MIMO-pipe Modeling and Scheduling for Efficient Interference Management in Multi-Hop MIMO Networks

Weiyang Ge, Junshan Zhang, and Guoliang Xue

Abstract—The multiple-input multiple-output (MIMO) technology, a recent breakthrough in wireless communications, has been shown to improve the channel capacity significantly in single-user systems. However, it remains largely open to obtain a rigorous understanding of the possible MIMO gains in multi-hop networks. One grand challenge is that multi-hop wireless networks are interference-limited and that the interference introduces coupling across various layers of the protocol stack, including the PHY (physical), MAC, network and transport layers. The fundamental differences between multi-hop networks and point-to-point settings dictate that leveraging the MIMO gains in multi-hop networks requires a domain change from high SNR regimes to interference-limited regimes.

In this paper, we develop a cross-layer optimization framework for effective interference management towards understanding fundamental tradeoffs among possible MIMO gains in multi-hop networks. We first take a bottom-up approach to develop a MIMO-pipe model based on PHY-interference, and extract a set of \( \{ (R_i, \text{SINR}_i) \} \), where each pair \((R_i, \text{SINR}_i)\) corresponds to a meaningful stream multiplexing configuration for individual MIMO links (with \(R_i\) being the rate and \(\text{SINR}_i\) the SINR requirement). Using this link abstraction model, we study MIMO-pipe scheduling for throughput maximization. Based on continuous relaxation via randomization, we study the structural property of the optimal scheduling policy. Our finding reveals that in an optimal strategy it suffices for each MIMO link to use one stream configuration only (although each individual MIMO link can have multiple stream configurations). In light of this structural property, we then formulate MIMO-pipe scheduling as a combinatorial optimization problem; and by using a multidimensional 0-1 knapsack approach, we devise centralized approximation algorithms, for both the dense network model and the extended network model, respectively. Next, we also develop a contention-based distributed algorithm, in which links update their contention probability based on local information only, and characterize the convergence and the performance of the distributed algorithm.

I. INTRODUCTION

A. Motivation

The past few years have witnessed a rapidly growing demand for ubiquitous communications and computing, and this is changing dramatically the ways we store, manage, search, and access information. Indeed, the wireless technology revolution is taking place at many frontiers: multiple antenna techniques, adaptive modulation/coding schemes, medium access control (MAC) design, routing protocols, security mechanisms, and clean-slate architecture design. Notably, there has been tremendous research effort on the multiple-input multiple-output (MIMO) technology. By using multiple antennas at both transmitters and receivers, this technique offers additional spatial degrees of freedom for information transmission. For single-user wireless channels, it has been shown that using the MIMO technique can lead to dramatic improvement, in terms of capacity and link reliability [7], [22].

The possible leverages offered by the MIMO technology for point-to-point transmissions can be summarized as follows: 1) spatial multiplexing gain increases spectral efficiency by opening up multiple spatial data pipes in the same frequency channel; 2) diversity gain improves link reliability by providing multiple faded signal paths between the transmitter and the receiver, and hence can result in a larger coverage by enhancing the received signal-to-noise ratio (SNR) through coherent combining of the signals arriving at the receive antenna array; and 3) interference suppression gain reduces co-channel interference by using the spatial degrees of freedom to mitigate undesired interfering signals. Recent studies have explored the fundamental tradeoffs between different gains in single-user/cellular MIMO systems [2], [24].

In stark contrast to the simpler point-to-point setting, there has been little work on exploring the MIMO technique in multi-hop wireless networks. Clearly, pure information-theoretic models for multi-hop MIMO networks are too complicated to implement. On the other hand, many of the existing approaches for protocol design in multi-hop MIMO networks use overly simplistic abstraction by treating the number of antennas as degrees of freedom to build graph models. However, the corresponding reliability of spatial channels embedded in the MIMO link is intrinsically coupled with the abstraction of degrees of freedom. Indeed, one grand challenge is that multi-hop wireless networks are interference-limited and that the interference introduces coupling across various layers of the protocol stack, including the PHY-layer, the MAC layer, the network layer and the transport layer. Clearly, leveraging MIMO gains is intimately related to effective interference management, and a key first step is to develop realistic PHY/MAC model abstraction of MIMO links in interference-limited networks.

As outlined in Table I, there are at least three major differences between single-user/cellular systems and multi-hop networks, which are crucial in leveraging MIMO gains in multi-hop networks. We elaborate further on this as follows.

Copyright (c) 2010 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.
TABLE 1

| Major Differences Between Single-User/Cellular Systems and Multi-Hop Networks |
|-----------------------------|-----------------------------|--------------------------------|
|                           | Multiplexing Gain/Feasible Rates | Diversity Gain/Reliability | Channel Matrix/Interference Structure |
| Single-User Systems        | High SNR Regime               | High SNR Regime              | Knowledge of Receiver CSI               |
| Cellular Systems           | High SNR Regime               | Interference Cancellation    | Knowledge of Interference               |
| Multi-Hop Networks         | Bounded SINR Regime           | Interference Tolerance       | No Knowledge of Interference            |

1) High SNR vs. bounded SINR. In a single-user setting, spatial multiplexing gain and diversity gain are defined in the high SNR regime (SNR → ∞) [24]. However, since multi-hop wireless networks are interference-limited, spatial multiplex gain should be revisited in terms of signal-to-interference-plus-noise ratio which is always bounded. Moreover, the notion of diversity gain defined in the high SNR regime is not applicable to multi-hop networks. In light of this, we will develop the MIMO-pipe model consisting of a set of feasible rates and reliability requirements.

2) Receiver channel state information (CSI) vs. no knowledge of interference structure. In a cellular network setting, receive antennas can be used to suppress interference, assuming that the interference structure is available at the receiver (base station). This assumption is reasonable since the base station is the intended receiver for all users in the cell. In a multi-hop MIMO network, however, different links have different receivers and the interference structure is often not attainable since it is very difficult to track the channel matrices corresponding to interfering links from interferers to different receivers.

3) Diversity gain vs. reliability enhancement/spatial reuse. Given that diversity gain is not directly applicable to multi-hop settings, we use the notion of reliability gain (interference tolerance enhancement) instead. We observe that MIMO diversity gain can enable the data to reach further at the same reliability level, and the hop length is intimately related to interference modeling, spatial reuse, and end-to-end routing delay. Intuitively speaking, the larger the hop length is, the smaller the spatial reuse factor would be.

Given that the MIMO technology has great potential in significantly enhancing the throughput and delay performance, there is an urgent need for developing useful models and analysis methodologies tailored towards optimal design for multi-hop MIMO networks. The fundamental differences between multi-hop networks and single-user/cellular settings dictate that advancing the MIMO technique in multi-hop networks requires a domain change from the high SNR regime to interference management. With this insight, a principal goal of this paper is to explore MIMO-pipe modeling and obtain a rigorous understanding of MIMO-pipe scheduling for throughput maximization.

We focus on single-channel multi-hop MIMO networks, in which the MIMO gains hinge heavily on effective interference management. Needless to say, accurate modeling of interference is of critical importance. In the existing literature, two main interference models have been used, namely the protocol model and the PHY-interference model [4]. In the protocol model, a transmission is successful if no other user within a certain transmission range from the receiver is transmitting. Due to its simplicity, the protocol model has been widely used (see, e.g., [11], [12], [25]). However, this model does not take the aggregated interference into consideration, which may significantly degrade the efficiency and reliability of MIMO transmissions. In light of this, we will use the PHY-interference constraint in developing the MIMO-pipe model, where a transmission is successful if the SINR at the receiver is above a certain threshold. The PHY-interference constraint requires a systematic study on MIMO-pipe scheduling to mitigate potential conflict and interference among links, and this is a main subject of this paper.

B. Summary of Main Results

In this paper, with the objective of maximizing the network throughput\(^1\), we first take a bottom-up approach to develop solid PHY/MAC interference models for MIMO-pipes. Specifically, given the number of antennas and the power constraint at each node, the more data streams there are at each MIMO link, the lower the reliability per stream would be and so is the interference tolerance capability. As noted before, it is more realistic to assume that the time-varying interference is unknown. Appealing to the fundamental theory of reliable communications [6], we first extract a new rate/reliability model, which consists of a set of feasible rates and SINR requirements, denoted by \( \{ (R_i, \text{SINR}_i) \} \), corresponding to meaningful stream multiplexing configurations. We note that the set of \( \{ (R_i, \text{SINR}_i) \} \) are a few points in the rate/reliability curve, and that for a given BER requirement, the rate/reliability curve can be obtained offline. Indeed, the (rate, SINR) model emerges naturally under PHY-interference and has been used for PHY-layer studies in [20], [23]. Clearly, this link abstraction model pass, to upper protocol layers, a set of feasible rates and SINR requirements, and this is in stark contrast to the traditional wisdom for wireless network design which abstracts the PHY-layer as a single number for upper layer optimization.

It is in this sense that we name the rate/reliability model “MIMO-pipe model”. In this paper, we investigate thoroughly the MIMO-pipe scheduling for throughput maximization, for both dense network models and extended network models.

We devise both centralized algorithms and distributed algorithms for MIMO-pipe scheduling. MIMO-pipe scheduling is essentially joint link scheduling and MIMO stream multiplexing. For the centralized case, we explore cross-layer...
optimization that can strike a balance between activating multiple interfering links and simultaneously achieving stream multiplexing gain for each link. Based on continuous relaxation via randomization, we investigate the structural property of the optimal scheduling policy. Our finding reveals that, although each individual MIMO link can have multiple stream configurations, in an optimal strategy it suffices for each link to choose one stream configuration only. Building on this structural property, we formulate the throughput maximization problem as a combinatorial optimization problem. We then devise polynomial time approximation algorithms by using multidimensional knapsack approach under two different network models, namely the extended network model and the dense network model. For the dense network model, where all links lie in a fixed area, we devise a polynomial-time-approximation-scheme (PTAS). For the extended network model, where the network scales with fixed node density, we develop a heuristic algorithm.

 Needless to say, it is more involved to devise distributed MIMO-pipe scheduling. To this end, we propose a suboptimal two-phase algorithm, where in the first phase each link contends for the channel with adaptive contention probability, and in the second phase each link determines whether to do stream multiplexing based on its local SINR requirement. Intuitively speaking, the contention phase is used to sort out, in a distributed manner, a “good” feasible set of links that can transmit simultaneously. Numerical results indicate that if the loss due to the signaling overhead is excluded, the throughput gap between the distributed algorithm and the optimal scheduling is not significant under some conditions.

C. Related Work

It has been shown that using multiple antennas at the wireless transmitter and the receiver, namely the MIMO technique, can boost up the channel capacity significantly [7], [22]. For point-to-point communications, it is shown that the spatial degrees of freedom (DoF) of a MIMO channel can be utilized to achieve spatial multiplexing gain, spatial diversity gain and interference suppression, and the tradeoffs between these gains have been studied in [2], [24]. Roughly speaking, assuming that the spatial channels across different antennas are more or less independent, a MIMO link with N antennas at the transmitter and N antennas at the receiver can provide N times the throughput of a single-antenna system in the same frequency channel; and this is so called “spatial multiplexing”. Alternatively, the spatial degrees of freedom can be exploited to obtain diversity gain (including both transmit diversity and receive diversity), by combining “coded replica” of the signal replica to combat fading ([21]); the reliability of wireless links can be improved significantly. The spatial degrees of freedom can also be used to null out undesired interfering signals to achieve interference suppression gain. The study in [12] investigated link scheduling in multi-hop MIMO networks under the protocol model. Recent work [15] studied joint optimal routing, power allocation and bandwidth allocation in a FDMA MIMO-based ad hoc network, assuming no interference across links. Another interesting paper [14] leveraged non-convex programming to investigate joint sub-carriers scheduling and power allocation in multi-carrier MIMO networks.

The rest of the paper is organized as follows. In Section II, we present the system model and develop stream configuration models with rate and SINR requirements for MIMO links. In Section III and Section IV, we study centralized and distributed schemes for joint stream multiplexing and scheduling, respectively. Numerical examples are provided in Section V to illustrate the theoretic findings. Finally, Section VI concludes the paper.

II. Problem Formulation

A. System Model

We consider a multi-hop MIMO network with M links (see Fig. 1), where all links transmit over a common channel, as shown in Fig 1. Assume a general model where each link employs $N_t$ transmit antennas and $N_r$ receive antennas. The received signal at the $i$th receiver denoted $y_i$, is given by

$$y_i = \sqrt{\frac{P}{N_t d_{ii}}} H_{ii} x_i + \sum_{j \neq i} \sqrt{\frac{P}{N_t d_{ji}}} H_{ji} x_j + n_i,$$  \hspace{1cm} (1)$$

where $P$ is the total transmission power at each transmitter, and is assumed to be fixed through the paper; $x_i$ is the $N_t \times 1$ transmitted signal from the $i$th transmitter, with normalized transmission power at each antenna array to be 1, in each symbol period; $\alpha$ is the path loss exponent; $d_{ji}$ is the distance from the $j$th transmitter to the $i$th receiver; $H_{ji}$ is the $N_r \times N_t$ channel matrix between the $j$th transmitter to the $i$th receiver, where the entries of each matrix are i.i.d. complex circular symmetric Gaussian with unit variance. Furthermore, the entries of $H_{ji}$ are independent from those of $H_{ji'}$ if $i \neq i'$; $n_i$ is the additive White Gaussian noise with $\sigma^2 = E[||n_i^2||]/N_r$. 

**Fig. 1.** A sketch of co-channel interference in a multi-hop network.

The first term in (1) is the desired data signal for link $i$, while the last two terms are co-channel interference and noise, respectively. As is standard, we assume that the channel matrix $H_{ii}$ is known at the receiver but unknown at the transmitter of link $i$ (CSI at the receiver) [16]. Furthermore, in practical systems, it is difficult, if not impossible, to obtain the MIMO channel matrices \{$H_{ji}, j \neq i$\} from interferers, simply because the signals are not intended for the desired link and it
is infeasible to estimate these complex channel matrices [1]. In light of this, we impose the following assumption on the knowledge of the interference.

A1) For each individual MIMO link, the interfering signals are unknown to its receiver and its transmitter.

It is clear that given a BER requirement, the rate/reliability curve can be easily obtained offline. To characterize the trade-off between stream multiplexing and reliability requirement, we shall choose rates of \( \{ kR, k = 1, \ldots, N \} \), where \( R \) is the base rate per stream. Observe that the SINR at each data stream depends on the time-varying interference that is critical to set transmission rates at each transmitter, but the instantaneous interference information is not available at the transmitter and the receiver. To simplify transmission scheduling, we set the transmission rate the same for independent data streams at each MIMO link, and the stream reliability is guaranteed by effective interference management. As will be shown below, only symmetric stream configurations will be meaningful (we will elaborate further in the next subsection). We emphasize that this assumption can be relaxed with modified SINR requirements.

Based on [16], the average energy of interference-plus-noise per receive antenna is given by

\[
T_i = \sum_{j \neq i} \frac{P}{N_i d_{ji}^2} \frac{E[\text{Tr}(H_{ji}H_{ji}^H)]}{N_r} + \sigma^2.
\]

(2)

Assuming the channel state information is available at the receiver only (receiver-CSI), the ergodic capacity of link \( i \), denoted \( C_i \) (SINR), is given by [16]

\[
E \left[ \log_2 \det \left( I + SINR_i \frac{H_{ii}H_{ii}^H}{N_r} \right) \right] = \frac{1}{\sigma^2} \left( \sum_{j \neq i} \frac{P}{N_i d_{ji}^2} H_{ji}H_{ji}^H \right) + 1,
\]

(3)

and \( SINR_i \) is the average SINR of link \( i \) along each of the principle directions of the interference-plus-noise space [16]:

\[
\text{SINR}_i = \frac{P}{N_i d_{ii}^2} \frac{E[||x_i||^2]}{T_i} = \frac{Pd_{ii}^{\alpha}}{\sum_{j \neq i} Pd_{ji}^{\alpha}} \frac{E[\text{Tr}(H_{ji}H_{ji}^H)]}{N_r N_i} + \sigma^2.
\]

(4)

Following the assumption that the entries of \( H_{ji} \) are identically distributed, and the power of each entry is 1, we have \( E[\text{Tr}(H_{ji}H_{ji}^H)] \) is \( N_i N_r \). It then follows from (4) that

\[
\text{SINR}_i = \frac{Pd_{ii}^{\alpha}}{\sum_{j \neq i} Pd_{ji}^{\alpha}} + \sigma^2.
\]

(5)

It is clear that \{SINR\} hinges heavily on the average power of interference-plus-noise \( (T_i) \), and plays a key role in determining the reliability per multiplexing stream in interference-limited communications. Based on the above derivations, we will first address a fundamental issue: How to extract solid PHY/MAC models that are realistic yet easy-to-use for optimizing upper protocol layers in multi-hop MIMO networks?

We take a bottom-up approach to extract a set of feasible rates and SINR requirements, which the PHY-layer can pass to upper protocol layers.

B. MIMO-pipe Rate/Reliability Abstraction

Every MIMO link can offer stream multiplexing by opening up multiple spatial data pipes in the same frequency channel, and achieve spatial multiplexing gain. However, given the number of antennas and the total transmission power at each node, the more data streams there are at each MIMO link, the lower the reliability per stream would be and so is the interference tolerance capability. Accordingly, the SINR requirement is more stringent. In the following, we will elaborate further, through an example, the tradeoff between stream multiplexing gain and interference tolerance capability (equivalently, SINR requirements). Note that in this study, the transmission power is assumed to be fixed. Dynamic power control is beyond the scope of this paper, and will be addressed in the future. Based on the fundamental requirements for reliable communications, we show that some configurations for stream multiplexing are meaningful (while others are not), and for each meaningful configuration, we develop the corresponding rate and the SINR requirement. In what follows, we assume that all nodes have the same number of antennas.

We assume a MIMO link with \( N_t = N_r = 4 \), and assume for now that the transmission power is equally split among 4 transmit antennas. The spatial degree of freedom is 4. There are five possible configurations for stream multiplexing. Since there is no interference information at the transmitter and the receiver, we assume that the transmission rate is the same for independent data streams. This is because the SINR at each data stream depends on the interference, and the interference information is critical to set transmission rates.

Stream Configuration I: All spatial degrees of freedom of this MIMO link are used for stream multiplexing to obtain full multiplexing gain, and accordingly each transmit antenna would transmit one data stream independently, as shown in Fig. 2(a). The overall data rate over this MIMO link is \( 4R_s \), where \( R_s \) is the data rate for each stream. Since all spatial degree of freedom are used for multiplexing, the SINR requirement per receive antenna, denoted by \( \beta_1 \), would be high, dictating that the interference tolerance capability is low.

Stream Configuration II: Alternatively, as shown in Fig. 2(b), two antennas can transmit the same data stream to improve reliability by exploiting spatial diversity, and let the other two antennas transmit different streams. The overall rate is \( 3R_s \). Clearly, the data stream \( x_3 \) is more reliable than \( x_1 \) and \( x_2 \) since the antennas provide spatial diversity gain for \( x_3 \). However, since the total interference from other users is the same for all 3 streams over the MIMO link, the SINR requirement for this configuration is still \( \beta_3 \), in order to guarantee the success of the transmission of \( x_1 \) and \( x_2 \). That is, \( x_1 \) and \( x_2 \) are the bottleneck streams. Observe that with the same SINR requirement, Configuration I can provide higher data rate than Configuration II, indicating that Configuration II is not meaningful, and can be ignored. 2

2This is due to the assumption that the transmitter equally splits its power over all transmit antennas. When optimal power allocation is possible, e.g., allocating more power on the top two antennas, and less on the bottom two antennas, the link can achieve data rate \( 3R_s \) with an SINR requirement smaller than \( \beta_3 \). In this scenario, the number of meaningful configurations is equal to 4.
Stream Configurations III and IV: Two configurations, shown in Fig. 2(c) and Fig. 2(d), can be used to obtain 2 multiplexing streams. Along the same line above, we observe that Configuration III achieves the data rate of $2R_s$ with SINR requirement $\beta_3$. In Configuration IV, each data stream is transmitted through two transmit antennas, which is more reliable than Configuration III. Therefore, the SINR requirement for this configuration, denoted as $\beta_2$, is smaller than $\beta_3$. Configuration III in Fig. 2(c) is inferior to Configuration IV shown in Fig. 2(d), and thus can be ignored.

Stream Configuration V: By letting all antennas transmit the same data stream, the link achieves full spatial diversity with data rate $R_s$, as shown in Fig. 2(e). Clearly, the SINR requirement for this configuration, denoted as $\beta_1$, is smaller than both $\beta_2$ and $\beta_3$.

In summary, for the MIMO link with $N_t = N_r = 4$, there are three different meaningful configurations under equal power allocation: Configuration I, Configuration IV and Configuration V, with overall data rate $4R_s$, $2R_s$, $R_s$, and SINR requirement $\beta_3 > \beta_2 > \beta_1$, respectively.

**Proposition 2.1:** a) For a MIMO link with $N$ transmit antennas and $N$ receive antennas, where each antenna is dedicated to one stream only, when equal power allocation is used across transmit antennas, the number of meaningful configurations is at most $\lceil \sqrt{4N+1} \rceil - 1$.

b) When optimal power allocation is used, the number of meaningful configurations is at most $N$.

**Proof:** Let $C(N)$ denote the number of meaningful configurations for a MIMO link with $N$ transmit antennas and $N$ receive antennas. It is clear that $C(N)$ is monotonically increasing as $N$. To prove part a), we first consider a special case with $N = n(n+1)$, where $n$ is a positive integer. Define $n_i$ as $n_i = \lceil N/i \rceil$ for all $i = 1, 2, \cdots, n$. It follows that $N = i*n_i + \gamma_i$, where $\gamma_i$ is the residual with $0 \leq \gamma_i < i < n_i$. It can be shown that by letting the MIMO link transmit with overall data rate $iR_s$, each data stream would have at least $n_i$ transmit antennas. The corresponding SINR requirement is $\beta_i$. Similarly, the link can also transmit with data rate $n_i R_s$, in which case each data stream has at least $i$ transmit antennas. In summary, when $N = n(n+1)$, $C(N) = 2n$. Furthermore, when $N < n(n+1)$, $C(N) < 2n$.

Using a similar argument, we can show that when $N = n^2$, $C(N) = 2n - 1$. It then follows from the monotonicity of $C(N)$ that when $n^2 \leq N < n(n+1)$, $C(N) = 2n - 1$. Note that when $n^2 \leq N < n(n+1)$, $2n < \sqrt{4n^2 + 1} \leq \sqrt{4N + 1} < \sqrt{4(n+1)^2 + 1} = 2n + 1$, which implies that $C(N) = \lceil \sqrt{4N+1} \rceil - 1$.

Similarly, we have that when $n(n+1) \leq N < (n+1)^2$, $C(N) = 2n$, and

$$2n + 1 \leq \sqrt{4N + 1} < \sqrt{(n+1)^2 + 1} < 2n + 2.$$  \hspace{1cm} (6)

It follows that $\lceil \sqrt{4N+1} \rceil = 2n + 1$, \hspace{1cm} (7)

Combining the above two cases, we have that for any $N$, $C(N) = \lceil \sqrt{4N+1} \rceil - 1$, which concludes the proof of part a).

**Discussions:** In a nutshell, there is a pair of parameters, i.e., (rate, SINR threshold), corresponding to each meaningful stream configuration. Then, the PHY-layer can pass, to upper protocol layers, a set of feasible pairs in the form of (rate, SINR threshold), corresponding to different configurations, for efficient interference management. This is in contrast to the traditional wisdom for wireless network design, which abstracts the physical layer as a single number and inputs it to upper protocol layers. We caution that it remains open to explore the equivalent MIMO-pipe protocol model (with a set
of possible rates and interference guard zones) corresponding to the MIMO-pipe SINR model.

We emphasize that the stream configurations here correspond to a set of points on the rate-reliability tradeoff curve, as shown in Fig. 3, and that the rates in this study are set to be multiplications of the basic rate to capture the multiplexing gain. More generally, one can choose any points on the rate-reliability tradeoff curve and develop a corresponding abstraction of the MIMO link. Needless to say, each point is associated with a pair of feasible parameters (rate, SINR threshold), and the more points chosen, the higher the complexity would be.

![Rate-reliability tradeoff for a MIMO link with 4 antennas.](image)

**Fig. 3.** Rate-reliability tradeoff for a MIMO link with 4 antennas.

### III. Centralized MIMO-Pipe Scheduling for Throughput Maximization

In the above, based on the PHY-interference model, we have extracted a set of rate and SINR requirements, \{\{R_i, SINR_i\}\}. In the following, we devise centralized algorithms for throughput maximization under the MIMO-pipe SINR model (similar studies can be carried out for the protocol model). Specifically, we study cross-layer optimization that can strike a balance between activating multiple interfering links and simultaneously achieving stream multiplexing gain for each link. Note that different links may use different numbers of streams for multiplexing, making the scheduling more challenging.

We consider a multi-hop MIMO network with \(M\) links, where each link has \(N\) transmit antennas and \(N\) receive antennas. Let \(K\) denote the number of meaningful configurations with \(K = \lfloor \sqrt{4N + 1} \rfloor - 1\). In such a network, each link can choose one stream configuration to transmit. Let \(k = 1, 2, \ldots, K\) denote the index of the configurations, \(d_k\) and \(\beta_k\) denote the corresponding data rate and SINR requirement, respectively. For example, when \(N = 4\) as in Sec. II-B, \(K = 3\), and \(d_1 = R_s, d_2 = 2R_s, d_3 = 4R_s\).

**A. Structural Properties of Optimal MIMO-Pipe Scheduling**

To characterize the structural property of the optimal scheduling policy, we first consider a continuous relaxation by using a randomized strategy in which each link probabilistically picks a stream configuration. More specifically, let \(p_{i,k}\) denote the probability link \(i\) chooses the \(k\)th configuration, and \(\sum_{k=0}^{K} p_{i,k} = 1\), where \(p_{i,0}\) is the probability that link \(i\) does not transmit. Let \(S\) denote a subset of links. Then, the probability that links in \(S\) are transmitting is given by

\[
q_S(P) = \prod_{i \in S} (1 - p_{i,0}) \prod_{j \notin S} p_{j,0},
\]

where \(P = [p_{i,k}, i = 1, 2, \ldots, M, k = 0, 1, \ldots, K]\) is the transmission probability matrix. For simplicity, we use \(q_S\) as a shorthand notation for \(q_S(P)\) in the following. Let \(\Phi(P)\) denote the average network throughput, which is given by

\[
\Phi(P) = \sum_S q_S \Phi(P|S) = \sum_S q_S \sum_{i \in S} \sum_{k=1}^{a_i(S)} d_k \frac{p_{i,k}}{1 - p_{i,0}},
\]

where \(a_i(S)\) is the index of the best configuration for link \(i\) when links in \(S\) are transmitting, i.e.,

\[
\beta_{a_i(S)+1} > \text{SINR}_i(S) = \frac{Pd_{a_i}^{\alpha}}{\sum_{j \in S \setminus i} Pd_{a_j}^{\alpha} + \sigma^2} \geq \beta_{a_i(S)}.
\]

We note that \(\{a_i(S)\}\) depend on the global network topology and hence require global information.

Then, the joint scheduling and stream multiplexing problem in a multi-hop MIMO network can be formulated as finding the optimal probability matrix that maximizes the network throughput:

\[
\max_{P} \Phi(P)
\]

s.t. \(\sum_{k=0}^{K} p_{i,k} = 1, \quad i = 1, 2, \ldots, M\)

(11)

The above problem is non-convex in general. To solve this problem, rewrite the throughput as

\[
\Phi(P) = \sum_S q_S \sum_{i \in S} \sum_{k=1}^{a_i(S)} d_k \frac{p_{i,k}}{1 - p_{i,0}} = \sum_S \prod_{j \notin S} p_{j,0} \prod_{i \in S \setminus j} \sum_{d_{m}} \sum_{k=1}^{a_i(S)} p_{i,m} d_k p_{i,k}.
\]

It can be shown that \(\Phi(P)\) is a monotonically increasing function of \(P\).

Accordingly, the optimization problem in (11) is a quasi-concave program, and can be solved by using the bisection method [3]. Let \(P^* = [p_{i,k}^*, i = 1, 2, \ldots, M, k = 0, 1, \ldots, K]\) denote the optimal probability matrix that maximizes the network throughput, i.e., \(P^* = \arg \max_{P} \Phi(P)\). Note that \(P^*\) may not be unique. We have the following property regarding \(P^*\).

**Proposition 3.1:** The optimal probability matrix \(P^*\) can be an integer matrix, i.e., \(p_{i,k}^* \in \{0, 1\}, \forall i, k.\)

**Proof:** Let \(P^*\) denote an optimal integer matrix that maximizes the network throughput among all integer matrices. Then, it is sufficient to show that for any fractional matrix

\[\Phi(P) \text{ is monotonically increasing if } \Phi(P_1) \geq \Phi(P_2) \text{ for any } P_1 \succ P_2,\]

where \(\succ\) denotes componentwise greater than or equal to.
Proposition 3.1 reveals that the optimal scheduling policy can have an integer matrix, but not necessarily unique. To characterize the maximum network throughput, however, it requires global information \( \{a_i(S)\} \), which are difficult (if not impossible) to obtain. In the following, based on the structural property of the optimal solution derived above, we take a combinatorial optimization approach to study throughput maximization in a multi-hop MIMO network.

Recall from (5) that the average SINR per receive antenna of link \( i \) is given by

\[
\text{SINR}_i = \frac{P d_{ii}^{-\alpha}}{\sum_{j \neq i} P d_{jj}^{-\alpha} + \sigma^2}.
\]

For convenience, define \( \{Z_{ji}, i, j = 1, 2, \ldots, M\} \) and \( \{c_i, i = 1, 2, \ldots, M\} \) as

\[
Z_{ji} = \begin{cases} 
\left( \frac{d_{ji}}{d_{ii}} \right)^{-\alpha}, & i \neq j, \\
0, & i = j.
\end{cases}
\]

\[
c_i = \frac{\sigma^2}{P (d_{ii})^{-\alpha}}.
\]

Let \( r_{i,k} \) be 1 if link \( i \) chooses the \( k \)th configuration and 0 otherwise. Then, the SINR constraints are equivalent to

\[
\sum_{k \in S} Z_{ji} \left( \sum_{k} r_{j,k} \right) + c_i \leq \sum_{k} r_{i,k}/\beta_k.
\]

Based on the above interference tolerance constraints, the one-shot scheduling problem can be recast as an integer programming problem with the objective of maximizing the total throughput, subject to interference constraints:

\[
\text{P1} \quad \underset{\{r_{i,k}\}}{\text{max}} \quad \sum_{i=1}^M \sum_{k=1}^K d_k r_{i,k}
\]

\[
\text{s.t.} \quad \sum_j Z_{ji} \left( \sum_k r_{j,k} \right) \leq \sum_k r_{i,k}/\beta_k - c_i, \quad i = 1, 2, \ldots, M
\]

\[
r_{i,k} \in \{0, 1\}, \quad \forall i, k.
\]

Following [17], we can show that the optimization problem P1 is NP-hard. With this, we convert the above optimization problem into a multidimensional 0-1 knapsack problem. To this end, we construct a new constraint

\[
\sum_j Z_{ji} \left( \sum_k r_{j,k} \right) + e \left( \sum_k r_{i,k} - 1 \right) \leq \sum_k r_{i,k}/\beta_k - c_i,
\]

where \( e \) is some positive constant. By choosing the constant \( e \) sufficiently large, we have that

- The case with \( \sum_k r_{i,k} = 0 \): in this case, link \( i \) does not transmit, the SINR constraint always holds;
- The case with \( \sum_k r_{i,k} = 1 \): in this case, link \( i \) chooses one configuration to transmit, it is equivalent to the first constraint in (19);
- The case with \( \sum_k r_{i,k} \geq 2 \): in this case, (21) is always invalid.

As a consequence, the two inequality constraints in (20) can be replaced by the constraint in (21), and P1 is equivalent to the following optimization problem

\[
\text{P2} \quad \underset{\{r_{i,k}\}}{\text{max}} \quad \sum_{i=1}^M \sum_{k=1}^K d_k r_{i,k}
\]

\[
\text{s.t.} \quad \sum_j Z_{ji} \left( \sum_k r_{j,k} \right) + e \left( \sum_k r_{i,k} - 1 \right) \leq \sum_k r_{i,k}/\beta_k - c_i, \quad r_{i,k} \in \{0, 1\}, \quad \forall i, k.
\]

Define the \( MK \times 1 \) vector of binary variables \( x \triangleq [r_1, r_2, \ldots, r_K]^T \), where \( r_k \triangleq [r_{1,k}, \ldots, r_{M,k}] \). Define the
profit vector \( p = [d_1, d_2, \cdots, d_K] \), where \( d_k \) is a \( 1 \times M \) row vector with all elements equal to \( d_k \). Let \( Z \) be the channel gain matrix with \( Z = [Z_{ij}, i, j = 1, 2, \cdots, M] \), and \( Z \triangleq [Z^T, Z^{T_2}, \cdots, Z^{T_K}] \). Let \( e \) be a sufficiently large real number.

Define a new matrix \( W \) by \( W = [w_{ij}, i, j = 1, 2, \cdots, M] \), where

\[
w_{ij} \triangleq \begin{cases} \hat{Z}_{ij} + e - 1/\beta_k, & \text{if } j = i + (k-1)M, k = 1, \cdots, K, \\ \hat{Z}_{ij}, & \text{otherwise}. \end{cases}
\]

Finally, define \( b_i \triangleq e - c_i \), and \( b \triangleq [b_1, i, j = 1, 2, \cdots, M] \). The optimization problem \( P_2 \) is then equivalent to the following multidimensional 0-1 knapsack problem.

\[
KP(W, b, p) : \max \sum_{j=1}^{MK} p_j x_j, \quad \text{s.t.} \sum_{j=1}^{MK} w_{ij} x_j \leq b_i, \quad i = 1, 2, \cdots, M \quad (22)
\]

Without loss of generality, we assume that \( w_{ij} \leq b_i \) for all \( i \) and \( j \), because otherwise \( x_j \) would always be zero in any feasible solution and can be removed from the problem formulation.

In a dense network, all links lie in a disk of a fixed area. As the number of links increases, the network gets saturated. In other words, the total number of concurrent transmitting links is upper-bounded by a constant, denoted by \( D \), regardless of the total number of links in the network [10]. We can investigate all \( O(M^D) \) size-\( D \) subsets of the \( M \) links, and for each size-\( D \) subset of the links, the corresponding knapsack problem has \( DK \) variables (rather than \( MK \) and \( D \) capacity constraints (rather than \( M \)). To distinguish from the original knapsack problem \( KP(W, b, p) \), we denote this size-constrained knapsack problem by \( DKP(W, b, p) \), formally defined in the following.

\[
DKP(W, b, p) : \max \sum_{j=1}^{DK} p_j x_j, \quad \text{s.t.} \sum_{j=1}^{DK} w_{ij} x_j \leq b_i, \quad i = 1, 2, \cdots, D \quad (23)
\]

Since \( D \) is a constant for dense networks, the solution to \( KP(W, b, p) \) boils down to the solution of \( DKP(W, b, p) \). Note that the problem \( DKP(W, b, p) \) is still NP-hard [9, 18]. When \( W \) and \( b \) both take integer values, the problem can be solved in \( O(KD(2D)^D) \) time, using dynamic programming.

For fixed \( D \), there is a polynomial time approximation scheme (PTAS) with running time \( O((DK)(2D)^{-D}) \). Let \( p^{opt}(W, b, p) \) denote the optimal value of \( DKP(W, b, p) \). The PTAS can compute a feasible solution to \( DKP(W, b, p) \) with a profit value no smaller than \( (1 - \epsilon) \times p^{opt}(W, b, p) \). The drawback of the known PTAS is its running time, which grows exponentially with \( 1/\epsilon \). More discouraging, it has been proved for \( D \geq 2 \) that there does not exist a fully polynomial time approximation scheme (FPTAS) unless \( P=NP \) [8].

In the following, we present two novel approaches to \( DKP \) with fully polynomial running times. We will use \( \epsilon > 0 \) to denote the desired approximation parameter. Without loss of generality, we assume that \( \epsilon < 1 \). We will use \( x^{opt}(W, b, p) \) to denote an optimal solution for \( DKP(W, b, p) \), and use \( p^{opt}(W, b, p) \) to denote the optimal value for \( DKP(W, b, p) \).

In the first approach, our goal is to compute an almost feasible solution \( x^A \) to \( DKP(W, b, p) \) with a profit value at least as good as \( p^{opt}(W, b, p) \). By almost feasible, we mean \( x^A \) is a feasible solution to \( DKP(W, (1 + \epsilon)b, p) \). Our algorithm is presented in Algorithm 1.

**Algorithm 1 Outer Approximation for \( DKP(W, b, p) \)**

1. Construct an instance \( DKP(W', b', p) \) by defining \( W' \) and \( b' \) in the following way:

\[
w'_{ij} = \left[ \frac{w_{ij}}{b_i} \right] \cdot \frac{DK}{\epsilon} + 1, 1 \leq i \leq D, 1 \leq j \leq DK. \quad (23)
\]

\[
b'_i = \left[ \frac{DK}{\epsilon} \right] + DK, i = 1, 2, \cdots, D. \quad (24)
\]

2: Using dynamic programming to compute an optimal solution \( x^A \) of \( DKP(W', b', p) \). If \( DKP(W', b', p) \) is infeasible, STOP--DKP(W, b, p) is infeasible.

3: OUTPUT \( x^A \) as an almost feasible solution to \( DKP(W, b, p) \).

**Theorem 3.1:** The time complexity of Algorithm 1 is \( O((KDZ)^D/\epsilon) \). If \( DKP(W, b, p) \) is feasible, \( DKP(W', b', p) \) is guaranteed to be feasible. In this case, the solution \( x^A \) computed by Algorithm 1 is a feasible solution of \( DKP(W, (1 + \epsilon)b, p) \). Furthermore, \( \sum_{j=1}^{DK} p_j x^A_j \geq p^{opt}(W, b, p) \).

**Proof:** Since \( b'_i = \left[ \frac{DK}{\epsilon} \right] + DK \) for each \( i \), the time complexity of Algorithm 1 is \( O((KDZ)^D/\epsilon) \). Let \( x^A \) be an optimal solution for \( DKP(W', b', p) \). We will prove that \( x^A \) has the claimed property.

First, we will prove that \( x^A \) is a feasible solution to \( DKP(W, (1 + \epsilon)b, p) \). From (23), we have

\[
w_{ij} \left[ \frac{DK}{\epsilon} \right] + DK < w'_{ij} \leq \frac{w_{ij}}{b_i} \cdot \frac{DK}{\epsilon} + 1. \quad (25)
\]

From the left-hand of (25), we obtain

\[
\frac{w_{ij}}{b_i} < \frac{\epsilon b_i}{DK} w'_{ij}, \quad (26)
\]

which implies

\[
\sum_{j=1}^{DK} w_{ij} x^A_j < \frac{\epsilon b_i}{DK} \sum_{j=1}^{DK} w'_{ij} x^A_j \leq \frac{\epsilon b_i}{DK} \left( \frac{DK}{\epsilon} + DK \right) \leq (1 + \epsilon) b_i. \quad (27)
\]

This proves that \( x^A \) is a feasible solution to \( DKP(W, (1 + \epsilon)b, p) \).

Next, we will prove that \( \sum_{j=1}^{DK} p_j x^A_j \geq p^{opt}(W, b, p) \). Toward this goal, it suffices to show that any feasible solution \( x \) to \( DKP(W, b, p) \) is also a feasible solution to \( DKP(W', b', p) \), since the value of an optimal solution for \( DKP(W', b', p) \) (\( x^A \) in this case) is at least as good as the value of a feasible solution \( (x^{opt}(W, b, p)) \) in this case.

Since \( x \) is a feasible solution to \( DKP(W, b, p) \), we have

\[
\sum_{j=1}^{DK} w_{ij} x_j \leq b_i. \quad (28)
\]

Therefore, we have

\[
\sum_{j=1}^{DK} w_{ij} x_j < \frac{DK}{\epsilon b_i} \sum_{j=1}^{DK} w'_{ij} x^A_j < \frac{DK}{\epsilon b_i} \sum_{j=1}^{DK} w_{ij} x_j \leq \frac{DK}{\epsilon} + DK. \quad (29)
\]

This proves that \( x^A \) is a feasible solution to \( DKP(W, b, p) \).
Since $\sum_{j=1}^{DK} w_{ij}' x_j$ is an integer, (28) also implies $\sum_{j=1}^{DK} w_{ij}' x_j \leq \left\lfloor \frac{DK}{\epsilon} \right\rfloor + DK$.

In the second approach, our goal is to compute a strictly feasible solution $x^B$ to DKP($W, b, p$) with a profit value at least as good as $p^{opt}(W, (1-\epsilon)b, p)$. Our algorithm is presented in Algorithm 2.

**Algorithm 2 Inner Approximation for DKP($W, b, p$)**

1: Construct an instance DKP($W', b', p$) by defining $W'$ by (23) and $b'$ in the following way:

$$b'_i = \left\lfloor \frac{DK}{\epsilon} \right\rfloor, i = 1, 2, \cdots, D. \quad (29)$$

2: Using dynamic programming to compute an optimal solution $x^B$ of DKP($W', b', p$). If DKP($W', b', p$) is infeasible, STOP.

3: OUTPUT $x^B$ as an approximate solution to DKP($W, b, p$).

**Theorem 3.2:** The time complexity of Algorithm 1 is $O(KD(D/\epsilon)^D)$. The solution $x^B$ computed by Algorithm 2 is a feasible solution of DKP($W, b, p$). Furthermore $\sum_{j=1}^{DK} p_j x_j^B \geq p^{opt}(W, (1-\epsilon)b, p)$.

**Proof:** The time complexity can be proved the same way as in Theorem 3.1. Let $x^B$ be an optimal solution for DKP($W', b', p$). We will prove that $x^B$ has the claimed property.

First, we will prove that $x^B$ is a feasible solution to DKP($W, b, p$). Following the proof of Theorem 3.1, we have (26), which implies

$$\sum_{j=1}^{DK} w_{ij} x_j^B < \frac{eb_i}{DK} \sum_{j=1}^{DK} w_{ij} x_j^B \leq \frac{eb_i}{DK} \left\lfloor \frac{DK}{\epsilon} \right\rfloor \leq b_i. \quad (30)$$

This proves that $x^B$ is a feasible solution to DKP($W, b, p$).

Next, we will prove that $\sum_{j=1}^{DK} p_j x_j^B \geq p^{opt}(W, (1-\epsilon)b, p)$.

Toward this goal, it suffices to show that any feasible solution $x$ to DKP($W, (1-\epsilon)b, p$) is also a feasible solution to DKP($W', b', p$).

Since $x$ is a feasible solution to DKP($W, (1-\epsilon)b, p$), we have

$$\sum_{j=1}^{DK} w_{ij} x_j \leq (1-\epsilon)b_i. \quad (31)$$

Following the right-hand of (25), we have

$$\sum_{j=1}^{DK} w_{ij} x_j \leq \frac{DK}{eb_i} \sum_{j=1}^{DK} w_{ij} x_j + \sum_{j=1}^{DK} x_j \leq \frac{DK}{\epsilon} (1-\epsilon) + DK = \frac{DK}{\epsilon}. \quad (32)$$

Since $\sum_{j=1}^{DK} w_{ij} x_j$ is an integer, (32) also implies $\sum_{j=1}^{DK} w_{ij} x_j \leq \left\lfloor \frac{DK}{\epsilon} \right\rfloor$. This proves the theorem.

**Discussions:** DKP($W, (1+\epsilon)b, p$) differs from DKP($W, b, p$) in that the constraints are relaxed from $b_i$ to $(1+\epsilon)b_i$. Our first approach essentially says that by sacrificing the constraints a little bit, we can compute a solution with a profit at least as good as that of the optimal solution. More importantly, this can be achieved in fully polynomial time.

DKP($W, (1-\epsilon)b, p$) differs from DKP($W, b, p$) in that the constraints are tightened from $b_i$ to $(1-\epsilon)b_i$. Our second approach essentially says that a feasible solution to the DKP problem can be computed (in fully polynomial time) whose corresponding profit value is at least as good as that of the optimal value of the tightened version of the DKP problem. The first approach is guaranteed to compute a solution when the original knapsack problem is feasible. However, the second approach may fail to compute a solution when the original knapsack problem is feasible.

C. MIMO-Pipe Scheduling in Extended Networks: A Greedy Approach

Next, consider MIMO-pipe scheduling in extended networks, where the network scales to cover an increasing area with fixed node density. In such a network, the number of concurrent transmissions increases as $M$ increases. In other words, the number of constraints in (22) scales as $M$, and it has been shown that heuristic approximation algorithms are the only possible approach to solve this problem [8]. Based on [17], we develop a heuristic approximation algorithm to find an approximated solution of the multidimensional knapsack problem, as outlined below:

**Algorithm 3** A heuristic algorithm for problem P2 in extended networks

1: Initialize: $r_j \leftarrow 1, j = 1, 2, \cdots, MK$ and $\lambda_i \leftarrow 0, i = 1, 2, \cdots, M$.
2: Normalize: $\hat{Z}_{ij} \leftarrow \hat{Z}_{ij}/b_i$ and $y_i \leftarrow \sum_{j=1}^{MK} \hat{Z}_{ij}$;
3: $t^* \leftarrow \arg \max_i y_i$ (determine the most violated link);
4: for each $j$ such that $r_j = 1$ do
5: $\hat{t}_{i,j} \leftarrow \left\{ \begin{array}{ll} a_j - \sum_{i=1}^{M} \lambda_i \hat{Z}_{ij} \end{array} \right\} / \hat{Z}_{i,j}, a_{i^*,j} > 0 \rightarrow \infty$,
6: end for
7: $t^* \leftarrow \min_j \{ t^* \mid r_j = 1 \};$
8: $r_{i,j} = 0, y_{i,j} = y_i - \hat{Z}_{i,j}, \hat{t}_{i,j}^* = \lambda_i^* + t^*_{i,j};$
9: if $\max y_{i,j} > 1$ then
10: go to 3
11: end if
12: (Improve the solution) Check to see whether any variable $r_j$ can be set back to 1 without violating the conditions $y_{i,j} \leq 1$. If there is more than one possibility, choose the variable with the greatest $c_j$.

It can be shown that the worst-case time complexity of Algorithm 3 is $O(M^3K^2)$ [17].

IV. CONTENTION-BASED DISTRIBUTED MIMO-PIPE SCHEDULING

In the above sections, we have focused on centralized algorithms based on long term average SINR, aiming to maximize the ergodic capacity. Next, we turn our attention to random-access based distributed algorithms that make use of instantaneous SINR. Specifically, we assume that all links use mini-slots to contend for the channel with probability $p_i, i = 1, 2, \cdots, M$; and after each contention, the links involved in the contention update their contention probabilities accordingly. The channel contention (probing) continues for $\Delta$ mini-slots, followed by data transmissions over a duration of $T$, as shown in Fig. 4.
In general, it is difficult to jointly optimize link scheduling and stream multiplexing in a distributed manner under the MIMO-pipe model. One unique challenge is that multiple links can transmit successfully through one common channel; furthermore, each link has to update its own contention probability based on local information only, because links have no knowledge of the channel condition and interference of other links, but the network throughput depends on the data rates of all transmitting links. Moreover, it is challenging to simultaneously characterize the optimal scheme for joint stream multiplexing and link scheduling in a distributed manner. In the following, we propose a sub-optimal two-phase algorithm.

**Phase I: Finding a feasible set of coexisting links by using random access.** To this end, let all links choose Configuration I initially, and contend for the channel with certain probability. After each mini-slot, each link updates its contention probability in the following way: if its received SINR is greater than $\beta_1$, it would increase its contention probability for the next mini-slot; if its SINR is smaller than $\beta_1$, it would decrease its contention probability; if it does not contend during this mini-slot, it would keep its contention probability. In summary, the contention probability of link $i$ at the $t$th mini-slot, denoted as $p_i(t)$, is given by

$$p_i(t + 1) = [p_i(t) + a(t)f_i(p(t))]_{0}^{1},$$

where $a(t)$ is the stepsize that follows

$$a(t) < 1; \lim_{t \to \infty} a(t) = 0; \sum_{t=1}^{\infty} a(t) = \infty; \sum_{t=1}^{\infty} a^2(t) < \infty,$$

and $f_i(p(t))$ is given by

$$f_i(p(t)) = \begin{cases} 1, & \text{SINR}_i(n - 1) \geq \beta_1; \\ -1, & \text{SINR}_i(n - 1) < \beta_1; \\ 0, & \text{link } i \text{ did not contend in mini-slot } t. \end{cases}$$

The purpose of using the contention phase is to sort out the optimal set of links that can simultaneously transmit under Configuration I in a fully distributed manner. Our intuition is that a link in the optimal set has small chance to be interfered by other links, and thus can increase its contention probability more often. On the contrary, a link not in the optimal set would be interfered by other links more frequently, which makes its contention probability decrease to zero eventually.

**Phase II: Local optimization of stream multiplexing.** Phase I is used to find a feasible set of links that can transmit simultaneously under the basic configuration. The objective of Phase II is to determine whether any of the links in the set can do stream multiplexing or not. Observe that for any given link, regardless of its configuration, the interference it creates is the same because the total transmission power is the same. That is to say, a link can switch to a configuration with higher data rate, provided that its own SINR requirement is met, and the interference it incurs to other links remains unchanged. In light of this, each link can determine whether to multiplex or not based on local SINR requirements only.

We have the following result regarding the convergence of the iterative algorithm in (33) for Phase I.

**Proposition 4.1:** For any positive initial value $p(0)$, the sequence $\{p(t)\}$, generated by the iterative algorithm in (33), converge to an optimal contention probability $p^*$, where $p^*$ is an integer vector, i.e., $p^*_i \in \{0, 1\}, \forall i$.

The proof follows directly from Chapter 5 of [13], and is omitted here.

**V. Numerical Examples**

In this section, we illustrate, via numerical examples, the performance by using the proposed algorithms in a multi-hop MIMO network. We randomly generate $M$ links in a square field of size $L \times L$. Each MIMO link in the network has 4 antennas. The corresponding number of meaningful configurations is $K = 3$. The rest parameters are listed in Table II.

We first consider a network where all links are deployed in a square of size $10 \times 10$, and depict in Fig. 5(a) the average network throughput in a dense network achieved by the multidimensional knapsack approach. For the sake of comparison, we also show the network throughput corresponds to the distributed algorithm and the exhaustive search (the optimum). It can be seen from Fig. 5(a) that as the number of links increases, the average network throughput of all approaches increase. As expected, the proposed distributed algorithm is inferior to the knapsack approach, mainly because of the overhead incurred by the contention phase. In Fig. 5(b), the network throughput is plotted for different network sizes. It can be seen that the throughput increases as the network area grows. Further, when the network size grows, the throughput of the knapsack algorithm gets closer to the optimum. Our intuition is that when the network area increases with the number of links being kept the same, there is less interference across links and hence the coupling becomes weaker, and then the link scheduling problem can be better approximated by the knapsack formulation.

**TABLE II**

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$d_l$</th>
<th>$d_m$</th>
<th>$d_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>$T$</td>
<td>$d_l$</td>
<td>$d_m$</td>
<td>$d_h$</td>
</tr>
<tr>
<td>10$T$</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

$^4$Configuration I, or the basic configuration indicates the configuration that all transmit antennas send the same data stream as shown in Fig 2(e). This configuration has the lowest SINR requirement, and thus can involve the maximal number of links in the optimal set.

$^5$Note that the optimal set may not be unique.
In Fig. 5(c), we examine the impact of the contention phase, and plot the network throughput for the distributed algorithm, with \( \Delta = 5, 10, \) and 100, respectively. To compare the throughput gain, we ignore the contention overhead in this figure. It can be seen that when \( \Delta \) is small, increasing \( \Delta \) would improve the network throughput significantly. However, when \( \Delta \) is large, the gain becomes marginal, and is offset by the overhead, resulting in a severe performance degradation.

Fig. 6 depicts the convergence behavior of the proposed distributed algorithm, where the stepsize is set to be \( \frac{1}{2} \). The contention probabilities of all links converge to either 0 or 1 eventually, corroborating Prop. 4.1.

We first took a bottom-up approach to develop solid PHY-interference model, and extracted a set of meaningful stream multiplexing configurations, and their corresponding rates and SINR requirements, denoted by \( \{(rate, SINR)\} \). We then investigated MIMO-pipe scheduling, with the objective of maximizing the network throughput. Based on continuous relaxation via randomization, we characterized the structural property of the optimal scheduling. Our findings reveal that in an optimal policy, each individual MIMO link can use one stream configuration only. In light of this structural property, we then formulated the problem as a combinatorial optimization problem, and devised, via the multidimensional knapsack approach, efficient centralized approximation algorithms for both the extended network model and the dense network model. We also developed a two-phase contention-based distributed algorithm, and characterized its performance and convergence.

End-to-end throughput is another meaningful metric for multi-hop networks. Note that link throughput optimization is carried out via rate control, whereas end-to-end throughput is studied in the context of flow control/congestion control. It is of significant interest to explore joint congestion control and MIMO-pipe scheduling, along the line of recent work [5], [19]. Specifically, at the link level, the link throughput is maximized subject to the constraint that the data rate is no larger than the total flow rate through this link; and at the network level, the source node adjust its flow rate based on the congestion level. This will be an important subject of future work.

VI. CONCLUSIONS

In this paper, we developed a cross-layer optimization framework for effective interference management in MIMO multi-hop networks. In sharp contrast to the point-to-point settings, where the spatial multiplexing gain and diversity gain are defined in the high SNR regime, multi-hop networks are interference limited, and the SINR of each link is always bounded. These fundamental differences require a domain change from high SNR regimes to interference-limited regimes, in order to characterize the tradeoffs among possible MIMO gains.

REFERENCES