A Method for Broadband Full-Duplex MIMO Radio

Yingbo Hua, Fellow, IEEE, Ping Liang, Yiming Ma, Ali Cagatay Cirik, and Qian Gao

Abstract—We present a time-domain transmit beamforming (TDTB) method for self-interference cancelation (SIC) at the radio frequency (RF) frontend of the receivers on broadband full-duplex MIMO radios. It is shown that the conventional frequency-domain transmit beamforming (FDTB) method along with the orthogonal frequency division multiplexing (OFDM) framework does not generally perform SIC in the prefix region of a transmitted frame. A hardware based test of the TDTB method shows a 50 dB SIC over a bandwidth of 30 MHz.

Index Terms—Full-duplex radar, full-duplex wireless communication, single channel full-duplex.

I. INTRODUCTION

A full-duplex radio (or radio station) is defined as a radio frequency (RF) transceiver that can transmit and receive signals at the same time and same frequency. All currently deployed radios for wireless communications are half-duplex which transmit and receive signals in two separate/orthogonal channels. A full-duplex radio can have twice as high spectral efficiency as a half-duplex radio. Possible applications of full-duplex radios include wireless base stations, wireless relays and personal-area wireless devices.

A fundamental issue for a full-duplex radio is known as self-interference cancelation (SIC). When a full-duplex radio transmits, it causes a strong self-interference to its receiver. There are three basic approaches to reduce the self-interference. The first is to apply RF attenuation between the transmit (Tx) antennas and the receive (Rx) antennas on the same radio unit. This can be done by increasing the distances between the Tx and Rx antennas, using RF absorbing materials, and/or designing and utilizing antenna null angles. For many practical settings, this approach can achieve 40–60 dB reduction of self-interference. But at the same time, the transmitted signal from a radio can be up to 100 dB stronger than a desired signal from a remote node. This approach alone typically still leaves a too large amount of self-interference. The second approach is baseband-only SIC at the receiver side, as in [1], which cancels the self-interference after the self-interference has been received and demodulated into baseband. When the self-interference is much stronger than the desired signal, this approach suffers from saturation of the analog-to-digital converter (ADC) of the receiver, i.e., the desired signal is corrupted by an additional quantization noise power which can be shown to be proportional to the power ratio of the interference over the desired signal.

The third approach, a focus of this paper, is to use transmit beamforming for SIC at the RF frontend of receivers. Together with the first two approaches, the third appears to hold the promise to make full-duplex radio a reality. The recent works on transmit beamforming for full-duplex radio are all formulated in the frequency-domain, e.g., see [2]–[6] and the references therein. We call these methods the frequency-domain transmit beamforming (FDTB). The experimental works shown in [7]–[12] either applied a FDTB method or assumed frequency-flat channels.

We will show that a FDTB method, when (typically) implemented with OFDM for broadband applications, does not generally remove the self-interference on a frame received from a remote node in a time region corresponding to the prefix region of a frame transmitted from the local node. To ensure the prefix region of the transmitted frame not to coincide with any useful part of the received frame is difficult for most full-duplex operations. The paper [12] did not recognize this problem although they reported difficulties with reception after transmission starts.

In this paper, we present a time-domain transmit beamforming (TDTB) method for broadband SIC at the RF frontend. The TDTB waveforms can be directly implemented at the circuit level without costly transformations between time and frequency domains at a sampling rate much higher than the baud rate applied for OFDM. A full-duplex radio based on TDTB removes self-interference at all time, which allows asynchronous full-duplex. We will also present a hardware based experimental result of TDTB for SIC.

II. CONDITION FOR SIC AT RF FRONTEND

We consider a generic MIMO radio unit equipped with \( n_r \) RF receivers and \( n_t \) RF signal generators/transmitters. Among all generators, there are \( n_a = n_r - n_r \) primary generators and \( n_r \) auxiliary generators. The primary generators are used to transmit up to \( n_a \) independent streams of data to a remote node. The auxiliary generators are used to generate RF waveforms for SIC at the RF frontend of the receivers on the same radio.

The primary generators can be directly connected to \( n_r \) primary transmit antennas, the auxiliary generators to \( n_r \) auxiliary transmit antennas, and the \( n_r \) receivers to \( n_r \) receive antennas. While the spacing between the primary transmit antennas and that between the receive antennas should be large enough for a high diversity, the spacing between an auxiliary transmit antenna and a receive antenna can be small for large coupling. For the strongest coupling, the auxiliary generators can be directly connected pairwise to the receivers at the RF frontend using RF power combiners. Furthermore, if a primary transmit antenna and a receive antenna have overlapping beam patterns in the directions of interest, they can be merged into one via a RF circulator. See Fig. 1 later. The following discussion is valid for any of the above situations.
We index the auxiliary generators by \( k = 1, \ldots, n_a \) and the primary generators by \( k = n_a + 1, \ldots, n_t \). Then, for each transmitted data packet subject to linear modulation, a RF signal stream transmitted from the \( k \)th generator ideally has the form

\[
x_k(t) = \sum_{i=1}^{L} g_k^{(i)}(t) s_i(t) p(t - nT)
\]

is the complex baseband form (also called I/Q waveform) of \( x_k(t) \). Here, \( g_k^{(i)}(t) \) is the complex impulse response of the \( k \)th transmit beamforming waveform for data stream \( i \) (of total \( I \) streams), \( s_i(t) \) is the complex symbol sequence for data stream \( i \), \( N + L \) is the number of complex symbols per stream (including the \( L \) prefixed symbols as used in OFDM system), and \( p(t) \) is the fundamental pulse waveform used for linear pulse modulation, which has the double-sided bandwidth \( W \) and the effective duration \( T \). For high spectral efficiency, it is typical that \( T \) is equal to or only slightly larger than \( 1/W \). The operator \( * \) denotes convolution.

The RF self-interference received by the \( l \)th receiver is

\[
\hat{u}_l(t) = R \{ e^{j2\pi f_c t} \} \text{ where } f_c \text{ is the carrier frequency and}
\]

\[
\hat{u}_l(t) = \sum_{i=1}^{I} g_l^{(i)}(t) s_i(t) p(t - nT)
\]

is the complex baseband form (also called I/Q waveform) of \( \hat{u}_l(t) \). Here, \( g_l^{(i)}(t) \) is the complex impulse response of the \( l \)th receive beamforming waveform, \( s_i(t) \) is the complex symbol sequence for the \( l \)th stream, \( p(t) \) is the fundamental pulse waveform used for linear pulse modulation, which has the double-sided bandwidth \( W \) and the effective duration \( T \). For high spectral efficiency, it is typical that \( T \) is equal to or only slightly larger than \( 1/W \). The operator \( * \) denotes convolution.

To cancel the RF self-interference \( \hat{u}_l(t) \) for all \( l \) and \( i \), it is sufficient to find \( g_l^{(i)}(t) \) for all \( k \) and \( i \) such that \( u_k(t) = 0 \) for all \( i \) or equivalently \( \sum_{k=1}^{N} h_{k, k}^{(i)}(t) g_k^{(i)}(t) = 0 \) for all \( i \) and \( l \). The matrix form of this condition is

\[
\begin{bmatrix}
h_{1,1}(t) & \cdots & h_{1,n_a}(t) \\
\vdots & \ddots & \vdots \\
h_{n_a,1}(t) & \cdots & h_{n_a,n_a}(t)
\end{bmatrix}
\begin{bmatrix}
g_1^{(i)}(t) \\
\vdots \\
g_{n_a}^{(i)}(t)
\end{bmatrix}
= 0
\]

(2)

or equivalently

\[
H(t) \ast \mathbf{g}^{(i)}(t) = 0
\]

(3)

in a more compact form. Although given in baseband, (3) ensures SIC even at the RF frontend. Also note that when all elements in a row of \( H(t) \) are corrupted by a common scalar due to receiver phase noise, the solution \( \mathbf{g}^{(i)}(t) \) to (3) is not affected.

III. SOLUTION OF BROADBAND TRANSMIT BEAMFORMING WAVEFORMS

A. A General Case

To find the solutions to the (2), we need to apply a known notion of vector space in the field of functions of time. We say that \( n \) vectors of functions of \( t: f_1(t), \ldots, f_n(t) \), are convolutively independent if \( \sum_{i=1}^{n} a_i(t) \ast f_i(t) = 0 \) implies \( a_1(t) = \cdots = a_n(t) = 0 \). The rank \( r_{H(t)} \) of the matrix \( H(t) \) is the largest number of columns (or rows) in \( H(t) \) that are convolutively independent. It follows that \( r_{H(t)} \leq \min\{n_r, n_a\} = n_r \). The dimension of the solution space of (2), which is also called the dimension of the (right) null space of \( H(t) \), is the number of convolutively independent solutions to (2), which is \( d_{n_a} = n_r - r_{H(t)} \geq n_a \). If \( d_{n_a} = n_r \), we call it a typical case (very likely in practice), or otherwise if \( d_{n_a} > n_r \), atypical case (not very likely in practice). The number \( I \) of the data streams in (1) must be no larger than \( d_{n_a} \).

In general, for \( n_a > 1 \) and \( n_r > 1 \), the \( n_r \)th in a set of \( n_a \) convolutively independent solutions to (3) can be written as

\[
\mathbf{g}^{(i)}(t) = -\text{adj} \{ \mathbf{A}(t) \} \ast \mathbf{b}_i(t)
\]

(4)

or equivalently

\[
\mathbf{g}^{(i)}(t) = -\text{det} \{ \mathbf{A}(t) \} \mathbf{b}_i(t)
\]

(5)

and \( \mathbf{g}_0^{(i)}(t) = \text{det} \{ \mathbf{A}(t) \} \). Both the adjoint \( \text{adj} \{ \mathbf{A}(t) \} \) and the determinant \( \text{det} \{ \mathbf{A}(t) \} \) can be obtained analytically in the same way as those of a matrix of numbers as shown in [14] except that all multiplications should be substituted by convolutions. It is important to note that the expression (5) does not involve any division but only convolutions and sums. This is a unique feature and very useful for hardware implementation. If \( n_r = 1 \) and \( n_a = 1 \), a solution for \( \mathbf{g}^{(i)}(t) \) is \( \mathbf{g}^{(i)}(t) = [-h_{1,2}(t), h_{1,1}(t)]^T \).

The solution shown in (4) is valid for arbitrary \( H(t) \) as long as \( \text{det} \{ \mathbf{A}(t) \} \neq 0 \). This condition can be met if \( h_{k,k}(t) \) for \( k = 1, \ldots, n_a \) have the largest norms among \( h_{k,k}(t) \) for all \( i \) and \( k \). To ensure that, we can either place the \( n_a \) auxiliary transmitting antennas close enough to the \( n_r \) receiving antennas or directly couple the \( n_a \) auxiliary generators to the \( n_r \) receivers at the RF frontend via RF power combiners. The property of RF power combiners can be found in [13].

B. Coupling via RF Power Combiners

If we connect (with proper connector shielding) the \( n_a \) auxiliary generators to the \( n_r \) receivers pairwise via RF power combiners, then only \( n_a \) transmit antennas are needed, and \( h_{k,k}(t) = 0 \) for \( k \neq i \) and \( k = 1, \ldots, n_a \). And hence

\[
H(t) = \begin{bmatrix}
h_{1,1}(t) & \cdots & h_{1,n_a}(t) \\
\vdots & \ddots & \vdots \\
h_{n_a,1}(t) & \cdots & h_{n_a,n_a}(t)
\end{bmatrix}
\]

and a set of \( n_a \) convolutively independent solutions of (2) are the columns of (see the equation at the bottom of the next page), where

\[
\psi(t) = \prod_{i=1}^{n_a} h_{i,i}(t) \ast \delta(t) \text{ and } \psi_i(t) = \prod_{i=1, i \neq j}^{n_a} h_{i,i}(t) \ast \delta(t).
\]
C. Issues of the Frequency-Domain Solution

Applying the Fourier transform to (3) yields

$$\hat{\mathbf{H}}(f) \hat{\mathbf{g}}(f) = 0 \quad (6)$$

where $\hat{\mathbf{H}}(f)$ and $\hat{\mathbf{g}}(f)$ are Fourier transforms of $\mathbf{H}(t)$ and $\mathbf{g}(t)$ respectively. The solutions $\hat{\mathbf{g}}(f)$ to (6) at a given frequency $f$ are the conventional form of the transmit beamforming vectors. For any fixed $f$, the basis vectors of the null space of $\hat{\mathbf{H}}(f)$ can be, and are commonly, computed numerically. The Fourier transform of $\hat{\mathbf{g}}(t)$ shown in (4) is also a valid solution to (6). But they are generally not orthogonal with each other. Furthermore, the condition $\det\{\mathbf{A}(t)\} \neq 0$ (i.e., not a zero function of $t$) does not necessarily imply $\det\{\hat{\mathbf{A}}(f)\} \neq 0$ for all $f$. Clearly, from the aspects of computation and implementation, the frequency-domain transform beamforming (FDTB) and the time-domain transmit beamforming (TDTB) are quite different.

All the prior works shown in [2]–[12] use the frequency-domain solutions except [10] where allpass channels are assumed. The frequency-domain solutions are naturally suited for narrowband applications for which the channel frequency responses must be completely flat within the bandwidth of interest around a given $f$. For broadband applications of the frequency-domain solution, an OFDM-based system supporting multiple subcarriers provides a natural platform. In [12], the authors applied the frequency-domain solution (assuming $n_s = n_\times - 1$) to an OFDM system with 64 subcarriers spanning 10 MHz bandwidth. However, as shown next, the frequency-domain solution with OFDM does not in general remove the self-interference at the RF frontend in the prefix region of a transmitted packet.

Consider the case of $n_s = n_\times - 1$ and the use of RF power combiner for cancelation. Denote the baud-rate channel response of the primary interference channel by $h_2[n] = h_{1,2}(nT)$ for $n = 0, 1, \ldots, L$ and the baud-rate channel response of the interference cancelation channel by $h_1[n] = h_{1,1}(nT) = \delta[n]$ where $T$ is the baud interval. (The baud rate is smaller than the sampling rate mentioned later.) For simplicity, we assume that the channel between the auxiliary generator and the receiver is completely flat, i.e., $h_2[n] = \delta[n]$. We will show below that the power of the self-interference in the prefix region of the transmitted frame is

$$P_f[n] = P_x \sum_{i = -n_\times + 1}^{L} |h_2[i]|^2; \quad n = 0, 1 \ldots, L - 1 \quad (7)$$

where $P_x$ is the power of the signal transmitted by the primary transmitter, and $L$ is the length of the prefix as determined by the OFDM system standard (which should be the upper bound on the maximum delay spread of all channels between radios in a given environment). We have assumed here that there is a gap of at least $L$ baud intervals between two adjacent OFDM frames transmitted by the primary transmitter. Otherwise, the self-interference in the prefix region is even stronger due to the contribution from the previous frame. We see that $P_f[n] = 0$ if and only if $h_2[i] = 0$ for $i \geq 1$ (i.e., if and only if the channel frequency response between the primary transmitter and the receiver is also completely flat over all subcarriers). Such a condition is difficult to hold to a precision required. The only coefficient of the channel impulse response that contributes zero interference is $h_2[0]$, which may or may not be the most dominant (depending on the channel environment between the primary transmit antenna and the receive antenna).

The result (7) should also be useful for many (beyond full-duplex radios) broadband interference cancelation applications using the OFDM-based FDTB method.

Proof of (7): Recall that the signal flow in an OFDM communication system has the following sequence of components: original symbols, IFFT, cyclic prefixing, RF modulation, transmission, RF demodulation, sampling, prefix discarding, and FFT. Let $\tilde{v}_0, \ldots, \tilde{v}_{N-1}$ be zero-mean and uncorrelated symbols to be transmitted from the radio in an OFDM frame. Its $N$-point IFFT is $v_0, \ldots, v_{N-1}$, which are also zero-mean and uncorrelated. After $L$-point cyclic prefixing, it becomes $s_0, \ldots, s_{N+L-1}$. This sequence is modulated into a baseband waveform, then modulated into an RF waveform by an RF carrier and then transmitted by the primary transmitter. At the local receiver, after RF demodulation and then baseband sampling (and before the stage of prefix-discarding and FFT), the interfering samples received by the local receiver in the prefix region of the transmitted frame are $y_1[n] = \sum_{i = 0}^{N} h_2[n - i] s_i = \sum_{i = 0}^{n} h_2[n - i v_{N-L+i}], 0 \leq n \leq N - 1$. For SIC, the auxiliary generator needs to transmit (concurrently with the primary generator) the following $N$ pre-IFFT symbols: $-\tilde{h}_2[0] \tilde{v}_0, \ldots, -\tilde{h}_2[N-1] \tilde{v}_{N-1}$ where $\tilde{h}_2[k] = FFT_N(h_2[2^k])$ as in [12]. Applying IFFT to these $N$ symbols yields $\tilde{g}_2[n] = -\sum_{i = 0}^{L} h_2[i] v_{(n-i)L}, 0 \leq n \leq N - 1$, where $(n-i)L$ denotes $n-i$ in modulo-$N$. Then, after cyclic-prefixing the sequence $\tilde{g}_2[n]$, the first $L$ samples to be transmitted by the auxiliary generator are $y_2[n] = -\sum_{i=0}^{L} h_2[i] v_{(n-i)L+N} = -\sum_{m=-L}^{n-L} h_2[n-m] v_{(N-L+m)n}$, where $0 \leq n \leq L - 1$. Since the channel from the auxiliary transmitter to the local receiver is assumed to be allpass, then the “net” self-interference at the local receiver in the prefix region is $y[n] = y_1[n] + y_2[n] = -\sum_{m=-L}^{n-L} h_2[n-m] v_{(N-L+m)n} + \sum_{i=0}^{L} h_2[i] v_{(n-i)L+N}$, with $0 \leq n \leq L - 1$. With the definition $P_f[n] = E\{y[n]^2\}$ and $P_x = E\{|v_0|^2\}$, (7) is hence proved.

$$\mathbf{G}(t) = [\mathbf{g}^{(1)}(t), \ldots, \mathbf{g}^{(n_x-n_\times)}(t)] = \begin{bmatrix} -v_1(t) * h_{1,2,n_\times + 1}(t) & \cdots & -v_1(t) * h_{1,2,n_\times}(t) \\ \vdots & \ddots & \vdots \\ -v_{n_\times-n_\times}(t) * h_{1,2,n_\times}(t) & \cdots & -v_{n_\times-n_\times}(t) * h_{1,2,n_\times}(t) \\ v(t) & \cdots & v(t) \end{bmatrix}$$
TABLE I
CANCELLATION OF A 30 MHZ SELF-INTERFERENCE USING TDTB

<table>
<thead>
<tr>
<th></th>
<th>INR before SIC</th>
<th>INR after SIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 antennas</td>
<td>76dB</td>
<td>26dB</td>
</tr>
<tr>
<td>1 antenna</td>
<td>72dB</td>
<td>25dB</td>
</tr>
</tbody>
</table>

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

In the previous section, we have ignored some of the practical limitations including thermal noise, quantization noise, phase noise, I/Q imbalance, and dynamic range of linear power amplifier.

To test our theory with \( n_p - n_r = 1 \), we first used two 2.4 GHz dipole antennas denoted by A1 and A2, two signal generators (Agilent MXG N5182A) denoted by SG1 and SG2, and one signal analyser (Agilent MXA N9020A) denoted by SA. See the left plot in Fig. 1. A1 is connected, via a RF power combiner (ZAPD-4+3+) denoted by PC, to SA (receiver). SG1 (auxiliary generator) is connected to PC for waveform cancellation. A2 is connected to SG2 (primary generator). SG1 and SG2 are connected with each other in a master-and-slave mode to ensure phase coherency. The LO (local oscillator) of SA is driven by a 10 MHz source signal from SG1. Yet, the phase of SA is not coherent with SG1 and SG2 (unlike the case when all components are on the same board). The sampling rate used in SG1 and SG2 is 125 MHz while the sampling rate used in SA is 50 MHz. When a waveform was transferred from SA to SG1 and SG2, a factor-of-5/2 upsampling was used. The maximum output power of SG1 and SG2 is limited to 17 dBm. Depending on the waveforms, the output power without causing unleveling can be even lower. In order to create the strongest possible self-interference, we placed A1 and A2 next to each other. We used a pair of identical Gaussian pulses of deviation 40 ns to form a basic I/Q waveform \( p(t) \) for transmission from SG1 and SG2 at two different times, we measured the two I/Q channel responses \( f_1(t) = h_1(t) * p(t) \) and \( f_2(t) = h_2(t) * p(t) \) at the receiver where \( h_1(t) \) represents the in-circuit cancellation channel from SG1 to SA, and \( h_2(t) \) the over-the-air self-interference channel from SG2 to SA. To reduce the effect of additive noise, \( f_1(t) \) and \( f_2(t) \) were estimated by averaging 250 periodical measurements. We then performed the SIC by transmitting \( f_2(t) \) and \( f_1(t) \) from SG1 and SG2, respectively, concurrently and repeatedly with period 1.6 \( \mu s \). The additional phase \( \theta \) was used to minimize the effect of phase noise on the measurements of \( f_1(t) \) and \( f_2(t) \). We fine tuned \( \theta \) such that the residual self-interference \( x(t) \) received at the receiver was minimized. We also used a single antenna for both reception and transmission via a RF circulator (CS-2.500) denoted by CR in Fig. 1. The maximum output power of SG1 and SG2 is 125 MHz while the sampling rate used in SA is 50 MHz. When a waveform was transferred from SA to SG1 and SG2, which is similar to the balun method in [10], the SIC (optimized over \( \alpha \) and \( \theta \)) was also within 20–30 dB.

The effect of noise and nonlinearity of power amplifier is well known. To understand the impact of I/Q imbalance (of the transmitters only), consider a RF signal from the \( k \)th generator: \( \tilde{x}_k(t) = a_k(t)(1 + \delta_k) \cos(2\pi f_c t + \phi_k) - b_k(t)(1 - \delta_k) \sin(2\pi f_c t - \phi_k) \) where \( \delta_k \) and \( \phi_k \) are the amplitude and phase imbalances and \( r_k(t) = a_k(t) + jb_k(t) \) is the complex I/Q waveform that drives \( \tilde{x}_k(t) \). But the actual I/Q waveform generated is \( x_k(t) = T_k \{ r_k(t) \} = a_k(t)(1 + \delta_k)e^{i\phi_k} + b_k(t)(1 - \delta_k)e^{-i\phi_k} \) or equivalently \( x_k(t) = T_k r_k \) where for example \( x_0(t) \) is the 2 \( \times \) 1 real vector form of the complex function \( x_k(t) \). The 2 \( \times \) 2 matrix \( T_k \) is identity if and only if \( \delta_k = \phi_k = 0 \). Now, if we ignore all noises and apply \( r_k(t) = p(t) \) for both transmitters, then the measured channel responses are \( f_1(t) = h_1(t) * T_k \{ p(t) \} \) and \( f_2(t) = h_2(t) * T_k \{ p(t) \} \). For cancelation, \( -f_2(t) \) and \( f_1(t) \) are applied at the transmitters 1 and 2, respectively and concurrently. Then the residual interference received is \( x(t) = -h_1(t) * T_k \{ f_2(t) \} + h_2(t) * T_k \{ f_1(t) \} \) or equivalently \( x(t) = -H_k(t) * T_k H_2(t) * T_2 F_2(t) + H_2(t) * T_k H_1(t) * T_1 F_1(t) \) where \( H_k(t) = \begin{bmatrix} Re\{h_k(t)\} & -Im\{h_k(t)\} \\ Im\{h_k(t)\} & Re\{h_k(t)\} \end{bmatrix} \). In other words, the I/Q interference destroys the permutability of cascaded operations—a cause of the residual interference.

V. CONCLUSION

We have presented the TDTB method for SIC at the RF front-end for a broadband full-duplex MIMO radio. This time-domain method can be directly implemented at the circuit level, which does not have the prefix-region problem associated with the OFDM-based FDTB method. Our experiment shows that the RF system theory shown in this paper is highly feasible with the current RF chip technology.

REFERENCES