variables (i.e., when interference are ignored). On the other hand, to take interference into account a game theoretic method was derived to approximately characterize the capacity region.

In wireless networks, it is reasonable to assume that there is no simultaneous packet transmission or reception by any transceiver. These properties of the wireless medium introduced a new cross-layer interactions that may not exist in the wired network. Sagduyu et al. [22] analyzed and designed wireless network codes in conjunction with conflict-free transmission schedules in wireless ad hoc networks. They studied the cross-layer design possibilities of joint medium access control and network coding. It was shown that when certain objectives such as throughput or delay efficiency are considered, then network codes must be jointly designed with medium access control. The joint design of medium access control and network coding [22] was formulated as a nonlinear optimization problem. In [23] the work reported in [22] was extended and a linear formulation was derived.

However, there are certain other considerations that must be taken into account in the search for a rate allocation vector in wireless networks. The wireless medium varies over time and suffers from fading channels due to multipath or shadowing, for example. In [24] the block fading model was introduced. In this model the channel gain is assumed to be constant over each coherence time interval. Typically, fading models are classified as fast fading or slow fading. In fast fading, the coherence time of the channel is small relative to a code block length and as a consequence the channel is ergodic with a well-defined Shannon capacity (also known as the ergodic capacity [25]). In slow fading, the code block length and the coherence time of the channel are of the same order. Hence, the channel is not ergodic and the Shannon capacity is not usually a good measure of performance. The notion of outage capacity was introduced in [24] for transmitting over fading channels when the channel gain is available only at the receiver. In this approach, transmission takes place at a certain rate and tolerates some information loss when an outage event occurs. An outage event occurs whenever the transmitted rate is not supported by the instantaneous channel gain; i.e., when the channel gain is too low for successful decoding of the transmitted message. It is assumed that outage events occur with low probability that reliable communication is available most of the time. A different strategy to deal with slow fading is

the broadcast channel approach [26]. In this approach different states of the channel are treated as channels toward different receivers (a receiver for each state). Hence, the same strategy as used for sending common and private messages to different users on the Gaussian broadcast channel can be applied here. When the channel gain is also available at the encoder, the encoder can adapt the power and the transmission rate as a function of the instantaneous state of the channel and thus can achieve a higher rate on average. Moreover, as regards the outage capacity, the transmitter can use power control to conserve power by not transmitting at all during designated outage periods.

When dealing with outage capacity for fading MAC, the common outage has a similar definition to the outage event in the point to point case. A common outage event is declared whenever we transmit with rates that are not supported by the instantaneous channel gains. If the channel gains are available at both the decoder and the encoders, additional notions of capacities for the fading MAC need to be taken into account. The throughput capacity region for the Gaussian fading MAC was introduced in [27]. In a nutshell, this is the Shannon capacity region where the codewords can be chosen as a function of the realization of the fading with arbitrarily long coding delays. However, as for the point to point case, this approach is not realistic in slow fading cases since it requires a very long delay to average out the fading effect. [28] derived the delay limited capacity for the Gaussian fading MAC (also known as the zero outage capacity). In the delay limited capacity, unlike the throughput capacity, the chosen coding delay has to work uniformly for all fading processes with a given stationary distribution. However, the delay limited capacity is somewhat pessimistic due to the demand to maintain a constant rate under any fading condition. The outage capacity region and the optimal power allocation for a fading MAC were described in [29]. As was pointed out in [29], in a slow fading environment, the decoding delay depends solely on the code-length employed and not on the time variation of the channel.

The demand for interference free channels at all nodes means that some level of orthogonality is required between different transmissions in the network. Avoiding interference between all nodes comes at the cost of loss of expensive bandwidth, or alternatively leads to rate degradation in band limited systems. The same argument can be applied to the limited interference model since some orthogonality at a certain radius is required. In [1], the MAC network coding model was introduced. In the MAC network model, in contrast to the wireless broadcast advantage based models, the superposition property of the wireless medium is exploited. The network is treated as a collection of multi access channels, such that each receiver simultaneously receives data from all its in-neighbors.

4

**Main contributions:** This paper explores the problem of rate allocation for multicasting over slow Rayleigh fading channels using network coding. The problem is examined in a model where the network is treated as a collection of Rayleigh fading multi access channels. In our network model, we assume that links on the network vary faster than the entire network can respond to the variations. Therefore, our goal is to find a rate allocation scheme that is based solely on the statistics of the channels which minimizes the outage probability. This paper differs from prior works at two major aspects. Prior works' models assume long time averaging of the instantaneous capacity (as in the ergodic capacity approach) or averaging of the packet arrival rate (see e.g., [9]). These assumptions are more suitable for fast fading model while in slow fading model this is unrealistic. Hence, in this paper we design a different rate allocation scheme which is more suitable to the slow fading model. Moreover, in this paper the design of the rate allocation scheme is based solely on the statistics which is desirable in many practical large scale networks, as will be emphasized in section 4.

The communication model is described in detail in section 2.

In section 3 we present lower and upper bounds for the outage probability of a fading MAC. In section 4 a suboptimal solution for the rate allocation problem is presented for the MAC network model. The solution is based on an upper bound on the probability of outage in the fading MAC. In section 5 a distributed solution is derived for the rate allocation problem in the MAC network model. In section 6 we report some simulation results. We end with concluding remarks.

## 2. Communication model

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph with a set of nodes $\mathcal{V}$ and directed edges $\mathcal{E} \subset \mathbf{V} \times \mathbf{V}$, where transceivers are nodes and channels are edges representing a wireless communication network. In this paper, scalars and random variables are denoted by lower case letters. Vectors and matrices are denoted by boldface lower and upper case letters, respectively. We are abusing of notation a bit by using the same letters to refer to random variables and their realizations. The cardinality of any set $\mathcal{A}$ is denoted by $|\mathcal{A}|$. All vectors are columns and inequalities between vectors are defined element-wise; i.e., $\mathbf{v} \leq \mathbf{u}$ implies $v_i \leq u_i$ for all $i$. For any node $j \in \mathcal{V}$, we denote the in-neighborhood and out-neighborhood of $j$ by $\mathcal{I}(j)$ and $\mathcal{O}(j)$ respectively, i.e., $\mathcal{I}(j) = \{i : (i, j) \in \mathcal{E}\}$ and $\mathcal{O}(j) = \{i : (j, i) \in \mathcal{E}\}$. The network is treated as a collection of multi access channels, such that each receiver simultaneously receives data from all its in-neighbors. For simplicity, it is assumed that there is no interference between transmissions toward different re-

ceivers (see Fig. 1(a)). This can be achieved by orthogonal transmissions e.g., by using a certain frequency reuse pattern or directional antennas. Clearly, this is an improvement over a model where all transmissions are orthogonal. If we consider the MAC network model with deterministic channel gains, a joint power control and rate allocation solution for a (convex) network utility can be found distributively [21]. This is due to the convexity of the capacity region of the multi access channels [30]. The deterministic model can be adapted to deal with fast fading channels in the case of a constant power allocation vector by using the ergodic MAC capacity region instead of the MAC capacity region. This ergodic capacity region is easily obtained by taking the expectations of the capacity constraints [30]. Here, we examine the MAC network model in the case of slow fading channels. We aim to find a rate allocation scheme that is based solely on the statistics of the channels which minimizes the outage probability.

The channel gain of link $(i,j)$ is denoted by $h_{i,j}$. $h_{i,j}$ is a zero mean circular complex normal random variable with variance of $v_{i,j}^2$. It is assumed that all $h_{i,j}$ are independent of each other. Denote by $\mathbf{h}_j := [h_{i,j} : i \in \mathcal{I}(j)]$ and by $\boldsymbol{\eta_j} := [|h_{i,j}|^2 : i \in \mathcal{I}(j)]$. The transmission on link $(i,j)$ is denoted by $x_{i,j}$ and it is transmitted with an average power $p_{i,j}$. We assume that $\sigma_j^2$ is the variance of $\xi_j$ - the zero mean Gaussian noise at node $j$. Hence, the received signal at node $j$ is given by:

$$y_j = \sum_{i \in \mathcal{I}(j)} h_{i,j} x_{i,j} + \xi_j. \tag{1}$$

Fig. 1(b) illustrates the MAC of node 1 in the network of Fig. 1(a). The rate transmitted on a link $(i,j)$ is denoted by $r_{i,j}$, the rate allocation vector is denoted by $\mathbf{r} = [r_{i,j} : (i,j) \in \mathcal{E}]$ and the local rate allocation vector is denoted by $\mathbf{r}_j = [r_{i,j} : i \in \mathcal{I}(j)]$. When the instantaneous channel gains $h_{i,j}$ are deterministic and known, this is the well-known Gaussian multi access channel [30]. Hence, the instantaneous MAC capacity region is given by:

$$\mathcal{V}_j^{\text{ins}}(\mathbf{h}_j) := \left\{ r_{i,j} : \begin{array}{c} \sum_{i \in \mathcal{M}(j)} r_{i,j} \leq \log_2\left(1 + \frac{P_{\mathcal{M}(j),j}}{\sigma_j^2}\right) \\ \forall \mathcal{M}(j) \subseteq \mathcal{I}(j) \end{array} \right\}, \tag{2}$$

where $P_{\mathcal{M}(j),j} = \sum_{i \in \mathcal{M}(j)} p_{i,j} |h_{i,j}|^2$. However, when dealing with Rayleigh channels, this capacity region may not be a good measure of performance and the outage capacity is a better and more practical alternative. A common outage event

6

Figure 1: The MAC network model: (a) An illustration of a wireless network with 12 transceivers positioned as a directed graph $\mathcal{G}$. In the MAC network model each receiver receives data from all its in-neighborhood nodes. For example the nodes in $\mathcal{I}(1) = \{2, 3, 6\}$ transmit toward node 1 and the nodes in $\mathcal{I}(4) = \{1, 6, 7\}$ transmit toward 4. However, it is assumed that (for example), there is no interference between the transmissions toward node 1 and the transmissions toward node 4. (b) The MAC of node 1.

is jointly declared for all links whenever we transmit toward a certain node with rates that are not supported by the instantaneous MAC capacity region.

**Definition 1.** *For a rate vector $\mathbf{r}_j$ and for the MAC associated with node $j$, the common outage event is*

$$\mathbf{r}_j \notin \mathcal{V}_j^{\text{ins}}(\mathbf{h}_j), \tag{3}$$

*where the (random) capacity region $\mathcal{V}_j^{\text{ins}}(\mathbf{h}_j)$ is defined in (2).*

**Definition 2.** *The probability of outage in the fading MAC of node $j$ is given by*

$$P_j^{\text{out}} = \Pr\left(\mathbf{r}_j \notin \mathcal{V}_j^{\text{ins}}(\mathbf{h}_j)\right). \tag{4}$$

Similar to these definitions, we define an outage event and outage probability for the MAC network model.

**Definition 3.** *The outage event for the MAC network model is the event for which there exists node $j \in \mathcal{V}$ such that $\mathbf{r}_j \notin \mathcal{V}_j^{\text{ins}}(\mathbf{h}_j)$.*

Hence, the probability of outage in the MAC network model is given by

$$P_{\text{MAC}}^{\text{out}} = \Pr\left(\bigcup_{j \in \mathcal{V}} \left\{\mathbf{r}_j \notin \mathcal{V}_j^{\text{ins}}(\mathbf{h}_j)\right\}\right). \tag{5}$$

To complete the description of the local communication model, we associate the codebooks, the encoders $(\mathcal{F}_{i,j} : i \in \mathcal{I}(j))$ and the decoder $g_j$ that establish the connection between $\mathcal{I}(j)$ and $j$ at rates $(r_{i,j} : i \in \mathcal{I}(j))$ to any node in the network. Obviously, node $j$ shares the appropriate codebooks and encoders with its in-neighbors.

The source node is denoted by $s \in \mathbf{V}$ and it is assumed that $\mathcal{I}(s) = \phi$. The set of all destinations (sinks) is denoted by $\mathcal{D}_s \subseteq \mathbf{V} \backslash \{\mathbf{s}\}$. Intermediate nodes are allowed to send out packets that are a combination of their received information and as a result they break the flow conservation by increasing/decreasing the outside rate. However, the main theorem of network coding for multicast is stated in terms of the max-flow (min-cut) between each source and its destinations. Therefore, we distinguish between the flow at an edge $(i, j)$ and the actual rate at that link. Let $f_{i,j}^d$ be the flow at edge $(i, j)$ destined for destination $d \in \mathcal{D}_s$, and let $r_{i,j}$ be the actual at edge $(i, j)$. The communication parameters are summarized in Table 1.

Table 1: A summary of communication model parameters

| | | |
|---|---|---|
| $x_{i,j}$ | Transmission on link $(i,j)$ | |
| $h_{i,j}$ | Channel gain of link $(i,j)$ | $h_{i,j} \sim \mathcal{CN}(0, v_{i,j}^2)$ |
| $\mathbf{h}_j$ | $\mathbf{h}_j = [h_{i,j} : i \in \mathcal{I}(j)]$ | |
| $\boldsymbol{\eta}_j$ | $\boldsymbol{\eta}_j = [|h_{i,j}|^2 : i \in \mathcal{I}(j)]$ | |
| $p_{i,j}$ | Average power of the transmission on link $(i,j)$ | |
| $\xi_j$ | Noise at node $j$ | $\xi_j \sim \mathcal{CN}(0, \sigma_j^2)$ |
| $y_j$ | Received signal at node $j$ | |
| $r_{i,j}$ | Rate at link $(i,j)$ | |
| $\mathbf{r}_j$ | Local rate allocation vector | $\mathbf{r}_j = [r_{i,j} : i \in \mathcal{I}(j)]$ |
| $\mathbf{r}$ | Rate allocation vector | $\mathbf{r} = [r_{i,j} : (i,j) \in \mathcal{E}]$ |
| $\mathcal{V}_j^{\text{ins}}(\mathbf{H})$ | Instantaneous MAC capacity region of node $j$ | |
| $P_j^{\text{out}}$ | Probability of outage in the fading MAC of node $j$ | |
| $P_{\text{MAC}}^{\text{out}}$ | Probability of outage in the MAC network model | |
| $s$ | Source node | |
| $\mathcal{D}_s$ | Set of all destinations | $\mathcal{D}_s \subseteq \mathbf{V} \backslash \{\mathbf{s}\}$ |
| $f_{i,j}^d$ | Flow at edge $(i,j)$ destined for destination $d \in \mathcal{D}_s$ | |

As was mentioned in section 1, there is no loss of optimality by first finding the optimal rate allocation solution and then designing the coding scheme that realizes the connection. In the following section the rate allocation vector for the MAC network model is given as the solution to an optimization problem and the coding scheme that realizes the connection is assumed to be given. For large scale networks, where global network information is not available, the random network coding shown in [14, 15] can be employed. In general, in random network coding, intermediate nodes store all their received packets in their memory and when a packet injection occurs on an outgoing link, the node forms a packet that is a random linear combination of the packets in its memory. In order to enable decoding at the destinations, the random coefficients of the linear combinations are included in the header of the packet as side information. These coefficients are called the global encoding vector of the packet. Decoding is possible if all destinations collect enough packets with linearly-independent global encoding vectors. The algorithm shown in [14] for random packet level network coding was adjusted to the MAC network model in [1].

## 3. Bounds on the probability of outage of a MAC

In this section we bound the outage probability of a fading MAC. To do so, we need the following notations and definitions. Consider a (slow) fading MAC with $n$ links each of which is a Rayleigh channel i.e., $h_i \sim \mathcal{CN}(0, v_i^2)$ $i = 1, 2, \cdots, n$. Denote the variance of the zero mean Gaussian noise at the receiver by $\sigma^2$. For any matrix $\mathbf{B}$, $b_{i,j}$ denotes the entry in the $i$'th row and $j$'th column of $\mathbf{B}$. Let $\mathbf{B}^{r*}$ be a submatrix of $\mathbf{B}$ constructed by deleting the $r$'th row of $\mathbf{B}$. For $r = 1$ we denote $\mathbf{B}^{1*} = \mathbf{B}^*$. For any $n \geq 1$ let $\mathbf{1}_n, \mathbf{0}_n$ be vectors with length $n$ of ones and zeros, respectively. For any $n \geq 1$, let $\mathbf{A}_n$ be a $(2^n - 1) \times n$ matrix, such that for $n = 1$ $\mathbf{A}_1 = 1$ and for $n \geq 2$

$$\mathbf{A}_{n+1} = \begin{bmatrix} \mathbf{0}_{2^n-1} & , & \mathbf{A}_n \\ 1 & , & \mathbf{0}_n^T \\ \mathbf{1}_{2^n-1} & , & \mathbf{A}_n. \end{bmatrix}, \tag{6}$$

i.e., each row of $\mathbf{A}_n$ is the binary representation of the row index (for example $\mathbf{A}_2 = \left[[0,1]^T, \ [1,0]^T, [1,1]^T\right]^T$). For any scalar $a$ and vector $\mathbf{v} \in \mathbb{R}^K$, $\mathbf{c} = a^{\mathbf{v}} - 1$ is calculated point-wise; i.e., $c_i = a^{v_i} - 1$.

The probability of outage of a fading MAC is given in definition 2. Obviously, the probability of outage can be expressed as

$$\mathrm{Pr}_{\mathrm{MAC_n}}^{\mathrm{out}} = 1 - \mathrm{Pr}\left(\mathbf{r} \in \mathcal{V}^{\mathrm{ins}}(\mathbf{h})\right), \tag{7}$$

where $\mathbf{h} := [\mathbf{h_1}, \mathbf{h_2}, \cdots, \mathbf{h_n}]$ and $\mathcal{V}^{\mathrm{ins}}(\mathbf{h})$ is the instantaneous capacity region. As can be seen from (2) the expression $\mathbf{r} \in \mathcal{V}^{\mathrm{ins}}(\mathbf{h})$ stands for a conjunction of $(2^n - 1)$ inequalities, each of which is in the form of

$$\sum_{i \in \mathcal{M}} r_i \leq \log_2\left(1 + \frac{P_{\mathcal{M}}}{\sigma^2}\right), \tag{8}$$

where $P_{\mathcal{M}} = \sum_{i \in \mathcal{M}} p_i |h_i|^2$ and $\mathcal{M}$ is a subset of $\{1, 2, \ldots, n\}$. Rewriting (8) in a matrix form yields

$$\mathbf{a}_{\mathcal{M}}^T \mathbf{r} \leq \log_2\left(1 + \mathbf{a}_{\mathcal{M}}^T \frac{1}{\sigma^2} \mathbf{P} \boldsymbol{\eta}\right), \tag{9}$$

where, $\mathbf{a}_{\mathcal{M}}$ is a vector with length $n$ such that

$$a_i = \begin{cases} 1 & i \in \mathcal{M} \\ 0 & \text{otherwise,} \end{cases}$$

10

$\mathbf{P}$ is an $n \times n$ diagonal matrix with $p_1, p_2, \ldots, p_n$ on the main diagonal and $\boldsymbol{\eta} = [|h_1|^2, |h_2|^2, \ldots, |h_n|^2]^T$. A simple algebraic operation yields that (9) is equivalent to

$$2^{\mathbf{a}_{\mathcal{M}}^T \mathbf{r}} - 1 \leq \mathbf{a}_{\mathcal{M}}^T \frac{1}{\sigma^2} \mathbf{P} \boldsymbol{\eta}. \tag{10}$$

Note that since $|h_i|^2 \quad i = 1, 2, \ldots, n$ are independent exponential random variables with an expectation of $2v_i^2$, the random variables $z_i = \frac{1}{2v_i^2}|h_i|^2 \quad i = 1, 2, \ldots, n$ are i.i.d exponential random variables with an expectation of $1$. Hence, the event in (10) is equivalent to the event

$$2^{\mathbf{a}_{\mathcal{M}}^T \mathbf{r}} - 1 \leq \mathbf{a}_{\mathcal{M}}^T \frac{1}{\sigma^2} \mathbf{P} \boldsymbol{\Upsilon} \mathbf{z}_n, \tag{11}$$

where $\boldsymbol{\Upsilon}$ is a diagonal matrix with $2v_1^2, 2v_2^2, \ldots, 2v_n^2$ on the main diagonal and $\mathbf{z}_n = [z_1, z_2, \ldots, z_n]^T$ is a vector of $n$ i.i.d standard exponential random variables; $E\{z_i\} = 1$. Therefore, from (7) and (11), it is implied that the outage probability in a MAC with $n$ links can be written as

$$\mathrm{Pr}_{\mathrm{MAC_n}}^{\mathrm{out}} = 1 - \mathrm{Pr}\left(\mathbf{A}_n \mathbf{D}_n \mathbf{z}_n \geq \mathbf{b}_n\right), \tag{12}$$

where $\mathbf{D}_n$ is a diagonal matrix with $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \cdots, \frac{1}{\lambda_n}$ on the main diagonal, $\lambda_i = \frac{1}{2v_i^2} \frac{\sigma^2}{p_i}$ and $\mathbf{b}_n = 2^{\mathbf{A}_n \mathbf{r}} - 1$. For example in a MAC with $n = 3$ links we have that

$$\mathbf{A}_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 2^{r_3} - 1 \\ 2^{r_2} - 1 \\ 2^{r_2+r_3} - 1 \\ 2^{r_1} - 1 \\ 2^{r_1+r_3} - 1 \\ 2^{r_1+r_2} - 1 \\ 2^{r_1+r_2+r_3} - 1 \end{bmatrix}.$$

Note that when the MAC is with i.i.d links we have that $\mathbf{D}_n = \frac{1}{\lambda} \mathbf{I}_n$ where $\mathbf{I}_n$ is an $n \times n$ identity matrix.

Hence, the probability of outage in a MAC is related to the joint distribution of linear combinations of exponential random variables. Huffer and Lin [31] presented an algorithm for the computation of the exact expression of the joint distribution of general linear combinations of spacings[2] by repeated uses of two

---

[2]Suppose $u_i \quad i = 1, 2, \cdots, n$ are independently and uniformly distributed on the interval

recursions that reduce the dimensionality of the problem. They also pointed out that the algorithm remains valid as well for linear combinations of exponential random variables. However, this is inaccurate and in this paper we revise result to handle exponential random variables. The new recursion is given in Lemma 6 in Appendix Appendix A.

By using the algorithm in [31] together with Lemma 6, an exact expression of the probability of common outage can be computed. However, the computation of a symbolic expression becomes extremely complicated in a MAC with more than 2 links. Therefore, we present an upper and a lower bound on that outage probability. To that end, we need the following lemma:

**Lemma 1.** *Let $\mathbf{z}$ be a vector of $n$ i.i.d exponential random variables with an expectation $E\{z_i\} = 1$. If there exists an entry $a_{r,i} > 0$ in $\mathbf{A}$ such that $a_{r,j} = 0$ for all $j \neq i$ and $b_r \geq 0$, the following holds*

$$\Pr\left(\mathbf{Az} > \lambda \mathbf{b}\right) = e^{-\lambda \frac{b_r}{a_{r,i}}} \Pr\left(\mathbf{A}^{r*}\mathbf{z} > \lambda\left(\mathbf{b}^{r*} - \frac{b_r}{a_{r,i}}\mathbf{a}^{r*}\right)\right).$$

*Proof:* This lemma is an immediate consequence of Lemma 6 (i.e., when $k = 1$ and $\Psi = \{j\}$). $\square$

In [32] it was pointed out that the outage probability of a MAC with $n$ i.i.d links is bounded from below by

$$\Pr_{\mathrm{MAC_n}}^{\mathrm{out}} \geq 1 - e^{-\lambda S_n} \frac{\Gamma\left(\mathrm{n}, \lambda\left(2^{\mathrm{R_n}} - 1 - \mathrm{S_n}\right)\right)}{(\mathrm{n} - 1)!}, \tag{13}$$

where $\Gamma(n, x)$ is the incomplete gamma function, $S_n = \sum_{k=1}^{n}\left(2^{r_n} - 1\right)$ and $R_n = \sum_{i=1}^{n} r_i$. In the case of a Rayleigh fading MAC with independent links but with different variances, Theorem 2 gives a lower bound on the probability of outage.

**Theorem 2.** *Let $\lambda_i = \frac{1}{2\nu^2}\frac{\sigma}{p_i}$ $i = 1, 2, \cdots, n$ have distinct values; i.e., $\lambda_i \neq \lambda_j$ for all $i \neq j$, then the probability of a MAC with $n$ independent $\mathrm{Rayleigh}(\upsilon_i)$ channel*

---

(0,1), and let $u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(n)}$ be the corresponding order statistics. The spacings $s_1, s_2, \cdots, s_{n+1}$ are defined by the successive differences between the order statistics: $s_i = u_{(i)} - u_{(i-1)}$, where $u_{(0)} := 0$ and $u_{(n+1)} := 1$.

$i = 1, 2, \cdots, n$ *is bounded below by*

$$\mathrm{Pr}_{\mathrm{MAC_n}}^{\mathrm{out}} \geq 1 - \sum_{i=1}^{n} \gamma_i e^{-\beta_n - \lambda_i \left(2^{R_n} - S_n - 1\right)}, \tag{14}$$

*where,* $\gamma_i = \prod_{j \neq i} \dfrac{\lambda_j}{\lambda_j - \lambda_i}$, $\beta_n = \sum_{i=1}^{n} \lambda_i (2^{r_i} - 1)$, $S_n = \sum_{i=1}^{n} (2^{r_i} - 1)$ *and* $R_n =$ $\sum_{i=1}^{n} r_i$.

*Proof:* As was explained earlier, the expression $\mathbf{A}_n \mathbf{D}_n \mathbf{z}_n \geq \mathbf{b}_n$ in equation (12) stands for a conjunction of $(2^n - 1)$ inequalities. The inequalities that are related to the rows of $\mathbf{A}_n$ indexed by $2^i \quad i = 0, 1, \ldots, (n-1)$ stand for the direct instantaneous capacity constraints $r_i \leq \log_2 \left(1 + \frac{|h_i|^2 p_i}{\sigma^2}\right)$, whereas all the other inequalities refer to constraints that involve more than one link of the MAC capacity region (see eq. (2)). In particular, the $(2^n - 1)^{\mathrm{th}}$ inequality refers to the constraint

$$\sum_{i=1}^{n} r_i \leq \log_2 \left(1 + \sum_{i=1}^{n} \frac{p_i |h_i|^2}{\sigma^2}\right).$$

Obviously,

$$\mathrm{Pr}\left(\mathbf{A}_n \mathbf{D}_n \mathbf{z}_n \geq \mathbf{b}_n\right) \leq \mathrm{Pr}\left(\tilde{\mathbf{A}}_n \mathbf{D}_n \mathbf{z}_n \geq \tilde{\mathbf{b}}_n\right), \tag{15}$$

where $\tilde{\mathbf{A}}_n$ is a submatrix of $\mathbf{A}_n$ constructed by taking the rows indexed by $\{2^i : i = 0, 1, \ldots, (n-1)\} \cup \{2^n - 1\}$ of $\mathbf{A}_n$ and $\tilde{\mathbf{b}}_n$ is a sub-vector of $\mathbf{b}_n$ constructed by taking the appropriate entries of $\mathbf{b}_n$. Note that up to a permutation of the rows, the first $n$ rows of $\tilde{\mathbf{A}}_n$ is the identity matrix. Therefore, we can eliminate these $n$ rows by $n$ uses of lemma 1. Hence, we have that

$$\mathrm{Pr}\left(\tilde{\mathbf{A}}_n \mathbf{D}_n \mathbf{z}_n \geq \tilde{\mathbf{b}}_n\right) = e^{-\beta_n} \mathrm{Pr}\left(\sum_{i=1}^{n} \frac{z_i}{\lambda_i} \geq x\right), \tag{16}$$

where $x = 2^{R_n} - S_n - 1$. The probability of a distinct coefficients linear combination of i.i.d exponential variables is given by [33]

$$\mathrm{Pr}\left(\sum_{i=1}^{n} \frac{z_i}{\lambda_i} \geq x\right) = \sum_{i=1}^{n} \gamma_i e^{-\lambda_i x}. \tag{17}$$

13

The claim now follows. □

Note that Theorem 2 is valid only when $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all distinct. When we have a MAC with a set of $K$ links with the same value of $\lambda_i$ and $n - K$ links with distinct values of $\lambda$, a similar bound can be computed by replacing the probability in (17) by integrating the expression of the pdf derived in [34]. We omit the calculation of this probability here, for the sake of brevity.

For simplicity, in the derivation of the upper bound we only consider a MAC with $n$ i.i.d links. Computing the bound for the case where there are independent links with nonidentical variances is much complicated and is considered as future work. The upper bound for a MAC with $3$ links is given in Lemma 3 and the upper bound for the general case is given in Theorem 4.

**Lemma 3.** *The probability of common outage of a MAC with $3$ i.i.d links is bounded by*

$$\Pr^{\text{out}}_{\text{MAC}_3} \leq 1 - e^{-\lambda\left(2^{\text{R}_3} - 1\right)} G(\lambda \alpha_3) \tag{18}$$

*where $G(x) = \frac{1}{2}x^2 + x + 1$, $\lambda = \frac{1}{2v^2}\frac{\sigma^2}{p}$, $R_3 = \sum_{i=1}^{3} r_i$ and $\alpha_3 = (2^{r_1} - 1)(2^{r_2} - 1)(2^{r_3} - 1)$.*

*Proof:* The probability of a successful (non-outage) transmission is given by

$$1 - \Pr^{\text{out}}_{\text{MAC}_3} = \Pr\left(\mathbf{A}_3 \mathbf{z}_3 \geq \lambda \mathbf{b}_3\right). \tag{19}$$

Define the following constants $\beta_i = 2^{r_i} - 1$ and $\beta_{i,j} = 2^{r_i + r_j} - 1$. Note that the rows of $\mathbf{A}_3$ and $\mathbf{b}_3$ indexed by $\{2^i : i = 0, 1, 2\}$ satisfy the conditions of lemma 1. Hence, by three uses of Lemma 1 we can eliminate these rows of $\mathbf{A}$ and $\mathbf{b}_3$. These three uses of Lemma 1 are legitimate, since after each use of the lemma the result matrix and vector still satisfy the conditions of Lemma 1 (since $\beta_i \geq 0$, $\beta_{i,j} - \sum_{i \in \mathcal{B}} \beta_i \geq 0$ and $(2^{R_3} - 1) - \sum_{i=1}^{j} \beta_i \geq 0$, for all $\mathcal{B} \subseteq \{i, j\}$ and $i, j \in \{1, 2, 3\}$).

Therefore, the probability of successful transmission can be rewritten as

$$1 - \Pr^{\text{out}}_{\text{MAC}_3} = e^{-\lambda S_3} \Pr\left(\mathbf{A}\mathbf{z} \geq \lambda \tilde{\mathbf{b}}\right), \tag{20}$$

where $\mathbf{A}$ is a submatrix of $\mathbf{A}_3$ constructed by deleting the rows indexed by $\{2^i : i = 0, 1, 2\}$, $S_3 = \sum_{i=1}^{3} (2^{r_i} - 1)$ and $\tilde{\mathbf{b}} = [\beta_{2,3} - \beta_2 - \beta_3 , \beta_{1,3} - \beta_1 - \beta_3 , \beta_{1,2} - $

14

$\beta_1 - \beta_2 , (2^{R_3} - 1) - \beta_1 - \beta_2 - \beta_3\big]^T$. It is easy to see that

$$\Pr\left(\mathbf{A}\mathbf{z} \geq \lambda\tilde{\mathbf{b}}\right) \geq \Pr\left(\tilde{\mathbf{A}}\mathbf{z} \geq \lambda\tilde{\mathbf{b}}\right), \tag{21}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad \tilde{\mathbf{A}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

since the eliminated $z_1, z_2$ and $z_3$ are non-negative random variables. Note that since $\tilde{b}_4 - \sum_{i=0}^{3} \tilde{b}_i = \alpha \geq 0$ and $\tilde{\mathbf{b}} \geq 0$, we have that $\tilde{b}_4 - \sum_{i=0}^{j} \tilde{b}_i \geq 0$ for all $j \in \{1, 2, 3\}$. Again, by three uses of Lemma 1, we can eliminate the first three rows of $\tilde{\mathbf{A}}$ and write

$$\Pr\left(\tilde{\mathbf{A}}\mathbf{z} \geq \lambda\tilde{\mathbf{b}}\right) = e^{-\lambda\tilde{\gamma}} \Pr\left(z_1 + z_2 + z_3 \geq \lambda\alpha_3\right), \tag{22}$$

where $\tilde{\gamma} = \tilde{b}_1 + \tilde{b}_2 + \tilde{b}_3$. Note that $Z = z_1 + z_2 + z_3$ has an $\mathrm{Erlang}(3, 1)$ distribution and therefore

$$\Pr\left(Z > z\right) = e^{-z}G(z). \tag{23}$$

Hence, (22) can be rewritten as

$$\Pr\left(\tilde{\mathbf{A}}\mathbf{z} \geq \lambda\tilde{\mathbf{b}}\right) = e^{-\lambda\tilde{\gamma}}e^{-\lambda\alpha_3}G(\lambda\alpha_3). \tag{24}$$

Combining (20),(21) and (24) yields

$$1 - \mathrm{Pr}_{\mathrm{MAC}_3}^{\mathrm{out}} \geq e^{-\lambda(\alpha_3+\gamma+\tilde{\gamma})}G(\lambda\alpha_3). \tag{25}$$

The claim now follows from the fact that

$$\alpha_3 + \gamma + \tilde{\gamma} = 2^{R_3} - 1.$$

$\square$

**Theorem 4.** *The probability of common outage of a MAC with $n \geq 3$ i.i.d $\mathrm{Rayleigh}(\upsilon)$ channels is bounded by*

$$\mathrm{Pr}_{\mathrm{MAC}_\mathrm{n}}^{\mathrm{out}} \leq 1 - e^{-\lambda\left(2^{\mathrm{R_n}}-1\right)}\tilde{G}(\lambda\alpha_\mathrm{n}) \tag{26}$$

*where $G(x) = \frac{1}{2}x^2 + x + 1$, $R_n = \sum_{i=1}^{n} r_i$ and $\alpha_n = \prod_{i=1}^{n}(2^{r_i} - 1)$.*

15

*Proof:* The proof proceeds by induction on $n$. The claim is true for $n = 3$ by Lemma 3. Let the statement be true for $n = k$. We will now prove the result for $n = k + 1$. From (12) we have that

$$1 - \text{Pr}^{\text{out}}_{\text{MAC}_{k+1}} = \text{Pr}\left(\mathbf{A}_{k+1}\mathbf{z}_{k+1} \geq \lambda \mathbf{b}_{k+1}\right). \tag{27}$$

Note that $[\mathbf{b}_{k+1}]_{2^k} = 2^{r_1} - 1$. Also note that by exploiting the structure of $\mathbf{A}_{k+1}$ (see equation (6)), the $2^k$'th row of $\mathbf{A}_{k+1}$ can be eliminated by using Lemma 1. Hence,

$$1 - \text{Pr}^{\text{out}}_{\text{MAC}_{k+1}} = e^{-\lambda(2^{r_1}-1)}\,\text{Pr}\left(\tilde{\mathbf{A}}_{k+1}\mathbf{z}_{k+1} \geq \lambda \tilde{\mathbf{b}}_{k+1}\right), \tag{28}$$

where

$$\tilde{\mathbf{A}}_{k+1} = \begin{bmatrix} \mathbf{0}_{2^k-1} & , & \mathbf{A}_k \\ \mathbf{1}_{2^k-1} & , & \mathbf{A}_k \end{bmatrix} \tag{29}$$

and $\tilde{\mathbf{b}}_{k+1} = \left(2^{\tilde{\mathbf{A}}_{k+1}\mathbf{r}} - 1\right) - (2^{r_1} - 1)\left[\mathbf{0}^T_{2^k-1}, \mathbf{1}^T_{2^k-1}\right]^T$. It can easily be verified that

$$\text{Pr}\left(\tilde{\mathbf{A}}_{k+1}\mathbf{Z}_{k+1} \geq \lambda \tilde{\mathbf{b}}_{k+1}\right) \geq \text{Pr}\left(\check{\mathbf{A}}_{k+1}\mathbf{Z}_{k+1} \geq \lambda \tilde{\mathbf{b}}_{k+1}\right), \tag{30}$$

where

$$\check{\mathbf{A}}_{k+1} = \begin{bmatrix} \mathbf{0}_{2^k-1} & , & \mathbf{A}_k \\ \mathbf{0}_{2^k-1} & , & \mathbf{A}_k \end{bmatrix},$$

since $z_1 \geq 0$. Note that for any two vectors $\mathbf{0}_{2^k-1} \leq \mathbf{x}_1 \leq \mathbf{x}_2$ the following holds

$$\text{Pr}\left(\check{\mathbf{A}}_{k+1}\mathbf{z}_{k+1} \geq \left[\mathbf{x}_1^T, \mathbf{x}_2^T\right]^T\right) = \text{Pr}\left(\mathbf{A}_k\mathbf{z}_k \geq \mathbf{x}_2\right). \tag{31}$$

Note that

$$2^{\tilde{\mathbf{A}}_{k+1}\mathbf{r}} = \begin{bmatrix} 2^{\mathbf{A}_k\boldsymbol{\gamma}} \\ 2^{r_1} \cdot 2^{\mathbf{A}_k\boldsymbol{\gamma}} \end{bmatrix}, \tag{32}$$

where $\boldsymbol{\gamma} = [r_2, r_3, \ldots, r_{k+1}]^T$. Therefore, we have that

$$\tilde{\mathbf{b}}_{k+1} = \begin{bmatrix} 2^{\mathbf{A}_k\boldsymbol{\gamma}} - 1 \\ 2^{r_1}\left(2^{\mathbf{A}_k\boldsymbol{\gamma}} - 1\right) \end{bmatrix}. \tag{33}$$

Since $2^{r_1} \geq 1$, combining (31) with (33) yields

$$\text{Pr}\left(\check{\mathbf{A}}_{k+1}\mathbf{z}_{k+1} \geq \lambda \tilde{\mathbf{b}}_{k+1}\right) = \text{Pr}\left(\mathbf{A}_k\mathbf{z}_k \geq \tilde{\lambda}\left(2^{\mathbf{A}_k\boldsymbol{\gamma}} - 1\right)\right), \tag{34}$$

where $\tilde{\lambda} = \lambda 2^{r_1}$. By the induction hypothesis for $k$, we have that

$$\text{Pr}\left(\mathbf{A}_k\mathbf{z}_k \geq \tilde{\lambda}\left(2^{\mathbf{A}_k\boldsymbol{\gamma}} - 1\right)\right) \geq e^{-\tilde{\lambda}\left(2^{\tilde{R}_2}-1\right)}G\left(\tilde{\lambda}\tilde{\alpha}_2\right), \tag{35}$$

16

where $\tilde{R}_2 = \sum_{i=2}^{k+1} r_i$ and $\tilde{\alpha}_2 = \prod_{i=2}^{k+1} (2^{r_i} - 1)$. Combining (28),(30),(34) and (35) yields

$$1 - \Pr_{\text{MAC}_{k+1}}^{\text{out}} \geq e^{-\lambda(2^{r_1}-1)} e^{-\tilde{\lambda}\left(2^{\tilde{R}_2}-1\right)} G\left(\tilde{\lambda}\tilde{\alpha}_2\right). \tag{36}$$

The claim now follows from the fact that

$$\lambda(2^{r_1} - 1) + \tilde{\lambda}\left(2^{\tilde{R}_2} - 1\right) = \lambda\left(2^{R_{k+1}} - 1\right),$$

$\tilde{\lambda}\tilde{\alpha}_2 \geq \lambda\alpha_{k+1}$ and that $G(x)$ is monotonically increased with $x$.  $\square$

Note that the lower bound (13) in the i.i.d case may also be expressed as

$$\Pr_{\text{MAC}_n}^{\text{out}} \geq 1 - e^{-\lambda\left(2^{R_n}-1\right)} \tilde{G}\left(\lambda\left(2^{R_n} - 1 - S_n\right)\right), \tag{37}$$

where, $\tilde{G}(x) = \sum_{k=0}^{n-1} \frac{1}{k!} x^k$. Hence,

$$\tilde{G}(\lambda\beta_n) \geq e^{\lambda\left(2^{R_n}-1\right)}\left(1 - \Pr_{\text{MAC}_n}^{\text{out}}\right) \geq G(\lambda\alpha_n), \tag{38}$$

where $\beta_n = 2^{R_n} - 1 - S_n$.

## 4. Rate allocation for the fading MAC network model

In this section we study the problem of finding the rate allocation vector for the fading MAC network model discussed in the previous sections. In our wireless model we assume a slow fading model with independent Rayleigh fading channels. For simplicity we only consider a network in which $\lambda_{i,j} = \lambda_j$ for all $j \in \mathcal{V}\backslash\{s\}$. In other words, the network is assumed to be a collection of multiple access channels with i.i.d links (note that $\lambda_j$ $j \in \mathcal{V}\backslash\{s\}$ may be distinct). This assumption comes down to the fact that we normalized the transmission power of nodes that connected to the same receiver appropriately to the statistics of the best channel. As was mentioned earlier, the case where there are independent links with nonidentical variances is much complicated and is considered as future work.

While it is assumed that the instantaneous channels gain may be available at both the encoders and the decoders, we assume that the rate allocation vector is determined a-priori, based solely on the statistics. The rationale for this assumption is that the rate allocation vector is determined based on network considerations,

whereas the instantaneous state of each component of the network varies faster than the entire network can respond to the variations. Note that this assumption is practical as well when the power constraints must be satisfied in each encoding block (e.g. when we are under Federal Communications Commission (FCC) regulations).

Consider the rate allocation graph $\tilde{\mathcal{G}} = (\mathcal{V}, \mathcal{E}, \mathbf{r})$ (the graph that is obtained by assigning a rate $r_{i,j}$ for each link $(i,j)$ in $\mathcal{G}$). The optimal rate allocation vector that minimizes the probability of outage in the MAC network model while maintaining a multicast rate of $R_s$ is a solution of the following optimization problem:

$$\min_{\mathbf{f}, \mathbf{r}} \ \Pr \left( \bigcup_{j \in \mathcal{V}} \{ \mathbf{r}_j \notin \mathcal{V}_j^{\mathrm{ins}}(\mathbf{h}_j) \} \right) \tag{39}$$

$$\text{subject to}$$

$$0 \leq f_{i,j}^d \leq r_{i,j} \quad \forall (i,j) \in \mathcal{E}, d \in \mathcal{D}_s \tag{39a}$$

$$\sum_{i \in \mathcal{I}(j)} f_{i,j}^d - \sum_{i \in \mathcal{O}(j)} f_{j,i}^d = \begin{cases} 0 & j \notin \{s,d\} \\ R_s & j = d \end{cases} \tag{39b}$$

$$\forall j \in \mathcal{V} \backslash \{s\}, d \in \mathcal{D}_s,$$

where the flow constraints (39a)-(39b) guarantee that any feasible solution of (39) provides a minimum min-cut of at least $R_s$ between the source and each destination. Therefore, a multicast rate of $R_s$ is achievable by network coding, see Theorem 1. in [13].

Unfortunately, as was mentioned earlier, the computation of the probability of a common outage becomes extremely complicated in a MAC with more than 2 links. Therefore, we present a suboptimal solution to the rate allocation problem in the fading MAC network model. We relax the problem and instead of using the exact expression of the probability of common outage, we minimize an upper bound on the outage probability of a multiple access channel. To that end, consider the following lemma.

**Lemma 5.** *The probability of common outage of a MAC with $n$ i.i.d* $\mathrm{Rayleigh}(\upsilon)$ *channels is bounded by* $\Pr_{\mathrm{MAC_n}}^{\mathrm{out}} \leq 1 - \mathrm{e}^{-\lambda \left( 2^{\mathrm{Rn}} - 1 \right)}$.

*Proof:* For $n = 1$, we have a Rayleigh fading Gaussian channel with outage probability given by:

$$\Pr_{\mathrm{MAC_1}}^{\mathrm{out}} = 1 - \mathrm{e}^{-\lambda(2^{\mathrm{r_1}} - 1)}. \tag{40}$$

18

For $n = 2$, a simple computation yields that

$$1 - \Pr_{\text{MAC}_2}^{\text{out}} = e^{-\lambda\left(2^{r_1+r_2}-1\right)} \left(1 + \lambda\left(2^{r_1}-1\right)\left(2^{r_2}-1\right)\right)$$
$$\geq e^{-\lambda\left(2^{r_1+r_2}-1\right)}. \tag{41}$$

For $n \geq 3$, the claim follows from Theorem 4 and the fact that for any non-negative $x$, we have $G(x) \geq 1$. $\quad\square$

The probability of an outage in the MAC network model is given in equation (5). By assumption, all $h_{i,j}$ are independent of each other. Therefore, the probability of an outage in the MAC network model can be rewritten as

$$P_{\text{MAC}}^{\text{out}} = 1 - \prod_{j\in\mathcal{V}\setminus\{s\}} \left(1 - P_j^{\text{out}}\right). \tag{42}$$

Obviously, if $P_j^{\text{out}}$ is bounded above by $\tilde{P}_j^{\text{out}}$ the following holds

$$P_{\text{MAC}}^{\text{out}} \leq 1 - \prod_{j\in\mathcal{V}\setminus\{s\}} \left(1 - \tilde{P}_j^{\text{out}}\right). \tag{43}$$

Although the bound in Lemma 5 is weaker than the one we get from Theorem 4, we used the weaker bound to find a rate allocation vector for the outage MAC model. For every $j \in \mathcal{V}\setminus\{s\}$ denote

$$\tilde{R}_j := \sum_{i\in\mathcal{I}(j)} r_{i,j}. \tag{44}$$

Hence, from Lemma 5 we have

$$P_{\text{MAC}}^{\text{out}} \leq 1 - \prod_{j\in\mathcal{V}\setminus\{s\}} e^{-\lambda_j\left(2^{\tilde{R}_j}-1\right)}. \tag{45}$$

Finally, since $e^{-\lambda\left(2^{R_j}-1\right)}$ is log-concave the problem becomes computationally tractable:

$$\min_{\mathbf{f},\mathbf{r}} \sum_{j\in\mathcal{V}\setminus\{s\}} \lambda_j 2^{\tilde{R}_j} \tag{46}$$

$$\text{subject to}$$

$$0 \leq f_{i,j}^d \leq r_{i,j} \quad \forall(i,j)\in\mathcal{E}, d\in\mathcal{D}_s \tag{46a}$$

$$\sum_{i\in\mathcal{I}(j)} f_{i,j}^d - \sum_{i\in\mathcal{O}(j)} f_{j,i}^d = \begin{cases} 0 & j\notin\{s,d\} \\ R_s & j=d \end{cases} \tag{46b}$$

$$\forall j\in\mathcal{V}\setminus\{s\}, d\in\mathcal{D}_s.$$

19

## 5. Distributed solution for MAC network model

In the previous section the rate allocation vector for the MAC network model was given as a solution to the convex optimization problem (46). This problem can easily be solved by a standard convex optimization technique in a centralized fashion. However, the centralized solution requires full knowledge of the network topology and statistics. In this section we discuss how to distributively solve (46).

As pointed out in e.g., [13, 35, 36], if certain conditions are satisfied, convex optimization problems may be solved distributively by a continuous time primal-dual method. This method can be described as follows. The optimization is studied through its Lagrangian where the primal and dual variables are updated simultaneously by a set of gradient laws (dynamic system). These laws define a trajectory in the direction of the respective partial gradients, starting from some initial point. The dynamic system is stated such that the saddle points of the Lagrangian are equilibrium points. Hence, if a strong duality holds for the original convex optimization problem, the algorithm stops updating the variables when it reaches the optimal solution. It is worth mentioning that in contrast to gradient method, in which convergence is guaranteed for convex problems from any initial point (see e.g., §9 in [37]), the asymptotic behavior of dynamic systems is not immediate in the general case (even though the problem is convex). In other words, convergence to an equilibrium point is not guaranteed in the general case. In this type of problem the existence of Lyapunov functions is used to prove the stability of the equilibrium points. There is no general technique for the construction of these functions. However, in some specific cases the construction of Lyapunov functions is known (see e.g., [38]).

As can be easily verified, the cost function in (46) is not strictly convex (and also is not separable in the decision variables $\mathbf{f}$ and $\mathbf{r}$). In problems with a non-strictly convex cost function, it is possible to have more than one optimum point. Hence, in this case the standard primal-dual solution may not converge. In [36] a modified primal dual gradient method was derived for non-strictly convex problems. In that method the solution will converge to one of the optimal points by modifying the constraint set of the convex optimization problem. Following [36],

we suggest the following modified convex optimization problem:

$$\min_{\mathbf{f},\mathbf{r}} \sum_{j \in \mathcal{V}\backslash\{s\}} \lambda_j 2^{\tilde{R}_j} \tag{47}$$

subject to

$$\phi\left(-f_{i,j}^d\right) \leq 0 \quad \forall (i,j) \in \mathcal{E}, d \in \mathcal{D}_s \tag{47a}$$

$$\phi\left(f_{i,j}^d - r_{i,j}\right) \leq 0 \quad \forall (i,j) \in \mathcal{E}, d \in \mathcal{D}_s \tag{47b}$$

$$\phi\left(q_j^d\right) \leq 0 \qquad \forall j \in \mathcal{V}\backslash\{s\}, d \in \mathcal{D}_s \tag{47c}$$

$$\phi\left(-q_j^d\right) \leq 0 \qquad \forall j \in \mathcal{V}\backslash\{s\}, d \in \mathcal{D}_s, \tag{47d}$$

where $\tilde{R}_j$ was defined in (44), $\phi(x) = e^x - 1$ and for all $j \in \mathcal{V}\backslash\{s\}, d \in \mathcal{D}_s$

$$q_j^d := \sum_{i \in \mathcal{I}(j)} f_{i,j}^d - \sum_{i \in \mathcal{O}(j)} f_{j,i}^d - \psi_j^d \tag{48}$$

$$\psi_j^d := \begin{cases} 0 & j \notin \{s,d\} \\ R_s & j = d \end{cases}. \tag{49}$$

Note that theorem 11 in [36] that guarantees convergence for the corresponding dynamic system was proved under the assumption that Slater's condition holds for the modified convex optimization problem. It is easy to see that Slater's condition does not hold for (47). However, it can be verified that their proofs remain valid as they are under any other constraint qualification (i.e., whenever strong duality for the modified optimization holds). In Appendix Appendix B, we show that strong duality holds for (47). Denote by $\rho_{i,j}^d$, $w_{i,j}^d$, $\varphi_j^d$ and $\mu_j^d$ the dual variables of (47) and define

$$\Delta_{i,j}^d := -\varphi_j^d e^{q_j^d} + I_{i \neq s}\varphi_i^d e^{q_i^d} + \mu_j^d e^{-q_j^d} - I_{i \neq s}\mu_i^d e^{-q_i^d}, \tag{50}$$

where $I_{i \neq s} = 1$ if $i \neq s$ and zero otherwise. The primal-dual gradient laws for

(47) are given by

$$\dot{r}_{i,j} = \tau_{i,j}\left(-\lambda_j e^{\tilde{R}_j} + \sum_{d\in\mathcal{D}_s} w_{i,j}^d e^{f_{i,j}^d - r_{i,j}}\right) \tag{51}$$

$$\dot{f}_{i,j}^d = k_{i,j}^d\left(\rho_{i,j}^d e^{-f_{i,j}^d} - w_{i,j}^d e^{f_{i,j}^d - r_{i,j}} + \Delta_{i,j}^d\right) \tag{52}$$

$$\dot{\rho}_{i,j}^d = \alpha_{i,j}^d\left[e^{-f_{i,j}^d} - 1\right]_{\rho_{i,j}^d}^+ \tag{53}$$

$$\dot{w}_{i,j}^d = \theta_{i,j}^d\left[e^{f_{i,j}^d - r_{i,j}} - 1\right]_{w_{i,j}^d}^+ \tag{54}$$

$$\dot{\varphi}_j^d = \beta_j^d\left[e^{q_j^d} - 1\right]_{\varphi_j^d}^+ \tag{55}$$

$$\dot{\mu}_j^d = \gamma_j^d\left[e^{-q_j^d} - 1\right]_{\mu_j^d}^+, \tag{56}$$

where $\tau_{i,j}, k_{i,j}^d, \theta_{i,j}^d, \alpha_{i,j}^d, \beta_j^d$ and $\gamma_j^d$ are some positive scalars and for any two scalars $x$ and $p$

$$[x]_p^+ = \begin{cases} 0 & x < 0, p < 0 \\ x & \text{otherwise.} \end{cases}$$

The dynamic (51)-(56) can be distributively implemented by associating a processor for each node in the network, excluding the source node. Each node $j$'s processor keeps track of the variables $\varphi_j^d$ and $\mu_j^d$ as well as the variables associated with node $j$'s ingoing links (i.e., the links in $\{(i,j) : i \in \mathcal{I}(j)\}$). Note that message passing is required only between direct neighbors.

## 6. Simulation results

In this section the probability of outage of the suboptimal algorithm for the MAC network model is presented. In the simulation we consider the networks shown in Fig. 1 where it was assumed that all links are i.i.d $\mathrm{Rayleigh}(\upsilon)$ channels, with $\upsilon = 1$. We solved (46) for various values of SNR $= \frac{P}{\sigma^2}$ and the results are shown in Fig. 2. The lower and upper bounds for the outage probability were obtained by calculating (13) and (26) (respectively) for each MAC associated with node $j$ if $\mathcal{I}(j) \geq 3$ and the exact expression of the outage probability for each receiver $j$ with $\mathcal{I}(j) \leq 2$. As can be seen, up to 6 bits/sec/Hz the bounds are quite tight.

22

We compared the performance of the fading MAC network model to the performance of the non-naive TDMA model. The optimal rate allocation scheme that minimizes the probability of outage for the TDMA based model was derived in [1]. The results for the non-naive TDMA model are shown in Fig. 3. Although for low multicast rate demands there is no significant gain in preferring the MAC network model over the Non-naive TDMA model, when we have a demand for a high multicast rate the MAC network model outperforms the TDMA based model. To emphasize this result, note that in the non-naive TDMA model (see Fig. 3) we achieved a multicast rate of $R_s \approx 5.5$ bits/sec/Hz with probability of outage of $P^{out} \approx 0.1$ with SNR $= 30dB$ whereas we obtained the same results (i.e., $R_s \approx 5.5$, $P^{out} \approx 0.1$) in the MAC network model with SNR $= 25dB$.

Finally, we simulated a discrete time version of the distributed algorithm shown in section 5. In this version we consider time steps $m = 1, 2, \ldots$ and the derivatives were replaced by differences. The scalars $\tau_{i,j}, k_{i,j}^d, \theta_{i,j}^d, \alpha_{i,j}^d, \beta_j^d$ and $\gamma_j^d$ can be thought as step sizes. We did not optimized these step sizes and they were randomly chosen at the initiation of the simulation ($\tau_{i,j}$ and $k_{i,j}^d$ where about $10$ times larger than the other step sizes). During the simulation, we considered the network sown in Fig. 1 and it was assumed that node $5$ was the source node and that the destinations were $\mathcal{D}_s = \{1, 4, 8, 10\}$. The convergence of the algorithm is shown in Fig. 4.

## 7. Conclusions

In this paper we studied the rate allocation problem for multicasting over slow Rayleigh fading channels using network coding. A rate allocation scheme based solely on the statistics of the channels was presented. In the MAC network model, where the network is treated as a collection of slowly Rayleigh fading multiple access channels, we proposed a suboptimal scheme as the solution to a convex optimization problem. This suboptimal solution is based on an upper bound on the probability of outage of a fading multiple access channel. A primal-dual gradient algorithm was derived to solve the problem distributively. In the simulation results, it is shown that the MAC network model outperforms the TDMA based model. The paper provides a practical solution to networks with slow fading channels in which long delays are unacceptable (e.g., in video streaming), with the objective of minimizing outage events throughout the network. As a potential future works, one should consider to derive a bound on the outage probability in the non i.i.d case and extend the problem of statistic-based rate allocation scheme to deal with other than Rayleigh fading model (e.g., Rician or Nakagami models).
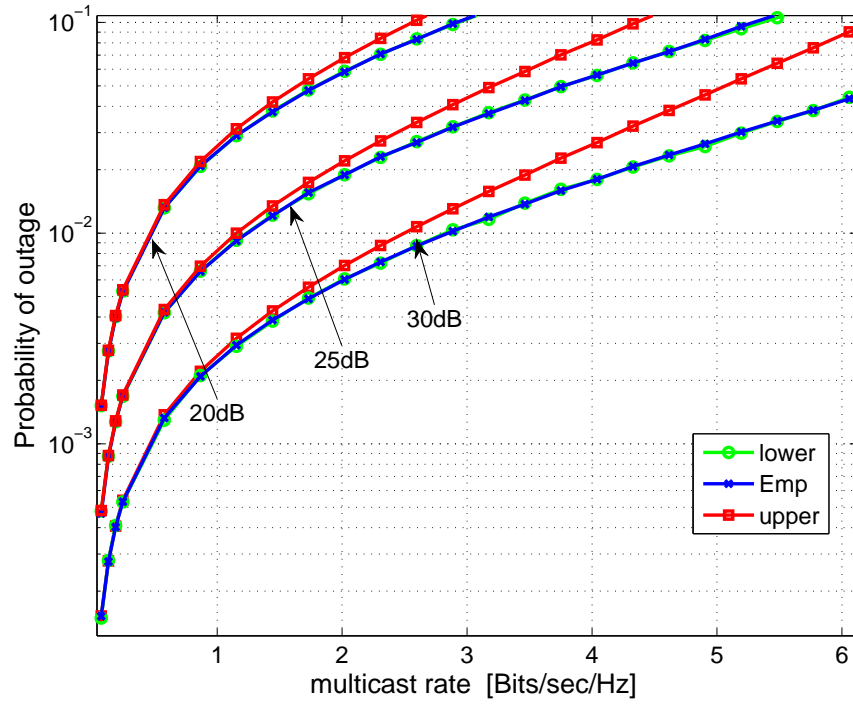
Figure 2: The probability of outage of the fading MAC network model for various values of SNR for the network shown in Fig. 1. It was assumed that node 5 was the source node and that the destinations were $\mathcal{D}_s = \{1, 4, 8, 10\}$.
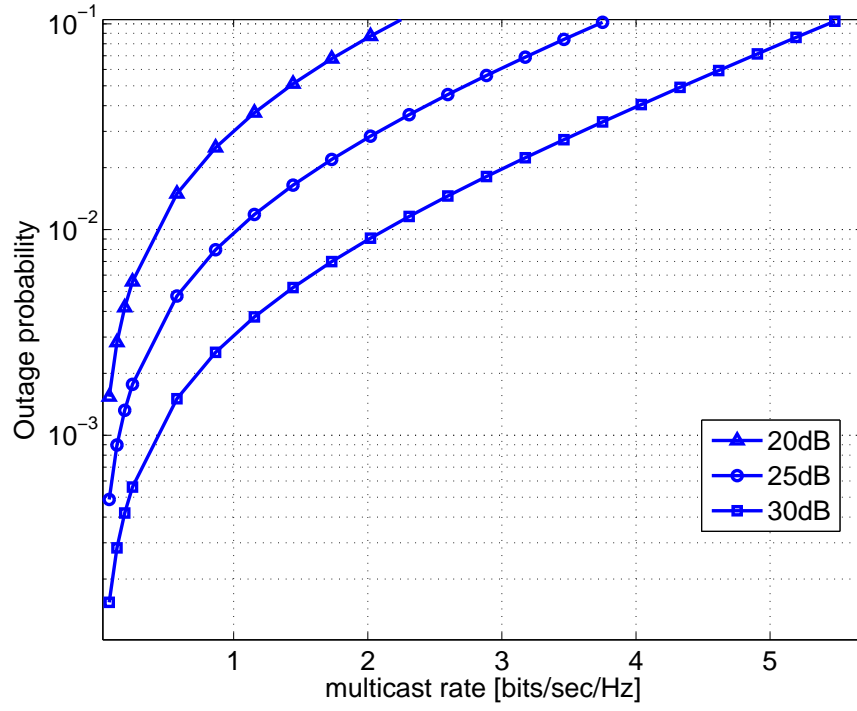
Figure 3: The probability of outage in the non-naive TDMA for various values of SNR for the network shown in Fig. 1. It was assumed that node 5 was the source node and that the destinations were $\mathcal{D}_s = \{1, 4, 8, 10\}$.
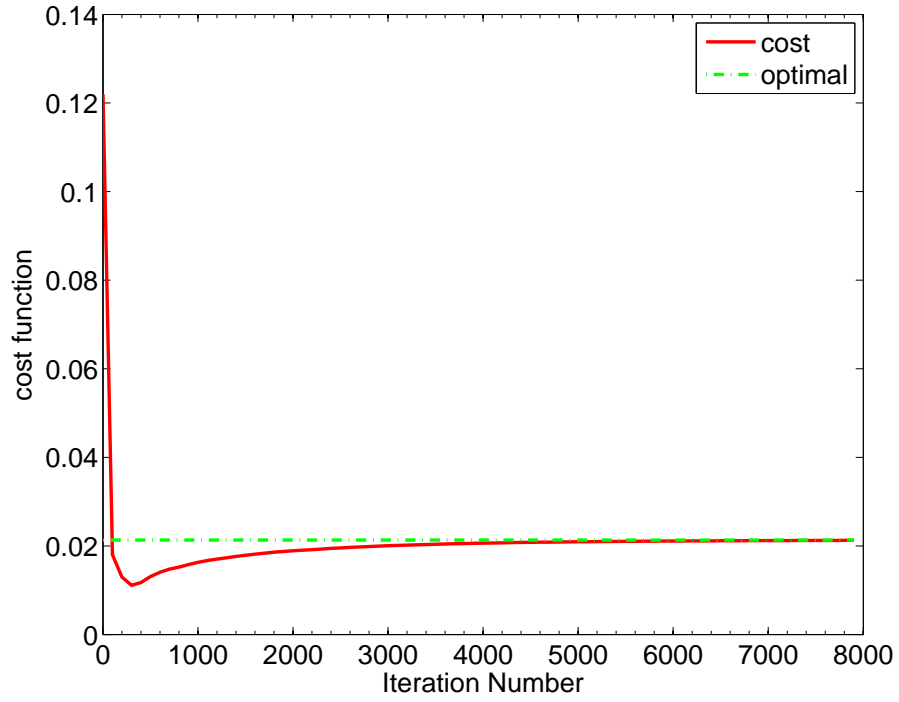
Figure 4: The convergence of the distributed algorithm for the network shown in Fig. 1. It was assumed that node $5$ was the source node and that the destinations were $\mathcal{D}_s = \{1, 4, 8, 10\}$. The dashed line represents the optimal solution of (46) and the solid line represents the value of the cost function in (46) over time.

### Appendix A. A recursion for computing the joint distribution of linear combinations of exponential random variables

In this section we present a new version of a recursion that appeared as equation (17) in [31]. The lemma in [31] gives a recursion for the computation of the joint distribution of linear combinations of spacings of uniform distribution. The authors in [31] remarked that the recursion remain valid as well for computing the joint distribution of linear combinations of exponential random variables. This is inaccurate and we revise the result to handle exponential random variables. To that end, we need the following notations. Let $\Psi = \{i_1, i_2, \ldots, i_k\}$  $k \geq 1$ be a set of indices of columns of a matrix $\mathbf{A}$ such that $i_\ell < i_{\ell+1}$ for all $\ell$ and let $\mathbf{A}_{-\Psi(m)}$ denote the submatrix of $A$ constructed by deleting the columns of $\mathbf{A}$ indexed by $\{i_1, i_2, \ldots, i_m\}$.

**Lemma 6.** *Let $z_1, z_2, \cdots, z_{N+1}$ be $(N+1)$ i.i.d exponential random variables with expectation $E\{z_i\} = 1$. Let $\Psi = \{i_1, i_2, \ldots, i_k\}$, $k \geq 1$ be a set of indices of identical columns of matrix $\mathbf{A}$ (without loss of generality $i_\ell < i_{\ell+1}$ for all $\ell$). If there exists a row $r$ in $\mathbf{A}$ such that a) $a_{r,i} > 0$ for $i \in \Psi$, b) $a_{r,i} = 0$ for $i \notin \Psi$ and c) $b_r \geq 0$, the following recursion holds*

$$\Pr\left(\mathbf{A}\mathbf{z} > \lambda\mathbf{b}\right) = \sum_{m=0}^{k-1} \frac{1}{m!}(\lambda\delta)^m e^{-\lambda\delta} \Pr\left(\mathbf{A}_{-\Psi(m)}^{r*}\mathbf{z} > \lambda\mathbf{c}\right), \qquad (A.1)$$

*where $\delta = \frac{b_r}{a_{r,i_1}}$ and $\mathbf{c} = \mathbf{b}^{r*} - \delta\mathbf{a}_{i_1}^{r*}$.*

*Proof:* As pointed out in [31], since the expression $\mathbf{A}\mathbf{z} > \lambda\mathbf{b}$ stands for a conjunction of inequalities involving i.i.d random variables, $\Pr\left(\mathbf{A}\mathbf{z} > \lambda\mathbf{b}\right) = \Pr\left(\boldsymbol{\pi}\mathbf{A}\mathbf{z} > \lambda\boldsymbol{\pi}\mathbf{b}\right)$ and $\Pr\left(\mathbf{A}\mathbf{z} > \lambda\mathbf{b}\right) = \Pr\left(\mathbf{A}\boldsymbol{\pi}\mathbf{z} > \lambda\boldsymbol{\pi}\mathbf{b}\right)$ hold for any permutation matrix $\boldsymbol{\pi}$ with the appropriate dimensions. Therefore, without loss of generality, we assume that $r = 1$ and $\Psi = \{1, 2, \ldots, k\}$ (See the illustration of such a matrix in Fig. A.5). Under these assumptions, the first inequality in $\mathbf{A}\mathbf{z} > \lambda\mathbf{b}$ is

$$D = \{a_{1,1}\sum_{\ell=1}^{k} z_\ell > \lambda b_1\}. \qquad (A.2)$$

Clearly,

$$\Pr\left(\mathbf{A}\mathbf{z} > \lambda\mathbf{b}\right) = \Pr\left(D \cap \{\mathbf{A}^*\mathbf{z} > \lambda\mathbf{b}^*\}\right). \qquad (A.3)$$

$$
\overset{R\times(N+1)}{A} =
\begin{bmatrix}
\overbrace{a_{1,1} \quad \cdots \quad a_{1,1}}^{k} & \overbrace{0 \quad \cdots \quad 0}^{N+1-k} \\
a_{2,1} \quad \cdots \quad a_{2,1} & a_{2,k+1} \quad \cdots \quad a_{2,N+1} \\
\vdots \quad \ddots \quad \vdots & \vdots \quad \ddots \quad \vdots \\
a_{R,1} \quad \cdots \quad a_{R,1} & a_{R,k+1} \quad \cdots \quad a_{R,N+1}
\end{bmatrix}
$$

Figure A.5: An illustration of a matrix $\mathbf{A}$ that satisfies the assumptions in Lemma 6, where $\Psi = \{1, 2, \ldots, k\}$ and $r = 1$.

The event $D$ can be written as the union of disjoint events $D = \cup_{m=0}^{k-1} D_m$, where

$$D_m = \{\sum_{\ell=1}^{m} z_\ell \le \lambda\delta \le \sum_{\ell=1}^{m+1} z_\ell\}. \tag{A.4}$$

Therefore,

$$\Pr\left(\mathbf{Az} > \lambda\mathbf{b}\right) = \sum_{m=0}^{k-1} \Pr\left(D_m\right)\Pr\left(\mathbf{A}^*\mathbf{z} > \lambda\mathbf{b}^*|D_m\right). \tag{A.5}$$

For $m < k$, the $r$th inequality in $\mathbf{Az} > \lambda\mathbf{b}$ can be rewritten as

$$\{\sum_{\ell=1}^{N+1} a_{r,\ell}z_\ell > \lambda b_r\} = \{\sum_{\ell=1}^{m+1} a_{r,\ell}z_\ell + \sum_{\ell=m+2}^{N+1} a_{r,\ell}z_\ell > \lambda b_r\} =$$

$$\{a_{r,1}\sum_{\ell=1}^{m+1} z_\ell + \sum_{\ell=m+2}^{N+1} a_{r,\ell}z_\ell - a_{r,1}\lambda\delta > \lambda b_r - a_{r,1}\lambda\delta\} =$$

$$\{a_{r,1}\left(\sum_{\ell=1}^{m+1} z_\ell - \lambda\delta\right) + \sum_{\ell=m+2}^{N+1} a_{r,\ell}z_\ell - > (b_r - \lambda\delta a_{r,1})\} =$$

$$\{\left[a_{r,1}, a_{r,m+2}, \cdots, a_{r,N+1}\right]^T \mathbf{T}^m > (b_r - \lambda\delta a_{r,1})\}, \tag{A.6}$$

where

$$\mathbf{T}^m = \left[\sum_{\ell=1}^{m+1} z_\ell - \lambda\delta, z_{m+2}, \cdots, z_{N+1}\right]^T. \tag{A.7}$$

Therefore,

$$\Pr\left(\mathbf{A}^*\mathbf{z} > \lambda\mathbf{b}^*|D_m\right) = \Pr\left(\mathbf{A}^*_{(-m)}\mathbf{T}^m > \lambda\left(\mathbf{b}^* - \delta\mathbf{a}^*\right)|D_m\right) \tag{A.8}$$

In the following we show that the event $\mathbf{T}^m|D_m$ has the same distribution as $(n+1-m)$ i.i.d exponential random variables. Obviously, $D_m$ is independent with $(\mathbf{T}^m)^*$. Therefore, in order to show that $\mathbf{T}^m|D_m$ has the same distribution as $(n+1-m)$ i.i.d exponential random variables, it suffices to show that $\sum_{\ell=1}^{m+1} z_\ell - \lambda\delta|D_m$ has the same distribution as $z_{m+1}$. Note that

$$\left\{\sum_{\ell=1}^{m+1} z_\ell - \lambda\delta|D_m\right\} = \left\{\sum_{\ell=1}^{m+1} z_\ell - \lambda\delta|0 \le \lambda\delta - \sum_{\ell=1}^{m} z_\ell \le z_{m+1}\right\}. \tag{A.9}$$

29

Therefore,

$$\Pr\left(\sum_{\ell=1}^{m+1} z_\ell - \lambda\delta > x \Big| 0 \le \lambda\delta - \sum_{\ell=1}^{m} z_\ell \le z_{m+1}\right) = \tag{A.10}$$

$$\Pr\left(z_{m+1} > x - \sum_{\ell=1}^{m} z_\ell + \lambda\delta \Big| 0 \le \lambda\delta - \sum_{\ell=1}^{m} z_\ell \le z_{m+1}\right). \tag{A.11}$$

Due to the memoryless property[3] of the exponential random variable we can write

$$\Pr\left(z_{m+1} > x - \sum_{\ell=1}^{m} z_\ell + \lambda\delta \Big| z_{m+1} \ge \lambda\delta - \sum_{\ell=1}^{m} z_\ell \ge 0\right) = \tag{A.12}$$

$$\Pr\left(z_{m+1} > x - \sum_{\ell=1}^{m} z_\ell + \lambda\delta - \left(\lambda\delta - \sum_{\ell=1}^{m} z_\ell\right)\right) = \tag{A.13}$$

$$\Pr\left(z_{m+1} > x\right). \tag{A.14}$$

Therefore, (A.8) becomes

$$\Pr\left(\mathbf{A}^*\mathbf{z} > \lambda\mathbf{b}^* | D_m\right) = \Pr\left(\mathbf{A}^*_{(-m)}\mathbf{z} > \lambda\left(\mathbf{b}^* - \delta\mathbf{a}^*\right)\right). \tag{A.15}$$

Combining (A.5) and (A.15) yields

$$\Pr\left(\mathbf{A}\mathbf{z} > \lambda\mathbf{b}\right) = \sum_{m=0}^{k-1} \Pr\left(D_m\right) \Pr\left(\mathbf{A}^*_{(-m)}\mathbf{z} > \lambda\left(\mathbf{b}^* - \delta\mathbf{a}^*\right)\right). \tag{A.16}$$

In order to complete the recursion we need an explicit expression of $\Pr(D_m)$.

$$\Pr(D_m) = \Pr\left(\sum_{\ell=1}^{m} z_\ell \le \lambda\delta \le \sum_{\ell=1}^{m+1} z_\ell\right) \tag{A.17}$$

$$= 1 - \Pr\left(\sum_{\ell=1}^{m} z_\ell > \lambda\delta\right) - \Pr\left(\lambda\delta > \sum_{\ell=1}^{m+1} z_\ell\right) \tag{A.18}$$

$$= \Pr\left(\sum_{\ell=1}^{m} z_\ell \le \lambda\delta\right) - \Pr\left(\sum_{\ell=1}^{m+1} z_\ell \le \lambda\delta\right). \tag{A.19}$$

---

[3]The memoryless property of an exponential variable means that for any $a, b \ge 0$, we have that $\Pr\left(Z > a + b | Z > a\right) = \Pr\left(Z > b\right)$, where $Z$ is exponential random variable.

Since $\sum_{\ell=1}^{m} z_\ell$ follows the $\mathrm{Erlang}(m, 1)$ distribution[4], we have that

$$\Pr(D_m) = \frac{1}{m!}(\lambda\delta)^m e^{-\lambda\delta}. \qquad (A.20)$$

The claim now follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

## Appendix B. Strong duality of (47)

**Lemma 7.** *Strong duality holds for (47).*

*Proof:* Consider the following optimization problem

$$\min_{\mathbf{f},\mathbf{r}} \sum_{j \in \mathcal{V}\setminus\{s\}} \lambda_j 2^{\tilde{R}_j} \qquad (B.1)$$

subject to

$$\phi\left(-f_{i,j}^d\right) \leq 0 \quad \forall (i,j) \in \mathcal{E}, d \in \mathcal{D}_s \qquad (B.1a)$$

$$\phi\left(f_{i,j}^d - r_{i,j}\right) \leq 0 \quad \forall (i,j) \in \mathcal{E}, d \in \mathcal{D}_s \qquad (B.1b)$$

$$\sum_{i \in \mathcal{I}(j)} f_{i,j}^d - \sum_{i \in \mathcal{O}(j)} f_{j,i}^d = \begin{cases} 0 & j \notin \{s,d\} \\ R_s & j = d \end{cases} \qquad (B.1c)$$

$$\forall j \in \mathcal{V}\setminus\{s\}, d \in \mathcal{D}_s.$$

Note that the refined Slater's condition holds for (B.1) and therefore (B.1) has zero duality gap [37], but it does not hold for (47). Obviously, the feasible sets of (46), (47) and (B.1) are all the same and therefore they have identical optimal solutions. We need to show that solving (47) through its Lagrangian yields the same solution. Denote the primal variables by $\mathbf{x} = (\mathbf{f}, \mathbf{r})$ and the dual variables by $\boldsymbol{\zeta}$. Denote the Lagrangians of (47) and (B.1) by $L(\mathbf{x}, \boldsymbol{\zeta})$, $L_M(\mathbf{x}, \boldsymbol{\zeta})$, respectively. The dual function of (47) is given by $q(\boldsymbol{\zeta}) = \min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\zeta})$. Assume that there exists $\tilde{\mathbf{x}}(\boldsymbol{\zeta}) = (\tilde{\mathbf{f}}, \tilde{\mathbf{r}})$ that is a minimizer of $L(\mathbf{x}, \boldsymbol{\zeta})$ that does not obey the flow conservation constraint (46b). Therefore, there exists node $j \in \mathcal{V}\setminus\{s\}$ such that

---

[4]The CDF of $\mathrm{Erlang}(m, \lambda)$ distributed random variable $Y$ is given by $\quad F_Y(y) = 1 - \sum_{\ell=0}^{m-1} \frac{1}{\ell!}(\lambda y)^\ell e^{-\lambda y}$.

$q_j^d \neq 0$. Hence, we have either $\phi(q_j^d) > 0$ or $\phi(-q_j^d) > 0$. Without loss of generality assume that $\phi(q_j^d) > 0$. In that case, since the cost function in (47) is bounded below by 0 we can always choose $\boldsymbol{\zeta}$ such that the dual solution $q(\boldsymbol{\zeta})$ is infinity (by setting all $\lambda_i$ to zero except the one with the positive coefficient $\phi(q_j^d) > 0$). This contradicts the feasibility of the primal (47). We conclude by noting that $L(\mathbf{x}, \boldsymbol{\zeta}) = L_M(\mathbf{x}, \boldsymbol{\zeta})$ for any $\mathbf{x}$ obeys the flow conservation constraint (46b). $\square$

## References

[1] A. Zanko, A. Leshem, E. Zehavi, Network coding for multicasting over Rayleigh fading multi access channels, in: Proc. 22nd Intl. Conference on Computer Communication and Networks (ICCCN), Nassau, Bahamas.

[2] R. Ahlswede, N. Cai, S.-Y. Li, R. Yeung, Network information flow, IEEE Trans. Information Theory, 46 (2000) 1204–1216.

[3] S.-Y. Li, R. Yeung, N. Cai, Linear network coding, IEEE Trans. Information Theory, 49 (2003) 371–381.

[4] S. Jaggi, P. Chou, J. Kamal, Low complexity algebraic multicast network codes, in: IEEE International Symposium on Information Theory (ISIT, pp. 36–8.

[5] A. Barbero, O. Ytrehus, Heuristic algorithms for small field multicast encoding, in: Information Theory Workshop, 2006. ITW '06 Chengdu. IEEE, pp. 428–432.

[6] S. Jaggi, P. Sanders, P. Chou, M. Effros, S. Egner, K. Jain, L. Tolhuizen, Polynomial time algorithms for multicast network code construction, IEEE Trans. Information Theory, 51 (2005) 1973–1982.

[7] T. Ho, M. Medard, R. Koetter, D. Karger, M. Effros, J. S., B. Leong, A random linear network coding approach to multicast, IEEE Trans. Information Theory, 52 (2006) 4413–4430.

[8] C. Fragouli, E. Soljanin, Network coding fundamentals, Foundations and Trends in Networking (2007).

[9] T. Ho, D. Lun, Network coding: an introduction, Cambridge University Press, 2008.

[10] D. Palomar, M. Chiang, A tutorial on decomposition methods for network utility maximization, IEEE Journal on Selected Areas in Communications 24 (2006) 1439–1451.

[11] M. Chiang, S. Low, A. Calderbank, J. Doyle, Layering as optimization decomposition: a mathematical theory of network architectures, Proceedings of the IEEE 95 (2007) 255–312.

[12] D. Lun, M. Medard, T. Ho, R. Koetter, Network coding with a cost criterion, Proc. IEEE Int. Symp. Information Theory and Its Applications (ISITA 2004), Parma, Italy (2004).

[13] D. Lun, N. Ratnakar, M. Medard, R. Koetter, D. Karger, T. Ho, E. Ahmed, F. Zhao, Minimum-cost multicast over coded packet networks, IEEE Trans. Information Theory, 52 (2006) 2608 –2623.

[14] D. Lun, M. Medard, R. Koetter, M. Effros, On coding for reliable communication over packet networks, Physical Communication 1 (2008) 3 – 20.

[15] P. Chou, Y. Wu, K. Jain, Practical network coding 41 (2003) 40–49.

[16] Y. Cui, Y. Xue, K. Nahrstedt, Optimal distributed multicast routing using network coding: theory and applications, SIGMETRICS Perform. Eval. Rev. 32 (2004) 47–49.

[17] S. Bhadra, S. Shakkottai, P. Gupta, Min-cost selfish multicast with network coding, IEEE Trans. Information Theory, 52 (2006) 5077 –5087.

[18] Y. Wu, M. Chiang, S. Kung, Distributed utility maximization for network coding based multicasting: a critical cut approach, in: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2006 4th International Symposium on, pp. 1 – 6.

[19] Y. Xi, E. Yeh, Distributed algorithms for minimum cost multicast with network coding, IEEE/ACM Trans. Netw. 18 (2010) 379 –392.

[20] D. Lun, N. Ratnakar, R. Koetter, M. Medard, E. Ahmed, H. Lee, Achieving minimum-cost multicast: a decentralized approach based on network coding, in: INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, volume 3, pp. 1607 – 1617 vol. 3.

[21] J. Yuan, Z. Li, W. Yu, B. Li, A cross-layer optimization framework for multihop multicast in wireless mesh networks, IEEE Journal on Selected Areas in Communications, 24 (2006) 2092–2103.

[22] Y. Sagduyu, A. Ephremides, On joint MAC and network coding in wireless Ad Hoc networks, IEEE Trans. Information Theory 53 (2007) 3697–3713.

[23] R. Niati, A. Banihashemi, T. Kunz, Throughput and energy optimization in wireless networks: joint MAC scheduling and network coding, IEEE Trans. Vehicular Technology 61 (2012) 1372–1382.

[24] L. Ozarow, S. Shamai, A. Wyner, Information theoretic considerations for cellular mobile radio, IEEE Trans. Vehicular Technology 43 (1994) 359 –378.

[25] A. Goldsmith, P. Varaiya, Capacity of fading channels with channel side information, IEEE Trans. Information Theory 43 (1997) 1986 –1992.

[26] S. Shamai, A broadcast strategy for the Gaussian slowly fading channel, in: Information Theory. 1997. Proceedings., 1997 IEEE International Symposium on, p. 150.

[27] D. Tse, S. Hanly, Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities, IEEE Trans. Information Theory 44 (1998) 2796 –2815.

[28] S. Hanly, D. Tse, Multiaccess fading channels. II. delay-limited capacities, IEEE Trans. Information Theory 44 (1998) 2816–2831.

[29] L. Li, N. Jindal, A. Goldsmith, Outage capacities and optimal power allocation for fading multiple-access channels, IEEE Trans. Information Theory 51 (2005) 1326–1347.

[30] A. El Gamal, Y. Kim, Network information theory, Cambridge University Press, 2011.

[31] F. Huffer, C. Lin, Computing the joint distribution of general linear combinations of spacings or exponential variates, Statistica Sinica 11 (2001) 1141–1158.

[32] B. Rao, R. K. Ganti, R. Aravind, Outage probability and goodput with ARQ in multiple access channels, ITG-Fachbericht-SCC 2013 (2013).

[33] F. Huffer, Divided differences and the joint distribution of linear combinations of spacings, Journal of applied probability (1988) 346–354.

[34] H. Van Khuong, H. Kong, General expression for PDF of a sum of independent exponential random variables, IEEE Communications Letters 10 (2006) 159–161.

[35] R. Srikant, The mathematics of Internet congestion control, Birkhauser Boston, 2004.

[36] D. Feijer, F. Paganini, Stability of primal-dual gradient dynamics and applications to network optimization, Automatica 46 (2010) 1974–1981.

[37] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2004.

[38] H. Khalil, Nonlinear systems, Prentice hall Upper Saddle River, 2002.