In-band Wireless Cut-through: Is It Possible?

Bo Chen, Gopi Krishna Tummala, Yue Qiao and Kannan Srinivasan
Department of Computer Science and Engineering
The Ohio State University, Columbus, OH 43210
{chebo, tummalag, qiaoyu, kannan}@cse.ohio-state.edu

Abstract

This paper explores if wireless cut-through is possible. Unlike wired cut-through, wireless cut-through, if realized, can reduce latency and improve throughput by upto 3x. This paper shows that one way to implement cut-through is for every node to forward the previously decoded packet while receiving the current packet. This strategy is called decode-and-forward (D&F). Another way is to have every node forward without decoding. This strategy is called amplify-and-forward (A&F). This paper shows that both D&F and A&F have many shortcomings. It proposes a new way to realize cut-through that reduces latency and increases throughput over traditional routing.

This paper implements the first cut-through routing modules. Through multiple emulations of existing operational networks, it shows that cut-through switching can improve sum network throughput by up to 2.2x compared to traditional routing.

1. INTRODUCTION

In wireless communication, end-to-end latency is a critical factor for multi-hop applications such as disaster recovery and cellular wireless backbone. The realization of cut-through transmission in wireless communication can potentially have a huge impact.

Cut-through switching is a method used in wired networks. When cut-through is employed, a switch directly forwards a packet (without decoding it) to the next neighbour, while it is still receiving that packet. The realization of cut-through switching in wired networks greatly reduces the end-to-end latency. Realizing cut-through in wireless is harder than in wired networks since wireless links are not independent of each other. However, the benefit is also significant. Other than the latency reduction, cut-through can also increase end-to-end throughput in wireless. Assume traditional routing is carried out for the wireless network in Figure 1. Here, Alice wishes to send packets to Duncan. When the link from Alice to Bob is active, Cathy to Duncan cannot be scheduled since, otherwise, it will lead to interference at Bob. Therefore, for traditional routing, only one transmission is possible in one slot. When cut-through feature is enabled, however, all links can be simultaneously activated. Thus, the throughput gain could be as high as **3x** for wireless cut-through.

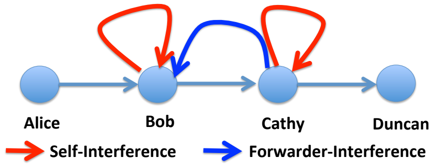


Figure 1: A Cut-Through Path. Alice wishes to send packets to Duncan through Bob and Cathy who are the forwarders.

What are the basic requirements to realize a wireless cut-through system? Full-duplex capability is the first technique required: Radios need to transmit and receive at the same time for all the nodes between the source and destination. In the past few years, several research groups [3, 4, 7, 5, 2] have already shown the feasibility of full-duplex wireless communications for Wi-Fi applications.

In addition to the self-interference caused by a transceiver to itself, the transmission from next hop also causes interference (noted as Forwarder-Interference in Figure 1). Therefore, in addition to self-interference, forwarder interference also needs to be canceled. A nice property of the forwarder interference is that the forwarders (Cathy) are sending the symbols that the previous nodes (Bob) along the path have already seen. In other words, the interference is known.

One obvious way to realize cut-through is to have every node along the cut-through path forward the previous packet while it is receiving the current one. We refer to this method as **Decode-and-Forward** (D&F) cut-through. There are many techniques in the literature supporting D&F [12, 8, 1, 6]. Thus, forwarder interference cancellation for D&F is also trivial. However, D&F has the following drawbacks. First, as packets need to be decoded before transmitted, the latency reduction based on cut-through feature disappears. Secondly, the throughput benefit would be limited if the cut-through path is used only for a short time. Consider an extreme case where only one packet needs to be delivered, which needs to go through one hop after another. The throughput performance would be the same as traditional transmission. Thus, D&F could be quite not useful.

Another way to realize cut-through is to have every node in the path forward their received signal without decoding it.

HotWireless'14, September 11, 2014, Maui, Hawaii, USA.

http://dx.doi.org/10.1145/2643614.2643623 .

We refer to this method as **Amplify-and-Forward** (A&F) cut-through. Existing packet-based cancellation techniques cannot be used to cancel forwarder interference in A&F. We should explore other ways to perform A&F forwarder-interference cancellation. For now, let's assume such a module is realizable. A&F cut-through will significantly reduce end-to-end latency. However, it has the following issues. First, any noise added at a node accumulates along the path since intermediate nodes don't decode. For a long path, this noise could deteriorate the throughput even below what traditional routing can achieve. Second, since intermediate nodes cannot change their symbols, the source needs to send symbols at the data rate supported by the bottleneck link in the path. This blocks all the links around this cut-through path for much longer than in traditional routing. Thus, A&F can reduce spatial efficiency and thus, reduce the overall network throughput. Therefore, A&F also looks not useful.

Should we then conclude that wireless cut-through is not worth exploring? This paper shows that cut-through is worth exploring and presents a viable design that improves both the latency and throughput compared to traditional routing. It combines D&F and A&F methods and takes advantage of both schemes.

How do we improve throughput and reduce latency at the same time? The intuition comes from the following observations. Firstly, to improve throughput over traditional routing, cut-through should be done over many consecutive hops along a path. It turns out that, if the number of consecutive hops is three or more, the desired 3x throughput gain is achievable. Secondly, to best weaken the noise accumulation effect and to allow cross-cutting flows, a D&F strategy, as in traditional routing, is desirable. D&F enables the decoding nodes to: (i) wait before forwarding so that other cross links can carry cross-cutting flows, (ii) choose the appropriate data rate at which to send to the next decoding node along the path, and (iii) drop any undecodable packets and stop error propagation. Thus, our proposal is to split the entire cut-through path into smaller cut-through paths. We refer to these smaller cut-through paths as ***Virtual Hops***. A virtual hop, thus, can have multiple nodes switching packets without decoding. The end nodes of a virtual hop will decode the packets and then decide the data rate at which to forward the packets.

This paper makes the following contributions:

(i) It presents a way to realize wireless cut-through. We present ***virtual-hop*** cut-through routing. Virtual-hop combines the basic principles of A&F and D&F to improve throughput and reduce latency. (ii) It presents a forwarder cancellation technique that supports A&F cut-through within a virtual-hop. (iii) Finally, it presents the first in-band cut-through system on the software-defined radio NI PXIe-1082 with NI 5791 RF frontend. The system achieves a 35dB cancellation in line-of-sight scenario and 30dB cancellation in non-line-of-sight scenario. Our testbed measurements show that this cancellation is enough to support forwarder-interference cancellation.

2. DESIGN OF CUT-THROUGH

This section explores the challenges and feasibility of cut-through, and also proposes the design of cut-through routing. In general, a cut-through path involves a source node, a destination and a sequence of forwarders as shown in Figure 1. The transmissions on all the links happen simultaneously as (1) Source is transmitting all the time, (2) Forwarders keep forwarding the received signal, and (3) Destination keeps decoding the signals.

2.1 Forwarder Cancellation Design

The source and destination nodes in a cut-through path are just operating basic transmitting and receiving functions, respectively. However, the other nodes operating in full-duplex mode need to cancel self-interference in order to forward while receiving. In addition, all the forwarders (except the last one) need to cancel the interference from the next-hop forwarder. Figure 2(a) illustrates the structure of a node when it's a forwarder. As shown in the figure, in the forwarding mode, a node only uses the digital-to-analog-converter (DAC), analog-to-digital-converter (ADC) and radio frequency (RF) parts of its radio physical layer. The only processing for data stream a node does in the digital domain is interference cancellation, and then amplification of the received signal. It bypasses all of the other traditional physical layer data processing such as decoding and error correction. Clearly, the MAC and higher layer processing are also bypassed.

In general, there are three signal streams that a forwarding radio deals with: (1) Data stream from the previous radio. It is the data the forwarder intends to receive and forward. (2) Self-interference signal stream. Forwarders need to transmit and receive on the same bandwidth at the same time. This leads to self-interference. (3) Forwarder-interference signal stream. It is the Data stream from the next hop radio.

To accurately receive the first data stream, the two interference streams need to be dealt with. In recent years, multiple works have been presented to deal with self-interference. The self-interference cancellation module shown in Figure. 2(a) is inherited from work [7] which separates the cancellation into two parts: an analog module cancels 70dB self interference and a digital module adds another 40dB cancellation. This is also the deployment of our experiment.

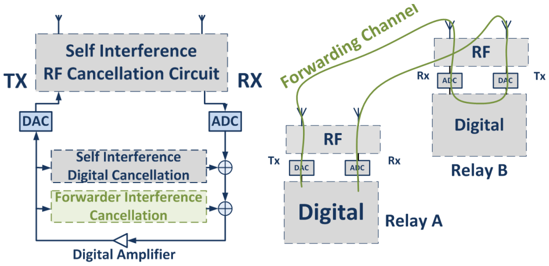
The forwarder-interference cancellation module is independent from the self-interference cancellation module. In order to develop an efficient forwarder-interference cancellation module, we need to understand the nature of the forwarded interference signal.

First, the forwarder interference signal is known because the forwarder gets the signal from its predecessor where the signal needs to be canceled.

Secondly, a forwarder interference signal stream comes from the next hop wireless transmission and the intended received data signal comes from the previous hop. Therefore, the forwarder interference signal power and the regular received signal are comparable: the two signals are within the resolution of analog to digital converter (ADC). This is different for self-interference which is many orders larger than the intended received signal.

Thirdly, the received signal, after going through self-interference and forwarder cancellation steps, is then directly amplified and sent to the transmitter side for forwarding. Thus, the received data signal stream lasts in a forwarder node for a fixed amount of time. Note that this forwarded data signal is observed as the forwarder-interference at the previous hop node. This means that the timing of the forwarder-interference signal is usually small and predictable.

The last aspect is the channel. Once this information is known, the forwarder cancellation module can simply apply this transformation to its previously sent data signals in order to replicate the forwarder interference signal. The cancellation performance will depend on how accurate the replication is. We refer to this "transformation" as the ***forwarding channel***, as shown in Figure 2(b). There are two nodes, Nodes A and B, along a cut-through path. Note that, same as self-interference, the origin and sink of the forwarder interference are the same node. Instead of the wireless channel from Node B to Node A, the ***forwarding channel*** encompasses three entities: (i) the wireless channel from Node A to Node B, (ii) the amplification circuitry within Node B, and (iii) the wireless channel back from Node B to Node A. Note that the amplification circuitry at Node B simply has a fixed gain. As the channel is predictable, the forwarding channel has a lot of similarities to the self-interference channel. Thus, the forwarder cancellation module in the system is designed following the structure of digital cancellation in self-interference cancellation: FIR filters are used to characterize the channel. Thus, even the multi-path components can be captured in the channel characterization.



(a) Structure of forwarder with our Forwarder Interference Cancellation Module

(b) Forwarding Channel Illustration

Figure 2: *Forwarding channel*

Is Frequency Offset an Issue for the forwarding channel? It's widely known that any two wireless nodes will have a frequency offset: the notion of 2.4GHz could be different as the crystals used to generate these frequencies have imperfections [6]. A frequency offset between two nodes, if uncorrected, will rotate the phase of a signal, continuously over time [10]. This continually changing phase can make channel estimation between two nodes extremely difficult: Even if the channel is relatively stable over time, the rotating phase will affect these estimates. To solve this problem, it is common to use a known sequence as a preamble for every packet. A receiver uses this preamble to estimate the frequency offset so that it can correct phase rotations.

Note that the digital cancellation for self-interference (in full-duplex radios) does not have this problem. This is because the transmit and receive radios belong to the same node and so there is no frequency offset. The forwarder interference in cut-through, however, is introduced by another node's –the nexthop node's– transmission. Therefore, in Figure 2(b), Node A's and Node B's crystals will have a frequency offset. Does this imply that a frequency offset correction mechanism is required at all the radios in the network? The answer is ***no***. Actually, the forwarding channel defined in our system is a ***frequency offset-free*** channel, even though there is frequency offset between two nodes.

Assume in Figure 2(b) that the actual center frequencies at Nodes A and B are f_A and f_B, respectively and $f_A \neq f_B$. Thus, the data received in forwarder B is affected by frequency offset $f_B - f_A$. Since nodes in cut-through don't decode signals, they don't correct for frequecy offset. Therefore, Node B directly forwards the signal with frequency offset. As it goes through the channel B to A, an additional frequency offset $f_A - f_B$ is applied to the signal when it's received by Node A. Thus, the two frequency offsets cancel out with each other and the forwarder-interference signal received in node A is frequency offset-free. It makes ***forwarding channel*** estimation at any nodes simple.

2.2 Virtual-Hop Structure

As Figure 1 shows, cut-through transmission can triple the throughput compared to traditional wireless routing. It seems that combining the source, destination and all the forwarder nodes in between into a single cut-through path is the right way to deploy cut-through in an ad hoc network. We call this cut-through as end-to-end cut-through. However, several realistic situations limit its performance:

(1) Worst link bottle neck: Cut-through routing does not take full advantage of the diversity of the link quality in a path. Since no intermediate node can modify the received symbols, the source needs to pick the transmission rate suitable for the bottleneck link along the path. This is the data rate at which all the links in that path will operate. It means that every link will carry the packets at a data rate supported by the bottleneck link even if the other links are far superior. In traditional routing, however, a packet can be transmitted at higher data rates on other links and sent at a low rate on the bottleneck. This frees up the medium sooner around (superior) links than around the bottleneck allowing other flows to start early.

(2) Noise accumulation : With cut-through, more number of links are enabled simultaneously. Therefore, throughput is improved when compared to traditional routing. However, since every node in cut-through needs to forward without decoding, noise accumulates in every hop. This leads to deteriorating SNR as the packet traverses towards the destination. If the deterioration is significant, then it can lead to a throughput even lower than traditional routing. Additionally, an erroneous packet is dropped at an intermediate switch in traditional routing. While in cut-through, errors go undetected until a packet is decoded by its destination. The propagation of errors wastes network resources.

It seems that, in a real scenario, cut-through routing over a long path is not efficient. To resolve this limitation, we introduce the concept of Virtual-Hop. Traditionally, one hop refers to a transmission from one node to its neighbouring node. Comparatively speaking, a virtual-hop represents a transmission from one node to a node several hops away. Within a Virtual-Hop, amplify and forward is carried out and the endpoint of a virtual hop decodes a packet. Thus, between a source and destination, there could be multiple virtual hops. We refer to such a cut-through structure as *virtual hop* structure.

Virtual hop cut-through limits the number of hops to do amplify-and-forward (A&F) over. Thus, virtual-hop mitigates the bottleneck link effect and error propagation problem. At the same time, when the hop count inside a virtual hop reaches three, the benefit of throughput improvement is already significant. That's the reason virtual-hop structure is the right way to the realization of wireless cut-through.

To realize the virtual hop structure, close cooperation among the neighbour nodes is needed. To reduce the MAC overhead, control messages need to be well designed. In the recent literature, several works like 802.11ec [9] have already shown that the size of the control messages could be really small. Thus, virtual hop structure can be realized without sacrificing too much overhead.

2.3 Cross-Hop Interference

Under the binary interference model assumption, links either interfere with each other or not at all. However, in real scenarios, any two links can interfere with each other. In traditional routing, if the interference is severe, one link can be disabled. However, for the cut-through system, such a setting is contradictory to the idea of enabling the links simultaneously. The forwarder-interference cancellation module only cancels interference from the immediate forwarder. We refer to the interference introduced by the radios more than one hop away as cross-hop interference(or k hop interference if the number of hops is known). Referring back to Figure 1, Alice can introduce cross-hop interference at Cathy and Duncan. Bob can affect the reception at Duncan. One possible way is to treat the signal as multi-path component. However, through our testbed evaluation, the average latency introduced in one relay node is around $1\mu s$. Compared with the 0.8 μs Cyclic Prefix (CP) length for the standard 802.11a protocol, even two hop interference cannot be covered by the cyclic prefix. Thus, the only solution without modifying the standard OFDM protocol is to treat the interference as noise. Under this setting, while the cut-through path is setup, the source should pick a lower datarate to account for this cross-hop interference. This could lower our throughput gain from 3x. This is part of the reason why our results show a 2.2x gain compared to traditional routing in Section 3.

3. IMPLEMENTATION AND EVALUATION

This section presents the implementation details of our cut-through modules. It also presents the evaluation results for our forwarder cancelltion module and the throughput results from our trace-driven emulations.

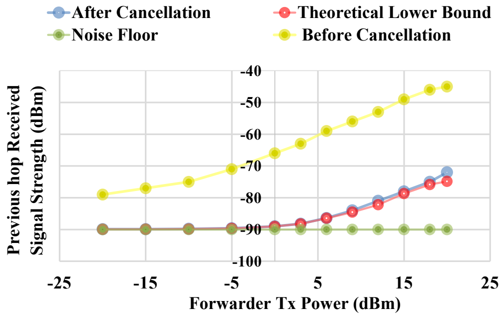
3.1 Setup

To verify the feasibility of cut-through transmission, we implement a system which can support both self-interference cancellation and forwarder cancellation modules. Our radio system has two RF chains and NI PXIe-1082, an RTOS-based controller. Each RF chain consists of NI-5791 (RF frontend and data converter module) and NI PXIe-7965R (Xilinx Virtex-5 FPGA). It can transmit and receive the OFDM signal with QPSK, QAM16 and QAM64 constellations. The communication system supports a large range of bandwidth up to 20MHz (same as Wi-Fi's bandwidth).

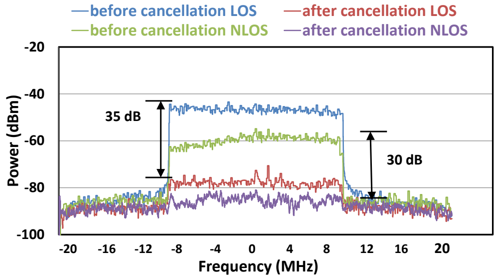
We implement the self-interference cancellation modules to support full duplex transmission. Our full duplex implementation follows the design in [7]. Our transmitting and receiving antennas are 15 cms apart. Together with the Balun cancellation circuit, our RF module can support 70dB self-interference cancellation for 20MHz bandwidth signal. The digital module can provide another 40dB cancellation. When combined, our full duplex system supports the 110dB cancellation needed to support Wi-Fi [2]. We operate in our system at a center frequency of 2.45 GHz.

3.2 Forwarder Cancellation Performance

To evaluate the forwarder cancellation performance, we conduct a sequence of experiments at a bandwidth of 20 MHz. We consider the worst case where the link between the previous hop and the forwarder is excellent since this will increase the forwarder interference at the previous hop node. The SNR at the previous hop is 30dB, and the white noise floor is at -90dBm. We tune the transmission gain at the forwarder. Figure 3a shows that as the transmission power at the forwarder increases from -20dBm to 20dBm, the residual power received by the previous hop node increases from -90dBm to -73dBm. Note that the forwarder interference not only has the signal originally sent by the previous hop, but also has the white noise from the forwarder node. Thus, even theoretically, the forwarder interference cannot be cancelled perfectly. Figure 3a shows this theoretical lower bound, which is not exactly reaching the noise floor. It also shows that the cancellation achieved by our module is within 3dB difference of the theoretical lower bound. This noise will also lower the datarate of cut-through transmissions.



(a) Received signal power level at the previous hop vs transmit power at forwarder. Our cancellation is within 3dB of the theoretical lowerbound.



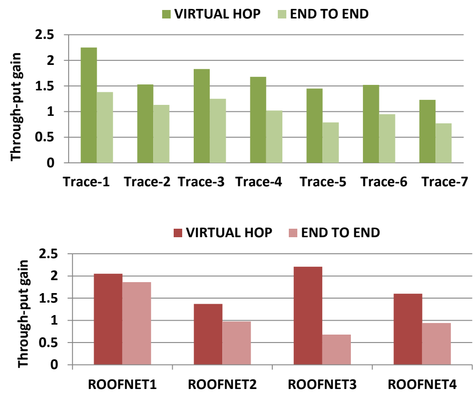
(b) Cancellation for LOS and NLOS links

Figure 3: Forwarder Cancellation Performance

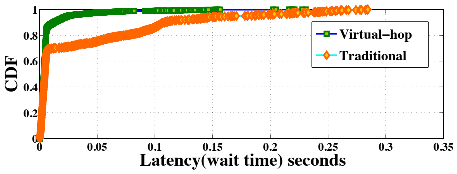
Figure 3b illustrates the cancellation performance under two settings: line-of-sight (LOS) and non-line-of-sight (NLOS). It shows that we can achieve 35dB cancellation in the LOS setting and 30dB cancellation in the NLOS setting. Note that the forwarder cancellation module is inherited from the digital cancellation module from the full duplex system. However, the performance is slightly worse: full duplex's digital cancellation module can achieve 40dB cancellation, while forwarder cancellation module can achieve only 35dB (in the LOS setting). Digging deeper, we find that this is due to the drift of the clock for two radios. With a better clock source at our nodes, better cancellation is achievable.

3.3 Trace-Driven Evaluation

We built an emulator to test the scalability and feasibility of cut-through routing. The emulator takes the SINR values and data rates across various links from our network and MIT Roofnet network [11] data traces. Trace-1 to Trace-7 are collected on our Wi-Fi testbed which comprises of live data taken from 40 nodes under 7 different transmit power levels. Roofnet data are taken from 39 nodes. Further, we feed the following measurements from our experiment to the simulator to emulate full-duplex cut-through physical layer. **(i) Interference cancellation:** The self-interference and forwarder-interference cancellation performance depends on the power of transmission and also the topology of the nodes. **(ii) Noise Accumulation:** As the number of hops for which a packet is being forwarded increases the throughput decreases. This increased BER is measured from our experiments for various hop-lengths and different SINR values. This effect is emulated by recalculating the effective SINR values across 2-hop, 3-hop, 4-hop routes with increased BER.



(a) Throughput gains of cut-through routing and end-to-end routing from emulator using the trace we collected and MIT Roofnet network data.



(b) CDF of Latencies of virtual-hop and traditional routing.

Figure 4: Through-put and Latency gains from emulation

Figure 4a shows the throughput gain achieved by two cut-through techniques over traditional routing. The end-to-end cut-through does not employ virtual-hop and carries out A&F along the entire path. The other cut-through uses virtual-hop. Figure 4a shows that virtual-hop can have gain of up to 2.2x. End-to-end, on the other hand, can do worse than traditional routing. This is because end-to-end can easily accumulate noise as it does not employ virtual-hop.

Figure 4b shows the average latencies of 4-hop and traditional routing for 2 to 6 hop lengths based on the emulation of Trace-1. For over 70% of time, the latency of traditional is same as virtual-hop. This is because 70% of packets received are only one hop away. Given that 70% of the packets have taken only 1-hop, it must be surprising that our throughput gain is still 2.2x. This is because the number of 1-hop packets is different for traditional routing and virtual-hop. The 2-hop, 3-hop, 4-hop virtual-hop paths clear their paths quickly for more number of 1-hop transmissions. Due to this, the number of 1-hop transmissions in virtual-hop routing is much more than in traditional routing, contributing to the gain achieved.

4. CONCLUSION AND FUTURE WORK

This paper showed the feasibility of wireless cut-through transmission. However, there are still several other challenges to be solved to make it fully practical:

Cut-through transmission, especially the A&F cut-through requires close cooperation among the neighbouring nodes. An efficient MAC protocol which can integrate the cut-through transmission into existing CSMA WiFi system is quite important to the realization of in-band wireless cut-through.

Our current system treated the cross-hop interference as noise. However, there should be a smarter way to deal with that. Given the interference can be predicted and estimated, replacing the current digital amplifier with an FIR filter would be a potential way to mitigate this interference.

5. ACKNOWLEDGMENTS

This material is based upon work partially supported by the National Science Foundation under Grant CNS-1254032.

References

[1] Bansal, T., Chen, B., Sinha, P., and Srinivasan, K. Symphony: Cooperative Packet Recovery over the Wired Backbone in Enterprise WLANs. In *Proc. ACM MOBICOM* (2013).

[2] Bharadia, D., Mcmilin, E., and Katti, S. Full Duplex Radios. In *Proc. ACM SIGCOMM* (2013).

[3] Choi, J. I., Jain, M., Srinivasan, K., Levis, P., and Katti, S. Achieving Single Channel, Full Duplex Wireless Communication. In *Proc. ACM MOBICOM* (2010).

[4] Duarte, M. and Sabharwal, A. Full-duplex Wireless Communications Using Off-the-shelf Radios: Feasibility and First Results. In *ASILOMAR* (2010).

[5] Duarte, M. Sabharwal, A., Aggarwal, V., Jana, R., Ramakrishnan, K. K., Rice, C. W., and Shankaranarayanan, N. K. Design and Characterization of a Full-duplex Multi-antenna System for Wifi Networks.. *IEEE Transactions on Vehicular Technology* 3 (2012).

[6] Gollakota, S., and Katabi, D. Zigzag Decoding: Combating Hidden Terminals in Wireless Networks. In *ACM SIGCOMM* (2008).

[7] Jain, M., Choi, J. I., Kim, T., Bharadia, D., Seth, S., Srinivasan, K., Levis, P., Katti, S., and Sinha, P. Practical, Real-time, Full Duplex Wireless. In *ACM MOBICOM* (2011).

[8] Katti, S., Gollakota, S., and Katabi, D. Embracing Wireless Interference: Analog Network Coding. In *ACM SIGCOMM* (2007).

[9] Magistretti, E., Gurewitz, O., and Knightly, E. 802.11 ec: Collision Avoidance Without Control Messages. In *ACM MOBICOMM* (2012).

[10] Rahul, H., Kumar, S., and Katabi, D. JMB: Scaling Wireless Capacity with User Demands. In *ACM SIGCOMM* (2012).

[11] RoofNet Data Traces. http://www.pdos.lcs.mit.edu/roofnet/, 2003.

[12] Rahul, H., Kumar, S., and Katabi, D. Sam: Enabling Practical Spatial Multiple Access in Wireless Lan. In *ACM MOBICOM* (2009).