# Cross-layer Optimization for Wireless Networks with Deterministic Channel Models

Ziyu Shao*, Minghua Chen*, Salman Avestimehr+, and Shuo-Yen Robert Li*
*Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
Email: {zyshao6, minghua, bobli}@ie.cuhk.edu.hk
+School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
Email: avestimehr@ece.cornell.edu

***Abstract*—Existing work on cross-layer optimization for wireless networks adopts simple physical-layer models, i.e., treating interference as noise. In this paper, we adopt a deterministic channel model proposed in [11, 12], a simple abstraction of the physical layer that effectively captures the effect of channel strength, broadcast and superposition in wireless channels. Within the Network Utility Maximization (NUM) framework, we study the cross-layer optimization for wireless networks based on this deterministic channel model. First, we extend the well-applied conflict graph model to capture the flow interactions over the deterministic channels and characterize the feasible rate region. Then we study distributed algorithms for general wireless multi-hop networks. The convergence of algorithms is proved by Lyapunov stability theorem and stochastic approximation method. Further, we show the convergence to the bounded neighborhood of optimal solutions with probability one under constant steps and constant update intervals. Our numerical evaluation validates the analytical results.**

## I. Introduction

Since Kelly's seminal work [1], the network utility maximization (NUM) framework [2] has attracted significant attentions for cross-layer optimization of wireless networks. In this framework, network protocols are understood as distributed algorithms that maximize aggregate user utility under network resource constraints.

Existing studies on wireless NUM [2] mainly base on a simple physical layer assumptions: treating interference as noise. However, it is known that "treating interference as noise" is in general sub-optimal and higher rates can be achieved by using more advanced physical layer coding techniques, such as superposition coding, successive interference cancelation and interference alignment.

Recently, a deterministic channel model is proposed in [11, 12]. This deterministic channel model simplifies the wireless flow interaction by eliminating the noise and allows studies to focus on interferences among transmissions. The deterministic channel provides an accurate, yet simple, abstraction of the physical layer that can be utilized effectively for the cross-layer protocol design [12].

On the other hand, the optimal scheduling is still a challenging problem for wireless networks [3, 4], irrespective of the underlying physical layer model. An adaptive CSMA based distributed scheduling algorithm is proposed in [6] recently. This algorithm has been shown to be throughput-optimal and utility-optimal asymptotically [6, 7, 8, 15]. Inspired by this series of work, we propose in [15] a Markov approximation framework for synthesizing distributed algorithms for general combinatorial network optimization problems, including the optimal scheduling problem as a special case.

Motivated by these progresses, in this paper, we study the cross-layer optimization for wireless networks with deterministic channel models, i.e., jointly optimizing flow control and scheduling. The key results and contributions are as follows:

- **Characterization of the feasible rate region by extended conflict graph based model:** Existing conflict-graph model [4, 10] are based on simple physical-layer assumptions. We extend the conflict-graph model to capture the flow interactions over the deterministic channels. In general, the resulted feasible rate region is a subset of the information-theoretic capacity region. In some special cases such as single-hop multiple-access networks, this feasible rate region is shown to be the same as the information-theoretic capacity region.
- **Distributed solutions for NUM over general wireless multi-hop networks with deterministic channel models:** By using Markov approximation framework [15], we construct Markov chains to approximately solve the optimal scheduling problem. Using Lagrange dual decomposition method to construct primal-dual algorithms to solve the optimal flow control problem, we solve the approximated cross-layer optimization problem within the NUM framework in a distributed way.
- **Convergence of the primal-dual flow control algorithm with or without time-scale separation assumption:** Existing work [6, 7, 8] focuses on the dual algorithm and prove its convergence. In this paper, we focus on the primal-dual algorithm because its smoother dynamics than that of the dual algorithm. With the time-scale separation assumption that Markov chain converges to its stationary distribution instantaneously, the proposed primal-dual algorithm is shown to converge to the optimal solutions globally asymptotically. Without such time-scale separation assumption, the resulted stochastic primal-dual algorithm is shown to converge to a bounded neighborhood of the same optimal solutions with probability one under constant step size and constant update interval.

Proofs and details are provided in [16].

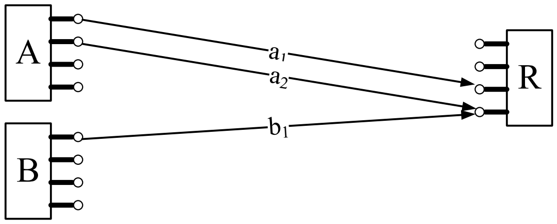


Fig. 1. One example of deterministic channel based wireless network with channel gains $\rho_{AR} = 2$ and $\rho_{BR} = 1$. Each rectangle represents a transmitter or receiver node, and each knob attaching to a rectangle represents a signal level and what is sent on it is a bit.

## II. System Model and Problem Formulation

We consider a wireless network with a set of $N$ users (source-destination pairs), denoted as $S = \{1, \ldots, N\}$. Each user is associated with a sending rate $x_s$, as well as a utility function $U_s(x_s)$ that measures the efficiency and fairness of resource allocation algorithms. We assume the utility function to be twice differentiable, increasing and strictly concave [2]. The utility maximization problem is as follows

$$\textbf{P1:} \quad \max_{\boldsymbol{x} \geq 0} \sum_{s \in S} U_s(x_s) \qquad (1)$$
$$\text{s.t.} \quad \boldsymbol{x} \in \{\text{feasible rate region}\}.$$

where $\boldsymbol{x}$ is the vector of user's rate (we use bold symbols to denote vectors through the whole paper) and the feasible rate region is a function of information transmission constraints.

### A. Deterministic Channel Model

In this paper, we study the problem **P1** over the feasible rate region based on the deterministic channel model. A formal definition of this channel model is as follows.

**Definition 1.** ***(Definition of the deterministic model [11, 12])*** *Consider a wireless network consisting of a set of nodes* V *and a set of channels, where $K = |V|$ is the number of nodes. Communication from node $i$ to node $j$ has a non-negative integer gain $\rho_{(i,j)}$ associated with it. This number models the channel gain in a corresponding Gaussian setting. At each time $t$, node $i$ transmits a vector $\mathbf{x}_i[t] \in \mathbb{F}_2^{\omega}$ and receives a vector $\mathbf{y}_i[t] \in \mathbb{F}_2^{\omega}$ where $\omega = \max_{i,j}(\rho_{(i,j)})$. The received signal at each node is a deterministic function of the transmitted signals at the other nodes, with the following input-output relation: if the nodes in the network transmit $\mathbf{x}_1[t], \mathbf{x}_2[t], \ldots \mathbf{x}_K[t]$ then the received signal at node j, $1 \leq j \leq K$ is: $\mathbf{y}_j[t] = \sum_{k=1}^{K} \mathbf{W}^{\omega - \rho_{k,j}} \mathbf{x}_k[t]$ for all $1 \leq k \leq V$, where $\mathbf{W}$ is the $\omega \times \omega$ shift matrix and the summation and multiplication is in $\mathbb{F}_2$.*

An example is shown in Fig. 1. In this figure, each signal is a sequence of bits at different signal levels, with the highest signal level being the most significant bit (MSB) and the lowest level being the least significant bit (LSB). The transmit and received signal levels are sorted from MSB (top) to LSB (bottom). The channel gain between two nodes *i* and *j* indicates how many of the first MSB transmitted signal levels of node *i* are received at destination node *j*.

There is a constraint on configurations of sub-links belonging to the same channel [12]: activated sub-links are consecutive from high signal levels to low signal levels, *not necessary* starting from MSB. When non-consecutive signal levels are used in the transmission, we need to use superposition code to encode the disjoint set of sublinks belonging to the same channel. Consequently, the complex code generation induces the loss of energy efficiency and the approximation accuracy to the Gaussian model [12], which depends on the number of "abruptiones" in signal levels. There is a tradeoff between 1) the deterministic model's rate region reduction caused by the consecutive signal level constraint, and 2) its approximation accuracy to the Gaussian model. This fundamental tradeoff deserves its own investigation [12]. In this paper, we focus on achieving a good network-wise performance with practically implementable schemes. As such, we cast the consecutive signal level constraint to get a higher approximation accuracy. So for the channel $l$ with the channel gain $\rho_l$, there are $\frac{\rho_l(\rho_l+1)}{2}$ *configurations* of channel $l$, i.e., possible combinations of consecutive sub-links within the channel $l$.

### B. Conflict Graph Based Model

The wireless network is represented as a graph *G=(V, L)* [1], where *V* is the set of nodes and *L* is the set of links(channels) between nodes. Each link(channel) $(i, j) \subset L$ consists of $\rho_{(i,j)}$ sub-links (each with one-unit capacity), where $i, j \in V$ and $\rho_{(i,j)}$ is the channel gain from node $i$ to node $j$.

Note that for a general wireless network with physical layer assumptions such as treating interference as noise, the existing conflict graph model [4, 10] is equivalent to the collision model, where if two interfering links transmit packets simultaneously, both packets are dropped. However, the interference itself actually carries information and has structure that can be potentially be exploited in mitigating its effect.

On the other hand, for wireless networks under deterministic channel modeling, we can exploit the structure of interference by superposition coding and interference cancelation [12]. Unlike the collision model where entire messages are lost when there is collision, the most significant bits of the stronger users remain intact. Then we propose an extended conflict graph model to capture the interference over the deterministic channels, shown as follows:

**Definition 2.** *The conflict graph $G_c$ of a graph $G$ is an undirected graph $G_c = (V_c, A)$, where vertex set $V_c$ correspond to possible configurations of all channels in G, and edge set $A$ represents the connection of adjacent vertices. Two vertices are adjacent if either they are two different configurations of the same channel or there exists at least two sub-links conflict with each other, i.e., intersecting at the same signal level of the same network node. All possible configurations of the same channel are conflict with others, since only one configuration can be activated at a time.*

**Definition 3.** *An independent set of the conflict graph $G_c = (V_c, A)$ is a subset $M \subset V_c$ of vertices that no two of which are adjacent (i.e., $(s, t) \notin A$ for all $s, t \in M$).*

[1]In broadcast scenarios, by modeling wireless networks as Hypergraphs, we can directly extend the definitions of conflict graph and independent set to hypergraph models.

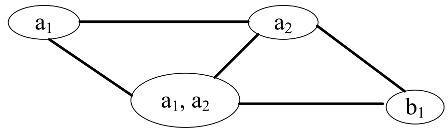


Fig. 2. Corresponding conflict graph for the network shown in Fig. 1. All nonempty independent sets are: $\{a_1\}$, $\{a_2\}$, $\{b_1\}$, $\{a_1, a_2\}$, and $\{\{a_1\}, \{b_1\}\}$.

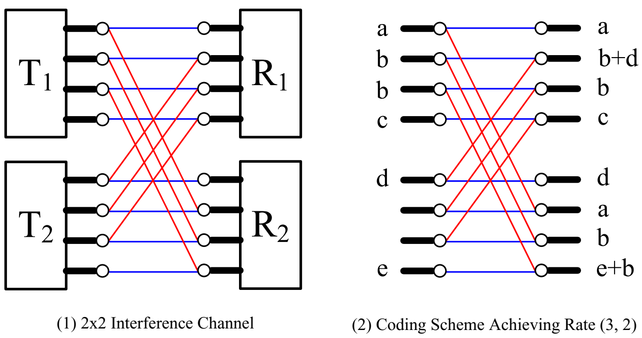


Fig. 3. [5]. Subfig 1 shows a 2x2 deterministic interference channel with channel gains $\rho_{T_1 R_1} = \rho_{T_2 R_2} = 4$ and $\rho_{T_1 R_2} = \rho_{T_2 R_1} = 3$. Subfig 2 shows a coding scheme achieving the rate tuple (3,2), where $T_1$ sends $a, b, b, c$ to $R_1$ and $T_2$ sends $d, \emptyset, \emptyset, e$ to $R_2$. $R_1$ can decode and obtain $a, b, c$, while $R_2$ can decode and obtain $d, e$.

Fig. 2 shows the conflict graph and associated non-empty independent sets for the example in Fig. 1.

The feasibility of simultaneous transmissions can be captured by a conflict graph associated with the wireless network. As a result, the feasible rate region is characterized as a convex hull of the feasible rates supported by possible individual independent sets on this conflict graph [10].

For some important cases such as deterministic multiple-access channels, the conflict graph based rate region is equal to the information-theoretic capacity region.

In general, the conflict graph based rate region is the subset of the information-theoretic capacity region. An example is shown in Fig. 3 [5]. It is not hard to see that the rate tuple (3,2) is out of the conflict graph based rate region, since in our conflict graph model, two bits arrive simultaneously at the same signal levels of the receiver are dropped.

For general wireless networks, the characterization of the information-theoretic capacity region is open. On the other hand, the rate region achieved by the conflict graph based model we presented is achievable by superposition coding and interference cancelation [12]. In this paper, we focus on the conflict graph based rate region.

## III. NUM Over General Multi-Hop Network

Consider a multi-hop network *G=(V, L)*, where each user is associated with a single path. Let $H$ be the set of all independent sets over the corresponding conflict graph $G_c$. Let $\mathbf{q} = [q_h, h \in H]^T$ be the vector of probability (or time fraction) of all independent sets. Let $\boldsymbol{x} = [x_s, s \in S]^T$ be the vector of sending rates of users. Let $\lambda_{l,h}$ be the capacity of link $l$ within the independent set $h$. $\lambda_{l,h} = 0$ means link $l$ is not activated within the independent set $h$. We also let $\{s : l \in s\}$ denote the set of users that sharing the link $l$.

Consider the following master utility maximization problem over $G_c$.

$$\mathbf{MP}: \max_{\boldsymbol{x} \geq 0, \boldsymbol{q} \geq 0} \sum_{s \in S} U_s(x_s) \tag{2}$$

$$\text{s.t.} \quad \sum_{s: l \in s} x_s \leq \sum_{h \in H} \lambda_{l,h} q_h, \quad \forall l \in L \tag{3}$$

$$\sum_{h \in H} q_h = 1.$$

Solving the master problem $\mathbf{MP}$ (2) is very challenging because the scheduling subproblem is NP-hard in general. To see that, first, by relaxing the first set of inequality constraints (3) in problem $\mathbf{MP}$, we get its partial Lagrangian:

$$L(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{r}) = \sum_{s \in S} U_s(x_s) - \sum_{l \in L} r_l \left( \sum_{s: l \in s} x_s - \sum_{h \in H} \lambda_{l,h} q_h \right),$$

where $\boldsymbol{r} = [r_l, l \in L]^T$ is the vector of Lagrange multipliers. Since the Slater constraint qualification conditions hold for the problem $\mathbf{MP}$ [13], the strong duality holds. Thus the problem $\mathbf{MP}$ can be solved by finding the saddle points of $L(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{r})$ via solving the following problem:

$$\begin{aligned} \mathbf{DP}: \min_{\boldsymbol{r} \geq 0} ( & \max_{\boldsymbol{x} \geq 0} ( \sum_{s \in S} U_s(x_s) - \sum_{s \in S} x_s \sum_{l: l \in s} r_l ) \\ & + \max_{\boldsymbol{q} \geq 0} ( \sum_{h \in H} q_h \sum_{l \in L} \lambda_{l,h} r_l )) \\ \text{s.t.} & \sum_{h \in H} q_h = 1. \end{aligned} \tag{4}$$

The above problem can be solved successively in $\boldsymbol{q}, \boldsymbol{x}, \boldsymbol{r}$. The key challenge lies in solving the sub-problem in $\boldsymbol{q}$, which is the Maximum Weighted Independent Set (MWIS) problem [14]:

$$\mathbf{MWIS}: \max_{\boldsymbol{q} \geq 0} \sum_{h \in H} q_h \sum_{l \in L} \lambda_{l,h} r_l \tag{5}$$

$$\text{s.t.} \quad \sum_{h \in H} q_h = 1.$$

The optimal value of the $\mathbf{MWIS}$ problem is given by computing the max function: $\max_{h \in H} \sum_{l \in L} \lambda_{l,h} r_l$.

### A. Markov Approximation

The $\mathbf{MWIS}$ problem is NP-hard [14] and hard to approximate even in a centralized manner [14]. Here we apply the Markov approximation framework [15] to solve the problem in a distributed way. First, we apply the log-sum-exp approximation

$$\max_{h \in H} \sum_{l \in L} \lambda_{l,h} r_l \approx \frac{1}{\beta} \log \left[ \sum_{h \in H} \exp \left( \beta \sum_{l \in L} \lambda_{l,h} r_l \right) \right], \tag{6}$$

where $\beta$ is a positive constant. As $\beta \to \infty$, there is no gap between log-sum-exp approximation and the max function [15, Propostition 1]. According to [15, Theorem 1], we are implicitly solving an approximated version of the problem $\mathbf{MWIS}$, off by an entropy term $-\frac{1}{\beta} \sum_{h \in H} q_h \log q_h$, as follows

$$\mathbf{MWIS} - \beta: \max_{\boldsymbol{q} \geq 0} \; -\frac{1}{\beta} \sum_{h \in H} q_h \log q_h + \sum_{h \in H} q_h \sum_{l \in L} \lambda_{l,h} r_l$$

$$\text{s.t.} \quad \sum_{h \in H} q_h = 1. \tag{7}$$

The unique optimal solution of $\mathbf{MWIS}-\beta$ is

$$q_h(\beta \boldsymbol{r}) = \frac{\exp\left(\beta \sum_{l\in L} \lambda_{l,h}\, r_l\right)}{\sum_{h\in H} \exp\left(\beta \sum_{l\in L} \lambda_{l,h}\, r_l\right)}, \forall h \in H. \quad (8)$$

.

By the log-sum-exp approximation, we are actually solving a problem close to the original problem $\mathbf{MP}$:

$$\begin{aligned} \mathbf{MP}-\beta: \max_{\boldsymbol{x}\geq 0, \boldsymbol{q}\geq 0} \quad & \textstyle\sum_{s\in S} U_s(x_s) - \frac{1}{\beta}\sum_{h\in H} q_h \log q_h \quad (9)\\ \text{s.t.} \quad & \textstyle\sum_{s:l\in s} x_s \leq \sum_{h\in H} \lambda_{l,h}\, q_h, \quad \forall l \in L \\ & \textstyle\sum_{h\in H} q_h = 1. \end{aligned}$$

Denote the optimal solution of master problem $\mathbf{MP}$ (2) by $\boldsymbol{x}^\star$, and the optimal solution of problem $\mathbf{MP}-\beta$ (9) by $\hat{\boldsymbol{x}}$, then $|\sum_{s\in S}(U_s(\hat{x}_s) - U_s(x_s^*))| \leq \frac{\log|H|}{\beta}$. As $\beta \to \infty$, $\hat{\boldsymbol{x}}$ approaches $\boldsymbol{x}^\star$.

$\mathbf{MP}-\beta$ is equivalent to the following problem:

$$\mathbf{DP}-\beta: \min_{\boldsymbol{r}\geq 0}\max_{\boldsymbol{x}\geq 0} L_\beta(\boldsymbol{x},\boldsymbol{r}) = \sum_{s\in S}(U_s(x_s) - x_s \sum_{l:l\in s} r_l) + \frac{1}{\beta}\log[\sum_{h\in H}\exp(\beta\sum_{l\in L}\lambda_{l,h}\, r_l)] \quad (10)$$

We explore algorithm design in the following subsections. The $q_h(\beta\boldsymbol{r}), h \in H$ in (8) can be interpreted as the stationary distribution of a time reversible Markov chain, whose states are the independent sets in $H$. We first discuss how to design and implement such a Markov chain in a distributed manner, then we design primal-dual algorithms to solve the problem $\mathbf{DP}-\beta$.

### B. Design and Implementation of Markov Chain

To construct a time-reversible Markov chain with its stationary distribution $q_h(\beta\boldsymbol{r}), h \in H$ in (8), we let $h \in H$ be the state of the Markov chain. We start by only allowing direct transitions between two "adjacent" states (independent sets) $h$ and $h'$ that differ by one and only one link. Note that doing so will not affect the stationary distribution for time-reversible Markov chains [15, Section II]. By this design, the transition from $h$ to $h' = h \cup \{l'\}$ corresponds to link $l'$ starting its transmission. Similarly, the transition from $h'$ to $h$ corresponds to link $l'$ finishing its on-going transmission.

Now, consider two states $h$ and $h'$ where $h' = h\cup\{l'\}$. We set $q_{h',h}$ to $\lambda_{l',h'}$, and

$$\begin{aligned} q_{h,h'} &= \lambda_{l',h'}\exp(\beta(\sum_{l\in h'}\lambda_{l,h'}\, r_l - \sum_{l\in h}\lambda_{l,h}\, r_l)) \\ &= \lambda_{l',h'}\exp\left(\beta\lambda_{l',h'}\, r_{l'}\right). \end{aligned}$$

To achieve transition rate $q_{h,h'}$, the transmitter of link $l'$ waits for a back-off time that follows exponential distribution with rate $\lambda_{l',h'}\exp\left(\beta\lambda_{l',h'}\, r_{l'}\right)$ before it starts to transmit. During the count-down, if the link $l'$ (in particular its transmitter) determines another interfering link is in transmission, link $l'$ will freeze its count-down process. This could be done in various ways, for instance by the receiver of link $l'$ communicating busy/idle notification to the transmitter using a dedicated low-rate feedback channel.

When the transmission is over, link $l'$ counts down according to the residual back-off time, which is still exponential distributed with the same rate, because of the memoryless property of exponential distribution.

The transition rate $q_{h',h}$ can be achieved by link $l'$ setting its transmission time to follow exponential distribution with rate $\lambda_{l',h'}$.

Similar to the proof in [15, Section II], we can show that the above procedure in fact implements a time-reversible Markov chain with stationary distribution in (8).

### C. Solving the Approximated Problem by the Primal-Dual Algorithm

The problem $\mathbf{DP}-\beta$ (10) can be solved by either a dual algorithm or a primal-dual algorithm. Existing work [6, 7] all focused on dual algorithms. We prefer the primal-dual algorithm because of of its fast convergence rate (only one time-scale) and smoothness of changes in parameters.

Define the user rates as $x_s, s \in S$ and the link prices as $r_l, l \in L$, we propose a primal-dual algorithm as follows:

$$\begin{cases} \dot{x}_s = \alpha_s\left[U_s'(x_s) - \sum_{l:l\in s} r_l\right]^+_{x_s} \\ \forall s \in S \ \text{ user rates updating} \\ \dot{r}_l = k_l\left[\sum_{s:l\in s} x_s - \sum_{h\in H}\lambda_{l,h}\, q_h\right]^+_{r_l} \\ \forall l \in L \ \text{ link prices updating} \end{cases}, \quad (11)$$

where $k_s (s \in S)$ and $\alpha_s (s \in S)$ are positive constants, and function $[b]^+_a = \max(0, b)$ if $a \leq 0$ and equals $b$ otherwise. With the time-scale separation assumption that Markov chain converges to its stationary distribution instantaneously compared to the time-scale of adaption of $\boldsymbol{x}$ and $\boldsymbol{r}$, we have the following result:

**Theorem 1.** *The primal-dual algorithm* (11) *is globally asymptotically stable.*

Since the equilibrium point of primal-dual algorithm (11) solves the problem $\mathbf{DP}-\beta$ (10) exactly, it also solves the problem $\mathbf{MP}-\beta$ (9) exactly. When $\beta \to \infty$, the primal-dual algorithm (11) solves the master problem $\mathbf{MP}$ (2) in a distributed way.

### D. Convergence of Stochastic Primal-Dual Algorithm

Without the time-scale separation assumption on Markov chain, the above primal-dual algorithm (11) turns to a stochastic primal-dual algorithm, given as follows:

$$\begin{cases} x_s(m+1) = \left[x_s(m) + \epsilon(m)\left(U_s'(x_s(m)) - \sum_{l:l\in s} r_l(m)\right)\right]_+ \\ \forall s \in S \quad \text{user rates updating} \\ r_l(m+1) = \left[r_l(m) - \epsilon(m)\left(\bar{\theta}_l(m) - \sum_{s:l\in s} x_s(m)\right)\right]_+ \\ \forall l \in L \quad \text{link prices updating} \end{cases}, \quad (12)$$

Where $[\cdot]_+ \triangleq \max(\cdot, 0)$, $\epsilon(m)$ is the step size, $\bar{\theta}_l(m)$ is the average link rate measured by link $l$ within the update interval $T_m$, and $T_m$ is the time interval between the system updating $(\boldsymbol{x}(m-1), \boldsymbol{r}(m-1))$ and $(\boldsymbol{x}(m), \boldsymbol{r}(m))$.

We omit convergence result of the stochastic primal-dual algorithm (12) under diminishing step sizes and increasing update intervals because we want to emphasis the following convergence result under constant step size and constant update intervals.

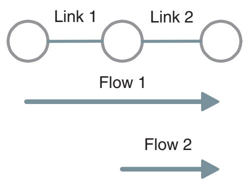


Fig. 4. A simple network with two links and two flows. The channel gain of link 1 is *2* and channel gain of link 2 is *1*. Flow 1 uses link 1 and link 2, and flow 2 uses only link 2.

**Theorem 2.** *Assume that* $U_s^{'}(0) < \infty, \forall s \in S$, $\max_{s,m} x_s(m) < \infty$ *and* $\max_{l,m} r_l(m) < \infty$. *If the sequence of step size* $\{\epsilon(m)\}$ *and the sequence of update interval* $\{T_m\}$ *satisfy the following conditions:*

$$T_m = T_0 > 0 \quad \forall m \tag{13}$$

$$\epsilon(m) = \epsilon > 0 \quad \forall m \tag{14}$$

*By running the stochastic primal-dual algorithm* (12), $(\boldsymbol{x}(m), \boldsymbol{r}(m))$ *converges with probability one to the bounded neighborhood of* $\hat{\boldsymbol{x}}$ *and* $\hat{\boldsymbol{r}}$, *i.e., optimal solutions of* $\mathbf{DP} - \beta$ (10) *as follows,*

$$\{(\boldsymbol{x}, \boldsymbol{r}) : |L_\beta(\boldsymbol{x}, \boldsymbol{r}) \ - \ L_\beta(\hat{\boldsymbol{x}}, \hat{\boldsymbol{r}})| \leq \frac{C_3}{T_0} + \epsilon\frac{(C_1 + C_2)}{2}\},$$

*where* $C_1, C_2, C_3$ *are positive constants.*

Our emphasis on the constant step size and constant update interval has two reasons. First, a diminishing step size usually leads to slow convergence near the optimal solutions. Second, it is convenient to implement the constant step size and constant update interval in practice.

Inspired by and similar to [8, 9], we also adopt the standard methods of stochastic approximation and Markov chain. The difference between our proof and [8, 9] is that, our proof studies the saddle points of Lagrangian function, while [8, 9] studies the optimal dual solutions directly.

## IV. Numerical Examples

In this section, we present numerical experiments to illustrate the performance of the primal-dual algorithms. We consider a simple network scenario shown in Fig. 4. By focusing on proportional-fairness, we choose utility function $U(\cdot) = \log(\cdot + 10^{-6})$. By running the primal-dual algorithm (12) with constant step size *0.05*, constant update interval *100* and approximating factor $\beta = 100$, we have the corresponding results shown in Fig. 5. We compare the numerical results with theoretically optimal values in Table I. The numerical results are close to the optimal solutions, within 2.75% of the optimal values. Further, for the summation of utility, the gap between the numerical result and the optimal value is negligible. Thus numerical results illusrate the convergence and optimality of our joint scheduling and primal-dual flow control algorithms.

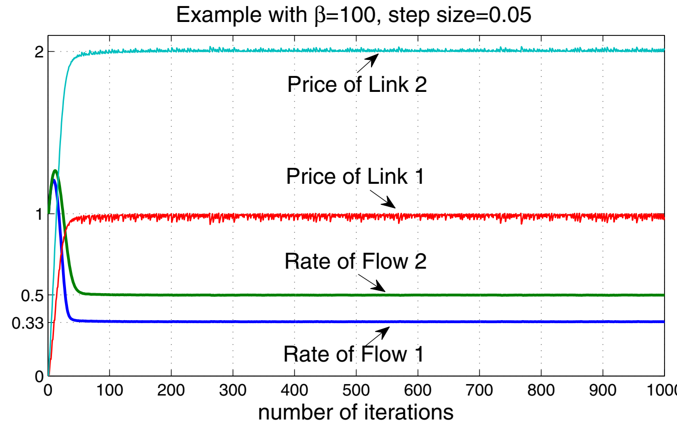


Fig. 5. Performance of the primal-dual algorithm on flow rates and link shadow prices (queue-lengths). Initial values of flow rates are both *1* and link prices are both *0*.

TABLE I
Performance of Primal-Dual Algorithms

| | Optimal | Approximation | Relative Error |
|---|---|---|---|
| Sum of Utility | -1.7918 | -1.7919 | 0.0056% |
| Rates of Flow 1 | 0.3333 | 0.3340 | 0.21% |
| Rates of Flow 2 | 0.5000 | 0.4990 | 0.20% |
| Prices of Link 1 | 1.0000 | 0.9725 | 2.75% |
| Price of Link 2 | 2.0000 | 2.0124 | 0.62% |

## References


[1] F. P. Kelly, A. Maulloo, and D. Tan, “Rate control for communication networks: shadow prices, proportional fairness, and stability,” *Journal of the Operationl Research Society*, pp. 237-252, 1998

[2] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, “Layering as optimization decomposition:a mathematical theory of network architectures,” *Proc. IEEE*, vol. 95, no. 1, pp. 255-312, Jan. 2007.

[3] L. Tassiulas and A. Ephremides, “Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks”, *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936-1949, Dec. 1992.

[4] X. Lin, N. B. Shroff and R. Srikant, “A Tutorial on Cross-Layer Optimization in Wireless Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, August 2006.

[5] G. Bresler and D. Tse, “The Two-User Gaussian Interference Channel: A Deterministic View”, *European Transactions in Telecommunications*, vol. 19, pp. 333-354, April 2008.

[6] L. Jiang and J. Walrand, “A distributed CSMA algorithm for throughput and utility maximization in wireless networks”, *Allerton* 2008.

[7] J. Liu, Y. Yi, A. Proutiere, M. Chiang and H. V. Poor, “Maximizing Utility via Random Access Without Message Passing”, Preprint:arXiv:0902.1996.

[8] L. Jiang and J. Walrand, “Convergence and Stability of a Distributed CSMA Algorithm for Maximal Network Throughput”, *Technical Report* UCB-EECS-2009-43.

[9] L. Jiang, D. Shah, J. Shin, and J. Walrand, “Distributed Random Access Algorithm: Scheduling and Congestion Control”. Preprint:arXiv:0907.1266.

[10] K. Jain, J. Padhye, V.N. Padmanabhan, and L. Qiu, “Impact of Interference on Multi-Hop Wireless Network Performance”, *Wireless Networks*, 2005.

[11] S. Avestimehr, S. Diggavi and D. Tse, “A deterministic approach to wireless relay networks” , *Allerton Conference*, Sept. 2007.

[12] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, “Wireless network information flow: a deterministic approach,” submitted to *IEEE Trans. on Inf. Theory*. Preprint: arXiv:0906.5394v1.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[14] V. Vazirani, *Approximation Algorithms*, Springer-Verlag, Berlin, 2001.

[15] M. Chen, S. Liew, Z. Shao, and C. Kai, “Markov Approximation for Combinatorial network Optimization”, *IEEE Infocom* 2010.

[16] Z. Shao, M. Chen, S. Avestimehr, and S. Y. Robert Li, “Cross-layer Optimization for Wireless Networks with Deterministic Channel Models”, *CUHK Technical Report 2009*. Available at http://personal.ie.cuhk.edu.hk/~zyshao6/numdtp.pdf