3D Pipeline Contention: Asymmetric Full Duplex in Wireless Networks

Shaohe Lv, Xuan Dong, Yong Lu, Xiaoli Du, Xiaodong Wang, Yong Dou, and Xingming Zhou
National Laboratory of Parallel and Distributed Processing
National University of Defense Technology, Changsha, Hunan, China
Email: {shaohelv, xdong, ylu8, xdu13, ydou, xdwang, xmzhou}@nudt.edu.cn

Abstract—**Coordination among users is an indispensable part in wireless networks for efficient medium access. Alone with the rapid increase of transmission rate, however, coordination time becomes insufferable. We present AFD, namely asymmetric full duplex, to achieve high coordination efficiency at nearly zero overhead. In AFD, channel contention is performed simultaneously with data transmission. We propose a 3D pipeline contention scheme where the contention process is divided into several parallel stages and executed in a pipelined manner in a 3D domain specified by time, frequency and spatial antenna. To mitigate the interference between the data packet and the contention signal, we adopt a singleton PN sequence as a contention pilot. AFD provides a novel network-scale full duplex capability. The performance is evaluated by both simulations and measurements in a testbed. AFD outperforms IEEE 802.11 significantly, i.e., the Jain's fairness index is around 0.95 with a throughput gain up to 120%.**

I. INTRODUCTION

Coordination among users is an indispensable part in wireless networks for efficient medium access. Due to the broadcast nature, what arrives at the receiver is a composite signal from all near-by transmissions. Only when specific constraints are satisfied, it is feasible to extract packets from collision by using the multi-packet reception techniques [1]. As different users can severely interfere with each other, node coordination is critical in wireless networks to avoid the harmful interference.

Node coordination is a costly operation taking a significant amount of valuable communication resource. Alone with the rapid increase of physical layer transmission speed, coordination time becomes insufferable, i.e., even several times higher than that for data transmission. In IEEE 802.11n, transmission time of a 1500 Bytes packet at 300Mbps is 40 μs. In comparison, the average coordination time, i.e., DIFS plus the random backoff, is larger than 100 μs [2].

In the conventional mechanisms such as CSMA or Carrier-Sense Multiple Access, access contention and data transmission are performed separately. CSMA has been widely adopted in many MAC protocols for implementing distributed access to shared communication medium. In a CSMA-based protocol (e.g., IEEE 802.11), a sender listens and contends for medium access when the channel is idle. The contention process is blocked when a nearby transmission is detected and activated after the end of the intervening transmission.

It is difficult for a separation-based mechanism to achieve high coordination efficiency as well as low overhead. Certain communication resource is consumed in the negotiation among the contending nodes in order to utilize the channel effectively and fairly. Nevertheless, the more resource node coordination occupies, the less available for data transmission. As shown in [2], the effect on the network performance becomes severe when the physical layer transmission rate is very high.

We explore a novel methodology, called asymmetric full duplex (AFD), for efficient distributed access control at nearly zero cost in wireless networks. The basic idea is to perform channel contention simultaneously with data transmission.

AFD provides a network-scale full duplex capability. *It allows uninterrupted data communication in a wireless channel and time for channel contention is completely hidden.* That is, the contending users compete for the future channel access when the channel engages in transmitting data packets. In comparison, conventional full duplex allows a single node to transmit and receive simultaneously. It cannot eliminate the coordination overhead as long as the access contention is performed only in an idle channel.

The requirement that coordination should be performed in a busy channel poses several technical challenges.

(1) *Efficiency*: A winner user should be selected in the contention process before the end of the ongoing transmission. Otherwise, channel has to be idle after the current transmission is completed. When data rate is high, the required time is very short (e.g., tens of millisecond) to transmit a data packet. In this situation, none of the available contention methods works well to pick out the winners in time.

(2) *Interference*: As the message for coordination is transmitted simultaneously with the data packet, the two types of signal can interfere with each other. Due to the mutual interference, it is challenging to detect reliably the contention signals as well as the data packets at the intended receivers.

To achieve the objective, we make the following contributions. We propose a 3D pipeline contention scheme where the contention process is divided into several parallel stages and executed in a pipelined manner in a 3D domain specified by time, frequency and spatial antenna. To mitigate the interference between the data packet and the contention signal, we adopt a singleton correlatable PN sequence as a contention pilot. Afterwards, a robust correlation-based method is used to detect the pilot signal in the presence of interference. As the PN sequence adopted by all competing users is the same and already known by all nodes, it is easy to remove the pilot signals at a receiver by interference cancellation [1, 3].

The performance is evaluated by both simulations and measurements in a testbed. It is shown that, the 3D pipeline contention scheme outperforms the conventional method such as IEEE 802.11 with much better fairness and a throughput gain as large as 120%.

The rest of this paper is organized as follows. Section II overviews the related work and Section III describes the motivation and the basic idea of our work. Section IV discusses the 3D pipeline contention scheme. Section V presents the experimental results. Finally, we conclude the research in Section VI.

II. RELATED WORK

It is a fundamental issue to design a access control method in wireless networks. Here we summarize only some of the works closely related to ours.

There are a lot of studies to reduce the coordination cost, for example, by introducing frequency domain contention [4], by using a 800ns short slot [2], and by avoiding the control messages [5, 6]. FICA [7] strives to amortize the control overhead over a number of simultaneous transmissions. Also, in IEEE 802.11n, frame aggregation is proposed to transmit multiple packets together. To reduce the collision in a contention-based access method, WIFI-BA [8] devises a bitwise arbitration mechanism to resolve the collision, and Idle Sense [9] adjusts an attempt probability rather than the contention window. Instead of collision avoidance, CSMA/CN [10] lets the receiver to notify the transmitter of a collision. In comparison, our way is to hide the coordination time rather than reduce it.

Researches on full duplex wireless [11, 12] study how to eliminate self-interference so that a node can transmit and receive simultaneously. There are several works to design a access control method at the full duplex mode, e.g., in-frame error recovery [13], WIFI-Nano [2], Back2F [4], and optimal spectrum access [14]. As each antenna at the full duplex mode can work as an independent user, our scheme also benefits by exploiting the antenna diversity.

It is a hot topic recently to provide a separate control plane on the same channel concurrently with data transmissions. In [15], a free coordination channel is built by generating intended interference patterns. In [16], Flashback carries the control messages by sending short high-powered flashes. Both of them require dedicated control of the senders and the concurrent interfering nodes. AFD relies on two widely-used techniques, e.g., correlation-based detection and interference cancellation, thereby can be adopted easily in practice.

III. MOTIVATION AND BASIC IDEA

Consider a single-hop wireless network of multiple stationary nodes. A node can have multiple antennas. Each antenna is omni-directional and operates in the half duplex mode. We discuss the implication of the full duplex mode in Section IV-C.

Orthogonal Frequency Division Multiplexing (OFDM) has become increasingly popular in modern wireless communications such as IEEE 802.11a/g/n, WiMax, and 3GPP LTE.

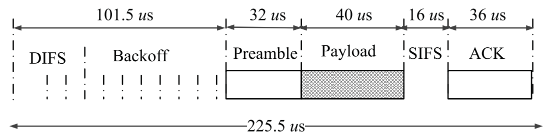

Fig. 1: Overhead in IEEE 802.11 at 300Mbps.

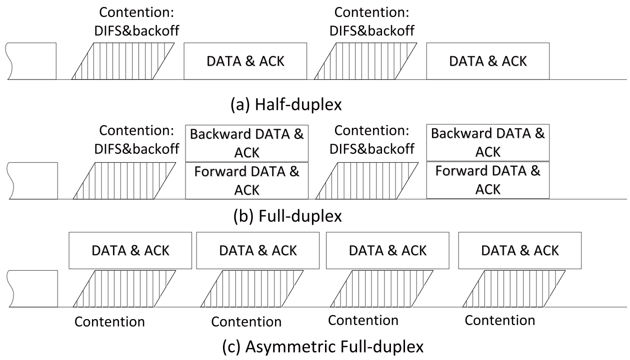

Fig. 2: Channel access at different modes: (a) half duplex; (b) full duplex; and (c) asymmetric full duplex.

OFDM can be abstracted as a PHY scheme that divides the wireless spectrum into multiple narrow band channels, called subcarriers. The subcarriers carry modulated data streams in parallel, but at a lower rate per-subcarrier. The 802.11a/g implementation of OFDM has 52 subcarriers, of which 48 are used for data transmission, and 4 for equalization. A transmitter stripes bits across all subcarriers, however, it is possible to transmit/receive only on a subset of them. Several multi-user access methods such as OFDMA and FICA [7] are proposed for OFDM-based wireless networks.

Motivation: While WiFi physical layer (PHY) data rates have increased from 1 Mbps in the original IEEE 802.11 to 1 Gbps in the recent IEEE 802.11ac standard, user level throughputs have not seen a commensurate increase. A key reason for this is the channel access overhead.

Fig. 1 shows the overhead in IEEE 802.11 at 300Mbps. As dictated by CSMA in 802.11, prior to transmission, the device must first sense that the channel is idle for the duration of DIFS. After DIFS, a node must typically defer its transmission for a random number of slots, generated from 0 to CW-1 (contention window size). The process is blocked when a nearby transmission is detected and activated only after the end of the intervening transmission. Given that the minimum value of CW as dictated by 802.11 is 16, the device will on average wait about 7.5 slots before transmission. As specified in the IEEE 802.11 standard, DIFS is 34 μs and the slot length is 9 μs. Therefore, the total coordination overhead is up to 101.5 μs. In comparison, transmission time of a 1500 bytes packet at 300Mbps is only 40 μs.

It is difficult to achieve high coordination efficiency as well as low overhead. The more resource node coordination occupies, the less available for data transmission. The effect on the network performance becomes severe when physical layer transmission rate is very high. As shown in Fig. 1, the percent of time for data transmission is less than 20%, while about 50% of the time is used for channel contention.

On the other hand, certain communication resource is required in the negotiation among the contending nodes in order to utilize the channel effectively and fairly. A very short slot (i.e., 800ns) is adopted in WIFI-Nano to reduce the control overhead. It is argued in [8] that, when the slot length is too short, a node cannot reliably detect the behaviors of the others during a slot even when the node operates at the full duplex mode. As a consequence, a distributed contention-based protocol is ineffective to prevent the harmful collisions.

Asymmetric full duplex (AFD): AFD is a novel methodology for distributed access control in wireless networks. The basic idea is to perform channel contention simultaneously with data transmission. Fig. 2 shows the channel access process at different modes. At the half duplex mode, channel contention is separated from data transmission. At the full duplex mode, the network throughput is doubled when the traffic demand is bidirectional. As long as contention is performed only in the idle channel, however, the coordination overhead cannot be eliminated. In comparison, AFD performs channel contention and data transmission in parallel. As a result, coordination time can be almost completely hidden.

The requirement that coordination should be performed in a busy channel poses several technical challenges. First, a winner user should be selected before the end of the ongoing transmission. Otherwise, channel has to be idle after the current transmission is completed. Second, the message for contention is transmitted simultaneously with the data packet. Due to the mutual interference, it is challenging to detect reliably the contention messages as well as the data packets at the intended receivers.

Next, we present a 3D pipeline contention scheme to select one or more winners at the end of each data packet transmission. In addition, a singleton PN sequence is adopted to carry the contention information in the presence of interference.

IV. 3D PIPELINE CONTENTION

We first present the basic pipeline contention scheme. Afterwards, we describe a 3D contention method and two advanced pipelines to improve the pipeline efficiency, and the design of the singleton PN sequence. Finally, we discuss the parameter settings and future extensions of pipeline contention.

A. Basic pipeline contention

We adopt the pipeline technique to accelerate the channel contention. Pipeline is a fundamental technique in computer architecture. Basically, a channel contention process is divided into several steps and executed in parallel. The access method of IEEE 802.11 DCF is not suitable for a pipeline implementation. As defer time is set randomly, the required time to complete a contention can vary drastically.

We adopt a new contention method with a fixed-length contention window, similar as that discussed in [8]. Algorithm 1 shows the channel access process with pipeline contention. A contention duration is partitioned into $\mathcal{D}$ slots (i.e., $t_1, \ldots, t_{\mathcal{D}}$) of a constant length (i.e., t_{slot}). A node S generates randomly a binary *contention vector* $\mathcal{V}_S$ with $|\mathcal{V}_S|$=$\mathcal{D}$. At the ith slot, the operation of node S is determined as follows.

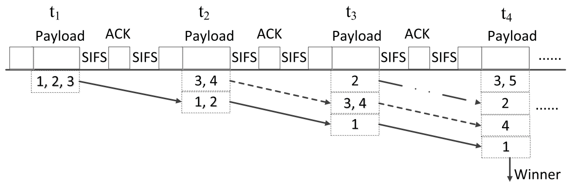

Fig. 3: Basic contention pipeline.

Algorithm 1: Channel access with pipeline contention.

```
repeat
    Generate a contention vector 𝒱;
    for i=1 to |𝒱| do
        if 𝒱[i]==1 then send a pilot message at slot t_i;
        else
            Sense the channel at slot t_i;
            if pilot is detected then continue;
        end
    end
until win the contention;
Transmit the data packet;
```

- $\mathcal{V}_S[i]$=1: send a pilot message;
- $\mathcal{V}_S[i]$=0: sense the channel, and withdraw from the contention if a pilot message is detected.

At the last stage of a pipeline, node S becomes the winner if, it does not detect a pilot at every t_i such that $\mathcal{V}_S[i]$=0. The winner can transmit data packets after the ongoing transmission is completed. When a node has to withdraw from the contention at a middle pipeline stage, it can generate a new binary vector and conduct a new contention pipeline.

Consider a network with five senders, i.e., S_1-S_5. Fig. 3 shows an example of 4-stage contention pipeline and Fig. 4 shows the used contention vectors. At slot t_1, S_1, S_2 and S_3 initiate a contention. S_3 withdraws at the end of t_1 from the contention as it detects the pilots from S_1 and S_2. At slot t_2, S_2 withdraws at the end as it detects a pilot from S_1. Meanwhile,

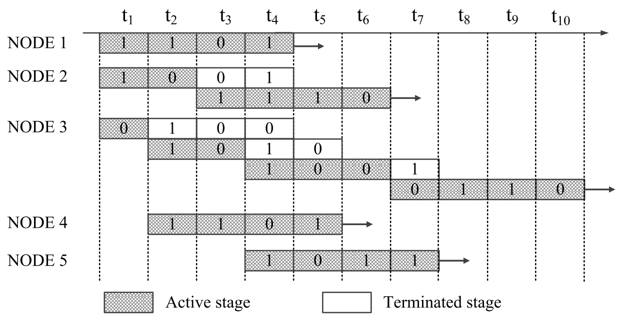

Fig. 4: Contention vectors used in the pipeline contention.

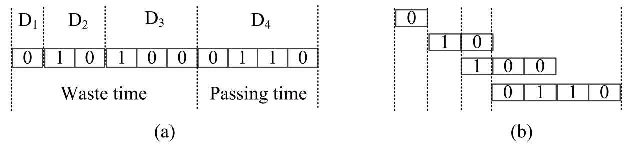

Fig. 5: Access delay of S_3: (a) in basic pipeline contention; (b) with antenna diversity.

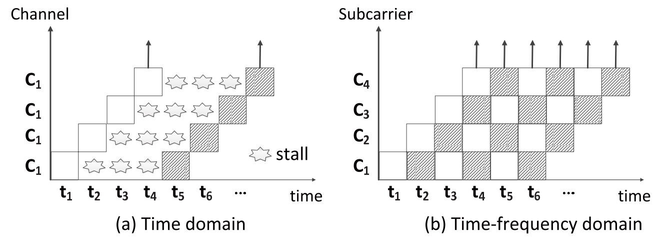

Fig. 6: Time-frequency diagram of a 4-stage pipeline.

S_3 and S_4 initiate a new contention. At slot t_3, S_1 executes the third pipeline stage, S_3 and S_4 contend at the second pipeline stage, and S_2 initiates a new contention. At t_4, S_1 wins in the contention at the end of the fourth pipeline stage. Afterwards, S_4, S_2, S_5 and S_3 become the winners in sequence.

In practice, as the detection of a packet preamble is susceptible to interference, channel contention is initiated at the beginning of a payload transmission. Also, the detection of a preamble can serve as a reference point so that the competing nodes can be loosely synchronized.

As shown in Fig. 4, except the first three slots, there is always a winner node at the end of each time slot. That is, with the pipeline contention, data transmission can become an steady and uninterrupted stream. To realize this, however, several challenges should be addressed.

Structural hazard: Ideally, all stages of a pipeline are executed in parallel. Structural hazard occurs when the operations in more than one stage require the same shared resource. At this time, only sequential execution is allowed for these operations. In the contention pipeline, wireless channel is usually a shared resource. At any given time, there are multiple groups of users compete at different stages. For example, at t_2, S_1 and S_2 contend at the second stage, while S_3 and S_4 contend at the first stage. If all of them operate at the same channel, the issued pilots will interfere with each other. To mitigate the interference effect, as shown in Fig. 6 (a), lots of stalls are introduced and the pipeline is underutilized.

Access delay: Access delay of a node is defined as the time to perform the pipeline contention until the node wins. It usually contains two parts : *waste time* to pass through all the (incomplete) pipelines except the last one and *passing time* of the last pipeline in which the node wins. As shown in Fig. 5 (a), for S_3, the waste time is the first 6 slots used to pass through the three incomplete pipelines (e.g., D_1-D_3) and the passing time is the last 4 slots used to perform the fourth pipeline. In total, 10 slots are required by S_3 for contention and 60% of them are wasted in the failed contentions.

B. Contention in a 3D space

To improve the pipeline efficiency, first, a time-frequency domain is adopted to allow the overlapping execution of all pipeline stages. Second, the redundant antennas are exploited to reduce the waste time, the first part of the access delay.

Time-frequency domain pipeline: To execute multiple pipeline stages in parallel, it is necessary to make sure that, at any time, the pilots from different users can be separated. We exploit the large number of available subcarriers provided by OFDM and assign different subcarrier to different pipeline stage. Due to the orthogonality, the signals at different subcarriers can be detected independently. For a $\mathcal{D}$-stage pipeline, $\mathcal{D}$ subcarriers are required.

Algorithm 2 shows the pipeline contention process in a time-frequency domain. Let $C_1, \ldots, C_{\mathcal{D}}$ denote the subcarriers. When a user enters a pipeline, it first contends at C_1. Afterwards, it turns to C_2 at the second pipeline stage, and so on. The node operation is the same as that in the basic pipeline (i.e., Algorithm 1) except that, pilot is sent or detected at a different subcarrier for different pipeline stage.

Algorithm 2: Contention in a time-frequency domain.

Data: Contention vector $\mathcal{V}$ ($|\mathcal{V}|$=$\mathcal{D}$)
Result: False-failed; True-success
for i=1 **to** $|\mathcal{V}|$ **do**
  if $\mathcal{V}[i]$==1 **then** send pilot at slot t_i by using subcarrier C_i
  else
    Sense the subcarrier C_i at slot t_i;
    if *pilot is detected* **then** return False;
  end
end
return True;

Fig.6 shows the time-frequency diagram of a 4-stage contention pipeline. In Fig.6 (a), there are lots of stalls in the contention pipeline when only the time domain is used. In comparison, in Fig.6 (b), all the pipeline stalls are eliminated by using a time-frequency contention domain. For example, though three pipelines are executed in parallel at t_3, there is no mutual interference as they use distinct subcarriers.

Antenna diversity: To reduce the waste time, a node is allowed to participate in multiple contention pipelines. When a node has multiple antennas, each of them can participate in a different contention pipeline. With a sufficient number of antennas, the waste time of a node S reduces to $(M-1)t_{slot}$ when S wins in the Mth contention. For example, in Fig. 5 (b), when S_3 has four antennas for contention, the number of waste slots is reduced by 50% (i.e., from 6 to 3).

It is easy to exploit the antenna diversity at the full duplex mode. Each full duplex antenna can compete or transmit as an independent user. That is, when a node has multiple antennas, it can use some of them to transmit data while the rest to contend for future channel access.

When each antenna operates at the half duplex mode, to avoid self-interference in the detection of pilots, first, every two of the antennas at a node cannot participate in the same contention pipeline. Note that, there is no mutual interference when two or more antennas are used for contention but stay at different pipeline stages. Second, if some antennas are used for data transmission, the rest cannot perform channel contention unless the subcarriers for data transmission and those for contention do not overlap.

How many contention pipelines should a node take part in? Let λ denote the number of currently involved pipelines. Obviously, a larger λ can reduce the waste time. However, consider a node with two antennas A_1 and A_2. After A_1 wins, A_2 will switch to transmit. If all the competitors except A_2 have already withdrawn, there will be no winner at the end of the contention that A_2 takes part in, leading to a pipeline efficiency loss.

We adopt a threshold-based adaptive control policy. When λ is less than a threshold τ, a node can take part in the incoming $(\tau - \lambda)$ contentions. Initially, τ=1. We then adjust τ according to the contention result:

- $\tau \leftarrow \tau+1$, when the node withdraws from a contention;
- $\tau \leftarrow \max\{1, \lceil \tau/2 \rceil\}$, when the node wins in a contention;

3D contention: In summary, the pipeline contention is performed in a 3D domain specified jointly by time, frequency and spatial antenna. An antenna occupies several consecutive time slots and a set of different subcarriers to complete a contention. Also, at any given time, there are several different sets of antennas and each of the sets uses distinct subcarriers for contention. Finally, at a subcarrier, there are different sets of antennas execute the same pipeline stage in sequence.

It is practically feasible to deploy the 3D contention method. First, as we will show later, a 12-stage pipeline is sufficient to resolve the contention among 50 nodes. The number of available subcarriers is 48 in 802.11a/g/n and becomes larger in the emerging standards. Second, as MIMO (Multiple Input Multiple Output) has been used widely (e.g., in IEEE 802.11n), it is expected that the number of antennas at a single device will increase persistently.

C. Advanced pipeline

Now consider the reduction of the passing time. For a pipeline of depth $\mathcal{D}$, the passing time is $\mathcal{D}t_{slot}$. In the basic contention pipeline, t_{slot} is the completion time of a packet, i.e., from the beginning of preamble transmission to the end of ACK transmission, which is also denoted by t_{pkt}.

We discuss two novel pipelines to reduce the passing time, i.e., compact pipeline and full pipeline. Basically, when multiple (e.g., n) pipeline stages are executed during a packet transmission, the pipeline passing time reduces to $\mathcal{D}t_{pkt}/n$.

In a *compact pipeline*, two or more stages are executed in a packet transmission duration. There is only one new pipeline issued during a packet transmission. A 4-stage compact pipeline is shown in Fig. 7, where there are two pipeline stages during the transmission of one packet. Though node S_2 has to withdraw from a contention in the middle of the first

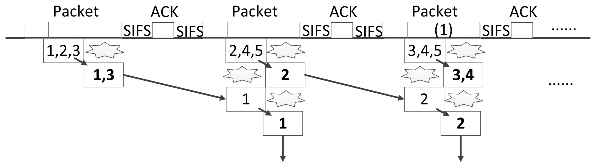

Fig. 7: Compact pipeline with stalls.

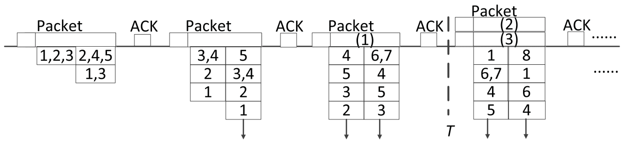

Fig. 8: Full pipeline with simultaneous data transmissions.

packet transmission, it cannot contend again immediately until the second packet is transmitted.

In a *full pipeline*, multiple pipelines are issued in a packet transmission duration. As there are several contention pipelines finished, more than one sender become the winners within the duration. They will transmit simultaneously after the ongoing transmission. A 4-stage full pipeline is shown in Fig. 8. After T, there are two winners in each transmission duration and the pipeline achieves the full capacity.

The effect of transmission rate should be taken into account. When transmission rate is low, transmission time of a packet is sufficient to execute multiple pipeline stages. For example, about 250 μs is required to transmit a 1500 Bytes payload at 54Mbps. When transmission rate is high, transmission time of a regular packet is very short, e.g., 12 μs for a 1500 Bytes packet at 1Gbps. To extend the transmission duration, frame aggregation in IEEE 802.11n allow us to transmit a very large packet. In comparison, a full pipeline works better with very high transmission rate. There are several techniques can be adopted by a full pipeline to perform concurrent transmissions in an OFDM network, such as OFDMA or FICA [7]. Basically, though the overall data rate is very high, to support simultaneous transmissions, a relatively lower transmission rate should be used for each packet, resulting in a longer transmission duration.

Define the number of concurrent transmissions in a full pipeline as *parallelism level*, $\mathcal{P}$. When $\mathcal{P}$=1, a full pipeline is the same as a compact pipeline. A compact pipeline is suitable when the node density or traffic demand is low. A full pipeline is the unique choice when the data rate is high. At this time, a large $\mathcal{P}$ should be chosen to make sure that a transmission duration is long enough to execute at least one pipeline stage. The effect of $\mathcal{P}$ is further discussed in Section IV-E.

D. Singleton correlatable pilot

At the ith pipeline stage, a competing user should send a pilot message when $\mathcal{V}[i]$ is 1. As the pilot message collides with data packets, it is required that: first, the detection of a pilot signal should be robust in the presence of interference;

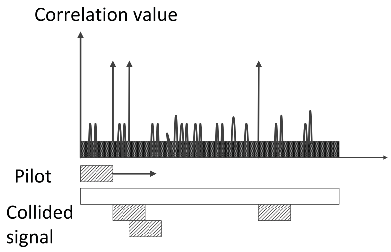

Fig. 9: Correlation-based detection.

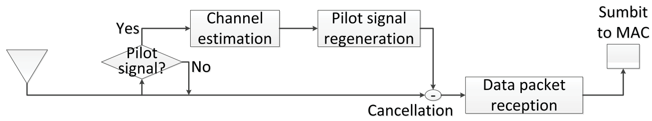

Fig. 10: Packet reception with pilot cancellation.

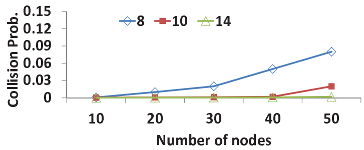

Fig. 11: Collision probability versus pipeline depth.

and second, the interference on the data packet decoding of the pilot must be removed.

We propose to use a singleton pilot message. First, a PN sequence is sent as a pilot. Based on the signal correlation technique, we can design a robust mechanism to detect whether or not there is a particular PN sequence in a collided signal. Second, the pilot is singleton. That is, the PN sequence adopted by all competing users is the same and already known by all nodes. For example, AP (access point) can select a pilot message and announce it to the clients at the association stage.

Let x denote the pilot sequence and H is the channel transfer function which includes effects such as attenuation, multipath, and Doppler shift. For a received signal $y = H_1x' + H_2I + No$, where I is the interfering signal (i.e., the data packet) and No is the noise, the correlation value is computed as below.

$$\Gamma(x) = y * \overline{x} = H_1x'\overline{x} + H_2I\overline{x} + No\overline{x}. \quad (1)$$

As x is a PN sequence, both $I\overline{x}$ and $No\overline{x}$ approaches 0. We have:

$$\Gamma(x) = H_1x'\overline{x} = \begin{cases} H_1|x|^2 & \text{if x=x'}, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Pilot detection: We use a sliding window to detect a pilot in a collided signal. The length of window is equal to that of a pilot. When there is no pilot, correlation value keeps at a very low level. A peak appears when the window aligns with a PN signal. Fig. 9 shows an example, where three pilots are inserted. According to the three peaks of the correlation value, one can locate all the pilots successfully.

It is necessary to choose a sufficiently long PN sequence to guarantee the detection robustness. The cross-correlation value between an arbitrary signal and a PN sequence is usually larger than 0. When the interfering signal is very strong, the cross-correlation value is so large to distort the detection. It is known that, a longer PN sequence is more robust. As shown in [17], a 20 Bytes PN sequence can be detected reliably even when the SINR is as low as -16dB.

Pilot cancellation: At a receiver node, to avoid the interference on the detection of a data packet, pilot signals should be removed from the collided signal. The cancellation process includes three major steps: (1) *Location*: We can locate the pilot signal precisely by signal correlation (see Fig. 9 for example). Energy detected can also facilitate the detection, i.e., the received power increases when a pilot appears and decreases at the end of the pilot. (2) *Estimation*: to estimate the channel parameters of the pilot in the presence of interference, we use the method described in [18]. (3) *Cancellation*: with the channel estimate, the arrived pilot signal can be generated and removed from the collided signal.

Recently, several studies questioned the effectiveness of interference cancellation (IC). It is argued in [19] that, to fully exploit the IC capability, one should need not be concerned with decoding but only with modeling and subtracting the interference. It is therefore critical to design the pilot as a singleton, which makes sure that, in our method, the pilot message is already known and no decoding is required.

E. Parameter analysis

There are three critical parameters of a contention pipeline:

- *slot length*, the time duration of a stage in a pipeline;
- *pipeline depth*, the number of stages of a pipeline;
- *parallelism level*, the number of concurrent transmission in a full pipeline.

Slot length: In a slot, the required operations include Tx/Rx switch, pilot transmission or detection. Therefore, the slot length is determined mainly by: (1) Tx/Rx turnaround time, i.e., t_{turn}; (2) propagation time, i.e., t_{prop}, which is equal to the inter-node distance divided by the light velocity; and (3) time required to detect the pilot, i.e., t_{dect}. In summary, slot length can be computed as

$$t_{slot} = t_{dect} + \max\{t_{turn}, t_{prop}\}. \quad (3)$$

In our implementation, we choose 20 Bytes PN sequence as a pilot. The detection time is $\frac{20*8}{4M}$=40 μs at 4Mbps with QPSK modulation. As recommended in IEEE 802.11 standard, both the turnaround time and the propagation time is less than 1 μs. In consequence, we set the slot length as 42 μs. In the first 2 μs, a short preamble is sent as a guard.

Pipeline depth: Collision occurs when there are more than one winner at the end of a contention pipeline. In general, less collisions occur when the pipeline depth is larger.

Note that, a contention vector is a binary sequence, thereby can be ordered from large to small. For a pipeline with depth

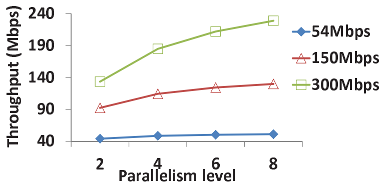

Fig. 12: Analytic throughput versus parallelism level.

$\mathcal{D}$, to choose a contention vector is equivalent to choose a number x $(0 \leq x < 2^{\mathcal{D}})$. For two users S_1 and S_2, the priority of S_1 is higher when the number chosen by S_1 is larger than that of S_2. Two users collides when they choose the same contention vector. Therefore, when there are N users, no collision occurs when: one user choose number x $(0 < x < 2^{\mathcal{D}})$, and the others choose numbers arbitrarily between 0 and x-1. The collision probability is

$$p_{col}^{3D} = 1 - \frac{N \sum_{1 \leq x < 2^{\mathcal{D}}} x^{N-1}}{2^{\mathcal{D}N}}. \tag{4}$$

Fig. 11 shows the collision probability of the 3D pipeline contention scheme in a network with different node number. First, collision probability is not high even for an 8-stage pipeline when the number of nodes is 50. The bitwise contention generates a sufficiently large space so that a collision event that two or more users choose the same minimum contention vector occurs rarely. Second, the collision probability reduces to 0.1% when the pipeline depth is 14 or more. We therefore do not need to use a very long contention pipeline.

Parallelism level: Let *Rate* be the original transmission rate. In a full pipeline, when channel is shared by $\mathcal{P}$ users, the transmission rate becomes $Rate/\mathcal{P}$ for a single user. The aggregate network throughput can be computed as:

$$Thr = \frac{payload}{t_{pkt}} = \frac{payload * \mathcal{P}}{t_{pre} + \frac{payload}{Rate/\mathcal{P}} + 2 * SIFS + t_{ACK}}, \tag{5}$$

where *payload* denotes the payload length, t_{pre} and t_{ACK} the transmission time of preamble and ACK, respectively.

The analytic results are plotted in Fig. 12 when *payload* is 1500 Bytes. A higher throughput is achieved with a larger parallelism level. Along with the increase of the transmission rate, a larger parallelism level is required to obtain a high throughput gain. However, when $\mathcal{P} = 6$, the throughput is already close to the nominal transmission rate, e.g., about 210Mbps at 300Mbps. In addition, it has many practical difficulties, i.e., synchronization, in realizing the simultaneous transmissions of a large number of users.

F. Discussions

In principle, AFD attempts not to reduce the control overhead, but to hide that under the data transmission. By using the interference cancellation and correlation-based detection techniques, the 3D pipeline contention scheme performs channel contention simultaneously with data transmission without introducing the harmful interference. In a conventional MAC protocol, it is critical to balance the time for coordination and that for transmission. In comparison, we can take a sufficiently long time to achieve high coordination efficiency while data packets are transmitted uninterruptedly.

Our scheme demonstrates a novel network-scale full duplex mode, which we called asymmetric full duplex (AFD). Conventional full duplex mode allows the same node to transmit and receive simultaneously. In our scheme, the whole network operates at the full duplex mode, i.e., transmission and contention is performed concurrently at different neighbor nodes. AFD provides a new way to exploit the communication resource near to an active link. In general, when a node is transmitting, the other nodes nearby must keep silent. That is, these nodes are idle but unavailable. By launching the contention among the idle nodes during the data transmission, our scheme can improve channel utilization significantly.

In addition, in a multihop network, when there are multiple neighborhoods, the node rankings do not obey any global order. Consider two neighborhoods (i.e., $\mathcal{N}_1$ and $\mathcal{N}_2$) and a common node S, assume there is no interference between the nodes in $\mathcal{N}_1 - \{S\}$ and those in $\mathcal{N}_2 - \{S\}$. When S is the winner in $\mathcal{N}_1$ but a loser in $\mathcal{N}_2$, all nodes in $\mathcal{N}_1$ including S cannot transmit. In fact, when a node in $\mathcal{N}_2$ except S transmits, there is another node in $\mathcal{N}_1$ can transmit. Finally, rate adaptation is widely deployed in modern wireless systems. It needs further investigation to design the pipeline contention scheme aware of rate adaptation.

V. PERFORMANCE EVALUATION

We evaluate the performance by simulations in Qualnet network simulator and measurements in a testbed. Each data point is obtained by averaging the results from 50 runs. Each run lasts for 5 minutes or until at least 5000 packets have been transmitted at every node.

A. Experimental Setup

We have implemented the proposed scheme on a software radio platform, the Universal Software Radio Peripheral (USRP) N210 [20]. To meet the stringent time requirement of MAC, we implement the pipeline contention process on FPGA while packet encoding/decoding are implemented in GNU Radio. When a packet is ready to send, it is buffered on FPGA and triggers the contention process.

We implement 802.11g (e.g., at data rate 54Mbps) in the USRP platform and simulate 802.11n modulation rates (e.g., 150/300Mbps) in Qualnet network simulator. Each USRP device has two antennas and a simulated node has four antennas. All nodes operate at the half duple mode and rate adaptation is disabled. The nodes generate fully backlogged CBR traffic, with packet size 1500 Bytes. Unless otherwise specified, all nodes are randomly positioned and able to sense each other. The performance metric is aggregate throughput and fairness. Throughput gain is defined as $(T_{3D} - T_{802})/T_{802}$, where T_{3D} is the throughput when the 3D pipeline contention scheme is used and T_{802} is that when 802.11 is deployed. We also evaluate the short-term fairness by the normalized Jain's index.

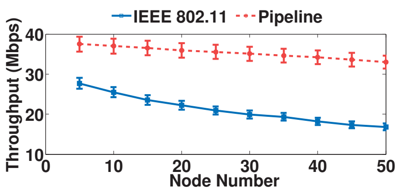

Fig. 13: Throughput versus the number of nodes at 54Mbps.

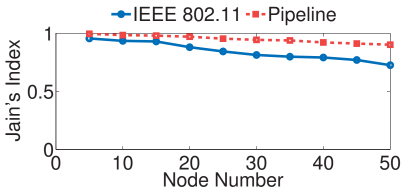

Fig. 14: Fairness versus the number of nodes at 54Mbps.

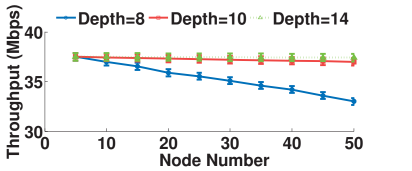

Fig. 15: Throughput versus the pipeline depth in a network with 10-50 nodes.

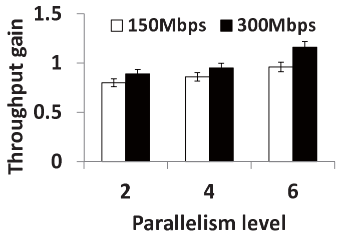

Fig. 16: Throughput gain versus parallelism level in a network of 30 nodes.

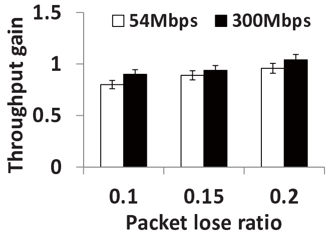

Fig. 17: Throughput gain in a 20-node network with channel fading.

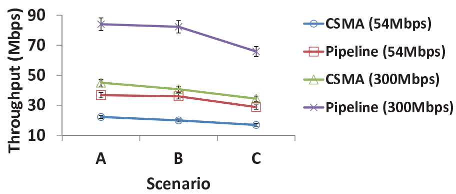

Fig. 18: Throughput in multihop networks at different transmission rates.

B. Results

Consider a single-hop 30m×30m network where a receiver is at center and all senders are randomly positioned. Fig. 13 and Fig. 14 show the throughput and the fairness of the 3D pipeline contention scheme at 54Mbps, respectively. The results of a 12-stage compact pipeline are shown. The node number varies from 10 to 50. It is clearly observed that, our scheme performs much better than IEEE 802.11. The throughput gain is on average 80%. Along with the increase of node number, the throughput of IEEE 802.11 decreases gradually as the collision probability grows up quickly. In comparison, with the 3D pipeline contention scheme, collision occurs rarely even when the number of nodes is up to 50. Moreover, though the fairness of IEEE 802.11 is relatively good (e.g., the Jain's index is larger than 0.8), our scheme further improves the system fairness, i.e., the Jain's index is around 0.95. We believe that, the nearly perfect fairness of our scheme results from the randomness of the contention vector chosen by each node.

We investigate the effect of pipeline depth in the USRP platform. Fig. 15 shows the aggregate throughput versus different pipeline depth in a network with 10-50 nodes. For a eight-stage pipeline, the throughput decreases along with the increase of node number. The collision probability is very low in small networks but grows up to about 10% in a 50-node network. When the pipeline depth increases, the network performance is improved drastically. For example, a 10-stage pipeline works much better than a 8-stage pipeline. As shown in Fig. 11, the collision probability of a 10-stage pipeline is lower than 2% in a 50-node network. It is unsurprisingly that, the performance of a 14-stage pipeline is almost the same as that of a 10-stage pipeline. Below, we only present the results when the pipeline depth is 12.

We also carry out simulations to investigate the effect of parallelism level. Fig. 16 shows the throughput gain at 150Mbps and 300Mbps in a 30-node network. The performance gain over IEEE 802.11 is between 80% and 120%. The results are consistent with the analysis in Section IV-E. With the increase of the parallelism level, a larger gain is achieved as the overhead (i.e., transmission time of preamble and ACK) is shared by more users. In the remaining experiments, the parallelism level is 6 for a full pipeline.

In practice, the received signal is affected by channel attenuation, multipath, and so on. The packet detection fails when the arrived signal is too weak. Fig. 17 shows the throughput gain in a network of 20 nodes. The packet loss ratio due to channel fading varies from 10% to 20%. The result of compact pipeline at 54Mbps and that of full pipeline at 300Mbps are shown. The performance gain is larger than 80% and up to 100% when the packet loss ratio is 20%. It is known that, contention window adjustment in IEEE 802.11 is easily distorted by the packet loss introduced by channel fading. When a packet is lost, contention window is doubled in IEEE 802.11 no matter the loss is induced by collision or channel fading. In comparison, our scheme is not affected by packet loss and robust in the presence of channel fading.

Finally, we present a preliminary result of the performance in a simulated multihop network. A large network topology is built where all senders can connect to the AP node but not all of them can sense each other. Consider three scenarios: (A) 15 nodes with no fading, (B) 30 nodes with no fading, and (C) 30 nodes with 10% packet loss. Fig. 18 shows the throughput gain in the three networks. We make two observations. First, our scheme outperforms IEEE 802.11 with a through gain as large as 80%. Second, the throughput is less than that in a full-connected network with the same size. The performance loss is about 20%. When a node belongs to more than one neighborhood, it cannot access the channel until it becomes the winner in all neighborhoods. Moreover, collision occurs frequently when there are hidden terminals. It is interesting to study how to integrate pipeline contention with the techniques to combat hidden terminal, which we leave as a future work.

In summary, the 3D pipeline contention scheme can effectively reduce the coordination overhead and the collision probability. It provides much better performance than IEEE 802.11, i.e., the Jain's fairness index is around 0.95 with a throughput gain between 80% and 120%. In a multihop network, further investigation is required to mitigate the efficiency loss due to the hidden terminal effect.

VI. Conclusions and Future Work

Coordination is critical in wireless networks for efficient medium access. This paper presents asymmetric full duplex (AFD) to achieve very high coordination efficiency at nearly zero overhead. We present a 3D pipeline contention scheme to realize AFD. It performs channel contention simultaneously with data transmissions. AFD demonstrates a network-scale full duplex capability. AFD also provides a novel way to make use of the communication resource close to an active link, which is idle but usually regarded as unavailable.

For future work, the pipeline contention mechanism can be extended from several aspects. For example, when rate adaptation is adopted, transmission time can be different even when all the data packets have the same length. It is our ongoing work to incorporate rateless codes with the pipeline contention to adapt to the varying channel condition. Also, in a multihop network, there is efficiency loss as the node rankings do not obey any global order. A similar problem is discussed in [10]. We plan to investigate whether or not the solution proposed in [10] can be adopted by our scheme.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under grants No. 61070203 and 61202484, and Doctoral Program Foundation of Ministry of Education of China under grant No.20124307120033.

References

[1] S. Lv, W. Zhuang, M. Xu, X. Wang, C. Liu, and X. Zhou, "Understanding the scheduling performance in wireless networks with successive interference cancellation," *IEEE Trans. Mob. Comput.*, vol. 12, no. 8, pp. 1625–1639, 2013.

[2] E. Magistretti, K. K. Chintalapudi, B. Radunovic, and R. Ramjee, "Wifi-nano: reclaiming wifi efficiency through 800 ns slots," in *Proc. ACM MOBICOM'11*, pp. 37–48, 2011.

[3] S. Lv, W. Zhuang, X. Wang, and X. Zhou, "Link scheduling in wireless ad hoc networks with successive interference cancellation," *Computer Netowrks*, vol. 55, no. 23, pp. 2929–2941, 2011.

[4] S. Sen, R. R. Choudhury, and S. Nelakuditi, "No time to countdown: migrating backoff to the frequency domain," in *Proc. ACM MOBICOM'11*, pp. 241–252, 2011.

[5] E. Magistretti, O. Gurewitz, and E. W. Knightly, "802.11ec: collision avoidance without control messages," in *Proc. ACM MOBICOM'12*, pp. 65–76, 2012.

[6] S. Yoon, I. Rhee, B. C. Jung, B. Daneshrad, and J. H. Kim, "Contrabass: Concurrent transmissions without coordination for ad hoc networks," in *Proc. IEEE INFOCOM'11*, pp. 1134–1142, 2011.

[7] K. Tan, J. Fang, Y. Zhang, S. Chen, L. Shi, J. Zhang, and Y. Zhang, "Fine-grained channel access in wireless lan," in *Proc. ACM SIGCOMM'10*, pp. 147–158, 2010.

[8] P. Huang, X. Yang, and L. Xiao, "Wifi-ba: Choosing arbitration over backoff in multicarrier wireless networks," in *Proc. IEEE INFOCOM'13*, pp. 771–779, 2013.

[9] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless lans," in *Proc. ACM SIGCOMM'05*, pp. 121–132, 2005.

[10] S. Sen, R. R. Choudhury, and S. Nelakuditi, "Csma/cn: Carrier sense multiple access with collision notification," *IEEE/ACM Trans. Netw.*, vol. 20, no. 2, pp. 544–556, 2012.

[11] S. S. Hong, J. Mehlman, and S. R. Katti, "Picasso: flexible rf and spectrum slicing," in *Proc. ACM SIGCOMM'12*, pp. 37–48, 2012.

[12] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 2012.

[13] J. Zhang, H. Shen, K. Tan, R. Chandra, Y. Zhang, and Q. Zhang, "Frame retransmissions considered harmful: improving spectrum efficiency using micro-acks," in *Proc. ACM MOBICOM'12*, pp. 89–100, 2012.

[14] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *Proc. IEEE INFOCOM'13*, pp. 1258–1266, 2013.

[15] K. Wu, H. Tan, Y. Liu, J. Zhang, Q. Zhang, and L. M. Ni, "Side channel: Bits over interference," *IEEE Trans. Mob. Comput.*, vol. 11, no. 8, pp. 1317–1330, 2012.

[16] A. Cidon, K. Nagaraj, S. Katti, and P. Viswanath, "Flashback: decoupled lightweight wireless control," in *Proc. ACM SIGCOMM'12*, pp. 223–234, 2012.

[17] S. Sen, R. R. Choudhury, and S. Nelakuditi, "Csma/cn: carrier sense multiple access with collision notification," in *Proc. ACM MOBICOM'10*, pp. 25–36, 2010.

[18] S. Gollakota and D. Katabi, "Zigzag decoding: combating hidden terminals in wireless networks," in *Proc. ACM SIGCOMM'08*, pp. 159–170, 2008.

[19] S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi, "Successive interference cancellation: A back-of-the-envelope perspective," in *Proc. ACM HotNets-IX*, pp. 11–15, 2010.

[20] USRP, "Ettus research," *http://www.ettus.com*, 2010.