Power Management of MIMO Network Interfaces on Mobile Systems

Hang Yu, Lin Zhong, *Member, IEEE* and Ashutosh Sabharwal, *Senior Member, IEEE*

***Abstract*—High-speed wireless network interfaces are among the most power-hungry components on mobile systems that rely on battery for energy supply and passive means for heat dissipation. This is particularly true for MIMO network interfaces which use multiple RF chains simultaneously. In this paper, we present a novel power management mechanism for MIMO network interfaces on mobile systems, called *antenna management*. The key idea is to adaptively disable some of the antennas as well as their RF chains to reduce circuit power when the capacity improvement of using a large number of antennas is small. The number of active antennas is judiciously determined to minimize energy per bit while satisfying the data rate constraint. We provide both theoretical framework and system design of antenna management: we first present an algorithm that efficiently solves the problem of minimizing MIMO energy per bit, and then offer its 802.11n-compliant system design. We employ both simulation and prototype-based experiment to validate the energy efficiency benefit of antenna management. The results show that antenna management can achieve 21% one-end energy per bit reduction to the front end of the MIMO network interface, compared to a static configuration that keeps all antennas active.**

***Index Terms*—Power management, MIMO, Antenna management, Energy per Bit**

I. INTRODUCTION

Multiple-input multiple-output (MIMO) technologies are considered as a leading candidate for the next-generation wireless broadband due to its potential in significantly increasing link capacity without additional usage of transmit power or spectrum [1-3]. The MIMO technologies have been adopted by cutting-edge and emerging mobile wireless network standards such as 802.11n, WiMAX and LTE.

The key idea of MIMO is to simultaneously use multiple antennas[1] at both the transmitter and receiver. By properly leveraging the multiple propagation paths between the transmitter and receiver, i.e. suppressing the correlation between different paths, a MIMO link can significantly boost channel capacity, yielding either higher data rate or higher communication reliability.

However, the simultaneous use of multiple antennas significantly increases the circuit power consumption of the MIMO network interface, due to multiple active RF chains. The circuit power increase is particularly problematic for short-range communication scenarios such as 802.11-based WLAN where circuit power is often more than comparable to transmit power. Existing work on MIMO mainly focus on improving the channel quality e.g. link data rate, given the transmit power budget; little published work has considered the dual problem of reducing power consumption including circuit power under a data rate constraint.

To address the power challenge, our solution is a novel power management mechanism, called *antenna management,* which dynamically determines the number of antennas and transmit power for each antenna. Antenna management delivers each data bit with minimum energy consumption, or achieves *minimum energy per bit*, while guaranteeing the required data rate.

Antenna management leverages the mobility of mobile systems. As mobile systems move around, they encounter different propagation environments, which can lead to different capacity benefit from MIMO. Since the circuit power cost of adding one active antenna in the MIMO network interface is fixed, different environments may lead to different numbers of antennas to achieve minimum energy per bit. For example, an indoor environment with rich multipath effect can provide a MIMO channel higher capacity improvement than an outdoor environment with a dominant line-of-sight (LOS) path. As a result, a larger number of antennas is more likely to be optimal for the indoor environment.

We make the following key contributions in this work.

- We analyze the MIMO energy per bit as a function of transmit power and antenna configuration, and subsequently formulate an optimization problem towards its minimization. We distinguish the energy per bit of a single end and both ends, denoted as one-end and two-end energy per bit, respectively.
- We offer a novel solution, antenna management, which efficiently solves the MIMO energy per bit minimization problem.
- We present 802.11n-compliant system designs of antenna management. Corresponding to one-end and two-end energy per bit, we present one-ended and two-ended antenna management, with the former suitable for communication between a mobile client and an access point while the latter for that between two mobile clients.

We evaluate antenna management first with a MATLAB-based simulation, and then a first-of-its-kind prototype of a 4x4 MIMO network interface using an open-access wireless research platform called WARP. We show the effectiveness of antenna management with simulation and experimental results:

[1] We use "antenna" to refer to the passive antenna and the corresponding RF chain that powers it.

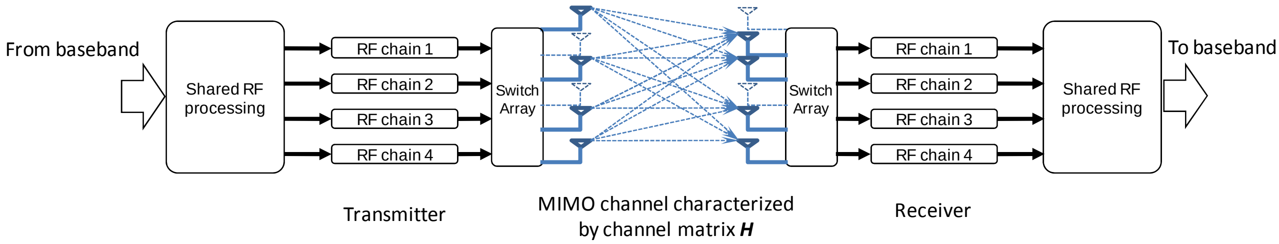


Fig. 1. A MIMO link with a transmitter and a receiver, each with four RF chains and six antennas

on average antenna management can save one-end and two-end power consumption of the front end of the MIMO network interface by 21% and 13% compared to a static MIMO link using all antennas constantly.

Antenna management can be easily extended to other multi-antenna technologies, e.g. beamforming, albeit our current design and implementation focus on MIMO. The fundamental rationale of antenna management is that in a multi-antenna system higher link performance, e.g., higher data rate or reliability, is often achieved at the cost of using more antennas and therefore higher circuit power consumption. As a result, one can improve energy efficiency when the performance is above the requirement.

The rest of the paper is organized as follows. Section II provides background on MIMO and Section III outlines related work. Section IV and V present the theoretical framework of antenna management. Section VI offers 802.11n-compliant system designs of antenna management. Section VII and VIII provide simulation-based and prototype-based evaluations, respectively. Section IX discusses antenna management in a broader context and Section X concludes the paper.

II. Background

We first provide background of MIMO technologies. In particular, we are interested in spatial multiplexing MIMO (SM-MIMO) that increases link data rate by sending independent spatial streams through multiple antennas. It has been adopted by 802.11n to support a WLAN link with a peak data rate of 300Mbps under a 20MHz channel. In the rest of the paper we use MIMO to refer to SM-MIMO unless otherwise specified.

A. *MIMO Link Architecture*

Fig. 1 illustrates a representative architecture of a MIMO link including a pair of transmitter and receiver. In practice, a MIMO transceiver can operate in a half-duplex manner and the transmitter and receiver in Fig. 1 actually stand for the MIMO transceiver in transmit and receive mode respectively. A MIMO transceiver can allow more passive antennas than RF chains and employ *antenna selection* techniques [4, 5] to determine the optimal subset of antennas, i.e., which four out of the six in Fig. 1. Each pair of transmit antenna and receive antenna forms a sub-channel between two ends, and the MIMO link is composed of these sub-channels.

B. *MIMO Channel Model*

The MIMO channel can be characterized by a $N_R \times N_T$ complex matrix $\boldsymbol{H}$, as illustrated in Fig. 1. The number of active antennas in the receiver and transmitter are denoted as N_R and N_T respectively. The channel model adopted by 802.11n [6] can be described by

$$\boldsymbol{H}(t) = \sqrt{\frac{K(t)}{K(t)+1}}\boldsymbol{H}_{LOS}(t) + \sqrt{\frac{1}{K(t)+1}}\boldsymbol{H}_{NLOS}(t).$$

In the model, $\boldsymbol{H}_{LOS}(t)$ and $\boldsymbol{H}_{NLOS}(t)$ denote the line-of-sight (LOS) component and non-line-of-sight (NLOS) component of the channel, respectively. The elements in $\boldsymbol{H}_{NLOS}(t)$ are independent normalized complex circularly symmetric Gaussian random variables. The elements in $\boldsymbol{H}_{LOS}(t)$ are all one multiplied by a phase shift $e^{j\varphi_0}$ where φ_0 is a random variable uniformly distributed within $[0,2\pi)$. In addition, $K(t)$ is the *Ricean* factor that indicates the propagation condition of the channel, i.e., how dominant is the LOS component compared to the NLOS component. By varying $K(t)$, the model can fit channels with various fading distributions. For example, $K = 0$ models ideal *Rayleigh* fading and $K = \infty$ models ideal *Ricean* fading. The change of both $\boldsymbol{H}_{LOS}(t)$, $\boldsymbol{H}_{NLOS}(t)$ and $K(t)$ can yield channel variation.

It is important to highlight that the above model has been normalized to the path loss of each sub-channel. In other words, given a fixed distance between the transmitter and receiver, the channel fluctuation due to small-scale node movement can be modeled as above. To eliminate ambiguity, we refer the change of $\boldsymbol{H}_{LOS}, \boldsymbol{H}_{NLOS}$ and K as *small-scale fading*, and the change of path loss as *large-scale fading*.

C. *MIMO Channel Capacity*

We consider a frequency-flat MIMO channel to calculate its capacity. The instantaneous channel capacity C can be calculated [1] as

$$C = \max_{\boldsymbol{R}_{SS}} B\log\det\left(\boldsymbol{I}_{N_R} + \frac{P_{TX}}{N_T N_0}\boldsymbol{H}\boldsymbol{R}_{SS}\boldsymbol{H}^H\right),$$

where B is the bandwidth, $\boldsymbol{R}_{SS} = E(\boldsymbol{s}\boldsymbol{s}^H)$ the covariance matrix of transmit signal $\boldsymbol{s}$ satisfying the transmit power constraint $Tr(\boldsymbol{R}_{SS}) = N_T$, $\boldsymbol{H}$ the channel matrix, P_{TX} the total transmit power across all transmit antennas, N_0 the noise level at the receiver, and $\boldsymbol{I}_{N_R}$ an $N_R \times N_R$ unit matrix. Note that the maximization is performed over all qualified covariance matrices $\boldsymbol{R}_{SS}$ with $Tr(\boldsymbol{R}_{SS}) = N_T$. The channel matrix $\boldsymbol{H}$ needs to be known by the transmitter to achieve the above capacity C by allocating the transmit power with the well-known water-filling algorithm [7].

The capacity is the upper bound on the error-free data rate allowed by the MIMO channel. The model clearly indicates that capacity C depends on both the channel and the communication parameters. That is, P_{TX}, N_T and N_R are dependent on the configuration of the transmitter and receiver, N_0 is dependent on the receiver, while $\boldsymbol{H}$ is determined by the channel.

Higher transmit power P_{TX} or more active antennas, i.e. larger N_T and N_R, can increase the channel capacity C, regardless of $\boldsymbol{H}$ [1]. However, both ways will meanwhile increase power consumption of the transceiver. A key observation that motivated this work is that under some particular circumstances more antennas can yield little improvement to the channel capacity, e.g., when the channel has a large *Ricean* factor K so that the sub-channels are highly correlated. Under these circumstances it may not be energy efficient to employ a large number of active antennas.

D. *MIMO Power Model*

A MIMO link consists of a transmitter and a receiver. Therefore, the power consumption of a MIMO link, P_{MIMO}, includes the power consumption of the transmitter, $P_{Transmit}$, and that of the receiver, $P_{Receive}$, or

$$P_{MIMO} = P_{Transmit} + P_{Receive}.$$

One can divide $P_{Transmit}$ into the power consumed by all the power amplifiers, P_{PA}, and that by all the other transmit circuit blocks $P_{Circuit\_T}$ [8], i.e.

$$P_{Transmit} = P_{PA} + P_{Circuit\_T}.$$

Note that we assume identical power amplifiers for all RF chains so that P_{PA} only depends on the total transmit power P_{TX}. Moreover, $P_{Circuit\_T}$ can be divided into that contributed by each active transmit RF chain, $P_{RF\_chain\_T}$, and that by the circuit shared by all active transmit RF chains, P_{Shared_T}. Therefore, we approximate $P_{Transmit}$ as

$$P_{Transmit} = \frac{P_{TX}}{\eta(P_{TX})} + N_T P_{RF\_chain\_T} + P_{Shared\_T},$$

where N_T is the number of active antennas in the transmitter, and η is the drain efficiency of the power amplifier depending on the transmit power. Usually the higher the transmit power, the higher the drain efficiency. We approximate η as a linear function of P_{TX} where $\eta_{min} \leq \eta \leq \eta_{max}$ and η_{min}, η_{max} correspond to the minimum and maximum transmit power respectively.

Similarly, we approximate $P_{Receive}$ as

$$P_{Receive} = N_R P_{RF\_chain\_R} + P_{Shared\_R},$$

where $P_{RF\_chain\_R}$ and P_{Shared_R} represent the power consumed by each receive RF chain and that by the shared receive circuit, respectively.

We note that a MIMO transceiver can stay idle and monitor the channel when it has nothing to transmit or receive. We represent the power consumption of idle MIMO transceiver as P_{Idle}. P_{Idle} is constant over time and proportional to the number of active antennas used for monitoring the channel.

III. Related Work

Next we discuss existing work related to antenna management in three orthogonal directions. First, we compare antenna management with existing MIMO capacity maximization techniques. All these techniques attempt to maximize the MIMO capacity under certain resource constraint, e.g. number of antennas or transmit power. Antenna management, however, intentionally underutilizes the resource for the purpose of power saving. Second, we show that antenna management can be considered as a rate adaptation protocol in MIMO systems. Previous rate adaption protocols, most of which target SISO systems, are mainly channel-aware. In contrast, antenna management is both channel-aware and energy-aware. Finally, we distinguish antenna management with other energy efficient MIMO system designs, by highlighting the special characteristics of antenna management.

A. *MIMO Capacity Maximization*

Existing MIMO capacity maximization works improve the capacity of the MIMO link with a given number of antennas and transmit power constraint, while antenna management solves its dual problem: find the optimal number of antennas to minimize energy per bit under a data rate constraint.

One popular approach for MIMO capacity improvement is to employ more antennas than RF chains and dynamically select a subset of antennas to maximize capacity, or *antenna selection* [4, 5, 9, 10]. That is, it selects the best N out of M antennas, where N is the number of RF chains and $N<M$. To the best of our knowledge, existing antenna selection techniques assume that the number of active antennas, or N, is constant. They can be considered to minimize energy per bit given the fixed number of active antennas. On the other hand, antenna management keeps only k active antennas and selects the best k out of M antennas to minimize energy per bit under a data rate constraint, where $1 \leq k \leq N < M$. From this perspective, antenna selection is inherently integrated into antenna management.

Another popular approach for MIMO capacity improvement is to allocate transmit power into each active antenna optimally, or *spatial power allocation* [7, 11, 12]. Spatial power allocation algorithms employ a fixed total transmit power budget and assume all antennas active. In contrast, antenna management dynamically optimizes the transmit power and manages the RF chains and passive antennas by activating only a subset of them. Inactive antennas can be considered as zero transmit power allocated. The optimized transmit power is further optimally allocated across all active antennas using the water-filling algorithm. From this perspective, spatial power allocation is also included in antenna management.

B. *Rate Adaptation*

Antenna management can be regarded as a realization of *SNR-triggered* rate adaptation techniques for MIMO links. Rate adaptation techniques [13-15] aim to find the most appropriate rate for the current channel condition to approach capacity. Due to channel fading, there may be a need to change the rate over time. *SNR-triggered* rate adaption is widely-used; it uses the received SNR as the basis to estimate channel capacity and then adapts rate. Similarly, antenna management alters data rate by changing the number of active antennas. It leverages the channel matrix to estimate the MIMO channel capacity, according to which the optimal antenna configuration is chosen.

C. *Energy Efficient MIMO System Design*

There exist previous works [16, 17] on the energy-efficient MIMO system design. We have presented early results from our work in [18]. In the following we selectively choose two of other previous works that are most related to our work and compare them with ours.

In [16], the authors suggest to dynamically adjust the transmission mode (SISO, SIMO, MISO or MIMO) and transmit power on each node in an ad-hoc network to maximize energy efficiency, according to the distance of each pair of transmitter and receiver. However, they used MIMO for diversity purpose and fixed the data rate of each link. Antenna management, in contrast, uses MIMO for spatial multiplexing and allows dynamic data rate while minimizes energy per bit. In addition, their work addresses the BER constraint while ours targets the data rate constraint.

Very recently, the authors of [17] propose an adaptive MIMO system that can switch between two modes (MIMO and SIMO). At the mobile end, either one or two transmit antennas are used according to the transmission rate at the mobile end and the load utilization at the base station. They seek to minimize energy per bit by choosing the better mode. However, our work departs from theirs in three ways. First, their work has addressed energy efficiency of the transmitter without that of the receiver, while we have taken into account both in two-ended management. Second, they have assumed the MIMO modes with no more than two RF chains and antennas. In contrast, antenna management can be effective for any MIMO configuration, irrespective of the number of RF chains and antennas. Last and most importantly, their work stays with theoretical analysis and simulation results, while we have extensively discussed system designs and implementation issues of antenna management with both simulation and experimental results.

IV. MIMO Energy per Bit Minimization

In this section we analyze the MIMO energy per bit minimization problem. As shown in Section II, for a MIMO link the circuit power increase by multiple active antennas is fixed while the capacity improvement can be dynamic depending on the propagation condition of the channel, or $\boldsymbol{H}$. An important question is: given $\boldsymbol{H}$, what is the optimal number of RF chains and subset of antennas that yields the minimum energy per bit? To answer the question, we first formulate it as an optimization problem.

A. *Objective Function*

We first embark on the objective function. The MIMO energy per bit E_b can be calculated as the power consumption P divided by the data rate R, or $E_b = P/R$. We make two observations on P and R before proceeding with the concrete formulation. First, while a MIMO link is composed of two ends, they can be either a pair of mobile nodes like smart phones or laptops, or a centralized node such as an access point or a base station and a mobile node. Usually a mobile node is power-constrained due to limited battery life or thermal concerns, while a centralized node is regarded as having unlimited power. Therefore, P can be either the power consumption of both ends or that of a single end. Second, the data rate R supported by the MIMO channel is always decided by both ends, i.e., R is function of the number of antennas of both ends, N_T and N_R, regardless of the specific expression of P. Nonetheless, depending on where the optimization takes places, N_T and N_R can be either dynamic as optimization variables or fixed as constants. Noticeably, the two ends need to choose their number of antennas cooperatively to achieve certain data rate when both N_T and N_R are dynamic. Such two observations motivate us to categorize the MIMO energy per bit minimization into multiple cases, elaborated as follows.

When both ends are mobile nodes, we have $P = P_{MIMO}$; when only a single mobile end is of energy efficiency interest, we have $P = P_{Transmit}$ and $P = P_{Receive}$ for the node as a transmitter and receiver respectively. In addition, when the optimization is performed in both ends in a cooperative manner, we have $E_b = E_b(P_{TX}, N_T, N_R)$; if a single end is considered, we have $E_b = E_b(P_{TX}, N_T)$ for the transmitter and $E_b = E_b(N_R)$ for the receiver. Therefore, there can be nine different cases due to three different expressions of the power consumption, as well as three different optimization variable combinations. TABLE I presents all nine cases: the rows refer to different power consumption expressions and columns to optimization variable combinations. We can further classify the nine cases according to their similarities, and Case 1, Case 5 and Case 9 are the only three cases that do not overlap each other and cannot be solved by existing techniques. We provide the detailed explanation in Appendix A. For clarity, we name Case 1 by *two-end* energy per bit minimization, and Case 5 and Case 9 by *one-end* energy per bit minimization. The formulation of their corresponding objective functions can be specified as follows:

Case 1:

$$E_b(P_{TX}, N_T, N_R) = \frac{(1+\alpha)P_{TX} + N_T P_{RF\_chain\_T} + N_R P_{RF\_chain\_R} + P_{Shared\_T} + P_{Shared\_R}}{R}$$

Case 5:

$$E_b(P_{TX}, N_T) = \frac{(1+\alpha)P_{TX} + N_T P_{RF\_chain\_T} + P_{Shared\_T}}{R}$$

TABLE I

CATEGORIZATION OF MIMO ENERGY PER BIT MINIMIZATION

		Optimizing both ends	Optimizing one end	
			Transmitter optimization	Receiver optimization
Energy per bit of both ends		***Case 1*** $E_b = P_{MIMO}/R = E_b(P_{TX}, N_T, N_R)$	***Case 2*** $E_b = P_{MIMO}/R = E_b(P_{TX}, N_T)$	***Case 3*** $E_b = P_{MIMO}/R = E_b(N_R)$
Energy per bit of one end	Transmit energy per bit	***Case 4*** $E_b = P_{Transmit}/R = E_b(P_{TX}, N_T, N_R)$	***Case 5*** $E_b = P_{Transmit}/R = E_b(P_{TX}, N_T)$	***Case 6*** $E_b = P_{Transmit}/R = E_b(N_R)$
	Receive energy per bit	***Case 7*** $E_b = P_{Receive}/R = E_b(P_{TX}, N_T, N_R)$	***Case 8*** $E_b = P_{Receive}/R = E_b(P_{TX}, N_T)$	***Case 9*** $E_b = P_{Receive}/R = E_b(N_R)$

Case 9:

$$E_b(N_R) = \frac{N_R P_{RF\_chain\_R} + P_{Shared\_R}}{R}$$

While the channel capacity C in bits/s is known as the maximum theoretically achievable data rate, practical coding schemes can only approach C with a certain gap. For the sake of tractability, we temporarily use channel capacity C to represent data rate R in this section.

B. *Constraint*

We add an important constraint to the optimization problem: minimum data rate R_{min}. Since the optimization variables, including P_{TX}, N_T and N_R, have a direct impact on the data rate R, we must minimize E_b under a data rate constraint, or $R \geq R_{min}$.

C. *Optimization Variables*

We next look into the optimization variables. Without loss of generality, we choose Case 1 for discussion.

First, we make a key observation on the transmit power P_{TX}: given $\boldsymbol{H}$ and R_{min} there exists an finite optimal transmit power, P_{TX_opt}, that yields minimum E_b. Such observation ensures the need of optimization for P_{TX} and we provide the detailed proof in Appendix B.

Second, we combine other optimization variables as one, named the antenna configuration ω. Note that ω consists of not only the number of antennas N_T, N_R, but also the subset of them which are not explicitly shown as optimization variables. Apparently, each ω yields an unique $\boldsymbol{H}(\omega)$ thereby a unique optimal transmit power $P_{TX\_opt}(\omega)$.

V. ALGORITHMIC DESIGN OF ANTENNA MANAGEMENT

Antenna management is intended to provide an efficient solution to the MIMO energy per bit minimization problem proposed in Section IV. In the following we first outline the key challenges of obtaining the solution and then detail how those challenges are addressed by elaborating the algorithmic design of antenna management.

A. *Key Challenges*

Here we show several important challenges of solving the MIMO energy per bit minimization problem:

- Given $\boldsymbol{H}(\omega)$ and R_{min}, no closed-form formulation of the optimal transmit power $P_{TX\_opt}(\omega)$ is obtainable. Heuristic methods to identify P_{TX_opt} will yield considerable complexity.
- The search complexity of finding the optimal antenna configuration ω_{opt} increases exponentially with the number of antennas, i.e., $O(2^{N_T N_R})$ searches are needed to find ω_{opt}.
- The optimal transmit power P_{TX_opt} and antenna configuration ω_{opt} should be jointly identified due to their dependency on each other.

B. *Solution*

Antenna management leverages two key techniques to tackle the above challenges: it employs a pre-built mapping to calculate P_{TX_opt}, and efficient antenna selection techniques to calculate ω_{opt}. The overall algorithm is summarized in Algorithm 1. We next describe each of the two techniques.

1) Pre-Built Mapping

First, we observe that, without considering R_{min}, P_{TX_opt} is primarily determined by the dimension of $\boldsymbol{H}$ under small-scale channel fading. In other words, the channel matrix $\boldsymbol{H}$ in typical indoor and outdoor environments follows certain patterns. The randomness of $\boldsymbol{H}$ is constrained to some level so that an approximated solution of P_{TX_opt} can yield close-to-optimal result. According to these observations, we build a mapping from (N_T, N_R) to P_{TX_opt} offline.

We employ both synthetic and measured channels to evaluate such offline mapping. The channel model in Section II is adopted for synthetic channels, and the setting for the measured channels is offered in Section VIII. We first compute P_{TX_opt} for each synthetic/measured channel matrix $\boldsymbol{H}(N_T, N_R)$ with different N_T and N_R, denoted as the training procedure. Then we take the median of those P_{TX_opt} as the approximated optimal transmit power, P_{TX_appx}, corresponding to each N_T, N_R and construct the mapping accordingly. To see how good P_{TX_appx} is, we compare the MIMO energy per bit under P_{TX_appx} and P_{TX_opt}, for 1000 synthetic and measured channels respectively. The results have shown that the MIMO energy per bit under P_{TX_appx} is only 1.3% and 1.6% higher than that under P_{TX_opt}, for synthetic and measured

Algorithm 1: Antenna Management

Input: MIMO channel matrix $\boldsymbol{H}$, minimum data rate constraint R_{min}

Output: optimal transmit power P_{TX_opt}, optimal antenna configuration ω_{opt}

```
E_{b,min} = +∞
for 1 ≤ n_t ≤ N_T, 1 ≤ n_r ≤ N_R
   identify H(n_t, n_r) using antenna selection algorithms
   P_TX = P_TX(n_t, n_r, R_min) using the pre-built mapping
   E_b = E_b(P_TX, n_t, n_r)
   if E_b < E_{b,min}
      E_{b,min} = E_b, P_TX_opt = P_TX, ω_opt = ω
   end
end
return P_TX_opt, ω_opt
```

channels respectively, which is negligible compared to the energy per bit reduction by antenna management. We note that no upper or lower bounds on P_{TX_opt} can be found theoretically. However, our results indicate that the probability that P_{TX_appx} is close to P_{TX_opt} is very high.

As proved in Appendix B, we can build a mapping from R to P_{TX}. Similarly, the mapping is approximately independent on $\boldsymbol{H}$ under small-scale channel fading thereby can be built offline.

2) Efficient Antenna Selection

In addition, we convert the procedure of finding ω_{opt} into multiple steps where each step can be solved by existing antenna selection algorithms. It is apparent that given fixed N_T and N_R, minimizing E_b is equivalent to maximizing R since P is constant. Therefore, it turns to be a capacity-maximization-based antenna selection problem where existing efficient algorithms can be straightforwardly leveraged, e.g. [10]. As a result, we only need to identify the optimal number of antennas, with a complexity of $O(N_T N_R)$.

VI. System Design of Antenna Management

Following the algorithmic description of antenna management, we next provide its system design. We choose the MIMO-based WLAN standard, IEEE 802.11n, as the operating protocol of antenna management due to its commercialization and popularity. Next we briefly outline 802.11n and then elaborate how antenna management can be implemented compatibly.

A. *MIMO-based 802.11n*

802.11n supports up to a 4x4 MIMO configuration. That is, up to four RF chains can be integrated in the MIMO network interface. More than one passive antenna can be attached to each RF chain to enable antenna selection. Each RF chain together with its selected passive antenna is responsible for sending a spatial stream. A single frame from MAC can be broken up and multiplexed across multiple spatial streams, and then reassembled at the other end. The number of spatial streams, or the number of active antennas is allowed to be dynamic over time. The PLCP header specifies this number so that the receiver is able to know it and correctly decode the original signals.

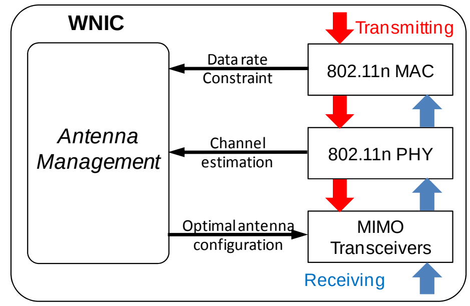


Fig. 2. Antenna management is implemented inside the WNIC. It obtains data rate constraint from MAC and channel estimation from PHY, and then selects the optimal antenna configuration.

Besides different number of active antennas, different modulation and coding schemes can be used for each active antenna to further diversify the overall data rate. We highlight that different data rates may leave different idle period between successive frames, if the traffic is not continuous. Generally more active antennas transmit the frame faster and the MIMO interface stays idle for longer time.

B. *Design Overview*

We will present two 802.11n-compliant system designs of antenna management. The first one is one-ended and the end with antenna management only considers its own energy efficiency. It suits the mobile node in the infrastructure-mode WLAN, since energy is not the primary concern for the access point. The access point can simply use legacy 802.11 and compatibly communicate with the antenna-managed mobile node. In contrast, the other design is two-ended. It implements antenna management at both ends and seeks the minimum energy per bit for the MIMO link through joint optimization on both ends. It is desirable when both ends are energy-constrained, e.g. two mobile nodes in an ad-hoc network. The one-end and two-end energy per bit minimization in Section IV are in fact the theoretical foundations for the one-ended and two-ended management, respectively.

We next elaborate the commonalities of both one-ended and two-ended antenna management and separately explain their uniqueness in the next two sections. For both designs, antenna management is intended to be implemented inside the WNIC, as illustrated by Fig. 2. It obtains the channel estimation from PHY and data rate constraint from MAC, and then identifies and subsequently selects the optimal MIMO configuration.

1) Setting the Data Rate Constraint

While 802.11n supports very high PHY data rates, wireless interfaces on mobile systems usually experience much lower rates, primarily due to the bottleneck in the rest of the network or system, e.g. DSL link or applications like VoIP with low rate or packet supply [19]. Therefore, the minimum constraint of data rate, R_{min}, should be set by upper layers that can as-

TABLE II
MAPPING FROM NUMBER OF ANTENNAS TO OPTIMAL TRANSMIT POWER

Number of transmit antennas → / *Number of receive antennas* ↓	1	2	3	4
1	27.4mW	31.1mW	39.0mW	47.9mW
2	33.2mW	53.4mW	59.6mW	62.5mW
3	35.7mW	55.3mW	73.8mW	79.2mW
4	37.5mW	61.1mW	79.4mW	95.6mW

sess the user/system/application requirements or performance bottleneck.

However, the required data rate from upper layers varies much slower than the rate supported by 802.11n, i.e., on a per-frame basis for the latter. Therefore, the data rate of each frame should be constrained in a way that the average data rate over a short period meets the requirement. As a result, we set the per-frame data rate constraint, $R_{min}(n)$, according to

$$R_{avg}(n) = w \cdot R_{min}(n) + (1 - w) \cdot R_{avg}(n - 1),$$

where n is the index of the frames to which antenna management is applied, $R_{avg}(n)$ the experienced average rate of frames 1 to n, and $0 < w < 1$ the weight given to current frame. Naturally, if each frame has approximately equal length, one should set $w = 1/n$.

2) Estimating the MIMO Channel

It is crucial that channel matrix $\boldsymbol{H}$ is known in order to estimate the channel capacity. For two-ended antenna management, both the transmitter and receiver need to have knowledge of $\boldsymbol{H}$, which imposes bigger challenge since accurate channel information is known to be hardly obtainable at the transmitter.

Nonetheless, there are indeed two ways to acquire $\boldsymbol{H}$ for the transmitter, namely *explicit channel estimation* and *implicit channel estimation*. Explicit channel estimation requires the transmitter to send training symbols to the receiver and the latter calculates $\boldsymbol{H}$ and then feeds it back to the former. This method maintains accuracy but incurs overhead because of considerable amount of information for feedback of $\boldsymbol{H}$. Implicit channel estimation, instead, leverages channel reciprocity and lets the transceiver estimate $\boldsymbol{H}$ in receive mode and assume it as the channel for transmitting next frame. It is effective only when the communication is by-directional.

802.11n supports both explicit and implicit channel estimation. The PLCP preamble of each frame contains training symbols that are used at the receiver for estimating the channel matrix. All antennas transmit the same preamble, only with time shifts relative to the others.

3) Identifying the Optimal Antenna Configuration

In Section VI we have presented an efficient algorithm to identify the optimal transmit power P_{TX_opt} and antenna configuration ω_{opt}. That is, we first establish a mapping from the number of antennas N_T, N_R to P_{TX_opt}, as well as a mapping

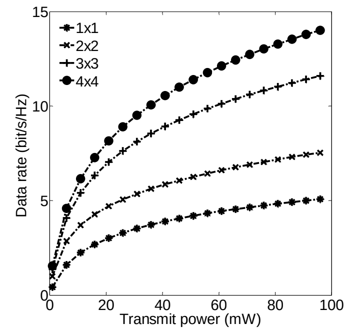


Fig. 3. Mapping from transmit power to data rate under different antenna configuration

from the data rate R to P_{TX}. To cope with large-scale channel fading by significant movement of the mobile node, the mappings should be updated when the path loss is found to have considerable variation. This can be accomplished by consistently examining the RSSI available in 802.11n since the RSSI variation due to small-scale fading is very small thereby distinguishable. TABLE II shows an example of the mapping from N_T, N_R to P_{TX_opt} and Fig. 3 shows an example of the correspondence between R and P_{TX}. Also note that the transmit power in realistic 802.11n transceivers is usually not continuously available so that the mapping can be further simplified.

To get ω_{opt}, we need to estimate the data rate for each number of antennas. While one can calculate the channel capacity using $\boldsymbol{H}$ as the upper bound of data rate so that minimize the lower bound of energy per bit, the solution can be only sub-optimal under non-continuously available rates. Therefore, specifically for 802.11n, we also need a mapping from the channel matrix to the achievable data rate. One can simply calculate the channel capacity and accordingly find the most appropriate data rate. Such procedure is already used in many rate adaptation works, e.g. [15, 20]. We also note that additional channel properties such as the BER or PER can be taken into account during this procedure and they are complementary to antenna management.

C. *One-ended Antenna Management*

Next we present the unique design principle of one-ended antenna management and target two-ended management in next section.

One-ended antenna management is triggered by the event that the MIMO transceiver receives a frame in the receive mode. The default configuration of receive mode has all antennas active, which is essential for the transceiver to implicitly estimate the channel matrix. The transceiver leverages the estimated channel matrix to obtain the achievable data rate, gets the optimal transmit power for each antenna configuration, and then identifies the one with minimum energy per bit.

The implicit channel estimation is effective because of the acknowledgement mechanism intrinsic to 802.11: a receiver immediately sends back an ACK frame to the transmitter to acknowledge the reception of a data frame, which provides the transmitter a free opportunity to estimate the channel matrix. Moreover, 802.11 provides a carrier sensing mechanism with RTS/CTS exchange before data frame transmission. While it is only used for large data frames, it guarantees that the channel estimation is valid, especially for non-continuous transmission where a long interval may exist between the last ACK frame and the current data frame.

D. *Two-ended Antenna Management*

Two-ended management leverages explicit channel estimation and is initiated by the transmitter which intends to transmit data frames to the receiver. Since each data frame actually grants the receiver an opportunity to estimate $\boldsymbol{H}$ with the inherent training symbols in the PLCP preamble (see Section A), no particular training frames are needed for explicit channel estimation. While the training symbols need to be sent and received at all transmit and receive antennas even though the actual data frame transmission involves only a subset of them, corresponding overhead is negligible since the training period is much shorter compared to frame duration.

Seeing the training symbols, the receiver estimates the channel matrix and subsequently identifies P_{TX_opt} and ω_{opt}. To reduce the amount of information needed for feedback, the receiver sends back ω_{opt} to the transmitter instead of $\boldsymbol{H}$. The optimal antenna configuration ω_{opt} can be encoded as indexes, i.e., a 4x4 MIMO link has 256 possible configurations and can be uniquely represented by 8 bits. Once the transmitter knows ω_{opt}, it can easily find the optimal transmit power P_{TX_opt} locally using the pre-built mapping. Such small amount of feedback information allows it to be included in the ACK frame sent immediately after the data frame. Therefore, two-ended antenna management does not incur additional channel occupation.

Moreover, in order to correctly perform two-ended antenna management, the receiver needs to know the power profile of the transmitter, i.e. the mapping for P_{TX_opt} as well as $P_{RF\_chain\_T}$ and $P_{RF\_chain\_R}$. These parameters can be exchanged by the two ends in advance of data transmission, and such exchange is needed only once.

E. *Overhead*

It is important to note that both one-ended and two-ended antenna management can be almost overhead-free. While powering on/off RF chains may produce additional latency and energy cost, such overhead is negligible. First, modern 802.11 transceivers incur very small latency for switching on/off a entire RF circuitry, e.g. 50μs for the MAX2829 transceiver [21] used in our prototype. Second, our measurement showed that the energy overhead can be readily compensated by powering off the transceiver for as short as 100ns, while frame transmission may last several milliseconds. Last and most importantly, antenna management does not incur additional switching besides the regular transition between transmit and receive mode, which is inevitable even without antenna management since 802.11n transceivers are half-duplex.

VII. Simulation-based Evaluation

In this section, we use a MATLAB-based simulation to evaluate antenna management under synthetic channels.

A. *MATLAB Simulation Setup*

We employ MATLAB to simulate a MIMO link that includes two identical 802.11n-like transceivers, denoted as Node 1 and Node 2. Each transceiver has four RF chains and antennas, in accordance to 802.11n. To maximally illustrate the potential benefit of antenna management, we go beyond current 802.11n and assume that both transmit power and data rate can be arbitrarily chosen. The 802.11n-compliant transmit power and data rate settings will be assumed for our prototype-based evaluation in Section VIII. Parameters in the power model are set as follows: $0.3 \le \eta \le 0.5$, $P_{RF\_chain\_T} = 48.2\text{mW}$, $P_{RF\_chain\_R} = 62.5\text{mW}$, $P_{Shared\_T} = P_{Shared\_R} = 50\text{mW}$. Those numbers are chosen according to the model in [8] and very close to recently reported realizations of 802.11n transceivers such as [22] (without baseband power).

We use two traffic patterns, namely continuous traffic and intermittent traffic, to represent different frame arrival rates. For the continuous traffic we assume that frames from upper layers arrive at an extremely high rate so that transceivers are always engaged in active transmitting or receiving, i.e., idle period never appears. This traffic represents applications which supply data at a much higher rate than the maximum rate achievable by the transceiver, e.g., FTP. Users can specify the minimum data rate which is acceptable and must be satisfied by antenna management. Intermittent traffic, in contrast, may introduce considerably long idle period between successive frames, during which the transceiver enters idle mode. Intermittent traffic represents applications which generate sparse traffic and relatively a lower data rate, such as VoIP. The data rate constraint can be determined by the application. While the distribution of inter-frame interval can be arbitrary, e.g. exponential, we assume a constant interval in our simulation for simplicity. Therefore, the intermittent traffic is completely periodic. It is important to note that some applications such as HTTP usually have both continuous and intermittent traffic patterns: data frames from application usually arrive in bursts. While each burst can be considered as a continuous traffic of multiple frames, the bursts come intermittently. It is worth noting that the inclusion of idle period energy does not change the structure of the MIMO energy per bit minimization problem so that antenna management is still valid under intermittent traffic. We provide elaborated explanation regarding this in Appendix C.

For both continuous and intermittent traffic, we explore five scenarios for evaluation using synthetic channels based on the channel model in Section II. For the first four scenarios we use a constant *Ricean* factor K which equals 0, 1, 10, 100 respectively, to represent different fading distributions; for the last scenario we assume K is random. We do not consider large-scale fading in the simulation. All the channels are assumed to be reciprocal, with a 20MHz bandwidth. We also assume the channel coherence time is longer than the frame length so that the per-frame adaption of antenna management can perfectly

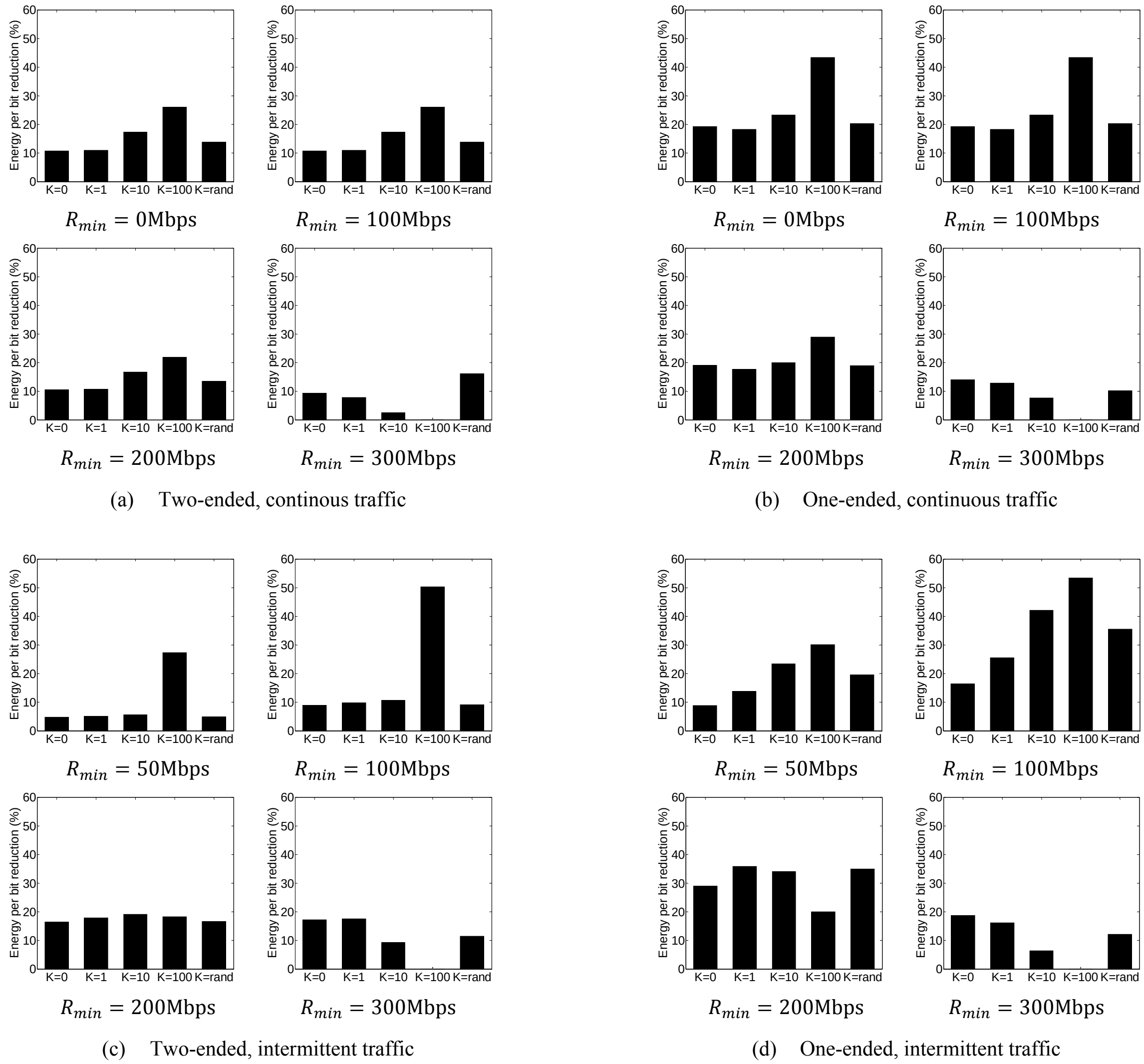


Fig. 4. MIMO energy per bit reduction by antenna management

cope with channel variation. In fact, real-world channels usually exhibit longer coherence time according to our channel traces in Section VIII.

B. *Simulation Results*

We use a static configuration with all antennas active all the time for both ends, i.e. a static 4x4 MIMO link, as the baseline to compare energy efficiency with antenna management. For two-ended management we measure the average energy per bit of the MIMO link for 1000 successive data frames, with the first 500 frames transmitted by Node 1 and the second 500 frames transmitted by Node 2. For one-ended management only Node 1 implements antenna management which transmits 1000 data frames to and receives ACK frames from Node 2. Instead of energy per bit of the MIMO link, we measure that of Node 1 only.

1) Continuous Traffic

We first evaluate the performance of antenna management under continuous traffic.

Fig. 4 (a) and (b) shows the MIMO energy per bit reduction by two-ended and one-ended antenna management, respectively. We employ four different data rate constraints for comparison: 0Mbps, 100Mbps, 200Mbps and 300M bps.

We make several key observations from Fig. 4 (a) and (b). First, the energy per bit achieved by antenna management is always no larger than that of the static configuration. Considering the scenario with a random K as the one where antenna management offers average performance, 13% and 21% energy per bit reduction can be achieved for two-ended and one-ended management, respectively. Since one-ended management only considers the energy per bit of a single end and the other end is supposed to use all antennas, the amount of energy per bit reduction can be potentially higher than that of two-ended management. Second, when the data rate constraint increases, the energy per bit reduction by antenna management is shrinking. This is because higher data rate requires more active antennas so that the optimal configuration given by antenna management gradually approaches the static one as data rate constraint increases. Third, under a relatively low data rate constraint, e.g., less than 200Mbps, when the LOS path is more dominant with a larger K, antenna management becomes more effective. This is because under channels with large K, adding more antennas brings marginal capacity improvement while incurs fixed additional power expense. Therefore, the static configuration using all antennas can be very inefficient. On the other hand, under high data rate constraint, the conclusion is on the contrary. That is, antenna management reduces energy per bit more for channels with smaller K. It can be explained by the fact that using the same number of antennas, a channel with smaller K is able to yield higher channel capacity so that more likely to meet the data rate constraint.

2) Intermittent Traffic

We then assess antenna management under intermittent traffic. As shown in Fig. 4 (c) and (d), while most observations for the continuous traffic can be similarly extended to the intermittent traffic, there are indeed unique characteristics for the latter. First, the peak energy per bit reduction for intermittent traffic is higher than that for continuous traffic. For example, under a random K, 18% and 34% average energy per bit can be reduced by two-ended and one-ended management under intermittent traffic, in contrast to 13% and 21% under continuous traffic. This is because antenna management under intermittent traffic also reduces energy consumption during the idle period. Recall that for intermittent traffic we assume fixed intervals between successive frames. Therefore, using fewer antennas to extend active transmitting or receiving duration can result in shorter idle period thereby less idle energy consumption. Second, under intermittent traffic there exists certain data rate constraint under which antenna management is most effective, e.g. R_{min}=200Mbps for two-ended management and R_{min}=100Mbps for one-ended management. Again, this is due to the consideration of idle energy. When the data rate constraint is low, i.e., the traffic is sparse, both transceivers spend very little time in transmitting or receiving. Therefore, the transmitting and receiving optimization by antenna management improves efficiency very little which is different from continuous traffic. For very high data rate constraint intermittent traffic indicates similar results as continuous traffic does. We must note that many have addressed the energy efficiency of the idle period, e.g., the authors of [19]. Such work is complementary to antenna management.

C. *Robustness against Channel Estimation Error*

In this section we use simulation results to demonstrate that antenna management still offers great energy saving even though the channel estimation is reasonably inaccurate. We assume one-ended management for simplicity while the methodology and answer also hold valid for two-ended management.

The effectiveness of antenna management depends on the channel estimation accuracy. One-ended management employs implicit channel estimation so that the estimated $\boldsymbol{H}$ may greatly deviate from its real value when channel reciprocity is poor or the received SNR is low. We examine the impact of channel estimation error on antenna management. Toward this end, we assume different received SNR in the simulation and compare the theoretically optimal configuration with the one identified by antenna management.

TABLE III

ROBUSTNESS OF ANTENNA MANAGEMENT UNDER INACCURATE CHANNEL ESTIMATION

Received SNR (dB)	*Probability of making correct decisions (%)*	*Increase of energy per bit (%)*
20	99.9	0.05
15	97.4	0.47
10	94.2	0.71
0	91.5	0.99

TABLE III shows the probability that antenna management offers the theoretically optimal configuration and the average energy per bit increase for four different SNR levels, from 0dB to 20dB. Even when the SNR is very low, i.e. 0dB, antenna management consumes only about 1% more energy per bit than the theoretically optimal configuration, which is completely negligible in comparison to the reduction achieved from a static configuration.

VIII. PROTOTYPE-BASED EVALUATION

To evaluate antenna management under realistic channels, in this section we measure its performance with a prototype implementation. The results have confirmed the energy efficiency benefit of antenna management and meanwhile demonstrated its feasibility in practical MIMO network interfaces.

A. *Prototype Implementation*

We have built a prototype of antenna management using an open-access wireless research platform, called WARP [23]. We used a 4x4 MIMO configuration consisting of two WARP boards each with four RF daughterboards. Each daughterboard has a MAX2829 transceiver [21] and is connected with one passive antenna. We used a Lenovo ThinkPad T400 laptop to interact with two WARP nodes, via WARPLab, a MATLAB-based interface that enables the laptop to send commands to and collects experimental data from WARP nodes. An Ethernet switch is used to enable interaction between two WARP nodes and the laptop. Due to space limitation, we only report the results for one-ended management, which is informative enough to show the feasibility and practical benefit of antenna management.

We have emulated an 802.11n WNIC in the MATLAB interface. Most PHY functionalities are realized, including the PLCP preamble and header, modulation and coding, as well as transmit power control. In accordance to 802.11n, we implement different modulation schemes including BPSK and MQAM (M=4, 16, 64), as well as different coding rates including 1/2, 2/3, 3/4, and 5/6. Therefore, data rate is not continuously available. In addition, we set the resolution of transmit power as 1dBm in the range of 0dBm-20dBm (1-100mW) without violating regulations on the maximal transmit power of 802.11n. Since our prototype-based evaluation is limited to

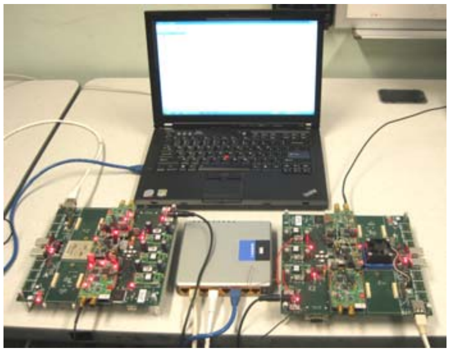


Fig. 6. Experimental setup for prototype-based evaluation: one WARP node with antenna management emulates the mobile node, the other WARP node with legacy 802.11n emulates the access point, and one laptop controls both nodes as well as collects data.

a single MIMO link, we omit most MAC functionalities and only emulate MAC frame generation and collection without real data in the frame body. We also highlight that our prototype uses a simple searching method to find the optimal antenna configuration without antenna selection algorithms. This is because one-ended antenna management with four RF chains and antennas used in our experiment only yields 16 possible configurations so that no advanced algorithm is needed. Nonetheless, for two-ended management and larger number of RF chains or antennas, more efficient algorithms can be employed as specified in Section VI.

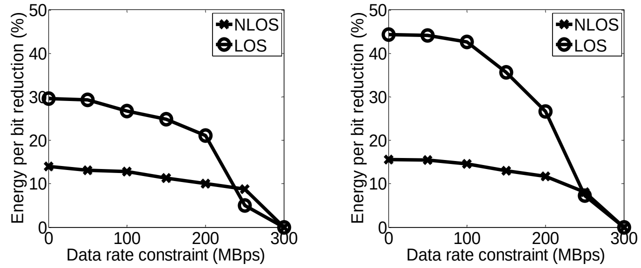
B. *Experimental Setup*

Fig. 5 shows our experimental setup. The evaluation involves two WARP nodes: one as the mobile node which implements antenna management and the other as the access point with all antennas active all the time. The radio is at 2.4GHz, and the channel is 20MHz. We choose two typical indoor environments: one has a dominant LOS path between two nodes (called the LOS scenario) and the other one does not have any LOS path (called the NLOS scenario). For the LOS scenario we put two WARP nodes in an office without any scatter between them, and for the NLOS scenario we put two WARP nodes in two adjacent offices apart from each other by a wall. While the WARP nodes are connected to the Ethernet switch via a wired cable thereby can be hardly moved in large-scale, we do introduce small-scale mobility to the nodes by manually moving them as well as the scatters. Therefore, the channel is supposed to have small-scale fading.

We generate null MAC frames at the mobile node and they are sent as data frames to the access point. A null frame has no meaningful data in the frame body and is only used to mimic realistic 802.11n data frames. In addition, since we have addressed the difference between continuous and intermittent traffic in the simulation, we only employ continuous traffic for the experimental evaluation.

C. *Measurement Results*

We report measurement results for two different MIMO links: 2x2 and 4x4. We employ the same experimental setup for both of them as specified above. For each of the two, we compare antenna management with a static configuration with



(a) 2x2 MIMO (b) 4x4 MIMO

Fig. 5. Energy per bit reduction by antenna management in prototype-based evaluation

all antennas active, as adopted in the simulation. To guarantee fair comparison, we evaluate antenna management online, while emulate the static configuration offline using the collected channel traces. Since the static configuration is transparent to energy awareness, its energy per bit is uniquely determined by the channel so that offline emulation is expected to offer the same results as an online evaluation.

The experimental results are presented in Fig. 6. We have shown the energy per bit reduction by antenna management with different data rate constraints for two MIMO links, i.e. 2x2 and 4x4, each in both LOS and NLOS scenario. While the experimental results for the 4x4 MIMO link echo the simulation results in Section VII, one can see that antenna management becomes less effective in energy efficiency improvement for the 2x2 MIMO link. This is consistent with the fact that a 2x2 MIMO link yields fewer choices of configuration so that the static configuration is more likely to be optimal. It is then apparent that as the number of RF chains and antennas in mobile systems becomes larger and larger with the progress of IC technology (Moore's Law), higher diversity of the antenna configuration can make antenna management increasingly energy efficient.

IX. DISCUSSIONS

In this section we discuss the limitations as well as possible extensions of antenna management.

A. *Limitations*

Our work is limited in the following aspects that can be addressed in future work. First, we only considered the energy impact of the MIMO transceiver front-end without the baseband. In fact, the number of active antennas also impacts the processing load in the baseband, especially for the receiver. Typically, the more antennas, the more processing and higher power consumption in baseband. Therefore, antenna management will conserve energy in baseband too by using a smaller number of active antennas. Second, our work only considers spatial multiplexing MIMO. Multiple antennas and RF chains are also used by other techniques, e.g. diversity-based MIMO or beamforming. Antenna management can be potentially used for these techniques should energy efficiency becomes a concern.

B. *Suitability for Cellular Networks*

Antenna management can be used in cellular networks as well, especially in fourth-generation (4G) mobile broadband communication [24], albeit our current design reported in this work targets WLAN. Current base stations in cellular networks have already embraced multi-antenna techniques, e.g. switched-beam systems [25] or beamforming systems [26]. MIMO is undergoing the technical examination and is very likely to be used in the foreseeable future [27]. Besides the base station, even the mobile client can employ multiple antennas, in order to increase uplink data rate.

One-ended antenna management can be suitable for the mobile client in cellular networks. Nonetheless, important dissimilarities between WLAN and cellular networks need to be taken into account. For example, uplink and downlink channels in cellular networks are usually not reciprocal, especially for networks operating in frequency-division-duplex (FDD) mode. Therefore, implicit channel estimation is no longer effective and an alternative design is needed. We also note that since cellular channels are more likely to be outdoors, where the propagation path is closer to LOS. Recall our results showed that antenna management reduces more energy per bit when the path is closer to LOS. Therefore, one can expect even higher energy efficiency benefit for the cellular network implementation of antenna management.

C. *Network Impact of Antenna Management*

Antenna management may induce network impact, by changing the data rate for a single link. Our current implementation targets a single link in the network without taking into account the energy efficiency or data rate of other links. In fact, antenna management can be potentially detrimental to other links, especially for CSMA-based networks. For instance, in a typical 802.11 network with one access point and multiple associated mobile nodes, those mobile nodes with antenna management may occupy the channel for longer time. Consequently the throughput of other nodes may decrease and their power consumption will increase, since they need stay idle for longer time to wait for the shared channel to become free. As a result, antenna management becomes less effective in heavily loaded networks.

However, antenna management can be considered as an end-to-end rate adaptation mechanism for MIMO links, and all end-to-end rate adaptation protocols will produce positive/negative impact on the network if using a higher/lower data rate [28].

X. Conclusion

In this paper, we presented a power-saving mechanism, *antenna management*, to maximize the energy efficiency of the MIMO network interface on mobile systems. Antenna management adaptively optimizes the transmit power and antenna configuration in order to achieve minimum energy per bit under a given data rate constraint.

We showed that antenna management can be realized with little change to the 802.11n protocol to maximize the energy efficiency of a single end or both ends of communication. Our evaluation using both MATLAB-based simulation and prototype-based experiment demonstrated that on average 13% two-end energy per bit reduction and 21% one-end energy per bit reduction can be achieved.

Acknowledgment

The work was supported in part by NSF awards ECCS/IHCS 0925942, CNS/NeTS-WN 0721894, CNS/CRI 0751173, CNS-0551692, CNS-0551692, and support from the TI Leadership University program. The authors would like to thank the anonymous reviewers for their constructive suggestions that improved the work.

Appendix A

We discuss the nine cases of the MIMO energy per bit minimization in TABLE I and show that we can reduce them to three non-overlapped cases that are of our interest.

Firstly, when the power consumption of both ends are considered, optimizing a single end becomes a special case of optimizing both, e.g. one can regard the transmitter optimization as the optimization for both ends with a fixed receiver antenna configuration, and similarly for the receiver optimization. Therefore, Case 2 and Case 3 are inherently included in Case 1. Analogically Case 4 can be broken down into Case 5 and Case 6, and Case 7 can be broken down into Case 8 and Case 9.

Secondly, when the power consumption of a single end is considered, optimization at the other end has no impact on the power consumption P and only changes the data rate R. For Case 6 and Case 8, the minimization of E_b is equivalent to the maximization of R, and this is in fact the capacity maximization problem without considering power consumption. It has been solved by antenna selection techniques and is out of the scope of this work. In other words, Case 4 is equivalent to Case 5, and Case 7 is equivalent to Case 9.

To summarize, Case 1, Case 5 and Case 9, are the only three cases that do not overlap with each and cannot be solved by existing techniques.

Appendix B

Here we prove the existence of unique P_{TX_opt} that minimizes E_b. Toward this, we only need to prove that there is a single minima of the objective function E_b. While calculating the second-order derivative of E_b is not straightforward, we can alternatively show the following two facts which equivalently prove our statement:

(1) The upper bound of E_b is infinite;

(2) $\frac{dE_b}{dP_{TX}} = 0$ has a unique solution in $(0, +\infty)$.

The fact in (1) is obvious given that $E_b(P_{TX} = +\infty) = +\infty$. For the fact in (2), we rewrite E_b as

$$E_b = \frac{(1+\alpha)P_{TX} + N_T P_{RF\_chain\_T} + N_R P_{RF\_chain\_R} + P_{Shared\_T} + P_{Shared\_R}}{B\sum_{i=1}^{r}\log\left(1+\frac{P_{TX}}{N_T N_0}\lambda_i\right)}$$

where r is the rank of $\boldsymbol{H}\boldsymbol{R}_{\mathrm{SS}}\boldsymbol{H}^H$ and $\lambda_i (i = 1,2, \cdots, r)$ denotes the positive eigenvalues of $\boldsymbol{H}\boldsymbol{R}_{\mathrm{SS}}\boldsymbol{H}^H$. Then $\frac{dE_b}{dP_{TX}} = 0$ implies

$$\sum_{i=1}^{r} \left[\frac{\lambda_i (P_{TX}+C)}{N_T N_0+\lambda_i P_{TX}} - ln\left(1+\frac{P_{TX}}{N_T N_0}\lambda_i\right)\right] = 0, \quad (1)$$

where

$$C = \frac{N_T P_{RF\_chain\_T} + N_R P_{RF\_chain\_R} + P_{Shared\_T} + P_{Shared\_R}}{1+\alpha}.$$

Then we show that equation (1) always has a unique solution of P_{TX} where $0 < P_{TX} < +\infty$. First we define

$$f(P_{TX}) = \sum_{i=1}^{r} \left[\frac{\lambda_i (P_{TX}+C)}{N_T N_0+\lambda_i P_{TX}} - ln\left(1+\frac{P_{TX}}{N_T N_0}\lambda_i\right)\right],$$

and to prove that $f(P_{TX}) = 0$ has a single solution of P_{TX} in $(0, +\infty)$, we can equivalently prove (i) $f(0) > 0$, (ii) $f(+\infty) < 0$, and (iii) $f'(P_{TX}) < 0$, since $f(P_{TX})$ is a continuous function. They are respectively proved as follows.

(i) $f(0) = \frac{C}{N_T N_0}\sum_{i=1}^{r}\lambda_i > 0$, because $C > 0$ and λ_i are the positive eigenvalues of $\boldsymbol{H}\boldsymbol{R}_{\mathrm{SS}}\boldsymbol{H}^H$;

(ii) $f(+\infty) = \lim_{P_{TX}\to+\infty} f(P_{TX}) = r[1 - ln(+\infty)] = -\infty < 0$;

(iii) $f'(P_{TX}) = \sum_{i=1}^{r}\left[\frac{\lambda_i(N_T N_0 - C\lambda_i)}{(N_T N_0+\lambda_i P_{TX})^2} - \frac{\lambda_i}{(N_T N_0+\lambda_i P_{TX})}\right] = \sum_{i=1}^{r}\left[\frac{-\lambda_i{}^2}{(N_T N_0+\lambda_i P_{TX})^2}(C+P_{TX})\right] < 0, 0 < P_{TX} < +\infty.$

It is apparent that R_{min} is a monotonic function with respect to P_{TX}. Therefore, given R_{min}, there exists a minimum constraint of P_{TX} denoted as $P_{TX,min}$, or $P_{TX} \geq P_{TX,min}$. However, the uniqueness of the optimal transmit power still holds when the domain of P_{TX} changes from $(0, +\infty)$ to $[P_{TX,min}, +\infty)$. We briefly show the proof in the following.

(i) If $P_{TX,min} \leq P_{TX,opt}$, the optimal transmit power is identified as $P_{TX,opt}$ since it is the single minima of E_b as elaborated above;

(ii) If $P_{TX,min} > P_{TX,opt}$, the optimal transmit power is identified as $P_{TX,min}$ since when $P_{TX} > P_{TX,opt}$, E_b is an increasing function of P_{TX} so that $P_{TX,min}$ corresponds to the minimum E_b.

APPENDIX C

Here we show that when considering idle period of the MIMO transceiver, the problem formulation is intrinsically identical as the one studied in Section IV.

We use a single 802.11n frame transmission as example to explain the above claim. Consider that a frame arrives from upper layers and is about to be sent by the transceiver. The transceiver can send it at different data rates R. We denote the frame length as L, and then the transceiver will spend L/R in the transmit mode. We further assume that it takes t_R for the receiver to acknowledge this frame and the next frame from upper layers arrives t_P after the last one came. Then the energy per bit for this frame transmission is

$$E_b = \frac{(L/R)P_{Transmit} + t_R P_{Receive} + (t_P - t_R - L/R)P_{Idle}}{L} = \frac{(L/R)(P_{Transmit} - P_{Idle}) + t_R(P_{Receive} - P_{Idle}) + t_P P_{Idle}}{L},$$

where R is the only parameter that the transceiver is able to specify, by tuning transmit power P_{TX} and antenna configuration ω. Therefore the optimal P_{TX} and ω is given by

$$\{P_{TX\_opt}, \omega_{opt}\} = \arg\min_{P_{TX},\omega} \frac{P_{Transmit}(P_{TX},\omega) - P_{Idle}}{R(P_{TX},\omega)}.$$

Since P_{Idle} is constant, it can be easily combined into the constant part of $P_{Transmit}$ as indicated in Section II. Therefore, the problem formulation with consideration of idle period is essentially identical to the one in Section IV.

REFERENCES

[1] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications - A key to gigabit wireless," *IEEE Proc.*, vol. 92, pp. 198-218, Feb 2004.

[2] D. Gesbert, M. Shafi, D.-s. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Journal on Selected Areas in Communications,* vol. 21, pp. 281-302, 2003.

[3] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Information Theory,* vol. 49, pp. 1073-1096, 2003.

[4] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Communications,* vol. 42, pp. 68-73, Oct 2004.

[5] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine,* vol. 5, pp. 46-56, 2004.

[6] V. Erceg, L. Schumacher, P. Kyritsi, A. Molisch, and D. S. Baum, "TGn channel models," in *IEEE document 802.11-03/940r2*, 2004.

[7] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Information Theory,* vol. 52, pp. 2203-2214, 2006.

[8] S. G. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications,* vol. 22, pp. 1089-1098, Aug 2004.

[9] R. W. Heath, Jr. and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int. Conf. Communications (ICC)*, 2001, pp. 2276-2280 vol.7.

[10] A. F. Molisch, M. Z. Win, and J. H. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int. Conf. Communications (ICC)*, 2001, pp. 570-574 vol.2.

[11] F. Rey, M. Lamarca, and G. Vazquez, "Robust power allocation algorithms for MIMO OFDM systems with imperfect CSI," *IEEE Trans. Signal Processing,* vol. 53, pp. 1070-1085, 2005.

[12] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommunications,* vol. 10, pp. 585-595, Nov-Dec 1999.

[13] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-Hop wireless networks," in *Proc. Int. Conf. Mobile Computing and Networking (MobiCom)* Rome, Italy: ACM, 2001.

[14] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," in *Proc. Int. Conf. Mobile Computing and Networking (MobiCom)* Atlanta, Georgia, USA: ACM, 2002.

[15] J. Camp and E. Knightly, "Modulation rate adaptation in urban and vehicular environments: cross-layer implementation and experimental

evaluation," in *Proc. Int. Conf. Mobile Computing and Networking (MobiCom)* San Francisco, California, USA: ACM, 2008.

[16] M. Z. Siam, M. Krunz, A. Muqattash, and S. Cui, "Adaptive multi-antenna power control in wireless networks," in *Proc. Int. Conf. Wireless Communications and Mobile Computing (IWCMC)* Vancouver, British Columbia, Canada: ACM, 2006.

[17] K. Hongseok, C. Chan-Byoung, G. de Veciana, and R. W. Heath, "A Cross-Layer Approach to Energy Efficiency for Adaptive MIMO Systems Exploiting Spare Capacity," *IEEE Trans. Wireless Communications,* vol. 8, pp. 4264-4275, 2009.

[18] H. Yu, L. Zhong, and A. Sabharwal, "Adaptive RF Chain Management for Energy-Efficient Spatial-Multiplexing MIMO Transmission," in *Proc. Int. Sym. Low Power Electronics and Design (ISLPED)* San Francisco, CA, USA ACM, 2009.

[19] J. Liu and L. Zhong, "Micro power management of active 802.11 interfaces," in *Proc. Int. Conf. Mobile Systems, Applications, and Services (MobiSys)* Breckenridge, CO, USA: ACM, 2008.

[20] H. Rahul, F. Edalat, D. Katabi, and C. G. Sodini, "Frequency-aware rate adaptation and MAC protocols," in *Proceedings of the 15th annual international conference on Mobile computing and networking* Beijing, China: ACM, 2009.

[21] Maxim Integrated Product, *Data sheet for MAX2829 Single/Dual-Band 802.11a/b/g World-Band Transceiver ICs*.

[22] L. Lin, N. Wongkomet, D. Yu, C.-H. Lin, M. He, B. Nissim, S. Lyuee, P. Yu, T. Sepke, S. Shekarchian, L. Tee, P. Muller, J. Tam, and T. Cho, "A fully integrated 2x2 MIMO dual-band dual-mode direct-conversion CMOS transceiver for WiMAX/WLAN applications," in *IEEE Int. Conf. Solid-State Circuits (ISSCC), 2009.*, 2009, pp. 416-417,417a.

[23] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine, ,* vol. 42, pp. 74-80, 2004.

[24] S. Y. Hui and K. H. Yeung, "Challenges in the migration to 4G mobile systems," *IEEE Communications Magazine,* vol. 41, pp. 54-59, 2003.

[25] B. Ming-Ju, G. L. Stuber, and M. D. Austin, "Performance of switched beam smart antennas for cellular radio systems," in *IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC'96)*, 1996, pp. 545-549 vol.2.

[26] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications,* vol. 16, pp. 1437-1450, 1998.

[27] A. Hottinen, M. Kuusela, K. Hugl, J. Zhang, and B. Raghothaman, "Industrial embrace of smart antennas and MIMO," *IEEE Wireless Communications,* vol. 13, pp. 8-16, 2006.

[28] A. J. Goldsmith, "The capacity of downlink fading channels with variable rate and power," *IEEE Trans. Vehicular Technology,* vol. 46, pp. 569-580, 1997.

Hang Yu received his B.S. from Tsinghua University in 2008. Since Fall 2008, he has been a Ph.D. student in the Department of Electrical & Computer Engineering, Rice University. His research area is energy-efficient wireless system design, including MIMO and Beamforming systems. He is a student member of IEEE.

Lin Zhong received his B.S. and M.S. from Tsinghua University in 1998 and 2000, respectively. He received his Ph.D. from Princeton University in September, 2005. He was with NEC Labs, America, for the summer of 2003 and with Microsoft Research for the summers of 2004 and 2005. He joined the Department of Electrical & Computer Engineering, Rice University as an assistant professor in September, 2005. He received the AT&T Asian-Pacific Leadership Award and the Harold W. Dodds Princeton University Honorific Fellowship. He coauthored a paper that is selected as one of the 30 most influential papers in the first 10 years of Design, Automation & Test in Europe conferences. He and his students received the best paper award from ACM MobileHCI 2007 and the Mark Weiser best paper award from IEEE PerCom 2009. His research interests include mobile & embedded system design, human-computer interaction, and nanoelectronics. His research has been funded by National Science Foundation, Motorola Labs, Texas Instruments, and Microsoft Research.

Ashutosh Sabharwal (S'90, M'99, SM'06) is a faculty of Electrical and Computer Engineering and Director of Center for Multimedia Communications at Rice University. He received the B.Tech. degree from Indian Institute of Technology, New Delhi in 1993 and the M.S. and Ph.D. degrees from the Ohio State University in 1995 and 1999 respectively. He was the recipient of Presidential Dissertation Fellowship award in 1998. His main research interests are information theoretic foundations, protocols and platforms for high performance wireless networks.