Broadband, High-Linearity Switches for Millimeter-Wave Mixers Using Scaled SOI CMOS

CAMERON HILL (Member, IEEE), AND JAMES F. BUCKWALTER (Fellow, IEEE)

Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93117, USA

CORRESPONDING AUTHOR: C. HILL (e-mail: cameron_hill@ucsb.edu)

This work was supported by the Defense Advanced Research Projects Agency (DARPA) SPAR Program managed by Dr. T. Hancock through Cooperative under Agreement HR0011-17-2-0003.

ABSTRACT This work demonstrates new circuit techniques in distributed-stacked-complimentary (DiSCo) switches that enable picosecond switching speed in RF CMOS SOI switches. By using series-stacked devices with optimized gate impedance and voltage swing, both high linearity and fast switching are possible. A theoretical analysis and design framework has been developed and verified through simulation and measurement through two broadband, high-linearity passive mixer designs, one optimized for linearity and the other for bandwidth, using a 45-nm SOI CMOS process. The mixers achieve $P_{1dB}$ of 16-22 dBm with $IIP_3$s of 25-34 dBm across a bandwidth from 1 GHz up to 30 GHz. This performance exceeds prior SOI RF and microwave mixer performance by more than an order of magnitude and is comparable to III-V device technologies. The mixers include integrated local oscillator (LO) driving amplifiers for high efficiency operation and low total power consumption. DC power consumption ranges from 250 mW to 1 W for the LO driver. The integrated LO drivers demonstrate a pathway to on-chip LO generation with simplified matching to maximize LO power delivered to the input of the switch.

INDEX TERMS Switch, mixer, microwave, millimeter-wave, broadband, intermodulation distortion, power handling, SOI CMOS.

I. INTRODUCTION

RF SWITCHES are critical to transceiver architectures and are typically used in the front-end for switching matrices, mixers, and samplers. Moreover, reconfigurable, multi-band radios are placing more demands on switching as a feature to tune individual components. Depending on the requirements, different device technologies are demanded; for example, SOI CMOS offers the capability of RF and digital integration while GaN and GaAs devices excel for high power handling and linearity. By stacking SOI CMOS devices in series and adding a large gate resistance on each stacked device, SOI CMOS switches with $P_{1dB}$ up to 40 dBm are possible [1], [2]. However, the gate resistance limits switching time and creates a trade-off between linearity and switching speed. A large commercial market exists for SOI-based RF switches that are used in front-end filter banks and transmit/receive switches where they operate in quasi-static modes and fast switching is not required.

High switching speed is necessary for microwave/millimeter-wave mixers [3], [4], [5], [6], [7], [8], [9], [10], [11], high dynamic range sample and hold [12], [13], or other emerging RF switching applications such as RF signal processing and spatio-temporal modulation systems [14], [15]. For these applications either the linearity needs to be sacrificed, or more advanced and costly technologies are needed [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. Several designs have previously attempted to increase the switching speed of high-linearity CMOS switches [29], [30], [31], [32], however, the theory and implementation of these designs imply that the fundamental ratio between switching frequency and RF frequency, or fractional bandwidth (FBW), is constrained.
to \( \sim \) 50\% and are not suitable for the applications such as mixers.

High linearity mixers are increasingly in demand with the growing use of arrays in commercial communication systems. In the presence of one or more strong signals at the antennas shown in Fig. 1, the combination of array and RF gain in the front end leads to as much as 15 dBm input power to the single mixer after combining depending on the size of the array [33]. In the case of digital arrays, no spatial filtering in the RF front-end demands additional linearity to prevent 3rd-order intermodulation distortion (IM3) products from appearing in-band and in-beam [34]. In addition, re-configurable, wide-band, and multi-standard arrays will be exposed to many interferers further driving the linearity requirements of array front ends [35].

This work proposes a gate capacitive element as shown in Fig. 1 (c), \( C_{ges} \), and device stacking similar to the approaches in power amplifiers [36], [37], [38], [39] to build a high-linearity SOI CMOS switches that are capable of fast switching necessary for passive mixers. A capacitive gate termination creates a frequency-independent voltage division which allows for high speed switching while maintaining linearity and achieving a FBW > 100\% that is only limited by the intrinsic (gate) resistances. The drawback of this approach is that the gate-drive voltage must be proportionally increased to achieve the correct gate voltage swing, \( V_{gs} \), leading to greater power consumption. An analysis of power consumption and linearity suggests an optimization of the design in Section II. The linearity of the switch is proportional to local oscillator (LO) power as with other semiconductors, e.g., GaAs or GaN, with higher drive/breakdown voltages.

With recent advances in SOI CMOS amplifiers, high efficiency can be leveraged from integrated driver amplifiers leading to overall improvements in system efficiency and integration. The high-efficiency broadband LO driver design is outlined in Section III. The application of the RF switch design to broadband, distributed-stacked-complimentary (DiSCO), passive mixers is presented in Section IV. Similar to other passive and active distributed mixers, [7], [40], [41], the proposed mixers integrate the transistors into a distributed artificial transmission line to improve the operational bandwidth. However, in this case the DiSCO devices are combined in series to increase the mixer linearity far beyond that of a single CMOS device, compared to previous work which leverages multiple devices to increase only isolation or conversion gain.

A prototype set of microwave and millimeter-wave (mmW) mixers were designed in a GlobalFoundries 45-nm SOI CMOS process with emphasis on comparing the linearity and bandwidth. A high-linearity version was presented recently and this work expands on earlier work to demonstrate a generalized design methodology and high-bandwidth variant that has not been previously reported [42]. Section V discusses the measured mixer performance and compares previous CMOS and III/V mixers where it is clear that the presented mixer outperforms other silicon approaches by more than an order of magnitude and competes with, or exceeds the performance of costly III-V materials.

II. BROADBAND SWITCH THEORY

Previous work has demonstrated the linearity of an RF switch can be considered independently for both the ON and OFF states [2]. In the conducting (ON) state, compression is avoided with a high impedance on the gate to induce voltage division allowing the gate voltage to swing in unison with the drain and source voltages. In the OFF state, the series, stacked devices distribute the drain-source voltage swing to prevent compression or breakdown. While device stacking is an established technique that has been previously handled comprehensively [43], [44], [45], it conventionally relies on resistive gate terminations which limit switching speed. The approach presented here investigates the choice of gate impedance that allows for maximum linearity and switching speed.

A. GATE IMPEDANCE AND FREQUENCY RESPONSE

The voltage response from the drain to the gate, \( V_{RF} \) to \( V_g \), and the LO voltage response to the gate, \( V_{LO} \) to \( V_g \), are shown in Fig. 2 (d) for an ON state switch model. The total \( V_g \) is characterized by

\[
V_g = \frac{j\omega C_{g,in}Z_g}{1 + j\omega C_{g,in}Z_g} V_{RF} + \frac{1}{1 + j\omega C_{g,in}Z_g} V_{LO},
\]  

where \( C_{g,in} \) is the intrinsic gate capacitance and \( Z_g \) is the external impedance placed at the gate.

This expression indicates that using a resistive or inductive external gate impedance produces an undesirable tradeoff as plotted in Fig. 2 (a) and (b) where \( |V_g,RF/V_{RF}| \) and \( |V_g,LO/V_{LO}| \) are plotted as a function of gate impedance.
that ranges from 0.1 to 10 times the intrinsic gate capacitance. The driving frequency, $f_{LO}$, is plotted for different cases where $f_{LO} < f_{RF}$. $f_{LO} = f_{RF}$, $f_{LO} > f_{RF}$ representing different modes of operation in a mixer.

To maintain high linearity, both the transfer functions from $V_{RF}$ to $V_{g,RF}$ and $V_{LO}$ to $V_{g,LO}$, plotted in Fig. 2 (a), (b), and (c), must remain close to 1 at the given modulation frequency. For the resistive gate termination in Fig. 2 (a), $V_{g,LO}$ declines rapidly as modulation frequency increases [2]. In Fig. 2 (b), using a low-Q inductive termination allows for both gate voltages to remain high at a much higher LO frequency and peaking is seen in the $f_{LO} = f_{RF}$ contour. However, as $f_{LO}$ exceeds $f_{RF}$, the modulated gate voltage drops rapidly, limiting inductive designs to applications where the LO frequency must be less than the RF frequency [46]. In these two cases, the response at the gate is frequency-dependent, limiting their for broadband applications.

Alternatively, an extrinsic gate capacitor produces a frequency-independent response in Fig. 2 (c). There is no extrinsic gate capacitance for which both $V_{g,RF}$ and $V_{g,LO}$ remain high. However, frequency independence allows the modulating gate voltage drop from a high $Z_{C_{g,ex}}$ to be compensated with increasing the modulating drive voltage uniformly across frequency, thereby, maintaining switch linearity.

B. DISTRIBUTED, STACKED SWITCHES

The capacitive gate approach is inherently wideband and will both linearize the device and reduce the shunt capacitance of the switch. However, stacking devices causes the total shunt capacitance to increase and introduces a lumped bandwidth constraint. For wideband designs, distributing the switch capacitance across a lumped-element transmission line, Fig. 3, and compensates for the bandwidth restriction.

Here, two design approaches for drive distribution are considered. First, a single amplifier drives the gate voltage along a distributed transmission line terminated in a resistive load as shown in Fig. 3 (a). Second, several smaller amplifiers drive each capacitive gate as shown in Fig. 3 (b). In Fig. 3, complimentary NMOS/PMOS devices are connected in parallel to cancel clock feedthrough in both cases.

The maximum tolerable input voltage before compression, $V_{max}$, is proportional to the number of stacked devices, $N$, and is typically,

$$V_{max} \approx N(V_{DD} - V_{th})$$  \hspace{1cm} (2)

where $V_{DD}$ is the nominal process voltage, i.e., 1 V for a 45 nm process. If, for a single device, the desired peak-to-peak voltage swing on the gate is $V_{g,pp} = 2V_{DD} (-1$ V to +1V) for optimum linearity, and $V_{LO,pp}$ is the peak-to-peak LO square-wave voltage, then

$$C_{g,ex} = \frac{V_{g,pp}}{V_{LO,pp} - V_{g,pp}}$$

and

$$C_{g,in} = \frac{2}{N - 1}$$

where $V_{LO,pp} = (N + 1)V_{DD}$ such that the LO drive voltage is proportional to $N$. The total gate capacitance, $C_{g}$, equal to $C_{g,in}$ and $C_{g,ex}$ in series is,

$$C_{g} = \frac{2NC_{g,in}}{N + 1}$$  \hspace{1cm} (4)

with $C_{g,in}$ scaled proportionally with $N$ because the devices need to be wider with more stacked devices to accommodate increased series resistance. The power consumption of the LO of the first case in Fig. 3 (a) can be determined directly by the characteristic impedance of the artificial line,

$$P_{LO,a} = \frac{V_{LO,pp}^2}{2R_L} = \frac{(V_{DD}(N + 1))^2}{2R_L}$$  \hspace{1cm} (5)

Using well-known expressions for the characteristic impedance and cutoff frequency, $f_c$, of an artificial transmission line,

$$P_{LO,a} = \pi V_{DD}^2(N + 1)f_c C_{g,in}$$  \hspace{1cm} (6)
Assuming that the amplifiers have sufficient gain such that the power of the input signal is negligible, similar calculations can be done for the switch described by Fig. 3 (b), for capacitive power dissipation.

\[
P_{LO,b} = NC_{gt}V_{LO,pp}^2
\]
\[
= 2V_D^2N^2(N+1)f_{LO}C_{g,in}
\]
\[
\text{(7)}
\]

Comparing (6) and (7) determines the frequency range where each approach is beneficial for LO power consumption and there is a crossover frequency defined by \( f_{XO} = \frac{\pi}{2N}f_c \). The design in Fig. 3 (a) will consume less power than Fig. 3 (b) over a certain range of frequencies so long as \( 2N > \pi \), regardless of amplifier efficiency. Fig. 4 confirms this insight by plotting the crossover frequency and cutoff frequency as a function of the number of stacked devices for a 45-nm process assuming 10 µm switches when \( N = 1 \). For larger stacks of devices, the crossover frequency is relatively low, in the range 5-10 GHz, so for millimeter wave designs, Fig. 3 (a) is highly favorable.

For the resistive optimized case in Fig. 3 (a), both the maximum input power (linearity) and the DC power consumption scale with the same factor, \( (NV_{proc})^2 \), a common relationship for diode and other III-V based mixers.

### III. DRIVER AMPLIFIER DESIGN

While the artificial transmission line approach lowers power consumption across a large bandwidth, the amplifier design is critical for efficient performance. Stacked class-D amplifiers have been used to produce high voltage swings across bandwidth in SOI CMOS [47], [48]. An example of such a stacked amplifier is shown in Fig. 5 (a).

The NMOS/PMOS are sized to match their internal resistances of the amplifier in pull-up and pull-down directions, and thus the external gate capacitance needs to be scaled accordingly following,

\[
C_{gex,n} = C_{gin,n} \frac{1}{2n-1}
\]
\[
\text{(8)}
\]

where \( N \) is the number of stacked devices in the amplifier, \( n \) is the index of the capacitor from 1 to \( N - 1 \), and like with the mixer, \( C_{gex,n} \) and \( C_{gin,n} \) are the intrinsic and extrinsic capacitances of each stacked device. Note that for \( n = 1 \), \( C_{gex,n} \) goes to infinity as this gate impedance should be an AC ground.

The theoretical efficiency of a switched amplifier is 100%, so the internal device resistances and the shunt capacitances are the primary sources of inefficiency. The maximum output voltage swing is realized with the minimum ON resistance, requiring large devices. However, this comes with the trade-off of larger parasitic capacitances, which degrade efficiency at high frequency. Smaller devices result in resistively-dominated losses and the voltage division that occurs within the amplifier is compensated with supply voltage. Because the breakdown of the devices will be determined by the output swing of the amplifier, the overall stress on the devices is only marginally increased. The appropriate supply voltage can be chosen using,

\[
V_{SUP} = V_{LO} \left( \frac{R_L + 2R_{amp,int}}{R_L} \right) > NV_{DD},
\]
\[
\text{(9)}
\]

where \( R_{amp,int} \) is the amplifier output resistance, as shown in Fig. 5 (a), and \( R_L \) is the load resistance. An example of how this supply voltage scaling impacts output power and efficiency for a five-stack amplifier is shown in Fig. 6. Smaller devices with a higher supply voltage offer better efficiency across a larger bandwidth while supplying the same saturated output power as larger devices, with an up to 15% improvement at 20 GHz.
bands with a lower LO drive voltage. As a two-stack design that operates up to millimeter-wave frequencies for maximum linearity as described in [42] while the second stack is a six-stack architecture that reduces cutoff frequency with increasing numbers of stacked devices. In Fig. 4. The first mixer is a six-stack architecture that provides the best results compared to III-V power handling CMOS SOI mixers with integrated LO drivers. Thus, the millimeter-wave demonstration of wide-band, high-linearity and high-power RF switch capable of fast switching is compelling to demonstrate limits compared to III-V technologies for larger scale integration.

IV. BROADBAND MIXER DESIGN

The wideband, high-linearity switch capable of fast switching can be demonstrated in a passive ring mixer. CMOS technologies typically have a maximum $P_{1dB}$ of 10 dBm [49] when using linearity optimized pass gates for switches, and other designs often have significantly lower $P_{1dB}$ [3], [50], [51]. Mixers designed in other technologies such as GaAs and GaN exhibit higher linearity with external drive amplifiers, adding significant burden to the LO generation chain, especially at high frequencies [25], [26]. Thus, the millimeter-wave demonstration of wide-band, high-power handling CMOS SOI mixers with integrated LO drivers is compelling to demonstrate limits compared to III-V technologies for larger scale integration.

In this section, two mixer designs are presented based on the trade-off between bandwidth and linearity shown by the reduced cutoff frequency with increasing numbers of stacked devices in Fig. 4. The first mixer is a six-stack architecture for maximum linearity as described in [42] while the second is a two-stack design that operates up to millimeter-wave bands with a lower LO drive voltage.

A. SIX-STACK DISTRIBUTED, COMPLEMENTARY MIXER

To remain competitive with state-of-the-art III-V mixers [16], [17], [18], simulations indicated that $N > 6$ to provide $P_{1dB} > 20$ dBm while keeping the insertion loss (IL) at 5-10 dB across 18 GHz of bandwidth. The width of the devices is chosen based on the cutoff frequency of the distributed lines and the desired IL due to series resistance of stacked devices. In a 50Ω environment, the conversion loss is

$$CL = 3dB - 20 \log_{10} \left( \frac{R_L}{R_{series} + R_L} \right), \quad (10)$$

which implies a total series resistance of about 15 Ω. For $N = 6$, 2.5Ω is allocated for each device, which corresponds to a width of 100μm, and allows the cutoff frequency to exceed 18 GHz. Because the mixer is fully differential and uses complimentary devices, there are eight LO and RF paths which forms an artificial ground at each gate such that eq. (1) holds true. To conserve space at the expense of bandwidth, many of the inductors with common LO signals are shared, and the differential inductors are coupled where possible to achieve higher flux density with the layout of the coupled inductors in Fig. 7 (c). The overall schematic of the DiSCo mixer is shown in Fig. 7 (a) and (d). To confirm that the distributed structure of the mixer improves the conversion loss at high frequency, Fig. 7 (b) simulates the conversion loss vs. inductor size in both the LO and RF paths, showing that conversion loss is improved by up to 3 dB even within the non-linear and time-varying circuit.

In this design, the RF and LO lines use the same sized inductors and are both matched to 50Ω as to match the time delay through the distributed lines as close as possible. However, this isn’t necessary for operation of the mixer as a miss-match in delay turns into a slight frequency shift in the LO line as expressed by $f'_{LO} = f_{LO} + (\theta_{LO} - \theta_{RF})/(2\pi \tau)$, where $f'_{LO}$ is the shifted frequency, and $\theta$ is the phase delay from one segment of the distributed RF and LO lines respectively, and $\tau$ is the time delay of one segment.

Within the DiSCo switch, the unit cells can be divided into two categories: edge elements and center elements. The edge elements require slightly different sizing as they interface with non-repeating components within the mixer. Since two transistors connect to a single input or output node, the size of the edge elements needs to be reduced to prevent excessive capacitive loading. $C_{comp}$ maintains the capacitive loading along the LO line. All of the inductors within the differential distributed switch are uniform throughout the device. While the ideal size of the inductors is 250 pH, coupling the differential inductors together adds significant shunt capacitance at each node, approximately equal to 50 pF. The additional shunt capacitance increases the desired inductor value to 360 pH which is then compensated for in the inductor layout. The sizes for each device in the six-stack mixer are listed in Table 1.

To drive the LO lines, the LO signal is brought into the chip differentially, then split to drive the four pseudo-differential amplifiers. A resistive feedback amplifier offers wideband input matching. Generalizing the schematic in

| Component | Parameter |
|-----------|-----------|
| Center Elements | Edge Elements |
| NFET W / L | 100μm / 40nm | 50μm / 40nm |
| PFET W / L | 100μm / 40nm | 50μm / 40nm |
| $C_{comp}$ | 55.5 pF | 21 pF |
| $C_{Comp}$ | N/A | 25 pF |
| $L_{RF}$ | 360 pH | 360 pH |
| $L_{LO}$ | 360 pH | 360 pH |
| $L_{in/out}$ | 130 pH |
| $C_{L}$ | 44.5 pF |
| $R_{L}$ | 50 Ω |
FIGURE 7. Block diagram of the DiSCo mixer core (a) with conversion loss plotted as a function of the inductor value (b) and layout of the distributed RF and LO inductances (c) with different metal layers indicated by different colors. The full schematic of the mixer core DiSCo switch is illustrated in (d).

TABLE 2. Five-stack driver parameters.

| Component | NFET Parameter | PFET Parameter |
|-----------|----------------|----------------|
| W / L     | 100μm / 40μm   | 120μm / 40μm   |
| C1        | 88.3 pF        | 1.06 pF        |
| C2        | 105 pF         | 123 pF         |
| C3        | 61 pF          | 74.2 pF        |
| Supply Voltage | 7.1 V        |                |

TABLE 3. Broadband DiSCo mixer parameters.

| Component | Parameters, Center/Edge Elements |
|-----------|----------------------------------|
| NFET W / L| 25μm / 40μm                      |
| PFET W / L| 25μm / 40μm                      |
| Cpe         | 68.4 pF                          |
| LRF         | 360 pH                            |
| LLO         | 260 pH                            |
| Lcin/out    | 130 pF                            |
| CL          | 22 pF                             |
| RL          | 50 Ω                              |

Fig. 5 to five stacked devices, the component sizing is given in Table 2. Fig. 8 plots performance of the driver amplifier. Note that both the third harmonic power and the fundamental power are shown as a measure of the overall efficiency and quality for the generation of the desired square wave.

B. TWO-STACK DISTRIBUTED MIXER

Equation (4) shows that the total gate capacitance is proportional to $N/(N + 1)$, implying that the bandwidth of the distributed switch is related to $(N + 1)/N$. The bandwidth of a single device, $N = 1$, will be double that of larger $N$ as $N \to \infty$. Much of the bandwidth limitations exist in the driving amplifier due to the parasitics that come with large devices. As such, a wider bandwidth mixer is realized with fewer series devices to extend the bandwidth above 30 GHz. In this section, the design of a two-stack version of the mixer is presented to explore the bandwidth-linearity trade-offs.

While the same fundamental framework is used for the distributed switch design in Fig. 7, only using two series devices removes the center elements and leaves edge elements. As a result, the sizing of the LO path passives can be reduced by removing $C_{comp}$ on each side, doubling the LO bandwidth. The new component values are shown in Table 3.
FIGURE 8. Simulated s-Parameters, saturated output power of both the fundamental and third harmonic, and efficiency of the total driving amplifier for the six-stack mixer.

FIGURE 9. Simulated s-parameters, saturated output power of both the fundamental and third harmonic, and efficiency of the total driving amplifier for the six-stack mixer.

FIGURE 10. The high-linearity and broadband mixers under a microscope with their dimensions displayed where (a) is the six-stack mixer and (b) as the two-stack mixer.

TABLE 4. Three-stack driver parameters.

| Component | Parameter | NFET | PFET |
|-----------|-----------|------|------|
| W / L     | 48μm / 40nm | 56μm / 40nm | |
| C0        | 358 fF      | 358 fF      | |
| C1        | 127 fF      | 127 fF      | |
| Supply Voltage | 4.5 V | | |

This design requires a three-stack driver amplifier schematic as shown in Fig. 9. This amplifier enables bandwidths of up to 40 GHz and the device parameters for the three-stack amplifier are listed in Table 4.

V. MEASUREMENT RESULTS

Measuring accurate 3rd-order input-intercept points (IIP3) and $P_{1dB} > 20$ dBm over 30 GHz requires attention to the test setup. This section details the test setups for $P_{1dB}$, IIP3, conversion loss, and LO feedthrough in the measurements of both up and down conversion. Micrographs of the six-stack and two-stack mixers are shown in Fig. 10. Measurements in this section use a rolling average smoothing function to reduce uncertainty.

A. SIX-STACK MEASUREMENTS

Since the six-stack mixer was designed to operate at <20 GHz RF frequencies, a Keysight N9030B 26.5 GHz spectrum analyzer was used to measure the output power. For upconversion, a high IF power was generated with a ZHL-10W-2G+ power amplifier which operates over 1-2 GHz and saturates above 40 dBm. Additional calibration was performed manually with an HP 437B power meter to confirm the input and output power. For IIP3, two tones were generated from combining the 40 dBm amplifier output with a Keysight N5183B MXG which has a maximum linear power output of 26.5 dBm at 1 GHz. Both sources were passed through CF1020 circulators before being combined to prevent source cross talk. This setup has an upper limit of 60 dBm IIP3 and a dynamic range of >100 dB.

Down conversion measurements require an additional wideband input amplifier because the cable losses become
significant at high frequency, and high RF power is needed across the entire bandwidth. An HMC998APM5E DC-20-GHz, 33-dBm amplifier minimizes distortion for down conversion $P_{1dB}$ measurements. Similarly, the output of the test setup was calibrated for power and the mixer output was measured on the spectrum analyser. Two tones were generated by the IM3 personality of the Keysight N5277A VNA and combined using a Marki Microwaves PR-OR636 power combiner/divider. The same Keysight N9030B spectrum analyzer was used to measure the output tones at the required dynamic range.

Figure 11 plots the simulated and measured up-conversion characterization for the six-stack mixer and Fig. 12 plots the down-conversion operation. In both cases the conversion loss is less than 10 dB from 2-18 GHz at the RF port where the IF frequency was 1 GHz. The mixer has a 1-dB compression power ranging from 18-22 dBm in both up and down conversion mode with IIP3 ranging from 25-33 dBm. There is some uncertainty in measuring IIP3 as the fitting of the IM3 tones has a large influence on the reported value. For this reason, Figs. 11 and 12 present upper and lower bounds for best and worst case extrapolation of the IM3 products, as well as a typical case. An example of this is shown in Fig. 13 where the same data is interpreted two ways. In Fig. 13 (a), the data is fit without constraining the slope of the IM3 products, representing a true interpretation of the data, while in Fig. 13 (b) the slope is constrained to be exactly 3, as is typically done. In Figs. 11 and 12, the upper bound represents fitting with a slope < 3, the lower bound represents fitting with slope > 3, and the data line in the middle forces the slope in the fit to be bounded between 2.7 and 3.3.

The simulated and measured isolation characteristics are shown in Fig. 14. LO-RF and LO-IF isolation remains remarkably high, better than 40 dB across the band and better than 50 dB below 10 GHz, owing to the triple balanced nature of the mixer. Because the mixer can be used in both up and down conversion modes, the LO-RF and LO-IF measurements are the same. Monte Carlo simulations of device mismatch and process variation show that
FIGURE 15. Simulated and measured power consumption for the six-stack mixer across the bandwidth.

FIGURE 16. LO input power calibrated to the pads of six-stack mixer for the previous measurements (Figs. 11–15) plotted alongside the simulated LO power required to saturate the on-chip amplifiers across the design frequency.

FIGURE 17. Simulated and measured insertion loss, compression point, and IIP3 for the two-stack mixer in up conversion mode to a 1 GHz IF.

B. TWO-STACK MEASUREMENTS

The measurement of the two-stack mixer required a slightly different setup because the bandwidth was larger but the linearity lower when compared with the six-stack device. The same setup as the six-stack measurements can be used when creating signals for up-conversion $P_{1dB}$ and IIP3, however the output was measured with the spectrum analyzer personality of the Keysight N5277A 67 GHz VNA. While the VNA has a lower dynamic range than the stand-alone spectrum analyser, it was sufficient to measure the outputs of the less linear two-stack mixer. The up-conversion metrics are shown in Fig. 17.

To measure down-conversion $P_{1dB}$, the HMC998APM5E amplifier was used from 2-20 GHz and an ADPA7007 18-44 GHz, 30 dBm amplifier generated sufficient power to cause compression. The IIP3 was measured with the VNA IM3 measurement tool and metrics are reported in Fig. 18.

LO isolation is shown in Fig. 19. Again, the isolation is better than 40 dB across the entire band. The power consumption is plotted as a function of frequency in Fig. 20.

To calculate the LO power reaching the mixers on chip, the external cables and balun losses were measured. The loss was subtracted from the output power of the microwave generator to determine the amount of power reaching the LO inputs on the mixer chip assuming good matching. Additionally, the LO input power required to saturate the on-chip LO amplifiers were simulated and the two values are plotted as a function of frequency in Fig. 16. In the case of the six-stack mixer, the LO power was sufficient to saturate the amplifiers across the bandwidth of 2-23 GHz. While the gain of the on-chip amplifiers rolls off, in all cases, < 5 dBm of input LO power is needed for optimum performance with this design.

C. COMPARISON TO OTHER MIXER TECHNOLOGIES

Finally, Table 5 compares the measured parameters of each of the two mixer variations, two-stack and six-stack, to
TABLE 5. Comparison of RF and MMW mixers.

| Technology       | Architecture | Bandwidth (GHz) | Conv. Loss (dB) | $P_{1\text{dB}}$ (dBm) | Typ. IIP3 (dBm) | Power Cons. (W) | LO Input (dBm) | LO Isolation (dB) |
|------------------|--------------|-----------------|-----------------|-------------------------|-----------------|----------------|----------------|------------------|
| [4] 180nm CMOS   | PET Ring     | 1-11            | 7               | 5                       | 12              | N/A            | 10             | N/A              |
| [5] 180nm CMOS   | PET Ring     | 15-50           | 14.5            | 10                      | N/A             | N/A            | 10             | N/A              |
| [6] 180nm CMOS   | PET Ring     | 2-30            | 18.5            | N/A                     | 4               | 0.02           | 3              | >50              |
| [7] 180nm CMOS   | Distributed | 5-45            | 11-13           | N/A                     | 16              | 0.0014         | 8              | >30              |
| [8] 28nm CMOS    | N-path       | 10-35           | 12.5 (NF)       | 1.5                     | 147/12.5        | N/A            | N/A            | N/A              |
| [10] 28nm CMOS   | N-path       | 0.1-2           | 6.3 (NF)        | -10                     | 5               | 3.0-3.96       | N/A            | N/A              |
| [16] SiC Schottky| Diode Ring   | 2-2.3           | 15              | 23                      | 35              | N/A            | 28             | 7                |
| [17] GaAs Schottky| Diode Ring  | 3-30/5-50       | 9.5             | 12                      | 25              | N/A            | 15             | 30               |
| [19] 0.25μm GaAs pHEMT | PET Ring | 3-10            | 9               | 18                      | 27              | N/A            | 16             | 30               |
| [20] 0.15μm GaAs pHEMT | PET Ring | 29              | 12              | N/A                     | 28              | N/A            | 17             | N/A              |
| [21] 0.25μm GaAs pHEMT | Dist. PET  | 2-20            | 6-10            | N/A                     | 33              | 0.026          | 18             | 34               |
| [22] 0.5μm AlGaN/GaN HEMT | Single FET | 5               | 7.3            | >20                     | 35              | N/A            | 30             | N/A              |
| [25] 0.7μm AlGaN/GaN HEMT | Single FET | 8-14           | 6.9            | ~10                     | 30              | N/A            | 23             | N/A              |
| [24] 100nm GaN-on-Si HEMT | Single Bal. | 26-31          | 11              | 11                      | 22              | N/A            | 12             | 29               |
| [50] 32nm SOI    | Harm. Rej.   | 0.05-6         | 6.6-10.8        | 7                       | 20              | 0.03-0.136     | N/A            | >50              |
| [51] 45nm SOI    | Harm. Rej.   | 0.8-3          | 7-10            | 6                       | 22              | 0.04-0.16      | N/A            | N/A              |
| Six-Stack       | DiSoC         | 1-18           | 6.5-10          | 22                      | 31              | 0.8-1         | <5             | >40              |
| Two-Stack       | DiSoC         | 1-30           | 6-10            | 25                      | 3-0.4          | <5             | >35            |                  |

When comparing to other silicon architectures, the presented mixers achieve record IIP3 across a wide bandwidth. A key consideration when comparing measured IIP3 values is the slope and extrapolation variance discussed earlier in this section. Most previous works do not discuss the extrapolation techniques, allowing for measured IIP3s > 9.6dB + $P_{1\text{dB}}$. Additionally, peaking IIP3 values that exceed this relationship are often highly bias/temperature/voltage dependent and
may not be stable across a variety of operating conditions. As such, measuring $P_{1dB}$ is a more robust measure of linearity for comparison as it does not depend on fine cancellation of 3rd harmonic terms. In this sense, the presented mixers exceed $P_{1dB}$ values of any other CMOS mixer by more than 15 dB. Historically, SOI over bulk CMOS has not provided a particular advantage in linearity, as seen when comparing [4] and [5] with [50] and [51], and therefore adapting the presented architectures to bulk CMOS could be valuable for reducing the cost of the designs in future work [52].

When comparing to other technology approaches, such as GaAs Schottky diodes, pHEMTs, or AlGaN/GaN HEMTs, the presented mixers exceed the performance of all the compared designs, other than [16], which has comparable performance. However, in the case of [16] the LO input power is 28 dBm. When considering the amplifier that would be required to generate such a high LO power, along with the losses of chip interconnects and cables, the presented mixers with integrated drivers will be significantly more efficient, consuming a maximum total power of 30 dBm in the LO amplifiers, even at 26 GHz. This advantage of integration will open many future opportunities of integrated, high efficiency systems.

CONCLUSION

This work improves the linearity of a SOI CMOS switch without impacting the fundamental device switching speed by using an extrinsic gate capacitor and device stacking along with co-designed on-chip driver amplifiers. These switches were incorporated into two versions of microwave and millimeter-wave mixers: one with compression points above 20 dBm and intermodulation distortion IIP3 greater than 30 dBm across 18 GHz of bandwidth, the other with compression points up to 18 dBm and IIP3 greater than 25 dBm across a bandwidth of 30 GHz. The linearity exceeds earlier work in CMOS or SOI by more than an order of magnitude better than existing CMOS or SOI for similar bandwidth, and matches or outperforms monolithic mixers designed in III-V processes. The high level of integration in CMOS SOI eliminates interconnects to improve the overall system efficiency compared to designs with external LO amplifiers. These results demonstrate the potential for high-performance microwave and millimeter-wave mixers.

ACKNOWLEDGMENT

The authors would like to thank the support of N. Cahoon and A. Sharma of GlobalFoundries for providing access to GF 45RFSOI chip fabrication. Additional thanks to Christopher Marki and Marki Microwaves for their support in test and measurement. Finally, the authors acknowledge Integrand Software (now part of Cadence) for access to EMX for electro-magnetic modeling.

REFERENCES

[1] H.-W. Kim et al., “Design and analysis of CMOS T/R switches with the impedance transformation technique,” IEEE Microw. Wireless Compon. Lett., vol. 27, no. 12, pp. 1137–1139, Dec. 2017.
[2] C. Hill, C. S. Levy, H. AlShammary, A. Hamza, and J. F. Buckwalter, “RF Watt-level low-insertion-loss high-bandwidth SOI CMOS switches,” IEEE Trans. Microw. Theory Techn., vol. 66, no. 12, pp. 5724–5736, Dec. 2018.
[3] M. Parlik and J. F. Buckwalter, “A passive IQ millimeter-wave mixer and switch in 45-nm CMOS SOI,” IEEE Trans. Microw. Theory Techn., vol. 61, no. 3, pp. 1131–1139, Mar. 2013.
[4] T. Chang and J. Lin, “1–11 GHz ultra-wideband resistive mixer in 0.18-μm CMOS technology,” in Proc. IEEE Radio Freq. Int. Circuits Symp., 2006, p. 4.
[5] J. Chen, C. Kuo, Y. Hsin, and H. Wang, “A 15–50 GHz broadband resistive FET ring mixer using 0.18-μm CMOS technology,” in IEEE MTT-S Int. Microw. Symp. Tech. Dig., 2010, pp. 784–787.
[6] H.-W. Wang, J.-H. Cheng, J.-Y. Zhong, T.-W. Huang, and J.-H. Tsai, “A 2–30 GHz ring mixer with active baluns in 0.18-μm CMOS technology for vital sign detection application,” in Proc. Eur. Microw. Conf., 2015, pp. 901–904.
[7] Y.-S. Lin, C.-L. Lu, and Y.-H. Wang, “A 5 to 45 GHz distributed mixer with Cascoded complementary switching pairs,” IEEE Microw. Wireless Compon. Lett., vol. 23, no. 9, pp. 495–497, Sep. 2013.
[8] S. Krishnamurthy and A. M. Niknejad, “10–35GHz passive mixer-first receiver achieving +14dBm in-band IIP3 for digital beam-forming arrays,” in Proc. IEEE Radio Freq. Integr. Circuits Symp., 2020, pp. 275–278.
[9] V. K. Purushothaman, E. A. M. Klumperink, R. Plompens, and B. Nauta, “Low-power high-linearity mixer-first receiver using implicit capacitive stacking with 3x voltage gain,” IEEE J. Solid-State Circuits, vol. 57, no. 1, pp. 245–259, Jan. 2022.
[10] Y. Lien, E. Klumperink, B. Tenbroek, J. Strange, and B. Nauta, “24.3 a high-linearity CMOS receiver achieving +44dBm IIP3 and +13dBm $P_{1dB}$ for SAW-less LTE radio,” in Proc. IEEE Int. Solid-State Circuits Conf., 2017, pp. 412–413.
[11] A. Ahmed, M.-Y. Huang, D. Munzer, and H. Wang, “A 43–97-GHz mixer-first front-end with quadrature input matching and on-chip image rejection,” IEEE J. Solid-State Circuits, vol. 56, no. 3, pp. 705–714, Mar. 2021.
[12] M. Sanduleanu, S. Reynolds, and J. Plouchart, “A 4GS/s, 8.45 ENOB and 5.71/2conversion, digital assisted, sampling system in 45nm CMOS SOI,” in Proc. IEEE Custom Integr. Circuits Conf., 2011, pp. 1–4.
[13] K. N. Madsen, T. D. Gathman, S. Daneshgar, T. C. Oh, J. C. Li, and J. F. Buckwalter, “A high-linearity, 30 GS/s track-and-hold amplifier and time interleaved sample-and-hold in an InP-on-CMOS process,” IEEE J. Solid-State Circuits, vol. 50, no. 11, pp. 2692–2702, Nov. 2015.
[14] C. Hill, A. Hamza, H. AlShammary, and J. F. Buckwalter, “Watt-level, direct RF modulation in CMOS SOI with pulse-encoded transitions for adjacent channel leakage reduction,” IEEE Trans. Microw. Theory Techn., vol. 67, no. 12, pp. 5315–5328, Dec. 2019.
[15] T. Dinc, A. Nagulu, and H. Krishnaswamy, “A millimeter-wave non-magnetic passive SOI CMOS circulator based on spatio-temporal conductivity modulation,” IEEE J. Solid-State Circuits, vol. 52, no. 12, pp. 3276–3292, Dec. 2017.
[16] M. Sudow, K. Andersson, P.-A. Nilsson, and N. Rorsman, “A highly linear double balanced Schottky diode S-band mixer,” IEEE Microw. Wireless Compon. Lett., vol. 16, no. 6, pp. 336–338, Jun. 2006.
[17] N. Drobotun, D. Danilov, and A. Drozdov, “A decade bandwidth mixers based on planar transformers and quasi-vertical Schottky diodes implemented in GaAs MMIC technology,” in Proc. Eur. Microw. Conf., 2021, pp. 957–960.
[18] S. Mohyuddin, G. H. Lee, D. H. Kim, I. B. Kim, H. C. Choi, and K. W. Kim, “A compact double-balanced diode ring mixer for wideband applications,” in Proc. Eur. Microw. Conf., 2018, pp. 648–651.
[19] Y. Pu, Z. Huang, S. Pan, and G. Wang, “A 3 GHz to 10 GHz GaAs double balanced mixer,” in Proc. IEEE Inf. Tech. Mechatronics Eng. Conf., 2017, pp. 1083–1086.
[20] M. S. Clements, A.-V. Pham, J. S. Sacks, and S. E. Avery, “Second-harmonic injection linearization of millimeter-wave FET resistive mixers,” IEEE Microw. Wireless Compon. Lett., vol. 29, no. 10, pp. 669–672, Oct. 2019.
[21] T. T. Nguyen, K. Fujii, and A.-V. Pham, “Highly linear distributed mixer in 0.25-μm enhancement-mode GaAs pHEMT technology,” IEEE Microw. Wireless Compon. Lett., vol. 27, no. 12, pp. 1116–1118, Dec. 2017.
K. K. W. Low, T. Kanar, S. Zihir, and G. M. Rebeiz, “A 17.7–20.2-GHz
CMOS 4-stack FET mixer,” IEEE Trans. Microw. Theory Techn., vol. 56, no. 10, pp. 2201–2206, Oct. 2008.

A. De Padova, P. E. Longhi, S. Colangeli, W. Ciccognani, and C. M. Thomas, “SiGe HBT N-path filter IC with fast discharge feature,” IEEE Solid-State Circuits Lett., vol. 6, no. 1, pp. 71–73.

K. Hill, C. S. Levy, H. Al Shammary, A. Hamza, and J. F. Buckwalter, “A single-ended resistive alGaN/GaN HEMT N-path mixer,” in Proc. IEEE Compound Semi. Integr. Circuits Symp., pp. 1–4.

K. Hill, C. S. Levy, H. Al Shammary, A. Hamza, and J. F. Buckwalter, “A 30.9 dBm, 300 MHz 45-nm SOI CMOS power modulator for spread-spectrum signal processing at the antenna,” in Proc. IEEE Int. Solid-State Circuits Conf., 2018, pp. 423–426.

R. Chen and H. Hashemi, “Passive coupled-switched-capacitor resonator-based reconﬁgurable RF front-end ﬁlters and duplexer,” in Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC), 2016, pp. 138–141.

K. K. W. Low, T. Kanar, S. Zihir, and G. M. Rebeiz, “A 17.7–20.2-GHz 1024-element K-band SATCOM phased-array receiver with 8.1-dB/K G/T, ±70° beam scanning, and high transmit isolation,” IEEE Trans. Microw. Theory Techn., vol. 59, no. 3, pp. 1769–1779, Mar. 2022.

B. Rupakula and G. M. Rebeiz, “Third-order intermodulation effects and system sensitivity degradation in receive-mode 5G phased arrays in the presence of multiple interferers,” IEEE Trans. Microw. Theory Techn., vol. 66, no. 12, pp. 5780–5795, Dec. 2018.

A. Alhamed, O. Kazan, G. Gültepe, and G. M. Rebeiz, “A multi-band/multistandard 15–57 GHz receive phased-array module based on 1 × 4 beamformer IC and supporting 5G NR FR2 operation,” IEEE Trans. Microw. Theory Techn., vol. 70, no. 3, pp. 1732–1744, Mar. 2022.

J. A. Jayamon, J. F. Buckwalter, and P. M. Asbeck, “Multigate-cell stacked FET design for millimeter-wave CMOS power amplifiers,” IEEE J. Solid-State Circuits, vol. 51, no. 9, pp. 2027–2039, Sep. 2016.

D. Manente et al., “A 28-GHz stacked power amplifier with 20.7-dBm output P_{1dB} in 28-nm bulk CMOS,” IEEE Solid-State Circuits Lett., vol. 3, no. 3, pp. 170–173, 2020.

K. Ning and J. F. Buckwalter, “An 18-dBm, 57 to 85-GHz, 4-stack FET power amplifier in 45-nm SOI CMOS,” in IEEE MTT-S Int. Microw. Symp. Tech. Dig., 2018, pp. 1453–1456.

P. M. Asbeck, N. Rostomyan, M. özem, B. Rabet, and J. A. Jayamon, “Power ampliﬁers for mm-Wave 5G applications: Technology comparisons and CMOS-SOI demonstration circuits,” IEEE Trans. Microw. Theory Techn., vol. 67, no. 7, pp. 3099–3109, Jul. 2019.

A. Safarian, A. Yazdi, and P. Heydari, “Design and analysis of an ultrawide-band distributed CMOS mixer,” IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 13, no. 5, pp. 618–629, May 2005.

P. Heydari, D. Lin, A. Shameli, and A. Yazdi, “Design of CMOS distributed circuits for multiband UWB wireless receivers [LNA and mixer],” in Proc. IEEE Radio Freq. Integr. Circuits Symp., 2005, pp. 695–698.

K. Andersson, V. Desmaris, J. Eriksson, N. Rorsman, and H. Zirath, “C-band linear resistive wide bandgap FET mixers,” in Proc. IEEE Int. Microw. Symp., vol. 2, 2003, pp. 1303–1306.

M. Sudow, K. Andersson, M. Fagerfink, M. Thorsell, P. Nilsson, and N. Rorsman, “A single-ended resistive alGaN/GaN HEMT MMIC mixer,” IEEE Trans. Microw. Theory Techn., vol. 56, no. 10, pp. 2201–2206, Oct. 2008.

A. De Padova, P. E. Longhi, S. Colangeli, W. Ciccognani, and L. E. Larson, “A 250 nm GaN N-path ﬁlter IC with +27 dBm blocker tolerance,” in Proc. IEEE Compound Semi. Integr. Circuits Symp., pp. 1–4.

K. Hill, C. S. Levy, H. Al Shammary, A. Hamza, and J. F. Buckwalter, “A 30.9 dBm, 300 MHz 45-nm SOI CMOS power modulator for spread-spectrum signal processing at the antenna,” in Proc. IEEE Int. Solid-State Circuits Conf., vol. 55, no. 11, pp. 790–793, Aug. 2020.

M. Ahn, B. S. Kim, C.-H. Lee, and J. Laskar, “A high power CMOS switch using substrate body switching in multistack structure,” IEEE Microwave. Compon. Lett., vol. 17, no. 9, pp. 682–684, Sep. 2007.

K. Hill and J. F. Buckwalter, “A 1-to-18GHz distributed-stacked-complementary triple-balanced passive mixer with up to 33dBm IIP3 and integrated LO driver in 45nm CMOS SOI,” in Proc. IEEE Int. Solid-State Circuits Conf., vol. 65, 2022, pp. 1–3.

K. Li, G. Freeman, M. Boenke, N. Cahoon, U. Kodak, and G. Rebeiz, “1W <0.9dB DC-20GHz T/R switch design with 45nm SOI process,” in Proc. IEEE Topical Meeting Silicon Monolithic Integ. Circuits RF Syst., 2017, pp. 57–59.

K. Li et al., “5G mm-wave front-end-module design with advanced SOI process,” in Proc. IEEE Int. Conf. ASIC, 2017, pp. 1017–1020.

H. Xu and K. Kenneth, “A 31.5-dBm bulk CMOS T/R switch using stacked transistors with sub-design-rule channel length in folded p-wells,” IEEE J. Solid-State Circuits, vol. 42, no. 11, pp. 2528–2534, Nov. 2007.

P. Katzin, B. Bedard, M. Shifrin, and Y. Ayasli, “High-speed, 100+W RF switches using GaAs MMICs,” IEEE Trans. Microw. Theory Techn., vol. 40, no. 11, pp. 1989–1996, Nov. 1992.

S. Nor et al., “A 1-bit digital transmit system using a 20-Gbps quadruple-cascade class-D digital power amplifier with 45nm SOI CMOS,” in Proc. IEEE Int. Solid-State Circuits Conf., 2019, pp. 734–737.

M. S. Dadash, D. Hame, and S. P. Voinigescu, “Large-swing 22nm Si/SiGe FDSOI stacked Cascades for 56GBaud drivers and 5G PAs,” in Proc. IEEE BiCMOS Compound Semic. Integr. Circuits Symp., 2018, pp. 267–270.

H. Alshammary, C. Hill, A. Hamza, and J. F. Buckwalter, “A code-domain RF signal processing front end with high self-interference rejection and power handling for simultaneous transmit and receive,” IEEE J. Solid-State Circuits, vol. 55, no. 5, pp. 1199–1211, May 2020.

K. Kibaroglu and G. M. Rebeiz, “A 0.05–6 GHz voltage-mode harmonic rejection mixer with up to 30 dBm in-band IIP3 and 35 dBc HRR in 32 nm SOI CMOS,” in Proc. IEEE Radio Freq. Integr. Circuits Symp., 2017, pp. 304–307.

O. Ei-Aaass, K. Kibaroglu, and G. M. Rebeiz, “A 16 path all-passive harmonic rejection mixer with watt-level in-band IIP3 in 45-nm CMOS SOI,” IEEE Microw. Wireless Compon. Lett., vol. 30, no. 8, pp. 790–793, Aug. 2020.

M. Ahn, B. S. Kim, C.-H. Lee, and J. Laskar, “A high power CMOS switch using substrate body switching in multistack structure,” IEEE Microwave. Compon. Lett., vol. 17, no. 9, pp. 682–684, Sep. 2007.