High-Throughput Beamforming Receiver for Millimeter Wave Mobile Communication

Bei Yin∗, Shadi Abu-Surra∗, Gary Xu∗, Thomas Henige∗, Eran Pisek∗, Zhouyue Pi∗, and Joseph R. Cavallaro∗

∗ECE Department, Rice University, Houston, TX, USA
Email: {by2, cavallar}@rice.edu
∗Dallas Technology Laboratory, Samsung Electronics, Dallas, TX, USA

Abstract—In this paper, we present a novel FPGA-based high-throughput beamforming MIMO receiver for millimeter wave mobile communication. With vast spectrum and small antenna element size, millimeter wave communication becomes very attractive and promising to support next generation mobile communication (5G). However, the high data rate requirement challenges both algorithm and architecture. In order to support the high data rate and to reduce the overhead of selecting the best beam pair, we propose a novel beamforming synchronization scheme more suitable for mobile communication. By further optimizing the algorithm and the architecture, we present a complete mobile receiver based on FPGA, which includes RF frontend, ADC, beamforming control, synchronization, channel estimator, soft MAP detector, and channel decoder. The design operates at 28 GHz carrier frequency with 500 MHz bandwidth. The throughput can reach 1.52 Gbps. We also performed the indoor and outdoor over-the-air transmission field tests. This work provides a platform for future millimeter wave mobile communication research.

Keywords—millimeter wave mobile communication; beamforming; FPGA.

I. INTRODUCTION

Mobile communication has been one of the most successful innovations in modern history. In recent years, the number of subscribers to mobile communication services has exceeded 5 billion and is growing fast. At the same time, new mobile communication technologies, such as LTE and WiMAX systems, have been developed to satisfy the increasing demand. However, as more and more users utilize mobile communication systems, there is an increasing need for a mobile communication system with larger capacity, higher throughput, and lower latency.

Conventionally, millimeter waves refer to radio waves with frequency of 3 GHz – 300 GHz. These bands exhibit unique propagation characteristics [1][2]. For example, compared with lower frequency bands, they suffer higher propagation loss, have poorer ability to penetrate objects, and are more susceptible to atmospheric absorption, deflection and diffraction due to particles (e.g., rain drops) in the air. On the other hand, due to their smaller wave lengths, more antennas can be packed into a relatively small area. In addition, because these bands have been less utilized than the lower frequency bands, they have less existing interference and lower spectrum cost. All these unique challenges and opportunities brought many researchers to investigate the utilization of millimeter waves for wireless communication. Power efficient RF integrated circuits were built for the unlicensed 60 GHz band in [3]. The authors in [4] built a 55 GHz system for TV studio usage. A few industrial consortiums also developed technologies and standards by using the 60 GHz band, such as IEEE 802.11ad [5], and Wireless Gigabit Alliance [6].

Although the 60 GHz band can provide very high data rate up to 10 Gbps within 1 meter and 3 Gbps within 10 meters range [7], the high atmospheric absorption limits the usage of 60 GHz for outdoor mobile communication [2]. Recently, in [2], authors showed that by choosing a lower unused band at 28 GHz, it is possible to achieve Gbps data rate in an urban mobile environment. By further measuring the channels in the outdoor environments, the authors in [8][9] showed that the 28 GHz band exhibits better propagation characteristics compared to 60 GHz and higher bands, and it has similar free space path loss to the widely used cellular bands. These facts enable millimeter-wave communication to be a strong candidate for the next generation mobile communication (5G).

However, the high data rate requirement of this new system challenges both algorithm and implementation. In this paper, we investigate the feasibility of a Gbps data rate receiver for millimeter-wave mobile communication. As far as we know, there is no related prior similar work. In order to support high data rate with low beam pairing overhead, which is critical for the millimeter-wave mobile communication, we propose a novel beamforming synchronization algorithm to rapidly select the best transmitter and receiver beam pair. Then by highly optimizing the algorithm and architecture, we further present a complete FPGA-based receiver. It consists of RF unit, ADC, beamforming control, automatic gain control (AGC), synchronization, channel estimator, soft MAP detector, and channel decoder. The receiver operates at 28 GHz carrier frequency with 500 MHz bandwidth. Compared to the existing LTE prototype using soft sphere detector at 220 Mbps [10], and 60 GHz prototype using linear equalizer at 228 Mbps [11] and 1.08 Gbps [12], our design can run up to 1.52 Gbps on a BEECube multi-FPGA system [13] using soft MAP detector. We further performed the indoor and outdoor over-the-air transmission. This platform enables many research opportunities for future millimeter wave mobile communication.

The following section II presents the system overview. In section III, the proposed beamforming synchronization algorithm is described. Architecture and implementation results are shown in section IV. We conclude the paper in section V.

The Rice University co-authors were supported in part by the NSF under grants CNS-1265332, ECCS-1232274, and EECS-0925942.
chains to transmit $N_T$ streams. The phase shifters of these RF chains are controlled by a beamforming processor at baseband to form analog beams. This process is called analog beamforming. By using analog beamforming, the number of required DACs is greatly reduced, which lowers the power consumption and cost.

At the receiver, analog beamforming with $kN_R$ RF-chains is used to adjust the direction of the receiver analog beam. The CP is removed from the received signal. Then a 2048-point FFT at each receiver baseband channel transforms the data into the frequency domain. The equivalent system in frequency domain for subcarrier $j$ is

\[ Y_j = B_j^X H_j X_j + N_j = H_j^{\text{map}} X_j + N_j \]

where $Y_j$ is an $N_R \times 1$ vector of the received baseband data; $B_j^X$ is an $N_R \times kN_R$ matrix corresponding to the receiver analog beamforming at subcarrier $j$; $H_j$ is a $kN_R \times kN_T$ channel response matrix of subcarrier $j$; $B_j^X$ is a $kN_T \times N_T$ matrix corresponding to the transmitter analog beamforming at subcarrier $j$; $X_j$ is an $N_T \times 1$ vector of the transmitted baseband data; $N_j$ is an $N_R \times 1$ vector of complex additive white Gaussian noise with zero mean and variance $\sigma^2_N$. The $B_j^X H_j X_j^T$ can be lumped together to be represented by an $N_R \times N_T$ equivalent channel matrix $H_j^{\text{map}}$ for subcarrier $j$. With the aid of pilots, this lumped channel is estimated at the channel estimator as $H_j^{\text{map}}$. The MIMO detection uses this estimated channel and the corresponding demapped received data $Y_j$ to compute the soft information as log-likelihood ratio (LLR) values. Then the LDPC channel decoder decodes the information bits from the LLR values.

### III. MILLIMETER WAVE BEAMFORMING RECEIVER

The key modules in the receiver are synchronization, beamforming control and AGC, channel estimation, and MIMO detection. In this section, a novel beamforming synchronization algorithm to support high data rate transmission is described.

#### A. Synchronization

The synchronization module is responsible for slice identification, timing, and frequency synchronization by detecting and processing the sequences in the SCH.
The SCH is designed to occupy only 1/8 of the system bandwidth and is centered around the carrier frequency. The location and the structure of the SCH in the frame are shown in Fig. 3. The BS transmits the synchronization signal on SCH in the first slot of every subframe. Inside that slot, one SCH OFDM symbol is transmitted for each slice. The SCH OFDM symbol also carries information about the slice-ID. For example, if there are four slices, SCH spans four OFDM symbols with \( Z_1, Z_2, Z_3, \) and \( Z_4 \), which is corresponding to each transmitter slice, respectively. However, because the receiver is not aligned with the transmitter OFDM symbol timing before acquiring a synchronization signal, it is possible for the receiver to start detecting the SCH OFDM symbols at the \( i \)-th transmitter slice \( Z_i \). And because the MS changes the receiver slice every subframe (1ms), MS would never be able to detect the synchronization signal. To solve this problem, by proposing a similar approach as the cyclic prefix, the first symbol in the SCH is copied and added to the end of SCH. By doing this, BS transmits SCH with 5 OFDM symbols as \( Z_1, Z_2, Z_3, Z_4, \) and \( Z_5 \) for the first subframe. In the following subframes, the BS transmits SCH with \( Z_2, Z_3, Z_4, Z_5, \) and \( Z_1 \), and so on.

Each SCH OFDM symbol is constructed using a unique Zadoff-Chu (ZC) sequence populating only the even subcarriers of the OFDM symbol. As a result, this SCH OFDM symbol, in time domain, has two identical halves. Thus, the SCH symbol is detected by scanning the air for a symbol with two identical halves using half-symbol correlation as

\[
D = \sum_{i=0}^{L/2-1} r_i^* r_{i+L/2}, \quad E = \sum_{i=0}^{L/2-1} r_i^* r_{i+L/2},
\]

\[
M = \frac{|D|^2}{E^2},
\]

where \( r_i \) is the received synchronization data at time \( i \); \( L \) is the length of an SCH symbol, which is 256, 1/8 of the system bandwidth, in our system; \( D \) is the correlation between two halves of the SCH symbol; \( E \) is the received power of half a SCH symbol. To tolerate the existing noise, the SCH symbol is set to be detected if \( M \) is larger than threshold. Then the transmitter slice-ID is determined by correlating the received SCH symbol with all reference SCH symbols and finding the maximum one, which is also used to confirm the SCH detection.

**B. Analog beamforming control and AGC**

We propose a two-stage beam pairing algorithm, which runs based on the state of the above synchronization module. In the first stage, the MS first records its receiver slice to the maximum correlated slice detected in the synchronization module. Then the MS sends the transmitter slice-ID to the BS in the uplink. Later the BS will use this slice to transmit data to this MS. However, the quality of the link can be further improved in a second stage by scanning a set of narrower beams within that selected slice in a similar procedure. Then the data is transmitted on the paired narrow beams.

Although it is straightforward for both transmitter and receiver to scan all combinations of transmitter and receiver beams to find the strongest beam direction, this direct search is very time consuming and not energy efficient. For example, in each round, the transmitter fixes its beam at one transmitting direction, and the receiver scans all receiving directions. This whole procedure in total uses \( q^2 \) SCH OFDM symbols to scan all combinations, where \( q \) is the number of narrow beams in each sector.

The better two-stage beamforming approach, proposed in this work, tremendously reduces the number of required SCH OFDM symbols to complete the beam pairing, which is much more suitable for mobile communication. In the first stage, the receiver and transmitter use a wide beam slice, each covering \( \alpha \) narrow beams. The receiver and transmitter can detect the strongest beam directions for this wide beam in \( q^2 / \alpha^2 \) OFDM symbols. Then the receiver and transmitter switch to narrow beam mode to scan inside that strongest wide beam. This takes \( \alpha^2 \) OFDM symbols to find the strongest beam direction for the narrow beam. In total, this algorithm uses \( q^2 / \alpha^2 + \alpha^2 \) OFDM symbols. When \( \alpha = \sqrt{q} \), the number of SCH OFDM symbols to detect the strongest beam pair is minimum, \( 2q \), which is much less than the previous \( q^2 \).

After finding the best beam pair, the scanning still continues running to track any changes due to mobility and obstructions. Once a better quality pair is found, the MS and BS will update the recorded pair for data transmitter.

As described above, in order to scan for the best beam pair, it is important to set the signal power at the input of the analog-to-digital converter (ADC) for each beam pair to a proper value in the ADC dynamic range. The AGC at the MS baseband adjusts the gain of the RF according to the received power \( E \), which is computed in the synchronization module.

**C. Channel estimation**

After pairing the transmitter and the receiver beams, pilots are sent out to estimate the MIMO channels. In the transmitter, as shown in Fig. 4, pilots from each transmitter chain are first interleaved in the frequency domain, and then are inserted into the data stream at specific subcarriers every 6 cycles. Although this scheme uses more bandwidth, it can achieve more stable channel estimation. At the receiver, with the received pilot, the
corresponding channel coefficient $\hat{h}_{mn,j}$ in the channel matrix $\tilde{H}_{j}^{Lump}$ for subcarrier $j$ is estimated by using the least square estimator $\hat{h}_{mn,j} = p_{mn,j}^{-1}y_{mn,j}$, where $p_{mn,j}$ is the reference pilot for subcarrier $j$ from transmitter chain $n$ to the receiver chain $m$; $y_{mn,j}$ is the received pilot. The estimated channel coefficients are then linearly interpolated to compute the channel coefficients of the subcarriers without the pilot.

![Fig. 4: Proposed pilots and data allocation.](image)

In addition to the channel coefficients, the complex noise variance $\sigma^2$ is also estimated by the channel estimator. This is computed by using the noise on the guard band.

### D. Soft MAP detection

In order to achieve the optimal detection performance, a soft MAP detector is designed. With estimated channel matrix $\hat{H}_{BS,j}$ for each subcarrier and noise variance $\sigma^2$, the a posteriori probability (APP) detector calculates the log-likelihood ratio (LLR) $L$ of bit $k$. By performing a maxLogAPP approximation [14], the LLR computation is reduced to

$$L_k \approx \min_{x_k, \bar{x}_{k-1}} \left\{ \frac{1}{2} (y - Hx)^T C^{-1}(y - Hx) \right\} - \min_{x_k, \bar{x}_{k-1}} \left\{ \frac{1}{2} (y - Hx)^T C^{-1}(y - Hx) \right\},$$

where $x_{k-1}$ is the set of symbols, where bit $k$ is +1, and $x_{k-1}$ is the set of symbols, where bit $k$ is -1; and $C$ is the covariance matrix.

By assuming the noise signals are independent from each other, in our 2x2 MIMO QPSK receiver, we further reduce the complexity of the first term

$$\min_{x_k, \bar{x}_{k-1}} \left\{ (y - Hx)^T C^{-1}(y - Hx) \right\}$$

to

$$\min_{x_k, \bar{x}_{k-1}} (-2Re(\frac{1}{\sigma^2_1} y_1^* h_{i1} + \frac{1}{\sigma^2_2} y_1^* h_{i2}) x_k) - 2Re(\frac{1}{\sigma^2_1} y_1^* h_{i2} + \frac{1}{\sigma^2_2} y_1^* h_{i1}) x_k - 2Re(\frac{1}{\sigma^2_1} y_2^* h_{i1} + \frac{1}{\sigma^2_2} y_2^* h_{i2}) x_k) \right\}$$

where $Re()$ is the real part of a complex number; $y^*$ is the conjugate of $y$; $h_{ij}$ is the element in row $i$ and column $j$ of the channel matrix; and $\sigma^2$ is the noise variance of receiver chain $i$.

The same simplification is also performed for the other term $\min_{x_k, \bar{x}_{k-1}} \left\{ C^{-1}(y - Hx) \right\}$.

### IV. FIXED POINT DESIGN AND IMPLEMENTATION

Based on above proposed beamforming synchronization algorithm, in this section we present a complete fixed point design. The system is highly pipelined and implemented in C and Verilog targeted at a BEECube system, which has multiple Xilinx Virtex6 LX550T FPGAs. Between these FPGAs, 6.6 Gbps GTX transceivers are connected for high speed data transmission. The RF and ADC units are separately designed and finally integrated into the BEECube. In our current system, QPSK is used for reducing the power consumption and dynamic range requirement in the ADC [2]. The architecture of each module is optimized for high speed data transmission.

#### A. Synchronization

The architecture of the synchronization module is depicted in Fig. 5. Because the SCH occupies only 1/8 of the system bandwidth, the received signal is first down sampled by a factor of 8. This is done using decimation filters, which are realized by using a polyphase architecture. This architecture can reduce the clock frequency by a factor of 8. The decimated output is then sent to the buffer and also fed to the symbol-halves correlation and energy estimate blocks. The latter two blocks use a moving-sum method to compute $D$ and $E$ in Eq. (1), respectively. With $D$ and $E$, the SCH symbol detection metric, $M$, is calculated in the SCH detection trigger block. Whenever the SCH symbol is in the moving-sum window, $M$ will be above the threshold. As a consequence of the CP in the SCH symbol, $M$ will reach a maximum value and stay at that value for the length of the CP. Thus, a slope detection circuit in the SCH detection trigger block can generate a trigger signal to pass the SCH symbol without CP from the buffer to the correlation engine. The correlation engine consists of a 256-point FFT, a 256-point IFFT, and the reference SCH symbols in the frequency domain. It computes all correlations and then the peak search block will compare them to find the maximum one with its index. These two values are sent to the controller, Xilinx Microblaze processor, by using an interrupt. The synchronization interrupt routine uses this information along with the time stamp to identify the transmitter and receiver slices to be used for data transmission. Because of usage of the polyphase architecture, the clock frequency of this module is 1/8 of the 552 MHz ADC frequency, which is 69 MHz.
B. Analog beamforming control and AGC

The analog beamforming, as shown in Fig. 6, consists of two essential blocks: specially designed antenna and the baseband control. With a dedicated antenna design, the phase of each antenna is able to be adjusted from the baseband. In the baseband, all of the phase shifters are controlled by a central beamforming processor. The beamforming algorithm is implemented in C code on a Xilinx Microblaze microprocessor in the FPGA. A predefined table is designed to store different antennas phases. Each entry in the table forms one analog beam. These beams have different directions and different widths. The entry is selected based on the beamforming algorithm, described in the previous section, and the synchronization state from the synchronization module.

The AGC is also implemented in C code on the microprocessor. It adjusts the variable gain amplifiers (VGA) in the RF to tune the received signal power to the ADC range.

C. Channel estimation

As described in the previous section, during the channel estimation stage, each receiver chain will receive one complete set of pilots. Thus, in total there are $N_s$ channel estimators working in parallel, one for each receiver chain. As shown in Fig. 7, each channel estimator consists of a least square estimator, $N_a$ linear interpolation filters, and one noise variance estimator, and generates the channel coefficient $\hat{h}_{mn}$. Because all receiver chains receive the same set of pilots, only one pilot generator needs to be designed. The estimated channels are then stored and used for MIMO detection. By utilizing the fact that pilots and data are received at different OFDM symbols, the channel estimator is further pipelined to reduce the clock frequency. This module runs at 190 MHz, which is the half of the MIMO detector clock frequency.

D. Soft MAP detector

The MIMO detector is a key module in the data path. Its throughput directly affects the receiver throughput. To optimize the soft MAP detector in Eq. (2), we pipelined the design into three stages for high data rate and low complexity. The architecture is shown in Fig. 8. Re() and Im() in the figure indicate the real and imaginary parts, respectively. To reduce complexity, the 1st stage computes all shared terms for $x_1$ and $x_2$. The 2nd stage computes the 2-norm values of different combinations of $x_1$ and $x_2$ in parallel. In a 2x2 MIMO and QPSK system, 16 terms are computed. In the 3rd stage, LLRs are computed by searching for the minimum value from these 16 computed 2-norms.

This highly pipelined architecture supports the design to run up to 380 MHz, which is close to the maximum limit of the FPGA. The throughput of this module is up to 1.52 Gbps.

E. LDPC decoder

The channel code used in our system is a rate-13/16 672-bit, Quasi-Cyclic LDPC code, which is adopted in the IEEE 802.11ad standard. We use the layered decoding algorithm to reduce the amount of required memory. To achieve high throughput with low complexity, the decoder is implemented with 8-bit LLR precision. As shown in Fig. 9, the fixed point implementation has negligible performance loss compared to the floating point cases in both BPSK and QPSK scenarios.

The maximum clock frequency in this module is 28.8 MHz, and the 1.61 Gbps throughput is achieved by using two LDPC decoders to run in parallel.
TABLE I. RESOURCE UTILIZATION OF MS RECEIVER ON BEECUBE WITH 2 XILINX VIRTEX6 LX550T FPGAS

<table>
<thead>
<tr>
<th>Module</th>
<th>#Slices</th>
<th>#LUTs</th>
<th>#Registers</th>
<th>#DSP48s</th>
<th>Max.Freq.</th>
<th>Max. Data rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFTs</td>
<td>2,124</td>
<td>5,030</td>
<td>7,484</td>
<td>40</td>
<td>450 MHz</td>
<td>1.8 Gbps</td>
</tr>
<tr>
<td>Channel estimator</td>
<td>544</td>
<td>1,761</td>
<td>1,014</td>
<td>0</td>
<td>190 MHz</td>
<td>-</td>
</tr>
<tr>
<td>Soft MAP detector</td>
<td>1,749</td>
<td>5,095</td>
<td>5,866</td>
<td>36</td>
<td>380 MHz</td>
<td>1.52 Gbps</td>
</tr>
<tr>
<td>Synchronization</td>
<td>6,223</td>
<td>13,967</td>
<td>17,106</td>
<td>136</td>
<td>69 MHz</td>
<td>-</td>
</tr>
<tr>
<td>Microblaze</td>
<td>-</td>
<td>1,022</td>
<td>959</td>
<td>0</td>
<td>150 MHz</td>
<td>-</td>
</tr>
<tr>
<td>LDPC decoder</td>
<td>87,528</td>
<td>252,486</td>
<td>46,958</td>
<td>0</td>
<td>28.8 MHz</td>
<td>1.61 Gbps</td>
</tr>
<tr>
<td>Total/Available</td>
<td>98,168/ (85,920 x 2)</td>
<td>279,361/(549,888 x 2)</td>
<td>79,387/(687,360 x 2)</td>
<td>212/(864 x 2)</td>
<td>-</td>
<td>1.52 Gbps</td>
</tr>
</tbody>
</table>

F. System integration with other modules

The FFT blocks from the Xilinx IP library are used in our system. This block can support up to a 2048 point complex FFT. To achieve high data rate, the pipelined streaming mode at 450 MHz is used.

In order to maximize the integrated system throughput, the system is partitioned into a few clock domains to maximize the throughput of each separate module. The FFT blocks run at 450 MHz, the MAP detector runs at 380 MHz, the channel estimator runs at 190 MHz, the synchronization block runs at 69 MHz, the Microblaze processor runs at 150 MHz, and the LDPC decoder runs at 28.8 MHz. To enable fast data transfer between the different clock domains, the double buffering technique is used between modules. The synthesis results of all modules are shown in Table 1. The current receiver is partitioned onto 2 FPGAs, one for two LDPC decoders, and the other for all the other modules. The data are transferred between the FPGAs through 6.6 Gbps GTX transceivers. The expected overall throughput of our integrated design can reach 1.52 Gbps for uncoded data. With the rate-13/16 LDPC code, this corresponds to 1.235 Gbps information bit rate. This throughput outperforms the reported rate in [10][11][12].

G. Over-the-air test

Indoor and outdoor over-the-air tests are performed on the integrated system between one BS and one MS. The BS has a 28 GHz RF unit, 552 MHz DACs, and data transmission module on FPGA. The MS has a 28 GHz RF unit, 552 MHz ADCs, and baseband running on FPGA.

First, an Error Vector Magnitude (EVM) test was performed on the BS and MS RF units at different distances. Fig. 10(a) illustrates the EVM test results using a QPSK modulation scheme. The received signal, represented by blue points, is better than -19 dB EVM at over 50 meters in distance. Then the integrated systems were tested in both indoor and outdoor environment. Fig. 10(b) shows the LLR values of the received QPSK signal at the output of the soft MAP detector in a 180 meters outdoor line-of-sight test. The results were captured by using Xilinx ChipScope at the MS baseband FPGA.

V. CONCLUSION

In this paper, we proposed a novel beamforming synchronization algorithm suitable for millimeter wave mobile communication. We further presented a complete high-throughput receiver supporting up to 1.52 Gbps data rate. By performing the over-the-air indoor and outdoor tests, we proved that it is feasible to design and implement a high data rate beamforming receiver for next generation mobile communication system (5G). With this platform, more millimeter-wave communication research work will be performed.

Fig. 10: (a) EVM test result for RF units in BS and MS. Blue points represent the received QPSK symbols. (b) LLR values of QPSK symbols at the output of MS soft MAP detector.

REFERENCES