Distributed Robust Optimization for Scalable Video Multirate Multicast Over Wireless Networks

Chenglin Li, Hongkai Xiong, *Senior Member, IEEE,* Junni Zou, *Member, IEEE,* and Chang Wen Chen, *Fellow, IEEE*

***Abstract*—This paper proposes a distributed robust optimization scheme to jointly optimize overall video quality and traffic performance for scalable video multirate multicast over practical wireless networks. In order to guarantee layered utility maximization, the initial nominal joint source and network optimization is defined, where each scalable layer is tailored in an incremental order and finds jointly optimal multicast paths and associated rates with network coding. To enhance the robustness of the nominal convex optimization formulation with nonlinear constraints, we reserve partial bandwidth for backup paths disjoint from the primal paths. It considers the path-overlapping allocation of backup paths for different receivers to take advantage of network coding, and takes into account the robust multipath rate-control and bandwidth reservation problem for scalable video multicast streaming when possible link failures of primary paths exist. Specifically, an uncertainty set of the wireless medium capacity is introduced to represent the uncertain and time-varying property of parameters related to the wireless channel. The targeted uncertainty in the robust optimization problem is studied in a form of protection functions with nonlinear constraints, to analyze the tradeoff between robustness and distributedness. Using the dual decomposition and primal-dual update approach, we develop a fully decentralized algorithm with regard to communication overhead. Through extensive experimental results under critical performance factors, the proposed algorithm could converge to the optimal steady-state more quickly, and adapt the dynamic network changes in an optimal tradeoff between optimization performance and robustness than existing optimization schemes.**

***Index Terms*—Multirate multicast, network coding, rate-distortion, robust optimization, scalable video coding.**

I. Introduction

VIDEO STREAMING over wireless networks has been compelling for a wide range of applications, from home entertainment and video surveillance to audiovisual communication [1]. To benefit the overall network utility, scalable video coding (SVC) is developed to adapt different user requirements and heterogeneous network conditions in a multirate multicast mode. SVC can allow rate adaptation not only at the encoder/decoder but also at the intermediate network nodes while achieving highly efficient rate-distortion (RD) performance [2]. An SVC stream comprising a base layer and one or multiple enhancement layers, with a flexible multidimension layer structure, can provide various operating points in spatial resolution, temporal frame rate, and video reconstruction quality. For multirate multicasting, different SVC layers are transported in different Internet protocol (IP) multicast groups and are subscribed by heterogeneous receivers with different computation and communication resources and capabilities. In communication networks, on the other hand, it has been demonstrated that coded network can achieve the capacity in single-source multiple-terminal multicast via algebraic operation at intermediate nodes [3]. Also, recent research has shown that network coding can significantly improve the network throughput and robustness to link or node failures [4] and packet losses [5], [6]. Distributed random linear coding schemes [7], [8] have been proposed for practical implementation of network coding. Chen *et al.* [9] developed adaptive rate-control algorithms for networks with and without coding subgraphs.

When both SVC and network coding are taken into account, layered multirate multicasting is equivalent to a generalized multisource problem where the progressive interlayer dependency is considered as fairness between different sources (layers) [10], and also network coding is implemented at relay nodes to enhance the communication efficiency of the network. In practical wireless networks, dynamic network changes (e.g., fluctuations of links or link failures) might occur due to wireless channels and mobility in wireless nodes [11]. Therefore, it is imperative to formulate the corresponding rate-control scheme in a robust way, to ensure service feasibility and availability against such uncertainties. Furthermore, the lack of centralized control in a wireless network requires that the scheme be performed in a distributed manner. In this paper, we aim to develop a robust and efficient optimization framework for scalable video streaming over wireless networks, where the layered SVC stream is generated at the source node, and distributed via network coding by relay nodes to different users through the wireless network. The objective is to maximize the overall video quality of all receivers with dynamic network changes.

Manuscript received April 3, 2011; revised October 17, 2011; accepted November 17, 2011. Date of publication February 3, 2012; date of current version May 31, 2012. This work was supported in part by the National Natural Science Foundation of China, under Grants 60802019, 60928003, 60772099, and 60632040, in part by the Program for New Century Excellent Talents in University, under Grant NCET-09-0554, and in part by the Shanghai Rising-Star Program, under Grant 11QA1402600. This paper was recommended by Associate Editor S. Battiato.

C. Li and H. Xiong are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: lcl1985@sjtu.edu.cn; xionghongkai@sjtu.edu.cn).

J. Zou is with the School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China (e-mail: zoujn@shu.edu.cn).

C. W. Chen is with the Department of Computer Science and Engineering, State University of New York, Buffalo, NY 14260 USA (e-mail: chencw@buffalo.edu).

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/TCSVT.2012.2186740

The contribution of this paper is twofold. First, we are motivated to make a robust formulation of jointly optimizing overall video quality and traffic performance for scalable video multirate multicast over practical wireless networks. To guarantee layered utility maximization, we define a nominal joint source and network flow optimization where each scalable layer is tailored in an incremental order and finds jointly optimal multicast paths and associated rates. At relay nodes, network coding is utilized for a decent multicast capacity. To enhance the robustness of the nominal convex optimization problem with nonlinear constraints, we reserve partial bandwidth for backup paths disjoint from the primal paths during the transmission of SVC multicast stream. It considers the path-overlapping allocation of backup paths for different receivers to take advantage of network coding, and takes into account the robust multipath rate-control and bandwidth reservation problem for SVC multicast streaming when possible link failures of primary paths exist. Impressively, an uncertainty set of the wireless medium capacity is introduced to represent the uncertain and time-varying property of parameters related to the wireless channel. The targeted uncertainty in the robust optimization problem is studied in a form of protection functions with nonlinear constraints, to analyze the tradeoff between robustness and distributedness. Second, using dual decomposition and the primal-dual update approach, we develop a fully decentralized algorithm where the globally optimal solution is achievable via distributed computation, by iteratively updating optimization variables. Extensive numerical and packet-level experimental results under critical factors demonstrate that the proposed algorithm (PA) can converge to the optimal steady-state more quickly, and adapt the dynamic network changes in an optimal tradeoff between optimization performance and robustness than existing schemes.

The remainder of this paper is organized as follows. Section II reviews related research on rate allocation schemes of scalable video streaming, as well as robust optimization frameworks. The notations, system models, and nominal flow control and resource allocation problem for SVC-based multirate multicasting with network coding-based routing over wireless networks is formulated in Section III. In Section IV, the robust problem formulation is presented, with consideration of uncertain factors in practical wireless networks. The fully distributed algorithm for the robust optimization formulation is proposed in Section V, where we also provide an efficient implementation scheme. Extensive results of both numerical experiments and packet-level simulations are presented in Section VI. Section VII concludes this paper.

II. Related Work

Over the past years, a few of rate-control schemes have been developed in literature to address the scalable video streaming problem over networks [10], [12]–[14]. The distributed rate allocation scheme in [10] addressed the problem of rate allocation for multiple SVC multicast sessions over wireless mesh networks, with the goal of minimizing total video distortion of all peers. To incorporate network coding as well as interlayer dependency into layered overlay multicast over networks, Zhao *et al.* [12] proposed a distributed heuristic algorithm with suboptimal performance. Recently, Zou *et al.* [13] explored the path cost and price of each layer as the priority parameters to capture interlayer dependency and developed a rigorous distributed algorithm proven to be stable and convergent. As a further improvement, a joint source and flow optimization scheme with mathematically rigorous layer dependency constraints has been shown the best overall performance for heterogeneous receivers via a fully distributed implementation supporting partial layer reception in [14]. This kind of optimization problems, although can achieve optimal or near optimal performance, would make the unrealistic assumption that data defining the optimization formulation can be obtained precisely, which may lead to poor or even infeasible solution in practical dynamic networks. Hereinafter, we call the corresponding optimization problems with deterministic parameters "nominal" problems.

As a framework of tackling optimization problems under data uncertainty, a couple of robust optimization schemes have emerged to seek a solution that remains feasible and near optimal under the fluctuation of parameters in the optimization formulation [15]– [18]. Each robust optimization problem is defined by three-tuple: a nominal formulation, a definition of robustness, and a representation of the uncertainty set. The process of making an optimization formulation robust can be viewed as a mapping from one optimization problem to another. In [15], it was demonstrated that optimal solutions of linear programming (LP) problems may become severely infeasible if the nominal data is slightly perturbed, also robust solutions of the above LPs which were in a sense immune against uncertainty were developed. A D-norm approach [16] has been proposed to model the data uncertainty, and has advantages such as guarantee of feasibility independent of uncertainty distributions and flexibility in tradeoff between robustness and performance. Yang *et al.* [17], [18] discussed several efficient models for describing parameter uncertainty sets that can lead to decomposable problem structures, and applied these models in distributively solving a robust flow control problem in wireline networks. These work mainly focus on the robust formulation of linear optimization problems and is thus not suitable for complicated video application.

III. System Models and the Nominal Optimization Problem Formulation

In this section, related system models are addressed. As a fundamental step for optimization formulation of robust SVC multirate multicast over wireless networks, in the sequel, we formulate the nominal optimization problem.

A. Wireless Network Model

Consider video content distribution over a wireless network $G = (V, E)$ with a set of wireless links E and a set of wireless nodes $V = \{s\} \cup N \cup D$, where s denotes the source node, and N and D represent the sets of relay nodes and destination nodes, respectively. The SVC-coded stream is encoded at the source s and then multicasted to destination nodes through a wireless network with network coding-based routing.

In wireless networks, the capacity of a wireless link is interrelated with other adjacent wireless links. Consequently, we should consider the wireless link contention in a shared transmission medium by introducing constraints of the location-dependent contention among the competing wireless data flows [19]. In the target problem formulation, the assumption is that the wireless medium capacity is shared among a wireless link l and the cluster of its competing links. A typical protocol model considers the spatial locations of the nodes and determines that transmission can be successfully received by its intended recipient [20]. It hypothesizes that any link originating from node k will interfere with link (i, j) if the link distance $d_{(k,j)} < (1+\gamma)d_{(i,j)}$, $\gamma \geq 0$ and define $\Psi_{(i,j)}$ for each link $(i, j) \in E$ as the cluster of links that cannot transmit when link (i, j) is active. As compared to individual links in traditional wired network, the notation of cluster can be treated as a basic resource unit. Wireless data flows compete for the capacity of individual cluster that is equivalent to the capacity of the wireless shared medium. Hence, the wireless network channel capacity constraint [21] is

$$0 \leq \frac{f_{(i,j)}}{1-\rho_{(i,j)}} + \sum_{(p,q)\in\Psi_{(i,j)}} \frac{f_{(p,q)}}{1-\rho_{(p,q)}} \leq C_{(i,j)}, \forall (i,j) \in E \quad (1)$$

where $f_{(i,j)}$ denotes the physical flow rate that is required to transmit through link (i, j), $C_{(i,j)}$ is defined as the maximum rate of link (i, j) and its corresponding cluster $\Psi_{(i,j)}$ supported by the wireless shared medium, and $\rho_{(i,j)}$ is assumed to be the packet loss probability at wireless link (i, j). Theoretically, this packet loss rate (PLR) can be derived from the Gilbert–Elliott model [22], [23].

B. SVC Coding Model

The layered scalability is adopted here, and assume that the SVC video stream is encoded into a set of M layers $\{L_1, L_2, \ldots, L_M\}$ with a predefined encoding rate based on the network condition. According to the encoding rates of layers, we can make the optimal adaptation decision in the scalability cube model illustrated in Fig. 1 by mapping from an SVC elementary stream with fully scalable representation into the layered representation. Correspondingly, the multicast of SVC video stream is divided into M multicast sessions. Each multicast session m has one source node s, a set of destination nodes D, and a set of relay nodes N. In order to successfully decode received SVC video streams, we should make sure that all destination nodes are able to subscribe to SVC multicast layers in an incremental order, since layer $m+1$ is not decodable without its previous layers 1 to m. In accordance with the SVC layer dependency constraint, lower layers need be received before higher layers for each destination node to promise the most efficient SVC decoding.

Considering both fluctuating network adaptation and optimization condition [27], [28], each layer is distributed over a multicast session at a variable transmission rate within a tolerable rate region $[r^m_{\min}, r^m_{\max}]$. Mathematically, the upper bound $r^m_{\max}$ (e.g., the encoding rate with a resilient margin) and the lower bound $r^m_{\min}$ (e.g., the minimum partial margin for layer m) are specified for a confidence interval of the layered transmission rate in layer m. It differentiates the layers with

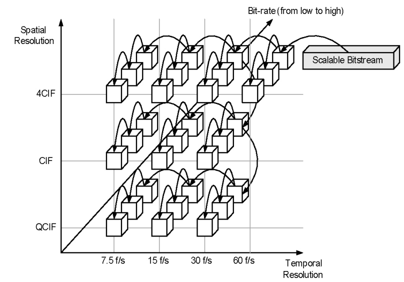


Fig. 1. Typical structure of scalable video bitstream with multiple dimensions.

the piecewise confidence intervals along the layer-dependent direction, namely, the achievable transmission rate for each layer is mathematically extended from an encoding rate point to a tolerable rate region. From the layered optimization, the fine-granular continuity of the targeted variables (rates) could specifically urge the convexity of optimization problem for developing a distributed solution, and support a strong notion of fairness.

C. Network Coding Model

To transmit multiple multicast sessions over a shared network, we might perform network coding across session to achieve the optimal throughput. However, combining data belonging to different layers makes it difficult to recover all original data for destination nodes that only receive partial layers. Thus, network coding is limited within each session in this paper. This approach is often referred to as intra-session coding or superposition coding [26].

For each multicast session, we find multiple paths from the source node to destination nodes using existing multipath routing schemes [29], [30], which are chosen based on general cost criteria that are independent of flow rates. Since each multicast session uses only a limited set of paths, it is shown that such approach may give lower rates compared with optimizing over the entire network, but it is much less complex [9]. For each destination node $d \in D$, we use a matrix $\mathbf{H_d} = \{h^l_{dj}\}$ to represent the relationship between its transmission paths and corresponding links. More specifically, suppose destination node d has $J(d)$ alternative paths from source node s, then $h^l_{dj} = 1$ if the path j of node d uses link l, and $h^l_{dj} = 0$ otherwise. In network coding-based routing, let R^m_{dj} denote the information flow rate of destination node ds jth path in multicast session m, and f^m_l represent the physical flow rate for link l in multicast session m.

With intra-session network coding, flows to different destinations of a multicast session are allowed to share network capacity by being coded together. Therefore, it is only required to set the actual physical flow on each link to be the maximum of the individual destinations' information flows, and accordingly the network coding constraint is

$$\sum_{j=1}^{J(d)} h^l_{dj} R^m_{dj} \leq f^m_l, \forall m \in M, \forall l \in E, \forall d \in D. \quad (2)$$

D. RD Model

From the perspective of application-layer quality of service, RD-related models [31] could be picked as the optimized targeted utility for video applications as follows:

$$D_e(R_e) = \frac{\theta}{R_e - R_0} + D_0 \tag{3}$$

where D_e is the distortion of the encoded video sequence and R_e is the encoded rate. The remaining variables θ, R_0, and D_0 are the parameters of the RD model, which can be fitted to empirical data from trial encodings using nonlinear regression techniques.

To characterize the video streaming performance of each layer m, we introduce a utility function $U_m(\cdot)$, which is continuously differentiable, increasing, and strictly concave with respect to the receiving rate. In this paper, we multicast the video streams to all destination nodes and attempt to maximize the total utility of all recipients, and accordingly the objective function is

$$\max \sum_{d\in D}\sum_{m\in M} U_m(R_d^m) = \max \sum_{d\in D}\sum_{m\in M} U_m\left(\sum_{j=1}^{J(d)} R_{dj}^m\right) \tag{4}$$

where R_d^m denotes the received rate at destination node d in multicast session m.

Within the context of SVC, the utility function $U_m(R_d^m)$ is defined for destination node d to represent the corresponding distortion decrement when a new layer m is successfully received and decoded as follows:

$$U_m\left(R_d^m\right) = -\left[D_e\left(\sum_{i=0}^{m-1} R_d^i + R_d^m\right) - D_e\left(\sum_{i=0}^{m-1} R_d^i\right)\right]. \tag{5}$$

E. Nominal Optimization Problem Formulation

In SVC streaming problems, layered multirate multicasting is equivalent to a generalized multisource problem where the interlayer dependency is considered. The proposed nominal optimization problem integrates the prior context of source decomposition into the layered multirate multicast optimization, which can be formulated as follows:

$$\mathbf{P1}: \quad \max_{\mathbf{R}} \quad \sum_{d\in D}\sum_{m\in M} U_m\left(\sum_{j=1}^{J(d)} R_{dj}^m\right) \tag{6}$$

s.t.

1) $\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m \leq f_l^m; \forall m \in M, \forall l \in E, \forall d \in D$

2) $\sum_{m\in M} \frac{f_l^m}{(1-\rho_l)} + \sum_{k\in\Psi(l)}\sum_{m\in M} \frac{f_k^m}{(1-\rho_k)} \leq C_l; \forall l \in E$

3) $r_{\min}^m \leq \sum_{j=1}^{J(d)} R_{dj}^m \leq r_{\max}^m,\ \ or\ \ \sum_{j=1}^{J(d)} R_{dj}^m = 0; \forall m \in M, \forall d \in D$

4) $\frac{\sum_{j=1}^{J(d)} R_{dj}^m}{r_{\min}^m} \geq \frac{\sum_{j=1}^{J(d)} R_{dj}^{(m+1)}}{r_{\max}^{(m+1)}}; \forall m \in \{1, 2, \ldots, M-1\}, \forall d \in D$

5) $R_{dj}^m \geq 0; \forall j \in J(d), \forall m \in M, \forall d \in D$

6) $f_l^m \geq 0; \forall l \in E, \forall m \in M.$

Constraint 1) specifies the required physical flow rate on each link for each layer under the network coding condition. With network coding, different destinations will not compete for link bandwidth within the same layer, therefore the required physical flow rate on link l for layer m is the largest information flow rate on link l consumed among all destination nodes. Note that the impact of network coding is embedded in this constraint. Constraint 2) characterizes the wireless link contention in a shared medium. For each wireless link l, the sum of ls actual physical flow rate and the actual physical flow rates of links in $\Psi(l)$ cannot exceed the maximum rate C_l. Constraint 3) gives the lower bound and upper bound of the transmission rates allocated for layer m, denoted by $r_{\min}^m$, and $r_{\max}^m$, respectively. The SVC layer dependency constraint is taken into account and promised by Constraints 3) and 4), according to [14, Prop. 1]. Constraints 5) and 6) specify that the allocated rates and required physical flow rates are nonnegative.

To make sure the convexity of the proposed optimization problem **P1**, Constraint 3) needs to be redefined to meet the convexity requirement. Based on the nonnegativity Constraint 5), it can directly imply $\sum_{j=1}^{J(d)} R_{dj}^m \geq 0$ from $R_{dj}^m \geq 0$. Hereby, we can simply extend the second equality term $\sum_{j=1}^{J(d)} R_{dj}^m = 0$ in Constraint 3) to $\sum_{j=1}^{J(d)} R_{dj}^m \leq 0$ because $\sum_{j=1}^{J(d)} R_{dj}^m = 0$ can be promised along with the nonnegativity Constraint 5). Therefore, Constraint 3) is formalized as $r_{\min}^m \leq \sum_{j=1}^{J(d)} R_{dj}^m \leq r_{\max}^m$, *or* $\sum_{j=1}^{J(d)} R_{dj}^m \leq 0$, and further simplified as a cubic inequality $(\sum_{j=1}^{J(d)} R_{dj}^m)(\sum_{j=1}^{J(d)} R_{dj}^m - r_{\min}^m)(\sum_{j=1}^{J(d)} R_{dj}^m - r_{\max}^m) \leq 0$.

We can see that the nominal optimization problem **P1** has a unique optimal solution since its objective function is strictly concave and the solution space defined by the constraints is convex. In other words, this is a convex optimization problem with either centralized or distributed feasible solutions under the assumption that the input parameters are precisely known and equal to some nominal values. This approach, however, does not take into account the influence of data uncertainties on the quality and feasibility of the practical problem. It is therefore noticeable that as the parameters take values different from the nominal ones, several constraints may be violated, which may further lead the optimal solutions obtained by the nominal problem to no longer optimal or even infeasible ones. To tackle the optimization problems under data uncertainty in practice, in the following sections, we will extend the nominal problem with deterministic parameters into a robust optimization problem that is still feasible in practical wireless networks, and accordingly develop a distributed solution to the proposed robust optimization problem.

IV. Robust Multipath Rate Control for SVC Multicast Streaming

In this section, we consider a practical wireless network where some link failures might occur due to wireless channels and mobility in wireless nodes [11]. To enhance the robustness of the wireless communication network, an efficient way is to reserve partial bandwidth for backup paths disjoint from the primal paths during the transmission of SVC multicast stream, such that a partial or all of the traffic would be rerouted to the corresponding disjoint backup paths when the primary paths fail. Here we consider the robust multipath rate-control and bandwidth reservation problem for SVC multicast streaming when possible link failures of primary paths exist, to ensure service feasibility and availability in the presence of link failures. In the meanwhile, an uncertainty set of the wireless medium capacity is introduced to represent the uncertain and time-varying property of parameters related to the wireless channel, in accordance with the fluctuation and perturbation in practical wireless networks.

A. Preliminaries to Distributed Robust Optimization

According to the form of the nominal optimization problem **P1**, we focus on a class of optimization problems with the following nominal form: maximization of a concave objective function over a given data set characterized by linear constraints as follows:

$$\begin{aligned} \max_{\mathbf{x}} \quad & U(x) \\ \text{s.t.} \quad & Ax \preceq c \end{aligned} \tag{7}$$

where A is an $M \times N$ matrix, x is an $N \times 1$ vector, c is an $M \times 1$ vector, and $Ax \preceq c$ denotes that $c - Ax$ is componentwise nonnegative.

The uncertainty of (7) may exist in the objective function $U(x)$, the matrix parameter A, and the vector parameter c. It is demonstrated by [16] that the uncertainty in the objective function $U(x)$ can be converted into uncertainty of the parameters defining the constraints. In the next section, we show that it is also possible to convert the uncertainty in c into uncertainty in A in certain cases. In the remainder of this section, therefore, we focus on studying the uncertainty in the matrix parameter A, the structures and physical meaning of which can readily lead to distributed algorithms in many networking problems. Furthermore, quantifying the tradeoff between robustness and distributedness is also studied in this paper.

In the robust counterpart of (7), the constraint $Ax \preceq c$ is required to be valid for any $A \in \mathcal{A}$, where $\mathcal{A}$ denotes the uncertainty set of A. In this case, the definition of robustness is the worst-case robustness [34], since the solution of the robust optimization problem is always feasible. However, this uncertainty definition may be too conservative. A more meaningful choice of robustness is the chance-constrained robustness, where the probability of infeasibility is upper bounded. By solving the worst-case robust optimization problem over a properly selected subset of the exact uncertainty set, the chance-constrained robustness of the robust solution can be flexibly adjusted.

If $\mathcal{A}$ is assumed to be an arbitrary uncertainty set, the robust optimization problem is difficult to solve even in a centralized manner [15]. Therefore, we focus on the study of constraint-wise (row-wise) uncertainty set, where the uncertainties between different rows in matrix parameter A are decoupled. Though restricted, this type of uncertainty set characterizes the data uncertainty in many practical problems, and also leads the robust optimization problem to a distributively solvable formulation.

To introduce uncertainty in the matrix parameter A, denote the jth row of A by a_j^T, which lies in a compact uncertainty set $\mathcal{A}j^T$. Then the robust optimization formulation of (7) can be written as follows:

$$\begin{aligned} \max_{\mathbf{x}} \quad & U(x) \\ \text{s.t.} \quad & a_j^T x \le c_j;\ \forall a_j \in \mathcal{A}_j,\ \forall 1 \le j \le M. \end{aligned} \tag{8}$$

It is shown by [17] that the uncertainty sets in robust optimization problem (8) can be equivalently written in the form of protection functions. Denote the matrix parameter in the nominal counterpart of (7) as a coefficient matrix $\bar{A}$ with the jth row's coefficient $\bar{a}_j \in \mathcal{A}_j$ when there is no uncertainty, we have the following proposition as proven in [17].

Proposition 1: Problem (8) is equivalent to the following convex optimization problem:

$$\begin{aligned} \max_{\mathbf{x}} \quad & U(x) \\ \text{s.t.} \quad & \bar{a}_j^T x + g_j(x) \le c_j;\quad \forall 1 \le j \le M \end{aligned} \tag{9}$$

where

$$g_j(x) = \sup_{a_j \in \mathcal{A}_j} (a_j - \bar{a}_j)^T x \tag{10}$$

is the protection function for the jth constraint depending on both the uncertainty set $\mathcal{A}_j$ and the nominal row coefficient $\bar{a}_j$. Each g_j is a convex function.

B. Robust Formulation

As the extension of the nominal problem with deterministic parameters as proposed in Section III, the robust formulation of multipath rate-control and bandwidth reservation problem for SVC multicast streaming is taken into account. First, to ensure robust data transmission against the wireless link failures and thus enhance the robustness of the wireless communication network, each destination node d also determines a backup path when it joins the network, in addition to the set of $J(d)$ primary paths. To take advantage of network coding, the path-overlapping allocation of backup paths for different receivers is considered. The nonnegative backup path choice matrix is denoted as

$$\mathbf{B_d} = \{b_d^l\},\ \begin{cases} b_d^l \in [0, B_d], & \text{link } l \text{ is on } d\text{s backup path} \\ b_d^l = 0, & \text{otherwise} \end{cases}$$

where $B_d > 0$ indicates the maximum percentage that destination node d allocates its rate to the backup path. The actual rate percentage of backup link b_d^l could be a random variable between 0 and B_d, depending on whether the primary paths fail. Furthermore, we assume that a path can only be selected as either a primary path or a backup path but not both for the same destination node.

Second, due to the fluctuation and perturbation in practical wireless networks, the rate of each wireless link and its corresponding cluster supported by the wireless shared medium are time-varying and cannot achieve the maximum rate C_l most of the time. To represent the uncertain property of parameters related to the wireless channel, we introduce a perturbation factor $\delta \in [\Delta, 1]$, which indicates the instant percentage of maximum rate C_l that can be practically achieved by the wireless channel. Therefore, the actual rate of wireless link l and its corresponding cluster $\Psi(l)$ fluctuates between the lower bound ΔC_l and the upper bound C_l. Accordingly, the robust multipath routing rate allocation problem for SVC multicast streaming is given by

$$\mathbf{P2}: \quad \max_{\mathbf{R}} \quad \sum_{d\in D}\sum_{m\in M} U_m \left(\sum_{j=1}^{J(d)} R_{dj}^m\right) \tag{11}$$

s.t.

1) $\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + g_l^m(\mathbf{B_v}, \mathbf{R}) \le f_l^m; \forall m \in M, \forall l \in E, \forall d \in D$

2) $\sum_{m\in M} \frac{f_l^m}{(1-\rho_l)} + \sum_{k\in\Psi(l)}\sum_{m\in M} \frac{f_k^m}{(1-\rho_k)} \le \delta C_l; \forall \delta \in [\Delta, 1],$

$\forall l \in E$and 3), 4), 5), 6) in **P1**

where $g_l^m(\mathbf{B_v}, \mathbf{R})$ is the protection function for the traffic from a set of destination nodes $\{v\} \subseteq D$ whose backup paths use link l.

Stipulating that nature will be restricted in such behavior that only a subset of destination nodes will suffer from link or path failures in order to adversely affect the solution, we propose a method similar to the D-norm approach [16], [35] to model the protection function $g_l^m(\mathbf{B_v}, \mathbf{R})$. Denote the notation $\mathcal{D}_\Gamma$ as the subset of destination nodes that might experience link or path failures, and Γ represents the number of destination nodes within this subset, i.e., $\mathcal{D}_\Gamma \subseteq D$ and $|\mathcal{D}_\Gamma| = \Gamma$. It is noted that the definition of robustness associated with the D-norm formulation is to maintain feasibility if at most Γ out of all possible $|D|$ destination nodes experience link or path failures. Considering network coding condition that different destination nodes will not compete for link bandwidth within the same layer, the protection function can be written as

$$g_l^m(\mathbf{B_v}, \mathbf{R}) = \max_{\mathcal{D}_\Gamma \subseteq D} \left\{ \max_{v\in\mathcal{D}_\Gamma} b_v^l (\sum_{j=1}^{J(v)} R_{vj}^m) \right\}. \tag{12}$$

In (12), Γ is a parameter instead of a optimization variable. If $\Gamma = 0$, then $g_l^m(\mathbf{B_v}, \mathbf{R}) = 0$, and the corresponding constraint is reduced to the nominal constraint. On the contrary, if $\Gamma = |D|$, then $g_l^m(\mathbf{B_v}, \mathbf{R}) = \max_{v\in D} b_v^l(\sum_{j=1}^{J(v)} R_{vj}^m)$ and the corresponding constraint becomes the worst-case formulation [16]. Therefore, the value of Γ theoretically takes control of the tradeoff between robustness and performance.

Centralized algorithms can be used to solve the convex optimization problem **P2**. In practice, however, such solutions require global information and coordination between all nodes and links, which is very costly and sometimes infeasible [34], [36]. In comparison, distributed algorithms can be used to dynamically adjust the rates in accordance with changes in the network. In the following section, we will develop a distributed solution of the proposed multipath rate-control problem. Since robustness, i.e., service reliability, is taken into account in this application, it is most likely that such a change and corresponding update are only required infrequently.

V. Distributed Algorithm

In this section, we develop a distributed solution to the proposed robust optimization problem **P2** that allows each node and link to control and update the transmission parameters by itself.

A. Equivalent Formulation

Observing **P2**, it is uncertain that Constraints 1) and 2) prevent the robust optimization problem from a fully distributed solution. We first show that the perturbation factor δ in Constraint 2) can be moved into the left part and thus the uncertain Constraint 2) is converted to a linear constraint with certain parameters.

Proposition 2: For any link $l \in E$, the constraint

$$\sum_{m\in M} \frac{f_l^m}{(1-\rho_l)} + \sum_{k\in\Psi(l)}\sum_{m\in M} \frac{f_k^m}{(1-\rho_k)} \le \delta C_l, \forall \delta \in [\Delta, 1] \tag{13}$$

is equivalent to the following constraint:

$$\sum_{m\in M} \frac{f_l^m}{\Delta(1-\rho_l)} + \sum_{k\in\Psi(l)}\sum_{m\in M} \frac{f_k^m}{\Delta(1-\rho_k)} \le C_l. \tag{14}$$

Proof: See Appendix A. ■

As for Constraint 1), it is shown by Proposition 3 that this nonlinear constraint with the protection function $g_l^m(\mathbf{B_v}, \mathbf{R})$ can be replaced by a set of linear constraints.

Proposition 3: For any link $l \in E$, any destination node $d \in D$, and any layer $m \in M$, the single constraint

$$\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + g_l^m(\mathbf{B_v}, \mathbf{R}) \le f_l^m \tag{15}$$

is equivalent to the following set of constraints:

$$\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + b_v^l \left(\sum_{j=1}^{J(v)} R_{vj}^m\right) \le f_l^m, \forall v \in \mathcal{D}_\Gamma, \forall \mathcal{D}_\Gamma \subseteq D. \tag{16}$$

Proof: See Appendix B. ■

Note that the number of constraints in (16) is $\binom{|D|}{\Gamma} \cdot \Gamma$, and increases quickly with $|D|$ and Γ. However, due to the special structure of the protection function that utilizes network coding, most of the constraints in (16) should be inactive and thus redundant. This motivates us to formulate the protection function $g_l^m(\mathbf{B_v}, \mathbf{R})$ into another set of linear constraints, the number of which is much less.

Proposition 4: For any link $l \in E$, any destination node $d \in D$, and any layer $m \in M$, the single constraint (15) is equivalent to the following set of constraints:

$$\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + b_v^l \left(\sum_{j=1}^{J(v)} R_{vj}^m \right) \leq f_l^m, \forall v \in D. \tag{17}$$

Proof: See Appendix C. ∎

It can be observed that the set of constraints (17) no longer relies on $\mathcal{D}_\Gamma$ and Γ. Propositions 2 and 4 together transform the robust optimization problem **P2** into a simplified formulation **P3** independent of uncertainty δ, Γ, and $\mathcal{D}_\Gamma$ as follows:

$$\mathbf{P3}: \quad \max_{\mathbf{R}} \quad \sum_{d\in D}\sum_{m\in M} U_m \left(\sum_{j=1}^{J(d)} R_{dj}^m \right) \tag{18}$$

s.t.

$$1) \sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + b_v^l \left(\sum_{j=1}^{J(v)} R_{vj}^m \right) \leq f_l^m; \forall v, d \in D, \forall m \in M, \forall l \in E$$

$$2) \sum_{m\in M} \frac{f_l^m}{\Delta(1-\rho_l)} + \sum_{k\in\Psi(l)}\sum_{m\in M} \frac{f_k^m}{\Delta(1-\rho_k)} \leq C_l; \forall l \in E$$

and 3), 4), 5), 6) in **P1**.

To analyze the impact of the protection parameters, B_d denotes the maximum backup percentage of destination node ds total rate on the backup path. If $B_d = 0$, then there is no rate allocated for the backup paths, and Constraint 1) in **P3** is reduced to the nominal constraint. As B_d increases from 0 to 1, the protection rate of backup path becomes larger, which may cause the overall allocated rates for destination nodes to decrease. Similarly, Δ reflects the extent of fluctuation of wireless links' maximum transmission rate. When $\Delta = 1$, the wireless channel condition is determined with known maximum transmission rate, and Constraint 2) in **P3** is reduced to the nominal constraint. As Δ decreases, the protection level for wireless channel's fluctuation becomes greater, and accordingly the overall allocated rates for destination nodes may decrease. Therefore, B_d and Δ can control the tradeoff between robustness and performance in **P3**.

B. Dual Decomposition

Decomposition theories provide a mathematical foundation for the design of modularized and distributed control of networks [36]. The decomposition procedure aims to decompose a large and complex optimization problem into a set of small subproblems, which can be then solved by distributed and often iterative algorithms that converge to the global optimum. It is observed that problem **P3** can be decoupled using dual decomposition. More specifically, by relaxing the coupling Constraints 1)–4) in **P3** with Lagrange multipliers λ, μ, η, and

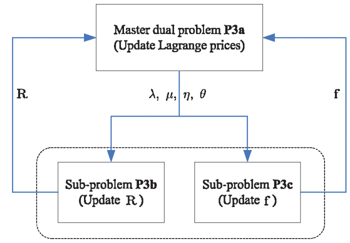


Fig. 2. Schematic diagram of dual decomposition.

θ, respectively, the Lagrangian of **P3** is obtained as follows:

$$\begin{aligned} L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta) = & \sum_{d\in D}\sum_{m\in M} U_m \left(\sum_{j=1}^{J(d)} R_{dj}^m \right) \\ & - \sum_{l\in E}\sum_{d\in D}\sum_{m\in M}\sum_{v\in D} \lambda_{d(v)}^{ml} \left[\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + b_v^l \left(\sum_{j=1}^{J(v)} R_{vj}^m \right) - f_l^m \right] \\ & - \sum_{l\in E} \mu_l \left[\sum_{m\in M} \frac{f_l^m}{\Delta(1-\rho_l)} + \sum_{k\in\Psi(l)}\sum_{m\in M} \frac{f_k^m}{\Delta(1-\rho_k)} - C_l \right] \\ & - \sum_{d\in D}\sum_{m\in M} \eta_d^m \left[\left(\sum_{j=1}^{J(d)} R_{dj}^m \right) \left(\sum_{j=1}^{J(d)} R_{dj}^m - r_{\min}^m \right) \left(\sum_{j=1}^{J(d)} R_{dj}^m - r_{\max}^m \right) \right] \\ & - \sum_{d\in D}\sum_{m=1}^{M-1} \theta_d^m \left[\frac{\sum_{j=1}^{J(d)} R_{dj}^{(m+1)}}{r_{\max}^{(m+1)}} - \frac{\sum_{j=1}^{J(d)} R_{dj}^m}{r_{\min}^m} \right] \end{aligned} \tag{19}$$

and the corresponding Lagrange dual function is

$$g(\lambda, \mu, \eta, \theta) = \sup_{\mathbf{R}\succeq 0,\ \mathbf{f}\succeq 0} L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta). \tag{20}$$

Then the Lagrange dual problem of **P3** can be formulated as

$$\min_{\lambda\succeq 0,\ \mu\succeq 0,\ \eta\succeq 0,\ \theta\succeq 0} g(\lambda, \mu, \eta, \theta). \tag{21}$$

According to convex optimization theories [36], [37], if the original problem **P3** is convex, it is equivalent to its Lagrange dual problem in (21). Then, **P3** can be decomposed into a master dual problem **P3a** and a set of subproblems **P3b–P3c** that can be solved in a distributed manner as follows:

$$\mathbf{P3a}: \quad \min_{\lambda,\mu,\eta,\theta} \quad g(\lambda, \mu, \eta, \theta) \tag{22}$$

$$\text{s.t.} \quad \lambda \succeq 0,\ \mu \succeq 0,\ \eta \succeq 0,\ \theta \succeq 0$$

$$\mathbf{P3b}: \quad \max_{\mathbf{R}} \quad U_b(\mathbf{R}) \tag{23}$$

$$\text{s.t.} \quad R_{dj}^m \geq 0; \forall j \in J(d), \forall m \in M, \forall d \in D$$

$$\mathbf{P3c}: \quad \max_{\mathbf{f}} \quad U_c(\mathbf{f}) \tag{24}$$

$$\text{s.t.} \quad f_l^m \geq 0; \forall l \in E, \forall m \in M$$

where $L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta) = U_b(\mathbf{R}) + U_c(\mathbf{f})$, $U_b(\mathbf{R})$ and $U_c(\mathbf{f})$ denote the summation of terms relating to $\mathbf{R}$ and $\mathbf{f}$ in (19), respectively.

As illustrated by the relationship in Fig. 2, subproblems **P3b** and **P3c** are controlled by the master dual problem

P3a through Lagrange prices. In the meanwhile, they can conversely affect the master dual problem through primal variables. At the lower level, subproblems **P3b** for each d, j, and m, and **P3c** for each l and m can be solved separately. Accordingly, the primal variables **R** and **f** are updated, respectively, and then sent to the master dual problem **P3a**. At the higher level, the master dual problem **P3a** aims to update Lagrange prices (dual variables) λ, μ, η, and θ, which are further used for the solution of **P3b** and **P3c**.

Since the objective functions of the master dual problem **P3a** and subproblems **P3b**–**P3c** are differentiable with respect to the dual variables λ, μ, η, and θ, and primal variables **R** and **f**, all problems can be solved by the gradient algorithm [38], [39]. Based on this observation, we propose the following primal-dual algorithm that updates the primal and dual variables simultaneously to solve the optimization problem **P3**:

$$R_{dj}^{m}(t+1) = \left[R_{dj}^{m}(t) + \alpha(t)\frac{\partial L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta)}{\partial R_{dj}^{m}}\right]^{+} \tag{25}$$

$$f_{l}^{m}(t+1) = \left[f_{l}^{m}(t) + \alpha(t)\frac{\partial L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta)}{\partial f_{l}^{m}}\right]^{+} \tag{26}$$

$$\lambda_{d(v)}^{ml}(t+1) = \left[\lambda_{d(v)}^{ml}(t) - \alpha(t)\frac{\partial L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta)}{\partial \lambda_{d(v)}^{ml}}\right]^{+} \tag{27}$$

$$\mu_{l}(t+1) = \left[\mu_{l}(t) - \alpha(t)\frac{\partial L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta)}{\partial \mu_{l}}\right]^{+} \tag{28}$$

$$\eta_{d}^{m}(t+1) = \left[\eta_{d}^{m}(t) - \alpha(t)\frac{\partial L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta)}{\partial \eta_{d}^{m}}\right]^{+} \tag{29}$$

$$\theta_{d}^{m}(t+1) = [\theta_{d}^{m}(t) - \alpha(t)\frac{\partial L(\mathbf{R}, \mathbf{f}, \lambda, \mu, \eta, \theta)}{\partial \theta_{d}^{m}}]^{+} \tag{30}$$

where t denotes the iteration index, $\alpha(t)$ are positive step sizes, and $[\cdot]^{+}$ denotes the projection onto the set of nonnegative real numbers.

In terms of physical meanings of Lagrange multipliers, λ represents the "congestion prices" of information flow at wireless links, i.e., $\lambda_{d(v)}^{ml}$ can be considered as the "congestion price" of information flow at link l for destination node d (l is on ds primary paths) and v (l is on vs backup path)s bandwidth requirement in layer m. At link l, if the total information flow bandwidth demand $\sum_{j=1}^{J(d)} h_{dj}^{l} R_{dj}^{m} + b_{v}^{l}(\sum_{j=1}^{J(v)} R_{vj}^{m})$ in layer m exceeds the supply f_{l}^{m}, then the "congestion price" $\lambda_{d(v)}^{ml}$ will increase. Accordingly, in problem **P3b**, R_{dj}^{m} will decrease in order to meet the link's bandwidth requirement of information flow, f_{l}^{m}, and vice versa, and μ specifies the "aggregated congestion prices" of physical flows at wireless links. At link l, if the total physical flow bandwidth demand $\sum_{m\in M} \frac{f_{l}^{m}}{\Delta(1-\rho_{l})} + \sum_{k\in\Psi(l)} \sum_{m\in M} \frac{f_{k}^{m}}{\Delta(1-\rho_{k})}$ exceeds the supply C_{l}, then the "aggregate congestion price" μ_{l} will increase. Consequently, f_{l}^{m} in problem **P3c** will decrease in order to meet the link's bandwidth supply of physical flow, C_{l}, and vice versa. Similarly, the other two Lagrange multipliers, η and θ, can be considered as the "SVC encoding prices" for each destination node in a multicast session. Furthermore, all updating operations are distributed and can be implemented at individual links and nodes using only local information.

Algorithm 1 Distributed optimization algorithm

Initialization

Set $t = 0$, and set $R_{dj}^{m}(0)$, $f_{l}^{m}(0)$, $\lambda_{d(v)}^{ml}(0)$, $\mu_{l}(0)$, $\eta_{d}^{m}(0)$, and $\theta_{d}^{m}(0)$ to some nonnegative value for all d, v, m, l, and j.

repeat

Update at wireless link $l \in E$**:**

1) Receives $f_{k}^{m}(t)$ from the cluster $\{k|k \in \Psi(l)\}$;
2) Receives $\mu_{k}(t)$ from the cluster $\{k|k \in \Phi(l)\}$;
3) Receives $R_{dj}^{m}(t)$ from $\{d \in D|h_{dj}^{l} = 1 \text{ or } b_{d}^{l} > 0\}$;
4) Fetches $\lambda_{d(v)}^{ml}(t)$, $\mu_{l}(t)$, $f_{l}^{m}(t)$ stored in local processors;
5) Updates $f_{l}^{m}(t)$, $\lambda_{d(v)}^{ml}(t)$, and $\mu_{l}(t)$ by (26)–(28), respectively;
6) Transmits the control packet (CP) comprising the updated physical flow rate $f_{l}^{m}(t+1)$ to the cluster $\{k|k \in \Phi(l)\}$;
7) Transmits CP that comprises the updated price $\mu_{l}(t+1)$ to the cluster $\{k|k \in \Psi(l)\}$;
8) Sends the CP that comprises updated price $\lambda_{d(v)}^{ml}(t+1)$ in the downstream direction to the subset $\{d \in D|h_{dj}^{l} = 1\} \cup \{v \in D|b_{v}^{l} > 0\}$.

Update at destination node $d \in D$**:**

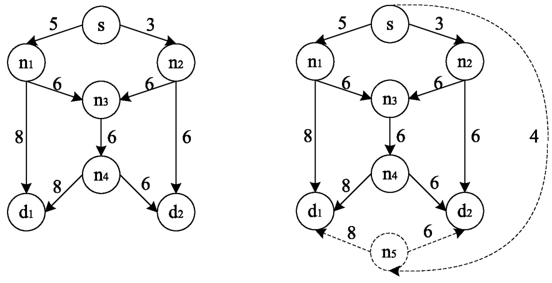
1) Receives $\lambda_{d(v)}^{ml}(t)$ from $\{l \in E|h_{dj}^{l} = 1 \text{ or } b_{v}^{l} > 0\}$;
2) Fetches $R_{dj}^{m}(t)$, $\eta_{d}^{m}(t)$, $\theta_{d}^{m}(t)$ stored in local processors;
3) Updates $R_{dj}^{m}(t)$, $\eta_{d}^{m}(t)$, and $\theta_{d}^{m}(t)$ by (25)–(30), respectively;
4) Sends the rate packet (RP) comprising $R_{dj}^{m}(t+1)$ in the upstream direction to $\{l \in E|h_{dj}^{l} = 1 \text{ or } b_{d}^{l} > 0\}$.

until All variables converge to the optimums or the maximum iteration number is achieved.

The convergence behavior of the PA can be analyzed by applying Lyapunov stability theorem [14], [38], [40]. Regarding the proposed primal-dual algorithm outlined in (25)–(30) as a nonlinear autonomous system, it can be shown in a similar way as in [14] that the equilibrium point of this dynamic system is asymptotically stable. Therefore, the global asymptotic stability of the primal and dual variables in (25)–(30) can be guaranteed, which leads to the convergence behavior of the distributed solution to the dual problem of **P3**. Since **P3** is convex, we can solve it through the equivalent dual problem using the proposed distributed algorithms [36].

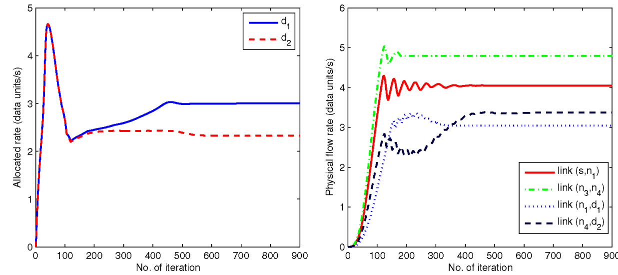
C. Implementation Issue

To implement the proposed distributed algorithm, each link l or destination node d is treated as an entity capable of processing, storing, and communicating information. In practice, each link $l = (i, j)$ is delegated to its sender node i, and all computations related to link $l = (i, j)$ will be executed on node i. Here, we assume that the processor for link l keeps track of variables f_{l}^{m}, $\lambda_{d(v)}^{ml}$, and μ_{l}, while the processor of



(a) (b)

Fig. 3. Network topologies associated with maximum transmission rates, where (a) is a typical butterfly topology for numerical experiment and (b) is with additional backup path.



(a) (b)

Fig. 4. Convergence performance of the PA. (a) Allocated rate for d_1 and d_2. (b) Physical flow rate for wireless links.

destination node d keeps track of variables R^m_{dj}, η^m_d, and θ^m_d. A distributed implementation of the proposed iterative algorithm is summarized in Algorithm 1.

In summary, the centralized approach requires all of the above primal and dual variables shared in the entire network and thus causes a great amount of communication overhead. Utilizing the proposed distributed algorithm, however, the communication overhead only comprises the sending overhead of $\lambda^{ml}_{d(v)}(t+1)$, $R^m_{dj}(t+1)$ and the transmitting overhead of $f^m_l(t+1)$, $\mu_l(t+1)$ at each iteration.

In terms of amount of message passing, the overhead of the proposed distributed algorithm consists of two parts: the network coding overhead and the communication overhead. According to [8], the side information required by network coding is the h-dimensional global encoding vector in the header of each packet, where h is the number of source packets that need to be transmitted. Therefore, the cost of the practical network coding scheme is the overhead of transmitting h extra symbols in each packet. Considering a large size packet with the number of both header and payload symbols far more than h, such overhead is approximately negligible. For example, in the Internet, a typical maximum packet size excluding headers is somewhat larger than 1400 bytes. Thus, each IP packet can carry about 1400 symbols if each symbol takes up 1 byte. If h is 50, then the overhead introduced by the side information of network coding is approximately $50/1400 \approx 3\%$.

On the other hand, the communication overhead is the CP and RP information at each iteration. Consider the implementation issues [27] and take Fig. 3(a), e.g., at the end of each iteration, wireless link (n_1, n_3) needs only to send its CP downward to destination nodes d_1, d_2 and to wireless link (n_2, n_3), while destination node d_1 needs only to send its RP upward to the links belonging to its two primal paths and one backup path. Supposing each updated primal or dual variable is float type that takes up 4 byte, and let $M = 3$, then the CP of link (n_1, n_3) requires $3 \times 2 \times 2 \times 4 + 1 \times 4 + 1 \times 3 \times 4 = 64$ bytes and the RP of node d_1 requires $3 \times 2 \times 4 = 24$ bytes. Thus, the total communication overhead sums up to 88 bytes. For the same configuration of IP packet with a packet size of 1400 bytes, the communication overhead introduced by the PA is $88/1400 = 6\%$. Furthermore, it can be noted that these packets (CPs and RPs) in practical implementation need not be communicated as separate packets; the CPs can be conveyed through a field in the video data packets, while the RPs can be conveyed through a field in the acknowledgement (ACK) packets. The maximal additional delay introduced by sending these packets is the one way propagation delay of the particular multicast destination node.

When a wireless network initializes multicast sessions, the proposed distributed algorithm would run and determine the optimal resource allocation rates and routes for all destination nodes through the convergence time. Within this duration, the total overhead comprises both network coding overhead and the communication overhead. Once it converges to the optimal values, the entire network will work with certain tolerance to dynamic network changes with only network coding overhead.

VI. Experimental Results

In this section, we will evaluate the overall performance of the PA. Both numerical and packet-level simulations are conducted over a typical butterfly network and a large-scale wireless network, respectively. The corresponding simulation results evaluate the convergence behavior of the proposed distributed algorithm, and further demonstrate that the PA can promise the optimal overall video quality for all destination nodes with appropriate protection against dynamic network changes.

A. Algorithm Behavior and Performance Evaluations

First, we implement numerical experiments and evaluate the PA in the wireless network with a simple but classical butterfly topology, as illustrated in Fig. 3(a). Here, s, n_i, and d_i represent source node, relay node, and destination node, respectively. The maximum transmission rates of wireless links are shown in Fig. 3(a). Suppose that the packet loss probability of each wireless link is $\rho_l = 0.1$, and set the protection parameters $B_d = 0.5$, $\Delta = 0.9$.

In the numerical experiments, we assume that the video bit stream has three layers, with the base layer at a rate of 3 (data units/s), the first enhancement layer at a rate of 2 (data units/s), and the second enhancement layer at a rate of 1 (data unit/s). In Fig. 3(a), it is obvious that each destination node has three alternative paths from the source node. We specify two primal paths for each destination node, which are $s \to n_1 \to d_1$, $s \to n_2 \to n_3 \to n_4 \to d_1$, and $s \to n_2 \to d_2$, $s \to n_1 \to n_3 \to n_4 \to d_2$, respectively. Accordingly, the rest two paths,

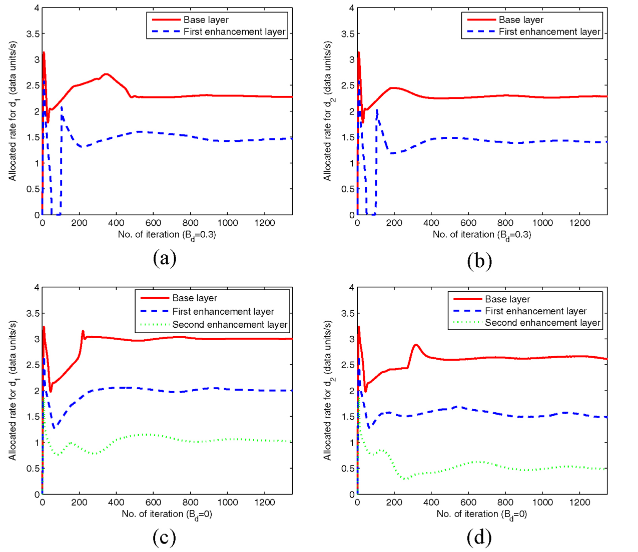


Fig. 5. Convergence performance of the PA, the maximum backup percentage B_d is set to 0.3 in (a) and (b), and 0 in (c) and (d).

$s \to n_1 \to n_3 \to n_4 \to d_1$ and $s \to n_2 \to n_3 \to n_4 \to d_2$, are set to be d_1s and d_2s backup paths.

1) *Convergence Behavior of the PA:* In accordance with subproblems **P3b** and **P3c**, Fig. 4 shows the convergence behavior of the two primal optimization variables $\mathbf{R}$ and $\mathbf{f}$ at a fixed step size 0.01. For example, Fig. 4(a) illustrates the iterations of allocated rates for both destination nodes and Fig. 4(b) represents the iterations of physical flow rates of four wireless links.

We observe that both variables $\mathbf{R}$ and $\mathbf{f}$ converge to the optimal values after 450 iterations, which implies that the PA converges very quickly. For example, the allocated rate for d_1 approaches within 10% of its optimal value after 400 iterations and achieves the optimal value 3.006 after 450 iterations. The physical flow rate of link (s, n_1) approaches within 10% of its optimal value after 100 iterations and achieves the optimal value 4.048 after 400 iterations. In fact, it can also be observed that in all cases the convergence to 90% optimality is much faster, with at least 20% fewer iterations than those required for convergence to optimality. Therefore, we can obtain a feasible solution to a certain degree of optimality in a much shorter convergence time.

Furthermore, it is noted in Fig. 4 that since the protection level is relatively high (i.e., $B_d = 0.5$), d_1 receives the base layer, while d_2 can only subscribe to the base layer partially with a rate of 2.3 (data units/s). To show the convergence behavior of the PA at different levels of protection, Fig. 5 presents the iterations of allocated rates for both destination nodes with $B_d = 0.3$ and $B_d = 0$, respectively. Similarly as in Fig. 4, all allocated rates in Fig. 5 also converge to the optimal values after several hundred iterations. As B_d reduces, the protection level becomes lower, and accordingly both destination nodes can receive more SVC video stream layers. For example, when $B_d = 0.3$, both destination nodes partially subscribe to base layer and the first enhancement layer. When $B_d = 0$, d_1 fully receives all the three layers, while d_2 partially subscribes to the three layers.

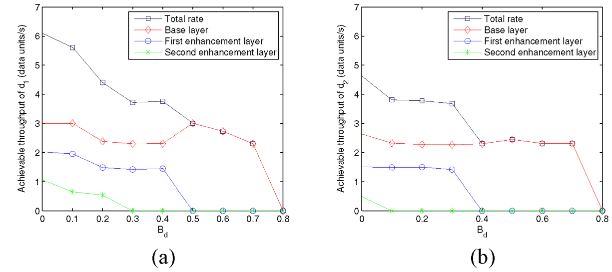


Fig. 6. Impact of maximum backup percentage B_d on allocated rate for (a) d_1 and (b) d_2.

2) *Impact of Maximum Backup Percentage Parameters:* Next, we study the impact of the maximum backup percentage parameter B_d and the corresponding results (in forms of the total allocated rate of all three layers, and the allocated rate for base layer, the first enhancement layer, the second enhancement layer) are shown in Fig. 6. Here, we vary B_d from 0 to 1. Fig. 6(a) and (b) illustrates the impact of B_d on the allocated rates for d_1 and d_2, respectively. When B_d increases from 0 to 1, the protection level of the PA against link failures becomes higher, meanwhile, the increment of protection rate of each backup path causes the capacities of links remaining for primal paths to decrease. Therefore, it can be seen from Fig. 6 that the total allocated rates of all SVC layers for both d_1 and d_2 decrease as a result of the increment of B_d.

It should be noted in Fig. 6(a) that the allocated rate for d_1s base layer dose not always decrease as the increment of B_d. For example, when B_d varies from 0.4 to 0.5, d_1s allocated rate of base layer increases from 2.3 to 3. This "unexpected" increasing is mainly because the PA ensures that every destination node receives as many SVC layers as possible to achieve the maximum total allocated rate. Therefore, at $B_d = 0.4$, d_1 subscribes to the base layer partially at a rate of 2.3 (data units/s) in order to make the rest bandwidth enough for receiving the first enhancement layer. As B_d increases to 0.5, the first enhancement layer cannot be received even if d_1 subscribes to the base layer at the minimum partial margin, however, there is still sufficient bandwidth for fully receiving the base layer at a rate of 3 (data units/s), and the total allocated rate of all layers for d_1 does decrease when B_d varies from 0.4 to 0.5.

Furthermore, in Fig. 6(b), as B_d varies from 0.1 to 0.3 (or 0.4–0.7), the total allocated rate of all layers for d_2 remains approximately the same value. This observation occurs when d_2s reception of lower layers is not affected by the bandwidth reduction caused by the increment of B_d. For example, at $B_d = 0.1$, d_2 receives the lower two layers since the remaining bandwidth cannot support the subscription to the second enhancement layer. As B_d increases to 0.3, this remaining bandwidth is accordingly reduced, but there is still sufficient bandwidth for the reception of the lower two layers. When B_d further increases to 0.4, d_2 can only receive the base layer since the reception of the first enhancement layer is not supported by the remaining bandwidth.

3) *Impact of Different Backup Path Selections:* As mentioned in Section IV-B, the backup path of one destination

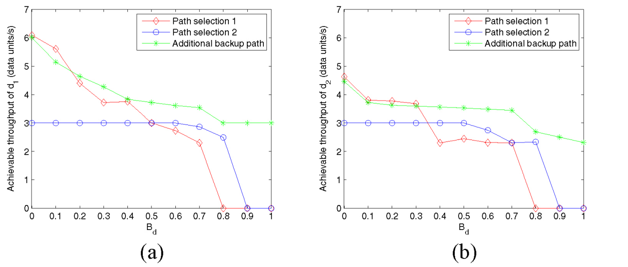


Fig. 7. Allocated rate for (a) d_1 and (b) d_2 with different backup path selections.

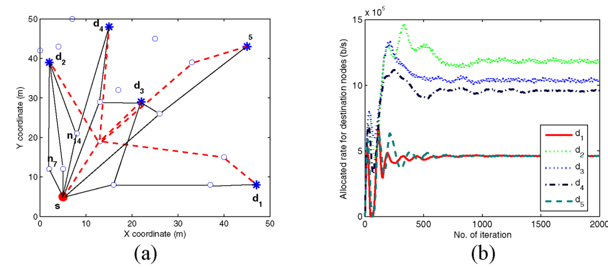


Fig. 8. (a) Network topology in packet-level simulation. (b) Convergence of allocated rate for d_1–d_5 with the PA.

node is selected in such way that the probability of overlapping with the backup paths of other destination nodes is high, so that network coding can be applied. We compare different choices of backup paths for d_1 and d_2, and their impact on the rate allocation is shown in Fig. 7. Here, the aforementioned primal and backup path selection is denoted as "path selection 1," and "path selection 2" represents a different choice of primal and backup paths, i.e., two primal paths for each destination node are $s \to n_1 \to n_3 \to n_4 \to d_1$, $s \to n_2 \to n_3 \to n_4 \to d_1$, and $s \to n_1 \to n_3 \to n_4 \to d_2$, $s \to n_2 \to n_3 \to n_4 \to d_2$, respectively, while $s \to n_1 \to d_1$ and $s \to n_2 \to d_2$ are set to be d_1 and d_2s backup paths without overlapping. With this path selection, network coding can be implemented neither on primal paths nor on backup paths. Therefore, compared with "path selection 2," "path selection 1" with overlapped backup paths generally leads to greater allocated rates for both destination nodes, especially when B_d is relatively small. When B_d increases to a larger value (e.g., greater than 0.4), the allocated rates with "path selection 1" may be less than "path selection 2," since at this time the backup paths of "path selection 1" cannot afford such protection rates.

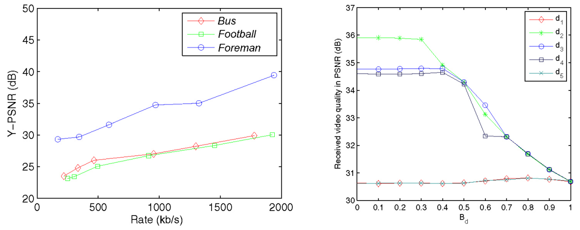
In order to further study the impact of different backup path selections. As illustrated in Fig. 3(b), we set a new relay node n_5 and specify the backup paths as additional paths $s \to n_5 \to d_1$ and $s \to n_5 \to d_2$, which are denoted by dashed lines. Note that these two additional backup paths are overlapped and totally disjoint from the primal paths. Accordingly, this choice of backup path selection is denoted as "additional backup path," whose primal paths are still the same as in "path selection 1." It can be seen from Fig. 7 that when B_d is relatively small, with both backup path selection ("additional backup path" and "path selection 1"), the allocated rates for both destination nodes are similar. However, when B_d increases to a larger value (e.g., greater than 0.4), the allocated rates with "additional backup path" are greater, since the two additional backup paths are disjoint from primal paths and thus have sufficient bandwidth for larger protection rates. In conclusion, to promise the application of network coding and sufficient bandwidth for protection rate allocation, it is encouraged to make the backup paths of different destination nodes overlapped and disjoint from primal paths.

B. Packet Level Simulations for Larger Scale Network

To evaluate the effect for a more general network, the packet-level simulation is conducted on the wireless network as shown in Fig. 8(a), which is generated by 20 randomly placed nodes in a 50 m-by-50 m area. Assume every two nodes within a distance of 30 m are able to communicate, and suppose the maximum transmission rate of each wireless link is 600 kb/s, and the wireless capacity in a shared medium is 3 Mb/s. Also set the protection parameters $B_d = 0.4$, $\Delta = 0.9$. In Fig. 8(a), one node is selected as source node s, and five nodes are set to be destination nodes d_1–d_5. For each destination node, the number of primal paths is randomly set between 1 and 3, and the one additional backup path is selected to be overlapped with other backup paths and disjoint from primal paths, which is signified with red dashed links.

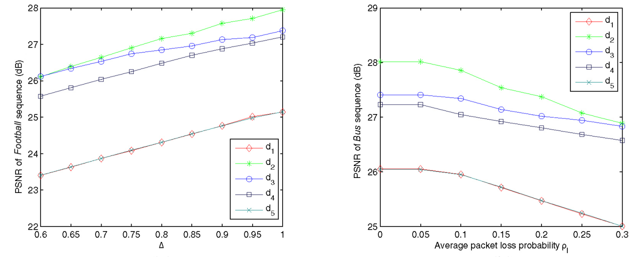
In the packet-level simulations, practical random network coding [8] is used to distribute the source packets of each layer. Here, we assume intra-session network coding is implemented within each layer to ensure easy operation. During data transmission, each relay node (as well as the source node) combines its received packets from different upstream links (or video source packets encoded by the source node) with random linear operations over a large Galois field and then sends the coded packets to its downstream links. Each destination node can correctly decode the original packets if it receives enough coded packets. To cope with asynchronous transmission, we use the buffer model [8] to synchronize the packet arrivals and departures. In the buffer model, packets that arrive at a node on any of the incoming links are put into a single buffer sorted by the layer number. Whenever there is a transmission opportunity at an outgoing link, the number of packets of every layer in the buffer is checked and a packet is generated containing a random linear combination of all the packets that belong to the layer with the largest number of packets. After the generated packet is transmitted to the outgoing link, certain old packets are flushed from the buffer according to the flushing policy. Specially, if two layers have the same number of packets in the buffer, the lower layer is prioritized to generate a packet for transmission.

We use Joint Scalable Video Model 7_10 reference codec of H.264/AVC extension standard [41], with three well-known test sequences (*Bus*, *Football*, and *Foreman*) at frame rate of 30 f/s, CIF (352 × 288) resolution, and a group of picture-length of 32 frames. They are encoded with 256 kb/s on the base layer, and 384 kb/s, 512 kb/s, and 1024 kb/s on the three enhancement layers by fine granularity scalable coding. Fig. 9(a) shows the RD performance [peak signal-to-noise ratio (PSNR)] of SVC for the three test video sequences obtained in the packet-level simulation.



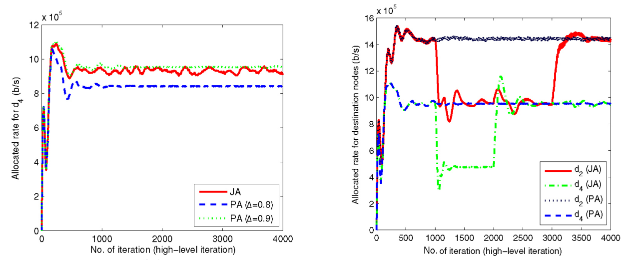
(a) (b)

Fig. 9. (a) PSNR performance achieved by SVC stream for three test video sequences. (b) Received video quality in PSNR of *Foreman* sequence for d_1–d_5 when B_d varies.



(a) (b)

Fig. 10. Impact of (a) Δ and (b) ρ_l on received video quality in PSNR for d_1–d_5.



(a) (b)

Fig. 11. Comparison of the convergence behavior when network fluctuation occurs. (a) Dynamic change of the maximum rate of wireless links. (b) Wireless link fails and reconnects.

Fig. 8(b) illustrates the iterations of allocated rates for destination nodes d_1–d_5. Theoretically, since the PA is implemented in a distributed manner that only requires local information, the communication overhead and thus convergence behavior would not be affected very much by larger scale of networks. In comparison with the convergence behavior of simple butterfly topology as shown in Figs. 4 and 5, for instance, although the number of iterations needed for convergence in Fig. 8(b) is moderately larger, it is still less than 1000 and only takes several hundred iterations to achieve convergence, which means the convergence speed slows down slightly as the scale of networks increases. Therefore, the impact on the system performance affected by larger scale networks is relatively small.

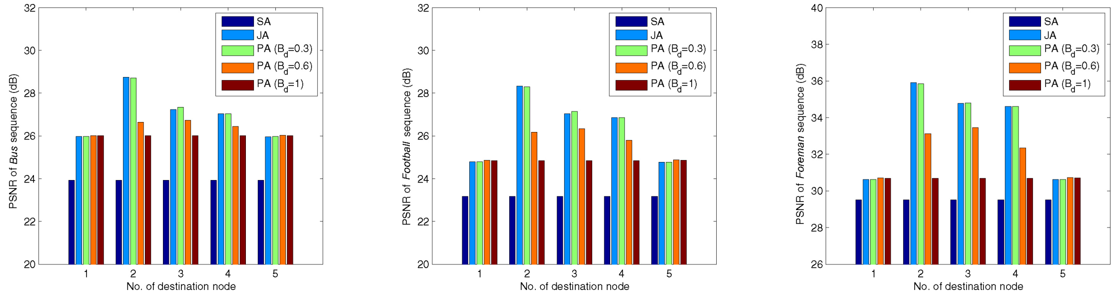
1) *Impact of Protection Parameters and PLR:* As analyzed in Section V-A, the protection parameters B_d and Δ would determine the protection level of the PA and thus adjust the tradeoff between robustness and optimization performance. In Fig. 9(b), we illustrate the received video quality of *Foreman* sequence of d_1–d_5 when the maximum backup percentage parameter B_d varies. It can be seen that the results in Fig. 9(b) conforms the same trend as that in Fig. 6, i.e., the overall received video quality of all destination nodes will decrease as the increment of B_d. In detail, when B_d is relatively small (e.g., not greater than 0.3), the backup path of each destination node can afford the corresponding protection rate and thus the received video quality would not be affected by varying B_d. Also, since only one primal path is specified for d_1 and d_5, and the protection rate on their backup paths can always be allocated when B_d varies from 0 to 1. Therefore, these two destination nodes would achieve approximately the same received video quality when B_d changes.

Fig. 10(a) reflects the impact of maximum wireless link fluctuation parameter Δ on the received video quality of *Football* sequence of d_1–d_5. Here, Δ changes from 0.6 to 1. When $\Delta = 1$, the wireless channel condition is determined with known maximum transmission rate, and all destination nodes achieve best overall video quality without any protection against wireless links' fluctuation. As the decrement of Δ, the protection level for wireless channel's fluctuation becomes higher, and accordingly the overall received video quality of all destination nodes will decrease.

It is noted that all the aforementioned results are obtained under an average PLR of each wireless link, and the impact of PLR of wireless links would be further validated here. By varying the average packet loss rate of each wireless link from 0% to 30%, the overall received video quality of all destination nodes solved by the PA is shown in Fig. 10(b). It can be seen that as the average packet loss rate increases, the received video quality of *Bus* sequence of each destination node will accordingly decreases, because of the reduction of available bandwidth for both primal and backup paths.

2) *Adaptation to Dynamic Network Changes:* Next, we study the adaptation of the proposed robust algorithm to practical networks with dynamic network changes, and further compare the PA with the Joint optimization Algorithm (JA) without protection in [14]. On the basis of Fig. 9(b), here B_d is set to 0.3 such that the proposed algorithm is implemented with both good performance and certain level of protection.

We study the convergence behavior of the proposed distributed algorithm under a dynamic fluctuation (perturbation factor) of C_l (the maximum transmission rates of wireless links) in Fig. 11(a). For each wireless link with theoretical maximum transmission rate C_l, suppose practically its maximum transmission rate fluctuates between $[0.8 \cdot C_l, C_l]$ (i.e., [480 kb/s, 600 kb/s]) in realistic situation. Both algorithms (PA and JA) start the implementation at iteration 0. It is noted that here one iteration of JA means one high-level iteration in its two level optimization structure. It can be observed that although the actual maximum transmission rate of each wireless link is time-varying, the allocated rate for d_4 by the proposed algorithm with two setups of Δ ($\Delta = 0.8$ and $\Delta = 0.9$) both converge smoothly to the corresponding steady states after several hundred iterations. In comparison, the allocated rate obtained by JA cannot converge to a steady state due to the perturbation of maximum transmission rates of wireless links, and is observed to fluctuate in accordance



(a) (b) (c)

Fig. 12. Comparison of received video quality in PSNR of (a) *Bus*, (b) *Football*, and (c) *Foreman* sequences for all destination nodes.

with such perturbation. The fluctuation of allocated rate solved by JA is upper bounded by the convergence rate achieved by PA with $\Delta = 0.9$ and lower bounded by the convergence rate achieved by PA with $\Delta = 0.8$. Therefore, although d_4s allocated rate by JA is slightly greater than by PA with $\Delta = 0.8$, the fluctuation property would make it inapplicable for smoothness requirement of video quality.

The adaptation of the proposed algorithm to a dynamic change of network topology is shown in Fig. 11(b), involving iterations (high-level iterations) of the allocated rates for d_2 and d_4. At iteration 1000, one of d_2s and d_4s primal paths fails, respectively, since the abrupt departure of node n_{14} and thus failure of wireless link (s, n_{14}). At iteration 2000, node n_{14} rejoins, however, node n_7 leaves the network, which causes one of d_2s primal paths to fail. At iteration 3000, node n_7 also rejoins the network. It can be seen that the allocated rates for d_2 and d_4 by PA both converge smoothly to the corresponding steady states after several hundred iterations. After convergence, both allocated rates would not affected by the sudden network topology changes (e.g., link or path failures) occurring at iteration 1000, 2000, and 3000. However, JA need recalculate the allocated rates for both d_2 and d_4 after either abrupt change, in order to reconverge to a new steady state. In some cases (e.g., d_2s allocated rate by JA between iteration 1000 and 2000), JA may even not approach to a steady state, since the time interval between two sudden changes is not sufficient for convergence. Therefore, JA is infeasible in practice where the sudden changes of network topology (e.g., link or path failures) occur frequently.

In conclusion, PA can adapt to dynamic changes and thus applicable in practical networks. Furthermore, it should be noted that in Fig. 11 we only show the high-level optimization iterations of JA's two level optimization structure. In fact, in order to implement one high-level iteration, it is required by JA to run a number of low-level iterations first. Therefore, even if in Fig. 11 it is seen that both PA and JA are with similar convergence speed, the actual convergence speed of JA is much slower than PA.

3) *Comparison of With or Without Protection:* By introducing protection against link failures and wireless links' fluctuation, and thus robustness to dynamic network changes, the proposed algorithm achieves the optimal tradeoff between robustness and optimization performance. We compare the proposed algorithm with algorithms without protection. The comparison result is shown in Fig. 12, by the received video quality of all three test sequences of all destination nodes. Within each destination node, from left to right, shown are the PSNR obtained by the shortest path algorithm (SA), the JA in [14] and the PA with different values of B_d. In accordance with the results shown in Fig. 9(b), here B_d is set to 0.3, 0.6, and 1, respectively. It can be seen that SA obtains the worst results while JA achieves the best performance. Generally, the performance of PA is lower bounded by SA, and upper bounded by JA. However, JA is a deterministic optimization formulation and thus not feasible for practical networks. In fact, the implementation of JA comprises two levels of optimization iterations, which cause the convergence speed to be much slower than PA. It can also be seen that when B_d is small, e.g., $B_d = 0.3$, PA obtains the same results as JA, but with certain reserved rates on backup paths in case of link failures of primal paths. Therefore, PA can be applied in practical networks with protection against dynamic network changes, and achieve the optimal tradeoff between optimization performance and robustness, according to requirements of different protection levels.

VII. Conclusion

Aiming at practical applications of wireless scalable video streaming, this paper investigated a robust optimization problem for the overall video quality and traffic performance of scalable video multirate multicast in wireless networks with network coding. To make the nominal convex optimization formulation robust, we reserved partial bandwidth for backup paths in case of possible link failures of primary paths. It considered the path-overlapping allocation of backup paths for different receivers to take advantage of network coding, and also an uncertainty set of the wireless medium capacity to represent the uncertain and time-varying property of the wireless channel. Through protection functions with nonlinear constraints for the targeted uncertainty, we studied the tradeoff between robustness and distributedness. A fully decentralized algorithm has also been provided to attain a practical solution from the dual decomposition and primal-dual update method. Through extensive experimental results under critical factors, the proposed algorithm has been validated to adapt to the dynamic network changes in an optimal tradeoff between optimization performance and robustness.

Appendix A
Proof of Proposition 2

If (13) is true, let $\delta = \Delta \in [\Delta, 1]$, then we have

$$\sum_{m\in M} \frac{f_l^m}{(1-\rho_l)} + \sum_{k\in\Psi(l)} \sum_{m\in M} \frac{f_k^m}{(1-\rho_k)} \le \Delta C_l \tag{31}$$

and thus we have (14). On the other hand, if $\mathbf{f}$ satisfies (14), then $\forall \Delta \le \delta \le 1$, we have

$$\sum_{m\in M} \frac{f_l^m}{(1-\rho_l)} + \sum_{k\in\Psi(l)} \sum_{m\in M} \frac{f_k^m}{(1-\rho_k)} \le \Delta C_l \le \delta C_l \tag{32}$$

i.e., it also satisfies (13). Therefore, these two constraints are equivalent.

Appendix B
Proof of Proposition 3

From the definition (12), we know that $b_v^l(\sum_{j=1}^{J(v)} R_{vj}^m) \le g_l^m(\mathbf{B_v}, \mathbf{R})$ for all $v \in \mathcal{D}_\Gamma \subseteq D$. If $\mathbf{R}$ satisfies (15), then $\forall v \in \mathcal{D}_\Gamma$ and $\forall \mathcal{D}_\Gamma \subseteq D$, we have

$$\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + b_v^l(\sum_{j=1}^{J(v)} R_{vj}^m) \le \sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + g_l^m(\mathbf{B_v}, \mathbf{R}) \le f_l^m \tag{33}$$

i.e., it also satisfies the set of constraints in (16). On the other hand, if (16) is true, then we have

$$\begin{aligned}
&\sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + g_l^m(\mathbf{B_v}, \mathbf{R}) \\
&= \sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + \max_{\mathcal{D}_\Gamma \subseteq D} \left\{ \max_{v\in\mathcal{D}_\Gamma} b_v^l(\sum_{j=1}^{J(v)} R_{vj}^m) \right\} \\
&\le \max_{\mathcal{D}_\Gamma \subseteq D} \max_{v\in\mathcal{D}_\Gamma} \left\{ \sum_{j=1}^{J(d)} h_{dj}^l R_{dj}^m + b_v^l(\sum_{j=1}^{J(v)} R_{vj}^m) \right\} \le f_l^m.
\end{aligned} \tag{34}$$

Therefore, these two constraints are equivalent.

Appendix C
Proof of Proposition 4

Given that the following two sets of destination nodes are equivalent: $\{v | \forall v \in \mathcal{D}_\Gamma, \forall \mathcal{D}_\Gamma \subseteq D\} = D$, we know (16) and (17) are equivalent. Therefore, according to Proposition 3, single constraint (15) is equivalent to (17).

References

[1] X. Zhu and B. Girod, "Distributed media-aware rate allocation for wireless video streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1462–1474, Nov. 2010.

[2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[3] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[5] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 789–804, Mar. 2006.

[6] D. S. Lun, M. Médard, and M. Effros, "On coding for reliable communication over packet networks," in *Proc. Allerton Conf. Commun., Contr., Comput.*, 2004, pp. 20–29.

[7] T. Ho, R. Koetter, M. Médard, D. R. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Proc. IEEE Int. Symp. Inform. Theory*, Jun.–Jul. 2003, p. 442.

[8] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. Allerton Conf. Commun., Contr. Comput.*, 2003, pp. 40–49.

[9] L. Chen, T. Ho, S. H. Low, M. Chiang, and J. C. Doyle, "Optimization based rate control for multicast with network coding," in *Proc. IEEE INFOCOM*, May 2007, pp. 1163–1171.

[10] X. Zhu, T Schierl, T. Wiegand, and B. Girod, "Video multicast over wireless mesh networks with scalable video coding (SVC)," in *Proc. Vis. Commun. Image Process.*, Jan. 2008, pp. 1–8.

[11] I. F. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Commun. Mag.*, vol. 43, no. 9, pp. 23–30, Sep. 2005.

[12] J. Zhao, F. Yang, Q. Zhang, Z. Zhang, and F. Zhang, "LION: Layered overlay multicast with network coding," *IEEE Trans. Multimedia*, vol. 8, no. 5, pp. 1021–1032, Oct. 2006.

[13] J. Zou, H. Xiong, L. Song, Z. He, and T. Chen, "Prioritized flow optimization with generalized routing for scalable multi-rate multicasting," in *Proc. IEEE ICC*, Jun. 2009, pp. 1–6.

[14] C. Li, H. Xiong, J. Zou, and Z. He, "Joint source and flow optimization for scalable video multi-rate multicast over hybrid wired/wireless coded networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 550–564, May 2011.

[15] A. Ben-Tal and A. Nemirovski, "Robust solutions to linear programming problems contaminated with uncertain data," *Math. Program.*, vol. 88, no. 3, pp. 411–424, 2000.

[16] D. Bertsimas and M. Sim, "The price of robustness," *Oper. Res.*, vol. 52, no. 1, pp. 35–53, Feb. 2004.

[17] K. Yang, J. Huang, Y. Wu, X. Wang, and M. Chiang. (2010). Distributed robust optimization (DRO) part I: Framework and example [Online]. Available: http://www.princeton.edu/~chiangm/DRO1.pdf

[18] Y. Wu, K. Yang, J. Huang, X. Wang, and M. Chiang. (2010). Distributed robust optimization (DRO) part II: Wireless power control [Online]. Available: http://www.princeton.edu/~chiangm/DRO2.pdf

[19] K. Yuen, B. Liang, and B. Li, "A distributed framework for correlated data gathering in sensor network," *IEEE Trans. Vehicul. Technol.*, vol. 57, no. 1, pp. 578–593, Jan. 2008.

[20] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[21] B. Hamdaoui and P. Ramanathan, "Rate feasibility under medium access contention constraints," in *Proc. IEEE GLOBECOM*, Dec. 2003, pp. 3020–3024.

[22] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, no. 9, pp. 1253–1265, Sep. 1960.

[23] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sep. 1963.

[24] S. Mao, Y. T. Hou, X. Cheng, H. D. Sherali, S. F. Midkiff, and Y. Q. Zhang, "On routing for multiple description video over wireless ad hoc networks," *IEEE Trans. Multimedia*, vol. 8, no. 5, pp. 1063–1074, Oct. 2006.

[25] H. Wu and A. A. Abouzeid, "Error robust image transmission in wireless sensor networks," in *Proc. 5th Workshop ASWN*, Jun. 2005, pp. 1–9.

[26] R. W. Yeung, "Multilevel diversity coding with distortion," *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 412–422, Mar. 1995.

[27] K. Kar, S. Sarkar, and L. Tassiulas, "A scalable low overhead rate control algorithm for multirate multicast sessions," *IEEE J. Sel. Areas Commun. (Special Issue Netw. Support Multicast Commun.)*, vol. 20, no. 8, pp. 1541–1557, Oct. 2002.

[28] K. Kar, S. Sarkar, and L. Tassiulas, "Optimization based rate control for multirate multicast sessions," in *Proc. IEEE INFOCOM*, Apr. 2001, pp. 123–132.

[29] I. Cidon, R. Rom, and Y. Shavitt, "Analysis of multi-path routing," *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 885–896, Dec. 1999.

[30] R. Ogier, V. Rutenburg, and N. Shacham, "Distributed algorithms for computing shortest pairs of disjoint paths," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 443–455, Mar. 1993.

[31] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jun. 2000.

[32] X. Lu, E. Erkip, Y. Wang, and D. Goodman, "Power efficient multimedia communication over wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1738–1751, Dec. 2003.

[33] Z. He and D. Wu, "Resource allocation and performance analysis of wireless video sensors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 590–599, May 2006.

[34] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[35] D. Bertsimas, D. Pachamanovab, and M. Simc, "Robust linear optimization under general norms," *Oper. Res. Lett.*, vol. 32, no. 6, pp. 510–516, Nov. 2004.

[36] D. Palomar and M. Chiang, "A tutorial on decomposition methods and distributed network resource allocation," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[37] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[38] S. Shakkottai and R. Srikant, "Network optimization and control," *Foundations Trends Netw.*, vol. 2, no. 3, pp. 271–379, 2007.

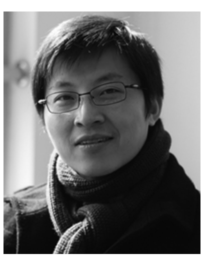
[39] A. Nedic and D. P. Bertsekas, "Incremental subgradient methods for nondifferentiable optimization," *SIAM J. Optimiz.*, vol. 12, no. 1, pp. 109–138, 2001.

[40] H. Khalil, *Nonlinear Systems*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[41] *JSVM 1.0 Software*, document JTC1/SC29/WG11/N6900, ISO/IEC, Hong Kong, China, Jan. 2005.

Chenglin Li received the B.S. and M.S. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2007 and 2009, respectively, where he is currently pursuing the Ph.D. degree.

His current research interests include network-oriented image/video processing and communication, and the network-based optimization for video sources.



Hongkai Xiong (M'01–SM'10) received the Ph.D. degree in communication and information systems from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2003.

Since 2003, he has been with the Department of Electronic Engineering, SJTU, where he is currently an Associate Professor. From December 2007 to December 2008, he was a Research Scholar with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA. At SJTU, he directs the Image, Video, and Multimedia Communications Laboratory and the Multimedia Communication Division of the Key Laboratory of the Ministry of Education of China, Intelligent Computing and Intelligent System, which is also co-granted by Microsoft Research. He has published over 90 refereed journal and conference papers. His current research interests include source coding/network information theory, signal processing, computer vision and graphics, and statistical machine learning.

Dr. Xiong was the recipient of the Top 10% Paper Award for "Super-resolution reconstruction with prior manifold on primitive patches for video compression" at the 2011 IEEE International Workshop on Multimedia Signal Processing, the SMC-A Excellent Young Faculty Award of SJTU, in 2010, and the New Century Excellent Talents in University Award from the Ministry of Education of China, in 2009. He a Technical Program Committee Member of several IEEE conferences. He currently is a Technical Committee Member of Signal Processing of the Shanghai Institute of Electronics.

Junni Zou (M'06) received the M.S. and Ph.D. degrees in communication and information systems from Shanghai University, Shanghai, China, in 2004 and 2006, respectively.

Since 2006, she has been with the School of Communication and Information Engineering, Shanghai University, where she is an Associate Professor. Currently, she is a Research Scholar with the Department of Electrical and Computer Engineering, University of California at San Diego, San Diego. She has published over 30 refereed journal and conference papers. Her current research interests include distributed resource allocation, multimedia communication, and network information theory.

Dr. Zou was a recipient of the Shanghai Rising-Star Award in 2011. She is a Technical Committee Member of Signal Processing with the Shanghai Institute of Electronics.

Chang Wen Chen (F'04) received the B.S. degree from the University of Science and Technology of China, Hefei, Anhui, China, in 1983, the M.S.E.E. degree from the University of Southern California, Los Angeles, in 1986, and the Ph.D. degree from the University of Illinois at Urbana-Champaign, Urbana, in 1992.

Since 2008, he has been a Professor of computer science and engineering with the State University of New York, Buffalo. From 2003 to 2007, he was an Allen S. Henry Distinguished Professor with the Department of Electrical and Computer Engineering, Florida Institute of Technology, Melbourne. He was on the Faculty of Electrical and Computer Engineering, University of Missouri-Columbia, Columbia, from 1996 to 2003, and with the University of Rochester, Rochester, NY, from 1992 to 1996. From 2000 to 2002, he was the Head of the Interactive Media Group, David Sarnoff Research Laboratories, Princeton, NJ. He has also consulted with Kodak Research Laboratories, Microsoft Research, Beijing, China, Mitsubishi Electric Research Laboratories, Cambridge, MA, NASA Goddard Space Flight Center, Greenbelt, MD, and the U.S. Air Force Rome Laboratories, Rome, NY.

Dr. Chen was the Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from January 2006 to December 2009. He was an Editor of the PROCEEDINGS OF THE IEEE, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE MULTIMEDIA, the *Journal of Wireless Communication and Mobile Computing*, the *EUROSIP Journal of Signal Processing: Image Communications*, and the *Journal of Visual Communication and Image Representation*. He has also chaired and served on numerous technical program committees for the IEEE and other international conferences. He was elected a fellow of the SPIE for his contributions in electronic imaging and visual communications.