An Approach to Sensor Network Throughput Enhancement by PHY-Aided MAC

Taejoon Kim, *Member, IEEE*, David J. Love, *Senior Member, IEEE*,
Mikael Skoglund, *Senior Member, IEEE*, and Zhong-Yi Jin, *Member, IEEE*

***Abstract*—Low power sensor networks with communication enabled by WiFi are expected to be widely deployed. A major challenge is collecting event-driven uplink data from a large number of low-power sensors with low latency. In WiFi, the access point (AP) typically polls nodes individually to schedule uplink transmission times, resulting in a large latency. In this paper, we present a physical (PHY) layer-aided medium access control (MAC) framework to enhance the uplink throughput of sensor data traffic. In the approach, the acknowledgements from the sensor nodes to the poll message are parallelized. By detecting the parallel acknowledgement, the AP knows which nodes have data to send and allocates channel resources by sending a pull message. This approach is referred to as the probe and pull MAC (PPMAC) mechanism. Our scheme is based on maximizing the achievable throughput of PPMAC by optimizing the PHY layer components. More precisely, we investigate the parallel acknowledgement detector design problem and develop a non-convex optimization framework that maximizes the PPMAC throughput by optimizing the parallel acknowledgement detection statistics. Numerical examples illustrate that PPMAC outperforms the point coordination function (PCF) and distributed coordination function (DCF) mechanisms, standardized in IEEE 802.11, in terms of the achievable throughput and the overhead.**

***Index Terms*—PHY-aided MAC, probe and pull MAC (PPMAC), cross-layer optimization, parallel acknowledgement.**

I. Introduction

THE recent growth in demand for smart environment technologies necessitates technical innovations in wireless sensor networks. One of the most prominent applications is in smart grid systems such as smart utility networks (SUN) which have already been standardized in the IEEE 802.15.4g wireless personal area networks (WPAN) standard [2]. Communications with the sensors are expected to be enabled by WiFi technology based on a modification of the current IEEE 802.11 standards family. A new sensor networking standard IEEE 802.11ah [3] is actively being developed. For instance, 802.11ah supports various applications requiring thousands of sensors spread in an area of radius up to 1 km.

One common requirement in sensor network applications is the need to support uplink sensor traffic with low latency. The key to attaining this goal is to maximize the achievable uplink throughput. However, sensor networks typically rely on battery-powered sensor nodes and can not support many of the power-hungry, but throughput boosting, operations commonly found in most wireless broadband systems. For example, packet retransmission is not suitable for sensor operations due to the large energy consumption from control signaling and retransmissions. Instead, energy efficient techniques that enhance reliability such as collaborative transmission [4], [5], efficient outer coding schemes [6], [7], and sensor data fusion [8], [9] are more suitable for sensors. These techniques show a higher tolerance to packet losses than packet retransmission when compared in a lower energy consumption setting.

Despite these physical layer enhancements, more must be done to improve sensor network performance. To obtain further improvement, it is critical to focus on how the sensor network performs channel utilization at the medium access control (MAC) level. Sensor networks in smart environments pose, however, unique communication challenges including dealing with a large numbers of nodes and unique traffic conditions (e.g., low duty-cycle, bursty, and event-driven traffic). This makes the MACs of conventional WiFi networks ill-suited for sensor applications. One popular contention-based MAC, the distributed coordinated function (DCF), has received considerable attention and been widely deployed [10]–[12]. We note, however, that DCF is unfair and inefficient in terms of energy efficiency and throughput, especially when hundreds of sensors contend for the channel simultaneously [13]–[15].

To address the limitations of a contention-based MAC, IEEE 802.11 standards provide an alternative contention free MAC/scheduling mechanism called the point coordination function (PCF) [10], [16]. The absence of collision alleviates the need to backoff and provides a bounded delay and energy saving [16]–[20], making it suitable for delay sensitive and battery-powered sensor applications. However, the sequential polling mechanism of PCF increases the waiting time for individual nodes when there is a large number of lightly loaded nodes in the network, incurring significant loss of efficiency. On the other hand, when the traffic load is light, the popular DCF scheduler can be an appropriate alternative. This demonstrates

Manuscript received January 1, 2014; revised May 5, 2014 and July 13, 2014; accepted August 25, 2014. Date of publication September 9, 2014; date of current version February 6, 2015. Parts of this work were previously presented at the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)Sydney, Australia, September 2012. The associate editor coordinating the review of this paper and approving it for publication was Z. Han.

T. Kim is with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: taejokim@cityu.edu.hk).

D. J. Love is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906 USA (e-mail: djlove@ecn.purdue.edu).

M. Skoglund is with the Communication Theory Laboratory, Royal Institute of Technology (KTH), 100 44 Stockholm, Sweden (e-mail: skoglund@ee.kth.se).

Z.-Y. Jin is with the Qualcomm Inc., San Diego, CA 92121 USA (e-mail: zhongyij@qti.qualcomm.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2014.2356507

TABLE I
SUMMARY OF SYMBOLS

Symbol	Remark	Symbol	Remark	Symbol	Remark
t_{beacon}	duration of Beacon	t_{pack}	duration of parallel ACK	t_{data}	duration of Data packet
t_{ack}	duration of ACK packet	$t_{data+ack}$	duration of Data+ACK packet	$t_{poll+ack}$	duration of Poll+ACK packet
t_{cfend}	duration of CF-end packet	t_{sifs}	short inter frame space (SIFS)	t_{pifs}	PCF inter frame space (PIFS)
T_s	duration of superframe	t_{cp}	duration of cyclic prefix (CP)	t_{rtt}	maximum round trip time
t_{poll}	duration of Poll packet	t_{seq}	net duration of parallel ACK	t_s	sample interval
t_{gi}	guard interval	t_d	maximum channel delay spread	$p_{e,poll}$	error rate of Poll packet
$p_{e,pull}$	error rate of Pull packet	$p_{e,data}$	error rate of Data packet	$p_{e,ack}$	error rate of ACK packet
$p_{e,cfend}$	error rate of CF-end	$p_{e,data+ack}$	error rate of Data+ACK packet	$p_{e,beacon}$	error rate of Beacon

the need for an efficient scheduler that alleviates the inefficiencies of DCF and PCF.

Recently, the probe and pull MAC (PPMAC), which uses a form of multiplexing, has been proposed [1], [21]. The central idea is to parallelize the polling operation by allowing the sensors with uplink data to concurrently transmit acknowledgements to the AP. We call this operation *parallel acknowledgment*. Various aspects of the implementation of parallel acknowledgement have been studied in [1]. On reception of the parallel acknowledgements, the AP obtains a list of indices of nodes having data to send, facilitating the channel resource allocation. PPMAC using parallel acknowledgement has the potential to alleviate inefficiencies of DCF and PCF.

In this paper, we consider a network in which a large number of sensors are connected with a single AP. The potential benefits of using PPMAC for sensor networks lie in the *parallel acknowledgement* mechanism. To correctly detect the parallel acknowledgements at the AP, some form of signal orthogonality among the nodes is required. We resolve the simultaneous transmissions in the code domain. In most multiple access schemes, knowledge of the uplink channel state information (CSI) at the receiver (i.e., at the AP) is critical for reliable coherent detection. This problem, however, is not found in PPMAC because the parallel acknowledgement is a non-coherent multiple access scheme. In the absence of side information, the parallel acknowledgement detection faces two major challenges: i) the design of a non-coherent, but robust, parallel acknowledgement detector and ii) adaptation to throughput performance. The performance of the parallel acknowledgement detection needs to be optimized so as to maximize the achievable throughput. However, the throughput optimization of PPMAC is non-convex and infeasible to be directly solved. The contributions of the paper are summarized as follows:

- We design a robust, non-coherent parallel acknowledgement detector. The proposed detector is a uniformly most powerful (UMP) detector [22] that maximizes the detection probability for a given target false alarm rate.
- We analyze the achievable throughput of PPMAC by integrating unsaturated traffic conditions, channel imperfections, and parallel acknowledgement detection statistics into a closed-form expression.
- We devise a cross-layer optimization framework for PPMAC which maximizes the achievable throughput by optimizing the parallel acknowledgement detection statistics. A discrete monotonic optimization technique, motivated by the branch-reduce-and-bound algorithms [23], is derived to solve the formulated optimization problem.
- We compare the performance of the optimized PPMAC with the PCF and DCF mechanisms under unsaturated traffic conditions. For the comparison, we also derive the achievable throughput of PCF.

The cross-layer optimized PPMAC is shown to outperform PCF and DCF. Furthermore, PPMAC exhibits an excellent ability to adapt the channel allocation to a wide spectrum of traffic conditions, making it suitable for bursty sensor applications. We are not aware of similar work that treats specific wireless access problems such as scheduling, medium assignment, and throughput maximization as part of the parallel acknowledgement optimization problem using discrete monotonic optimization tools like those in [23], [24]. Though we propose the framework to optimize the channel utilization for wireless sensor networks, by being able to solve the optimization problem formulated in the paper, we hope that our approach will inspire the use of monotonic optimization techniques in other related wireless communication problems.

The rest of the paper is organized as follows. In Section II, the system model is explained. Section III designs the parallel acknowledgement detector and analyzes its performance. Section IV investigates useful properties of the detector statistics and introduces the monotonic optimization framework. Section V derives a closed-from expression of the achievable throughput of PPMAC, and the throughput optimization technique is developed in Section VI. In Section VII, the achievable throughput of PPMAC is compared with that of PCF and DCF, before concluding the paper in Section VIII.

II. SYSTEM OVERVIEW

In this section, we present the system model and comment on the assumptions made. We start by presenting the backgrounds of the contention free MACs, PCF, and PPMAC. The symbols summarized in Table I are used throughout the paper.

A. Point Coordination Function (PCF) Scheduler

Consider a network with one AP and $U_{\max}$ sensors as shown in Fig. 1(a). When the PCF scheduler in IEEE 802.11 is implemented, the channel access within a frame alternates between PCF and DCF [17]. The contention free period (CFP) begins with the beacon message and ends with the CF-end message[1] being sent by the AP as shown in Fig. 1(a). On receiving CF-end

[1]The beacon message contains information about the rate of the CFP, a maximum duration of the CFP, and a list of nodes served during the CFP (polling list). The CF-end indicates the end of the CFP.

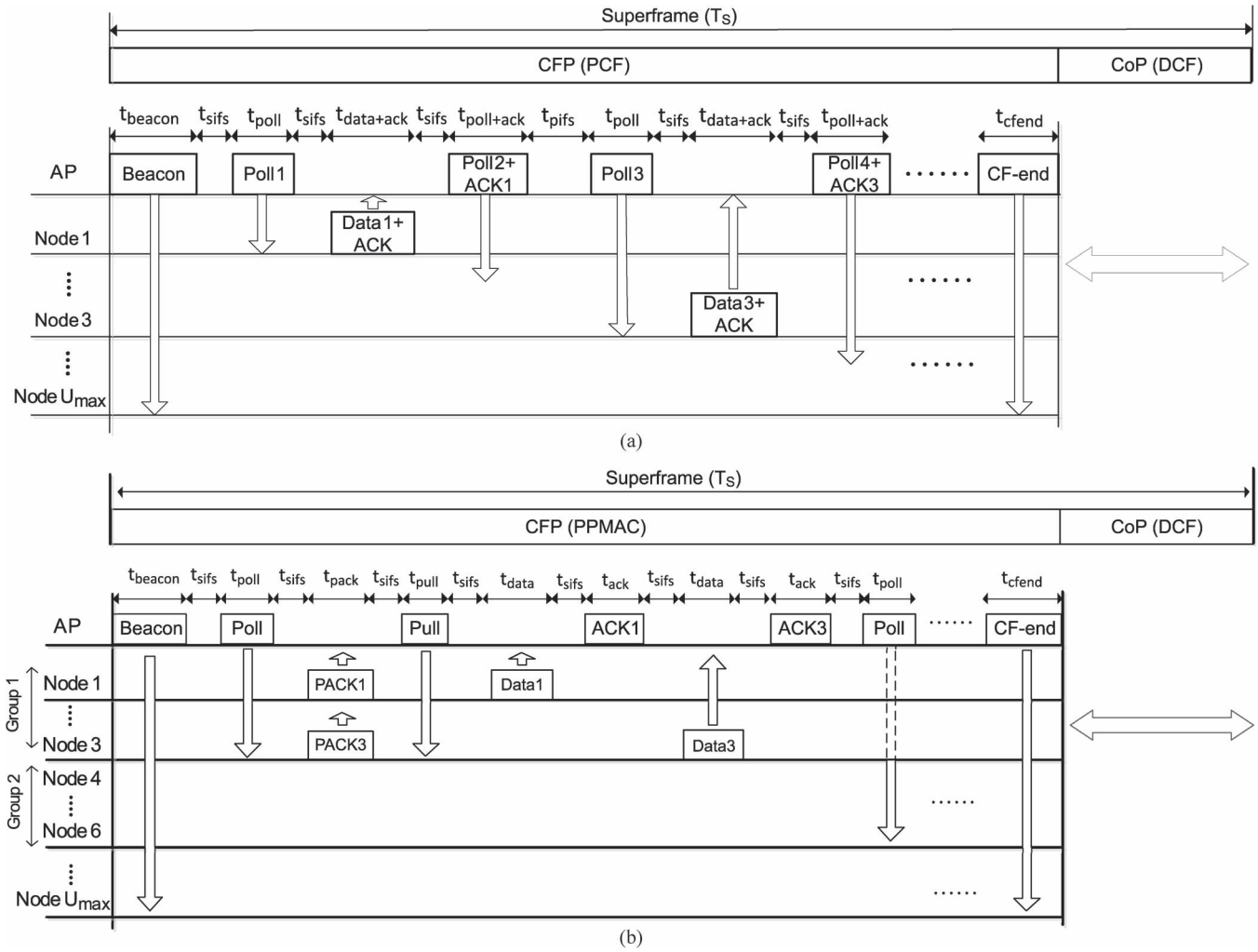

Fig. 1. Contention free medium access: PCF and PPMAC. (a) PCF-based uplink medium access. (b) PPMAC-based uplink medium access.

the contention period (CoP) starts, and the CoP is utilized for DCF to acquire data from other WiFi devices. We call a frame duration where DCF and PCF alternate a *superframe* [17]. The unit for time duration is seconds, but we omit those time units for simplicity.

In Fig. 1(a), Node1 and Node3 are assumed to have uplink sensor data to send. The AP polls each node sequentially after the t_{sifs}, and in response to the poll message, the node transmits the piggybacked Data+ACK containing the payload along with the acknowledgement to the poll message. If the AP receives Data+ACK, it can send Poll+ACK where the poll message is to poll the next node and the ACK part is used to acknowledge the previous data frame. In either cases of no data to transmit or failure to receive the poll message, the node responds with a null frame and the AP proceeds to the next node after the t_{pifs} in Fig. 1(a). In these cases, the channel remains idle, making the protocol considerably inefficient when there is a large number of low traffic nodes.

B. Probe and Pull MAC (PPMAC)

To resolve the inefficiency of the PCF scheduler, the multiplexing-based PPMAC has been proposed [1], [21]. Similar to the PCF scheduler, the channel access alternates between PPMAC and DCF as shown in Fig. 1(b). In PPMAC, $U_{\max}$ sensors can be divided into multiple groups, for which we use U to represent a *group* size. The value $U = 3$ is assumed in Fig. 1(b). The first poll message probes the first three nodes simultaneously in the Fig. 1(b) example. Once the nodes with data to send (i.e., Node 1 and Node 3 in Fig. 1(b)) receive the poll message, they concurrently send acknowledgements to the AP after t_{sifs}, which we call *parallel acknowledgement*. We emphasize here that no packet transmission is required for the parallel acknowledgement, but the node simply transmits a *sequence* (or a signature) to the AP. After resolving the parallel acknowledgement, the AP has information about which node has data to send and allocates time slots to the client nodes via the pull message. After the first group, the next group is processed by the AP. The ability to resolve the parallel acknowledgement is the key to PPMAC throughput enhancement.

Throughout the paper, we assume that energy efficient and reliable PHY techniques (e.g., [4]–[9]) for the battery-powered sensors are employed and packet retransmission is not used during the uplink sensor data delivery phase. We also assume each sensor employs a pre-determined energy efficient and reliable modulation scheme. For instance, the optimal modulation can be determined by employing an energy-constrained modulation

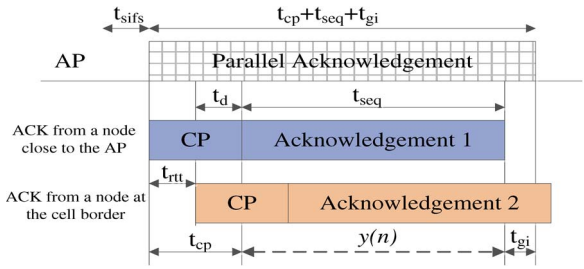

Fig. 2. Signal representation of the received parallel acknowledgement.

optimization technique, e.g., [25], [26].[2] We first model the physical parallel acknowledgement link and formulate the parallel acknowledgement detection problem.

C. Parallel Acknowledgement

We focus on the parallel acknowledgement session in Fig. 1(b) for one group, since other groups follow the same procedure. The received parallel acknowledgements from the closest and farthest nodes at the AP are illustrated in Fig. 2. The time notations used in Fig. 2 are summarized in Table I. The parallel acknowledgement has a cyclic prefix (CP) to combat both *multipath fading* (related to t_d) and *random propagation delay* (related to t_{rtt}) and, at the same time, to maintain a cyclic property across multiple transmissions. Hence, the duration of CP is $t_{cp} \geq t_{rtt} + t_d$. The guard interval t_{gi} avoids interference into the subsequent transmissions. The duration of a parallel acknowledgment session is $t_{pack} = t_{cp} + t_{seq} + t_{gi}$.

We use $\mathcal{U} \subseteq \{1, 2, \ldots, U\}$ to denote the set of indices of the *active nodes* with sensor data to send. Capturing the channel output during t_{seq} in Fig. 2 and sampling it at the sampling rate $1/t_s$ yield the block output[3] $\mathbf{y} = [y(0), \ldots, y(N-1)]^T \in \mathbb{C}^N$ with the block length $N = \left\lceil \frac{t_{seq}}{t_s} \right\rceil$. We write $\mathbf{y}$ as

$$\mathbf{y} = \sum_{u \in \mathcal{U}} \mathbf{v}_u + \mathbf{w} \tag{1}$$

where $\mathbf{w} = [w(0), \ldots, w(N-1)]^T \in \mathbb{C}^N$ is zero-mean complex Gaussian noise with covariance $E[\mathbf{w}\mathbf{w}^*] = \mathbf{I}_N$. If we denote the sequence $\mathbf{s}_u = [s_u(0), \ldots, s_u(N-1)]^T \in \mathbb{C}^N$ as Node u's channel input subject to $\|\mathbf{s}_u\|^2 = N$, the vector $\mathbf{v}_u = [v_u(0), \ldots, v_u(N-1)] \in \mathbb{C}^N$ in (1) is the channel output (excluding noise) of $\mathbf{s}_u$.

The parallel acknowledgement employs Zadoff-Chu (ZC) sequences as the signatures of nodes. Chu showed in his 1972 paper [27] that the sequences defined by

$$z_q(n) = e^{-j\pi q n(n+\bmod(N,2))/N}, \quad n = 0, \ldots, N-1 \tag{2}$$

with $q \in \{1, \ldots, N-1\}$ being the root index of the sequence have *zero cyclic autocorrelation property* such that

$$\mathbf{z}^* \mathbf{Z}_0 = N \mathbf{e}_1^T,$$

$$\mathbf{Z}_0 = \begin{bmatrix} z(0) & z(N-1) & \cdots & z(2) & z(1) \\ z(1) & z(0) & z(N-1) & \cdots & z(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z(N-1) & z(N-2) & z(N-3) & \cdots & z(0) \end{bmatrix}.$$

Here, we treat $\mathbf{z}_q = \mathbf{z} = [z(0), \ldots, z(N-1)]^T \in \mathbb{C}^N$ and suppress the dependency of q on $z_q(n)$ for notational simplicity. The matrix $\mathbf{Z}_0 \in \mathbb{C}^{N \times N}$ is a Toeplitz matrix formed by cyclically shifting $\mathbf{z}$. The zero cyclic autocorrelation property of ZC sequence provides a succinct way of generating multiple orthogonal sequences by applying different *cyclic shifts* to one reference sequence $\mathbf{z}$. Furthermore, each element of (2) is *unit-modulus* and its discrete Fourier transform is also another ZC sequence (for a prime N) [28], meaning ZC sequences exhibit a minimal peak-to-average power ratio.

D. Sequence Dimensioning and Signatures

Given a ZC sequence $\mathbf{z}$, U orthogonal signatures $\{\mathbf{s}_u\}_{u=1}^U$ for each node in a group are generated by applying $0, \ell, \ldots, (U-1)\ell$ cyclic shifts to $\mathbf{z}$ according to $\mathbf{s_u} = \mathbf{Z}_0 \mathbf{e}_{(u-1)\ell+1}$. Each group can share the same set of signatures $\{\mathbf{s}_u\}_{u=1}^U$. We refer to $\ell \geq 1$ as the unit for *cyclic shift dimension* such that $U\ell \leq N$, i.e., $\ell = \lfloor N/U \rfloor$. Notice that the number of cyclic shifts $(u-1)\ell$ identifies Node u.

We express the t_{rtt} and t_d in discrete-time forms as $M = \lceil t_d/t_s \rceil$ and $L = \lceil t_{rtt}/t_s \rceil$, respectively. Hence, the cyclic shift dimension must satisfy $M + L \leq \ell$ for the AP to detect the node (i.e., the signature) by measuring the number of cyclic shifts within the window $[(u-1)\ell \; u\ell - 1]$. Denoting $\ell_{\min} \triangleq M + L$, ℓ must meet

$$\ell_{\min} \leq \ell = \left\lfloor \frac{N}{U} \right\rfloor \leq N, \tag{3}$$

and the U congruent with the condition (3) is

$$1 \leq U \leq U_{\max} = \left\lfloor \frac{N}{\ell_{\min}} \right\rfloor. \tag{4}$$

The value of N satisfying both (3) and (4) is $N = \ell_{\min} U_{\max}$.

E. Signal Model

We rewrite $\mathbf{v}_u$ in (1) as

$$\mathbf{v}_u = \sqrt{\rho} \mathbf{Z}_0 \mathbf{P}_u \mathbf{h}^u, \tag{5}$$

where $\mathbf{P}_u = [\mathbf{e}_{d_u+(u-1)\ell+1} \; \mathbf{e}_{d_u+(u-1)\ell+2} \; \cdots \; \mathbf{e}_{d_u+(u-1)\ell+M}] \in \mathbb{R}^{N \times M}$ and $\mathbf{h}^u = [h_0^u, \ldots, h_{M-1}^u]^T \in \mathbb{C}^M$ is the impulse response of the multipath fading channel of node u with independent and identically distributed (i.i.d.) $h_m^u \sim \mathcal{CN}(0, \sigma_{u,m}^2)$, $\sigma_{u,i}^2 \neq \sigma_{u,j}^2 \geq 0$, $\forall i \neq j$. We, moreover, assume that the channel is normalized in the sense that $\sum_{m=0}^{M-1} \sigma_{u,m}^2 = 1$. The ρ is the signal-to-noise ratio (SNR).

[2] It is shown in [25], [26] that the energy optimal modulation for long-range transmission (e.g., a distance above 30 m) is the lowest modulation order (e.g., binary phase shift keying (BPSK)) with the maximum transmission duration.

[3] A bold capital letter $\mathbf{A}$ is a matrix, a bold lowercase letter $\mathbf{a}$ is a vector, a calligraphic capital letter $\mathcal{A}$ is a set, $\mathbf{I}_N$ is the $N \times N$ identity matrix, $\mathbf{e}_i \in \mathbb{R}^N$ denotes the ith column of $\mathbf{I}_N$, $|\mathcal{A}|$ is the cardinality of $\mathcal{A}$, $\mathbf{A}^T$ and $\mathbf{A}^*$ are the transpose and conjugate transpose of $\mathbf{A}$, respectively, $\|\mathbf{a}\|$ is the vector 2-norm, $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ are the ceiling and flooring operators, respectively, $\mathcal{CN}(a, b)$ is a complex Gaussian distribution with mean a and variance b, $\bmod(a, b)$ is a modulo b, $\delta(n)$ denotes the Kronecker delta function,and $a_k \searrow a$ denotes convergence from above of sequence a_k as $k \to \infty$.

The $\mathbf{P}_u$ models the effects of the number of cyclic shifts and the random propagation delay d_u applied to Node u. The d_u is uniformly distributed such that $d_u \in \{0, 1, \ldots, L\}$. This means that $\mathbf{v}_u$ in (5) is the cyclic cross correlation between $\mathbf{z}$, shifted by $d_u + (u-1)\ell$, and $\mathbf{h}^u$. The nth element of $\mathbf{v}_u$ is, therefore, written as

$$v_u(n) = \sqrt{\rho} \sum_{m=0}^{M-1} \left(z\left(n - d_u - (u-1)\ell - m\right)\right)_N h_m^u$$

with $\left(z(n-d_u-(u-1)\ell-m)\right)_N \triangleq z(\mathrm{mod}(n-d_u-(u-1)\ell-m, N))$ being the cyclic shift. Denoting $\mathcal{U} \triangleq \{u_{u_1}, \ldots, u_{u_{|\mathcal{U}|}}\}$ and plugging (5) in (1) yield

$$\mathbf{y} = \mathbf{Z}_0 \mathbf{v} + \mathbf{w}, \tag{6}$$

where $\mathbf{v} = \sqrt{\rho}[\mathbf{P}_{u_1} \ldots \mathbf{P}_{u_{|\mathcal{U}|}}][(\mathbf{h}^{u_1})^T \ldots (\mathbf{h}^{u_{|\mathcal{U}|}})^T]^T \in \mathbb{C}^N$. The channel state information, including d_u, $\mathbf{h}^u$, and $\mathcal{U}$, is unknown to the AP. We, however, allow that the second order statistics $\{\sigma^2_{u,m}\}_{m=0}^{M-1}$ of the channel are known to the AP. The unknown $\mathbf{d}_u$, $\mathbf{h}^u$, and $\mathcal{U}$ are abstracted out in $\mathbf{v}$ in (6).

III. Parallel Acknowledgement Detection

Based on the signal model presented in Section II, we investigate the parallel acknowledgement detector design problem.

A. Detector Design

The receiver first projects $\mathbf{y}$ in (6) onto its signature subspace $\mathbf{Z}_0$ yielding[4]

$$\hat{\mathbf{v}} = \frac{1}{N}\mathbf{Z}_0^* \mathbf{y} = \mathbf{v} + \bar{\mathbf{w}}, \tag{7}$$

where $\hat{\mathbf{v}} = [\hat{v}(0), \ldots, \hat{v}(N-1)]^T$. The nth element of $\bar{\mathbf{w}}$ is the sample mean $\bar{w}(n) = (1/N)\sum_{i=0}^{N-1}(z^*(n-i))_N w(i)$ that follows zero mean complex Gaussian with $E[\bar{\mathbf{w}}\bar{\mathbf{w}}^*] = \left(\frac{1}{N}\right)\mathbf{I}_N$.

Let $\mathcal{A}_u \triangleq [(u-1)\ell \; u\ell - 1]$ with $|\mathcal{A}_u| = \ell$ be the cyclic shift dimension regime of Node u. The detector tests ℓ observations in $\{\hat{v}(n) : n \in \mathcal{A}_u\}$. For the sake of formality, we define $\{\hat{v}_u(n)\} \triangleq \{\hat{v}(n) : n \in \mathcal{A}_u\}$ and a set of indices corresponding to the delayed channel taps as

$$\mathcal{B}_u \triangleq \{d_u + m + (u-1)\ell\}_{m=0}^{M-1} \subseteq \mathcal{A}_u = [(u-1)\ell \; u\ell - 1]$$

with $|\mathcal{B}_u| = M$. With the above definitions in mind, the nth element of $\hat{\mathbf{v}}$ in (7), when $u \in \mathcal{U}$, can now be rewritten as

$$\hat{v}_u(n) = \sqrt{\rho} \sum_{m=0}^{M-1} h_m^u \delta_{n \in \mathcal{B}_u} + \bar{w}(n), \quad u \in \mathcal{U} \tag{8}$$

where $\delta_{n\in\mathcal{B}_u} \triangleq \delta(n - d_u - m - (u-1)\ell)$. On the other hand, when $u \notin \mathcal{U}$, $\hat{v}_u(n)$ is

$$\hat{v}_u(n) = \bar{w}(n), \quad u \notin \mathcal{U}. \tag{9}$$

[4]For normal Gaussian $\mathbf{w}$, the projection is indeed the maximum likelihood estimate of $\mathbf{v}$, i.e., $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \|\mathbf{y} - \mathbf{Z}_0\mathbf{v}\|^2$ for any fixed $\mathbf{h}^u$.

Define the binary hypotheses

$\mathcal{H}_0 : u \notin \mathcal{U}$, Node u is *inactive* (null hypothesis)

$\mathcal{H}_1 : u \in \mathcal{U}$, Node u is *active* (alternative hypothesis).

The detector declares Node u is active ($\mathcal{H}_1$) by testing the maximum likelihood ratio between $\hat{v}_u(n)$ under $\mathcal{H}_1$ and $\hat{v}_u(n)$ under $\mathcal{H}_0$ against threshold α', i.e.,

$$\max_{n\in\mathcal{A}_u} \left\{ \frac{f\left(\hat{v}_u(n) | u \in \mathcal{U}\right)}{f\left(\hat{v}_u(n) | u \notin \mathcal{U}\right)} \right\} \geq \alpha', \tag{10}$$

where the probability density functions (PDFs) are

$$f\left(\hat{v}_u(n) | u \in \mathcal{U}\right) = \frac{1}{\pi\sigma^2_{u\in\mathcal{U}}} \, e^{-\frac{|\hat{v}_u(n)|^2}{\sigma^2_{u\in\mathcal{U}}}} \tag{11}$$

with $\sigma^2_{u\in\mathcal{U}} \triangleq \rho\sigma^2_{u,m}\delta_{n\in\mathcal{B}_u} + (1/N)$ and

$$f\left(\hat{v}_u(n) | u \notin \mathcal{U}\right) = \frac{1}{\pi\sigma^2_{u\notin\mathcal{U}}} e^{-\frac{|\hat{v}_u(n)|^2}{\sigma^2_{u\notin\mathcal{U}}}}, \; \sigma^2_{u\notin\mathcal{U}} \triangleq \frac{1}{N}. \tag{12}$$

Notice that the PDF $f(\hat{v}_u(n)|u \in \mathcal{U})$ contains the unknown parameter d_u. Unlike typical hypothesis testing where the PDFs under both hypotheses are completely known, we must handle the nuisance (i.e., unknown) parameter to implement (10), raising challenges of the detector design.

B. Uniformly Most Powerful (UMP) Detector and Performance Analysis

A detector is called a *uniformly most powerful (UMP)* detector if a detector maximizes the detection probability over *all possible tests* for a given target false alarm rate [29]. This is, if the detector is a UMP detector, the dependence on nuisance parameters is *illusory*. We, moreover, know from *Karlin-Rubin's theorem* in detection theory [22] that if the test is *one-sided* (i.e., if the likelihood ratio is a nondecreasing function of observations) and the test statistics under $\mathcal{H}_0$ are independent of the unknown parameters, the UMP detector exists and it is equivalent to the Neyman-Pearson detector. It can readily be shown by Karlin-Rubin's theorem that the detector in (10) is indeed a UMP detector, whose proof is relegated to the Appendix A.

As shown in Appendix A, the equivalent detector to (10) is

$$\max_{n\in\mathcal{A}_u} \lambda_u(n) \geq \alpha, \tag{13}$$

where $\lambda_u(n) = |\hat{v}_u(n)|^2$ and the false alarm rate (the probability of detecting Node u when no sequence is transmitted by Node u) is

$$p_f(\alpha | u \notin \mathcal{U}) = 1 - (1 - e^{-N\alpha})^\ell. \tag{14}$$

Now to formalize, we define a mis-detection as the event of failing to detect the presence of the sequence. Its rate is defined as $p_{md}(\alpha | u \in \mathcal{U}) = \Pr(\max_{n\in\mathcal{A}_u} \lambda_u(n) \leq \alpha | u \in \mathcal{U})$. The variable $\lambda_u(n)$ under $u \in \mathcal{U}$ has different statistics depending on whether $n \in \mathcal{B}_u$ or $n \in \mathcal{A}_u \setminus \mathcal{B}_u$. If $n \in \mathcal{B}_u$, the

$\lambda_u(n)$ is $\lambda_u(n \in \mathcal{B}_u) = |\sqrt{\rho} h_m^u + \bar{w}(n)|^2$ from (8), and it is exponentially distributed with the PDF $f_{\lambda_u(n\in\mathcal{B}_u)}(\lambda|u \in \mathcal{U}) = \frac{1}{\rho\sigma_{u,m}^2+1/N} e^{-\frac{\lambda}{\rho\sigma_{u,m}^2+1/N}}$. If $n \in \mathcal{A}_u \setminus \mathcal{B}_u$, $\lambda_u(n)$ is reduced to $\lambda_u(n \in \mathcal{A}_u \setminus \mathcal{B}_u) = |\bar{w}(n)|^2$. Finally, we have

$$p_{md}(\alpha|u \in \mathcal{U}) = \prod_{m=0}^{M-1} \left(1 - e^{-\frac{N\alpha}{\rho N\sigma_{u,m}^2+1}}\right) \left(1 - e^{-N\alpha}\right)^{\ell - M}. \tag{15}$$

Comparing different detection schemes, in the high SNR region, can be based on characterizing the *diversity gain*. A detection scheme is said to achieve diversity gain ξ if the mis-detection rate $\lim_{\rho\to\infty} \frac{\log p_{md}(\alpha|u\in\mathcal{U})}{\log \rho} = \lim_{\rho\to\infty} \log_\rho p_{md}(\alpha|u \in \mathcal{U}) = -\xi$, and we write $p_{md}(\alpha|u \in \mathcal{U}) \doteq \rho^{-\xi}$.

Theorem 1: The UMP detector (10) constrained to have fixed $p_f(\alpha|u \notin \mathcal{U})$ in (14) has diversity order $\xi = M$, i.e., $p_{md}(\alpha|u \in \mathcal{U}) \doteq \rho^{-M}$.

Proof: Since we are only interested in the ρ exponent, we can ignore $\left(1 - e^{-N\alpha}\right)^{\ell - M}$ in (15). Therefore,

$$\lim_{\rho\to\infty} \log_\rho p_{md}(\alpha|u\in\mathcal{U}) = \lim_{\rho\to\infty} \sum_{m=0}^{M-1} \log_\rho \left(1 - e^{-\frac{N\alpha}{\rho N\sigma_{u,m}^2+1}}\right)$$
$$= \sum_{m=0}^{M-1} \lim_{\rho\to\infty} \log_\rho \left(\frac{N\alpha}{\rho N\sigma_{u,m}^2+1}\right) = -M$$

where the second step follows from the fact that the Taylor series expansion of $\log_\rho \left(1 - e^{-\frac{N\alpha}{\rho N\sigma_{u,m}^2+1}}\right)$ is dominated by the first order term $\log_\rho \left(\frac{N\alpha}{\rho N\sigma_{u,m}^2+1}\right)$ as $\rho \to \infty$. ■

Remark 1: The diversity gain M in Theorem 1 is the number of independently fading channel taps, implying the UMP detector achieves the full multipath diversity inherited in the channel $\{h_m^u\}_{m=0}^{M-1}$.

Remark 2: Up to this point, we investigated per node detection statistics. From (14) and (15), one can write

$$p_{md}(\alpha|u\in\mathcal{U}) = \prod_{m=0}^{M-1} \left(1 - e^{-\frac{N\alpha}{\rho N\sigma_{u,m}^2+1}}\right) \left(1 - p_f(\alpha|u\notin\mathcal{U})\right)^{\frac{\ell-M}{\ell}}.$$

This reveals that the lower the false alarm rate, the more frequent the mis-detection event will be, i.e., a lower false alarm rate comes at the price of sacrificing the detection performance. Seen from the operation of PPMAC, correctly identifying both the *active nodes* (related to the *mis-detection rate*) and *inactive nodes* (related to the *false alarm rate*) for a given *traffic condition* determines the achievable throughput of PPMAC. This means the tradeoff between $p_{md}(\alpha|u \in \mathcal{U})$ and $p_f(\alpha|u \notin \mathcal{U})$ must be resolved by maximizing the PPMAC throughput for a given traffic load. This problem is the main focus for the work presented in the following sections.

IV. Preliminaries

Prior to optimizing the PPMAC, we investigate useful properties of the parallel acknowledgement detection statistics and introduce the basic concept of *monotonic optimization*.

A. Monotonicity

For any two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^l$ we write $\mathbf{a} \le \mathbf{b}$ to mean $a_i \le b_i$ for every $i = 1, \ldots, l$. If $\mathbf{a} \le \mathbf{b}$ then the *box* $[\mathbf{a}\ \mathbf{b}]$ is the set of all $\mathbf{x}$ such that $\mathbf{a} \le \mathbf{x} \le \mathbf{b}$. For notational simplicity, we drop the $u \notin \mathcal{U}$ and $u \in \mathcal{U}$ from (14) and (15), respectively. We first define a discrete set

$$\mathcal{S} = \left\{[U,\ \alpha]^T \in \mathbb{R}^2 | U \in I_+, \alpha \in [0\ \alpha_{\max}]\right\}, \tag{16}$$

where $I_+ = \{1, 2, \ldots, U_{\max}\}$. It is useful in the later derivation to extend the discrete subset I_+ to the real domain $\tilde{I}_+ = \{\widetilde{U} \in \mathbb{R} | 1 \le \widetilde{U} \le U_{\max}\}$ and $\widetilde{U} \in \tilde{I}_+$. Defining

$$\tilde{\mathcal{S}} = \left\{[\widetilde{U},\ \alpha]^T \in \mathbb{R}^2 | \widetilde{U} \in \tilde{I}_+,\ \alpha \in [0\ \alpha_{\max}]\right\}, \tag{17}$$

the false alarm probability extended on $\tilde{\mathbf{x}} \in \tilde{\mathcal{S}}$ is expressed as

$$\tilde{p}_f(\tilde{\mathbf{x}}) = 1 - \left(1 - e^{-N\alpha}\right)^{\frac{N}{\widetilde{U}}}.$$

It is straightforward to verify that the first partial derivatives of $\tilde{p}_f(\tilde{\mathbf{x}})$ with respect to $\widetilde{U}$ and α, which we denote here as $\tilde{p}_f^{(\widetilde{U})}(\tilde{\mathbf{x}}) = \frac{\partial \tilde{p}_f(\tilde{\mathbf{x}})}{\partial \widetilde{U}}$ and $\tilde{p}_f^{(\alpha)}(\tilde{\mathbf{x}}) = \frac{\partial \tilde{p}_f(\tilde{\mathbf{x}})}{\partial \alpha}$, respectively, are $\tilde{p}_f^{(\widetilde{U})}(\tilde{\mathbf{x}}) \le 0$ and $\tilde{p}_f^{(\alpha)}(\tilde{\mathbf{x}}) \le 0$, verifying $\tilde{p}_f(\tilde{\mathbf{x}})$ is monotonically decreasing in $\tilde{\mathbf{x}} \in \tilde{\mathcal{S}}$.

For notational simplicity, we define $\beta_m = \rho N\sigma_{u,m}^2 + 1$ in (15) and write the mis-detection rate extended on $\tilde{\mathbf{x}} \in \tilde{\mathcal{S}}$ as

$$\tilde{p}_{md}(\tilde{\mathbf{x}}) = \prod_{m=1}^{M} \left(1 - e^{-N\frac{\alpha}{\beta_m}}\right) \left(1 - e^{-N\alpha}\right)^{\frac{N}{\widetilde{U}} - M}.$$

It is not difficult to see that $\tilde{p}_{md}^{(\widetilde{U})}(\tilde{\mathbf{x}}) \ge 0$. Taking the first partial derivative of $\tilde{p}_{md}(\tilde{\mathbf{x}})$ with respect to α yields

$$\tilde{p}_{md}^{(\alpha)}(\tilde{\mathbf{x}}) = \sum_{i=1}^{M} \frac{N}{\beta_i} e^{-\frac{N\alpha}{\beta_i}} \prod_{j\ne i}^{M} \left(1 - e^{-\frac{N\alpha}{\beta_j}}\right) \left(1 - e^{-N\alpha}\right)^{\frac{N}{\widetilde{U}} - M}$$
$$+ \left(\frac{N}{\widetilde{U}} - M\right) N e^{-N\alpha} \prod_{i=1}^{M} \left(1 - e^{-\frac{N\alpha}{\beta_i}}\right) \left(1 - e^{-N\alpha}\right)^{\frac{N}{\widetilde{U}} - M - 1}.$$

Since $\frac{N}{\widetilde{U}} \ge M + L - 1 \ge M$ by (3), we have $\tilde{p}_{md}^{(\alpha)}(\tilde{\mathbf{x}}) \ge 0,\ \forall \alpha \in [0\ \alpha_{\max}]$, verifying that $\tilde{p}_{md}(\tilde{\mathbf{x}})$ is monotonically increasing on $\tilde{\mathbf{x}} \in \tilde{\mathcal{S}}$.

Overall, we have for $\tilde{\mathbf{x}} \in \tilde{\mathcal{S}}$

$$\tilde{p}_f^{(\alpha)}(\tilde{\mathbf{x}}) \le 0,\ \tilde{p}_f^{(\widetilde{U})}(\tilde{\mathbf{x}}) \le 0,\ \tilde{p}_{md}^{(\alpha)}(\tilde{\mathbf{x}}) \ge 0,\ \tilde{p}_{md}^{(\widetilde{U})}(\tilde{\mathbf{x}}) \ge 0. \tag{18}$$

Note that if a function $f(\mathbf{x}) : \mathbb{R}^l \to \mathbb{R}$ is monotonic on $\mathbf{x} \in [\mathbf{a}\ \mathbf{b}]$, then $f(\mathbf{x})$ defined on $\mathbf{x} \in \mathcal{A} \subseteq [\mathbf{a}\ \mathbf{b}]$, regardless of whether $\mathcal{A}$ is a connected subset or not, is also monotonic in $\mathcal{A}$. Since $\tilde{\mathcal{S}} \cap \mathcal{S} = \mathcal{S}$, the above monotonicity results also hold for $\mathcal{S}$.

Remark 3: For any two $\mathbf{x}, \mathbf{y} \in \mathcal{S}$, if $\mathbf{x} \le \mathbf{y}$, we have $p_f(\mathbf{x}) \ge p_f(\mathbf{y})$ and $p_{md}(\mathbf{x}) \le p_{md}(\mathbf{y})$.

With the monotonicity established in Remark 3, we adopt discrete monotonic optimization techniques [23], [24] to design the PHY parameter vector $\mathbf{x} = [U, \alpha]^T \in \mathcal{S}$ that maximizes the PPMAC throughput. To this end, we need the general concept of discrete monotonic optimization framework.

B. Discrete Monotonic Optimization

A discrete monotonic optimization problem has a general form

$$\max\{f(\mathbf{x}) = f_+(\mathbf{x}) - f_-(\mathbf{x}) | \mathbf{x} \in [\mathbf{a}\ \mathbf{b}] \cap \mathcal{A}\}. \quad (19)$$

The $f(\mathbf{x})$ is a difference of monotonic functions where $f_+, f_- : \mathbb{R}^l \to \mathbb{R}$ are increasing functions and $\mathcal{A} = \{\mathbf{x} \in \mathbb{R}^l | (x_1, \ldots, x_s) \in \mathcal{D}, s \leq l\}$ is the discrete constraint with $\mathcal{D}$ being a discrete subset. The problem (19) is non-convex programming problem and non-trivial to solve.

The high level description of the iterative branch-reduce-and-bound approach [23] that solves (19) is given. We denote the set of *boxes* at iteration k as $\mathcal{M}_k = \{M_{k_1}, \ldots, M_{k_{|\mathcal{M}_k|}}\}$ such that $\mathcal{M}_k \subset [\mathbf{a}\ \mathbf{b}]$. The algorithm locally approximates the objective function $f(\mathbf{x})$ in each box $M_k \in \mathcal{M}_k$ to an upper bound $\mu(M_k)$ satisfying

$$\mu(M_k) \geq \gamma(M_k) = \max\{f(\mathbf{x}) | \mathbf{x} \in M_k \cap \mathcal{A}\}. \quad (20)$$

It then iteratively refines the upper bounds by constructing a nested sequence of boxes $\{M_k \in \mathcal{M}_k\}$ containing the global maximizer $\mathbf{x}^\star$, i.e.,

$$M_1 = [\mathbf{a}\ \mathbf{b}] \supset M_2 \supset \cdots \supset M_k \supset \cdots \ni \mathbf{x}^\star \quad (21)$$

in such a way that $\mu(M_k) \searrow f(\mathbf{x}^\star)$. In particular, the algorithm starts with the initial box $\mathcal{M}_1 = M_1 = [\mathbf{a}\ \mathbf{b}]$. At each iteration, the box $M \in \mathcal{M}_k$ with the largest upper bound $\sup_{M \in \mathcal{M}_k} \mu(M)$ is selected and branched into two smaller boxes. The sizes of the branched boxes are reduced and the vertices of the reduced boxes are adjusted to align within the discrete constraint $\mathcal{A}$. At each iteration, the algorithm updates two quantities, i.e., the set of upper bounds $\{\mu(M)\}_{M \in \mathcal{M}_k}$ and the current objective function value γ_k satisfying $\gamma_k \leq \sup_{M \in \mathcal{M}_k} \gamma(M)$. To narrow down the search space, some boxes in $M \in \mathcal{M}_k$ yielding $\mu(M) < \gamma_k$ are dropped from the set $\mathcal{M}_k$. In this manner, the algorithm converges as $\mu(M_k) \searrow f(\mathbf{x}^\star)$.

We call $\mathbf{x}$ *feasible* at iteration k if $\mathbf{x} \in M_k$ and $f(\mathbf{x}) \geq \gamma_k$ for $M_k \in \mathcal{M}_k$. The implementation of the algorithm that solves the PPMAC throughput maximization problem is presented in Section VI.

V. PPMAC Throughput Analysis

A closed-form expression of the achievable throughput of PPMAC is of interest. The derived expression is central to deriving the cross-layer optimization framework in Section VI.

A. Assumptions

Given the *group size* U, PPMAC requires a total of $G = \lceil \frac{U_{\max}}{U} \rceil$ rounds of polling and pulling sessions to acquire sensor data from the G groups. When the AP completes the exchange of packets with nodes in the network, we say that an *uplink transaction* is completed.

Let the probability that a node has uplink data to send be $\theta (0 \leq \theta \leq 1)$ independent of u. It is a parameter indicating stationary traffic load (e.g., $\theta = 0.9$ implies heavy traffic, while $\theta = 0.1$ implies light traffic).[5] The packet error rate, defined in Table I, is determined by the channel conditions, modulation/coding scheme, and packet size where these effects are represented by the raw channel bit error rate (BER) p_b. In a slow-fading environment, p_b is assumed to be constant during a frame. Hence, the packet error rate (PER) is

$$p_{e,\eta} = 1 - (1 - p_b)^{z_\eta} \quad (22)$$

where z_η is the length (in bits) of packet $\eta \in \{beacon, poll, pull, data, ack, cfend\}$. Each sensor transmits one fixed size of uplink data packet per frame. After the packet transmission, the sensor simply goes into the sleep mode until the scheduled next beacon is received.

B. Achievable Throughput Analysis

Let J, I, and K be integer-valued joint random variables. Let $J \in \{0, \ldots, U_{\max}\}$ be the number of nodes replying to the poll message by transmitting the parallel acknowledgement, $I \in \{0, \ldots, J\}$ be the number of nodes successfully decoding the pull message, and $K \in \{0, \ldots, I\}$ be the number of nodes declaring successful uplink transaction. In Fig. 1(b), only sensors with data to send that have correctly decoded the beacon and poll messages transmit acknowledgements to the AP. Hence, the probability that the AP correctly decodes the acknowledgement from a node is

$$p_{pa}(\mathbf{x}) = \theta(1 - p_{e,beacon})(1 - p_{e,poll})\,(1 - p_{md}(\mathbf{x})). \quad (23)$$

The probability that I out of $J = j$ nodes correctly receive the pull message and transmit uplink data is

$$P[I = i | j] = \binom{j}{i} (1 - p_{e,pull})^i p_{e,pull}^{j-i}. \quad (24)$$

A successful data transaction is declared once the node receives the acknowledgement (from the AP) to its uplink data transmission. Hence, the probability that K out of $I = i$ nodes declare successful uplink transaction is

$$P[K = k | i] = \binom{i}{k} p_{da}^k (1 - p_{da})^{i-k}$$

where $p_{da} = (1 - p_{e,data})(1 - p_{e,ack})$. The average payload length (in bits) of the uplink transaction of PPMAC is therefore formulated by

$$\mathcal{L}_{PPM} = \nu \sum_{j=0}^{U_{\max}} P[J = j] \sum_{i=0}^{j} P[I = i | j] \sum_{k=0}^{i} k P[K = k | i]$$

where ν is the uplink data payload size (in bits) and $P[J = j] = \binom{U_{\max}}{j} (p_{pa}(\mathbf{x}))^j (1 - p_{pa}(\mathbf{x}))^{U-j}$. Since a binomial random variable X with parameters l and p has $E[X] = \sum_{x=0}^{l} x \binom{l}{x} p^x (1-p)^{l-x} = pl$, the average payload length is simplified to

$$\mathcal{L}_{PPM}(\mathbf{x}) = \nu U_{\max} p_{pa}(\mathbf{x})(1 - p_{e,pull}) p_{da}.$$

[5] In Section VII-A, we model the traffic load θ incorporating a poisson traffic with finite length of queue.

Now, the expected time duration taken to deliver $\mathcal{L}_{PPM}(\mathbf{x})$ is of interest. We use the following notations. $\mathcal{T}_{PP} = t_{poll} + t_{sifs} + t_{pack} + t_{sifs} + t_{pull} + t_{sifs}$ is the combined durations for the poll, parallel acknowledgement, and pull sessions, $\mathcal{T}_D = t_{data} + t_{sifs} + t_{ack} + t_{sifs}$ is the duration of one uplink data packet transaction, and $\mathcal{T}_O = t_{beacon} + t_{sift} + t_{cfend} + t_{sifs}$ denotes the duration of beacon and CF-end. In the event when the AP allocates a time slot $\mathcal{T}_D$ but the node does not transmit data, the channel is wasted. The allocation of the null $\mathcal{T}_D$ slot is caused by one of two independent events: 1) the AP correctly schedules the time slot, but the node misses the pull message, 2) no acknowledgement is transmitted by a node, but the AP allocates the time slot for that node. We shall use two integer-valued random variables J_1 and J_2 to indicate the numbers of null time slots allocated due to events 1) and 2), respectively.

In the event 1), provided $J = j$, we have $J_1 = j - I$ for $0 \le I \le j$. The expected J_1 given $J = j$ is $E_{J_1|j}[J_1 = j - I|j] = \sum_{i=0}^{j} (j-i) \binom{j}{j-i} p_{e,pull}^{j-i} (1-p_{e,pull})^i = \sum_{i=0}^{j} (j-i) P[I = i|j]$ where we use the fact $\binom{j}{j-i} = \binom{j}{i}$ and the equality in (24). Taking the expectation with respect to J to $E_{J_1|j}[J_1 = j - I|j]$ yields

$$E[J_1] = \sum_{j=0}^{U_{\max}} P[J=j] \sum_{i=0}^{j} (j-i) P[I=i|j] = U_{\max} p_{pa}(\mathbf{x}) p_{e,pull}.$$

The event 2) is a network level *false alarm* event, and it happens when a node has no data to send or a node with uplink date to send fails to decode the beacon or the poll messages. Hence, the probability of event 2) is

$$p_{fa}(\mathbf{x}) = \left(1 - \theta(1 - p_{e,poll})(1 - p_{e,beacon})\right) p_f(\mathbf{x}). \quad (25)$$

The expected J_2 is therefore given by

$$E[J_2] = \sum_{j_2=0}^{U_{\max}} \left(p_{fa}(\mathbf{x})\right)^{j_2} \left(1 - p_{fa}(\mathbf{x})\right)^{U_{\max} - j_2} = U_{\max} p_{fa}(\mathbf{x}).$$

Consequently, the expected number of null (i.e., wasted) time slots $\mathcal{N}_{loss} = E[J_1 + J_2] = U_{\max}(p_{pa}(\mathbf{x}) p_{e,pull} + p_{fa}(\mathbf{x}))$. The expected number of time slots in which the uplink data transmissions are attempted is $\mathcal{N}_{use} = E[I] = \sum_{j=0}^{U_{\max}} P[J = j] \sum_{i=0}^{j} i P[I = i|j] = U_{\max} p_{pa}(\mathbf{x})(1 - p_{e,pull})$. Therefore, the total expected time duration for delivering the payload $\mathcal{L}_{PPM}(\mathbf{x})$ is given by

$$\begin{aligned} \mathcal{T}_{PPM}(\mathbf{x}) &= \mathcal{T}_O + G\mathcal{T}_{PP} + \left(\mathcal{N}_{use} + \mathcal{N}_{loss}\right) \mathcal{T}_D \\ &= \mathcal{T}_O + G\mathcal{T}_{PP} + U_{\max} \left(p_{pa}(\mathbf{x}) + p_{fa}(\mathbf{x})\right) \mathcal{T}_D, \end{aligned}$$

The expected achievable throughput $\mathcal{R}_{PPM}(\mathbf{x}) = \mathcal{L}_{PPM}(\mathbf{x}) / \mathcal{T}_{PPM}(\mathbf{x})$ is now formulated by

$$\mathcal{R}_{PPM}(\mathbf{x}) = \frac{U_{\max} \nu p_{pa}(\mathbf{x})(1 - p_{e,pull}) p_{da}}{\mathcal{T}_O + \left\lceil \frac{U_{\max}}{U} \right\rceil \mathcal{T}_{PP} + U_{\max} \left(p_{pa}(\mathbf{x}) + p_{fa}(\mathbf{x})\right) \mathcal{T}_D}. \quad (26)$$

where we use the fact that $G = \left\lceil \frac{U_{\max}}{U} \right\rceil$. Notice that $\mathcal{R}_{PPM}(\mathbf{x})$ is a function of $p_{pa}(\mathbf{x})$, $p_{fa}(\mathbf{x})$, and U, where $p_{pa}(\mathbf{x})$ and $p_{fa}(\mathbf{x})$ are again functions of mis-detection rate $p_{md}(\mathbf{x})$ in (15), false alarm rate $p_f(\mathbf{x})$ in (14), and traffic load θ. In the light traffic regime (i.e., at low θ regime) decreasing the frequency of false alarm event by choosing a high threshold α value increases the channel use efficiency and thereby may improve $\mathcal{R}_{PPM}(\mathbf{x})$. However, choosing too large of an α leads to near zero achievable throughput as the AP is almost unable to identify the active nodes (i.e., high mis-detection rate). A similar analogy can be made in the high θ regime. Furthermore, given N, how to choose the cyclic shift dimension value $\ell = \left\lfloor \frac{N}{U} \right\rfloor$ in (3) also affects the tradeoff between $p_{md}(\mathbf{x})$ and $p_f(\mathbf{x})$ and thereby impacts $\mathcal{R}_{PPM}(\mathbf{x})$. Hence, the PHY parameters α and U must be optimized to maximize $\mathcal{R}_{PPM}(\mathbf{x})$.

VI. Throughput Maximization: Discrete Monotonic Optimization

We first formulate the global optimization problem and identify useful properties of $\mathcal{R}_{PPM}(\mathbf{x})$.

A. Problem Formulation

We rewrite $\mathcal{R}_{PPM}(\mathbf{x})$ in (26) as

$$\mathcal{R}_{PPM}(\mathbf{x}) = \omega \underbrace{\frac{p_{pa}(\mathbf{x})}{\frac{\mathcal{T}_O + \left\lceil \frac{U_{\max}}{U} \right\rceil \mathcal{T}_{PP}}{U_{\max} \mathcal{T}_D} + \left(p_{pa}(\mathbf{x}) + p_{fa}(\mathbf{x})\right)}}_{\triangleq \widehat{\mathcal{R}}_{PPM}(\mathbf{x})} \quad (27)$$

where $\omega = U_{\max} \nu (1 - p_{e,pull}) p_{da} / U_{\max} \mathcal{T}_D$. We then focus on the problem

$$\mathbf{x}^\star = \underset{\mathbf{x} \in \mathcal{S}}{\operatorname{argmax}} \mathcal{R}_{PPM}(\mathbf{x}) = \underset{\mathbf{x} \in \mathcal{S}}{\operatorname{argmax}} \widehat{\mathcal{R}}_{PPM}(\mathbf{x}) \quad (28)$$

where $\mathcal{S}$ is given in (16). The problem in this expression is non-convex and does not admit a closed-form solution. Intuitively, only a grid search over $\mathbf{x} \in \mathcal{S}$ could approximately find the globally optimal points, but it does not guarantee a certain precision. To solve (28) within a given accuracy in a finite number of steps, we first investigate useful properties of objective function $\widehat{\mathcal{R}}_{PPM}(\mathbf{x})$.

B. Properties of the Objective Function

Rewriting $p_{pa}(\mathbf{x})$ in (23) and $p_{fa}(\mathbf{x})$ in (25) as $p_{pa}(\mathbf{x}) = \kappa(1 - p_{md}(\mathbf{x}))$ and $p_{fa}(\mathbf{x}) = (1 - \kappa) p_f(\mathbf{x})$, respectively, where $\kappa = \theta(1 - p_{e,beacon})(1 - p_{e,poll})$ returns

$$\widehat{\mathcal{R}}_{PPM}(\mathbf{x}) = \frac{\kappa \left(1 - p_{md}(\mathbf{x})\right)}{\frac{\mathcal{T}_O + \left\lceil \frac{U_{\max}}{U} \right\rceil \mathcal{T}_{PP}}{U_{\max} \mathcal{T}_D} + \kappa \left(1 - p_{md}(\mathbf{x})\right) + (1 - \kappa) p_f(\mathbf{x})}.$$

Denote the function $g_+(\mathbf{x})$ as

$$g_+(\mathbf{x}) = \frac{\kappa}{\frac{\mathcal{T}_O + \left\lceil \frac{U_{\max}}{U} \right\rceil \mathcal{T}_{PP}}{U_{\max} \mathcal{T}_D} + \kappa \left(1 - p_{md}(\mathbf{x})\right) + (1 - \kappa) p_f(\mathbf{x})}. \quad (29)$$

Lemma 1: The function $g_+(\mathbf{x})$ defined according to (29) is monotonically increasing on $\mathbf{x} \in \mathcal{S}$, i.e., for $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ if $\mathbf{x}_1 \le \mathbf{x}_2$, $g_+(\mathbf{x}_1) \le g_+(\mathbf{x}_2)$.

Proof: See Appendix B. ■

Based on Lemma 1, the next lemma shows that the problem in (28) is a discrete monotonic optimization problem.

Lemma 2: The problem $\max_{\mathbf{x}\in\mathcal{S}} \widehat{\mathcal{R}}_{PPM}(\mathbf{x})$ in (28) is equivalent to the maximization of difference of monotonically increasing functions with the discrete constraint $\mathcal{S}$.

Proof: See Appendix C. ■

The problem at hand is now formulated as

$$\mathbf{x}^\star = \underset{\mathbf{x}}{\operatorname{argmax}} \left\{ \widehat{\mathcal{R}}_{PPM}(\mathbf{x}) = g_+(\mathbf{x}) - g_-(\mathbf{x}) | \mathbf{x} \in \tilde{\mathcal{S}} \cap \mathcal{S} \right\}, \quad (30)$$

where $g_-(\mathbf{x}) = g_+(\mathbf{x}) p_{md}(\mathbf{x})$, $\tilde{\mathcal{S}}$ is given in (17), and $\mathbf{x}^\star = [U^\star \ \alpha^\star]^T$. Without the discrete constraint, the problem (30) can be solved by the *polyblock* framework [30]. However, with the discrete constraint $\mathcal{S}$, the polyblock technique only computes an approximate solution even with infinite number of iterations. Motivated by the branch-reduce-and-bound approach [23], we propose a discrete monotone optimization algorithm that directly solves (30). We devise new features that are necessary to solve our problem and thereby, provide a complete presentation and proofs.

C. Branch-Reduce-and-Bound Algorithm

We use the general notations adopted in Section IV-B. The first step to construct the sequence of nested boxes in (21) is to reduce a box $M \in \mathcal{M}_k$ to M' at each iteration k such that $M' \subset M$.

1) Reduction: Let $M = [\mathbf{p} \ \mathbf{q}]$ with $\mathbf{x} \in M$ and satisfy $f(\mathbf{x}) \geq \gamma$, i.e., $\mathbf{x}$ is a *feasible solution*, where we abstract out the index k in γ_k for simplicity. The reduction $M' = [\mathbf{p}' \ \mathbf{q}'] \subset [\mathbf{p} \ \mathbf{q}]$ is defined by new vertices

$$\mathbf{p}' = \mathbf{q} - \sum_{i=1}^{l} \varphi_i (q_i - p_i)\mathbf{e}_i \text{ and } \mathbf{q}' = \mathbf{p}' + \sum_{i=1}^{l} \beta_i (q_i - p_i')\mathbf{e}_i, \quad (31)$$

where for $\forall i = 1, \ldots, l$,

$$\varphi_i = \sup\{\varphi | 0 \leq \varphi \leq 1, f_+ (\mathbf{q} - \varphi(q_i - p_i)\mathbf{e}_i) - f_-(\mathbf{p}) \geq \gamma\} \quad (32)$$

and

$$\beta_i = \sup \{\beta | 0 \leq \beta \leq 1, f_+(\mathbf{q}) - f_- (\mathbf{p}' + \beta (q_i - p_i')\, \mathbf{e}_i) \geq \gamma\}. \quad (33)$$

The reduction defined above is a relaxation of the canonical form in Lemma 16 in [23]. The next lemma, however, shows that the relaxation in (32) and (33) still preserves the feasibility of $\mathbf{x} \in [\mathbf{p}' \ \mathbf{q}']$. The lemma is based on the compactness of $\varphi, \beta \in [0 \ 1]$ and the monotonicity of f_+ and f_-.

Lemma 3: If there exists a feasible solution $\mathbf{x} \in [\mathbf{p} \ \mathbf{q}]$ with $f(\mathbf{x}) \geq \gamma$, then $\mathbf{x}$ is also contained in $[\mathbf{p}' \ \mathbf{q}']$, i.e., $\mathbf{x} \in [\mathbf{p}' \ \mathbf{q}'] \subset [\mathbf{p} \ \mathbf{q}]$, where the box $[\mathbf{p}' \ \mathbf{q}']$ is formed according to (31).

Proof: See Appendix D ■

Remark 4: The contradiction of Lemma 3 reveals that if $\mathbf{x}' \notin [\mathbf{p}' \ \mathbf{q}']$, i.e, $\mathbf{x}' \in \{[\mathbf{p} \ \mathbf{p}'), (\mathbf{q}' \ \mathbf{q})\}$, this $\mathbf{x}'$ results in $f(\mathbf{x}') < \gamma$ and one can simply cut off $\{[\mathbf{p} \ \mathbf{p}'), (\mathbf{q}' \ \mathbf{q})\}$ from the box $[\mathbf{p} \ \mathbf{q}]$, resulting in the nested box $[\mathbf{p}' \ \mathbf{q}'] \subset [\mathbf{p} \ \mathbf{q}]$.

Remark 5: Since the functions $f_+(\mathbf{q} - \varphi(q_i - p_i)\mathbf{e}_i) - f_-(\mathbf{p})$ and $f_+(\mathbf{q}) - f_-(\mathbf{p}' + \beta(q_i - p_i')\mathbf{e}_i)$ are monotonically decreasing in φ and β, φ_i in (32) and β_i in (33) can be efficiently found, for instance, by Bolzano's bisection method [31].

Next, the $\mathbf{x} \in [\mathbf{p}' \ \mathbf{q}']$ is required to be adjusted to meet the discrete constraint $\mathcal{S}$.

2) Closer $\mathcal{S}$-Adjustment: Now, we define the *closer $\mathcal{S}$-adjustment* of a feasible solution $\mathbf{x} \in [\mathbf{p} \ \mathbf{q}]$. The *closer $\mathcal{S}$-adjustment* of $\mathbf{x} \in [\mathbf{p} \ \mathbf{q}]$ is the vector

$$[\mathbf{x}]_{\mathcal{S}} = [c_1, \ldots, c_s, x_{s+1}, \ldots, x_n]^T \text{ such that}$$
$$c_i = \underset{\tilde{c}_i}{\operatorname{argmin}} \{|\tilde{c}_i - x_i| | \tilde{c}_i \in \{\lceil x_i \rceil_{\mathcal{S}}^*, \lfloor x_i \rfloor_{\mathcal{S}}^*\}\} \quad (34)$$

where $\lceil x_i \rceil_{\mathcal{S}}^* = \min\{c_i | \mathbf{c} \in \mathcal{S} \cap [\mathbf{x} \ \mathbf{q}], \ c_i \geq x_i\}$ and $\lfloor x_i \rfloor_{\mathcal{S}}^* = \max\{c_i | \mathbf{c} \in \mathcal{S} \cap [\mathbf{p} \ \mathbf{x}], \ c_i \leq x_i\}$. The next lemma elucidates that the *closer $\mathcal{S}$-adjustment* preserves the feasibility of $\mathbf{x}$.

Lemma 4: If there exists a feasible $\mathbf{x} \in [\mathbf{p} \ \mathbf{q}]$ with $f(\mathbf{x}) \geq \gamma$, then the closer $\mathcal{S}$-adjustment of $\mathbf{x}$ satisfies $[\mathbf{x}]_{\mathcal{S}} \in [\mathbf{p} \ \mathbf{q}] \cap \mathcal{S}$ with $f([\mathbf{x}]_{\mathcal{S}}) \geq \gamma$.

Proof: See Appendix E. ■

The branch-reduce-and-bound approach in [23] defines the *lower* and *upper $\mathcal{S}$-adjustments* to align the reduced vertices $\mathbf{p}'$ and $\mathbf{q}'$ with the discrete constraint $\mathcal{S}$. Our framework, however, does not require the vertex adjustments, but rather it is based on the newly devised *closer $\mathcal{S}$-adjustment* that directly compute the $\mathcal{S}$-adjusted feasible solution of $\mathbf{x}$.

3) Bounding: At each iteration k, we have a set of boxes $\mathcal{M}_k$ and a current value γ. To ensure the convergence of the algorithm, the choice of the upper bound $\mu(M_k)$ in (20) must be consistent such that $\lim_{k\to\infty} \mu(M_k) = f(\mathbf{x}^\star)$. Any upper bound satisfying

$$\gamma(M_k) \leq \mu(M_k) \leq f_+(\mathbf{q}) - f_-(\mathbf{p}) \quad (35)$$

will ensure $\lim_{k\to\infty} \mu(M_k) = f(\mathbf{x}^\star)$. At each iteration k, any box $M \in \mathcal{M}_k$ with $\mu(M) \leq \gamma$ is deleted because we have at the very most $f(\mathbf{x}) < \gamma$, $\forall \mathbf{x} \in M$, for that box.

4) Branching: The last operation that accelerates the convergence is dividing a box $M = [\mathbf{p} \ \mathbf{q}]$ into two boxes

$$M_+ = \{\mathbf{x} \in M | x_{i_M} \geq \lceil r_{i_M} \rceil\} \text{ and } M_- = \{\mathbf{x} \in M | x_{i_M} \leq \lfloor r_{i_M} \rfloor\} \quad (36)$$

where $r_{i_M} = (p_{i_M} + q_{i_M})/2$, $i_M = \arg\max_{i=1,\ldots,l}(q_i - p_i)$, $\lceil r_{i_M} \rceil = \min\{y_{i_M} | \mathbf{y} \in \mathcal{S} \cup \{\mathbf{q}\}, y_{i_M} \geq r_{i_M}\}$ and $\lfloor r_{i_M} \rfloor = \max\{y_{i_M} | \mathbf{y} \in \mathcal{S} \cup \{\mathbf{p}\}, y_{i_M} \leq r_{i_M}\}$ [23]. Notice that if $i_M > s$ then $\lceil r_{i_M} \rceil = \lfloor r_{i_M} \rfloor = r_{i_M}$.

Based on the above four basic operations, the branch-reduce-and-bound algorithm solving the discrete monotone optimization problem in (30) is now described in Algorithm 1. We set $f(\mathbf{x}) = \hat{\mathcal{R}}_{PPM}(\mathbf{x})$, $f_+ = g_+$, $f_- = g_-$, and $\mathcal{P}_1 = \tilde{\mathcal{S}}$. In the implementation of Line 4, one can simply compute $\mu(M) = f_+(\mathbf{q}) - f_-(\mathbf{p})$. Given a tolerance $\epsilon \geq 0$, a feasible solution $\mathbf{x}^{(k)}$ such that $f(\mathbf{x}^{(k)}) \geq \mu(M_k) - \epsilon$, where $\mu(M_k)$ is computed based on the box M_k in Line 11, is called *ϵ-optimal solution* (Line 12). The general convergence proof in Theorem 17 in [23] verifies that any branch and bound approach satisfying $\lim_{k\to\infty} \mu(M_k) = f(\mathbf{x}^\star)$ with the necessary condition (35) achieves ϵ-optimal solution after finite number of steps. Hence, the programming problem in (30) is solved using Algorithm 1 within ϵ-accuracy and in a finite number of steps.

TABLE II
PARAMETER SETTING

Parameter	Value	Remark	Parameter	Value	Remark
Data rate	$650kbps$		Slot time	52μsec	
SIFS	$160\mu sec$	$t_{sifs} = 160\mu sec$	PIFS	$212\mu sec$	$t_{pifs} = 212\mu sec$
DIFS	$500\mu sec$	$t_{difs} = 500\mu sec$	PLCP header size	8 symbols	
Service bits	16bits		Tail bits	6bits	
MAC header size	28bytes		Beacon size (Z_{beacon})	36bytes	$t_{beacon} = 800\mu sec$
CF End size (Z_{cfend})	20bytes	$t_{cfend} = 600\mu sec$	Poll size (Z_{poll})	26bytes	$t_{poll} = 680\mu sec$
Data size (Z_{data})	256bytes	$t_{data} = 3.9msec$	Pull size (Z_{pull})	38bytes	$t_{pull} = 840\mu sec$
ACK size (Z_{ack})	14bytes	$t_{ack} = 560\mu sec$	Data+ACK size ($Z_{data+ack}$)	270bytes	$t_{data+ack} = 4.0msec$
Sample duration	$0.5\mu sec$	$t_s = 0.5\mu sec$	Delay spread	$1.5\mu sec$	$t_d = 1.5\mu sec, M = 3$
Round trip time	$3\mu sec$	$t_{rtt} = 3\mu sec, L = 6$	Min. cyclic shift dim.	9 samples	$\ell_{min} = 9$

Algorithm 1: Branch-reduce-and-bound algorithm

Initialization: Let $\mathcal{P}_1 = \{M_1\}$, $M_1 = [\mathbf{a}\ \mathbf{b}]$, $\mathcal{M}_1 = \phi$. Compute the current value $\gamma = f(\mathbf{x}^{(1)}), \mathbf{x}^{(1)} = [(\mathbf{a}+\mathbf{b})/2]_{\mathcal{S}}$. If γ is not available, set $\gamma = -\infty$. Set $k = 1$;
while *ϵ-accuracy is not reached* **do**
  Compute $\mathcal{P}'_k = \{[\mathbf{p}'\ \mathbf{q}'] | [\mathbf{p}\ \mathbf{q}] \in \mathcal{P}_k\}$ in (31);
  Compute a bound $\mu(M)$ for each box $M \in \mathcal{P}'_k$;
  $\mathbf{x}^{(k)} = \text{argmax}\{f(\mathbf{x}) > \gamma | M = [\mathbf{p}\ \mathbf{q}] \in \mathcal{P}'_k, \mathbf{x} = [(\mathbf{p}+\mathbf{q})/2]_{\mathcal{S}}\}$. If there exists $\mathbf{x}^{(k)}$, then $\gamma = f(\mathbf{x}^{(k)})$ and otherwise $\mathbf{x}^{(k)} = \mathbf{x}^{(k1)}$;
  Let $\mathcal{M}'_k = \mathcal{P}'_k \cup \mathcal{M}_k$. Delete $\forall M \in \mathcal{M}'_k$ if $\mu(M) < \gamma$ and let $\mathcal{M}_{k+1}$ be the collection of remaining boxes;
  if $\mathcal{M}_{k+1} = \phi$ **then**
    Terminate;
  else
    if $\gamma = -\infty$ **then**
      The problem is infeasible;
    else
      Let $M_k = \text{argmax}\{\mu(M) | M \in \mathcal{M}_{k+1}\}$;
      if $\gamma \geq \mu(M_k)\,\epsilon$ **then**
        ϵ-accuracy is achieved. $\mathbf{x}^\star = \mathbf{x}^{(k)}$;
      else
        Divide M_k into M_+ and M_- as in (36);
        Let $\mathcal{M}_{k+1} = \mathcal{M}_{k+1} \backslash M_k$ and $\mathcal{P}_{k+1} = \{M_+, M_-\}$;
      end
    end
  end
  $k = k + 1$;
end
output: Solution $\mathbf{x}^\star$ to (30).

VII. PERFORMANCE BENCHMARK AND NUMERICAL EVALUATIONS

We present numerical results comparing the optimized PPMAC with PCF and DCF when the number of sensors grows by one or two orders of magnitude. We adopt system parameters in IEEE 802.11ah for sensor applications [3], [10].

The DCF throughput with imperfect channel and capture effect under unsaturated traffic condition is studied in [11], [12], and they are adopted for the comparison. Previous work [18], [20] in PCF assumes saturated traffic ($\theta = 1$) and the analysis for unsaturated traffic condition is not available. Hence, we derive an achievable throughput expression of the PCF scheduler, incorporating both the traffic condition and the channel imperfection, which is given by (43) in Appendix F.

A. Numerical Simulation Setup

The parameters set to carry out the numerical evaluations are summarized in Table II. The sensors in 802.11ah [3], [10] operate in a 2 MHz band[6] (i.e., sampling interval $t_s = 1/2$ MHz $= 0.5$ μs) with 900 MHz carrier frequency. The control packets (i.e., $beacon, poll, pull, ack, cfend$) and the data packets (i.e., $data, data + ack$) are OFDM-modulated [3]. The OFDM symbol duration is 40 μs including 8 μs of CP. Thus, the net symbol duration is 32 μs corresponding to 64 points FFT size (i.e., 32 μs/0.5 μs $= 64$) and 52 subcarriers (out of 64 subcarriers) are used for data modulation. Both the control and data packets are BPSK-modulated with a rate 1/2-convolutional coding. This means one OFDM symbol can carry a total of 26 bits. Hence, the network capacity is 26 bits/40 μs $=$ 650 Kbps. Let Z_η be the size in bytes of packet $\eta \in \{beacon, poll, pull, data, ack, data + ack, cfend\}$, as shown in Table II. Then, the durations of the control packets are $t_\eta = \left(\left\lceil \frac{(Z_\eta \times 8 + 6 + 16)\text{ bits}}{26\text{ bits}} + \text{PLCP header size} \right\rceil\right) \times 40$ μs for $\eta \in \{beacon, poll, pull, ack, cfend\}$ and the duration for the data packets are $t_\eta = \left(\left\lceil \frac{(Z_\eta \times 8 + 6 + 16 + (\text{MAC header size}) \times 8)\text{ bits}}{26\text{ bits}} + \text{PLCP header size} \right\rceil\right) \times k40$ μs for $\eta \in \{data, data + ack\}$, where 6 bits and 16 bits are the tail and service bits, respectively. Throughout the simulation, a BER of $p_b = 10^{-5}$ (for BPSK 1/2-convolutional coding) is assumed, which corresponds to the operating SNR $\rho = 7.5$ dB after adding 3 dB of multipath fading loss in the link budget analysis of 802.11ah channel in [32]. The PER is calculated based on the formula (22).

We assume poisson traffic with packet arrival rate ϕ (packets/sec), and the traffic is assumed to be independent among sensors. Since each sensor with data loaded gets to transmit in every superframe duration T_s in Fig. 1(a) and (b), the utilization is ϕT_s. Assuming an $M/M/1/l_Q$ queue of length l_Q, the θ is calculated as $\theta = \frac{\phi T_s - (\phi T_s)^{l_Q+1}}{1-(\phi T_s)^{l_Q+1}}$. For a network size $U_{\max}$, the duration of superframe is determined by the maximum available frame durations, i.e., $T_s = \mathcal{T}_O + U_{\max}(\mathcal{T}_{PP} + \mathcal{T}_D)$ and $T_s = \mathcal{T}_O + U_{\max}\mathcal{T}_B$ for PPMAC and PCF, respectively. This means that for a fixed packet arrival rate ϕ, θ grows as

[6]IEEE 802.11ah is basically a 10 times down-sampled version of regular WiFi for a low-power operation.

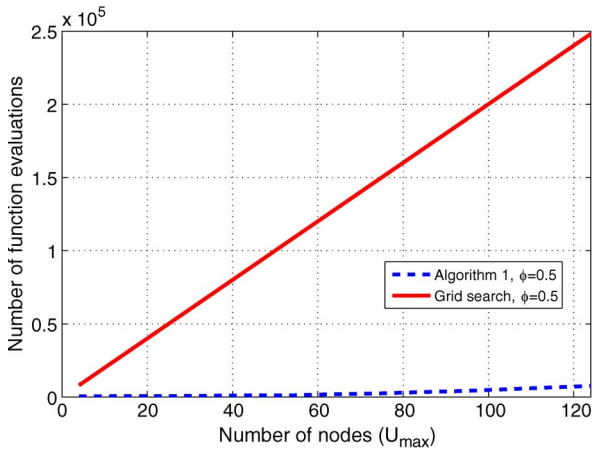

Fig. 3. Number of objective function evaluations: Algorithm 1 vs. grid search.

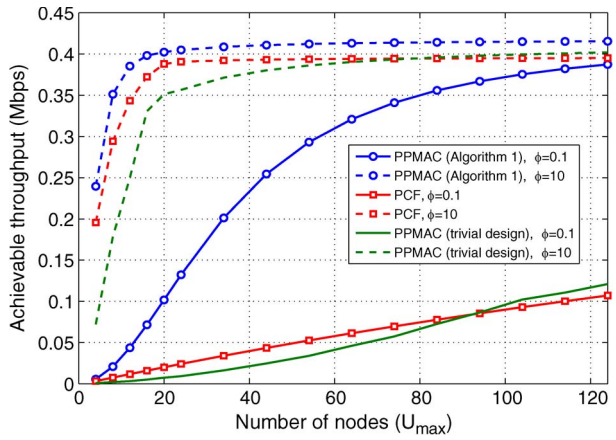

Fig. 4. Achievable throughput across different network sizes: PPMAC vs. PCF.

the network size $U_{\max}$ increases. For a small (large) ϕ, θ grows slowly (fast) as $U_{\max}$ increases. The length of the queue is set to $l_Q = 30$.

The $U_{\max}$ stations are uniformly distributed in a circular area of radius 500 m and connected to the AP placed in the center. The transmission and reception of the parallel acknowledgement follow the procedure presented in Section II.[7] Given the 500 m radius, the round trip time is roughly $t_{rtt} = 3$ μs, i.e., $L = 3$ μs$/0.5$ μs $= 6$ samples. The number of channel taps is $M = 3$. Thus, the minimum cyclic shift dimension in (3) is $\ell_{\min} = M + L = 9$. In the numerical evaluation, each curve for PPMAC is averaged over 100 random channel realizations for $\{\mathbf{h}^u\}_{u=1}^{U_{\max}}$. For each channel realization, Algorithm 1 jointly optimizes $\mathbf{x} = [U\ \alpha]$ for a given network size $U_{\max}$.

B. Numerical Results

1) Complexity: Grid Search vs. Algorithm 1: We validate the computational efficiency of the proposed Algorithm 1 compared to a grid search. Fig. 3 shows the total number of objective function $\widehat{\mathcal{R}}_{PPM}$ evaluations for both the grid search and Algorithm 1 to achieve the same ϵ-precision ($\epsilon =$ 10 bps accuracy). The packet arrival rate $\phi = 0.5$ is assumed. Given the $\mathcal{S} = \{[U,\ \alpha]^T \in \mathbb{R}^2 | U \in I_+, \alpha \in [0\ \alpha_{\max}]\}$ for $I_+ = \{1, \ldots, U_{\max}\}$, a $U_{\max} \times \frac{\alpha_{\max}}{\Delta}$ grid search with the quantization step $\Delta = 5 \times 10^{-4}$, where Δ is set to achieve $\epsilon = 10$ bps-accuracy, is evaluated and compared to Algorithm 1. Fig. 3 demonstrates the high computational efficiency of the proposed algorithm. For instance, with $U_{\max} = 20$ and $\alpha_{\max} =$ 1.0, a 20×2000 grid search (4×10^4 function evaluations) gives the optimum as $\mathcal{R}_{PPM} = 291.376$ Kbps, while the solution found by Algorithm 1 returns $R_{PPM} = 291.378$ Kbps with 616 function evaluations. With $U_{\max} = 124$, a grid search requires 2.5×10^5 function evaluations to achieve $\mathcal{R}_{PPM} =$ 412.581 Kbps, while Algorithm 1 evaluates 7575 times to achieve the same precision $\mathcal{R}_{PPM} = 412.584$ Kbps.

2) Achievable Throughput: PPMAC vs. PCF Scheduler: Fig. 4 shows the performance in terms of the achievable throughput across the network size $U_{\max}$ for the PPMAC and PCF scheduler. The figure shows the performance for $\phi = 0.1$ (light traffic) and $\phi = 10$ (heavy traffic). For each fixed packet arrival rate ϕ, the θ increases as the network size $U_{\max}$ grows. The figure also displays the performance curves of the optimized PPMAC and the trivially designed PPMAC. The trivial design simply set $U = U^\star/2$ and $\alpha = \alpha^\star/2$ and evaluates the performance.

As shown in Fig. 4, the optimized PPMAC results in a significant performance enhancement, compared to the trivial design, in the light traffic regime. The trivial PPMAC design is even worse than PCF when $U_{\max} \leq 94$ for $\phi = 0.1$ and $U_{\max} \leq 74$ for $\phi = 10$. The gain from the optimization is relatively small for heavy traffic loads. When $\phi = 0.1$, the throughput of our PPMAC outperforms that of the PCF scheduler as $U_{\max}$ tends to large, and it has around 5 times throughput improvement over the PCF scheduler for $U_{\max} \geq 60$. The price paid is the increased complexity from adapting the medium allocation to the current traffic condition. In the rest of the numerical study, PPMAC performance is evaluated based on running Algorithm 1.

3) Achievable Throughput: PPMAC vs. DCF: We next compare the performance of PPMAC with that of the DCF, while we also display the performance of the PCF scheduler. When the traffic load at each node is light, DCF can also provide performance benefit. We adopt the m-retry DCF model with the channel capture in [12] to compare them with our PPMAC. In the DCF design, we also assume the same $M/M/1/l_Q$ queue model. The utilization is given by ϕd_c where d_c denotes the channel access delay, yielding $\theta = \frac{\phi d_c - (\phi d_c)^{l_Q+1}}{1-(\phi d_c)^{l_Q+1}}$. The d_c is found by numerically solving a set of nonlinear system equations in [12]. The channel capture [11], [12] is an advanced PHY technology that enhances the efficiency of DCF by correctly receiving packets even when more than one frame collide. For this, a spreading sequence with sufficient processing gain is necessary and in our evaluation an 11 chip Barker sequence is used. In general, the use of a chip-rate spreading sequence increases power consumption. The channel capture threshold 8 dB is assumed. The retry limit is set to $m = 5$. The maximum contention window size is thus 2^5 W for $W = 32$ and the contention window size at the ith retry is given by 2^i W, $i = 0, \ldots, 5$.

[7] In practice, as shown in [1], the parallel acknowledgement using ZC sequences can be implemented without adding any additional PHY modules to the conventional 802.11 OFDM PHY specifications [10].

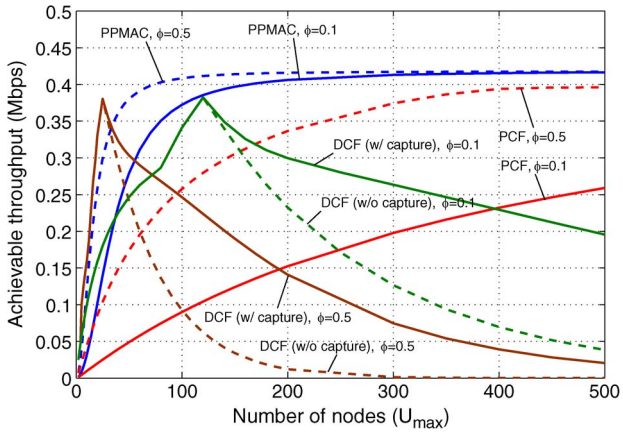

Fig. 5. Achievable throughput across different network sizes for $\phi = 0.1, 0.5$: PPMAC vs. DCF vs. PCF.

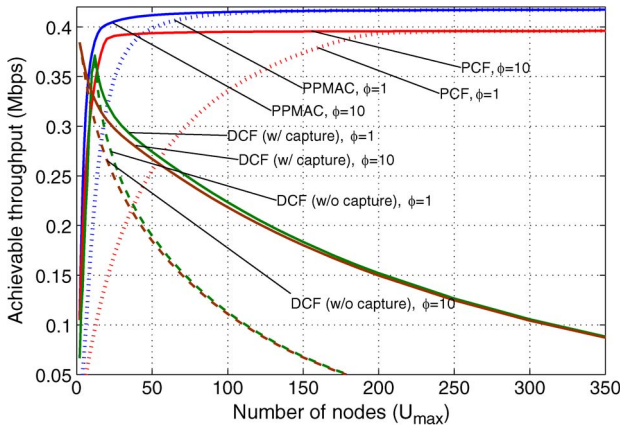

Fig. 6. Achievable throughput across different network sizes for $\phi = 1, 10$: PPMAC vs. DCF vs. PCF.

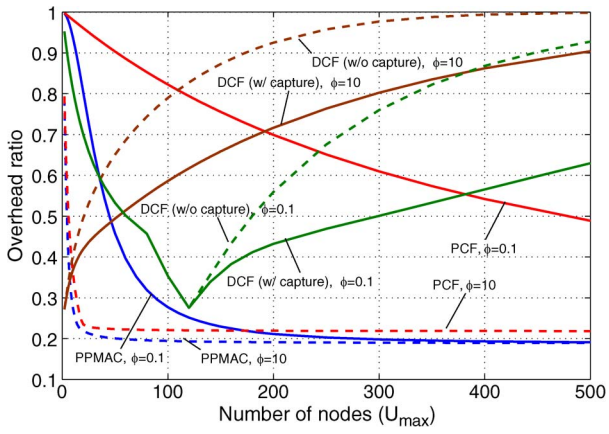

Fig. 7. Overhead ratio across different network sizes: PPMAC vs. DCF vs. PCF.

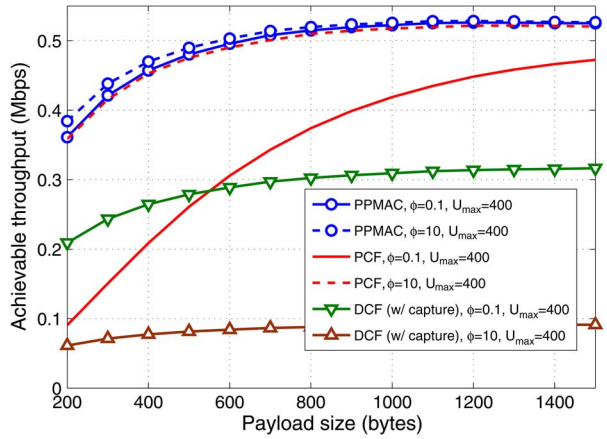

Fig. 8. Achievable throughput across different payload sizes for $U_{\max} = 400$: PPMAC vs. DCF vs. PCF.

Fig. 5 and 6 display the throughput curves versus the network size $U_{\max}$ for the packet arrival rate $\phi = 0.1, 0.5$ and $\phi = 1$, 10, respectively. This means that Fig. 5 emulates slowly increasing traffic, while Fig. 6 simulates fast growing traffic as $U_{\max}$ grows. When the network size is small, DCF outperforms both PPMAC and PCF, as expected. For instance, DCF outperforms the other approaches for $U_{\max} \leq 35$ when $\phi = 0.1$ and $U_{\max} \leq 5$ when $\phi = 10$, respectively. However, as the number of nodes increases the DCF throughput immediately saturates, and after the saturation, the throughput starts to decreases. This is because the major portion of channel uses is wasted for channel contention and backoff operation for a large $U_{\max}$. As a result, our PPMAC employing the parallel acknowledgement substantially outperforms DCF in the most of network sizes of interest. The channel capture of DCF can efficiently alleviate the need for backoff and increase the efficiency, but the gain from it is substantially low compared to the PPMAC. Bursty traffic means the traffic loads go from low data rate to saturated data rate and then back to low data rate. As seen from the numerical studies, the major benefit of PPMAC is that it works for both low and high traffic conditions.

4) Overhead Ratio and Payload Sizes: Let the overhead ratio be $\mathcal{O} = \frac{\text{Average frame overhead duration to deliver the data payload}}{\text{Duration of the data payload}}$, quantifying the amount of overhead wasted to deliver the data payload. Fig. 7 demonstrates the channel use efficiency of PPMAC, PCF, and DCF. When $\phi = 0.1$, more than 70% of the channel uses of PPMAC are utilized for actual data transmission for $U_{\max} \geq 80$, while the efficiency of the PCF scheduler and DCF is substantially lower than that of the PPMAC. When $\phi = 10$, both PPMAC and PCF have high efficiency, but PPMAC still shows improvement. The PPMAC achieves more than 80% of efficiency for $U_{\max} \geq 50$. When $\phi = 10$, the overhead ratio of DCF drastically increase as $U_{\max}$ grows.

In the previous evaluations, the payload size is fixed to $Z_{data} = 256$ bytes. The achievable throughput with different payload sizes is shown in Fig. 8 for $U_{\max} = 400$. Since the throughput of PCF for $\phi = 0.1$ and $U_{\max} = 400$, as seen from Fig. 5, is unsaturated, increasing the payload size results in an almost linear increase of throughput. However, the increase eventually saturates due to the increased packet error rate. Since the throughput of PPMAC for $\phi = 0.1, 10$ and PCF for $\phi = 10$ at $U_{\max} = 400$ is already saturated in Fig. 5 and Fig. 6, the throughput increase as the payload size increases is dominated by the increased packet error rate. The same is observed for DCF with $\phi = 0.1, 10$ at $U_{\max} = 400$, but the further degradation is caused by the increased backoff operation. The performance gap between $\phi = 0.1$ and $\phi = 10$ is narrow for

PPMAC while the others show large deviation. This basically validates the robust traffic adaptation ability of PPMAC for different traffic loads and packet sizes.

VIII. Conclusion and Discussions

We have enhanced the uplink throughput for sensor networks based on a PHY-aided MAC optimization framework. We addressed the problem of designing reliable non-coherent parallel acknowledgement detector and optimizing the detection statistics to maximize the PPMAC throughput for the given network traffic and channel conditions. The efficiency of the proposed optimization is compared with an exhaustive grid search. Moreover, the performance of our scheme is benchmarked with the PCF scheduler and PHY-enhanced DCF scheduler. We were able to demonstrate the superior performance of our approach. In particular, PPMAC outperforms PCF and DCF, and it is able to adapt a wide spectrum of traffic conditions for most of the network sizes of interest. The addressed concept and PHY-aided MAC design can be extended to broadband and multiple antenna systems.

One limitation of the work is that we did not address strategies of grouping and group scheduling for PPMAC. The grouping procedure could incur additional overheads, especially when nodes leave or join the network dynamically. Investigation of practical ways of grouping and implementing group scheduler to reduce potential overheads and improve fairness issue is very interesting topics and is subject to further research. The numerical results in Section VII are based on the evaluations of the analytical formulae derived in this paper and the devised optimization techniques with Monte Carlo simulation. Compared to numerical approaches, empirical performance evaluations with high-fidelity discrete-event simulators could definitely provide additional insights that are closer to actual measurement from physical deployments. Another aspect for further research is thus to model and investigate PCF and PPMAC with network simulators (e.g., NS3) under unsaturated traffic conditions. One main issue here is to investigate what kind of reasonable models (due to the lack of existing models) should be developed and how to incorporate the developed optimization framework (i.e., the PHY-aided MAC optimization) into system-level simulations.

Appendix A
The Proposed Detector is UMP Detector

To verify that the proposed detector is UMP, we need to show that the test is one-sided and the resulting test statistic under the null hypothesis $\mathcal{H}_0$ is independent of the unknown parameter.

Plugging (11) and (12) in the likelihood ratio in (10), for arbitrary $n \in \mathcal{A}_u$, we have

$$\frac{f\left(\hat{v}_u(n)|u \in \mathcal{U}\right)}{f\left(\hat{v}_u(n)|u \notin \mathcal{U}\right)} = \frac{\sigma^2_{u\notin\mathcal{U}}}{\sigma^2_{u\in\mathcal{U}}}\, e^{|\hat{v}_u(n)|^2\left(\frac{1}{\sigma^2_{u\notin\mathcal{U}}}-\frac{1}{\sigma^2_{u\in\mathcal{U}}}\right)} \geq \alpha', \quad (37)$$

and it is clear to conclude that the likelihood ratio is a monotonically increasing function of the observation $|\hat{v}_u(n)|$ for every possible pairs of $\{\sigma^2_{u\in\mathcal{U}}, \sigma^2_{u\notin\mathcal{U}}\}$ such that $\sigma^2_{u\in\mathcal{U}} \geq \sigma^2_{u\notin\mathcal{U}} \geq 0$. Hence the test is *one-sided*. Now, taking the logarithm of (37) and manipulating it, we have

$$\lambda_u(n) = |\hat{v}_u(n)|^2 \geq \frac{\ln(\alpha') - \ln\left(\sigma^2_{u\notin\mathcal{U}}/\sigma^2_{u\in\mathcal{U}}\right)}{1/\sigma^2_{u\notin\mathcal{U}} - 1/\sigma^2_{u\in\mathcal{U}}}. \quad (38)$$

Notice that since $\lambda_u(n) = |\hat{v}_u(n)|^2$ is also monotonically increasing with respect to $|\hat{v}_u(n)|$, an equivalent test statistic to the left hand side of (10) is $\max_{n\in\mathcal{A}_u} \lambda_u(n)$. Hence, Node u is declared, without loss of any optimality, as *active*$(\mathcal{H}_1)$ if

$$\max_{n\in\mathcal{A}_u} \lambda_u(n) \geq \alpha, \quad (39)$$

and it is *inactive* $(\mathcal{H}_0)$ otherwise, where α is determined to meet the target false alarm rate.

We are left to show that the test statistic $\max_{n\in\mathcal{A}_u} \lambda_u(n)$ under $\mathcal{H}_0$ is independent of the unknown parameters. The false alarm is the event of detecting Node u when no sequence is transmitted by Node u, and we write it as $p_f(\alpha|u \notin \mathcal{U}) = \Pr(\max_{n\in\mathcal{A}_u} \lambda_u(n) \geq \alpha | u \notin \mathcal{U})$. The variable $\lambda_u(n)$ under $u \notin \mathcal{U}$ is $\lambda_u(n) = |\bar{w}(n)|^2$ from (9), and it is exponentially distributed with the PDF $f(\lambda|u \notin \mathcal{U}) = Ne^{-N\lambda}$. Hence, the false alarm rate is given by

$$p_f(\alpha|u \notin \mathcal{U}) = 1 - \left(1 - e^{-N\alpha}\right)^{\ell} \quad (40)$$

which is clearly independent of the unknown channel parameters. Thus, (39) is the UMP detector (equivalent to the detector (10)) that maximizes the detection probability for a given target false alarm rate.

Appendix B
Proof of Lemma 1

The first partial derivative of $g_+(\mathbf{x})$ with respect to α is computed as

$$g^{(\alpha)}_+(\mathbf{x}) = \kappa \frac{\kappa p^{(\alpha)}_{md}(\mathbf{x}) - (1-\kappa)p^{(\alpha)}_f(\mathbf{x})}{\Lambda^2} \geq 0 \quad (41)$$

where $\Lambda = \frac{\mathcal{T}_O + \lceil U_{\max}/U\rceil \mathcal{T}_{PP}}{U_{\max}\mathcal{T}_D} + \kappa(1 - p_{md}(\mathbf{x})) + (1-\kappa)p_f(\mathbf{x})$. Notice that since $p_{md}(\mathbf{x})$ and $p_f(\mathbf{x})$ are continuous in α, the first order conditions $\tilde{p}^{(\alpha)}_f(\tilde{\mathbf{x}}) \leq 0$ and $\tilde{p}^{(\alpha)}_{md}(\tilde{\mathbf{x}}) \geq 0$ in (18) are equivalent to $p^{(\alpha)}_f(\mathbf{x}) \leq 0$ and $p^{(\alpha)}_{md}(\mathbf{x}) \geq 0$. The last step in (41) follows from the latter.

Suppose $\tilde{g}_+(\tilde{\mathbf{x}}) = \kappa/\tilde{\Lambda}$, that is, $g_+(\mathbf{x})$ extended on $\tilde{\mathcal{S}}$ in (17) where $\tilde{\Lambda} = \frac{\mathcal{T}_O + (U_{\max}/\tilde{U})\mathcal{T}_{PP}}{U_{\max}\mathcal{T}_D} + \kappa(1 - \tilde{p}_{md}(\tilde{\mathbf{x}})) + (1-\kappa)\tilde{p}_f(\tilde{\mathbf{x}})$. The first order derivative of $\tilde{g}_+(\tilde{\mathbf{x}})$ with respect to $\widetilde{U}$ is

$$\tilde{g}^{(\widetilde{U})}_+(\tilde{\mathbf{x}}) = \frac{\kappa\frac{\mathcal{T}_{PP}}{\mathcal{T}_D}\frac{1}{\widetilde{U}^2} + \kappa^2\tilde{p}^{(\widetilde{U})}_{md}(\tilde{\mathbf{x}}) - \kappa(1-\kappa)\tilde{p}^{(\widetilde{U})}_f(\tilde{\mathbf{x}})}{\tilde{\Lambda}^2} \geq 0.$$

The last inequality is due to the fact that $\widetilde{U} \geq 1$, $\tilde{p}^{(\widetilde{U})}_{md}(\tilde{\mathbf{x}}) \geq 0$, and $\tilde{p}^{(\widetilde{U})}_f(\tilde{\mathbf{x}}) \leq 0$ in (18). Since $\mathcal{S} \subset \tilde{\mathcal{S}}$, the monotonicity of $\tilde{g}_+(\tilde{\mathbf{x}})$ with respect to $\widetilde{U} \in \tilde{I}_+$ also implies the monotonicity of $g_+(\mathbf{x})$ in $U \in I_+$. This concludes the proof.

APPENDIX C
PROOF OF LEMMA 2

Objective function $\widehat{\mathcal{R}}_{PPM}(\mathbf{x})$ is rewritten as $\widehat{\mathcal{R}}_{PPM}(\mathbf{x}) = g_+(\mathbf{x}) - g_-(\mathbf{x})$ where $g_-(\mathbf{x}) = g_+(\mathbf{x})p_{md}(\mathbf{x})$. Since $p_{md}(\mathbf{x})$ and $g_+(\mathbf{x})$ are both monotonically increasing in $\mathbf{x} \in \mathcal{S}$ (from Remark 3 and Lemma 1, respectively), the product of them $g_1(\mathbf{x})p_{md}(\mathbf{x}) = g_-(\mathbf{x})$ in $\mathbf{x} \in \mathcal{S}$ is also monotonically increasing. This concludes the proof.

APPENDIX D
PROOF OF LEMMA 3

We first show that with the feasibility condition of $\mathbf{x}$, i.e., $\mathbf{x} \in [\mathbf{p}\ \mathbf{q}]$ with $f(\mathbf{x}) \geq \gamma$, there exists $\mathbf{p}'$ such that $\mathbf{p} \leq \mathbf{p}' \leq \mathbf{x}$. Suppose the equality $\mathbf{x} = \mathbf{q} - \sum_{i=1}^{l}(q_i - x_i)\mathbf{e}_i = \mathbf{q} - \sum_{i=1}^{l} \tilde{\varphi}_i(q_i - p_i)\mathbf{e}_i$ where $\tilde{\varphi}_i = \frac{q_i - x_i}{q_i - p_i} \in [0 \quad 1]$. For any feasible $\mathbf{x} = \mathbf{q} - \sum_{i=1}^{l} \tilde{\varphi}_i(q_i - p_i)\mathbf{e}_i \in [\mathbf{p}\ \ \mathbf{q}]$ with $f(\mathbf{x}) \geq \gamma$, we have $\gamma \leq f_+(\mathbf{x}) - f_-(\mathbf{x}) \leq f_+(\mathbf{x}) - f_-(\mathbf{p})$. Next, notice that $f_+(\mathbf{x} = \mathbf{q} - \sum_{i=1}^{l} \tilde{\varphi}_i(q_i - p_i)\mathbf{e}_i) - f_-(\mathbf{p})$ is monotonically decreasing in $\tilde{\varphi}_i$. Thus, the compactness of $\tilde{\varphi}_i$ implies there exists $\varphi_i' \geq \tilde{\varphi}_i$ such that $\gamma \leq f_+(\mathbf{q} - \sum_{i=1}^{l} \varphi_i'(q_i - p_i)\mathbf{e}_i) - f_-(\mathbf{p}) \leq f_+(\mathbf{x} = \mathbf{q} - \sum_{i=1}^{l} \tilde{\varphi}_i(q_i - p_i)\mathbf{e}_i) - f_-(\mathbf{p})$, that is, the bound $\gamma \leq f_+(\mathbf{q} - \sum_{i=1}^{l} \varphi_i'(q_i - p_i)\mathbf{e}_i) - f_-(\mathbf{p})$ can be made arbitrarily tight for some $\varphi_i' \geq \tilde{\varphi}_i$, verifying the existence of φ_i in (32). Therefore, φ_i in (32) satisfies $\mathbf{x} = \mathbf{q} - \sum_{i=1}^{l} \tilde{\varphi}_i(q_i - p_i)\mathbf{e}_i \geq \mathbf{q} - \sum_{i=1}^{l} \varphi_i(q_i - p_i)\mathbf{e}_i = \mathbf{p}'$ with $\tilde{\varphi}_i \leq \varphi_i' \leq \varphi_i$, i.e., $\mathbf{p} \leq \mathbf{p}' \leq \mathbf{x}$ for any feasible $\mathbf{x} \in [\mathbf{p}\ \mathbf{q}]$. In a similar manner, we can also show the existence of $\mathbf{q}'$ such that $\mathbf{x} \leq \mathbf{q}' \leq \mathbf{q}$ for $\mathbf{x} \in [\mathbf{p}'\ \mathbf{q}]$.

APPENDIX E
PROOF OF LEMMA 4

By Lemma 3, if $\mathbf{x} \in [\mathbf{p}\ \mathbf{q}]$ with $f(\mathbf{x}) \geq \gamma$, then $\mathbf{x} \in [\mathbf{p}'\ \mathbf{q}']$ with $f(\mathbf{x}) \geq \gamma$. Since the statement holds for an arbitrary $\mathbf{x} \in [\mathbf{p}\ \mathbf{q}]$ with $f(\mathbf{x}) \geq \gamma$, it does also hold with the discrete constraint $\mathcal{S}$, i.e., if $\mathbf{x} \in [\mathbf{p}\ \mathbf{q}] \cap \mathcal{S}$ with $f(\mathbf{x}) \geq \gamma$, then $\mathbf{x} \in [\mathbf{p}'\ \mathbf{q}'] \cap \mathcal{S}$ with $f(\mathbf{x}) \geq \gamma$. Hence the lemma immediately follows from (34) implying if $\mathbf{x} \in [\mathbf{p}'\ \mathbf{q}'] \cap \mathcal{S}$ with $f(\mathbf{x}) \geq \gamma$, then such $\mathbf{x}$ is $\mathbf{x} = [\mathbf{x}]_{\mathcal{S}}$ with $f([\mathbf{x}]_{\mathcal{S}}) \geq \gamma$.

APPENDIX F
ACHIEVABLE THROUGHPUT OF PCF UNDER UNSATURATED TRAFFIC

We assume $U_{\max}$ nodes are polled as shown in Fig. 1(a), and K nodes complete the uplink transaction. The probability that $K = k$ out of $U_{\max}$ nodes successfully complete the transaction is $P[K = k] = \binom{U_{\max}}{k} p_t^k (1 - p_t)^{U_{\max} - k}$ with $p_t = \theta(1 - p_{e,beacon})(1 - p_{e,poll})(1 - p_{e,data+ack})$. The $p_{e,data+ack}$ is the packet error rate of the piggybacked DATA+ACK. The expected payload size (in bits) is therefore formulated as

$$\mathcal{L}_{PCF} = \nu \sum_{k=0}^{U_{\max}} kP[K = k] = \nu U_{\max} p_t.$$

In either cases when the polled node has no data to send or when the node has data to send but it misses the beacon or the poll, the node does not respond. As a result, the AP proceed to the next node after t_{pifs}, for which a total time of $\mathcal{T}_{ID} = t_{poll} + t_{pifs}$ is wasted. The probability of the channel idled per node is then $p_{id} = 1 - \theta(1 - p_{e,beacon})(1 - p_{e,poll})$. If we denote K_1 as the numbers of $\mathcal{T}_{ID}$ slots during a frame, the average number of waisted $\mathcal{T}_{ID}$ slots is given by $E[K_1] = U_{\max} p_{id}$. On the other hand, when the poll message is correctly received but the upstream is in error, the channel is still busy until the uplink transmission ends. In this case, a total of $\mathcal{T}_B = t_{poll} + t_{sifs} + t_{data,ack} + t_{sifs}$ is spent. Since the same amount of time $\mathcal{T}_B$ is spent for the successful uplink data delivery, the probability of spending a $\mathcal{T}_B$ time slot per node is $1 - p_{id}$. Let K_2 be the number of $\mathcal{T}_B$ slots. Then, the expected K_2 is given by $E[K_2] = U_{\max}(1 - p_{id})$. Hence, the total expected duration of time taken to deliver $\mathcal{L}_{PCF}$ is

$$\begin{aligned} \mathcal{T}_{PCF} &= \mathcal{T}_O + E[K_1]\mathcal{T}_{ID} + E[K_2]\mathcal{T}_B \\ &= \mathcal{T}_O + U_{\max}\left(p_{id}\mathcal{T}_{ID} + (1 - p_{id})\mathcal{T}_B\right). \end{aligned} \tag{42}$$

This yields the expected achievable throughput of the PCF,

$$\mathcal{R}_{PCF} = \frac{U_{\max}\nu p_t}{\mathcal{T}_O + U_{\max}\left(p_{id}\mathcal{T}_{ID} + (1 - p_{id})\mathcal{T}_B\right)}. \tag{43}$$

Notice that the expression in (43) is in fact an upper bound of the actual achievable throughput of the PCF scheduler because the additional overhead in the piggybacked poll message is ignored for simplicity.

REFERENCES

[1] T. Kim, C. Sayantan, Z. Jin, K. Doppler, and C. Ghosh, "Simultaneous polling mechanism for low power sensor networks using Zadoff-Chu sequence," in *Proc. IEEE Int. Symp. PIMRC*, Sep. 2012, pp. 1–5.

[2] *Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 3: Physical Layer (PHY) Specifications for Low-Data-Rate, Wireless, Smart Metering Utility Networks*, 2012.

[3] *IEEE Specification Framework for Tgah Rev.14*, IEEE 802.11-11/1137r14, 2013.

[4] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of mimo and cooperative mimo techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.

[5] H. Ochiai, P. Mitran, H. Poor, and V. Tarokh, "Collaborative beamforming for distributed wireless ad hoc sensor networks," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4110–4124, Nov. 2005.

[6] S. Kim, R. Fonseca, and D. Culler, "Reliable transfer on wireless sensor networks," in *Proc. 1st Annu. IEEE Commun. SECON*, Oct. 2004, pp. 449–459.

[7] H. Wen, C. Lin, F. Ren, Y. Yue, and X. Huang, "Retransmission or redundancy: Transmission reliability in wireless sensor networks," in *Proc. IEEE Int. Conf. Mobile Adhoc Sensor Syst.*, Oct. 2007, pp. 1–7.

[8] J. Xiao and Z. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.

[9] J. Fang and H. Li, "Power constrained distributed estimation with cluster-based sensor collaboration," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3822–3832, Jul. 2009.

[10] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2010.

[11] F. Daneshgaran, M. Laddomada, F. Mesiti, and M. Mondin, "Unsaturated throughput analysis of IEEE 802.11 in presence of non ideal transmission channel and capture effects," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1276–1286, Apr. 2008.

[12] M. Siddique and J. Kamruzzaman, "Performance analysis of m-retry beb based dcf under unsaturated traffic condition," in *Proc. IEEE WCNC*, Apr. 2010, pp. 1–6.

[13] M. Balakrishnan, D. Benhaddou, Y. Xiaojing, and D. Gurkan, "Channel preemptive EDCA for emergency medium access in distributed wireless networks," *IEEE Trans. Wireless Communi.*, vol. 8, no. 12, pp. 5743–5748, Dec. 2009.

[14] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in *INFOCOM 2004*, 2004, vol. 2, pp. 1404–1413.
[15] T. Kim *et al.*, Uplink Performance Comparison of PP-MAC and DCF, IEEE 11-12/0326r1, Mar. 2012. [Online]. Available: https://mentor.ieee.org/802.11/documents
[16] S. Mangold, S. Choi, G. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS support in wireless LANs," *IEEE Wireless Commun.*, vol. 10, no. 6, pp. 40–50, Dec. 2003.
[17] B. Sikdar, "An analytic model for the delay in IEEE 802.11 PCF MAC-based wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1542–1550, Apr. 2007.
[18] X. Ma, Y. Wu, Z. Niu, and T. Saito, "Performance analysis of the packetized voice transmission with PCF in an IEEE 802.11 infrastructure wireless LAN," in *Proc. APCC*, 2003, vol. 2, pp. 571–575.
[19] M. Siddique and J. Kamruzzaman, "VoIP capacity over PCF with imperfect channel," in *Proc. IEEE GLOBECOM*, 2009, pp. 1–6.
[20] M. Siddique and J. Kamruzzaman, "Performance analysis of PCF based WLANs with imperfect channel and failure retries," in *Proc. IEEE GLOBECOM*, 2010, pp. 1–6.
[21] Z. Jin *et al.*, MAC Consideration for 802.11ah (Prove and Pull MAC), IEEE 11-11/1512r4, Nov. 2011. [Online]. Available: https://mentor.ieee.org/802.11/documents
[22] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, Time Series Analysis*. New York, NY, USA: Addison-Wesley, 1990.
[23] H. Tuy, M. Minoux, and N. T. Hoai-Phuong, "Discrete monotonic optimization with application to a discrete location probelem," *SIAM J. Optim.*, vol. 17, no. 1, pp. 78–97, 2006.
[24] H. Tuy, N. D. Nghia, and L. S. Vinh, "A discrete location problem," *Acta Math. Vietnamica*, vol. 28, no. 2, pp. 185–199, 2003.
[25] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.
[26] Q. Chen and M. Gursoy, "Energy-efficient modulation design for reliable communication in wireless networks," in *Proc. 43rd Annu CISS*, 2009, pp. 811–816.
[27] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
[28] S. Beyme and C. Leung, "Efficient computation of DFT of Zadoff-Chu sequences," *Electron. Lett.*, vol. 45, no. 9, pp. 461–463, Apr. 23, 2009.
[29] S. Kay, *Fundamentals of Statistical Signal Processing: Detection theory*, *Prentice Hall Signal Processing Series*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
[30] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM J. Optim.*, vol. 11, no. 2, pp. 464–494, Oct. 2000.
[31] R. L. Burden and J. D. Faires, *Numerical Analysis*, 7th ed. Pacific Grove, CA, USA: Brooks/Cole, 2001.
[32] R. Porat, Link Budget, IEEE 802.11-11/0552r2 2011. [Online]. Available: https://mentor.ieee.org/802.11/documents

Taejoon Kim (S'08–M'11) received the B.S. degree (with highest honors) from Sogang University, Korea and the M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST), in 2002 and 2004, respectively. From 2004 to 2006, he was with Electronics and Telecommunications Research Institute (ETRI), Korea. In 2007, he joined Purdue University, West Lafayette, Indiana and earned Ph.D. degree in 2011. From 2011 to 2012, he was with the Nokia Research Center (NRC), Berkeley, California as a Senior Researcher. Before joining City University of Hong Kong as an Assistant Professor in 2013, he was a postdoctoral researcher in the Communication Theory group at Royal Institute of Technology (KTH) in Stockholm, Sweden. He is a member of the State Key Laboratory of Millimeter Wave, City University of Hong Kong.

Dr. Kim was the recipient of a 2012 IEEE PIMRC best paper award. He has served as a guest editor for special issues of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS. His research interests are in the design and analysis of adaptive communication systems, array signal processing, and MIMO wireless systems.

David J. Love (S'98–M'05–SM'09) received the B.S. (with highest honors), M.S.E., and Ph.D. degrees in electrical engineering from the University of Texas at Austin in 2000, 2002, and 2004, respectively. During the summers of 2000 and 2002, he was with Texas Instruments, Dallas, TX. Since August 2004, he has been with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, where he is now a Professor and recognized as a University Faculty Scholar. He has served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and as a guest editor for special issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the EURASIP Journal on Wireless Communications and Networking. His research interests are in the design and analysis of communication systems and MIMO array processing. He is recognized as a Thomson Reuters Highly Cited Researcher. Dr. Love has been very involved in commercialization of his research with more than 25 U.S. patent filings, 21 of which have issued.

Dr. Love has been inducted into Tau Beta Pi and Eta Kappa Nu. Along with co-authors, he was awarded the 2009 IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY Jack Neubauer Memorial Award for the best systems paper published in the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in that year and a 2013 IEEE Globecom best paper award. He was the recipient of the Fall 2010 Purdue HKN Outstanding Teacher Award and the Fall 2013 Purdue ECE Graduate Student Association Outstanding Faculty Award. He was an invited participant to the 2011 NAE Frontiers of Engineering Education Symposium. In 2003, Dr. Love was awarded the IEEE Vehicular Technology Society Daniel Noble Fellowship.

Mikael Skoglund (S'93–M'97–SM'04) received the Ph.D. degree in 1997 from Chalmers University of Technology, Sweden. In 1997, he joined the Royal Institute of Technology (KTH), Stockholm, Sweden, where he was appointed to the Chair in Communication Theory in 2003. At KTH, he heads the Communication Theory Division and he is the Assistant Dean for Electrical Engineering. He is also a founding faculty member of the ACCESS Linnaeus Center and director for the Center Graduate School.

Dr. Skoglund has worked on problems in source-channel coding, coding and transmission for wireless communications, Shannon theory and statistical signal processing. He has authored and co-authored more than 125 journal and about 300 conference papers, and he holds six patents.

Dr. Skoglund has served on numerous technical program committees for IEEE sponsored conferences. During 2003–2008 he was an associate editor with the IEEE TRANSACTIONS ON COMMUNICATIONS and during 2008–2012 he was on the editorial board for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Zhong-Yi Jin received the B.A. degree in computer science from Boston University, Boston, MA, USA, and the M.S. and Ph.D. degrees in computer science from University of California San Diego, San Diego, CA, USA. He is currently affiliated with Qualcomm Research San Diego working on expanding LTE into unlicensed bands. Previously, he conducted research in the areas of energy efficient wireless MAC, context-aware energy management, and Internet of Things at Samsung Research and Nokia Research. His current research interests include the coexistence and cooperation of different wireless technologies, cross-layer optimization for wireless MAC, and the security and privacy of Internet of Things.