Cross-Layer Design of Congestion Control and Power Control in Fast-Fading Wireless Networks

Nguyen H. Tran, *Member*, *IEEE*, Choong Seon Hong, *Senior Member*, *IEEE*, and Sungwon Lee, *Member*, *IEEE*

Abstract—We study the cross-layer design of congestion control and power allocation with outage constraint in an interference-limited multihop wireless networks. Using a complete-convexification method, we first propose a message-passing distributed algorithm that can attain the global optimal source rate and link power allocation. Despite the attractiveness of its optimality, this algorithm requires larger message size than that of the conventional scheme, which increases network overheads. Using the bounds on outage probability, we map the outage constraint to an SIR constraint and continue developing a practical near-optimal distributed algorithm requiring only local SIR measurement at link receivers to limit the size of the message. Due to the complicated complete-convexification method, however the congestion control of both algorithms no longer preserves the existing TCP stack. To take into account the TCP stack preserving property, we propose the third algorithm using a successive convex approximation method to iteratively transform the original nonconvex problem into approximated convex problems, then the global optimal solution can converge distributively with message-passing. Thanks to the tightness of the bounds and successive approximations, numerical results show that the gap between three algorithms is almost indistinguishable. Despite the same type of the complete-convexification method, the numerical comparison shows that the second near-optimal scheme has a faster convergence rate than that of the first optimal one, which make the near-optimal scheme more favorable and applicable in practice. Meanwhile, the third optimal scheme also has a faster convergence rate than that of a previous work using logarithm successive approximation method.

Index Terms—Cross-layer design, convex optimization, congestion control, power control

1 Introduction

Wireless ad hoc networks have drawn much of attention with their extensive practical applications, rapid development and deployment in various fields such as wireless sensors, mesh and cognitive networks. Due to their limited resources (i.e., bandwidth, energy, spectrum, etc.) and lack of a central controller, the main challenge is the demand of efficient and fair resource allocation requiring distributed feedback control mechanisms that can react quickly to changing network conditions. In wired networks, distributed control has been applied to Internet congestion control in the form of message-passing dynamic response. Many of works in this track are devoted to the formulation of a network utility maximization (NUM) framework which can be solved implicitly using congestion avoidance mechanism of various transmit control protocols (TCPs) (e.g., [1], [2], [3]) in which link capacities are usually assumed to be known and fixed. However, these results cannot be applied directly to wireless multihop networks because the wireless link capacity depends on the received signal and interference level, which complicates the NUM design and distributed solutions due to the newly added degree of freedom: power control.

Hence, the congestion control and power control have a mutual relationship in wireless multihop network. The congestion control regulates the source rates to avoid overwhelming any link capacity which depends on interference levels, which in turn decided by link transmit power control. The first NUM-based joint congestion control and power control (JCPC) problem was characterized by Chiang [4]. The critical point of Chiang's proposed solution lies in the high-SIR approximation, which enables the transformation of an original nonconvex problem into a convex optimization problem. Using the gradient-based algorithm, the author showed that the optimal end-to-end rate control and power allocation could be achieved in a distributed fashion with message-passing. Each source s receives the aggregate congestion state of all links on its route and then relies on this to adjust its data rate. Link receivers also periodically broadcast the perceived noise measurement, which helps other link transmitters use this information to update their transmit powers.

However, in addition to the unfavorable high-SIR assumption, this work also assumes static fading wireless channels, which means that such an algorithm should be able to update source rates and link powers whenever the fading state changes. This assumption restricts its applicable scope to a slowly varying wireless channel. If we consider a realistic case of fast-fading channel, the update rate must be fast enough to keep track of changing fading

- *The authors are with the Department of Computer Engineering, Kyung Hee University, 1 Seocheon, Giheung, Yongin, Gyeonggy 446-701, Korea. E-mail: {nguyenth, cshong, drsungwon}@khu.ac.kr.*

Manuscript received 17 Oct. 2011; revised 24 Jan. 2012; accepted 14 Mar. 2012; published online 28 Mar. 2012.
Recommended for acceptance by E. Leonardi.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDS-2011-10-0774.
Digital Object Identifier no. 10.1109/TPDS.2012.118.

states. This leads to the extravagant message-passing overhead and the excessive waste of signal processing energy due to frequent iterative updates.

In this paper, we investigate the JCPC problem in an interference-limited multihop network in a dynamic fading environment and with no assumption of high-SIR. Our objective is to maximize the aggregate utilities and minimize the total expended power. We aim to design a resource allocation scheme that does not have to keep track of the instantaneous fading state of the wireless channel. Instead, we allow outages to occur between successive updates; as a result, the updates can proceed on a much slower time scale (i.e., the same time scale as log-normal shadowing variations). We explicitly include the fading-induced outage constraint into the underlying cross-layer NUM problem, where we account for the statistical variation in each link's SIR and allow the SIR to drop below a prescribed threshold with a predetermined probability. Our main contributions are summarized as follows:

- We pose a NUM framework of joint congestion control and power allocation with explicit outage constraint as a nonconvex programming problem. Even though the outage constraint adds complexity to the nonconvex property of this cross-layer NUM, we first transform it to a new equivalent nonlinear programming problem and then show that the new problem is a convex optimization problem.
- Using a complete-convexification method, we propose two message-passing distributed algorithms that solve the newly formulated convex optimization problem.
  - The first algorithm can attain the global optimal source rates and link powers using a dual-based algorithm, which involves dual decomposition and a gradient-type algorithm. However, in order to update the dual variables, this scheme either requires receivers to measure individual power noise from other transmitters, which is impractical, or enlarges the size of the control message, which increases the network overhead. This algorithm serves as a benchmark to access the performance of our second algorithm.
  - By mapping the outage constraint into the SIR constraint using the outage probability bounds in a Rayleigh fading model, we design a second algorithm which is near-optimal but practically implementable. It alleviates the first algorithm's overhead-induced drawbacks in that it needs only the local SIR measurement and a small-size control message as in [4].
- The congestion control mechanism of the first and second algorithms no longer preserves the existing TCP stack like that of the conventional work [4]. We continue developing the third algorithm using a successive convex approximation method to account for the TCP stack preserving property. This method iteratively transforms the original nonconvex problem into approximated convex problems, then the global optimal solution can converge distributively with message-passing.
- In practical systems, the feedback signal is transmitted over a wireless channel and is error-prone due to the channel variations in link quality. Hence, we also examine the convergence behavior of proposed algorithms with regard to the random-error message passing.
- Extensive numerical results show that the performance of three algorithms is almost indistinguishable. Despite the same complete-convexification method, the second design demonstrate a faster convergence rate than the first one. A numerical comparison also shows that the third algorithm converge faster than a previous work using logarithm successive approximation method.

The rest of this paper is organized as follows: in Section 2, we discuss related work and in Section 3 we describe the system model. We present the first proposal which can achieve the optimal solution in Section 4. The second scheme is a near-optimal algorithm and is proposed in Section 5. The third algorithm using successive convex approximation method is presented in Section 6. The convergence analysis with random errors is presented in Section 7. Illustrative numerical results are presented in Section 8, and our concluding remarks are provided in Section 9.

2 Related Work

In the literature, distributed algorithms for cross-layer design have been widely recognized as robust and practical methods to provide the efficiency and fairness of resource allocation in wireless multihop networks (e.g., [5], [6], [7], [8], [9]). Realizing the importance of convexity in this field, many works have employed the transformations of variables to convert the underlying nonconvex problems into the convex counterparts to facilitate the optimal algorithm design [5], [7], [9], [10], [11], [12], [13], [14], [15].

The idea of using outage probability constraint to updates network operations on a slower time scale was first considered in [16] to solve a power control problem in a single hop network, where the authors employed the centralized interior-point method for numerical implementation. Another study [17] illustrates the convergence of a power control scheme coupled with an outage constraint and multiuser detection in single hop networks by employing standard interference function proposed by Yates [18]. Our works instead study joint congestion control and power control for multihop networks by designing distributed algorithms with message-passing. Literature [19] may be the first work utilizing an outage constraint to address cross-layer JCPC problem. The authors first consider the rate-outage constraint, then reformulate it as a conventional source-rate constraint with a link outage capacity [20] using the upper bound on outage probability, which turns out to be an approximated optimization problem. Tran and Hong [12] tackled the nonconvexity of JCPC using a successive approximation method without high-SIR assumption, but this work also assumed static fading channel as in [4].

Finally, we summarize the key related models and compare the existing cross-layer JCPC designs in the literature with our three proposals, namely, Algorithm 1 (Alg. 1), Algorithm 2 (Alg. 2), and Algorithm 3 (Alg. 3) in Table 1. All of the compared properties (i.e., high-SIR assumption, implicit or explicit outage constraint and message size) of the existing schemes as well as proposed

TABLE 1
Comparisons of Various JCPC Schemes, Where CC and SA Stand for the Complete-Convexification and Successive Approximation Methods, Respectively

Method	High-SIR assumption		Outage constraint		Message size	
	Yes	No	Implicit	Explicit	Small	Large
CC	[4]	[9], Alg. 1, Alg. 2	[9]	Alg. 1, Alg. 2	[4], [9], Alg. 2	Alg. 1
SA	Unknown	[9], [12], Alg. 3	[9]	Alg. 3	[9], [12]	Alg. 3

schemes are mainly based on the complete-convexification and successive approximation methods. The successive approximation method can preserve the TCP stack via its congestion control mechanism, which cannot be achieved by complete-convexification method.

3 System Model and Problem Formulation

3.1 Network Model

We consider a wireless multihop network with $\mathcal{L} = \{1, \ldots, L\}$ logical links shared by $\mathcal{S} = \{1, \ldots, S\}$ sources. We assume that each source s emits a flow using a fixed set of links $L(s)$ on its route. The set of sources using link l is denoted by $S(l) = \{s \mid l \in L(s)\}$.

In this context, each source s always has data to transmit and it obtains a utility $U_s(x_s)$ when transmitting a flow at data rate x_s. We denote the vector of source rates $\mathbf{x} = [x_1, \ldots, x_S]^T$. The utility function $U_s(x_s)$ is assumed to be twice continuously differentiable, nondecreasing and strictly concave in x_s. A utility can be interpreted as the level of satisfaction attained by a user as a function of the resource allocation. A large class of user fairness can be characterized by the following general α-fair utility function [21]

$$U_s^\alpha(x_s) = \begin{cases} (1-\alpha)^{-1} x_s^{1-\alpha}, & \text{if } \alpha \geq 0, \alpha \neq 1, \\ \log x_s, & \text{if } \alpha = 1. \end{cases} \tag{1}$$

For example, it provides proportional fairness with $\alpha = 1$, harmonic mean fairness with $\alpha = 2$ and max-min fairness with $\alpha \to \infty$.

At the physical layer, we use a similar CDMA physical model to that in [4] where simultaneous communications can occur, resulting in multiple-access interference. The instantaneous capacity of link $l \in \mathcal{L}$ is a global and nonconvex function of link power vector $\mathbf{P} = [P_1, \ldots, P_L]^T$

$$c_l(\gamma_l(\mathbf{P})) = W \log(1 + K\gamma_l(\mathbf{P})), \tag{2}$$

where W is the baseband bandwidth and K is a constant depending on modulation, coding scheme and bit-error rate (BER) [20]. Unless otherwise stated, we assume $W = K = 1$ without loss of generality. $\gamma_l(\mathbf{P})$ is the *instantaneous* SIR of link l which is defined as

$$\gamma_l(\mathbf{P}) = \frac{P_l G_{ll} F_{ll}}{\sum_{k \neq l} P_k G_{lk} F_{lk} + n_l}, \tag{3}$$

where the gain G_{lk} represents a large-scale, slow-fading channel (e.g., distance-dependent path-loss or log-normal shadowing), the gain F_{lk} models a small-scale, fast-fading channel from the transmitter on link k to the receiver on link l, and n_l is the thermal noise power at each receiver of link l.

We assume a nonline-of-sight radio transmission environment among transmitters and receivers. In this case, we can employ a Rayleigh fading model, where exponential random variables F_{lk} are i.i.d. Over the considered time scale, G_{lk} is assumed to be constant. Then, the *certainty-equivalent* SIR is

$$\begin{aligned} \bar{\gamma}_l(\mathbf{P}) &= \frac{E[P_l G_{ll} F_{ll}]}{E\left[\sum_{k \neq l} P_k G_{lk} F_{lk} + n_l\right]} \\ &= \frac{P_l G_{ll}}{\sum_{k \neq l} P_k G_{lk} + n_l}, \end{aligned} \tag{4}$$

which can be interpreted as the link l's SIR by assuming fading-free channels with normalized $E[F_{lk}] = 1, \ \forall k$ [16].

Before proceeding, we introduce the notations relating to the operating ranges of vectors of source rates $\mathbf{x}$ and link powers $\mathbf{P}$ as follows:

$$\mathcal{X} = \{x_s, s \in \mathcal{S} \mid x_s^{min} \leq x_s \leq x_s^{max}\}, \tag{5}$$

$$\mathcal{P} = \{P_l, l \in \mathcal{L} \mid P_l^{min} \leq P_l \leq P_l^{max}\}. \tag{6}$$

3.2 Problem Formulation: JCPC with an Explicit Outage Constraint

It is implicitly understood that the NUM problem of [4] is linked directly with a determined fading state (i.e., both F_{lk} and G_{lk} are fixed). Hence, for every new channel state, any algorithm must be recalculated to determine the new optimal solutions. From a practical viewpoint, this will prohibit the effectiveness of such a message-passing iterative algorithm in a fast-fading channel environment. For example, when the fading rate increases, the iteration rate must also increase in order to keep track of new channel states, thus producing a considerable message-passing overhead so that the scheme is no longer able to track the channel states and collapses. In order to alleviate this problem, instead of seeking optimal source rates and powers based on instantaneous link capacities, we allow the network to experience a tolerable level of outage so that resources can be allocated on a much slower time scale [9], [16]. To account for this issue, we incorporate the outage constraint into the underlying NUM as follows:

$$\begin{aligned} &\underset{\mathbf{x} \in \mathcal{X}, \mathbf{P} \in \mathcal{P}}{\text{maximize}} \sum_s U_s(x_s) - \sum_l P_l \\ &\text{subject to} \sum_{s \in S(l)} x_s \leq c_l(\bar{\gamma}_l(\mathbf{P})) \quad \forall l, \\ &\qquad Pr\left[\gamma_l \leq \gamma_l^{th}\right] \leq \xi_l \quad \forall l, \end{aligned} \tag{7}$$

where $c_l(\bar{\gamma}_l(\mathbf{P})) = \log(1+\bar{\gamma}_l(\mathbf{P}))$, $Pr[\gamma_l \le \gamma_l^{th}]$ is the outage probability defined as the proportion of time that some SIR threshold γ_l^{th} is not met for a sufficient reception at link l's receiver, and $\xi_l \in (0,1)$ is the outage probability threshold on link l. The objective is to maximize the network utility while minimizing the total power. For a Rayleigh fading channel, as in [16], the closed-form outage probability is

$$\Pr[\gamma_l \le \gamma_l^{th}] = 1 - \exp\left(-\frac{n_l \gamma_l^{th}}{P_l G_{ll}}\right) \prod_{k \ne l} \left(1 + \gamma_l^{th} \frac{P_k G_{lk}}{P_l G_{ll}}\right)^{-1}. \quad (8)$$

Then problem (7) can be rewritten as

$$\begin{aligned} &\underset{\mathbf{x}\in\mathcal{X},\mathbf{P}\in\mathcal{P}}{\text{maximize}} \sum_s U_s(x_s) - \sum_l P_l \\ &\text{subject to} \sum_{s\in S(l)} x_s \le c_l(\bar{\gamma}_l(\mathbf{P})) \quad \forall l, \\ &\qquad \prod_{k\ne l}\left(1+\gamma_l^{th}\frac{P_k G_{lk}}{P_l G_{ll}}\right) \le \Omega_l(P_l) \quad \forall l, \end{aligned} \quad (9)$$

where

$$\Omega_l(P_l) = \frac{\exp\big(-n_l\gamma_l^{th}/P_l G_{ll}\big)}{1-\xi_l}. \quad (10)$$

We further assume that γ_l^{th} and ξ_l are chosen such that there exist feasible points in problem (9).

3.3 Prior Treatment: JCPC with an Implicit Outage Constraint

To address the same issue, Papandriopoulos et al. [9] focus on the rate-outage probability $O_l^{rate} = \Pr[\sum_{s\in S(l)} x_s > c_l(\mathbf{P})]$ and apply the upper bound derived in [16] to reformulate their outage constraint as follows:

$$O_l^{rate} \le 1 - \exp\left(-\frac{\exp(\sum_{s\in S(l)} x_s) - 1}{\bar{\gamma}_l(\mathbf{P})}\right) \le \xi_l. \quad (11)$$

Manipulating the second inequality, they form a new source-rate constraint included into the following NUM:

$$\begin{aligned} &\underset{\mathbf{x}\in\mathcal{X},\mathbf{P}\in\mathcal{P}}{\text{maximize}} \quad \sum_s U_s(x_s) - \sum_l P_l \\ &\text{subject to} \quad \sum_{s\in S(l)} x_s \le \log(1-\log(1-\xi_l)\bar{\gamma}_l(\mathbf{P})) \quad \forall l. \end{aligned} \quad (12)$$

The right-hand side of the above constraint is called ξ-outage capacity [20]. Intuitively, in order to deal with the fast-varying channel, the original link capacity has been "lowered" to the ξ-outage capacity to ensure the source-rate control achieves the rate-outage probability target ξ_l.

3.4 Discussion

Technically, problem (12) seems easier to solve than (9), and its algorithm is well presented in [9]. However, there are two reasons advocating us to seek the optimal solutions of (9).

- In problem (9), we can characterize a large class of network QoS by tuning parameter γ_l^{th} (see [22], which showed that a minimum successful frame rate can be converted to an appropriate γ_l^{th} for a specific modulation and coding scheme). With different values of γ_l^{th} on different links, different optimal solutions exist. Problem (12) does not have this property.
- The approximated constraint of problem (12) reduces the size of the original constraint set. Hence, there is no guarantee that its feasible region contains the true optimal points, which may result in suboptimal solutions. Conversely, in problem (9), the right-hand size of the first constraint is the link capacity in true form (Shannon sense), not in approximation form like that of problem (12). And the second constraint is the outage constraint in explicit form using Rayleigh fading model. We clearly see that both constraints of problem (9) are in rightly and explicitly closed-form, so this model guarantees the true optimal solutions with an optimal algorithm.

4 Optimal Algorithm: Complete-Convexification Method

Problem (9) is generally a nonconvex and nonseparable optimization problem. In this section, we first transform (9) into an equivalent convex problem, decompose it into separable congestion and power control problems, and finally present the distributed optimal algorithm.

4.1 Equivalent Convex Formulation

The new variables and sets are denoted as follows:

$$\begin{aligned} &\hat{P}_l = \log P_l, \quad \hat{x}_s = \log x_s, \\ &\hat{\mathcal{X}} = \{\hat{x}_s,\ \forall s\in\mathcal{S}|\ \log x_s^{min} \le \hat{x}_s \le \log x_s^{max}\}, \\ &\hat{\mathcal{P}} = \{\hat{P}_l,\ \forall l\in\mathcal{L}|\ \log P_l^{min} \le \hat{P}_l \le \log P_l^{max}\}. \end{aligned}$$

Also, in order to simplify the notation, henceforth we denote $\bar{\gamma}_l = \bar{\gamma}_l(\mathbf{P})$ and $\hat{\bar{\gamma}}_l = \bar{\gamma}_l(e^{\hat{\mathbf{P}}})$. Then problem (9) is transformed into the following equivalent nonlinear programming problem:

$$\begin{aligned} &\underset{\hat{\mathbf{x}}\in\hat{\mathcal{X}},\hat{\mathbf{P}}\in\hat{\mathcal{P}}}{\text{maximize}} \sum_s U_s(e^{\hat{x}_s}) - \sum_l e^{\hat{P}_l} \\ &\text{subject to} \log\left(\sum_{s\in S(l)} e^{\hat{x}_s}\right) \le \log c_l\big(\hat{\bar{\gamma}}_l\big) \quad \forall l, \\ &\qquad \sum_{k\ne l} \log\left(1+e^{\hat{P}_k-\hat{P}_l}\frac{\gamma_l^{th}G_{lk}}{G_{ll}}\right) \le \log\Omega_l\big(e^{\hat{P}_l}\big) \quad \forall l. \end{aligned} \quad (13)$$

We furthermore assume that the utility function in this context satisfies

$$\frac{d^2U_s(x_s)}{dx_s^2}x_s + \frac{dU_s(x_s)}{dx_s} \le 0, \quad (14)$$

then $U_s(\exp(.))$ is a concave function [23]. We note that this assumption is not restrictive and holds for the large class of α-fair utility functions (1) when $\alpha \ge 1$. Finally, we have the following theorem.

Theorem 1. *Problem (13) is a convex optimization problem.*

Proof. See Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPDS.2012.118. □

4.2 Dual Decomposition and Optimal Solution

Thanks to the separable nature of problem (13), its Lagrangian can be decomposed into two separate partial functions as follows:

$$L(\hat{\mathbf{x}}, \hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda}) + L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu}), \tag{15}$$

where

$$L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda}) = \sum_s U_s(e^{\hat{x}_s}) - \sum_l \lambda_l \log\left(\sum_{s\in S(l)} e^{\hat{x}_s}\right), \tag{16}$$

$$\begin{aligned} L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_l \Bigg(&\lambda_l \log c_l(\hat{\bar{\gamma}}_l) + \nu_l \log \Omega_l(e^{\hat{P}_l}) - e^{\hat{P}_l} \\ &- \nu_l \sum_{k\neq l} \log\left(1 + e^{\hat{P}_k - \hat{P}_l}\frac{\gamma_l^{th} G_{lk}}{G_{ll}}\right)\Bigg). \end{aligned} \tag{17}$$

Here $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_L]^T$ and $\boldsymbol{\nu} = [\nu_1, \ldots, \nu_L]^T$, the Lagrange multipliers of the first and second constraint, are considered the link *congestion price* and *outage price*, respectively, following the spirit of [2]. The partial dual functions can be represented as

$$D_1(\boldsymbol{\lambda}) = \max_{\hat{\mathbf{x}}\in\hat{\mathcal{X}}} \quad L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda}), \tag{18}$$

$$D_2(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \max_{\hat{\mathbf{P}}\in\hat{\mathcal{P}}} \quad L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu}), \tag{19}$$

which are separate maximization problems. We denote $D(\boldsymbol{\lambda}, \boldsymbol{\nu}) = D_1(\boldsymbol{\lambda}) + D_2(\boldsymbol{\lambda}, \boldsymbol{\nu})$. The dual problem of (13) is

$$\min_{(\boldsymbol{\lambda}, \boldsymbol{\nu})\geq 0} D(\boldsymbol{\lambda}, \boldsymbol{\nu}). \tag{20}$$

From Theorem 1, we know that the objective of the primal problem (13) is a concave function and that the constraints are convex. The Slater's constraint qualification holds (which will be shown later), leading to strong duality (i.e., zero duality gap). This allows us to solve the primal (13) via the dual (20) problem using the following iterative algorithm.

Algorithm 1. Optimal JCPC with an Outage Constraint using Complete-Convexification Method
All primal and dual variables update iteratively as follows until the termination criterion is satisfied
Congestion control. The source rate updates

$$x_s(t+1) = \left[U_s'^{-1}(\lambda_s(t))\right]_{x_{min}}^{x_{max}}, \tag{21}$$

where $U_s'^{-1}$ is the inverse of the first derivative of utility and

$$\lambda_s(t) = \sum_{l\in L(s)} \frac{\lambda_l(t)}{\sum_{f\in S(l)} x_f(t)}.$$

Power Control: The link power updates

$$P_l(t+1) = \left[\frac{\delta_l(t) - \nu_l(t)\tilde{m}_l(t)\frac{n_l}{\log(1-\xi_l)}}{1 + \sum_{k\neq l}\left(G_{kl} m_k(t) + \nu_k(t)\frac{G_{kl}\tilde{m}_k(t)}{1+G_{kl}\tilde{m}_k(t)P_l(t)}\right)}\right]_{P_{min}}^{P_{max}}, \tag{22}$$

where

$$\delta_k(t) = \lambda_k(t)\frac{1}{\log(1+\bar{\gamma}_k(t))}\frac{\bar{\gamma}_k(t)}{(1+\bar{\gamma}_k(t))}, \tag{23}$$

$$m_k(t) = \delta_k(t)\frac{\bar{\gamma}_k(t)}{G_{kk}P_k(t)}, \tag{24}$$

$$\tilde{m}_k(t) = \frac{\gamma_k^{th}}{G_{kk}P_k(t)}. \tag{25}$$

Link Congestion Price Update:

$$\lambda_l(t+1) = \left[\lambda_l(t) - \kappa(t)\left(\log c_l(\bar{\gamma}_l(t)) - \log\left(\sum_{s\in S(l)} x_s(t)\right)\right)\right]^+. \tag{26}$$

Link Outage Price Update:

$$\begin{aligned} \nu_l(t+1) = \Bigg[&\nu_l(t) - \kappa(t)\Bigg(\log\Omega_l(P_l(t)) \\ &- \sum_{k\neq l}\log\left(1 + \gamma_l^{th}\frac{G_{lk}P_k(t)}{G_{ll}P_l(t)}\right)\Bigg)\Bigg]^+. \end{aligned} \tag{27}$$

Here, $[a]^+ = \max\{a, 0\}$, $[a]_b^c = \max\{\min\{a, c\}, b\}$ and $\kappa(t)$ is the step size.

Proposition 1. *The dual problem (20) has the strong duality property.*

Proof. Problem (13) is a convex problem according to Theorem 1. There exist strictly feasible points; for example, if we choose

$$\begin{cases} \hat{x}_s = -\infty \quad \forall s \in \mathcal{S}, \\ 0 < \hat{P}_1 = \hat{P}_2 \cdots = \hat{P}_L < \infty, \\ \Big\{(\gamma_l^{th}, \xi_l) \in \mathrm{IR}_+^2, \ \forall l \in \mathcal{L} \big| \sum_{k\neq l}\log\Big(1 + \frac{\gamma_l^{th} G_{lk}}{G_{ll}}\Big) \\ \qquad < \log\Omega_l\Big(e^{\hat{P}_l}\Big)\Big\}. \end{cases}$$

So the Slater's constraint qualification holds [24, pp. 226-227]. □

Proposition 2. *The source rate update (21) solves the maximization problem (18) for a fixed* $(\boldsymbol{\lambda}, \nu)$.

Proof. Because $L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda})$ is a strictly concave function with respect to $\hat{\mathbf{x}}$ for a fixed $\boldsymbol{\lambda}$, the application of the first-order optimal condition results in

$$\frac{\partial L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda})}{\partial \hat{x}_s} = 0 = e^{\hat{x}_s}\left(U_s'(e^{\hat{x}_s}) - \sum_{l\in L(s)}\frac{\lambda_l}{\sum_{f\in S(l)} e^{\hat{x}_f}}\right). \tag{28}$$

The result is then obtained via transformation back to the x-space. □

Proposition 3. *The power update (22) solves the maximization problem (19) for a fixed* $(\boldsymbol{\lambda}, \nu)$.

Proof. See Appendix B, which is available in the online supplemental material. □

Because both the partial Lagrangians $L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda})$ and $L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ are strictly concave, their optimal solutions are unique for a specific $(\boldsymbol{\lambda}, \boldsymbol{\nu})$. The dual functions $D_1(\boldsymbol{\lambda})$ and $D_2(\boldsymbol{\lambda}, \boldsymbol{\nu})$ are differentiable everywhere according to [25, Prop. 6.1.1]. Based on this, we applied the projected gradient method to solve the dual problem (20) using the dual variable updates (26) and (27). We address the convergence of Algorithm 1 in the following theorem.

Theorem 2. *For any initial power $\mathbf{P}(0) \in \mathcal{P}$, source rate $\mathbf{x}(0) \in \mathcal{X}$ and prices $(\boldsymbol{\lambda}(\mathbf{0}), \boldsymbol{\nu}(\mathbf{0})) \geq 0$, the sequence of $\{\mathbf{x}(t), \mathbf{P}(t), \boldsymbol{\lambda}(t), \boldsymbol{\nu}(t)\}$ generated via Algorithm 1 converges to the global optimal points if the step size satisfies $\kappa(t) > 0, \sum_{t=0}^{\infty} \kappa(t)^2 < \infty$ and $\sum_{t=0}^{\infty} \kappa(t) \to \infty$.*

The convergence of Algorithm 1 can be proved using the standard technique of gradient algorithm convergence analysis [26]. Due to the limited space, the proof is not included here.

For the sake of convenience, we use the same step size $\kappa(t)$ for both congestion and outage price updates without loss of generality. If the step sizes are different, the prices can be rescaled with no effect on the algorithm convergence.

Remarks:

1. The congestion control can be implemented distributively using message passing. The destination sends a message back to the source to adjust its rate according to (21), where the message accumulates the congestion prices $\frac{\lambda_l(t)}{\sum_{f \in S(l)} x_f(t)}$ of every intermediate link l along its path to produce a total price $\lambda_s(t)$ at source node s.
2. Link power can also be updated in a distributed fashion through message passing, analogous to the algorithm in [4]. In this manner, each receiver of link k broadcasts its control message containing three real-value fields, $m_k(t)$, $\tilde{m}_k(t)$, and $\nu_k(t)$. Each transmitter of link l then receives these values, estimates G_{kl} by using the training sequences and updates its power according to (22) using both congestion price and outage price.
3. The link congestion price update (26) only needs the link's local information, including the ingress rate and SIR measurement.
4. The link outage price update (27) needs not only the information of its local link power but also the individual received powers of other interfering transmitters. This requires that the receiver be able to individually measure each interfering power, which might be impractical. Another way to solve this issue is to reserve another field (i.e., the fourth field) which contains $P_k(t)$ in the control message broadcast by the receiver of link k.
5. Due to the explicit outage constraint nature of (13), the messages broadcast by receivers contain much information. As a result, the overhead increases and transmitters require more energy to receive these messages and extract their information. In the next section, we eliminate this problem by proposing a near-optimal scheme.

5 Near-Optimal Algorithm: Complete-Convexification Method

Due to the explicit outage constraint nature of (13), the messages broadcast by receivers contain much information. In this section, we eliminate this issue by proposing a near-optimal scheme.

The upper and lower bounds on the outage probability of link l, which were shown in [16], can be obtained as follows:

$$\frac{\gamma_l^{th}}{\bar{\gamma}_l + \gamma_l^{th}} \leq Pr\left[\gamma_l \leq \gamma_l^{th}\right] \leq 1 - \exp\left(-\frac{\gamma_l^{th}}{\bar{\gamma}_l}\right). \tag{29}$$

Applying these bounds to the outage probability constraint, we have

$$Pr\left[\gamma_l \leq \gamma_l^{th}\right] \leq 1 - \exp\left(-\frac{\gamma_l^{th}}{\bar{\gamma}_l}\right) \leq \xi_l, \tag{30}$$

$$\frac{\gamma_l^{th}}{\bar{\gamma}_l + \gamma_l^{th}} \leq Pr\left[\gamma_l \leq \gamma_l^{th}\right] \leq \xi_l, \tag{31}$$

which correspond to these following average SIR constraints:

$$\bar{\gamma}_l \geq -\frac{\gamma_l^{th}}{\log(1-\xi_l)}, \tag{32}$$

$$\bar{\gamma}_l \geq \gamma_l^{th}\left(\frac{1}{\xi_l} - 1\right). \tag{33}$$

Hence, problem (7) can be reformulated as

$$\begin{aligned} &\underset{\mathbf{x}\in\mathcal{X},\mathbf{P}\in\mathcal{P}}{\text{maximize}} \sum_s U_s(x_s) - \sum_l P_l \\ &\text{subject to} \sum_{s \in S(l)} x_s \leq c_l(\bar{\gamma}_l) \quad \forall l, \\ &\bar{\gamma}_l \geq \eta_l \quad \forall l, \end{aligned} \tag{34}$$

where η_l is the value of the right-hand side of either inequality (32) or (33). The power allocation with lower bound SIR constraint (33) is more aggressive than that with the upper bound one (32), which will be illustrated in the following section. Using a similar technique of log-changed variables, (34) can be transformed into the following equivalent optimization problem:

$$\begin{aligned} &\underset{\hat{\mathbf{x}}\in\mathcal{X},\hat{\mathbf{P}}\in\hat{\mathcal{P}}}{\text{maximize}} \sum_s U_s(e^{\hat{x}_s}) - \sum_l e^{\hat{P}_l} \\ &\text{subject to} \log\left(\sum_{s \in S(l)} e^{\hat{x}_s}\right) \leq \log c_l\left(\hat{\bar{\gamma}}_l\right) \quad \forall l, \\ &-\log \hat{\bar{\gamma}}_l \leq -\log \eta_l \quad \forall l. \end{aligned} \tag{35}$$

This problem is also a convex programming problem. While the objective function and the first constraint are the same as in the convex problem (13), the second constraint

$$-\log \hat{\bar{\gamma}}_l = -\log\left(G_{ll} e^{\hat{P}_l}\right) + \log\left(\sum_{k \neq l} G_{lk} e^{\hat{P}_k} + n_l\right), \tag{36}$$

is clearly a convex function of $\hat{\mathbf{P}}$ due to the sum of linear and log-sum-exp terms. Using the same approach as in Section 4, the partial Lagrangians of (35) are

$$L_{\hat{x}}(\hat{\mathbf{x}}, \boldsymbol{\lambda}) = \sum_s U_s(e^{\hat{x}_s}) - \sum_l \lambda_l \log\left(\sum_{s\in S(l)} e^{\hat{x}_s}\right), \quad (37)$$

$$L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_l \lambda_l \log c_l(\hat{\bar{\gamma}}_l) + \nu_l \log \hat{\bar{\gamma}}_l - e^{\hat{P}_l}. \quad (38)$$

Making use of the projected gradient algorithm to solve the dual problem analogously to these in the previous section, we design the second iterative algorithm as follows.

Algorithm 2. Near-Optimal JCPC with an Outage Constraint using a Complete-Convexification Method
All primal and dual variables are updated iteratively as follows until the termination criterion is satisfied
Congestion control: The source rate updates

$$x_s(t+1) = \left[U_s'^{-1}(\lambda_s(t))\right]_{x_{min}}^{x_{max}}, \quad (39)$$

where $U_s'^{-1}$ is the inverse of the first derivative of utility and

$\lambda_s(t) = \sum_{l\in L(s)} \frac{\lambda_l(t)}{\sum_{f\in S(l)} x_f(t)}$.

Power control: The link power updates

$$P_l(t+1) = \left[\frac{\delta_l(t) + \nu_l(t)}{1 + \sum_{k\neq l} G_{kl} m_k(t)}\right]_{P_{min}}^{P_{max}}, \quad (40)$$

where

$$\delta_k(t) = \lambda_k(t) \frac{1}{\log(1+\bar{\gamma}_k(t))} \frac{\bar{\gamma}_k(t)}{(1+\bar{\gamma}_k(t))}, \quad (41)$$

$$m_k(t) = (\delta_k(t) + \nu_k(t)) \frac{\bar{\gamma}_k(t)}{G_{kk} P_k(t)}. \quad (42)$$

Link Congestion Price Update:

$$\lambda_l(t+1) = \left[\lambda_l(t) - \kappa(t)\left(-\log\left(\sum_{s\in S(l)} x_s(t)\right) + \log c_l(\bar{\gamma}_l(t))\right)\right]^+. \quad (43)$$

Link Outage Price Update:

$$\nu_l(t+1) = [\nu_l(t) - \kappa(t)(\log \bar{\gamma}_l(t) - \log \eta_l)]^+. \quad (44)$$

Proposition 4. *The dual problem of (35) has the strong duality property.*

Proof. We know that (35) is a convex problem and we can choose any strictly feasible points such as

$$\begin{cases} \hat{x}_s = -\infty \quad \forall s \in \mathcal{S}, \\ \{(\hat{P}_l, \gamma_l^{th}, \xi_l) \in \mathbb{R}_+^3,\ \forall l \in \mathcal{L} | \log \hat{\bar{\gamma}}_l > \log \eta_l\}. \end{cases}$$

So the Slater's constraint qualification holds [24, pp. 226-227]. □

From (37), we see that the congestion control mechanism is the same as Algorithm 1. We focus on the power control in the following result.

Proposition 5. *The power update (40) solves the maximization problem* $\max_{\hat{\mathbf{P}}\in\hat{\mathcal{P}}}(38)$ *for a fixed* $(\boldsymbol{\lambda}, \boldsymbol{\nu})$.

Proof. See Appendix C, which is available in the online supplemental material. □

Similarly, with the same step-size condition as that in Algorithm 1, the convergence of Algorithm 2 also can be proved using gradient-based standard techniques.

Theorem 3. *For any initial power* $\mathbf{P}(0) \in \mathcal{P}$, *source rate* $\mathbf{x}(0) \in \mathcal{X}$, *and prices* $(\boldsymbol{\lambda}(0), \boldsymbol{\nu}(0)) \geq 0$, *the sequence of* $\{\mathbf{x}(t), \mathbf{P}(t), \boldsymbol{\lambda}(t), \boldsymbol{\nu}(t)\}$ *generated via Algorithm 2 converges to the global optimal points if the step size satisfies* $\kappa(t) > 0, \sum_{t=0}^{\infty} \kappa(t)^2 < \infty$ *and* $\sum_{t=0}^{\infty} \kappa(t) \to \infty$.

Remarks:

1. The congestion control mechanism is the same as in Algorithm 1.
2. Link power control is much more simplified than that of Algorithm 1, where the control message broadcast by each receiver of link k only contains $m_k(t)$ with locally measurable quantities. Link power update (40) also uses both of link congestion and outage prices.
3. The link outage price update requires only its link's local SIR measurement.
4. We note that Algorithm 2 converges to the global optimal points of (35), which is an approximation of (13). Hence, an optimal solution of (35) is considered as a near-optimal solution of (13) (due to the tightness of the bounds).

6 Optimal Algorithm: Successive Approximation Method

In previous section, a generalized convexity has been established for the original problem (9) which allowed us to propose Algorithm 1 that can achieve a globally optimal solution through messaging passing without high-SIR assumption. Due to the complicated convexification, however the rate allocation of Algorithm 1 with explicit message passing no longer preserves the existing TCP stack like that of [4], which makes it less favorable. To avoid high-SIR assumption yet preserve TCP stack, in this section we propose an algorithm using a novel successive approximation method to iteratively transform the original nonconvex problem of JCPC into approximated convex problem, then the global optimal solution can converge distributively with message passing.

6.1 Approximated Convex Optimization Problem

In order to turn the original nonconvex problem (9) to an approximated convex problem, we begin to form a new lower bound approximation to the first constraint of problem (9)

$$\sum_{s\in S(l)} x_s \leq \hat{c}_l(\mathbf{P}) \leq c_l(\bar{\gamma}_l(\mathbf{P})). \quad (45)$$

We note that $c_l(\bar{\gamma}_l(\mathbf{P}))$ can be rewritten in the form

$$c_l(\bar{\gamma}_l(\mathbf{P})) = \log\left(\sum_{k\in\mathcal{L}} G_{lk}P_k + n_l\right) - \log\left(\sum_{k\neq l} G_{lk}P_k + n_l\right).$$

The arithmetic-geometric mean inequality states that $\sum_i \theta_i u_i \geq \prod_i u_i^{\theta_i}$ with $u_i \geq 0, \theta_i > 0 \forall i$ and $\sum_i \theta_i = 1$. By letting $v_i = \theta_i u_i$, the inequality becomes $\sum_i v_i \geq \prod_i (v_i/\theta_i)^{\theta_i}$ and the equality happens when $\theta_i = v_i/\sum_i v_i$. Supposed each link l have a vector $\boldsymbol{\theta}^l = [\theta_1^l, \theta_2^l, \ldots, \theta_{L+1}^l]$, employing the arithmetic-geometric mean inequality to each link l we have

$$\sum_{k\in\mathcal{L}} G_{lk}P_k + n_l \geq \prod_{k=1}^{L}\left(\frac{G_{lk}P_k}{\theta_k^l}\right)^{\theta_k^l}\left(\frac{n_l}{\theta_{L+1}^l}\right)^{\theta_{L+1}^l}, \tag{46}$$

and the equality happens when

$$\begin{aligned}\theta_k^l &= \frac{G_{lk}P_k}{\sum_{k\in\mathcal{L}} G_{lk}P_k + n_l}, \quad k = 1, \ldots, L,\\ \theta_{L+1}^l &= \frac{n_l}{\sum_{k\in\mathcal{L}} G_{lk}P_k + n_l}.\end{aligned} \tag{47}$$

Because $\log(.)$ is an increasing function of positive variables, by taking logarithm on both sides of (46) we have

$$\begin{aligned}&\log\left(\sum_{k\in\mathcal{L}} G_{lk}P_k + n_l\right)\\ &\geq \sum_{k=1}^{L}\theta_k^l \log\left(\frac{G_{lk}P_k}{\theta_k^l}\right) + \theta_{L+1}^l \log\left(\frac{n_l}{\theta_{L+1}^l}\right) \triangleq f(\mathbf{P}, \boldsymbol{\theta}^l).\end{aligned} \tag{48}$$

Letting $\hat{c}_l(\mathbf{P}, \boldsymbol{\theta}^l) = f(\mathbf{P}, \boldsymbol{\theta}^l) - \log(\sum_{k\neq l} G_{lk}P_k + n_l)$, we have

$$\hat{c}_l(\mathbf{P}, \boldsymbol{\theta}^l) \leq c_l(\bar{\gamma}_l(\mathbf{P})), \tag{49}$$

and the equality happens when (47) holds.

Letting $\hat{P}_l = \log P_l$, it is easy to see that

$$f(\hat{\mathbf{P}}, \boldsymbol{\theta}^l) = \sum_{k=1}^{L}\theta_k^l \hat{P}_k + \sum_{k=1}^{L}\theta_k^l \log\left(\frac{G_{lk}}{\theta_k^l}\right) + \theta_{L+1}^l \log\left(\frac{n_l}{\theta_{L+1}^l}\right),$$

is a linear function of $\hat{\mathbf{P}}$, so

$$\hat{c}_l(\hat{\mathbf{P}}, \boldsymbol{\theta}^l) = f(\hat{\mathbf{P}}, \boldsymbol{\theta}^l) - \log\left(\sum_{k\neq l} G_{lk}e^{\hat{P}_k} + n_l\right), \tag{50}$$

is a concave function of $\hat{\mathbf{P}}$ (recall that log-sum-exponent is convex). Hence, we have the approximated convex optimization problem of the original one (9) with variables $\mathbf{x}$ and $\hat{\mathbf{P}}$(θ^l is fixed) as following:

$$\begin{aligned}&\underset{\mathbf{x}\in\mathcal{X}, \hat{\mathbf{P}}\in\hat{\mathcal{P}}}{\text{maximize}} \sum_s U_s(x_s) - \sum_l e^{\hat{P}_l}\\ &\text{subject to} \sum_{s\in S(l)} x_s \leq \hat{c}_l(\hat{\mathbf{P}}, \boldsymbol{\theta}^l) \quad \forall l,\\ &\sum_{k\neq l}\log\left(1 + e^{\hat{P}_k - \hat{P}_l}\frac{\gamma_l^{th}G_{lk}}{G_{ll}}\right) \leq \log\Omega_l\left(e^{\hat{P}_l}\right) \quad \forall l.\end{aligned} \tag{51}$$

6.2 Optimal Solution of Approximated Convex Problem

Using the same approach as in Section 4, the partial Lagrangians of (51) are

$$L_x(\mathbf{x}, \boldsymbol{\lambda}) = \sum_s U(x_s) - \sum_s \sum_{l\in L(s)} \lambda_l x_s, \tag{52}$$

$$\begin{aligned}L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_l \Bigg(&\lambda_l \hat{c}_l(\hat{\mathbf{P}}, \boldsymbol{\theta}^l) + \nu_l \log\Omega_l(e^{\hat{P}_l}) - e^{\hat{P}_l}\\ &- \nu_l \sum_{k\neq l}\log\left(1 + e^{\hat{P}_k - \hat{P}_l}\frac{\gamma_l^{th}G_{lk}}{G_{ll}}\right)\Bigg).\end{aligned} \tag{53}$$

The maximization problem $\max_{\mathbf{x}\in\mathcal{X}} L_x(\mathbf{x}, \boldsymbol{\lambda})$ of (52) is the conventional rate control problem which is implicitly solved by the congestion control mechanism for different U_s [4], hence preserving the existing TCP stack. The second maximization problem $\max_{\hat{\mathbf{P}}\in\hat{\mathcal{P}}} L_{\hat{P}}(\hat{\mathbf{P}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ of (53) is the power control problem.

Similarly to previous sections, with utility function's assumption and strong duality because of Slater condition qualification, we will use the projected gradient algorithm to solve the dual problem. The following procedure solves the approximated convex problem (51) of the original nonconvex problem (9).

Procedure 1. Finding Optimal Solution of Approximated JCPC with Outage Constraint

All primal and dual variables are updated iteratively as follows until the termination criterion is satisfied

Congestion control: The source rate updates

$$x_s(t+1) = \left[U_s'^{-1}\left(\sum_{l\in L(s)} \lambda_l(t)\right)\right]_{x_{min}}^{x_{max}}, \tag{54}$$

where $U_s'^{-1}$ is the inverse of the first derivative of utility.

Power control: The link power updates

$$\begin{aligned}&P_l(t+1)\\ &= \left[\frac{\Delta_l(t)}{1 + \sum_{k\neq l}\left(G_{kl}\hat{m}_k(t) + \nu_k(t)\frac{G_{kl}\tilde{m}_k(t)}{1+G_{kl}\tilde{m}_k(t)P_l(t)}\right)}\right]_{P_{min}}^{P_{max}},\end{aligned} \tag{55}$$

where

$$\Delta_l(t) = \lambda_l(t)\theta_l^l - \nu_l(t)\tilde{m}_l(t)\frac{n_l}{\log(1-\xi_l)}, \tag{56}$$

$$\hat{m}_k(t) = \frac{\lambda_k(t)\bar{\gamma}_k(t)}{G_{kk}P_k(t)}, \tag{57}$$

$$\tilde{m}_k(t) = \frac{\gamma_k^{th}}{G_{kk}P_k(t)}. \tag{58}$$

Link Congestion Price Update:

$$\lambda_l(t+1) = \left[\lambda_l(t) - \kappa(t)\left(\hat{c}_l(\hat{\mathbf{P}}(t), \boldsymbol{\theta}^l) - \sum_{s\in S(l)} x_s(t)\right)\right]^+. \tag{59}$$

Link Outage Price Update:

$$\begin{aligned}\nu_l(t+1) = \Bigg[&\nu_l(t) - \kappa(t)\Bigg(\log\Omega_l(P_l(t))\\ &- \sum_{k\neq l}\log\left(1 + \gamma_l^{th}\frac{G_{lk}P_k(t)}{G_{ll}P_l(t)}\right)\Bigg)\Bigg]^+.\end{aligned} \tag{60}$$

Theorem 4. *For any initial power $\mathbf{P}(0) \in \mathcal{P}$, source rate $\mathbf{x}(0) \in \mathcal{X}$, and prices $(\boldsymbol{\lambda}(0), \boldsymbol{\nu}(0)) \geq 0$, the sequence of $\{\mathbf{x}(t), \mathbf{P}(t), \boldsymbol{\lambda}(t), \boldsymbol{\nu}(t)\}$ generated via Procedure 1 will converge to the global optimal points of problem (51) if the step size satisfies $\kappa(t) > 0, \sum_{t=0}^{\infty}\kappa(t)^2 < \infty$ and $\sum_{t=0}^{\infty}\kappa(t) \to \infty$.*

Proof. Analogously, by setting $\frac{\partial L_x(\mathbf{x},\boldsymbol{\lambda})}{\partial x_s} = 0$ and $\frac{\partial L_{\hat{P}}(\hat{\mathbf{P}},\boldsymbol{\lambda},\boldsymbol{\nu})}{\partial \hat{P}_l} = 0$, we have the congestion control update (54) solves the maximization problem $\max_{\mathbf{x}\in\mathcal{X}}(52)$ and the power update (55) solves the maximization problem $\max_{\hat{\mathbf{P}}\in\hat{\mathcal{P}}}(53)$ respectively, for a fixed $(\boldsymbol{\lambda}, \boldsymbol{\nu})$. The updates of (59) and (60) shows that we apply the projected gradient-descent method to solve the dual problem (20), which guarantees the convergence of dual variable with the appropriately chosen step size $\kappa(t)$ [26]. □

Remarks:

1. The congestion control update (54) is well known in the literature [3]. It has been shown in [3] that we can reuse existing distributed TCP algorithms and different TCP algorithms solve for different utility functions. For example, $U_s(x_s) = \alpha_s d_s \log x_s$ is shown to be associated with TCP Vegas, where α_s is the Vegas parameter and d_s is the propagation delay. TCP Reno-1 and Reno-2 are associated with the utility functions $U_s(x_s) = \frac{\sqrt{3/2}}{D_s}\tan^{-1}(\sqrt{2/3}x_s D_s)$ and $U_s(x_s) = \frac{1}{D_s}\log\frac{x_s D_s}{2x_s D_s + 3}$, respectively, where D_s is propagation delay plus congestion-induce queuing delay.
2. Link power control is analogous to that of Algorithm 1. Each receiver of link $k$broadcasts its control message containing three real-value fields, $\hat{m}_k(t)$, $\tilde{m}_k(t)$, and $\nu_k(t)$. Each transmitter of link l then receives these values, estimates $G_{kl}$by using the training sequences and updates its power according to (55) using both congestion price and outage price.
3. The link outage price update is similar to Algorithm 1.
4. The congestion price update (59) only needs its link local information including the ingress rate and received signal measurement.

6.3 Successive Convex Approximations: Algorithm and Optimality

We continue presenting an algorithm that can achieve the globally optimal solutions of the original nonconvex problem (9) by solving successively the approximated problem (51).

Algorithm 3: Optimal JCPC with Outage Constraint using a Successive Convex Approximation Method

1. Initialize $(\mathbf{x}, \mathbf{P}) = 0$, $\tau = 1$.
2. Form the τ-th approximated convex problem (51) of the original problem (9) by updating $\boldsymbol{\theta}^{l^{(\tau)}}$, $\forall l$ with (47).
3. Solve the τ-th approximated convex problem (51) for optimal solution $(\mathbf{x}^{*(\tau)}, \mathbf{P}^{*(\tau)})$ using Procedure 1.
4. Increment τ and go to step 2 until convergence.

Theorem 5. *The series of approximations of Algorithm 3 converge to the stationary points satisfying the Karush-Kuhn-Tucker (KKT) conditions of the original problem (9).*

Proof. Letting $h(\mathbf{x}, \mathbf{P}) = \frac{\sum_{s:l\in L(s)} x_s}{c_l(\bar{\gamma}_l(\mathbf{P}))}$ and $\hat{h}(\mathbf{x}, \mathbf{P}) = \frac{\sum_{s:l\in L(s)} x_s}{\hat{c}_l(\mathbf{P})}$, we need to prove that this series of approximations satisfies the following properties according to [27]

1. $h(\mathbf{x}, \mathbf{P}) \le \hat{h}(\mathbf{x}, \mathbf{P})$.
2. $h(\mathbf{x}^{\mathbf{o}}, \mathbf{P}^{\mathbf{o}}) = \hat{h}(\mathbf{x}^{\mathbf{o}}, \mathbf{P}^{\mathbf{o}})$.
3. $\nabla h(\mathbf{x}^{\mathbf{o}}, \mathbf{P}^{\mathbf{o}}) = \nabla \hat{h}(\mathbf{x}^{\mathbf{o}}, \mathbf{P}^{\mathbf{o}})$, where $(\mathbf{x}^{\mathbf{o}}, \mathbf{P}^{\mathbf{o}})$ is the optimal solution of the previous iteration.

Conditions 1 and 2 are clearly satisfied with (47) and (49). It is straightforward to verify condition 3 by taking derivative. Then, the globally optimal convergence of Algorithm 3 can be proved similarly as in [9]. □

7 Convergence Analysis with Random Errors

In order to implement Algorithms 1, 2, and 3, information feedback (message passing) is crucial for computing gradients at each link. However, in practical systems the feedback signal is transmitted over a wireless channel and is error-prone due to the channel variations in link quality. The main objective in this section is to study the convergence behavior of Algorithms 1, 2, and 3 with regard to the random-error message passing.

Here we choose Algorithm 1 as the candidate for the convergence analysis which applies similarly to Algorithm 2 and 3. Let the vectors $\mathbf{g}(t) = (g_1(t), \ldots, g_L(t))$ and $\mathbf{h}(t) = (h_1(t), \ldots, h_L(t))$ be the gradient vectors of dual function $D(\boldsymbol{\lambda}(t), \boldsymbol{\nu}(t))$ with respect to $\boldsymbol{\lambda}(t)$ and $\boldsymbol{\nu}(t)$. From Algorithm 1, their lth elements are as follows:

$$g_l(t) = \log c_l(\bar{\gamma}_l(t)) - \log\left(\sum_{s\in S(l)} x_s(t)\right), \tag{61}$$

$$h_l(t) = \log \Omega_l(P_l(t)) - \sum_{k\neq l} \log\left(1 + \gamma_l^{th} \frac{G_{lk} P_k(t)}{G_{ll} P_l(t)}\right). \tag{62}$$

In the presence of random-error message passing, the gradients $\mathbf{g}(t)$ and $\mathbf{h}(t)$ are stochastic. Let $\tilde{\mathbf{g}}(t)$ and $\tilde{\mathbf{h}}(t)$ be the corresponding estimators, then the stochastic versions of (26) and (27) can be written as

$$\lambda_l(t+1) = [\lambda_l(t) - \kappa(t)\tilde{g}_l(t)]^+, \tag{63}$$

$$\nu_l(t+1) = [\nu_l(t) - \kappa(t)\tilde{h}_l(t)]^+. \tag{64}$$

Denote the random vector $\mathbf{R}(t) = (\tilde{\mathbf{g}}(t), \tilde{\mathbf{h}}(t), \boldsymbol{\lambda}(t), \boldsymbol{\nu}(t))$ for $t \ge 0$. Let $\mathcal{F}_t$ be the σ-field generated by $\mathbf{R}(0), \mathbf{R}(1), \ldots, \mathbf{R}(t)$ and denoted by

$$\mathcal{F}_t = \sigma(\mathbf{R}(0), \mathbf{R}(1), \ldots, \mathbf{R}(t)). \tag{65}$$

Then, each lth element of $\tilde{\mathbf{g}}(t)$ and $\tilde{\mathbf{h}}(t)$ can be decomposed into three parts

$$\tilde{g}_l(t) = g_l(t) + (E[\tilde{g}(t)|\mathcal{F}_t] - g_l(t)) + (\tilde{g}(t) - E[\tilde{g}(t)|\mathcal{F}_t]), \tag{66}$$

$$\tilde{h}_l(t) = h_l(t) + (E[\tilde{h}(t)|\mathcal{F}_t] - h_l(t)) + (\tilde{h}(t) - E[\tilde{h}(t)|\mathcal{F}_t]). \tag{67}$$

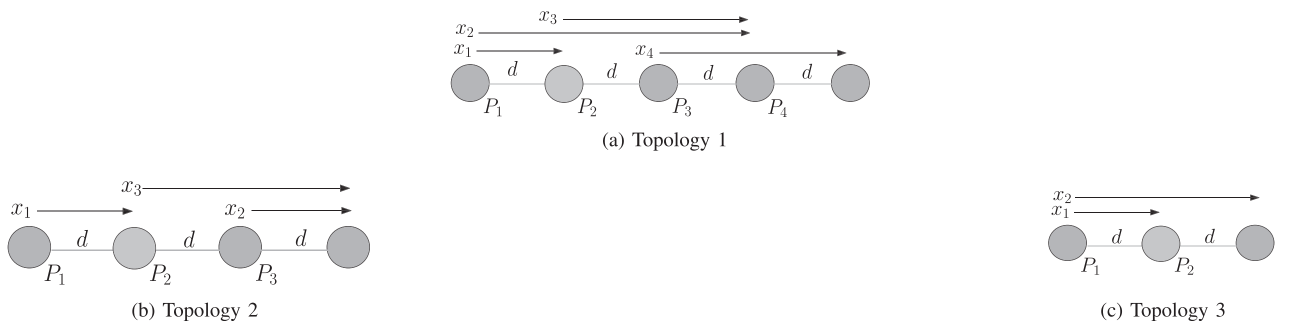


Fig. 1. Three different linear topologies for multihop wireless networks where nodes are placed equidistantly. (a) Topology 1 with four links and four flows. (b) Topology 2 with three links and three flows. (c) Topology 3 with two links and two flows.

The first part is the exact gradients $g_l(t), h_l(t)$. The second part is the biased estimation errors of $g_l(t)$ and $h_l(t)$, denoted by

$$B_l^{\lambda}(t) = E[\tilde{g}_l(t)|\mathcal{F}_t] - g_l(t),\ \mathbf{B}^{\lambda}(t) = \big(B_1^{\lambda}(t), \ldots, B_L^{\lambda}(t)\big), \tag{68}$$

$$B_l^{\nu}(t) = E[\tilde{h}_l(t)|\mathcal{F}_t] - h_l(t),\ \mathbf{B}^{\nu}(t) = \big(B_1^{\nu}(t), \ldots, B_L^{\nu}(t)\big). \tag{69}$$

And the third part is a zero-mean martingale difference noises, denoted by

$$N_l^{\lambda}(t) = \tilde{g}_l(t) - E[\tilde{g}(t)|\mathcal{F}_t],\ \mathbf{N}^{\lambda}(t) = \big(N_1^{\lambda}(t), \ldots, N_L^{\lambda}(t)\big), \tag{70}$$

$$N_l^{\nu}(t) = \tilde{h}_l(t) - E[\tilde{h}(t)|\mathcal{F}_t],\ \mathbf{N}^{\nu}(t) = \big(N_1^{\nu}(t), \ldots, N_L^{\nu}(t)\big). \tag{71}$$

Therefore,

$$\tilde{g}_l(t) = g_l(t) + B_l^{\lambda}(t) + N_l^{\lambda}(t), \tag{72}$$

$$\tilde{h}_l(t) = h_l(t) + B_l^{\nu}(t) + N_l^{\nu}(t). \tag{73}$$

Under suitable conditions of step sizes and biased errors, we address the convergence of Algorithm 1 (similarly to Algorithm 2) in the following theorem.

Theorem 6. *Assume that*

$$\lambda_{max} = \max_{l,t} \lambda_l(t) < \infty,\ \nu_{max} = \max_{l,t} \nu_l(t) < \infty, \tag{74}$$

$$\sum_t \kappa(t)|B_l^{\lambda}(t)| < \infty,\ \sum_t \kappa(t)\big|B_l^{\nu}(t)\big| < \infty, \tag{75}$$

and the step size is chosen such that

$$\kappa(t) > 0,\ \sum_{t=0}^{\infty} \kappa(t)^2 < \infty,\ \sum_{t=0}^{\infty} \kappa(t) \to \infty, \tag{76}$$

then Algorithm 1 with stochastic dual variable updates (63) and (64) converges to the optimal solutions of (13) with probability 1.

Proof. See Appendix D, which is available in the online supplemental material. □

8 Simulation Results

8.1 Simulation Setting

Figs. 1 and 2 show network topologies that we use for the simulations in this section. Following the conventional work [4], unless otherwise stated, we mainly consider the linear network topology 1 as in Fig. 1a with four flows and five nodes placed at d meters equidistantly. The baseband bandwidth W is set to 32 kHz, and we use $K = -1.5/\log(5\mathrm{BER})$ with $\mathrm{BER} = 10^{-3}$ for MQAM modulion [20]. The slow-fading channel gain is assumed to be $h(d) = h_o(\frac{d}{100})^{-4}$, where h_o is a reference channel gain at a distance of 100 m. The maximum power, noise and h_o are selected so that the average receive SNR at 100 m is 30 dB. We choose $P_l^{min} = 1$ mW and $P_l^{max} = 100$ mW, while $x_s^{min} = 0$, and x_s^{max} is adjusted dynamically with respect to link capacities. The step sizes of both algorithms are chosen to be $0.01/t$. The α-fair utility function (1) is set to all users, which means that the congestion control (21) is $x_s(t+1) = [\lambda_s(t)^{-\frac{1}{\alpha}}]_{x_{min}}^{x_{max}}$. The outage probability thresholds ξ_l of Topology 1, 2, 3, and 4 are (0.2 0.3 0.3 0.2), (0.2 0.3 0.2), (0.2 0.2), and (0.3 0.3 0.3 0.2 0.2), respectively. The SIR thresholds γ_l^{th} of Topology 1, 2, 3, and 4 are set to (0.6 0.2 0.2 0.6), (0.6 0.2 0.6), (0.6 0.6), and (0.2 0.2 0.2 0.6 0.6) dB, respectively.

8.2 Optimal Gap

Henceforth, we denote the objective function $\mathcal{U}(\mathbf{x}, \mathbf{P}) = \sum_s U_s(x_s) - \sum_l P_l$ for the ease of numerical presentation, and we choose the lower bound SIR constraint (33) for Algorithm 2. We evaluate and compare the objective value $\mathcal{U}(\mathbf{x}, \mathbf{P})$ of three proposed algorithms and the conventional one [4]. All of results are averaged out of 100 random scenarios.

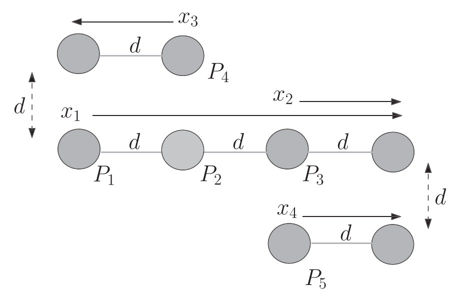
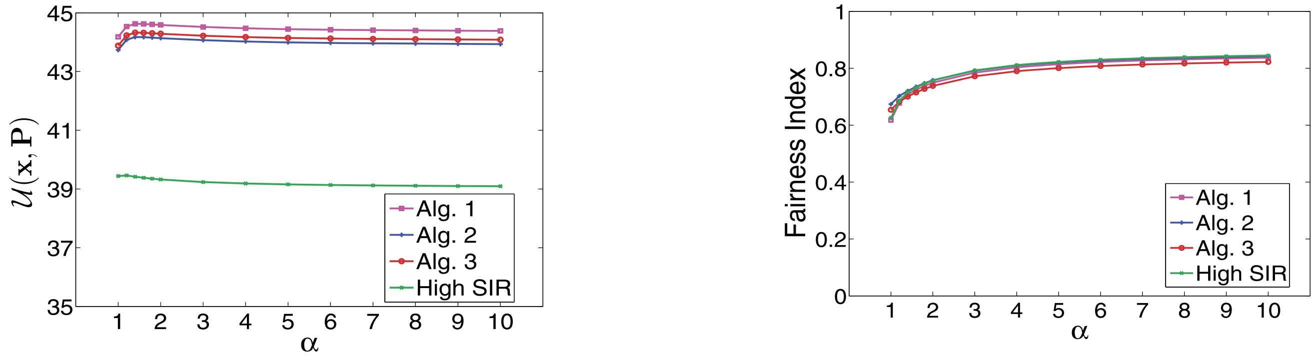
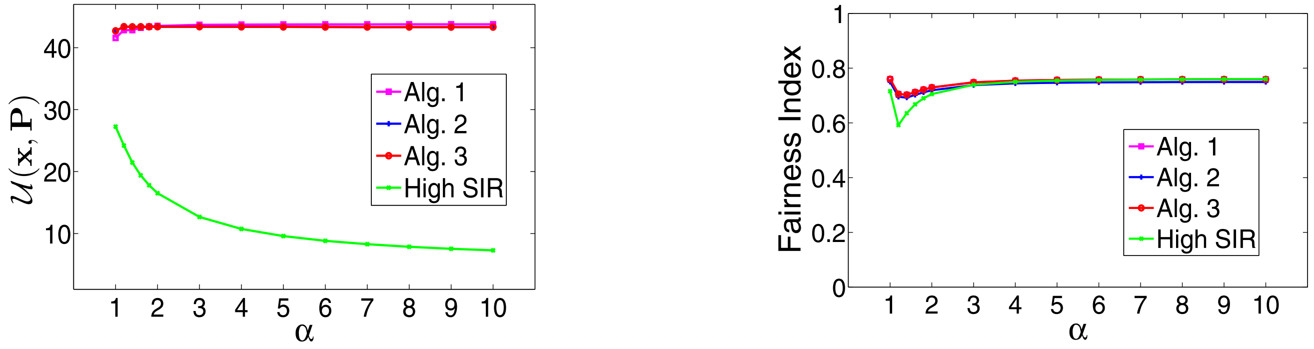


Fig. 2. Topology 4 is a nonlinear network topology with four flows (from x_1 to x_4) and five links (link transmit powers are from P_1 to P_5). Flow 1 passes through link 1, 2, and 3. Flow 2, flow 3, and flow 4 are on link 3, link 4, and link 5, respectively.



(a) (b)

Fig. 3. Impact of a variable utility parameter α on network efficiency and fairness of three proposed algorithms and Chiang's algorithm (high-SIR approximation) in a linear topology. (a) Objective value $\mathcal{U}(\mathbf{x},\mathbf{P})$. (b) Jain's fairness index.



(a) (b)

Fig. 4. Impact of a variable utility parameter α on network efficiency and fairness of three proposed algorithms and Chiang's algorithm (high-SIR approximation) in a nonlinear topology. (a) Objective value $\mathcal{U}(\mathbf{x},\mathbf{P})$. (b) Jain's fairness index.

First, we investigate the impact of the utility parameter α on the network performance. This parameter can act as a *knob* to control the tradeoff between network efficiency and fairness in a general NUM problem [5], [23]. We fix $d = 80$ m and vary α from 1 to 10 to compare the network efficiency (objective value $\mathcal{U}(\mathbf{x},\mathbf{P})$) and fairness, where we use the Jain's fairness index [28] as the standard fairness measurement: $(\sum_s x_s)^2/(S\sum_s x_s^2)$. Fig. 3 shows the network efficiency and fairness comparisons performed in a linear network topology as in Fig. 1a. We can see in Fig. 3 that when α increases, the objective value achieves the maximum value at $\alpha = 1.5$ and then becomes less efficient. The fairness of the system increases when α increases. We observe from Fig. 3a that the performances of Algorithms 1, 2, and 3 are almost identical. Even though the convergent objective values of both Algorithms 2 and 3 are a bit lower than that of Algorithm 1 because of the outage bounds usage and the approximated nature, the relative errors are small, only 1.57 and 1.12 percent, respectively. Moreover, three proposed algorithms clearly outperform the conventional scheme [4] which spent higher power transmission due to the high SIR approximation. From Fig. 3, all of the compared schemes achieve nearly the same fairness performance. This can be explained that they are in the same manner of proportional allocation of the congestion control. Fig. 4 shows the network efficiency and fairness comparisons performed in a nonlinear network topology as in Fig. 2. In Fig. 4a, we see that when α increases, the network efficiency of three proposed algorithms also increases a little, while that of the conventional scheme [4] decreases sharply, which shows the inefficiency of high-SIR approximation method. Fig. 4b shows that when α starts to increase from the value 2, the fairness performance of all compared algorithms also increases and is the same. When α varies from 1 to 2, all schemes decreases the performance to the minimum values at $\alpha = 1.2$ and then starts to increase afterward. However, within this range of α, the fairness performance of three proposed algorithms is higher than that of the conventional scheme due to the error effects in fair source rate allocation of high-SIR approximation method.

Next, we vary the distance parameter d to evaluate the network performance. We fix $\alpha = 1$ (i.e., proportional fairness) for all d. As can be seen from Table 2, the objective values of three proposed schemes are almost identical. We also observe that when d increases, the objective tends to decrease with small variations because the extra power consumed by nodes to communicate is also very small due

TABLE 2
The Average Objective $\mathcal{U}(\mathbf{x},\mathbf{P})$ Comparisons between Proposals when Varying the Distance Parameter d

d(m)	40	60	80	100	120	140
Algorithm 1	44.548	44.502	44.539	44.258	43.861	43.747
Algorithm 2	43.865	43.854	43.812	43.336	43.139	43.026
Algorithm 3	44.089	44.069	44.022	43.943	43.845	43.732

TABLE 3
Convergence Speed Comparison between Algorithms 1 and 2 with Different Topologies

ϵ	Algorithm 1			Algorithm 2		
	Topo 1	Topo 2	Topo 3	Topo 1	Topo 2	Topo 3
10^{-3}	93.97	65.75	25.87	33.34	30.04	27.67
10^{-4}	141.26	87.56	36.77	58.29	46.63	33.55
10^{-5}	139.01	89.08	37.41	105.23	73.66	39.66

TABLE 4
Convergence Speed Comparison between Algorithm 3 and Log Successive Approximation Algorithm of [9]

ϵ	Algorithm 3		log successive convex	
	Inner Convg.	Outer Convg.	Inner Convg.	Outer Convg.
10^{-3}	33.97	7.26	94.01	8.44
10^{-4}	41.97	9.78	143.46	10.52
10^{-5}	77.44	10.94	186.67	12.38

to the mutual interference in this interference-limited environment.

8.3 Algorithm Convergence without Random Errors

The criterion used to evaluate the convergence speed is

$$\max_{l\in\mathcal{L}} \frac{|P_l(t) - P_l^{(t-1)}|}{P_l^{(t-1)}} < \epsilon,$$

where ϵ is an arbitrary small number. We fix $\alpha = 1$ and $d = 80$ m for these scenarios.

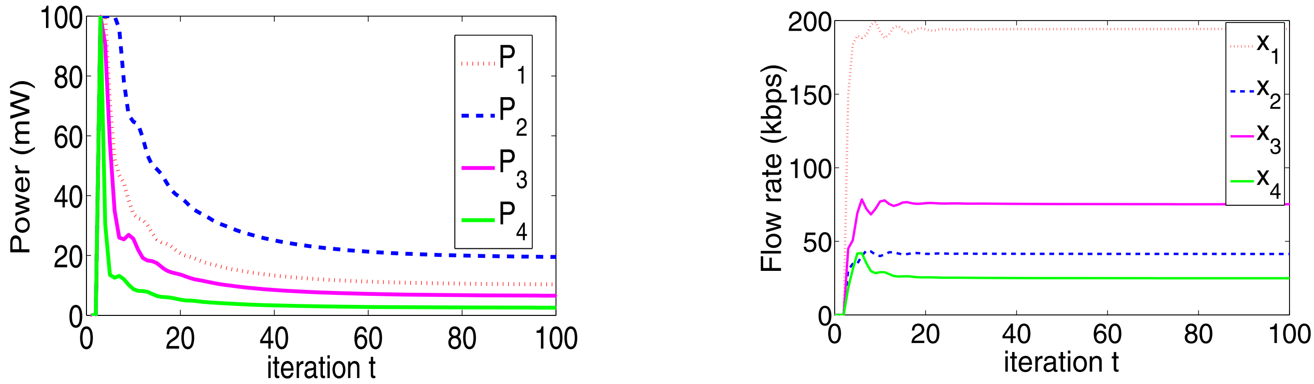
We first compare the convergence speed between Algorithms 1 and 2 because they employ the same complete-convexification method. Table 3 shows the average number of iterations over 100 realizations with various values of ϵ under three different topologies as in Figs. 1a, 1b, and 1c. We see that the near-optimal scheme converges faster than does the optimal scheme clearly in the topology with many nodes and flows (i.e., Topology 1 and 2). With a simple topology like Topology 3, we see that the convergence rate of both Algorithms 1 and 2 are nearly the same. This can be explained that due to the complex updates of Algorithm 1 requiring many parameters from large-size control message, the "direction" to the optimal points of Algorithm 1 has some minor error effects, which entail longer convergence time. This is a significant point, as Algorithm 2, which can achieve a close-to-optimal solution with smaller control message size and faster convergence, would be efficiently practical.

We continue comparing the convergence speed of Algorithm 3 and [9, Alg. B] where both utilized the successive approximation method. Two criteria are used to evaluate the convergence-speed performance. The first one is the convergence condition of solving step 3 (i.e., inner convergence) and the second one is convergence at step 4 (i.e., outer convergence) of Algorithm 3. Both are represented by $\max_{l\in\mathcal{L}}|P_l(t) - P_l(t-1)| < \epsilon$ and $\max_{l\in\mathcal{L}}|P_l^{*(\tau)} - P_l^{*(\tau-1)}| < \epsilon$, respectively, where ϵ is a small number. Table 4 shows the average number of iterations over 100 realizations with various values of ϵ. We see that our scheme converge faster than [9, Alg. B] (i.e., log successive approximation), especially with inner convergence.

Figs. 5, 6, and 7 show the comparison between three proposed algorithms regarding the convergence of source rates and link powers with $\epsilon = 10^{-5}$. It can be observed that the source-rate allocations of three schemes are the same, while the power allocation of Algorithm 2 is somewhat more aggressive than that of Algorithms 1 and 3 due to the constraint approximation. The outage probabilities of three schemes also converge to the desired values similarly as in Fig. 8a. Fig. 8 shows a convergence realization of objective values of three proposed schemes and the conventional scheme [4]. Again we see that three proposed algorithms have nearly the same performance and outperform the high-SIR approximation algorithm [4].

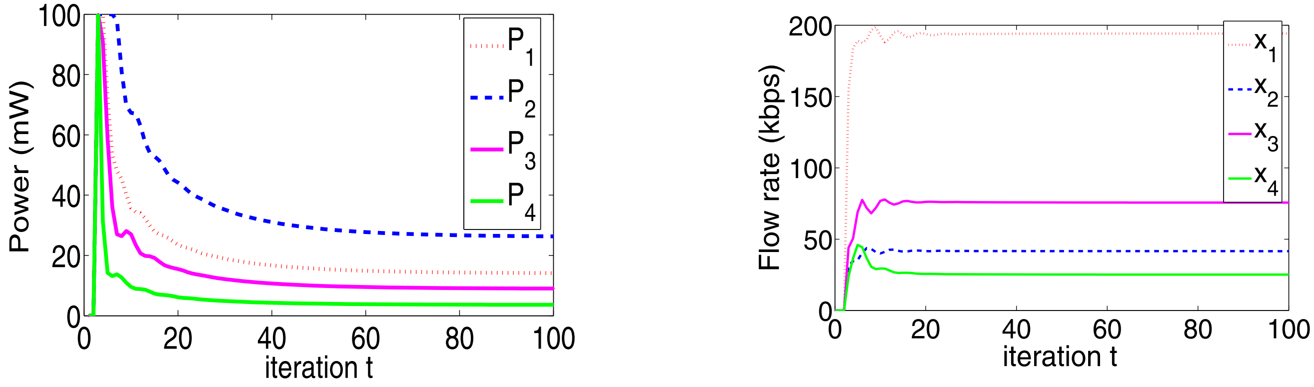
8.4 Algorithm Convergence with Random Errors

We continue investigating the algorithm convergence with random-error message passing. We set $B_l^{\lambda}(t) = B_l^{\nu}(t) = 1/t$, together with the chosen diminishing step size $\kappa(t) = 0.01/t$, to satisfy conditions (75) and (76). The random noise is approximated by Gaussian random variables, where $N_l^{\lambda}(t) \equiv N_l^{\nu}(t) \sim \mathcal{N}(0, 100)$ for all link l. Fig. 9a shows the flow rate convergence of Algorithm 1 using diminishing step size (the other convergence cases of Algorithms 1 and 2 can be reasoned similarly but were not included here due to the limited space). Even though the feedback is affected by noise, the iterations are robust to the random errors. We almost see no fluctuation effect on the algorithm convergence, which supports Theorem 6.



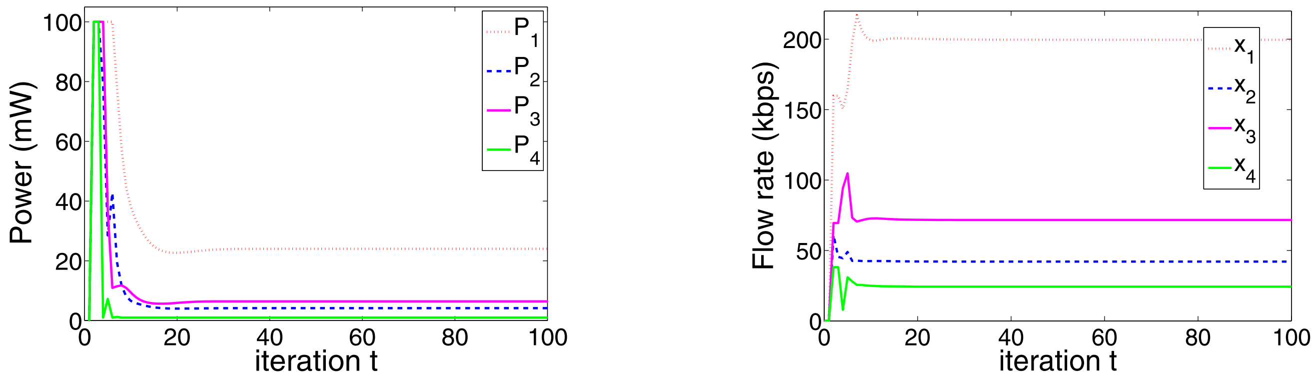
(a) (b)

Fig. 5. Convergence of the Algorithm 1. (a) Link powers. (b) Flow rates.



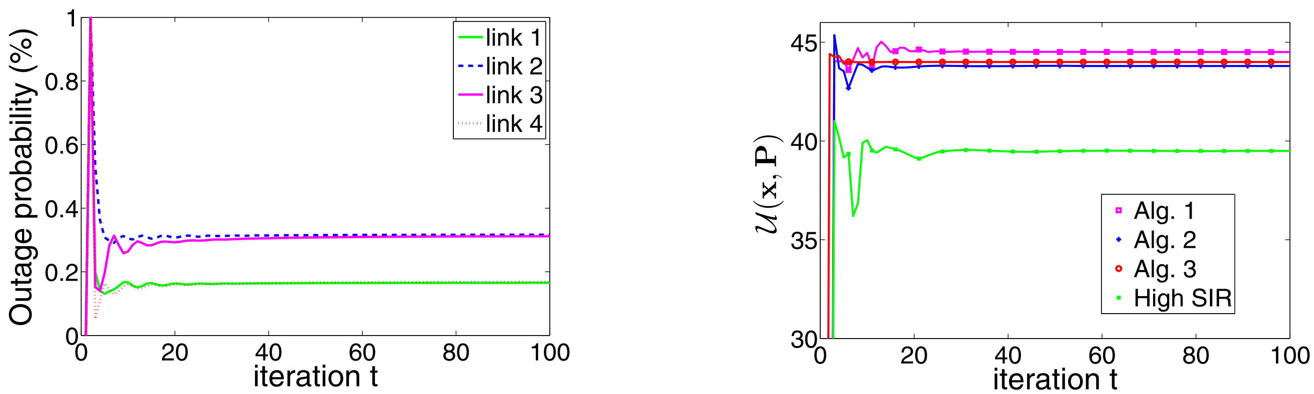
(a) (b)

Fig. 6. Convergence of Algorithm 2 with lower bound SIR constraint. (a) Link powers. (b) Flow rates.



(a) (b)

Fig. 7. Convergence of Algorithm 3. (a) Link powers. (b) Flow rates.



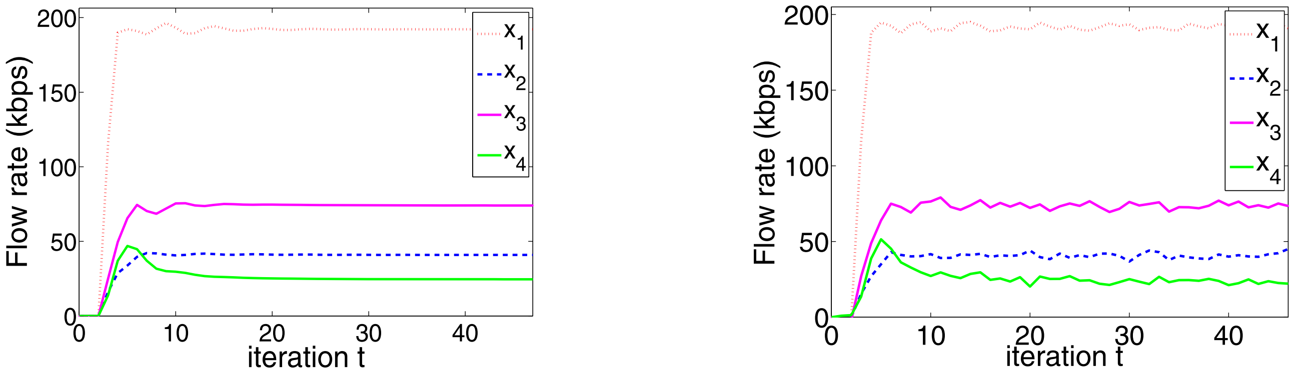
(a) (b)

Fig. 8. (a) Outage probabilities' convergence of three proposed algorithms. (b) Convergence realizations of the objective values of three proposed algorithm and high-SIR approximation algorithm [4].

On the other hand, Fig. 9b shows the convergence of Algorithm 1 with the same parameter settings except a constant step size $\kappa(t) = 0.001$. We observed that the iterations using constant step size (which not satisfy (76)) are sensitive to random errors. The iteration fluctuation remains significant over time, which only guarantees the convergence to a neighborhood of the optimal points.

9 Conclusions

We reconsidered joint rate control and power allocation in wireless multihop networks with an additional constraint on outage probability. We chose the explicit approach to deal with outage constraint because the implicit constraint used in previous works cannot characterize network QoS and may be suboptimal. We designed three algorithms for this cross-layer issue. The first algorithm is the optimal scheme which suffers from high overhead since its control message contains a large amount of information. The second design with a small-size control message is a near-optimal scheme based on a tight bound approximation on outage probability. However, because of the complicated complete-convexification method, both schemes cannot preserve the TCP stack for its congestion control mechanism. We used successive approximation method to propose the third algorithm that can take the TCP stack preservation



(a) (b)

Fig. 9. Flow rate convergence of Algorithm 1 with random errors. (a) Diminishing step size. (b) Constant step size.

into account. Numerical experiments showed that the network performances of three schemes were nearly identical and outperform the conventional works.

Acknowledgments

This research was supported by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0020518). Dr. C.S. Hong is the corresponding author.

References

[1] S.H. Low and D.E. Lapsley, "Optimization Flow Control. I. Basic Algorithm and Convergence," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 861-874, Dec. 1999.

[2] F. Kelly, A. Maulloo, and D. Tan, "Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability," *The J. Operational Research Soc.*, vol. 49, no. 3, pp. 237-252, 1998.

[3] S. Low, "A Duality Model of TCP and Queue Management Algorithms," *IEEE/ACM Trans. Networking*, vol. 11, no. 4, pp. 525-536, Aug. 2003.

[4] M. Chiang, "Balancing Transport and Physical Layers in Wireless Multihop Networks: Jointly Optimal Congestion Control and Power Control," *IEEE J. Selected Areas Comm.*, vol. 23, no. 1, pp. 104-116, Jan. 2005.

[5] J.-W. Lee, M. Chiang, and R.A. Calderbank, "Jointly Optimal Congestion and Contention Control Based on Network Utility Maximization," *IEEE Comm. Letters*, vol. 10, no. 3, pp. 216-218, Mar. 2006.

[6] X. Lin, N.B. Shroff, and R. Srikant, "A Tutorial on Cross-Layer Optimization in Wireless Networks," *IEEE J. Selected Areas Comm.*, vol. 24, no. 8, pp. 1452-1463, Aug. 2006.

[7] C. Long, B. Li, Q. Zhang, B. Zhao, B. Yang, and X. Guan, "The End-to-End Rate Control in Multiple-Hop Wireless Networks: Cross-Layer Formulation and Optimal Allocation," *IEEE J. Selected Areas Comm.*, vol. 26, no. 4, pp. 719-731, May 2008.

[8] D.P. Palomar and M. Chiang, "Alternative Distributed Algorithms for Network Utility Maximization: Framework and Applications," *IEEE Trans. Automatic Control*, vol. 52, no. 12, pp. 2254-2269, Dec. 2007.

[9] J. Papandriopoulos, S. Dey, and J. Evans, "Optimal and Distributed Protocols for Cross-Layer Design of Physical and Transport Layers in Manets," *IEEE/ACM Trans. Networking*, vol. 16, no. 6, pp. 1392-1405, Dec. 2008.

[10] M. Chiang, C. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power Control by Geometric Programming," *IEEE Trans. Wireless Comm.*, vol. 6, no. 7, pp. 2640-2651, July 2007.

[11] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QOS and Fairness Constrained Convex Optimization of Resource Allocation for Wireless Cellular and Ad Hoc Networks," *Proc. IEEE INFOCOM*, vol. 2, pp. 477-486, June 2002.

[12] N.H. Tran and C.S. Hong, "Joint Rate and Power Control in Wireless Network: A Novel Successive Approximations Method," *IEEE Comm. Letters*, vol. 14, no. 9, pp. 872-874, Sept. 2010.

[13] A. Ghasemi and K. Faez, "Jointly Rate and Power Control in Contention Based Multihop Wireless Networks," *Computer Comm.*, vol. 30, pp. 2021-2031, 2007.

[14] B. Dogahe, M. Murthi, X. Fan, and K. Premaratne, "A Distributed Congestion and Power Control Algorithm to Achieve Bounded Average Queuing Delay in Wireless Networks," *Telecomm. Systems*, vol. 44, pp. 307-320, 2010.

[15] H.-J. Lee and J.-T. Lim, "Cross-Layer Congestion Control for Power Efficiency over Wireless Multihop Networks," *IEEE Trans. Vehicular Technology*, vol. 58, no. 9, pp. 5274-5278, Nov. 2009.

[16] S. Kandukuri and S. Boyd, "Optimal Power Control in Interference-Limited Fading Wireless Channels with Outage-Probability Specifications," *IEEE Trans. Wireless Comm.*, vol. 1, no. 1, pp. 46-55, Jan. 2002.

[17] J. Papandriopoulos, S. Dey, and J. Evans, "Optimal Power Control for Rayleigh-Faded Multiuser Systems with Outage Constraints," *IEEE Trans. Wireless Comm.*, vol. 4, no. 6, pp. 2705-2715, Dec. 2005.

[18] R. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE J. Selected Areas Comm.*, vol. 13, no. 7, pp. 1341-1347, Sept. 1995.

[19] J. Papandriopoulos, S. Dey, and J. Evans, "Distributed Cross-Layer Optimization of Manets in Composite Fading," *Proc. IEEE Int'l Conf. Comm.*, June 2006.

[20] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, Aug. 2005.

[21] J. Mo and J. Walrand, "Fair End-to-End Window-Based Congestion Control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 556-567, Oct. 2000.

[22] S. Dey and J. Evans, "Optimal Power Control in Wireless Data Networks with Outage-Based Utility Guarantees," *Proc. IEEE 42nd Conf. Decision and Control*, Dec. 2003.

[23] J.-W. Lee, M. Chiang, and A. Calderbank, "Utility-Optimal Random-Access Control," *IEEE Trans. Wireless Comm.*, vol. 6, no. 7, pp. 2741-2751, July 2007.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, Mar. 2004.

[25] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, Sept. 1999.

[26] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation*. Prentice-Hall, 1989.

[27] R. Marks and G. Wright, "A General Inner Approximation Algorithm for Nonconvex Mathematical Programs," *Operations Research*, vol. 26, pp. 681-683, 1978.

[28] R. Jain, J. Hawe, and D. Chiu, "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems," Technical Report DEC-TR-301, Sept. 1984.

[29] H. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[30] J. Zhang, D. Zheng, and M. Chiang, "The Impact of Stochastic Noisy Feedback on Distributed Network Utility Maximization," *IEEE Trans. Information Theory*, vol. 54, no. 2, pp. 645-665, Feb. 2008.

[31] D. Bertsekas and J. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, May 1996.

Nguyen H. Tran received the BS degree in electrical engineering at Hochiminh City University of Technology in 2005. He was then awarded the Korean IITA Fellowship for his graduate study at Kyung Hee University, where he received the PhD degree in computer engineering in 2011 with Best Foreign Student Award. In 2011, he was a postdoctoral research associate in Kyung Hee University. He joined the Department of Computer Engineering, Kyung Hee University in 2012, where he is currently an assistant professor. His research interest is mainly on employing various applied mathematical tools such as queueing theory, optimization theory, control theory and game theory to design, analyze and optimize the cutting-edge applications in stochastic communication networks. He is a member of the IEEE.

Choong Seon Hong received the BS and MS degrees in electronic engineering from Kyung Hee University, Seoul, Korea, in 1983 and 1985, respectively, and the PhD degree at Keio University in March 1997. In 1988, he joined KT, where he worked on Broadband Networks as a member of the technical staff. In September 1993, he joined Keio University, Japan. He had worked for the Telecommunications Network Lab at KT as a senior member of the technical staff and as a director of the networking research team until August 1999. Since September 1999, he has been working as a professor of the Department of Computer Engineering, Kyung Hee University. He has served as a program committee member and an organizing committee member for International conferences such as SAINT, NOMS, IM, APNOMS, ICOIN, CSNM, ICUIMC, E2EMON, CCNC, ADSN, ICPP, DIM, WISA, BcN, ManFI, TINA, etc. His research interests include future Internet, wireless networks, network security, and network management. He is a senior member of the IEEE and a member of the ACM, IEICE, IPSJ, KICS, KIISE, KIPS, and OSIA.

Sungwon Lee received the PhD degree from Kyung Hee University, Korea. He is a professor of the Computer Engineering Departments at Kyung Hee University, Korea. He was a senior engineer of the Telecommunications and Networks Division at Samsung Electronics Inc. from 1999 to 2008. He is an editor of the *Journal of Korean Institute of Information Scientists and Engineers: Computing Practices and Letters*. He is a member of the IEEE.