Achievable Rate and Optimal Physical Layer Rate Allocation in Interference-Free Wireless Networks

Tao Cui and Tracey Ho
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, USA
Email: {taocui, tho}@caltech.edu

Jörg Kliewer
Klipsch School of Electrical and Computer Engineering
New Mexico State University
Las Cruces, NM 88003, USA
Email: jkliewer@nmsu.edu

Abstract—We analyze the achievable rate in interference-free wireless networks with physical layer fading channels and orthogonal multiple access. As a starting point, the point-to-point channel is considered. We find the optimal physical and network layer rate trade-off which maximizes the achievable overall rate for both a fixed rate transmission scheme and an improved scheme based on multiple virtual users and superposition coding. These initial results are extended to the network setting, where, based on a cut-set formulation, the achievable rate at each node and its upper bound are derived. We propose a distributed optimization algorithm which allows to jointly determine the maximum achievable rate, the optimal physical layer rates on each network link, and an opportunistic back-pressure-type routing strategy on the network layer. This inherently justifies the layered architecture in existing wireless networks. Finally, we show that the proposed layered optimization approach can achieve almost all of the ergodic network capacity in high SNR.

I. INTRODUCTION

In wireless communication networks, channel impairments such as fading, shadowing and path loss limit the capacity. Determining the capacity region and corresponding achievable strategy for general multiterminal networks is a long-standing open problem. An outer bound for the capacity region is known to have cut-set interpretation, but this cut-set bound is not always achievable. For example, there is no known scheme to achieve this outer bound for the simple relay channel. In contrast, in wireline networks with orthogonal channels, the cut-set bound is achievable even in the multicast scenario, where multiple destinations demand the same information from several sources. In this case, the capacity of wireline networks can be achieved via a separation of physical layer channel coding and higher layer network coding. Furthermore, in [1] it is shown that the cut-set bound is tight in wireless erasure networks. Also, related work in [2] shows that the capacity of networks with deterministic channels and broadcasting has a cut-set formulation. The optimal rate trade-off between physical and network layer is found analytically in [3] for a point-to-point channel under an overall delay constraint, and by simulation in [4][1] for minimum energy consumption in cellular broadcast networks. However, similar achievability results do not yet exist for more general wireless networks with fading channels.

In this paper, we study the achievable rate and optimal physical layer rate allocation for a special class of wireless networks, where for each node there is no interference between the received signals from its neighbors. This could be

realized by using orthogonal access strategies such as time-, frequency-, or code-division multiple access. Further, the channels between each node and one of its neighbors are modeled as slow fading channels. For the case where channel state information (CSI) is only known at the receiver, each node simply transmits at a constant rate regardless of the fading state. Therefore, each node's transmitted information gets lost when the transmission rate is high so that the current channel cannot support it, which results in outage. For the point-to-point case we present a throughput-efficient coding scheme by partitioning the transmit power across multiple virtual users with different rates. The corresponding information is then transmitted using superposition coding. This approach outperforms the constant rate scheme in terms of achievable rate. For an interference-free network with fading channels we consider the outage probability as the link erasure probability in [1]. Different from [1] where the erasure probability is given, the outage probability depends on the transmission rate at each node. We first obtain the achievable rate of this network and its outer bound based on a cut-set formulation. Further, by using a flow formulation, we derive a distributed optimization algorithm based on a modified dual decomposition approach to maximize the achievable rate. As a byproduct, the algorithm also returns the optimal physical layer rate at each node, and an optimal opportunistic back-pressure-type routing strategy on the network layer. Finally, we bound the gap between the achievable rate and the outer bound. The result reveals that the proposed layered structure can achieve almost all of the ergodic network capacity in high SNR, which justifies the pervasive layered architecture in existing wireless networks.

II. PRELIMINARIES

A. Network Model

We model the wireless network by a directed acyclic graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$, where $\mathcal{V}$ is the vertex set and $\mathcal{E}$ is a directed edge set. For each node $i\in\mathcal{V}$, $\mathcal{N}_{\mathcal{O}}(i)$ and $\mathcal{N}_{\mathcal{I}}(i)$ denote the set of in-neighbors and out-neighbors of i, i.e., $\mathcal{N}_{\mathcal{I}}(i)=\{j|(j,i)\in\mathcal{E}\}$, $\mathcal{N}_{\mathcal{O}}(i)=\{j|(i,j)\in\mathcal{E}\}$. A cut for $x,y\in\mathcal{V}$ is a partition of $\mathcal{V}$ into two sets $\mathcal{V}_x$ and $\mathcal{V}_y=\mathcal{V}_x^c$ such that $x\in\mathcal{V}_x$ and $y\in\mathcal{V}_y$. Furthermore, we define $\mathcal{V}_x^*$ to be the set of nodes in $\mathcal{V}_x$ that have at least one outgoing edge in $\mathcal{V}_x^c$, i.e., $\mathcal{V}_x^*=\{i|\exists j, \text{ such that } (i,j)\in\mathcal{E}\}$. We focus on acyclic graphs in the following.

In this paper, each edge $(i,j)\in\mathcal{E}$ represents a memoryless Gaussian channel from node i to node j. Also, we assume

---

that each node has a unit bandwidth. Let x_i be the transmitted signal by node i and $y_{i,j}$ be the received signal at node j from node i, where the average power of x_i is P_i, e.g., $E\{|x_i|^2\}=P_i$. We have the channel model

$$y_{i,j}=\sqrt{h_{i,j}}x_j+v_{i,j}, \tag{1}$$

where $\sqrt{h_{i,j}}$ is the channel fading between node i and node j and $v_{i,j}$ is the additive white Gaussian noise at node j for the signal reception from node i. Without loss of generality, we assume that $v_{i,j}$ has zero mean and unit variance. The variance $\sigma_{i,j}^2$ of $h_{i,j}$ is given and normalized such that $0\leq\sigma_{i,j}^2\leq1$. The channel gain $\sqrt{h_{i,j}}$ is assumed to be constant over each packet of size n and to vary identically and independently between different packets. We consider a Rayleigh fading channel, where the probability density function (pdf) of $h_{i,j}$ is

$$f(h_{i,j})=\frac{1}{\sigma_{i,j}^2}e^{-\frac{h_{i,j}}{\sigma_{i,j}^2}}, \tag{2}$$

and $h_{i,j}$ and $h_{i,j'}$ are independent $\forall j,j'\in\mathcal{N}_{\mathcal{O}}(i),j\neq j'$. However, the results extend to arbitrary fading distributions and correlated fading channels in a straightforward way.

In the following we address multicast problems where a set of sinks $\mathcal{D}=\{d_1,...,d_{|\mathcal{D}|}\}\subset\mathcal{V}$ demands all of the information from a set of sources. In this paper, we focus on the case of a single source $s\in\mathcal{V}$. The case of multiple sources can be generalized in the same way as in [1].

Our encoding scheme is given by layered scheme based on a product code [5], where a horizontal channel code is used for hop-by-hop physical layer error correction and a vertical erasure correcting code is used for end-to-end error correction on the network layer. This provides an implicit block interleaving as an outage for a physical layer block leads to individual symbol erasures on the network layer.

III. ACHIEVABLE RATE IN POINT-TO-POINT CHANNELS

We first study the point-to-point channel, where the system model $y=\sqrt{h}x+v$ is as described in the previous section but with the subscripts i,j removed for brevity. Our objective is to maximize the average throughput by optimizing the rate allocation for both network and physical layer.

When CSI is only known at the receiver the transmitter cannot adapt its transmission rate according to h. In this case the capacity is [6]

$$C=E_h\left\{\frac{1}{2}\log(1+hP)\right\}, \tag{3}$$

if maximum likelihood decoding is used [6]. Simple encoding and decoding techniques for AWGN channels cannot be applied directly. Without transmitter CSI, one method is to let the transmitter transmit at a constant rate R. Therefore, if $R\leq C(h)=\frac{1}{2}\log(1+hP)$, the receiver can successfully decode the transmitted signal. To maximize the average throughput, we need to solve

$$\max_R R\int_{\frac{2^{2R}-1}{P}}^{+\infty}f(h)dh=\max_R Re^{-\frac{2^{2R}-1}{P\sigma^2}}=\max_R F(R), \tag{4}$$

where we have used the Rayleigh fading pdf in (2). In (4), $F(R)$ denotes the overall rate, R the rate on the physical layer, and the term $\exp(-\frac{2^{2R}-1}{P\sigma^2})$ corresponds to the network layer rate as a function of R. By taking the derivative of $F(R)$

with respect to R and setting the result to zero, we obtain the optimal physical layer rate as

$$R^*=\frac{1}{2\ln2}W(P\sigma^2), \tag{5}$$

where $W(\cdot)$ is the Lambert W function, the inverse function of $g(w)=we^w$. By substituting R^* into $F(R)$ in (4), we obtain the maximal average overall throughput as

$$F(R^*)=R^*e^{-\frac{2^{2R^*}-1}{P\sigma^2}}=\frac{1}{2\ln2}W(P\sigma^2)e^{-\frac{1}{W(P\sigma^2)}}e^{\frac{1}{P\sigma^2}}. \tag{6}$$

We now discuss a coding scheme which is able to improve the throughput compared to the constant rate case. The idea was first presented in [7] and is based on superposition coding by introducing an infinite number of virtual users at the transmitter. We then propose a modification of this approach by considering the practically more relevant case for a finite number of virtual users.

Suppose that each of the users transmits with a power dz on the physical layer. At the decoder we can employ successive interference cancellation to decode the virtual users' messages, where the virtual user with a higher interference level z, $0\leq z\leq P$, is decoded first by treating all the virtual users below as interference. Thus, different virtual users incur different interference levels z at the receiver. In particular, the virtual user at interference level z transmits at rate $r(z)dz$, where $r(z)$ denotes the marginal rate. Note that at channel fading state h, a virtual user with power dz and Gaussian interference $1+hz$ can achieve the rate

$$\frac{1}{2}\log\left(1+\frac{hdz}{1+hz}\right)\overset{dz\to0}{\longrightarrow}\frac{1}{2\ln2}\frac{h}{1+hz}dz. \tag{7}$$

Therefore, the message sent by the virtual user at interference level z can be decoded if $\frac{1}{2\ln2}\frac{h}{1+hz}\geq r(z)$, and all the virtual users with higher interference levels are decoded successfully, if $\frac{1}{2\ln2}\frac{h}{1+hz'}\geq r(z')$, $\forall z'\geq z$, which is denoted as Condition A in the following. If $\frac{1}{2\ln2}\frac{h}{1+hz}=r(z)$ has a unique solution for z between 0 and P and $r(z)$ is a strictly decreasing function, Condition A is satisfied. Then, the probability that the virtual user at interference level z can be decoded is given as

$$\mathcal{P}_d(r(z),z)=Pr\left(\frac{1}{2\ln2}\frac{h}{1+hz}\geq r(z)\right). \tag{8}$$

Note that in this case the network layer rate is obtained by integrating (8) over all z, $0\leq z\leq P$. Thus, the achievable average overall throughput for Rayleigh fading is

$$F(r(\cdot))=\int_0^P\mathcal{P}_d(r(z),z)r(z)dz=\int_0^P e^{-\frac{r(z)}{(1-r(z)z)\sigma^2}}r(z)dz. \tag{9}$$

We need to maximize (9) over all possible marginal rate functions $r(\cdot)$ that satisfy Condition A. Clearly, the marginal rate function $r(z)=\frac{1}{2\ln2\left(\frac{P}{2^{2R}-1}+z\right)}$ satisfies Condition A, in which case (9) becomes the single fixed rate problem (4). This means that the superposition approach contains the fixed rate scheme as a special case.

By maximizing (9) for each z individually, we obtain the optimal marginal rate function as

$$r^*(z)=\frac{1}{(2\ln2)\sigma^2z^2}\left(1+2\sigma^2z-\sqrt{1+4\sigma^2z}\right). \tag{10}$$

We can see that $\frac{1}{2\ln2}\frac{h}{1+hz}=r^*(z)$ has a unique solution $z=\frac{\sigma^2-h}{h^2}$ and that $r^*(z)$ satisfies Condition A. Therefore, $r^*(z)$ is

the optimal solution. By integrating $r^*(z)$ over all z, $0 \leq z \leq P$, we get the optimal physical layer rate.

We now propose a modification of the above scheme for a finite number of virtual users. For example, for two virtual users we choose the marginal rate function as

$$r(z) = \begin{cases} \frac{1}{2\ln 2 \left(\frac{P}{2^{2R_1}-1} + z \right)}, & \text{if } 0 \leq z \leq \alpha P, \\ \frac{1}{2\ln 2 \left(\frac{P}{2^{2R_2}-1} + z \right)}, & \text{if } \alpha P \leq z \leq P. \end{cases} \quad (11)$$

To satisfy Condition A, we must have $R_2 \leq R_1$. Substituting (11) into (9), from (10) we obtain the physical layer rate

$$F(R_1, R_2, \alpha) = \frac{1}{2}\log\left(1 + \alpha(2^{2R_1}-1)\right)e^{-\frac{2^{2R_1}-1}{P\sigma}} + \frac{1}{2}\log\left(1 + \frac{(1-\alpha)}{\frac{1}{2^{2R_2}-1} + \alpha}\right)e^{-\frac{2^{2R_2}-1}{P\sigma}}. \quad (12)$$

Fixing α, we can find the optimal $R_1(\alpha)$ and $R_2(\alpha)$ by maximizing the two summands in (12) separately. Substituting $R_1(\alpha)$ and $R_2(\alpha)$ back into (12), we can find the optimal α by a line search. We show experimentally in Section V that the performance for two virtual users is not far from the infinite user bound. The approach extends to the case with more virtual users in a straightforward way.

IV. ACHIEVABLE RATE IN WIRELESS NETWORKS

In this section, we consider general wireless networks without interference. As above, we assume that CSI is only available at the receiver.

A. Achievable Rate Region and Upper Bound

For the sake of brevity we only consider an extension of the fixed rate case in Section III; results similar to those presented below can also be obtained for the virtual user-based scheme. Each node i simply broadcasts to its neighbors at a common rate R_i due to the lack of CSI at the transmitter. The probability that node $j \in \mathcal{N}_\mathcal{O}(i)$ can receive the packets sent by node i is given as

$$\mathcal{P}_{i,j}(R_i) = \int_{\frac{2^{2R_i}-1}{P_i}}^{+\infty} f(h_{i,j})dh_{i,j} = e^{-\frac{2^{2R_i}-1}{P_i\sigma_{i,j}^2}}, \quad (13)$$

where we have used the Rayleigh pdf from (2). Due to the ergodicity of the channel, we can associate each network layer packet transmitted by node i with an erasure probability $\epsilon_{i,j}(R_i) = 1 - \mathcal{P}_{i,j}(R_i)$ at node j. Therefore, at the network layer, we see an erasure network, which is similar to that in [1]. It can be readily verified that the results in [1] can be extended to our case with only a slight modification in the proof, which leads to the following Theorem.

Theorem 1. *The capacity of the wireless network with fixed rate transmission and Rayleigh fading on each link is given by the minimum value of the cuts between the source node and any of the destinations, i.e.,*

$$C = \max_{\{f_i(\cdot)|i\in\mathcal{V}\}} \min_{\substack{\{\mathcal{V}_s|\mathcal{V}_s\subset\mathcal{V}, s\in\mathcal{V}_s \\ \text{and } \mathcal{D}\cap\mathcal{V}_s^c\neq\emptyset\}}} \sum_{i\in\mathcal{V}_s^*} \int_0^{+\infty} f_i(R_i)R_i$$
$$\times \left(1 - \prod_{\{j|j\in\mathcal{V}^c\cap\mathcal{N}_\mathcal{O}(i)\}} \epsilon_{i,j}(R_i)\right)dR_i$$

subject to $\int_0^{+\infty} f_i(R_i)dR_i = 1,$ (14)

In Theorem 1, the function $f_i(R_i)$ plays the role of time sharing between different rates. Since $R_i\left(1 - \prod_{\{j|j\in\mathcal{V}^c\cap\mathcal{N}_\mathcal{O}(i)\}} \epsilon_{i,j}(R_i)\right)$ is generally not a convex function in R_i, time sharing may achieve some rate that is not achievable without using time sharing. Note that a similar result as in Theorem 1 can also be obtained for the virtual-user based scheme from Section III.

Since Theorem 1 only addresses the fixed rate transmission scheme for the receiver-only CSI case, it is instructive to also consider the upper bound on the achievable rate for any transmission scheme. Define $\underline{Y}_j = \{Y_{i,j}|i\in\mathcal{N}_\mathcal{I}(j)\}$. By using the cut-set bound from [8, Theorem 14.10.1] we obtain

$$R_s \leq I\left(X^{\mathcal{V}_s}; \underline{Y}^{\mathcal{V}_s^c}, H^{[\mathcal{V}_s,\mathcal{V}_s^c]}|X^{\mathcal{V}_s^c}\right), \forall \mathcal{V}_s\subset\mathcal{V}, s\in\mathcal{V}_s \text{ and } \mathcal{D}\cap\mathcal{V}_s^c\neq\emptyset, \quad (15)$$

where $X^{\mathcal{V}_s} = \{X_i|i\in\mathcal{V}_s\}$, $Y^{\mathcal{V}_s} = \{\underline{Y}_j|j\in\mathcal{V}_s^c\}$, $H^{[\mathcal{V}_s,\mathcal{V}_s^c]} = \{H_{i,j}|i\in\mathcal{V}_s, j\in\mathcal{V}_s^c, (i,j)\in\mathcal{E}\}$, and $H_{i,j}$ denotes the corresponding random variable for $h_{i,j}$. As CSI is known at each receiver, we can interpret the CSI as an additional received quantity at the receiver. Since we assume that different nodes have orthogonal channels and that the network is acyclic, an upper bound on the network capacity of receiver CSI only can be obtained from (15) as

$$C \leq \min_{\substack{\{\mathcal{V}_s|\mathcal{V}_s\subset\mathcal{V}, s\in\mathcal{V}_s \\ \text{and } \mathcal{D}\cap\mathcal{V}_s^c\neq\emptyset\}}} \sum_{i\in\mathcal{V}_s^*} E_h\left\{\frac{1}{2}\log\left(1 + P_i\sum_{j\in\mathcal{V}_s^c\cap\mathcal{N}_\mathcal{O}(i)} h_{i,j}\right)\right\} \quad (16)$$

Note that the upper bound is difficult to achieve in a general network because the receivers in the cut-set are assumed to have full cooperation to obtain the cut-set rate. In particular, this holds for an achievable scheme. However, in Section IV-C we will show that the achievable rate is not far away from this upper bound in high SNR.

B. Opportunistic Routing and Distributed Algorithm

To solve (14) directly, we need to consider all the cuts which may be exponential in the number of nodes in the network. As an alternative, we consider a flow-based formulation where the constraints are only polynomial in $|\mathcal{E}|$ and $|\mathcal{V}|$. At each node i in the network let the rate on each outgoing link be denoted as $x_{i,j}^d$ for each destination node $d\in\mathcal{D}$. Then, in order to find the maximum multicast rate, we need to solve

$$\max_{C,\mathbf{x},f} C$$
$$\text{s.t.} \sum_{\{j|j\in\mathcal{N}_\mathcal{O}(i)\}} x_{i,j}^d - \sum_{\{j|j\in\mathcal{N}_\mathcal{I}(i)\}} x_{j,i}^d = \begin{cases} C, & \text{if } i=s, \\ -C, & \text{if } i=d, \\ 0, & \text{otherwise}, \end{cases}$$
$$\sum_{j\in\mathcal{Z}} x_{i,j}^d \leq \int_0^{+\infty} f_i(R_i)R_i\left(1 - \prod_{j\in\mathcal{Z}} \epsilon_{i,j}(R_i)\right)dR_i, \forall\mathcal{Z}\subseteq\mathcal{N}_\mathcal{O}(i),$$
$$\int_0^{+\infty} f_i(R_i)dR_i = 1, \quad (17)$$

where $x_{i,j}^d$ is the flow rate on link (i,j) for destination $d\in\mathcal{D}$ and the vector $\mathbf{x}$ contains the flow rates for all d and for all edges in the network. The right hand side of the last constraint in (17) represents the total amount of information that can be decoded by at least one node in $\mathcal{Z}$. It can be readily verified that (14) and (17) have the same optimal value.

Note that (17) is similar to the problem in [9], where a dual decomposition based cross layer optimization is proposed.

Our problem differs from that in [9] in the realization of the physical layer rate, which also results in a different routing scheme compared to [9] as shown in the following. By adopting a similar dual decomposition approach as in [9], we can write the Lagrange dual of (17) as

$$D(\mathbf{q}) = \max_{\mathbf{x},f,C} \left(1 - \sum_{d\in\mathcal{D}} q_s^d\right)C + \sum_{(i,j)\in\mathcal{E}}\sum_{d\in\mathcal{D}}\left(q_i^d - q_j^d\right)x_{i,j}^d$$

$$\text{s.t. } \sum_{j\in\mathcal{Z}} x_{i,j}^d \le \int_0^{+\infty} f_i(R_i)R_i\left(1 - \prod_{j\in\mathcal{Z}}\epsilon_{i,j}(R_i)\right)dR_i, \forall \mathcal{Z}\subseteq\mathcal{N}_\mathcal{O}(i),$$

$$\int_0^{+\infty} f_i(R_i)dR_i = 1, \tag{18}$$

where χ_i^d is defined as the right hand side of the first constraint in (17), and q_i^d denotes the Lagrangian multiplier at node i for destination d. The collection of all q_i^d is given by the vector $\mathbf{q}$. Due to the convexity of (17) there is no duality gap.

To find $D(\mathbf{q})$ in (18), we find that the maximization problem separates in two individual maximizations over $\mathbf{x}$ and f, resp., and C for given $\mathbf{q}$. In order to update C we use a primal subgradient method. Let C^t be the value of C at the t-th iteration. We update C using

$$C^{t+1} = \left[C^t + \gamma^t\left(1 - \sum_{d\in\mathcal{D}} q_s^d\right)\right]^+, \tag{19}$$

where $[\cdot]^+$ denotes the mapping to non-negative numbers and $\gamma^t > 0$ is a stepsize. For the second maximization over $\mathbf{x}$, let π^d be a permutation of $j\in\mathcal{N}_\mathcal{O}(i)$ such that $w_{i,j}^d = \left[q_i^d - q_j^d\right]^+$ satisfies $w_{i,\pi_1^d} \ge w_{i,\pi_2^d} \ge \cdots \ge w_{i,\pi_{|\mathcal{N}_\mathcal{O}(i)|}^d}$. Due to the polymatroid structure of the constraint in (18), (18) is maximized when

$$x_{i,\pi_1^d}^d = \int_0^{+\infty} f_i(R_i)R_i\left(1 - \epsilon_{i,\pi_1^d}(R_i)\right)dR_i,$$

$$x_{i,\pi_k^d}^d = \int_0^{+\infty} f_i(R_i)R_i\prod_{j=1}^{k-1}\epsilon_{i,\pi_j^d}(R_i)\left(1 - \epsilon_{i,\pi_k^d}(R_i)\right)dR_i, \tag{20}$$

$$k = 2,\ldots,|\mathcal{N}_\mathcal{O}(i)|.$$

We can see that the solution (20) corresponds to prioritizing the neighbors of node i for destination d according to $w_{i,j}^d$. Thus, if a higher priority neighbor has received the message from node i for destination d, all the lower priority neighbors will drop the same message even though they also have received it. We can interpret q_i^d as the queue length of the flow for destination d at node i and $w_{i,j}^d$ as the aggregate queue length difference between i and j for destination d, respectively. Different from opportunistic routing [10] where neighbors are prioritized heuristically, we optimize the prioritization by taking into account the queue information at each node. Note that each node's queue length is affected by both the incoming traffic rate and the depletion rate. The depletion rate implicitly counts the optimization over the network beyond node i. Thus, this prioritization approach can be seen as special variant of a back-pressure-based routing scheme [11].

Substituting (20) into (18), we obtain an optimization problem which depends only on $f_i(\cdot)$. The solution is $f_i(R_i) = \delta(R_i - R_i^*)$, where $\delta(\cdot)$ is the Dirac delta function, leading to

$$R_i^* = \arg\max_{R_i} R_i \sum_{j=1}^{|\mathcal{N}_\mathcal{O}(i)|}\sum_{d\in\mathcal{D}}\left(w_{i,\pi_j^d}^d - w_{i,\pi_{j+1}^d}^d\right)\left(1 - \prod_{k=1}^j \epsilon_{i,\pi_k^d}(R_i)\right). \tag{21}$$

As (21) is not convex, we can solve it using a line search. Note that at different iterations t the chosen rate R_i^* in (21) can be different due to oscillations of $\mathbf{q}$. Time sharing is then realized by averaging R_i^* over all iterations.

After solving (18), we obtain the dual function $D(\mathbf{q})$, which needs to be minimized. Instead of minimizing $D(\mathbf{q})$ directly, we use the dual subgradient method to update $\mathbf{q}$. Let $q_i^{d,t}$ be the dual variable at the t-th iteration. We update $q_i^{d,t}$ via

$$q_i^{d,t+1} = \begin{cases} \left[q_i^{d,t} - \eta^t\left(\sum_{\{j|j\in\mathcal{N}_\mathcal{O}(i)\}} x_{i,j}^d - \sum_{\{j|j\in\mathcal{N}_\mathcal{I}(i)\}} x_{j,i}^d - C^t\right)\right]^+, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } i=s, \\ 0, \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{if } i=d, \\ \left[q_i^{d,t} - \eta^t\left(\sum_{\{j|j\in\mathcal{N}_\mathcal{O}(i)\}} x_{i,j}^d - \sum_{\{j|j\in\mathcal{N}_\mathcal{I}(i)\}} x_{j,i}^d\right)\right]^+, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise,} \end{cases} \tag{22}$$

where $\eta^t > 0$ is a constant stepsize, and C^t and $x_{i,j}^d$ are obtained from (19) and (20), respectively. The updated $q_i^{d,t+1}$ is then used in (18) again for the next iteration. By the subgradient optimization theory and the convexity of the problem, the proposed primal-dual subgradient algorithm in (19)-(21) converges to the optimal solution in (18) when stepsizes γ^t and η^t go to zero as $t\to\infty$ and converges to a small neighborhood of the optimal solution with fixed stepsizes [9].

An important benefit of the presented algorithm is that (19), (20), and (21) can all be solved locally at each node i given the queue length information of its neighbors, which only involves local information exchange. This algorithm can not only be used to determine the maximal multicast rate and the corresponding physical layer rate on each link but also suggests the optimal routing protocol to achieve this rate. We can see that a layered structure can be used, where channel coding is applied at the physical layer and opportunistic routing and erasure correction coding are applied at the networking layer.

C. Asymptotic Behavior

In [12] we prove the following theorem which quantifies the capacity gap between the achievable scheme proposed in Section IV-B and the upper bound on the achievable rate from Section IV-A.

Theorem 2. *Let* $C^\dagger(\mathbf{P}) = C^\dagger([P_1, P_2, \ldots, P_{|\mathcal{V}|}])$ *be the capacity region of the fixed size wireless network without interference and* $C(\mathbf{P})$ *be the achievable rate with fixed rate transmission at each node. Further, let* $\varsigma = \max_i\max_{j\ne j'}\sigma_{i,j}^2/\sigma_{i,j'}^2$. *Then,*

$$\lim_{\mathbf{P}\to\infty} C^\dagger(\mathbf{P}) - C(\mathbf{P}) \le \frac{1}{2}|\mathcal{D}||\mathcal{V}|\log(|\mathcal{V}|\varsigma) + 0.7588|\mathcal{D}||\mathcal{V}|$$

$$+ o\left(|\mathcal{D}|\sum_{i\in\mathcal{V}}\ln\ln(P_i)\right), \tag{23}$$

where the limit is taken for each entry of $\mathbf{P}$ *individually. Moreover,* $\lim_{\mathbf{P}\to\infty} C^\dagger(\mathbf{P})/C(\mathbf{P}) = 1$.

Theorem 2 indicates that the fixed rate transmission scheme can achieve almost all of the ergodic network capacity asymptotically in high SNR. Hence, in high SNR the proposed layered structure is close to optimal which justifies the layered structure in existing wireless network protocols. However, different from existing protocols, in the proposed scheme the queue length information determines not only the opportunistic routing but also the physical layer data transmission rate at each node.
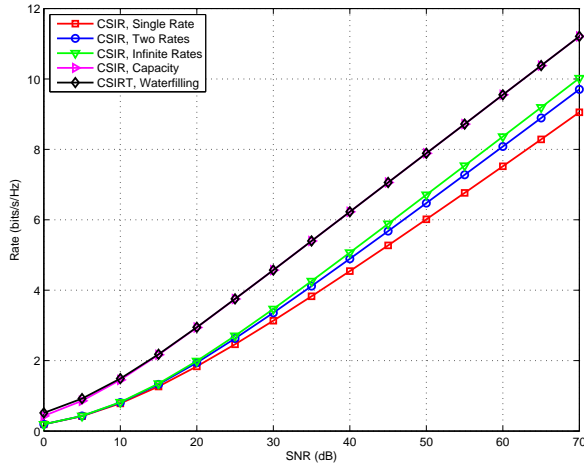
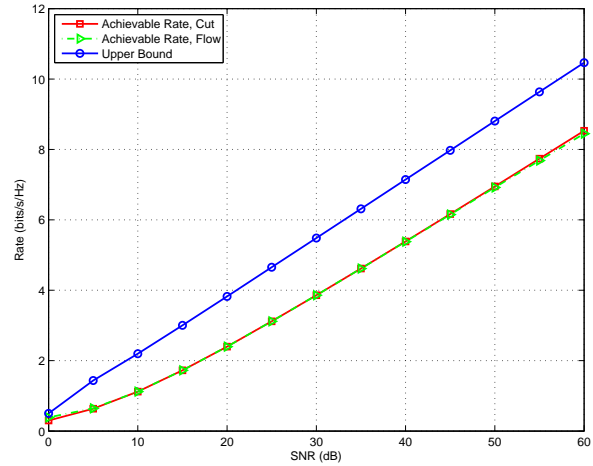Fig. 1. Comparison of achievable rates for different schemes in point-to-point Rayleigh fading channels.

Fig. 2. Comparison of rates for a diamond network by maximizing (14) directly and by using the proposed distributed algorithm and the upper bound (16).

V. SIMULATION RESULTS

We assume in the following that all the nodes have the same transmit power P and that the channel gains have the same unit variance.

Fig. 1 compares the achievable rate of different schemes for a point-to-point Rayleigh fading channel. The fixed (single) rate transmission without transmitter CSI (5) is denoted as "CSIR, One Rate", the scheme with two virtual users and superposition coding in (11) is denoted as "CSIR, Two Rates", 'CSIR, Infinite Rates" represents the approach for infinitely many users and the achievable rate (10), and "CSIR, Capacity" denotes the capacity with receiver-only CSI in (3). Also, the capacity with both transmitter and receiver CSI obtained by waterfilling [6] is shown as "CSIRT, Waterfilling". We can see that the CSIR and CSIRT capacities are the same in high SNR, indicating that transmitter CSI is not necessary in this case. We also observe that the rate loss of single rate fixed transmission to the capacity increases as SNR increases. For two virtual users, the achievable rate improves at high SNR, however, still with an increasing rate loss to capacity. In contrast, the infinite virtual user case exhibits a constant rate loss to capacity.

We now consider a diamond network with a single source node s, a single destination node d, and two relay nodes r_1 and r_2. There are no direct links between source and destination node and between the relay nodes, resp., and source and relays transmit in orthogonal channels. Fig. 2 displays the achievable rate and the upper bound (16) for different P. As the network has only four cuts, we can compute the achievable rate for fixed rate transmission by using the cut-set formulation (14) directly. We also include the achievable rate using the flow-based formulation (17) optimized via the distributed algorithm in Section IV-B. We can see that the flow-based formulation achieves the same rate as the cut-set formulation. The rate gap between the achievable rate and the upper bound increases as P increases but diminishes relatively to the total rate.

VI. CONCLUSION

We have analyzed the achievable average rate in interference-free networks along with the optimal rate allocation on network and physical layers. For the point-to-point fading channel, we presented a superposition coding approach based on multiple virtual users to improve the throughput of fixed rate transmission when only receiver CSI is available. For networks, a distributed cross-layer optimization algorithm was proposed to obtain the maximum achievable rate using a fixed rate transmission scheme on the physical layer. It was shown that the proposed layered approach asymptotically achieves almost all of the ergodic network capacity in high SNR.

REFERENCES

[1] F. A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 789–804, Mar. 2006.
[2] N. Ratnakar and G. Kramer, "Multicast capacity of deterministic relay networks wth no interference," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2424–2432, June 2006.
[3] M. Vehkaperä and M. Médard, "A throuput delay trade-off in paketized systems with erasures," in *Proc. IEEE Int. Symposium on Inform. Theory*, Adelaide, Australia, Sept. 2005, pp. 1858–1862.
[4] M. Luby, T. Gasiba, T. Stockhammer, and M. Watson, "Reliable multimedia download delivery in cellular broadcast networks," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 235–246, Mar. 2007.
[5] S. Lin and D. J. Costello, *Error Control Coding*. Prentice Hall, 2004.
[6] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1986–992, Nov 1997.
[7] S. Shamai, "A broadcast strategy for the Gaussian slowly fading channel," in *Proc. IEEE Int. Symposium on Inform. Theory*, Ulm, Germany, June 1997, p. 150.
[8] T. Cover and J. Thomas, *Elements of Information Theory*, 1991.
[9] L. Chen, S. H. Low, M. Chiang, and J. C. Doyle, "Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks," in *Proc. of IEEE Infocom*, Apr. 2006.
[10] S. Biswas and R. Morris, "ExOR: Opportunistic routing in multi-hop wireless networks," in *Proc. ACM SIGCOMM*, Aug. 2005, pp. 133–144.
[11] T. Ho and H. Viswanathan, "Dynamic algorithms for multicast with intra-session network coding," in *Proc. of Allerton Conf. on Comm., Contr. and Comput.*, Sept. 2005.
[12] T. Cui, T. Ho, and J. Kliewer, "Achievable rates in interference-free fading networks," in preparation, 2009.