A Two-Beam Eight-Element Direct Digital Beamforming RF Modulator in 40-nm CMOS

Boyi Zheng, Lu Jie, John Bell, Yan He, *Member, IEEE*, and Michael P. Flynn, *Fellow, IEEE*

***Abstract*— A digital eight-element beamforming RF modulator allows accurate steering of multiple independent beams. The key to the efficiency is the pairing of area-efficient bandpass $\Delta\Sigma$ modulation with N-path filtering to suppress quantization noise. Beamforming digital signal processing (DSP) is enabled by careful frequency management, which facilitates efficient digital phase shifting, upsampling, and upconversion. A 40-nm CMOS prototype generates two 1.2-GHz beams, with 40-MHz RF bandwidth. The prototype consumes 128 mW and occupies an active area of only 0.19 mm^2, which is only 16 mW and 0.02 mm^2 per element. This represents an order-of-magnitude improvement in area and power consumption per element compared to the state-of-the-art digital to RF modulators.**

***Index Terms*— $\Delta\Sigma$ modulation, digital beamforming, digital direct to RF (DDRF), N-path filter.**

I. Introduction

BEAMFORMING is essential for high-speed wireless communication, especially for multiple-input multiple-output (MIMO) and 5G wireless. Transmit beamforming is very promising since it allows a large reduction in power amplifier (PA) power. However, for a large number of $\lambda/2$ spaced antennas, the beam becomes very narrow, and therefore, the accurate beamforming is essential. Although narrow beamwidth is good in terms of energy density and directivity, the accurate beamforming is not easy, especially when the target is moving. The need for accuracy and fast switching favors digital beamforming. Digital beamforming can also easily support multiple beams.

In conventional analog beamforming, the multiple analog stages in all the transmit signal chain contribute to nonlinearity, mismatch, and noise, which are big challenges to beam accuracy. In particular, the mixer and PA stages are significant contributors to nonlinearity in a conventional transmitter. For example, as shown in Fig. 1(a), the amplitude imbalance of I/Q modulated signal may lead to variation in the transconductance in the analog mixer [1]. Further challenges are the finite dynamic range of the V-to-I stage and switch pair, as well as the phase noise and spur added by local oscillator (LO) signal [2]. Insertion loss in the signal path and the difficulty in forming multiple simultaneous beams are additional concerns with analog beamforming.

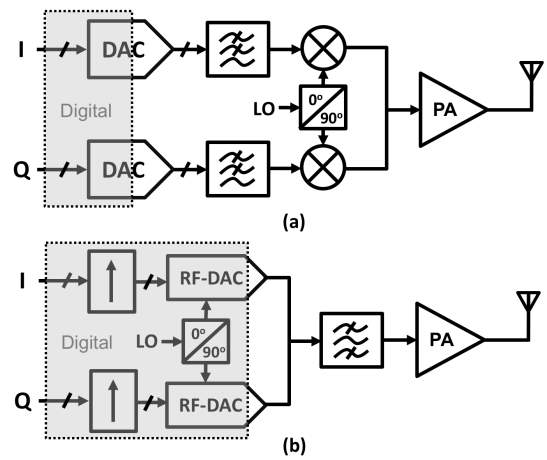


Fig. 1. Conventional TX versus DDRF TX.

The requirements for beam accuracy and multiple beams are best met by digital beamforming. The challenges of analog transmission motivate mixerless digital-to-RF transmit schemes, especially so for beamforming systems. Digital direct to RF (DDRF) transmission avoids analog nonidealities, since the upmixing, along with the phase shifting, can be digital. Digital phase shifting also facilitates multiple simultaneous beams. Furthermore, mostly digital architecture is scaling friendly. However, the conventional digital transmit beamforming requires extensive digital signal processing (DSP), large DACs, upconverters, and filters, and so tends to be large and power hungry.

Although impressive progress has been made on standalone DDRF transmitters [3]–[5], existing approaches require too much die area and consume too much power to be practical for transmit beamforming. Efficiency is very challenging for practical DDRF. Most current DDRF architectures require a high DAC resolution for high signal to noise (SNR). This high DAC resolution complicates the DAC design and results in a large die area. In other cases, off-chip filtering or calibration is needed. Roverato *et al.* [3] use a smaller 10-bit DAC and programmable $\Delta\Sigma$ modulator with out-of-band notch filtering for a 0.82 mm^2 area. Paro Filho *et al.* [4] deliver charge through a 12b resistive DAC and occupy 0.22 mm^2, but require off-chip DSP for charge calculation. Sommarek *et al.* [5] use bandpass digital $\Delta\Sigma$ modulation but require off-chip analog

Manuscript received November 24, 2018; revised March 6, 2019 and May 6, 2019; accepted May 7, 2019. Date of publication June 6, 2019; date of current version July 1, 2019. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) through the Arrays at Commercial Timescales (ACT) Program. This paper is an expanded version of a paper presented at the 2018 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium. *(Corresponding author: Boyi Zheng.)*

The authors are with the Department of Electrical and Computer Engineering (ECE), University of Michigan, Ann Arbor, MI 48109 USA (e-mail: boyizh@umich.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMTT.2019.2916803

filtering to suppress the substantial shaped quantization noise. Existing DDRF schemes have high power consumption in the range of 150–380 mW. The die area of these DDRF implementations ranges from 0.2 to 1.3 mm^2 per element. The large amount of DSP, the high DSP speed, and the significant DAC size restrict these schemes to single-element use and none support transmit beamforming.

Our digital beamforming approach combines efficient digital-to-analog conversion with compact and efficient digital beamform processing. As already stated in our previous publication [6], the key to our approach is the combination of bandpass digital-to-analog conversion and N-path filtering. A bandpass $\Delta\Sigma$ modulator directly generates the analog RF signal. The 1.5-bit DAC, driven by the $\Delta\Sigma$ modulator, is both very compact and inherently linear. The out-of-band quantization noise is suppressed using a small on-chip N-path filter, eliminating the need for off-chip filtering. The combination of a digital bandpass $\Delta\Sigma$ modulator along with a 1.5-bit DAC facilitates an order-of-magnitude reduction in DAC area.

The power and area of the beamforming DSP are minimized through frequency planning and through uncomplicated and innovative upsampling and upconversion strategies. Conventionally, a bandpass $\Delta\Sigma$ modulator runs at four times the analog output frequency; however, this would lead to an excessive clock rate of 4.8 GS/s in our case. Instead, we run the $\Delta\Sigma$ modulator at the reduced rate of 4/3 times the RF center frequency and make use of the image frequency.[1] Furthermore, in our frequency planning, the digital phase shifter runs eight times slower, leading to an order-of-magnitude reduction in DSP power. Upsampling is done through simple linear interpolation. Upconversion is with simple MUX-based processing.

The combination of digital phase shifting with digital bandpass $\Delta\Sigma$ modulation and N-path filtering enables accurate beamforming and a prototype eight-element digital beamforming RF modulator with record power and area efficiency. Moreover, unlike other DDRF approaches, which are restricted to single-element use and do not support transmit beamforming [3]–[5], our proposed DDRF modulator not only provides high beam accuracy but also supports simultaneous multiple beams, each potentially with its own unique transmit data. The 40-nm CMOS prototype generates two 0.8–1.2-GHz beams,[2] consumes 128 mW, and occupies an active area of 0.19 mm^2. The 16 mW and 0.02 mm^2 consumed per element represent an order-of-magnitude improvement compared to the state of the art.

This paper expands on the conference paper [6], providing new material and additional explanation. To give context for digital transmit phase shift beamforming, Section II-A derives the phase shift for each element. Section II-B summarizes the architecture and outlines of the signal flow. Section III explains the implementation in detail. In addition to providing more detail compared to [6], Section III adds a discussion on the advantages and implementation of linear interpolation. It also explains the design and operation of the digital $\Delta\Sigma$ modulator in detail. Furthermore, Section III discusses the digital word-length allocation throughout a single digital transmit stripe, as this is critical to balancing speed and precision. Finally, Section III considers the nonidealities of the DAC and N-path filter. Section IV presents measured results.

[1]In [7], a similar technique is presented. Unlike our design methodology, which customizes the NTF for target sampling frequency and bandwidth, Frappe *et al.* [7] switch to the image frequency when the carrier reaches a certain frequency. Since the NTF of the $\Delta\Sigma$ modulator is fixed, this causes a discontinuity in the bandwidth between various carrier frequencies. In addition, the N-path filter in our design tracks the image frequency as its center frequency.

[2]In our prototype, we use the same data for both channels—this is simply due to the limited digital input bandwidth to the chip. The modulation data are through a 200-MS/s, 10-bit wide digital bus. We use a single digital input channel so that the package size is not excessive.

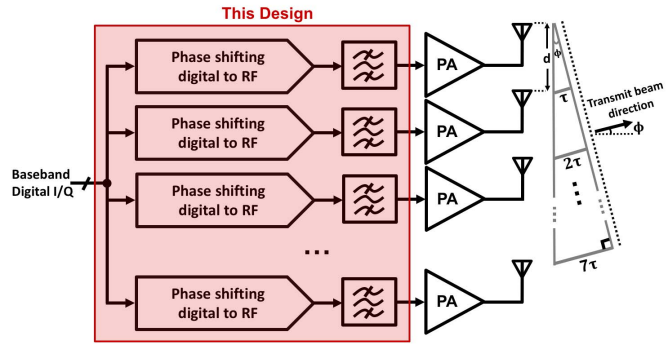


Fig. 2. Complete beamforming TX model used for calculating the desired phase shift.

II. System Architecture

We begin with an overview of the architecture. We consider the phase shift for each element in the array. We review the digital beamforming architecture.

A. Beamforming Transmit Model

A general transmit beamforming model is shown in Fig. 2. The model has eight identical DDRF transmitters each with a different phase shift. For a desired beam angle of ϕ, we derive the desired phase shift for each channel as follows.

Assuming the element spacing is d, then the time delay difference between adjacent channels should be $\tau = d \cdot \sin\phi / c$.

The delay for each channel can be represented as

$$\tau_k = k \cdot d \cdot \sin\phi / ck = 0, \ldots, 7 \tag{1}$$

in which k stands for the antenna element index and c is the speed of light.

The antenna spacing, d, is usually chosen to be $\lambda/2$ for a reasonable beamwidth and to avoid a grating lobe [8]. Then, with $d = \lambda/2$ and $f_c = c/\lambda$ (f_c is carrier frequency), we get

$$\tau_k = 0.5k \cdot \sin\phi / f_c, \quad k = 0, \ldots, 7. \tag{2}$$

Our desired transmit signal at the kth element can be represented as

$$x_k(t) = \cos(2\pi f(t - \tau_k)) = \cos(2\pi f t - 2\pi f \tau_k). \tag{3}$$

For narrowband beamforming, we can approximate the incident wave frequency with the carrier frequency, which

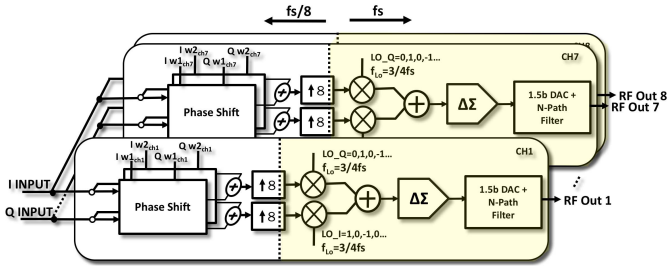


Fig. 3. System architecture.

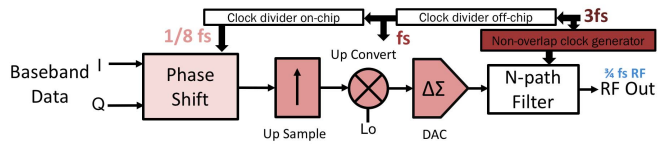


Fig. 4. Frequency planning and synchronization.

yields $f \approx f_c$. Then, we obtain the desired phase for each channel to be shifted as

$$\theta_k = 2\pi f \tau_k \approx 2\pi f_c \tau_k = \pi k \cdot \sin\phi, \quad k = 0, \ldots, 7. \quad (4)$$

This phase shift can ultimately be implemented using complex weight multiplication (CWM), as described later.

B. System Block Diagram

Fig. 3 shows our overall system architecture. The eight-element RF modulator is comprised of eight identical stripes. Each stripe phase shifts, upsamples, upconverts, and converts to analog, and finally filters each per-element output RF transmit signal.

Fig. 4 illustrates the overall clock-rate distribution in the prototype chip. We use f_s to represent the upconversion and $\Delta\Sigma$ modulator processing rate. The other frequency domains are all represented in terms of f_s for clarity. The RF center frequency is chosen as $3f_s/4$ for the simple representation of the digital LO and to reduce the power consumption of the DSP. This is elaborated in Section III-C. The N-path filter requires a clock rate of four times of the RF center frequency for a four-path nonoverlap clock generation, which yields $3f_s$. The phase shifter is operated at $f_s/8$, which is chosen in consideration of DSP efficiency, simplicity of the implementation of linear interpolator, and reasonable digital I/O speed. This frequency planning greatly eases practical hardware implementation. The simplified processing and the relaxed DSP sampling rate make digital synthesis of the entire DSP practical in 40-nm CMOS technology.

A single off-chip clock generator ensures that all operations and frequencies are synchronized. The RF center frequency is determined directly by tuning the clock rate, and then N-path filter inherently tracks the center frequency. The RF bandwidth is chosen as $f_s/40$, ensuring an oversampling rate [OSR, defined as $f_s/(2*$ bandwidth)] of 20 for the $\Delta\Sigma$ modulator. The bandwidth of the N-path filter, which is also $f_s/40$, complements the noise transfer function (NTF) of $\Delta\Sigma$ modulator. In the rest of this paper, f_s is 1.6 GS/s. The RF bandwidth, the RF carrier, and input sampling frequency are

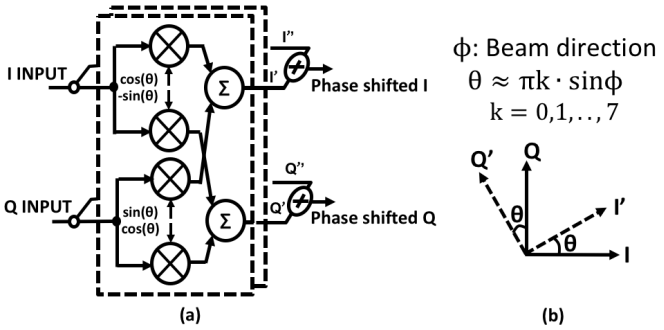


Fig. 5. Phase shifting using CWM. (a) Implementation of CWM. (b) Illustration of phase shifting in complex plane.

40 MHz, 1.2 GHz, and 200 MS/s, respectively, and all are derived from f_s.

Beginning with phase shifting in each stripe, time-delay is approximated by phase shifting the baseband I/Q transmit data. The input baseband I and Q transmit data are multiplied with I and Q coefficients. We form two simultaneous beams by multiplying with two sets of complex weights. After multiplication, the phase shifted I/Q data from different beams are summed to form single I and Q streams.

Fig. 5 shows the phase shifting operation in more detail. The desired phase shift, θ, for each channel is calculated from the desired angle for the main transmit lobe(s) as given in (4). The phase shifting is implemented with CWM.[3] In CWM, we multiply the I and Q inputs by cos θ and sin θ, respectively, and add to form the shifted, I', data stream. Similarly, we multiply by sin θ and cos θ, respectively, and add to form the shifted Q' data stream. We multiply with two sets of weight coefficients to form two independent beams and then add these together to form the multibeam phase shifted I/Q baseband data.

After phase shifting, the phase-modulated I/Q streams are upsampled by 8 to the f_s (i.e., $4/3f_c$) sampling-rate using linear interpolation. Linear interpolation has spectral advantages over interpolation with nonreturn-to-zero, while the hardware complexity is still kept low. This is explained in Section III-B. The upsampled baseband streams are upmixed from baseband to RF by multiplication with digital I/Q LO sequences. The ratio between sampling rate and RF center frequency ($f_{\mathrm{LO}} = 3/4f_s$) allows the digital LO signals to be represented by -1, 0, 1, 0..., which greatly simplifies multiplication and summing and makes DSP processing a lot more efficient. We also use the image frequency to reduce the final sampling rate by a factor of 3—Section III-C elaborates on this.

The upconverted, I and Q streams are combined into a single digital stream and then noise-shaped by a digital bandpass $\Delta\Sigma$ modulator to drive a 1.5-bit DAC, which directly generates the analog RF signal. This avoids the need for a high-resolution DAC which would need a large additional die area and often requires extra calibration circuitry to reduce

[3]The fractional bandwidth and the number of elements are small enough so that phase shift is a good approximation to time delay. A rule of thumb is that phase shifting is sufficient if the bandwidth is less than 10% of the carrier frequency [9].

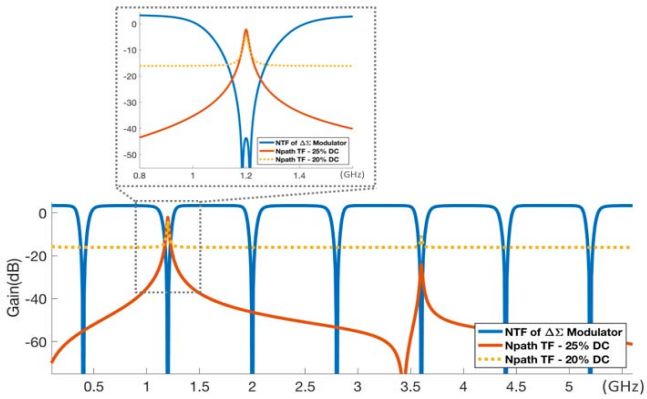


Fig. 6. Fourth-order inverse Chebyshev NTF and N-path filter frequency responses with different N-path duty cycles (dc).

mismatch. The use of a minimal resolution DAC, driven by a $\Delta\Sigma$ modulator, relaxes the DAC linearity problem and greatly reduces analog complexity [10]. Noise-shaping and oversampling with the low-resolution DAC deliver the equivalent in-band precision of a high-resolution DAC. The OSR is the ratio between the sampling rate and signal bandwidth.

We take advantage of the inherent linearity of a three-level or 1.5-bit DAC. In our work, we use a 1.5-bit DAC, driven by a digital bandpass $\Delta\Sigma$ modulator for its very small area and immunity to mismatch. As in the case with a 1-bit DAC, a 1.5-bit DAC is inherently linear; however, a three-level quantizer generates less quantization noise than a two-level one, potentially providing a higher SNR. We also solve the problem of spurs caused by the low resolution (three-level) quantizer in the modulator by efficiently introducing near-Gaussian dither in the modulator. The dither scheme is elaborated in Section III-D.

Counteracting the shaped quantization noise of DAC with on-chip bandpass filtering is another highlight of our scheme. The $\Delta\Sigma$ modulator translates the 15-bit wide data stream into a three-level digital signal to drive the 1.5-bit DAC. The modulator shapes quantization noise to sides of the RF band centered at f_c, thereby providing a high SNR within the RF bandwidth. We add an N-path filter to reduce out-of-band quantization noise. The NTF of the modulator and the transfer function of the N-path filter are complementary. For this reason, an N-path filter is a natural match for the bandpass modulator since it can remove the shaped quantization noise produced at the output of the modulator. The transfer function of the N-path filter is tailored to the NTF. Although the N-path filter generates harmonics at multiples of the center frequency, the $\Delta\Sigma$ modulator also generates NTF notches in the corresponding image bands, as shown in Fig. 6. A further advantage is that both the modulator and N-path clocking are derived from the same clock and inherently track the center frequency. An N-path filter is also small and efficient and does not require inductors, which are needed in other on-chip RF filters.

C. Specification Targets

Although this work is not specifically designed to meet a particular standard, it does meet the specifications of some of the local thermal equilibrium (LTE) and WiFi standards. The frequency range is designed to be 0.8–1.2 GHz, and the RF bandwidth is 40 MHz, which matches intermediate LTE and WiFi specifications. For LTE, an 8% in-band transmit EVM requirement is stated for 64-quadratic-amplitude modulation (QAM) for the wide area base station according to TS36.104 [11]. For WiFi, the 64-QAM EVM requirement is -22 dB, or 7.94% [12], which is close to LTE specification.[4]

Calculating using $\mathrm{SNR} = -[3.7 + 20 \times \log10\,({}^{\mathrm{EVM}}/_{100\%})]$ for 64-QAM, a minimal SNR of 18.3 dB is required for each independent element output. Choosing an OSR of 20, and three-level quantizer, gives considerable design margin. The peak in-band SNDR of our RF modulator is simulated to be 55 dB with a 700-mV peak-to-peak input amplitude, when the DAC is not loaded by the N-path filter.

As we discuss in Section III, the N-path filter provides some suppression of fundamental tone. Although this tone is still significant, it can be easily further filtered since it is far away from the RF center frequency.

III. Implementation

We now discuss the processing stages in detail, beginning with the digital phase shifting.

A. Digital Phase Shifting

As illustrated in the system architecture (Fig. 3), the baseband I/Q transmit data are phase shifted by multiplying with programmable 5-bit I and Q weights. CWM is used to implement the digital phase shifting. CWM involves multiplying the 10-bit I and Q data streams with 5-bit coefficients. In general, a 5-bit weight resolution is sufficient for a precise beam direction since the quantization errors from different elements average leading to good beam accuracy in an eight-element system.[5] Even 3-bit coefficients are commonly used in industry for commercial products. In the prototype, the 5-bit coefficients are fully programmable over a serial peripheral interface (SPI).

We take advantage of simple processing and optimal frequency planning for compact, low-power beamforming DSP. Digital beamforming necessitates the placement of the phase shifter in the digital domain; however, we still have flexibility in the placement of the CWM within the digital transmit chain. We apply CWM at the 200-MS/s baseband rate and then upsample the phase shifted data streams. Simulations show that operating this phase shifting at $f_s/8$ (200 MS/s) consumes only 2 mW per element. This represent a 90% reduction in power consumption as simulations indicate a 20-mW power consumption if operated at f_s (1.6 GS/s).

B. Digital Linear Interpolation

Applying linear interpolation, we upsample the 200-MS/s phase shifted streams by 8 to the $f_s = 4/3 f_c$ frequency. We implement linear interpolation by first calculating the difference between adjacent samples by subtracting the value of the previous sample from the current sample. A single interpolation step is calculated by dividing this difference by

[4]Coding Rate = 2/3, for OFDM 802.11a/g legacy devices and 802.11n devices.

[5]The measured beam accuracy in [6] and [9] supports this argument.

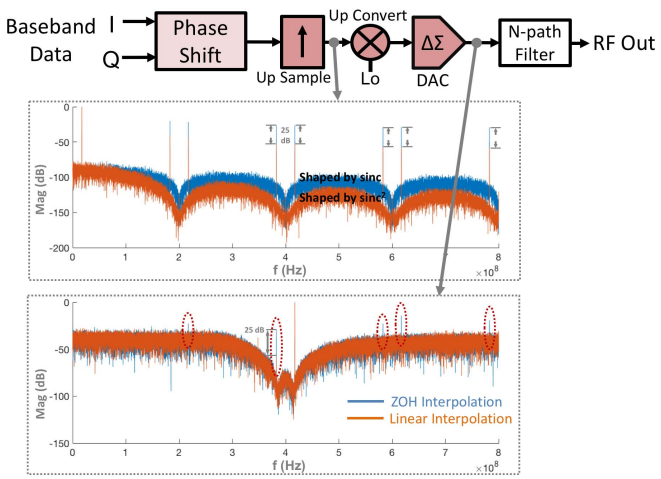


Fig. 7. Shaping effect and alias suppression of a conventional interpolator (sinc) compared with a linear interpolator (sinc^2).

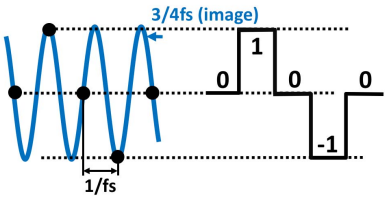


Fig. 8. Illustration of subsampling of the 3/4 f_s LO and the resulting state transition sequence of the digital LO (i.e., 1, 0, −1, 0,. . .).

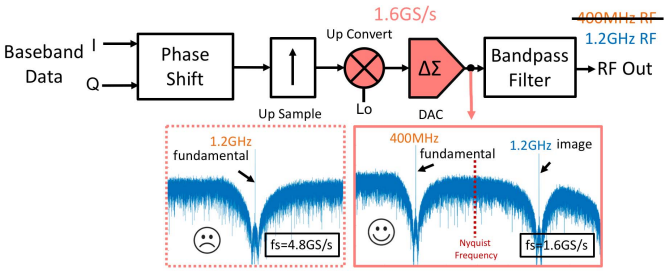


Fig. 9. Comparison of the output spectrum of the bandpass $\Delta\Sigma$ modulator operating at $f_s = 4.8$ GS/s and at $f_s = 1.6$ GS/s in a single stripe.

8. This division by 8 is simply implemented by 3-bit word shifting. Incremental addition of this single interpolation step yields the seven interpolated values.

Linear interpolation is advantageous over conventional upsampling using the zero-order hold technique. Interpolation with a zero-order hold results in aliases that are shaped by a sinc function [13]. Linear interpolation can be seen as a convolution of the signal with a triangular pulse, and we can show that the signal is enveloped by a sinc^2 function in the frequency domain. Therefore, aliases caused by upsampling are filtered by the sinc^2 transfer function. This significantly improves alias suppression. Fig. 7 shows the spectra with an 18-MHz input signal applied to the system. The spectra of the output of the interpolator and the output of $\Delta\Sigma$ modulator for a single stripe are shown to illustrate the shaping effect of different types of interpolation and the advantages of linear interpolation in alias suppression. In addition, linear interpolation only leads to a slight in-band attenuation. The total RF bandwidth of our design is 40 MHz, and therefore, the 20-MHz I or Q bandwidth is only 1/80 f_s so there is very little attenuation of the desired signal.

C. Digital Upmixing

Sampling at four times (4×) the output frequency greatly simplifies digital upconversion but the high sampling speed is challenging to implement. When the sampling rate is four times of the RF center frequency, we can represent the digital sinewave LO sequence as $0, -1, 0, 1, \ldots,$ making LO multiplication trivial. This digital-domain I/Q mixing eliminates I/Q mismatch problems. However, for the desired 1.2-GHz RF signal frequency, this would lead to a sampling rate of 4.8 GS/s. Such a high DSP speed is a great challenge for digital beamform processing, especially for the digital $\Delta\Sigma$ modulator. In fact, the digital $\Delta\Sigma$ modulator becomes a critical path for DSP, because the digital filter within the modulator requires high-resolution multiplications. This 4.8-GS/s speed is very challenging in 40-nm CMOS and is not practical for Verilog-based design. The power consumption of the DSP would also be very high—at least 100 mW per channel.

Instead, we operate at one-third of the 4× sampling frequency (i.e., $f_s = 4/3 f_c$) and make use of the image to produce the RF output. As indicated by sample dots in Fig. 8, with the reduced sampling rate, the sampled digital LO sequence can still be represented as $0, 1, 0, -1, \ldots,$. We exploit the image frequency at $3/4 f_s$, which is shown overlaid as a sine wave on the same plot. The output spectrum of the $\Delta\Sigma$ DAC is shown in Fig. 9. For a sampling rate of 4.8 GS/s, the fundamental output frequency is 1.2 GHz. Instead, operating at 1.6 GS/s, the fundamental output frequency of the modulator is at 400 MHz and its image is at 1.2 GHz. Using the image allows us to run the upconverter and the modulator at 1.6 GS/s, consuming only 16 mW per element and saving more than 85% in digital power consumption.

We digitally upmix the phase shifted, upsampled baseband signals to the RF frequency of 1.2 GHz by multiplying the upsampled I and Q signals with digital I and Q LO sequences. We emphasize that the relationship between the sampling frequency f_s and the RF carrier frequency ($f_c = 3/4 f_s$) allows the digital LO to be represented by the same sequence as when $f_c = 1/4 f_s$. Then, as shown in Fig. 10, we sum the upmixed I and Q sequences to form the digital RF output. In practice, we can further simplify this implementation because the I and Q LO sequences are orthogonal and only one is nonzero at any time. Ignoring the 0 values, the resulting summed-up digital RF sequence is equivalent to MUXing between the I and Q streams and alternately multiplying by 1 and −1. This process greatly simplifies the upmixing circuitry and reduces hardware complexity.

It is worth mentioning that mixing I and Q baseband data with orthogonal I and Q LO sequences, respectively, and then adding them together, eliminates the high-frequency

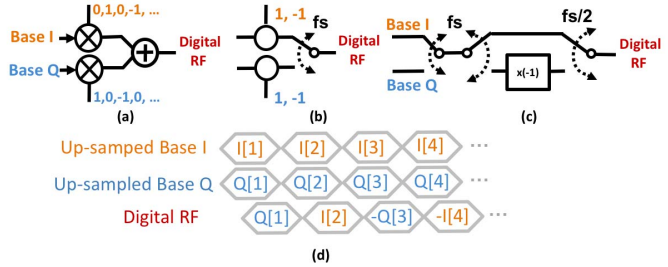


Fig. 10. (a) Digital upconversion and summing. (b) Simplified with MUXing. (c) Final implementation. (d) Timing diagram.

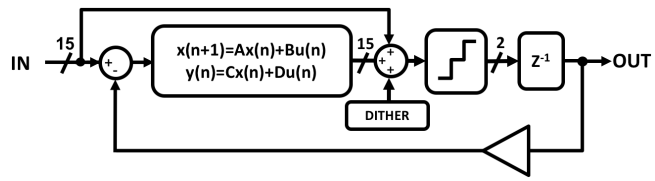


Fig. 11. Block diagram of the fourth-order digital bandpass $\Delta\Sigma$ modulator.

image. This single sideband (SSB) modulation reduces power consumption and is bandwidth efficient.

Using the image in the second Nyquist zone attenuates the signal power by around 12 dB, and this is the primary disadvantage of this approach. As an example in [7], a 12-dB in-band power degradation and a 9-dB maximum ACPR degradation are observed in measurement. However, our design is not intended for direct RF power generation, and a preamplifier is required.

D. $\Delta\Sigma$ Modulator

In our approach, bandpass digital $\Delta\Sigma$ modulation is vital in transforming the 15-bit word-length of upmixed digital RF signal to the 1.5-bit digital signal needed to drive the 1.5-bit DAC. In this way, we can use a 1.5-bit DAC and get the same in-band SNR ratio as with a higher resolution (and larger) conventional DAC. The digital bandpass modulator provides noise-shaped three-level copies of the digital signal at 1/4 and 3/4 of the sampling rate, as already shown in Fig. 9. As discussed, our choice of sampling frequency greatly simplifies the upconversion mixer.

The bandpass $\Delta\Sigma$ modulator is formed as a single loop, with a single feedback path and a feedforward path around a fourth-order Chebyshev band reject NTF filter, as shown in Fig. 11. The fourth-order digital filter is an integration of four independent feedback loop and four feedforward path in the canonical form with single time delay. The additional feedforward path around the digital filter is designed to ensure that the signal transfer function (STF) is 1, as illustrated in Fig. 12.

Since the stability of the modulator is sensitive to the peak out-of-band gain, choosing a filter that is maximally flat out-of-band allows the greatest in-band noise reduction without instability. This, combined with an even distribution of in-band

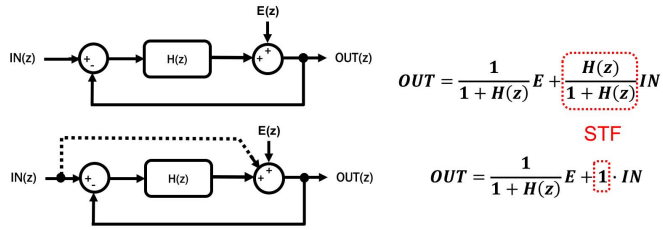


Fig. 12. Effect of feedforward path on STF.

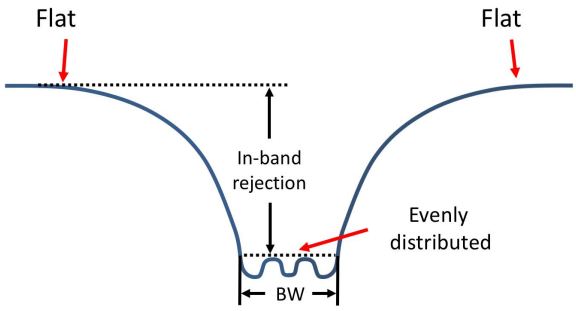


Fig. 13. General inverse Chebyshev band-reject NTF.

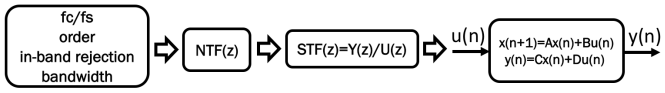


Fig. 14. Design flow for the digital filter of the digital bandpass $\Delta\Sigma$ modulator.

noise, makes the inverse Chebyshev (aka Chebyshev Type II) the most spectrally efficient choice, as shown in Fig. 13.

We now describe the design flow for the fourth-order Chebyshev digital filter. Starting with the in-band signal attenuation and the order of the filter, one can derive the equivalent bandwidth. The center frequency (f_c) and the sampling frequency are scalable with $f_c/f_s = 3/4$ as mentioned for the simplest digital upconversion. In our case, for a fourth-order Chebyshev NTF, an in-band rejection of 47 dB and the sample-to-signal ratio yield a bandwidth of 38 MHz, which meets our specification. We place the poles and zeros of the inverse Chebyshev NTF based on this information. The STF without the feedforward path is derived from the NTF using $\mathrm{STF}(z) = ((1/\mathrm{NTF}) - 1) \cdot z$. The matrix values in the fourth-order discrete state space (DSS) are formulated from the STF using the Matlab ssdata function. This design process is readily configurable in Matlab and the DSS model of the digital filter can be easily implemented in Verilog, as shown in Fig. 14.

We add dither at the output of the digital filter before the signal is quantized to improve the performance of the bandpass modulator [14]. Generally, in the design of a $\Delta\Sigma$ modulator, the quantizer is modeled as additive white Gaussian noise that is spread across the entire frequency spectrum. However, this assumption is only true if the quantizer resolution is large. When the resolution of the quantizer is low (in our case, only 1.5 bits), the low quantizer resolution leads to spurious tones at certain frequencies. To eliminate these spurs, we add one

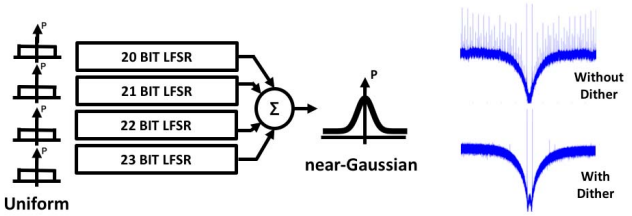


Fig. 15. Generation of near-Gaussian dither and the improvement in the bandpass modulator spectrum thanks to the introduction of dither.

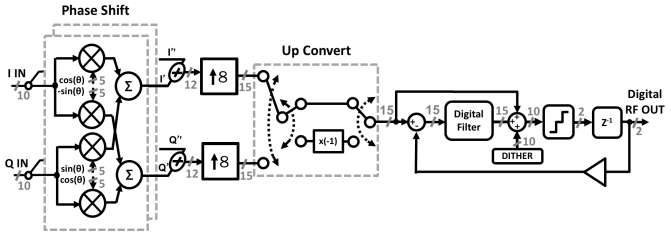


Fig. 16. Detailed implementation of the digital circuitry for a single stripe annotated with the word-length allocation.

LSB wide near-Gaussian dither to the input of the quantizer to enforce the additive white Gaussian noise assumption. Since we have a high effective resolution at the output, the benefit of adding a dither far outweighs the cost of adding random LSB-level noise.

We efficiently generate near-Gaussian random dither on-chip by summing the outputs of four 20b–23b linear-feedback shift registers (LFSRs), as shown in Fig. 15. Each LFSR generates uniformly distributed noise. Each LFSR is different in length to ensure that they are not correlated with each other. When four uncorrelated uniformly distributed noise signals are added, a near-Gaussian noise is derived. The digital near-Gaussian noise is scaled to ensure that the variance of the dither added to the circuit is around $\mathrm{LSB}^2/12$, which is equivalent to 1 LSB of quantization noise.

E. Digital Word-Length Allocation

Word-length allocation optimization is an important factor in designing a precise and efficient digital circuit to satisfy the desired function. The digital filter in the $\Delta\Sigma$ modulator requires a high resolution but this digital filter is also a bottleneck limiting the sampling rate of the whole signal processing chain, so word length must be carefully managed. Fig. 16 shows the word length allocation throughout a single transmit stripe.

In our design, the baseband digital I and Q inputs are 10-bits wide—the MSB is the sign bit, MSB-1 is the integer bit, and the remaining bits are fractional bits. All signals are in two's complement signed-value representation. A saturation check feature in the phase shift block deals with unexpected large signals. If the MSBs of the internal signals (represented with 15 bits) are not identical, the signal saturates to a designed threshold value. In the interpolator, the signal is upsampled by 8 and linearly interpolated, which naturally yields three

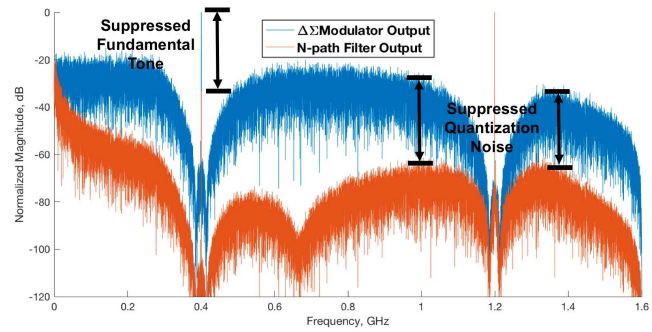


Fig. 17. Spectra of the output of $\Delta\Sigma$ modulator and output of the N-path filter.

additional fractional bits. The signals are all scaled by 4 to maintain the input of the $\Delta\Sigma$ modulator in the desired range. In all, 5 more fractional bits are added to yield 13 fractional bits. This value is selected considering both the requirement for precision in implementing the desired NTF function for the $\Delta\Sigma$ modulator and to meet the combinational logic timing limitations of 40-nm CMOS. This word length restricts the speed of the critical path in the digital filter, which is also the critical path of the whole digital chain.

The implementation of the upconverter is a simple MUX selecting from the I or Q data stream and multiplying by 1 or -1 and does not increase the word length. The hardware consumption of the upconverter is kept to the minimum. For the $\Delta\Sigma$ modulator, with a 15-bit input and reasonable approximation of the matrix value in the fourth-order DSS, the NTF shape is generated without sacrificing much precision and the $\Delta\Sigma$ modulator block can still run at gigahertz sampling rates. The output of the digital filter, dither, and fed-forward input is added together and quantized by a three-level quantizer. A 2-bit word length represents the final 1.5-bit digital RF output of the modulator.

F. DAC and N-Path Filter

The upmixed, noise-shaped 1.6-GS/s 1.5-bit digital signal directly drives the 1.5-bit DAC to form the 1.2-GHz RF signal. An N-path filter in each transmit stripe suppresses the shaped quantization noise caused by the digital $\Delta\Sigma$ modulator, as illustrated in Fig. 17. The N-path filter also suppresses the 400-MHz fundamental tone more than 30 dB with ideal duty cycle and zero switch resistance. The combination of a 1.5-bit $\Delta\Sigma$ DAC and N-path filtering leads to a small and efficient DAC because of the small size of 1.5-bit current steering DAC compared with a high-resolution DAC. Compared to a single-bit DAC, a three-level DAC generates less quantization noise and thanks to its single current source, it still avoids the problem of unit mismatch seen in higher resolution DACs. The digital bandpass modulator and 1.5-bit DAC also ensure flexible center frequency and direct modulation to RF without need for analog upconversion.

An N-path filter also contributes to the compactness and efficiency of analog circuitry since an N-path filter does not need inductors as in some RF bandpass filters. An N-path filter requires only switches, resistors, and a capacitor array.

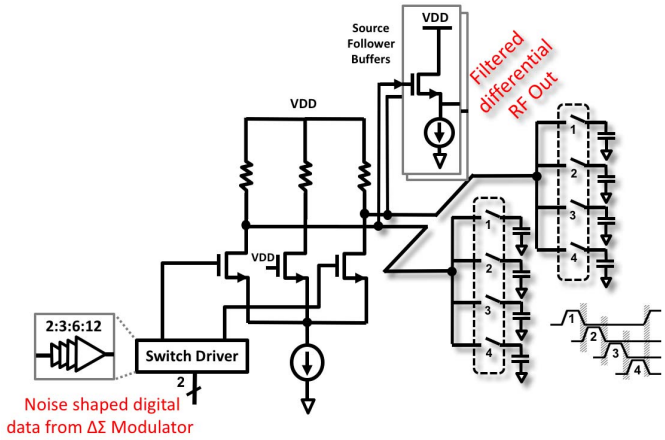


Fig. 18. Circuit implementation of the 1.5-bit DAC and N-path filter.

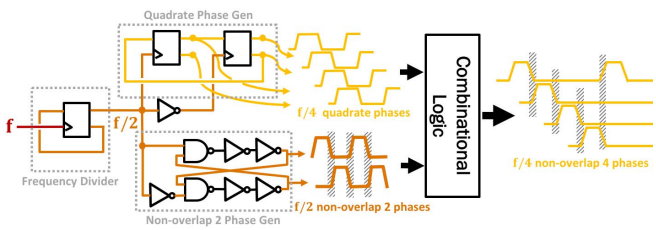


Fig. 19. Implementation of four-phase nonoverlap clock generator.

An N-path filter is a compact, highly configurable, high-Q bandpass filter that tracks the center frequency of the modulator and effectively suppresses the shaped quantization noise.

Fig. 18 shows the N-path filter and DAC in detail. The DAC is directly driven by the three-level 1.6-GS/s digital signal from the bandpass $\Delta\Sigma$ modulator. The switch driver is implemented as cascaded buffers with increasing sizes. The three-level signal is represented with 2 bits as 01, 00, and 10 for the high, middle, and low levels, respectively. The three-level DAC has a single current source and three switches connecting to the two differential load resistors and the dummy load. The dummy branch ensures that there is always a current path from the source. Filter capacitors are connected sequentially to the load resistors to provide a filter passband centered on the 1.2-GHz RF carrier, which is a quarter of the N-path filter clock rate. There are two copies of the filter capacitor array since it is fully differential.

The load resistors are shared by the DAC and N-path filter. The load resistance is 1 kΩ and the filter capacitors are 2-pF MOM caps, both integrated on-chip. Theoretically, this RC value sets the bandwidth $f_{rc} = 1/(2\pi RC) = 79$ MHz. Using the more accurate four-path filter model [15], including other nonidealities, such as reduction in duty cycle and switch resistance, a periodical ac analysis simulation predicts a 3-dB filter bandwidth of approximately 40 MHz, which matches the bandwidth of digital $\Delta\Sigma$ modulator.

The four-phase 1.2-GHz nonoverlapping clocks for the N-path filter are derived from a 4.8-GHz clock using a frequency divider and simple combinational logic, as shown in Fig. 19. The 4.8-GHz clock is three times of the modulator sampling rate, and clocking is shared across the transmit array

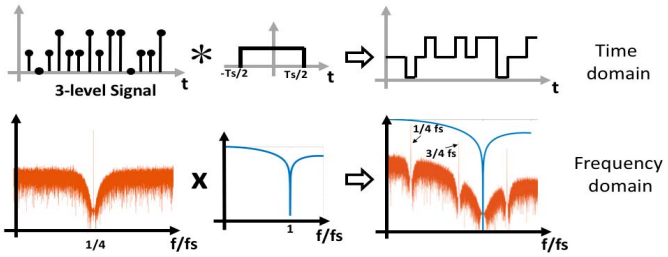


Fig. 20. Shaping effects of DAC in the frequency domain.

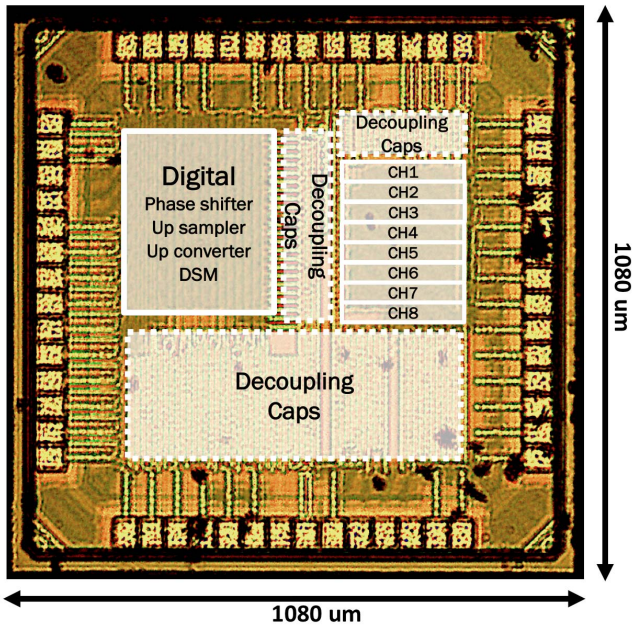


Fig. 21. Die micrograph.

since the clock of N-path filter inherently tracks the modulator sampling frequency. This tracking of the RF center frequency provides better filtering performance and also significantly improves the beam accuracy compared with fixed-center-frequency passive filters.

A current-steering DAC does not have a rail-to-rail output swing. However, for a three-level signal, current steering is still advantageous because it makes use of a third dummy branch and avoids the mismatch inherent in a DAC with multiple cells. As shown in Fig. 20, the conventional nonreturn-to-zero type of DAC has a shaping effect on the original signal. By holding the digital sampled value during the whole sampling period, the signal is convolved with a single square pulse of width $1/f_s$. In the frequency domain, the signal is shaped by $\sin c(\pi f/f_s)$, as also shown in Fig. 20. At the image center frequency of 3/4 f_s, the signal is attenuated by 12 dB. This power loss is a drawback of using the image, but it facilitates our mostly digital approach. As an alternative, the use of a mixing DAC [16] ensures less signal loss by moving the maximum power region to the second Nyquist zone.

For an N-path filter, the finite number of paths (in our case, four) and reduction in duty cycle all contribute to loss [17]–[19].

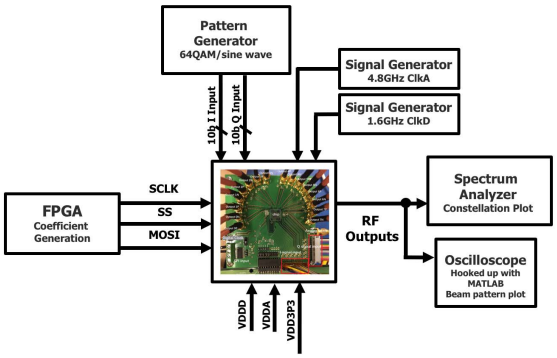


Fig. 22. Test setup.

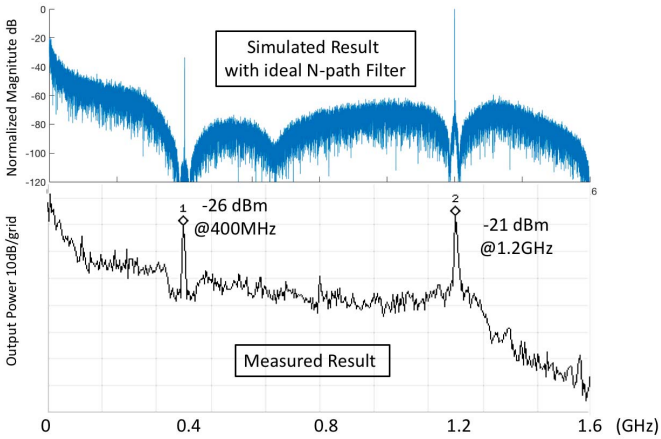


Fig. 23. Comparison of simulated and measured output spectra.

With four paths, the theoretical loss is about 2 dB. Switch resistance reduces the maximum attenuation of the sideband, limiting suppression of the fundamental tone. We use low V_T devices to keep the switch ON-resistance low. A source follower at the output of each stripe, also shown in Fig. 18, ensures 50-Ω output matching.

IV. Measurement Results

A. Die Photograph and Test Setup

The prototype is implemented in 40-nm CMOS, and a die photograph is shown in Fig. 21. The active area is 0.19 mm^2 and the digital core size is 0.12 mm^2. The overall power consumption is 128 mW for a 1.2-GHz output, corresponding to 16 mW per channel. The test setup is shown in Fig. 22.

B. Single-Channel Spectrum

Fig. 23 shows the comparison of the simulated and measured output spectra. The simulated result is based on an ideal N-path filter with ideal duty-cycle switching and ideal zero-resistance switches. Also, this simulation does not consider the power loss caused by the output matching circuit. From the measured spectrum, we can notice the shaped noise around the fundamental tone, while for the image tone, where we apply the N-path filter, we see that the quantization noise is suppressed. However, it is noteworthy that the noise-shaping around the fundamental is not significant since it is attenuated by the N-path filter. Since the power of the fundamental tone is originally 12 dB higher than the image tone, the suppressed fundamental tone is still significant.[7]

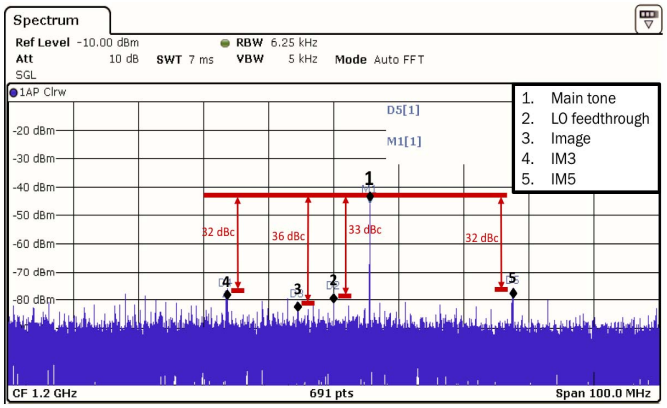


Fig. 24. Zoomed-in measured output spectrum.

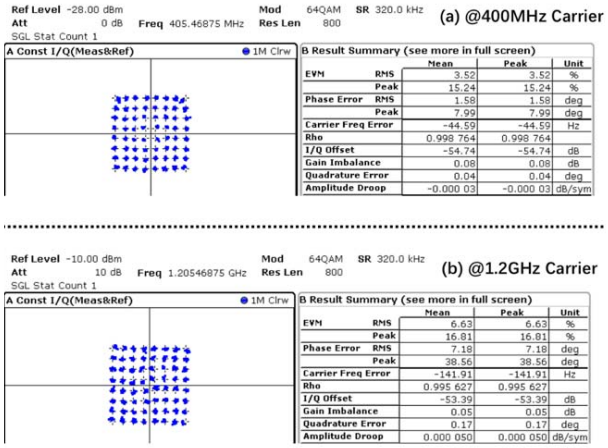


Fig. 25. 64-QAM constellation plot with 320k symbol/s rate, EVM = 3.52% at 400-MHz carrier (top) and EVM = 6.63% at 1.2-GHz carrier (bottom).[6]

Fig. 24 shows a zoomed-in spectrum around the 1.2-GHz carrier frequency. The main tone is at 1.2055 GHz. The LO feedthrough, image, IM3, and IM5 are marked in the figure. The measured IM3 is 32 dBc, which corresponds to an in-band SFDR of 32 dB. The measured image rejection is 36 dBc. The estimated SNR is 24 dB.[8]

C. Modulation and Beamforming Test

Fig. 25 shows the measured constellation when a 64-QAM modulated digital baseband signal is applied to the prototype. The constellation plot is generated from the output of a single channel. EVM is further improved when combining the outputs of the eight channels because uncorrelated noise from different channels is averaged out. The digital transmit data are filtered with a Gaussian filter with a rolloff factor of 0.22. With a 64-QAM signal at 1.205 GHz, an EVM of −23.6 dB is measured for a symbol rate of 320 kSymbols/s. For the same conditions, the measured EVM is −29.0 dB for a 405-MHz carrier.

[6]The measured EVM values are primarily limited by the output power level. The prototype does not include the PA design.

[7]The internal digital noise-shaped signals are not available for measurement since the digital signals are not driven off chip.

[8]For 64 QAM, this is equivalent to a theoretical EVM of 4.1%.

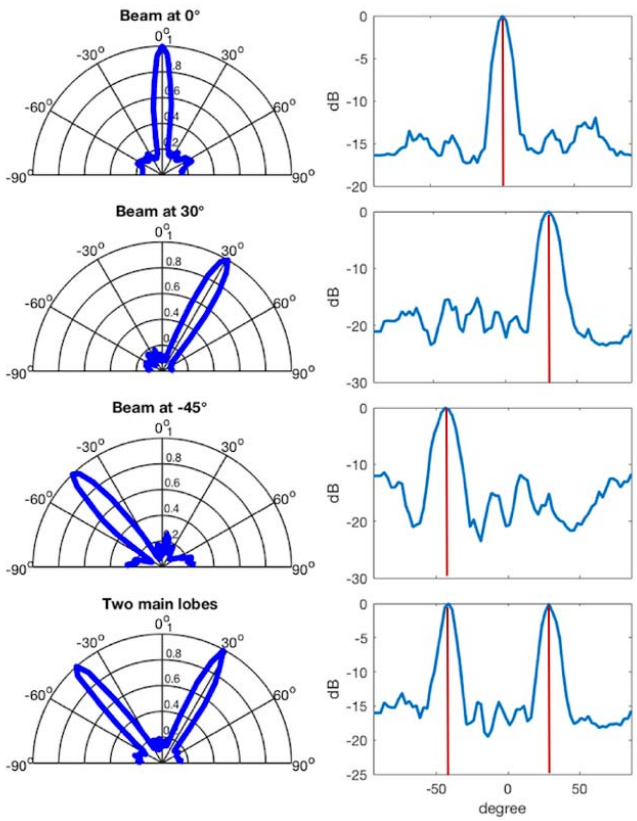


Fig. 26. Measured transmit beam patterns at 0°, 30°, and −45°, and with two simultaneous main lobes. The ρ-axis of the polar plot represents the normalized amplitude.

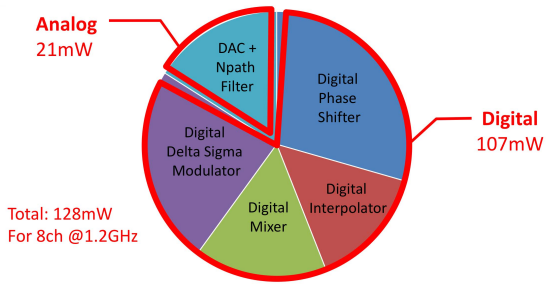


Fig. 27. Power consumption breakdown.

The measured beam patterns for 0°, 30°, and 45° and for two separate beams are shown in Fig. 26. These measurements assume eight in-line dipolar antennas, spaced at half wavelengths. The measurement step size is 2.5°. These measurements show good beam patterns. The side lobes are around −15 to −12 dB with various steering angles (the ideal is −12 dB). The beam direction is accurate within 2°, as indicated in Fig. 26 where the ideal beam direction marked in red.

A power breakdown is given in Fig. 27. Except for the DAC and N-path filter, all the other blocks are digital and these digital blocks consume most of the total power. It is also clear that the digital phase shifting is a significant part of the power consumption, highlighting the advantage of reducing the digital phase shifting frequency by a factor of eight, as well as the benefits of our approach to frequency planning.

TABLE I

DESIGN SUMMARY

Architecture	Digital transmit beam forming
Technology	40-nm CMOS
RF bandwidth	40 MHz
Carrier frequency	0.8-1.2 GHz
Input data	200 MS/s 10-bit I/Q baseband
Beamforming	Eight channels
Coefficient resolution	10 bits, 5-bit I and Q
# of simultaneous beams	2
EVM for 64 QAM	3.5% @400 MHz, 6.6% @ 1.2 GHz
Area	0.19 mm^2 (0.02 mm^2 /channel)
Power consumption	83 mW @ 0.8 GHz (**10 mW / channel**) 128 mW @ 1.2 GHz (**16 mW / channel**)

TABLE II

COMPARISON WITH PREVIOUS WORK

	[3]	[7]	[20]	This work[a]
ΔΣ Type	2nd order BP	3rd order LP	2nd order MASH	4th order BP
Carrier	900 MHz	1.95 GHz	5.25 GHz	1.2 GHz
BW [MHz]	20	50	200	40
EVM/SNR	36 dB	3.42%	30 dB	6.6%
Peak output power	3 dBm	-15.8 dBm	-13.73 dBm	-21 dBm
Power consumption	150 mW	69 mW	120 mW (digital Modulator only)	16 mW
Process	28 nm	90 nm	130 nm	40 nm
Area	0.82 mm^2	0.15 mm^2	0.16 mm^2 (digital Modulator only)	0.02 mm^2
Power amplification	1.5V MS DAC	Output stages	DRFC (2.5V)	No
BPF?	no	no	LC	N-path

The simple implementation of the digital upmixer ensures that it is a minor contributor to the total power consumption.

V. CONCLUSION

Table I summarizes the measured performance. Table II shows a comparison with previous works. Similar to our RF modulator, Roverato *et al.* [3] and Frappe *et al.* [7] do not include a PA. For [20], the power consumption and area of the digital-RF converter (DRFC), in which the output driver is integrated, are excluded to make the comparison meaningful.

The die area per antenna element of our work is an order of magnitude smaller than the state of the art. Apart from the DAC and the N-path filter, the design is generated with digital synthesis and place and route tools. This work demonstrates the promise of digital techniques for highly compact, flexible, and efficient direct-to-RF digital beamforming.

ACKNOWLEDGMENT

The authors would like to thank Dr. L. Sattler for her advice on Gaussian noise generation. The authors would also like to thank Dr. H.-S. Kim for discussions on signal modulation and demodulation and the anonymous reviewers for very helpful comments and suggestions.

REFERENCES

[1] S. Balasubramanian *et al.*, "Ultimate transmission," *IEEE Microw. Mag.*, vol. 13, no. 1, pp. 64–82, Jan./Feb. 2012.

[2] P. Eloranta, P. Seppinen, S. Kallioinen, T. Saarela, and A. Parssinen, "A multimode transmitter in 0.13 μm CMOS using direct-digital RF modulator," *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2774–2784, Dec. 2007.

[3] E. Roverato *et al.*, "All-digital RF transmitter in 28 nm CMOS with Programmable RX-band Noise Shaping," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, Feb. 2017, pp. 222–223.

[4] P. E. P. Filho, M. Ingels, P. Wambacq, and J. Craninckx, "A 0.22 mm^2 CMOS resistive charge-based direct-launch digital transmitter with-159 dBc/Hz out-of-band noise," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, Jan./Feb. 2016, pp. 250–252.

[5] J. Sommarek, J. Vankka, J. Ketola, J. Lindeberg, and K. Halonen, "Digital modulator with bandpass delta-sigma modulator," in *Proc. 30th Eur. Solid-State Circuits Conf.*, Leuven, Belgium, 2004, pp. 159–162.

[6] B. Zheng, J. Bell, Y. He, L. Jie, and M. Flynn, "A 0.19 mm^2 128 mW 0.8-1.2 GHz 2-beam 8-element digital direct to RF beamforming transmitter in 40 nm CMOS," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Philadelphia, PA, USA, Jun. 2018, pp. 128–131.

[7] A. Frappe, A. Flament, B. Stefanelli, A. Kaiser, and A. Cathelin, "An all-digital RF signal generator using high-speed $\Delta\Sigma$ modulators," *IEEE J. Solid-State Circuits*, vol. 44, no. 10, pp. 2722–2732, Oct. 2009.

[8] S. K. Garakoui, "Wideband RF beamforming: Architectures, time-delays and CMOS implementations," Universiteit Twente, Enschede, The Netherlands, Tech. Rep., 2015. doi: 10.3990/1.9789036539012.

[9] S. Jang, J. Jeong, R. Lu, and M. P. Flynn, "A 16-element 4-beam 1 GHz IF 100 MHz bandwidth interleaved bit stream digital beamformer in 40 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 53, no. 5, pp. 1302–1312, May 2018.

[10] S. Luschas, R. Schreier, and H.-S. Lee, "Radio frequency digital-to-analog converter," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1462–1467, Sep. 2004.

[11] *Technical Specification Group Radio Access Network; Base Station(BS) Radio Transmission and Reception*, document Release 11.2; 3GPP TS 36.104 V11.2, 2012.

[12] *Comments on EVM and Required Study*, M. A. Rahman, Standard IEEE 802.15-08/554r1, 2008.

[13] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and Systems*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

[14] S. R. Norsworthy, "Effective Dithering of Sigma-delta Modulators," in *Proc. IEEE Int. Symp. Circuits Syst.*, San Diego, CA, USA, vol. 3, May 1992, pp. 1304–1307.

[15] A. Ghaffari, E. A. M. Klumperink, M. C. M. Soer, and B. Nauta, "Tunable high-Q N-path band-pass filters: Modeling and verification," *IEEE J. Solid-State Circuits*, vol. 46, no. 5, pp. 998–1010, May 2011.

[16] C. Erdmann *et al.*, "A 330 mW 14b 6.8 GS/s dual-mode RF DAC in 16 nm FinFET achieving-70.8 dBc ACPR in a 20 MHz channel at 5.2 GHz," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, Feb. 2017, pp. 280–281.

[17] M. Darvishi, R. van der Zee, E. A. M. Klumperink, and B. Nauta, "Widely tunable 4th order switched G_m-C band-pass filter based on N-path filters," *IEEE J. Solid-State Circuits*, vol. 47, no. 12, pp. 3105–3119, Dec. 2012.

[18] A. Mirzaei, H. Darabi, and D. Murphy, "Architectural evolution of integrated M-phase high-Q bandpass filters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 1, pp. 52–65, Jan. 2012.

[19] S. Hameed, M. Rachid, B. Daneshrad, and S. Pamarti, "Frequency-domain analysis of N-path filters using conversion matrices," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 1, pp. 74–78, Jan. 2016.

[20] A. Jerng and C. G. Sodini, "A wideband $\Delta\Sigma$ digital-RF modulator for high data rate transmitters," *IEEE J. Solid-State Circuits*, vol. 42, no. 8, pp. 1710–1722, Aug. 2007.

Boyi Zheng received the B.S. degree in microelectronics from Shanghai Jiao Tong University, Shanghai, China, in 2016, and the M.S. degree in electrical and computer engineering from the University of Michigan, Ann Arbor, MI, USA, in 2018, where she is currently pursuing the Ph.D. degree.

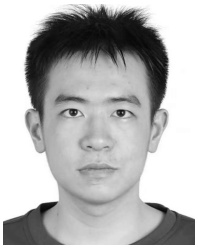
Her current research interests include digital intensive communication systems, and digital direct to RF and beamforming techniques.

Ms. Zheng was a recipient of the Analog Device Outstanding Student Designer Award in 2017.

Lu Jie was born in Guangzhou, China, in 1995. He received the B.Eng. degree in electrical and electronic engineering from Zhejiang University, Hangzhou, China, in 2017. He is currently pursuing the Ph.D. degree at the University of Michigan, Ann Arbor, MI, USA, with a focus on hybrid architecture ADC design.

John Bell received the B.S.E. degree in electrical engineering and the B.S. degree in sound engineering from the University of Michigan, Ann Arbor, MI, USA, in 2010, where he is currently pursuing the Ph.D. degree.

From 2011 to 2014, he was a Product Engineer with Connor Winfield, while starting up a digital-violin synthesis company called Weinreich Labs and working as a Freelance Live Sound Engineer. His current research interests include analog/mixed-signal circuits and sigma–delta modulation.



Yan He (M'18) received the B.S degree in electronic science and technology from Zhejiang University, Hangzhou, China, in 2018. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Rice University, Houston, TX, USA.

He was a Summer Intern with the University of Michigan, Ann Arbor, MI, USA, in 2017. His current research interests include analog and mixed-signal integrated circuit design.

Michael P. Flynn (M'95–SM'98–F'15) received the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA, in 1995.

From 1995 to 1997, he was a Technical Staff Member with Texas Instruments Incorporated, Dallas, TX, USA. From 1997 to 2001, he was with Parthus Technologies, Cork, Ireland. Since 2001, he has been a Professor with the University of Michigan, Ann Arbor, MI, USA. His current research interests include RF circuits, data conversion, serial transceivers, and biomedical systems.

Dr. Flynn is a 2008 Guggenheim Fellow. He formerly served on the Technical Program Committees for the Asian Solid-State Circuits Conference and the Symposium on VLSI Circuits. He serves on the Technical Program Committees for the European Solid-State Circuits Conference and the International Solid-State Circuits Conference. He was a recipient of the 2016 University of Michigan Faculty Achievement Award, the 2011 Education Excellence Award, the 2010 Ted Kennedy Family Team Excellence Award from the College of Engineering, University of Michigan, the 2005–2006 Outstanding Achievement Award from the Department of Electrical Engineering and Computer Science, University of Michigan, the 2004 NSF Early Career Award, and the 1992–1993 IEEE Solid-State Circuits Pre-Doctoral Fellowship. He was the Editor-in-Chief of the IEEE JOURNAL OF SOLID-STATE CIRCUITS from 2013 to 2016. He has served as an Associate Editor for the IEEE JOURNAL OF SOLID-STATE CIRCUITS and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS. He is a former Distinguished Lecturer of the IEEE Solid-State Circuits Society.