Mobile RF Front End Integration

James P. Young
Senior Member, IEEE, Skyworks Solutions Inc., Cedar Rapids, IA 52411, USA,
james.young@skyworksinc.com, 319-743-4033

Key words – RF integration, SOC, SIP, Power amplifier, RF switch, handset

Abstract – The mobile phone has become a big part of our daily lives, which has created an exponential growth in network data and an ever-increasing complex RF Front End to support this growth. This in turn has demanded more complex requirements for the filters, switches and power amplifiers (PA) in the front end. In order to address this increase in complexity, there has been a migration from discrete solutions to highly integrated solutions, but this increase in complexity and integration creates significant design challenges. Plus, there are two integrated implementations competing for market share: a system on chip (SOC) and system in package (SIP) solution. This presentation will overview the state-of-the-art in front end module design and integration, comparing SOC and SIP implementations, offering high performance, small and low cost solutions to the identified problems.

I. INTRODUCTION

The mobile RF Front End design has significantly increased in complexity over the past 3 years due to the ever increasing demand for speed and data by the consumer. The growth of mobile handset smart phone data has been exponential as shown in Fig. 1.[1] This has led to the expansion of new standards and bands which is shown in Fig. 2. While 4G has almost tripled the mobile handset data rate, this would be consumed in about 3 years without the addition of new bands to further increase the data capacity of the network. Fig. 3 demonstrates the band growth by showing the typical number of bands supported in a high end smart phone over time. Then, in addition to the band growth, carrier aggregation allows simultaneous LTE channels to be used across multiple bands to further increase the download speeds. Fig. 2 notes the increasing number of carrier aggregated bands over time. All of these upgrades add to the complexity of the mobile handset RF Front End, and challenges the RF Front End designer to support the increased complexity while maintaining a small footprint, acceptable battery life and low cost.

A Front End block diagram which can support thirteen to twenty-four 4G bands and carrier aggregation is shown Fig. 4. This paper will discuss various technologies that can be used to implement the three major block diagram functions of Fig. 4. Section II will discuss the RF switch, which includes both the antenna and band switch. Section III will discuss the technology options for the filters, which include band pass filters, duplexers and diplexers. Section IV will discuss the 3/4G power amplifier. Finally, section V will summarize the best technology of choice for each block, and based on performance, size and cost, will conclude whether the SOC or SIP solution is best for smart phone applications.

II. RF Switch

When considering the RF switch function in the block diagram, the key performance parameters to optimize are insertion loss, isolation, linearity, cost and size. When designing the switch, the first design parameter to address is the standing wave RF voltage across the switch at maximum Pout. When transmitting a low band GSM signal, the max. power level at the antenna is typically specified at 33dBm. The operational voltage across the switch is determined with an antenna VSWR of up to 2.5:1, and a no damage condition of up to a 10:1 VSWR. The 10:1 VSWR voltage has a 70.5V peak RF voltage. A single active MOS transistor typically cannot handle these operating conditions without either generating harmonic levels, which are unacceptable [2], or being damaged.

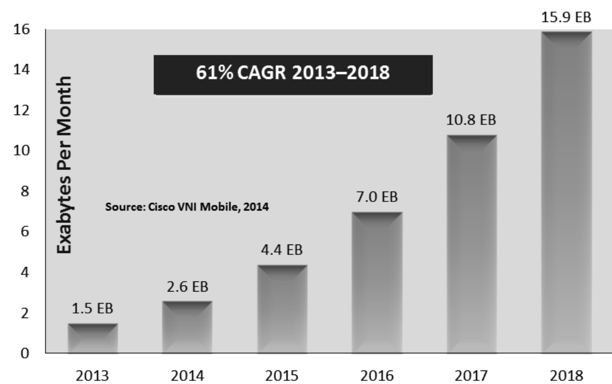

Fig. 1. Mobile Data Traffic

	2012	2014	2016E	2020E
LTE Rel	LTE Rel-11	LTE Rel-12	LTE Rel-X	5G
# CA bands	2	3	3	5
MIMO	8x8	8x8	8x8	64X8
Proposed New Bands		5+	24	50
CA band combos	25	75+	172	300
Peak/Max DL	1.2Gbps	3Gbps	6Gbps	18Gbps

Fig. 2. Mobile Standards, Bands & Data Rate

In order to overcome these limitations, typically each switch is composed of a stack of multiple devices in series, which splits the voltage across multiple FETs. For example, in a 0.18u SOI process a stack of between 8 to 12 devices will typically be used for the switch. In contrast, a dual gate PHEMT device has a larger breakdown voltage, so typically only 3 devices are required in the stack.

Once the proper stack is determined, the width of each FET in the stack is set based on the desired off capacitance which determines the off RF isolation. An off isolation between 25 and 30dB is required to assure off arms do not disturb the performance of the on arm. This design process determines the stack and the maximum FET width, which provides compliance to reliability and isolation.

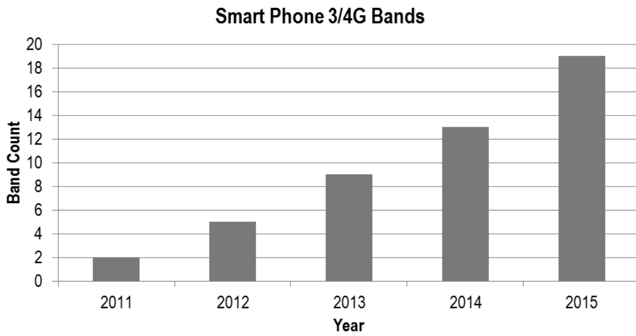

Fig. 3. Number of 3/4G Bands in a Smartphone

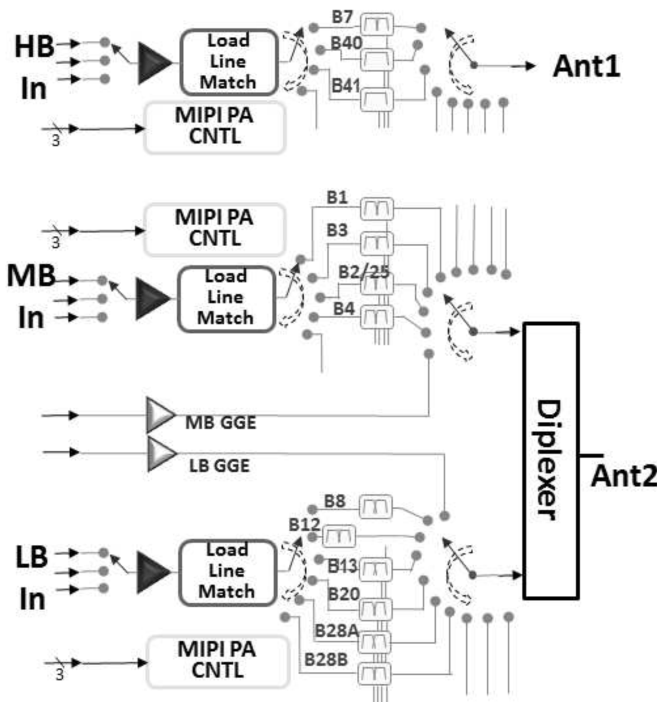

Fig. 4. Mobile Front End Block Diagram

The FET width determines the switch Ron, which is the predominate factor in setting the RF insertion loss. So by applying this design sequence for each potential switch technology, we can compare the insertion loss for a nine throw antenna switch, as shown in Fig. 5. While MEMS provides the lowest insertion loss, SOI is today a fraction of the cost and provides the highest level of integration. So SOI is the preferred technology.

III. Filter Technology

Filters are a significant contributor to the overall font end insertion loss between the PA output and the antenna. So for filters, the key design parameter is to maintain out of band isolation while minimizing the passband insertion loss. A typical passband response for a duplexer is shown in Fig. 6.

As the spacing between the transmit passband and the receive passband decrease, the number of resonators required to achieve the out of band attenuation increases, and this in turn increases the insertion loss. For a high performance duplexer, you typically want the Tx isolation in the Rx passband to be >55dB and the Tx or Rx passband insertion loss to be <2dB.

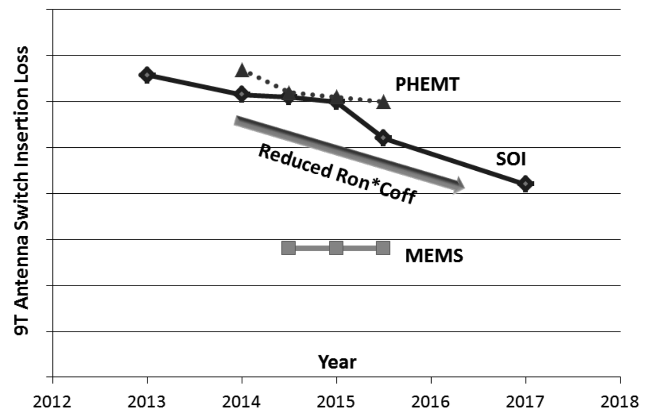

Fig. 5. 9T Antenna Switch Insertion Loss

Fig. 7 is a plot of the passband insertion loss vs. unloaded resonator Q for a band 2 Tx filter designed to maintaining the specified Rx band isolation. The Q of potential technologies is also noted on the plot. It is clear that discrete filter implementations with inductors implemented in PCB, LTCC or chip form with Qs around 40−50 would have far too high an insertion loss. Also noted are tunable technologies. But their low Q also results in an unacceptable passband insertion loss. In order to have an insertion loss of <2dB, a high Q SAW or FBAR process is required.

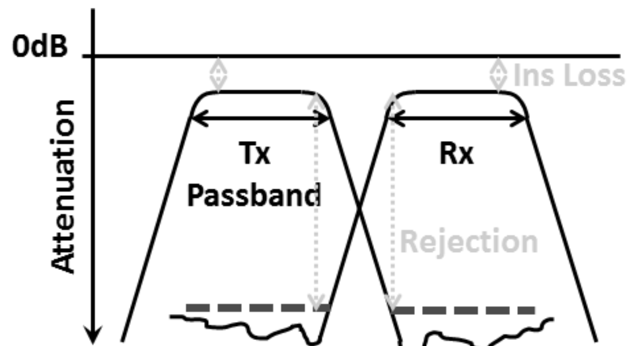

Fig. 6. Duplexer Passband Response

Carrier aggregation allows multiple bands to simultaneously receive or transmit a signal. Previously it was acceptable to have a switch which would connect one band at one time to the antenna. With carrier aggregation it is necessary now to connect two or three bands simultaneously to the antenna through a diplexer, triplexer, or quadplexer. In the example block diagram of Fig. 4, the mid and low frequencies are

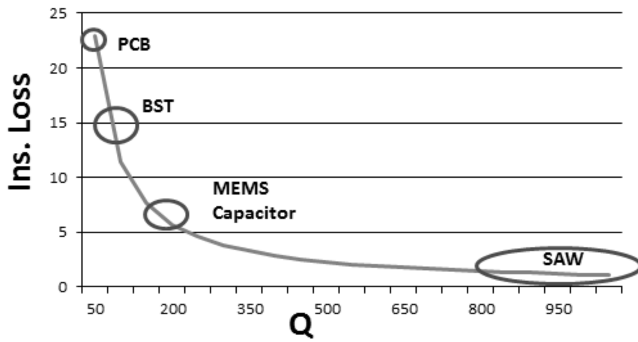

Fig. 7. Filter Resonator Q vs. Insertion Loss

combined through a diplexer, and the high band frequencies have a separate antenna. This allows up to three bands to be active at one time. When implementing this diplexer in a PCB or LTCC technology the insertion loss is ≈0.4dB. However, this implementation requires two antennas. One high and one mid/low antenna, which consumes added space in the phone. It would seem apparent to triplex the three bands, reducing the number of antennas to one. But the small percentage frequency spacing between the mid and high bands causes the implementation loss to be high. A mid-high band diplexer implemented in a low cost PCB, lumped element or LTCC technology would have ≈2dB of insertion loss. To overcome this added front end loss, the PA output power must increase 2dB, and this would result in the PA consuming 58% more battery current. This is a high price to pay and is the reason why a two antenna implementation is often used.

In summary, the high Q SAW and FBAR technologies are required for band filters and duplexers. Carrier aggregation capability can be added to the front end with a low cost lumped element technology. But to minimize the implementation loss the number of antennas may increase. [3]

IV. PA Technology

For the PA design, the competitive advantage is power added efficiency (PAE), size, and cost while maintaining other system specifications. The Class of operation, load line impedance (R_l), transistor size, and bias are set to optimize the efficiency of the PA. There are primarily two types of PAs used today in the handsets, linear PAs and ET PAs.

For linear PAs, the saturated Pout is set high enough to pass the peaks of the waveform with minimal distortion. As the peak-to-average power increases, the average power is set further below the maximum PA saturated output power. The power delta between the average power and the saturated power is referred to as the PA "back-off" power. The higher the data rate, the higher the peak-to-average waveform. Therefore, the higher the back-off power the lower the efficiency of the PA.

In contrast to the linear PA, the ET PA operates based on two fundamental concepts: 1.) A PA's maximum saturated output power is proportional to VCC, and 2.) A PA's maximum PAE occurs at the maximum saturated output power. The load line is calculated and fixed based on the maximum peak output power to be transmitted at the maximum rated supply voltage. Once the load line R_l is fixed, the saturated Pout is varied by changing Vcc to track the envelope of the signal. Note that the range of power which can be tracked in an ET system is limited by the minimum voltage which can be placed on the drain or collector of the transistor. Typically the minimum voltage is around 1V, below which the gain of the transistor drops to unacceptable levels. So the range of output power which the ET system can track is typically limited to the top 6 to 10dB of output power. Below this output power range the system converts back to a linear PA operation.

Class of Operation: RF PAs can be designed using various classes of operation A–F. For mobile handset linear PA operation, a class F load is typically used because it produces a higher efficiency at a given linearity [4].

The class F PA is achieved by shaping the collector voltage and current waveforms with harmonic traps or resonators in the output network. The voltage waveform is shaped to approximately a square wave, and the current waveform approximates a half-sine wave. This is accomplished by presenting a short to the collector at the even harmonics and an open for the odd harmonics. As the number of harmonic terminations increase, the efficiency of an ideal class-F PA increases. For high efficiencies the output network can become complex but in practice, only the first few harmonics are required to be correct [4, 5]. Typically the 2nd and 3rd harmonics are optimally terminated for the sake of size and cost.

Class F and class E PAs have both been used for ET PAs. Class E provides improved efficiency because with a simple output network the collector waveforms are properly shaped. Class F is used because it provides better linearity when the system drops back to APT mode.

From a transistor technology perspective, there is an ongoing debate in the market and in literature[6] primarily between GaAs HBT PAs, which currently hold over 95% of the market, and SOI or CMOS PAs.[7] It is straightforward to compare these technologies once we establish the PAE operating points of APT and ET systems.[8] By plotting the PAE vs. output power and noting the average operating points, we can see and compare the efficiency difference between the linear and ET PAs, as well as HBT and SOI/CMOS PAs.

Fig. 8 is a plot of PAE vs. Pout for a two stage PA with various Vcc voltages. For the rest of this paper, the transistor

drain or collector voltage will be referred to as a collector voltage for simplicity, but when a FET transistor is used, the “collector” reference should be replaced with “drain”. Since for an ET system the PA is saturated, and the collector voltage is modulated to vary the saturated output power. The ET operating curve is the solid black line labeled “ET PA”. This family of curves assumes that the maximum peak voltage which can be applied to the drain or collector is 3.5V. The saturated output power at 3.5V will be equal to the maximum peak output power of the PA. So when transmitting a WCDMA voice waveform with a 3.5dB peak to average power ratio, the ET average operating point will be ≈3.5dB below Point A, which is labeled the “Voice Mode ET” operating point. From the plot the average ET voice operating point at max. Pout would be about 4% PAE below point A, and has an average modulated collector voltage of about 2.3V.

For the APT operation, the collector voltage must remain at 3.5V to support the WCDMA voice waveform peaks, but the average operating point of the PA follows the black dotted line labeled “Linear PA”. So the “voice mode APT” operating point is about 21% below Point A, and is 17% below the “Voice Mode ET” point. For a WCDMA voice signal the difference in PA PAE between ET and APT operation is about 17%. It must be noted that while the data for this example was measured on a specific PA, the relative change in efficiency from point A to the ET or linear operating points will remain similar for changes in transistor technology or load lines. So the absolute efficiency or starting Point A will vary from PA design to PA design, but the relative difference in efficiency from Point A will remain close to the same. Also, one can estimate the PA efficiency for various waveforms by moving down the ET or Linear PA operating curve from point A by the peak to average power of the new waveform.

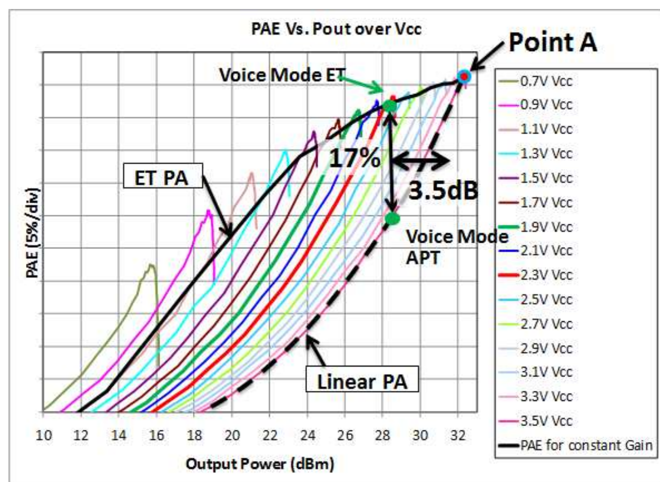

Fig. 8 PA PAE vs. Pout over Vcc – Voice Mode OP

Fig. 9 is the same plot as Fig. 8, but adds a LTE data operating point with a peak to average waveform of about 6.25dB. In this case the ET PA performance will be about 8% below Point A, and the APT operating point will be 30% below Point A. While the curves plotted are for a HBT transistor, this plot is generally true for all transistor types. The SOI or CMOS transistor and some HBT transistor types can have a higher efficiency of up to 10% at drain or collector voltages below 1.5V, but this has no more than a 2% improvement in PAE at the maximum Pout operating point.

Now that we know the relative change in PAE from Point A for various operating points, we can now simplify the transistor technology comparison by comparing the performance at Point A. Also, to compare published or measured results, there must be a consistent PA block used for the comparison. Since most handset PA designs are a two stage PA, this block diagram will be used and is shown in Fig. 10.

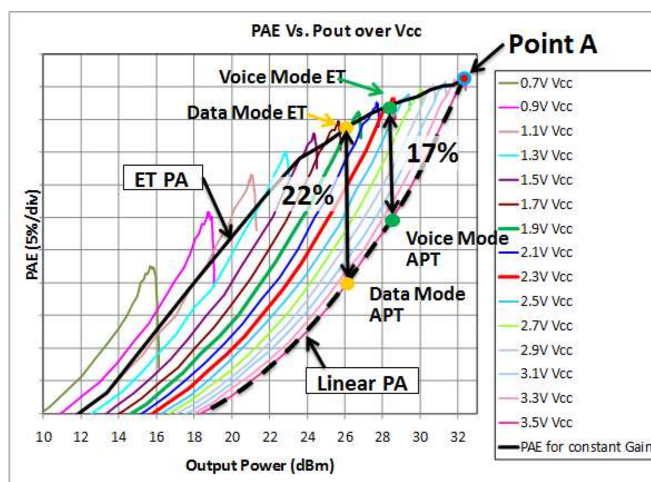

Fig. 9 PA PAE vs. Pout Over Vcc – Data Mode OP

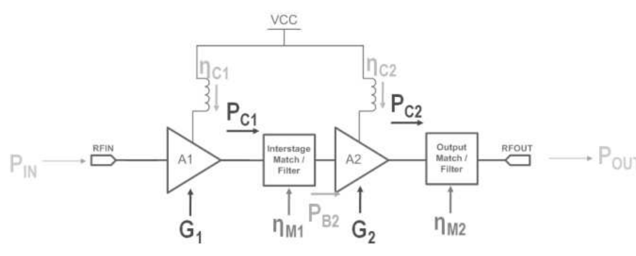

Fig. 10 Multi Stage PA

This block diagram identifies the key parameters that affect the overall PA PAE.

P_{IN} = RF input power

P_{OUT} = RF output power

η_M = Match efficiency

G= Gain

η_C = Collector or Drain efficiency

P_C = RF Collector or Drain power

P_B = RF Base or Gate power

The final stage dominates the efficiency of the PA PAE and can be calculated as follows. The second equation is the PAE for the 2 stage PA.

$$PAE_{FINAL} = \eta_{C2}(\eta_{M2} - 1/G_2)$$

$$PAE_{2stage} = \frac{G_1 G_2 \eta_{M1} \eta_{M2} \eta_{C1} \eta_{C2} - \eta_{C1} \eta_{C2}}{G_1 \eta_{C2} + G_1 G_2 \eta_{M1} \eta_{C1}}$$

With these equations, we can calculate and compare various published PA or transistor PAE results. [8] Table 1 provides this comparison of HBT, SOI and CMOS technologies.

Table 1
PA PAE vs. Pout Over Vcc – Data Mode OP

Technology	Drain or Col. Eff. η_{C2}	Output Match Loss η_{M2} (dB)	G_2	Inter-stage loss η_{M2} (dB)	Driver eff. η_{C1}	PA PAE Point A
GaAs HBT - Simulation	89%	-0.65	14	-0.5	50%	71%
Measured HBT 2013 Product						66%
Measured HBT 2015 Product						78%
SOI - Simulation	75%	-0.65	14	-0.5	40%	60%
CMOS Amalfi[9]						50%
CMOS Nujira paper[10]	57%		14	-0.5	40%	54%
SOI Carrara, Presti, et al [11]	72%	-0.65	14	-0.5	40%	57%
CMOS 65nm [12]	70%	-0.65	14	-0.5	40%	56%
Peregrine SOI PA [13]	72%	-0.65	14	-0.5	40%	57%
ST H9SOI [14]	80%	-0.65	14	-0.5	45%	64%

From this data it is clear that the HBT still has an advantage in PAE, especially when one compares products that are shipping in volume production.

V. CONCLUSION

The best technology of choice for each section is clear from a performance perspective. For the best performance, a combination of SOI, SAW, and HBT technologies are required. If we are to pick the best SOC technology, SOI is the leader with the exception of filter integration. While it is typically assumed that a SOC solution has the advantage in cost and size based on digital trends, this is not true for the RF front end. By moving the PA output match into the low cost laminate and all logic and bias into a low cost CMOS process, the resulting product is both lower in cost and smaller in size. This is because the HBT die size becomes a small fraction of the SOI SOC solution, lowering cost. The logic and bias functions are in CMOS, rather than SOI, further lowering cost. Plus, with these two separate die, we can stack one on top of the other, reducing the overall size of the implementation. The result is that by choosing the best performing technology and proper partitioning of circuit blocks, the SIP solution becomes the clear leader in performance, cost and size.

ACKNOWLEDGEMENT

The author would like to acknowledge the work and effort of Ed Lawrence, Greg Blum, Dave Ripley and Phil Lehtola that contributed to this paper, among many other colleagues at Skyworks.

REFERENCES

[1] Cisco VNI Mobile 2014.

[2] 3GPP Standard TS 36.101 release 12, available online at ftp://ftp.3gpp.org/specs/2014-12/Rel-12/36_series/

[3] J. Young, “Carrier Aggregation (CA) Quantifying Front End Loss,” IWPC Workshop, Chicago, Sept. 16, 2014.

[4] F. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Pothecary, J. Sevic, N Sokal, “Power Amplifiers and Transmitters for RF and Microwave,” IEEE Transactions on Microwave Theory and Techniques, Vol. 50 No. 3, March 2002.

[5] F. Raab, “Maximum Efficiency and Output of Class-F Power Amplifiers,” IEEE Transactions on Microwave Theory and Techniques, Vol. 49, NO. 6, June 2001.

[6] M. Lapedus, “CMOS And SOI Invade RF Front End,” Semiconductor Engineering, April 18, 2013.

[7] H. Jones, “Market Overview and Opportunities With Emphasis On RF,” 2014 International RF SOI Workshop, Shanghai, China, Sept. 23, 2014.

[8] J. Young, “Mobile Handset PA Performance, ET vs. APT & GaAs HBT vs. SOI/CMOS,” 2014 International RF SOI Workshop, Shanghai, China, Sept. 23, 2014.

[9] Hee-Soo Lee, Andy Howard, “RF Power Amplifier Design Series Part 4: RF Module Design using Amalfi CMOS PA,” 2012 Agilent Technologies, Inc., Webcast.

[10] “Envelope Tracking: Unlocking The Potential Of CMOS PAs In 4G Smart Phones,” Nujira White Paper, Feb. 2013.

[11] F. Carrara, C.D. Presti, G. Palmisano, A Scuderi “Power Transistor Design Guidelines and RF Load-Pull Characterization of a 0.13-um SOI CMOS Technology”.

[12] S. Leuschner, et al. “A 31dBm, High Ruggedness Power Amplifier in 65nm Standard CMOS with High-Efficiency Stacked-Cascode Stages,” 2010 IEEE Radio Frequency Integrated Circuit Symposium, RTU1C-3.

[13] N. Comfoltey, D Kelly, D Nobbe, . Olson, “State-of-the-Art of RF Front-End Integration in SOI CMOS,” 2013 IEEE S3S Conference.

[14] L. Formenti, “ST H9SOI_FEM: 0.13 RFSOI Technology for Front End Module Monolithic Integration,” 2014 International RF SOI Workshop, Shanghai, China, Sept. 23, 2014.

ACRONYMS

SOC: System on Chip
SIP: System in Package
PA: Power Amplifier
PAE: Power Added Efficiency
HBT: Heterojunction Bipolar Transistor
GaAs: Gallium Arsenide
SOI: Silicon on Insulator
MEMS: Microelectromechanical Systems
TDD: Time Domain Duplexed
SAW: Surface Acoustic Wave
TC-SAW: Temperature Compensated SAW
FBAR: Thin Film Bulk Acoustic Resonator
BAW: Bulk Acoustic Wave
APT: Average Power Tracking
ET: Envelope Tracking