# Challenges in Compact Modeling for RF and Microwave Applications

A. M. Niknejad
M. Dunga, B. Heydari, H. Wan, C.-H. Lin, S. Emami, C. Doan, X. Xi, J. He and C. Hu

University of California, Berkeley

## ABSTRACT

Commercial CMOS chips routinely operate at frequencies up to 5 GHz and exciting new opportunities exists in higher frequency bands such as 3-10 GHz, 17 GHz, 24 GHz, and 60 GHz. Many research groups have already demonstrated CMOS circuit operation beyond 100 GHz. Will circuit design and compact modeling continue along the same course, or is a new microwave design methodology required? This paper will highlight the design and modeling challenges in moving up to these higher frequencies. A merger of RF and microwave design perspectives will be used to offer insight into the problem. The paper will discuss requirements for a next generation compact model to meet these challenges and offer potential solutions.

***Keywords***: compact modeling, cmos rf, bsim5, rf models, microwave models

## 1 INTRODUCTION

Silicon technology has all but displaced GaAs and other technologies for RF applications in the low GHz regime. A few niche applications, such as power amplifiers, remain as a stronghold but are also under threat by several upstarts. For those with faith in Moore's Law, this was an inevitable consequence in scaling. Transistors became small enough, and consequently fast enough, to operate into the GHz frequencies, thus vying for countless communication applications in these frequency bands. This was not simply a choice between two technologies, but rather a philosophical revolution.

Engineers trained in the art and science of Si analog design carried with them a design methodology based on compact models and SPICE-based transient computer simulation. The microwave community, on the other hand, relied on a rich theory based on two-port parameters and extensive measurements to characterize devices. Non-linear analysis was performed with harmonic balance simulation rather than SPICE time based transient simulation. Active devices were often treated as black boxes with a given $N$-port response characterized by measurements or simple equivalent circuits, rather than compact models based on physical equations. Which design methodology is most appropriate for emerging new applications in the microwave and mm-wave bands?

Economics were the deciding factor in the low-GHz decision. Silicon RF technology was inexpensive, and in the case of CMOS, practically free. The burden and major cost to develop CMOS technology was provided by a large digital microprocessor market. Early RF and analog designers using CMOS were hampered by modeling problems, as the models were often extracted for digital applications. Today the situation is quite different. CMOS is the technology of choice for many RF applications, including low-range low-cost radios, short-range high data-rate wireless LAN, cellular radios, and even power amplifiers. Today communication equipment is a large and perhaps the largest fraction of the world's semiconductor manufacturing and thus RF compact models and design kits are more common. Most commercial applications require operation up to 6 GHz, but a plethora of new standards are emerging with applications up to 10 GHz for UWB, 24 GHz [1], and even 60 GHz [2]. Other emerging applications include automotive radar at 77 GHz.

In this paper we will review the traditional CMOS RF compact modeling approaches and compare it to a microwave design methodology used for a 60 GHz CMOS front-end. As we shall see, successful CMOS design at these higher frequencies requires more careful modeling, and more attention to passive devices. In order to retain the flexibility of a geometry-scalable and bias independent model, a hybrid approach is proposed. The requirements for the next generation compact model are then discussed, and many features of the new BSIM5 model are elaborated.

## 2 ROLE OF COMPACT MODELS

Compact models are the interface between technology and design. A circuit designer learns about a process by experimenting with the compact model, rather than running expensive and time consuming experiments. Therefore compact models should be scalable with geometry, accurate across a wide temperature and bias voltage range. It is increasingly important for compact models to be rooted on a physical foundation in order to account for the effects of statistical fluctuations inherent in the fabrication process. Good examples include

dopant fluctuations and gate oxide variations. With aggressive device scaling, the corner-based design strategy is over pessimistic and statistically robust designs are now more common.

Present day models are not necessarily "compact," often occupying about 10,000+ lines of C code, but should be fast enough to allow "real time" simulation. The emergence of Verilog-A compact models has the potential to simplify the description of compact models, by divorcing the compact model from the circuit simulator [3]. In theory, a simple simulator independent compact model will also allow fast testing and incorporation of the models into circuit simulators.

## 3 CMOS TECHNOLOGY FOR MM-WAVE APPLICATIONS

As Si technology continues to improve, there are new microwave and mm-wave frequency bands to conquer. Most people have no doubt that Si technology is the right choice for conquering this new frequency spectrum, but doubts about the right design approach remain. Already we have seen that 130 nm digital CMOS technology has an $f_{max} = 135\,\text{GHz}$ [2], and 90 nm CMOS has even better performance, with demonstrations of $f_{max} = 280\,\text{GHz}$ [4].

The main disadvantages of CMOS technology stem of the use of a polysilicon gate, which results in high gate resistance, a conductive substrate, which has a tremendous impact on the quality of passive devices, and a relatively low surface mobility due to surface scattering and dopant scattering. The gate resistance is a relatively minor issue, since advanced lithography allows very short gate lengths to be employed, and thus the gate resistance is minimized by using a multifinger (or folded) transistor [5][2]. The aspect ratio of the MOS device is awkward, though, requiring a relatively narrow transmission line feed to taper out in order to feed a large transistor.

As CMOS technology continues to scale, though, the structure of the core device is undergoing rapid changes. For instance, metal gates may replace polysilicon for work function engineering [6]. In a highly scaled transistor, threshold control by doping is impractical due to small atomic dimensions and dopant variations cause large threshold shifts. The lossy bulk silicon substrate may be displaced by an insulator, an option available today through SOI technology. New device structures such as the FinFET [7] are also fabricated over an insulating substrate. Strain by the introduction of Ge in the device also produces enhanced carrier mobility[8]. The introduction of thick copper metal along with 6-10 metal layers and low-K dielectrics is also benefiting passive devices in other ways. Fine lithography allows metal finger (comb) capacitor structures which achieve high density capacitance without any additional masks. Overall, CMOS technology is becoming increasingly suitable for RF applications, especially when moving to higher speeds.

MOSFET devices play an increasingly important role in RF and microwave circuits. Not only as a small-signal amplifier, but as a large-signal power device, low-loss switching devices (mixers), and as a varactor. As such, the compact modeling of CMOS devices for RF and microwave devices has become increasingly important. On one front the high frequency parasitics and NQS behavior needs to be modeled. On another front, though, a high $f_T$ device can be applied to low-GHz applications in sub-threshold mode, exploiting the higher $g_m/I$ bipolar mode of operation. Accumulation mode varactors are replacing PN junctions in many applications for VCOs. Thus a transistor model that is accurate over the entire region of operation, from accumulation, depletion, sub-threshold, and strong inversion is needed.

## 4 REVIEW OF TRADITIONAL RF COMPACT MODELS

Most RF compact models emerged as a result of discrepancy between measured RF components and simulated performance. Typical problems included impedance mismatches at the input of amplifiers, higher noise in LNAs, higher phase noise in VCOs, and lower output power in CMOS power amplifiers. Other issues included mismatches in device distortion, especially when devices were operated as switches about a $V_{ds} = 0$ operating point. Spiral inductors presented another challenge, as most fabricated inductors had lower quality factors than predicted by models, even full three-dimensional electromagnetic simulation.

The source of the modeling errors was quickly identified and simple enhancements to the core compact models were proposed [9] [10]. Much of the mismatch between measured and simulated two-port parameters were attributed to the gate resistance at the input and the substrate resistance at the drain of a transistor. The gate resistance included the bias-independent distributed poly gate and the distributed bias-dependent channel resistance. Problems with noise were quickly attributed to the lack of the gate noise and proper correlation of gate noise to the channel thermal noise of a device. In essence, both of these issues were related to the non-quasi static behavior of a MOS device. An accurate and consistent NQS model would correctly predict the two-port parameters as well as the enhanced noise of the device. A sectional model based on a quasi-static core transistor model has been demonstrated to accurately predict NQS issues [11].

Today the problems with RF compact modeling are relatively well understood, but an elegant and accurate solution remains the focus of intense research. In BSIM4, the channel thermal noise is partitioned to a

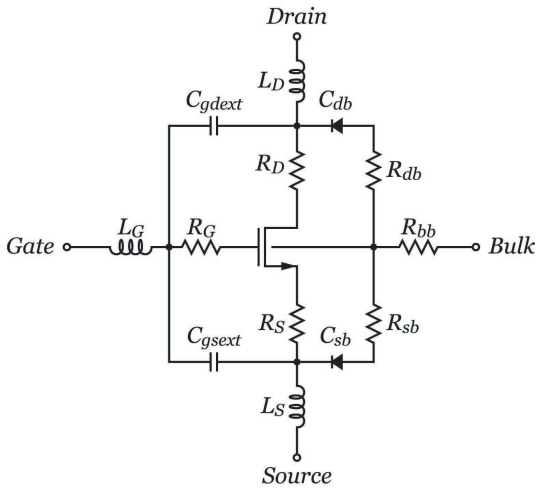


Figure 1: A core BSIM3.3 model enhanced with frequency independent lumped components to model high frequency effects.

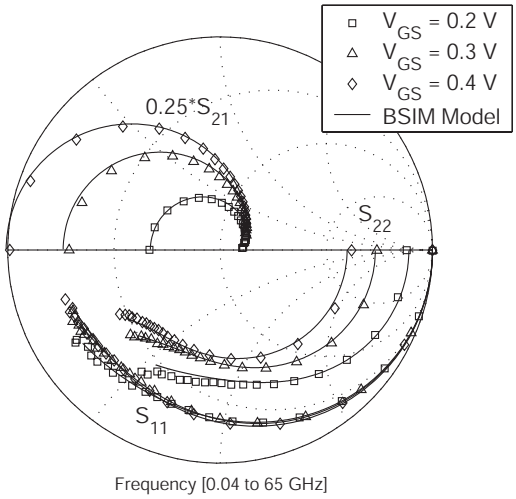


Figure 2: Measured and simulated $s$-parameters up to 65 GHz versus bias.

drain current noise and a source side noise, thereby automatically producing gate-induced noise with a correlation to the drain noise [12]. The strength of the BSIM model is the simplicity, but the model has some shortcomings. For instance, the holistic thermal noise model in BSIM4 incorporates the magnitude of the correlation coefficient for the gate and drain noise, but not the complex phase. Fortunately, most practical circuits are not sensitive to the phase of the coefficient.

In BSIM4, the NQS effect is modeled with an extra node, modeling the deficit or surplus charge [13]. The BSIM4 NQS model is also very simple and effective, capturing bias-dependent gate resistance in a MOSFET. To first order, the frequency dependence of two-port parameters is also captured with this approach, and ongoing research is underway to improve the accuracy of the NQS model. A new geometry scalable five-resistor substrate resistance network [14] of BSIM4 is demonstrated to match two-port parameters very accurately up to high frequencies.

Large area MOS devices also find wide application in RF power amplifiers [18] [19]. CMOS power devices are notorious for their underperformance, with 5%-10% worse efficiency and measuring 1 dB -2 dB below expected output power. These multi-finger devices require careful modeling of transistor loss, especially the substrate resistance, in order to correctly predict the power gain and efficiency. Furthermore, due to high power dissipation, temperature dependent effects such as mobility degradation need accurate modeling.

# 5 MICROWAVE COMPACT MODELS

## 5.1 Active Devices

Microwave design is distinctly different from the CMOS approach. The design methodology relies much more heavily on $s$-parameter measurements. Since GaAs and other microwave substrates are inherently insulating, passive devices have inherently higher $Q$ and are simpler to model. Non-linear large signal models are much simpler, using truly compact equations to capture the behavior of the device. Otherwise a table based $y$-parameter approach is used to model devices. The Root Model is a popular hybrid technique [15].

The classic microwave approach treats an active device as a two-port black box. A rich body of theory is available to analyze the gain, stability, and noise of a linear two-port in terms of its two-port parameters ($s$-parameters, $y$-parameters, etc.). It is interesting to note that a relatively simple equivalent circuit can fit the data over a wide frequency range. By adding a few additional frequency-independent lumped elements to a core BSIM model, shown in Fig. 1, the model can capture the AC behavior of a transistor over a wide frequency range (Fig. 2) and bias. The core BSIM model is optimized using DC measurements and the parasitics are added to fit $s$-parameters. A particularly good test of the compact model is Mason's Unilateral gain $U$. Mason's gain is defined as the maximum gain of a transistor that is unilaterized by lossless passive feedback. Mason's gain is very sensitive to transistor loss and a good fit between measurement and theory is a sure indication of the quality of the model.

But this modeling approach is limited to a single transistor as the geometry of the transistor is fixed.

Even though a fixed layout with optimal finger width is employed, one cannot simply scale the model with finger count due to the important role of the wiring parasitics at high frequency. A scalable model is very desirable for microwave and RF design. For instance, varying the bias voltage and geometry of a FET can achieve a simultaneous noise and power match. Co-simulation of the transistor parasitics using an EM solver with a compact model of the intrinsic transistor finger yields a scalable model.

To characterize the large signal performance, harmonic power is measured and compared to the predictions of the enhanced core BSIM model [16]. The results indicate that the DC non-linearity of the device, in addition to the $C$-$V$ non-linearity inherent in the model, can predict non-linear behavior reasonably well up to 65 GHz. This has been verified at the device level and at the circuit level, where a mixer conversion gain and amplifier compression are verified experimentally up to 60 GHz[2].

## 5.2 Passive Devices

Passives play an equally important role at high frequencies. Since we operate closer to $f_T$ or $f_{max}$, we need passive elements to tune the input/output of a MOSFET to the correct optimal impedance values. Coupling capacitors are commonly used to isolate DC levels. Bypass capacitors are important since high frequency signals must be kept localized. PN-junction or MOS varactors are common for frequency tuning. As long as the self-resonant frequency (SRF) of these structures is captured, simple compact models can be employed.

One of the more difficult elements to model is the multi-turn inductor. But often a high frequency circuit needs very small amounts of reactance to tune out a small MOS capacitor. Reactances can be as low as tens of pH. For instance, 16 pH has 10 Ω of reactance at 100 GHz. A ring inductor is one option to realize a small inductance value and simple compact models work well.

We can also view the ring inductor as a variable spacing two-line transmission line. An ideal T-line is inherently scalable in length and it fully characterized by $Z_0$, $\alpha$, and $\beta$. The layout of a real transmission line must inherently suppress higher order and undesired modes from propagation. This is easy in a modern CMOS process as multiple metal layers "bridges" can be employed to suppress unwanted modes. The dominant quasi-TEM mode can be modeled as a simple electrical lossy transmission line. A compact model which incorporates the physical losses over a CMOS substrate as a function of frequency and line geometry can be developed if the electric substrate currents and the magnetic substrate eddy currents are included in the analysis [17].

Transformers are now more commonly employed in RF and microwave circuits. Transformers are used for power summing in CMOS power amplifiers [18] [19] and in VCOs as high-Q resonators [20]. Transformer modeling is more difficult than inductor modeling since the interwinding capacitance between the primary and secondary plays an important role in setting the upper frequency limit of operation.

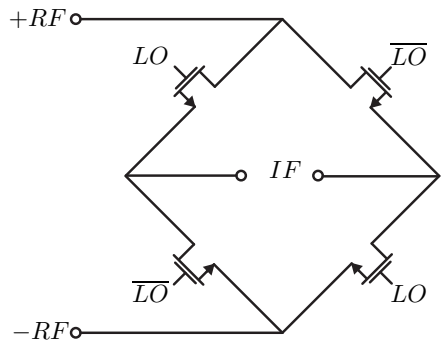


Figure 3: Simplified schematic of a passive ring CMOS mixer.

One key problem with integrated passive devices is their close proximity and interaction through the substrate. Since most design kit models are for single isolated structures whereas real chips contain several passive elements packed in closely to minimize area, there can be a discrepancy in the performance of the devices. The parasitic coupling between the elements can also lead to unexpected and difficult to trace errors such as I/Q mismatches and spurious outputs. The only way to properly account for such coupling is a full-chip EM extraction of the entire chip. This is prohibitevly expensive using full-wave simulation, and faster techniques are sorely needed to address such applications.

# 6 REQUIREMENTS FOR THE NEXT GENERATION MOS MODEL

Due to the complexity of real MOSFET devices, the only accurate way to simulate the structure is with 2D or 3D device simulators (solve PDEs for transport, quantum and electromagnetics simultaneously using numerical techniques). "Physical" compact models of MOSFETs require unphysical assumptions such as the gradual channel approximation (vertical charge control, lateral current flow), the charge sheet approximation, an assumption of uniform doping profiles, and simplified mobility models. Rather, compact models should have physical behavior (consistent $I$-$V$/$C$-$V$, positive conductances, non-negative capacitances, etc.).

## 6.1 General Requirements

The next generation compact model must be built around a continuous and accurate current/voltage ($I$-$V$) core model. The models should be flexible and allow

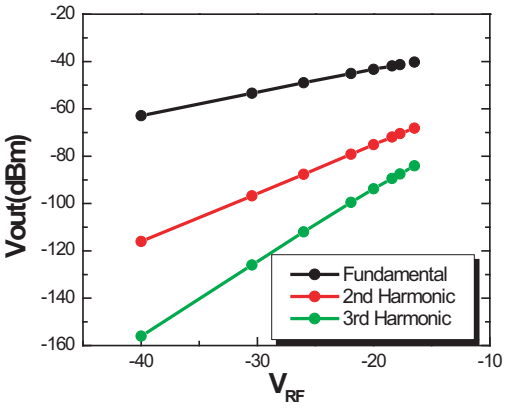


Figure 4: The fundamental and harmonic tones generated by a passive CMOS mixer.

advanced device structures (halo doping, velocity saturation, field dependent mobility, etc). The structure of the model should also easily incorporate advanced structures (FinFETs, Ultra-Thin Body, SOI, etc.). It is important to perform extensive testing of a compact model. Important metrics include an accurate fit of the model and all derivatives of the $I$-$V$ curves. This includes the output resistance model $r_o$, and the transconductance $g_m$ versus bias. The current normalized transconductance $g_m/I_d$ versus $V_{gs}$ and $V_{bs}$ is an important test of the DC characteristics of the model.

The charge/voltage model ($C$-$V$) model is also an equally important component for high frequency applications. A continuous and self-consistent $C$-$V$ model that accurately represents the behavior from accumulation, depletion, weak-inversion, and through strong inversion is required. The model must include poly depletion, quantum effects, and extrinsic capacitance effects. BSIM5 is built around a core charge-voltage model that results in a smooth non-regional model. The charge-voltage relationship is valid from accumulation to strong inversion. The $I$-$V$ and $C$-$V$ model is built on top of the charge model, resulting in a self-consistent model. All equations are bulk referenced to realize a symmetric model. The Gummel symmetry test can verify the accuracy of the DC model. Plot of the trans-capacitances versus bias demonstrates a proper $C$-$V$ model [21].

Simulation of the $HD_{2,3}$ of a passive mixer is shown in Fig. 3, where the transistors switch around a $V_{ds} = 0$ operating point. A non-symmetric model, such as BSIM4, cannot match the theoretical trend of $HD$ shown in Fig. 4, due to the source reference in the model. The BSIM5 curve, though, matches theoretical prediction as expected. A symmetric model is also needed for other important circuits such as CMOS attenuators [22].

The small-signal model should likewise be self-consistent with the DC model and include a good approximation to distributed effects such as the induced gate resistance (NQS) and substrate resistance. The noise model in particular should be very accurate with bias and device scaling. The noise model should be accurate in the subthreshold region as many low current RF circuits operate in this region. The compact model must be capable of simultaneously predicting accurate small-signal metrics (noise) in the presence of a large signal.

A typical situation with RF circuits is the near-far situation (distant weak desired signal, nearby strong interferer), placing strong demands on the time-varying noise behavior of a compact model. High frequency large signal performance should accurately reproduce distortion (e.g. $HD$ and $IM$ products). Good fitting of DC $I$-$V$ and $C$-$V$ curves will help a great deal in fitting high frequency $IM$ curves. Accurate measurements of $IM_3$, though, can show faults in the mobility model [23].

## 6.2 Advanced Transistor Structures

While bulk devices dominate the market today, in the near future SOI and multi-gate devices such as FinFETS may play an important role. These devices have unique characteristics that require additional effort for compact modeling.

In particular, self-heating effects are of integral importance whenever the transistor body is isolated from the bulk. The self-heating effects have a drastic effect on the DC performance, but a negligible effect on the RF device. Thus, these thermal effects should be carefully extracted from the model, requiring special measurement techniques.

For fully depleted SOI devices (FDSOI), the body resistance network may not needed for RF fitting whereas for partially depleted devices (PDSOI), the body resistance network is needed similar to a bulk MOSFET. However this body resistance network has a strong bais dependence. The body resistance network should also capture the transition from FD to PD as a function of bias.

For a multi-gate device, the gate resistance has strong geometry dependence, especially for the multi-fin devices (for large W). The gate resistance is no longer a linear function of the sheet resistance multiplied by square ratio, but instead, the model should include the height, width, and distance of the fins. In addition, the amplification of the channel nose through the body effect ($g_{mb}$) is no longer present, but amplification of noise by the back and top gate needs further research.

# 7 CONCLUSION

Compact modeling for RF and microwave applications plays an increasingly important role. Scaling trends in CMOS provide new opportunities in the microwave and mm-wave spectrum. A new hybrid compact modeling approach is proposed, with emphasis on microwave

characterization of devices while retaining the flexibility of geometry scaling from RF CMOS compact models.

## Acknowledgment

The authors would like to acknowledge the contributions of the SRC member companies, ST Microelectronics and IBM for chip fabrication, and Agilent for with assistance with measurements.

## REFERENCES

[1] G. Xiang, H. Hashemi, A. Hajimiri, "A fully integrated 24-GHz eight-element phased-array receiver in silicon," *IEEE J. of Solid-State Circuits*, vol. 39, Dec. 2004, pp. 2311-2320.

[2] C. H. Doan, S. Emami, A. M. Niknejad, R. W. Brodersen, "Design of CMOS for 60 GHz applications," *2004 ISSCC Digest of Technical Papers*, pp. 440-538.

[3] L. Lemaitre, G. Coram, C. McAndrew, K. Kundert, "Extensions to Verilog-A to support compact device modeling," *Proceedings of the 2003 International Workshop on Behavioral Modeling and Simulation*, pp. 134-138.

[4] L. F. Tiemeijer et al., "Record RF performance of standard 90 nm CMOS technology," *2004 IEDM Technical Digest*, session 17, paper 6.

[5] L. F. Tiemeijer et al., "A record high 150 GHz fmax realized at 0.18 m gate length in an industrial RF-CMOS technology," *2001 IEDM Technical Digest*, pp. 10.4.1 - 10.4.4.

[6] R. Lin, L. Qiang, P. Ranade, T.-J. King, C. Hu, "An adjustable work function technology using Mo gate for CMOS devices," *IEEE Electron Device Letters*, vol. 23, Jan. 2002, pp. 49-51.

[7] S. H. Tang et al., "FinFET-a quasi-planar double-gate MOSFET," *2001 ISSCC Digest of Technical Papers*, pp. 118-119,437.

[8] T. Ghani et al., "A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors," *2003 IEDM Technical Digest*, pp. 11.6.1-11.6.3.

[9] C. Enz, Y. Cheng, "MOS transistor modeling for RF IC design," *IEEE J. of Solid-State Circuits*, vol. 35, Feb. 2000, pp. 186-201.

[10] X. Jin et al., "An effective gate resistance model for CMOS RF and noise modeling," *1998 IEDM Technical Digest*, pp. 961-964.

[11] R. van Langevelde et al., "New compact model for induced gate current noise [MOSFET]," *2003 IEDM Technical Digest*, pp. 36.2.1-36.2.4.

[12] *The BSIM 4.4.0 Manual*, www-device.eecs.berkeley.edu/ bsim3

[13] M. Chan, K.Y. Hui, C. Hu, P.K. Ko, "A robust and physical BSIM3 non-quasi-static transient and AC small-signal model for circuit simulation," *IEEE Transactions on Electron Devices*, vol. 45, April 1998, p. 834-841.

[14] M. V. Dunga, *A scalable MOS device substrate resistance model for RF and microwave circuit simulation*, Master's Thesis, 2004.

[15] J. Wood and D. E. Root, "Bias-dependent linear scalable millimeter-wave FET model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48 , Dec. 2000, pp. 2352-2360.

[16] S. Emami, C.H. Doan, A.M. Niknejad, R.W Brodersen, "Large-signal millimeter-wave CMOS modeling with BSIM3," *2004 RFIC Symposium Digest of Papers*, pp.163-166.

[17] A. M. Niknejad, R. G. Meyer, "Analysis of eddy-current losses over conductive substrates with applications to monolithic inductors and transformers," *IEEE Trans. on Microwave Theory and Techniques*, vol. 49, Jan. 2001, pp. 166 - 176.

[18] I. Aoki, S.D. Kee, D.B. Rutledge, A. Hajimiri, "Fully integrated CMOS power amplifier design using the distributed active-transformer architecture," *IEEE J. of Solid-State Circuits*, vol. 37, March 2002, pp. 371 - 383.

[19] T.-S. D. Cheung, J. R. Long, D. L. Harame, "A 21 to 26GHz SiGe Bipolar PA MMIC," *2005 ISSCC Digest of Technical Papers*, pp. 530-531.

[20] M. Straayer, J. Cabanillas, G. M. Rebeiz, "A low-noise transformer-based 1.7 GHz CMOS VCO," *2002 ISSCC Digest of Technical Papers*, pp. 286 - 287.

[21] X. Xi, J. He, M. Dunga, C.-H. Lin, B. Heydari, H. Wan, M. Chan, A.M. Niknejad, and C. Hu, "The next generation BSIM for sub-100nm mixed-signal circuit simulation," *2004 CICC Proceedings*, pp. 13-16.

[22] H. Dogan, R. G. Meyer, A. M. Niknejad,"A DC-10GHz linear-in-dB attenuator in 0.13 $\mu$ CMOS technology," *2004 CICC Proceedings*, pp. 609-612.

[23] R. van Langevelde, R. M. Klaassen, "Effect of gate-field dependent mobility degradation on distortion analysis in MOSFETs," *IEEE Transactions on Electron Devices*, vol. 44, Nov. 1997, pp. 2044-2052.