

Resource Allocation Strategies for Network-Coded Video Broadcasting Services Over LTE-Advanced

Andrea Tassi, Chadi Khirallah, Dejan Vukobratović, Francesco Chiti, John S. Thompson, and Romano Fantacci

Abstract—Video service delivery over 3GPP Long Term Evolution-Advanced (LTE-A) networks is gaining momentum with the adoption of the evolved Multimedia Broadcast Multicast Service (eMBMS). In this paper, we address the challenge of optimizing the radio resource allocation process so that heterogeneous groups of users, according to their propagation conditions, can receive layered video streams at predefined and progressively decreasing service levels matched to respective user groups. A key aspect of the proposed system model is that video streams are delivered as eMBMS flows that utilize the random linear network coding (NC) principle. Furthermore, the transmission rate and NC scheme of each eMBMS flow are jointly optimized. The simulation results show that the proposed strategy can exploit user heterogeneity to optimize the allocated radio resources while achieving desired service levels for different user groups.

Index Terms—Evolved Multimedia Broadcast and Multicast Service (eMBMS), Long-Term Evolution (LTE)-Advanced (LTE-A), multimedia communication, network coding (NC), resource allocation.

I. INTRODUCTION

Video content delivery over fourth generation (4G) mobile cellular networks, namely, Long-Term Evolution (LTE) and LTE-Advanced (LTE-A), is estimated to exponentially grow due to the surge in demand for bandwidth-intensive applications based on video streaming [1].

To support video multicasting and broadcasting, LTE and LTE-A offer the functionality of managing point-to-multipoint (PtM) communications through the evolved Multimedia Broadcast and Multicast Service (eMBMS) [2]. Two transmission modes have been defined [3]: the single cell (SC) and the single frequency network (SFN) eMBMS. Unlike SFN-eMBMS, the SC-eMBMS design allows each eNodeB (eNB) to independently select service delivery parameters. Regardless of the transmission mode, it is infeasible that a large number of user equipment (UE) devices can explicitly provide feedback to the eNB about their propagation conditions for the eMBMS services.

This paper deals with SC-eMBMS deployment, where the eNB delivers broadcast video services to all UE devices that belong to one cell. In particular, the main goal of this paper is to define an efficient resource allocation strategy suitable for scalable video broadcasting. We consider video flows encoded by using the H.264 Scalable Video

Coding (H.264/SVC) [4] codec, which provides video streams formed by multiple video layers, namely, the *base layer* and several *enhancement layers*. The base layer provides a basic reconstruction quality that is gradually improved by decoding subsequent layers [5].

One of the key points of resource allocation strategies for PtM communications is the possibility of exploiting user heterogeneity (in terms of propagation conditions) to maximize the level of satisfaction of each user. Several allocation strategies have been proposed for multicast/broadcast communications over orthogonal frequency-division multiple access (OFDMA)-based systems [6]. Among them, we have the least channel gain (LCG) strategy, which delivers services such that they can be recovered by UE devices experiencing the worst propagation conditions in the network [7], [8]. In this case, the maximum PtM service transmission rate is limited by UE devices experiencing the worst reception quality.

Multirate transmission (MrT) schemes promise to overcome this problem by 1) splitting the set of users targeted by the delivered PtM service into subsets and 2) differentiating the service delivery into subflows (one per subset), which are optimized according to the propagation conditions of each subset [9], [10]. Although MrT schemes can better exploit user heterogeneity, they usually assume that UE devices provide feedback to the transmitting node reporting their propagation conditions [9], [11], [12] or positioning information [10]. In addition, these schemes do not address the resource allocation problem by taking into account the tight constraints imposed by 3GPP on the scheduling and structure of LTE radio frames containing eMBMS subframes [2]. It is worth noting that, although there are allocation strategies that aim to minimize the average/instantaneous user dissatisfaction [9], [11], none had proposed a resource allocation framework that ensures a predefined service level to a certain group of users.

Reliable packet-loss-resilient multimedia service broadcasting via eMBMS has been considered a challenging problem [13]. In particular, 3GPP has foreseen the adoption of application-level forward error correction (AL-FEC) schemes based on raptor codes to improve the reliability of broadcast and multicast eMBMS communications [3]. However, a major concern about AL-FEC coding strategies is that they lead to a large communication delay [14]. To overcome that issue, link-level random network coding (RNC)-based strategies have been recently proposed as a valuable and affordable (from a computational point of view) alternative to fountain code-based AL-FEC schemes [14], [15].

Several works dealing with the definition and optimization of network coding (NC) applications to data broadcasting over a multihop network have been proposed [16], [17]. Among the proposals, the NC principle is usually utilized by intermediate communication nodes (which linearly combine incoming data streams and forward them as a single stream) and communication-ends (which have to decode the network-coded streams they receive) [18]. Among recent works in this field, Xi and Yeh [16] proposed a utility-based optimization model where the multicast scheme used to deliver a set of independent multicast sessions is optimized to maximize the overall delivery utility function and minimize the network cost (namely, the cost metric associated with a multicast session delivery). In addition, Zhang and Mandayam [17] proposed a multicast scheme aiming at minimizing the total transmission power associated with the delivery of a multicast data session over a multihop OFDMA-based network. However, it is worth noting that both [16] and [17] refer to independent sets of multicast services that are not connected by any coupling constraints (this usually happens in the case of layered video communications).

NC-based strategies have been also proposed for delivering PtM layered services over multihop network topologies [19]–[22]. In particular, Dumitrescu *et al.* [19] designed an NC-based multicast scheme,

Manuscript received July 22, 2013; revised December 7, 2013 and April 15, 2014; accepted June 25, 2014. Date of publication July 8, 2014; date of current version May 12, 2015. The work of J. S. Thompson was supported by the Engineering and Physical Sciences Research Council under Grant EP/J015180/1. The review of this paper was coordinated by Dr. M. Dianati.

A. Tassi is with the School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, U.K. (e-mail: a.tassi@lancaster.ac.uk).

C. Khirallah and J. S. Thompson are with the Institute for Digital Communications, School of Engineering, University of Edinburgh, Edinburgh EH9 3JL, U.K. (e-mail: c.khirallah@ed.ac.uk; john.thompson@ed.ac.uk).

D. Vukobratović is with the Department of Power, Electronics and Communication Engineering, University of Novi Sad, Novi Sad 21000, Serbia (e-mail: dejanjav@uns.ac.rs).

F. Chiti and R. Fantacci are with the Department of Information Engineering, University of Florence, 50139 Florence, Italy (e-mail: francesco.chiti@unifi.it; romano.fantacci@unifi.it).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2014.2336751

where intermediate communication nodes can linearly combine data streams associated with different service layers. The transmission model proposed in [19] aims at maximizing the sum of video layers recovered by all the multicast users. Likewise, Supittayapornpong *et al.* [20] proposed a resource allocation framework aiming at maximizing the overall number of recovered service layers such that given sets of users can recover (at least) a predetermined number of layers. The aforementioned goal is also fulfilled in [21] and [22] by means of a two-stage message-passing and Edmonds–Karp maximum flow algorithm, respectively. The theoretical frameworks presented in [19]–[22], as well as those proposed in [16] and [17], mainly refer to code design issues related to the multihop nature of the networks they consider.

In contrast to [16], [17], and [19]–[22], this paper deals with the application of RNC as a way to improve reliability of communications over a one-hop broadcast network [23]–[26]. Furthermore, this paper draws inspiration from [15] where Khirallah *et al.* proposed to modify the standard LTE media access control (MAC) by adding a coding sublayer (called MAC-RNC). In particular, this provides improved resilience to packet loss of delivered services by using RNC. This paper enhances the work presented in [15] by extending the MAC-RNC design to deliver H.264/SVC video streams as eMBMS broadcast traffic flows. In addition, the authors in [15] investigated the performance of the MAC-RNC-based delivery strategy by comparing it with 3GPP-standardized Hybrid Automatic Repeat-reQuest (HARQ) strategies. However, they do not try to optimize the system design under investigation. To this end, this paper proposes a novel MrT-based strategy aiming at jointly optimizing the modulation and coding scheme (MCS), the transmission rate, and the RNC scheme used to deliver each H.264/SVC video layer to heterogeneous sets of users. We would like to highlight that, unlike the aforementioned MrT schemes [9]–[12], the provided allocation strategy 1) does not require any feedback from the UE devices and 2) ensures that each service layer is successfully received with a given probability by the corresponding user group.

This paper is organized as follows. Section II describes the extension to the MAC layer we considered and the theoretical framework used to evaluate the service level of a H.264/SVC video service transmission. In addition, Section II describes the proposed optimal resource allocation model. Numerical results are presented in Section III, and finally, Section IV concludes this paper.

II. SYSTEM MODEL AND RESOURCE ALLOCATION MODEL FOR RNC-BASED eMBMS VIDEO DELIVERY

Consider a H.264/SVC video stream that is delivered by an eNB as an SC-eMBMS flow to all the UE devices in a cell. Moreover, assume that the service is composed by the set of layers $\{v_0, v_1, \dots, v_L\}$, where v_0 and $\{v_1, \dots, v_L\}$ are the base video layer and the L enhancement layers, respectively.

Fig. 1 shows the LTE protocol stack, proposed in [15], which we refer to. Assuming that each video layer is associated with an independent IP packet stream, the figure shows the stream composed of $L + 1$ video layers that enter the communication stack at the packet data conversion protocol (PDCP) layer. The PDCP protocol data units (PDUs) are concatenated/segmented in the radio link control (RLC) layer and then forwarded to the MAC layer [2]. Because the MAC-RNC sublayer should improve the reliability of data broadcasting, we have that 1) the stream of RLC PDUs related to a video layer is segmented into *information symbols* of L_S bits; 2) information symbols are grouped into sets of K_l items, namely, $\{p_1, \dots, p_{K_l}\}$ as the so-called *information messages* [18]; and 3) according to the RNC principle, the MAC-RNC sublayer produces a stream of *coded symbols* $\{c_1, c_2, \dots\}$ from each information message. Finally, the i th coded symbol is obtained as a linear combination of information symbols (forming an information

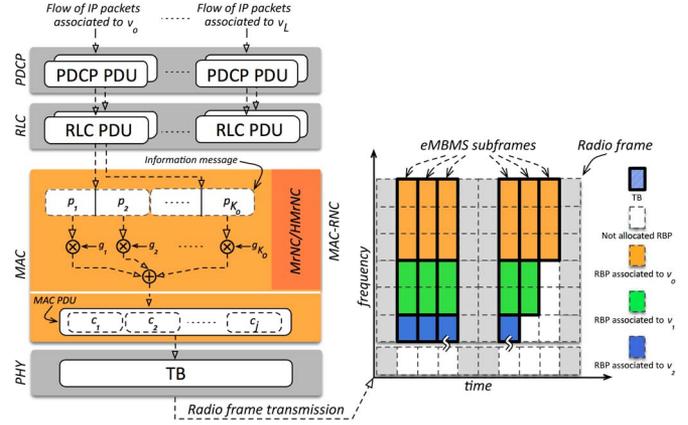


Fig. 1. LTE/LTE-A protocol stack and a part of the radio frame (for $L = 2$).

TABLE I
COMMONLY USED NOTATION

v_l	l -th H.264/SVC video layer
L_S	bit length of an information/coded symbol
K_l	no. of information symbols forming an information message associated with the v_l
N_l	no. of TB transmissions associated with v_l
m_l	MCS used for transmitting TBs of v_l
$N_{\text{RBP},l}$	no. of RBPs forming TBs related to v_l
$n(m_l)$	no. of coded symbols per RBP as a function of m_l
P_{e,m_l}	TBLER related to the MCS m_l
U_l	number of UEs belonging to MA_l
$P_{\text{UE},l}$	probability that a user of MA_l recovers the first $l + 1$ layers
$P_{\text{MA},l}$	delivery probability of the first $l + 1$ layers over MA_l

message), i.e., $c_i = \sum_{j=1}^{K_l} g_j \cdot p_j$, where each *coding coefficient* g_j is taken at random from a uniform distribution over a finite field of size q . A stream of coded symbols associated with an information message is mapped (by the MAC layer) in N_l MAC PDUs. Each MAC PDU is mapped onto a transport block (TB) and broadcast to the UE devices. Hence, depending on the TB size and MCS in use¹, a TB holds a variable number of coded symbols. A UE recovers the delivered information message as soon as K_l linearly independent coded symbols are collected. In this paper, as proposed in [15], we assume that both the eNB and UE devices are equipped with synchronized random number generators (RNGs) such that they can recompute coding coefficients by using RNG seeds. In particular, the RNG seed associated with the first coded symbol in a TB is delivered as part of LTE signaling information. The RNG seeds associated with the remaining coded symbols in the TB are then incrementally computed (from the initial one).

For the sake of clarity, we summarized in Table I the list of symbols that are extensively used in this paper. Table II lists the MCSs that are eligible for TB transmission. In particular, we consider the set of MCSs that corresponds to channel quality indicator (CQI) values m_l that the UE devices feed back for point-to-point (PtP) services indicating their channel conditions [2]. Finally, we assume that all of the TBs containing data associated with the l th video layer are delivered by means of the same MCS m_l .

The transmission time duration of a TB is fixed and equal to transmission time interval (TTI), namely, 1 ms [2]. In addition, a TB

¹The LTE standard imposes that a MAC PDU has to be mapped to a TB. Hence, the MAC PDU size depends on the MCS used for the TB transmission. Again, according to the LTE standard, the MAC layer selects the MCS used for the TB transmission [2].

TABLE II
NUMBER OF CODED SYMBOLS PER RBP VERSUS.
 m_l (FOR $L_S = 32$ bits) [15]

m_l	Mod.	Code Rate	$n(m_l)$	m_l	Mod.	Code Rate	$n(m_l)$
1-3	No Tx	-	-	10	64QAM	0.45	10
4	QPSK	0.3	2	11	64QAM	0.55	12
5	QPSK	0.44	3	12	64QAM	0.65	14
6	QPSK	0.59	4	13	64QAM	0.75	16
7	16QAM	0.37	5	14	64QAM	0.85	18
8	16QAM	0.48	6	15	64QAM	0.93	20
9	16QAM	0.6	8				

may consist of $N_{\text{RBP},l}$ resource block pairs² (RBPs). Fig. 1 shows the time–frequency structure of an LTE radio frame. It consists of ten subframes (each with a transmission time duration of one TTI). The figure reports also the maximum number of subframes that can be used for delivering eMBMS data flows, namely, six out of ten subframes per radio frame. As shown in the figure, we assume that during each eMBMS-capable subframe, the eNB can deliver (at most) one TB per video layer. Hence, a subframe holding eMBMS data can deliver (at most) $L + 1$ TBs (namely, the base layer and L enhancement layers).

The TBs that contain coded symbols associated with the l th video layer are delivered using the MCS m_l and contain $n(m_l)$ coded symbols per RBP. Hence, the total number of coded symbols that can be placed in a TB is $C(m_l, N_{\text{RBP},l}) = n(m_l) \cdot N_{\text{RBP},l}$. Table II lists all the possible values of $n(\cdot)$, for $L_S = 32$ bits [15].

Moreover, we define the l th multicast area (MA) MA_l as the fraction of the cell area where every UE can recover the first $l + 1$ video layers with a given probability. In this paper, we assume that the relation $m_0 \leq m_1 \leq \dots \leq m_L$ holds, i.e., the MCS index of the l th video layer cannot be smaller than that of the $(l - 1)$ th video layer. Let us approximate MA_l (for $l = 0, \dots, L$) as a circle of radius r_l equal to the maximum distance between the eNB and the furthest point where the TB error rate (TBLER) Pe_{m_l} (characterizing the reception of TBs associated with the l th video layer) is not greater than 0.1%.³ For these reasons, we have that $r_l \leq r_{l-1}$. Assuming that UE distribution follows a Poisson point process of average density λ , the average (rounded up) number of UE devices belonging to MA_l is given by $U_l = \lceil \lambda \pi r_l^2 \rceil$ [27]. Hence, the average number of UE devices in the cell is $U_e = \lceil \lambda \pi r_e^2 \rceil$, where r_e is the maximum distance between the eNB and the cell edge.

An H.264/SVC encoded video stream is divided into groups of pictures (GoPs) that consist of g_{GoP} video frames. The video frame rate is given by f_{GoP} frames/s, and the time duration of a GoP is $t_{\text{GoP}} = g_{\text{GoP}}/f_{\text{GoP}}$. Moreover, we can express the time duration of a GoP in terms of the number of TTIs as: $d_{\text{GoP}} = \lceil t_{\text{GoP}}/t_{\text{TTI}} \rceil$, where a t_{TTI} is the LTE TTI (namely, 1 ms).

Since the decoding process of an H.264/SVC video takes place on a per-GoP basis, we define an RNC information message of the l th SVC video layer as the set of information symbols forming the l th layer of a GoP. Hence, K_l is defined as $K_l = \lceil (R_l \cdot t_{\text{GoP}})/L_S \rceil$, where R_l is the bitrate of the l th SVC video layer.⁴

In this paper, the term quality of service (QoS) refers to the received video quality expressed in terms of the number of reconstructed

video layers. For an information message of the l th video layer, the probability that a UE device recovers it (i.e., the probability that a UE device collects K_l linearly independent coded symbols) after N_l TB transmissions as a function of N_l , m_l , and $N_{\text{RBP},l}$ can be expressed as follows [28]:

$$P_{\text{UE},l} \doteq P_{\text{UE}}(N_l, m_l, N_{\text{RBP},l}) = \sum_{t=\chi_l}^{N_l} \left[\binom{N_l}{t} \text{Pe}_{m_l}^{N_l-t} [1 - \text{Pe}_{m_l}]^t \cdot \prod_{i=0}^{K_l-1} \left(1 - \frac{1}{q^{tC(m_l, N_{\text{RBP},l)} - i}} \right) \right] \quad (1)$$

where $\chi_l = \lceil K_l/C(m_l, N_{\text{RBP},l}) \rceil$ is the minimum (integer) number of TB transmissions needed to deliver at least K_l coded symbols. Let us assume that TB reception errors occur as statistically independent events among UE devices of the same MA. From (1), the probability that U_l UE devices recover the l th SVC video layer of a GoP is $(P_{\text{UE},l})^{U_l}$. Hence, the probability that U_l UE devices belonging to MA_l recover the basic and the first l enhancement video layers is (at least) equal to

$$P_{\text{MA},l} \doteq P_{\text{MA}}(N_0, \dots, N_l, m_0, \dots, m_l, N_{\text{RBP},0}, \dots, N_{\text{RBP},l}) = \prod_{i=0}^l P_{\text{UE},i}^{U_i} \quad (2)$$

A. Rate-Optimized and Coverage-Aware Resource Allocation Strategy

The novel resource allocation strategy we propose, which we call “multirate network coding” (MrNC), is embedded into the MAC-RNC sublayer (see Fig. 1), implemented at the eNB side and does not rely on any information related to UE devices in the given cell. The proposed strategy aims at allocating resources to ensure that heterogenous QoS levels are achieved for different MAs. That goal is achieved, for each video layer, by jointly optimizing 1) the TB sizes (in terms of the number of RBPs per TB) $N_{\text{RBP},l}$ and 2) the number of TB transmissions N_l within a GoP time interval and by 3) selecting the MCS m_l of each MA. In particular, the proposed strategy aims at optimizing the number of transmitted coded symbols per video layer. The MrNC model can be stated as follows:

$$(\text{MrNC}) \quad \min_{\substack{m_0, \dots, m_L \\ N_0, \dots, N_L \\ N_{\text{RBP},0}, \dots, N_{\text{RBP},L}}} \sum_{l=0}^L N_l N_{\text{RBP},l} \quad (3)$$

$$\text{subject to} \quad \frac{U_l}{U_e} \geq U_{\text{TH},l} \quad l = 0, \dots, L \quad (4)$$

$$m_l < m_{l+1} \quad l = 0, \dots, L - 1 \quad (5)$$

$$P_{\text{MA},l} \geq \hat{P}_{\text{TH},l} \quad l = 0, \dots, L \quad (6)$$

$$N_{\text{RBP},l} \leq \hat{N}_{\text{TH}} \quad l = 0, \dots, L \quad (7)$$

$$N_l \leq \lfloor \text{TTI}_e d_{\text{GoP}} \rfloor \quad l = 0, \dots, L \quad (8)$$

where constraint (4) ensures that the average number of UE devices per MA is not smaller than a certain value, and (5) avoids the overlapping of any two MAs, since it would be pointless to deliver the same video service characterized by two different QoS levels across the same fraction of the cell area. Using constraint (6), the v_0, \dots, v_l video layers will be recovered with a probability that is at least equal to $\hat{P}_{\text{TH},l}$. It is worth mentioning that the value of $P_{\text{MA},l}$ in (6) has been evaluated by setting $\text{Pe}_{m_l} = 10^{-1}$ (i.e., we set Pe_{m_l} to the greatest TBLER value) in (1) and (2). As for (7), it ensures that the frequency span of each TB cannot be greater than \hat{N}_{TH} . The TB transmissions associated with each video layer have to be completed (at most) in d_{GoP} subframes.

²This is a fixed frequency–time unit of resource allocation within LTE that consists of 12 OFDM subcarriers (180 kHz) \times 1 ms [2].

³In LTE/LTE-A systems, transmitting by using a given MCS is permitted as long as the TBLER experienced by a UE device is equal to or smaller than 10^{-1} [2]. We assume that r_l can be estimated 1) during the network deployment phase or 2) by the eNB itself, which uses CQI values reported by UE devices for standard PtP services.

⁴It is worth mentioning that if the value of K_l is too large for the Gaussian elimination decoder in use, the complexity of the decoding process can be reduced by referring to the systematic version of RNC [15].

Due to the fact that only 60% ($\text{TTI}_e = 0.6$) of subframes per frame are eMBMS capable, constraint (8) states that N_l cannot be greater than $0.6 \cdot d_{\text{GoP}}$. The objective function (3) will minimize the overall radio resource footprint associated with the transmission of all the video layers of a GoP, conditioned so that the QoS constraints, as defined in (4)–(8), are met. In particular, (3) minimizes the sum of RBPs ($N_l \cdot N_{\text{RBP},l}$) associated with each video layer.

Unfortunately, the MrNC model is a complex nonlinear integer optimization problem. Although the solution of MrNC can be found by means of a genetic strategy [29], it cannot be considered a viable alternative in a practical scenario because 1) the number of iterations after (with a good approximation) the optimum solutions of the problem are found cannot be evaluated in advance, and 2) the time required to find the solution quickly becomes prohibitive as the number of variables increases [30]. For this reason, the rest of this section proposes an efficient heuristic strategy to solve the MrNC model, i.e., the heuristic MrNC (HMrNC) strategy. HMrNC comprises three sequential steps that aim to 1) optimize the MCSs of each MA, 2) choose the TB sizes, and 3) optimize the number of TB transmissions.

Considering Procedure 1 that is in charge of the first step, namely, 1) it iterates over the MCS values (starting from 15, see Table II), and 2) for each video layer, it identifies the smallest MA such that constraints (4) and (5) hold. For the second step of HMrNC, we decided to set $N_{\text{RBP},l}$ equal to the maximum possible value (\hat{N}_{TH}) and then optimize the number of TBs transmitted to each MA⁵ (the third step). Let us define⁶ $\tilde{P}_{\text{MA},l}(N_0, \dots, N_l) \doteq P_{\text{MA}}(N_0, \dots, N_l | m_0, \dots, m_l, N_{\text{RBP},0}, \dots, N_{\text{RBP},l})$. Since m_0, \dots, m_L and $N_{\text{RBP},0}, \dots, N_{\text{RBP},L}$ are given, the MrNC problem can be restated as follows:

$$(H1) \quad \min_{N_0, \dots, N_L} \sum_{l=0}^L N_l \quad (9)$$

$$\text{subject to } N_l \leq \lfloor \text{TTI}_e d_{\text{GoP}} \rfloor \quad l = 0, \dots, L \quad (10)$$

$$\tilde{P}_{\text{MA},l}(N_0, \dots, N_l) \geq \hat{P}_{\text{TH},l} \quad l = 0, \dots, L. \quad (11)$$

Once again, H1 is a noninteger and nonlinear problem, but in this case, it can be efficiently solved. To this end, considering $\tilde{P}_{\text{MA},l}(N_0, \dots, N_l)$, from (2), we can see that the probability value cannot decrease when N_l increases, and the remaining variables are kept constant. Furthermore, let N_l^* (for $l = 0, \dots, L$) be the smallest value of N_l such that $\tilde{P}_{\text{MA},l}(N_0, \dots, N_l) \geq \hat{P}_{\text{TH},l}$ (for $l = 0, \dots, L$) holds. Likewise, the approach presented in [28] and starting from $l = 0$, the value of N_l^* can be efficiently found by testing all the possible values of N_l from χ_l until $\tilde{P}_{\text{MA},l}(N_0, \dots, N_l) \geq \hat{P}_{\text{TH},l}$ holds. In particular, Proposition 1 proves that the objective function (9) is minimized by $\{N_0^*, \dots, N_L^*\}$. Finally, Procedure 2 proposes a possible implementation of the proposed strategy.

Procedure 1: Definition of the MAs.

```

t ← 15
for l = 0 → L do
  repeat
    m_l ← t
    t ← t - 1
  until U_l/U_e ≥ U_TH,l or t < 4
end for

```

⁵This method will tend to reduce the transmission time duration of a GoP rather than optimize the TB sizes. In addition, the latter aspect can be indirectly addressed during the service deployment phase by tuning the value of \hat{N}_{TH} .

⁶In this paper, we define $f(x|t_0, \dots, t_w)$ as the parametric function where x is the variable, and t_0, \dots, t_w are parameters.

Procedure 2: Minimization of the time duration of the process.

```

for l = 0 → L do
  N_l* ← χ_l
  while P_MA,l(N_0, ..., N_l) |_{N_0=N_0*, ..., N_{l-1}=N_{l-1}*} < P_TH,l do
    N_l* ← N_l* + 1
  end while
end for

```

Proposition 1: $\{N_0^*, \dots, N_L^*\}$ is an optimum solution of H1.

Proof: The probability $\tilde{P}_{\text{MA},l}(N_0, \dots, N_l)$ is a nondecreasing function with respect to the variable N_l (for any $l = 0, \dots, L$). Considering Procedure 2, it starts from $l = 0$ and minimizes the functions $\tilde{P}_{\text{MA},l}(N_0), \dots, \tilde{P}_{\text{MA},l}(N_0, \dots, N_l) |_{N_0=N_0^*, \dots, N_{l-1}=N_{l-1}^*}$, etc. Let us assume the existence of another solution $\{N'_0, \dots, N'_L\}$ of H1 such that $\sum_{l=0}^L N'_l < \sum_{l=0}^L N_l^*$. Hence, there is at least one term N'_l such that $N'_l < N_l^*$. However, because of the definition of N_l^* , constraint (11) would not hold. This completes the proof by reductio ad absurdum. ■

Consider Procedure 2, it can solve H1 in a finite number of steps. In particular, we can note that N_l^* belongs to the interval $I = [\chi_l, \lfloor \text{TTI}_e \cdot d_{\text{GoP}} \rfloor]$. During one iteration, the procedure tests just one value of I . Hence, N_l^* is found in a number of iterations that are less than or equal to the number of items in I . For this reason, Procedure 2 returns (at most) after Q iterations such that $Q \leq \sum_{l=0}^L (\lfloor \text{TTI}_e d_{\text{GoP}} \rfloor - \chi_l + 1)$.

III. NUMERICAL RESULTS

This section investigates the system performance in terms of the resource load index η defined as

$$\eta = \frac{1}{\lfloor \text{TTI}_e d_{\text{GoP}} \rfloor \hat{N}_{\text{TH}}} \sum_{l=0}^L N_{\text{RBP},l} N_l \quad (12)$$

where $\sum_{l=0}^L N_{\text{RBP},l} \cdot N_l$ represents the *radio resource footprint* of the allocation strategy (namely, the objective function of the MrNC problem). In addition, we consider the probabilities⁷ $P_{\text{MA},l}$ that a reference group of ten UE devices can recover each service QoS level [see (2)], and hence, the maximum achievable peak signal-to-noise ratio (PSNR) defined as

$$\bar{p} = \max_{l=0, \dots, L} \{\hat{p}_l P_{\text{MA},l}\} \quad (13)$$

where, \hat{p}_l is the PSNR obtained after recovery of the video layers v_0, \dots, v_l .

We provide performance comparisons between the resource allocation solutions obtained by the HMrNC heuristic approach and, for the sake of comparison, by directly solving the MrNC model.⁸ We also consider the allocation model proposed in [11], which is named hereafter as the *video rate allocation* (VRA) strategy, that tries to maximize the sum of the video quality perceived by each UE. To make a fair comparison among the MrNC (directly solved), HMrNC, and VRA methods, we impose that the eNB cannot skip the transmission

⁷Here, we referred to Pe_{m_l} (for $l = 0, \dots, L$) values computed by averaging TBBLER values obtained by 10^3 iterations of the datalink simulation framework presented in [15].

⁸It is worth mentioning that the MrNC problem has been solved by means of a genetic strategy [29] (see Section II-A). Throughout this section, we will refer to that kind of solution as the “direct solution” of the MrNC problem.

TABLE III
SIMULATION PARAMETERS CONSIDERED

Parameter	Value	
Inter-Site-Distance (ISD)	500 m	
System Bandwidth	20 MHz	
Transmission Scheme	SISO	
Duplexing Mode	FDD	
Carrier Frequency	2.0 GHz	
Transmission Power	40 W per-sector	
eNB and UE Antenna Gains	see Table A.2.1.1-2 [31]	
Pathloss and Penetration Loss	see Table A.2.1.1.5-1 [31]	
$\hat{P}_{TH,l}$	0.9, for $l \in 0, \dots, L$	
\hat{N}_{TH}	$[4, \dots, 12]$ RBPs	
Stream A [11]	$\{\hat{z}_0, \dots, \hat{z}_2\}$ [kbps]	$\{117.1, 402.5, 1506.3\}$
	$\{\hat{p}_0, \dots, \hat{p}_2\}$ [dB]	$\{29.94, 34.78, 40.73\}$
	$\{U_{TH,0}, \dots, U_{TH,2}\}$	$\{0.4, 0.5, 0.9\}$
Stream B	$\{\hat{z}_0, \dots, \hat{z}_3\}$ [kbps]	$\{160.0, 300.0, 560.0, 1150.0\}$
	$\{\hat{p}_0, \dots, \hat{p}_3\}$ [dB]	$\{29.45, 32.30, 34.52, 38.41\}$
	$\{U_{TH,0}, \dots, U_{TH,3}\}$	$\{0.4, 0.55, 0.75, 0.9\}$
Stream A, B	g_{GoP}	16 frames
	f_{GoP}	30 fps
	q	2^8

of any video layer; hence, we restate the VRA objective function as follows⁹:

$$\begin{aligned} \text{(VRA)} \quad & \max_{m_0, \dots, m_L} \sum_{l=0}^L U_l \hat{p}_l & (14) \\ \text{subject to} \quad & m_l < m_{l+1} \quad l = 0, \dots, L-1. & (15) \end{aligned}$$

Furthermore, we compare both the direct solution of MrNC and that obtained by HMrNC with an MrT-based strategy [6]. For the latter strategy, we draw inspiration from [12], where UE devices are split into multiple multicast groups (MGs); the transmission rate used to deliver data to an MG is constrained by the UE experiencing the worst propagation conditions (in the MG). This means that the MrT optimization problem can be restated, for any $l = 0, \dots, L$, in the following equivalent form:

$$\text{(MrT)} \quad \arg \max_{m_l \in [4, \dots, 15]} \left\{ m_l \mid \frac{U_l}{U_e} \geq U_{TH,l} \right\}. \quad (16)$$

This tries to deliver the l th video layer over MA_l by using the maximum MCS such that relation (4) holds (namely, an LCG-based approach is used within an MG).

Due to the fact that neither the VRA nor MrT strategies explicitly address the TB sizing problem, we assume that each TB consists of \hat{N}_{TH} RBPs. In addition, both the VRA and MrT strategies assume that UE devices can report to the eNB CQI feedback; however, one of the key points of both MrNC and HMrNC is that they do not rely on any UE feedback. Hence, for the sake of comparison, we assume that the actual number of UE devices, which, on average, report the CQI value m_l , is equal to U_l . Finally, in the case of both the VRA- and MrT-based video delivery, transmissions take place through the standard LTE MAC layer (namely, a communication stack without the MAC-RNC sublayer).

We consider a network of 19 macrocell eNBs, each managing three hexagonal sectors. eNBs are organized in two concentric circles centered on the target eNB. In addition, TBLER values experienced by a UE device, as a function of a given MCS and distance from the eNB, are estimated by the finite-state Markov model approach presented in [15]. Table III summarizes both the main simulation parameters and the two H.264/SVC video streams [32] that we considered. In particular, for each video layer, we report also the bitrate \hat{z}_l obtained after recovery of the first $l + 1$ video layers.

⁹The original formulation of the VRA model aims at jointly optimizing the set of delivered layers and MCSs used in the transmissions [11].

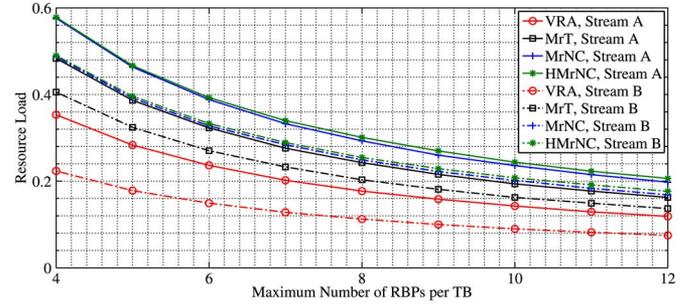


Fig. 2. Resource load index as a function of \hat{N}_{TH} .

Results reported here will clearly show that both resource allocation solutions obtained by directly solving the proposed MrNC or by using the HMrNC strategy meet predefined service constraints (4) and (6), with the minimum resource footprint (3). Furthermore, in spite of the fact that the radio resource footprint of VRA and MrT (required to achieve their respective goals) is smaller than those associated with the direct MrNC and HMrNC solutions, they cannot ensure that a predefined video QoS level is maintained over the targeted fractions of the cell area.

In Fig. 2, we compare the value of η , as a function of \hat{N}_{TH} , characterizing all the considered resource allocation strategies, in the case of video streams A and B. From (12), we have that the overall number of RBPs used to deliver a stream increases as the value of η enlarges. Considering the figure, the performance gap between the solution obtained by directly solving MrNC (indicated in all the figures in this section as “MrNC”) and that derived by HMrNC is negligible (at most, it is less than 0.01). It is worth noting that that is caused by the fact that, in the latter case, both the MCS selection and TB sizing processes are separated from the optimization of the number of TB transmissions.¹⁰ On the other hand, both the VRA and MrT strategies deliver video streams A and B by using a fraction of the radio resources, which is smaller than that we have by directly solving MrNC (or by using the HMrNC strategy) of, at most, 1.63 and 1.19 (2.18 and 1.20) times, respectively.

In spite of the larger radio resource footprint for resource allocation obtained by directly solving MrNC or derived by means of HMrNC, it is worth noting that the proposed resource allocation framework *can deliver a service with the desired QoS level over a given fraction of the cell area*. Considering stream A, Fig. 3 compares (for $\hat{N}_{TH} = 6$) the $P_{MA,l}$ values of each QoS level and \bar{p} as a function of the distance of the considered reference group from the eNB. For each MA, the figures report the value of r_l (the dashed lines). Unlike the direct MrNC and HMrNC-based resource allocation solutions, both the VRA and MrT strategies cannot deliver the service over the desired fractions of the cell area. For instance, MA_0 (MA_2) defined by the VRA and MrT strategies extends up to a distance that is 81.9 and 14.9 m (20.2 m for both strategies) smaller than the minimum required, respectively. Finally, Fig. 4 shows (for $\hat{N}_{TH} = 6$) similar behavior for video stream B. In this case, we note that the MA_0 (MA_3) provided by the VRA and MrT allocation strategies spans up to a distance that negatively diverges from the minimum required one of 71.9 and 15.9 m (20.2 m for both strategies), respectively.

It is worth noting that the resource allocation solution derived by the HMrNC strategy is 1) a feasible solution of the MrNC problem but

¹⁰In particular, as expected, η values associated with the HMrNC strategy are (slightly) greater than those relative to the solutions obtained by directly solving the MrNC problem. As a consequence, the number of coded symbols used to deliver a video layer might be slightly bigger than that we have by directly solving MrNC.

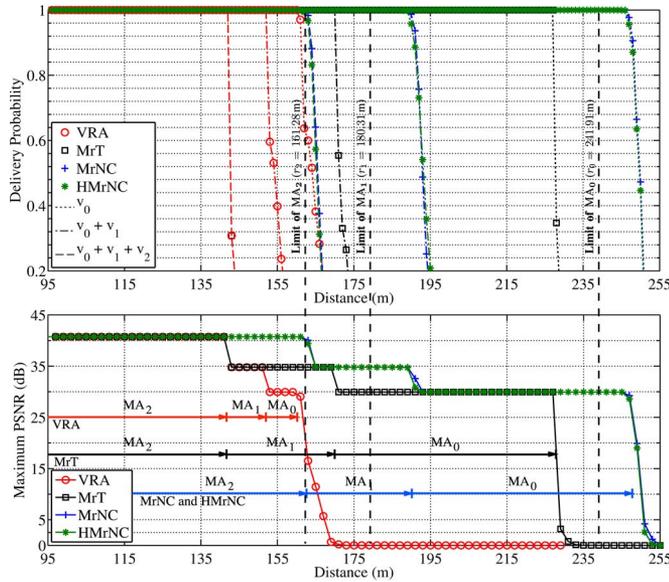


Fig. 3. Video delivery probabilities and maximum PSNR of stream A versus distance from eNB.

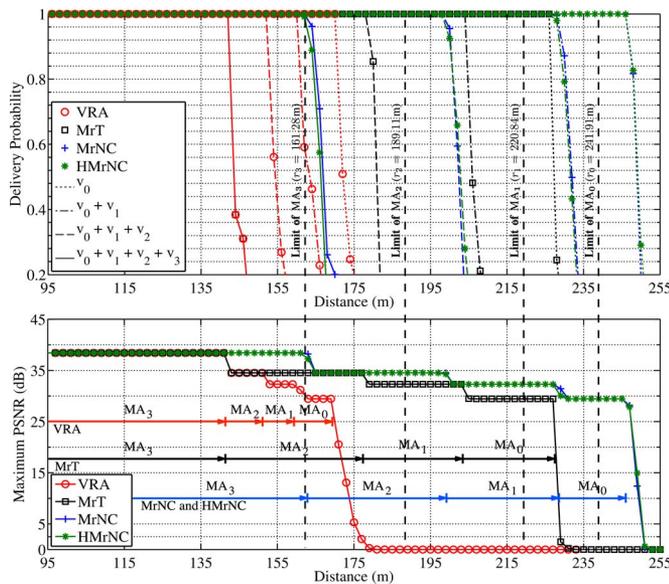


Fig. 4. Video delivery probabilities and maximum PSNR of stream B versus distance from eNB.

2) may be characterized by a slightly greater resource footprint than that of the direct solution of MrNC (which, with a good approximation, approaches the optimum solution of MrNC). Hence, the HMrNC strategy may provide a resource allocation solution that leads to delivering more coded symbols per video layer than the corresponding direct solution. For this reason, the $P_{MA,l}$ values (2) may be slightly greater than those associated with the direct solution of MrNC. This means that the HMrNC strategy could be able to deliver a video stream, at a certain QoS level, over an MA, which is slightly greater than that associated with the direct solution of MrNC. In particular, this effect can be noted (see Figs. 3 and 4) by considering the delivery probability values associated with the reception of v_0 and v_1 (v_0 , v_1 , and v_2).

IV. CONCLUSION

In this paper, we have proposed an optimal (the MrNC model) and heuristic resource allocation strategy (the HMrNC procedure)

suitable for SC-eMBMS broadcast communications delivered through the MAC-RNC sublayer. We demonstrated that HMrNC can efficiently derive feasible solutions of MrNC with a reduced computational load. Unlike VRA and MrT strategies, both MrNC and HMrNC ensure the desired service coverage. In particular, the VRA and MrT strategies can deliver the considered video streams at the maximum (minimum) QoS level over MAs, which, at least, are 22% (50% and 12%, respectively) smaller than the desired service levels.

REFERENCES

- [1] "Cisco visual networking index—Forecast and methodology, 2013–2018," San Jose, CA, USA, Tech. Rep., 2014.
- [2] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, 2011.
- [3] D. Lecomte and F. Gabin, "Evolved Multimedia Broadcast/Multicast Service (eMBMS) in LTE-advanced: Overview and Rel-11 enhancements," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 68–74, Nov. 2012.
- [4] "Advanced video coding for generic audiovisual services," Geneva, Switzerland, Tech. Rep. ITU-T H.264, Nov. 2007.
- [5] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.
- [6] R. Afolabi, A. Dadlani, and K. Kim, "Multicast scheduling and resource allocation algorithms for OFDMA-based systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 240–254, 2013.
- [7] S. Li, X. Wang, H. Zhang, and Y. Zhao, "Dynamic resource allocation with precoding for OFDMA-based wireless multicast systems," in *Proc. IEEE VTC Spring*, 2011, pp. 1–5.
- [8] J. Liu, W. Chen, Z. Cao, and K. Letaief, "Dynamic power and sub-carrier allocation for OFDMA-based wireless multicast systems," in *Proc. IEEE ICC*, 2008, pp. 2607–2611.
- [9] D. Munaretto and M. Zorzi, "Robust opportunistic broadcast scheduling for scalable video streaming," in *Proc. IEEE WCNC*, 2012, pp. 2134–2139.
- [10] R. Radhakrishnan and A. Nayak, "An efficient video adaptation scheme for SVC transport over LTE networks," in *Proc. IEEE ICPADS*, 2011, pp. 127–133.
- [11] D. Munaretto, D. Jurca, and J. Widmer, "Scalable video broadcast in cellular networks: Impact on QoS and network resources," in *Proc. IEEE ISCC*, Riccione, Italy, 2010, pp. 969–974.
- [12] C. Tan, T. Chuah, and S. Tan, "Adaptive multicast scheme for OFDMA based multicast wireless systems," *Electron. Lett.*, vol. 47, no. 9, pp. 570–572, Apr. 2011.
- [13] J. Vella and S. Zammit, "A survey of multicasting over wireless access networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 718–753, 2013.
- [14] E. Magli, M. Wang, P. Frossard, and A. Markopoulou, "Network coding meets multimedia: A review," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1195–1212, Aug. 2013.
- [15] C. Khirallah, D. Vukobratović, and J. Thompson, "Performance analysis and energy efficiency of random network coding in LTE-advanced," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4275–4285, Dec. 2012.
- [16] Y. Xi and E. Yeh, "Distributed algorithms for minimum cost multicast with network coding," *IEEE/ACM Trans. Netw.*, vol. 18, no. 2, pp. 379–392, Apr. 2010.
- [17] D. Zhang and N. Mandayam, "Resource allocation for multicast in an OFDMA network with random network coding," in *Proc. of IEEE INFOCOM*, Apr. 2011, pp. 391–395.
- [18] C. Fragouli, "Network coding: Beyond throughput benefits," *Proc. IEEE*, vol. 99, no. 3, pp. 461–475, Mar. 2011.
- [19] S. Dumitrescu, M. Shao, and X. Wu, "Layered multicast with inter-layer network coding," in *Proc. IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009, pp. 442–449.
- [20] S. Supittayapornpong, P. Saengudomlert, and W. Kumwilaisak, "Qos aware layered multi-cast with network coding in lossy networks," in *Advances in Multimedia Information Processing—PCM 2009*, vol. 5879, P. Muneesawang, F. Wu, I. Kumazawa, A. Roeksabutr, M. Liao, and X. Tang, Eds. Berlin, Germany: Springer-Verlag, 2009, ser. Lecture Notes in Computer Science, pp. 428–439.
- [21] K. Minji, D. Lucani, X. Shi, F. Zhao, and M. Médard, "Network coding for multi-resolution multicast," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–9.
- [22] J. Widmer, A. Capalbo, A. Anta, and A. Banchs, "Rate allocation for layered multicast streaming with inter-layer network coding," in *Proc. IEEE INFOCOM*, Orlando, FL, USA, Mar. 2012, pp. 2796–2800.

- [23] M. Ghaderi, D. Towsley, and J. Kurose, "Reliability gain of network coding in lossy wireless networks," in *Proc. IEEE INFOCOM*, Phoenix, AZ, USA, Apr. 2008, pp. 196–200.
- [24] A. Eryilmaz and L. Ying, "Scaling laws with network coding," in *Network Coding: Fundamentals and Applications*. New York, NY, USA: Academic, 2012, pp. 235–265.
- [25] N. Xie and S. Weber, "Network coding broadcast delay on erasure channels," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, Feb. 2013, pp. 1–8.
- [26] A. Tassi, F. Chiti, R. Fantacci, and F. Schoen, "An energy efficient resource allocation scheme for RLNC-based heterogeneous multicast communications," *IEEE Commun. Lett.*, 2014. [Online]. Available: <http://dx.doi.org/10.1109/LCOMM.2014.2331069>
- [27] J. Andrews, R. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 156–163, Nov. 2010.
- [28] F. Chiti, R. Fantacci, F. Schoen, and A. Tassi, "Optimized random network coding for reliable multicast communications," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1624–1627, Aug. 2013.
- [29] K. Deep, K. Pratap Singh, M. Kansal, and C. Mohan, "A real coded genetic algorithm for solving integer and mixed integer optimization problems," *Appl. Math. Comput.*, vol. 212, no. 2, pp. 505–518, Jun. 2009.
- [30] D. Goldberg, *Genetic Algorithms*. New Delhi, India: Pearson Publishing, 2013.
- [31] *Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects*, 3GPP TR 36.814 v9.0.0 (Release 9), 2010.
- [32] YUV Video Sequences. [Online]. Available: <http://trace.eas.asu.edu/yuv/>

Robust Joint Beamforming and Jamming for Secure AF Networks: Low-Complexity Design

Chao Wang and Hui-Ming Wang, *Member, IEEE*

Abstract—This paper investigates the physical-layer security issue of an amplify-and-forward (AF) relay network. We propose a joint cooperative beamforming (CB) and cooperative jamming (CJ) design that is robust to the imperfect channel state information (CSI) of the multiple multiantenna eavesdroppers. The objective is the worst-case secrecy rate maximization (WCSR). Based on the inner convex approximation and semidefinite relaxation (SDR) technique, the original nonconvex problem is transformed into a sequence of approximate convex problems, which can be solved conveniently. The proposed robust design guarantees to converge to a Karush–Kuhn–Tucker (KKT) solution and is efficient in computational complexity. Furthermore, we prove that SDR always has a rank-1 solution, which identifies that SDR is tight. Simulations demonstrate the validity of the proposed strategy.

Index Terms—Cooperative beamforming (CB), inner convex approximation (ICA), jamming, secrecy rate, semidefinite relaxation (SDR).

Manuscript received October 28, 2013; revised April 23, 2014; accepted June 26, 2014. Date of publication July 2, 2014; date of current version May 12, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 61102081 and Grant 61221063, by the Specialized Research Fund for the Doctoral Program of Higher Education of China under Grant 20110201120013, by the New Century Excellent Talents Support Fund of China under Grant NCET-13-0458, by the Industrial Research Fund of Shaanxi Province under Grant 2012GY2-28, by the Fok Ying Tong Education Foundation under Grant 141063, and by the Fundamental Research Funds for the Central University under Grant 2013jdgz11. The review of this paper was coordinated by Prof. H.-F. Lu. (*Corresponding author: Hui-Ming Wang.*)

The authors are with the School of Electronic and Information Engineering and Ministry of Education Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: wangchaoxuzhou@stu.xjtu.edu.cn; xjbswhm@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2014.2334640

I. INTRODUCTION

Physical-layer secrecy exploits the physical characteristics of the wireless channels to improve the secure transmissions [1]. Cooperative transmission utilizing multiple distributed relay nodes to improve the *secrecy rate* of the wireless networks has received increasing attention [2]–[12]. In both decode-and-forward (DF) and amplify-and-forward (AF) relay networks, two efficient ways are proposed: cooperative beamforming (CB) and cooperative jamming (CJ). For the CB, through designing the complex weight at each relay node, the received signal at the legitimate destination is strengthened, and the information leakage to the eavesdroppers can be reduced [2]–[12]. For the CJ, cooperative nodes transmit jamming signals to confuse the eavesdroppers [2], [7].

DF-based CB and CJ for the secrecy rate maximization (SRM) and the transmit power minimization with a secrecy rate constraint have been separately investigated in [2] and [3]. Comparing with the DF-based CB, the AF-based CB design has lower implementation complexity. However, due to the noise amplification effects, it is even more difficult to design the optimal AF-based CB in terms of SRM [2], [4]–[7]. Aiming to improve its secrecy rate, two alternatives have been proposed by the existing works for the AF network [2]–[7]. The first is to consider some suboptimal schemes. For example, "Null-Space Beamforming" has been proposed in [2] and [4]–[6] to do beamforming in the nullspace of the eavesdroppers' channels. Obviously, this scheme can only obtain the strictly suboptimal solution. When the eavesdroppers' channel state information (CSI) is imperfect, its performance degrades significantly. To achieve a better secrecy rate, the other technique is to do a two-level optimization, which is an exhaustive 1-D search involving a sequence of semidefinite programming (SDPs) [6], [7]. With this technique and the semidefinite relaxation (SDR) [8], Yang *et al.* investigated the CB and CJ designs in [6], [7]. The common problems of the proposed strategies in [2]–[7] are three-fold: 1) The optimality of the 1-D search algorithm heavily depends on the search interval, and the computational complexity is very high. 2) The eavesdroppers' CSI is assumed to be perfect. When the eavesdroppers are passive, their perfect CSI is difficult to get. 3) The assumptions that each eavesdropper has only a single antenna and they are non-colluding may underestimate the capability of the eavesdropper and is over optimistic, which may result in confidential information leakage.

To consider the more practical case that the perfect CSI of the eavesdroppers is not known, in our previous work [9], we propose a security quality-of-service-based joint CB and CJ design without the eavesdroppers' CSI. Recently, assuming only the imperfect eavesdroppers' CSI is available, based on the worst-case SRM (WCSR), Huang and Swindlehurst [10] have investigated the robust secure transmission in multi-input–single-output channels, and Vishwakarma and Chockalingam proposed a robust joint design of the CB and CJ for securing AF networks in [11]. Similarly, a two-level optimization is required in [11], and in particular, a 1-D search is involved as well. With imperfect CSI, Wang *et al.* in [12] investigated the zero-forcing and match-and-forward-based secure relay beamforming with a multiantenna relay and a single eavesdropper, which is strictly suboptimal as well.

In this paper, we consider the secrecy of an AF relay network where there are multiple eavesdroppers equipped with multiple antennas each and their CSI is imperfectly known. We propose a joint robust optimization of CB and CJ for WCSR under both individual and global power constraints of the relay nodes. The worst-case design based on the deterministic model has a long history in signal processing community [10], [11], [13], [14], which optimizes the performance under the worst-case channel condition [14]. In this case, the absolute