Full-duplex (Two-way) Wireless: Antenna Design and Signal Processing

Amir K. Khandani
Electrical and Computer Engineering Department, University of Waterloo, Waterloo, ON, Canada

***Abstract*—Current wireless systems are one-way (similar to walkie-talkies), meaning that disjoint time or frequency segments are used to transmit and to receive. Realization of two-way wireless has challenged the research community for many years. This article[1] establishes the theory and presents practical realization of two-way (true full-duplex) wireless. In contrast to the widely accepted beliefs, it is shown that two-way wireless is not only feasible, but is fairly simple, with virtually no degradation in signal-to-noise-ratio[2]. The innovation is in the antenna design and multiple levels for cancelling self-interference. Methods are developed to support Multiple-Input Multiple-Output (MIMO) two-way transmission (increasing multiplexing gain, and/or diversity order). The developed hardware (operating over 2.4 and 5Ghz unlicensed 802.11 bands with 20 or 40Mhz bandwidth, and 400-800MHhz white space band with 6Mhz bandwidth), uses off-the-shelf components, antennas have a simple structure, are omnidirectional (can be directional, if needed), do not suffer from bandwidth limitations, have a small size/spacing, and the increase in overall complexity vs. legacy one-way systems is minimal. The setup is extensively tested in harsh (significant reflections from surroundings) indoor and outdoor environments and the achieved performance in each link is virtually the same as the corresponding one-way system.**

I. Introduction

A communication link with capability to support connections in both transmit and receive directions at the same time and over the entire frequency band is called full-duplex, or two-way. In contrast, a link that can support connection in only one direction at a time (over a given frequency band) is called one-way or half-duplex. Current wireless systems are one-way and rely on either separate time slots (time division duplex) or separate frequency bands (frequency division duplex) to transmit and to receive. These alternatives have their pros and cons, but both suffer from lack of ability to transmit and to receive concurrently over entire frequency band. Even in the context of Orthogonal Frequency Division Multiple Access (OFDMA), where different frequency tones are used to simultaneously service multiple users, there is no method known to use the OFDM tones in opposite directions. A similar shortcoming exists in the context of Code Division Multiple Access (CDMA). Although two-way wireless is theoretically possible, its implementation is difficult due to an excessive amount of self-interference, i.e., the interference each transmitter generates to its own receiver(s).

Full-duplex communication is currently used in many applications, e.g., wired telephones, digital subscriber line, wireless with directional antennas, and free-space optics. The impact of full-duplex links in these earlier applications is limited to doubling the rate by providing two symmetrical pipes of data flowing in opposite directions. In contrast, in multi-user wireless systems, due to the broadcast nature of transmission (everyone hears everyone else), full-duplex capability has the potential to do more than merely doubling the rate, e.g., it facilitates networking, collaborative transmission, and security.

To cancel the self-interference in analog domain, an Auxiliary Transmit signal (ATX) is generated from the Primary Transmit signal (PTX) and added to the received signal in the analog domain. Prefiltering, e.g., by pre-weighting coefficients applied to OFDM tones, are calculated for the ATX signal to cancel the self-interference. ATX can be radio frequency (RF) modulated and added to (i.e., coupled with) the received signal in the Radio Frequency (RF) domain prior to Low Noise Amplifier (LNA). It can be also added to the received signal in analog base-band prior to Analog-to-Digital converter (A/D), at the cost of using LNA with a larger dynamic range. In addition to cancellation in the analog domain, digital subtraction is deployed at the receive base-band to further reduce the self-interference. In particular, linearity of the Digital-to-Analog converter (D/A) is exploited to subtract the remaining amount of self-interference from the base-band received signal (while maintaining and benefiting from underlying OFDM structure).

Symmetrical transmit and receive antennas are relatively positioned to reduce self-interference. In two dimensions, pair-wise symmetric antennas are proposed which have (theoretically) zero coupling over entire frequency range. The idea of symmetry is generalized to three dimensions. It is shown there exist triple-wise symmetric antennas with zero coupling between any pair. For Multiple-Input Multiple-Output (MIMO) transmission, two sets of such antennas (to be used for transmit and receive) can be arranged in three dimensions such that any antenna in one set is decoupled from all the antennas in the other set. Such three dimensional structures can be also implemented in 2.5 dimensions using layers of a Printed Circuit Board (PCB), e.g., by using patch antennas where one arm of antenna is generated through reflection of the other arm in the ground plane.

As an alternative to pair-wise symmetrical structures, methods are developed to place one set of antennas in the plane of symmetry of another set, which is shown to be an equipotential surface (in the absence of the scattering due to the placement of the second set of antennas). Examples of such constructions are presented where the same patch is used as the transmit antenna, the receive antenna, and the ATX coupler. This

[1]Supported by Ontario Ministry of Research and Innovation (ORF-RE).
[2]Due to space limitations, see [1] for details on performance measures.

construction is also generalized to MIMO.

Implementation: RF transmission is based on 802.11 using a 20MHz channel at 2.4 GHz and 5Ghz bands. Transmission power is about 20dbm which is typical for cellular applications. The basic physical layer follows 802.11 in terms of OFDM structure, preamble, synchronization, etc. For hardware implementation, the software defined radio platform by Lyrtech (now Nutaq) is used, and the final outcome has been tested in outdoor and indoor environments, and it essentially works as reliably as a one-way system. A second implementation is for White Space applications, using 6Mhz TV channels selectable over a band of 300 to 800Mhz.

A. Literature Review

Two-way wireless has been of interest over a relatively long period of time and there have been some other works addressing this problem [1]-[11]. Author's initial interest in this topic started in 2004, followed by a provisional patent in 2005, actual patent filed in 2006, which was issued in 2010 [2]. The starting point for the author's work was to use multiple transmit antennas with transmit beam-forming to create a null at the position of a receive antenna. In particular, using two transmit antennas with 180 degree phase shifts to create a null at the position of a receive antenna which is positioned in the middle of the two transmit antennas. The same antenna structure was later rediscovered in [10] [11]. Current article presents a more advanced design. The contents of this article have been publicized on-line in April 2012 [1]. There are several critical components contributing to the excellent performance of the method reported here as compared to the results reported by others, and in particular by research teams from Rice [6]-[9] and Stanford [10][11]:

- **Antennas** are designed to provide (theoretically) zero self-interference over the entire frequency range, including support for MIMO. [3].
- **Analog Active Cancellation** exploiting the linearity of D/A with proper training for channel measurement. Overall, active cancellation is done in a way that it does not contradict linearity in the cancellation path. As a result, active cancellation does not need to be precise and any such lack of precision, which is unavoidable, will be accounted for (subsequently measured) and compensated in the next step in the digital base-band cancellation.
- **Power Amplifier (PA) modeling and compensation**.
- Methods to deal with other imperfections, in particular: (i) computational errors caused by numerical inaccuracies (e.g., rounding) in FFT/IFFT and filtering operations, (ii) dealing with non-idealities in RF modulation/demodulation, e.g., phase jitter, and (iii) methods to optimize accuracy in fixed point arithmetic prior to D/A.

In addition, compared to other research works, this work includes support for MIMO.

The rest of the paper is organized as follows. Section II describes the full-duplex wireless network under consideration, including the proposed self-interference cancellation techniques. Antenna design is presented in Section III. Methods for modeling and compensation of PA nonlinearity are presented in Section IV. Finally, Section V presents some concluding remarks.

[3] [6]-[9] do not discuss antenna design and [10][11] rediscovered the same antenna structure as in the author's issued patent [2]

II. Two-Way Connectivity

Consider the full-duplex communication network shown in Fig. 1, in which an access point is communicating with multiple clients. The access point employs OFDMA to service multiple users with full-duplex connectivity over each OFDM tone. The access point can also support multiple transmit and multiple receive antennas to further exploit spatial degrees of freedom to increase rate and/or diversity. In addition, the access point supports new incoming clients which can asynchronously join the network (without prior time/frequency synchronization). The key challenge in the operation of this network is the self-interference between each component's transmit and receive chains. Thus, the ability of self-interference cancellation is the main requirement for the nodes of this network.

Consider a main signal composed of a "desired part" contaminated with "interference". Active cancellation, which is a well developed and widely used technique, is based on forming a corrective signal that, when combined with (e.g., added to) the main signal, will cancel the effect of the interference. Active cancellation is an obvious building block for cancelling self-interference in two-way wireless. However, two main issues make the application of active cancellation in two-way wireless different from conventional scenarios:

Active cancellation is based on linearity assumption, i.e., interference is added to the main signal through linear filters. Imperfections in the transmit signal, such as those due to nonlinearity and/or noise of the PA, or due to mathematical imprecisions, cannot be compensated through linear active cancellation. Even the nonlinear compensation methods proposed here will fail if the level of nonlinear leakage is high. Due to this reason, it is crucial to minimize the coupling between transmit and receive chains prior to active cancellation stage, as a lower coupling at this stage will reduce both the linear and the nonlinear portions in the transmit-to-receive leaked signal. This enables a level of overall isolation that would not be feasible through linear signal processing techniques. As an example, our study shows that the numerical errors in a 14 bits 64-points IFFT operation, with an optimized fixed point arithmetic design, is about -80dB below the signal level. This is also in accordance with the theoretical resolution of a 14 bits digital-to-analog (D/A) converter, which is widely available and a reasonable D/A choice for cost effective implementation. Under this condition, a transmit signal at 30dBm results in numerical errors at -60dBm. If RF isolation between transmit and receive chain is -50dB, then the numerical errors (at -110dBm) will be below the thermal noise level. In practice, we require such imperfections to be comparable to the noise level, which enables subsequent signal processing to account for and compensate the nonlinear effects. This is possible only through the antenna design techniques propose here. It is important to realize that, although there may be research works

reporting higher RF isolations in simulation or in controlled laboratory environments, in practice, -50dB RF isolation is extremely difficult to achieve. This article introduces novel antenna design and symmetry design conditions that will in theory being the RF isolation to $-\infty$dB, and in practice, an RF isolation of about -50dB is obtained due to reflections from neighboring environment.

It is important to be able to have both Analog and Digital active cancellations, where the analog cancellation is prior to A/D, and digital cancellation is in digital base-band. In general, if the analog active cancellation is not properly designed, the benefits of digital cancellation can disappear, or even add more degradation due to various imperfections present in the corrective digital signal. This article presents methods to addresses this issue such that digital active cancellation is always helpful. This is based on exploiting the fact that unlike A/D, D/A operation is (in theory) linear. In practice, D/A operation is not entirely linear either, but the degradation from linearity will be significantly lower than the power of D/A input signal, which in this case is a corrective signal with relatively low power[4]. This enables us to form an analog corrective signal for active cancellation and add it to the incoming signal prior to A/D without violating linearity which is essential for OFDM operation. The corrective signal is formed in baseband using some weights for each OFDM tone, or the equivalent realization of the filtering operation in time. Filter weights are obtained by sending training signals and measuring the self-interference. It is not possible to have accurate measurements of these weights. The reason is that we are dealing with large quantities, while interested in measuring error terms that are relatively small. As a result, various imperfections, including additive noise, affect the accuracy. This work exploits the property that, as D/A operation is linear, measurement of filter structure for analog active cancellation does not need to be accurate. An error in this measurement will act as an additional parasitic linear system (referred to as "equivalent channel hereafter), which is subsequently measured and accounted for in the digital cancellation. In other words, the equivalent channel for the residual self-interference remains linear and can be used in conjunction with OFDM. Consequently, as long as we fix the weights computed for active cancellation (with whatever error they may contain) and then accurately measure the equivalent channel for the residual self interference, we can successfully apply the final stage of digital cancellation in base-band. To improve accuracy of measuring the residual self interference channel, we rely on sending training sequences with higher power, and repeat the measurement several times and average the results. In addition, the first stage of self-interference reduction using active cancellation may result in cancelling most of the interference, and consequently, the residual self-interference (and weights for its equivalent base-band channel) will become too small, in which case the final stage of digital cancellation is bypassed.

ATX signal, which is added at the receive chain to reduce the self-interference in analog domain prior to A/D (preferably prior to LNA), can be constructed by weighting each OFDM tone of TX signal by a proper value to cause cancellation. The filtering operation to construct ATX from PTX can be also implemented in time domain; in which case only one IFFT block is used. In any event, measurement of the filter coefficients can be performed in the frequency domain. ATX chain is designed to have a high coupling with the RX chain. This avoids the use of a Power Amplifier (PA) for the ATX chain and consequently helps to maintain linearity in the ATX path. In this case, the non-linearity of the PA in the PTX chain is modeled in time, using measurements in frequency domain. Due to the linearity of the ATX chain and the fact the PA non-linearity is invertible, one can construct a proper base-band ATX signal such that the overall effects of the PA non-linearity and the filtering operations due to H_1 and H_2 are compensated.

[4]The corresponding power is determined by the power of the residual self-interference which at this stage is relatively low due to antenna isolation and prior analog active cancellation.

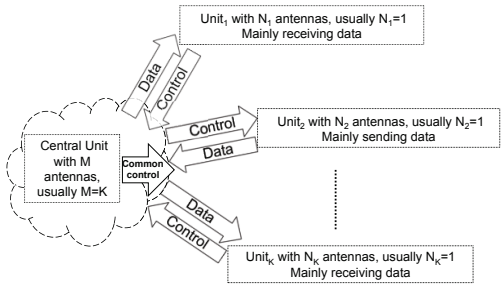

Fig. 1: $2K$ pipes of data/control are established over the same time/frequency where control includes reference for time/frequency/clock synchronization, channel gain/phase, information for user selection and channel inversion in SDMA, channel matrix in MIMO, ARQ, power control, instruction for adaptive coding and modulation, etc. In SDMA down-link, main data flow is out of central unit, while in SDMA uplink most data flow is into the central unit. Common control includes reference for time/frequency/clock synchronization.

Figures 2 and 3 illustrate abstract views of the system. PTX and ATX signals are pre-weighted in each OFDM tone such that they cancel each other at the RX chain. The weights are obtained by sending two separate (in time or frequency) pilots from PTX and ATX chains to measure the PTX to RX and ATX to RX base-band channels. These channels are denoted by H_1 and H_2, respectively. To measure H_1, transmit power of the training signal is reduced to keep the PA in linear regime. In addition, as mentioned earlier, the ATX chain is designed to have a high coupling with the RX chain. This avoids the use of a Power Amplifier (PA) for the ATX chain and consequently helps to maintain linearity in the ATX path. In spite of these provisions, it is not possible to measure H_1 and H_2 accurately, as various imperfections, including additive noise, affect the accuracy of the measurement. Let ΔH_1 and ΔH_2 respectively denote the possible error terms in the measurement of H_1 and H_2. The weighting factors applied to TX and ATX are $(H_2 + \Delta H_2)$ and $-(H_1 + \Delta H_1)$, respectively.

The remaining self-interference after analog active cancellation, referred to as residual self-interference is subsequently canceled digitally at the base-band. To this aim, the equivalent transmit to receive base-band channel (considering both TX and ATX chains) should be measured. The measurement is

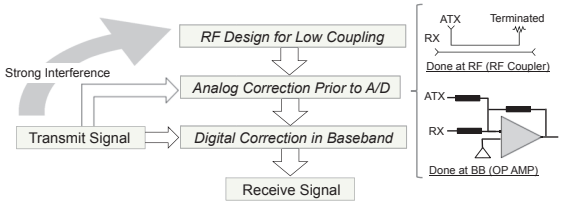

Fig. 2: Main components involved in self-interference cancellation.

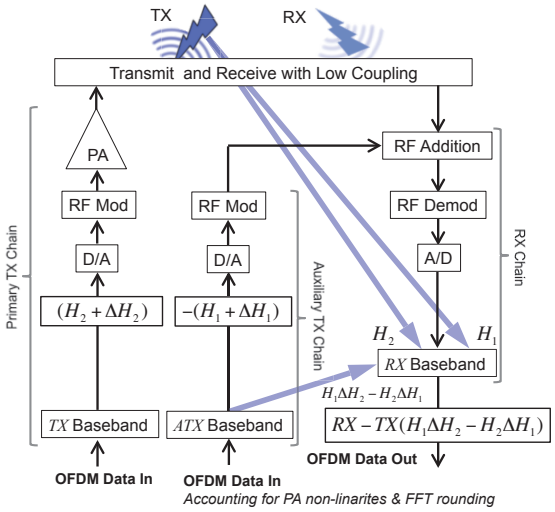

Fig. 3: Details of the self-interference cancellation.

performed by sending two weighted pilots simultaneously as TX and ATX signals using the weights computed in the earlier phase. This second training phase results in the following signal in the base-band.

$$(H_1\Delta H_2 - H_2\Delta H_1)P + (H_1\Delta H_2 - H_2\Delta H_1)\Delta P \quad (1)$$

where P and ΔP respectively denote the pilot signal and its possible error term. Assuming

$$(H_1\Delta H_2 - H_2\Delta H_1)\Delta P \approx 0, \quad (2)$$

the equivalent transmit to receive base-band channel, denoted by H_E, is consequently obtained as

$$H_E = H_1\Delta H_2 - H_2\Delta H_1. \quad (3)$$

This means (ignoring second order terms), the self-interference remaining in the base-band signal after analog active cancellation is a linear combination of PTX, ATX, which due to maintaining the linearity can be modeled as the PTX signal passed through a linear system. In other words, ignoring second order terms, errors in analog active cancellation will act as an additional parasitic linear system. As a result, the equivalent channel for the residual self-interference remains linear and can be handled relying on its OFDM structure. This linear system is measured through training and its OFDM structure is used to subtract the remaining self-interference in the digital domain (digital active cancellation).

As long as H_E is measured accurately, the residual self-interference can be successfully compensated at the base-band. Since H_E is of lower magnitude as compared to H_1 and H_2, it can be measured more precisely (in terms of relative error). The accuracy of the measurement of H_E can be also improved by repeating the measurement several times and averaging the values. In practice, in the measurement of H_E, several training signals are averaged to reduce the effect of the measurement error.

Now, let us assume the OFDM data frame Γ, including a possible error term $\Delta\Gamma$ representing computational errors due to finite precision arithmetic, is passed through TX and ATX chains. The corresponding RI_{BB} at the receiver is:

$$RI_{BB} = H_E(\Gamma + \Delta\Gamma). \quad (4)$$

The term $H_E\Gamma$ in (4) is digitally subtracted at the base-band. The term $H_E\Delta\Gamma$ is compensated digitally. Note that the above expressions symbolize the operations in order to explain the different terms and their effects. In practice, the error term due to imperfections in the D/A path, including numerical errors, is more sophisticated, and in particular depends on the method used for the actual implementation of filtering operations. In general, two factors contribute to such errors, namely "finite precision in intermediate computations", and "rounding effects to cast the result to the limited number of bits of D/A". Details of implementation are omitted due to space limitations, however, it is crucial to manage such errors as the effects any remaining residue after nonlinear compensations will be a noise with a power scaling with the power of self-interference.

It is well known that OFDM signals exhibit occasional large peak values. In most cases, analog active cancellation brings the level of the self-interference to values smaller or comparable to the incoming signal, however, this is not always the case. Let us consider situations that the self-interference after analog active cancellation is significantly higher that the level of the incoming signal. In such cases, peaks in the transmitted signal, when leaked into the receive chain, limit the dynamic range of the A/D and can cause occasional overflows. Similarly, there are occasional peaks in the incoming OFDM signal, however, it is very unlikely that both transmit and receive OFDM signals have a large peak at the same time. As a result, it can be beneficial to clip the overall receive signal prior to A/D and compensate for the clipped part in digital base-band using only the location and magnitude of the peaks of the transmit signal. Note that the clipping of an analog signal with typical bandwidths encountered in wireless transmission is relatively easy, e.g., can be implemented using a simple operational amplifier. Figure 4 shows such intentional OFDM signal clipping.

Note that Fig. 3 is just a high level abstraction aimed to capture various possibilities to realize the filtering operations. For example, for simplicity, only the ATX signal can be filtered in time domain, while filter coefficients (channels' impulse responses) are measured in the frequency domain. Another option is to apply separate filtering (in time) to PTX and ATX. Note that if the ATX chain has a direct coupling to the receive chain, the ATX will have a flat frequency response and consequently the filtering applied to PTX will not affect the magnitude of the OFDM tones aimed at distant users.

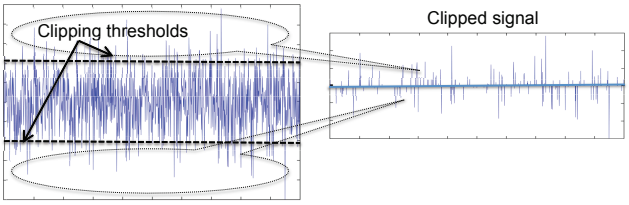

Fig. 4: Intentional clipping of analog signal prior to A/D. Location and magnitude of the clipped are subsequently compensated at digital based-band using solely the peaks of the transmitted signal.

$$\nabla \times D = -j\omega B \qquad D = \varepsilon E: \text{ Electric Field}$$
$$\nabla \times B = J + j\omega D \qquad B = \mu H: \text{ Magnetic Field}$$
$$\nabla \cdot D = \rho \qquad J: \text{ Current}$$
$$\nabla \cdot B = 0 \qquad \omega: \text{ Frequency}$$

Fig. 5: Maxwell equations.

Although above derivations depend on the particular method used for filtering, the general argument that the first order approximation of the remaining self-interference forms a linear channel, which can be measured to be used in digital cancellation, applies to a wider class of implementations. Note that unlike the first stage of analog cancellation, the energy of the remaining self-interference (to be digitally cancelled in the second stage) is comparable to that of the signal received from distant user and consequently the error due to using a first order approximation, such as shown in Eq. 2, can be ignored.

III. Antenna Design

A full-duplex node can be considered as a two-port network described in terms of scattering parameters S_{11}, S_{12}, S_{21}, and S_{22}. The objective is to reduce coupling between TX and RX chains, i.e., $S_{12} = S_{21}$ should be small. It is also desirable to have small S_{11} and S_{22} for better antenna efficiency. Moreover, the above conditions should be satisfied over the entire operating frequency range.

Note that the low coupling requirement in a full-duplex node is different from that of MIMO systems. In MIMO, it is desirable that the channels between transmit and receive antennas in distant nodes are independent. This is achieved by spacing antennas sufficiently far apart, otherwise, entries of the MIMO channel matrix will be correlated, which reduces the channel capacity. For low coupling in a full-duplex node, however, transmit and receive antennas within the same node should induce small power on each other. Unlike the case of MIMO mentioned above, this requirement does not impose any immediate restriction on antenna spacing. Using techniques proposed in this section, such antennas can be placed close to each other while having a small coupling over the desired frequency band.

Due to vicinity of transmit and receive antennas, near-field effects will be significant and dominate the system behavior. This feature is indeed beneficial. According to Maxwell equations (see Fig. 5), geometrical symmetry in structure (shape, material, boundary conditions) and excitation (feed terminals) of an antenna lead to geometrical symmetry in electric and magnetic fields. The geometrical symmetry in antenna fields can be used to cancel the self-interference. The following definitions are useful in subsequent theorems.

Definition 1: An antenna is called *self-symmetrical* if its two arms are image of each other with respect to a plane of symmetry. This includes the symmetry of construction, excitation, and parasitic elements.

We will refer to the geometrical refection of a vector in such a plane of symmetry as mirror image. Any self-symmetrical antenna can support two forms of current distribution: (1) even-symmetric current, where current distributions at any two symmetric points are mirror image of each other, and (2) odd-symmetric current, where current distributions at such symmetric points are mirror image with a sign change. The odd mode is used in practical antenna design, which results in high current flowing through antenna terminal, resulting in radiation (propagation of real/active power). Our following discussions assume the odd-symmetric current mode.

Definition 2: An antenna is called *doubly-symmetrical* if it has two planes of symmetry, one intersecting the terminals (primary plane of symmetry) and one orthogonal to it (secondary plane of symmetry).

Definition 3: Two doubly-symmetrical antennas are called *pair-wise symmetrical* if the primary plane of symmetry of one antenna overlaps with the secondary plane of symmetry of the other antenna, and vice versa.

Poynting's vector is defined as

$$\vec{P} = \vec{E} \times \vec{H}^*, \tag{5}$$

where $\vec{E}$ and $\vec{H}$ denote electric and magnetic fields, respectively, and $*$ is the complex conjugate. According to Poynting's theorem, the power flowing out of a surface is equal to the integration of Poynting's vector over that surface. Real component of the Poynting's vector, corresponding to the flow of real/active power, will be zero if the phase difference between electric and magnetic fields is equal to $\pi/2$. Real component of the Poynting's vector does not change with time, and its imaginary component has double the frequency of the sinusoidal excitation. If the real component is zero in one frequency, it will be zero for all frequencies. It follows (see Fig. 6):

Theorem 1: For a doubly-symmetrical antenna (with odd-symmetric current), we have (i) under reflection in the primary plane of symmetry: H-field is mirrored, while E-field is mirrored with a sign change, and (ii) under reflection in the secondary plane of symmetry: E-field is mirrored, while H-field is mirrored with a sign change (see Fig. 6).

Proof: Proof follows noting linearity and geometrical symmetry of Maxwell equations, together with the Ampere's law concerning the direction of magnetic field. ■

Corollary 1: In a doubly-symmetrical antenna with (with odd-symmetric current), Poynting's vector will be mirrored under reflection in the primary and/or secondary planes of symmetry (see Fig. 6).

Theorem 2: If the transmit and receive antennas of a full-duplex node are pair-wise symmetrical, then they have zero coupling, i.e., $S_{12} = S_{21} = 0$ independent of frequency.

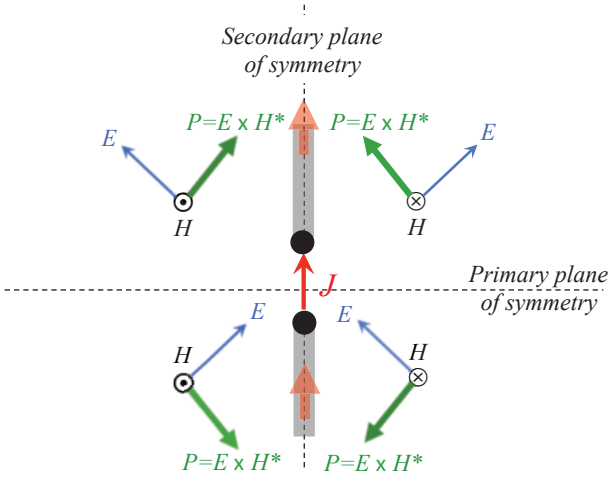

Fig. 6: A doubly symmetric antennas with input current in the y-axis direction.

Proof: First of all, due to reciprocity, in the proof of this theorem the roles of the two antennas can be exchanged. The setup is shown in Fig. 7 where the terminals of the receive antenna are connected with a lumped load element capable of receiving energy, such as a matched resistor. Note that in terms of the geomtery of the load, only the length (connecting the two terminals) has to be non-vanishing (lumped element), and the other two dimensions can be made arbitrary small.

A simple proof is based on a direct integration of the E-field along the shortest path connecting terminals of the receive antenna (see Fig. 7), and noting (see theorem 1) that E-field will be mirrored under reflection in the secondary plane of symmetry. This results in zero voltage between terminals of the receive antenna. A similar argument applies to the current through the lumped load element, which, due to even symmetry, has to be zero to satisfy the condition for the continuity of current through the lumped element (Kirchhoff's Current Law, KCL).

To have a more rigorous proof, referring to Figs. 7, let us consider a symmetric region around the receive antenna which does not include any part of the transmit antenna. The intension is to show that the integration of the real component of the Poyting's vector over this surface is zero. For this purpose, it is enough to show that integration of the real component of the Poyting's vector over the small cylindrical region surrounding the lumped element is zero (see Fig. 7). This region can be divided into symmetrical pairs of surfaces, e.g., 1 and $1'$, 2 and $2'$. Noting the symmetries in Poynting's vector given in Corollary 1, it follows that the integration of the real component over the top and bottom surfaces will be mirror image. On the other hand, dimensions of the cylinder, except for its length, can be made arbitrarily small. This means the top and bottom surfaces can be brought arbitrarily close to each other. In this case, imposing the conditions for the continuity of the Poyting's vector, the two mirror image vectors should be equal to zero, which means the net real power flowing in/out the whole region is zero. This in turn implies that RX antenna does not absorb any energy from TX antenna, i.e., $S_{12} = S_{21} = 0$.

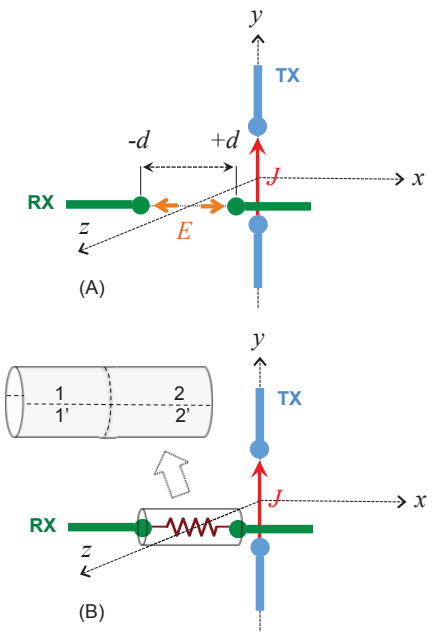

Fig. 7: Pairwise symmetrical antennas with input current (I) in the y-axis direction. (A): Integration of E-field between terminals al of the RX antenna is zero which results in $S_{12} = S_{21} = 0$. (B): Net energy flowing out the region around the RX antenna terminals is zero which results in $S_{12} = S_{21} = 0$.

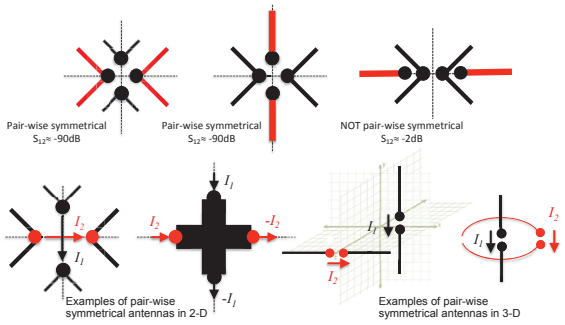

Fig. 8: All structures are pair-wise symmetric except for the one on top right corner which shows significant coupling.

Note that the simple proofs mentioned in terms of zero voltage and zero current across receive terminals, which are, respectively, reflections of the Kirchhoffs' Voltage and Current Laws for the lumped load element, are indeed analogous to the condition for the continuity of the Poyting's vector in the detailed proof.

■

Fig. 8 illustrates some examples of pair-wise symmetric antennas. Figure 8 also shows that the coupling between antennas can be very strong ($S_{12} = -2$ dB) due to the near-field effect unless it is canceled relying on pair-wise symmetry. The values of coupling are obtained by using high frequency structural simulator (HFSS) at 2.4 GHz band.

Low coupling in RF is crucial, and it directly helps in brining down any remaining non-linear residues to a level below the AWGN floor. Otherwise, as the power of such remaining residues scales with the power of self-interference, they may raise the effective noise floor to an unacceptable

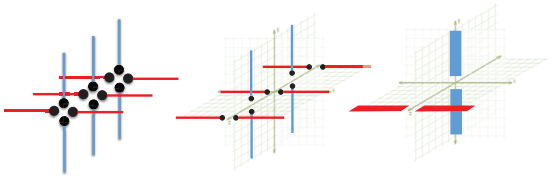

Fig. 9: Three-dimensional antenna structures with zero coupling for MIMO structure.

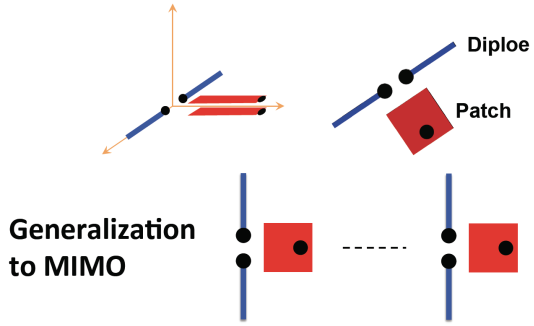

Fig. 10: Realizing pair-wise symmetrical antenna structures in 2.5 dimensions.

level. This feature is one of the main reasons for the superior performance of the methods discussed here as compared to other implementations [3]- [11]. In practice $S_{12} = S_{21}$ is close to zero due to reflections from surrounding environment, and sources of imperfections in hardware realization. As shown in [1], small movements of the antenna structure or surrounding environment can cause large fluctuations in the coupling. In practice, we have observed that the antenna design rules provided here guarantee a "worst case" coupling of -40 to -50 dB.

The idea of symmetry is generalized to obtain triple-wise symmetric antennas in three dimensions in Fig. 9(A). Note that pair-wise symmetric structures require two dimensions effectively, and as a result, it is possible to generate more transmit/receive antennas along the third dimension. As a result, MIMO structures with zero coupling can be realized in three dimensions. Fig. 9(A) shows an example where every antenna in TX set (say antennas parallel with the x-axis) is decoupled from all the antennas in RX set (antennas parallel with the y-axis).

Three-dimensional antennas with pair-wise symmetry can be implemented by using opposite sides or different layers of a printed circuit board (PCB). Such constructions are referred as 2.5 dimensional in RF literature. Figure 10 shows an example for such a construction where a patch and a dipole are implemented on a multi-layer PCB. Note that in practice, the second arm of the patch structure in obtained by reflection in the ground plane, and this contributes to violating the perfect symmetry. Figure 10 also includes a generalization to a MIMO structure in 2.5 dimensions.

Although pair-wise symmetrical structures can be realized in three dimensions, in some applications, it will be of interest to have low coupling antenna structures in two dimensions. Following theorem provides the basis to realize such antenna structures with low (but theoretically non-zero) coupling.

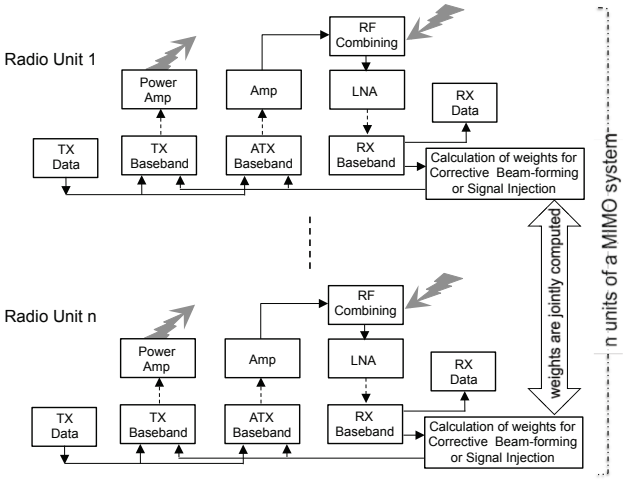

Fig. 11: Cancellation of self-interference in MIMO setups is based on representing the coupling between transmit to receive chains as a matrix and forming the corrective signals by relying on linear combinations of different transmit signals.

Theorem 3: In a self-symmetrical antenna (with odd-symmetric current), electric at the location of the primary plane of symmetry will be orthogonal to this plane.

Proof: Let us consider the two half spaces on the two sides of the primary plane of symmetry. Noting the symmetry of current with respect to the primary plane of symmetry, the role of one of this half spaces can be replaced by a conducting plane along the primary plane of symmetry. It follows that E-field will be orthogonal to such a conducting plane. ■

Note that such an orthogonal E-field will integrate to zero if integrated over any line inside the primary plane of symmetry. This means primary plane of symmetry will be an equipotential surface. i.e., points on this surface will be of equal electrical potential. This theorem motivates us to place a second antenna (set of antennas) along such an equpooteial surface. However, such a placement will cause disturbance to the conditions of theorem 3. As a result, this configuration results in low, but theoretically non-zero, coupling. To further minimize the disturbance caused by the placement of the second set, these can be symmetrically placed. Figure 12 shows some numerical results using HFSS for such a construction. It is observed that very low coupling of about -90 to -100 can be achieved.

In all our designs, we have aimed to maintain the symmetry in the entire circuit structure. To study the effect of symmetry, in Fig. 13 two parasitic objects are places close to the antenna structure given in Fig. 12, and as a result the coupling has dropped from about -100dB to about -30dB. However, as shown in Fig. 14, by placing similar parasitic objects on the opposite side to make the system symmetrical, the coupling has reduced to about -80dB. This observation is used as a guideline in designing the PCBs, in the sense that the footprint and placement of components has been designed to maintain the symmetry as much as possible. Fig. 15 shows such an arrangement for a 4×4 MIMO.

Shape of arms and spacing between antennas can be also adjusted to compensate for lack of perfect symmetry, and for

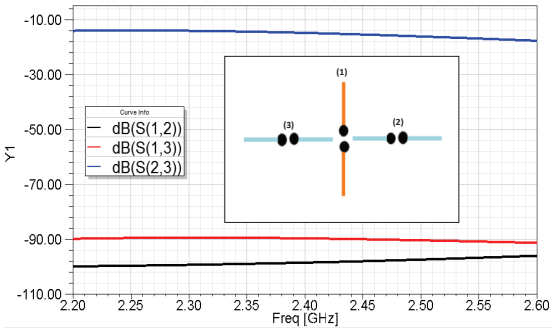

Fig. 12: HFSS results for antenna structures with low, but non-zero coupling obtained by placing one set of antennas (horizontal dipoles) in the plane of symmetry of other set (vertical dipole). Note that the difference between the curves showing S_{12} (black curve) and S_{13} (red curve) is due to numerical inaccuracies in HFSS.

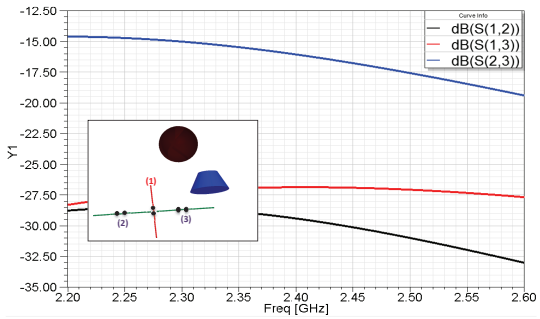

Fig. 13: Effect of parasitic objects on the coupling.

non-zero width of antennas' arms. Fig. 16 shows that the coupling between two antennas can be improved from -70dB to -110dB by modifying shape of arms. Results are obtained at 2.4 GHz band using HFSS.

Figure 17 shows a MIMO configuration in which antenna arms are merged into a single arm above the ground plane, resulting in small (but non-zero) coupling. ATX coupling is achieved using a third terminal for the same arm (with a high coupling to the RX terminal). Figure 18 shows a different

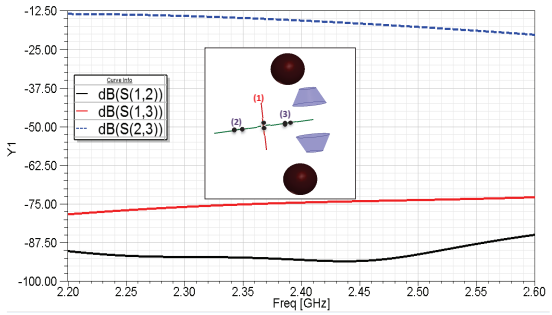

Fig. 14: Effect of adding more parasitic objects to compensate for the lack of symmetry and improve the coupling.

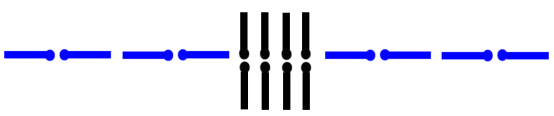

Fig. 15: 4×4 MIMO full-duplex node with low coupling implemented in two dimensions.

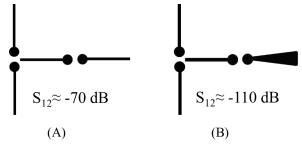

Fig. 16: Shape of arms is adjusted to compensate the lack of symmetry in a low coupling structure.

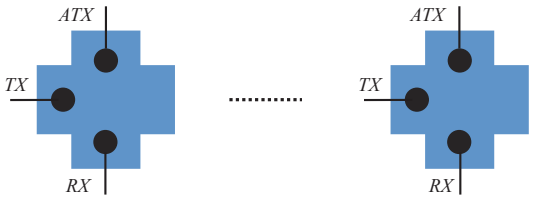

Fig. 17: Pair-wise symmetry in 2.5 D with generalization to MIMO.

configuration of TX, ATX, and RX antennas in 2.5 dimensions, where ATX is a fully functional transmit antenna and has zero coupling with RX antenna.

IV. Power Amplifier (PA) Modeling and Compensation

In previous sections, the arguments were based on linearity of transmit path. This path is composed of D/A and PA which can contradict this linearity condition. D/A is theoretically linear, and our investigations show that any deviations from this assumption, as it will be the same in PTX and ATX signals, will not have a noticeable effect. On the other hand, the PA can be highly nonlinear, which will make the ATX signal to be different from the self-interference (note that measurements, and compensation techniques explained earlier are based on linearity assumption). For the ATX signal, as the coupling to the receive path is intentionally set to be adequately high, there is no need for a PA, and consequently, the ATX path can be kept to act in linear region. As a result, it is important to include the effect of the PA nonlinearity in the construction of the ATX signal. This is achieved in the baseband of the ATX chain, in time domain, by passing the constructed linear signal through a nonlinear curve (for both magnitude and phase), which follows the PA nonlinear model. This is different from the pre-compensation methods used in PA design, which aim to pre-adjust the baseband signal going through the nonlinear PA. Note that in pre-compensation techniques, the primary path will be composed on a nonlinear system (pre-compensation) concatenated with the linear system corresponding to RF modulation, then concatenated

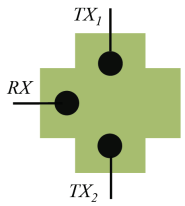

Fig. 18: Another configuration for pair-wise symmetry in 2.5 D.

% of PA dynamic range	SNR loss with compensation	SNR loss without compensation
95%	0	10dB
90%	0	9.5dB
85%	0	4.5dB

TABLE I: Degradation in effective SNR due to nonlinear PA, measured as the Mean Square Error (MSE) between the ideal (linear signal) signal and the signal received with nonlinear PA. Ratio of the total signal power (linear and nonlinear part) to AWGN noise power is about 40dB.

with the nonlinear system corresponding to the PA, and finally concatenated with the linear system from the output of PA to the baseband of the receive chain. As, due to nonlinearity, the order of these four systems cannot be freely changed, it will not be possible to measure each them and compensate their combined effect through pre-compensation. Note that in full-duplex applications, any error in the compensation will scale with the power of the transmit signal and can have a much more damaging effect as compared to the case of ordinary PA design in legacy one-way systems. To study this effect, we present the results in conjunction with a Long Training Sequence (LTS) of 802.11. This is a 64 tone OFDM signal in which the 6 middle tones are zero, and the rest of the tones have equal power. The PA is SZA-2044 by RF Micro Devices which, at 5 volts, has: POUT=22dBm at 3% EVM, and P1dB=29.5dBm. The phase and magnitude of the PA, obtained through using a sinusoidal signal corresponding to tone number 4 of the LTS (1.25MHz), are shown in Figs. 19, and 20, respectively. To show the effect of nonlinearity, the PA is driven close to its full dynamic range, and is connected to the receiver RF front-end through a cable with adjustable attenuation. Figure 21 provides comparisons between the values with and without compensation. Table I contains relative improvement in effective noise level (combination of thermal and nonlinear noise) due to the proposed compensation (for different coverage of PA dynamic range). To observe the relative changes in the SNR loss compared to the AWGN noise floor in Table I, the attenuation between transmitter and receiver is adjusted to reduce the level of the signal. Otherwise, the SNR loss due to nonlinearity, which scales with the input power, would completely dominate the total noise. This has caused the received signal power to be about 40dB above the noise level. In practice, the level of self-interference will be significantly higher. For example, assuming 20dBm transmit power, RF isolation of 50dB, and noise floor of -100dBm, the power of self-interference will be 70dB above noise floor, which means the SNR loss will be 30dB higher as compared to the values in the third column of Table I. In such a setup, even if the PA is driven less into nonlinearity, without the methods described here, the degradation in effective SNR would be unacceptable.

V. Conclusion

This paper proposed new self-interference cancellation techniques for practical implementation of full-duplex (two-way) wireless networks. Various methods in both analog and digital domains were presented to mitigate the self-interference

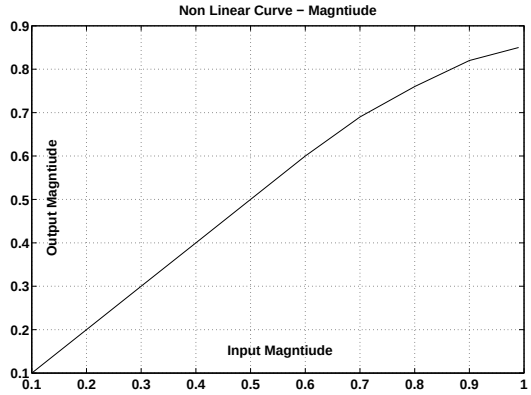

Fig. 19: Nonlinear curve corresponding to the input/output magnitude of the PA.

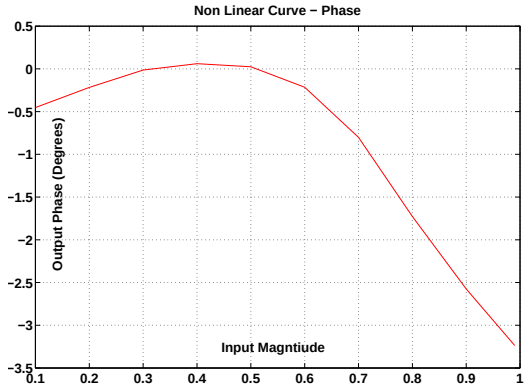

Fig. 20: Nonlinear curve corresponding to output phase change vs. input magnitude of the PA.

.

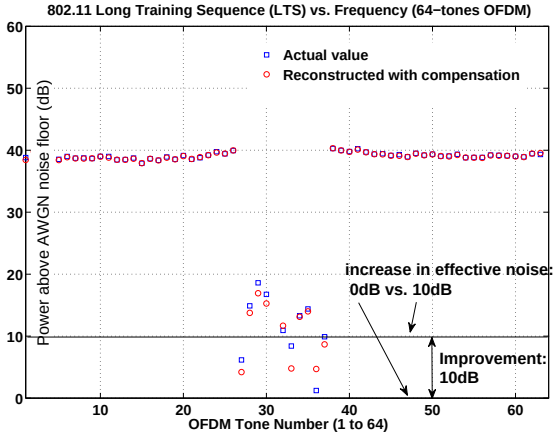

Fig. 21: Power of 802.11 LTS (6 tones in the middle of the 64 tones are zero and the other ones have equal power) vs. tone number with and without compensation (90% of the PA dynamic range is covered). Noise level (used as reference): 0dB, Mean Square Error (MSE) with compensation : +1.9 dB, MSE without compensation: +13.4dB.

effects. In a first aspect, antenna design was considered at a full-duplex node to reduce the coupling between transmit and receive chains. In a second aspect, an auxiliary transmit signal was generated and combined with the receive signal in analog domain to cancel the self-interference. The auxiliary signal was constructed by using the primary transmit signal and instantaneous measurement of the equivalent transmit to receive base-band channel. Digital cancellation techniques were also used at the receive base-band to eliminate the residual self-interference.

Acknowledgements: This work would not be possible unless through generous support provided by the Ontario Ministry of Research and Innovation, through ORF-RE. An equipment grant from Canada Foundation for Innovation (CFI), matched with Ontario Ministry of Research and Innovation (ORF-RI) was critical in realizing hardware implementation. HFSS simulations reported in Figs. 7 and 9 were performed by Dr. A. Attia.

REFERENCES

[1] A. K. Khandani, on line (video) presentation of two-way wireless posted April 2012. www.cst.uwaterloo.ca

[2] A. K. Khandani, "Methods for spatial multiplexing of wireless two-way channels," US Patent number: 7817641, Filing date: Oct 17, 2006, Issue date: Oct 19, 2010, Application number: 11/581,427

[3] K. Tsubouchi, H. Nakase, A. Namba, K. Masu, Full duplex transmission operation of a 2.45-GHz asynchronous spread spectrum using a SAW convolver, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Sept. 1993 (Res. Inst. of Electr. Commun., Tohoku Univ., Japan)

[4] S. Chen, M. A. Beach, J.P. McGeehan, Division-free duplex for wireless applications, Electronics Letters, Jan. 1998, (Centre for Commun. Res., Bristol Univ.)

[5] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, G. Dejean, Rethinking Indoor Wireless: Low Power, Low Frequency, Full-duplex, "Microsoft Technical Report MSR-TR-2009-148", July 2009

[6] M. Duarte and A. Sabharwal, Full-Duplex Wireless Communications Using Off-The-Shelf Radios: Feasibility and First Results, Asilomar Conference on Signals, Systems, and Computers, Nov. 2010

[7] A. Sahai, G. Patel and A. Sabharwal, "Pushing the Limits of Full-duplex: Design and Real-time Implementation", Rice tech report, Feb. 2011, see also announcements by Rice University (Sept. 2011): http://www.youtube.com/watch?v=tXMwn2mm0VY

[8] M. Duarte, C. Dick and A. Sabharwal, "Experimental-driven Characterization of, Full-Duplex Wireless Systems", submitted to IEEE Transactions on Wireless Communications, June 2011

[9] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Exploiting Directional Diversity in Full-duplex Communication, Asilomar Conference, Nov. 2011

[10] J. Choiy, M. Jainy, K. Srinivasany, P. Levis, S. Katti, "Achieving Single Channel, Full Duplex Wireless Communication," Mobicom 2010, Sept. 2010. See also announcement by Stanford University (Feb. 2011): http://www.youtube.com/watch?v=RiQb5NdDWgk

[11] M. Jain, J. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, P. Sinha, "Practical, Real-time, Full-Duplex Wireless", Mobicom, Sept. 2011