Analog Processing to Enable Scalable High-Throughput mm-Wave Wireless Fiber Systems

(Invited Paper)

Mahmoud Sawaby(1), Babak Mamandipoor(2), Upamanyu Madhow(2), Amin Arbabian(1)
(1) EE Department, Stanford University, Stanford, CA, USA
(2) ECE Department, University of California, Santa Barbara, CA, USA

Emails: {msawaby, arbabian}@stanford.edu, {bmamandi, madhow}@ece.ucsb.edu

Abstract—This work explores challenges in silicon integration of scalable high-throughput “Wireless Fiber” links that exploit the increase in spatial and spectral degrees of freedom at higher carrier frequencies due to LOS MIMO spatial multiplexing and higher bandwidths. In order to utilize these increased degrees of freedom, however, hardware must scale in dynamic range, speed and number of antenna elements. To this end, we examine tradeoffs in the partitioning of functionality between the transmitter and receiver, as well as between the analog and digital domains, and investigate a new scalable analog processing architecture for the receiver.

Index Terms—Line of Sight, MIMO, antenna array, spatial multiplexing, broadband communication systems, mm-wave, analog, digital.

I. INTRODUCTION

Today’s demand for mobile multimedia and connectivity present unprecedented data rate requirements on wireless networks [1]. Multi-Gigabit per second (Gbps) links are needed for many applications, including but not limited to streaming high-fidelity multimedia to mobile users, implementing over-the-air backhaul links for 5G wireless networks, high speed interconnects inside data centers [2], and in supplying internet through the sky [3]. Moreover, the availability of high-throughput low-cost wireless provide efficient links for moderate range applications, which fills the gap between current technologies, such as copper cable (inexpensive moderate speed links at limited range), and optical fiber (high speeds, large range, but with high integration and deployment complexity and power consumption overhead).

Recent improvements in silicon implementation open up the possibility of exploiting spectral and spatial Degrees of Freedom (DoF) available in mm-wave frequencies [4]. In addition to capacity gains due to large swaths of bandwidth available at mm-wave, we can achieve high spatial multiplexing gains, even in pure line-of-sight (LoS) environments with reasonable array sizes: the spatial DoF for a transmit array of aperture $A_T$, and a receiver array of aperture $A_R$, in a LoS environment, with link range of $D$ and carrier wavelength $\lambda$, is given by $\frac{A_T A_R}{D^2 \lambda^2}$ [5]. That is, the spatial DoF is inversely proportional to $\lambda^2$. Putting this together with the roughly linear scaling of bandwidth with carrier frequency, we can hope for a cubic increase in capacity with frequency.

Many authors have studied optimal geometrical array placement in order to achieve maximum spatial multiplexing gains using parallel linear and planar arrays in LoS environments

II. HARDWARE IMPLEMENTATION CHALLENGES

Due to the super-linear increase of the DoF with frequency, the capacity of a mm-wave link can be orders of magni-
In this section, we briefly describe some challenges in high-speed
mm-wave circuits, and provide insight on why replicating the
conventional low frequency architectures can significantly
limit utilization of the DoF available on mm-wave channels.

1) Limited Transmit Power: In a LoS wireless link, the
signal-to-noise ratio (SNR) at the input of a receiver is given by:

\[ \text{SNR} \propto G_{TX} G_{RX} \frac{P_{TX}}{\sigma^2 f_{BW}} \left( \frac{\lambda}{\pi D} \right)^2, \]  

(1)

where \( G_{TX} \) and \( G_{RX} \) are the transmit and receive antenna
gains, \( P_{TX} \) is the transmit power, \( \sigma^2 \) is the noise power density
at the receiver input, \( f_{BW} \) is the noise bandwidth.

As we scale up the frequency into the mm-wave regime
while fixing the range (\( D \)) and the transmitter and receiver
apertures, we notice the following:

- \( G_{TX} \) and \( G_{RX} \) increase with frequency (\( \propto 1/\lambda^2 \)) [12].
- \( f_{BW} \) increases linearly with the signal bandwidth (\( \propto 1/\lambda \)).
- \( P_{TX} \) decreases due to circuit limitations (\( \propto \lambda \)) [13].

Substituting these findings in (1), it is clear that the received
SNR does not benefit from the high antenna gain that come
with scaling the carrier frequency into the mm-wave regime.
In order to increase the communication range, one of the
remaining few options is to operate the transmitter as close
as possible to the peak power, which means to transmit
signals with the smallest possible peak-to-average power ratios
(PAPR), restricting the amount of precoding that can be done at
the transmitter.

2) Analog-to-Digital Conversion Complexity: It is impor-
tant to note that using spatial multiplexing over a LoS MIMO
channel has a similar effect on the received signal as having
a more complicated multipath channel. To illustrate the con-
sequences for implementation complexity, consider the LoS
MIMO system shown in Figure 1(c) with QPSK constellations
used at the transmitter side. The independent streams arrive at
each receiver element with random phases and relative delays,
causings the received signal to have a more complex structure,
requiring ADCs with larger dynamic range to limit information
loss due to digitization.

With today’s technologies and circuit techniques, the resolu-
tion capabilities of high speed ADC running at tens of Giga-
symbols per second (GSps) is limited to only a few bits [14].
Moreover, the power consumption of such high speed ADCs
scales with speed and resolution as follows [15]:

\[ P_{ADC} \propto f_s^2 \cdot 2^{#Bits}. \]

The strong dependency of ADC power consumption on
sampling frequency limits practical implementations to close-
to-Nyquist-rate sampling frequencies. However, not employing
fractional sampling can cause performance losses in fully
digital architectures [10]. Also, the complexity and power con-
sumption of the ADC scales exponentially with its resolution,
doubling for each additional bit.

III. REALIZATION OF MOSTLY-ANALOG ARCHITECTURES

In a fully digital architecture, the higher dynamic range
requirements in a LoS MIMO receiver does not affect the
ADC complexity alone, but is also reflected on the complexity
of digital processing after the ADC. In conventional systems,
digital processing is traditionally more appealing due to many
benefits, such as (1) the higher noise margins to crosstalk and
other forms of noise coupling from the environment, (2) the
continued benefits from device size scaling, (3) the advances
in automated digital implementation tools and methodologies,
and (4) the flexibility of programming and adapting the
hardware to different functionalities. However, from a Shannon
capacity point of view, this form of signal representation
wastes most of the capacity of circuits and wires by loading
them with only binary voltages, trading off throughput per
transistor count with reliability and modularity.

As we scale the speed requirements, the benefits that digital
implementations enjoy start to fade in comparison to their
analog counterparts. In a sense, we can view analog processing
as a way to compress digital signals into fewer nets, potentially
leading to simpler routing between system blocks and fewer
transistors at the inputs and outputs of each block, which can
be very beneficial in high throughput MIMO detection as we
show latter in this paper.

A. Channel Decomposition

For the purpose of this study, we consider a uniform square
antenna array with four antennas at each of the TX and RX
sides. The antenna configuration is shown in Figure 2(a). The
array spacing is given by [7]:

\[ d_{H,V}^{opt} = \sqrt{\frac{D \lambda}{N_{H,V}}}, \]

(2)

where \( d_{H}^{opt} \) and \( d_{V}^{opt} \) are the optimum horizontal and vertical
spacing, \( N_{H} \) and \( N_{V} \) are the number of array elements in the
horizontal and vertical dimensions, respectively, \( D \) denotes
the link distance, and \( \lambda \) is the carrier wavelength. The values used
in this study are annotated to Figure 2(a). The carrier frequency
is chosen to be 130 GHz. In theory, \( 4 \times 4 \) MIMO system can
give up to four fold more capacity over the corresponding
SISO system. It can be shown from Equation (2) that antenna
spacing for LoS MIMO is only practical in the case of mm-
wave links, were the wavelength is small enough for the
needed antenna spacing to be feasible.
For a two dimensional $4 \times 4$ MIMO array with the geometry defined by Equation (2), the channel matrix is given by:

$$H_{\text{LoS}} = \begin{bmatrix}
1\angle 0^\circ & 1\angle 180^\circ & 1\angle 90^\circ & 1\angle 90^\circ \\
1\angle 90^\circ & 1\angle 180^\circ & 1\angle 90^\circ & 1\angle 90^\circ \\
1\angle 90^\circ & 1\angle 90^\circ & 1\angle 0^\circ & 1\angle 180^\circ \\
1\angle 180^\circ & 1\angle 90^\circ & 1\angle 90^\circ & 1\angle 0^\circ 
\end{bmatrix},$$  \hspace{1cm} (3)

which represents a full-rank channel, hence it can be inverted using zero-forcing (ZF) detection. The ZF detector used is memory-less (single tap) spatial equalizer. The memory-less assumption is not valid if the TX or RX array is slightly tilted. We can assume a tilt as shown in Figure 2(b), where TX and RX elements have extra delay $z^{-\alpha_i}$ and $z^{-\beta_i}$, and random complex phase shifts of $\alpha_i$ and $\beta_i$. The channel matrix accounting for a tilt in the TX and RX side is given by

$$H_{\text{tilted}} \approx \text{diag} \left[ \alpha_1 z^{-\alpha_1} \alpha_2 z^{-\alpha_2} \alpha_3 z^{-\alpha_3} \alpha_4 z^{-\alpha_4} \right]$$

$$\times H_{\text{LoS}} \times \text{diag} \left[ \beta_1 z^{-\beta_1} \beta_2 z^{-\beta_2} \beta_3 z^{-\beta_3} \beta_4 z^{-\beta_4} \right]$$

$$= T_{\text{RX}} \times H_{\text{LoS}} \times T_{\text{TX}}$$

$$= T_{\text{RX},e_{\text{T}}} \times T_{\text{RX},n_{\text{T}}} \times H_{\text{LoS}} \times T_{\text{TX},e_{\text{T}}} \times T_{\text{TX},n_{\text{T}}},$$  \hspace{1cm} (4)

where $T_{\ldots,e_{\text{T}}}$ and $T_{\ldots,n_{\text{T}}}$ are the diagonal delay matrices representing sub-bit period and full-bit period delays, respectively, at each of the TX and RX sides. The diagonalization done here assumes that the small tilting perturbations does not cause rank reduction (and thus the matrix $H_{\text{LoS}}$ is unchanged), which is a good approximation for angles around $10^\circ$, especially since the number of array elements is small [7]. At such tilting angles, the terms $z^{-\alpha_i}$ and $z^{-\beta_i}$ can be as large as four symbol period delays for symbol rates around 20 GSps.

### B. Analog Components

Using the channel decomposition described in section III-A, we can see that MIMO detection need two main operations: (1) complex multiply-and-add to invert the LoS part of the channel ($H_{\text{LoS}}$), and (2) true-time delay to compensate for the excess memory due to misalignments. In this subsection we investigate the performance trade-offs of analog components that can be used to implement these functionalities.

#### 1) Analog Multiply and Add

Mathematically, the multiplication of two complex row vector $V$ and a column vector $S$ can be represented as:

$$\sum_k (v_k, \text{real} s_k, \text{real} - v_k, \text{imag} s_k, \text{imag}) + j(v_k, \text{imag} s_k, \text{real} + v_k, \text{real} s_k, \text{imag}),$$

where $j$ is the imaginary square root of -1. If we denote each individual iteration on $k$ by $i_k = i_{k, \text{real}} + j i_{k, \text{imag}}$, then one possible implementation to this product using analog hardware is shown in Figure 3(a).

The output current $i_k$ from such circuit is proportional to the product $s_k$ and $v_k$. The two main performance metrics of such multiplier are the operating speed and the signal dynamic range, and are given by:

$$\text{Dynamic Range} \propto \frac{I_{ds}}{f_{BW}},$$  \hspace{1cm} (5)

and

$$\text{Speed} \propto \frac{I_{ds}}{C_{load}},$$  \hspace{1cm} (6)

where $I_{ds}$ is the DC current through each device, $f_{BW}$ is the noise bandwidth, and $C_{load}$ is the effective load at the output node. Equations (5) and (6) indicate that the current consumption of the circuit has to be large enough to satisfy the dynamic range and speed requirements, simultaneously. Using the process parameters of an advanced CMOS technology, Figure 3(b) compares the current consumption derived from Equations (5) and (6) as we sweep the signal dynamic range. The speed limited part of the plot (for small to moderate dynamic range) is very interesting for the following reasons:

1) The power consumption of the analog circuit is independent from dynamic range scaling (for example using a larger constellation size modulation). This is not true for digital circuits where every added bit increases the complexity in virtually all of the operating regions.

2) Optimizing an analog circuit for speed with less dynamic range considerations usually yields small device sizes and shorter wires, which results in compact silicon area that can be competitive to high speed digital implementations.

#### 2) Analog True-Time Delay

Analog true-time delay (TTD) implementations are generally an approximation to the exponential $e^{-ST_d}$, where $T_d$ is the amount of delay in seconds and is usually programmable in discrete steps. There are two common approximations to this exponential in the literature: Taylor and Padé approximations [17]. These approximations are usually implemented in a form of a filter approximation [18], or by tuning the electrical length of the delay line the signal passes through. Although passive implementations of TTD are possible [19], [20], active implementations are usually preferred since they provide better gain stability for the same delay tuning range.

Table I summarizes some recent implementations of active programmable delay lines. An interesting note is that the ratio between delay-bandwidth product to the power consumption of the delay line is almost constant independent of the technology and the architecture used in the implementation. Using this observation, we can conclude that the power consumption of an analog delay line increases linearly with the required delay.
range for a fixed operating bandwidth. Comparing this finding to digital implementations requires us first to distinguish between two types of digital delays:

1) Digital symbol-period delays: which are relatively simple to implement using low-cost digital flip-flops.
2) Digital sub-symbol-period delays: which would generally require oversampling at the ADC, leading to increased complexity and super-linear increase in the power consumption [15].

With this in mind, if we consider the channel decomposition described in Section III-A, it is clear from Equation (4) that the maximum amount of delay required to invert the channel is determined by sum of the two longest delays in the diagonal matrices $T_{TX}$ and $T_{RX}$. Approximately half the amount of this delay can be compensated for in the digital domain (for example at the transmitter side by pre-aligning the bit streams and sending $X = T_{TX}^{-1} X_{s} + U$, where $U$ is the transmitter message) reducing the complexity of the detector. On the other hand, the sub-symbol-period delays in the matrices $T_{TX . eT_s}$ and $T_{RX . eT_s}$ require small amount of delays and can be implemented in the analog domain more efficiently.

### C. Performance Comparison

Using the ideas described in the previous sections, we now describe our proposed analog LoS MIMO detector and evaluate its performance. Figure 4 shows the proposed analog channel separation network (CSN). The complex multiply-and-add block is composed of an array of the analog multiply-and-add circuit shown in Figure 3(a), that is used to invert the $H_{LoS}$ part of the channel and separate the individual streams before feeding them to the input of the slicer (or more generally, the ADC). This aligns perfectly with the desire to reduce the dynamic range of the signal at the ADC input, thus reducing its complexity.

Excess channel delays due to misalignments at the receiver are corrected using digitally controlled analog delays. Figure 5 compares the performance of our proposed architecture with the optimized fully digital implementation described in [10]. For both systems, we assume the same LoS MIMO channel described in Section III-A, with 20 GSPs QPSK bit loading on each stream. The symbol error rate (SER) for each of the two systems is shown in Figure 5 along with the ideal ISI-free QPSK performance. The hardware components for both systems are designed to minimize the SNR penalty in the SER performance. The mostly-analog architecture uses simple single-bit slicers instead of the power hungry 5-bit ADCs required in the fully digital architecture, cutting down the power consumption in the full receiver to almost one third compared to the fully digital architecture. The benefits in terms of reduced power consumption become more significant as we increase the modulation order to medium sized constellations, since ADC requirements are tightly coupled to the modulation size in a conventional architecture. This is not the case for analog processing as long as it operates in the dynamic-range-insensitive regime described in Section III-B.

### D. Analog Channel Identification

We now show that analog techniques can also be extended to MIMO channel identification. Traditionally, adaptive filtering techniques, such as least-mean-square algorithms [24], are used to iteratively tune the detector and learn the channel coefficients. Analog channel adaptation techniques are usually based on sending orthogonal low frequency pilots at the transmitter and recording their magnitudes and phases at the receiver to learn the channels [9]. Both techniques can be efficient in estimating the memory-less part of the channel ($H_{LoS}$), but require high resolution ADCs running at high speeds to be able to program the analog delay lines. In order to avoid the complexity of using high-speed ADCs, we need to first identify the information that can be collected using simple analog operations. Let us first introduce one extra analog operation to obtain the peak values for each of the received signals, which can be easily achieved using simple analog peak detectors, as shown in Figure 6. If we turn on one transmitter at a time, and assuming we first run the transmitters at low speeds (so that the memory part of the channel is irrelevant), we can record all the individual quadrature channel gains $|h_{i . I}|$ and $|h_{i . Q}|$. The relative sign of the coefficients can be found by mixing two channels, once by addition and once by subtraction; the maximum between these two indicate the relative sign between the streams. For example, if $|h_{i . I} + h_{i . Q}|$
is greater than $|h_{i, T} - h_{i, Q}|$, then $h_{i, T}$ and $h_{i, Q}$ have the same sign. This idea of maximizing the sum can be extended to learn the relative delays between the quadrature channels, this time with the transmitters running at high rate. The flow chart shown in Figure 7 describes the proper sequence to apply this method for channel identification.

IV. CONCLUSIONS

In this paper, we have presented new methodologies and architectures for implementing high throughput wireless LoS MIMO transceivers. We propose efficient and scalable analog processing techniques that can dramatically reduce the complexity of such systems without sacrificing performance. The trade-offs of the analog components required to realize such architectures is studied, and an example architecture utilizing those ideas is introduced and compared to a conventional fully digital architecture. We extend analog processing ideas to channel identification, and show that channel state information can be obtained with minimal hardware overhead added to the proposed analog architecture.

ACKNOWLEDGMENT

This work was supported in part by Systems on Nanoscale Information Fabrics (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA, and by the National Science Foundation under grants CNS-1518632 and CNS-1518812.

REFERENCES


