Low-Resolution Architectures for Power-Efficient Scaling of mmWave Phased Array Receivers

Maryam Eslami Rasekh*, Navid Hosseinzadeh†, Upamanyu Madhow‡, Mark Rodwell§
Department of Electrical and Computer Engineering
University of California Santa Barbara
Email: {*rasekh, †navid, ‡madhow, §rodwell}@ece.ucsb.edu

Abstract—Utilizing mmWave massive MIMO frontends for base station to mobile communication promises unprecedented throughput gains in cellular networks. Power efficiency is a significant bottleneck in scaling to the large array sizes required for closing the link at high frequencies, particularly in the battery-powered handset. Conventional phased array architectures at 100+ GHz require close to 100 mW of power per receive antenna, making them ill-suited for scaling to massive arrays. In this paper, we propose drastically simplified receiver frontend designs that are predicted to slash this per-channel wattage by an order of magnitude or more at 140 GHz. These power savings come at the cost of imperfect beamforming and under-utilization of the array elements and die area. We quantify these tradeoffs for our baseline on-off architecture, as well as 1-bit and 2-bit phase switched arrays. In the lowest resolution setting, on-off beamforming, we show that beamforming efficiency of 70% is achieved with 40% utilization of the aperture. This tradeoff improves significantly with the adoption of even the most coarse (1 or 2 bit) phase control, motivating us to pursue power-optimized designs for low-precision mmWave phase shifters.

Index Terms—Low-power frontends, millimeter wave, Terahertz, massive MIMO, phased array antennas, analog beamforming, on-off beamforming.

I. Introduction

The ever increasing demand for mobile data has driven industry towards utilization of higher frequency bands in the millimeter wave (mmWave) and Terahertz (THz) range, where large available bandwidth and immense potential for spatial multiplexing provide the required orders of magnitude increase in throughput [1]. At these high frequencies, electronically large yet physically small antenna arrays may be deployed to overcome the higher path loss and atmospheric losses in the channel. Fixing the 2-dimensional aperture size (and, therefore, die area), and the $\lambda/2$ (half wavelength) spacing, the number of antenna elements scales quadratically with the carrier frequency. Consequently, sub-THz transceivers with hundreds or even thousands of elements easily fit in small form factors such as that of a hand-held device or compact picocellular base station. A fully digital or hybrid array architecture may be more suitable for the base station to allow spatial multiplexing between different mobile users, whereas analog beamforming with a phased array is sufficient for a handset forming a single, narrow beam towards the base station.

Power consumption is a major bottleneck in the deployment of THz frontends when the number of array elements grows large [2]–[5], especially in the battery-powered handset, which
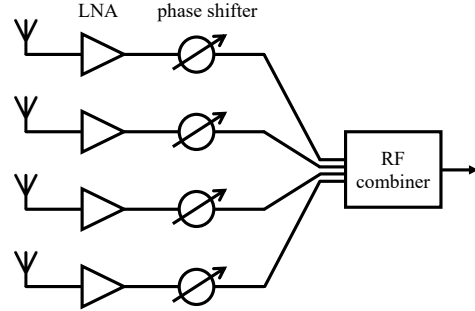

Fig. 1. The conventional phased array architecture.

is our focus in this paper. In a conventional phased array architecture, depicted in Figure 1, the signal of each antenna goes through a low noise amplifier (LNA) and phase shifter (passive or active) to provide sufficient SNR and phase alignment in each channel before summation in the combiner. The amount of power burned in an LNA increases proportionally with its gain. If high gain is required in the LNA (e.g., to compensate for a noisy phase shifter), this consumption quickly adds up as the array grows large. In this paper, we propose alternative architectures that allow significant reduction in the required LNA gain, providing an order of magnitude or more reduction in per-channel power consumption relative to the state of the art for 140 GHz phased array receivers [6], [7].

Figure 2 shows three such architectures. The first is an on-off architecture wherein phase shifters are replaced with a simple switch on each channel that can turn that channel ON or OFF, without the possibility of tuning its phase. We argue that by taking advantage of the abundance of antennas, sufficient beamforming gains for closing the link can be achieved by activating only a *subset* of channels whose signals happen to be sufficiently aligned in phase at the combiner. Thus, we invoke a core principle of massive MIMO which allows relaxing per-element sophistication by leveraging scale. We take this as our baseline architecture with the lowest power consumption per active channel.

We note, however, that there is room for improvement in the noise characteristics of phase shifters at 100+ GHz, as the LNA-phase shifter block has yet to be optimized for low power performance at THz frequencies. Our preliminary analysis
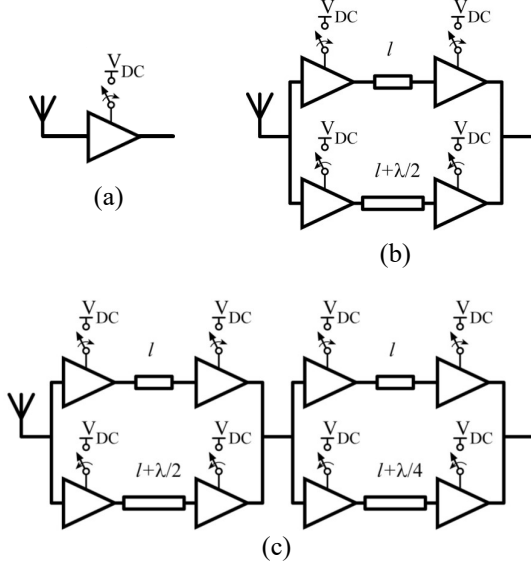
Fig. 2. Schematic of one channel in (a) the baseline on-off architecture, and benchmark architectures with (b) 1-bit, and (c) 2-bit phase control. Phase control is done by switching from the default path of length l to a path of length $l + \lambda/2$ or $l + \lambda/4$ to add a phase shift of $180°$ or $90°$, respectively.

shows that low-power designs that allow 1 or 2 bits of phase control in each channel, similar to those depicted in Figure 2b-c, may in fact be possible at 140 GHz. We therefore include, in this paper, an analysis of the power-utilization tradeoffs and beamforming performance of low-resolution phased arrays (with 1-bit and 2-bit phase control) as a benchmark for comparison with the proposed on-off architecture. It is interesting to note that a similar strategy is pursued for improving the power efficiency of *digital* arrays, namely, by employing low-resolution analog to digital converters (ADCs) to cut down the power consumption of each channel [8]–[10].

Contributions: We first propose a low-power on-off array architecture for reducing power consumption of phased array receivers at the cost of hardware redundancy and under-utilization of antenna die area. We provide a coarse comparison of power consumption in our architecture with that of a conventional phased array, as well as analytical and numerical assessment of the performance of this scheme, laying out the inherent power-redundancy tradeoff that it entails. Our calculations predict that 7X power savings (10X per-channel power reduction with 70% beamforming efficiency) can be attained with 40% utilization of die area or, equivalently, 2.5X hardware redundancy. We then provide benchmark comparisons with 1-bit and 2-bit phased array architectures, in hopes of motivating the design of low-power phase shifters optimized for mmWave and THz carrier frequencies. We show that the 1-bit architecture provides 80% beamforming efficiency with 60% utilization, while a 2-bit architecture promises 90% efficiency with 95% utilization, or 85% efficiency at full utilization.

II. SYSTEM MODEL

A conventional phased array architecture is depicted in Figure 1a. The signal of each antenna is amplified by the RF LNA before going through the phase shifter, such that the phase-aligned signals are combined constructively in the analog combiner. The phase shifters may be *active* or *passive*. Passive phase shifters, especially at mmWave and THz frequencies, impose considerable attenuation on the signal, placing more stringent gain and noise figure requirements on the LNA [11]. Active phase shifters, on the other hand, offload part of the amplification requirements from the LNA and, consequently, reduce LNA wattage. However, as active components, they consume additional power which adds up as the array size grows. To date, no CMOS prototypes of low-noise active phase shifters, or low-attenuation and hence low-noise *passive* phase shifters, have been demonstrated at 140 GHz. Most existing designs downconvert the RF signal to an intermediate frequency (IF) band where phase correction and combining is performed [6], [7], [12]–[15]. This approach requires incorporating an additional mixer and LO driver *on each channel* which further adds to the per-channel power consumption of the phased array. In fact, a similar dynamic arises with the RF to IF mixer: It is either passive and noisy, and drives up the LNA power consumption, or it is active and contributes to the power consumption of each channel. Thus, removing the phase shifter altogether has the potential to reduce power consumption significantly – by an order of magnitude or more – both directly by eliminating active components, and indirectly by relaxing the gain and noise figure requirements of the LNA.

We note that, in our analysis, we exclude the power consumption of components downstream of the combiner, i.e., "shared" components, and only account for the power consumption of components that are duplicated for each channel. Since we do not assume a specific array size, this approach simplifies our analysis and facilitates an informative comparison between the different architectures. Thus, by "per-channel power consumption" we refer to the consumption of LNAs, phase shifters, RF to IF converters, and any block that the signal passes *before* arriving at the combiner.

A. A Low-Power Architecture

Our proposed architecture, depicted in Figure 2a, replaces the phase shifter in each channel with a simple switch that, in the ON position, activates the channel and allows its signal to propagate to the combiner, and deactivates it when turned OFF. Deactivated channels draw near zero current while active channels consume a small amount of power, relative to a conventional phased array channel, enough to provide the gain required to overcome the loss of the combiner. We note that altering of the switch positions is done at the channel coherence rate which is much slower than the symbol rate, therefore switching speed is not a bottleneck for the system.

In the designs published to date at 140 GHz, the LNA typically provides 17 to 20+ dB of gain and burns around 30 to 60+ mW of DC power. The mixer and IF beamformer consume
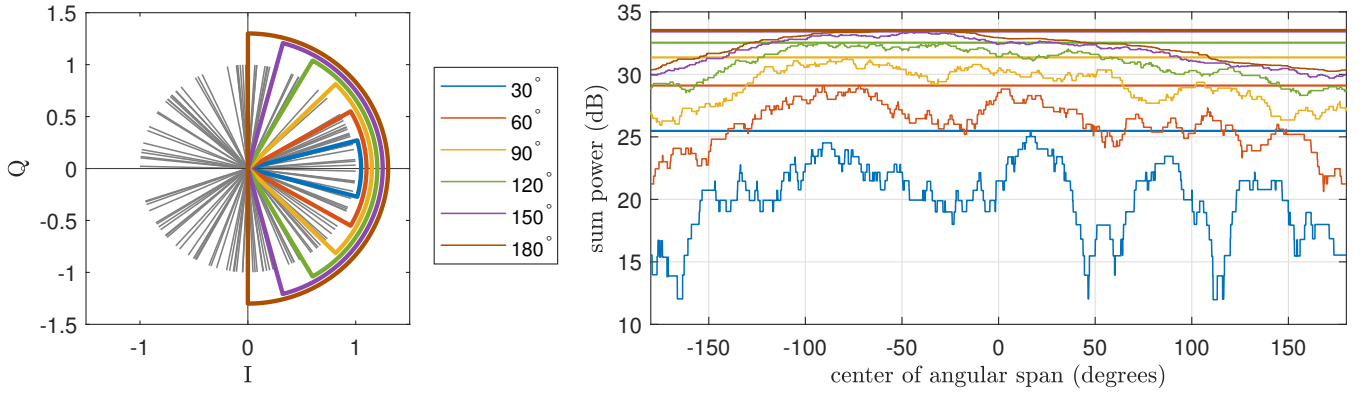
Fig. 3. Typical example of phasor distribution, $P(\theta)$, and output power with different choices for angular span, ψ, of phasors activated. The maximum possible output for each choice of ψ is marked by a solid horizontal line. Extending the active span beyond $120°$ provides negligible gain in output power.
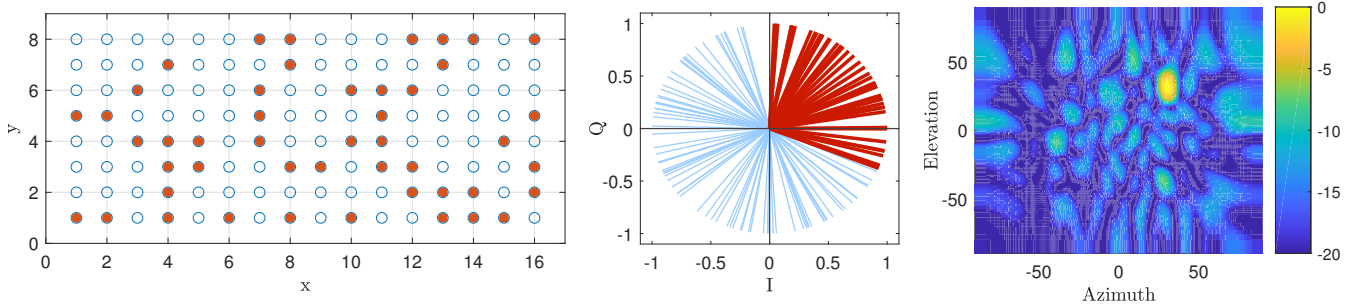

Fig. 4. Example of on-off beamforming toward $30°$ azimuth and $30°$ elevation in a 16×8 element rectangular array. Active elements are marked in red (left), and corresponding phasors of inactive and active channels are depicted in the I-Q plane, in blue and red, respectively (center). The resulting radiation pattern is depicted in the azimuth-elevation plane (right).

around 14 mW and 30 mW, respectively, bringing the overall per-channel consumption to approximately 76-100+ mW [6], [7].[1] By avoiding the phase shifter, only a few dB of gain is required, reducing LNA power consumption to well below 10 mW. The power consumption per active channel is therefore at least an order of magnitude lower than conventional designs. Since this simplified architecture allows us to increase the number of channels with ease, the overall consumption per active channel would be even lower, as the (relatively fixed) shared power consumption is divided between a larger number of channels.

Thus, we predict that with the on-off configuration, the per-channel power consumption is reduced by a factor of 10X or more. We note, however, that, due to the lack of phase control, a portion of the power of each active channel is effectively "wasted" due to imperfect phase alignment between channels. In the following section, we discuss the beamforming procedure and gain-efficiency tradeoffs in the

[1]These numbers do not include the power consumption of the LO that drives the IF mixers, as the values reported in these references were lumped up with that of the LO multiplier, which is a shared component. With the optimistic assumption of only a few mW consumption by each channel's LO driver, the state of the art design can be assumed to burn 80+ mW per channel (excluding consumption of shared components).

on-off architecture.

B. On-Off Beamforming in the Switched Array

Without phase shifters, it is not possible to align all of the channels to maximize array gain. In this setup, "beamforming" constitutes choosing the switch position of each antenna such that the largest possible number of "sufficiently aligned" signals make their way to the combiner. Since some elements are switched off, and phase alignment is not perfect between those that are active, the overall beamforming gain provided by an N element on-off receiver is a fraction of N, and, in fact, a fraction of the number of active channels. Thus, this architecture forces a tradeoff between die area utilization (and hardware redundancy) and power efficiency: We can get away with switches instead of phase shifters at the cost of leaving a portion of our aperture area unused, while also allowing some of the power used to activate the ON channels to be wasted due to phase misalignment.

We assume here that each channel includes a constant but random phase shift in its path to the combiner. This will most likely occur naturally due to variability in the path lengths and layout of the RF circuits, but it can also be imposed by design. As a result, the phasors arriving at the combiner from the N channels – which include the circuit path phase shifts

plus the phases of the channel response on the array – are randomly distributed in the I-Q plane. This is shown for a typical example in Figure 4, assuming a single-path line of sight (LoS) channel. The actual phasor directions vary with the angle of arrival, but their distribution remains uniformly random on the unit circle. In order to beamform toward a given direction, we look at the positions of these phasors for that angle of arrival and activate the largest possible set of channels whose signals add up with sufficient coherence. We may find this set by circular-convolving a box function,

$$b(\theta) = I_{[-\frac{\psi}{2}, \frac{\psi}{2}]}(\theta) \tag{1}$$

with the function

$$P(\theta) = \sum_{i=1}^{N} \delta(\theta - \phi_i) \tag{2}$$

where ϕ_i is the phase of channel i at the combiner and $\delta(.)$ is the Dirac function. Assuming this convolution is maximized (in magnitude) at a box shift of θ_0, i.e.,

$$\theta_0 = \text{argmax}_\theta (P \circledast b)(\theta) \tag{3}$$

we select the set of active channels as

$$\{i : \phi_i \in [\theta_0 - \frac{\psi}{2}, \theta_0 + \frac{\psi}{2}]\}.$$

This procedure is sketched in Figure 3, while Figure 4 shows the selected set for an example scenario with box width $\psi = 120°$, along with the resulting beamformed pattern.

C. Benchmark: low-resolution phase control

We also consider the benchmark architectures shown in Figure 2, where each channel is equipped with a 1-bit or 2-bit phase shifter. With 1 bit of phase control in each channel, i.e., the choice of adding a $0°$ or $180°$ phase shift, all phasors arriving at the combiner can be folded onto a single half-plane in I-Q space. Similarly, with 2 bits of phase control (phase shift options $0°$, $180°$, $90°$, and $-90°$), all phasors can be folded onto a single quarter-plane (quadrant). In this case, beamforming requires optimizing the orientation of the reference half-plane or quadrant, as well as choosing the activated phasors after folding for maximum power combining efficiency. This can be done in a manner similar to (1-3), by convolving with two or four equi-spaced boxes, i.e.,

$$b^{\text{1-bit}}(\theta) = I_{S_{\text{1-bit}}}(\theta), \quad b^{\text{2-bit}}(\theta) = I_{S_{\text{2-bit}}}(\theta)$$

where

$$S_{\text{1-bit}} = [-\frac{\psi}{2}, \frac{\psi}{2}] \cup [\pi - \frac{\psi}{2}, \pi + \frac{\psi}{2}]$$

$$S_{\text{2-bit}} = [-\frac{\psi}{2}, \frac{\psi}{2}] \cup [\pi - \frac{\psi}{2}, \pi + \frac{\psi}{2}]$$
$$\cup [\frac{\pi}{2} - \frac{\psi}{2}, \frac{\pi}{2} + \frac{\psi}{2}] \cup [-\frac{\pi}{2} - \frac{\psi}{2}, -\frac{\pi}{2} + \frac{\psi}{2}]$$

and activating the shifted span that maximizes the convolution output. Figure 5 shows examples of the 1-bit and 2-bit box functions with an angular span of $\psi = 60°$.
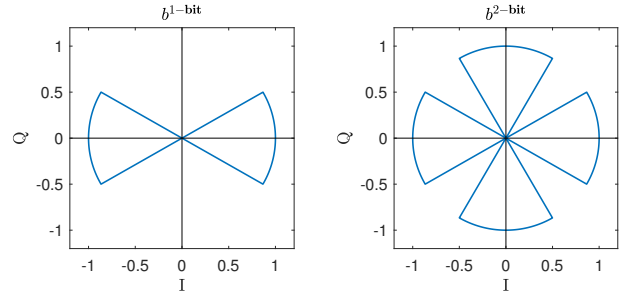

Fig. 5. Convolved activation spans for 1-bit and 2-bit beamforming.

In the next section, we provide analytical insight into the beamforming capacity and gain-efficiency tradeoff of this architecture, as well as comparisons with the 1-bit and 2-bit phase control benchmarks. We also discuss some of the open questions and practical challenges that we hope to address in the future.

III. PERFORMANCE ANALYSIS AND NUMERICAL RESULTS

With adequate array calibration and channel estimation, the relative phases of the signals arriving at the combiner are known, allowing us to determine the phasor response signal $P(\theta)$ defined in (2). The details of array calibration and channel estimation are left for future work, and we focus here on the performance tradeoffs obtained with on-off beamforming as described in Section II-B, as well as that of 1-bit and 2-bit beamforming as described in Section II-C. By tuning the activation span, ψ, we control the level of phase coherence between signals arriving at the combiner, which determines the beamforming efficiency. The activations may further be tweaked to steer one or more nulls in the direction of undesired interferers or jammers, which may also degrade phase coherence to some extent. Imperfect phase alignment causes part of the power used to activate channels to be wasted and not contribute to the overall beamformed power, dictating a tradeoff between utilization or total beamforming gain and power efficiency, which we proceed to quantify in this section and validate via Monte Carlo simulations. Our nominal simulation settings assume a 128-element array with calibration phases distributed uniformly between 0 and 2π, and a single-path channel with angle of incidence distributed uniformly between $-\pi/2$ and $\pi/2$. While the array configuration has no effect on efficiency-utilization tradeoffs, we provide example radiation patterns for both linear and rectangular arrays in all three architectures.

A. The gain-efficiency tradeoff

As mentioned above, the angular range of activated phasors, ψ, decides a tradeoff between the total beamforming gain and power efficiency, i.e., beamforming gain per unit of power burned. The level of phase coherence, and hence power efficiency, improves as ψ decreases, and, as shown in Figure 3, expanding the angular range of active phasors beyond $120°$ yields diminishing returns.

In order to quantify our activation efficiency, we define the contribution of each channel, or the "per-channel gain" (PCG) as the normalized inner product between its phasor and the sum of all other active phasors. This value captures the effective benefit (in amplitude) we get from burning power to activate that channel, conditioned on the set of already active channels. Consequently, the *minimum* per-channel gain (MPCG) over all active channels is a suitable criterion for beamforming efficiency. MPCG of ϵ can be translated to angular span ψ of activated phasors using the formula,

$$\epsilon = \cos\frac{\psi}{2}, \qquad \psi = 2\cos^{-1}\epsilon, \qquad (4)$$

since any phasor within $[-\frac{\psi}{2}, \frac{\psi}{2}]$ of the sum phasor provides a contribution larger than or equal to

$$\epsilon = \cos\frac{\psi}{2}$$

to the sum phasor. With the pessimistic assumption that the box shift is chosen uniformly at random, i.e., without maximizing the convolution described in 3 in order to focus on the area with highest phasor density, the statistical average of the number of active elements, N_{active}, and absolute sum of active phasors, G, is given by

$$N_{\text{active}} = \frac{\psi}{2\pi}N, \qquad (5)$$

$$G = \frac{N}{2\pi}\int_{-\frac{\psi}{2}}^{\frac{\psi}{2}}\cos\varphi\, d\varphi = \frac{\sin\frac{\psi}{2}}{\pi}N, \qquad (6)$$

bringing the average power efficiency of the scheme to

$$\eta = \left(\frac{G}{N_{\text{active}}}\right)^2 = \text{sinc}^2\frac{\psi}{2\pi}. \qquad (7)$$

In practice, the activation span will be shifted to maximize this gain, therefore, these predictions are a lower bound for the true beamforming performance.

With 1-bit phase control, all phasors can be aggregated in one half-plane of the I-Q space by selecting the π phase shift position for channels that fall outside the half-plane. Thus, for the 1-bit phase shifted architecture, the number of active channels and resulting gain are doubled, i.e., for the same ϵ and ψ, we have,

$$N_{\text{active}}^{\text{1-bit}} = \frac{\psi}{\pi}N, \qquad (8)$$

$$G^{\text{1-bit}} = \frac{N}{\pi}\int_{-\frac{\psi}{2}}^{\frac{\psi}{2}}\cos\varphi\, d\varphi = 2\frac{\sin\frac{\psi}{2}}{\pi}N. \qquad (9)$$

Note that the beamforming efficiency, η, as a function of ϵ is the same as that of on-off beamforming since both parameters are scaled by a factor of 2. This is to be expected, since the beamforming efficiency is a function of the phase coherence between active channels. Keeping ϵ, and therefore ψ, fixed, going from on-off to 1-bit phase control increases the number of phasors that fall within the activation span, but does not improve their phase alignment.
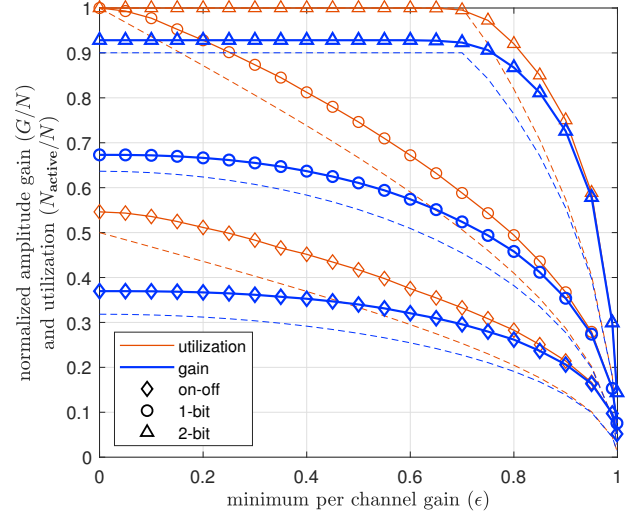
(a)
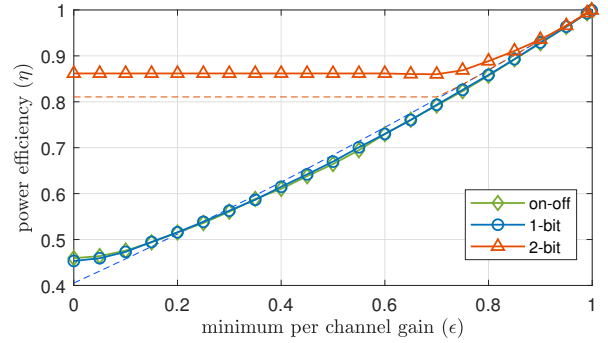
(b)

Fig. 6. Beamforming performance for the proposed and benchmark architectures as a function of ϵ, (a) amplitude of on-off beamformed signal (as a fraction of N) and fraction of active antennas, (b) power efficiency, i.e., ratio of beamformed power to that of a standard phased array with the same number of active antennas. Solid lines are numerical results averaged over 100 realizations, dashed lines are theoretical lower bounds.

For 2-bit phase control, N_{active} and G are scaled by a factor of 4 relative to (5) and (6), with the caveat that, since all phasors are in a $\pi/2$ angular range in the I-Q plane, for any

$$\psi \geq \pi/2, \quad \epsilon \leq \frac{1}{\sqrt{2}},$$

we have,

$$N_{\text{active}}^{\text{2-bit}} = N,$$
$$G^{\text{2-bit}} = \frac{2\sqrt{2}}{\pi}N \approx 0.9N,$$
$$\eta^{\text{2-bit}} = \frac{8}{\pi^2} = 0.81. \qquad (10)$$

Since extending ψ beyond $180°$ is never beneficial, this saturation behavior is not seen with 1-bit beamforming.

Figure 6 depicts numerical results, averaged over 100 realizations of a 128 element array, for the variation of on-off beamforming gain (G), utilization (N_{active}/N), and power
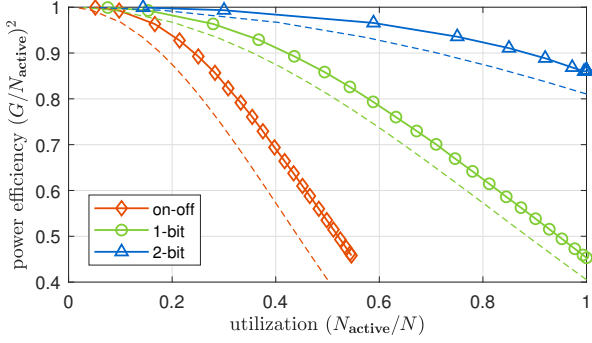
Fig. 7. The efficiency-utilization tradeoff: beamforming power efficiency, as a function of aperture utility (portion of active channels). Numerical results and theoretical lower bounds depicted by solid and dashed lines, respectively.

efficiency (η) as functions of MPCG (ϵ), along with benchmark values of low-resolution beamforming with 1-2 bits of phase control. This figure shows that, as ϵ increases from 0 up to around 0.5, there is very little loss in the overall beamforming gain but significant improvement in power efficiency. Thus we choose this threshold for on-off beamforming and performance analysis in our system. With $\epsilon = 0.5$, we have $\psi = 2\pi/3$ by (4). That is, all active phasors are guaranteed to reside within a $120°$ span in I-Q space. Although phasors are distributed uniformly, the selection process favors areas with higher density than average[2], and we may conclude from (5) that *at least* $1/3$ of antennas are activated. Thus, a *lower bound* for the combined phasor amplitude is calculated by the statistical average described in (6) as

$$G = \frac{N}{2\pi} \int_{-\frac{\pi}{3}}^{\frac{\pi}{3}} \cos\varphi \, d\varphi = \frac{\sqrt{3}}{2\pi} N \approx 0.28N. \quad (11)$$

Thus, by setting the MPCG to $\epsilon = 0.5$, we activate at least $1/3$ of the antennas, providing at least a fraction 0.28 of the full beamforming potential, bringing the effective power efficiency to $(0.28 \times 3)^2 \approx 70\%$. From the Monte Carlo results shown in Figure 6 for a 128 element array, we see that, in fact, 40% of channels are active on average, and the mean normalized gain is 0.34, providing the same power efficiency as the analytical prediction, 70%. Note that without setting a threshold for ϵ (i.e., setting a threshold of 0), this efficiency would be around 40%. Thus, with a choice of $\epsilon = 0.5$, the on-off switched architecture provides a factor of $10 \times 0.7 = 7X$ reduction in power consumption relative to a conventional phased array (for providing the same beamforming gain) at the cost of 40% utilization of hardware and die area (or 2.5X hardware redundancy).

For 1-bit and 2-bit beamforming, we see the same efficiency curves, except for the saturation of 2-bit beamforming as quantified in (10). Utilization and gain curves for 1-bit beamforming show the expected $\sim$2X improvement described in (8)

[2]Finding the best possible phasor selection can be done with $\mathcal{O}(N)$ computational complexity, and only needs to be repeated on the time scale of channel coherence.
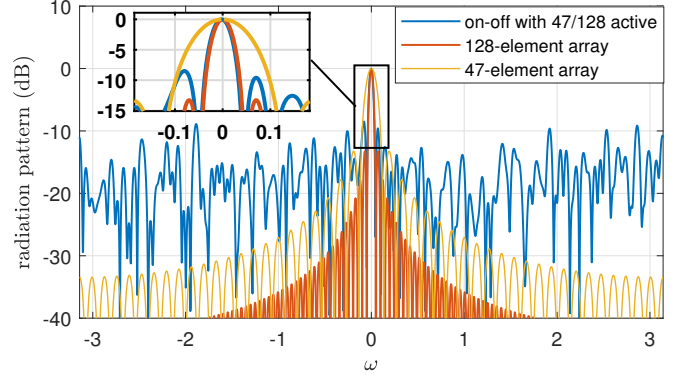

Fig. 8. Normalized radiation pattern of a linear 128-element on-off array with 47 active channels, alongside that of standard 128- and 47-element phased arrays. Patterns are plotted as a function of normalized spatial frequency in radians. We observe that the main beam tracks a 128 element array closely because of equal aperture size, but side lobe levels are, on average, higher than the 47-element array: with equal radiated power (number of active elements), a narrower main beam results in more power spread onto the sidelobes.
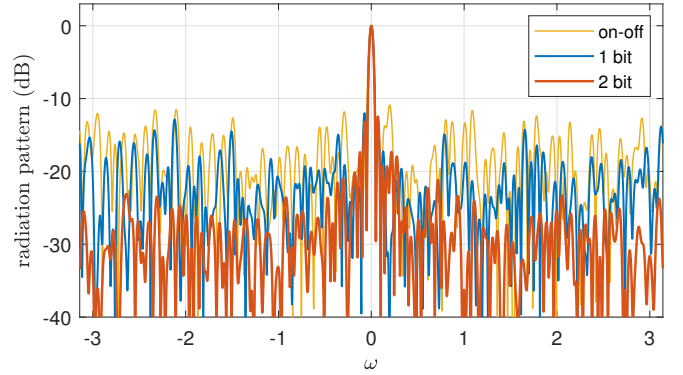

Fig. 9. Radiation pattern of a 128 element linear array with on-off, 1-bit, and 2-bit beamforming at $\epsilon = 0.5$.

and (9), while 2-bit beamforming provides the approximately 4X improvement up to the saturation point of $N_{\text{active}} = N$ and $G \approx 0.9$, as predicted in (10). The full tradeoff curves between power efficiency and hardware utilization are depicted in Figure 7, alongside the 1-bit and 2-bit benchmarks. These results show the extent to which, for a given utilization ratio, the beamforming efficiency improves as the level of phase control increases. If low-resolution phase shifters with comparably low power consumption are realized in the future, they may be exploited to increase aperture utilization or improve beamforming efficiency (or both) as described by this graph.

Based on our analysis thus far, a switched architecture is a promising approach for low-power scaling of the handset array. In order to effectively deploy such a system, however, several issues in signal processing and hardware must be addressed.

B. On-off nullforming for interference suppression

Since active elements are scattered across the entire aperture, array under-utilization as shown in Figure 4 does not
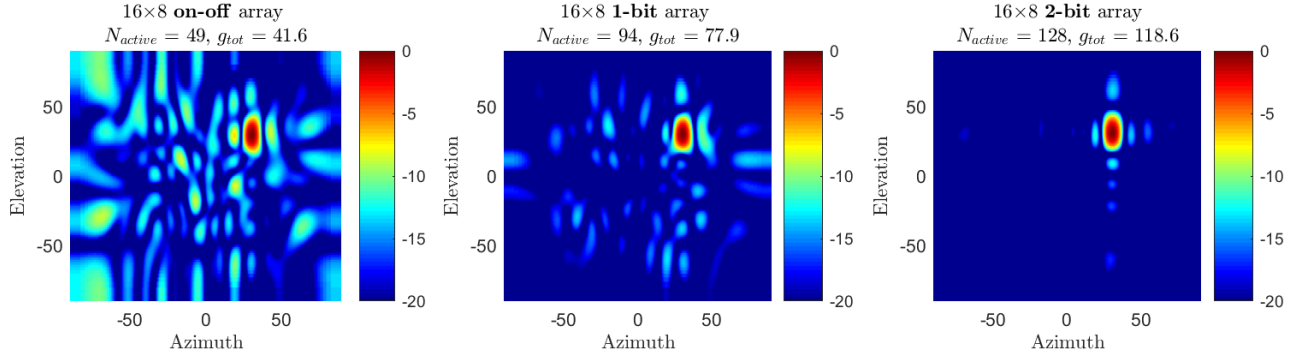
Fig. 10. 2D radiation pattern of a 16×8 array with on-off, 1-bit, and 2-bit beamforming at $\epsilon = 0.5$.

result in a wider beam or lower angular resolution, but instead increases the radiation outside the main beam causing *higher side lobe levels*. In fact, as shown in Figure 8, a 128 element on-off array (with 47 elements active, in this instant) has the same beamwidth as a fully sampled 128 element array, but exhibits much higher gain outside the main beam, even relative to a fully sampled 47 element array. Thus, if the array is impacted by interference from undesired radiation sources in the environment, active *nullforming* is likely to be necessary. Assuming sufficient separation between the desired and undesired channels, nullforming is a straightforward task in conventional phased arrays (beamforming weights are easily tuned to project onto the nullspace of the interferer's channel, with little cost to beamforming gain). When limited to switches instead of phase shifters, however, conventional techniques are not applicable, and finding the optimal switch positions for beam and null steering has combinatorial complexity. The problem of null steering in switched arrays with a discrete, finite set of positions for each channel is therefore an interesting open problem that we plan to address in future work.

The level of undesired sidelobe power is reduced in a 1-bit architecture, and even more so with 2-bit beamforming, as shown in Figure 9 for a 128 element linear array, and in Figure 10 for a 2-dimensional 16×8 array. Thus, interference cancellation is likely to be less of a problem in these architectures. Furthermore, the added phase control increases interference rejection capacity and reduces main beam degradation by increasing the degrees of freedom that can be applied to the nullforming process. This may be considered as yet another reason to pursue the realization of low-power array architectures with low-resolution phase control.

C. Calibration and channel estimation

A prerequisite for any beam or null formation, of course is sufficiently accurate knowledge of the response from each antenna at the combiner. In order to determine these phasors, we not only need to identify the channel matrix from the transmitter and interferers on the array (channel estimation), but we also need to know the relative phase offsets induced by the signal chain of each channel (calibration). Assuming the calibration offsets are known, compressive channel estimation
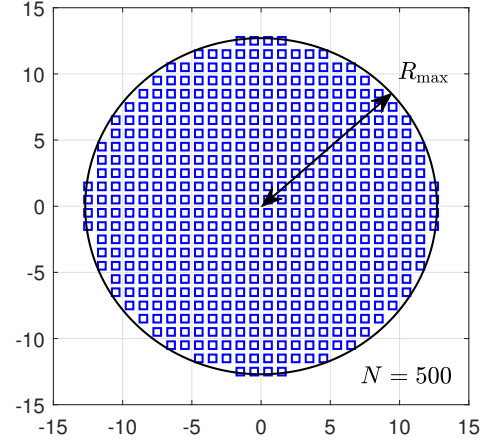

Fig. 11. Number of elements in 2D array limited by maximum length of transmission line to chip (center). At 0.1 dB/mm, maximum allowable radius for ≤ 1.5 dB loss is around 13 mm, corresponding to around 500 elements at 140 GHz assuming (near) half-wavelength spacing between elements in each dimension.

techniques developed in prior work may be applied to the switched array for low-overhead channel estimation. The conventional Newtonized orthogonal matching pursuit (NOMP) algorithm [16] can be applied if the link maintains coherence between switch updates, otherwise a noncoherent scheme may be adopted for this purpose [17], [18].

Calibrating the array in a *controlled* environment (i.e., inside an anechoic chamber where the channel vector is known) and with *coherent* measurements is fairly straightforward: After taking $K \geq N$ measurements of the channel with different (random) switch positions, we solve (or find the MMSE fit to) the system of linear equations to obtain the overall array response and, thereby, calibration coefficients. It may, however, be desirable to calibrate arrays *in the field* [19] with unknown channels and, more importantly, without relying on phase coherence across measurements. This is also an open issue that will be the focus of future work.

D. The limits of scaling

While the switched architecture is a promising approach to low-power scaling of arrays in size and frequency, other practical limitations may hamper the growth of array size beyond some point. Let us consider the 2D circular array of Figure 11 to understand these practical limits. As the array grows larger, the length of the paths that route outer elements to the centrally located combiner becomes large, and the propagation loss of these paths becomes a bottleneck. Our electromagnetic simulations of a high performance laminate circuit board show that at 140 GHz, the transmission line exhibits propagation loss of 0.1 dB/mm. In order to ensure routing loss below 1.5 dB, a maximum radius of 15 mm is imposed on the array which, assuming half-wavelength spacing between elements, limits the number of antennas that can efficiently be connected to one RF combiner to around 500 elements for our case study. Interestingly, the loss per unit distance tends to scale linearly with carrier frequency, f. Therefore, as we scale up in frequency, the allowable array diameter reduces linearly (and its area quadratically) with f. The element spacing, $\lambda/2$, also reduces proportionally to $1/f$. Thus, the maximum number of elements in a 2D circular array remains *relatively constant* as the carrier frequency moves higher in the spectrum! This imposes a natural limit on the number of antennas that can be routed to a single RF combiner, irrespective of frequency. In order to get past this limit, we may devise a more complex hierarchical combining strategy or adopt a modular approach of cascading, or "tiling", multiple on-off sub-arrays to form a larger super-array. A tiled architecture with a separate RF chain on each tile would allow *hybrid* beamforming with the possibility of spatial multiplexing as well as improved beam and null steering capability via spatial processing in the baseband digital domain.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we motivate and propose scalable, low-power architectures for mmWave massive MIMO frontends that rely on switching in place of noisy RF phase shifters for each array element. In the simplest case of on-off switching, per-channel power consumption is reduced by an order of magnitude relative to state of the art designs at 100+ GHz, making this architecture a promising approach for scaling to large arrays at high frequencies. These power savings come at the cost of under-utilization of the array aperture and, consequently, die area. Our results show that maintaining beamforming efficiency of 70% is possible with approximately 40% utilization of array elements, or 2.5X hardware redundancy relative to a conventional phased array providing the same beamforming gain. On each channel, the switched architecture provides 10X reduction in power consumption, which is reduced to 7X savings in *beamformed* power consumption due to the reduced power efficiency of on-off beamforming. This efficiency can be increased by placing more stringent requirements on the phase alignment of active channels, which would result in lower utilization, higher sidelobes, and greater hardware redundancy.

We also showed the extent to which these tradeoffs could be improved by incorporating binary or quaternary phase shifting on each channel. With the realization of low-power phase shifters in the future, these architectures would allow increasing the aperture utilization by up to 4X and/or improving the beamforming efficiency. In future work, we plan to develop efficient algorithms for calibration and channel estimation in the on-off, 1-bit, and 2-bit arrays, and to extend the beamforming strategies to incorporate nullforming for interferer rejection.

V. ACKNOWLEDGMENTS

REFERENCES

[1] M. J. Rodwell, "100-340GHz spatially multiplexed communications: IC, transceiver, and link design," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2019, pp. 1–5.

[2] M. Rodwell, A. A. Farid, A. Ahmed, M. Seo, U. Soylu, A. Alizadeh, and N. Hosseinzadeh, "100-300 GHz wireless: ICs, arrays, and systems," in *2021 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*. IEEE, 2021, pp. 1–4.

[3] A. Puglielli, A. Townley, G. LaCaille, V. Milovanović, P. Lu, K. Trotskovsky, A. Whitcombe, N. Narevsky, G. Wright, T. Courtade *et al.*, "Design of energy- and cost- efficient massive MIMO arrays," *Proceedings of the IEEE*, vol. 104, no. 3, pp. 586–606, 2015.

[4] P. Skrimponis, S. Dutta, M. Mezzavilla, S. Rangan, S. H. Mirfarshbafan, C. Studer, J. Buckwalter, and M. Rodwell, "Power consumption analysis for mobile mmwave and sub-THz receivers," in *2020 2nd 6G Wireless Summit (6G SUMMIT)*. IEEE, 2020, pp. 1–5.

[5] P. Skrimponis, N. Hosseinzadeh, A. Khalili, E. Erkip, M. J. Rodwell, J. F. Buckwalter, and S. Rangan, "Towards energy efficient mobile wireless receivers above 100 GHz," *IEEE Access*, vol. 9, pp. 20 704–20 716, 2020.

[6] S. Li and G. M. Rebeiz, "A 134-149 GHz IF beamforming phased-array receiver channel with 6.4-7.5 dB NF using CMOS 45nm RFSOI," in *2020 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. IEEE, 2020, pp. 103–106.

[7] S. Nicolson, A. Tomkins, K. Tang, A. Cathelin, D. Belot, and S. Voinigescu, "A 1.2 V, 140GHz receiver with on-die antenna in 65nm CMOS," in *2008 IEEE Radio Frequency Integrated Circuits Symposium*. IEEE, 2008, pp. 229–232.

[8] W. B. Abbas, F. Gomez-Cuba, and M. Zorzi, "Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution ADCs," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8131–8146, 2017.

[9] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, "Achievable uplink rates for massive MIMO with coarse quantization," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 6488–6492.

[10] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 4038–4051, 2017.

[11] O. Bakr, M. Johnson, J. Park, E. Adabi, K. Jones, and A. Niknejad, "A scalable-low cost architecture for high gain beamforming antennas," in *2010 IEEE International Symposium on Phased Array Systems and Technology*. IEEE, 2010, pp. 806–813.

[12] T. Heller, E. Cohen, and E. Socher, "A 102–129-GHz 39-dB gain 8.4-dB noise figure I/Q receiver frontend in 28-nm CMOS," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 5, pp. 1535–1543, 2016.

[13] A. Hajimiri, H. Hashemi, A. Natarajan, X. Guan, and A. Komijani, "Integrated phased array systems in silicon," *Proceedings of the IEEE*, vol. 93, no. 9, pp. 1637–1655, 2005.

[14] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, and A. Hajimiri, "A 77-GHz phased-array transceiver with on-chip antennas in silicon: Receiver and antennas," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12, pp. 2795–2806, 2006.

[15] E. Cohen, M. Ruberto, M. Cohen, O. Degani, S. Ravid, and D. Ritter, "A CMOS bidirectional 32-element phased-array transceiver at 60 GHz with LTCC antenna," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 3, pp. 1359–1375, 2013.

[16] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Frequency estimation for a mixture of sinusoids: A near-optimal sequential approach," in *3rd IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2015.

[17] M. E. Rasekh, Z. Marzi, Y. Zhu, U. Madhow, and H. Zheng, "Noncoherent mmWave path tracking," in *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*, 2017, pp. 13–18.

[18] M. E. Rasekh and U. Madhow, "Noncoherent compressive channel estimation for mm-wave massive MIMO," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2018, pp. 889–894.

[19] M. Rasekh, B. Puranik, U. Madhow, and M. Rodwell, "In-the-field calibration of all-digital mimo arrays," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2022.