

A 39GHz Balanced Power Amplifier with Enhanced Linearity in 45nm SOI CMOS

by

Haien Ma

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2022

© Haien Ma 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

With the high data rate communication systems that come with fifth-generation (5G) mobile networks, the shift of operation to millimeter-wave frequency becomes inevitable. The expected data rate in 5G is significantly improved over 4G by utilizing the large available channel bandwidth at millimeter wave frequencies and complex data modulation schemes. With this increase in operation frequency, many new challenges arise and research efforts are made to tackle them. Among them, the phased array system is one of the hottest topics as it can be made use of to improve the link budget and overcome the path loss challenge at these frequencies. As the last circuit component in the transmitter's front-end right before the antenna, the power amplifier (PA) is one of the most crucial components with significant effects on overall system performance. Many of the traditional challenges of CMOS PA design such as output power and efficiency, are now compounded with the additional challenges that are imposed on complementary metal-oxide semiconductor (CMOS) PAs in millimeter wave phased array systems.

This thesis presents a balanced power amplifier design with enhanced linearity in GlobalFoundries' 45nm silicon-on-insulator (SOI) CMOS technology. By using the balanced topology with each stage terminating with a differential 2-stacked architecture, the PA achieves saturated output power of over 21 dBm. Each of the two identical sub-PAs in the balanced topology uses 2-stage topology with driver and PA co-design method. The linearity is enhanced through careful choice of biasing point and a strategic inter-stage matching network design methodology, resulting in amplitude-to-phase distortion below 1 degree up to the output 1dB compression level of over 19 dBm. The balanced amplifier topology significantly reduces the PA performance variation over mismatched load impedance at the output, thus improving the PA performance over different antenna active impedance caused by varying phased array beam-steering angles. In addition to this, the balanced topology also optimizes the PA input and output return loss, giving a better matching than -20 dB at both input and output, and minimizing the risk of potential issues and performance degradation in the system integration phase. Lastly, the compact transformer based matching networks and quadrature hybrids reduce the chip area occupation of this PA, resulting in a compact design with competitive performance.

Acknowledgements

I would like to thank my supervisor Professor Slim Boumaiza for his mentorship and feedback during the past two years, also for giving me the opportunity to study and work in the Emerging Radio Systems Group (EmRG).

I also would like to thank all EmRG colleagues who have participated in conversations with me.

Table of Contents

List of Figures	viii
List of Tables	xii
1 Introduction	1
1.1 Motivation	1
1.2 Problem Statement	3
1.3 Thesis Organization	4
2 Background and Literature Review	6
2.1 Output Power Enhancement	6
2.1.1 Active FET-Stacking	6
2.1.2 PA-level Passive Power Combining	9
2.2 Circuit Level Linearization Techniques	10
2.2.1 Cold-FET and Passive Pre-distorters	11
2.2.2 Complementary NMOS/PMOS Linearization	15
2.2.3 Linearity Compensation in Multi-stage PA	16
2.2.4 Inductive Degeneration	18

2.2.5	Harmonic manipulation	19
2.3	Efficiency Enhancement Techniques	20
2.3.1	Load Modulated PA	20
2.3.2	Dual-Drive PA	22
2.4	VSWR Resilient PA	23
2.5	Literature Review Summary	25
3	Transistor-Level Design	27
3.1	Transistor Layout Topologies and Trade-offs	27
3.2	Study of Transistor-level Stability	34
3.2.1	Differential Stability and Stabilization Techniques	34
3.2.2	Common-Mode Stability and Stabilization Techniques	36
3.3	Stacked-FET Design and Considerations	39
3.3.1	Comparison of Stacked-FET and Cascode Topology	39
3.3.2	Design of Stacked-FET Circuit	41
3.3.3	Analysis and load-pull of the 180um PMOS neutralized Stacked-FET	43
4	37-43 GHz Balanced PA Design	45
4.1	PA Specification and High-level Design Decisions	45
4.2	Design Methodology	47
4.3	Biasing Considerations	48
4.4	Matching Network Design	49
4.4.1	Input and Output Matching Network	50
4.4.2	Inter-stage Load-Pull	53
4.5	Transformer-based Hybrid 90-Degree Coupler Design	55

4.5.1	Output Single-Ended 90-Degree Coupler	55
4.5.2	Input Differential 90-Degree Coupler	59
4.6	Full BPA Simulation Results	62
4.6.1	Small Signal S-Parameter Simulation	64
4.6.2	Large Signal Performance under Continuous-Wave Simulation	64
4.6.3	Large Signal CW Simulation Over Load VSWR	66
5	Conclusion, Reflection and Future Work	69
5.1	Conclusion	69
5.2	Future Works	70
	References	72

List of Figures

1.1	Attenuation per kilometer of millimeter frequencies[32]	2
1.2	Major wireless frequency bands for 5G Frequency Range (FR) 1 and 2 [1] .	2
2.1	Saturated Output Power of PAs reported in literature vs Frequency	7
2.2	Example 3-Stacked Circuit Architecture [12]	8
2.3	Combiner Topologies: a) parallel combining b) series combining [25]	9
2.4	Schematic of a PA with 4-way hybrid power combining [40]	10
2.5	Schematic of the PA in [42] using a cold-FET linearizer with feedback . . .	12
2.6	Schematic of the PA in [34] using a shunt inter-stack cold-FET linearizer .	13
2.7	Schematic of the PA in [6] using a varactor based impedance-tuning linearizer with envelop tracking	14
2.8	Schematic of the non-linear inductor in [3]	15
2.9	Example plot of cgs versus sweep of VG with fixed VDS, for an NMOS device (dashed) and a PMOS device (solid) [2]	16
2.10	Schematic of a 38 GHz Power Amplifier with input PMOS varactor for AMPM linearization [10]	16
2.11	Schematic of NMOS/PMOS push-pull pair[19]	17
2.12	Schematic of the V-band doubly-hybrid NMOS/PMOS linearized PA [24] .	17
2.13	Schematic of a inductively degenerated PA for improved linearity [45] . . .	18

2.14	Schematic of the PA in [REF] with 2nd-harmonic short circuit[33]	19
2.15	A Doherty-like LMBA [38]	21
2.16	Dual-Drive PA Technique [15]	22
2.17	Performance parameters under mismatch for different PA topology[5]	24
3.1	Ladder-like source and drain interconnect [37]	28
3.2	A layout technique to minimize impact of source parasitic resistance in PA [26]	29
3.3	Layout in [13]	30
3.4	Layout in [47]	31
3.5	Layout in [38]	32
3.6	Ladder-like source and drain interconnect [11]	33
3.7	Gmax at 40GHz plotted vs neutralization capacitance in a 150um differential common-source	35
3.8	Stability factor vs neutralization capacitance in a 150um differential common-source across 10-100 GHz	36
3.9	Stability factor from DC-100GHz before adding stablization resistor at input	37
3.10	Stability factor from DC-100GHz after adding stablization resistor at input	38
3.11	Maximum available gain of a 2-stacked 180um amplifier vs varying capacitance at the gate of the stacked transistor	41
3.12	Load-pull results of the differential 180um 2-stacked transistor core	44
4.1	Magnitude and phase of the large signal transconductance vs input power for the 180um core with various biasing points	48
4.2	Schematic model of vertically coupled transformer with parasitics included	50
4.3	Primary winding inductance of schematic model vs layout EM	51

4.4	Secondary winding inductance of schematic model vs layout EM	52
4.5	Coupling factor of schematic model vs layout EM	53
4.6	Load impedance presented by matching networks to the drains of two amplifiers over 37-43 GHz	54
4.7	Schematic of the single-ended transformer based quadrature hybrid [14] . .	56
4.8	Layout of the single-ended transformer based quadrature hybrid design . .	56
4.9	Insertion loss and isolation for the single-ended transformer based quadrature hybrid	57
4.10	Relative phase difference between S21 and S31 for the single-ended transformer based quadrature hybrid	57
4.11	Matching of each port for the single-ended transformer based quadrature hybrid	58
4.12	Original layout of the differential transformer based quadrature hybrid . .	58
4.13	Updated layout of the differential transformer based quadrature hybrid . .	59
4.14	Insertion loss and isolation for the original differential transformer based quadrature hybrid	60
4.15	Relative phase difference between S21 and S31 for the original differential transformer based quadrature hybrid	60
4.16	Insertion loss and isolation for the improved differential transformer based quadrature hybrid	61
4.17	Relative phase difference between S21 and S31 for the improved differential transformer based quadrature hybrid	61
4.18	Matching of each port for the improved differential transformer based quadrature hybrid	62
4.19	Final PA schematic	63
4.20	Bumped version of the PA layout	63

4.21 BPA Small signal s-parameter simulation results	64
4.22 Gain vs output power in large signal continuous wave simulation	65
4.23 Amplitude-to-phase distortion in large signal continuous wave simulation	65
4.24 Overall power-added-efficiency vs output power in large signal continuous wave simulation	66
4.25 Gain vs output power at 39GHz with matched load (in blue) and over 2:1 VSWR (in red) in large signal continuous wave simulation	67
4.26 Amplitude-to-phase distortion at 39GHz with matched load (in blue) and over 2:1 VSWR (in red) in large signal continuous wave simulation	67
4.27 Overall power-added-efficiency vs output power at 39GHz with matched load (in blue) and over 2:1 VSWR (in red) in large signal continuous wave simulation	68

List of Tables

2.1	Comparison of Recent Ka-Band Power Amplifiers in CMOS Technology . .	25
-----	--	----

Chapter 1

Introduction

1.1 Motivation

In the recent era, technology is advancing at a more rapid pace than ever. Wireless communication is starting to play a critical role in numerous applications such as wearable devices, mobile data, augmented reality (AR), virtual reality (VR), ubiquitous health care, smart home, cloud office, etc[27]. All of these have triggered an exponential rise in mobile data traffic demand. Although techniques such as carrier aggregation and advanced multiple-input multiple-output (MIMO) has sustained the fourth-generation (4G) Long Term Evolution (LTE) to meet the increasing data traffic demand for an extended period of time, the low available bandwidth in sub-6 GHz frequency ranges will undoubtedly fall short when the technologies are expected to meet data rate of gigabits per second and beyond. These demands see an evolution of wireless communication technology from 4G to fifth-generation (5G).

With this increase in demand of signal bandwidth, conventional sub-6GHz starts to see its limitations due to the limited available frequency bands in the sub-6GHz frequency ranges. As a result, millimeter wave frequency bands will see more frequent usage in 5G services because of the gigahertz of available bandwidth. Due to the relatively lower attenuation level at the Ka band frequencies, as illustrated in Figure 1.1 and the large

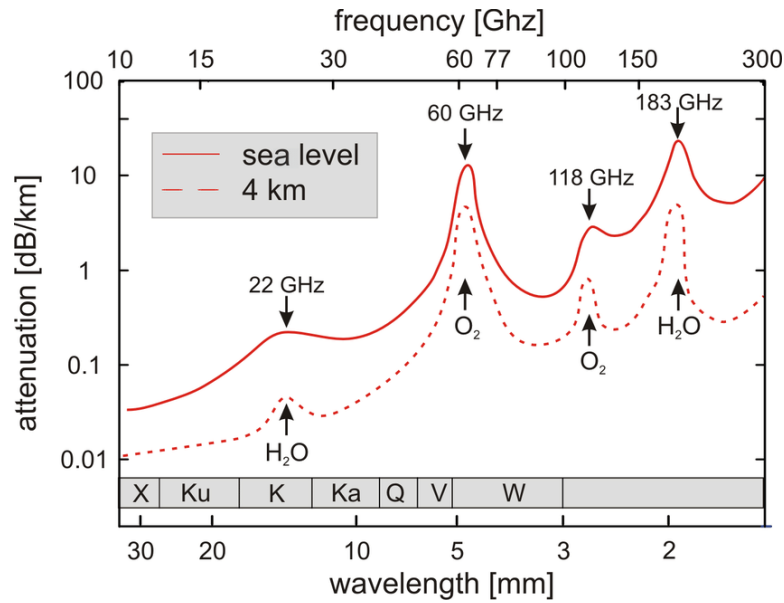


Figure 1.1: Attenuation per kilometer of millimeter frequencies[32]

available bandwidths, various bands in this frequency range are chosen for high data rate 5G communication.

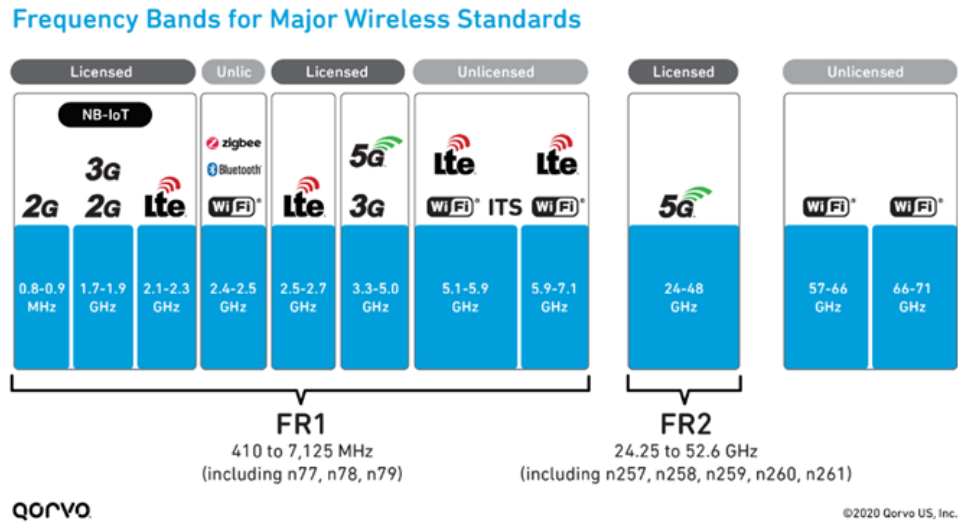


Figure 1.2: Major wireless frequency bands for 5G Frequency Range (FR) 1 and 2 [1]

In addition to the wider bandwidths, 5G utilizes the massive MIMO technique. Compared to the MIMO technique in 4G, it essentially combines the concept of beamforming with MIMO and enables data link between multiple users and a base station. This, combined with the increased path loss at millimeter wave frequencies compared to sub-6 frequency ranges, results in the hardware solution of phased array systems. Meanwhile, low-cost chips with high integration capability are enabled as the modern complementary metal-oxide semiconductor (CMOS) technologies advances in terms of high frequency performance. These factors drive the need for a phased array system with beam-forming front-end circuitry implemented in highly integrated beam-forming chips.

1.2 Problem Statement

Many challenges exist for modern PA designers. First, CMOS technology is known for its advantages in low cost and high integration capability. However, as the device size reduces in more advanced technology nodes for improved speed, the resulting low breakdown voltage limits the output power achievable of the PA. The contrast is especially stark when compared to some of the III-V technologies such as gallium nitride, which could operate under supply of tens of volts without breaking down. Another problem that comes with high frequency design is the high capacitive reactance at these frequencies. Simply increasing the dimensions of the transistor could result in deteriorated performance. Thus, to make a CMOS PA viable and reach a satisfactory level of output power, techniques to enhance the output power of a PA without causing significant performance degradation should be carefully considered.

Second, to maximize data throughput, 5G millimeter wave wireless links will make use of complex modulation schemes and techniques, such as high-order quadrature amplitude modulation (QAM), Orthogonal Frequency-Division Multiplexing (OFDM). These signals with complex modulation schemes come with large Peak-to-Average-Power-Ratio (PAPR). For example, the 64 QAM single-carrier signal typically has a PAPR of 7.1 dB, whereas the 3GPP 5G NR FR2 400 MHz 2-CC 64 QAM signal has a PAPR of more than 11.7 dB at a 0.01% of its complementary cumulative distribution function (CCDF). Enveloped

signals with high PAPR in range of 6 to 12 dB puts strong emphasis on both linearity and the back-off efficiency of the PA. In addition, the increased power consumption of digital pre-distortion (DPD) with the wide signal bandwidth compounded with the low power of CMOS PA lead to the low feasibility of DPD on individual PA due to the significant degradation to overall transmitter efficiency. This places further limitation on tolerable PA non-linearity.

In addition to these traditional PA design challenges, antenna coupling in phased array systems presents further trouble for PA designers. The poor matching of the antenna elements as well as load impedance mismatch that comes from beam-steering are both factors that could introduce significant degradation to PA performance as the impedance deviates from the designed impedance.

This thesis presents a linearity enhanced balanced power amplifier design as a solution to these aforementioned problems. Using 2-way differential 2-stacked PA stages, the PA achieves saturated output power of over 20.5 dBm, with a peak power-added efficiency (PAE) of over 37%. The linearity enhanced 2-stage sub-PA design in each branch of the balanced PA makes the PA suitable to operate with high-PAPR signals. With 2-stage class-AB driver and PA co-design, the PA has an output 1dB compression point (OP1dB) of over 19dBm and a amplitude to phase distortion of up to 1 degree until OP1dB. The balanced amplifier improves the input and output return loss to over 20dB, and improves PA resiliency to VSWR variation as verified in the final results.

1.3 Thesis Organization

The thesis is organized as follows. Chapter 2 goes over the main challenges of millimeter-wave power amplifier designs and discusses the approaches and techniques used in state-of-the-art power amplifiers.

Chapter 3 will start off with a brief comparison of different transistor layout topology. This is followed by a discussion of stability in power amplifiers. Lastly, the chapter wraps up by discussing the design of stacked-FET PA as well as showcasing some load-pull results of different transistors available in the archive.

Chapter 4 will present the design specification, design consideration and general design methodology of the millimeter wave 2-stage balanced power amplifier. Complete results of EM simulation are included, and PA performance under variation of VSWR will be tested and presented.

Chapter 5 concludes the thesis with a summary of findings and will mention some possible directions that future design work could take.

Chapter 2

Background and Literature Review

2.1 Output Power Enhancement

Due to the low breakdown voltage, achieving a high saturated output power has been one of the major challenges of CMOS PAs compared to other technologies, as shown in Figure 2.1. The increased parasitic and lower power gain at millimeter frequencies further complicates the issue. The parasitic capacitance introduced from increasing the device widths will reduce the transition frequency of the device as well as cause degradation to the power efficiency achievable by the PA.

There are generally two ways to increase the output power of the amplifier: FET stacking and passive power combining.

2.1.1 Active FET-Stacking

A device level technique of improving output power known as stacking has been frequently adopted in CMOS SOI technology, example in [12]. In this topology, the impedance seen at the drain and gate of the each transistor stacked in series are designed such that the voltage across any two given nodes of a transistor in the topology follows the reliability requirements.

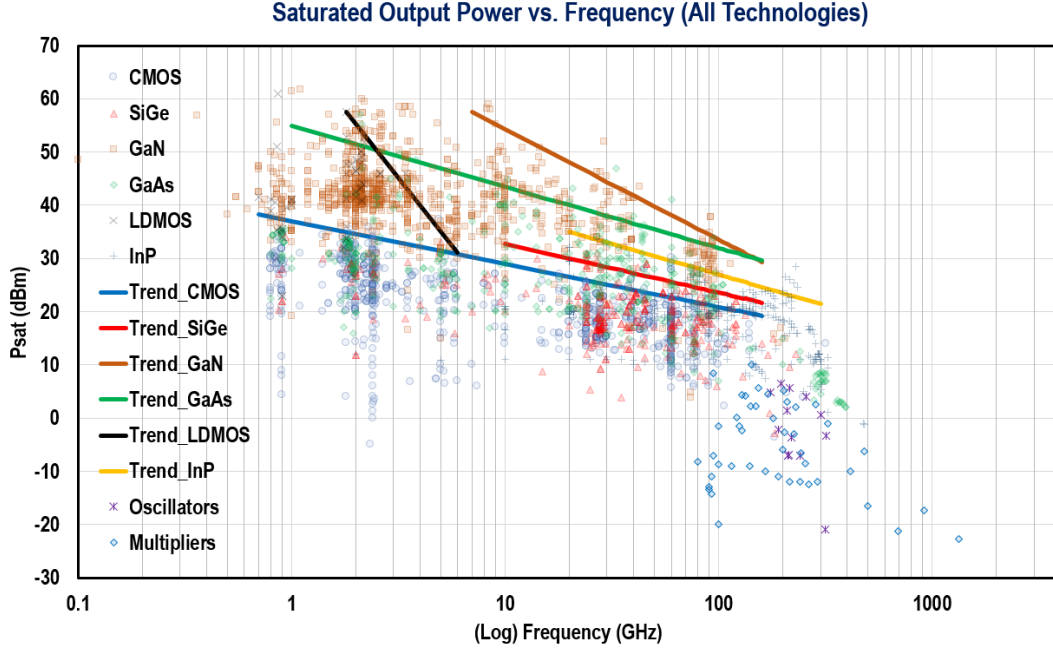


Figure 2.1: Saturated Output Power of PAs reported in literature vs Frequency

This method of boosting output power of the PA has obvious advantages. Firstly, one advantage with stacked-FET PA is that the chip area occupied by the stacked-FET structure is small compared to typical multi-way power combining methods, therefore significantly reducing the area cost needed to reach a certain output power threshold. Secondly, in device with low breakdown voltages, stacked-FET architecture can effectively boost the low R_{opt} in the basic common-source topology.

However, the issues with stacked-FET are also apparent. With the increasing number of stacks, the parasitics become more and more significant. As a result, efficiency and linearity both suffer. This also leads to a diminishing return of output power as number of stacked-FET increases, due to the leakage current in the parasitic capacitance as well as inter-stack matching networks.

Overall, stacking of transistor beyond two stack is less frequently seen in literature because of the linearity and efficiency degradation, and stacking sees a further reduction

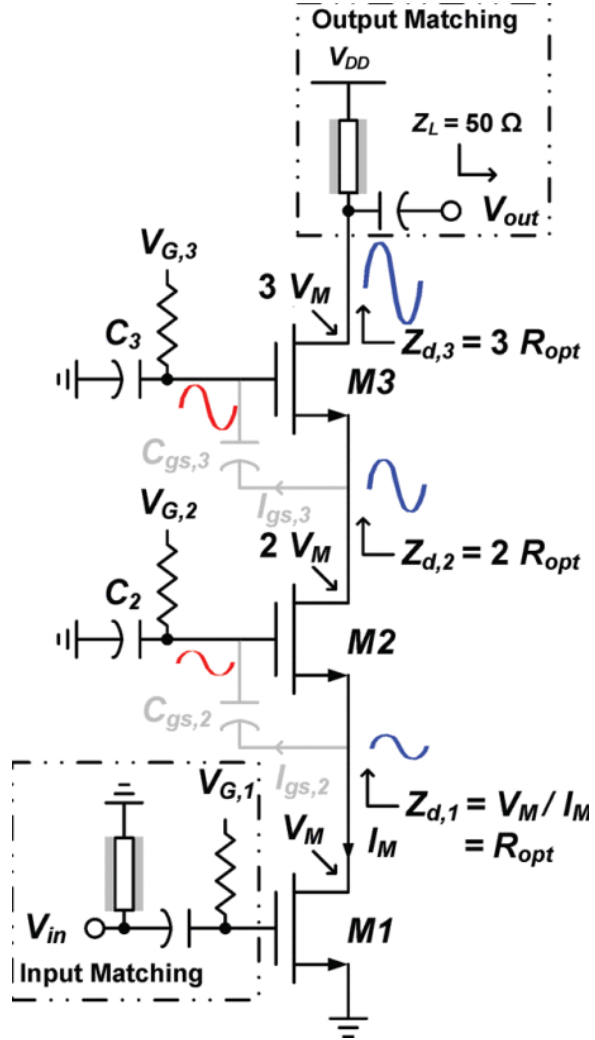


Figure 2.2: Example 3-Stacked Circuit Architecture [12]

in usage as frequency of applications increases, particularly beyond 100GHz as parasitics become increasingly significant.

2.1.2 PA-level Passive Power Combining

The most common method of increasing output power of a power amplifier is through multi-way power combining using passive combiner networks. At higher frequencies, particularly above 100GHz, distributed combiners are more common, with the frequent usage of coupled-line based combiner networks seen in literature for their low loss combining. However, in 5G millimeter wave frequency bands at 28GHz and 40GHz, distributed architectures occupy an overly large area and is very costly to implement. Lumped passives, in particular, transformers are commonly adapted for power combining and matching in Ka-band frequency ranges because of its compact size and wide achievable bandwidth. There are two types of power combining topologies, parallel combining (or current combining) and series combining (or voltage current combining).

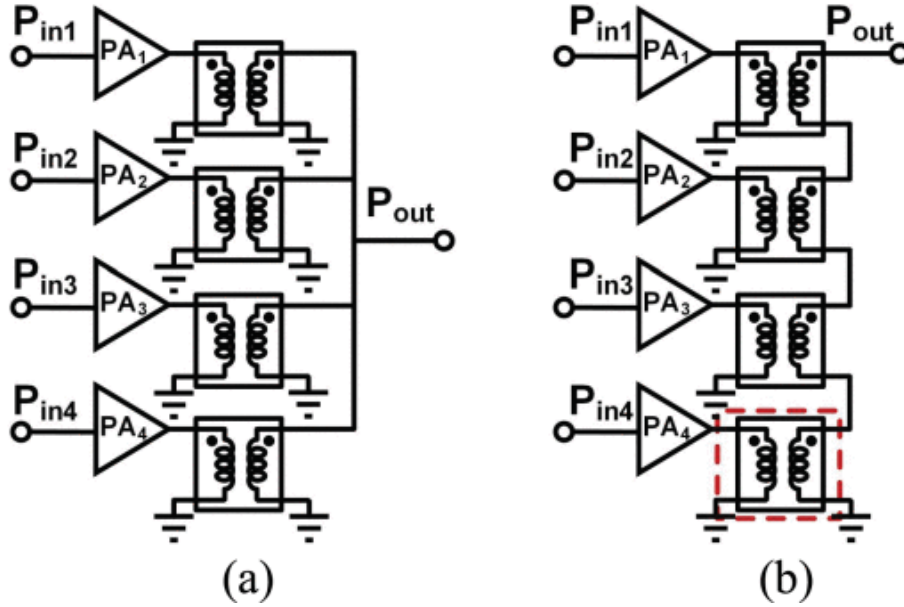


Figure 2.3: Combiner Topologies: a) parallel combining b) series combining [25]

The biggest difference in these two topologies lies in their impedance transformation. In parallel combining structure, the the load resistance seen by each branch is higher because they combine the current in shunt; whereas in series combining the load resistance seen

by each branch is lower because they combine the voltages in series. Depending on the R_{opt} of the PA transistors, designer may choose the combiner topology careful for a low impedance transformation ratio in their output matching networks.

One common method to alleviate such an issue is to through the usage of hybrid combining topology, where it includes both parallel and series combining. As a result, the load resistance does not see any significant transformation before being matched to the load presented to the transistors. In addition, mixing the structure makes it easier to achieve structural symmetry and it reduces the complexity of overall layout. An example of such combining network is shown in [40], where two pairs of the total four sub-PAs are combined in series before they combine in parallel at the output.

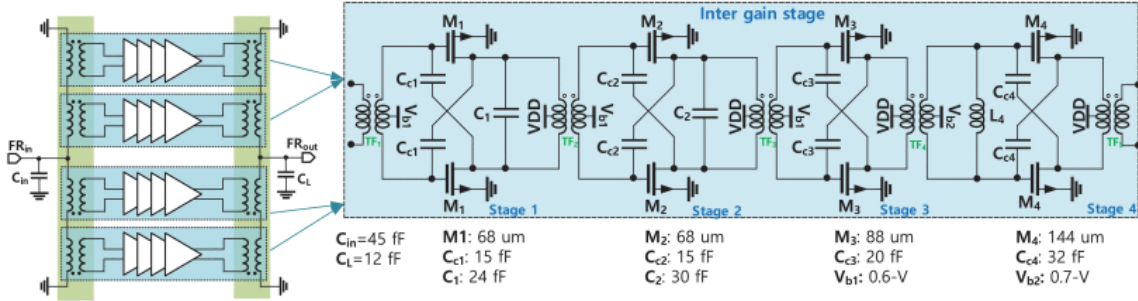


Figure 2.4: Schematic of a PA with 4-way hybrid power combining [40]

In addition, series combining suffers from being more susceptible to impedance and phase mismatch due to parasitics, whereas parallel combining typically do not suffer from this issue to the same degree and have lower losses[4]. However, parallel combining typically results in a more complex routing and thus resulting in a larger layout compared to series combining.

2.2 Circuit Level Linearization Techniques

In practical operation, the PA would be transmitting modulated signal, the transmitted signal need to meet the emission mask requirement in the frequency spectrum as well as having a non-distorted constellation to ensure the information being transmitted is as

intended. The linearity of the power amplifier has a significant impact on the final output signal. More importantly in 5G, the complex QAM modulation schemes used have high signal PAPR, which places stringent linearity requirement on the PA. Additionally, Digital-Pre-Distortion (DPD) has an increased power consumption for millimeter wave 5G bands due to the large signal bandwidth. This combined with the relatively low output power of CMOS PAs lead to the inevitable conclusion that it is best to operate the PA without DPD. As such, PAs with enhanced linearity has been a popular topic for research. Some major causes to PA non-linearity include the various harmonic components of the transistor output current and voltage, as well as the transistors' non-linear parasitic capacitance. It is important to analyze what factors can affect the linearity of the PA linearity in order to improve it. Methods of linearization can be categorized based on their approaches.

2.2.1 Cold-FET and Passive Pre-distorters

[42] utilizes a cold-FET linearizer with feedback at the input to the driver stage, serving as a pre-distorter to correct for AMAM distortion of the PA. Compared to using an extra driver stage for AMAM correction, these cold-FET based linearizers are smaller, easier to implement, and consumes less power. However, as it is passive, the pre-distorter effectively reduces the gain at lower input power level as opposed to improving the gain at higher power levels. This reduction in small signal power gain may become difficult to tolerate at higher frequencies where power gain is more precious.

In [34], a shunt cold-FET based inter-stack linearizer is implemented. The FET again behaves like a variable resistor, where the impedance seen into the linearizer varies with power level. The linearity of the cold-FET network compensates for the gain and phase non-linearity experienced by the original circuit. However, it may be worth to look into the effect this inter-stack linearizer has on the voltage swing across top and bottom transistors for reliability concerns.

The load impedance presented to the drain of a transistor affect its equivalent input capacitance. As a result, one can affect the PA linearity by changing the load impedance based on input power level. [6] proposes a varactor based network at the gate of the common-gate device, and an envelop-tracking circuit is used to track the power level and

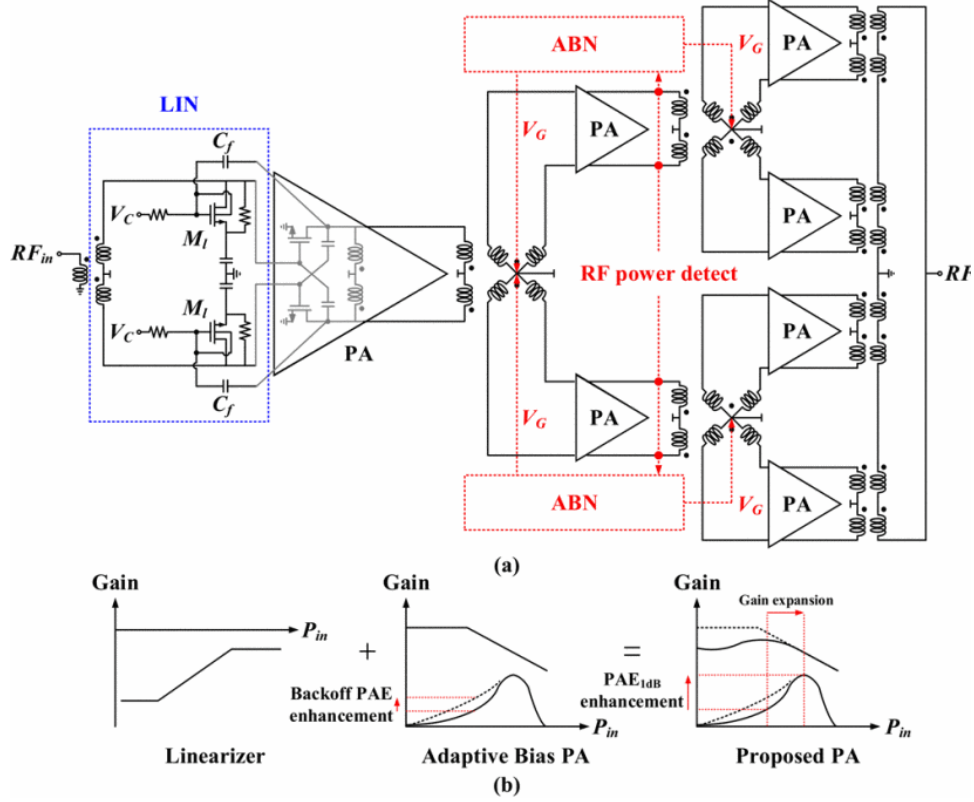


Figure 2.5: Schematic of the PA in [42] using a cold-FET linearizer with feedback

control capacitance of the varactor in order to alter the impedance seen by the drain of common-source transistor, as seen in Figure 2.7. This reduces the miller capacitance at the input of the device as input power level increases, as such effectively linearizing the input capacitance. Overall, the PA sees a significant reduction in both AMPM and AMAM distortion compared to the case where no varactor is present. However, it is worth noting that the result compares between the performance achieved with and without varactor implemented, whereas a more appropriate comparison might be to the case where the stacked architecture is implemented with regular capacitor in place of the varactor-based tuning network. The performance enhancement shown in the results is massive, to the degree that it is questionable whether the reference case represents a properly designed PA without proposed the linearization network.

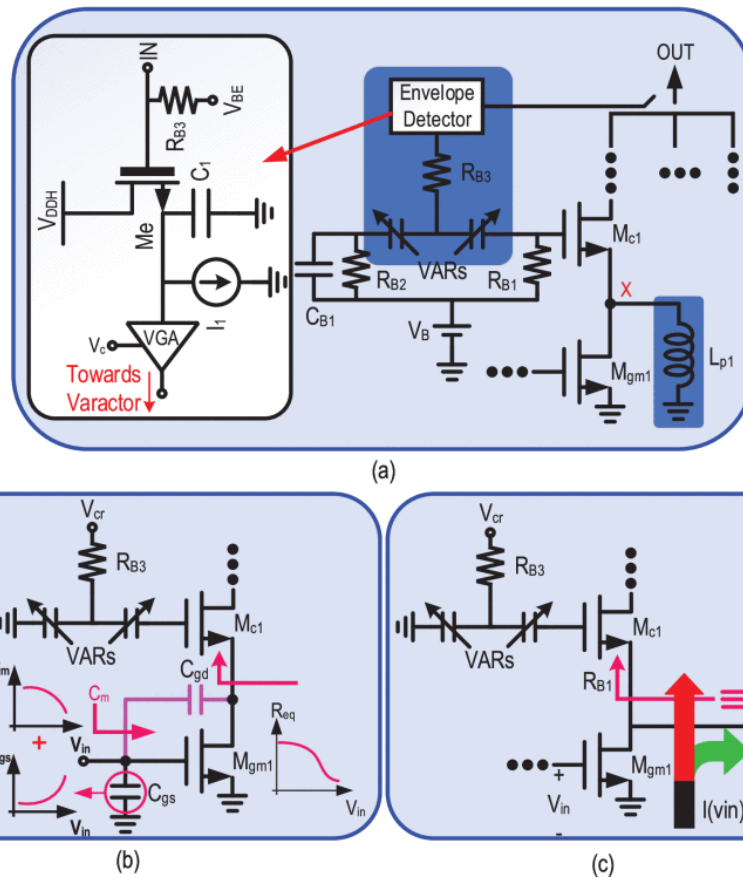


Figure 2.7: Schematic of the PA in [6] using a varactor based impedance-tuning linearizer with envelop tracking

network. One advantage of implementing the transformer based frequency re-configurable network is that with the implementation of a switch, the effect of the extra winding could be largely removed with minimal extra losses, which is beneficial for on-bench comparison of cases with or without the additional tuning network. In addition, the benefits in common-mode rejection of a differential PA cannot be ignored. as such, investigations in the implementation of this methodology in differential circuit configuration should be considered. Although, network in that case might be limited by the number of accessible low-loss metal layers in the process.

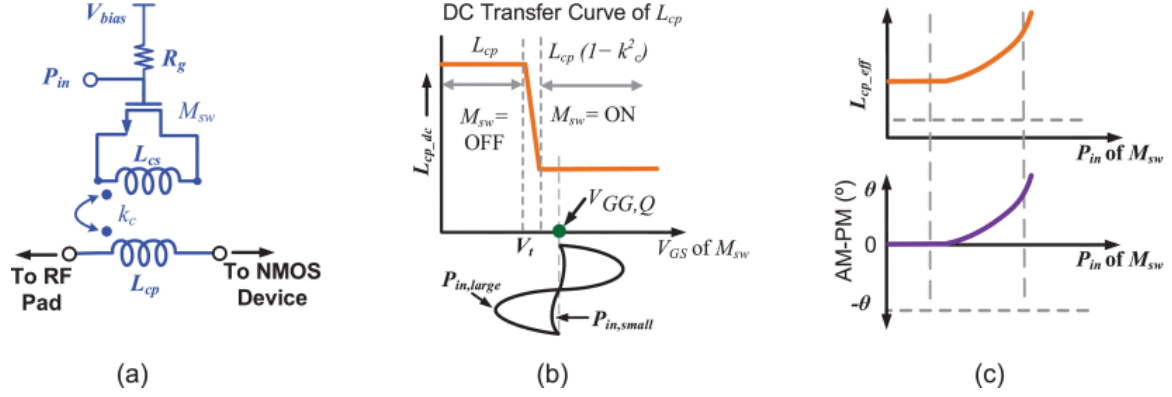


Figure 2.8: Schematic of the non-linear inductor in [3]

2.2.2 Complementary NMOS/PMOS Linearization

As mentioned before, the amplitude-to-phase (AMPM) distortion is mainly caused by the non-linear capacitance at input of the transistors. The major contributor to this is the non-linear gate-to-source capacitance in transistors. One method of linearizing this non-linear C_{gs} is to utilize the complementary behavior of PMOS transistors. The opposite trend of C_{gs} to gate-to-source voltage in PMOS will compensate that of the NMOS transistors and resulting in an overall improved AMPM linearity for the whole PA.

One of the most popular techniques is to use PMOS capacitor or varactor in the NMOS based PA. [10][41][43] use PMOS varactors or capacitors at the gate input of the NMOS transistors as a part of the input matching network. In [2], PMOS varactors are not used at the input, but instead they are used to replace the neutralization capacitors.

Another technique that involves the use of PMOS is not through PMOS varactors, but instead PMOS transistors are used together with NMOS transistors in a push-pull PA architecture in Figure 2.11 [19]. The downside of using PMOS transistor in PA is that their RF performance, such as transition frequency, trans-conductance are inferior compared to those in the NMOS transistors. Even though in more advanced technology nodes, the performance of PMOS transistors have been improved to a closer match to the NMOS transistors, the difference nonetheless exists in many.

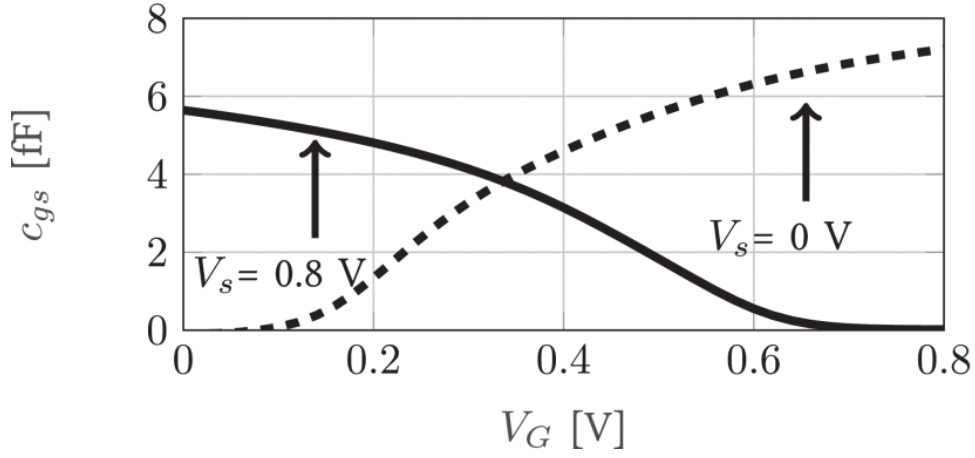


Figure 2.9: Example plot of c_{gs} versus sweep of V_G with fixed V_{DS} , for an NMOS device (dashed) and a PMOS device (solid) [2]

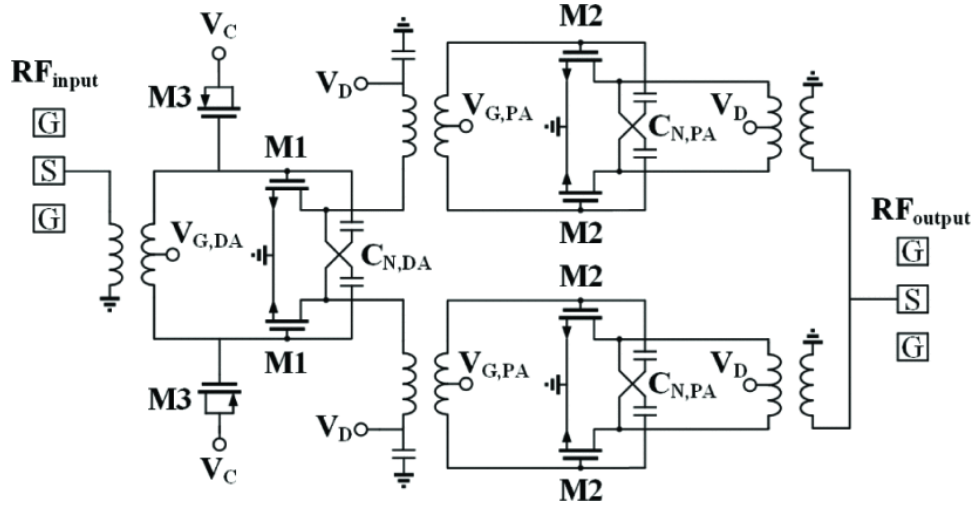


Figure 2.10: Schematic of a 38 GHz Power Amplifier with input PMOS varactor for AMPM linearization [10]

2.2.3 Linearity Compensation in Multi-stage PA

Another approach to PA linearization is through the usage of multi-stage or multi-way power combining with different sub-PAs or different stages exhibiting different linearity

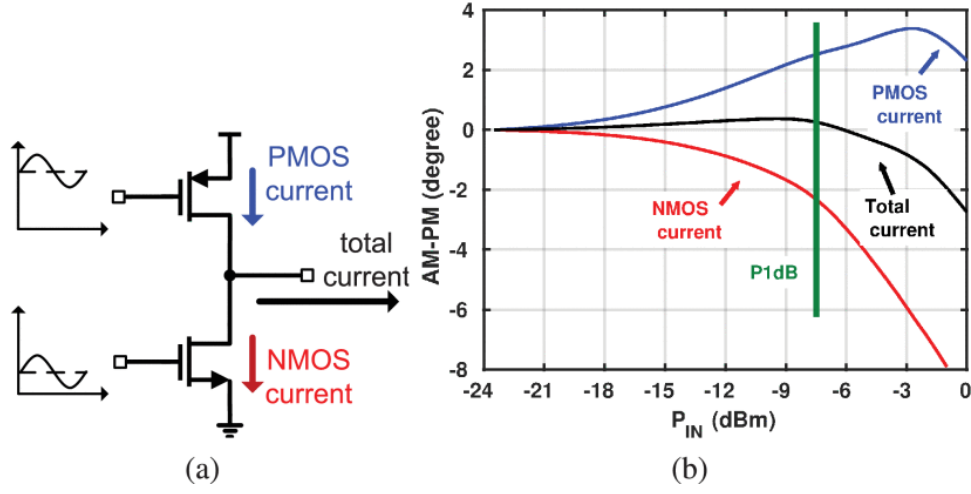


Figure 2.11: Schematic of NMOS/PMOS push-pull pair[19]

behaviors. With strategical design approaches in different sub-PA or different stages, the overall PA could combine to show a linearized response.

In [24], a three-stage four-way combining PA is designed and measured at 55GHz. The PA uses both NMOS and PMOS drivers as well as sub-PAs together for a combined PA with improved linearity. The author designs the sub-PAs with transistor of small widths and avoid stacking, resulting in high overall efficiency. This approach suffers from classic trade-off between power-combining and FET-stacking, leveraging chip area for efficiency.

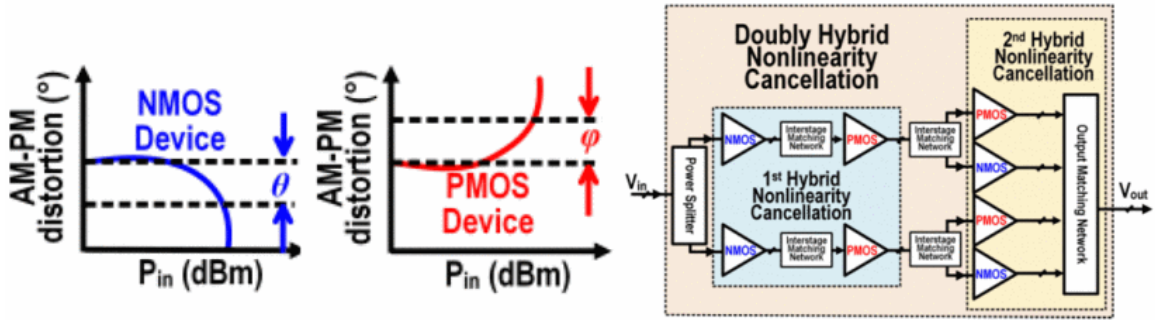


Figure 2.12: Schematic of the V-band doubly-hybrid NMOS/PMOS linearized PA [24]

A top-down design approach is taken in [44], through an inter-stage load-pull design

approach. The inter-stage matching network parameters here are chosen for an improved linearity in a systematic approach for the entire 2-stage PA instead of designing the stages separately in the traditional approach.

2.2.4 Inductive Degeneration

[45] make use of inductive degeneration in the PA stage for AMAM linearity improvement. The biasing of driver and PA stage are chosen separately for an improved 2-stage AMPM response while making minimum sacrifice in efficiency with the trade-off in biasing point selection.

It is known that parasitic losses in the source terminal reduces the power gain as well as the efficiency achievable by the PA, however it is argued here that the degeneration inductor improves the AMAM and pushes the OP1dB to a higher input power level, resulting in a higher PAE at OP1dB and power-back off compared to plain CS pair. In addition, inductive degeneration technique can assist with the input matching, making it an attractive option in 2-stage designs where the inter-stage impedance transformation ratio may limit the bandwidth.

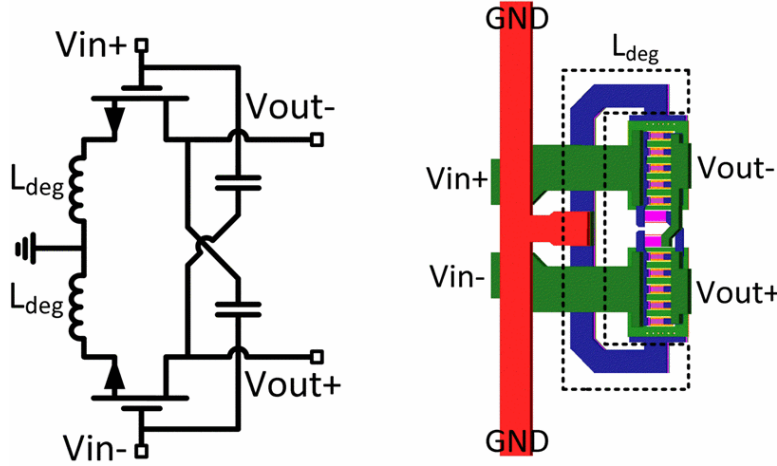


Figure 2.13: Schematic of a inductively degenerated PA for improved linearity [45]

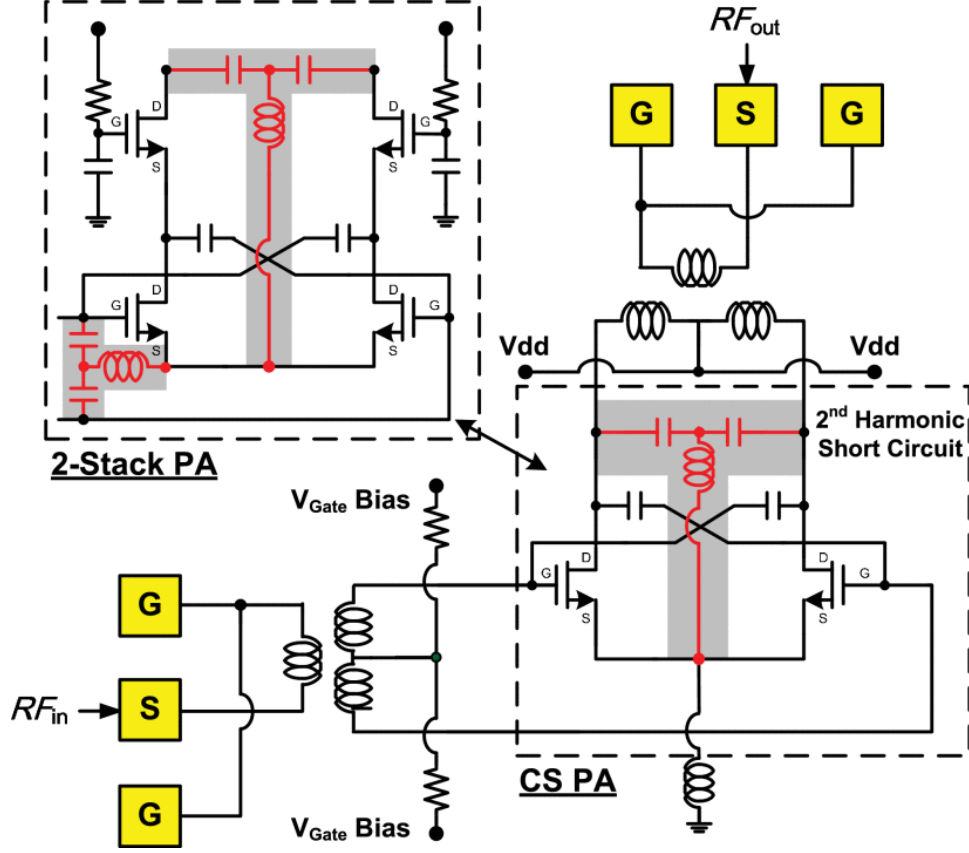


Figure 2.14: Schematic of the PA in [REF] with 2nd-harmonic short circuit[33]

2.2.5 Harmonic manipulation

The technique known as sweet-spot biasing is often used in class AB PAs in order to minimize the effect of third-order inter-modulation (IM3) product on PA linearity. However, there is usually a non-negligible second harmonic component. In [33], a technique of second harmonic short circuit is used to minimize the impact of this current component at second harmonic on the PA performance.

As seen in Figure 2.14, a second harmonic series LC resonator connects from drain of both output to the virtual ground at the source, shorting the second harmonic component of current flow. With the application of proper biasing and the implementation of drain-

to-source second harmonic short circuit, the PA shows an improvement in PAE, AMAM and IMD3 distortion in two-tone simulation. The downside of only applying this approach for linearization is that the AMPM non-linearity is not improved, which can lead to more severe memory effect and degrade the output signal EVM.

2.3 Efficiency Enhancement Techniques

2.3.1 Load Modulated PA

In 5G cellular communications, the complex modulation schemes typically come with high signal PAPR. This results in a need for back-off efficiency enhancement in order to improve the average power efficiency of the PA.

The Doherty technique has proven to be the most popular technique for PA active-load modulation and back-off efficiency enhancement. Novel combiner network is one of the most popular areas of research in Doherty PA publications, a few examples are listed as follows. [30] uses coupler-based baluns for output combiner synthesis. [16] proposes a transformer-based combiner synthesis technique which could absorb the transistor output and pad capacitance in a dual band design. [18] uses LC tuning circuits in transformer-based output combiner to minimize the loading from auxiliary amplifier in its off-state. [29] uses on-chip antenna as a series doherty combiner.

Adaptive biasing technique is used to counter-act the effect of the early turn-on and reduce the sensitivity of Peaking PA to threshold voltage variations. The adaptive biasing network ensures that the auxiliary amplifier remains in class-C for low voltages, and the biasing is elevated to class-AB as the input power increases past the designed back-off level. [9] proposes a novel adaptive biasing circuit with envelop detection and careful attention to the RC time constant. Whereas the example in [17] uses adaptive biasing to achieve a dual-mode wide-band PA which can swap the main and auxiliary PA. The drawback of the adaptive biasing technique include the complexity of the control and the discontinuity in the gain characteristic of the PA, which degrades linearity.

Although there exists challenges to implement DAC-based PA at millimeter wave frequencies due to linearity and efficiency degradation under high modulation bandwidths, their presence and implementation is not to be overlooked. In both [8] and [7], mm-wave DAC based PA is used for its reconfigurability in frequency and back-off efficiency enhancement.

[36] incorporates the Doherty PA into the balanced amplifier topology, effective designing a series-Doherty PA on both paths of the balanced PA, exploiting the advantages of the balanced amplifier structure in its resiliency to load variations and antenna mutual coupling effect.

In [38], a Doherty-like load modulated balanced amplifier(LMBA) is designed with coupled-line based hybrid 90 degree couplers. However, both the LMBA topology and the use of coupled-line based hybrid structure contribute to a large chip area of the PA.

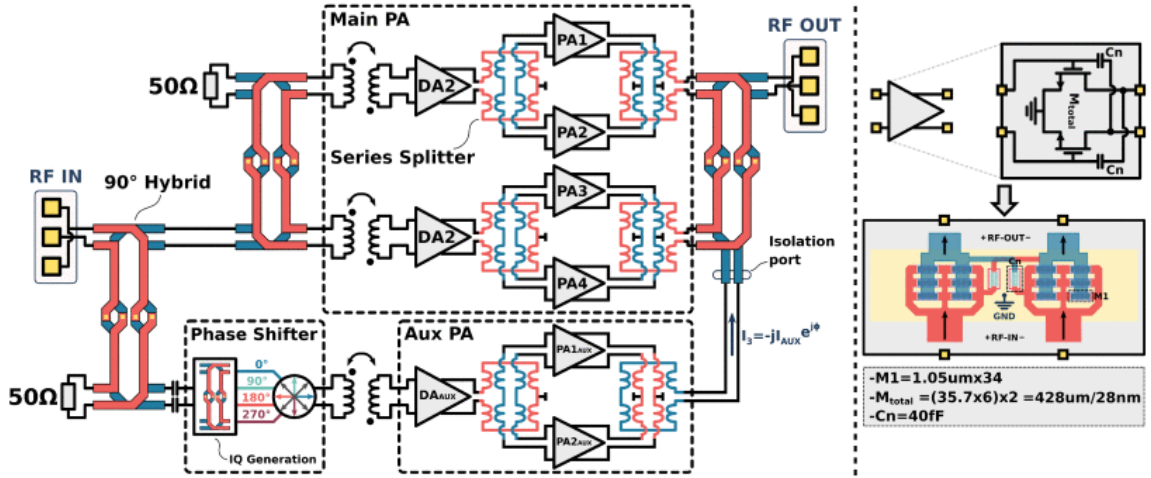


Figure 2.15: A Doherty-like LMBA [38]

Although not as popular as the Doherty PA, there are also research efforts made into outphasing PA. Outphasing PA topology in theory eliminates the linearity dependence on the amplifier sub-blocks due to their constant envelop voltage amplitude[46]. Many recent outphasing PA make use of Chireix combiner for back-off efficiency enhancement. [21] uses on-antenna combiner network for active load modulation, whereas [39] designs its combiner

with a triaxial balun. In order to address the performance issues that comes with operating the PA in over-driven class-B for voltage-mode approximation, research effort is made to design the PA in a current-mode inverse outphasing architecture [31][22].

2.3.2 Dual-Drive PA

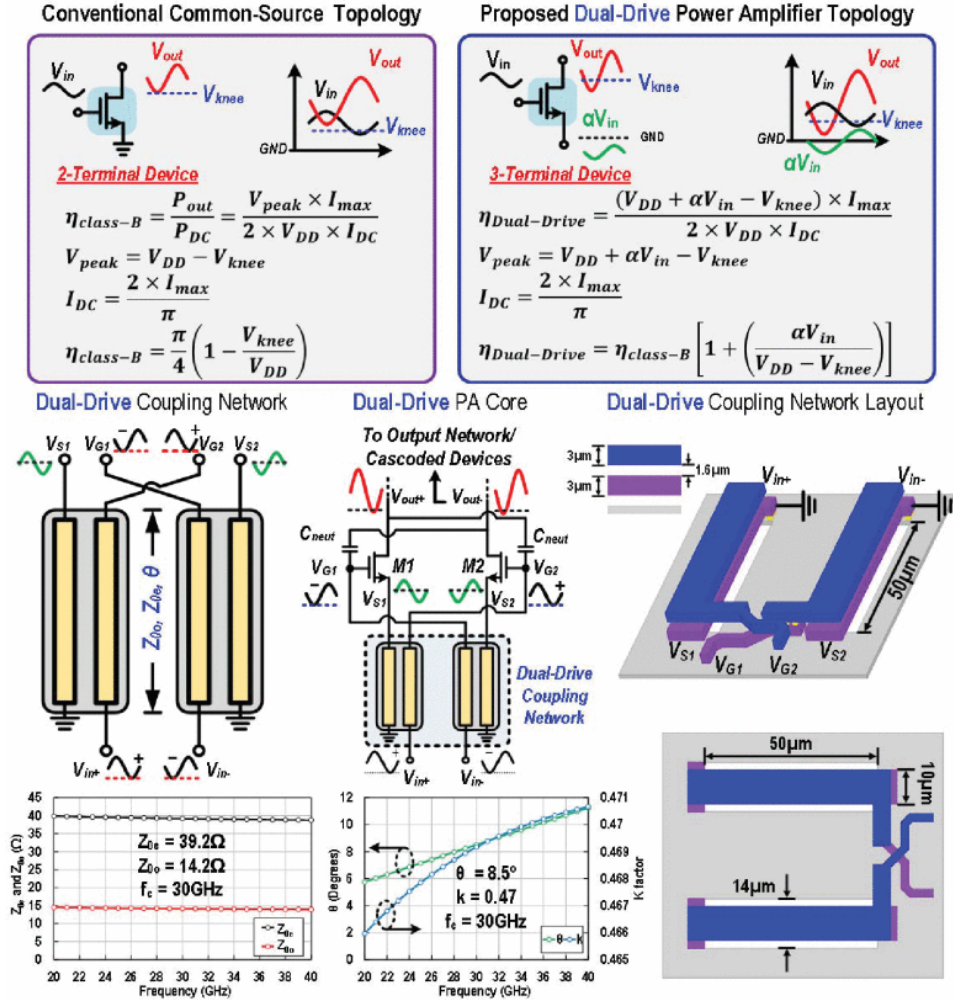


Figure 2.16: Dual-Drive PA Technique [15]

[15] proposes a novel PA operation mode by designing the PA input with a coupler

that also feed into its source terminals named "dual-drive" PA as seen in Figure 2.16. This enables the voltage at the source to swing together with the input, this increasing the transistor output voltage swing, improving the PA drain efficiency. However, compared to traditional PA topology, the dual-drive PA suffers from lower power gain, particularly in regular common-source configuration. This degrades power-added efficiency and limits its merits in single-stacked architectures.

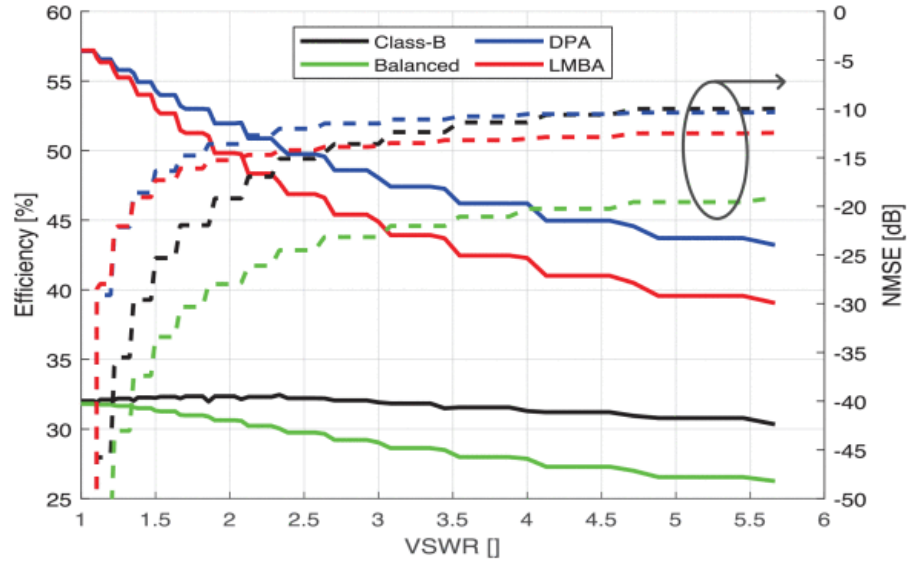
2.4 VSWR Resilient PA

When it comes to designing PAs for millimeter wave bands in 5G applications, it is expected that they will be operating in phased array systems with large number of elements. It is a known problem there that the impedance presented by the antenna elements of the phased array system shows significant variation for different steering angles. Therefore, for practical PA usage in such a system, it is important that the amplifier maintains an acceptable level of performance over loads with different phases on a VSWR circle.

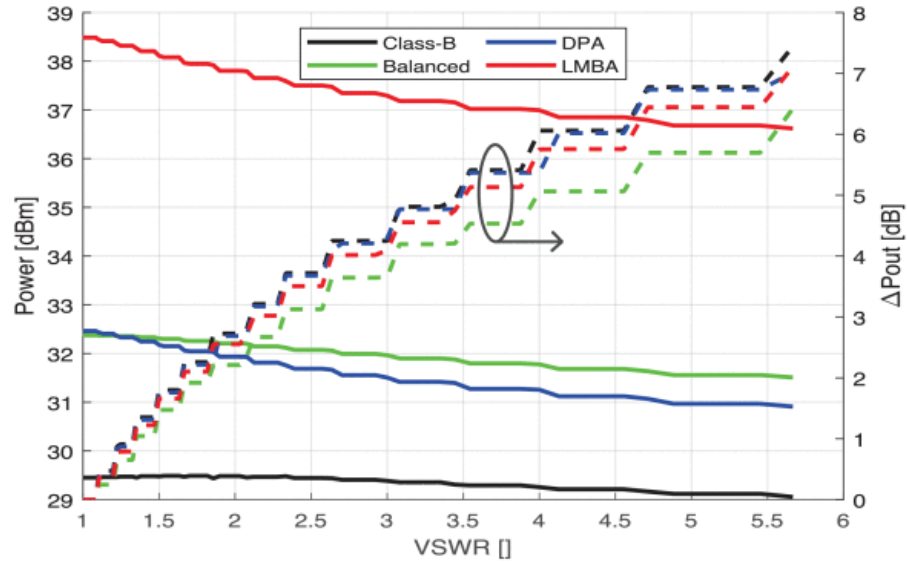
Several approaches have been adopted to combat the performance variation over different load impedance. [7] exploits the usage of DAC-based PA with a multi-port network synthesis technique to improve the performance of PA with mismatched load. However, this technique requires prior knowledge of the load impedance.

In [28], it is shown that series and parallel Doherty PA topology demonstrates different degradation to their output power for different phases of load impedance under VSWR variation. To exploit such observation, a re-configurable topology series/parallel Doherty PA is designed with an improved output power and efficiency over 3:1 VSWR circle.

The above technique require load-sensing for VSWR resiliency which introduces extra complexity and losses at the PA output. To avoid such a problem, there exists PA architectures that have inherent mismatch tolerance such as Balanced PA topology. Figure 2.17[5] studies and plots the behavior of different PA topologies under load mismatch conditions, and the advantages of Balanced PA is apparent.



(a) Worst case linearity and average of average Drain Efficiency



(b) Average and ΔP_{OUT} output power

Figure 2.17: Performance parameters under mismatch for different PA topology[5]

2.5 Literature Review Summary

Reference	[15]	[34]	[38]	[20]	[48]	[3]	[6]	[10]
Year	2021	2020	2021	2021	2021	2019	2020	2019
Frequency (GHz)	30	31	36	26.5/ 37	28	28	28	39
Technology	45nm SOI CMOS	65nm CMOS	28nm CMOS	28nm CMOS	22nm FD SOI CMOS	65nm CMOS	65nm CMOS	65nm CMOS
Topology	Dual Drive	2-stack class-AB	Doherty-like LMBA	2-stack class-AB	3-stack Doherty	single-ended CS	single-ended 2-stack class-AB	CS class-AB
Gain (dB)	20.4	18.9	18	13/ 10.7	26.1	15.8	18	14.4
OP _{1dB} (dBm)	19.0	15	19.6	18.7/ 18.6	21.1	14	22.7	16.8/ 20.2
P _{sat} (dBm)	20.1	17.1	22.6	20.2/ 19.1	22.5	15.6	23.2	20.7
PAE _{peak} (%)	49.7	38.2	32	33.6/ 32	28.5	41	35.5	35
PAE _{6dBBO} (%)	N/A	N/A	24.2	N/A	22.1	25	N/A	N/A
AMPM @OP _{1dB} (degree)	N/A	2	3-4	N/A	N/A	1	1.6	7/ 1

Table 2.1: Comparison of Recent Ka-Band Power Amplifiers in CMOS Technology

A number of state-of-the-art Ka-band CMOS PAs from publications in recent 3 year period are selected and compared in Table 2.1. The most commonly reported topics of choice are Doherty PA and linearity-enhanced class-AB PA. The saturated output power achieved with CMOS in this frequency range is generally around 20-23 dBm, with peak PAE just over 40%. An exception being [15], which achieves an outstanding peak efficiency of nearly 50% with dual drive technique discussed in a previous section. Highest achieved

AM-PM distortion regardless of frequency or technology is generally below 1 degree up to OP1dB, but simultaneous AM-AM and AM-PM enhancement is often difficult to achieve. The most common transistor topology among the recent PA in literature is common-source and 2-stacked. [48] uses 3-stacked transistor for a Doherty design, achieving a saturated output power of 22.5dBm in a chip area of only 0.2 square-millimeters. However, the peak efficiency is not as high as other state-of-the-art Doherty designs, partially due to this high FET-stacking.

Chapter 3

Transistor-Level Design

3.1 Transistor Layout Topologies and Trade-offs

The main parasitics of concern when it comes to active devices in SOI CMOS are the following: R_g , R_s , R_d , C_{gs} , C_{gd} , and C_{ds} . Transistor directly corresponds to the inherent performance limit a circuit can achieve in a given technology and it is desirable to maintain as close to the theoretical maximum achievable performance as possible. The transistor level parasitics in the lower metal layers cause a performance deterioration from the original FET model, and it is highly desirable to minimize these parasitics to boost the performance. The choice of active device layout directly impacts the parasitic resistances and capacitances of the transistor. In general, more parallel interconnects result in less parasitic resistances, however, that increases the parasitic capacitance in the layout. As a result, there is a trade-off in choosing the desired layout topology in order to minimize the parasitics that are the most detrimental to circuit performance while sacrificing the less important ones. Thus, the choice of device layout varies depending on the application as well as circuit topology in the context of a specific design.

A few popular transistor layout topologies are listed and discussed qualitatively in this section. Layout decisions at the finger level can be classified by how the contacts at each transistor finger are connected to the surrounding metal layers. Whether the gate, drain

and source interconnect wires make the contact on one or both sides of the transistors is one of the biggest factor affecting the parasitic resistance and capacitance to all the terminals, as well as how the transistor unit cell connects to each other when scaling to bigger transistor widths.

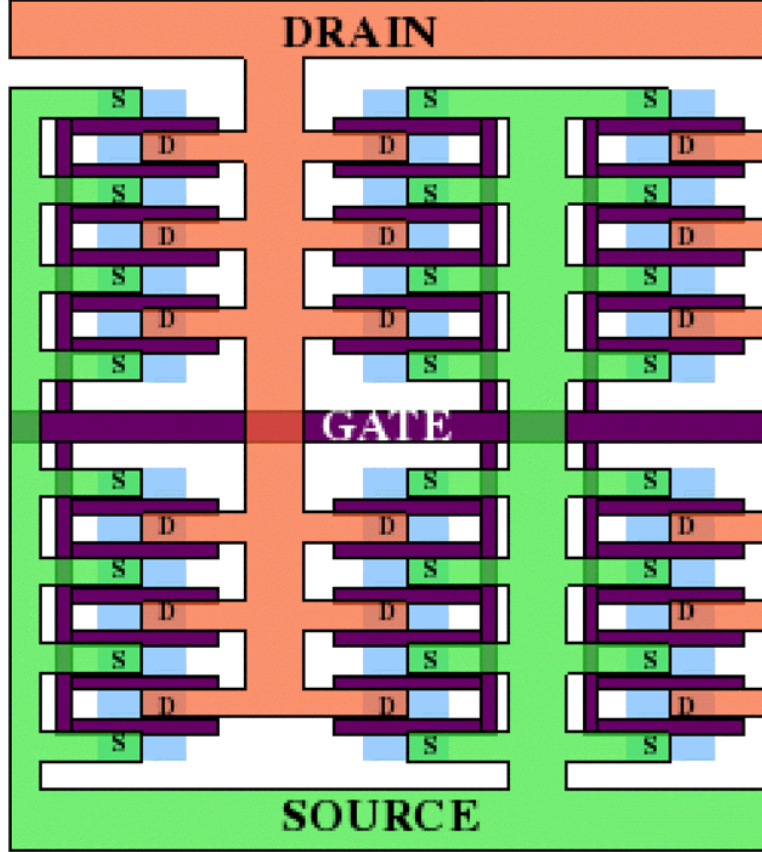


Figure 3.1: Ladder-like source and drain interconnect [37]

The example in [37] adopts the "zipper" layout topology, as shown in Figure 3.1. This layout structure has a single-sided contact for the gate, source, and drain. The feed-lines are perpendicular to the finger orientation. This structure inherently comes with high parasitic resistance and high interconnect inductance due to the long single-side metal interconnect on the lower layers. This increase in R_g and R_s can cause a significant deterioration in f_{max} . As mentioned in the paper, the post-layout peak f_{max} was cut in

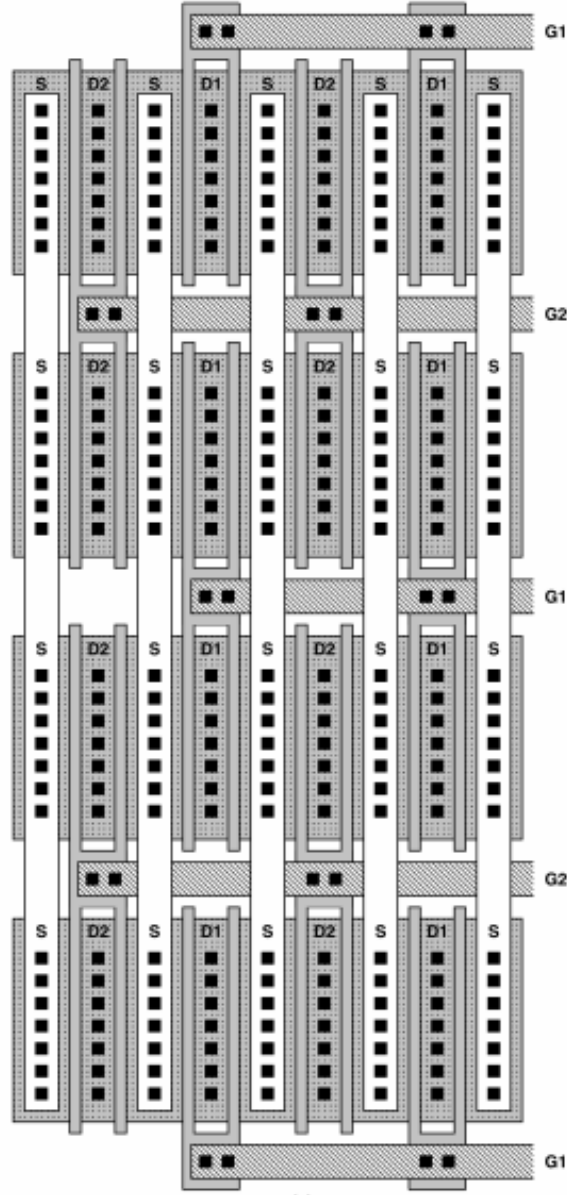


Figure 3.2: A layout technique to minimize impact of source parasitic resistance in PA [26]

half, dropping from 500GHz to 240GHz in a 155-um device. However, the capacitances between the interconnects are relatively lower compared to other layout topologies. The reduced C_{gd} and C_{ds} in this structure improve the transistor unilaterally and reduce the

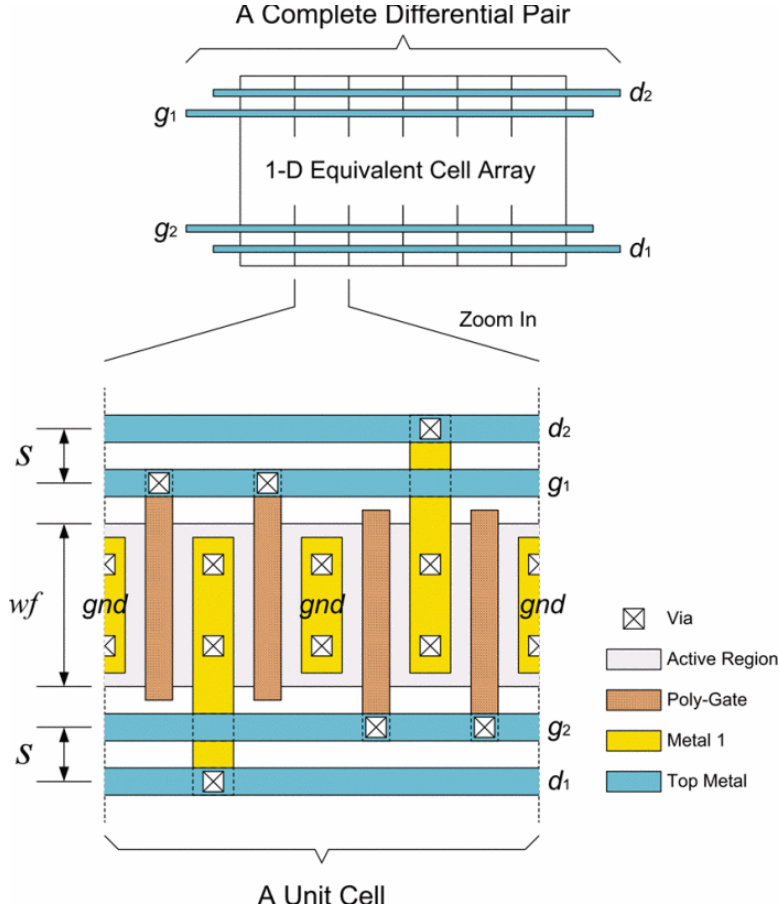
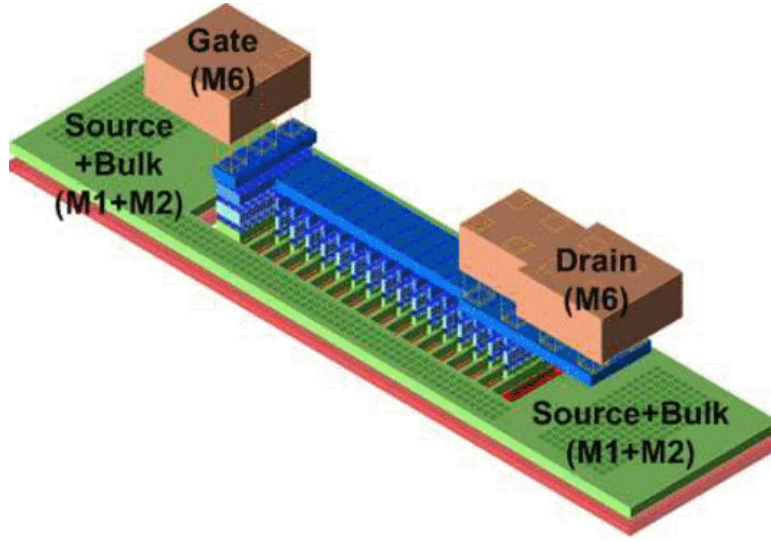


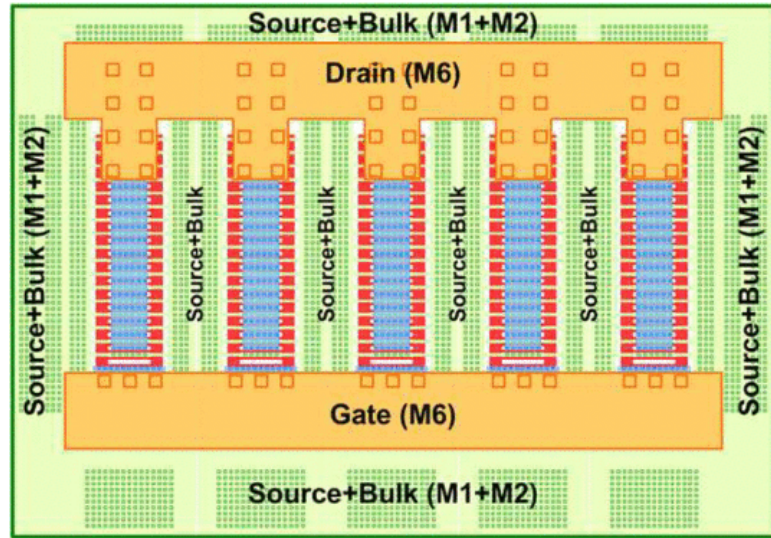
Figure 3.3: Layout in [13]

output capacitance. The former affects stability and gain, whereas the latter is a key factor to consider at high frequencies, where capacitive reactance is low enough to effectively short circuit most of the harmonics, and possible even at fundamental frequency, if the transistor size is large enough and the frequency of operation is high enough.

Some layout topologies make significant improvements in source and gate resistances compared to the previously discussed "zipper" topology. For example, Figure 3.2[26] presents a layout topology to try to reduce R_s to increase efficiency and output power for PA application. Through source sharing, neighboring transistor fingers share their source terminals, and parasitic resistance in the source is minimized. In addition, the lengths of



(a)



(b)

Figure 3.4: Layout in [47]

source interconnects are reduced. Figure 3.3[13] proposes a novel layout that does this in an interesting way. Author not only adopts the source sharing technique but also interleaves the transistor fingers in a differential pair for a layout parasitic-based neutralization

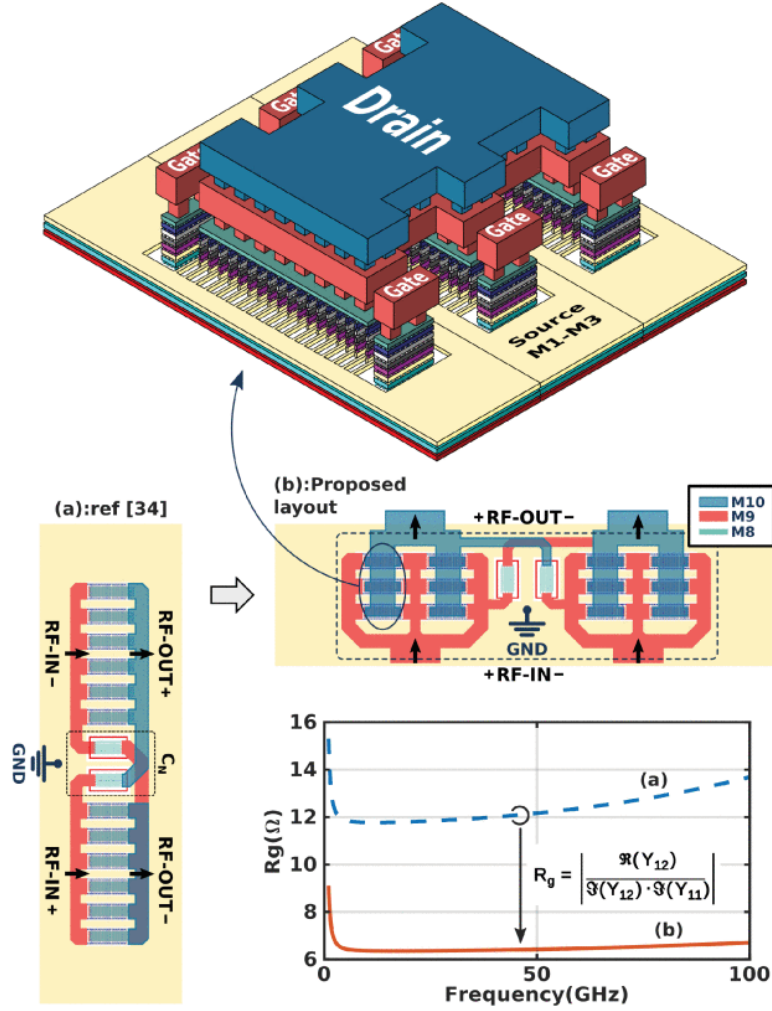


Figure 3.5: Layout in [38]

technique. The trade-off in these two topologies is that no improvement is made to the gate resistance, so there will still be a large decrease in gain. An additional problem exist in [13] that is its scalability. The way in which the unit transistor cells are connected to the signal lines in the upper metal layers is one thing to consider here. As the gate and drain interconnects are running next to each other on both sides, they need to be connected out in lower metal layers first before going through vias to upper metal layers. As such, the parasitic resistance and inductance in both gate and drain are expected to further increase

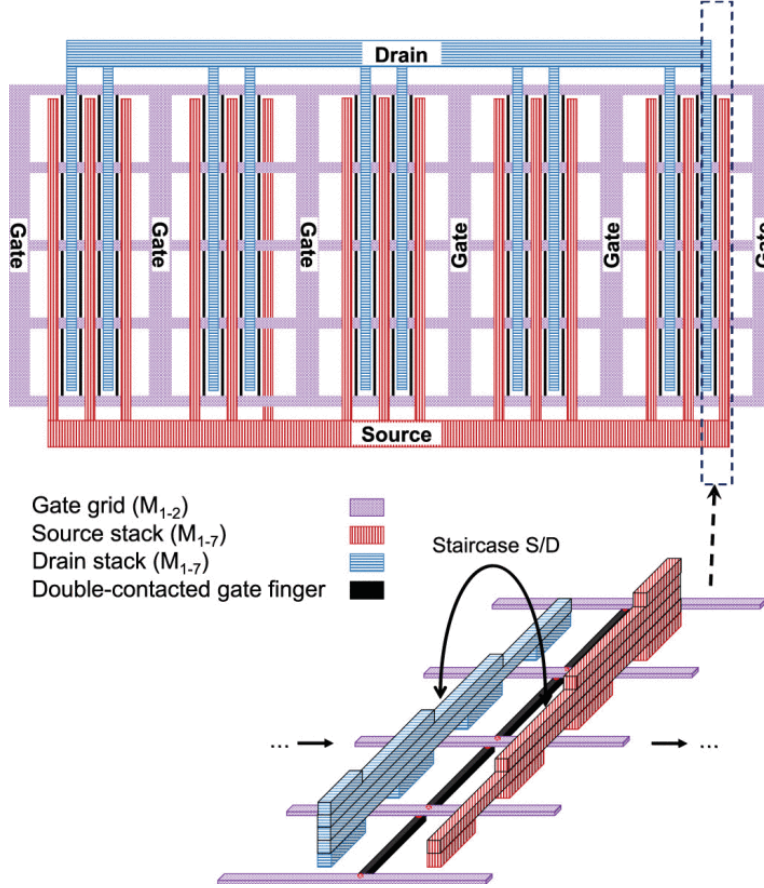


Figure 3.6: Ladder-like source and drain interconnect [11]

especially when adopting this layout for large transistor sizes.

To improve the power gain, a low R_g is needed. The first change here is to adopt a layout topology with dual side gate contact. Signal can be fed to the fingers either in series (perpendicular to finger orientation) or in shunt (parallel to the finger orientation). The layout topology shown in Figure 3.4[47] shows double sided contact for both gate and source terminals. The gate signal feed makes it easy for large transistor sizes when scaling to high number of multiplicity count. 3.5[38] increases the gate signal path to two, still feeding the transistor fingers in series, but adding another terminal at the other end of the unit cell, attempting to further reduce the gate resistance. For smaller transistor sizes,

parallel feeding can be adopted for further reduction in gate parasitic. In [23], parallel feeding can be seen. Compared to series feeding, this further reduces the gate resistance because the signal travels less distance in the lower metal layers. However, scalability may pose a challenge for large transistor sizes. A grid-like structure is used in Figure 3.6[11] for a low-noise amplifier (LNA) design. This topology also focus on minimizing the gate resistance, by adopting double-side gate contact while having transistors in small unit cells with low number of finger count. The source and drain connections are made in a ladder-like structure to reduce the Cds between the wiring.

The transistor layout chosen for the PA in this thesis is the topology shown in Figure 3.4[47], offering low parasitic resistance on all terminals, while making it easy to scale to large multiplicity of transistor unit cells.

3.2 Study of Transistor-level Stability

3.2.1 Differential Stability and Stabilization Techniques

The capacitive neutralization technique cancels out the effect of gate-drain transistor gate-drain capacitance when the amplifier is operated in the differential configuration. Generally, a properly neutralized amplifier can show unconditional stability in the differential configuration across a fairly wide bandwidth.

An example is shown with two 150um transistors in differential common-source configuration. S-parameter sweeps are done, while varying the neutralization capacitance, the change in the maximum achievable gain, and the stability factor are plotted. Transistor parasitic extraction results give that the 30um transistor unit cell used in the layout results in parasitic Cgd of approximately 9 fF. Estimating the parasitic capacitance in a 150um transistor gives 45 fF of Cgd. Figure 3.7 shows the graph of the maximum achievable gain at 40GHz with increasing neutralization capacitance. Figure 3.8 shows the plot of the stability factor against the increasing neutralization capacitance, with 10 different traces at 10 evenly spaced frequencies between 10GHz and 100GHz. As seen in the figures, the improvement in both the power gain and stability at the point of unilaterality is clearly

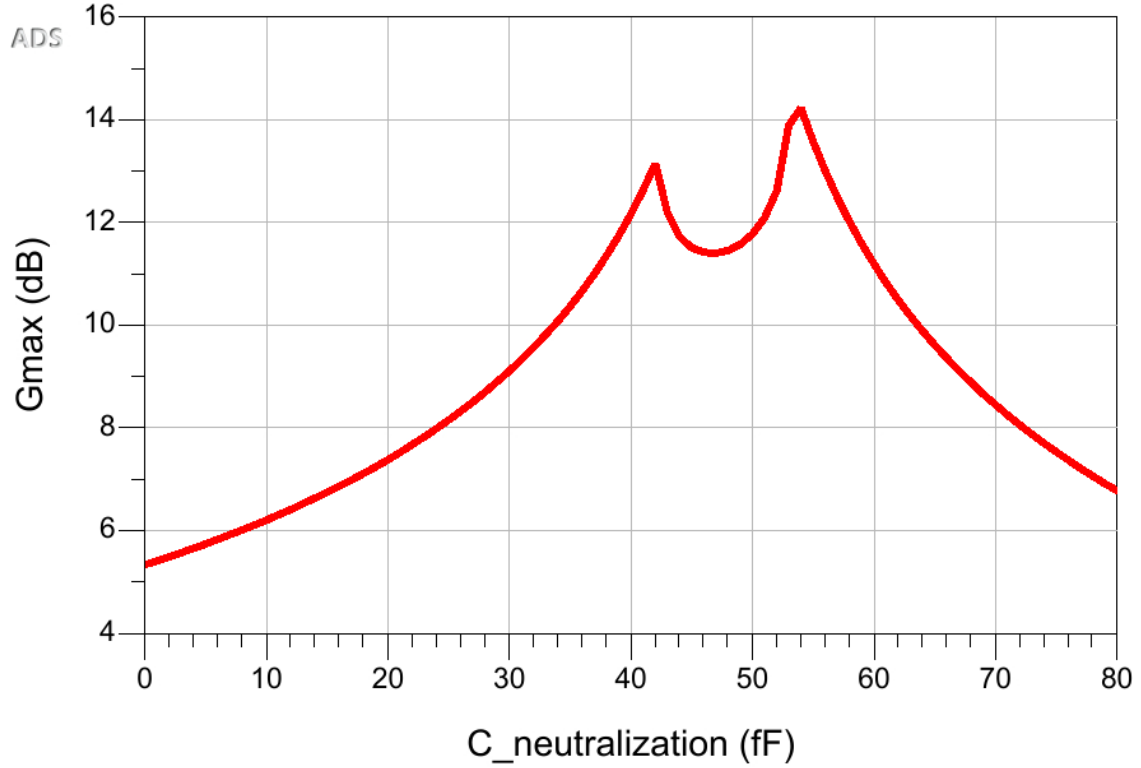


Figure 3.7: Gmax at 40GHz plotted vs neutralization capacitance in a 150um differential common-source

observable. With a choice of neutralization capacitance of 47 fF, the simulation shows a Gmax of 11.5dB at this bias point, with a stability factor above 1 in the wide frequency ranges.

However, one issue remains if one plots the differential stability below 10GHz. At low frequencies, stability issue remains, as shown in Figure 3.9. A simple method of adding loss at the input resolves the issue, in form of a pair of resistors in the gate biasing feed-line. With a pair of 50 Ohm resistors added, the stability factor is shown in Figure 3.10. One can observe unconditional stability in differential configuration up to at least 100GHz. Above 100GHz, the transistor has poor gain, and the chance of oscillation gets further reduced with the increase in frequency.

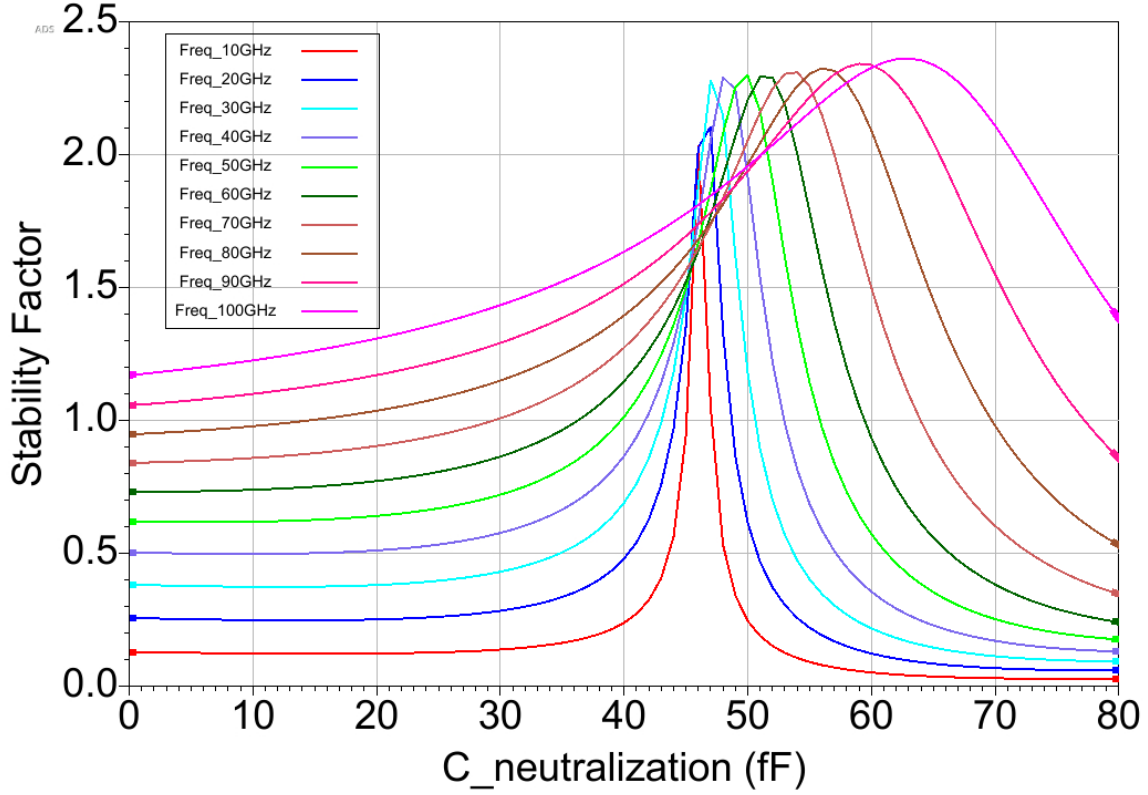


Figure 3.8: Stability factor vs neutralization capacitance in a 150μm differential common-source across 10-100 GHz

3.2.2 Common-Mode Stability and Stabilization Techniques

As discussed above, the capacitive neutralization technique provides significant improvements in stability. However, as the technique requires differential operation in order to function properly, this technique as expected does not work on common-mode stability. In order to improve common-mode stability of the amplifiers, a closer look back at the theory of stability and oscillation is necessary. With proper supporting theory, the corresponding techniques of stability enhancement can be proposed. Firstly, one can recognize that the load impedance of the transistor in the drain will not be purely real for most of the frequency ranges. It is worth recalling that one configuration to generate negative resistance is to look at the gate of a common-source amplifier with inductance in the drain.

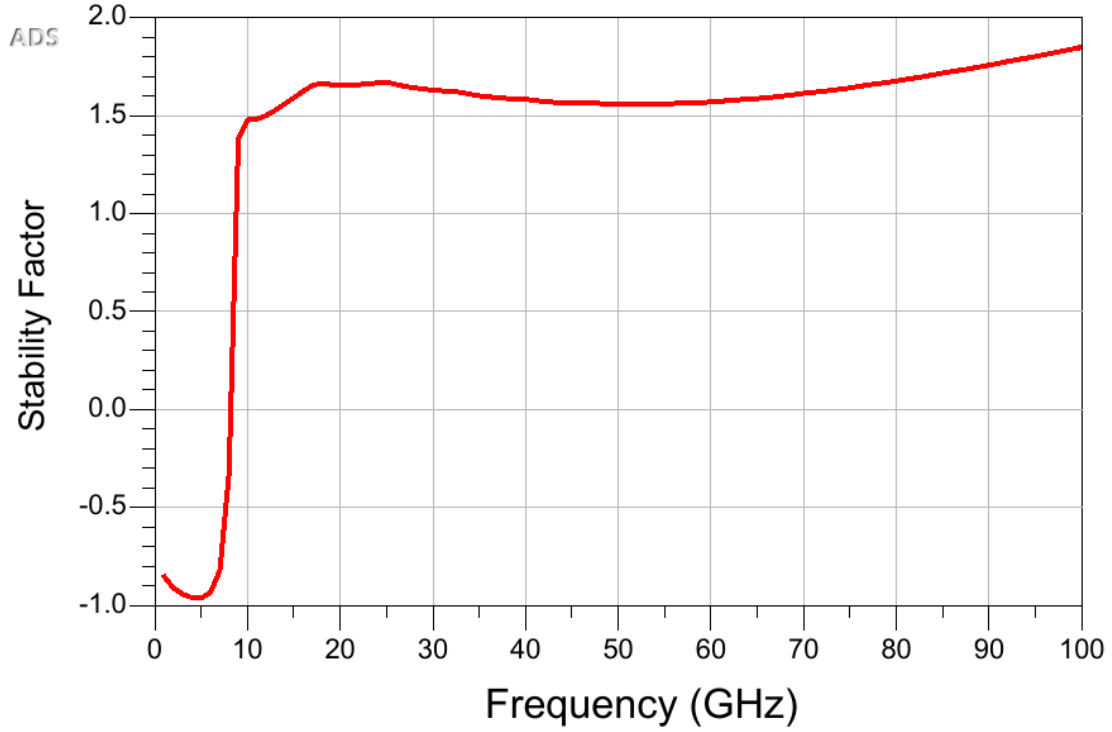


Figure 3.9: Stability factor from DC-100GHz before adding stablization resistor at input

If one models the output load with shunt LC-network, then it is not hard to deduce that at frequencies below the LC tank's resonance frequency, the load will behave inductively. This in theory produces a risk for instability. From theory, the input impedance of a common-source transistor with inductive load can be approximated as follows:

$$Z_{in} = -\frac{1}{\omega^2 g_m C_{gd} L_L} \quad (3.1)$$

In order to create an oscillation, the magnitude of the negative input conductance needs to be larger than the one looking into the source. Hence, to improve the stability, one can reduce the transconductance, reduce gate to drain capacitance, or reduce the load inductance. This is in agreement with the intuition. This also provides an intuition for when stability could deteriorate. Transistor transconductance increases with increasing bias, in-

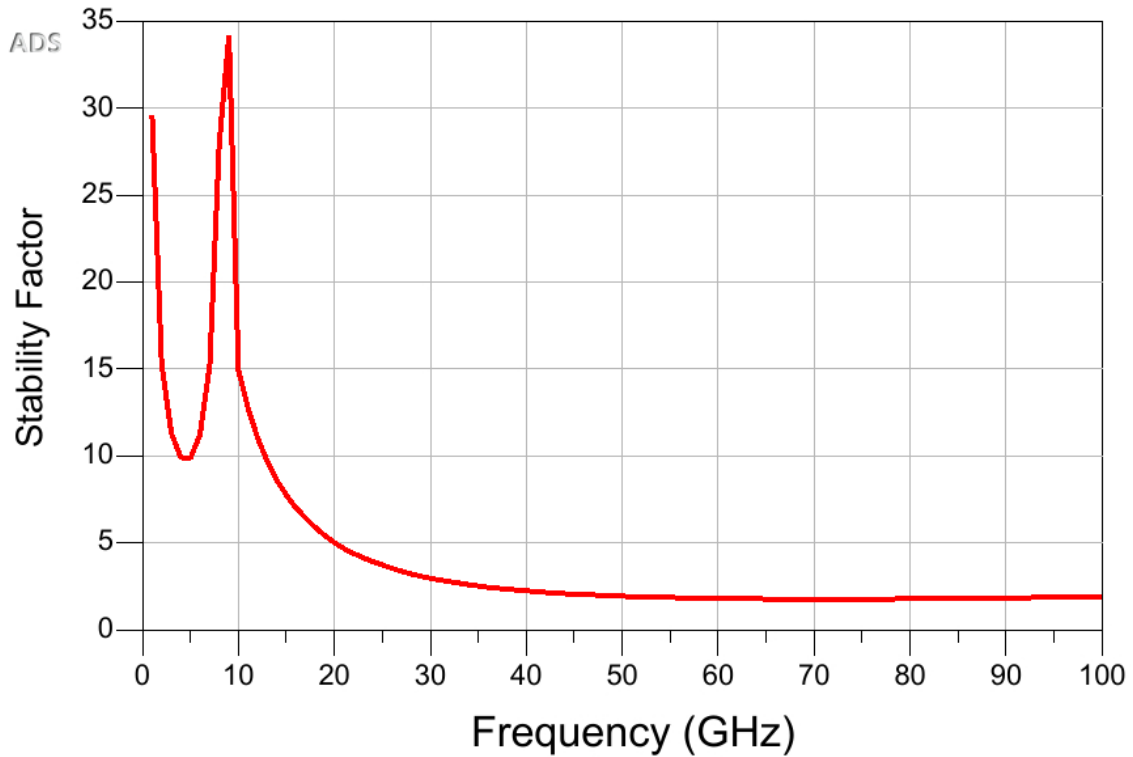


Figure 3.10: Stability factor from DC-100GHz after adding stablization resistor at input

creasing supply voltage, increasing transistor width, and decreasing frequency. Firstly, it can be noted that larger transistor not only results in increased transconductance, but also in an increase in gate-to-drain capacitance. Secondly, improper decoupling on the feedline at output transformer center-tap where the drain supply is connected could lead to an increase in drain inductance.

When it comes to reducing the transconductance and thus reducing the gain, one popular and proven technique is through inductive degeneration. When extra inductance is added between the common source node and the ground, the transconductance in the common mode configuration can be reduced while not affecting the power gain in the differential configuration. However, one issue here is that when the common mode issue takes place at lower frequency ranges, the required inductance to make a visible difference

in stability may become too large for chip area considerations. One method was proposed to use an shunt LC resonator at the frequency of interest in order to generate a high impedance at that frequency. However, this resonator at high frequency would behave capacitively. Recalling that a common-source configured transistor with capacitance in the source is also a way to generate negative resistance. Thus, because of the concern about this added capacitive reactance in the source at higher frequencies, this method was not implemented. The common-mode rejection ratio of the matching network transformers were also examined. The center tap was terminated with a capacitor to ground by design, and this improves the common mode rejection of the transformer-based matching network.

Lastly, if it turns out that chip does turn out to be unstable, additional methods regarding bias and supply voltage can be taken. By slowly increasing both from 0, one may find a bias and supply voltage below that of the original designed ones in simulation. Because of the lower transconductance at this lower biasing point, stability is expected to be improved in this way. However, this should be avoided if possible because a different bias and supply voltage would cause a significant change to PA performance, particularly large signal linearity. The pre-designed load impedance would also deviate from the ideal impedance. This method might find an operating point without oscillation given an unstable amplifier, but it would nonetheless pose a significant concern to deteriorating the overall 2-stage PA performance, which would render the design meaningless.

3.3 Stacked-FET Design and Considerations

3.3.1 Comparison of Stacked-FET and Cascode Topology

The main goal of trying to adopt a cascode-like structure for power amplifiers is their merit in providing increased output power by increasing the peak voltage swing at the drain output. However, the issue with CMOS technology is that the breakdown voltage for individual transistor is low. If the drain voltage swings across common-source and common-gate transistors are out of phase, it is likely that one or both of these transistors will experience a drain-to-source voltage swing amplitude higher than the tolerable level

specified for reliability concerns. Thus, the designer would need to keep the drain-to-source voltage swing across different transistors in the stack balanced. The second consideration for implementation of the topology in power amplifier applications is the voltage swing at the gate of the common-gate transistor. At high output power levels, the output terminal sees a full voltage swing. However, in a conventional cascode structure, the gate of the common-gate devices sees an equivalent RF ground. This means that the voltage swing at the gate is approximately 0. This poses an issue on the reliability of the gate-to-drain junction due to the high voltage swing amplitude across these two nodes.

In order to overcome these aforementioned issues, several changes are made in the stacked architecture compared to the conventional cascode topology. The first being a finite impedance seen at the gate of the stacked transistor. This forms a capacitive voltage divider with drain-to-source capacitance and results in a nonzero voltage swing at the gate node, thus alleviating the reliability problem across gate-drain junction of this transistor. The second change being that the impedance presented to the drain of the bottom common source transistor is designed to be R_{opt} , compared to the load impedance at the drain of the stacked structure, which is $2R_{opt}$. This leads to in-phase series combining of the voltage swings across each transistor.

One trade-off of the stacked-FET structure compared to conventional cascode is its lower gain. In the stacked structure, it can be shown that due to the existence of a capacitor at the gate of the stacked transistor, the output impedance of the stacked structure is reduced. This reduces the voltage gain of the stacked-FET amplifier. This can be verified with a simple S-parameter sweep versus the gate capacitance. As shown in Figure 3.11, the maximum available gain of a given stacked-FET amplifier is swept against the gate capacitance and the resulting monotonically increasing trend is apparent. With increasing capacitance, the circuit behaves more and more like the classic cascode, and the gain increases. However, the increase in output power without reliability issues in the stacked-FET architecture is worth this sacrifice in gain.

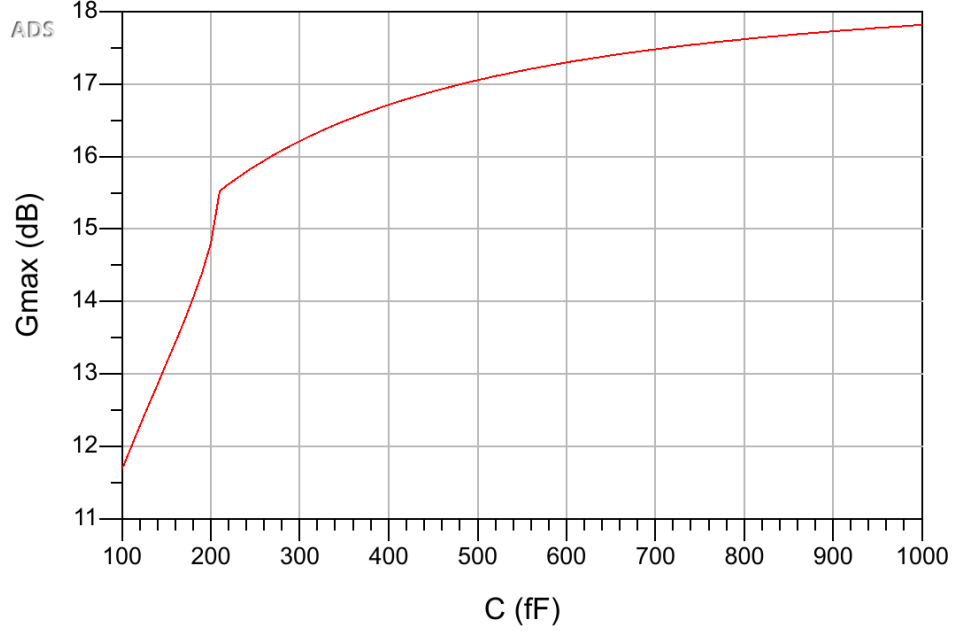


Figure 3.11: Maximum available gain of a 2-stacked 180um amplifier vs varying capacitance at the gate of the stacked transistor

3.3.2 Design of Stacked-FET Circuit

In Figure 2.2, neglecting transistor small-signal output resistance and drain-source capacitance, the impedance Z_d looking up towards the source of the stacked transistor can be expressed as follows:

$$Z_{d,k-1} = \frac{C_{gs,k} + C_k + C_{gd,k}(1 + g_{m_k}Z_{d,k})}{(g_{m_k} + sC_{gs,k})(C_{gd,k} + C_k)} \quad (3.2)$$

$C_{gd,k}$ and $C_{gs,k}$ are the gate-to-drain and gate-to-source capacitances corresponding to the k -th transistor in the stack. $Z_{d,k}$ is real when frequency of operation is much lower than the transition frequency of the transistor. To ensure that the voltage swings are equally distributed and are added in series to the stack, the impedance $Z_{d,k}$ should be designed to be k -multiple of transistor's optimal load resistance R_{opt} .

Assume that the transconductance is much larger than sC_{gs} , if $Z_{d,k}$ is real and chosen to be kR_{opt} , setting the real part of $Z_{d,k-1}$ to $(k-1)R_{opt}$ gives the capacitance C_k to the expression below:

$$C_k = \frac{C_{gs,k} + C_{gd,k}(1 + g_{m_k}R_{opt})}{(k-1)g_{m_k}R_{opt} - 1} \quad (3.3)$$

In reality, the finite output impedance of transistors and the non-zero drain-to-source capacitance are sources of errors to this approximation, but it can be used as a first guess nevertheless. In the actual design test bench, it is always necessary to sweep the value of this capacitance to determine what the most suitable choice of this capacitance is to account for the non-idealities. In addition, the imaginary part of impedance $Z_{d,k-1}$ will not be 0, hence an additional matching network is needed to match this impedance to R_{opt} . This could take the form of a series or shunt inductance in this particular technology and frequency of interest.

Voltage distributions at various nodes of the circuit is also an important factor to pay attention to. Due to the high supply voltage to the stacked PA, gate bias of the stacked transistors need to be properly biased such that the various voltage swings across the different junction pairs on the transistors are less than voltage limit for reliable transistor operations. When biased under class-AB, increasing the RF input power will cause the DC component of the output current to increase. The fixed DC biasing voltages at the transistor gates should be considered for the maximum operating power level in order to ensure safe operation. As such, the DC gate biasing voltage can be expressed as follows:

$$V_{G,k} = \frac{k-1}{k}V_{DD} + V_{GS-sat,k} \quad (3.4)$$

where $V_{GS-sat,k}$ is the DC component of the gate-to-source voltage of k th stacked transistor.

3.3.3 Analysis and load-pull of the 180um PMOS neutralized Stacked-FET

In order to determine the optimal transistor choice for driver and PA stage, study is done to consider the performance trade-off. Power transistors of various size in both common-source and 2-stack architecture are analyzed based on their load-pull results, mainly based on its output power and linearity.

In order to maximize the output power while maintaining an acceptable level of linearity, 2-stacked transistors are considered for PA stage. After extensive load-pull and large-signal CW simulations with available transistors, a differential 180um 2-stacked transistor design was chosen for the PA. The original circuit was designed following the theory of stacked-FET PA design discussed above, except for some minor changes and additions. Firstly, the inter-stage node is matched with a shunt inductor followed by a series capacitor to ground. The shunt inductance matches impedance seen by drain of the bottom transistor to the desired impedance in differential mode in the main frequency band of interest. The series capacitor helps modify the impedance seen at second harmonic frequency, where the signal is in common mode. Secondly, PMOS capacitor is used for neutralization to reduce the AM-PM non-linearity by linearizing the input capacitance as mentioned in [10] and discussed in the literature review section of this thesis. Although it was observed that gate voltage swing in this core is smaller than desirable levels in large signal simulations, this is due to a relatively large gate capacitance choice at the gate of the stacked transistor. For future work, potential improvements can be made here for better reliability and long-term operation, especially considering the stacked transistor.

The load-pull results of the the differential 180um 2-stacked circuit with 180um PMOS varactor neutralization is shown below in Figure 3.12. The input power is set at 5dBm, with a gate biasing voltage of 0.3V and drain supply of 2.4V.

The output power here in this load-pull figure is around the 2 dB compression point. Which indicates that with extra linearization from driver stage, this core should be able to reach an output 3dB compression point of over 20 dBm without a problem. With the enhanced AM-PM linearity as an added bonus, this core is chosen for this PA design.

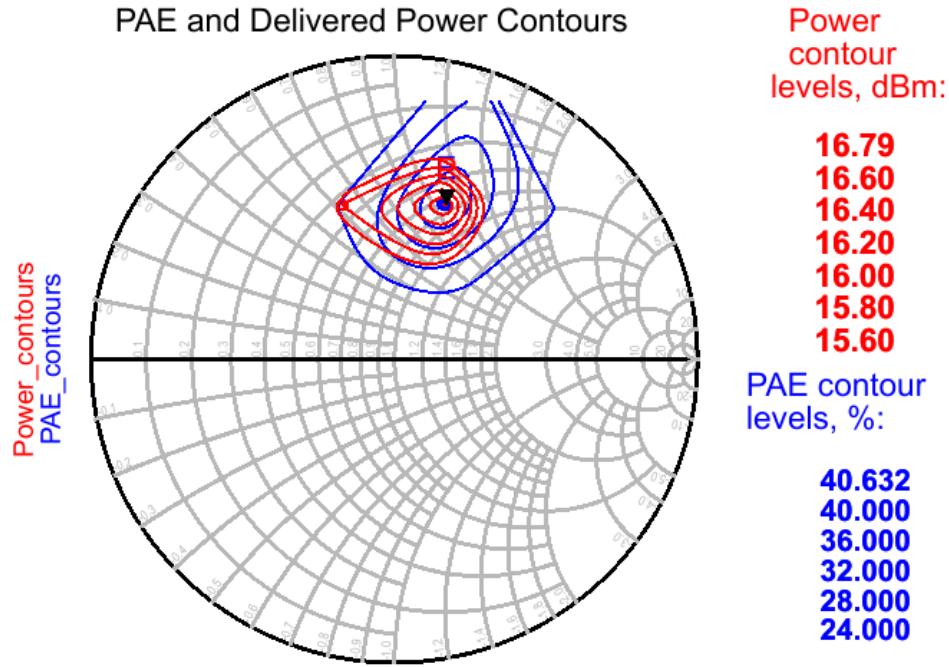


Figure 3.12: Load-pull results of the differential 180um 2-stacked transistor core

Similar analysis is performed with transistors for the driver stage. In this design, priority is given to linearity as that is the focus of this project, hence a relatively large-sized differential CS 150um driver was picked. This driver can maintain operation in the linear region up to the power level required to drive the main PA stage to its 3dB compression point. Despite the extra power consumption due to a large driver size, the degradation of the overall PAE of the 2-stage PA is minimal because of the power gain from the last PA stage.

Chapter 4

37-43 GHz Balanced PA Design

4.1 PA Specification and High-level Design Decisions

The PA architecture, transistor sizing, and various design trade-offs as well as considerations are crucial to the eventual performance the PA is capable of achieving, and thus a good step the designer should take is to analyze of the initial performance specifications for an initial plan.

The most fundamental performance parameter of the power amplifier is the saturated output power that the PA is capable of achieving, given a certain chip area to design with. This can also be quantified with the power density in units of watt per square millimeter. In this design, the initial requirement indicates a saturated output power of 20 dBm within an estimated chip area of 300 x 800 square micrometers. This gives an idea of roughly what topology the PA will allow. The individual circuits will be differential, given the advantage they have in terms of simple implementation of the cross-coupled capacitor neutralization technique, improved common-mode rejection ratio, and the inherent DC-blocking capabilities of compact transformer-based matching networks. Thus, assuming differential operation, the given width of 300um means that at most 2-way power combining can be used before the final layout gets unreasonably crowded. And from prior basic load-pull simulations of various transistors at this frequency range, it is known that the

gain requirement cannot be achieved with a basic single stage class-AB PA. Hence it is determined that the area will be fully exploited with a differential 2-stage PA with 2-way power combining. Knowing this, each individual sub-PA can be determined to be 2-stack in order to have a good balance between high output power and low degradation to linearity and power efficiency. Then, with a 17dBm requirement on each sub-PA, a differential 180um 2-stacked core is chosen for this design.

As the PA is planned to operate in a large phased array system, it is important to consider the challenges the system would place on the PA. As discussed before, this comes mainly in two ways: good output matching and good performance resiliency against load VSWR variation. It is aimed that the PA would have a better than 10dB return loss across the entire frequency band of interest at both input and output, as well as maintaining reasonable performance over various phases on the 2:1 VSWR circle at the output. In order to deal with these challenges, the balanced amplifier topology is chosen for this PA. The balanced topology comes with two identical amplifiers in each of its two branches, and a total of two quadrature hybrids at input and output, respectively. Due to the inherent operating principle of the balanced amplifier, ideally the balanced PA would see no reflection to the matched port at input and output as a result of the reflections canceling out at the port of interest and combining at the isolated port terminated with the matched load. This, combined with the improved VSWR resiliency as discussed in the previous chapter, makes the balanced amplifier topology a very attractive topology of choice. It is structurally simple to design with, offers low complexity and simple design flow. It is estimated that the final design will exhibit a return loss of 20dB, much exceeding that specified in the initial plan.

The other performance specifications including gain, efficiency and linearity will be mainly considered in the following design process, with trade-off being carefully considered and made. The gain target of 15-20dB is completely reasonable with 2 stages. The peak PAE target of 40% and linearity target will be considered following the design methodology described in the following section.

4.2 Design Methodology

To achieve a well-balanced trade-off between power efficiency, gain, and linearity, class AB biasing is considered for both the driver and PA stages. The exact biasing points of 2 stages are crucial to overall PA's performance trade-off, and will be discussed more in detail in the next subsection. But initially, with the maximum PAE targeted at 40% and the OP1dB aimed to be close to saturated output power, it can be certain that the design will be in a deep class-AB biasing point, not too far away from the optimal sweet-spot bias based on experience and simulations performed in load-pull and prior harmonic balance simulations.

In a 2-stage PA, the design of the second stage will have the largest impact on overall output power and efficiency. Among the various design considerations and parameters, the load impedance is crucial to the final achievable efficiency and output power. Hence the output matching network of the final PA stage that presents the load impedance to the last stage, should be designed with a heavy focus on output power and efficiency. The impact of first driver stage has on overall 2-stage PAE will be reduced by the gain of the last stage, hence the driver can have a heavier design emphasis on linearity.

Starting off with a rough bias point, the output matching network of the second stage is designed on the basis of the load-pull, and the input to the first stage is designed assuming the conjugate match condition.

As for the inter-stage matching network, careful considerations must be made, as it is crucial to the performance of both stages and the most complex in terms of design considerations. Here, the method of inter-stage load-pull as discussed in [44] is tweaked and adopted. The traditional approach of designing the load and source of the two stages separately before synthesizing the final inter-stage matching network is overly complex and consumes a large amount of time for tuning iterations following the initial design. The method proposed here use achievable numbers of the final transformer-based matching network as design parameters to sweep the possible combinations of them while comparing the resulting large signal behavior of the overall 2-stage PA directly. This saves on design time and makes for a faster, more strategic design approach.

4.3 Biasing Considerations

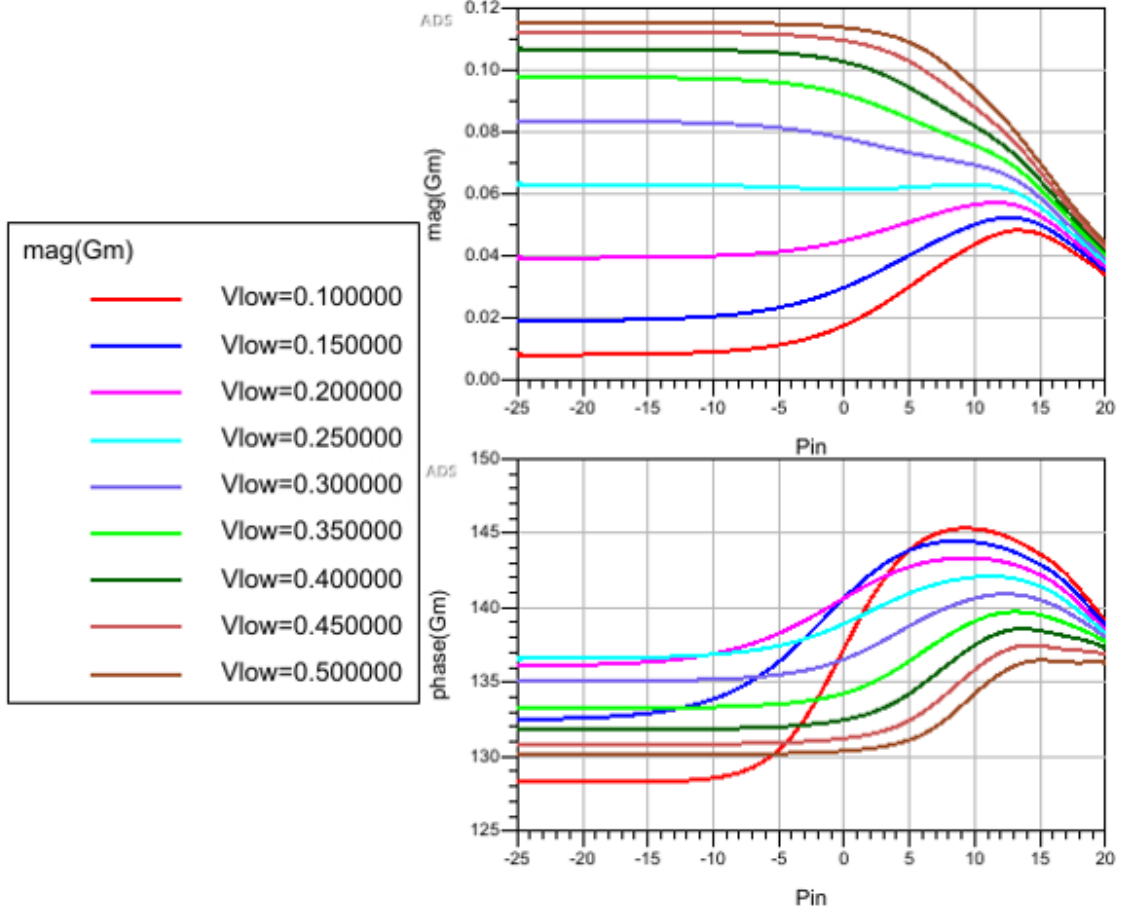


Figure 4.1: Magnitude and phase of the large signal transconductance vs input power for the 180um core with various biasing points

Biasing level of the PA affects a few performance specifications. For example, as the biasing point of a class-AB PA reduces and approaches a class-B bias, the drain efficiency of the PA increases. Another important factor to consider is the linearity of the PA. Both AM-AM and AM-PM nonlinearity is affected by the biasing point. For AM-AM linearity, the biasing level affects the magnitude of different orders of nonlinearity that the transistor generates. For instance, the 3rd-order transconductance changes sign in deep class-AB bias,

and this biasing point is commonly known as the "sweet-spot" biasing. Operating the PA near this biasing point leads to best AM-AM linearity and a low third-order harmonic distortion level. However, the biasing point also affects the PA AM-PM distortion. As seen in Figure 4.1, the phase distortion of the transconductance can be seen to increase as the bias level is reduced and the optimal AM-AM biasing point is approached. Thus, a trade-off needs to be made regarding the exact biasing point in terms of AM-AM, AM-PM and efficiency.

Note that biasing also affects power gain; however, it is much more difficult to boost the power-added efficiency and linearity given a poorly designed PA than it is to compensate for small signal gain. It is possible to add extra drivers in the TX chain to boost the total gain to acceptable levels with some extra chip area as a trade-off. Thus gain is not a factor that influence the decision making in choosing the biasing point.

4.4 Matching Network Design

The transformer is often used in sub-100GHz millimeter wave IC designs for its compactness and wide achievable bandwidth. Inherently it can be modeled like a T-coil network and thus offers the benefits of a multisection LC matching network. Hence, transformers offer better bandwidth than simple inductors over the same chip area.

For this design, a non-ideal transformer model with parasitics that has been used in the past is chosen for schematic design phase of this PA. The non-idealities are modeled as follows. A parasitic capacitance of the shunt to the substrate followed by a shunt RC network is used to model the loss of the substrate. A capacitor between primary and secondary coil is used to model the coupling capacitance. Lastly, a resistor is used to model the resistive loss in the metal conductors. It is worth noting that the skin effect is not modeled in this circuit model; hence one potential improvement that could be made here is to include the skin effect circuit model in order to provide a better match for inductor quality factor. The transformer elements models include an ideal transformer and the leakage inductance which takes into effect of non-ideal coupling.

Therefore, the main design parameters in the design process are the desired primary

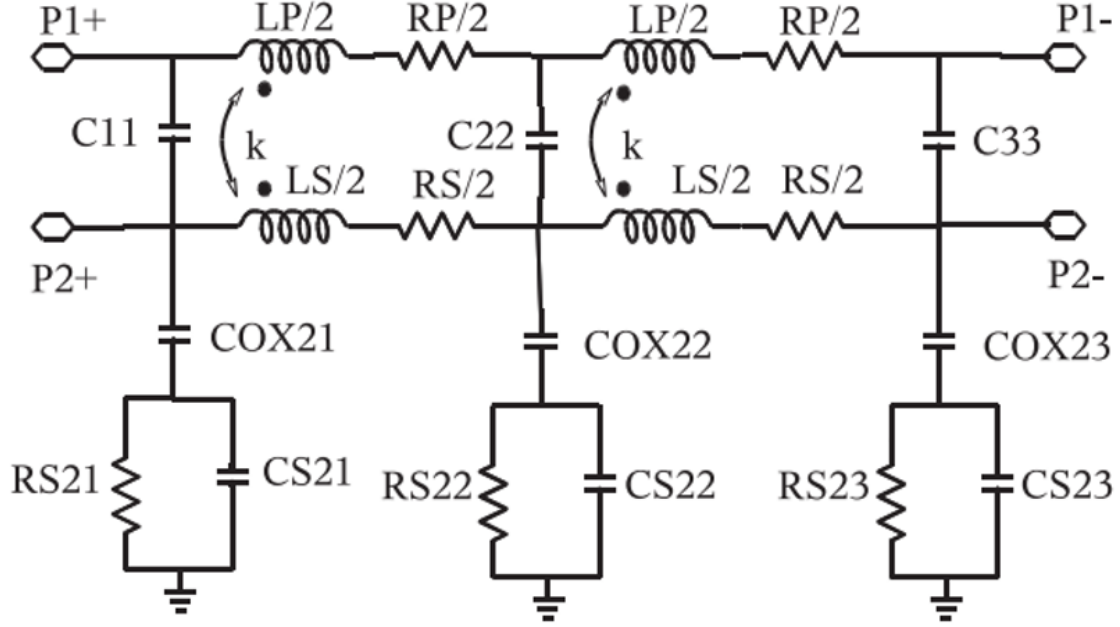


Figure 4.2: Schematic model of vertically coupled transformer with parasitics included

winding inductance L_p , the secondary winding inductance L_s , and the coupling factor k . The parasitic values are extracted to have the schematic model give a closer match to the EM simulation results.

4.4.1 Input and Output Matching Network

The design process of a typical matching network in this PA generally goes as follows. Firstly, the desired impedance that need to be presented by the matching network is determined. When an appropriate choice of winding inductance and coupling factor is found, an initial schematic design of a transformer-based matching network is done to match the desired impedance. The initial values of the input transformers need to conjugate match the inputs to the transistors, and the output transformers should present the optimal load impedance to the PA identified from the load-pull contours.

Then, layout is done to give the desired primary winding inductance, secondary winding inductance, and coupling factor according to these initial schematic values. Afterwards, the EM simulation results of this transformer is used to extract and verify the loss and coupling parasitics. Then, the schematic circuit is updated with the newly extracted parasitic values, and three main design parameters are adjusted to give the desired impedance from before. Using this model with better accuracy, an updated layout is done to match the behavior of the schematic transformer circuit model. This iteration is typically repeated and the layout is fine-tuned afterwards to achieve the desired circuit performance. This initial input and output matching network design iterations are fairly straight forward, as these are linear design processes to provide desirable impedance expanding from basic lumped component matching theory.

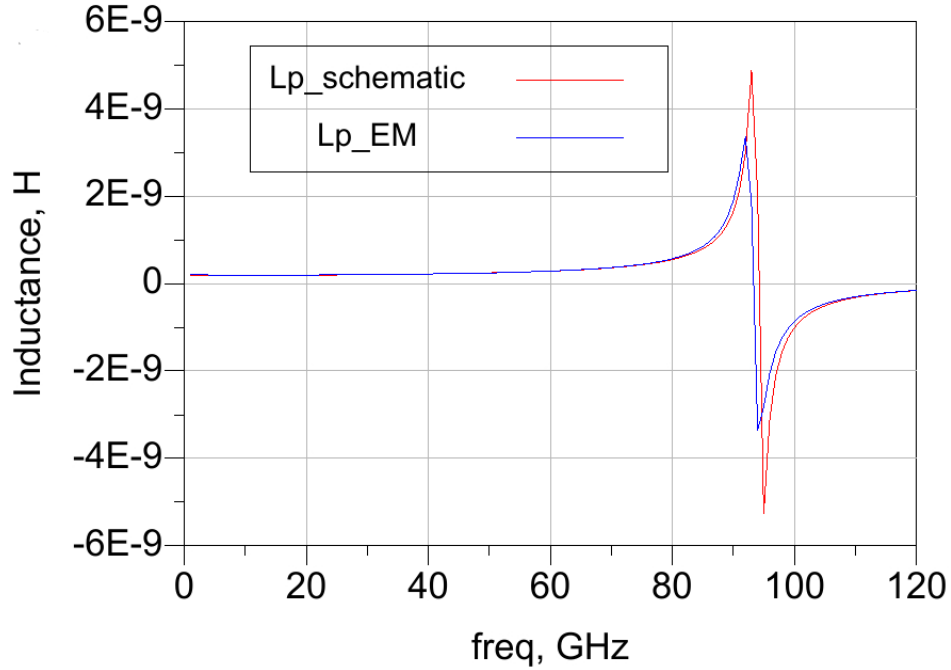


Figure 4.3: Primary winding inductance of schematic model vs layout EM

After the initial design iterations, these matching networks are swept and optimized to improve the 2-stage PA's performance. Both input and output matching affect the overall

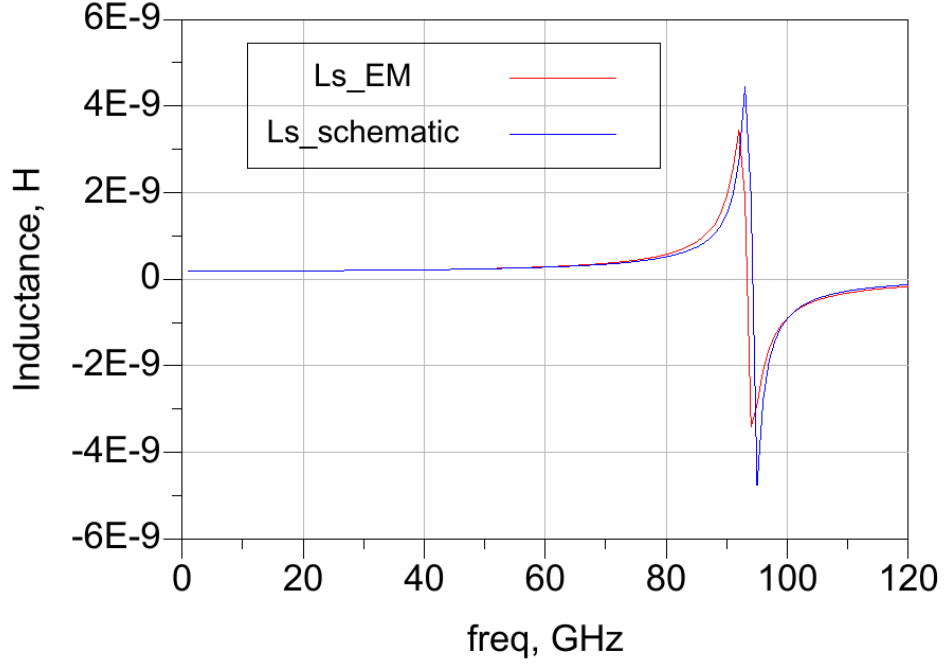


Figure 4.4: Secondary winding inductance of schematic model vs layout EM

2-stage linearity and can be tuned in small ranges to make some adjustments based on large signal simulation results. However, the changes are not significant, as that may cause significant deterioration in gain, output power, or efficiency.

As an example, the output matching network is designed and optimized in schematic, then the layout is done to match the results of the schematic model. Comparison of the inductances and coupling factor is shown in Figure 4.4, Figure 4.5 and Figure 4.6. Relative difference between schematic and layout EM is below 1 percent for all three parameters within the frequency of interest. The self-resonant frequency of the two inductors also see minimal differences between layout and schematic model. The final load impedance presented by the output matching networks to the two branches of PA is shown as follows in Figure ???. Referring to Figure 3.12, it is apparent that the load impedance presented to the final PA is around the optimal impedance shown in the load pull results. The input matching network is verified in the same way and will not have its results shown here.

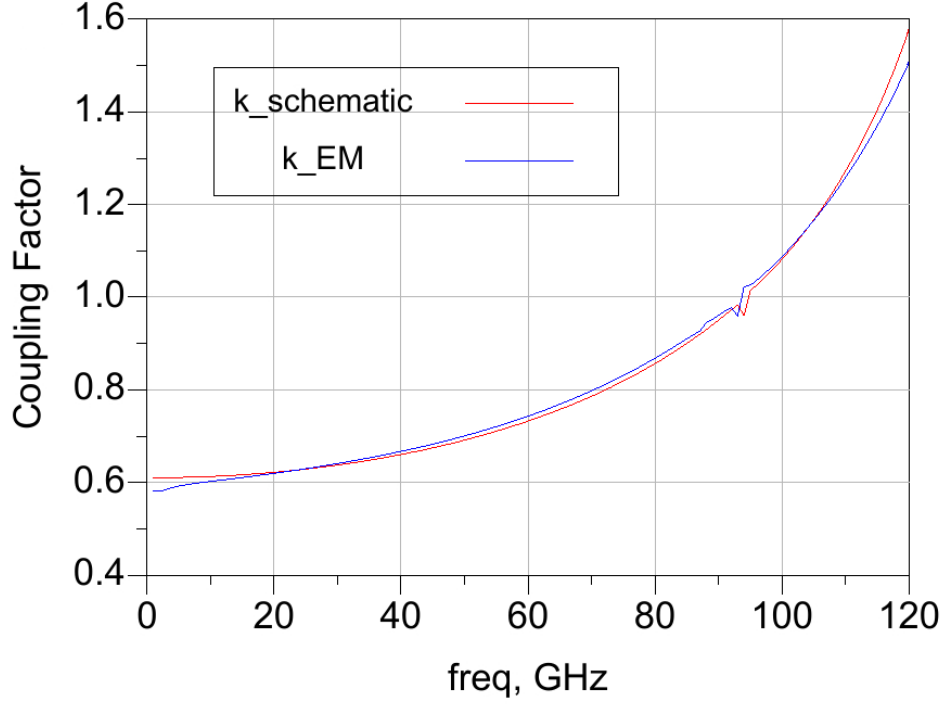


Figure 4.5: Coupling factor of schematic model vs layout EM

4.4.2 Inter-stage Load-Pull

The original method of interstage load-pull as discussed in [44] chooses the matching network parameters based on results at a certain input power level in large signal CW simulation. This results in clean and easily readable plots showcasing the performance. However, one disadvantage is that some key information is missing from these plots. Examples of these include how the non-linearity behaves and changes at other power levels, as well as how the small-signal gain behave across frequency.

One simple change is adopted in this design. By sweeping the key design parameters in the interstage transformer, the AM-AM, AM-PM and overall PAE are plotted against a range of input power level and frequencies. This gives a better picture of most, if not all, of the PA performance criteria rather than a limited number of them.

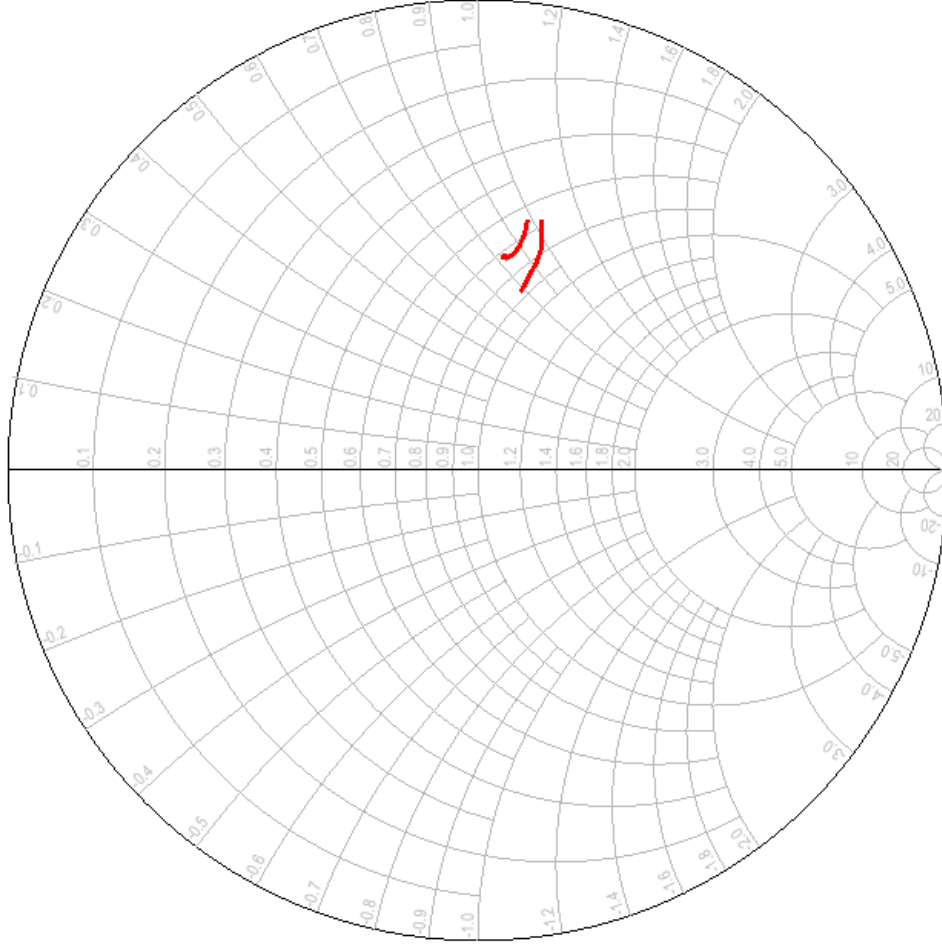


Figure 4.6: Load impedance presented by matching networks to the drains of two amplifiers over 37-43 GHz

The design process is described as follows. Firstly, an ideal inter-stage network is used by presenting ideal impedance to the driver stage output and PA stage input. Then, the input matching network to the driver and the output matching network to the PA is designed as described in section 4.3.1. Then, with a working input and output matching network for the 2-stage PA, a schematic transformer model including the parasitic values is added in the interstage to replace the ideal impedance blocks. Then, the interstage sweep is performed by sweeping the wideband PA large-signal CW performance against

the interstage against the transformer design parameters L_p , L_s , and k .

One issue with this method is the high number of resulting plots, as complete set of data needs to be plotted for every combination of potential design parameter values. Therefore plots of the inter-stage loadpull process is not included in the thesis itself. Rather, the effectiveness of this method is verified indirectly with the final results.

4.5 Transformer-based Hybrid 90-Degree Coupler Design

The theory of a transformer-based hybrid is described in [14]. With proper choice of capacitance, inductance and coupling factor, a single-ended quadrature hybrid can be designed. In [35], the layout of a compact differential quadrature hybrid coupler is proposed, based on the theory of the original single-ended hybrid. Normally, the structure of a differential hybrid would simply include two symmetric single-ended quadrature hybrids. However, one downside with this is its large chip area occupation. With double the area, the layout would not be compact. Instead, by orienting the two transformers in the way that the magnetic field generated by the current adds constructively, the design can become compact with the added benefits of enhanced magnetic coupling. The final layout proposed in this paper gives a symmetrical folded structure where the coils are adjacently coupled in the same chip area.

4.5.1 Output Single-Ended 90-Degree Coupler

The output hybrid is designed to be ultra-compact with a size of 80um by 80um shown in Figure 4.8 above. Simulation results are shown below, same port reference is used for both single-ended hybrid and differential hybrid, with port 1 as input, port 2 as through, port 3 as coupled and port 4 as the isolated port. EM simulation was performed with momentum Virtuoso, and the results of the S-parameter are included below. As shown in Figure 4.10, the hybrid achieves a phase difference of 90 degrees between the phase change from the input port to the through port and the phase change from the input port to the

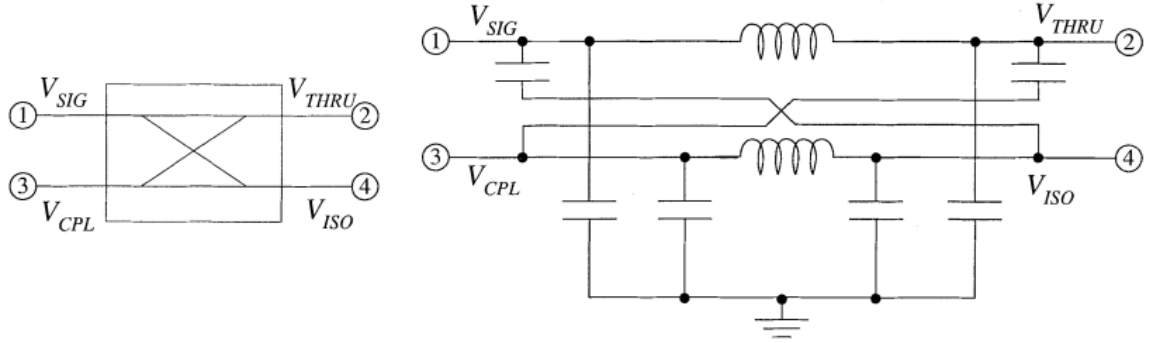


Figure 4.7: Schematic of the single-ended transformer based quadrature hybrid [14]

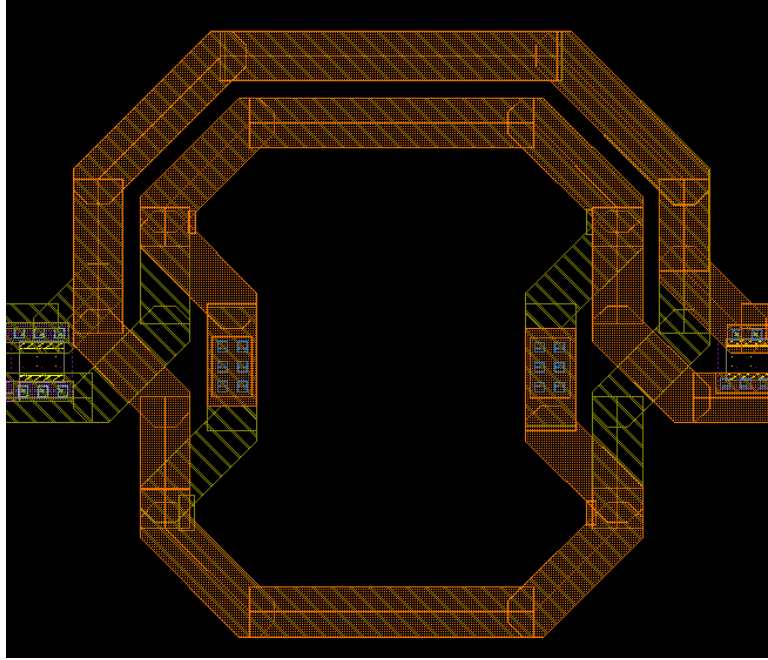


Figure 4.8: Layout of the single-ended transformer based quadrature hybrid design

coupled port, resulting in a phase error of less than 1 degree across the band. However, the amplitude mismatch is high outside of the operation frequency range, making the design narrow-band, as shown in Figure 4.9. All ports have return losses of more than 20dB, as shown in Figure 4.11, as the matching of these passive hybrids does not seem to be an

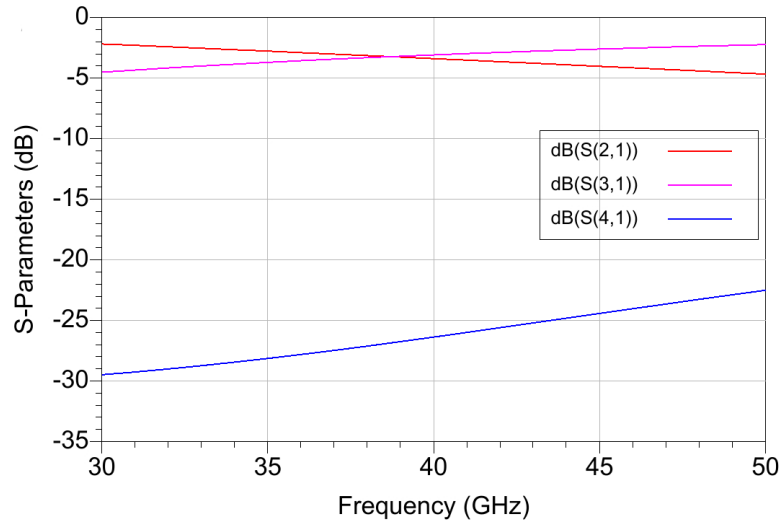


Figure 4.9: Insertion loss and isolation for the single-ended transformer based quadrature hybrid

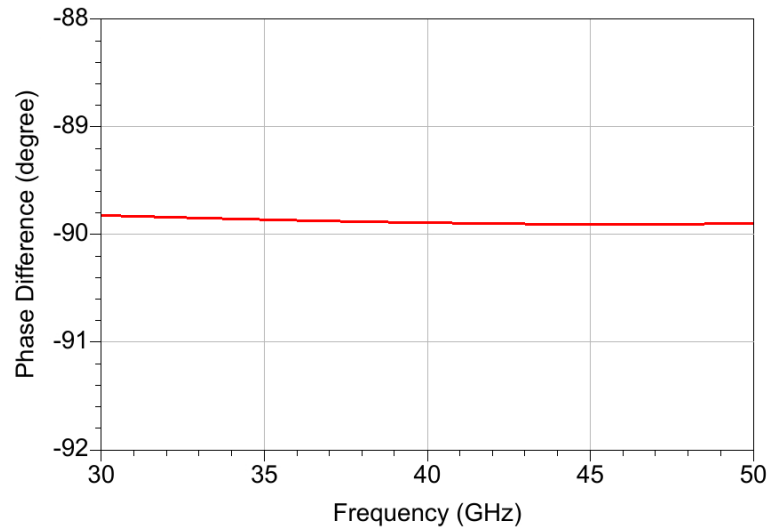


Figure 4.10: Relative phase difference between S21 and S31 for the single-ended transformer based quadrature hybrid

issue in this technology.

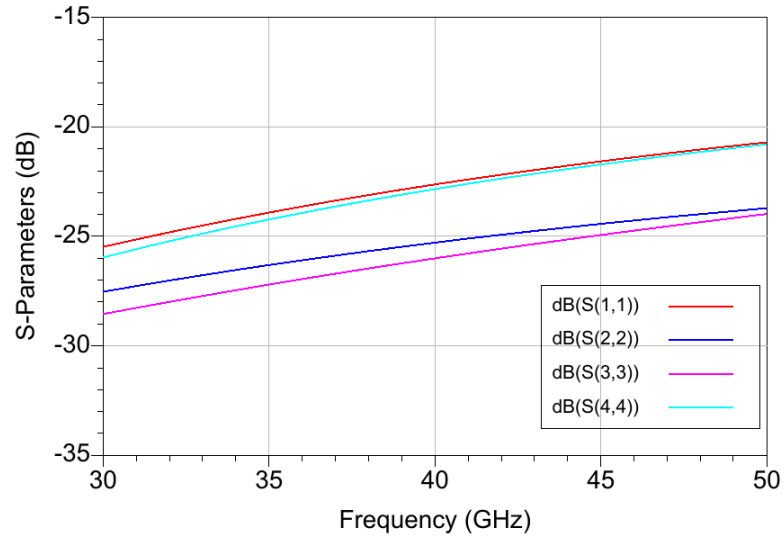


Figure 4.11: Matching of each port for the single-ended transformer based quadrature hybrid

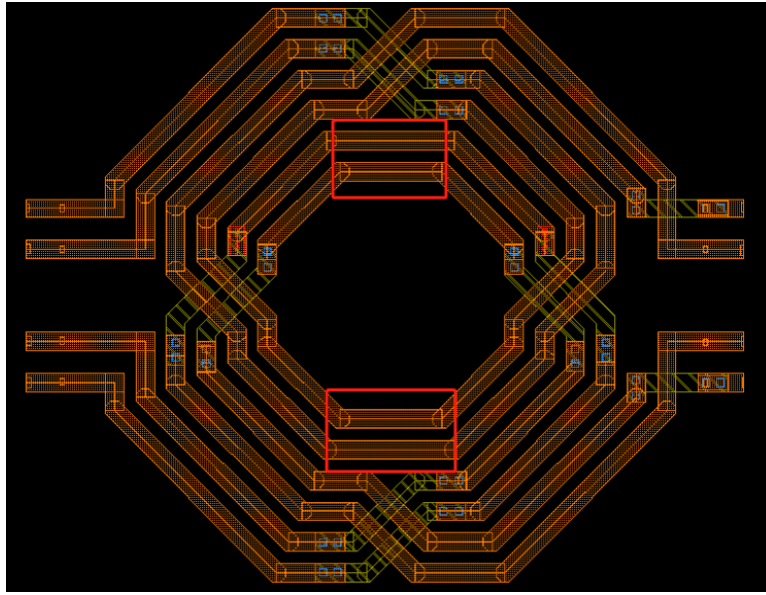


Figure 4.12: Original layout of the differential transformer based quadrature hybrid

4.5.2 Input Differential 90-Degree Coupler

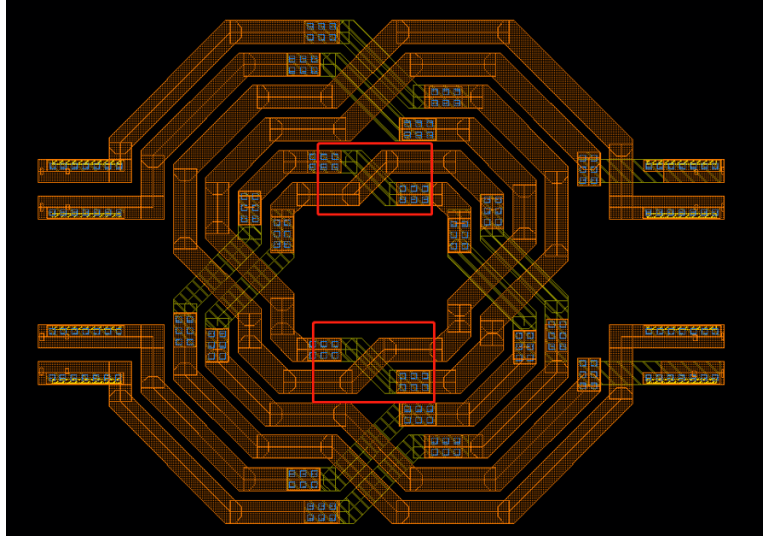


Figure 4.13: Updated layout of the differential transformer based quadrature hybrid

A 28GHz hybrid with structure shown in 4.12 has been designed before by past group members. However, upon updating the layout of this hybrid and attempting to design one at 39GHz, it was noticed that there was a significant phase error of over 5 degrees and was difficult to tune out. A change was made to the layout to improve its structural symmetry as seen in Figure 4.13. The two neighboring coils are changed to have a cross-over point in the center such that one side of the coil does not run around the other coil with a different coupling factor for the entire length of the routing. This improved structural symmetry significantly eased the design process and reduced the phase error. Observation was made here that the structural symmetry of the transformer-based quadrature hybrid is important and has significant affect on its performance as well as ease of design.

The simulation results of the final differential quadrature hybrid is shown as follows. The amplitude mismatch is minimal with 0.3dB or lower across a wide 35-50GHz frequency range, as shown in Figure 4.16. Figure 4.17 shows the relative phase difference between the through port and the coupled port, with the phase error kept below 0.7 degrees across the band of interest, between 37 and 43GHz. Return loss of each port is above 25dB as shown

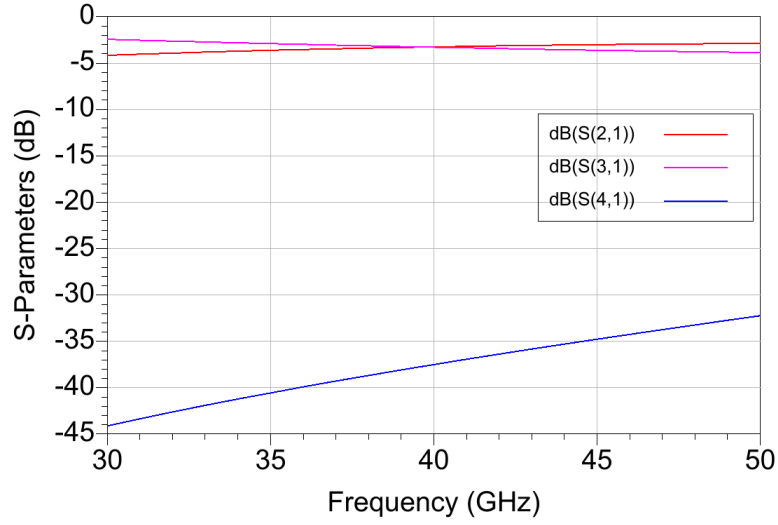


Figure 4.14: Insertion loss and isolation for the original differential transformer based quadrature hybrid

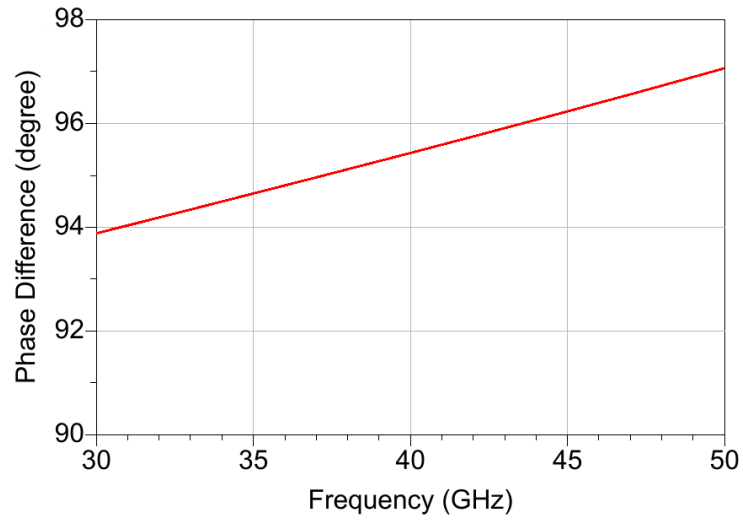


Figure 4.15: Relative phase difference between S21 and S31 for the original differential transformer based quadrature hybrid

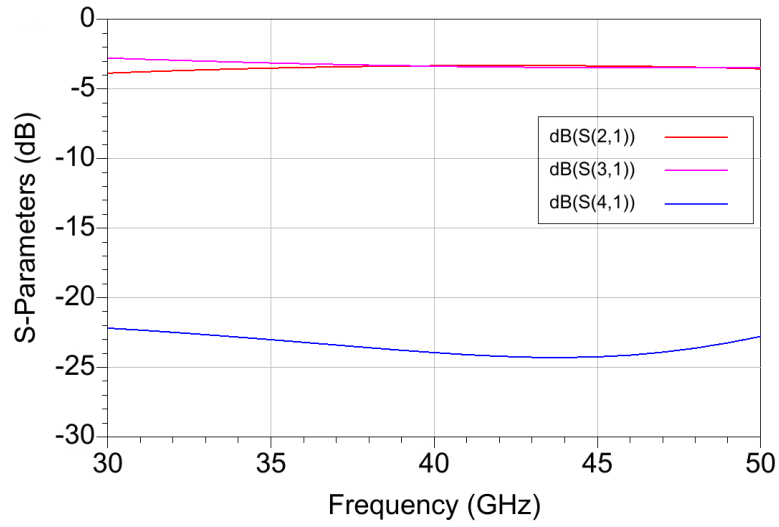


Figure 4.16: Insertion loss and isolation for the improved differential transformer based quadrature hybrid

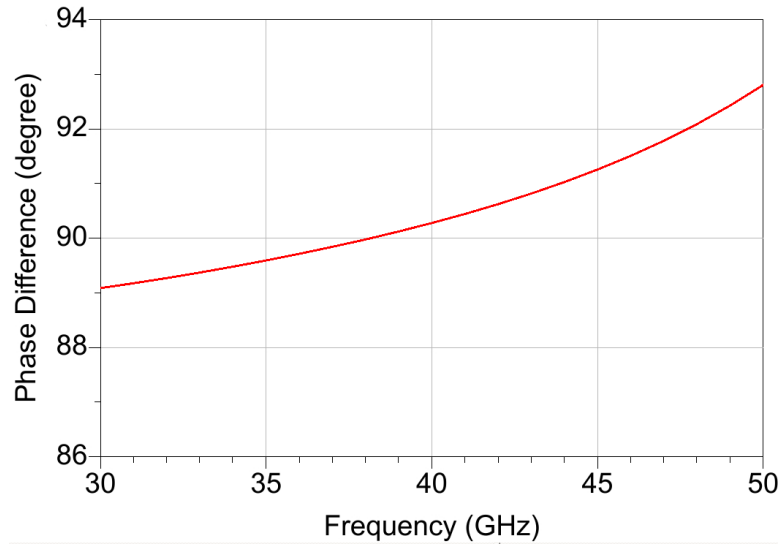


Figure 4.17: Relative phase difference between S21 and S31 for the improved differential transformer based quadrature hybrid

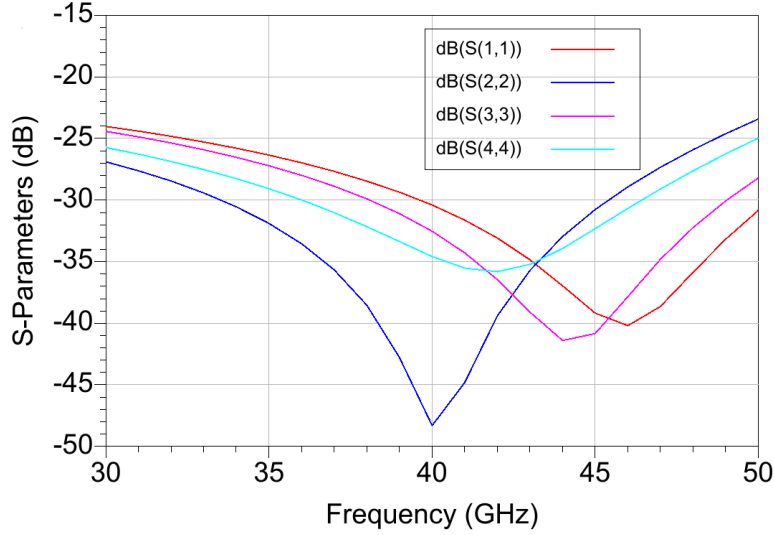


Figure 4.18: Matching of each port for the improved differential transformer based quadrature hybrid

in Figure 4.18. Results of a hybrid optimized using the original structure is shown in 4.14 and 4.15, the altered structure sees a reduction of phase error from 5 degrees down to 0 degrees in the frequency of interest, and the insertion loss of input to through port and input to coupled port sees an added inflection point which minimizes the gain imbalance in higher frequencies. Although the latter observation is out of band for this design and thus not of high importance in this case.

4.6 Full BPA Simulation Results

The final full-EM simulation results of the PA is shown in this section. All passive structures above layer C1 are simulated together in one single block. The transistor interconnects were performed in separate EM simulations and later in the final ADS test bench. The schematic of the final design is shown in Figure 4.19, while the bumped version of the layout is shown in Figure 4.20 for illustration purposes.

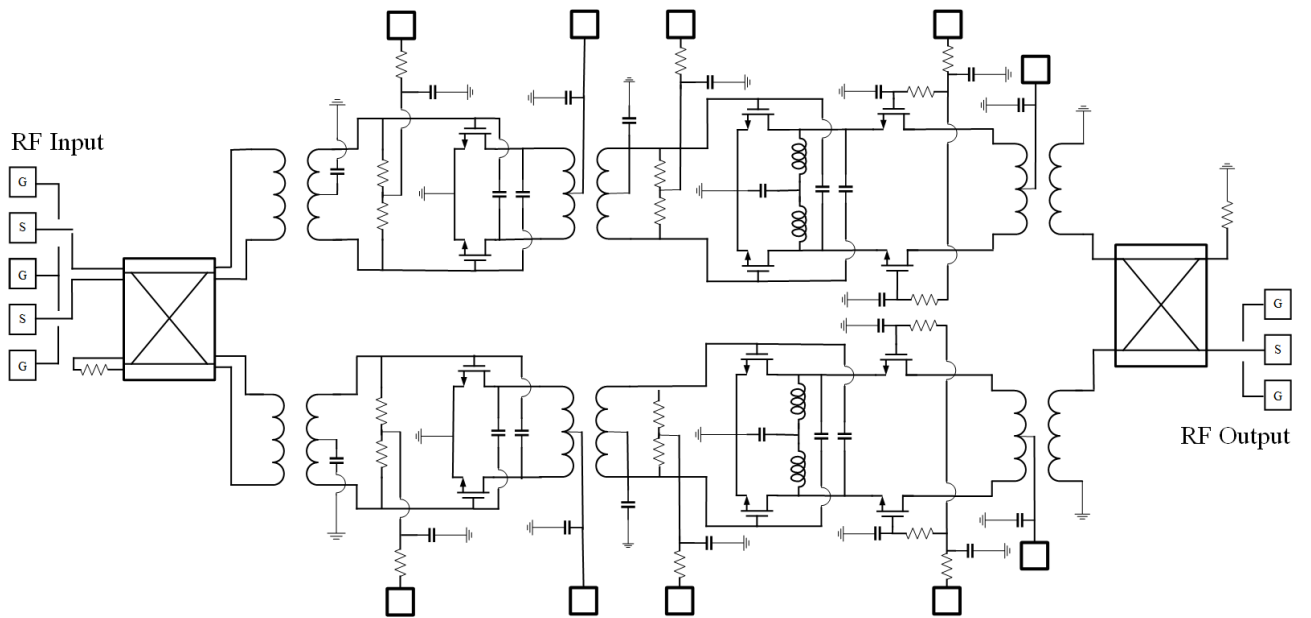


Figure 4.19: Final PA schematic

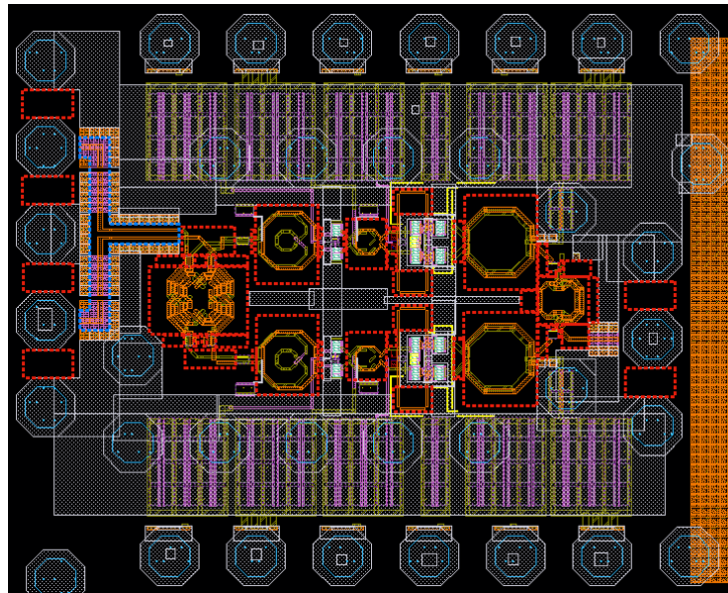


Figure 4.20: Bumped version of the PA layout

4.6.1 Small Signal S-Parameter Simulation

It is known that the balanced amplifier structure cancels the reflected power from the two identical PA branches at input and output given the matched source and load impedance. This is verified with the s-parameter simulation in 4.21. Input and output return loss are kept above 20dB across the relevant frequencies. Small signal gain variation is under 0.5dB per GHz, which meets the target specification.

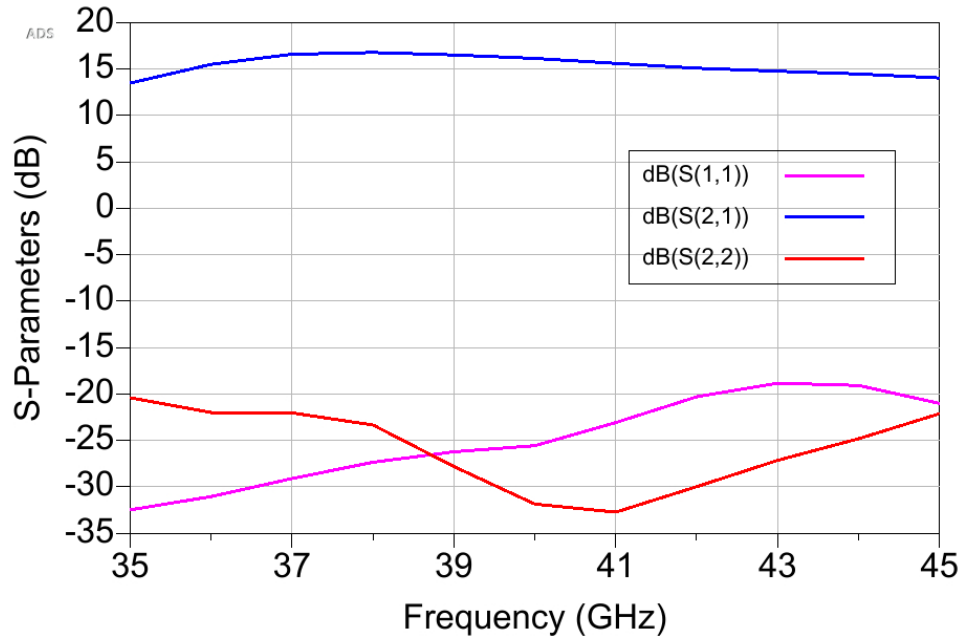


Figure 4.21: BPA Small signal s-parameter simulation results

4.6.2 Large Signal Performance under Continuous-Wave Simulation

Large-signal CW simulation is performed against input power level at discrete frequencies across 37 to 43 GHz with 1 GHz spacing. All three figures use the same color code for data at different frequencies, from red to dark green, with red representing the data at 37GHz and dark green corresponding to 43GHz.

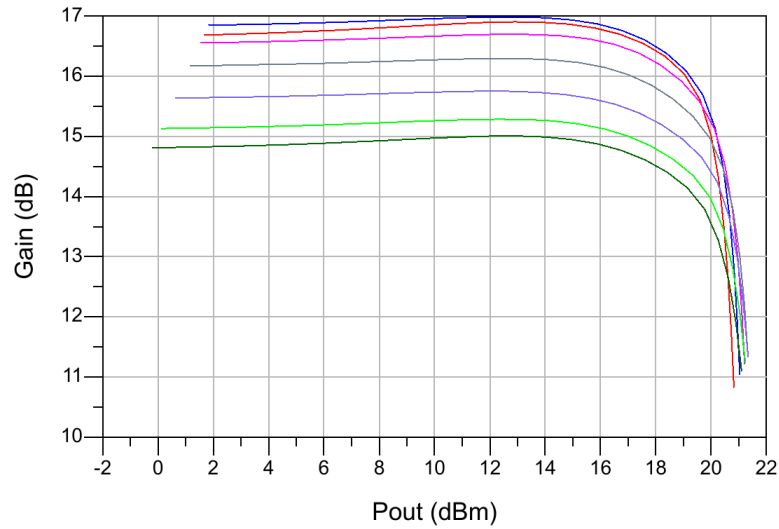


Figure 4.22: Gain vs output power in large signal continuous wave simulation

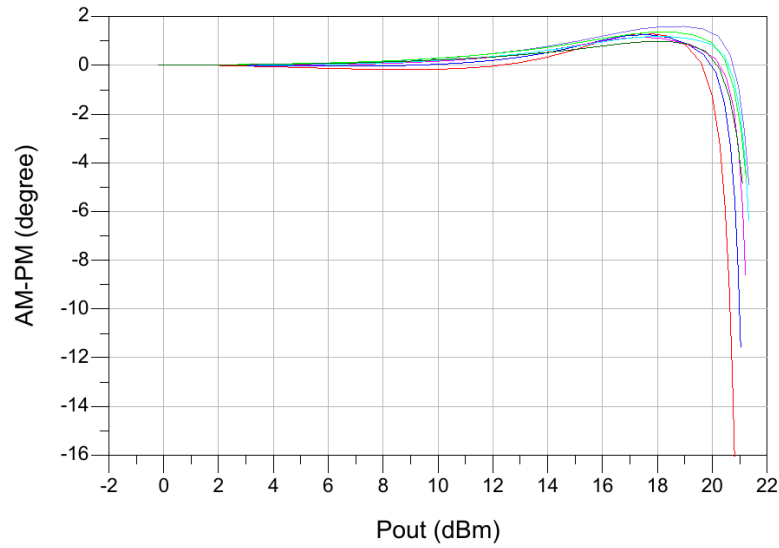


Figure 4.23: Amplitude-to-phase distortion in large signal continuous wave simulation

In Figure 4.22, the output 1dB compression point is between 19 and 19.5 dBm, and the output 3dB compression point is around 21 dBm in the frequency band. Phase distortion up to the output 1dB compression point is well maintained below approximately 1 degree

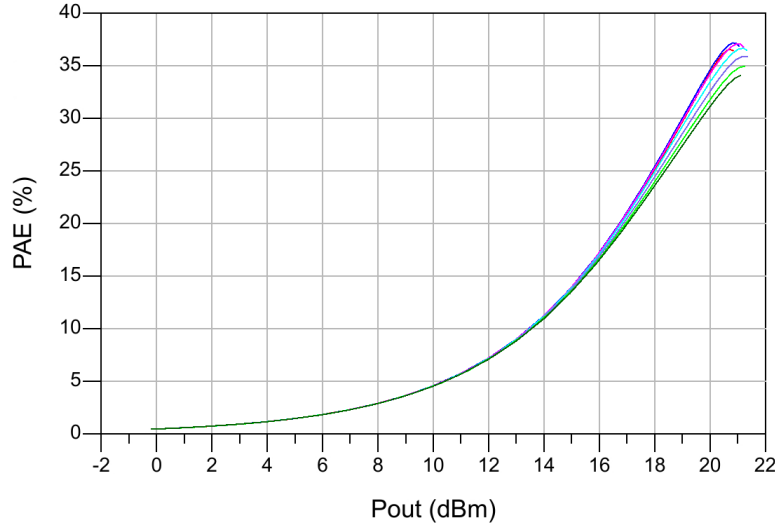


Figure 4.24: Overall power-added-efficiency vs output power in large signal continuous wave simulation

across the operating frequencies, as shown in Figure 4.23. The overall PAE is plotted in 4.24, with the peak PAE being 37% in the 37-40GHz band, followed by a slow drop in efficiency beyond 41GHz, although still above 33%.

4.6.3 Large Signal CW Simulation Over Load VSWR

PA large signal performance over 2:1 VSWR is verified in this section. Harmonic balance simulation is performed at 39GHz over 8 evenly spaced out impedance points over the 2:1 VSWR circle. The results of the large signal CW results are shown in Figure 4.25, Figure 4.26, and Figure 4.27. As seen in the figures, the gain of this BPA shows under 0.2dB variation across the different phases on the 2:1 VSWR circle, and the absolute degradation from the reference matched load scenario is only 0.5dB. Consistency in linearity is maintained as well, showing within 0.5 dB of variation and less than 1 dB of degradation across the simulated impedances. The variation of AM-PM distortion is generally more challenging to minimize; however, it is still kept within 2 degrees. Overall PAE of this PA is maintained above 30%.

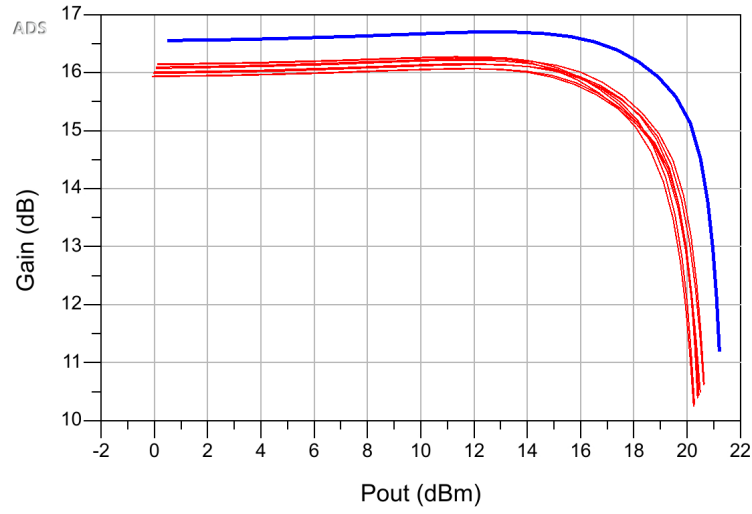


Figure 4.25: Gain vs output power at 39GHz with matched load (in blue) and over 2:1 VSWR (in red) in large signal continuous wave simulation

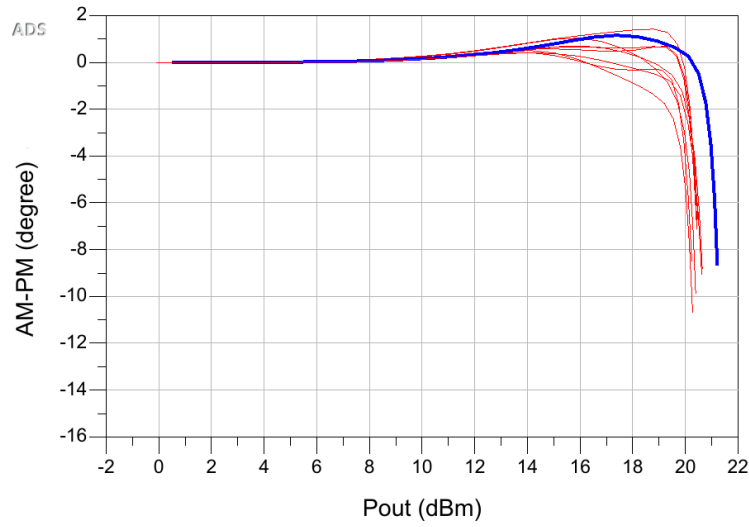


Figure 4.26: Amplitude-to-phase distortion at 39GHz with matched load (in blue) and over 2:1 VSWR (in red) in large signal continuous wave simulation

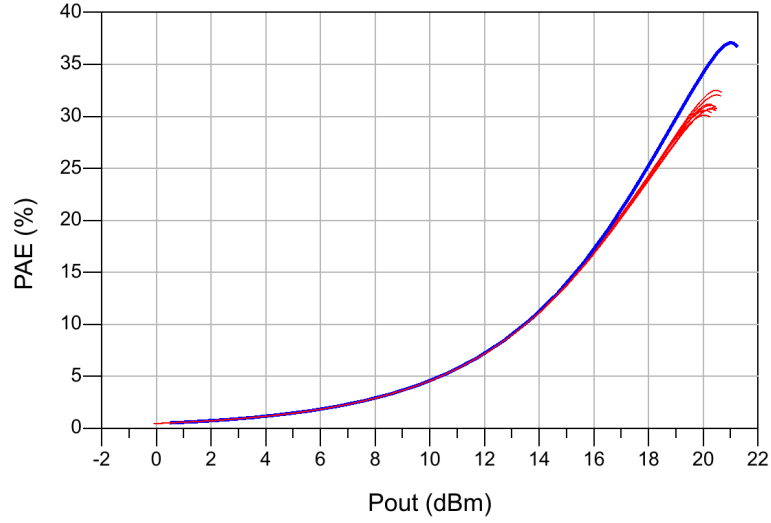


Figure 4.27: Overall power-added-efficiency vs output power at 39GHz with matched load (in blue) and over 2:1 VSWR (in red) in large signal continuous wave simulation

Overall, this balanced PA design has shown competitive performance when compared to the state-of-the-art class-AB PAs in section 2.5. Furthermore, the PA maintains good linearity under load mismatch, as demonstrated in the large-signal simulation in this section.

Chapter 5

Conclusion, Reflection and Future Work

5.1 Conclusion

This thesis has showcased the advantages of balanced power amplifier topology in highly integrated millimeter wave phased array systems with a design done using cutting-edge SOI CMOS technology. Chapter 2 provided an overview of state-of-the-art PA in the literature and the approaches that are taken to resolve various challenges in millimeter-wave PA design. Chapter 3 started off by discussing various common transistor layout topologies and their application to PA circuits. Then stability in PA is discussed, including various methods to stabilize an amplifier. Lastly, this chapter concludes with discussion on transistor choices for this specific design work. Chapter 4 discusses the general methodology and the main considerations in the design process, as well as presenting the final EM simulation results of the PA.

A 37-43 GHz balanced power amplifier design was demonstrated in this thesis. By making use of the 2-stacked FET in the main PA stage, the amplifier achieves state-of-the-art large signal CW performance up to a high saturated output power of over 21 dBm, while occupying a core area of only 0.26 square millimeters. With linearity as a strong

design focus, the 2-stage driver and PA co-design methodology yields a resulting AM-PM of 1 degree up to OP1dB of over 19 dBm over the 37-43 GHz targeted frequency band. Efficiency is also a key design target. By operating in deep class AB, the PA achieves a power-added-efficiency of 37%. These numbers are competitive with the performance seen in state-of-the-art linearity-enhanced class-AB PA designs in the literature. The balanced power amplifier is implemented with ultra-compact transformer based quadrature hybrid couplers while also minimizing their impact on PA performance. The advantages of adopting the balanced amplifier topology are also clearly demonstrated in the results. The improved return loss of the PA as well as the resiliency against mismatched load impedance are both highly desirable features for PA to be used in phased array systems. The PA is matched at input and output with over 20dB return loss. This level of S22 is typically impossible to achieve with other PA architectures aiming for state-of-the-art efficiency and output power level. In addition, the variation in terms of PA gain, output power and AM-PM are kept below 0.5 dB, 1 dB and 1 degree respectively, thus expected to lead to significant improvement in terms of PA performance against different phased array steering angles. Overall, this design is expected to prove its value in a phased array system and there is no doubt that this topology will see further uses in other practical phased array systems.

5.2 Future Works

One problem that remains unresolved in this design is the lack of improvement to the PA's power efficiency. As 5G systems are expected to operate with signals that have high PAPR, it is highly desirable to improve the back-off efficiency of the PA in order to reduce overall power consumption. One potential direction of improvement could be to adopt a balanced-Doherty topology, with each of the two branches of the balanced amplifier being a complete Doherty PA. This approach would likely gain the merits of balanced amplifier topology as well as having the improved back-off efficiency of the Doherty.

Another potential functionality that could be implemented is reconfigurability. Due to the various millimeter wave frequency bands available for 5G applications in different

countries and regions, it is highly desirable to have one PA that could work for multiple frequency bands. This introduces the need for wide-band PA, dual-band PA, and frequency re-configurable PAs. Secondly, re-configurable linearity could also be an interesting topic for investigation, as demonstrated to be feasible by some existing publications in the literature. This could be implemented in the form of re-configurable linearizers or re-configurable matching networks.

References

- [1] An essential of the wi-fi tri-band system. <https://www.qorvo.com/design-hub/blog/wifi-5-point-2-ghz-rf-filters>.
- [2] Mohammed Abdulaziz, Halil Volkan Hünerli, Koen Buisman, and Christian Fager. Improvement of am–pm in a 33-ghz cmos soi power amplifier using pmos neutralization. *IEEE Microwave and Wireless Components Letters*, 29(12):798–801, 2019.
- [3] Sheikh Nijam Ali, Pawan Agarwal, Srinivasan Gopal, and Deukhyoun Heo. Transformer-based predistortion linearizer for high linearity and high modulation efficiency in mm-wave 5g cmos power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 67(7):3074–3087, 2019.
- [4] Kyu Hwan An, Ockgoo Lee, Hyungwook Kim, Dong Ho Lee, Jeonghu Han, Ki Seok Yang, Younsuk Kim, Jae Joon Chang, Wangmyong Woo, Chang-Ho Lee, Haksun Kim, and Joy Laskar. Power-combining transformer techniques for fully-integrated cmos power amplifiers. *IEEE Journal of Solid-State Circuits*, 43(5):1064–1075, 2008.
- [5] Roger Arguez-Ramirez, Jose-Ramon Perez-Cisneros, and Christian Fager. Investigation of power amplifier performance under load mismatch conditions. In *2021 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pages 41–43, 2021.
- [6] Alireza Asoodeh, Hossein Miri Lavasani, Mengye Cai, and Shahriar Mirabbasi. A highly linear and efficient 28-ghz pa with a psat of 23.2 dbm, p1 db of 22.7 dbm, and

- pae of 35.5% in 65-nm bulk cmos. *IEEE Journal of Solid-State Circuits*, 56(3):914–927, 2021.
- [7] Chandrakanth Reddy Chappidi, Tushar Sharma, and Kaushik Sengupta. Multi-port active load pulling for mm-wave 5g power amplifiers: Bandwidth, back-off efficiency, and vswr tolerance. *IEEE Transactions on Microwave Theory and Techniques*, 68(7):2998–3016, 2020.
 - [8] Chandrakanth Reddy Chappidi, Xue Wu, and Kaushik Sengupta. Simultaneously broadband and back-off efficient mm-wave pas: A multi-port network synthesis approach. *IEEE Journal of Solid-State Circuits*, 53(9):2543–2559, 2018.
 - [9] Shichang Chen, Gaofeng Wang, Zhiqun Cheng, Pei Qin, and Quan Xue. Adaptively biased 60-ghz doherty power amplifier in 65-nm cmos. *IEEE Microwave and Wireless Components Letters*, 27(3):296–298, 2017.
 - [10] Yu-Chun Chen, Tsung-Ching Tsai, Jeng-Han Tsai, and Tian-Wei Huang. A 38-ghz-band power amplifier with analog pre-distortion for 1600-mhz transmission bandwidth 64-qam ofdm modulated signal. In *2019 IEEE MTT-S International Microwave Symposium (IMS)*, pages 312–315, 2019.
 - [11] Bolun Cui and John R. Long. A 1.7-db minimum nf, 22–32-ghz low-noise feedback amplifier with multistage noise matching in 22-nm fd-soi cmos. *IEEE Journal of Solid-State Circuits*, 55(5):1239–1248, 2020.
 - [12] Hayg-Taniel Dabag, Bassel Hanafi, Fatih Golcuk, Amir Agah, James F. Buckwalter, and Peter M. Asbeck. Analysis and design of stacked-fet millimeter-wave power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 61(4):1543–1556, 2013.
 - [13] Zhiming Deng and Ali M. Niknejad. A layout-based optimal neutralization technique for mm-wave differential amplifiers. In *2010 IEEE Radio Frequency Integrated Circuits Symposium*, pages 355–358, 2010.
 - [14] R.C. Frye, S. Kapur, and R.C. Melville. A 2-ghz quadrature hybrid implemented in cmos technology. *IEEE Journal of Solid-State Circuits*, 38(3):550–555, 2003.

- [15] Edgar Felipe Garay, David Joseph Munzer, and Hua Wang. 26.3 a mm-wave power amplifier for 5g communication using a dual-drive topology exhibiting a maximum pae of 50% and maximum de of 60% at 30ghz. In *2021 IEEE International Solid-State Circuits Conference (ISSCC)*, volume 64, pages 258–260, 2021.
- [16] Song Hu, Fei Wang, and Hua Wang. A 28-/37-/39-ghz linear doherty power amplifier in silicon for 5g applications. *IEEE Journal of Solid-State Circuits*, 54(6):1586–1599, 2019.
- [17] Tzu-Yuan Huang, Naga Sasikanth Mannem, Sensen Li, Doohwan Jung, Min-Yu Huang, and Hua Wang. 26.1 a 26-to-60ghz continuous coupler-doherty linear power amplifier for over-an-octave back-off efficiency enhancement. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, volume 64, pages 354–356, 2021.
- [18] Ercan Kaymaksut, Dixian Zhao, and Patrick Reynaert. Transformer-based doherty power amplifiers for mm-wave applications in 40-nm cmos. *IEEE Transactions on Microwave Theory and Techniques*, 63(4):1186–1192, 2015.
- [19] Shailesh Kulkarni and Patrick Reynaert. A 60-ghz power amplifier with am–pm distortion cancellation in 40-nm cmos. *IEEE Transactions on Microwave Theory and Techniques*, 64(7):2284–2291, 2016.
- [20] Jaehun Lee, Ji-Seon Paek, and Songcheol Hong. Frequency reconfigurable dual-band cmos power amplifier for millimeter-wave 5g communications. In *2021 IEEE MTT-S International Microwave Symposium (IMS)*, pages 846–849, 2021.
- [21] Sensen Li, Taiyun Chi, Jong-Seok Park, Huy Thong Nguyen, and Hua Wang. A 28-ghz flip-chip packaged chireix transmitter with on-antenna outphasing active load modulation. *IEEE Journal of Solid-State Circuits*, 54(5):1243–1253, 2019.
- [22] Sensen Li, Min-Yu Huang, Doohwan Jung, Tzu-Yuan Huang, and Hua Wang. 24.3 a 28ghz current-mode inverse-outphasing transmitter achieving 40%/31% pa efficiency at psat/6db pbo and supporting 15gbit/s 64-qam for 5g communication. In *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 366–368, 2020.

- [23] Siwei Li and Gabriel M. Rebeiz. High efficiency d-band multiway power combined amplifiers with 17.5–19-dbm psat and 14.2–12.1% peak pae in 45-nm cmos rfsoi. *IEEE Journal of Solid-State Circuits*, 57(5):1332–1343, 2022.
- [24] Tso-Wei Li, Sensen Li, Hossein Miri Lavasani, and Hua Wang. A v-band doubly hybrid nmos/pmos four-way distributed-active-transformer power amplifier for non-linearity cancellation and joint linearity/efficiency optimization. In *2021 IEEE MTT-S International Microwave Symposium (IMS)*, pages 382–385, 2021.
- [25] Xingcun Li, Wenhua Chen, Shuyang Li, Yunfan Wang, Fei Huang, Xiang Yi, Ruonan Han, and Zhenghe Feng. A high-efficiency 142–182-ghz sige bicmos power amplifier with broadband slotline-based power combining technique. *IEEE Journal of Solid-State Circuits*, 57(2):371–384, 2022.
- [26] ChuanKang Liang and Behzad Razavi. A layout technique for millimeter-wave pa transistors. In *2011 IEEE Radio Frequency Integrated Circuits Symposium*, pages 1–4, 2011.
- [27] Guangyi Liu and Dajie Jiang. 5g: Vision and requirements for mobile communication system towards year 2020. *Chinese Journal of Engineering*, 2016:1–8, 01 2016.
- [28] Naga Sasikanth Mannem, Min-Yu Huang, Tzu-Yuan Huang, and Hua Wang. A re-configurable hybrid series/parallel doherty power amplifier with antenna vswr resilient performance for mimo arrays. *IEEE Journal of Solid-State Circuits*, 55(12):3335–3348, 2020.
- [29] Huy Thong Nguyen, Taiyun Chi, Sensen Li, and Hua Wang. A linear high-efficiency millimeter-wave cmos doherty radiator leveraging multi-feed on-antenna active load modulation. *IEEE Journal of Solid-State Circuits*, 53(12):3587–3598, 2018.
- [30] Huy Thong Nguyen and Hua Wang. A coupler-based differential mm-wave doherty power amplifier with impedance inverting and scaling baluns. *IEEE Journal of Solid-State Circuits*, 55(5):1212–1223, 2020.

- [31] Kang Ning, Yihao Fang, Navid Hosseinzadeh, and James F. Buckwalter. A 30-ghz cmos soi outphasing power amplifier with current mode combining for high backoff efficiency and constant envelope operation. *IEEE Journal of Solid-State Circuits*, 55(5):1411–1421, 2020.
- [32] Dimitrios Ntaikos, Bobby Gizas, Georgios Papageorgiou, and Constantinos Papadias. Channel model simulator for multi-antenna terrestrial links. pages 699–704, 05 2017.
- [33] Byungjoon Park, Sangsu Jin, Daechul Jeong, Jooseung Kim, Yunsung Cho, Kyunghoon Moon, and Bumman Kim. Highly linear mm-wave cmos power amplifier. *IEEE Transactions on Microwave Theory and Techniques*, 64(12):4535–4544, 2016.
- [34] Jinseok Park, Seunghoon Kang, and Songcheol Hong. Design of a ka-band cascode power amplifier linearized with cold-fet interstage matching network. *IEEE Transactions on Microwave Theory and Techniques*, 69(2):1429–1438, 2021.
- [35] Jong Seok Park, Shouhei Kousai, and Hua Wang. A fully differential ultra-compact broadband transformer based quadrature generation scheme. In *Proceedings of the IEEE 2013 Custom Integrated Circuits Conference*, pages 1–4, 2013.
- [36] Masoud Pashaeifar, Leo C. N. de Vreede, and Morteza S. Alavi. 14.4 a 24-to-30ghz double-quadrature direct-upconversion transmitter with mutual-coupling-resilient series-doherty balanced pa for 5g mimo arrays. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, volume 64, pages 223–225, 2021.
- [37] Sataporn Pornpromlikit, Hayg-Taniel Dabag, Bassel Hanafi, Joohwa Kim, Lawrence E. Larson, James F. Buckwalter, and Peter M. Asbeck. A q-band amplifier implemented with stacked 45-nm cmos fets. In *2011 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, pages 1–4, 2011.
- [38] Valdrin Qunaj and Patrick Reynaert. A ka-band doherty-like lmbs for high-speed wireless communication in 28-nm cmos. *IEEE Journal of Solid-State Circuits*, 56(12):3694–3703, 2021.

- [39] Bagher Rabet and James Buckwalter. A high-efficiency 28ghz outphasing pa with 23dbm output power using a triaxial balun combiner. In *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, pages 174–176, 2018.
- [40] Van-Son Trinh and Jung-Dong Park. An 85-ghz power amplifier utilizing a transformer-based power combiner operating beyond the self-resonance frequency. *IEEE Journal of Solid-State Circuits*, 57(3):882–891, 2022.
- [41] Marco Vigilante and Patrick Reynaert. A wideband class-ab power amplifier with 29–57-ghz am–pm compensation in 0.9-v 28-nm bulk cmos. *IEEE Journal of Solid-State Circuits*, 53(5):1288–1301, 2018.
- [42] Shih-Min Weng, Yi-Chun Lee, Tse-Hung Chen, and Jenny Yi-Chun Liu. A 60-ghz adaptively biased power amplifier with predistortion linearizer in 90-nm cmos. In *2018 IEEE/MTT-S International Microwave Symposium - IMS*, pages 651–654, 2018.
- [43] Tianzuo Xi, Sherry Huang, Shita Guo, Ping Gui, Daquan Huang, and Sudipto Chakraborty. High-efficiency e-band power amplifiers and transmitter using gate capacitance linearization in a 65-nm cmos process. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 64(3):234–238, 2017.
- [44] Jingjing Xia, Xiaohu Fang, Ahmed Ben Ayed, and Slim Boumaiza. Millimeter wave soi-cmos power amplifier with enhanced am-pm characteristic. *IEEE Access*, 8:8861–8875, 2020.
- [45] Yang Zhang and Patrick Reynaert. A high-efficiency linear power amplifier for 28ghz mobile communications in 40nm cmos. In *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pages 33–36, 2017.
- [46] Dixian Zhao, Shailesh Kulkarni, and Patrick Reynaert. A 60-ghz outphasing transmitter in 40-nm cmos. *IEEE Journal of Solid-State Circuits*, 47(12):3172–3183, 2012.
- [47] Dixian Zhao and Patrick Reynaert. A 60-ghz dual-mode class ab power amplifier in 40-nm cmos. *IEEE Journal of Solid-State Circuits*, 48(10):2323–2337, 2013.

- [48] Zhiwei Zong, Xinyan Tang, Khaled Khalaf, Dongyang Yan, Giovanni Mangraviti, Johan Nguyen, Yao Liu, and Piet Wambacq. A 28-ghz soi-cmos doherty power amplifier with a compact transformer-based output combiner. *IEEE Transactions on Microwave Theory and Techniques*, 69(6):2795–2808, 2021.