A Review of Technologies and Design Techniques of Millimeter-Wave Power Amplifiers

Vittorio Camarchia, Senior Member, IEEE, Roberto Quaglia, Member, IEEE, Anna Piacibello, Member, IEEE, Duy P. Nguyen, Member, IEEE, Hua Wang, Senior Member, IEEE, Anh-Vu Pham, Senior Member, IEEE

Abstract—This paper reviews the state-of-the-art of millimeter-wave power amplifiers, focussing on broadband design techniques. An overview of the main solid-state technologies is provided, including Si, GaAs, GaN and other III-V materials, and both field-effect and bipolar transistors. The most popular broadband design techniques are introduced, before critically comparing through the most relevant design examples found in the scientific literature. Given the wide breadth of applications that are foreseen to exploit the millimeter-wave (mm-wave) spectrum, this contribution will represent a valuable guide for designers who need a single reference before adventurous in the challenging task of mm-wave power amplifier design.

Index Terms—Broadband, millimeter wave, power amplifiers.

I. INTRODUCTION

The millimeter-wave (mm-wave) spectrum is attracting a great interest for applications such as 5G and future satellite communications that go well beyond the traditional niche of military and scientific use in terms of investments and potential revenues. The most attractive feature of mm-waves compared to the RF and microwave band is the huge spectrum availability that gives a great advantage in terms of capacity for telecommunication systems. Other advantages of mm-wave systems are the compact size of circuits, especially antennas, and the intrinsic easiness of frequency reuse thanks to the high free-space attenuations. There are also some advantages that are band-specific; for example, the very high attenuation in the 60GHz band due to oxygen absorption that enables intrinsically secure communications.

On the other hand, the very high frequency of operation poses significant challenges to system and circuit design. Similarly to what happens at lower frequency, one of the most critical circuit components is the power amplifier (PA). Its main figure of merit (FoM) is the output power that must be reached while respecting other requirements such as linearity and power consumption. At mm-wave, the power losses, the sensitivity to components’ tolerance, and the physical limitations of transistor technologies make the PA design even more challenging than at lower frequencies.

While these challenges represent an obstacle to an immediate deployment of functional and affordable mm-wave systems, they are very welcomed by the high-frequency PA community who has found an interesting space for the research and development of innovative technologies and techniques that can supplement the knowledge developed for lower frequency PAs.

This paper reviews the evolution and State-of-the-Art (SoA) of mm-wave solid-state PAs, with a particular emphasis on broadband solutions. The term broadband is relative and somehow subjective; this paper focusses on PA examples that generally achieve at least a 10% fractional bandwidth.

The paper is organized as follows. Section II gives a brief overview of the available solid state technologies for power amplification at mm-wave frequency, while Section III summarizes the most common broadband PA design techniques and describes their advantages and limitations. Sections IV to VII critically describe some relevant examples of PAs published in scientific journals or conferences, each section covering a portion of the mm-wave spectrum, according to IEEE radar bands [1]. Finally, Section VIII concludes the paper.

II. TECHNOLOGIES

One of the most important characteristics of a transmitter is the covered range, which depends on the product of the antenna gain and the PA output power and translates into circuit specifications in terms of PA performance. Although the total transmitter power can be combined from a number of PAs, it is still strictly related to the power density of the transistor technology adopted in the PA. For this reason, power density measured at the mm-wave frequency of interest is a crucial FoM for technology comparison. The other key characteristic of a PA is its ability to amplify, meaning that the gain the transistors can achieve while respecting other requirements such as linearity and maximum oscillation frequency (fMAX). However, power density plays again an important role, since displacement current effects result in an equivalent capacitance that is roughly proportional to the transistor active area. This means...
that for the same output power target the transistor with better power density will have lower capacitance hence, assuming the real part of the load is the same in both cases, it will be easier to use in high frequency, broadband PAs with good gain.

Power efficiency and linearity are notably linked; their trade-off, that arises when linearity is a strict system constraint, is difficult to optimize at transistor level, so circuit level solutions must be sought. However, if considered independently, linearity is deeply linked to the transistor characteristics such as the transconductance profile, the presence of traps, and the dynamic non-linearities. At the same time, the efficiency is also deeply affected by some transistor characteristics such as the ratio between drain bias and knee voltage. Moreover, the adoption of circuit solutions to improve efficiency, such as harmonic tuning, relies on the ability of the transistor to 1) generate drain current profiles with sufficient harmonic content, i.e., that are sharp enough, meaning high speed transistors, and 2) sustain the resulting voltage waveforms, i.e. with sufficiently high breakdown voltage.

By keeping in mind the above considerations, the survey result shown in Fig.1 is not surprising. It reports the output power versus frequency performance of an extended database of published PAs, extracted from the survey maintained at Georgia Tech [2]. In particular, Fig.1 focusses on PAs in Ka-band and above, while the original survey includes microwave and RF bands, as well as multipliers and oscillators. The different markers and colors refer to different technologies, with corresponding trend lines. To be noted; the trend lines are not regression lines, but a tentative to capture the best efforts for each technology, not considering however the single cases which exceed considerably the general trend.

What makes a successful technology or product is delivering value to the end-user, and no single semiconductor technology achieves the optimal balance in every application. III-V compound semiconductor technologies (sketch of typical device cross-sections in Fig.2), particularly GaAs and GaN, offer exceptional power density, gain and efficiency over a broad frequency range, providing system/design engineers with a wide trade-space to achieve optimal PA performance and economic value. In addition to their inherent performance advantages, and driven by smartphone demand, these technologies have achieved the economies of scale to satisfy the expected volume requirements of emerging mm-wave applications [3], [4]. On the other hand, silicon technologies (typical device cross-sections in Fig.3) can rely on great advantages in terms of integration and they are the technology of choice for medium and lower power levels.

### A. III-V technologies

GaN devices for mm-wave are High Electron Mobility Transistors (HEMTs) with aggressive gate scaling, ranging from 150/100 nm in commercially available processes down to 20 nm at research level [27], corresponding to $f_T$ between 50-70 GHz and up to 300-400 GHz, respectively. This scaling is achieved by adopting T-gate processes, and supported by ad-hoc design of the epitaxial layers and Schottky contacts. Comparable performance with InP and GaAs in terms of gain and noise figure can be observed while maintaining high breakdown voltage, of particular interest in transceiver integration [32], [33]. Fig.1 shows that the trends of GaN and GaAs converge at around 150 GHz, with the GaN trend line dropping faster than GaAs. The main reason for this is likely to be the higher maturity of GaAs technology that allows for a better current density. In our opinion, GaN technology still has margins for growth and will become the technology of choice when raw power at mm-wave is the main target, while HBTs scaling may make these technologies more attractive than HEMTs beyond 100 GHz. Substrate selection is a critical
issue in GaN. High performance GaN devices are grown on SiC substrate that is known for its high cost. As a consequence, several groups have focussed on GaN on Si substrate as viable solution for mass-market applications, and commercial products are already available. On the other hand, Si substrates have worse electrical and thermal performance compared to SiC substrates, and the compromise performance versus cost must be evaluated at system/product level. Also, the choice between high- or low-resistivity substrates is source of debate and research [34]. Low-voltage GaN, although still at research level, is interesting for mobile user applications since it would bring some of the advantages of III-V technology but at a reduced cost compared to GaAs, mainly thanks to the possibility of integrating on a Si substrate.

GaAs-based processes can still be considered the prime choice in mm-wave applications when a significant amount of power is needed. Fig. 1 shows how spread in both power and frequency the presence of GaAs PAs in literature is, a clear indicator of the great flexibility offered by GaAs that enables its use in different scenarios. The two main epitaxial structures used are pseudomorphic HEMTs (pHEMT) and metamorphic HEMTs (mHEMT). The gate lengths vary from 150 nm for K-bands applications, to 100 nm to cover applications up to 100 GHz for pHEMTs, while mHEMT-based PAs with the same gate length are found up to 160 GHz [35]. mHEMTs with gate length of 35 nm allow PA design up to 300 GHz [36].

InP- based technology is the only one currently able to provide PAs with >20 dBm output power above 200 GHz [37–41]. Bipolar transistors (Heterojunction Bipolar Transistors (HBT), and Double-Heterojunction (DHBTs)) have been used to design PAs up to 325 GHz, while HEMT-based PAs have been presented up to 185 GHz, with gate length of 80 nm [13]. The DHBT 500 nm and 250 nm technology nodes provide $f_T$ around 330 GHz and 400 GHz, respectively, the latter with an $f_{MAX}$ up to 700 GHz. HEMTs with gate length of 50 nm give $f_T$ around 600 GHz, however with similar $f_{MAX}$. The main limitation of InP is its cost, due to the difficulty in sourcing In and because it is a very brittle material, making its manufacturing and handling extremely perilous.

### B. Si and SiGe technologies

Si-based PAs clearly struggle to achieve the same power versus frequency performance of the III-V competition, but they are unbeatable in terms of cost (when mass-produced) and integrability. Also, they are competitive in terms of frequency alone, with examples up to 230 GHz in SiGe HBT [42] and up to 210 GHz in Si on Insulator (SOI) Complementary Metal-Oxide-Semiconductor (CMOS) [43].

CMOS technology for mm-wave PAs relies on extreme gate scaling, from 180 nm down to 28 nm and beyond. However, the frequency of operation is not always the main drive for choosing the gate length, but sometimes other considerations such as cost and the ability to integrate with the rest of the transceiver are guiding the selection. A clear example are 40 nm CMOS PAs and 45 nm CMOS SOI PAs that are being developed for 5G at around 28 GHz [30, 44], where longer gate CMOS PAs are still capable of operating, although with lower performance [45]. The main conundrum to be solved is related to the most evident drawback of scaling down the technology that is the breakdown voltage reduction. Although circuit design techniques might help mitigating this problem, such as transistor stacking, they never come for free and lead to increased parasitics that affect maximum frequency as well as achievable energy efficiency and linearity.

BiCMOS technology is intrinsically an example of integration, since it combines bipolar and field-effect transistors on
the same substrate. The technology nodes typically used at mm-wave range from 250 nm with $f_T$ around 220 GHz [46], down to 90 nm with $f_T$ of 300 GHz [47]. BiCMOS can be used for analogue Intermediate Frequency (IF) and baseband functions as well, and the CMOS part supports Digital Signal Processing (DSP) functions.

Integration is, in the view of the authors, the main advantage of Si technology. With up to ten metal levels, the capability of producing very complex circuits is unrivalled. Also, the performance of passive elements in terms of losses or quality factor is quite competitive [47]. The cost advantage however becomes clear only for mass produced parts, since the initial costs are much higher than in III-V technologies.

C. The 5G array scenarios

The high capacity services of 5G will use the mm-wave spectrum with the so called 5G New Radio (5G-NR) [48]. Fig. 4 reports the bands assigned by 3GPP (TS 38.101-2) to the Frequency Range 2 (FR2) of the 5G-NR. Two distinct set of bands can be observed. The first one, around 27 GHz, includes the n257, n258 and n261 bands; therefore, a transmitter aiming to cover this set will need to be designed over 24–30 GHz, corresponding to a 22% fractional bandwidth. The n260 band, around 38.5 GHz, extends from 37 to 40 GHz, requiring a 8% fractional bandwidth. The biggest change compared to existing systems at mm-wave is however the unprecedented requirement for very large instantaneous bandwidths. In fact, the channel widths supported are, for FR2, up to 400 MHz per carrier, with the option of having multiple carriers and with practical requirements for the base-station side up to 1 GHz being considered at the moment. This puts quite a lot of stress on the PA design from an equalisation point of view, since it cannot be expected from predistorters to correct excessive amounts of in-band gain variations. However, even more important is to guarantee uniform non-linear characteristics (AM-AM and AM-PM) on such a large bandwidth. At the same time, the baseband impedance must be controlled accurately up to a frequency of 3-5 times the channel bandwidth to eliminate memory effects. This means that bias networks must be resonant free potentially up to 5 GHz, stressing the requirements for on- and off-chip by-pass circuitry.

From a transmitter architecture point of view, base-stations will be based on active antenna arrays with a considerable number of elements to allow for beam-steering and/or beam-forming. The type of array architecture will be a deciding factor in the selection of the PA technology. In fact, if many antenna elements will be clustered and fed by a single PA, or in case of analog beam-forming, higher power technologies will be preferred, and GaN could be regarded as the winning solution. On completely the other hand, if each single radiator will be fed by a separate transceiver, as in digital beam-forming arrays, then Si-based technology makes more sense. In the short term, however, the system cost for the latter case is still too high due to its enormous complexity in both front-end and back-end, making this solution unpractical. Therefore, while current real-field 5G systems employ CMOS based active array configurations [49], hybrid beam-forming arrays [50] are likely to become the short term solution for 5G NR, and GaAs might actually be the winning solution for the PA (and perhaps the TX/RX switch and LNA), with the rest of the transceiver made on BiCMOS technologies.

III. DESIGN TECHNIQUES

High frequency PAs at these frequencies are mostly based on reduced current conduction angle modes, typically class AB to maintain a good compromise between linearity and efficiency. Class C is used on its own when linearity is not a concern, for example in radar and imaging, or in complex architectures such as the Doherty PA [51], [52].

Reactive matching is the standard solution for presenting the optimum load at the transistors. Due to the displacement current, each transistor will present an output capacitance; for FR2, up to 400 MHz per carrier, with the option of having multiple carriers and with practical requirements for the base-station side up to 1 GHz being considered at the moment. This puts quite a lot of stress on the PA design from an equalisation point of view, since it cannot be expected from predistorters to correct excessive amounts of in-band gain variations. However, even more important is to guarantee uniform non-linear characteristics (AM-AM and AM-PM) on such a large bandwidth. At the same time, the baseband impedance must be controlled accurately up to a frequency of 3-5 times the channel bandwidth to eliminate memory effects. This means that bias networks must be resonant free potentially up to 5 GHz, stressing the requirements for on- and off-chip by-pass circuitry.

From a transmitter architecture point of view, base-stations will be based on active antenna arrays with a considerable number of elements to allow for beam-steering and/or beam-forming. The type of array architecture will be a deciding factor in the selection of the PA technology. In fact, if many antenna elements will be clustered and fed by a single PA, or in case of analog beam-forming, higher power technologies will be preferred, and GaN could be regarded as the winning solution. On completely the other hand, if each single radiator will be fed by a separate transceiver, as in digital beam-forming arrays, then Si-based technology makes more sense. In the short term, however, the system cost for the latter case is still too high due to its enormous complexity in both front-end and back-end, making this solution unpractical. Therefore, while current real-field 5G systems employ CMOS based active array configurations [49], hybrid beam-forming arrays [50] are likely to become the short term solution for 5G NR, and GaAs might actually be the winning solution for the PA (and perhaps the TX/RX switch and LNA), with the rest of the transceiver made on BiCMOS technologies.
at RF and microwave frequencies [55], [56], it has already found application at mm-wave frequencies, as for example in [57], [58]. While for an application of textbook continuous mode theory the harmonics must be controlled precisely, the practical advantage that they bring is the understanding that matching requirements can be relaxed by accepting a slight and controlled reduction of performance [59]. One drawback of continuous modes, and of most of harmonic manipulation techniques, is that they require the active devices to sustain voltage waveforms with high peaks. While this might be acceptable in technologies like GaN where the breakdown is much larger than twice the bias voltage, it might represent a limitation for low voltage technologies where bias voltage reduction will lead to lower output power density.

Independently of the technology adopted, single transistors are mostly unable to provide the required level of output power, therefore some form of power combining is necessary. Multi-finger transistors are the obvious first step. GaN technology has the potential to provide Watt-level output power with a single multi-finger transistor cell. There is a limit, however, on the number of fingers that can be fed by a single gate access, due to the need of keeping parasitics under control, especially due to phasing and long source connections. Corporate combiners are the most common in III-V technology [35], [41], [60] but used also in Silicon [61], and they are a form of parallel (or current) combining. Based on fork combiners, normally in a binary arrangement, they provide both power combining and matching functions. A common way of designing these even-mode combiners is to start from a multi-stage ladder matching network that matches an equivalent device with the total periphery, as if the N individual devices to be combined were in parallel, and then split the components in symmetric arrangements to achieve the desired combiner (see Fig. 5(a)-(c)). To combine N identical devices having optimum load resistance $R_{\text{OPT}}$ and equivalent output capacitance $C_{\text{OUT}}$, the equivalent $N \times 1$ device with the total periphery will have optimum load resistance $R_{\text{OPT}}/N$ and equivalent output capacitance $N \times C_{\text{OUT}}$. This approach does not affect bandwidth, at least in principle, but it requires optimisation when real component will be placed in. A semi-lumped implementation is often preferred in MMICs, where the series inductors are replaced by equivalent distributed elements (TLs in Fig. 5(d)). This, however, has an impact on bandwidth, and requires a re-tuning of the values since the transmission lines will also absorb some capacitance. The main drawback of this solution is the extensive use of shunt capacitors whose accuracy and repeatability becomes a key requirement for the MMIC process. Off-chip power combining is also possible, and gives great flexibility. Bulk waveguide combiners are the most common choice, but sometimes off-chip combination on a substrate is also used (mainly to save area on precious MMIC substrates). An example using both waveguide and microstrip combiners can be seen in [62] for a V-band PA module. This solution is very interesting at higher mm-wave bands, where the dimensions of the air combiners is very small.

An interesting alternative for broadband design is the use of balanced PAs. While at microwave frequencies the use of Lange couplers is sometimes seen as a problem due to their size and aspect ratio, at mm-wave they are very small and they provide a much better design environment than corporate combiners. There is a technology challenge, however; wafers for mm-wave applications above Ka-band are usually thinned to 50 $\mu$m, meaning that width and spacing of Lange fingers must be brought down to a challenging 4 $\mu$m to achieve the right ratio between odd- and even-mode impedance in broadband design. Alternatives exist, and broadside coupling can be adopted instead of edge coupling to obtain 3 dB couplers; in III-V technology this might require to use soft dielectric layers to build the spacer between the coupled lines [36]. Normally, a combination between even mode combining and balanced structures is adopted, as depicted in Fig. 6. An advantage of this configuration if compared with the corporate combiner of Fig. 6 is the reduced impedance transformation ratio required, with possible improvement of bandwidth. This is true if the Fano limit is not being approached. The Fano limit depends solely on the $Q$–factor of the load to be matched that is almost independent on periphery, since the resistive part decreases, but the capacitive part increases, with increasing device size. The other advantages are related to the use of balanced structures in general, such as the inherent isolation between stages and the possibility of achieving simultaneous flat gain and good matching over a wide bandwidth. On the other hand, couplers generally introduce more loss than corporate combiners, and normally complicate the routing of the bias voltage since they naturally block DC between the stages. Push-pull is a differential power combining technique that ideally combines two class-B PAs to reconstruct a full-wave

![Simplified diagrams showing the design procedure of even-mode corporate combiners for the case $N = 4$, where the individual device has optimum load resistance $R_{\text{OPT}}$ and equivalent output capacitance $C_{\text{OUT}}$.](image-url)
at the output, therefore eliminating the need of even harmonic resonators and improving the bandwidth. Since most systems require a single-ended output, the output signal is obtained through a balun transformer. This normally brings other advantages since it theoretically provides a 25 $\Omega$ impedance at the amplifier ports, favouring the impedance ratio for matching. In reality, due to the limited coupling coefficient and self-loading, a 25 $\Omega$ impedance can be achieved but is not readily present. Numerous push-pull PA examples are found at mm-wave frequencies, in particular in GaAs [63], [64] and CMOS [65], [66]. On the other hand, power combining through transformers is a common feature in CMOS and BiCMOS design [67]–[69] thanks to the exploitation of the many metal layers available to realize 3D broadside transformers whose parasitics effects are reduced compared to planar transformers. Series combination through transformers is used also in III-V PAs [70].

The main limitation with parallel combining in low-voltage technologies is that optimum impedance decreases almost linearly with the increasing number of transistors, affecting the achievable bandwidth due to a larger impedance ratio to match. An intuitive strategy to reduce the impedance ratio is transistor stacking that ideally corresponds to voltage combining. In particular, the transistors are stacked, starting from a bottom stage with grounded source/emitter, and connecting on top a number of transistors. Parallel and stacking configurations are compared to a single stage configuration in the simplified schematics of Fig. 7. It can be noted how the idealised matching network connected at the output scales differently in the two cases, with the advantage of lowered impedance ratio and capacitance compared to the parallel combining case. This technique is almost a standard in CMOS [71] and BiCMOS PAs (with better stacking possible in SOI technologies), but it is used for GaAs as well [72], [73]. GaN stacking might be employed effectively for short gate transistors. For example, for 100 nm HEMTs, that work at around 15 V bias, stacking might become useful to reach an output power higher than 10 W.

Distributed amplifiers allow to achieve very broadband design by exploiting a very simple but powerful concept. When considering a capacitively-loaded transmission line, the bandwidth on which the structure behaves like a normal transmission line can be enhanced by increasing the number of loaded sections per unit length. The distributed amplifier uses this concept by loading two transmission lines, at input and output, with the transistors’ gate and drain capacitances, respectively. By equalizing the delay at input and output between each transistor, the distributed amplifier can achieve amplification on a very broad frequency band. As a consequence, they are the topology of choice for applications such as radar countermeasures and optoelectronic modulator drivers. Schematic
diagrams of three different distributed amplifier configurations are shown in Fig. 8. The uniform distributed amplifier provides good bandwidth but generally does not show good efficiency and does not maximise the output power. The reason is that the devices are not seeing a uniform voltage swing along the structure, since the voltage builds-up along the output line-up, with the one further away from the load having the lower voltage swing and therefore lower output power and efficiency. The non-uniform distributed power amplifier (NDPA), on the other hand, uses pieces of transmission line with different impedance (decreasing towards the load) to achieve a more uniform voltage swing along the structure and therefore a better output power and efficiency for all devices. Also, it is very common to substitute the transistors with Cascode cells in order to decrease the input capacitance of each amplifying unit, thus increasing gain and bandwidth of the distributed amplifier. Examples of distributed amplifiers can be found both in III-V [9], [74]–[78] and Silicon-based PAs [79]–[82].

The NDPA [83], [84] was initially proposed in the 1980s adopting GaAs devices. However, the low output power (typically lower than 30 dBm) limited its application and favoured reactively matched topologies. The advent of GaN MMICs increased the achievable output power on one side, but limited the bandwidth achievable by reactive matching due to the high-Q terminations required. This paved the way for the re-discovery of the NDPA topology for wideband GaN PAs.

There are also variants to the traditional distributed amplifier, for example the constructive wave amplifier, where drain and gate are loading the same transmission line [85].

IV. K-BANDS (18–40 GHz)

The low portion of the mm-wave spectrum (more properly cm-wave/mm-wave) is covered by the K-bands (K-band 18–26.5 GHz and Ka-band 26.5–40 GHz). In the past, these bands were exploited mainly for radar and backhaul applications, while more recently they have become attractive also for other strategic applications relevant to a large number of emerging mass-markets like satellite communications and, most importantly, 5G mobile communication.

Regarding 5G, the bands around 28 and 39 GHz have been identified (see e.g. the 3rd Generation Partnership Project (3GPP), released December 15, 2017 [86] and following) as the best candidates for the initial fixed deployment (commercial deployments expected to start in 2020).

Normally, conservative design approaches are used at these frequencies, resulting in mm-wave PAs with poor efficiency. The reason is that low-frequency well-established design and characterisation techniques, such as harmonic load-pull and waveform engineering, are still being explored at mm-wave. In terms of signal bandwidth, the present trend is toward bandwidths from 800 MHz to more than 1 GHz. This opens up the potential to utilize the same portion of the spectrum, and therefore the same equipment, not only for mobile broadband access, but also for backhaul, to simplify small cells deployment. On the other hand, the power consumption of individual wireless devices and back/base stations in 5G networks must be minimised to cope with the billions of connected devices and enable a reduction in energy usage over existing 4G networks [87].

The applications identified differ mainly in terms of power requirements, driving the choice of the adopted technology and the complexity of the PA architecture. They span from around 10–20 dBm for radars up to 37–50 dBm required by satellite applications. For 5G and backhaul, the power levels are still partially under definition, but two sub-ranges can be identified, depending on the transmitter architecture in terms of the analog or digital beam-forming scenarios. Table II reports, for the identified applications, the technologies that can achieve the required power at MMIC level, without the need of off-chip combiners.

A. Standard PAs

In 1986, Kim et al. presented the first millimeter-wave MMIC PAs at 44 GHz: a single stage amplifier with 10 dB gain and a parallel-combined topology with over 17 dBm saturated power and a power density of 23 dBm/mm adopting 300 nm GaAs FETs [90].

The first examples of K-bands MMIC PAs appeared around the early 90s do not target a specific application, but are suitable for radar, communications, smart munitions or local multipoint distribution services (LMDS). The technology node of choice in the first decade was 150 nm GaAs, which enabled to achieve from 18 dBm up to roughly 38 dBm above 30 GHz. The architectures of these examples are based on combined or balanced devices, mostly adopting two stages to increase gain. In [91], a complete on-chip transceiver working up to 40 GHz developed by TRW Inc. is reported. The PA achieves 18 dBm output power with 12 dB gain from 37 to 40 GHz. Fig. 9 reports, still by TRW Inc., a power module combining two GaAs MMIC power amplifiers (two stages) and a driver (two stages) to achieve the record output power of 37.7 dBm at 34.5 GHz, with a corresponding efficiency of 24% and a small-signal gain of 22 dB [92]. The technology was not mature enough to realize a single chip, mainly due to the losses of the combining and matching networks, which were realised off-chip as highlighted in Fig. 9.

Full on-chip power combining in GaAs is achieved in [93], [94]. Both PAs, developed by TRW Inc., adopt the same two-stage architecture reported in Fig. 10, where both driver and power stages are balanced amplifiers. To achieve the required

<table>
<thead>
<tr>
<th>Application</th>
<th>Freq. (GHz)</th>
<th>Power (dBm)</th>
<th>Technology/ies</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backhaul</td>
<td>18–26</td>
<td>23–36</td>
<td>GaAs, GaN</td>
<td>[88]</td>
</tr>
<tr>
<td>5G NR</td>
<td>24.25–27.50</td>
<td>10–20</td>
<td>CMOS, SiGe, GaAs, GaN</td>
<td>[86]</td>
</tr>
<tr>
<td>Radar</td>
<td>24.05–24.25</td>
<td>10–20</td>
<td>CMOS, SiGe</td>
<td>[1]</td>
</tr>
<tr>
<td>Satellite</td>
<td>17.3–22.0</td>
<td>37–50</td>
<td>GaAs, GaN</td>
<td>[89]</td>
</tr>
</tbody>
</table>

TABLE II
PA saturated output power requirements for commercial applications in K- and Ka-bands.
power level, devices are parallel-combined both in the driver and the power stages. In [93] the target application is LMDS, covering the 27.5–29.5 GHz band, where the PA maintains a saturated output power and PAE higher than 32 dBm and 32%, respectively. The MMIC in [94] is one of the first examples targeting ground terminal satellite applications. It achieves a saturated output power and PAE higher than 32 dBm and 22% over a 3 GHz bandwidth, from 29 to 32 GHz, demonstrating the feasibility to fully cover the uplink satellite band with this technology. This PA shows a substantial bandwidth increment compared to [93] without compromising the performance.

Broadband PAs achieving more than 10 GHz bandwidth can be found since 2000. Even if different design strategies are applied, e.g. distributed in [95], three-stage PA adopting single devices in [96] and two-stage balanced in [97], a common feature of these examples in 250 nm GaAs is the limited output power (below 23 dBm), possibly requiring off-chip combining to achieve Watt-level operation. A novel strategy to achieve broadband operation is proposed in [98], where the parasitic drain capacitance of the active devices is absorbed into the output matching/combining network, which also provides some immunity to process variations. The single-stage PA adopting two 150 nm GaAs pHEMTs parallel-combined with coupled resonators demonstrates 22.5 dBm output power, 30% PAE and 10 dB gain from 17 to 35 GHz.

A lower power (17 dBm), though on a wider bandwidth (26–40 GHz), is achieved in [17] on 130 nm SiGe BiCMOS.

This two-stage differential PA features 10 dB gain and is compatible with satellite and ground-based communications and radar applications. However, its PAE is limited to 5% over the whole bandwidth due to the adoption of a linear class of operation (class-A). Conversely, the highest power in SiGe is demonstrated in [15], a combined PA which can ensure around 30 dBm output power with PAE higher than 13% from 20 to 28 GHz. Even wider bandwidth (12–40 GHz) is achieved by adopting the distributed topology in 130 nm SiGe BiCMOS [18], which achieves 19 dBm output power while maintaining PAE higher than 8%.

The advantages of the rising GaN technology in terms of power density directly reflect on the performance of [99] (Fig. 11). This two-stage distributed PA in 100 nm GaN/Si covers the 6–37 GHz band with 10 dB gain and output power as high as 32 dBm. Even if the technology was still novel at the time, it achieved a PAE (7%) comparable with the more consolidated SiGe examples. GaN MMIC PAs in K-bands were first demonstrated in 2004 [25], where a 150 nm GaN/SiC ($f_T > 80$ GHz, $f_{MAX} > 120$ GHz) process was adopted to realize two similar single-stage prototypes, one in coplanar waveguides (CPW) and one in microstrip, with 4 parallel-combined devices achieving around 32 dBm output power with more than 15% PAE and 8 dB gain at 33 GHz. The evolution of the GaN technology allowed, some years later, to realise more complex structures, both in terms of number of stages and architecture. In 2012, TriQuint presented a 3-stage single-ended PA and its corresponding balanced structure [100] (Fig. 12). The balanced PA maintains output power in excess of 39 dBm, PAE higher than 23% and small-signal gain higher than 20 dB in the range 25–30 GHz, almost doubling the power of the single-ended PA with a limited impact on PAE over the same bandwidth.

In 2015, Northrop Grumman achieved a significant improvement in performance adopting a two-stage structure combining 16 devices in 200 nm GaN/SiC HEMT technology, as shown in Fig. 13. The PA achieves 45 dBm between 26 and 30 GHz, with a PAE higher than 30% and a small-signal gain around 20 dB [23]. The highest power achieved in GaN is in excess of 46 dBm [101], with similar gain and slightly lower PAE over the same bandwidth. In less than 10 years GaN filled the technological gap with respect to GaAs. In fact, the highest power achieved in K-bands by GaAs MMIC PAs is of the order of 38 dBm at 40 GHz [102]. To reach power levels compatible with applications currently covered by travelling wave tubes, of the order of 45 dBm, which is feasible at MMIC level with GaN, GaAs technology requires off-chip combination of
several MMICs, as in [103].

For several years, since its introduction, GaN HEMTs have been fabricated on SiC substrates. Despite their significant cost and limited availability, SiC substrates have enabled complete modules with superior performance with respect to any competitors (mainly GaAs), especially for applications where absolute performance is more important than cost, such as military, satellite and backhaul. A recent example of a broadband, general purpose PA is presented in [26] (see Fig. 14). The PA shows output power in excess of 37.8 dBm, PAE of the order of 25% and small-signal gain of 22 dB between 26 and 35 GHz, thus covering most of the Ka-band. Lately, especially in the framework of analog beam-forming for 5G networks, Si substrates have been explored due to the lower cost and easier integrability of the PA with the digital part of the transceiver, despite having worse thermal properties. In 2018, OMMIC presented a 3-stage combined PA targeting 33 dBm output power and 20 dB gain at 40 GHz adopting its 100 nm GaN/Si process [22], shown in Fig. 15.

For applications requiring power levels limited to around 20 dBm, K-bands PAs in Si technology appeared around 2005. Before the advent of 5G, the absence of a high-volume application limited it to few examples demonstrated mainly at research level, with PAE mostly limited to below 20% at saturation. Around 2015, the great interest for 5G active antenna arrays exploiting digital beam-forming has boosted the number of solutions proposed based on Si technology, which is favourable in terms of integration and unit cost for mass-applications, and compatible in terms of performance. Stacked transistor cells have been often adopted instead of single common-source and cascode topologies for the PA basic cell.

In [107], the comparison of a common source and a two-stacked PAs designed in 28 nm bulk CMOS at 28 GHz is presented. The two-stacked PA allows for higher gain (13.6 dB) and saturated power (19.8 dBm) in the same chip area (0.28 mm²) as the common source stage. High linearity is achieved using a deep class-AB bias point and appropriate harmonic control circuit, demonstrating 25% PAE and -33 dBc adjacent channel leakage ratio (ACLR-UTRA) at 14.6 dBm average output power with an LTE signal with 16-quadrature amplitude modulation (QAM), 7.5 dB PAPR and 20 MHz bandwidth.

The best performance in terms of saturated power (of the order of 25–26 dBm) and bandwidth is achieved by combined structures [108], [110]. In [108], 4 cascode cells in 90 nm CMOS are combined with transformers as 2 differential amplifier cells, resulting in an extremely compact structure, with a saturated PAE of 27% and gain of 15 dB from 22 to 30 GHz. In [110], a stacked topology in 45 nm SOI CMOS is adopted in a two-stage configuration for each basic cell of the 8-combined PA. The resulting efficiency in the range 33–45 GHz is of the order of 10%.

The inherent low efficiency of Si-based PAs has an even higher impact when non-constant envelope modulation schemes are adopted, like in 5G systems, where the PA is operated most of the time around 6–9 dB of back-off. The
TABLE III
SOA K-bands standard PAs.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>Technology</th>
<th>Freq. (GHz)</th>
<th>Topology</th>
<th>N. of stages</th>
<th>P_{SAT} (dBm)</th>
<th>PAE_{SAT} (%)</th>
<th>SS Gain (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[104]</td>
<td>2003</td>
<td>100 nm GaAs mHEMT</td>
<td>32</td>
<td>combined</td>
<td>2</td>
<td>35</td>
<td>40</td>
<td>24</td>
</tr>
<tr>
<td>[105]</td>
<td>2004</td>
<td>150 nm GaAs mHEMT</td>
<td>18–40</td>
<td>single-ended</td>
<td>2</td>
<td>9.5</td>
<td>-</td>
<td>20</td>
</tr>
<tr>
<td>[102]</td>
<td>2010</td>
<td>300 nm GaAs pHEMT</td>
<td>40</td>
<td>combined</td>
<td>2</td>
<td>32.9</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>[98]</td>
<td>2014</td>
<td>150 nm GaAs pHEMT</td>
<td>17–35</td>
<td>combined</td>
<td>1</td>
<td>22.5</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>[12]</td>
<td>2007</td>
<td>500 nm InP DHBT</td>
<td>38</td>
<td>cascode combined</td>
<td>2</td>
<td>30.6</td>
<td>28.6</td>
<td>15</td>
</tr>
<tr>
<td>[106]</td>
<td>2008</td>
<td>InP DHBT</td>
<td>20</td>
<td>combined</td>
<td>2</td>
<td>32</td>
<td>37.8</td>
<td>17</td>
</tr>
<tr>
<td>[17]</td>
<td>2007</td>
<td>130 nm SiGe BiCMOS</td>
<td>25–40</td>
<td>balanced</td>
<td>2</td>
<td>17</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>[15]</td>
<td>2016</td>
<td>1800 nm SiGe HBT</td>
<td>20–28</td>
<td>combined</td>
<td>2</td>
<td>29.5</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>[18]</td>
<td>2018</td>
<td>130 nm SiGe BiCMOS</td>
<td>12–40</td>
<td>distributed</td>
<td>4</td>
<td>19</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>[19]</td>
<td>2014</td>
<td>100 nm GaN HEMT</td>
<td>6–37</td>
<td>distributed</td>
<td>5</td>
<td>30</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>[23]</td>
<td>2015</td>
<td>200 nm GaN/SiC HEMT</td>
<td>26–30</td>
<td>combined</td>
<td>2</td>
<td>45.5</td>
<td>32</td>
<td>22</td>
</tr>
<tr>
<td>[22]</td>
<td>2018</td>
<td>100 nm GaN/SiC HEMT</td>
<td>37–43</td>
<td>combined</td>
<td>3</td>
<td>40</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>[26]</td>
<td>2019</td>
<td>100 nm GaN/SiC HEMT</td>
<td>26–35</td>
<td>combined</td>
<td>3</td>
<td>37.8</td>
<td>-</td>
<td>22</td>
</tr>
<tr>
<td>[107]</td>
<td>2016</td>
<td>28 nm bulk CMOS</td>
<td>28</td>
<td>stacked</td>
<td>2</td>
<td>19.8</td>
<td>43</td>
<td>13.6</td>
</tr>
<tr>
<td>[108]</td>
<td>2018</td>
<td>90 nm CMOS</td>
<td>22–30</td>
<td>cascode combined</td>
<td>2</td>
<td>25</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td>[109]</td>
<td>2020</td>
<td>45 nm CMOS SOI</td>
<td>24–40</td>
<td>differential cascode</td>
<td>2</td>
<td>19</td>
<td>36.6</td>
<td>12</td>
</tr>
</tbody>
</table>

*a* Distributed amplifier with 4 cells, each consisting of a cascode structure with one additional stacked transistor.

*b* Two cascaded distributed amplifiers with 5 cells each.

---

Fig. 15. Broadband 3-stage combined PA in 100 nm GaN/Si [22] © IEEE 2018.

---

press the memory effect generated by the non-linear mixing between the envelope and the fundamental. In CW operation, it achieves 14 dBm saturated power with a corresponding PAE of 20%, and a 22 dB small-signal gain. Remarkably, the third order intermodulation products are maintained under -30 dBc across 1 GHz bandwidth at an output power of 5 dBm.

The state of the art in terms of output power, efficiency and bandwidth for the PAs in the various analysed technologies is reported in Table III.

### B. Doherty PAs

The Doherty PA is a well-established back-off efficiency enhancement technique based on the load modulation principle [112], [113]. While it is widely and successfully applied at communications frequencies (sub-6 GHz), its exploitation for mm-wave bands is still at research level. This is especially critical for CMOS, where complex power combining strategies are typically required to reach the target power levels. On the other hand, the larger power densities of compound semiconductors technologies lead to simpler cells (no need of stacking/combining), thus allowing to exploit more complex topologies.

Around 2000, led by two factors, i.e. the III-V technology improvements and the introduction of complex modulation schemes for backhaul and satellite applications, the necessity and possibility for the first mm-wave MMIC Doherty PAs arose. In 1999, TriQuint [114] presented the first 17 GHz MMIC Doherty PA adopting a 250 nm GaAs process. Two single devices are employed as shown in Fig. 16 to provide 25 dBm output power with a PAE in excess of 40% and a 8 dB small-signal gain. In 2000, the first demonstration of a 20 GHz Doherty PA was given in [115]. It adopts a 150 nm GaAs pHEMT process and achieves similar performance to [114], namely 23 dBm output power, 32% PAE and 8.4 dB small-signal gain. These first works kept at a minimum the stage complexity, without inserting power combiners or driver
stages, leading to limited gain and output power levels. The promising back-off efficiency, higher than 20% at 6 dB OBO in both cases, demonstrated the feasibility of the Doherty concept at Ku- and K-bands, at least for narrowband operation. To develop these initial proofs-of-concept into amplifiers with gain and power levels compatible with the real-field applications, an advancement in the technology was required. At the same time, no strong push from the market existed yet. Around ten years later, these two conditions were verified and brought to multi-stage wideband K-band DPAs.

GaN technology reached the required maturity to develop wideband K-band DPAs around 2012, when TriQuint presented a 37 dBm two-stage DPA adopting 150 nm GaN/SiC HEMTs in the 21–24 GHz band [24]. Although the technology was still under development, this prototype showed very promising performance, achieving higher output power compared to the more mature GaAs technology without the need of power combining, thanks to the higher power density of GaN as well as the better thermal management of SiC substrates. Furthermore, from the design point of view, GaN devices present more favourable impedance levels allowing for simpler matching networks with lower losses and potentially wider bandwidth [116]. As a result, this preliminary example already showed a peak PAE higher than 42% over the whole bandwidth and in excess of 30% at 6 dB OBO at 23 GHz.

In [117] a two-stage Doherty adopting the TriQuint 150 nm GaAs is presented in the band 22.8-25.2 GHz for backhaul applications. The MMIC is one of the first GaAs DPAs in which drivers are embedded in both main and auxiliary branches of the Doherty configuration, as shown in Fig. 18. The paper clearly demonstrates that the position of the driver (in front of the Doherty stage (a) or inside each branch (b)) has a significant impact on the overall PAE, as reported in Fig. 17. In particular, a two-stage DPA in which the drivers are embedded within the Doherty architecture (b) shows a better PAE compared to the case where a single driver feeds the Doherty power stage (a), regardless of the driver efficiency. Moreover, even drivers with limited efficiency (around 20% in this case) are able to improve the PAE compared to the single-stage DPA, besides increasing the gain. On the other side, the design of the interstage matching networks between the drivers and the power stage is critical, especially when wideband operation is targeted, due to their effect on the signal de-phasing among the branches, which is a key aspect of the DPA operation. In this work no power combining at device level is implemented, but the device periphery of the final stage is significant.

Another work that aims at assessing the feasibility of the Doherty architecture in GaN towards mm-wave frequencies was reported in 2017 [118]. Because at that time the 150 nm GaN technology was not openly available, this work pushes the more stable 250 nm GaN/SiC UMS process at its limits realizing a single-stage 14–15 GHz Doherty PA for point-to-point radio applications. The resulting gain is clearly limited, but the efficiency is comparable to that achieved by state-of-the-art GaAs DPAs (both in saturation and in back-off) with an output power which is four times higher.

In the last few years, the 5G application of the 28 GHz band has given momentum to several new DPA contributions at research level. The two possible scenarios foreseen for 5G active antenna arrays, introduced in Sec. II-C, will require power levels of the order of 30–40 dBm for analog beam-forming and 10–20 dBm in the case of digital beam-forming for the single PA, thus making III-V compounds and Si techn-
TABLE IV
SOA K-BANDS DOHERTY PAs.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>Technology</th>
<th>Freq. (GHz)</th>
<th>Final stage topology</th>
<th>N. of stages</th>
<th>PSAT (dBm)</th>
<th>PAE SAT (%)</th>
<th>PAE OBO (%)</th>
<th>SS Gain (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[117]</td>
<td>2014</td>
<td>150 nm GaAs pHEMT</td>
<td>22.8-25.2</td>
<td>single device</td>
<td>2</td>
<td>29.9</td>
<td>25</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>[119]</td>
<td>2017</td>
<td>150 nm GaAs pHEMT</td>
<td>29.25-30.25</td>
<td>single device</td>
<td>2</td>
<td>27</td>
<td>35</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>[120]</td>
<td>2018</td>
<td>150 nm GaAs pHEMTs</td>
<td>29-31.8</td>
<td>single device</td>
<td>2</td>
<td>25.7</td>
<td>31</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>[121]</td>
<td>2018</td>
<td>150 nm GaAs pHEMTs</td>
<td>27.25-28.5</td>
<td>combined stacked</td>
<td>2</td>
<td>28</td>
<td>34</td>
<td>24</td>
<td>13</td>
</tr>
<tr>
<td>[122]</td>
<td>2019</td>
<td>150 nm GaAs pHEMTs</td>
<td>26.8-29.5</td>
<td>combined</td>
<td>2</td>
<td>25</td>
<td>35</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>[24]</td>
<td>2012</td>
<td>150 nm GaN/SiC HEMT</td>
<td>21-24</td>
<td>single device</td>
<td>2</td>
<td>37</td>
<td>42</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[123]</td>
<td>2019</td>
<td>100 nm GaN/Si HEMT</td>
<td>27.5-28.35</td>
<td>single device</td>
<td>2</td>
<td>32</td>
<td>25</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>[31]</td>
<td>2018</td>
<td>45 nm CMOS SOI</td>
<td>25-31</td>
<td>stacked</td>
<td>1</td>
<td>20</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>[30]</td>
<td>2019</td>
<td>45 nm CMOS SOI</td>
<td>27</td>
<td>cascode, mixed-signal</td>
<td>2</td>
<td>23.3</td>
<td>40.1</td>
<td>33.1</td>
<td>19.1</td>
</tr>
<tr>
<td>[20]</td>
<td>2019</td>
<td>130 nm SiGe BiCMOS</td>
<td>29-31</td>
<td>differential</td>
<td>2</td>
<td>16.8</td>
<td>20.3</td>
<td>13.9</td>
<td>15</td>
</tr>
<tr>
<td>[21]</td>
<td>2020</td>
<td>130 nm SiGe BiCMOS</td>
<td>24-30</td>
<td>differential</td>
<td>2</td>
<td>28</td>
<td>30</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

*PAE at 6 dB OBO, unless otherwise noted.

During the same years, Si-based DPAs (CMOS and SiGe BiCMOS) have become feasible for the 5G digital beam-forming scenario [19], [20], [30], [31], which requires power levels (10–20 dBm) low enough to be reached by low-complexity power combining structures. Until recently, the record bandwidth (> 5 GHz) belonged to CMOS PAs [31], with a resulting efficiency on the average comparable or lower than their compounds counterparts. On the other side, DPAs have much higher back-off efficiency than the standard PAs presented in Sec. IV-A in the same technology. Finally, a very recent 2-stage DPA in 130 nm SiGe BiCMOS is presented in [21], achieving 28 dBm saturated power, 30% peak PAE, and 20% PAE at 6 dB back-off from 24 GHz to 30 GHz. This represents the SoA output power among 28 GHz SiGe/C莫斯 K-bands PAs, effectively competing with compounds DPAs in terms of both output power and efficiency.

To complete the K-bands scenario, two other technologies that have been exploited at the early stages are the GaAs mHEMTs and InP-based devices. Both have been tested when the other technologies were still unable to provide the required power or gain, but they are now less researched.

One of the first examples of GaN DPA for 5G-NR has been recently presented in [123], and based on a GaN on Si technology. In the band 27.5-28.35 GHz, it shows 32 dBm output power, 25% saturated PAE and 28% back-off PAE. In this work, the advantages of GaN in terms of power density are not yet completely exploited due to the partial immaturity of the technology but the back-off efficiency is already promising, suggesting that in a short while the technology improvement will lead to interesting results.

One of the first examples of InP PA in K-band was presented in 1994 [124], with the state-of-the-art performance of 28 dBm output power and 31% PAE at 44.5 GHz. After the advent of the double heterojunction structure, several examples became...
available, such as [12], [106] which achieved output power in excess of 30 dBm with peak PAE higher than 30%. It progressively disappeared after 2010, moving to higher frequencies, mainly due to its high production costs, as shortly mentioned in Sec. II.

GaAs mHEMTs have also been used to take advantage of the higher gain of such technology. The first multi-watt mHEMT PA was presented by BAE Systems in [104], which demonstrated 35 dBm output power, 39% PAE and 24 dB linear gain with only two stages. A very wide bandwidth PA for radar applications is reported in [105]. It adopts a two-stage configuration where parallel resistive feedback is used to provide a flat gain response and to reduce the sensitivity to process variations. It achieves 20 dB of small-signal gain and around 10 dBm output power from 18 to 40 GHz.

The state of the art DPAs in K-bands are summarised in Table IV.

V. V-BAND (40–75 GHz)

The 40–75 GHz spectrum, commonly known as the “V-band” [125], is one of the most widely exploited mm-wave bands for high capacity mm-wave communications. Like other mm-wave bands, V-band wireless links also require unobstructed line-of-sight channels and exhibit considerable path loss that demands high antenna gain for compensation. In addition to the significant rain fade effect in V-band [126], it is noteworthy that the atmospheric oxygen absorption peaks at 60 GHz with a sea-level attenuation of 20 dB/km [125]. As a result, the V-band spectrum is primarily used for high-capacity and short-distance (less than 2 kilometres) communications. Over the 57–71 GHz FCC unlicensed band, the WiGig or IEEE 802.11ad standard was announced in 2009 with up to six contiguous 2.16 GHz-channels, targeting applications like wireless transmission of uncompressed UHD videos with up to 7 Gbit/s over 10 meters [127]. Following that, IEEE 802.11ay is proposed as the next-generation 60 GHz WiFi and the second WiGig standard. It supports a maximum bandwidth of 8.64 GHz by carrier aggregation/bonding as well as multi-stream MU-MIMO for up-to 100 Gbit/s and 300–500 meters of extended range for various applications, including mm-wave backhaul, fixed wireless access for homes/businesses [128]. In addition, the 37–50 GHz spectrum is proposed for the next-generation communication for non-geostationary satellite constellations.

These high-data-rate high-capacity communication applications largely drive the design of V-band power amplifiers, necessitating a balanced performance of output power, bandwidth, linearity, and energy efficiency. For compound semiconductor PAs, GaAs is a very popular technology of choice for V-band PAs. In the range 40–50 GHz sub-band, the best reported GaAs PA achieves a saturated power in excess of 34.5 dBm with peak PAE of around 25% and 13 dB power gain at 45 GHz [129]. To connect power devices with large output peripheries, [129] uses a transmission-line (T-line) based in-phase combiner on four PA cells, as shown in Fig. 21, with lower passive loss compared to a conventional Wilkinson combiner. In the range 60–75 GHz, reported GaAs PAs support 32.5 dBm saturated power and 25 dB gain at 60 GHz [130], and saturated power higher than 28 dBm and 26 dB power gain from 71 to 76 GHz [131]. T-line power combiners are extensively used in these high power PAs.

GaN has also been widely employed for V-band PAs. The recent PA in [22] that covers the upper portion of Ka-band and the lower portion of V-band, which has already been presented in Sec. IV (Fig. 15), shows remarkable performance. The PA in [132] realizes 33 dBm saturated power with 24.6% peak PAE and 12 dB power gain over the 70–86 GHz sub-band, covering both the up-link (71–76 GHz) and down-link (81–86 GHz) satellite communication bands.

Various InP technologies have been used for V-band PA designs targeting medium-power but high-efficiency applications. The PA in [124] uses InP HEMT technology and pushes the output power to 28 dBm at 44.5 GHz with a high PAE of 31%. At 60 GHz, an InP HEMT PA supports 23.5 dBm output with 43% PAE [133], which is also the highest reported PAE. At 71–76 GHz, the InP HBT PA shown in Fig. 22 realizes 26 dBm output power, 23% PAE and 10 dB gain [134]. It combines four output PA cells, each of which is formed by two multi-finger InP devices and output passive network that can carry high DC biasing current.

The increasing need for medium-power PAs to address array-based high-capacity short-range applications also stimulates the research and development on silicon-based V-band PAs. At upper V-band frequencies (60–75 GHz), advanced silicon processes still offer adequate power gain (e.g., $f_{MAX}$ of 350 GHz for Globalfoundries 45 nm CMOS SOI), while distributed passives and antennas become commensurate with...
on-chip active circuits due to the reduced signal wavelength; this interesting combination has open the door to a plethora of design innovations.

Boosting the output power of silicon PAs at this frequency is of a primary interest to ensure sufficient link budget and coverage. The on-chip passive power combiner is therefore often a key focus in V-band PA designs. The major consideration is how to best manage the impedance transformation, passive loss, and bandwidth of on-chip passive power combiners. Similar to lower frequency PAs, on-chip transformers are popular choices for medium-power applications (~ 20 dBm) with a compact form-factor [69]. To achieve even higher output power, the combiner’s scalability becomes critical, and on-chip T-line combiners are demonstrated for 60 GHz and above [136], [137]. In particular, [136] reports a 16-way T-line power combiner on 3-stage single-ended PA cores in 90 nm SiGe which achieves 27.3 dBm saturated power, 22.3 dBm power at 1 dB compression and 12.4% PAE with a 1.8 V supply. There are also explorations on the distributed active transformer (DAT) for series power combining that is especially useful for PAs using low-voltage silicon devices [138], [139]. However, using transformer magnetic coupling for DAT combiners fundamentally needs “circular-shape” output arrangement and results in substantially complicated input signal distribution, low area-efficiency, and limited scalability on the number of combining paths. Recently, a 60 GHz PA using an asymmetric coupler-based series power combiner is proposed and demonstrated [135] in 45 nm CMOS SOI process (Fig. 23). Cascading asymmetric couplers achieve DAT-based power combining by providing equal load impedance on identical sub-PA. The coupler structure also supports differential PA cores with broadband capacitive neutralization for higher PA core gain, stability, and output power. Moreover, this architecture achieves DAT-based series power combining with a “line-shaped” non-circular layout, and multiple coupler-based DAT PAs can be further power combined using zero-degree T-line combiner, essentially as hybrid series-parallel combining, to boost the PA output power. This 60 GHz CMOS SOI PA realizes on-chip power combining of 24 sub-PAs and achieves 30.1 dBm output power, 24.7 dB power gain, and 20.8% peak PAE, as well as 12 Gb/s 64-QAM and 8 Gb/s 16-QAM modulations with long-term reliability, leading the state of the art of high-power V-band PAs.

Besides power combiners, there is also extensive research on stacking Si PA devices to sustain a high supply voltage with minimum impedance transformation for high output power. Unlike conventional cascode PAs, stacked transistor PAs require appropriate voltage swings at the gate/base terminals of the stacked devices to ensure reliability [71], [110], [140]. In addition to reliability, managing the amplitude and phase of these gate/base voltage swings is highly critical yet non-trivial at V-band, which often limits the achievable PA efficiency and linearity in practice. The stacked PA topology has been explored in both SiGe and CMOS SOI technologies, with the CMOS SOI particularly conducive to stacking more devices due to the floating body. Moreover, stacked PAs have been reported for both class-AB linear PAs and quasi-switching class-E PAs for further efficiency enhancement.

Another technology direction is to improve the PA bandwidth and intrinsic linearity. Such broadband linear medium-power PAs are critical components for large-array systems with massive spatial power combining [65], [142], [143]. Like broadband PAs at lower frequency [53], high-order passive networks are typically used to provide a large-signal optimum load to the power devices over a wide frequency range [142]. For PA linearity, static biasing and dynamic biasing can be optimized to ensure flat AM-AM linearity. In CMOS PAs, using NMOS and PMOS to compensate for the device AM-PM nonlinearity has been a common practice, while for multi-stage PAs, using drivers and output stages for nonlinearity compensation is also widely exercised [144], [145].

There are also increasing explorations of boosting the back-off efficiency for V-band PAs. This is essential for high capacity wireless communication that support complex and spectrally efficient modulations with large PAPR. Reported approaches include Doherty PAs, Outphasing PAs, and an asymmetrical non-isolating combiner PA [143]. However, due to the high loss in passive networks and degraded active devices, e.g., gain, efficiency, and output impedance, most of these V-band PAs show very limited back-off efficiency enhancement compared to class-B PA operations and often exhibit degraded linearity, requiring extensive digital pre-
TABLE V
SOA V-BAND PAs.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>Technology</th>
<th>Freq. (GHz)</th>
<th>Topology</th>
<th>N. of stages</th>
<th>$P_{SAT}$ (dBm)</th>
<th>PAE$_{SAT}$ (%)</th>
<th>SS Gain (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[129]</td>
<td>2005</td>
<td>150 nm GaAs pHEMT</td>
<td>45</td>
<td>4-way combined</td>
<td>2</td>
<td>34.5</td>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>[130]</td>
<td>2005</td>
<td>150 nm GaAs pHEMT</td>
<td>60</td>
<td>8-way combined</td>
<td>2</td>
<td>32.5</td>
<td>25</td>
<td>-</td>
</tr>
<tr>
<td>[131]</td>
<td>2013</td>
<td>100 nm GaAs pHEMT</td>
<td>71–76</td>
<td>8-way combined</td>
<td>4</td>
<td>28</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>[132]</td>
<td>2018</td>
<td>70 nm GaN HEMT</td>
<td>70–86</td>
<td>8-way combined</td>
<td>4</td>
<td>30</td>
<td>36</td>
<td>16</td>
</tr>
<tr>
<td>[134]</td>
<td>1994</td>
<td>150 nm InP HEMT</td>
<td>1.44–1.5</td>
<td>single-stage</td>
<td>1</td>
<td>28</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>[135]</td>
<td>2015</td>
<td>150 nm InP HEMT</td>
<td>71–76</td>
<td>4-way combined</td>
<td>2</td>
<td>26</td>
<td>23</td>
<td>19</td>
</tr>
<tr>
<td>[136]</td>
<td>2016</td>
<td>90 nm SiGe</td>
<td>68–91</td>
<td>16-way combined</td>
<td>3</td>
<td>27.3</td>
<td>12.4</td>
<td>19.3</td>
</tr>
<tr>
<td>[137]</td>
<td>2014</td>
<td>28 nm bulk CMOS LP</td>
<td>40–67</td>
<td>differential</td>
<td>2</td>
<td>13</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>[139]</td>
<td>2015</td>
<td>45 nm CMOS SOI</td>
<td>56–63</td>
<td>24-way combined</td>
<td>3</td>
<td>28.5</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>[146]</td>
<td>2019</td>
<td>45 nm CMOS SOI</td>
<td>60</td>
<td>differential Doherty</td>
<td>3</td>
<td>20.1</td>
<td>26</td>
<td>13</td>
</tr>
</tbody>
</table>

*a* and 16.6% PAE at 7 dB OBO.

Fig. 25. 60 GHz multi-feed radiator in 45 nm CMOS SOI for on-antenna Doherty active load modulation and power combing [147] ©IEEE 2019.

distortion (DPD) in practice. A recent 60 GHz Doherty PA in 45 nm CMOS SOI uses on-chip coupler baluns to construct a low-loss Doherty load modulation network [146]. It measures 20.1 dBm saturated power and 26% PAE, as well as 16.6% PAE at 7 dB back-off with 1.45× PAE enhancement over a class-B PA, presently leading the V-band PAs with back-off efficiency enhancement.

Recently, there is a new trend on mm-wave frontend innovations that merges antennas together with frontend electronics to achieve both desired radiations and various new “on-antenna” functionalities. For example, antennas can be designed and driven by multiple electronics paths as “multi-feed antennas” (Fig. 24). It is shown that multi-feed antennas can achieve on-antenna power combining and inherent impedance transformation among the feeds before using any on-chip passive networks, which greatly improves the combining efficiency over conventional on-chip passive combiners. A 4-feed on-chip slot antenna is demonstrated to realize parallel power combining of total 16 sub-PAs at 60 GHz and achieve 27.9 dBm saturated power and 23.4% PAE on a 45 nm CMOS SOI process [141]. A near-field coupled slot antenna array achieves power combining of 8 sub-PAs for 27.4 dBm saturated power and 30.8% drain efficiency at 74 GHz in a 65 nm bulk CMOS process [148]. Moreover, multi-feed antennas serve as part of the Doherty or outphasing active load modulation network [147], [149], [150] (Fig. 25). [147] shows a 2-way Doherty radiator using a dual-feed on-chip loop antenna at 65 GHz, achieving 19.4 dBm saturated power and PAE of 28.3% and 20.1% at saturation and 6 dB back-off, realizing 1.46× PAE enhancement at 6 dB back-off over an ideal class-B PA. [149] shows that two near-field coupled multi-feed slot antennas realizes a 3-way Doherty radiator with 21.2 dBm saturated power, 20.7 dBm power at 1 dB compression, and PAE of 21.8%/20.6%/20.1%/19.3% at Psat/OP1dB/4dB-PBO/8dB-PBO, achieving 1.42× and 2.34× PAE enhancement at 4 dB-PBO and 8 dB-PBO over an ideal class-B PA. Both Doherty radiators/PAs realize the best back-off efficiency enhancement among reported V-band PAs. Moreover, compared to spatial polar combining with a very limited field-of-view (FoV) [151], both Doherty radiators/PAs first combine the output power of main and auxiliary paths before radiating it out, which supports full antenna FoV (beyond -45° to +45°) with multi-Gbit/s 64-QAM and 16-QAM modulations and no DPD.

The state of the art in terms of output power, efficiency and bandwidth for the various technologies previously mentioned is reported in Table V.

VI. W-BAND (75–110 GHz)

W-band frequency spectrum (75–110 GHz) has been deployed in a number of applications thanks to its low atmospheric attenuation and the capability of providing high data rate throughput. In particular, the 77 GHz band is used for automotive cruise control radars. The 94 GHz window is widely employed in mm-wave imaging applications, including defense, astronomy, and security systems. More importantly, the 71–76 GHz and 81–86 GHz segments are allocated for satellite services and long-haul transmission [152]. As the low-frequency spectrum becomes more crowded and suffers from orbit congestion, W-band satellite systems have recently gained much interest from research and industry. Furthermore, the recent 5G development has pushed the frequency of many applications such as automotive radars, long-haul optical
communications, and measurement instrumentation into the W-band spectrum [153]. One of the biggest challenges in solid-state W-band power amplifiers is the ability to provide high power with good efficiency. The difficulty is mainly due to the limitation of the semiconductor processes. Within the same semiconductor technology, the breakdown voltage of the transistors is typically reduced when the cut-off frequency $f_T$ increases. Moreover, high passive losses at W-band significantly degrade the amplifier efficiency. Nonetheless, in the past decades, W-band PAs have been demonstrated in various technologies: GaAs, GaN, InP, SiGe, and CMOS.

The very first W-band amplifiers were designed in GaAs and InP processes around the early 90s. In 1990, Advantek introduced the first single-chip GaAs amplifier operating from 85 to 95 GHz [155]. The work presented two circuits using a sub 200 nm MESFET and pHEMT devices with an $f_T$ of 100 GHz and $f_{MAX}$ of 200 GHz. Although the gain was still low, and efficiency was not reported due to measurement capability, the works laid out a solid foundation for W-band power amplifier design. The 0.1-$\mu$m GaAs T-gate technology, which was first developed by TRW Inc., is the most popular GaAs technology deployed in W-band. Using this T-gate technology, multi-stage amplifiers have been reported. Among these, the push-pull amplifier presented in 1995 in [63] achieved a record (at that time) 13 dB small-signal gain and 13.3% PAE at 18.8 dBm at 90 GHz. Most W-band power amplifiers reported in the 90s mainly focused on the 94 GHz band for military radar applications. Other frequencies in W-band have been later explored for automotive radar at 77 GHz [157], broadband mm-wave imaging 72–95 GHz [158], and the far-infrared and sub-millimeter telescope [159]. In particular, the authors in [159] demonstrate three amplifier modules that cover three different sub-bands: 72–81, 90–101, and 100–113 GHz. Regarding wideband applications, Fig. 26 presents a fully integrated MMIC that covers the whole W-band from 75 to 100 GHz [154]. The single chip can provide an average of 14.5 dBm output power across the entire bandwidth. Four identical MMICs in Fig. 26 are then combined using a 4-way septum power combiner to achieve up to 20 dBm output power with an associated gain of 15 dB over 75–110 GHz. Recently, an ultra-wideband amplifier has been presented in a 50 nm InAlAs/InGaAs mHEMT technology developed by Fraunhofer IAF with an $f_T$ of 375 GHz and $f_{MAX}$ of 670 GHz [156].

The schematic and chip photo of the 1.75 $\times$ 1 mm$^2$ MMIC amplifier based on stacked-FET unit cells are illustrated in Fig. 27. The prototype achieves 22 dBm output power, 16.8 dB small-signal gain, and a peak PAE of 10.7%. Most notably, the bandwidth is extended beyond W-band, covering from 65 to 125 GHz. Using a 0.1-$\mu$m GaAs pHEMT process with solder hot-via RF transition, the authors in [160] demonstrate a 28 dBm output power and a 22 dB small-signal gain from 81–86 GHz, which is the highest output power achieved by a single GaAs MMIC in W-band to date. Fig. 28 presents a 5-stage MMIC that exhibits a 27 dBm maximum power and 27 dB of gain from 92–102 GHz [161]. The measured 12.5% peak PAE is the highest PAE reported above 100 GHz using GaAs technologies.

Besides GaAs, InP HEMT and HBT processes have also been deployed in W-band PAs since the early 90s. In 1991, an InP MMIC presented by Device Laboratory using a 0.1-$\mu$m T-gate InP HEMT process achieves a maximum gain of 12 dB, and the bandwidth covers from 75 to 100 GHz [163]. As compared to GaAs, InP technologies generally provide slightly lower output power, but at higher efficiency. 20% PAE with an associated 26.3 dBm output power recorded in [164] is among the highest power reported to date using InP processes at W-band. Furthermore, using a 0.1-$\mu$m HEMT process, a finite-ground coplanar waveguide design, which
is shown in Fig. 29, exhibits an 18.2% PAE with 18.6 dBm output power and 14.3 dB gain at 94 GHz [14]. Recently, double heterojunction bipolar transistor (DHBT) technology, which has the advantage of providing better linearity and efficiency over HEMT devices, has become more popular among InP processes. Most notably, Northrop Grumman Corporation presents an InP MMIC amplifier using an advanced 250 nm InP DHBT [162]. As shown in Fig. 30, the work reports a remarkable 41.7% PAE with a saturated power of 12.4 dBm at 93 GHz, which is the highest efficiency W-band PA reported in open literature.

Although recent development in GaAs and InP technologies provides a great enhancement in power, bandwidth, and efficiency in W-band PAs, the maximum output power of a single MMIC is still limited below 1W. To achieve watt level performance at W-band, GaN MMIC is the only candidate. High-frequency GaN on SiC processes have been developed by HRL Lab since 2006 with the first W-band three-stage GaN MMIC reporting 25 dBm output power with 14% PAE at 80.5 GHz [165]. Recently, Fujitsu developed an 80 nm GaN HEMT process that observes among the highest power density of 3.6W/mm at W-band [166]. A $2 \times 1.8$ mm$^2$ MMIC shows a measured output power of 30.6 dBm with a small-signal gain of 18.9 dB and a peak PAE of 12.3% at 86 GHz. One of the most advanced GaN processes is the 40 nm T-gate process developed by HRL Lab which has an $f_T$ of 200 GHz and $f_{MAX}$ of 400 GHz with the breakdown voltage up to 40 V. The fabricated prototype achieves over 20 dB gain from 79–95 GHz, an output power of 31.3 dBm at 27% PAE [27]. In addition to conventional reactive matching amplifiers, several techniques have been introduced in GaN MMIC PAs to extend the bandwidth. For instance, Fig. 31 shows a traveling-wave three-stage GaN amplifier presented by Quinstar in [167] achieves more than 33 dBm power and at least 16 dB gain from 75 to 100 GHz. On the other hand, in [168], a broadband radial stub is deployed in a W-band PA using a 100 nm Fraunhofer IAF GaN HEMT process to cover the completed W-band from 70 to 110 GHz. The MMIC in Fig. 32 demonstrates a maximum power and gain of 28.6 dBm and 8.6 dB, respectively. Higher output power can also be achieved by employing off-chip waveguide combiners. For instance, an outstanding high power 45.6 dBm GaN amplifier module has been demonstrated from 75 to 100 GHz using septum combiners and 12-way radial combiners [22] [169].

Despite the capability to provide high output power over a wide frequency range, GaAs and GaN MMICs typically consume a large amount of dc power, have large chip size, and low level of integration. On the other hand, Silicon-based circuits, which have been dominated in commercial applications at RF frequencies thanks to their high level of integration and low cost, are more suitable for applications requiring below 20 dBm power. In particular, the 65 nm and
45 nm CMOS processes prove themselves to be better candidates for high gain and high efficiency W-band amplifiers. In particular, the authors in [172] present a 65 nm CMOS amplifier that exhibits a measured output power of 19.3 dBm with a gain of 24.2 dB and peak PAE of 19.2%. To overcome the low breakdown limitation of CMOS processes, stacked-FET is utilized in [173] to provide 19 dBm at 90 GHz. Fig. 33 presents a 45 nm SOI CMOS W-band amplifier using an 8-way zero-degree combiner [170]. The circuit achieves a maximum output power of 21.1 dBm with a gain of 10 dB. However, the peak PAE is only 5.2% due to the high losses from the power combiner. To minimize the losses, a broadband parallel-series power combiner is employed in a 40 nm CMOS PAs [174]. The PA biased at 0.9 V can deliver 20.9 dBm power at 22.3% PAE, which is among the highest PAE reported to date in W-band CMOS PAs. In terms of bandwidth, the full W-band CMOS PA in 65 nm CMOS has been reported in [171]. The MMIC PA shown in Fig. 34 exhibits an 18 dB small-signal gain with the bandwidth covers from 77–110 GHz. Similarly, 40 nm CMOS PA in [175] employs transformers and capacitor coupling as shown in Fig. 35 to achieve 18 dB gain from 76–106 GHz. Nonetheless, the former circuit only has a peak PAE of 4.5% while the latter shows 10% PAE. Spatially power-combined has also been demonstrated on a 2 x 4 SOI CMOS amplifiers array [176]. The total output power recorded is 24 dBm, which is the highest W-band output power from a single CMOS chip reported to date.

SiGe BiCMOS technology provides a good balance solution that can achieve higher breakdown voltage than CMOS processes, utilize the good linearity and efficiency BJT devices while still preserve the high level of integration and maintain a compact chip size.

A 68–91 GHz PA using an advanced 90 nm 9HP SiGe process from Global Foundries (previously known as IBM) has been demonstrated [136]. The prototype shown in Fig. 36 demonstrates a measured output power of 27.3 dBm with an associated PAE of 12.4% and a small-signal gain of 19.3 dB. The design is the highest power silicon-based W-band PA reported to date. It uses a 16-way T-line power combiner on 3-stage single-ended PA cores in 90 nm SiGe. To achieve high efficiency, the class-E amplifier design is investigated in [31]. The MMIC shown in Fig. 37 demonstrates an outstanding 40.4% PAE at 17.7 dBm output power at 93 GHz. Although the class-E switching technique can provide high PAE performance, the voltage swing typically exceeds the breakdown voltage of the transistor, posing an issue for long-term reliability.

Besides conventional amplifiers and power combining techniques, Doherty amplifiers have also been explored to improve efficiency at back-off power. In [68], the authors present a 40 nm CMOS Doherty amplifier at 77 GHz for base-station backhaul communications. Despite high losses of the silicon substrate, the circuit achieves a peak PAE of 12% at 16.2 dBm output power and a 5.7% PAE at 6-dB power back-off. Table VI summarizes the power, gain and efficiency performances of state-of-the-art W-band PAs using different technologies.

VII. ABOVE 110 GHz

In this review section, we focus on power amplifier technologies beyond 110 GHz, mostly between 110 GHz and 300 GHz. In this upper mm-wave frequency regime, the short
TABLE VI
SOA W-BANDS PAs.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>Technology</th>
<th>Freq. (GHz)</th>
<th>Topology</th>
<th>N. of stages</th>
<th>P_{SAT} (dBm)</th>
<th>PAE_{SAT} (%)</th>
<th>SS Gain (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[156]</td>
<td>2018</td>
<td>50 nm GaAs mHEMT</td>
<td>65–125</td>
<td>stacked</td>
<td>1</td>
<td>22</td>
<td>10.7</td>
<td>16.8</td>
</tr>
<tr>
<td>[161]</td>
<td>2018</td>
<td>100 nm GaAs pHEMT</td>
<td>92–102</td>
<td>single-ended</td>
<td>5</td>
<td>27</td>
<td>12.5</td>
<td>27</td>
</tr>
<tr>
<td>[164]</td>
<td>1999</td>
<td>150 nm InP HEMT</td>
<td>85–95</td>
<td>single-ended</td>
<td>2</td>
<td>26.3</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>[162]</td>
<td>2017</td>
<td>250 nm InP DHFET</td>
<td>95</td>
<td>single-ended</td>
<td>1</td>
<td>12.4</td>
<td>41.7</td>
<td>10</td>
</tr>
<tr>
<td>[27]</td>
<td>2014</td>
<td>40 nm GaN HEMT</td>
<td>65–110</td>
<td>single-ended</td>
<td>3</td>
<td>31.1</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>[161]</td>
<td>2015</td>
<td>100 nm GaN HEMT</td>
<td>75–100</td>
<td>travelling wave</td>
<td>3</td>
<td>34.7</td>
<td>12.1</td>
<td>16</td>
</tr>
<tr>
<td>[169]</td>
<td>2016</td>
<td>100 nm GaN HEMT</td>
<td>75–110</td>
<td>oilt chip</td>
<td>3</td>
<td>45.7</td>
<td>-</td>
<td>14</td>
</tr>
<tr>
<td>[172]</td>
<td>2012</td>
<td>65 nm CMOS</td>
<td>79</td>
<td>8-way combined</td>
<td>4</td>
<td>19.3</td>
<td>15.2</td>
<td>24.2</td>
</tr>
<tr>
<td>[174]</td>
<td>2014</td>
<td>40 nm CMOS</td>
<td>71–86</td>
<td>4-way combined</td>
<td>2</td>
<td>20.9</td>
<td>22</td>
<td>18.1</td>
</tr>
<tr>
<td>[156]</td>
<td>2016</td>
<td>90 nm SiGe BiCMOS</td>
<td>68–91</td>
<td>16-way combined</td>
<td>5</td>
<td>27.3</td>
<td>12.4</td>
<td>19.3</td>
</tr>
<tr>
<td>[177]</td>
<td>2015</td>
<td>130 nm SiGe BiCMOS</td>
<td>93</td>
<td>class-E</td>
<td>2</td>
<td>17.7</td>
<td>40.4</td>
<td>15</td>
</tr>
</tbody>
</table>

Fig. 38. Five-stage amplifiers in 150 nm T2 GaN process of HRL Laboratories: 27 dBm 98–122 GHz PA (top), and 29.5 dBm 102–118 GHz PA (bottom) [181] ©IEEE 2018.

Fig. 39. SSPA module based on the combination of four GaN MMIC power amplifiers in a hybrid assembly with H-tee waveguide combiner [181] ©IEEE 2018.

wavelength of the electromagnetic waves is attractive for high resolution radar, imaging systems to detect concealed weapons, synthetic aperture radar (SAR) systems to track maneuvering targets and image scenes on the battlefield [40], [178]. One example is the DARPA Video Synthetic Aperture Radar operating at 235 GHz and providing real-time high resolution video SAR imagery [179]. In March 2019, the Federal Communication Commission (FCC) voted to open up 21.2 GHz unlicensed spectrum between 116 GHz and 246 GHz [180]. The frequency bands above 100 GHz can potentially be used for 6G networks and wireless applications.

In this frequency regime, scaled field-plate GaN HEMTs are the leading technology that provides the highest power density and breakdown voltage [33], [181].

The process parameters of the scaled GaN HEMT processes available from HRL Lab [33] are provided in the comparison Table I. In addition to the 40 nm gate length T3 process already mentioned in Sec VI, the 20 nm T4 process features $f_T$ and $f_{MAX}$ of 329 GHz and 558 GHz, respectively.

More importantly, the breakdown voltage of the T4 process is 17 V, which is more than 10 times of those of scaled GaAs, InP, CMOS and SiGe technologies in the same frequency range [182]. Therefore, it is expected that GaN power amplifiers can provide watt-level output power in the frequency regime of 110 to 220 GHz. While GaN HEMT MMICs have been shown to have stellar parameters that are highly favorable for power amplifiers of greater than 30 dBm power beyond 110 GHz, only a few power amplifiers have been reported in open literature. In 2014, the 1st G-band 180–200 GHz GaN MMIC PA reported in [33] only achieves 14 dBm output power with an associated 4.5 dB gain. Recently, a PA at 115 GHz has also been reported using a 100 nm GaN HEMT process from Fraunhofer IAF [183]. The four-stage design achieves a very high gain of 34.8 dB and a maximum power of 20.3 dBm.

The highest output power single die GaN amplifiers that have been reported in literature in 2018 are the 29.5 dBm output power PA from 102 to 118 GHz, and the 27 dBm output power one from 98 to 122 GHz [181]. The two highest output power amplifiers, which are shown in Fig. 38, were designed by Quinstar and fabricated by HRL Laboratories, Malibu, CA using their 150 nm T2 GaN process on a 50 μm thick SiC substrate [33], [181]. The T2 process features a double-heterostructure field-effect transistor (DHFET) with a typical $I_{max}$ of 1 A/mm. The main DHFET used in the final stage has a total gate width of 100 μm consisting of four 25 μm unite fingers. The 100 μm DHFET achieves a measured gain of 8.44 dB at 87 GHz with 12 V drain voltage operation and a
power density of 1.8 W/mm. In order to achieve 20 dB gain, the authors have developed two five-stage power amplifiers (Fig. 38). The second PA has twice the total gate width from stages 3 to 5 to produce higher output power. For instance, the final stage of the 29.5 dBm PA employs eight 100 µm transistors. Four transistors are combined using T-junctions, and two four-transistor power cells are combined using a Wilkinson power combiner. Inter-stage matching networks use quarter-wave transformers separated by a series capacitor Y-inverter. The GaN MMIC has a chip size of 3.3 × 1.94 mm² and achieves a measured small-signal gain of 20 dB from 105 to 115 GHz. The measured output power at P5dB is around 28 to 29 dBm from 102 to 118 GHz with a PAE of 8%. The lower power PA achieves a measured power at 6 dB back-off of more than 25 dBm from 98 to 122 GHz. Furthermore, four GaN MMIC PAs of around 29 dBm output power were also used in a hybrid assembly with H-tee waveguide combiner (Fig. 39). The size of the solid-state power amplifier (SSPA) is 5.4 × 4.4 × 1.9 cm³. The GaN SSPA achieves a measured output power of 34.4 dBm at 114 GHz and greater than 33 dBm from 102 to 116 GHz. The associated measured gain is 23 dB, and efficiency is around 6%. The measurements were done with waveguide fixtures and included losses of the integrated waveguide to microstrip transitions. Fig. 39 presents the SSPA module.

Likewise, a number of papers have been reported on InP PAs from 110 to 260 GHz. The dominant technology in this frequency regime is the InP HBT although a few advanced InP HEMT processes are also being used [184]. Some of the high frequency amplifiers up to 260 GHz [11] are also summarised in Table I. The highest fMAX of the 128 nm InP HBT is 1.3 THz with a 3.3 V breakdown voltage, and a current density of 2.3 mA/µm emitter length.

Single chip power amplifiers from 17 dBm to 23 dBm up to 260 GHz have been widely reported in these processes [40], [184]-[186]. In particular, a sub-50 nm InP HEMT process is used in a 207–230 GHz MMIC that delivers a 17 dBm output power with 11.5 dB small signal gain and an associated 2.3% PAE [185]. On the other hand, a 22.5 dBm MMIC with 22 dB small signal gain has also been demonstrated in [182] using a 250 nm InP HBT process. Due to low current density and low voltage operation, InP MMIC amplifiers employ as many as 16 parallel HBTs in the output stage. Specifically, authors in [40] have demonstrated a 17–24 dBm output power PA from 180-265 GHz in a single InP chip. The PA has been designed in a 250 nm HBT process. The common-emitter and common base HBTs are used to form a cascode cell with a total emitter periphery of 4-fingers × 6 µm (Le = 24 µm) for each device. The 3-stage amplifier is shown in Fig. 40 in which the last stage has a total of 16 parallel transistor cells. The 2.1 and 4.1 on-chip power combiners are used in the last stage, having insertion losses of 0.5 dB and 0.55 dB, respectively. The amplifier has a measured gain of 30 dB at 220 GHz and maintains more than 20 dB of gain from 190 to 260 GHz. The collector voltage operation is from 1.65 V to 2.2 V. The amplifier achieves a maximum output power of 23.4 dBm and higher than 20 dBm from 185 to 255 GHz.

The measured PAE is between 1.3% to 4.1% within the same band. Using the same 250 nm InP HBT process, the latest D-band MMIC demonstrates a measured gain of 29.5 dB with an outstanding PAE of 10.5% at 140 GHz [186]. However, the output power is still limited to 150 mW.

The InP MMICs have also been used to implement a multi-deck PA. In particular, Fig. 41(a) shows a 8-way waveguide combiner employed in a module that combines 8 InP MMICs to form one amplifier deck [193]. A phase shift is incorporated at each input of the single MMIC to adjust the phase for maximizing the combining efficiency. Four amplifier decks are then combined using a 4:1 WR-4 waveguide combiner/divider. The four-deck power amplifier module (32 MMICs) shown in Fig. 41(b) achieved 28.5 dBm saturated power at 230 GHz, and more than 32 dB small signal gain from 200 to 260 GHz. Amplifiers using InP HEMTs with 25 nm gate length have also been demonstrated to have 9 B of gain at 1 THz [194].

Besides InP, GaAs mHEMT is another technology that can provide very high frequency operation. GaAs mHEMT power
TABLE VII
SOA PAs above 110 GHz.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Year</th>
<th>Technology</th>
<th>Freq. (GHz)</th>
<th>Topology</th>
<th>N. of stages</th>
<th>$P_{SAT}$ (dBm)</th>
<th>PAE$_{SAT}$ (%)</th>
<th>SS Gain (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[181]</td>
<td>2018</td>
<td>150 nm GaN HEMT</td>
<td>102–118</td>
<td>8-way combined</td>
<td>5</td>
<td>29.5</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>[183]</td>
<td>2018</td>
<td>100 nm GaN HEMT</td>
<td>115</td>
<td>cascode</td>
<td>4</td>
<td>20.3</td>
<td>2.6</td>
<td>34.8</td>
</tr>
<tr>
<td>[40]</td>
<td>2017</td>
<td>250 nm InP HBT</td>
<td>200–255</td>
<td>16-way combined</td>
<td>3</td>
<td>17–24</td>
<td>4.1</td>
<td>24</td>
</tr>
<tr>
<td>[187]</td>
<td>2015</td>
<td>35 nm GaAs mHEMT</td>
<td>250</td>
<td>4-way combined</td>
<td>3</td>
<td>10</td>
<td>2.9</td>
<td>32.8</td>
</tr>
<tr>
<td>[188]</td>
<td>2018</td>
<td>35 nm GaAs mHEMT</td>
<td>240</td>
<td>3-stacked</td>
<td>2</td>
<td>10.8</td>
<td>5</td>
<td>21.3</td>
</tr>
<tr>
<td>[189]</td>
<td>2017</td>
<td>130 nm SiGe BiCMOS</td>
<td>160</td>
<td>differential</td>
<td>5</td>
<td>15.5</td>
<td>7.2</td>
<td>30</td>
</tr>
<tr>
<td>[190]</td>
<td>2012</td>
<td>65 nm CMOS</td>
<td>140</td>
<td>8-way combined</td>
<td>4</td>
<td>13.2</td>
<td>14.6</td>
<td>16</td>
</tr>
<tr>
<td>[191]</td>
<td>2018</td>
<td>40 nm CMOS</td>
<td>140</td>
<td>2-way combined</td>
<td>5</td>
<td>14.8</td>
<td>8.9</td>
<td>20.3</td>
</tr>
<tr>
<td>[192]</td>
<td>2018</td>
<td>65 nm CMOS</td>
<td>114–131</td>
<td>differential</td>
<td>4</td>
<td>14.5</td>
<td>10.2</td>
<td>22.3</td>
</tr>
</tbody>
</table>

amplifiers have been demonstrated up to 330 GHz. Amplifiers using a 100 nm mHEMT process and a 50 nm mHEMT process have been demonstrated at 186–212 GHz [195] and 210–235 GHz [196], respectively. Both MMICs exhibits 7.5 dBm output power with around 2.5% PAE. Operating at frequencies higher than 200 GHz, the 35 nm mHEMT is the most up to date technology node of which the $f_T$ and $f_{MAX}$ are 515 GHz and 1000 GHz, respectively. The breakdown voltage of the process is around 2 V with a power density of 0.6 W/mm. Therefore, the GaAs mHEMT process is more suitable for developing low noise amplifiers at high frequency. The authors in [187] present a 32.8 dB gain PA at 250 GHz using the 35 nm mHEMT GaAs process. To improve the power, triple stacked-FET technique are used [36], [188]. The 240 GHz MMICs has a measured power of 10.8 dBm with 21.5 dB of gain, and the 300 GHz MMICs achieves 4.5 dBm output power with 11.6 dB of gain. Recently, a 238–292 GHz amplifier has also been presented with 6.7 dBm power with an associated maximum gain of 14.7 dB [197].

Silicon-based technologies have also been developed and deployed in mm-wave frequencies above 110 GHz. While SiGe and Si-CMOS have higher level of integration, their power amplifiers yield much lower power as compared to GaN MMICs. The 130 nm SiGe BiCMOS process is popular and has been used for developing power amplifiers to 255 GHz. For the 130 nm SiGe process, its $f_T$ and $f_{MAX}$ are 295 GHz and 400 GHz, respectively. The breakdown voltages BVCEO and BVCBO are 1.55 V and 6.0 V. A 170 GHz amplifier using the 130 nm SiGe process has been presented with up to 32 dB gain and 10 dBm output power [198]. To achieve high power, 4-way and 8-way on-chip power combiners are usually used in SiGe PAs. The typical output power is in the range of 7 to 15.5 dBm at different frequency ranging from 110 to 255 GHz [42], [189], [199], [200]. The highest output power achieved by SiGe amplifiers is 22 dBm at around 120 GHz [201]. The MMIC shown in Fig. 42 is fabricated in an advanced 90 nm SiGe BiCMOS process and demonstrates up to 22 dBm power with 7.7 dB gain and 3.6% PAE.

Si CMOS amplifiers have been demonstrated from 110 to 170 GHz. The output power is the lowest among all technologies and is in the range of 5 to 15 dBm. The most popular technology node in this frequency regime is 65 nm [137]. The authors in [190] presented a high efficiency D-band PA using 8-way power combiner in a 65 nm CMOS process. The 4-stage circuit exhibits a measured output power of 13.2 dBm with an associated peak PAE of 14.6% at 140 GHz, which is among the highest efficiency achieved in the similar frequency range. Up to 22.3 dB gain and 14.5 dBm output power has been presented in a 114–131 GHz amplifier using the same 65 nm process [192]. Other CMOS processes such as FD SOI 28 nm and 40 nm CMOS have also been deployed in D-band PAs [191], [202]. The highest output power CMOS PA reported to date (shown in Fig.43) was implemented in the 40 nm process. The design employing 2-way transformer-based power combining technique demonstrates a maximum power of 14.8 dBm with very high gain of 20.3 dB and peak PAE of 8.9% [191].
The state of the art in terms of output power, efficiency and bandwidth for the various technologies previously mentioned is reported in Table VII.

VIII. CONCLUSION

This paper has provided an overview of the current state-of-the-art in power amplifiers in the whole millimeter-wave frequency range. The relevance of this topic is increasing steadily thanks to the drive from large scale applications such as 5G and satellite communications, providing a fertile environment for the exploration of new technology solutions and design techniques, as well as the adaptation of methods successfully adopted at lower frequency by the microwave community.

REFERENCES


T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, 
IEEE 802.11ad Microwave Wi-Fi/WiGig Tutorial, June 2018, pp. 228–231.


Duy P. Nguyen (M’17) received the B.S. degree (with highest honors) in electrical and electronics engineering from Vietnam National University, Ho Chi Minh City, Vietnam, in 2011, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Davis, CA, USA, in 2014 and 2017, respectively. In 2017, he joined MACOM Technology Solutions Inc., Santa Clara, CA, USA, as a Senior Design Engineer where he designs optical linear drivers, and other RFICs and analog circuits. He has developed and designed a number of RFICs and MMICs in GaAs, GaN, InP, SiGe, and CMOS including power amplifiers, wideband distributed amplifiers, Doherty amplifiers, LNAs, T/R switches, phase shifters, fully integrated front-end modules, and linear drivers for optical modulators. He has authored and co-authored more than 30 patents, journals and conference papers. Dr. Nguyen was a recipient of Anil Jain Award for the best Ph.D. dissertation from UC Davis, Spot Recognition Award from MACOM in 2018, and Vietnam Education Foundation (VEF) fellowship.

Hua Wang (M’05–SM’15) is an associate professor at the School of Electrical and Computer Engineering (ECE) at Georgia Institute of Technology and the director of Georgia Tech Electronics and Micro-System (GEMS) lab. Prior to that, he worked at Intel Corporation and Skyworks Solutions on mm-Wave integrated circuits and RF front-end modules. He received his M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 2007 and 2009, respectively. Dr. Wang is interested in innovating analog, mixed-signal, RF, and mm-Wave integrated circuits and hybrid systems for wireless communication, sensing, and bioelectronics applications. He has authored or co-authored over 160 peer-reviewed journal and conference papers. Dr. Wang received the DARPA Young Faculty Award in 2018, the National Science Foundation CAREER Award in 2015, the IEEE MTT-S Outstanding Young Engineer Award in 2017, the Georgia Tech Sigma Xi Young Faculty Award in 2016, the Georgia Tech ECE Outstanding Junior Faculty Member Award in 2015, and the Lockheed Dean’s Excellence in Teaching Award in 2015. He held the Demetrius T. Paris Professorship from 2014 to 2018. His GEMS research group has won multiple academic awards and best paper awards, including the 2019 Marconi Society Paul Baran Young Scholar, the IEEE RFIC Best Student Paper Awards (1st Place in 2014, 2nd Place in 2016, and 2nd Place in 2018), the IEEE CICC Outstanding Student Paper Awards (2015, 2018, and 2019), the IEEE CICC Best Conference Paper Award (2017), the 2016 IEEE Microwave Magazine Best Paper Award, and the IEEE SENSORS Best Live Demo Award (2nd Place in 2016). Dr. Wang is a Technical Program Committee (TPC) Member for IEEE ISSCC, RFIC, CICC, and BCTCS conferences. He is a Steering Committee Member for IEEE RFIC and CICC. He is the Conference Chair for CICC 2019 and Conference General Chair for CICC 2020. He is a Distinguished Lecturer (DL) for the IEEE Solid-State Circuits Society (SSCS) for the term of 2018-2019. He serves as the Chair of the Atlanta’s IEEE CAS/SSCS joint chapter that won the IEEE SSCS Outstanding Chapter Award in 2014.

Anh-Vu Pham (SM’03) received the B.E.E. (with highest honors), M.S., and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1995, 1997, and 1999, respectively. Anh-Vu joined the University of California at Davis in 2002 as an Assistant Professor and was promoted to full Professor in 2008. From 1999 to 2002, he was an Assistant Professor at Clemson University. Anh-Vu is conducting research in RF to THz hybrid and integrated circuits, especially power amplifiers, antennas, radar and sensors, and energy harvesting systems. His research has been supported by DARPA, National Science Foundation, Office of Naval Research, Air Force Research Laboratory, Lawrence Livermore National Laboratory, and numerous companies. He has published 200 peer-reviewed papers, several book chapters, and two books and has graduated 21 Ph.D. students. Anh-Vu served as the Co-chair of Technical Program Committee of the IEEE International Microwave Symposium, 2016 in San Francisco, Co-Chair of the Technical Program Committee for the IEEE Asia Pacific Microwave Conference, 2017 in Malaysia, an IEEE MTT-S Distinguished Microwave Lecturer (2010-2012), and member/chair of International Microwave Symposium Technical Program Review Committee on power amplifiers. He also served as a guest editor of the IEEE Transactions on Microwave Theory and Techniques Special Issues on Broadband Millimeter-wave Power Amplifiers (2020) and Asia-Pacific Microwave Conference (2017). He received the National Science Foundation CAREER Award in 2001 and Outstanding Young Engineer Award from the IEEE Microwave Theory and Techniques Society, 2008. Anh-Vu is currently a co-director of the Davis Millimeter Wave Research Center. In 1997, Anh-Vu co-founded RF Solutions, a fabless semiconductor company providing power amplifiers and RFICs for WiFi applications. In 2003, RF Solutions was acquired by Anadigics. In 2008, he co-founded Planarmag, Inc and served as the CTO. In 2010, Planarmag, Inc was successfully acquired by TE Connectivity.