

行政院國家科學委員會專題研究計畫 期中進度報告

高速多天線分頻正交多工無線區域網路的整合介質存取控制及實體層設計(1/3)

計畫類別：整合型計畫

計畫編號：NSC93-2219-E-002-015-

執行期間：93年08月01日至94年07月31日

執行單位：國立臺灣大學電信工程學研究所

計畫主持人：蘇炫榮

計畫參與人員：林士駿、林秉勳、劉炫佃、劉浩然、枋立璋、余仲晟、顏瑞霖、邱文賢、朱祚仁

報告類型：完整報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 94 年 5 月 31 日

行政院國家科學委員會補助專題研究計畫

成果報告
期中進度報告

高速多天線分頻正交多工無線區域網路的整合介質存取控制及實體層設計(1/3)

Integrated MAC and Physical Layer Design for a High Speed Multi-Antenna OFDM WLAN (1/3)

計畫類別： 個別型計畫 整合型計畫

計畫編號：93-2219-E-002-015-

執行期間：2004年 8月 1日至 2005年 7月 31日

計畫主持人：蘇炫榮

共同主持人：

計畫參與人員：林士駿、林秉勳、劉炫佃、劉浩然、枋立瑋、余仲晟、顏瑞霖、邱文賢、朱祚仁

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：國立臺灣大學電信工程學研究所

中 華 民 國 94 年 5 月 27 日

Integrated MAC and Physical Layer Design for a High Speed Multi-Antenna OFDM WLAN (1/3)

英文摘要

Wireless LAN has enjoyed a growing popularity in recent years due to its unlicensed, low-cost and high-rate nature compared to cellular networks. This success has triggered a change in the end-user demand toward high quality, rate-hungry multimedia applications such as wireless high definition digital videos and games. As a result, the current 2.4 GHz and 5 GHz wireless LAN bands are becoming inadequate to fulfill the demand, partly due to insufficient bandwidth, but also due to the inefficient medium access control (MAC) protocols that are being used in the current wireless LAN. In view of this, it is the purpose of this project to investigate the physical layer (PHY) and MAC issues involved in deploying wireless LANs with wideband transmission in the 60 GHz band where ample unlicensed bandwidth is available. In addition to wide transmission bandwidth, multiple-input multiple-output (MIMO) antenna techniques will be used to further boost the transmission rate. To ease the complexity of equalization in such a wide bandwidth, orthogonal frequency division multiplexing (OFDM) is adopted. A coordinated, cellular-like network structure is considered to enhance the MAC throughput and frequency reuse.

In the first year of the project, some PHY and MAC issues have been considered for the proposed MIMO OFDM system. Specifically, we considered two multiplexing schemes for multiple users to share the downlink channel capacity increased by MIMO: grouped transmission, and overlapped transmission. In the former, for simplicity, the available transmission antennas are divided into a few groups, and each user is assigned a group of antennas. For such a scheme, each receiver only receives useful data from its corresponding transmission antennas, therefore it may know only the channel gains from these antennas through training. In order to achieve a good detection performance, the receiver will need an adaptive mechanism to measure the statistics of the interference from the other transmission antennas in order to suppress it. We investigate in this report an adaptive maximum signal-to-noise ratio (SNR) filter bank receiver for such a grouped transmission. The overlapped transmission is the optimal capacity sharing scheme in terms of maximizing the total transmission rate. For such a scheme, all users use all the transmission antennas, thus interfering with one another. We apply the Dirty Paper Coding (DPC) method to suppress mutual interference among users. The optimal design parameters for Vector (i.e., MIMO) DPC are derived for a specific superposition coding realization.

To improve the network-wide throughput, a power control and allocation method was adopted for multi-cell OFDM networks to achieve interference avoidance between cells. Such a method relies on a simple power allocation mechanism autonomously run by each co-channel cell to achieve multi-cell water filling. As the mechanism runs, each cell will snatch sub-carriers that are good for it, and gradually release the sub-carriers with bad channel conditions to reduce interference to other cells. Such a method has another inherent benefit of lower peak-to-average power ratio as compared to conventional equal power allocation method. The advantage is also investigated in this report.

Finally, as adaptive modulation and coding (AMC) and Hybrid ARQ (HARQ) are becoming popular in new high speed wireless network standards for their ability to maximize the end-to-end throughput, we will also consider them in the proposed MIMO OFDM system. In the first year of the project, we proposed a simple adaptive method to find the optimal parameters for operating AMC and HARQ. This method is very general that it can be easily ported to any PHY structure, including MIMO OFDM.

Keywords: WLAN, 60 GHz, OFDM, MIMO, diversity, smart antenna, MAC, Hybrid ARQ, adaptive modulation and coding (AMC), power control, scheduling, multiple access, dirty paper coding.

中文摘要

無線區域網路因其所在頻段無需使用執照，且較行動通訊成本低、傳輸率高的優勢，於近年來蓬勃發展。然而隨著無線多媒體影音及遊戲等應用的風行，無線區域網路所使用的 2.4 GHz 及 5 GHz 頻段已逐漸有擁塞、不敷使用的現象。追根究底，其原因除了現有頻段不夠容納高品質的多媒體需求外，介質存取控制協定效率低也是主要原因之一。本子計畫的目的為研究在具有極大可用頻寬的 60 GHz 頻段設置無線區域網路所可能遭遇的實體層及介質存取控制上的問題。除了使用寬頻外，多天線技術 (MIMO) 將被採用來進一步增加傳輸速率。至於因寬頻而引發的高複雜度等化器問題，我們將採用正交分頻多工調變 (OFDM) 的架構來解決。此外，我們將採用具協調性、類似 (行動通訊系統的) 蜂巢網路的架構來增加介質存取控制層的產率及系統的頻率重複使用率。

在第一年的計畫裡，我們探討了一些 MIMO OFDM 的實體層及介質存取控制設計上的問題。首先我們考慮了兩種可以讓多用戶分享因 MIMO 而增加的下行通道容量的方式：第一種是分群傳輸、第二種是重疊傳輸。分群傳輸是將發射端天線分為幾群，每個用戶使用一群的簡單辦法。當使用這種方式時，接收端同常只在乎從本群天線傳來的信號，因此只會估測本群天線的通道參數。這時候抑制由其他發射天線造成的干擾、以增加資料傳輸的正確率通常必須藉由可適性統計及信號處理來完成。在這個報告中，我們探討了用於分群傳輸的可適性最大訊噪比濾波方式。重疊傳輸則可達到最佳的通道容量分享及最大的總和傳輸率。這種分享方式讓所有用戶都使用所有的發射天線。用戶間的干擾則藉由「髒紙編碼」來降低。在這個報告中，我們針對疊加性 MIMO 髒紙編碼推導出最佳的设计參數。

蜂巢狀 OFDM 網路的系統產率則是藉由電力分配及控制來提升。我們用了一個簡單、可由各基地台及所屬用戶獨立執行、而不需額外全網路協調的電力分配機制來避免不同基地台間的相互干擾。隨著通道狀況改變，這個機制能讓每個基地台逐漸佔用對它來說比較有利的 OFDM 子通道，而釋放出不好的通道以降低對其他佔用這些通道的基地台的干擾。這個電力控制機制還有低峰均值的優點。這個優點在報告中一併探討。

最後，我們考慮使用最近許多高速無線通信系統都具備的可適性調變編碼及混合自動重新傳輸技術於 MIMO OFDM 中。在計畫的第一年，我們提出了一個間單的可適性機制讓系統能自動搜尋最佳的可適性調變編碼及混合自動重新傳輸設計參數。這個方法具有普遍性，可以被使用在任何有可適性調變編碼及混合自動重新傳輸的通信系統。

關鍵詞：無線區域網路、60 GHz、正交分頻多工調變、多輸入多輸出通道、多重性、智慧型天線、介質存取控制、混合自動重新傳輸、可適性調變編碼，傳輸功率控制、排時、多重存取、髒紙編碼。

Integrated MAC and Physical Layer Design for a High Speed Multi-Antenna OFDM WLAN (1/3)

Abstract

Wireless LAN has enjoyed a growing popularity in recent years due to its unlicensed, low-cost and high-rate nature compared to cellular networks. This success has triggered a change in the end-user demand toward high quality, rate-hungry multimedia applications such as wireless high definition digital videos and games. As a result, the current 2.4 GHz and 5 GHz wireless LAN bands are becoming inadequate to fulfill the demand, partly due to insufficient bandwidth, but also due to the inefficient medium access control (MAC) protocols that are being used in the current wireless LAN. In view of this, it is the purpose of this project to investigate the physical layer (PHY) and MAC issues involved in deploying wireless LANs with wideband transmission in the 60 GHz band where ample unlicensed bandwidth is available. In addition to wide transmission bandwidth, multiple-input multiple-output (MIMO) antenna techniques will be used to further boost the transmission rate. To ease the complexity of equalization in such a wide bandwidth, orthogonal frequency division multiplexing (OFDM) is adopted. A coordinated, cellular-like network structure is considered to enhance the MAC throughput and frequency reuse.

In the first year of the project, some PHY and MAC issues have been considered for the proposed MIMO OFDM system. Specifically, we considered two multiplexing schemes for multiple users to share the downlink channel capacity increased by MIMO: grouped transmission, and overlapped transmission. In the former, for simplicity, the available transmission antennas are divided into a few groups, and each user is assigned a group of antennas. For such a scheme, each receiver only receives useful data from its corresponding transmission antennas, therefore it may know only the channel gains from these antennas through training. In order to achieve a good detection performance, the receiver will need an adaptive mechanism to measure the statistics of the interference from the other transmission antennas in order to suppress it. We investigate in this report an adaptive maximum signal-to-noise ratio (SNR) filter bank receiver for such a grouped transmission. The overlapped transmission is the optimal capacity sharing scheme in terms of maximizing the total transmission rate. For such a scheme, all users use all the transmission antennas, thus interfering with one another. We apply the Dirty Paper Coding (DPC) method to suppress mutual interference among users. The optimal design parameters for Vector (i.e., MIMO) DPC are derived for a specific superposition coding realization.

To improve the network-wide throughput, a power control and allocation method was adopted for multi-cell OFDM networks to achieve interference avoidance between cells. Such a method relies on a simple power allocation mechanism autonomously run by each co-channel cell to achieve multi-cell water filling. As the mechanism runs, each cell will snatch sub-carriers that are good for it, and gradually release the sub-carriers with bad channel conditions to reduce interference to other cells. Such a method has another inherent benefit of lower peak-to-average power ratio as compared to conventional equal power allocation method. The advantage is also investigated in this report.

Finally, as adaptive modulation and coding (AMC) and Hybrid ARQ (HARQ) are becoming popular in new high speed wireless network standards for their ability to maximize the end-to-end throughput, we will also consider them in the proposed MIMO OFDM system. In the first year of the project, we proposed a simple adaptive method to find the optimal parameters for operating AMC and HARQ. This method is very general that it can be easily ported to any PHY structure, including MIMO OFDM.

Index Terms

WLAN, 60 GHz, OFDM, MIMO, diversity, smart antenna, MAC, Hybrid ARQ, adaptive modulation and coding (AMC), power control, scheduling, multiple access, dirty paper coding.

I. INTRODUCTION

With the boom of the Internet, and the fact that the services provided by the current wireless cellular systems failed to catch up with the growing demand of high data rate Internet access, the wireless LAN has enjoyed its success in filling this gap in recent years. Since wireless LAN is low-cost and easy to deploy, it has gained increasing popularity which in turn has changed the end user behavior in accessing the Internet. More and more data-rate-hungry multimedia and gaming applications are being launched routinely. Many vendors are also envisioning a “digital home” in which the wires that used to connect home appliances will be replaced by wireless links to allow central control, remote monitoring and intelligent operation. In that scenario, home entertainment and education will be more far-reaching, with multimedia contents from the Internet being played on the high definition digital televisions using home computers or wireless LAN access points as gateways. As a result, even the transmission rate provided by the current wireless LAN (up to 11 Mb/s for 802.11b, and up to 54 Mb/s for 802.11a and g) is becoming inadequate to meet the demand. The grip of this triggered the recent proposition to enhance the data rate of the wireless LAN.

One straightforward approach to increasing the data rate of wireless LANs is to look for a wider transmission bandwidth. On this front, the 60 GHz band is of particular interest because this is the band in which a massive amount of spectral space has been allocated worldwide for wireless LAN communications. For example, in Europe the frequency ranges 62-63 GHz and 65-66 GHz are reserved for wideband mobile networks (Mobile Broadband System, MBS), whereas 59-62 GHz range is reserved for unlicensed wideband wireless LAN. In the United States the frequency range 59-66 GHz is a generally unlicensed range. In Japan 59-66 GHz is reserved for wireless communications [1] [2]. Due to the extremely high carrier frequency, the 60 GHz band poses several challenges on the transceiver design such as high phase noise, high Doppler shift, high atmosphere attenuation and shadowing effect, just to name a few. Understanding the 60 GHz channel in various usage scenarios and environments, and to design efficient and feasible transceivers to be used in a high throughput wireless LAN are the major goals of the group project. The group project will also investigate the additional throughput and performance boost offered by multiple input multiple output (MIMO) antennas [3], as the short wavelength at 60 GHz may allow many antennas to be packed tightly and still providing diversity with uncorrelated transmissions and receptions.

The goal of this project is to exploit the 60 GHz channel characteristics and the physical layer (PHY) transceiver design aspects, together with the other two sub-projects, and use these properties in the medium access control (MAC) protocol design to achieve a bandwidth efficient wideband wireless LAN. The focus will be more on the information-theoretic understanding of the system and environments, and how the protocols, system architectures and parameters can be enhanced to close the gap between the practical throughput and the theoretical capacity. Since the single-link transmission bandwidth has to be very large in order to accommodate multimedia services, it is generally agreed that orthogonal frequency division multiplexing (OFDM) is a low-complexity modulation solution under such constraints. Thus, in addition to the MIMO technology, OFDM will be adopted as the transceiver structure of this project.

In the first year of the project, some PHY and MAC issues have been considered for the proposed MIMO OFDM system. Specifically, we considered two multiplexing schemes for multiple users to share the downlink channel capacity increased by MIMO: grouped transmission, and overlapped transmission. In the former, for simplicity, the available transmission antennas are divided into a few groups, and each user is assigned a group of antennas [4]. For such a scheme, each receiver only receives useful data from its corresponding transmission antennas, therefore it may know only the channel gains from these antennas through training. In order to achieve a good detection performance, the receiver will need an adaptive mechanism to measure the statistics of the interference from the other transmission antennas in order to suppress it. We investigate in this report an adaptive maximum signal-to-noise ratio (SNR) filter bank receiver for such a grouped transmission.

The overlapped transmission is the optimal capacity sharing scheme in terms of maximizing the total transmission rate [5]. For such a scheme, all users use all the transmission antennas, thus interfering with

one another. We apply the Dirty Paper Coding (DPC) [6] method to suppress mutual interference among users. The optimal design parameters for Vector (i.e., MIMO) DPC are derived for a specific superposition coding realization.

To improve the network-wide throughput, a power control and allocation method was adopted for multi-cell OFDM networks to achieve interference avoidance between cells [7]. Such a method relies on a simple power allocation mechanism autonomously run by each co-channel cell to achieve multi-cell water filling. As the mechanism runs, each cell will snatch sub-carriers that are good for it, and gradually release the sub-carriers with bad channel conditions to reduce interference to other cells. Such a method has another inherent benefit of lower peak-to-average power ratio (PAPR) as compared to conventional equal power allocation method. The advantage is also investigated in this report.

Finally, as adaptive modulation and coding (AMC) and Hybrid ARQ (HARQ) are becoming popular in new high speed wireless network standards for their ability to maximize the end-to-end throughput, we will also consider them in the proposed MIMO OFDM system. In the first year of the project, we proposed a simple adaptive method to find the optimal parameters for operating AMC and HARQ. This method is very general that it can be easily ported to any PHY structure, including MIMO OFDM.

This report is organized as follows. In Section II, an adaptive maximum SNR array processing algorithm will be considered for the grouped MIMO transmission. In Section III, the optimal design parameters for vector superposition DPC used in the overlapped MIMO transmission will be derived. Section IV demonstrates an autonomous power allocation algorithm for a multi-cell OFDM network operating at 60 GHz. Section V then shows an adaptive throughput maximization algorithm for AMC systems with HARQ. The conclusion and self progress evaluation are given in Section VI.

II. ADAPTIVE MAXIMUM SNR ARRAY PROCESSING

In this section, we consider adaptive antenna array processing for a simple grouped MIMO transmission [4] enabling multiple users to share the downlink capacity.

A. Background

For the grouped MIMO transmission, the transmit antennas are divided into groups each serving a specific user, and independently space-time coded. The space-time coding advantages are explored by the antennas within the same group, while the system throughput is increased via independent transmission among groups. For this type of systems, signal processing at the receiver is required to suppress the interferences from other transmission groups when decoding a group. Even in the absence of the interfering signals, signal processing can be applied to reduce the maximum likelihood decoding complexity. In [8], a maximum SNR method was used for interference suppression. The derivation can be seen as a generalized minimum mean squared error (MMSE) approach when the desired signal has more than one dimension.

The approach in [8] assumed the signal covariance matrix \mathbf{R}_s and the interference covariance matrix \mathbf{R}_n were available. In adaptive signal processing systems, the exact values of them are unknown and need to be estimated. In this section we develop a stochastic gradient algorithm which recursively updates the array filter bank coefficients for each new set of received signal. The succeeding sections are organized as follows. In Section II-B the system model and problem formulation are described. The adaptive algorithm is shown in Section II-C. Some preliminary simulation results and discussions are presented in Section II-D. Section II-E surveys some channel estimation algorithms for obtaining \mathbf{R}_s .

B. System model

The communication system considered consists of a transmitter which is equipped with n antennas and a receiver which has m antennas. The signal seen at each receive antenna is a noisy superposition of the n transmitted signals corrupted by the fading channel which is assumed to be flat, Rayleigh, and spatially independent, meaning fading is statistically independent from one transmit-receive antenna pair to another. The signal r_t^j received by antenna j ($1 \leq j \leq m$) at time t is given by

$$r_t^j = \sum_{i=1}^n \alpha_{i,j} c_t^i \sqrt{E_i} + \eta_t^j \quad (1)$$

where i is the transmit antenna index, E_i is the symbol energy of the i^{th} transmit antenna, c_t^i is the symbol transmitted from antenna i at time t , and $\alpha_{i,j}$ is the path gain from the i^{th} transmit antenna to the j^{th} receive antenna. The noise η_t^j is assumed to be additive white Gaussian (AWGN) with variance $N_0/2$ for both real and imaginary components.

For notational simplicity, we redefine $\alpha_{i,j}$ as the product of the channel gain $\alpha_{i,j}$ and the symbol amplitude $\sqrt{E_i}$ in (1). With this new notation, we rewrite (1) in the vector form

$$\mathbf{r}_t = \mathbf{\Omega} \mathbf{c}_t + \boldsymbol{\eta}_t$$

where

$$\mathbf{\Omega} = \begin{pmatrix} \alpha_{1,1} & \alpha_{2,1} & \cdots & \alpha_{n,1} \\ \alpha_{1,2} & \alpha_{2,2} & \cdots & \alpha_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1,m} & \alpha_{2,m} & \cdots & \alpha_{n,m} \end{pmatrix},$$

and $\mathbf{c}_t = (c_t^1, c_t^2, \dots, c_t^n)^{\mathbf{T}}$, $\mathbf{r}_t = (r_t^1, r_t^2, \dots, r_t^m)^{\mathbf{T}}$, $\boldsymbol{\eta}_t = (\eta_t^1, \eta_t^2, \dots, \eta_t^m)^{\mathbf{T}}$.

The data to be transmitted is divided into q groups, each space-time coded and transmitted with one group of antennas. These groups of data are successively decoded (and cancelled) at the receiver. Let n_j denote the number of transmit antennas in group j , $1 \leq j \leq q$, we have $n_1 + n_2 + \dots + n_q = n$. In decoding

a specific group, the uncanceled groups of signals are considered as interferences, therefore antenna array processing is necessary to suppress them. Without loss of generality, we consider decoding of the first group code denoted by C_1 , and assume that the channel gain matrix Ω is known to the receiver.

We now introduce the maximum SNR array processor proposed in [8]. For a given time t , since the space-time coded symbols are unknown at the array processing stage, the SNR can only be maximized statistically. For this purpose, the signal and interference covariance matrices are defined, considering decoding the first group code, as

$$\begin{aligned}\mathbf{R}_s &= \Omega(C_1)E \left[\mathbf{c}_t^1 (\mathbf{c}_t^1)^\dagger \right] \Omega^\dagger(C_1) \\ \mathbf{R}_n &= \Lambda(C_1)E \left[\mathbf{c}_t^o (\mathbf{c}_t^o)^\dagger \right] \Lambda^\dagger(C_1) + N_0\mathbf{I},\end{aligned}\quad (2)$$

respectively, where $\mathbf{c}_t^1 = (c_t^1, c_t^2, \dots, c_t^{n_1})^\mathbf{T}$, $\mathbf{c}_t^o = (c_t^{n_1+1}, c_t^{n_1+2}, \dots, c_t^n)^\mathbf{T}$, and

$$\Omega(C_1) = \begin{pmatrix} \alpha_{1,1} & \alpha_{2,1} & \cdots & \alpha_{n_1,1} \\ \alpha_{1,2} & \alpha_{2,2} & \cdots & \alpha_{n_1,2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1,m} & \alpha_{2,m} & \cdots & \alpha_{n_1,m} \end{pmatrix}, \quad \Lambda(C_1) = \begin{pmatrix} \alpha_{n_1+1,1} & \alpha_{n_1+2,1} & \cdots & \alpha_{n,1} \\ \alpha_{n_1+1,2} & \alpha_{n_1+2,2} & \cdots & \alpha_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n_1+1,m} & \alpha_{n_1+2,m} & \cdots & \alpha_{n,m} \end{pmatrix}.$$

The maximum SNR array processor will require the same number of linear filters as the dimension of the signal space in order to collect all energy of the desired signal. Each of these linear filters can be decomposed into two components: one in the signal space and the other in the orthogonal space. The maximum SNR array processor is obtained using the following theorem [8].

Theorem 1: Let m be the number of receive antennas, n_1 be the number of transmit antennas of the space-time code group in consideration; \mathbf{R}_s and \mathbf{R}_n be the signal and interference covariance matrices, respectively, defined by (2). The linear maximum SNR array processor for this group, denoted by $\Theta(C_1) = (\mathbf{w}^1 \ \mathbf{w}^2 \ \cdots \ \mathbf{w}^K)$, consists of a set of K linearly independent eigenvectors corresponding to the nonzero eigenvalues, counting multiplicities, of the generalized eigenvalue problem

$$\mathbf{R}_s \mathbf{w} = \lambda \mathbf{R}_n \mathbf{w}, \quad (3)$$

where $K \leq \min(m, n_1)$ is the rank of \mathbf{R}_s . The filtering outputs of these eigenvectors $\{(\mathbf{w}^k)^\dagger \mathbf{r}_t\}_{k=1}^K$ are uncorrelated with one another.

C. Adaptive Algorithm

The goal is to find \mathbf{w} in (3) adaptively. Here we adopt the stochastic gradient algorithms. The first step is to apply Newton's method on (3), assuming \mathbf{R}_s and \mathbf{R}_n are perfectly known. The solution is provided in [9]. We first rewrite (3) as

$$\mathbf{R}_s \mathbf{w} = \lambda \mathbf{R}_n \mathbf{w}, \quad \mathbf{e}_k^\mathbf{T} \mathbf{w} = 1,$$

where \mathbf{e}_k is a vector with all entries zero except the k th which is one. To find multiple eigenvectors, an additional equation is added because we need one degree of freedom to differentiate eigenvectors. Let $\mathbf{v} = [\mathbf{w}^\mathbf{T}, \lambda]^\mathbf{T}$, an equivalent nonlinear minimization problem is formed as

$$F(\mathbf{v}) = \begin{bmatrix} (\mathbf{R}_s - \lambda \mathbf{R}_n) \mathbf{w} \\ \mathbf{e}_k^\mathbf{T} \mathbf{w} - 1 \end{bmatrix}.$$

Newton's iteration generates a sequence \mathbf{v}_T produced by

$$\mathbf{v}_{T+1} = \mathbf{v}_T - J(\mathbf{v}_T)^{-1} F(\mathbf{v}_T),$$

where $J(\mathbf{v}_T) = F'(\mathbf{v}_T)$ is the Jacobian. And according to [9], the above equation is equal to

$$\mathbf{M}_T \Delta_T = \xi(\mathbf{v}_T),$$

where the residual is

$$\xi(\mathbf{v}_T) = \lambda_T \mathbf{R}_n \mathbf{w}_T - \mathbf{R}_s \mathbf{w}_T, \quad (4)$$

the Newton step size is Δ_T , and the matrix \mathbf{M}_T is formed by replacing the k^{th} column of $\mathbf{R}_s - \lambda \mathbf{R}_n$ with $-\mathbf{R}_n \mathbf{w}_T$, respectively. One feature of this method is that it does not assume the matrices in (3) to be symmetric, which is favorable when approximated covariance matrices are used.

The Newton method described above is indeed an infeasible start Newton method [10, Chapter 10.3]. It works with initial points and iterates which are not feasible. To speed up the convergence, we adopt the backtracking line search [10] on the residual two-norm $\|\xi(\cdot)\|_2$. That is, we first choose step size $\mu = 1$ and multiply it by β until the following condition is not satisfied

$$\|\xi(\mathbf{v}_T + \mu \Delta_T)\|_2 > (1 - \alpha \mu) \|\xi(\mathbf{v}_T)\|_2,$$

where the constants are bounded by $0 < \alpha < 0.5$ and $0 < \beta < 1$.

After introducing the calculation of Newton step, the remaining task is to find the stochastic approximation of (2). Assuming we have perfect channel estimation mechanism to find $\Omega(C_1)$, we can form accurate \mathbf{R}_s . As for \mathbf{R}_n , we use the following as the approximation for the T^{th} iterations,

$$\hat{\mathbf{R}}_n = \frac{1}{T} \sum_{t=1}^T \mathbf{r}_t (\mathbf{r}_t)^\dagger - \mathbf{R}_s. \quad (5)$$

The detail adaptive algorithm calculating the filter coefficients \mathbf{w}_T for the T^{th} iterations is summarized in Algorithm 1. Here we choose the constants $\alpha = 0.5$ and $\beta = 0.8$ in determining the step size.

Algorithm 1 Adaptive Maximum SINR array processing algorithm

Initial condition

1.1 Find the rank of \mathbf{R}_s as the total number of filters K

1.2 Assign $\mathbf{w}_1^k = \mathbf{e}_k$ and proper λ_1^k for the k th filter

Adaptation For the k^{th} filter and the T^{th} iterations, to find the filter coefficient \mathbf{w}_T^k

2.1 $\hat{\mathbf{R}}_n = \frac{1}{T} \sum_{t=1}^T \mathbf{r}_t (\mathbf{r}_t)^\dagger - \mathbf{R}_s$

2.2 Form the residual error $\xi(\mathbf{v}_T) = \lambda_T^k \hat{\mathbf{R}}_n \mathbf{w}_T^k - \mathbf{R}_s \mathbf{w}_T^k$

2.3 Form \mathbf{M}_T by replacing the k th column of $\mathbf{R}_s - \lambda_T^k \hat{\mathbf{R}}_n$ with $-\hat{\mathbf{R}}_n \mathbf{w}_T^k$

2.4 Find the increments by solving $\mathbf{M}_T \Delta_T = \xi(\mathbf{v}_T)$ then $\Delta \lambda_T^k = \mathbf{e}_k^T \Delta_T$ and $\Delta \mathbf{w}_T^k = \Delta_T - \Delta \lambda_T^k \mathbf{e}_k$

2.5 Choosing the step size μ

$\mu \leftarrow 1$

while $\|\xi(\mathbf{v}_T + \mu \Delta_T)\|_2 > (1 - \alpha \mu) \|\xi(\mathbf{v}_T)\|_2$ **do**

$\mu \leftarrow \beta \mu$

end while

2.6 Updating: $\lambda_{T+1}^k \leftarrow \lambda_T^k + \mu \Delta \lambda_T^k$, $\mathbf{w}_{T+1}^k \leftarrow \mathbf{w}_T^k + \mu \Delta \mathbf{w}_T^k$

D. Preliminary Simulation Results and Discussions

We use quasi-static channel to test our algorithm. It assumes that the fading remains constant within a frame and varies from one frame to another. The coefficients after convergence are used to decode data sequences. The 32-state space-time trellis code with 4-PSK modulation [4] is used. Here we show two convergence results under two different channel realizations. There are four antennas at the transmitter and four antennas at the receiver. Two groups each having two transmit antennas are simulated and the

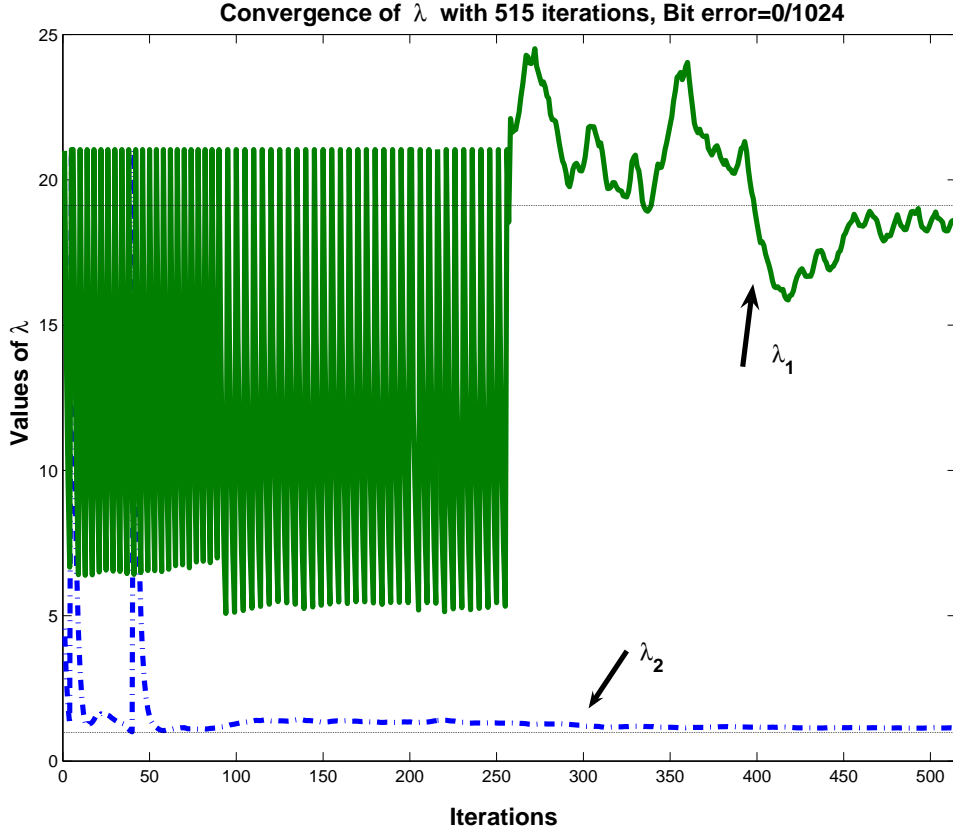


Fig. 1. Convergence of λ under channel realization 1 and 515 iterations , $\lambda_1=19.1$ and $\lambda_2=0.98$

results focus on the first group. The SNR per receiver antenna is 11 dB. In Fig. 1, a total of 1024 bits are transmitted in a frame and the bit error is zero. Fig. 2 shows the performance under another channel realization and the number of bits per frame is doubled.

However, the bit error rate (BER) is 1.4×10^{-2} for a 100-frame simulation with channel independent from frame to frame, which is much worse than the non-adaptive result in [8]. Although our algorithm does not require symmetric $\hat{\mathbf{R}}_n$, the performance still suffers from the inaccurate ensemble average estimation $\hat{\mathbf{R}}_n$ in (5). In [11] and [12], \mathbf{R}_n is assumed known, and symmetric. Although this assumption might not hold for adaptive systems, we may try to modify these algorithms and test the performances. In addition, if only the largest eigenvector needs to be found, [13] gives an adaptive algorithm to find its coefficient. We may also consider this algorithm and compared it with our algorithm.

E. Channel Estimation Issues

To obtain the promising advantage of the MIMO system, a good channel coefficient estimator is needed. In [14]–[16], a zero-padded OFDM (ZP-OFDM) with pseudo-random-postfix modulation is proposed. We will first discuss the single-antenna case then the MIMO case. This modulation replaces the zero-padded part in ZP-OFDM with a pseudo-randomly weighted deterministic sequence to perform channel estimation. Compared with the conventional cyclic-prefix OFDM, this scheme does not impact the achieved useful data rate when adding known sequence to do channel estimation. Once the channel estimation procedure is done, we can subtract the impact of this known postfix with the estimated channel impulse response from the received signal. Then the low-complexity equalization of ZP-OFDM can still be performed by using the overlap-and-add method. In the MIMO case, since for each received antenna the various postfixes are

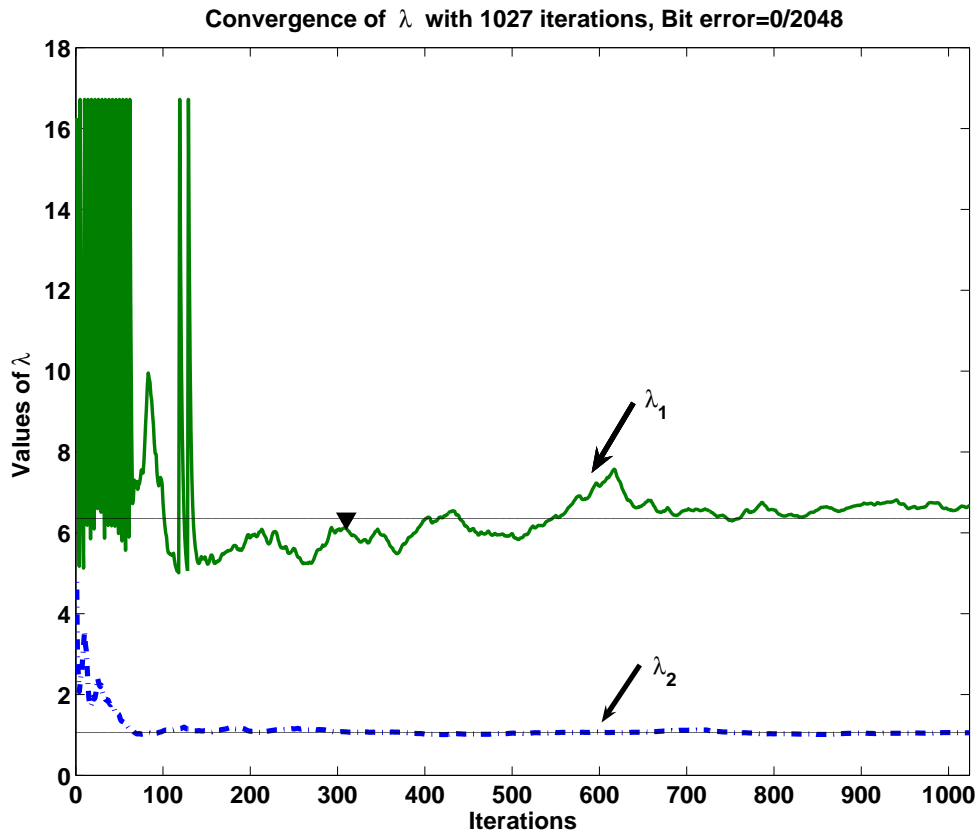


Fig. 2. Convergence of λ under channel realization 2 and 1027 iterations, $\lambda_1=6.35$ and $\lambda_2=1.06$

superimposed, we need to separate them. For this goal, an orthogonal matrix is added in the postfixes generating process [14].

The channel estimation procedure in [15] is derived by a simple averaging of the sum of the postfix with the head of the time domain OFDM received block. Since the useful data symbols are assumed to be zero-mean, this averaging can cancel the interference of the samples carrying them on the postfix. Since the postfix is known, we can then find the channel response. This method can be further combined with the soft-information of the data sequences available at error-correcting decoder output [16]. With the channel estimated by previous iteration, we can obtain the log-likelihood ratios of data symbols at the decoder. The soft-information of data can be used to find an estimation of the aforementioned interference in MMSE sense. This new interference estimation can be subtracted from the received block, then more precise channel response and lower symbol error rate can be obtained in the next iteration.

Based on the results in [14]–[16], we will focus on migrating it into MIMO space-time coded system. Powerful space-time repeat-accumulate or turbo code will be used. More design considerations will be considered in choosing postfix in MIMO system, such as spectrum flatness and channel spatial correlation.

III. VECTOR SUPERPOSITION DIRTY PAPER CODING FOR MIMO GAUSSIAN BROADCAST CHANNELS

In this section, we consider overlapped MIMO downlink transmission which is the optimal capacity sharing scheme in terms of maximizing the total transmission rate. For such a scheme, all users use all the transmission antennas, thus interfering with one another. We apply the DPC method to suppress mutual interference among users. The optimal design parameters for Vector Superposition DPC are derived.

A. Background

A MIMO broadcast channel (BC) system consists of one transmitter sending information to many receivers, all equipped with multiple antennas. The capacity region of a general MIMO BC is still unknown, but that of the MIMO Gaussian BC (GBC) has recently been shown to coincide with the achievable rate region when DPC is utilized [5]. With DPC [6], [17], the multi-user interference, known at the transmitter, can be removed for a desired receiver to achieve the interference-free rate performance.

The goal of this section is to find a vector DPC structure for MIMO GBC that can achieve the capacity. We extend the scheme proposed in [18] to the MIMO scenario. In addition, we provide some insights on how this coding scheme achieves capacity. The section organization is as follows. In Section III-B we will introduce the system model and problem formulation. The vector superposition DPC is presented in Section III-C. Section III-D provides the code parameters selection by performing a random-coding analysis. Finally, Section III-E concludes this section.

B. System Model

In this section, the matrices and vectors are denoted in boldface. For matrix \mathbf{G} , $\text{Tr}(\mathbf{G})$ denotes the trace; \mathbf{G}^T and \mathbf{G}^\dagger denote the transpose and conjugate transpose, respectively. \mathbf{G}_s^{-1} and $|\mathbf{G}_s|$ are the inverse and determinant of a square matrix \mathbf{G}_s . And \mathbf{I}_t denotes the identity matrix of dimension t .

Consider a two-user MIMO GBC system with M transmitter antennas whose total transmission power is constrained. The channel gains from the transmitter antennas to the receiver antennas are assumed known perfectly at both the transmitter and the receivers. The optimal coding scheme for this channel [17] will first specify the vector DPC achievable rate for each user by determining the channel input covariance matrices and encoding order for all users, and then apply the vector DPC on each user's message. The sum of the coded signals of all users will be sent to all receivers. At the receivers, whether or how a user's signal will be interfered by the other users' signals is governed by the vector DPC encoding order. In general, the signals encoded earlier will be invisible to the signals encoded later, while the former will be interfered by the latter. In this section, we will focus on the coding scheme for a user who already knows the previously encoded users' signals and tries to achieve the specified interference-free rate.

Without loss of generality, we consider the two-user scenario and assume that the first user is assigned the interference-free rate. The received signal at the first receiver, which has N receiver antennas, can be written as

$$\mathbf{y}_t^c = \mathbf{H}^c(\mathbf{x}_{1,t}^c + \mathbf{x}_{2,t}^c) + \mathbf{n}_t^c, \quad t = 1 \dots T, \quad (6)$$

where t is the time index, T is the number of symbols in the code block; $\mathbf{y}_t^c \in \mathbb{C}^{N \times 1}$ is the t th received symbol; $\mathbf{x}_{k,t}^c \in \mathbb{C}^{M \times 1}$, $k = 1, 2$, is the t th vector symbol for the message of user k ; $\mathbf{H}^c \in \mathbb{C}^{N \times M}$ is the MIMO channel gain matrix which is assumed to be constant over a code block, and $\mathbf{n}_t^c \in \mathbb{C}^{N \times 1}$ is the additive Gaussian noise at receiver where $\mathbf{n}_t^c \sim N_{\mathbb{C}}(0, \mathbf{I}_N)$.

In the sequel, we will show that with the transmission power constraint P_1 for the first user, when $T \rightarrow \infty$, there exists a coding structure with rate equal to the single user MIMO capacity [17]

$$C_{MIMO}^S = \log |\mathbf{H}^c \Sigma_1 (\mathbf{H}^c)^\dagger + \mathbf{I}_N|, \quad (7)$$

where Σ_1 is the optimal channel input covariance matrix satisfying the power constraint, namely, $\text{Tr}(\Sigma_1) = P_1$. To simplify the presentation, we rewrite (6) in the equivalent real channel representation by concatenating the real and imaginary parts of the complex vectors for all T symbols

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{s} + \mathbf{n}, \quad (8)$$

where $\mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_T^T)^T$ with $\mathbf{x}_t^T = [\text{Re}\{\mathbf{x}_{1,t}^c\}^T, \text{Im}\{\mathbf{x}_{1,t}^c\}^T]^T$. The noncasually known side-information at the transmitter \mathbf{s} , and the additive noise $\mathbf{n} \sim N_{\mathbb{R}}(0, \frac{1}{2}\mathbf{I}_{2NT})$, are the real channel equivalents of the multi-user interference $\mathbf{H}^c\mathbf{x}_{2,t}^c$ and receiver noise \mathbf{n}_t^c , respectively, obtained similarly as \mathbf{x} from $\mathbf{x}_{1,t}^c$. The $2NT \times 2MT$ block-diagonal real channel matrix \mathbf{H} consists of the same $2N \times 2M$ diagonal-block repeated T times

$$\mathbf{H} = \mathbf{I}_T \otimes \left(\begin{bmatrix} \text{Re}\{\mathbf{H}^c\} & -\text{Im}\{\mathbf{H}^c\} \\ \text{Im}\{\mathbf{H}^c\} & \text{Re}\{\mathbf{H}^c\} \end{bmatrix} \right),$$

where \otimes denotes the Kronecker product.

C. Vector Superposition Dirty Paper Coding

Unlike the channel in [18], the channel in (8) is the Gaussian ‘‘colored-paper’’ channel [19] with the transmission power constraint P_1T . Following the approach in [18], we consider a vector DPC for this channel consisting of a $(2^{TR_s}, m)$ shaping code C_s , which makes the transmitted signal meet the total energy constraint P_1T , and a $(2^{TR_c}, m)$ channel code C_c , which boosts up the code rate R_c to C_{MIMO}^s in (7). Here we use the notation m to replace $2MT$ for conciseness. Before introducing the encoding/decoding procedures, we first define some useful operations. For a vector \mathbf{g} , the mod- A operation $(\mathbf{g} \bmod A) = \mathbf{g}'$ is applied componentwise to each \mathbf{g}_i such that $\mathbf{g}'_i = \mathbf{g}_i - Q_A(\mathbf{g}_i) \forall i$, where $Q_A(\mathbf{g}_i)$ is the nearest multiple of A to \mathbf{g}_i . The modulo operation associated with the shaping code, $\lfloor \mathbf{g} \rfloor_{C_s}$, is defined as $\mathbf{g} - Q_{C_s}(\mathbf{g})$, where the nearest neighbor quantizer $Q_{C_s}(\mathbf{g})$ associated with C_s is defined by

$$Q_{C_s}(\mathbf{g}) = \mathbf{c} \in C_s \quad \text{if } \|\mathbf{g} - \mathbf{c}\|^2 \leq \|\mathbf{g} - \mathbf{c}'\|^2, \quad \forall \mathbf{c}' \in C_s$$

where $\|\cdot\|$ denotes the Euclidean norm.

Encoder: The encoder selects a codeword $\mathbf{c}_c \in C_c$ according to the message index and transmits the vector

$$\mathbf{x} = \lfloor (\mathbf{c}_c - \mathbf{W}\mathbf{s} - \mathbf{u}) \bmod A \rfloor_{C_s} \bmod A,$$

where \mathbf{u} , uniformly distributed in the m -dimensional cube $[-A/2, A/2]^m$ and independent of the channel, is a dither signal known to both the transmitter and the receiver. The role of the matrix \mathbf{W} will be clarified when we discuss the design of the decoder.

As in [18], the transmitted signal \mathbf{x} can be expressed as

$$\mathbf{x} = (\mathbf{c}_s + \mathbf{c}_c - \mathbf{W}\mathbf{s} - \mathbf{u}) \bmod A, \quad (9)$$

where $\mathbf{c}_s = Q_{C_s}(-\mathbf{c}_c + \mathbf{W}\mathbf{s} + \mathbf{u} \bmod A)$. Note that the transmitted signal is the superposition of codewords from the two codes C_s and C_c , therefore we call the coding scheme vector superposition DPC. The superposition codeword $\mathbf{c}_s + \mathbf{c}_c$ belongs to the coset of C_s and is consistent with the design guidelines in [6]. The transmitter should make the transmitted signal \mathbf{x} meet the energy constraint P_1T , which can be done by properly choosing the code rate R_s of C_s as will be shown in III-D.1.

Decoder: After passing \mathbf{x} through the channel in (8), the decoder first performs some signal processing on the received signal to get

$$\begin{aligned} \hat{\mathbf{y}} &= (\mathbf{W}\mathbf{y} + \mathbf{u}) \bmod A \\ &= (\mathbf{c}_s + \mathbf{c}_c + \mathbf{e}) \bmod A, \end{aligned} \quad (10)$$

where $\mathbf{e} = -(\mathbf{I} - \mathbf{W}\mathbf{H})\mathbf{x} + \mathbf{W}\mathbf{n}$ and the second equality comes from (9) and the distributive property of the mod- A operation, and \mathbf{W} acts as an equalizer. Successive decoding is performed on $\hat{\mathbf{y}}$. That is, we first decode $\hat{\mathbf{c}}_s$ from $\hat{\mathbf{y}}$, then subtract it from $\hat{\mathbf{y}}$ and decode the desired codeword $\hat{\mathbf{c}}_c$ from $(\hat{\mathbf{y}} - \hat{\mathbf{c}}_s) \bmod A$. In the sequel, we will demonstrate the procedures to obtain the highest R_c possible.

D. Code Parameters Selection and the Interference-Free Rate Achievement

We now perform the random-coding analysis to find the design parameters R_s , \mathbf{W} and A that can make R_c equal to C_{MIMO}^s with an arbitrary low probability of error, when the energy constraint P_1T is satisfied. The main difference between our analysis and that of [18] is that the random codebooks must now satisfy certain covariance properties optimized according to the channel gain matrix. In other words, the codewords are vectors whose components are no longer i.i.d. To that end, we define the following m -dimensional random vectors:

- U_A^m is a random vector uniformly distributed over $[-A/2, A/2]^m$, which is the Voronoi region [20] of the m -dimensional cubic lattice $A\mathbf{Z}^m$.
- X_G^m is a real zero-mean colored Gaussian vector with covariance matrix

$$\Sigma_G = \frac{1}{2}\mathbf{I}_T \otimes \left(\begin{bmatrix} \text{Re}\{\Sigma_1\} & -\text{Im}\{\Sigma_1\} \\ \text{Im}\{\Sigma_1\} & \text{Re}\{\Sigma_1\} \end{bmatrix} \right),$$

where Σ_1 is defined in (7)

- U_c^m is also a real zero-mean colored Gaussian vector with covariance matrix $\mathbf{W}\mathbf{H}\Sigma_G$. The choice of this covariance will be clear later.
- All other capital random vectors represent their corresponding signals (denoted in lower case previously) in the random-coding analysis. For example, X^m represents \mathbf{x} .

In the above, the first three random vectors are independent of one another and the the interference S^m , which is not necessarily Gaussian as in [19]. The random codebook for C_c is formed by generating the 2^{TR_c} codewords according to the same distribution as U_c^m , with each codeword generated independently. Similarly, the random codebook for C_s is obtained by generating the 2^{TR_s} codewords according to a distribution to be determined.

1) *Parameters Selection for the Encoder:* We first select R_s that meets the power constraint, namely, $\mathbb{E}\|X^m\|^2 \leq P_1T$. From (9), we have, equivalently,

$$\mathbb{E}\|((-U_c^m + \mathbf{W}S^m + U_A^m) \bmod A - U_s^m) \bmod A\|^2 \leq P_1T, \quad (11)$$

where U_s^m denotes a codeword from the random codebook for C_s . Using the definition of \mathbf{c}_s , this inequality gives the maximum allowed distortion for the vector quantizer output U_s^m when the input is $(-U_c^m + \mathbf{W}S^m + U_A^m) \bmod A$. Thus the rate distortion pair (R_s, P_1T) for the shaping code C_s must be achievable as defined in [21].

To this end, we choose U_s^m as

$$U_s^m = [(-U_c^m + \mathbf{W}S^m + U_A^m) + X_G^m] \bmod A, \quad (12)$$

and specify the corresponding code rate R_s as

$$R_s = \frac{1}{T}I(U_s^m; -U_c^m + \mathbf{W}S^m + U_A^m \bmod A), \quad (13)$$

where $I(\cdot)$ denotes the mutual information between random vectors. And we have the following result

Lemma 1: The rate distortion pair (R_s, P_1T) of the shaping code C_s with the mod- A distortion measure (given in (11)) is achievable.

Proof: First note that the mod- A distortion measure in (11) is additive, thus the rate-distortion theory in [22, Sec. 4.4] can be applied and we know that R_s must satisfy

$$R_s > \min_{U_s^m : U_s^m \in \mathbb{D}_s} \frac{1}{T}I(U_s^m; -U_c^m + \mathbf{W}S^m + U_A^m \bmod A), \quad (14)$$

where \mathbb{D}_s is the set of all possible quantizer outputs U_s^m whose joint distribution with the quantizer inputs satisfy (11). We first focus on the satisfaction of the distortion criterion. Substituting U_s^m in (11) with (12), we have $X^m = (X_G^m \bmod A)$ and the distortion is equal to $\mathbb{E}\|X_G^m \bmod A\|^2$. According to the following inequality obvious from the definition of the mod- A operation

$$\mathbb{E}\|X_G^m \bmod A\|^2 \leq \mathbb{E}\|X_G^m\|^2 = \text{Tr}(\Sigma_G) = P_1 T,$$

the distortion criterion is satisfied.

Since the selection of U_s^m in (12) does not guarantee the minimization of the right hand side of (14), the rate distortion pair $(R_s, P_1 T)$ is achievable. Indeed, the distortion (11) belongs to the mod- Λ difference-distortion measure defined in [23] if $\Lambda = AZ^m$. The rate-distortion bound is achieved if X_G^m in (12) is chosen as the truncated Gaussian distribution [23]. ■

Furthermore, since the channel input X^m is equal to $(X_G^m \bmod A)$, and according to the independence between X_G^m and S^m by selection, it is independent of the multi-user interference S^m . With the parameters of the shaping code C_s designed to meet the requirements at the transmitter, we focus on the decoder in the next subsection.

2) *Parameters Selection for the Decoder and the Maximum Code Rate Calculation:* The goal of this section is to choose parameters to make the successive decoding applied on (10) work well, and boost R_c as high as possible. First, we check whether the choice of R_s can make successful decoding of \mathbf{c}_s from $\hat{\mathbf{y}}$ by treating $\mathbf{c}_c + \mathbf{e}$ as noise. We focus on the reliably decodable rate $\frac{1}{T}I(U_s^m; \hat{Y}^m)$. Surprisingly, reliable decoding is possible according to the following lemma.

Lemma 2: With A large enough and the equalizer \mathbf{W} being the linear minimum mean square error (LMMSE) equalizer of the interference-free channel

$$Y_f^m = \mathbf{H}X_G^m + N^m, \quad (15)$$

the rate R_s of the shaping code C_s equals to $\frac{1}{T}I(U_s^m; \hat{Y}^m)$, and is given by

$$R_s = 2M \log A - \log((\pi e)^M |\Sigma_1|). \quad (16)$$

Proof: To prove this lemma, we first calculate R_s in (13). Since the mod- A operation equals to the modulo- Λ operation defined in [20] with $\Lambda = AZ^m$, by observing the definitions of the dither U_A^m and U_s^m in (12), we have that U_s^m is uniformly distributed over $[-A/2, A/2]^m$ with differential entropy $m \log A$ according to Lemma 1 of [20]. Thus we have

$$\begin{aligned} R_s &= \frac{1}{T} \{h(U_s^m) - h(U_s^m | -U_c^m + \mathbf{W}S^m + U_A^m \bmod A)\} \\ &= 2M \log A - \frac{1}{T} h(X_G^m \bmod A), \end{aligned} \quad (17)$$

where $h(\cdot)$ denotes the differential entropy of its argument. The second equality of this equation comes from the independence between X_G^m and the quantizer input by definition.

We now compute $\frac{1}{T}I(U_s^m; \hat{Y})$. As A increases, $X^m = (X_G^m \bmod A)$ approaches X_G^m [24]. Also, according to the channel (10), we have

$$\hat{Y}^m = (U_s^m + U_c^m + E^m) \bmod A, \quad (18)$$

where $E^m = -(\mathbf{I} - \mathbf{W}\mathbf{H})X_G^m + \mathbf{W}N^m$. From (12), with the independence between U_c^m , S^m , U_A^m and X_G^m by definition, it is easy to show that U_s^m is independent of $U_c^m + E^m \bmod A$ by applying Lemma 1 of

[20] again. The same lemma can also be used to show from (18) that \hat{Y}^m is uniformly distributed over $[-A/2, A/2]^m$, by treating U_s^m as the dither. Therefore,

$$\begin{aligned} \frac{1}{T}I(U_s^m; \hat{Y}^m) &= \frac{1}{T}\{h(\hat{Y}^m) - h(U_c^m + E^m \bmod A | U_s^m)\} \\ &= 2M \log A - \frac{1}{T}h(U_c^m + E^m \bmod A). \end{aligned} \quad (19)$$

By the definition of E^m , the LMMSE equalizer \mathbf{W} makes $-E^m$ the LMMSE estimation error for estimating X_G^m in channel (15). By the orthogonality principle [25] we have that the covariance matrix of E^m is $(\mathbf{I} - \mathbf{W}\mathbf{H})\Sigma_G$. Furthermore, E^m is also Gaussian and independent of U_c^m . The distribution of the sum $U_c^m + E^m$ is thus Gaussian with covariance matrix $(\mathbf{I} - \mathbf{W}\mathbf{H})\Sigma_G + \mathbf{W}\mathbf{H}\Sigma_G = \Sigma_G$ which is the same as the distribution of X_G^m . From (17) and (19), we conclude that R_s is equal to $\frac{1}{T}I(U_s^m; \hat{Y}^m)$.

With A large enough, $h(X_G^m \bmod A)$ approaches $h(X_G^m)$. By using the definition of Σ_G in (17) we have (16). \blacksquare

An insight to the above derivation can be drawn by viewing R_s in (13) as the mutual information between the input and output of the test channel in source coding specified in (12). The duality between source and channel coding [21] can be invoked between the test channel and the channel between U_s^m and \hat{Y} . Since the additive noise in the test channel is the quantization error $(X_G^m \bmod A)$, the choices of code design parameters should make the additive noise in decoding \mathbf{c}_s equal to it to validate the rate equivalence. We note that this selection of the LMMSE filter \mathbf{W} is consistent with [19]. Moreover, it is equal to $(\mathbf{H}^T\mathbf{H} + \frac{1}{2}\Sigma_G^{-1})^{-1}\mathbf{H}^T$, which is block-diagonal and operates only in the spatial dimension.

We are now ready to compute the maximum achievable rate R_c of the channel code C_c . From [25], by assuming perfect decoding of \mathbf{c}_s , the maximum rate for successful decoding of \mathbf{c}_c is $\frac{1}{T}I(U_c^m; \hat{Y} | U_s^m)$. With the aid of Lemma 1 and 2, we have

Theorem 2: With A large enough, and the shaping code rate R_s and the LMMSE equalizer \mathbf{W} specified in Lemma 2, the achievable rate for successful decoding of the channel code $R_c = \frac{1}{T}I(U_c^m; \hat{Y} | U_s^m)$ is equal to the interference-free single-user MIMO capacity.

Proof: We first compute $I(U_c^m; \hat{Y} | U_s^m)$ from (18) as

$$\begin{aligned} I(U_c^m; \hat{Y} | U_s^m) &= h(\hat{Y} | U_s^m) - h(\hat{Y} | U_s^m, U_c^m) \\ &= h(U_c^m + E^m \bmod A) - h(E^m \bmod A) \\ &= h(X_G^m \bmod A) - h(E^m \bmod A). \end{aligned}$$

As in the proof of Lemma 2, the first equality comes from the independence between these random vectors, and the second equality from the fact that the distribution of $U_c^m + E^m$ is the same as X_G^m . With A large enough, the aliasing effects of the modulo operation are negligible, we have

$$I(U_c^m; \hat{Y} | U_s^m) = h(X_G^m) - h(E^m) = \frac{1}{2} \log \frac{|\Sigma_G|}{|\Sigma_E|},$$

where Σ_E is the covariance matrix of the LMMSE error $-E^m$, and this conditional mutual information is the maximum achievable rate of the backward-channel $X_G^m = \mathbf{W}Y_f^m - E^m$ from the forward-channel (15).

By invoking the information-lossless property [25] we have the maximum achievable rate

$$\frac{1}{2T} \log \frac{|\Sigma_G|}{|\Sigma_E|} = \frac{1}{T}I(X_G^m; Y_f^m) = \frac{1}{T}\{h(Y_f^m) - h(N^m)\}.$$

Using the fact that the covariance matrix of N^m in (15) is $\frac{1}{2}\mathbf{I}_{2NT}$, we arrive at

$$\frac{1}{T}I(U_c^m; \hat{Y} | U_s^m) = \frac{1}{2T} \log |\mathbf{H}(2\Sigma_G)\mathbf{H}^T + \mathbf{I}_{2NT}|.$$

It is easy to show that this rate equals to the C_{MIMO}^s in (7) with complex notations. ■

Indeed, we select the parameters to make the channel input-output mutual information in successive decoding \mathbf{c}_c equal to that of a modulo backward-channel of the MMSE estimation [25]. The capacity of the corresponding interference-free forward channel (15) is thus achievable from the information-lossless property of the MMSE-estimation.

E. Conclusion

In this section, we have proposed a vector dirty paper coding structure for the MIMO GBC channel. In particular, we extended the superposition coding approach developed by Bennatan *et al.* [18] to the vector channel. Through a theoretical random coding analysis, the resultant coding structure was shown to be able to achieve the promised rate of the vector DPC. For simplicity, in the above derivation we have focused on coding and decoding for the user assigned the interference-free rate in a two-user system. These results, however, can be easily extended to cancel the previously encoded signals for any user in a MIMO GBC system with any number of users and any DPC encoding order. Although the optimal code parameters have been determined in this section, it remains to develop practical coding and decoding schemes that can benefit from these theoretical results.

IV. AUTONOMOUS POWER ALLOCATION FOR A MULTI-CELL OFDM NETWORK AT 60 GHz BAND

In this section, we consider a coordinated, cellular-like WLAN network using OFDM and operating at the 60 GHz band. For such a network, more coordinated multiple access schemes, such as TDMA or a mixture of TDMA and Aloha with slot reservation, are employed. In order to increase frequency reuse, cells that use the same frequency band are usually tightly packed and may interfere with one another. Thus power control becomes very important to ensure that all the co-channel cells meet the required performance while not causing excessive interference to other cells. In [7], an efficient autonomous power allocation and control algorithm was proposed for multi-carrier networks employing soft decoding. It was shown that the algorithm has the advantages of fast convergence for power control, low network outage probability, and low power consumption. These advantages are a result of multi-cell water filling. In this section, we adopt the algorithm in [7] to evaluate the multi-cell WLAN performance at the 60 GHz band. The multi-cell water filling usually causes the cells to shut down ill-conditioned sub-channels, thus it may also help reduce the peak-to-average power ratio (PAPR).

A. 60 GHz Channel Characteristics

The 60 GHz band has a high path loss and high attenuation. In addition, it is susceptible to obstruction. For the WLAN application, the coverage of an access point (AP) is usually within a room. Thus, with walls in between them, the cells are quite isolated. In [26], it was shown through measurement that, considering an ensemble of different locations, the channel model regularly used for lower frequency band can still be applied to the 60 GHz band. In particular, the path loss exponent for the 60 GHz band in an office with cubicals is 3.83. The shadowing effect can also be modeled as lognormally distributed, with zero mean and standard deviation between 6 and 9 dB. The typical standard deviation in offices is 7.6 dB. According to [26], the multipath effect at the 60 GHz band can also be well modeled with the Saleh-Valenzuela (S-V) model [27] with only the first multipath cluster. The power delay profile of this cluster decays exponentially, and the mean excess delay (MED) is between 5 and 75 ns, with a typical value of 23 ns for office environments.

B. Simulation Parameters and Results

We used the following parameters for the 60 GHz multi-cell WLAN simulation: A hexagonal cellular system with frequency reuse 4 was considered in the simulation. Around a central cell, we considered up to the fourth tier. That is, 6 co-channel cells in the second tier and 12 co-channel cells in the fourth tier. Therefore a total of 19 co-channel cells were simulated. In each hexagonal cell, an AP was situated at the center. Co-channel users were generated as uniformly distributed in the corresponding cells for each simulation instance. Path distance loss obeyed a decay law with exponent 3.83, and its gain was normalized to unity at (1/100)th the distance from the base station to a cell corner. We avoided in the simulation to have users situated closer than this 1/100 reference distance. The frequency selective fading was generated using the S-V model with one cluster whose power delay profile was exponential with MED 23 ns. The shadowing standard deviation was between 6 and 9 dB. Fadings between different user-AP pairs were generated independently.

The channel coding scheme was a turbo code with $(1, 5_8/7_8, 5_8/7_8)$ constituent codes. So the overall code rate was $1/3$. The interleaver length of the turbo code was 1000 bits. After the encoder, a random channel interleaver with a block length of 3000 coded bits was placed before the OFDM modulator. We defined the required $E_b/N_0 = 4$ dB, which guaranteed the BER smaller than 10^{-6} . Thus, the corresponding constant $B = \frac{1}{N} \sum_{k=1}^N \exp\left(-\frac{G_k P_k}{I_k}\right)$ (N is the number of sub-channels, k is the sub-channel index, G is the channel gain, P is the transmitted power, and I is the interference power) in the following distributed soft decoding assisted power allocation (DSDPA) algorithm was 0.4329 [7].

DSDPA:

For every cell i at every iteration n :

- 1) Measure the sub-channel signal-to-interference ratios (SIR) $\gamma_{i,k}$, gains $G_{i,k}$ and noise powers $\frac{1}{2}I_{i,k}$; form the effective noise powers $I'_{i,k} = \frac{I_{i,k}}{G_{i,k}}$. Let the set of inactive sub-channels $f_i = \emptyset$, and n_{f_i} be the number of elements in this set.
- 2) Set target SIR $\Gamma_{i,k}$ according to

$$\Gamma_{i,k} = \begin{cases} 2 \left(\log \left(\frac{\sum_{l \notin f_i} I'_{i,l}}{BN - n_{f_i}} \right) - \log I'_{i,k} \right), & k \notin f_i \\ 0, & k \in f_i \end{cases}.$$

If there is any $\Gamma_{i,k} < 0$, add the active sub-channel with maximum $I'_{i,k}$ to the inactive set f_i and repeat. Otherwise go to 3.

- 3) Power control:

$$P_{i,k}(n+1) = P_{i,k}(n) \frac{\Gamma_{i,k}}{\gamma_{i,k}},$$

which means that the receiver sends the power increment command $\frac{\Gamma_{i,k}}{\gamma_{i,k}}$ (in dB) back to the transmitter, and the transmitter adjusts the power according to this command. Note that in case $P_{i,k}(n) = 0$, the actual power requirement is transmitted in the next iteration. That is,

$$P_{i,k}(n+1) = \frac{\Gamma_{i,k} I'_{i,k}}{2}.$$

A maximum average transmission power constraint of 10 W per cell was set in order to preclude the situations that need excessive transmission power, while still being high enough not to affect capturing the convergence properties given by the channel. In reality, the maximum power constraint is defined according to hardware limitations. In that case, numerical results will alter accordingly, the trends shown here however remain true. On each sub-channel, BPSK modulation with Nyquist pulse shape was used. The receiver AWGN had $N_0 = -120$ dBm W/Hz and the transmission bandwidth was 200 MHz.

We compare the DSDPA and a conventional scheme which has equal power allocation for all sub-channels. The distributed equal power (DEP) allocation used binary search in the range: $\frac{I_m}{G_m} \log \frac{1}{B} \leq P \leq \frac{I_n}{G_n} \log \frac{1}{B}$ (m, n are the indices of the sub-channels having the lowest and the highest effective noise powers ($\frac{I_k}{G_k}$), respectively.) to find the required transmission power and hence the target SIRs, then applied the same closed-loop power control algorithm as in the DSDPA.

To reduce the amount of power control feedback, the number of sub-channels considered was 32. On each of the sub-channels we assumed that the fading was flat and fixed within the period of convergence. Since the transmission bandwidth is very wide, it is possible to have more OFDM sub-channels to minimize the throughput loss due to the guard time. In that case, each sub-channel considered here can be seen as a group of consecutive sub-channels. Since the sub-channels within a group are closely spaced, they are often subject to very similar channel gains, thus will be allocated the same power.

In the simulation, the algorithms were executed independently and synchronously by the co-channel cells. 3000 sample instances were generated for each fading condition. For fair comparison, each time the two algorithms started with the same equal power allocation which guaranteed the required BER in the absence of co-channel interference. The convergence was considered reached whenever the performance measure B of all cells were within $\pm 0.1\%$ of the target and the transmission powers of individual cells did not exceed the maximum power constraint. We noticed through simulation that the DSDPA sometimes oscillated between different power allocations smaller than the maximum constraint, therefore a maximum number of 30 iterations was set to terminate each simulation instance. Based on the requirements and constraints defined, an outage (nonconvergence) was claimed whenever the target B could not be achieved by all co-channel cells within 30 iterations; or the target achieving power for any of the co-channel cells

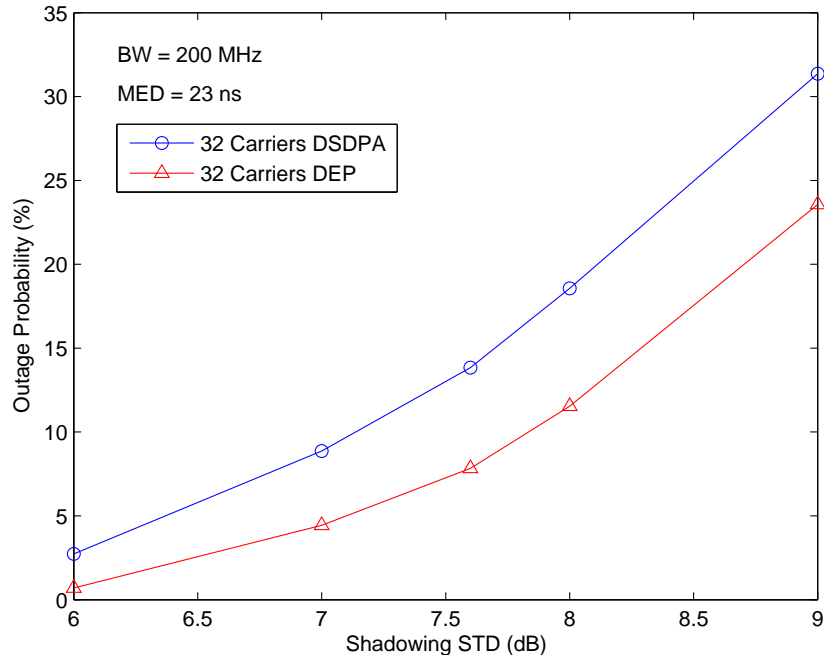


Fig. 3. Outage probability vs. shadowing standard deviation.

exceeded the maximum power constraint. Except for the outage probability, other performance results were averaged only over the instances where outage was not claimed.

Fig. 3 shows the outage probabilities of the two schemes for various shadowing standard deviation. Although it is shown that DEP has 2% to 7% lower outage probability than the DSDPA, we note that the DSDPA sometimes oscillates between feasible power allocations. Simulation showed that the DSDPA had an oscillation probability ranging from 3% to 15%, while the DEP had zero oscillation probability. This suffices to show that the DSDPA actually has a lower outage probability than the DEP.

Fig. 4 shows the relation between the average power consumption per cell and the degree of shadowing. As the fading becomes worse, more power is required. It is shown that the DSDPA always saves about 1 dB power over the DEP.

In the following, we consider the situation where a peak power constraint is imposed to avoid the nonlinear effect of the power amplifier. As the average transmission power constraint is 10 W, and the number of sub-channels is 32, we can easily see that the peak power subject to the average power constraint will never exceed 320 W. In reality, this peak may be much larger than the linear range of the power amplifier. In the sequel, we define a factor α such that, in addition to the average power constraint, a peak power constraint of $10/\alpha$ is also imposed on the algorithms. A larger α will apparently set a lower peak power constraint.

In Fig. 5 the outage probabilities with the peak power constraint are shown. It is clear from this figure that a lower peak power constraint causes a higher outage probability. Again, DSDPA has higher outage probabilities due to oscillation.

To see the effect of the peak power constraint, we plot in Fig. 6 the portion of the outage probability caused by the peak power constraint. It is now clear that the DSDPA has lower peak power, as a result of the multi-cell water filling.

In Fig. 7 we show another experiment in which only the peak power constraint is applied, and the average power constraint is removed. From this figure, it is again shown that DSDPA has a much lower outage probability than DEP.

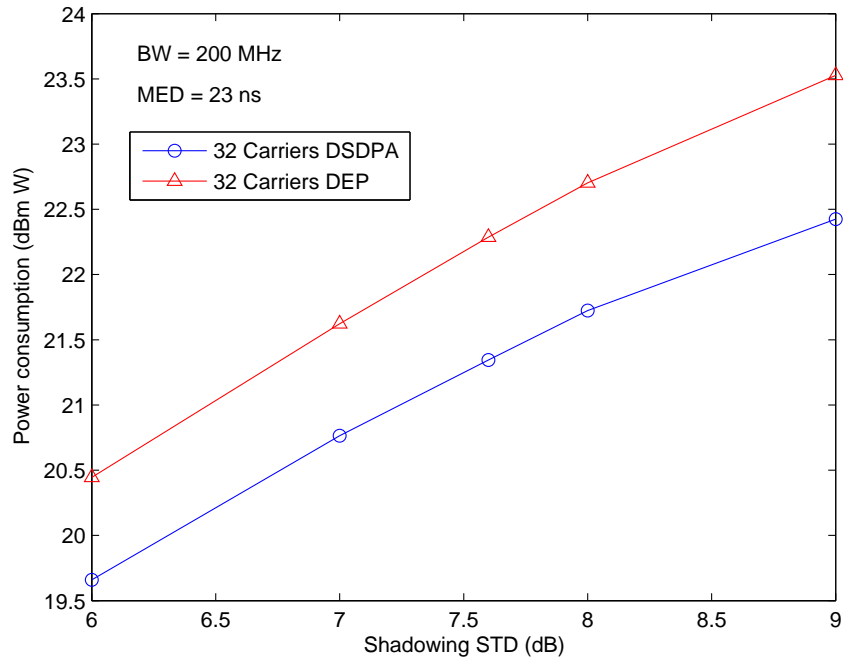


Fig. 4. Power consumption per cell vs. shadowing standard deviation.

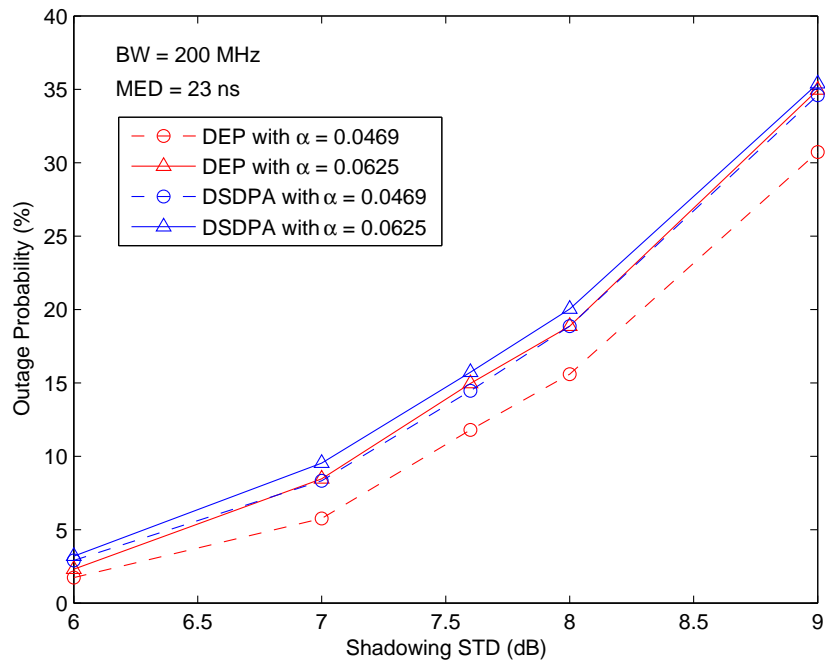


Fig. 5. Outage probability vs. shadowing standard deviation (with both the peak power constraint and the average power constraint).

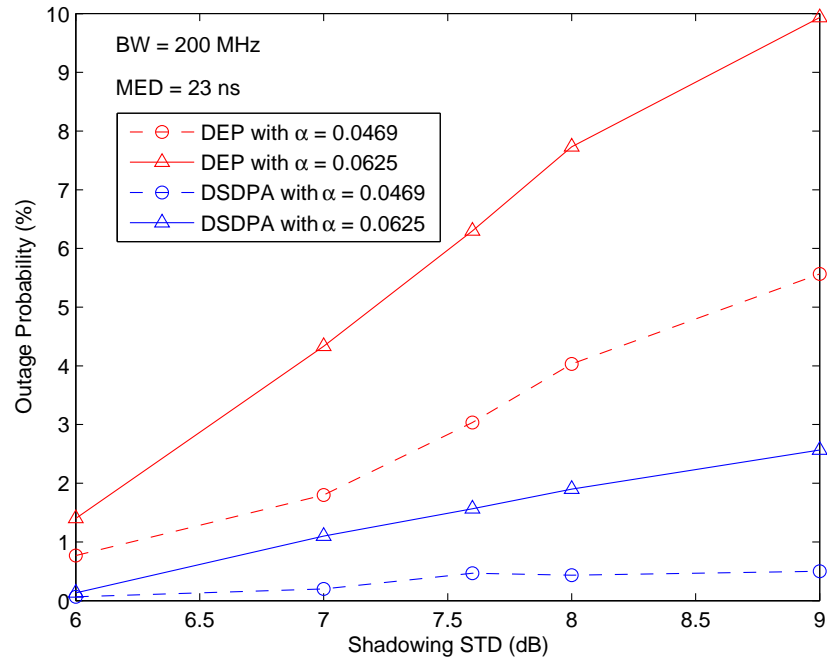


Fig. 6. Outage probability caused by the peak power constraint vs. shadowing standard deviation.

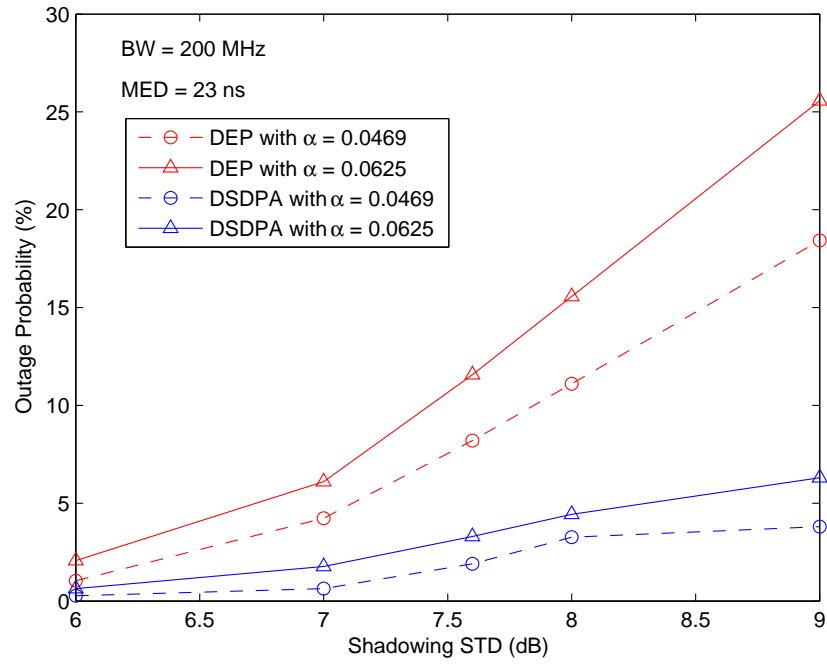


Fig. 7. Outage probability vs. shadowing standard deviation (with the peak power constraint but not the average power constraint).

V. ADAPTIVE THROUGHPUT MAXIMIZATION FOR ADAPTIVE MODULATION AND CODING SYSTEMS WITH HYBRID ARQ

In this section, we consider an adaptive threshold adjustment algorithm that can adaptively maximize the throughput for AMC systems with HARQ. For simplicity, the algorithm considers only single-antenna channels. Proper channel condition metric can be developed to apply this algorithm to MIMO systems.

A. Background

Contrary to the traditional voice application that has both delay and quality constraints, data applications are usually more tolerant of delay, thus allowing exploitation of the theoretical time-domain water-filling benefit. The most common water-filling approach is AMC which adjusts the transmission power, modulation and coding rate according to the instantaneous channel condition to maximize the average throughput. In new wireless systems such as the High Speed Downlink Packet Access (HSDPA), AMC is jointly implemented with HARQ [28] to offer more efficient transmission.

One particular challenge for AMC is appropriate selection of modulation and coding schemes (MCS). For single-antenna systems, MCSs are usually selected according to the SNR. Nonetheless, determination of the MCS switching thresholds is a complex process that requires the knowledge of the dependence of each MCS's performance on factors such as SNR, channel time-varying rate, system configuration and impairments, etc. As such, it is quite difficult to determine the thresholds analytically. In addition, as the channel statistics may vary with time, the optimal thresholds also change with time. These practical considerations make an adaptive threshold adjustment algorithm desirable.

Adaptive threshold adjustment has been considered previously for AMC only. And the goal of the adjustment was mostly to meet certain error rate requirements. Specially, a stochastic learning automaton was proposed to adaptively select, among a number of given values, the best value for each threshold [29]. Taking a more effective approach, [30] proposed to increase or decrease the thresholds according to the frame acknowledgement feedback. And the amounts to increase or decrease were derived according to certain error rate conditions. [31] further proposed modifications aiming to overcome the problem of unbalanced adjustment seen in the algorithm in [30]. As will be shown in this section, the error rate conditions used in [30] and [31] do not necessarily reflect the actual error rates (and throughputs) achieved by the MCSs. In addition, the algorithm in [31] may react too slowly to the change in channel characteristics. Through an analysis that also considers HARQ and the channel statistics, we provide further conditions and modifications that can mitigate these problems and offer more precise control of the system.

Since the new adaptive threshold adjustment algorithm can control precisely the error rates (and throughput) of the MCSs, we can combine it with an adaptive error rates setting algorithm that can maximize the total throughput. In this section, such an adaptive throughput maximization mechanism is also provided.

B. System Model and Notations

The system considered includes a transmitter, a receiver, and a single-input single-output forward channel and a feedback channel between them. The receiver keeps measuring the received SNR, and reporting back to the transmitter. We assume that the transmission power is constant, while the transmitter can select different MCSs for different transmission time intervals (TTI) according to the SNR feedback. In each TTI, a decodable and error-detectable coded frame is transmitted. Upon receiving a frame, the receiver decodes the frame and checks whether the decoding is successful. If so, it sends a positive acknowledgement (ACK) to the transmitter. Otherwise a negative acknowledgement (NACK) is sent. Both the SNR and ACK/NACK feedbacks may be subject to processing and propagation delay.

For simplicity, we assume a constant TTI, \mathcal{L} , for all MCSs and initial/re-transmissions, and the retransmissions of a frame use the same MCS as the initial transmission.¹ Assume that there are \mathcal{M} possible MCSs with nominal transmission rates $\mathcal{R}_1, \dots, \mathcal{R}_{\mathcal{M}}$, in increasing order. The MCS having rate \mathcal{R}_i , denoted MCS_{*i*}, is selected if the SNR feedback falls in the region Λ_i . With careful design of the nominal transmission rates and the error rate requirements, it is possible to arrive at a simplified situation where $\Lambda_i \equiv [\Gamma_{i-1}, \Gamma_i)$, $i = 1, \dots, \mathcal{M}$, with $\Gamma_0, \dots, \Gamma_{\mathcal{M}}$, in increasing order, being the thresholds for selecting MCS. No matter whether retransmission is allowed, and whether frame combining (such as Chase combining and incremental redundancy [28]) is implemented for HARQ, for a given MCS_{*i*} that has been used in N_i TTIs, assuming that the ACK/NACK feedbacks are error-free, we have the associated average throughput

$$\mathcal{T}_i = \frac{\mathcal{R}_i \mathcal{L} N_{\mathcal{A}_i}}{\mathcal{L} N_i} = \mathcal{R}_i \frac{N_{\mathcal{A}_i}}{N_i}, \quad (20)$$

where $N_{\mathcal{A}_i}$ is the number of ACKs generated for the N_i MCS_{*i*} transmissions. Define the event that an ACK is generated for a TTI using MCS_{*i*} as \mathcal{A}_i , and the corresponding NACK event as $\bar{\mathcal{A}}_i$. Clearly, the probability $P_{\mathcal{A}_i} \equiv \lim_{N_i \rightarrow \infty} (N_{\mathcal{A}_i}/N_i)$ determines the efficiency of MCS_{*i*}. In particular, for systems that do not allow retransmission, $P_{\mathcal{A}_i} = 1 - P_{\bar{\mathcal{A}}_i} = 1 - P_i$, where P_i is the average frame error rate (FER) of MCS_{*i*}. In general, $P_{\mathcal{A}_i}$ depends on the SNR and channel variation rate encountered by each retransmission, as well as the HARQ, modulation, coding parameters, and the SNR distribution. In any event, the SNR feedback right before the initial transmission of a frame, denoted as $\hat{\rho}$, will be the major factor based on which MCS is selected. Let $p_{\bar{\mathcal{A}}_i}(\hat{\rho})$ be the probability of a NACK generated for a TTI given $\hat{\rho}$ for the corresponding frame, and $p_{\mathcal{A}_i}(\hat{\rho})$ be the corresponding ACK probability, we have

$$P_{\bar{\mathcal{A}}_i} \equiv 1 - P_{\mathcal{A}_i} = E [p_{\bar{\mathcal{A}}_i}(\hat{\rho}) | \hat{\rho} \in \Lambda_i] = \frac{\int_{\hat{\rho} \in \Lambda_i} p_{\bar{\mathcal{A}}_i}(\hat{\rho}) dF(\hat{\rho})}{\int_{\hat{\rho} \in \Lambda_i} dF(\hat{\rho})},$$

where $F(\hat{\rho})$ is the cumulative distribution function of $\hat{\rho}$. The goal of the adaptive threshold adjustment algorithm is to maximize the throughput while maintaining certain error rate requirement β_i for MCS_{*i*} such that $P_{\bar{\mathcal{A}}_i} = \beta_i$, $i = 1, \dots, \mathcal{M}$.

C. Adaptive Threshold Adjustment Subject to Error Rate Requirements

Since $p_{\bar{\mathcal{A}}_i}(\hat{\rho})$'s are all decreasing functions of $\hat{\rho}$, and $\mathcal{R}_1 < \mathcal{R}_2 < \dots < \mathcal{R}_{\mathcal{M}}$, maximizing the throughput (by letting $\Gamma_{\mathcal{M}} = \infty$) subject to the error rate requirements determines the thresholds almost uniquely (with the exception that when a threshold falls in an SNR region of zero probability, it can assume any value in that region without changing the error rates and throughput). To derive an algorithm that adaptively finds these thresholds, we take the independent adjustment approach in [30], meaning ACK/NACK for MCS_{*i*} only affects Λ_i . An event $\bar{\mathcal{A}}_i$ implies that $P_{\bar{\mathcal{A}}_i}$ may be too high. Therefore, Λ_i should be extended toward higher $\hat{\rho}$ and/or shrunken from the lower end. Thus we may make the following adjustments

$$\text{If } \bar{\mathcal{A}}_i : \begin{cases} \Gamma_{i-1} \leftarrow \Gamma_{i-1} + \Delta_{\bar{\mathcal{A}}_i, i-1} \\ \Gamma_i \leftarrow \Gamma_i + \Delta_{\bar{\mathcal{A}}_i, i} \end{cases} \quad (21a)$$

$$(21b)$$

where $\Delta_{\bar{\mathcal{A}}_i, i-1}$ and $\Delta_{\bar{\mathcal{A}}_i, i}$ are nonnegative step sizes to be determined. On the opposite, for an event \mathcal{A}_i we may make the following adjustments

$$\text{If } \mathcal{A}_i : \begin{cases} \Gamma_{i-1} \leftarrow \Gamma_{i-1} - \Delta_{\mathcal{A}_i, i-1} \\ \Gamma_i \leftarrow \Gamma_i - \Delta_{\mathcal{A}_i, i} \end{cases} \quad (22a)$$

$$(22b)$$

¹Note that there are some recent applications that use variable TTIs and/or MCSs for retransmissions. By letting \mathcal{L} be the construction unit of TTIs, the methodology in this section can be applied to systems with variable TTIs. In that case, N_i in (20) is the total number of such time units used by MCS_{*i*}. Each time a NACK for a transmission with TTI equal to $n\mathcal{L}$ is generated, n $\bar{\mathcal{A}}_i$'s are triggered. On the other hand, if an ACK is generated and the corresponding frame's initial TTI, which determines the amount of data contained in that frame, was $p\mathcal{L}$, p \mathcal{A}_i 's are triggered, along with some compensating adjustments which make the total number of \mathcal{A}_i 's and $\bar{\mathcal{A}}_i$'s equal to the number of \mathcal{L} 's occupied by that frame. In addition, if retransmissions can use different MCSs, there should be a separate threshold adjustment mechanism for each retransmission's MCS selection.

where $\Delta_{\mathcal{A}_i, i-1}$ and $\Delta_{\mathcal{A}_i, i}$ are also nonnegative.

These adjustments give rise to an average up-adjustment per TTI, $Pr(\Lambda_i)P_{\bar{\mathcal{A}}_i}\Delta_{\bar{\mathcal{A}}_i, i-1} + Pr(\Lambda_{i-1})P_{\bar{\mathcal{A}}_{i-1}}\Delta_{\bar{\mathcal{A}}_{i-1}, i-1}$, for threshold Γ_{i-1} , and a corresponding down-adjustment $Pr(\Lambda_i)P_{\mathcal{A}_i}\Delta_{\mathcal{A}_i, i-1} + Pr(\Lambda_{i-1})P_{\mathcal{A}_{i-1}}\Delta_{\mathcal{A}_{i-1}, i-1}$, where $Pr(\Lambda_i) \equiv \int_{\hat{\rho} \in \Lambda_i} dF(\hat{\rho})$ is the probability of using MCS_{*i*}. In the steady state, and when the error rate requirements are met, the adjustments for all thresholds should balance out. A necessary condition for $\Delta_{\bar{\mathcal{A}}_i, i-1}$, $\Delta_{\bar{\mathcal{A}}_i, i}$, $\Delta_{\mathcal{A}_i, i-1}$, $\Delta_{\mathcal{A}_i, i}$ is

$$\begin{aligned} Pr(\Lambda_i)\beta_i\Delta_{\bar{\mathcal{A}}_i, i-1} + Pr(\Lambda_{i-1})\beta_{i-1}\Delta_{\bar{\mathcal{A}}_{i-1}, i-1} = \\ Pr(\Lambda_i)(1-\beta_i)\Delta_{\mathcal{A}_i, i-1} + Pr(\Lambda_{i-1})(1-\beta_{i-1})\Delta_{\mathcal{A}_{i-1}, i-1}. \end{aligned} \quad (23)$$

As the steady state $Pr(\Lambda_i)$'s are unknown a priori, they should be avoided in determining the step sizes. We may let

$$\Delta_{\mathcal{A}_i, i-1} = \frac{\beta_i}{1-\beta_i}\Delta_{\bar{\mathcal{A}}_i, i-1}, \quad (24)$$

$$\Delta_{\mathcal{A}_i, i} = \frac{\beta_i}{1-\beta_i}\Delta_{\bar{\mathcal{A}}_i, i}, \quad \forall i. \quad (25)$$

Then (23) is always satisfied no matter what the steady state $Pr(\Lambda_i)$'s are. In the situation where the steady state thresholds all fall in the support of the SNR distribution, conditions (24) and (25), together with either of the conditions for steady boundary thresholds: $P_{\bar{\mathcal{A}}_{\mathcal{M}}}\Delta_{\bar{\mathcal{A}}_{\mathcal{M}}, \mathcal{M}} = P_{\mathcal{A}_{\mathcal{M}}}\Delta_{\mathcal{A}_{\mathcal{M}}, \mathcal{M}}$ for $\Gamma_{\mathcal{M}}$, and $P_{\bar{\mathcal{A}}_1}\Delta_{\bar{\mathcal{A}}_1, 0} = P_{\mathcal{A}_1}\Delta_{\mathcal{A}_1, 0}$ for Γ_0 , suffice to make all MCSs meet the error rate requirements. In reality, $\Gamma_{\mathcal{M}}$ (and/or Γ_0) may lie outside the support of the SNR distribution. In that case, $P_{\bar{\mathcal{A}}_{\mathcal{M}}}$ ($P_{\bar{\mathcal{A}}_1}$) might reach a steady state higher (lower) than $\beta_{\mathcal{M}}$ (β_1) even if the boundary threshold $\Gamma_{\mathcal{M}}$ (Γ_0) kept moving up (down). The steady $P_{\bar{\mathcal{A}}_{\mathcal{M}}}$ ($P_{\bar{\mathcal{A}}_1}$) might in turn make the interior thresholds steady without meeting the error rate requirements. This is especially true considering that, in practice, we usually let $\Gamma_{\mathcal{M}} = \infty$ to maximize the throughput, and $\Gamma_0 = -\infty$ to enforce nonstop transmission. Thus, it will be better if each threshold is controlled by only one MCS. Then a threshold will become steady only when the controlling MCS meets the error rate requirement. In other words, using (21a) and (22a) alone, with positive $\Delta_{\bar{\mathcal{A}}_i, i-1}$ and $\Delta_{\mathcal{A}_i, i-1}$ that satisfy (24), can offer better control of the error rates.

However, if we only use (21a) and (22a) for an MCS, its upper threshold will never change if the next higher MCS is never used. This may cause the upper threshold to stay too high up and prevent immediate use of the next higher MCS when the channel condition suddenly becomes better. To enhance the adaptability, we can add the following conditional adjustment:

$$\text{If } \mathcal{A}_i \text{ and } Pr(\Lambda_{i+1}) = 0: \quad \Gamma_i \leftarrow \Gamma_i - \Delta_{\mathcal{A}_i, i} \quad (26)$$

where $Pr(\Lambda_i)$'s can be obtained through real-time measurement at the transmitter, and $\Delta_{\mathcal{A}_i, i}$ is positive. We can, for example, let $\Delta_{\mathcal{A}_i, i} = \Delta_{\mathcal{A}_{i+1}, i}$. Although this conditional adjustment violates (23), it happens only when $Pr(\Lambda_{i+1}) = 0$, and will only slightly affect the throughputs of MCS_{*i*} and MCS_{*i+1*} when MCS_{*i+1*} is selected as an immediate consequence of it. Once this happens, $Pr(\Lambda_{i+1})$ becomes positive again, and the algorithm gets back to its normal operation. In addition, in the extreme case when the channel is static, the constant SNR may be too good for MCS_{*i*} to keep its error rate at the requirement, but not good enough for MCS_{*i+1*}. In this situation the lower threshold of MCS_{*i*} will keep decreasing. When the channel statistics change, it may take a while to bring this threshold back. This problem can be alleviated with the following condition:

$$\text{If } \mathcal{A}_i \text{ and } \sum_{j < i} Pr(\Lambda_j) = 0: \quad \Gamma_{i-1} \text{ remains unchanged.} \quad (27)$$

In summary, we have the following single-ended adjustment (SEA) algorithm: with the step sizes satisfying (24) and $\Delta_{\mathcal{A}_i, i} = \Delta_{\mathcal{A}_{i+1}, i}$, if $\bar{\mathcal{A}}_i$, then (21a); if \mathcal{A}_i , then (22a) with condition (27), and conditionally (26).

In [30] and [31], similar adjustment algorithms were used, but both the error rate requirements and threshold adjustments were applied only to the initial transmission of a frame. Apparently, for a system

that allows retransmissions, the error rate of the initial transmission does not necessarily tell what the throughput might be. Even if we generalize the algorithms in [30] and [31] according to our analysis (i.e., set requirements for $P_{\bar{\mathcal{A}}_i}$'s and make adjustments in every TTI, as done in Section V-C.1), they still do not offer precise control of the error rates. In [30], (21a), (22a) satisfying (24), and (22b) with $\Delta_{\mathcal{A}_i,i} = \Delta_{\mathcal{A}_i,i-1}$ were used. With the additional down-adjustment (22b), (23) is not satisfied for the steady state. Thus all MCS regions are moved lower, and the error rates for all MCSs are higher than the requirements. To avoid this unbalanced threshold adjustment, [31] proposed to set a narrow band around each threshold, and to decrease/increase a threshold only when the SNR lies within its band and an ACK/NACK is received. A probabilistic condition was also set to determine the step sizes for each threshold. That condition, however, is not directly related to the error rate requirement for any MCS. Therefore it does not correctly indicate the throughput of an MCS.

1) *Numerical Results:* The algorithms were tested on a HSDPA system with the following parameters. The spreading factor was 16, and one out of the 16 orthogonal codes was used. Flat Rayleigh fading with Doppler frequency 55.56 Hz (30 km/h mobility) and 10 dB average symbol SNR was assumed. Decoding errors were assumed always detected, so the ACK/NACK feedback was always correct, but had a delay of one TTI (2 ms). The instantaneous SNR measurement was assumed perfect, and its feedback delay was two TTIs. The transmission was nonstop with the assumptions of single user with infinite backlog and multiple stop-and-wait HARQ processes. Four MCSs were considered, with MCS₁ using a rate 1/2 code and QPSK, MCS₂ rate 1/2 and 16-QAM, MCS₃ rate 3/4 and 16-QAM, MCS₄ rate 3/4 and 64-QAM. All codes were turbo codes as specified in HSDPA, with Max-Log MAP decoding and eight iterations. For MCS₂ to MCS₄, the error rate requirements were 0.2. There was no requirement for MCS₁. At each TTI, if applicable, the transmitter first adjusted the thresholds, then used the adjusted thresholds to select MCS for new frames. Assumed to have no prior knowledge of the channel statistics, the transmitter set, as an example, initial thresholds $\Gamma_0 = -\infty$, $\Gamma_1 = 5$ dB, $\Gamma_2 = 6$ dB, $\Gamma_3 = 7$ dB, $\Gamma_4 = \infty$. Initially, the transmitter assumed that the MCSs had equal probabilities. Then, for each TTI that used MCS_{*i*}, the probabilities were measured at the transmitter with a moving average having memory factor 0.98: $Pr(\Lambda_i) \leftarrow 0.98Pr(\Lambda_i) + 0.02$; and $Pr(\Lambda_j) \leftarrow 0.98Pr(\Lambda_j)$, $\forall j \neq i$. These probabilities were used in conditions (26) and (27), where they were deemed equal to zero if smaller than 10^{-5} . The results shown here are for Chase combining. Other HARQ schemes with various fading speeds were also tested and similar trends observed.

Fig. 8 shows the transient behavior of the thresholds and the ACK probabilities of the MCSs, for SEA which had, as an example, $\Delta_{\bar{\mathcal{A}}_i,i-1} = 0.2$ dB, $\forall i$, and $\Delta_{\mathcal{A}_i,i-1}$ satisfying (24). The ACK probabilities $P_{\mathcal{A}_i}$ were initialized to zero, and measured with the following moving average: upon \mathcal{A}_i , $P_{\mathcal{A}_i} \leftarrow 0.98P_{\mathcal{A}_i} + 0.02$; upon $\bar{\mathcal{A}}_i$, $P_{\mathcal{A}_i} \leftarrow 0.98P_{\mathcal{A}_i}$; otherwise $P_{\mathcal{A}_i}$ remained unchanged. Also plotted in the background of Fig. 8(a) is the SNR feedback. From Fig. 8(a) we observe that for the simulated channel, MCS₁ was the most selected with $P_{\mathcal{A}_1}$ (not shown for conciseness), which depends on the SNR distribution in Λ_1 , about 0.9. The SEA effectively kept $P_{\mathcal{A}_2}$ at the target of 0.8. While MCS₃ and MCS₄ were almost never selected in the steady state because the channel did not allow them to achieve the error rate requirements.

Fig. 9 gives the performance of the unbalanced algorithm of [30] under the same conditions ($\Delta_{\bar{\mathcal{A}}_i,i-1} = 0.2$ dB). Due to the additional down-adjustment $\Delta_{\mathcal{A}_i,i}$, this algorithm has all its steady-state thresholds lower. And $P_{\mathcal{A}_2}$, which was kept at the target by the SEA, also becomes lower than the requirement. The discrepancies between $P_{\mathcal{A}_i}$'s and the requirements vary with the SNR distribution. Therefore, the specified error rate conditions are somewhat irrelevant to the actual throughput.

Fig. 10 shows the performance of the ‘‘banded’’ algorithm of [31] whose bands were ± 0.5 dB around each threshold. To be comparable with the above two algorithms, this algorithm's condition on the probability of up-adjustment for all bands was 0.2, and each up-adjustment was 0.2 dB. With adjustments only when the SNR falls in one of the bands, this algorithm adapts very slowly to the channel statistics. In addition, the probabilistic conditions specified for the adjustments also do not reflect precisely the actual $P_{\mathcal{A}_i}$'s achieved.

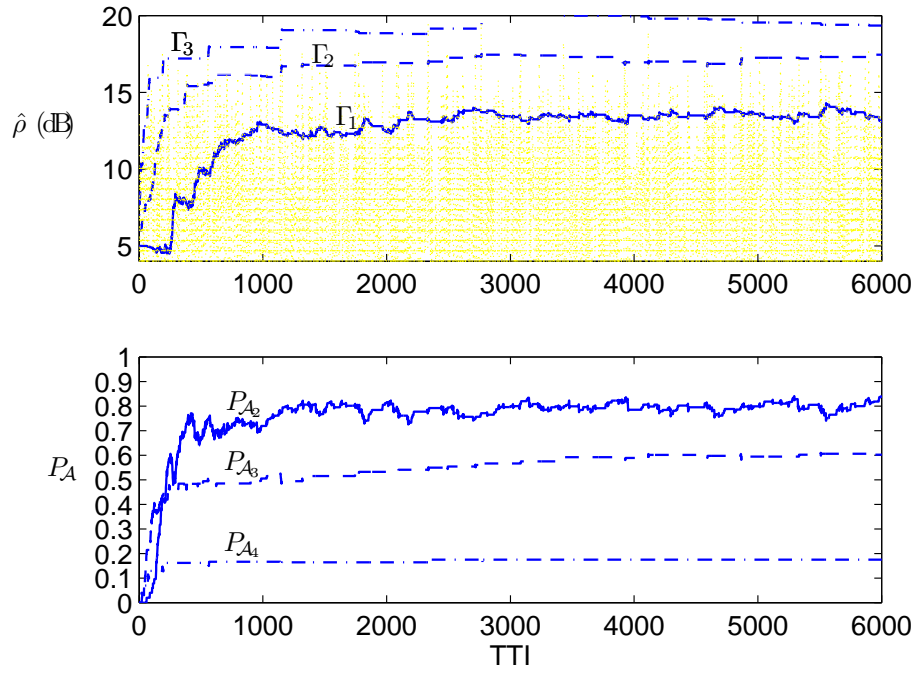


Fig. 8. The single-ended algorithm: (a) thresholds (b) probabilities of ACK.

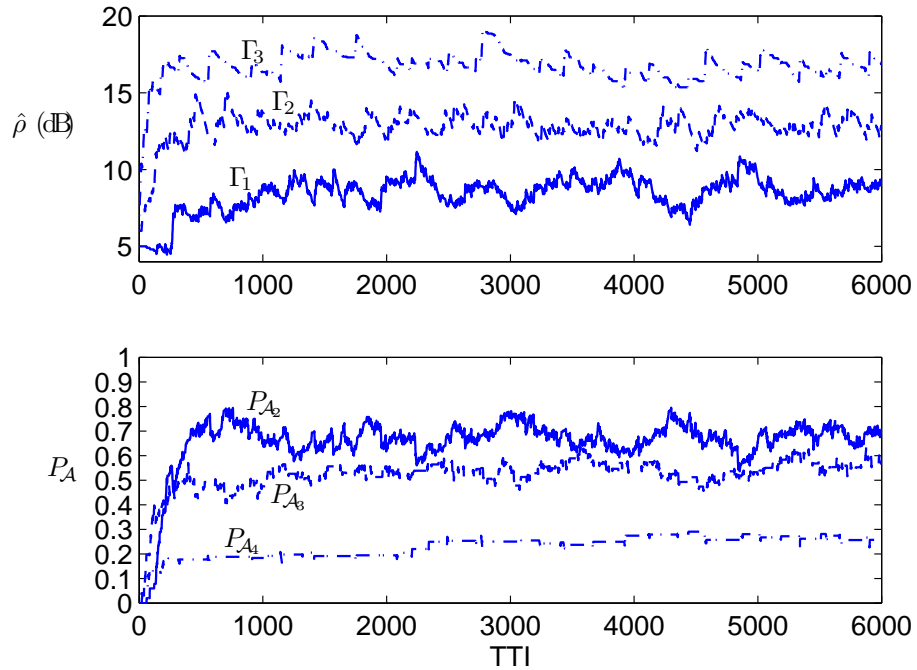


Fig. 9. The algorithm in [30]: (a) thresholds (b) probabilities of ACK.

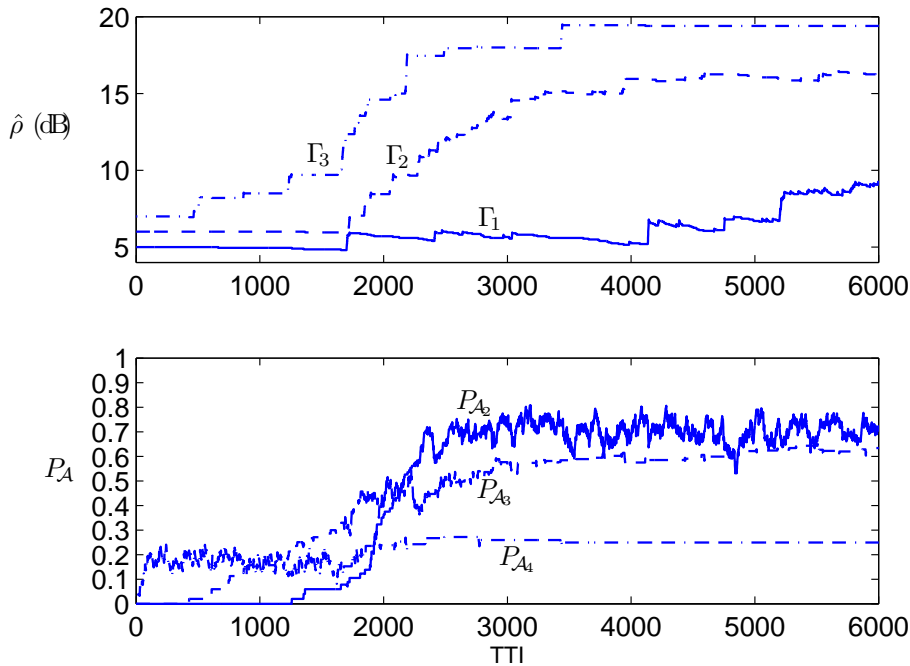


Fig. 10. The algorithm in [31]: (a) thresholds (b) probabilities of ACK.

D. Adaptive Throughput Maximization

From the above analysis, the average throughput of the system is

$$\mathcal{T} = \sum_{i=1}^{\mathcal{M}} \mathcal{R}_i \int_{\hat{\rho} \in \Lambda_i} p_{\mathcal{A}_i}(\hat{\rho}) dF(\hat{\rho}). \quad (28)$$

For any distribution of $\hat{\rho}$, it can be shown that the throughput \mathcal{T} is maximized if $\Lambda_i = \{\hat{\rho} : \mathcal{R}_i p_{\mathcal{A}_i}(\hat{\rho}) \geq \mathcal{R}_j p_{\mathcal{A}_j}(\hat{\rho}), \forall j \neq i\}$, where in case of a tie, any of the MCSs with the highest $\mathcal{R}_i p_{\mathcal{A}_i}(\hat{\rho})$ can be selected as long as the Λ_i 's are kept disjoint. Thus, the conditions for throughput maximization is

$$\mathcal{R}_i p_{\mathcal{A}_i}(\Gamma_i) = \mathcal{R}_{i+1} p_{\mathcal{A}_{i+1}}(\Gamma_i), \quad i = 1, \dots, \mathcal{M} - 1. \quad (29)$$

Since $p_{\mathcal{A}_i}(\hat{\rho})$ depends on the SNR and channel variation rate encountered by each retransmission, as well as the HARQ, modulation and coding parameters, determination of the thresholds satisfying (29) is usually not analytically feasible. In the sequel, we adopt a self-learning approach to find these throughput-maximizing thresholds.

Consider an interval $[\Gamma_i - \gamma_i^-, \Gamma_i + \gamma_i^+)$, where γ_i^- and γ_i^+ are positive values. For γ_i^- and γ_i^+ small enough such that $p_{\mathcal{A}_i}(\hat{\rho})$ and $p_{\mathcal{A}_{i+1}}(\hat{\rho})$ are almost constant for $\hat{\rho} \in [\Gamma_i - \gamma_i^-, \Gamma_i)$ and $\hat{\rho} \in [\Gamma_i, \Gamma_i + \gamma_i^+)$, respectively, and let \mathcal{B}_i^i represent the event that a TTI using MCS_{*i*} occurs with its corresponding frame's $\hat{\rho} \in [\Gamma_i - \gamma_i^-, \Gamma_i)$, \mathcal{B}_{i+1}^i the event that a TTI using MCS_{*i+1*} occurs with its corresponding frame's $\hat{\rho} \in [\Gamma_i, \Gamma_i + \gamma_i^+)$; also let \mathcal{A}_i^i be the event that an ACK is generated for \mathcal{B}_i^i , and \mathcal{A}_{i+1}^i the event that an ACK is generated for \mathcal{B}_{i+1}^i , we may approximate the conditions (29) as $\mathcal{R}_i E[p_{\mathcal{A}_i}(\hat{\rho}) | \mathcal{B}_i^i] \approx \mathcal{R}_{i+1} E[p_{\mathcal{A}_{i+1}}(\hat{\rho}) | \mathcal{B}_{i+1}^i]$, or equivalently,

$$\mathcal{R}_i \frac{Pr(\mathcal{A}_i^i)}{Pr(\mathcal{B}_i^i)} \approx \mathcal{R}_{i+1} \frac{Pr(\mathcal{A}_{i+1}^i)}{Pr(\mathcal{B}_{i+1}^i)}. \quad (30)$$

As $p_{\mathcal{A}_i}(\hat{\rho})$'s are all increasing functions of $\hat{\rho}$, we find that if the left-hand side of (30) is larger than the right-hand side, it means that Γ_i is smaller than its optimal value and has to be increased. Similarly, if

the right-hand side is larger, Γ_i has to be decreased. Therefore, we may adjust Γ_i upon the occurrences of \mathcal{A}_i^i and \mathcal{A}_{i+1}^i , respectively, by

$$\text{if } \mathcal{A}_i^i: \Gamma_i \leftarrow \Gamma_i + \frac{Pr(\mathcal{B}_{i+1}^i)}{Pr(\mathcal{B}_i^i) + Pr(\mathcal{B}_{i+1}^i)} \delta_{\mathcal{A}_i^i}, \quad (31)$$

$$\text{if } \mathcal{A}_{i+1}^i: \Gamma_i \leftarrow \Gamma_i - \frac{Pr(\mathcal{B}_i^i)}{Pr(\mathcal{B}_i^i) + Pr(\mathcal{B}_{i+1}^i)} \delta_{\mathcal{A}_{i+1}^i}, \quad (32)$$

where $Pr(\mathcal{B}_i^i)$ and $Pr(\mathcal{B}_{i+1}^i)$ can be easily measured at the transmitter, and the step sizes $\delta_{\mathcal{A}_i^i}$ and $\delta_{\mathcal{A}_{i+1}^i}$ should be designed such that when Γ_i assumes its optimal value, the up and down adjustments for Γ_i will balance out. In other words,

$$\frac{Pr(\mathcal{A}_i^i)Pr(\mathcal{B}_{i+1}^i)}{Pr(\mathcal{B}_i^i) + Pr(\mathcal{B}_{i+1}^i)} \delta_{\mathcal{A}_i^i} = \frac{Pr(\mathcal{A}_{i+1}^i)Pr(\mathcal{B}_i^i)}{Pr(\mathcal{B}_i^i) + Pr(\mathcal{B}_{i+1}^i)} \delta_{\mathcal{A}_{i+1}^i}. \quad (33)$$

Comparing (30) and (33), we need

$$\frac{\delta_{\mathcal{A}_i^i}}{\delta_{\mathcal{A}_{i+1}^i}} = \frac{\mathcal{R}_i}{\mathcal{R}_{i+1}}. \quad (34)$$

As demonstrated previously, adjustment only when $\hat{\rho} \in [\Gamma_i - \gamma_i^-, \Gamma_i + \gamma_i^+)$ results in very slow adaptation. Observe that when all thresholds assume their optimal values except Γ_i which is smaller than its optimal, the error rates for MCS_i and MCS_{i+1} will be higher than their optimal values. We may thus adjust the thresholds by changing the error rate requirements.

Let $\kappa_i \equiv (1 - \beta_i)/\beta_i$. As κ_i is a decreasing function of β_i , (31), (32) and (34) can be readily applied to κ_{i+1} to give equivalent effects. Thus, we can use the following additional procedures to adjust the error rate requirements for the SEA.

$$\begin{aligned} & \text{If } \mathcal{A}_i^i: \kappa_{i+1} \leftarrow \kappa_{i+1} + P_{\mathcal{B}_{i+1}^i} \delta_{\mathcal{A}_i^i}; \\ & \text{if } \mathcal{A}_{i+1}^i: \begin{cases} \text{if } \sum_{j \leq i} Pr(\Lambda_j) > 0: \kappa_{i+1} \leftarrow \kappa_{i+1} - P_{\mathcal{B}_i^i} \delta_{\mathcal{A}_{i+1}^i} \\ \text{otherwise: } \kappa_{i+1} \text{ remains unchanged,} \end{cases} \end{aligned}$$

where $P_{\mathcal{B}_{i+1}^i}$ and $P_{\mathcal{B}_i^i}$ are shorthands for the corresponding coefficients in (31) and (32). A similar condition as in the SEA is used here to prevent the error rate requirement of an MCS from keeping increasing when the SNR distribution is such that the MCSs lower than it are never selected.

By construction, this modified SEA algorithm, termed SEA-MAX from now on, has the error rate requirements adjusted at a slower pace. While the MCS switching thresholds are adjusted at a much faster pace using the SEA. We further impose the following conditions to handle the situation where the channel statistics cause some MCSs to be eliminated, and make Λ_i 's with nonconsecutive indices adjacent to each other

$$\frac{\delta_{\mathcal{A}_i^{i-1}}}{\delta_{\mathcal{A}_{i+1}^i}} = \frac{\mathcal{R}_i}{\mathcal{R}_{i+1}}, \quad i = 2, \dots, \mathcal{M} - 1. \quad (36)$$

We may also use approximations for the throughput-maximizing β_i 's as the initial conditions. One simple approximation is to assume that $p_{\mathcal{A}_i}(\hat{\rho})$ decreases linearly from $p_{\mathcal{A}_i}(\Gamma_i) = 1$ to $p_{\mathcal{A}_i}(\Gamma_{i-1}) = \mathcal{R}_{i-1}/\mathcal{R}_i$, and $F(\hat{\rho})$ increases linearly in Λ_i . This yields the initial error rate requirements

$$\beta_i = \frac{\mathcal{R}_i - \mathcal{R}_{i-1}}{2\mathcal{R}_i}, \quad i = 2, \dots, \mathcal{M}. \quad (37)$$

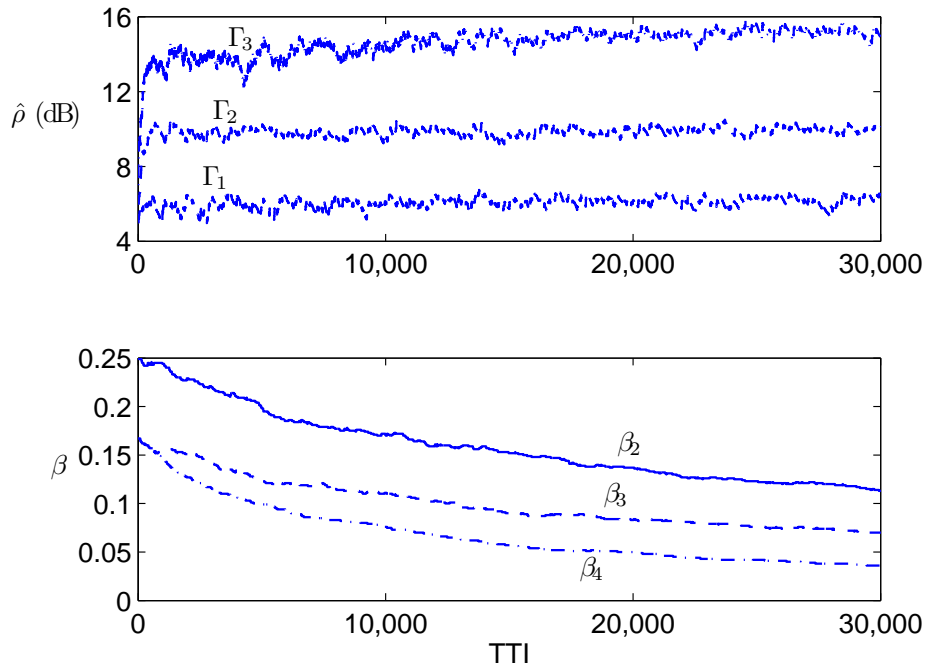


Fig. 11. Transient performance of SEA-MAX with $\hat{\rho}$ uniformly distributed between -10 dB and 30 dB: (a) thresholds (b) error rate requirements.

1) *Numerical Results:* The simulation parameters are the same as in the previous section. For MCS₂ to MCS₄, the initial error rate requirements were set using (37). For every threshold, $\gamma_i^+ = \gamma_i^- = 0.5$ dB. $P_{\mathcal{B}_{i+1}^i}$ and $P_{\mathcal{B}_i^i}$ were both initialized to 0.5, and measured with the following moving average at the transmitter. If \mathcal{B}_{i+1}^i : $P_{\mathcal{B}_{i+1}^i} \leftarrow 0.98P_{\mathcal{B}_{i+1}^i} + 0.02$ and $P_{\mathcal{B}_i^i} \leftarrow 0.98P_{\mathcal{B}_i^i} + 0.02$; if \mathcal{B}_i^i : $P_{\mathcal{B}_i^i} \leftarrow 0.98P_{\mathcal{B}_i^i} + 0.02$ and $P_{\mathcal{B}_{i+1}^i} \leftarrow 0.98P_{\mathcal{B}_{i+1}^i}$. At each TTI, if applicable, the transmitter first adjusted the error rate requirements, then the thresholds, then used the adjusted thresholds to select MCS for new frames. For the adjustment of the error rate requirements, $\delta_{\mathcal{A}_4^3} = 0.2$, and the other step sizes were determined according to (34) and (36). For the threshold adjustments, $\Delta_{\bar{\mathcal{A}}_{i,i-1}} = 0.2$ dB, $\forall i > 1$, and $\Delta_{\bar{\mathcal{A}}_{1,0}} = 0$. Assumed to have no prior knowledge of the channel statistics, the transmitter set initial thresholds $\Gamma_0 = -\infty$, $\Gamma_1 = 5$ dB, $\Gamma_2 = 6$ dB, $\Gamma_3 = 7$ dB.

We first tested the SEA-MAX in a quasi-static, flat fading channel with the SNR varying independently, and uniformly between -10 dB and 30 dB, from TTI to TTI. Retransmissions were not allowed, so ACK/NACK was used solely for threshold adjustment. The SNR feedback was assumed delay-free. Otherwise AMC would have been meaningless. In Fig. 11 we clearly see that while the error rate requirements converged slowly to their optimal values, the initial approximation of the error rate requirements was effective, and the SEA algorithm very quickly made the thresholds satisfy the error rate requirements. For constant-SNR channels, in general, the initial error rate requirements given by (37) are higher than the optimal, due to turbo codes' steep decreasing of error rate at the water-fall region. For the same reason, we also notice that the 0.5 dB upper band of a threshold does not approximate precisely the boundary throughput of the higher MCS. Thus even after the convergence, the thresholds were slightly lower than the optimal values of 6.5, 10.5, and 15.7 obtained through $p_{\mathcal{A}_i}(\hat{\rho})$'s via offline simulation. For its much larger selection region $\Lambda_4 = [\Gamma_3, 30]$, most of which had $p_{\mathcal{A}_4}(\hat{\rho}) \approx 1$, compared to MCS₃, MCS₄'s optimal error rate requirement is lower. This is evident from Fig. 11(b).

In a Rayleigh fading channel, the SNR distribution does not spread as wide as in the above example. Therefore the adaptation of error rate requirements is faster. On the other hand, such a concentrated SNR distribution may prevent certain MCSs from achieving their optimal error rate requirements and leave their

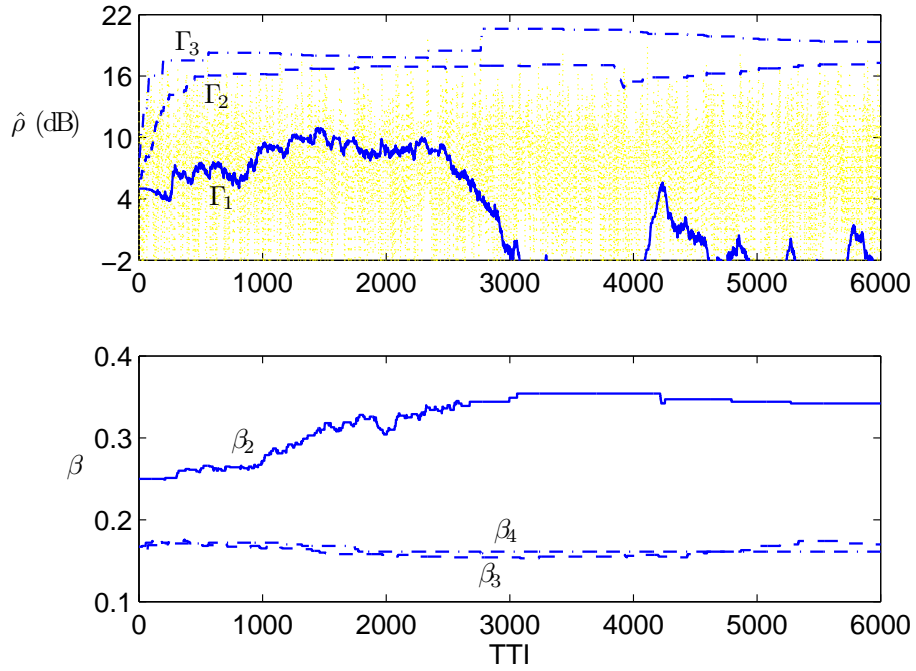


Fig. 12. Transient performance of SEA-MAX for a flat Rayleigh fading channel with 10 dB average SNR: (a) thresholds (b) error rate requirements.

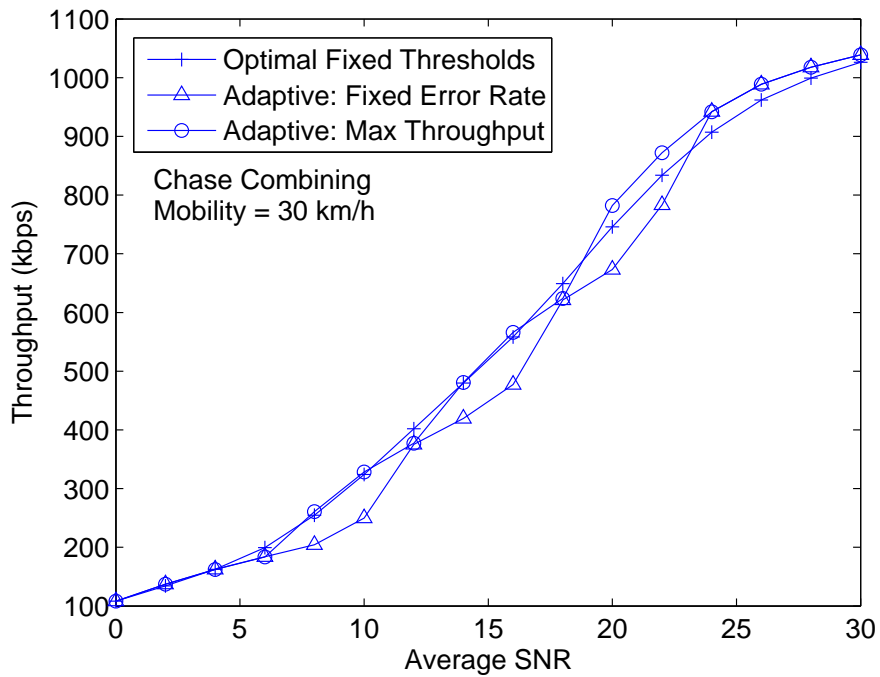


Fig. 13. Average throughput comparison for SEA-MAX, SEA, and a regular AMC system with fixed but optimal thresholds.

thresholds drifting. The drifting, however, is in regions of almost no occurrence of $\hat{\rho}$, and will not affect the system throughput much. We illustrate in Fig. 12 the transient performance of the SEA-MAX in a flat, Rayleigh fading channel with 10 dB average SNR and Doppler frequency 55.56 Hz (30 km/h mobility). Also plotted in the background of Fig. 12(a) is the SNR feedback. The HARQ was Chase combining, and the SNR feedback delay was 2 TTIs. As shown in Fig. 12(a), the SNR distribution does not allow MCS₃ and MCS₄ to achieve the initial error rate requirements. Thus Γ_2 and Γ_3 were quickly pushed out of the possible SNR range. The initial β_2 was lower than its optimal. As β_2 was adjusted higher, the corresponding Γ_1 quickly turned from increasing to decreasing. Eventually, the SNR distribution became too good for the optimal β_2 , and Γ_1 was driven lower and kept oscillating at the lower end of the SNR concentration region by the conditional procedures of the SEA-MAX.

To show that the drifting of thresholds on the edge of the SNR distribution does not affect the throughput much, we repeated the simulation in Fig. 12 for different average SNRs, and let each simulation run for 20000 TTIs (40 sec) to average out the transient behavior. We compare this system with a regular AMC system using fixed, but optimal thresholds obtained from offline evaluation; and a system which had the same initial thresholds and error rate requirements, but only adjusted the thresholds with the SEA using fixed error rate requirements. It is shown in Fig. 13 that SEA-MAX yields the maximum throughput most of the time. In fact, it sometimes outperforms the regular AMC system with optimal thresholds for its capability of tracking the short-term SNR statistics. The effectiveness of the adjustment of error rate requirements is evident from the fact that SEA-MAX improves over SEA.

VI. CONCLUSION

In the first year of this project, we have investigated some PHY and MAC issues for the proposed MIMO OFDM system. Specifically, we considered two multiplexing schemes for multiple users to share the downlink channel capacity increased by MIMO: grouped transmission, and overlapped transmission. A power control and allocation method has also been evaluated for a multi-cell OFDM WLAN operating at the 60 GHz band. In addition, a simple adaptive method to find the optimal parameters for operating AMC and HARQ is proposed.

For the grouped MIMO transmission, we have developed an adaptive maximum SNR filter bank receiver. However, this method needs to be further evaluated and optimized. And the impact of channel estimation for the antennas within the desired group on the adaptive algorithm has to be investigated. We expect to complete these first-year work items in the next three months.

Evaluation of power control and allocation for multi-cell OFDM WLANs operating at the 60 GHz band was supposed to be a work item of the second year. Thus we are ahead of the schedule on this front. In the next year, algorithms combining power allocation and PAPR reduction will be studied.

As for the optimization of AMC and HARQ, we are also ahead of the schedule. This was supposed to be a third-year work item. In the coming two years, we will extend the methodology for the single-antenna channel to MIMO channels. This extension is not trivial because the metric for determining MIMO modulation and coding schemes is very complicated.

REFERENCES

- [1] P. Smulders, "Exploiting the 60 GHz band for local wireless multimedia access: Prospects and future directions," *IEEE Commun. Mag.*, pp. 140 – 147, Jan. 2002.
- [2] J. Mikkonen, "Emerging wireless broadband networks," *IEEE Commun. Mag.*, pp. 112 – 117, Feb. 1998.
- [3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311 – 335, Mar. 1998.
- [4] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, vol. 45, no. 4, pp. 1121 – 1128, May 1999.
- [5] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," submitted to *IEEE Transactions on Information Theory*, Jul. 2004.
- [6] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439 – 441, May 1983.
- [7] H.-J. Su and E. Geraniotis, "Power allocation and control for multi-carrier systems with soft decoding," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1759 – 1769, Oct. 1999.
- [8] H. J. Su and E. Geraniotis, "Maximum signal-to-noise ratio array processing for space-time coded systems," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1419 – 1422, Sept. 2002.
- [9] P. I. Davies, "Solving the symmetric definite generalized eigenvalue problem," Ph.D. dissertation, University of Manchester, Dec. 2000.
- [10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [11] G. Mathew and V. Reddy, "A quasi-newton adaptive algorithm for generalized symmetric eigenvalue problem," *IEEE Trans. Signal Processing*, vol. 44, no. 10, pp. 2413 – 2422, Oct. 1996.
- [12] P. Strobach, "Fast orthogonal iteration adaptive algorithms for the generalized symmetric eigenproblem," *IEEE Trans. Signal Processing*, vol. 46, no. 12, pp. 3345 – 3359, Dec. 1998.
- [13] D. R. Morgan, "Downlink adaptive array algorithms for cellular mobile communications," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 476 – 488, Mar. 2003.
- [14] M. Muck, A. R. Dias, M. Debbah, and P. Duhamel, "A pseudo random postfix OFDM based modulator for multiple antennae systems," in *Communications, 2004 IEEE International Conference on*, vol. 4, Jun. 2004, pp. 2392 – 2396.
- [15] M. Muck, M. de Courville, M. Debbah, and P. Duhamel, "A pseudo random postfix OFDM modulator and inherent channel estimation techniques," in *Telecommunications Conference, 2003. GLOBECOM, IEEE*, vol. 4, Dec. 2003, pp. 2380 – 2384.
- [16] M. Muck, M. de Courville, X. Miet, and P. Duhamel, "Iterative interference suppression for pseudo random postfix OFDM based channel estimation," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference*, vol. 3, Mar. 2005, pp. 765 – 768.
- [17] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658 – 2668, Oct. 2003.
- [18] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, "Superposition coding for side-information channels," submitted to *IEEE Transactions on Information Theory*, 2004.
- [19] W. Yu, A. Sutivang, D. Julian, T. Cover, and M. Chiang, "Writing on colored paper," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Jun. 2001, p. 302.
- [20] U. Erez and R. Zamir, "Achieving $1/2 \log(1+\text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.
- [21] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

- [22] T. Berger, *Rate distortion theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [23] S. Y. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, Massachusetts Institute of Technology, Sep. 2000.
- [24] G. D. Forney, Jr., M. D. Trott, and S. Y. Chung, "Sphere-bound-achieving coset codes and multilevel coset codes," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 820 – 850, May 2000.
- [25] G. D. Forney, Jr., "Shannon meets Wiener II: On MMSE estimation in successive decoding schemes," in *42th Annual Allerton Conf. on Comm. Control, and Comput., Monticello, IL*, Sep. 2004.
- [26] M. Fiaco, M. Parks, and H. Radi. (1988, Aug.) Final report - indoor propagation factors at 17 ghz and 60 ghz. Radio Communication Agency, UK., University of Surrey. [Online]. Available: <http://www.radio.gov.uk>
- [27] A. A. M. Saleh and R. A. Valenzuela, "A statistical model for indoor multi-path propagation," *IEEE J. Select. Areas Commun.*, vol. 5, no. 2, pp. 128 – 127, Feb. 1987.
- [28] S. Kallel and D. Haccoun, "Generalized type II hybrid ARQ scheme using punctured convolutional coding," *IEEE Trans. Commun.*, vol. 38, no. 11, pp. 1938–1946, Nov. 1990.
- [29] C. Tang, "An intelligent learning scheme for adaptive modulation," in *Proc. IEEE Vehicular Technology Conference (VTC 2001 Fall)*, Atlantic City, New Jersey, Oct. 2001, pp. 141–148.
- [30] J. Lee, R. Arnott, K. Hamabe, and N. Takano, "Adaptive modulation switching level control in high speed downlink packet access transmission," in *Proc. Third International Conference on 3G Mobile Communication Technologies*, May 2002, pp. 156–159.
- [31] M. Nakamura, Y. Awad, and S. Vadgama, "Adaptive control of link adaptation for high speed downlink packet access (HSDPA) in W-CDMA," in *Proc. The 5th International Symposium on Wireless Personal Multimedia Communications*, Oct. 2002, pp. 382–386.