RF POWER AMPLIFIER AND OSCILLATOR DESIGN FOR RELIABILITY AND VARIABILITY

by

SHUYU CHEN

B.S. Zhejiang University, 2006
M.S. Fudan University, 2009

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical Engineering and Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2013

Major Professor: Jiann S. Yuan

ABSTRACT

CMOS RF circuit design has been an ever-lasting research field. It gained so much attention since RF circuits have high mobility and wide band efficiency, while CMOS technology has the advantage of low cost and better capability of integration. At the same time, IC circuits never stopped scaling down for the recent many decades. Reliability issues with RF circuits have become more and more severe with device scaling down: reliability effects such as gate oxide break down, hot carrier injection, negative bias temperature instability, have been amplified as the device size shrinks. Process variability issues also become more predominant as the feature size decreases. With these insights provided, reliability and variability evaluations on typical RF circuits and possible compensation techniques are highly desirable.

In this work, a class E power amplifier is designed and laid out using TSMC 0.18 μm RF technology and the chip was fabricated. Oxide stress and hot electron tests were carried out at elevated supply voltage, fresh measurement results were compared with different stress conditions after 10 hours. Test results matched very well with mixed mode circuit simulations, proved that hot carrier effects degrades PA performances like output power, power efficiency, etc.

Self- heating effects were examined on a class AB power amplifier since PA has high power operations. Device temperature simulation was done both in DC and mixed mode level. Different gate biasing techniques were analyzed and their abilities to compensate output power were compared. A simple gate biasing circuit turned out to be efficient to compensate self-heating effects under different localized heating situations.

Process variation was studied on a classic Colpitts oscillator using Monte-Carlo simulation. Phase noise was examined since it is a key parameter in oscillator. Phase noise was modeled using analytical equations and supported by good match between MATLAB results and ADS simulation. An adaptive body biasing circuit was proposed to eliminate process variation. Results from probability density function simulation demonstrated its capability to relieve process variation on phase noise. Standard deviation of phase noise with adaptive body bias is much less than the one without compensation.

Finally, a robust, adaptive design technique using PLL as on-chip sensor to reduce Process, Voltage, Temperature (P.V.T.) variations and other aging effects on RF PA was evaluated. The frequency and phase of ring oscillator need to be adjusted to follow the frequency and phase of input in PLL no matter how the working condition varies. As a result, the control signal of ring oscillator has to fluctuate according to the working condition, reflecting the P.V.T changes. RF circuits suffer from similar P.V.T. variations. The control signal of PLL is introduced to RF circuits and converted to the adaptive tuning voltage for substrate bias. Simulation results illustrate that the PA output power under different variations is more flat than the one with no compensation. Analytical equations show good support to what has been observed.

To my parents and my little sunshine.

ACKNOWLEDGMENTS

First, I would like to express my special thanks to my advisor, Professor Jiann S. Yuan, for his warm, strong support, consideration, and encouragement throughout my graduate studies. His passion and dedication to research has sound effects on me and will guide me in my future life. His consideration and support has won my greatest respect and sincere friendship. He contributed continuous guidance and suggestions in this work. At the meantime, I have been provided with the essential lab resources and software tools to conduct the research work. His technical and editorial experiences were very critical to the completion of this dissertation too.

My thanks also go to my other four dissertation committee members: Dr. Kalpathy B. Sundaram, Dr. John Z. Shen, and Dr. Xun Gong and Dr. Morgan C. Wang for attending my proposal and dissertation defense, reading previous drafts of this dissertation and providing valuable comments that improved the presentation and contents of this dissertation.

I am grateful to all my colleagues: Yidong Liu, Karan Kutty, Hongxia Tang, Yiheng Wang, Jason Steigner, Gabriel Vazquez, Yuying Zhang, Chengcheng Xiao, Alex(Ekavut Kritchanchai). In particular, Yidong Liu helped me with LNA circuit design and ADS Monte Carlo simulation. Karan and I worked together on oscillator IC design as well as Cadence simulation and layout environment set up. Yiheng gave me lots of suggestions on the RF PA design; especially on load pull and source pull simulations. I have collaborated with Hongxia, Jason, Yuying, Chengcheng and Alex in many other issues, they gave me precious inspirations.

Last, but not least, I would like to thank my family, especially my husband Guowei Zhao, for their understanding and support during the past few years. It was their love and support that

made this dissertation possible. My parents – Xizheng Chen and Xucui Zhuang have my deepest gratitude and love for their dedication and many years of support during my studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

ADS	Advanced Design System
ASIC	Application Specific Integrated Circuits
BD	Breakdown
BTI	Bias Temperature Instability
CDMA	Code Division Multiple Access
CHE	Channel Hot Electron
CMOS	Complementary Metal Oxide Semiconductor
DC	Direct Current
FET	Field Effect Transistor
HC	Hot Carrier
HCI	Hot Carrier Injection
HCD	Hot Carrier Degradation
HF	High Frequency
IC	Integrated Circuit
IIP3	Input Third Harmonic Intercept Point
IP3	Third Harmonic Intercept Point
KCL	Kirchhoff's Current Law
KVL	Kirchhoff's Voltage Law
LNA	Low Noise Amplifier
MC	Monte Carlo
MOS	Metal Oxide Semiconductor

MOSFET	MOS Field Effect Transistor
NBTI	Negative Bias Temperature Instability
NF	Noise Figure or Noise Factor
NMOSFET	N-type MOS Field Effect Transistor
OPAMP	Operation Amplifier
PA	Power Amplifier
PAE	Power Added Efficiency
PBTI	Positive Bias Temperature Instability
PLL	Phase-Locked Loop
PMOSFET	P-type MOS Field Effect Transistor
RF	Radio Frequency
RFIC	Radio Frequency Integrated Circuit
SBD	Soft Breakdown
SH	Self-Heating
SNR	Signal-to-Noise Ratio
SOC	System on Chip
SOI	Silicon-on- Insulator
STD	Standard Deviation
SV	Statistic Variation
TDDB	Time Dependent Dielectric Breakdown
μ_n	Mobility of N-Type MOS Field Effect Transistor
VLSI	Very-Large-Scale Integration
V_{DD}	DC Supply Voltage

V_T	Threshold Voltage

CHAPTER ONE: INTRODUCTION

1.1 Motivation

For centuries, semiconductor industry has followed a rule called “Moore’s Law”. That the device size will scale down by a factor of two every three years and transistor density would double every one or two years. This trend has been proved in digital IC field. Analog and RF circuit systems are a little bit behind this trend, but generally device sizes in these fields are also shrinking with time. Ideally, technology scaling will keep the “constant field scaling”, with both lateral and vertical dimensions decrease by the same percentage, threshold voltage and supply voltage decrease by the same percentage, and doping level increase by the same amount too. In practice, this is hard to realize, mostly because supply voltage and threshold voltage can’t shrink with the same pace, and also due to inevitable short channel caused second order effects. As a result, a mixture of “constant field scaling” and “constant voltage scaling” is what we have. This increases the electric field and thus raises the possibility of device breakdown and other reliability issues.

CMOS technology has its advantage in low cost and high integrity and finds its applications in high frequency ICs. With the rapid growth of IC industry, CMOS RFICs are widely used in wireless communication systems, like mobile phone and TV, Bluetooth, WLAN, wireless sensing system, etc. Due to aggressive scaling in device dimensions for improving speed and functionality, CMOS transistors in the deep sub-micrometer to nanometer regime has resulted in major reliability issues including gate oxide breakdown and channel hot electron degradation, NBTI, and variability.

Many RF circuits are vulnerable to RF reliability issues mentioned above, they suffer from different reliability problems since they have different operation schemes and different structures. It is very urgent to study the sources and reasons of their reliability and variability issues, compare possible compensation circuits and their effects on circuit performances, thus aiming for the goal of high yield and better reliability performance. Unfortunately, there have not been any universal rules developed on the relationship between RF circuits and susceptible reliability issues yet. Each RF circuits should be studied individually according to its unique features.

Power amplifier and oscillator are essential parts in RF transceivers. Power amplifier is the last one before antenna in a transmitter and serves to amplifier the power to be transmitted. Oscillators are used to provide signal sources for frequency conversion and carrier generation. It is of great significance to keep them working stably over variations of temperature, process, voltage supply, and other stress and degradation conditions.

1.2 Goal of research

This work is mainly focused on solving issues listed below:

1. Principle and theoretical study of typical reliability issues (Impact Ionization, self-heating effects) and verification by device and circuit level simulations.
2. Circuit design and chip implementation of RF PA, trippler, oscillator circuits.
3. Reliability analysis based on experimental results on class E PA.
4. Propose and compare possible compensation circuits, like adaptive gate biasing circuit, adaptive body biasing circuits, on chip PLL, etc.

5. Set up equations and modeling circuit behavior in Matlab, compare with ADS simulation results.
6. Monte-Carlo simulation to demonstrate variability issues and compensation effects of adaptive body biasing circuit.

1.3 Results outline

To summarize, chapter two gives an overview of current reliability and variability issues remained in RF circuits design. The author evaluated hot electron and oxide stress effects on Class E PA by designed chip and experiments, details are shown in chapter three. Temperature compensation technique is established for RF PA, verified with analytical equations, this is described in chapter four. Chapter five examined Process variations and reliability on Colpitts oscillator using Monte Carlo simulation and mixed-mode simulation, and supported by analytical equations. A robust, adaptive design technique to reduce PVT variation effects on RF circuits is developed in chapter six. Chapter seven is the final conclusion and future work.

CHAPTER TWO: RF CIRCUITS AND RELIABILITY ISSUES OVERVIEW

2.1 Breakdown

Breakdown refers to the destruction at the gate oxide of a MOSFET due to overstressed electrical field. It is claimed that a complete evolution of gate breakdown is composed of four stages as shown in Figure 1[1]: SILC (Stress Induced Leakage Current) stage, Di-BD (Digital Breakdown) stage, An-BD (Analogue Breakdown) stage and HBD (Hard Breakdown) stage.

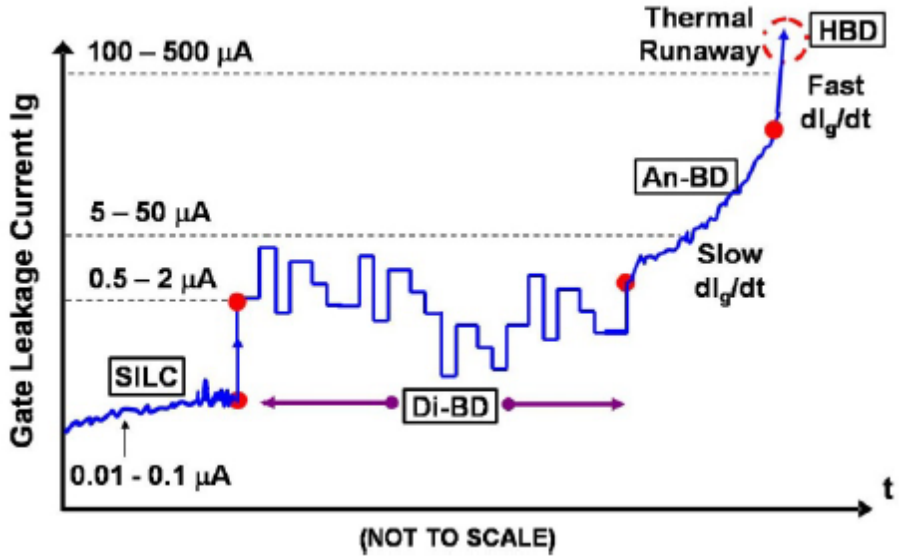


Figure 1: Illustration of a complete evolution of gate breakdown[1].

There are some random weak spots within the dielectric which comes from the uneven or poor dielectric growth process, which is described in Figure 2. They may contain sodium ions, contaminations or crystalline Si defects. Defects within the gate oxide are usually called traps because the degraded oxide can trap charges. Traps allow for creation of conduction path. Different factors that induce the formation of traps within the gate oxide include the electric field,

hot carriers, and radiation. In the first stage, which is also called Pre-TDDB, Trap assisted tunneling (TAT) is initiated and results in a power law increase of the leakage current with time. Leakage current at this stage is quite low, typically below 0.1μA.

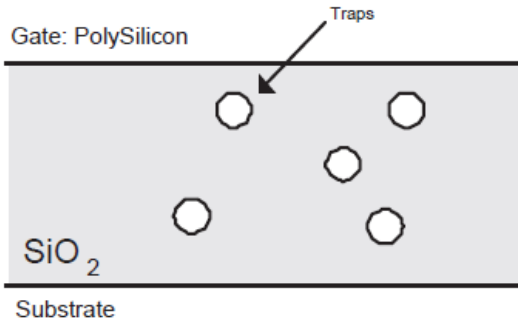


Figure 2: Formation of traps within dielectric.

If the stress lasts longer, more and more defects are accumulated; the average distance between defects becomes smaller. Some defects may now get connected and form a complete conducting path bridging the gate and substrate, as in Figure 3. Once this conduction path is created we have Soft Breakdown (SBD). Soft breakdown is a chronic, long term effects usually caused by relatively low gate oxide stress. The property of this MOSFET changes from the normalized value, The threshold voltage and channel mobility will shift, the current-voltage characteristics will also vary, as proved in many works[2].In order to further explain the mechanism of soft breakdown in details, many models have been proposed, including Anode Hole Injection Model[3], Electron Trap Generation Model[4,5], and Percolation Model [6,7].

This phenomenon is also termed as “percolation” of traps, evolves the device into post-BD regime, two distinct stages occur, which are Digital breakdown and Analog breakdown. In Di-BD, gate current random telegraph noise (RTN) fluctuates between certain fixed discrete current levels. These fluctuations are attributed to the Poissonian trapping-detrapping mechanism

of the electrons at the oxygen vacancy defect region. An-BD takes place when the average leakage current starts to increase monotonically, usually in an exponential manner. The transition from Di-BD to An-BD is governed by a critical voltage, V_{CRIT}.

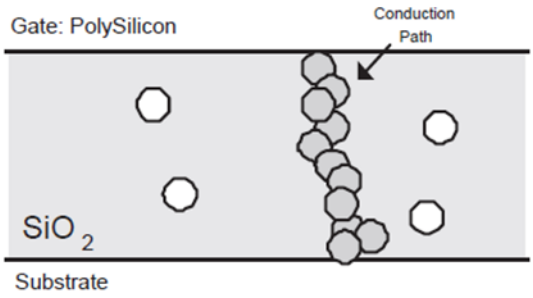


Figure 3: Soft Breakdown.

The last stage of dielectric wear out is hard breakdown (HBD), as shown in Figure 4. SBD and HBD are typically believed to have the same physical origin, related to the same kind of defects and BD triggering conditions [8].There is an apparent continuity from SBD to HBD. The energy dissipation amount determines the severity of the BD, either SBD (weaker) or HBD (stronger) [9].However, There haven't been any clear definition of BD hardness and boundary between SBD and HBD [10].

Each dielectric material has a maximum electric field it can intrinsically sustain (dielectric strength). To make a MOSFET function properly, it must operates within the insulating limit of the dielectric layer. As the stress becomes higher, the electrical field inside the gate oxide goes far beyond the insulating limitation of dialectics. More traps will be generated and thus better conduction is formed. As the gate tunneling current becomes stronger, more heat is generated, resulting more thermal damages inside the gate, which turn out as gate traps. This self-sustaining mechanism goes on and will amplifier itself, finally avalanches into HBD. Once

the avalanche happens, it happens very quickly, the transistor lost its characteristics and there is no more gate control, the damage caused is permanent and device is non-reversibly ruined.

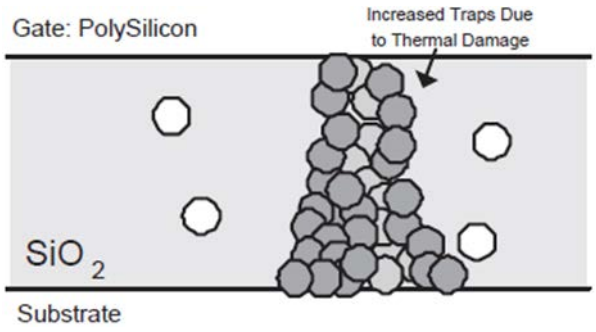


Figure 4: Hard Breakdown.

Figure 5 is a picture of after-breakdown from Emission Microscopy (EMMI).[11] Light regions are the areas of gate-oxide breakdown where the Silicon has melted. After a HBD, usually silicon in the breakdown spots melts, oxygen is released, and silicon filament is formed from gate to substrate. Typically, not only the gate oxide is ruptured after HBD, but also the Si substrate channel is severely damaged by gate oxide BD-induced thermal effect. In some cases, a direct short in the channel between source and drain is observed from TEM in HBD MOSFETs.

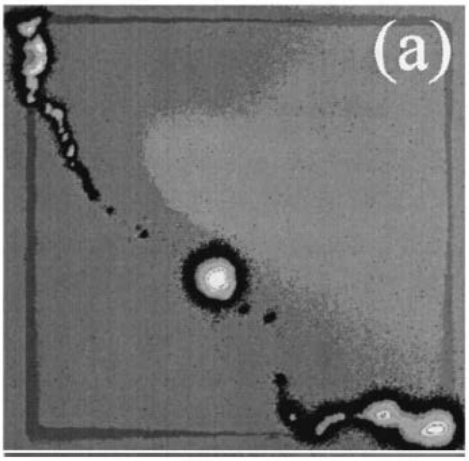


Figure 5: EMMI image of the gate after breakdown.

2.2 Hot carrier injection

2.2.1 A review of hot carrier injection mechanism

Silicon crystal breaks at the interface between silicon and gate oxide, which gives rise to many dangling bonds, as well as extra energy states. When this MOSFET is operated in saturation region, the lateral electrical field can be high because of the large source–drain voltage. As we know that the mobility of carriers depends on the lateral electrical field of the channel, the velocity of carriers will approach a saturation value, which is about 10^7cm/s when electrical field is sufficiently high. However, there exist some carriers whose instantaneous velocity and kinetic energy continues to increase, which accelerate themselves near the drain region under the influence of high lateral field. These carriers (electrons and holes) are what we called “Hot carriers”, illustrated in Figure 6. High field-induced impact ionization accelerates the generation of hot electrons and holes at the drain region of the MOS transistor [12].

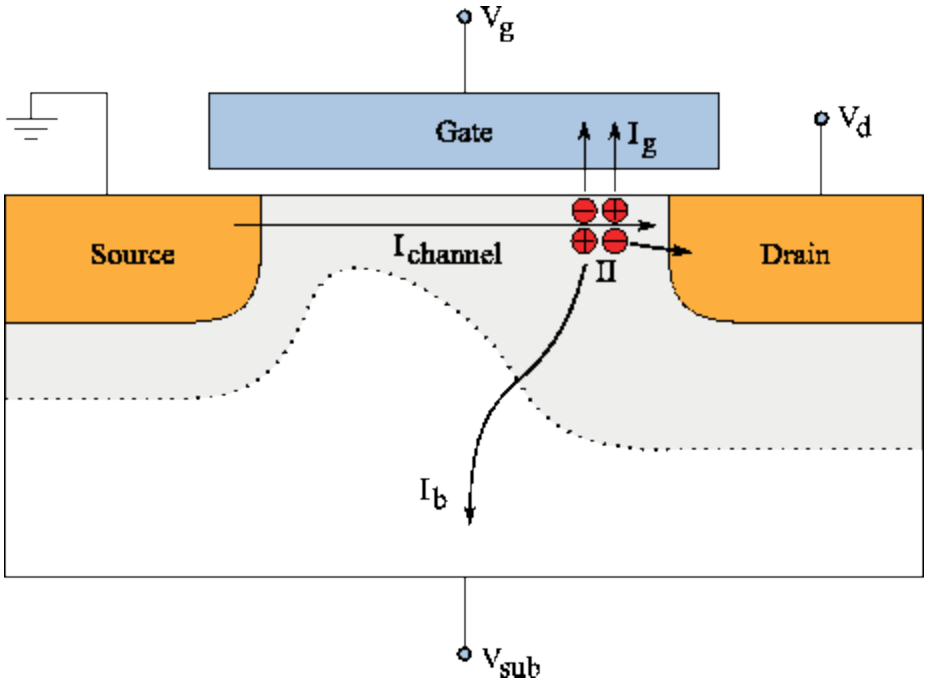


Figure 6: Hot carriers inject into gate, drain and substrate.

Hot carrier effects result from heating and subsequent carriers' tunneling into gate oxide. Localized and non-uniform interface states (N_{it}) and oxide charges (N_{ot}) near the drain junction of the transistor will be built up [13]. Traps will be generated both at the interface and in the oxide. Hot electrons may inject into the gate oxide and increase the oxide trap charge, while hot holes could generate excess interface states between the SiO_2 and Si interface. These interface traps can capture or release charge carriers from or to the channel. Electrons can be absorbed by the drain and holes by the sbustrate, causing substrate current which may further induce latch up effects. Carriers can also inject into gate oxide and flow out of the gate terminal, causing gate current. All these defects will leads to threshold voltage shift, trans-conductance degradation, drain current reduction, etc., and eventually lead to device failure.

There are four injection mechanism distinguished by the way hot carriers injection into dielectric. They are namely channel hot-electron (CHE) injection, drain avalanche hot-carrier

(DAHC) injection, substrate hot-electron/hole (SHE/SHH) injection, and secondary generated hot-electron (SGHE) injection.

CHE happens mostly when V_{GS} is equal to V_{DS}. Electrons that have gained high kinetic energy across the channel overcome the barrier between Si-SiO_2 and inject into gate at the drain end of the channel. Gate current can be measured with moderate V_{GS} and V_{DS}. When V_{GS} is not high engough, it cann't generate strong enough vertival field to draw the electrons to gate terminal.

DAHC is mainly caused if V_{DS} is too high, higher than V_{GS}, impact ionization is triggered at the drain area. Both electrons and holes are injected into gate and current flow will be smaller. Some of the carriers can inject into substrate and leads to a bulk current.

SHE/SHH is due to high bias voltage, either positive or negative. Carriers in the substrate generated either from optical generation or electrical injection from a buried p-n junction are driven into the Si-SiO_2 interface, and later on they gain high kinetic energy in the depletion region and inject into oxide finally.

SGHE originates from photon induced generation process when both V_{sub} and high V_{DS} exist. As a combined effect of DAHC and SHE/SHH, electron –hole pairs are generated due to impact ionization, photons are generated in the high field area near drain, secondary electron – hole pairs are generated all these resulted carriers of avalanch multiplication can inject into oxide, get trapped and degrade the device performance.

2.2.2 Life time determination

It is important to investigate the life time of a device in reliability study. The life time is how long it can function before it degrades to meet a certain criteria under normal operation conditions. The life time is usually determined by experiment approach when the failure

mechanism is accelerated, like higher temperature, high operation voltage or current, etc. The criteria to judge is often selected as a degradation parameter that is relevant to the damage caused. As electrical parameter variation is based on both the amount of damage and the influence the damage has on the electrical characteristics, charge pump current allows to separate these two effects and provides only the effect of real damage.

For hot carrier effects, there are many life time determination methods been proposed during the years of research, most of them are based on the exaggerated drain voltage, since this mechanism is not a strong function of temperature. One of the most famous is the lucky electron model proposed by Simon Tam, etc.[14] They analyzed three probabilities that are critical to form channel hot electron injection current. These are: a) probability of a hot electron to gain enough kinetic energy. b) Probability of keeping away from inelastic collision during transport to Si-SiO2 interface. c) Probability of suffering no collision in the oxide image-potential well. They reaffirmed a correlation between the gate current and the substrate current both by theory and experiments.

Here, we use generated interface traps ΔN_{it} to evaluate how much damage is caused, thus to derive the life time τ as when ΔN_{it} reach a certain amount. ΔN_{it} can be expressed using this equation[15]

$$\Delta N_{it}(t) = C_1[\frac{I_d}{W} * t * exp(-\frac{\emptyset_{it,e}}{q\lambda_e E_m})]^n \quad (1)$$

Where W is the width of the device, $\emptyset_{it,e}$ is the energy an electron needs to generate interface trap and λ_e is the hot electron mean free path. Electrical field E_m can be derived by a measurement of drain and substrate current, given that $\frac{I_{sub}}{I_d} = C_2 \exp(-\frac{\emptyset_i}{q\lambda_e E_m})$.$C_1$ and C_2 are coefficients, $\emptyset_i$ is the impact ionization energy.

Besides interface trapΔN_{it}, several conventional parameters have been commonly used, like electrical parameters which measure the change of electrical characteristics:$\Delta\, V_{TH}$, $\Delta\, g_m/g_m$, $\Delta\, I_d/I_d$, and physical parameters that measure the real damage of the interface: charge pump current.

2.2.3 Strategies to improve hot carrier reliability

Hot carrier effect is one of the main reasons that cause RF circuit failure. It has become a hot field of study that even over shadows gate oxide breakdown effect. It is well known that HCD is a strongly dependent on internal electric field distribution of the MOSFET. The lateral electric field near the drain junction is responsible for carrier heating and avalanche (monitored by substrate current), and the vertical electric field influences carrier injection into the gate oxide (monitored by gate current). As CMOS device dimension keep scaling down, the internal electric field distributions are changing and so is the carrier heating and injection processes. This makes the hot carrier reliability increasingly important for digital and RF circuit performance. So it's also important to know how to improve the hot electron reliability.

The first way to think about is to increase channel length, this is very straightforward since it reduce the lateral electrical field at a given source drain voltage.

Another way is through drain engineering, specifically through double diffusion of source /drain or also called lightly doped drain to reduce the E-field near the drain. The generation of hot carrier at a given supply voltage is limited by reducing the lateral electric field. In these devices the drain and depletion area is connected by a lightly doped area, so the transition near the drain is smoother.

The third way is to improve gate dialectics by using better quality materials for gate, like nitride oxides and/or oxy-nitride. This is important since it minimize the damage in the oxide for a given hot carrier injection influences.

People also suggest to move the maximum drain electric field deeper in the channel, so that the position of hot-carriers generated are away from the Si-SiO_2 interface, their chances to be injected into the oxide is reduced [16].

2.3 NBTI Effects

2.3.1 What is NBTI

Negative Bias Temperature Instability (NBTI) is a main reliability concern in modern digital and analog/RF circuit design. It is a key issue for PMOS transistors since they are mostly operated with negative gate bias condition. It causes the PMOS characteristics like threshold and linear and saturation current, gate-drain capacitance to shift from normal value due to the build-up of positive interface charges. Device scaling as well as elevated temperature will make the degradation more severe. As the thickness of gate oxide becomes smaller than 4 nm, the threshold degradation caused by NBTI effects of PMOS transistor has more effects on device lifetime, even greater than what is caused by hot carrier injection effects. What's more, since hot carrier effects only take place when there is current flow in the channel, NBTI can happen as long as there is static stress on the gate, NBTI degradation is exacerbated in most modern appliances during standby mode.[17] The degradation exhibits logarithmic dependence on time. For NMOS transistors, positive gate bias leads to the corresponding degradation mechanism called PBTI (Positive Bias Temperature Instability).

While the exact scheme of NBTI mechanism is still not yet fully discovered, it is widely accepted that interface traps are the main reason to cause the degradation. Charge pumping and

gate leakage current measurements revealed that NBTI under moderate oxide fields is purely due to interface traps N_{it} and the generated oxide traps N_{ot} can be neglected[18] . Interface traps are generated since Silicon-Hydrogen bonds break at the interface of Si-SiO2. The broken Silicon bonds become the interface Trap near the Si-SiO2 interface and the Hydrogen atoms and molecules diffuse into the gate oxide. The inverse process called annealing re-connects the interface trap and floating Hydrogen and reduces the number of interface traps as well as threshold voltage.

Time dependence of trap generation can be described by a simple power law:[19]

$$N_{it} - {N_{it}}^{0} = \frac{{N_{hb}}^{0}}{(1+vt)^{\alpha}} \quad (2)$$

N_{it} is interface trap concentration, ${N_{it}}^{0}$ and ${N_{hb}}^{0}$ are the initial concentrations of interface traps and Si-H bonds (or the concentration of hydrogen on Si bonds) respectively. The power α is stress dependent and varies between 0 and 1. $v = v_A e^{\frac{-\epsilon_A}{kT}}$ is a reaction constant in the Arrhenius approximation, ϵ_A is the Si-H activation energy, and T is the Si-H temperature.

Another model that can interpret negative bias temperature instability (NBTI) phenomena is the Reaction and Diffusion (R-D) model, as shown in Figure 7. This model describes how the Si-H bond break at the interface, electrically active interface state and a mobile, hydrogen related species are formed, subsequently, how the hydrogen species travel away from the interface into the dielectric. The reverse process: transport of a diffusing hydrogen species back to the interface and re-passivation of a Si-H dangling bond is called recovery.

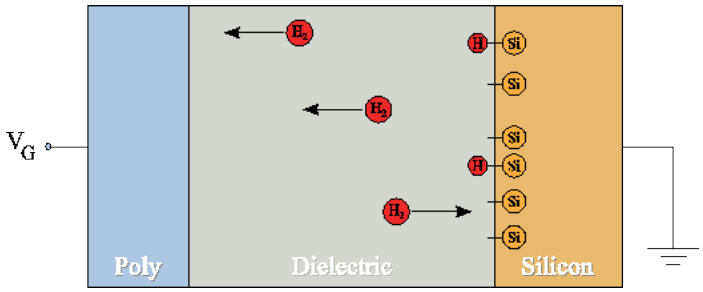


Figure 7: Illustration of Reaction and Diffusion Model.

The diffusion of hydrogen in oxide can be expressed as follows:

$$\frac{dN_H}{dx} = \frac{dN_{hb}}{dt} \qquad x = 0$$

$$\frac{dN_H}{dt} = D\frac{d^2N_H}{dx^2} \qquad 0 < x < x_p \tag{3}$$

$$D\frac{dN_H}{dx} = -k_p\left(N_H - N_H{}^0\right) \qquad x = x_p$$

N_H is a concentration of hydrogen in oxide, $D = D_0 e^{\frac{-\epsilon_H}{kT}}$ is its diffusion coefficient, $x = 0$is a coordinate of the silicon–oxide interface, $x = x_p$is the coordinate of the oxide–polysilicon gate interface (which is equal to the oxide thickness), k_p is the surface recombination velocity at the oxide–polysilicon gate interface, and $N_H{}^0$ is an equilibrium (initial) concentration of hydrogen in the oxide.

2.3.2 Lifetime estimation and ways to mitigate NBTI effects in RF circuits

There is no specific standard how to determine a device's life time over NBTI effects up to now. Many test conditions and projection methods are used in industry. However it is common to set a 10% shift criteria for a key parameter like threshold voltage, saturate drain current, or trap concentration at 125 ^oC, this is the edge to decide whether it's a working or failed device.

Two methods are used to predict the lifetime in simulation:

a): Simulate a normal-operating device using transient directly for a time long enough (like, 30 years), wait till the criteria parameter reaches the limit pre-defined (10% shift). The lifetime τ_D is determined by the point the criteria is met.

b): The second way is to calculate the de-passivation constants ratio between stressed and unstressed conditions v^{stress}/v, then apply extrapolation to get the degradation. In Quasi-stationary, if the previous transient statement result of the device lifetime is ${\tau_D}^{stress}$, device life time can then be predicted as $\tau_D = (v^{stress}/v){\tau_D}^{stress}$.

As previously mentioned, negative gate biasing and high temperature will accelerate the aging of the device due to NBTI effect, so the most straightforward way to extend the device life time is to operate the device at lower temperature with less negative gate biasing on PMOS transistor.

At the same time, there are many solutions proposed for robust design aspect: adjust V_{dd} and V_{th} to compensate for NBTI degradation over the years of usage. It is found that there exist an optimum V_{dd} value to reach the minimum performance of degradation.[20] Also traditionally there is gate sizing technique to allow for enough product reliability in sacrifice of more design overhead, this can be realized by oversizing of the entire gate, especially oversizing PMOS. By properly tuning the duty cycle of the circuit, it gives more recovery time on the gate compared to the stress time, thus less V_{th} shift can be expected. [21]

2.4 Variability issues

As the characteristic dimensions of device becomes smaller and smaller, it becomes harder and harder to precisely control the physical dimensions and dopant levels during the fabrication process. As a result, these growing uncertainties lead to more and more statistic variations in circuit performance and behaviors from designed circuit. Traditionally, designers tend to think in a deterministic way, while with these variability issues become too big to ignore, designers got more problem to solve. As shown in Figure 8, initially, process variation has been treated mainly as die to die variation, that is the difference originated from different die environments, but devices from the same die share the same properties. With the device size shrinks, intra-die variations have become the main concern for design since it will cause local mismatch even if chips are cut from the same die.

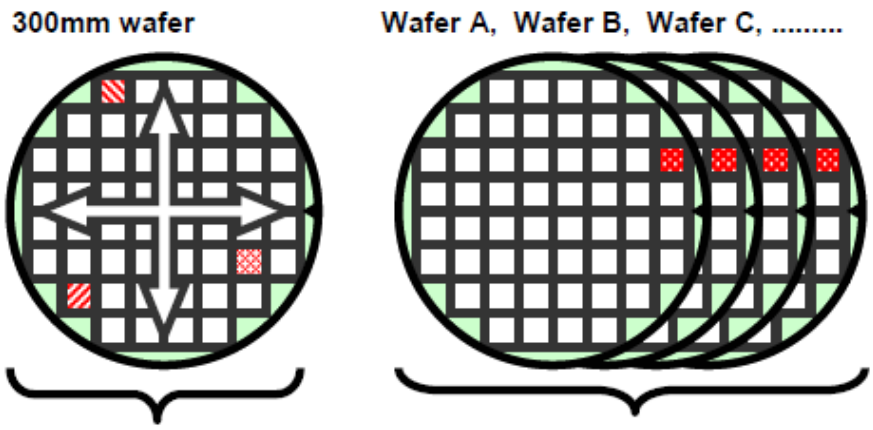


Figure 8: Variability issue inside and among dies.

It is generally believed that three sources contribute to process variations, namely: random dopant fluctuation (RDF), line edge roughness (LER), and poly gate granularity (PGG). RDF has been the main contributor to process variation for a long period. Shown in Figure 9, within each single device, there are many regions needed to be doped to different levels

according to the specific device design. However, as shown in Figure 10, with technology node becomes smaller and smaller, the average number of dopant atoms become less and less, giving more challenges to process control[22]. At the same time, certain amount of dopant fluctuation will have more effects on device properties thus circuit performances.

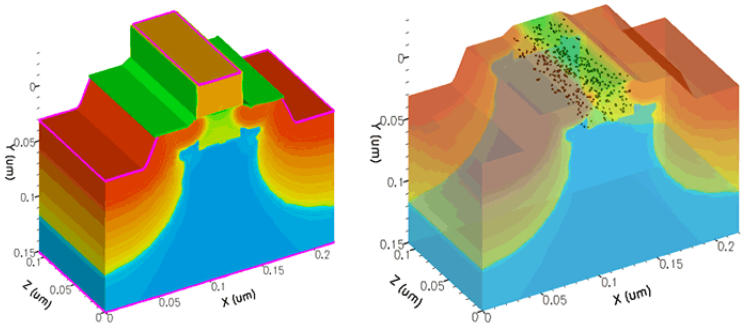


Figure 9: Different doping levels inside device.

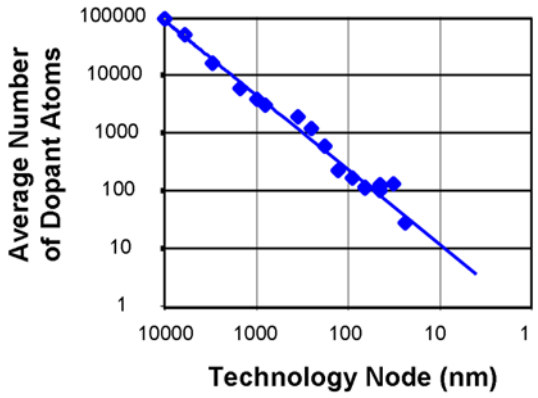


Figure 10: Average Number of dopant atoms with technology nodes.

Line edge roughness is another source of process variations and is illustrated in Figure 11, it is defined a random deviation of line edges from gate definition, it is mostly dependent on poly gate patterning. LER is associated with the variations in sub-threshold current as well as

threshold voltage of a device. Fabrication process like photo-resist (PR) type, thickness, substrate reflectivity, image contrast, etching type and condition will greatly affect the amount of LER, and improvement in such factors will minimize process variations caused by LER. LER does not scale with line width easily.

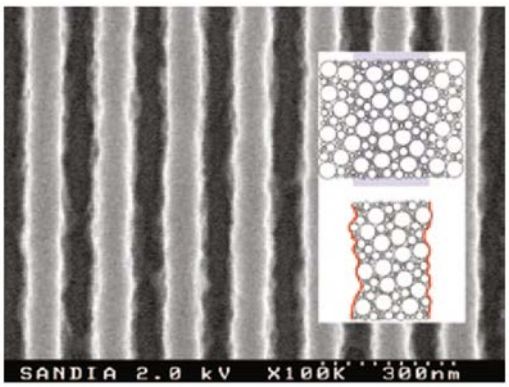


Figure 11: Illustration of line edge roughness.

Poly gate granularity (PGG) is another important source of variability, illustrated in Figure 12. One direct reason is gate dielectric thickness variations from the defined value. There are other reasons caused by granularity of matter: Faster diffusion along the grain boundaries lead to doping non-uniformity. Besides these two reasons, as a results of granularity of poly gate, the density of defect states are high between the grain boundaries, which will cause Fermi level pinning at these boundaries. Fermi level pinning will then cause surface potential fluctuation within MOSFET channel, thus results in variations in threshold voltage and current characteristics.[23]

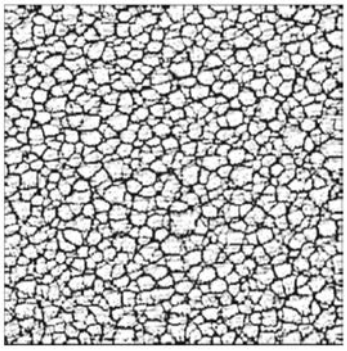


Figure 12: Granularity of poly gate.

Figure 13 compares the channel length dependence of σV_T introduced by these three sources according to ITRS.[24] The average size of the poly silicon grains was kept at 40 nm for all channel length that will keep PGG constant. One can see that σV_T change becomes larger and larger with device size shrinks in three cases. The line with diamond shows the effect caused by random discrete dopant only. While the one with square shows three with LER scales down according to ITRS. The one with triangle shows the total effects if LER is constant, which is 4nm. In the last case σV_T increases rapidly with reduced L since LER becomes the dominant source of variability for shorter channel length.

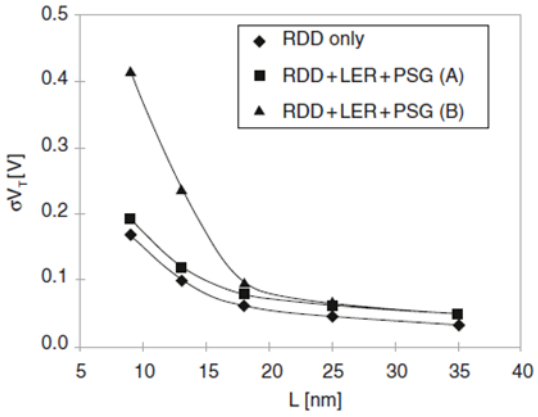


Figure 13: Channel length dependence of σV_T, ITRS.

2.5 Second Order and Short Channel effects

In analyzing the behavior of a MOSFET, most likely we treat it as ideal for simplicity. In fact, there are many essential effects that MOSFET suffers from during actual MOSFET operation. Among which the most common ones are second order effects and short channel effects.

2.5.1 Second order effects

Second order effects include body effect, channel length modulation effect and sub-threshold conduction. Body effects take place when the body terminal has the different electric potential with the source terminal, the threshold voltage will shift from the normal value. If a negative body bias is applied to a NMOS, a higher gate supply is required to compensate the holes that were drawn by a negative body bias. Two equations are cited here to describe the relationship:

$$V_{TH} = V_{TH0} + \gamma(\sqrt{|2\emptyset_F + V_{SB}|} - \sqrt{|2\emptyset_F|}), \quad (4)$$

$$\text{And, } V_{TH0} = \emptyset_{\square S} + 2\emptyset_F + \frac{Q_{dep}}{C_{ox}}. \quad (5)$$

V_{TH0} is defined as the gate voltage at which the interface turns from the depletion region to inversion region, that is the interface is "as much n-type as the substrate is p-type". $\emptyset_{MS}$ is the difference of work function between the gate and the substrate, $\emptyset_F = \frac{kT}{q}\ln(\frac{N_{sub}}{n_i})$, N_{sub} is the doping concentration of the substrate, n_i is the intrinsic carrier(either hole or electron) concentration, for Silicon at 300K, it is $1.5\times10^{10}/cm^3$. $Q_{dep} = \sqrt{4q\epsilon_{si}|\emptyset_F|N_{sub}}$, is the charge in the depletion region and ϵ_{si} denotes the dielectric constant of silicon. C_{ox} is the gate oxide capacitance per unit area. (Another way to define V_{TH0} is to plot the inverse on-resistance of the device as a function of V_{GS} and extrapolate it to zero, as $R_{on}^{-1} = \mu C_{ox}\frac{W}{L}(V_{GS} - V_{TH0})$ the V_{GS}

value is defined as V_{TH0}) $\gamma = \frac{\sqrt{2q\epsilon_{s\square}N_{sub}}}{C_{ox}}$ is the body effects coefficient lies in the range of 0.3 to 0.4 $V^{1/2}$. $V_{SB}=V_S-V_B$ is the source bulk potential difference.

The Body terminal can be utilized as a second gate in terms of control the drain current. That is for a NMOS, when Body bias is more positive, the threshold voltage will be lowered, thus increase the drain current.

Channel length modulation is triggered when "pinch off" is started as the drain source voltage is greater than over-drive voltage. The effective inverted channel length becomes shorter and the drain current is increased by a factor of λV_{DS} in saturation region. As a result, the role for a MOSFET as a current source in saturation region is no more ideal. The effect of channel-length modulation is less for a long-channel MOSFET than for a short-channel MOSFET. λ is larger with device size scaling down, that is the slope of I_D vs. V_{DS} curve is proportional to $1/L^2$. The effect of channel-length modulation or DIBL (which cause I_D to increase linearly with V_{DS}) is modeled by the transistor output resistance, r_o as shown in Figure 14.

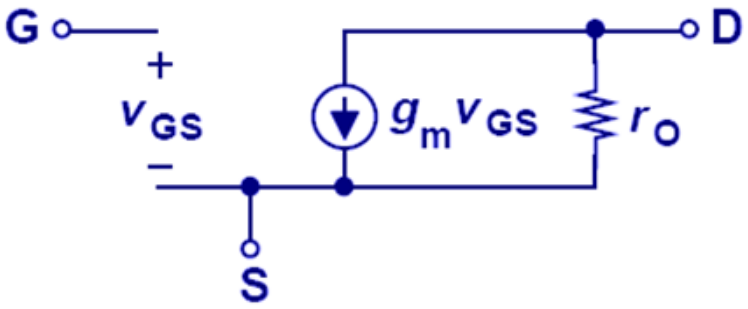


Figure 14: Small signal model for MOSFET in saturation region.

Sub-threshold conduction is another non-ideal effect in MOSFET operation. That is the switch cannot shut down thoroughly, drain current remains noticeable even when gate voltage is below threshold voltage. This is called sub-threshold region or week inversion region. The drain

current is exponentially dependent on V_{GS}, for V_{DS} greater than 0.2 V, $I_D = I_0 \exp \frac{qV_{GS}}{kT\zeta}$, and ζ is non-ideal factor larger than 1, I_0 is a process dependent constant. At room temperature with typical value of ζ, I_D drops by one decade when V_{GS} drop by 80 mV. This is an obvious obstacle for low power circuit design, especially for large memories.

2.5.2 Short channel effects

Short channel effects are becoming more and more severe and drawing increasing attract as device keeps scaling down. Ideally the scaling should follow these three rules to keep the electric field remain constant. A) all the vertical and lateral dimensions reduce by α (Source /Drain junctions' depth is hard to reduce). B) threshold and supply voltage decrease by α (Both turn out to be hard to scale) C) all doping levels increase by α .

In practice, since a)power supply is not scaled down proportionally and electrical field increased, b) Threshold voltage is not scaled easily since sub-threshold slope is not saclable, c) mobility decreases as increased doping level in substrate scaling down d) The built in potential $\emptyset_B = \frac{kT}{q} \ln(N_A N_D / {n_i}^2)$ increases with N_A and N_D. Thus the total width of depletion region $W_d = \sqrt{\frac{2\epsilon_{si}}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(\emptyset_B + V_R)}$ as well as the depletion region capacitance $C_d = \sqrt{\frac{\epsilon_{si} q N_{sub}}{4\emptyset_B}}$ is not scaled by α. e) the depth of source/ drain junction can't scale easily. Technology scaling has followed a mixture of constant field and constant voltage trends, inviting more reliability issues and design challenges.

Short channel effects find their expression mainly on threshold voltage variation, mobility degradation due to vertical field strengthen, velocity saturation, drain output impedance variation, and hot carrier effects which was described in this chapter.

Sub-threshold slope S means how much lower V_{GS} needs to be to make the drain current becomes ten times less. A small S is desirable to scale down threshold voltage V_{TH}. S is defined as $S = 2.3\,\frac{kT}{q}\left(1 + \frac{C_d}{C_{ox}}\right)\,V/dec$, and its magnitude is relatively constant. This sub-threshold behavior makes V_{TH} hard to scale.

V_{TH} has negative temperature dependence, about -1mV/^{o}K. That brings threshold voltage variation when the operating environment is unstable. Research also proved that longer channel length has higher threshold voltage since the effects of charge sharing in less evident in longer devices, shown in Figure 15. And since channel length cannot be accurately controlled during fabrication, causing another uncertainty of threshold voltage. DIBL (Drain Induced Barrier Lowering) is also related to threshold voltage variation. The lift in drain voltage helps the gate voltage to increase the potential at the interface and draw carriers from source, the barrier to the flow of charge and threshold voltage is lower as a result.

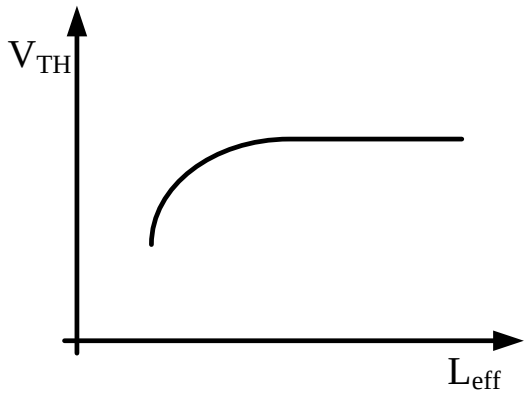


Figure 15: Threshold voltage roll off.

Small-geometry devices suffer from more severe mobility degradation as gate oxide – channel electrical field becomes higher. The charge carriers are confined to a narrower region and more carrier scattering leads to lower mobility rate. An empirical equation to model this

effect is $\mu_{eff} = \frac{\mu_0}{1+\theta(V_{GS}-V_{TH})}$, here, μ_0 is the low-field mobility and θ is a fitting parameter roughly equal to $10^{-7}/t_{ox}$ V^{-1},note here that θ is inverse proportional to t_{ox} (electrical field in oxide is larger, more scattering) and makes the degradation even stronger. The results of mobility degradation include decrease in current capability and trans-conductance and higher even harmonics in the drain current and more non-linearity.

The traveling speed of carriers is determined by the lateral electric field. While increase with lateral electric field, it starts to saturate at around 10^5m/s when the electric field is sufficiently high, like 10^6V/m. The carrier velocity no longer increases because the carriers lose energy through increased levels of interaction with the lattice. The saturated drain current can be expressed in this equation: $I_D = v_{sat}WC_{ox}(V_{GS} - V_{TH})$, and is linearly proportional to over-drive voltage and does not depend on length. That is if V_{GS} is increased by a certain amount, the drain current increases less compared to without saturation effect. The consequence is that current saturation occurs before pinch-off and lower trans-conductance as predicted by the square law.

Drain output impedance can be approximated as [25]

$$r_o = \frac{2L}{1-\frac{\Delta L}{L}}\frac{1}{I_D}\sqrt{\frac{qN_B}{2\epsilon_{si}}(V_{DS} - V_{DS,sat})} \quad (6)$$

$V_{DS,sat}$ is the drain-source voltage at the onset of pinch-off.Also, $r_o \equiv \frac{\partial V_{DS}}{\partial I_D} \approx \frac{1}{\lambda I_D}$. The trend of r_o with V_{DS} is displayed in Figure 16. As V_{DS} increases, the channel length modulation effects dominate at the first phase. Since I_D increase less for the same amount of V_{DS} increase, output impedance will increase. At the second phase, DIBL (Drain Induced Barrier Lowering), which makes output impedance smaller, becomes significant as V_{DS} increases further. These two effects co-exist and the output impedance curve stays flat in this phase. As V_{DS} increment continues into the third phase, impact ionization near the drain area generates large current

flowing from drain to substrate, the output impedance drops dramatically, this overcomes the aforementioned two effects, and the curve droops finally.

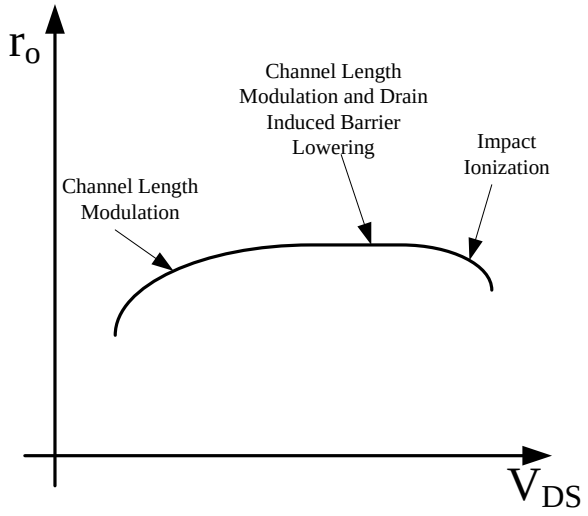


Figure 16: Overall variation of Ro vs. $V_{DS.}$

CHAPTER THREE: EXPERIMENTAL VERIFICATION ON A CASODE CLASS E PA

3.1 PA overview

3.1.1 Classifications

By definition, a power amplifier is an amplifier which is used to greatly amplify the input signal in power, that is, the output power of a PA should be much higher than the input power. Typically, it is used at the final amplification stage of a transmission system. Power amplifiers distinguished themselves from small signal amplifiers in a way that signal currents of PAs are highly dependent on DC bias level on the gate, thus serious distortion/low linearity is expected.

Power amplifiers are traditionally categorized in an alphabetical order according to their circuit configurations and operation mode (current or switch). Like Class A, B, C, AB, D, E, F, different classes are distinguished with tradeoff between linearity and power efficiency. There is no "better" or "worse" class than any other class, people can choose the type of operation by the specific design spec in different amplifying circuit.

Here are the advantages and disadvantages of each class of PAs:

Class A PA: Usually there is one single transistor being used and it is turned on the whole cycle of the input signal, there is current flowing also. Thus the output signal faithfully follows the input signal but the power efficiency can be less than 40%. It is widely used in high quality audio instruments and rarely used in high power designs.

Class B PA: The output transistor is biased so that it is on for half cycle. Typically this class has two different types of transistors. Each transistor is turned on for half cycle and the output signal is combined by these two outputs. Class B operation has no DC bias voltage; instead the transistor only conducts when the input signal is greater than the transistor’s threshold

voltage. When the lower part of the input waveform is below this threshold voltage, which is during the time when both transistors are "OFF", the output signal will not be reproduced accurately, resulting in a distorted area of the output waveform. So there is a small part of the output waveform around the zero voltage cross over point distorted. This causes a problem named 'Crossover distortion'. Class B PA is twice as efficient as class A amplifiers with a maximum theoretical efficiency of about 70% because the amplifying device only conducts (and uses power) for half of the input signal. They are mostly used in low cost designs and where the signal quality requirements are not so high.

Class AB PA has an efficiency rating between that of Class A and Class B but poorer signal reproduction than class A amplifiers. They also use pairs of transistors, but with both of them being biased slightly ON so that the crossover distortion (associated with Class B amps) is largely eliminated. They are called Class AB since they combine the merits of both Class A and Class B. If the drain current of a Class AB, $i_D = I_{DC} + i_{rf}\cos\omega_0 t$, then the conduction angle $2\emptyset = 2*\cos(-\frac{I_{DC}}{i_{rf}})$. Class AB is probably the most common amplifier class currently used in home stereo and similar amplifiers.

Class C PA, the transistor conducts for less than half cycle and thus it is the most efficient amplifier class as only a very small portion of the input signal is amplified therefore the output signal bears very little resemblance to the input signal. Class C amplifiers have the worst signal reproduction. It is not used for audio systems but widely used in RF circuits.

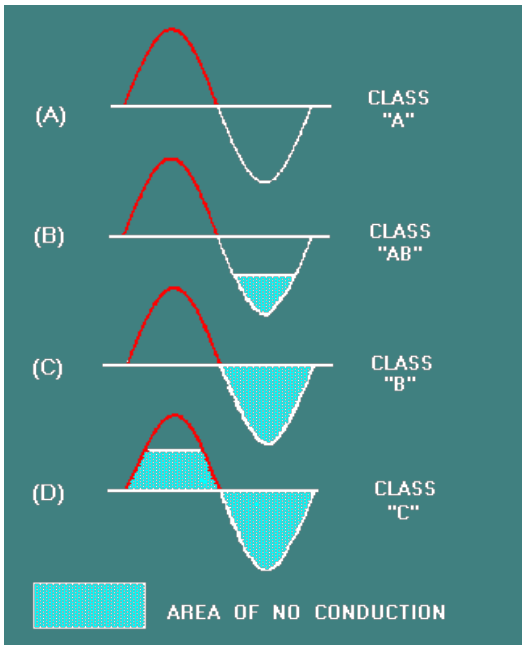


Figure 17: Summary of current mode PA operation.

Figure 17 summarized the characteristics of four current mode power amplifiers. To maintain the active device working as a voltage controlled current source, drain voltage must be precisely controlled so that it does not enter triode region. This requirement makes efficiency highly dependent on supply voltage, component values, and Q of matching network.[26] To summarize these four classes in a more clear way, we get table 1 as shown below:

Table 1: Summary of four current mode Pas.

Class	A	B	C	AB
Conduction Angle	360^{o}	180^{o}	Less than 90^{o}	180 to 360^{o}
Position of the Q-point	Centre Point of the Load Line	Exactly on the X-axis	Below the X-axis	In between the X-axis and the Centre Load Line
Overall Efficiency	Poor, 25 to 30%	Better, 70 to 80%	Higher than 80%	Better than A but less than B 50 to 70%
Signal Distortion	None if Correctly Biased	At the X-axis Crossover Point	Large Amounts	Small Amounts

On the other hand, there are switch mode power amplifiers, like Class D, Class E, Class F.

Class D amplifier is one kind that use the transistor as a switch instead of current source. Ideally there is either zero voltage across or zero current through the main transistor, so that its power efficiency can reach 100%

Among all these switch mode Pas, class E PAs are nonlinear amplifiers with high power efficiency while delivering full power. As switching power amplifier, the voltage across and the current through the switch do not overlap near the transition region, and the slope when the transistor is turned on is zero, therefore class E PA serves as a good candidate for low-cost, high integration portable communication systems such as cell phones, wireless local area networks, wireless sensor networks, global positioning systems, and Bluetooth applications. As a tradeoff of its high efficiency, output harmonic can be high, thus proper design of high Q output network is required and filtering circuit can be added at the expense of power loss. In the off state, the voltage across the main transistor can be very high, approximately 3.6Vdd-2.56 Vs, where Vs is the minimum voltage across the transistor [27]. This property makes class E PA vulnerable to breakdown.

Based on the idea of harmonic termination, Class F power amplifier is another nonlinear amplifier extended from Class A stage. The load network is properly selected so that it provides high termination impedance at the second and the third harmonics. Voltage waveform across the transistor reduces the power loss since it has sharper edge than a sinusoid, a semi-rectangular shape. With a half-wave rectified sinusoid drain current, Class F can reach a peak efficiency of 88% for third harmonic peaking and 85% for second harmonic peaking. [28]

3.2 Class E PA Reliability Issues

CMOS technology for radio frequency applications has its advantages in integrated, low-cost RF power amplifiers (PAs) for wireless communications ICs. Class E PA has become popular due to its high power efficiency[29], and therefore widely applied in low-cost, high integration portable communication systems such as cell phones, wireless local area networks, wireless sensor networks, global positioning systems, and Bluetooth applications. As a result of aggressive scaling in device dimensions for improving speed and functionality, CMOS transistors have progressed into the deep sub-micrometer to nanometer regime, leading to major reliability issues including gate oxide breakdown [30] and channel hot electron degradation [31,32].

Class E power amplifier with a shunt capacitor was first introduced by N.O. Sokal and A.D. Sokal [27] and was examined by Rabb in an analysis of idealized operation [33].The voltage through the switch transistor and current through are trimmed by the LC tuning network such that they do not overlap at the turn on point, ideally its power efficiency can reach 100%. At the same time, Class E PA is very vulnerable to oxide stress because its drain source potential can approach more than three times of supply voltage easily. To ensure the reliability of class E PA operation, V_{DD} is typically selected conservatively with the cost of reduced output power and power efficiency.

Cascode topology is adopted in class E PA to divide the output voltage and decrease gate oxide stress effect [34,35, 36]. The cascode topology over-performs non-cascode structure due to reduced drain-gate voltage stress on the output transistor. In addition, a thick oxide of the cascode transistor may be used to alleviate oxide stress at the expense of reduced output voltage swing [37]. Note that existing publications [34-37] on class E PA reliability issues focused on gate oxide stress, not channel hot electron injection. This motivates us to evaluate hot electron effect

on class E power amplifier degradations using experiments and mixed-mode device and circuit simulation.

In this work, a cascode class E power amplifier is designed at 5.2 GHz using ADS simulation [38].The circuit performances are examined for post layout simulation. Measured PA performances before and after RF stress are analyzed. RF stress experiments and class E PA performance degradations with high input power and increased V_{DD} stress are presented. The reliability of cascode class E PA subjected to hot electron effect is discussed through the examination of impact ionization and lattice heating using the mixed-mode device and circuit experiment analysis in the following sections.

3.3 Design of Class E PA

A cascode class E PA is designed for fabrication to evaluate the class E PA reliability by experiments, Figure 18 shows the circuit schematics of a cascode class E PA.

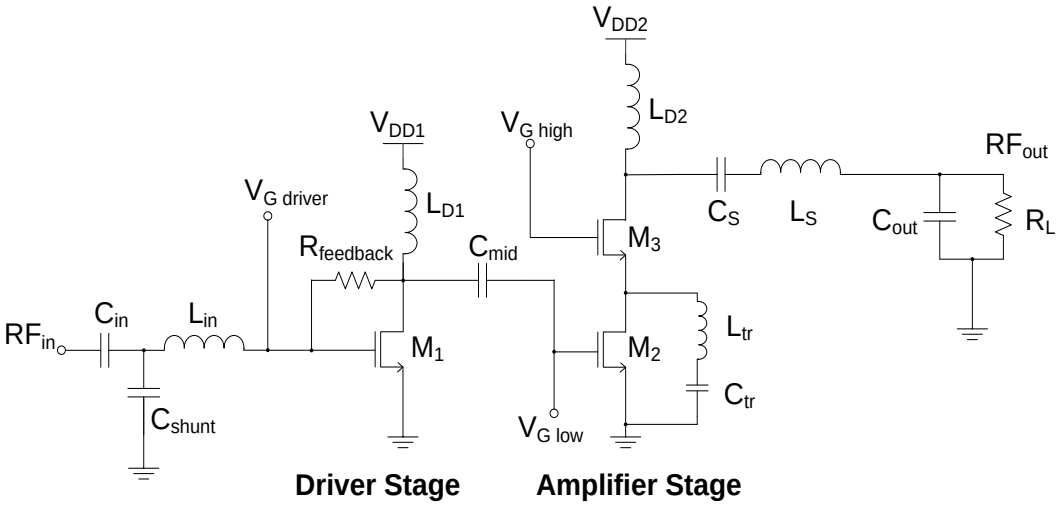


Figure 18: Schematic of a cascode class E power amplifier.

The class E PA is designed at 5.2 GHz using TSMC 0.18 µm mixed-signal CMOS technology and is evaluated in ADS simulation. The DC supply voltage V_{DD1} for the driver stage is set at 1 V. The supply voltage V_{DD2} for the main amplifier is selected to be at 2.4 V. To reduce power consumption, the gate of M1 is biased at 0.1 V (class-C mode of biasing). The gate DC voltages of M2 and M3 are at 0.7 V and 1 V, respectively. Multi-finger transistors with n-channel length of 0.18 µm are used. The driver transistor M1 has the channel width of 256 µm. The main transistor M2 and cascode transistor M3 have the channel width of 512 µm and 544 µm, respectively. The feedback resistance $R_{feedback}$ is 600 Ω.The spiral inductor and capacitor values used in this design are L_{in} = 3.61 nH, L_{D1} = 1.47 nH, L_{D2} = 4.56 nH, L_{tr} = 0.27 nH, L_S = 3.61 nH, C_{in} = 398 fF, C_{shunt} = 1.79 pF, C_{mid} = 1.68 pF, C_{tr} = 804 fF, C_S = 398 fF, and C_{OUT} = 35.6 fF.

The cascode class E power amplifier was laid out using Cadence Virtuoso software [39], followed by successful Calibre DRC for design rule checking and LVS for layout versus schematic verification. The finished layout is displayed in Figure 19 with DC biasing marked.

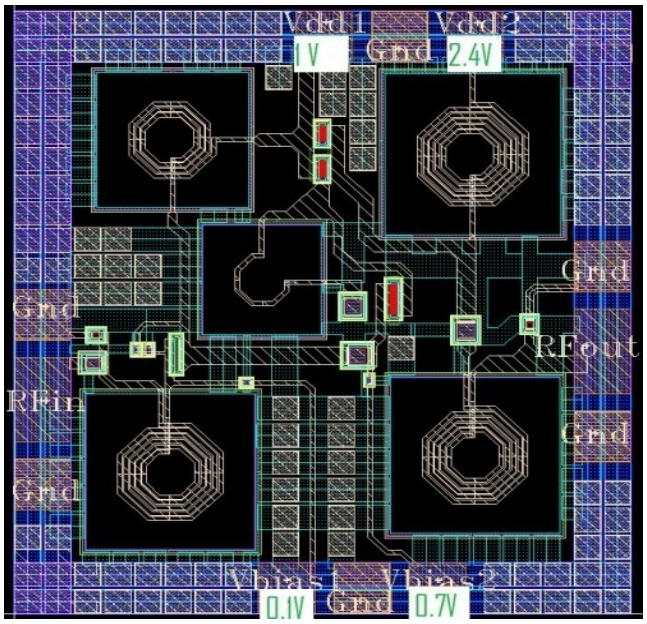


Figure 19: Class E PA Layout.

Afterwards, the layout parasitic effects were extracted using ADS Momentum (EM) simulation. Post layout simulation results of output power and power-added efficiency (η_{add}) as a function of input power are shown in Figure 20. Noticeable yet reasonable degradation is observed compared to pre-layout simulation results.

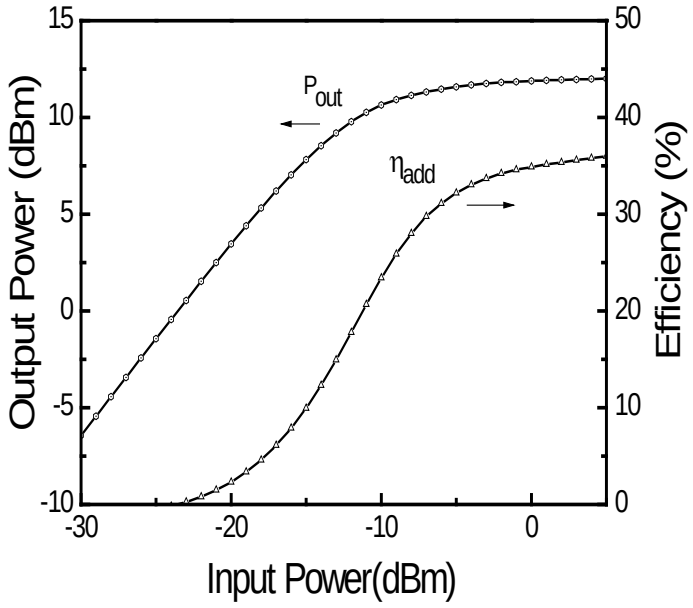


Figure 20: Output power and power-added efficiency versus input power after post layout simulation.

To examine the electrical stress effect on this cascode class E design, the gate-source voltage and drain-source voltage as a function of time for the cascode transistor and main transistor are shown in Figure 21 and Figure 22. As seen in these two figures the cascode transistor bears more voltage stress at the drain of M3 than that of the main transistor M2. The cascode transistor could suffer hot electron effect when both gate-source voltage and drain-source voltage are high during switching point (see Figure 22).

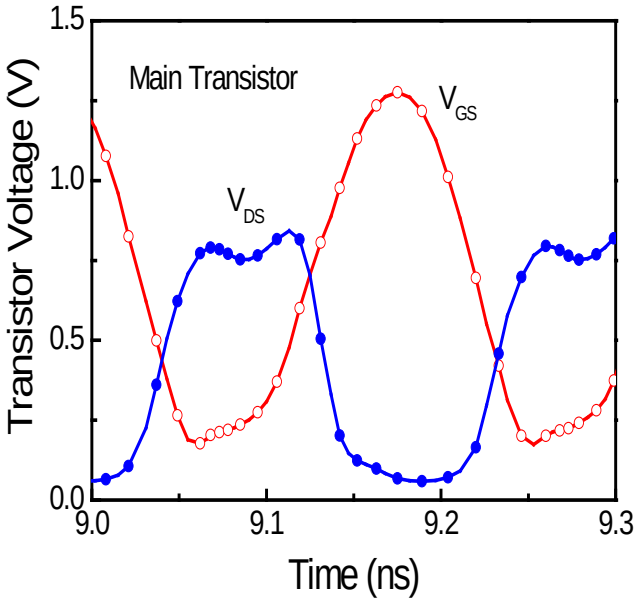


Figure 21: Simulated gate-source and drain-source voltage of main transistor. Pin = 0 dBm and VDD2 = 2.4 V.

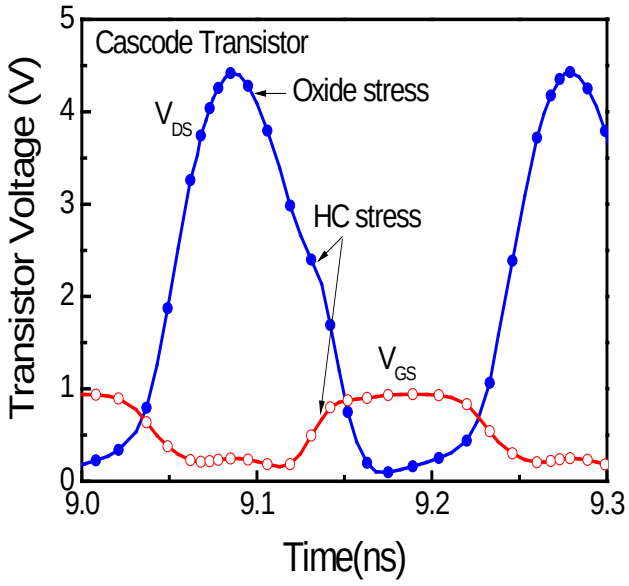


Figure 22: Simulated gate-source and drain-source voltage of cascode transistor. Pin = 0 dBm and VDD2 = 2.4 V.

3.4 RF stress experiments

A silicon chip of the designed PA was fabricated using TSMC 0.18 µm mixed-signal CMOS technology. The silicon chip is shown in Figure 23 and its size is 820×887 µm2. One can distinguish spiral inductors, capacitors, transistors, GSG RF input pads and output pads, DC supply voltage pads, and gate bias pads in this figure.

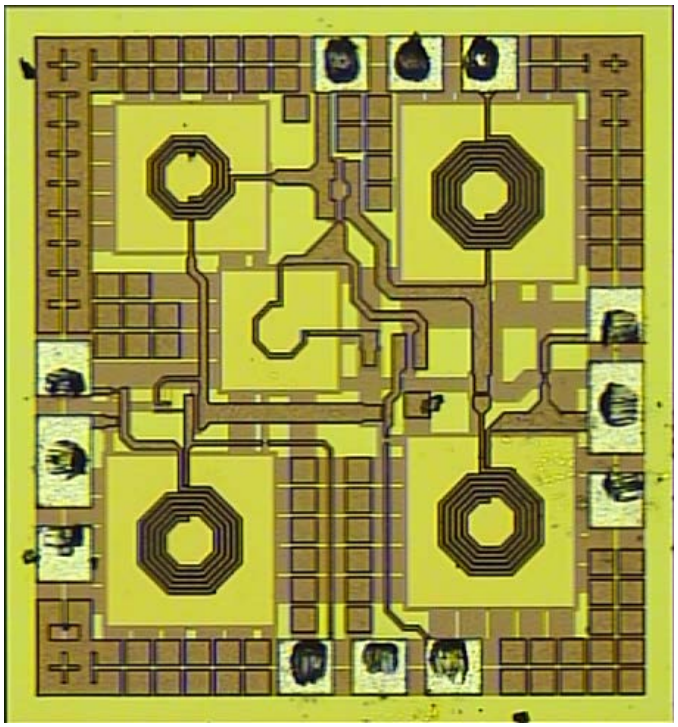


Figure 23: Chip view of the cascode class E power amplifier used for RF stress and measurement.

The PA's performances before and after RF stress are measured freshly at room temperature. DC biases of V_{G1} = 0.1 V, V_{G2} = 0.7 V, V_{DD1} = 1 V, and V_{DD2} = 2.4 V were used for the circuit at the normal operation. The PA was applied with continuous stress for 10 hours with an RF input power of 0 dBm and different V_{DD2} stress level at 3.5, 4.0, or 4.5 V. After

continuous RF and elevated DC stresses, the RF parameters were measured again to compare with the experimental data obtained at the fresh circuit condition.

Figure 24 shows the measured small-signal gain S21 versus frequency at different stress conditions. The solid line represents the fresh circuit result, the squares represent the measured result after 10 hours of RF stress at V_{DD2} = 4.5 V, the triangles represent the data after 10 hours of RF stress at V_{DD2} = 4.0 V, the circles represent the PA's experimental data after 10 hours of RF stress at V_{DD2} = 3.5 V. As seen here the S21 degradation is larger over a wide range of frequencies when higher V_{DD2} brings elevated stress.

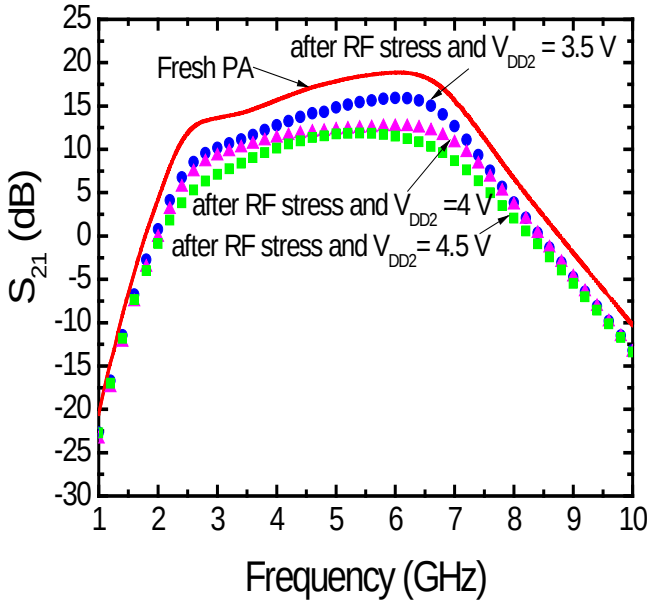


Figure 24: Measured S21 before and after RF stress. During the RF stress Pin is at 0 dBm and VDD2 was kept at 3.5, 4, or 4.5 V.

At 5.2 GHz the measured output power and power gain are plotted in Figure. 25. The output power increases with input power and saturates at high input power, thus reduces the power gain at high input power. Both the output power and power gain decrease after RF stress, especially when V_{DD2} is increased. The degradations of RF circuit performances are attributed to

gate oxide stress as well as hot electron effects on the output transistor. Please refer to Sec. 3.4. for more detailed physical explanations.

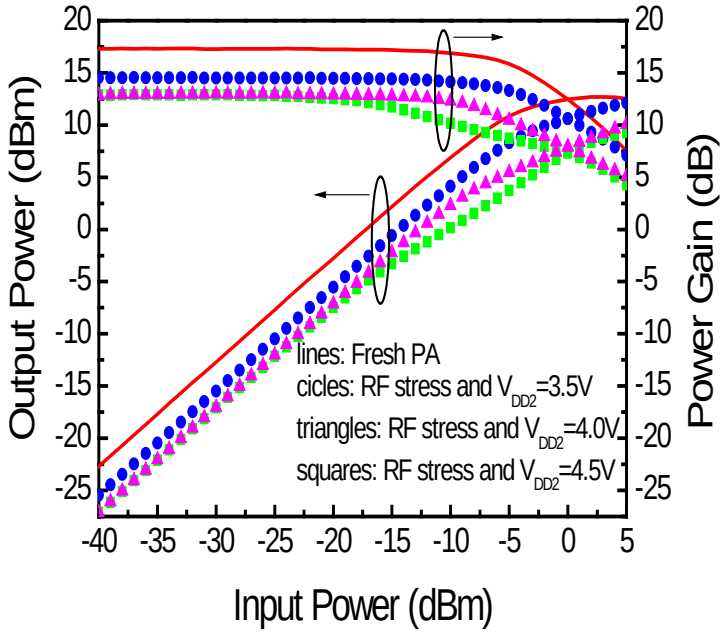


Figure 25: Measured output power and power gain versus input power before and after RF stress. During this RF stress Pin is at 0 dBm and V_{DD2} was kept at 3.5, 4, or 4.5 V.

The measured power-added efficiency is illustrated in Figure 26. Power-added efficiency increases with input power, reaches saturation, and then decreases with input power due to reduced output power and increased DC power dissipation when input power goes high. The power-added efficiency is defined as (RF output power - RF input power)/total DC power dissipation where both the power stage's and driver stage's power consumption are counted in. Note that the power-added efficiency is lower than the drain efficiency since additional power dissipation from the driver stage. At 5.2 GHz the peak power-added efficiency of the fresh PA approaches 25% (a value slightly lower than expected due to layout parasitic effect and additional DC power dissipation in the driver stage). The peak power efficiency decreases significantly after RF stress, especially when the V_{DD2} is higher.

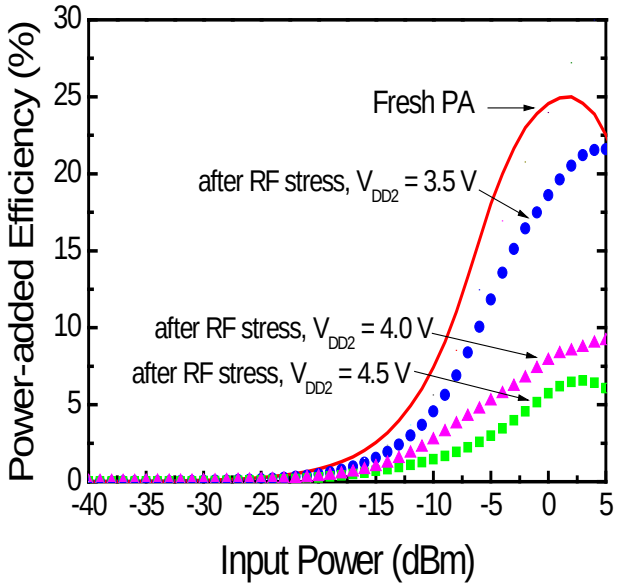


Figure 26: Measured power-added efficiency versus input power before and after RF stress @ 5.2 GHz. During this RF stress Pin is at 0 dBm and V_{DD2} was kept at 3.5, 4, or 4.5 V.

Table 2 lists the small-signal gain S21 at 5.2 GHz, output power p_o when the input power is 0 dBm, power gain (p_o/p_i) at the input power of -20 dBm, and maximum power-added efficiency before and after RF stress. Their normalized parameter shifts such as $\Delta S21/S21$, $\Delta p_o/p_o$, $\Delta(p_o/p_i)/(p_o/p_i)$, $\Delta\eta_{add}/\eta_{add}$ $\times 100\%$ from the fresh condition are also shown in Table 2.

Table 2: RF parameter degradations.

RF parameters	S_{21} @5.2 GHz	p_o @ p_i=0dBm	gain @-20dBm	peak η_{add}
Fresh	18.2 dB	12.5 dBm	17.3 dB	25%
After RF stress[1]	15.2 dB	10.6 dBm	14.5 dB	21.6%
After RF stress[2]	12.3 dB	7.9 dBm	12.9 dB	9.1%
After RF stress[3]	11.9 dB	7.3 dBm	12.5 dB	6.6%
Change [1]	-16.5%	-15.2%	-16.2%	-13.6%
Change [2]	-32.4%	-36.8%	-25.4%	-63.6%
Change [3]	-34.6%	-41.6%	-27.7%	-73.6%

[1] RF stress at p_i = 0 dBm and V_{DD2} = 3.5V for 10 hours
[2] RF stress at p_i = 0 dBm and V_{DD2} = 4.0V for 10 hours
[3] RF stress at p_i = 0 dBm and V_{DD2} = 4.5V for 10 hours

3.5 Mixed mode simulation

To get more physical insight of hot electron effect in the cascode PA, the amplifier stage of the cascode power amplifier in Figure 18 is simulated in Sentaurus TCAD [40]. Note that the mixed-mode device and circuit simulation reflects what happens to the real circuit, thus provides the examination of device physical insight under the practical circuit operation condition. Impact ionization and self-heating are monitored specifically. Figure 27 shows impact ionization rates for the cascode transistor and main transistor at different transient points, supply voltage V_{DD2} is set at 3.5 V. As seen in Figure 27 the cascode transistor has much higher impact ionization rates than those of the main transistor due to higher electric field at the drain. Larger drain-source voltage also makes impact ionization rates at the peak of output voltage transient (left figure) higher than those during the output switching (right figure) as seen in Figure 27 High impact ionization rates near the drain of MOS transistor (~ 6.3×1026 /cm^3/s) suggest that a large amount of hot electrons may inject into the gate of the MOSFET. Some hot electrons could be trapped within the oxide without reaching the gate contact. As time goes by, the accumulated trapped electron charges increases the threshold voltage of the MOSFET. In addition, the interfacial layer between the SiO2 and Si interface near the drain region may be damaged or degraded since more interface states are generated by the channel hot holes. Thus, the effective electron mobility of the MOSFET decreases. Consequently, the drain current and transconductance of the MOSFET decrease. Finally the output power and efficiency of the power amplifier is reduced due to the reduction in drain current as demonstrated by the experimental data in Figure 25 and Figure 26.

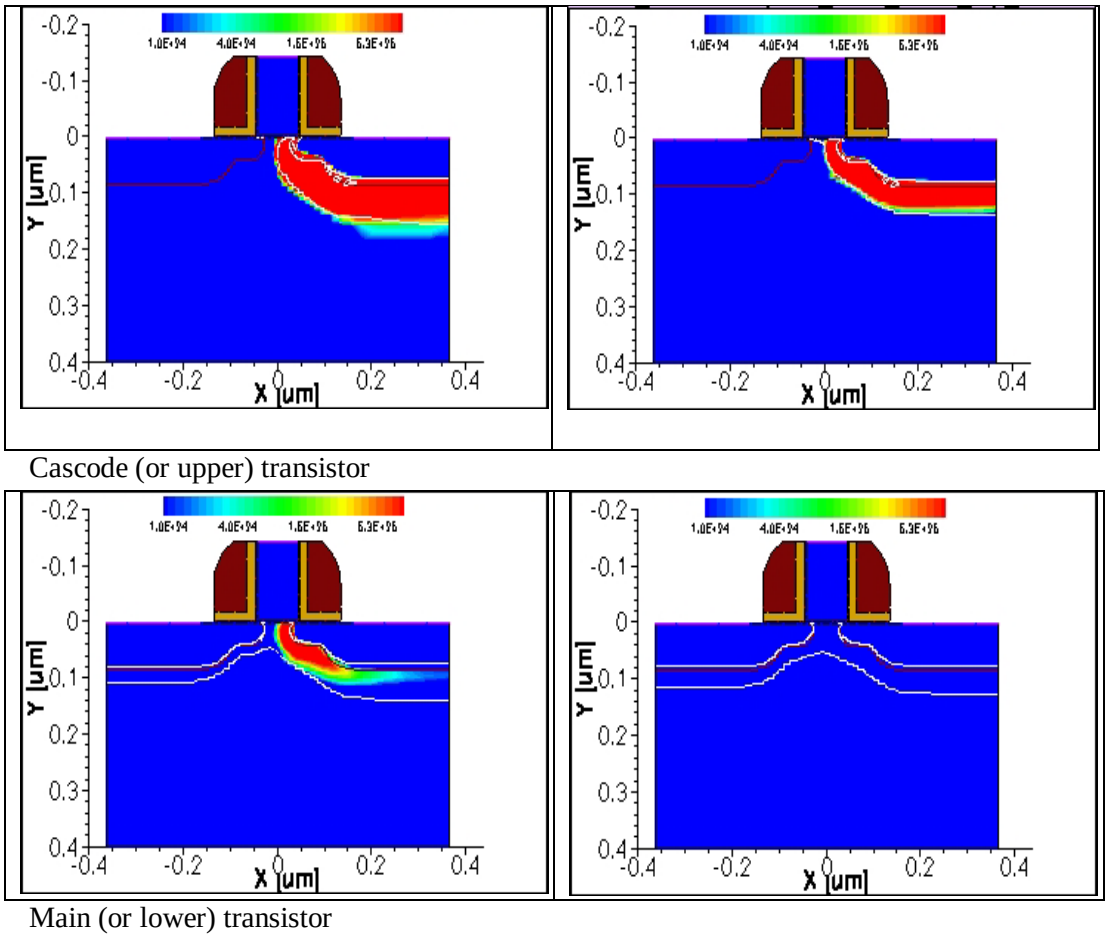


Figure 27: Impact ionization rates of the cascode transistor (upper plots) and main transistor (lower plots) at the maximum (left figures) and middle (right figures) of the output voltage transient. In this mixed-mode device and circuit simulation, VDD2 = 3.5 V.

Lattice temperature of the cascode transistor and main transistor is simulated and presented in Figure 28. The basic lattice temperature of the nMOS substrate is set to be at 300 K. Many keywords, like Thermode, RecGenHeat, Thermodynamic, and AnalyticTEP models are applied to account for lattice heating. A Thermode is a boundary where the Dirichlet boundary condition is set for the lattice. RecGenHeat includes generation-recombination heat sources. The thermodynamic model extends the drift-diffusion approach to account for electro-thermal effects. AnalyticTEP gives analytical expression for thermoelectric power. As shown in Figure 28, the cascode transistor has a higher peak lattice temperature (~ 310 K) than that in the main transistor

because of larger power dissipation in the cascode transistor. The self-heating effect is enhanced during the output voltage switching (right figures in Figure 28) because of relatively high drain-source voltage and high drain current simultaneously. High temperature increase resulting from lattice self-heating could lead to further drain current reduce of the power amplifier. Consequently, the output power and power-added efficiency of the power amplifier degrade even more. It is worth mentioning that class E power amplifier is vulnerable to the gate oxide breakdown due to very high drain-gate field stress. In this study, however, we have demonstrated that the cascode class-E power amplifier is degraded by hot electron effect during high output voltage switching with the experimental data in Section 3.3. The mixed-mode device and circuit simulation of high impact ionization rates for the cascode transistor here supports the experimental finding: PA performance degradation due to hot carrier effects subjected to DC supply voltage for 10 hours of continued RF stress at the input power of 0 dBm. The impact ionization leads to the formation of electron-hole pairs: electrons can be trapped in the gate oxide, whereas holes can generate interface states. Trapped electrons increase the threshold voltage of the n-channel MOSFET, while interface states may degrade the interface as well as the effective channel electron mobility. For the power amplifier performance degradation, threshold voltage shift is more important than mobility degradation [41]. Note more degradation can be caused with high input power RF stress in hot electron effect than that under pure DC stress [42].

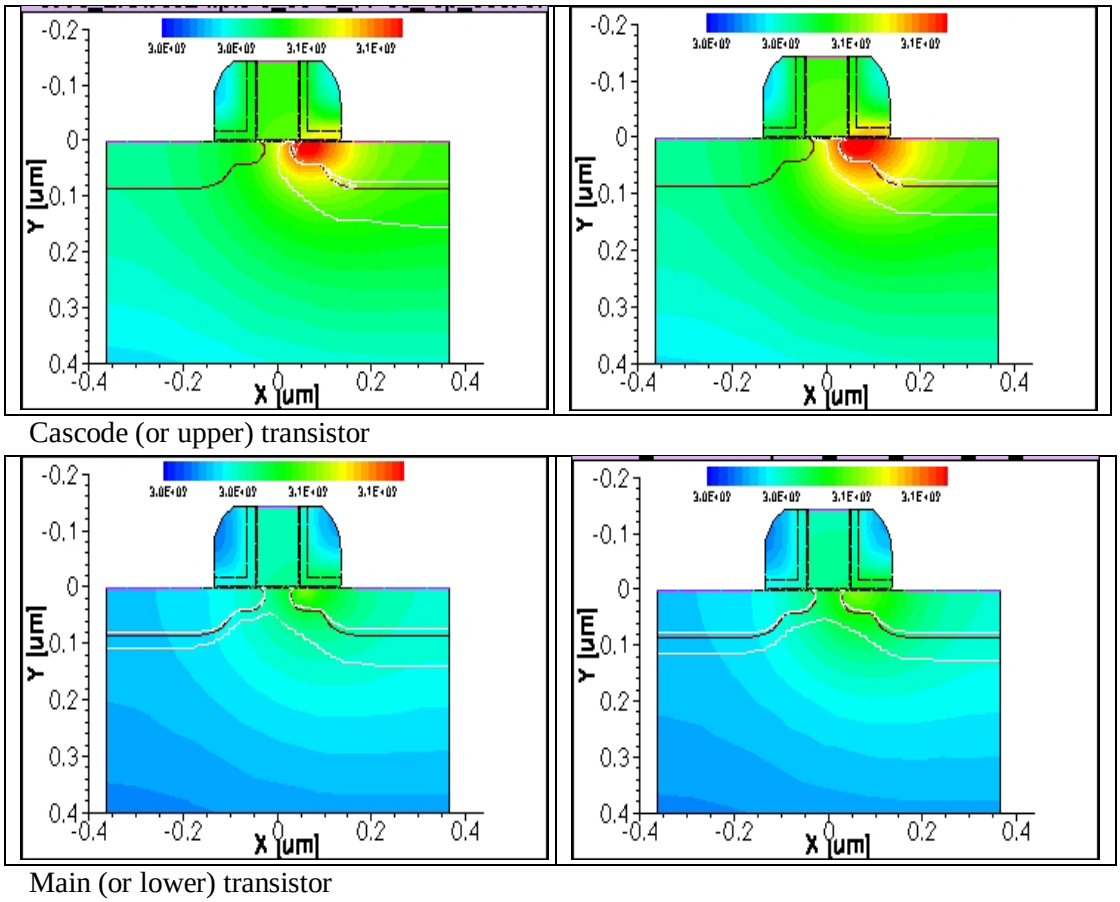


Figure 28: Lattice temperature of of the cascode transistor (upper plots) and main transistor (lower plots) at the maximum (left figures) and middle (right figures) of the output voltage transient. In this mixed-mode device and circuit simulation, VDD2 = 3.5 V.

Impact ionization rates have a peak which reaches 6.3×10^{27} /cm^3/s and the maximum lattice temperature of the cascade transistor is about 320 K, according to additional mixed-mode simulation at RF stress with V_{DD2} equal to 4.5 V. This suggests that when V_{DD2} is higher, drain electric field is higher and the hot electron effect and lattice heating are enhanced. High temperature from lattice heating, however, could reduce the hot electron effect compared to that without self-heating [43]. At the same time, high temperature accelerates gate oxide breakdown which is easily influenced by temperature and electric field [44]. In our stress experiments, however, no noticeable increase in gate leakage current was detected when VDD2 was biased at 3.5, 4.0, and 4.5 V. This implies that no transistor oxide hard breakdown occurred since hard

breakdown typically features a sudden surge of gate current [45, 46], and could destroy RF performances. In addition, ADS circuit simulation indicates that the peak drain-gate voltage of the cascode transistor with the oxide thickness of 4.08 nm has a smaller electric field than the critical field for oxide breakdown [47]. Consequently, the requirements for a hard breakdown is not satisfied, however, the oxide under this high RF and elevated DC stress may experience some kind of soft breakdown [47] which deteriorates the PA circuit performances further. Soft breakdown increases the gate leakage current noise due to formulation of random defects and conducting path within the oxide [48]. After soft breakdown trapped charge or defects are accumulated in the oxide, the nMOS transistor's threshold voltage is increased [49,50] as a result. Drain current decreases as a result of increased threshold voltage. Consequently, the PA's output power and power efficiency decrease after soft breakdown (to the first order, $\Delta P_o/P_o$ is proportional to $\Delta I_D/I_D$ [41]).

3.6 Summary

A cascode class E power amplifier has been designed at 5.2 GHz. According to the mixed-mode device and circuit simulation in the same circuit environment, which is carried out to examine impact ionization rates and lattice heating of the cascode and main transistors, the cascode transistor suffers more impact ionization and self-heating than main transistor. The design was fabricated using TSMC 0.18μm RF technology and measured freshly and with 10 hour elevated DC stress with 0 dBm RF input afterwards. The measured PA circuit performances after RF stress at different elevated V_{DD2} conditions are examined and compared with the experimental data obtained from the fresh circuit condition. Test results show that measured power gain, output power, power-added efficiency, and linearity are degraded after RF and

increased DC stresses. The circuit performance degradations are larger at higher V_{DD2} stress level. Hot electron effect turns out to be the dominate reliability resource for the degradation of cascode class E power amplifier evaluated at high input power and elevated supply voltage stress for 10 hours. Soft breakdown may contribute additional degradation to the output power and power efficiency of cascode class E PA, according to increased cascode transistor's supply voltage from 3.5 V to 4.5 V at high input power RF stress.

CHAPTER FOUR: TEMPERATURE EFFECTS STUDY ON A CLASS AB PA

4.1 Self-heating effects and RF circuits

4.1.1 Self -heating and reliability

As the feature sizes of transistors become smaller and smaller, self-heating and its impacts on the device performance and reliability are expected to become increasingly important. Circuit speed could be slowed down and interconnect delay increased for scaled device size and increasing circuit density. For scaled technology with low thermal conductivity materials such as SOI or SiGe and new device structures which are physically confined like FinFET, thermal problems can be even worse. SOI MOSFETs suffer from severe self-heating problems since thermal conductivity of the buried oxide is poor [51]. Gate-all-around silicon nanowire MOSFETs have comparable self-heating problems to that of SOI devices although the nanowire device is built on the bulk substrate [52].The thermal conductivity of thin semiconductor films is much lower than that of silicon bulk as a result of phonon confinement and boundary scattering. The increased temperature slows the transistor speed, deteriorates the interconnect delay, and causes reliability concerns. Device with feature size smaller than 32 nm has larger transistor current and power density which means worse self-heating.

Self-heating aids impact ionization, more hot carriers are generated, and device and circuit reliability is reduced accordingly. Hot-carriers are subsequently injected into the gate oxide, and giving rise to a localized and non-uniform pileup of oxide defects and leaves interface states near the drain-channel junction. At the same time, hot carriers interact with the lattice and energy is transferred to the phonon batch, thus increasing the lattice vibrations and temperature. On the other hand, the hot electron degradation for the bulk Si transistors is improved at higher

temperature since electron mean free path is reduced from phonon scattering. As a result of these two effects, the temperature coefficient of impact ionization (I.I.) turns out to be dependent on voltage. A negative temperature coefficient is typically observed for high drain voltage. When the drain voltage drops to lower region (~< 1.5 V), a positive temperature coefficient arise [53]. Furthermore, for SOI and Si/SiGe MOSFETs, positive temperature coefficient of impact ionization has been observed [54,55].The localized hot spot is channel length dependent for nano-scale transistors [56].Three-dimensional electro-thermal simulation results show that self-heating effects degrade the FinFET on-current significantly. A detailed thermal analysis of a 30 nm gate length n-channel FinFET was presented in [57].

4.1.2 Self- heating and RF circuits

Historically, Power amplifier (PA) design has been a critical design subject in RF part for the key role it plays in modern communication systems. PA has a self-heating problem born with its high power operation feature which degrades the power amplifier performances. Maintaining the stable output power over a wide range of temperatures is desirable in many applications. A practical example can be found in the wideband code-division multiple access wireless communications system, where multiple users share the same carrier frequency. In order to get equal power from each user at the base station, the transmitter gain has to be regulated.

The effect of the transistor self-heating phenomenon is discovered to be more severe under narrow-band signal[58]. The temperature effect on a Ku-band NMOS common-gate low-noise amplifier has been examined by Chen et al. [59]. Yamauchi et al. [60] proposed an X-band monolithic-microwave integrated-circuit power amplifier in which a simplified on-chip temperature compensation circuit composed of diodes and a resistor was utilized. Process and temperature compensation technique for RF low-noise amplifiers and mixers was presented by

Gómez et al. [61]. Filanovsky and Allam [62] pointed out that mutual compensation of mobility and threshold voltage variations on temperature may be achieved by proper bias point of a MOS transistor. Gain variation caused by temperature-dependent parameters of transistor leads to unavoidable electro-thermal memory effects. Boumaiza and Ghannouchi proposed a dynamic electro-thermal behavior modelon power amplifiers and used the temperature-compensated pre-distortion function to compensate for self-heating effects[58].

4.1.3 Related work

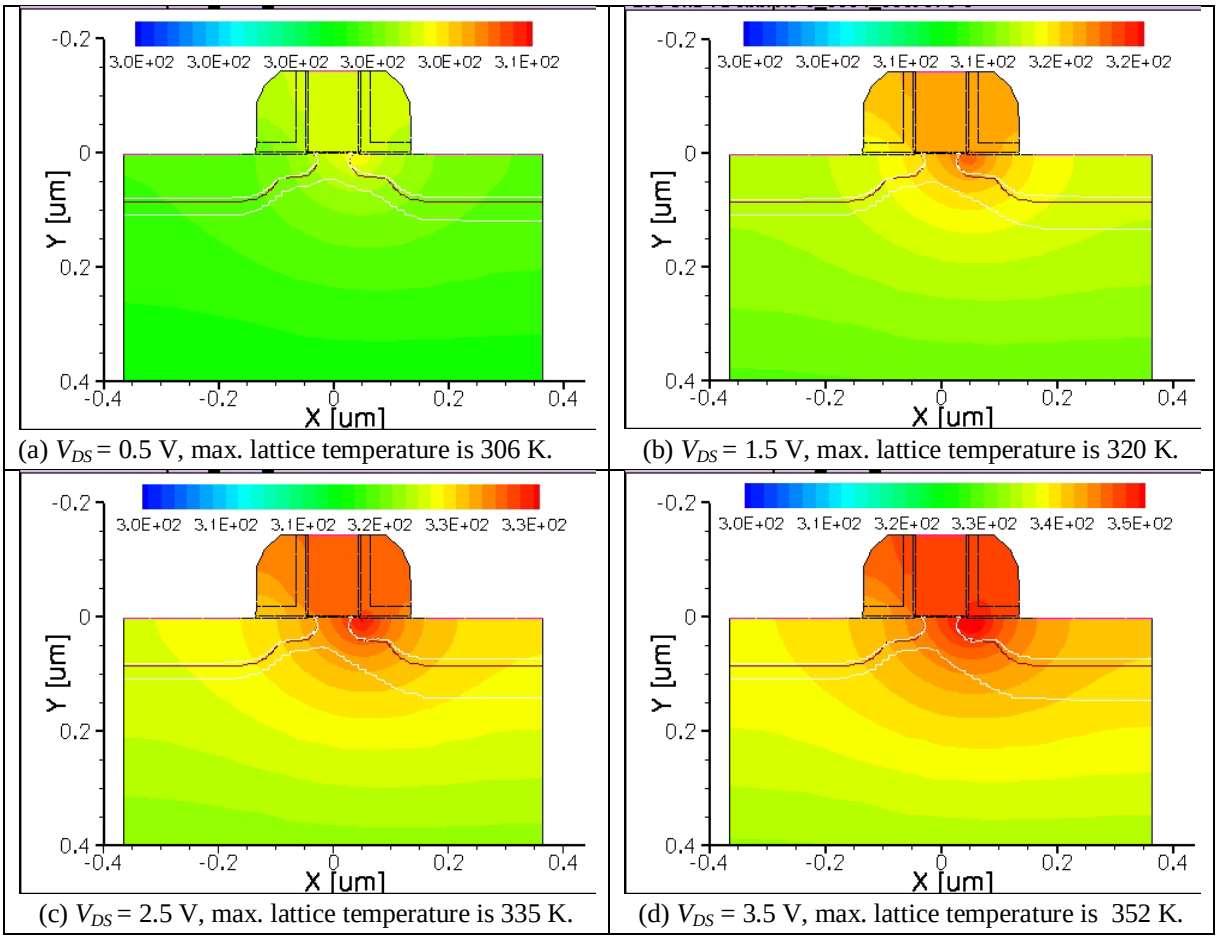
TCAD tools are known to analyze device performances such as self-heating, current density, field distribution etc. Besides the publications in [52], [56], [57], a 2-D drift-diffusion electro-thermal simulation was applied by Fiegna et al. to analyze the thermal effects on nano-scale SOI nMOSFETs[63]. Choi et al., investigated the strained Si nMOS ESD protection behavior including device self-heating effects [64]. Shrivastave et al., used the device simulation to examine a novel bottom spacer FinFET structure for short channel, power-delay, and self-heating performances [65].

The self-heating effect of the NMOS power amplifier is studied in this work. Lattice temperature is examined on a single NMOS transistor for DC sweep as well as transient simulation in circuit environment. Self-heating also affects impact ionization rates and influences device and circuit reliability, so the I.I effects are also evaluated. A class AB RF power amplifier operating at 5.2 GHz is designed in ADS. Different gate bias circuits for power amplifier temperature compensation are examined. Comparison was made among the output power and power-added efficiency, η_{add}, of the power amplifier for the different biasing schemes.

4.2 Mixed mode Simulation on a Class AB PA

4.2.1 DC Device Simulation

Self-heating effect of a silicon nMOSFET is evaluated in the circuit environment using Sentaurus TCAD software [66]. The channel length of the NMOS transistor simulated is 90 nm and channel width is 250 µm. Single device simulation was performed at first and results presented in Figure 29. The gate is biased at 0.8 V and drain voltage is swept from 0 to 3.5 V. The physical models of Shockley-Read-Hall recombination, Auger recombination, are adopted for impact ionization. The impact ionization model assumes the impact ionization coefficient to be a function of the local field. Poisson's and Hole electron continuity equations with drift-diffusion transport mechanisms are selected. Mathiessen's rule is used to calculate the low field mobility the bulk- and surface mobility incorporated. The bulk mobility model is Philips unified mobility model [67], which takes into account electron-hole scattering, screening of ionized impurity by charge carriers, and clustering of impurities. To account for lattice heating, Thermodynamic, Thermode, RecGenHeat, and AnalyticTEP models are included. The thermodynamic model extends the drift-diffusion approach to account for electrothermal effects. A Thermode is a boundary where the Dirichlet boundary condition is set for the lattice. RecGenHeat includes generation-recombination heat sources. AnalyticTEP gives analytical expression for thermoelectric power. Figure 1 shows the transistor lattice temperature contours at different bias conditions. The hottest spot is near the drain edge and temperature increases with V_{DS} due to higher DC power dissipation or Joule heating. The maximum lattice temperature is 308 K at V_{DS} = 0.5 V while climes up to 352 K at V_{DS} = 3.5 V.



(a) V_{DS} = 0.5 V, max. lattice temperature is 306 K.

(b) V_{DS} = 1.5 V, max. lattice temperature is 320 K.

(c) V_{DS} = 2.5 V, max. lattice temperature is 335 K.

(d) V_{DS} = 3.5 V, max. lattice temperature is 352 K.

Figure 29: Lattice temperature at (a) V_{DS}= 0.5 V, (b) V_{DS}= 1.5 V, (c) V_{DS}= 2.5V, (d) V_{DS}= 3.5 V.

4.2.2 Mixed mode circuit transient simulation on Class AB PA

The mixed-mode device and circuit simulation is used in Sentaurus to examine the nMOS transistor in the power amplifier operation. The class-AB PA in Figure 30 is operating at 5.2 GHz. The impact ionization rates as well as temperature variation as a function of time during PA operation is investigated.

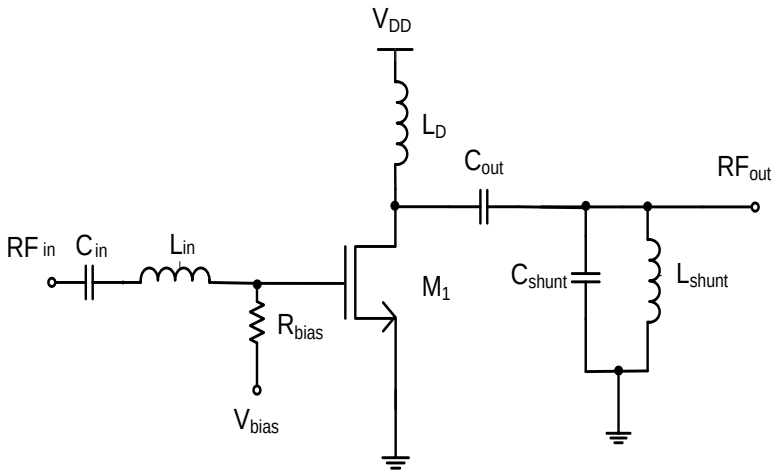
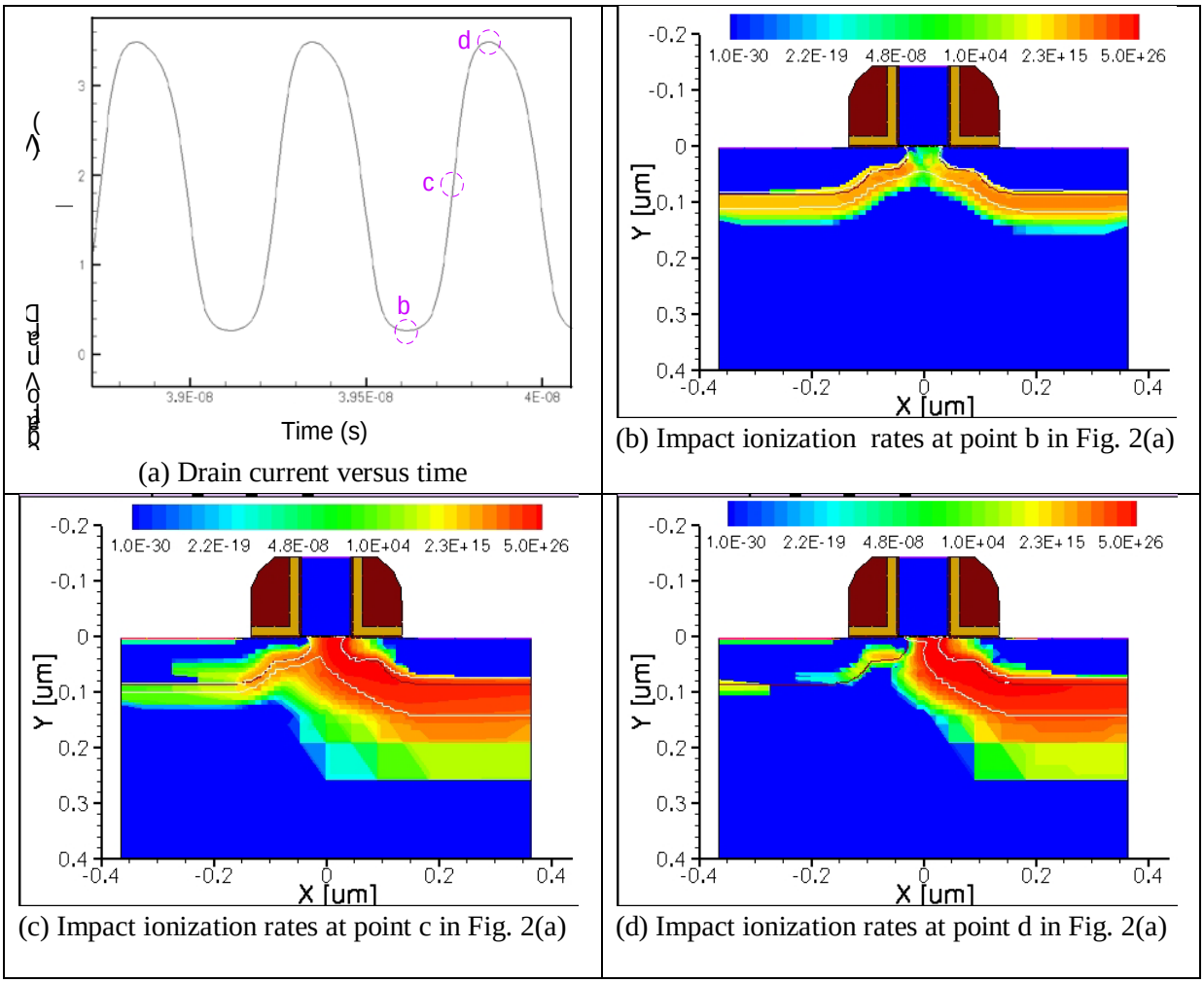


Figure 30: Circuit schematics of a class-AB power amplifier.

The impact ionization rates are probed at three different time points (bottom of VDS, middle point of VDS, and top of VDS) of the drain voltage waveform in Figure 31(a). As can be observed from Figures 31(b), (c), and (d), the I.I. rates at the top of the V_{DS} point is highest due to largest local electrical field, which may cause more hot electrons injection into the gate oxide near the drain region.



(a) Drain current versus time

(b) Impact ionization rates at point b in Fig. 2(a)

(c) Impact ionization rates at point c in Fig. 2(a)

(d) Impact ionization rates at point d in Fig. 2(a)

Figure 31: Impact ionization rates corresponding to bottom, middle point, or top of VDS.

Transient lattice temperature simulation results are displayed in Figure 32. The localized lattice temperature hot spot rises from initial substrate temperature 300 K to about 360 K. If simulating for longer time, the maximum lattice temperature will saturate at around 360 K, which reaches good agreement with the DC simulation results in Figure 29.

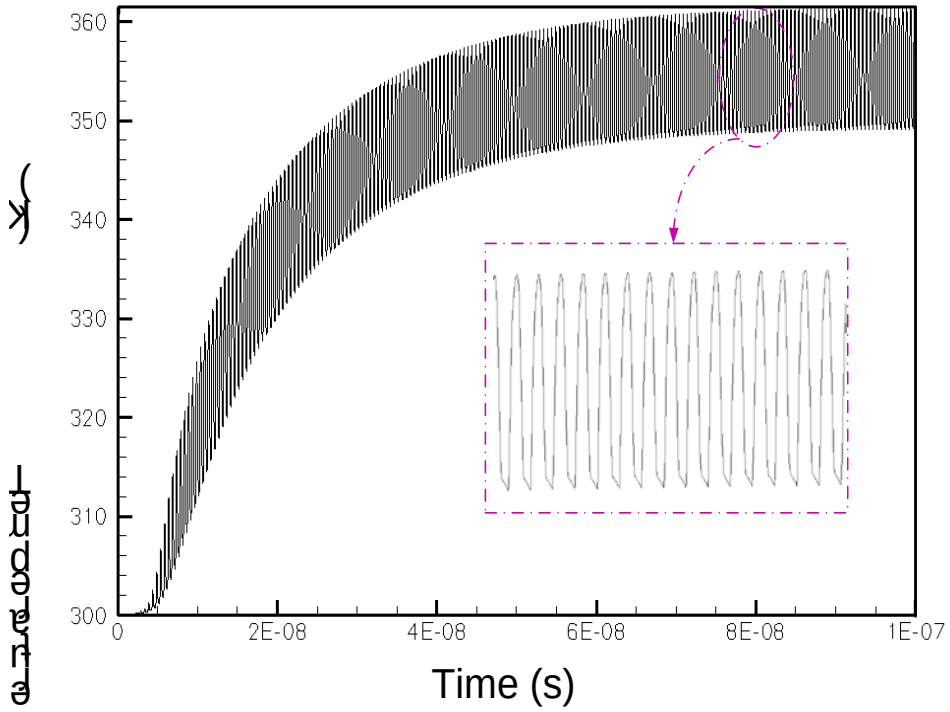
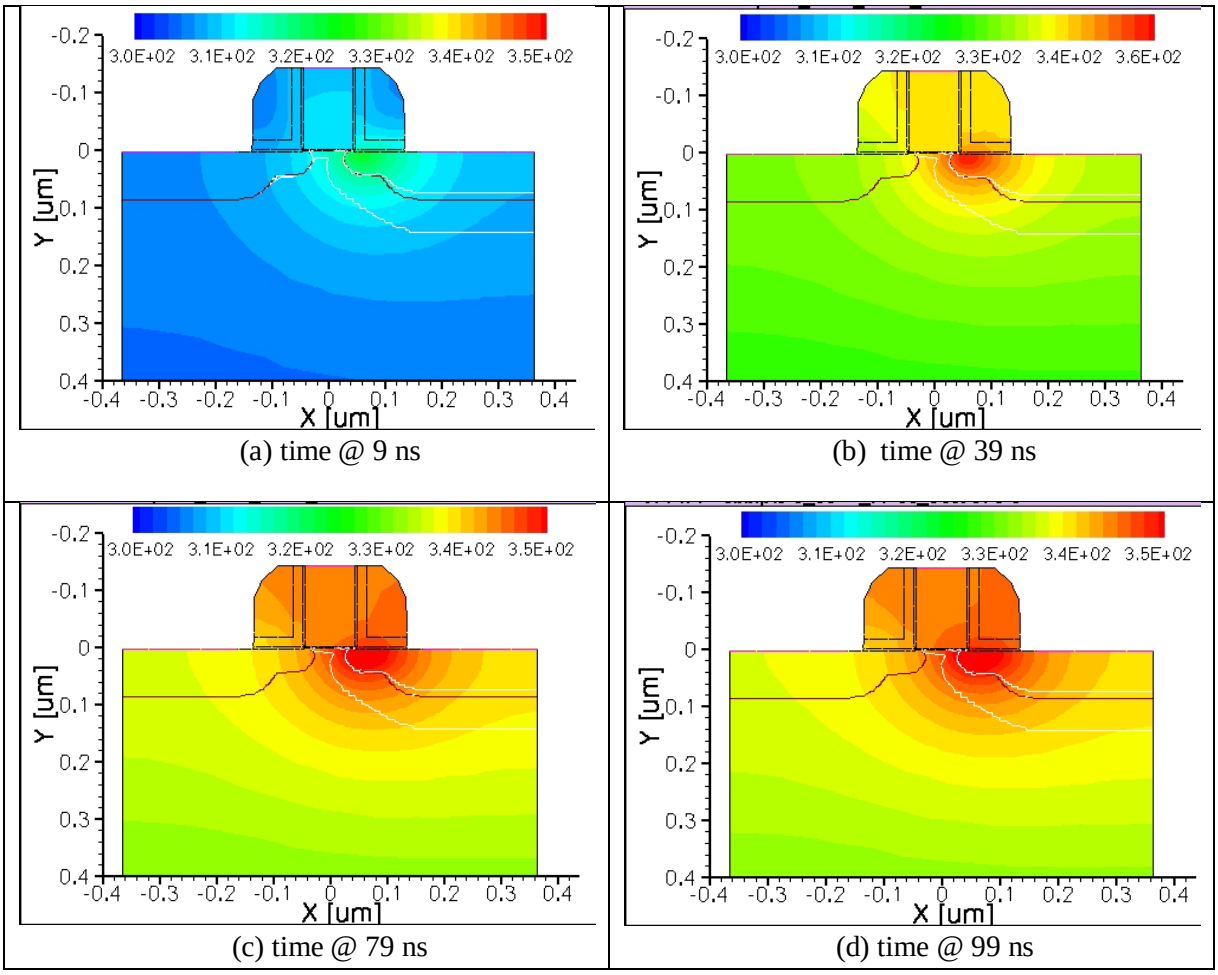


Figure 32: Maximum lattice temperature versus time.

Figure 33 displays the lattice temperature at four time points respectively 9 ns, 39 ns, 79 ns, and 99 ns. For better comparison the maximum lattice temperature of the plots is set as 355 K. As a result of self-heating, the color near the drain region becomes redder/ hotter as time goes longer.



(a) time @ 9 ns (b) time @ 39 ns

(c) time @ 79 ns (d) time @ 99 ns

Figure 33: Lattice temperature versus time @ 9 ns, 39 ns, 79 ns, and 99 ns, respectively.

4.3 Temperature compensation techniques

As a rule of thumb, RF performance degrades as result of self-heating effects. Power amplifier performances are especially sensitive to temperature variations. Output power decreases at higher operation temperature. Temperature increase could be attributed to device self-heating or ambient temperature rise. It is desirable to have a temperature compensation

circuit to stabilize PA performance over a wide range of temperatures. Various gate bias circuits are examined in this section.

4.3.1 Compensation circuit-- classic constant G_m

A classical constant-gm circuit as shown in Figure 34 is investigated. In this circuit the transconductance of M3 can be written as [61]

$$g_{m3} = \frac{2(1-\eta^{-0.5})}{R} \tag{7}$$

where η is the channel width ratio between M3 and M4, R is the resistance.

Using the drain current equation of $I_{D3} = \frac{\mu_n C_{ox} W_3}{2L_3}(V_{GS3} - V_T)^2$, the transconductance g_{m3} can also be written as

$$g_{m3} \equiv \frac{\partial I_{D3}}{\partial V_{G3}} = \frac{\mu_n C_{ox} W_3}{L_3}(V_{GS3} - V_T) \tag{8}$$

here μ_n is the electron mobility, V_T is the threshold voltage, C_{ox} is the oxide capacitance per unit area, W_3 is the channel width, L_3 is the channel length, and V_{GS3} is the gate-source voltage of M_3. This drain current equation is merely used for illustration of this adaptive gate bias technique and does not include the secondary effects such as channel length modulation, drain induced barrier lowing (DIBL), etc. Combining (7) and (8) gives the bias voltage

$$V_{bias} = V_{GS3} = V_T + \frac{2(1-\eta^{-0.5})L_3}{RC_{ox}W_3\mu_n}. \tag{9}$$

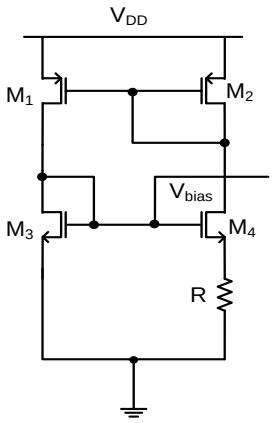


Figure 34: Constant-gm bias circuit.

Using (9) the temperature sensitivity of bias voltage is derived to be

$$\frac{\partial V_{bias}}{\partial T} = \frac{\partial V_T}{\partial T} - \frac{2(1-\eta^{-0.5})L_3}{RC_{ox}W_3\mu_n^2}\frac{\partial \mu_n}{\partial T} . \quad (10)$$

Since both $\frac{\partial V_T}{\partial T}$ and $\frac{\partial \mu_n}{\partial T}$ are negative, the first term of (10) is a negative number (or V_{bias} decreases with temperature) and the second term is positive (or V_{bias} increases with temperature). Thus, V_{bias} can be designed to have a positive trend with temperature to compensate the drain current decreasing resulted from temperature increasing.

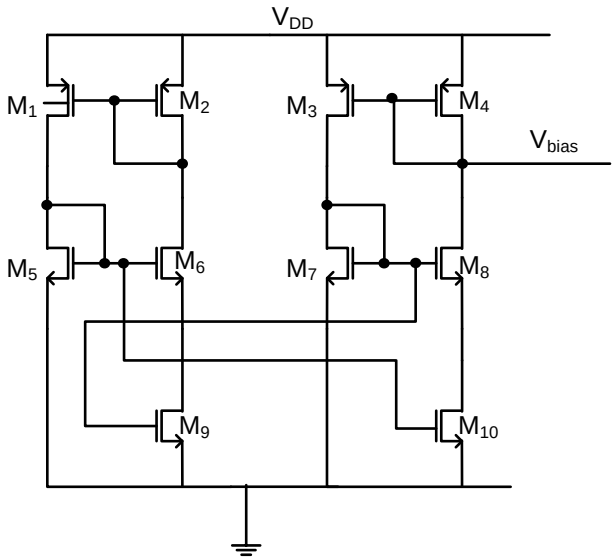


Figure 35: Improved current-source bias circuit.

An improved current-source gate bias circuit [61] is shown in Figure. 35, they share the same constant biasing scheme and the only difference is that this is a two-stage structure. Similarly, the equations derived above can also be used to illustrate this biasing technique. This structure is also examined for its capability of temperature compensation, results are compared later.

4.3.2 Compensation circuit---simple gate biasing

For minimum design overhead, a simple adaptive biasing scheme that produces a stable output power over a wide temperature range is desirable. This simple circuit is illustrated in Figure 36, the KCL equation yields

$$V_{bias} = V_{DD} - I_D R1. \tag{11}$$

Using the drain current expression $I_D = \frac{\mu_n C_{ox} W}{2L}\left(V_{GS} - V_T\right)^2$, the temperature sensitivity of V_{bias} is derived to be

$$\frac{\partial V_{bias}}{\partial T} = -R1\frac{\partial I_D}{\partial T}$$

$$= -R1\left(\frac{\partial I_D}{\partial V_T}\frac{\partial V_T}{\partial T}\right) - R1\left(\frac{\partial I_D}{\partial \mu_n}\frac{\partial \mu_n}{\partial T}\right)$$

$$= R1\left(\frac{2I_D}{V_{GS}-V_T}\frac{\partial V_T}{\partial T}\right) - R1\left(\frac{I_D}{\mu_n}\frac{\partial \mu_n}{\partial T}\right) . \quad (12)$$

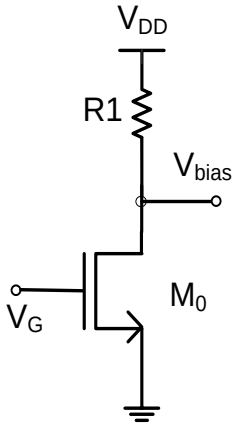


Figure 36: A simple gate bias circuit.

The mobility and threshold voltage of the MOS transistor decrease as a result of temperature increases [61]. The drain current of the MOSFET will then decrease with temperature if the mobility term dominates over the threshold voltage term in (12). Thus, V_{bias} will increase with temperature. In (12), as temperature rises, the first term will result in a decrease of V_{bias} while second term results in an increase of V_{bias}. If V_{GS}-V_T is high enough, the second term in (6) will dominate the V_{bias} temperature performance.

4.4 Temperature effects modeling

To provide more insights on heating effects, the threshold voltage and mobility as a function of temperature is modeled in this section. Typically the temperature-dependent threshold voltage is given as

$$V_T = V_T(T_0) + \alpha_V (T - T_0) \tag{13}$$

where α_V is the temperature coefficient for the threshold voltage. From (13) one obtains

$$\frac{\partial V_T}{\partial T} = \alpha_V . \tag{14}$$

α_V lies in the range of -0.5 to -4 mV/K. In our analytical model, α_V = -5×10^{-4} V/K and $V_T(T_0)$ = 0.36 V. The temperature-dependent electron mobility can be expressed as

$$\mu_n = \mu_n(T_0) + (T - T_0)^{\alpha_u} \tag{15}$$

where α_μ is the temperature coefficient for the electron mobility. Using (15), the temperature sensitivity of the electron mobility is derived as

$$\frac{\partial \mu_n}{\partial T} = \alpha_\mu (T - T_0)^{\alpha_\mu - 1} . \tag{16}$$

α_μ is between the range of -1.5 to -2 [62]. α_μ = -1.5 means the mobility model is determined by the carrier-carrier scattering mechanism [68]. To demonstrate the validity of analytical equations above and provide insight to bias temperature compensation, the curves of bias voltage as a function of temperature for different mobility model parameters (α_μ = -1.5, -1, and -0.7) are shown in Figure 37 and Figure 38. Comparison was made among the simple bias circuit in Figure 36 and the constant-g_m bias circuit in Figure 34. Here, $\mu_n(T_0)$ = 258 cm^2/V·s. It is obvious from this comparison that α_μ = -0.7 gives better fit to the ADS simulation result, although not

consistent with common knowledge. Another way to model the temperature-dependent mobility is polynomial form:

$$\mu_n = \mu_n(T_0) + \alpha_{\mu 1}(T - T_0) + \alpha_{\mu 2}(t - T_0)^2 \tag{11}$$

where $\alpha_{\mu 1}$ and $\alpha_{\mu 2}$ are the first-order and second-order temperature coefficients. Thus,

$$\frac{\partial \mu_n}{\partial T} = \alpha_{\mu 1} + 2\alpha_{\mu 2}(T - T_0). \tag{12}$$

The polynomial modeling results are also shown in Figure 37 and Figure 38. In Figure 37 $\alpha_{\mu 1}$ = -0.48, $\alpha_{\mu 2}$ = -1×10^{-5}, V_{DD} = 1.8 V, R_1 in Figure 36 is 165 Ω, and the nMOS transistor using the TSMC 0.18 μm process has the channel width of 24.8 μm. In Fig. 38 $\alpha_{\mu 1}$ = -0.48 and $\alpha_{\mu 2}$ = -1×10^{-3}, V_{DD} = 1.8 V, and R in Figure 34 is 5 Ω. As seen in both Figure 37 and Figure 38, a better fit to the ADS simulation results can be get from the second-order polynomial expression for both bias circuits over a wide range of temperatures (from -40 °C to 120 °C).

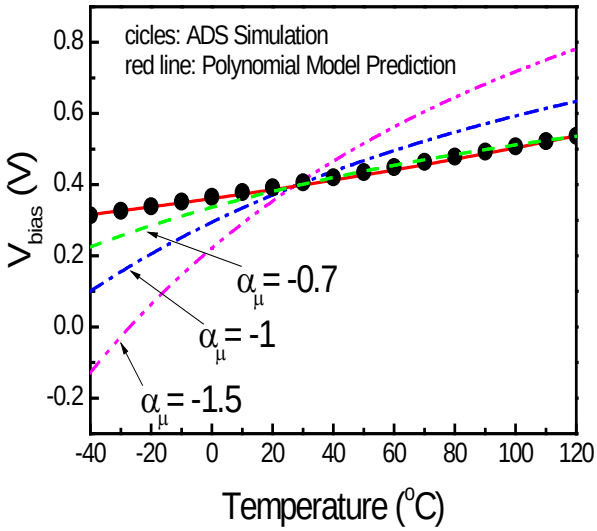


Figure 37: Bias voltage versus temperature for the simple bias circuit in Fig. 34.

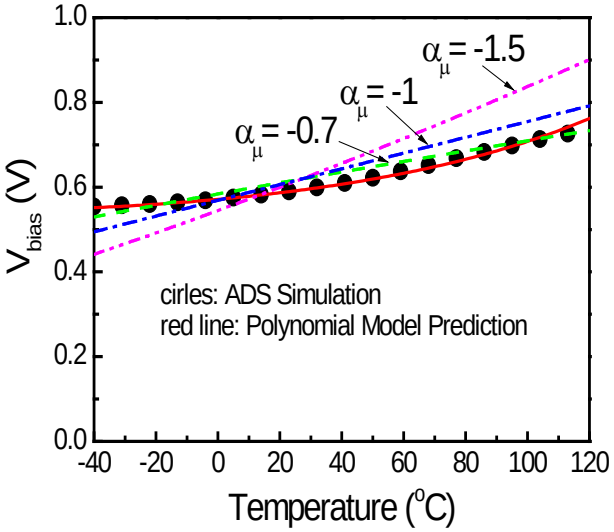


Figure 38: Bias voltage versus temperature for the constant-gm bias circuit in Fig. 36.

4.5 Class AB PA Temperature performances

A class AB single-stage PA with the same architecture as shown in Figure 30 is designed in order to evaluate the PA performances over wide temperature variations. It has exactly the same operation condition as used in Sentaurus mixed mode simulation. The PA is operated at 5.2 GHz.

To gain the stressed MOSFET transistor model parameters, nMOS transistors were stressed at V_{GS} = V_{DS} = 2.8 V for 1800, 3600, 5400, and 7200 seconds, *I-V* measurement was done at 87 °C (≈ 360 K) and 127 °C (≈ 400 K) respectively. Hot electron effect analysis was done on the power amplifier based on the data obtained.

The normalized output power ($\Delta P_o/P_o$ ×100%, ΔP_o and P_o are in mW) and power-added efficiency ($\eta_{add} \equiv (P_o\text{-}P_i)/P_{DC}$ ×100%, P_o P_i, and P_{DC} are in mW) versus hot electron stress time are presented in Figure 39 and Figure 40. Both output power and power-added efficiency decrease with stress time. When temperature is increased, the drain current and PA's conduction

angle will decrease, which results in faster normalized output power and power-added efficiency decrease at 127 °C than at 87 °C.

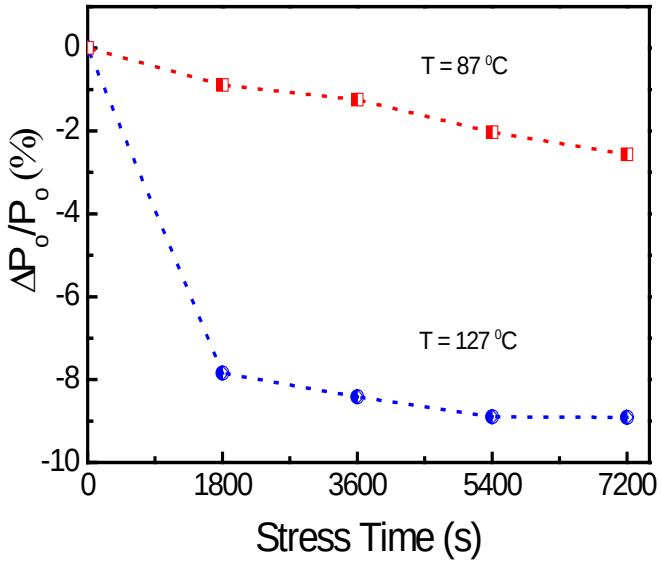


Figure 39: Normalized output power versus hot electron stress time.

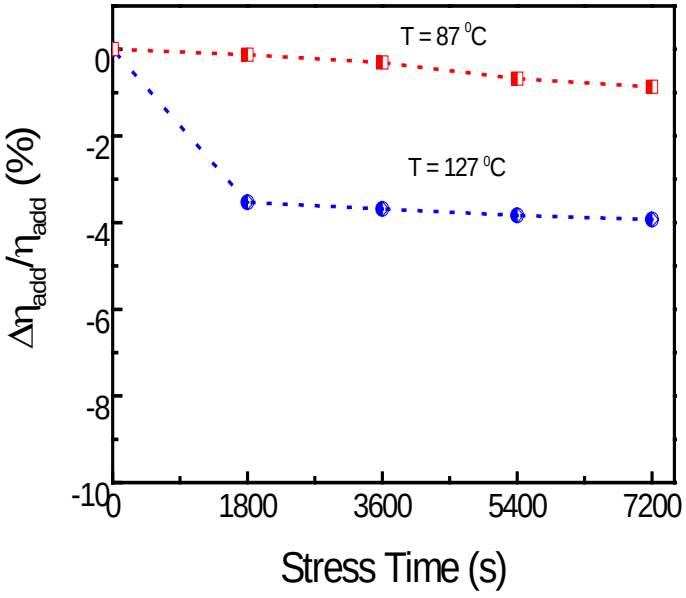


Figure 40: Normalized power-added efficiency versus hot electron stress time.

A wide range temperature sweep from -40 ^{o}C to 120 ^{o}C is performed with harmonic balance simulation. The temperature increase may be due to device self-heating and/or ambient temperature rises. Results from ADS simulation, as what is expected, show that both output power and power-added efficiency decrease with temperature. Drain current of the main transistor also decreases at higher temperature. The output power of decreases as a result. In this technology used, which is BSIM4 model, secondary effects, including nonlinear output conductance characteristics (channel length modulation, DIBL, substrate-current induced body effect, etc.) and temperature dependence are included in the drain current equation. Normalized output power and power-added efficiency as a function of temperature are displayed in Figure 41 and Figure 42. Dash lines are the output power and power-added efficiency of the PA with constant gate bias and the solid lines stands for the results from the power amplifier with the gate bias circuits, described previously. It can be shown that constant-g_m bias circuit does not provide much temperature compensation. However, the improved double stage current-source gate bias circuit [61] shown in Figure 35 reduces the output power temperature drift as seen in Figure 41 and Figure 42. The gate voltage of the double stage gate bias circuit is also demonstrated in Figure 41. V_G increases with temperature to compensate the decrease of I_{DS} due to the thermal effect.

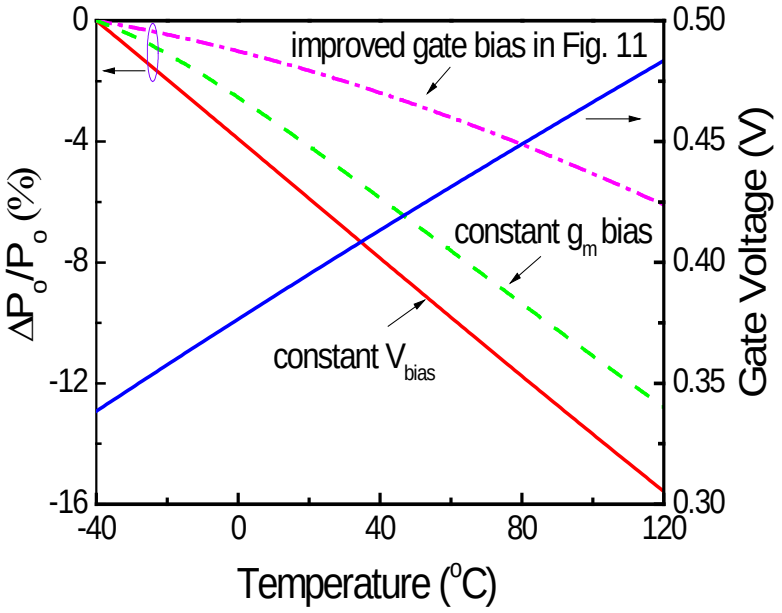


Figure 41: Normalized output power versus temperature.

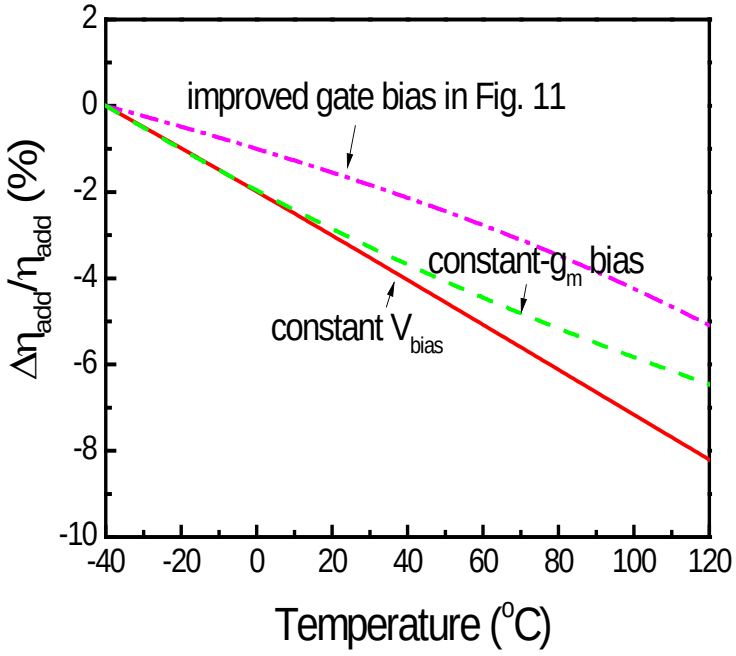


Figure 42: Normalized power-added efficiency versus temperature.

To get more comparison, temperature compensation abilities on normalized output power and power-added efficiency of the adaptive gate biasing in Figure 36 are demonstrated in Figure 43 and Figure 44. The performance variations of constant gate biasing power amplifier are also shown both figures for better contrast. Since typically the main transistor dissipates more power compared to the bias transistor and they may be spatially separated, different temperature increase on the main transistor and the bias transistor are considered. Three different conditions are simulated and the output power and power-added efficiency are shown in both figures: the ambient temperature increase for main transistor and the bias transistor are both zero (dash line), the main transistor has higher temperature rise (T_{rise}) of 15 ˚C due to self-heating, while the bias transistor has lower temperature rise of 5 ˚C (dash and dotted line), and the main transistor has even higher temperature rise of 25 ˚C, while the bias transistor remains a 5 ˚C temperature rise (dotted line). Note that the values of T_{rise} for the main transistor and the bias transistor can be independently set in ADS simulation. Two independent BSIM4 models are used for these two transistors to account for different temperature rises [69].

If one can observe from Figure 43, over temperature range that is concerned, the output power of the PA with adaptive gate bias is much less sensitive to temperature variations than that with a constant gate bias. PA performances are even more stable when the main transistor has more temperature rise compared to the bias transistor, due to enhanced gate bias temperature compensation effect.

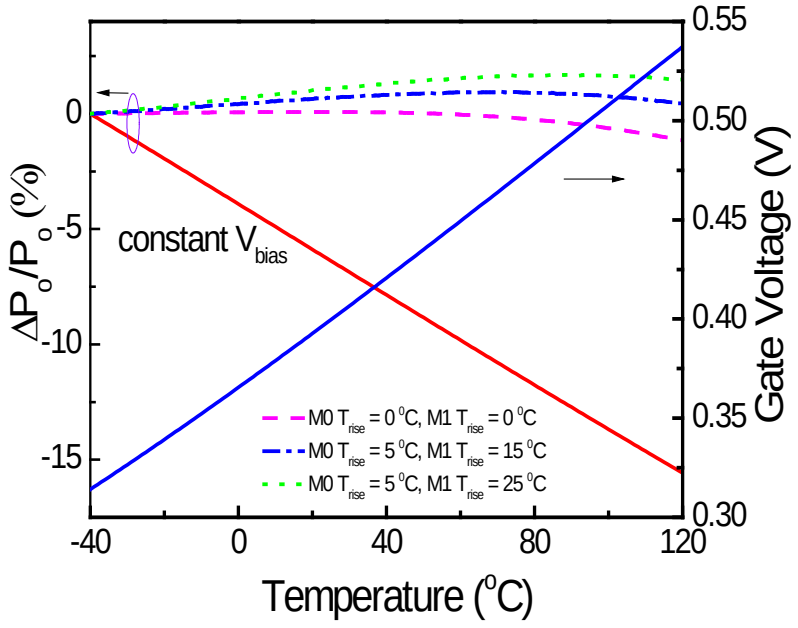


Figure 43: Normalized output power versus temperature (adaptive gate bias). nMOS transistor is biased at 1.45 V.

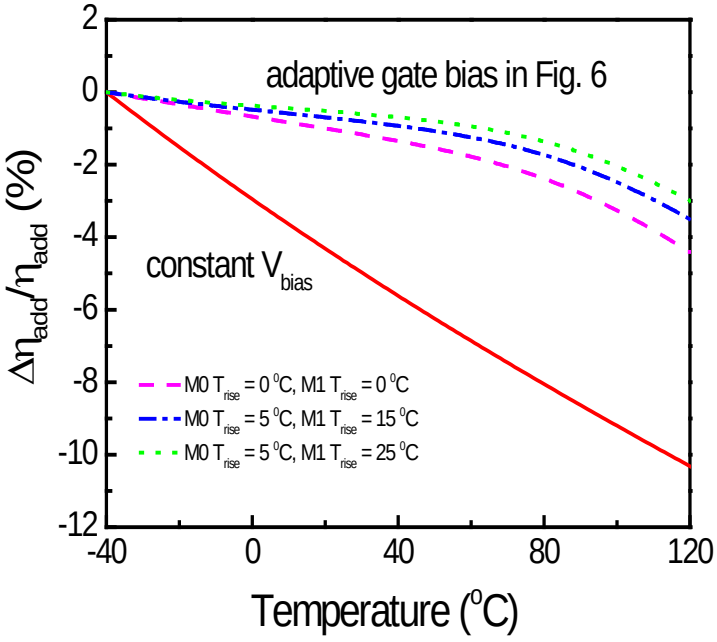


Figure 44: Normalized power-added efficiency versus temperature (adaptive gate bias).

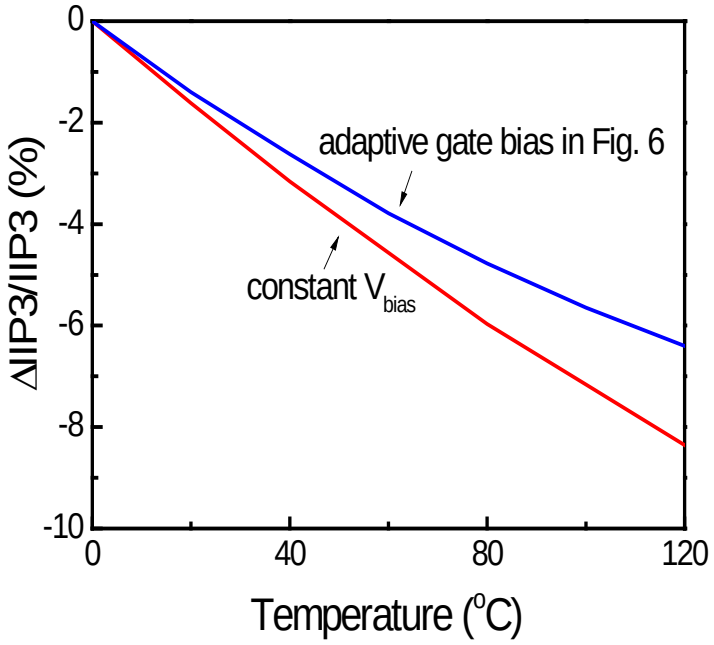


Figure 45: Normalized IP3 referring to the input versus temperature.

Figure 45 shows normalized third-order intermodulation referred to the input ($\Delta IIP_3/IIP_3$ ×100%, ΔIIP_3 and IIP_3 are in dBm) versus temperature for the power amplifier. Comparison is made between constant gate bias and adaptive gate bias. IIP3 decreases with temperature for both. However, less temperature sensitivity is observed from the power amplifier with simple adaptive gate bias.

4.6 Summary

The author has examined self-heating effect on a RF power amplifier using the mixed-mode device and circuit simulation. The hottest lattice temperature region is near the drain region of the n-channel MOSFET and is associated with drain voltage and time during PA operation. Different gate bias circuits have been evaluated for temperature compensation

capability on PA. The adaptive gate bias scheme, although simple, can effectively reduce the effect of temperature variations on PA over a wide range of temperatures (from -40 to 120 ˚C).

CHAPTER FIVE: PROCESS VARIATION STUDY: MONTE CARLO SIMULATION

5.1 Process Variations

5.1.1 New trend of reliability study

To gain better speed and further reduce coat, silicon CMOS are scaled down to 22 nm [70] and beyond. The well-known reliability mechanisms such as gate oxide breakdown (GOB), hot carrier injection (HCI), and negative bias temperature instability (NBTI) remain very important for the design of digital and RF circuits.

HCI is associated with hot carrier get trapped in gate oxide which is deteriorated by high lateral field induced impact ionization (I.I.), while GOB is related to field -induced oxide traps or defects due to oxide vertical scaling. NBTI occurs due to a build-up of positive charges occurs either at the Si/SiO2 interface or in the oxide layer of p-channel MOSFETs under negative gate bias at higher temperatures. The reaction-diffusion model [71] illustrates the holes in the inversion layer of pMOSFETs react with the Si-H bonds at the SiO2/Si interface. The hydrogen species diffuse away from the interface toward the polysilicon gate. This causes the threshold voltage shift of p-MOSFETs.

Originally, process variations were considered in die to die variations. However, with transistors progress into nanoscale regime, intra die variations are posing the major design challenge as technology node scales. Fluctuations with intrinsic device parameters that result from process uncertainties have substantially affected the device characteristics. For state-of-the-art nano-scale circuits and systems, device variation and uncertainty of signal propagation time between dies and inside die have become crucial in the variation of system timing and the determination of clock speed. Yield analysis and optimization, which takes into account the

manufacturing tolerances, model uncertainties, variations in the process parameters, and aging factors are known as indispensable components of the circuit design procedure.

5.1.2 Process variability details

Statistical variability (SV), which is mainly caused by the discrete nature of charge and the granularity of matter, is one of the crucial limitations of device scaling. Process variability comes from random dopant fluctuation (RDF), line edge roughness (LER), and poly gate granularity (PGG). Inside each device, there are many regions needs to be doped specifically according to the device design. With device size becomes smaller, the average number of dopant atoms becomes less and less. It is extremely difficult to control the accuracy of the dopant amount, resulting fluctuations of dopants between devices. RDF [72] remains the dominant source of statistical variability and is mainly caused by silicon dopant fluctuations during fabrication process. It becomes more severe as device size shrinks. LER, is a random deviation of line edges from gate definition. It is notoriously difficult to scale with line width due to the molecular structure of photo-resist. Study shows that in bulk MOSFETs beyond gate lengths of 20nm it can overtake RDD in becoming the dominant IPF source[73]. PGG is basically attributed to gate dielectric thickness variations which contribute to threshold voltage variations; it is also caused by faster diffusion speed along the gate oxide grain boundaries which leads to uneven doping. High density of defect states along the boundaries of gate grains can cause Fermi level pinning, as a result, surface potential fluctuates within the MOSFET channel, which also contributes variation of threshold voltage and other device parameters as part of the effects of PGG. All the above mentioned process variations cause fluctuation of threshold voltage, mobility, and oxide thickness, which in turn affect the device and circuit performance. Furthermore, reliability issue could widen the standard derivation of process variation in Gaussian distribution [74].

5.1.3 Related work

There have been numerous papers on reliability and process variability and their impacts on circuit performances published recently. To illustrate, NBTI is a major contributor to CMOS ring oscillator propagation delay [75]. GOB reduces the static noise margin of the SRAM cell [76]. Hot electron effect increases noise figure of low noise amplifier [77], decreases the output power and power efficiency of power amplifier [78], and increases phase noise of cross-coupled oscillator [79]. For process variability, random-dopant-induced variability was studied by Li et al. in nano-scale device cutoff frequency and CMOS inverter gate delay [80]. Hansson and Alvandpour demonstrated that the delay variation in the master-slave flip flops is 2.7 times larger than the delay variation in a 5-stage inverter chain [81]. Rao et al. described a on-chip technique to measure local random variation of FET current which is completely digital [82]. Mukhopadhyay et al. presented that large variability and asymmetry in threshold-voltage distribution significantly increase leakage spread and degrade stability of fully depleted SOI SRAM cell due to random dopant fluctuation [83].

What's more, to provide solutions of process variation, Didac Gómez [84] presented a circuit compensation technique to analyze and reduce temperature and process variation effects on low noise amplifiers and mixers. Han et al. [85] addressed a post-manufacturing self-tuning technique that aims to compensate for multi-parameter variations. However, the effects of aging and process variations on RF oscillator and the circuit technique to reduce variability effect on oscillator have not been well studied. Liu and Yuan [86] developed an adaptive body bias technique for power amplifier resilient to reliability aging and process variations.

5.2 Mixed Mode Simulation on Colpitts Oscillator

5.2.1 Colpitts oscillator circuit design

The Colpitts oscillator used in the mixed-mode device and circuit simulation is described in Figure 46. Device physical insight can be extracted from the mixed-mode simulation and it mimics the practical circuit environment.

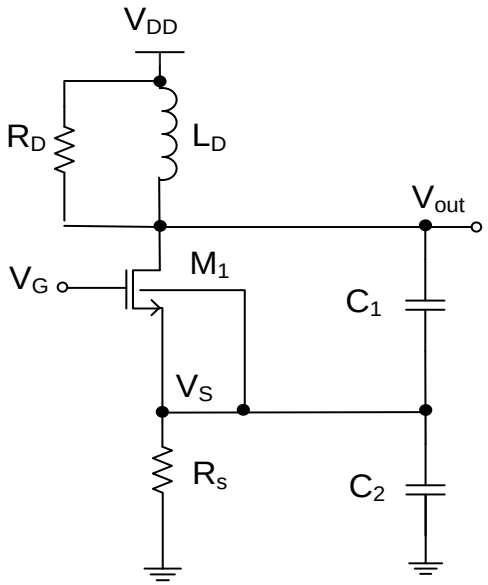


Figure 46: Schematics of an oscillator used in the mixed-mode device and circuit simulation.

The MOSFET has the channel length of 65 nm and the channel width of 64 μm in Sentaurus simulation. The circuit parameters used are C_1 = 22 pF, C_2 = 27.2 pF, L_D = 0.15 nH, R_D = 2900 Ω, R_S = 40 Ω, V_G = 1.8 V, and V_{DD} = 3.3 V.

The simulated oscillator output response from Sentaurus is displayed in Figure 47. The oscillator has a sinusoidal oscillating output waveform from 0.5 V to 2.9 V. To analyze the reliability effect on the Colpitts oscillator, transient gate-source stress and drain-source stress are depicted in Figure 48. Examining the voltage waveforms in Figure 47 and Figure 48, one can define three key points a, b, and c (i.e., the bottom, middle, and top of the output voltage) to

probe impact ionization, field, current density, and lattice temperature at these three critical time points for later use.

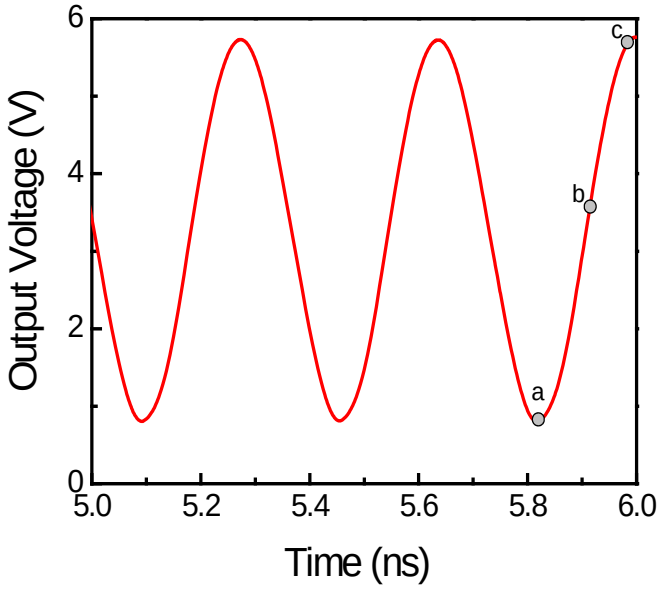


Figure 47: Oscillator output response from mixed-mode device and circuit simulation.

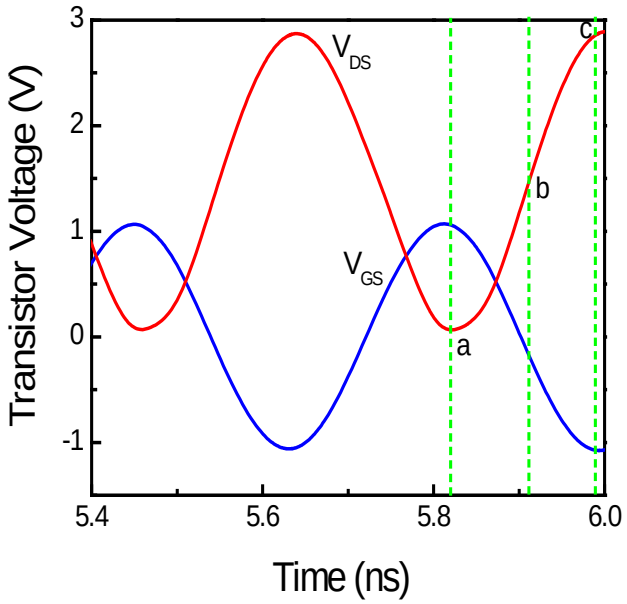


Figure 48: Gate-source and drain-source voltages versus time.

5.2.2 Mixed mode simulation

Hot carrier reliability is increasingly important for digital and RF circuit performance as result of channel length scaling of CMOS devices. Hot carrier injection originates from electrons or holes with excessive kinetic energy in the channel entering oxide layer, producing interface states and oxide trap charges. The MOS transistor's threshold voltage is increased and RF circuit performances degraded eventually.

In order to evaluate the physical insight into the Colpitts oscillator circuit operation, the mixed-mode device and circuit simulation using Sentaurus TCAD software is adopted. In Sentaurus device simulation, Poisson's equation and the electron and hole continuity equations are implemented for charge transport with drift-diffusion model. The Shockley-Read-Hall carrier recombination, Auger recombination, and impact ionization models are used. University of Bologna impact ionization model is applied in this work as the physical model for impact ionization based on impact ionization data generated by the Boltzmann solver[87]. A wide range of electric fields (50 kV/cm to 600 kV/cm) and temperatures (300 K to 700 K) is covered. Calibration against impact ionization measurements was done in the whole temperature range. [88] The low field mobility is calculated by Mathiessen's rule and incorporates the bulk- and surface mobility. Thermodynamic, Thermode, RecGenHeat, and AnalyticTEP models in Sentaurus are included to account for lattice heating,. The thermodynamic model extends the drift-diffusion approach to account for electrothermal effects. A Thermode is a boundary where the Dirichlet boundary condition is set for the lattice. RecGenHeat includes generation-recombination heat sources. AnalyticTEP gives analytical expression for thermoelectric power.

To draw more physical insight into hot electron injection, impact ionization rates at the three different time points are shown in Figure 49, time points are probed as shown in Figure 48.

At point a the electric field is low since drain-source voltage reaches the minimum; however, the current density is very high due to large V_{GS} (see Figure 50). The impact ionization rates at point b are higher than those at points a and c. This is attributed to relatively high drain-source voltage with simultaneously reasonable drain current density at point b, as indicated in Figure 48. The impact ionization rates at point b reach to a peak 1×10^{26} /cm3/s, a precursor of hot carrier effect. Higher drain current enhances I.I. generated carriers under high electric field. At point c the drain-source voltage reaches the maximum and the resulting I.I. rates are high, while the current density is low since the transistor works in cutoff (see Figure 51).

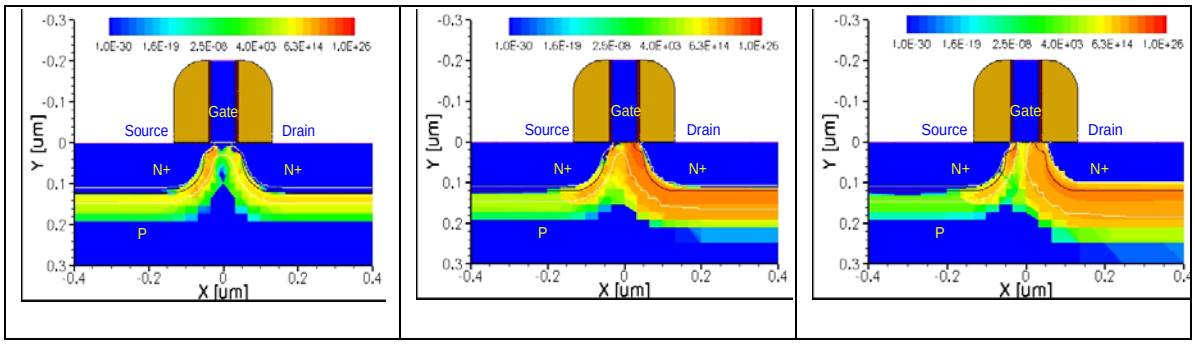


Figure 49: Impact ionization rates at points a, b, and c in Figure 48.

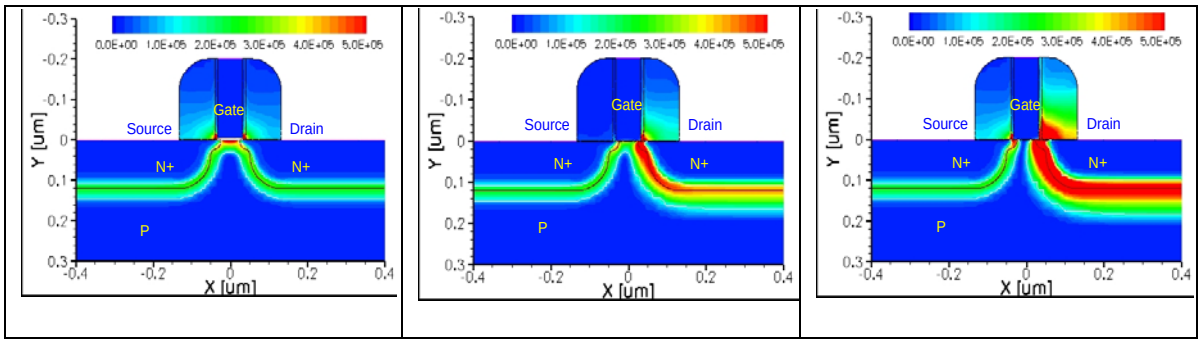


Figure 50: Electric field at points a, b, and c in Figure 48.

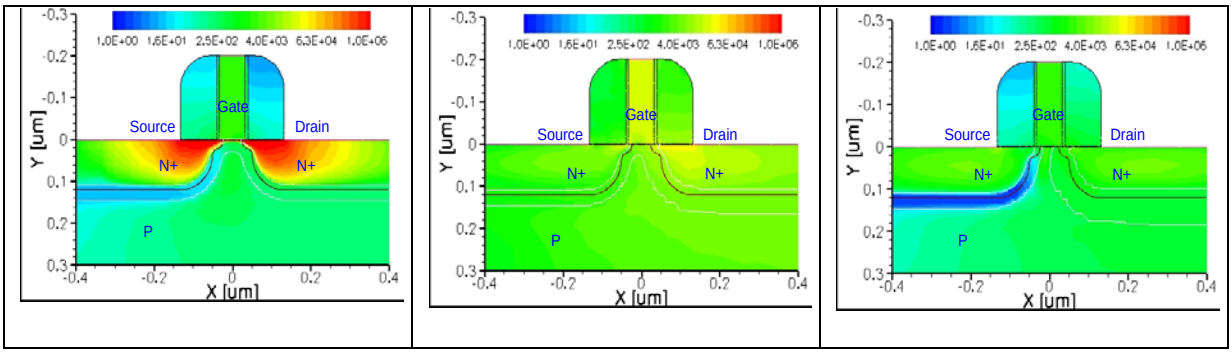


Figure 51: Total current density at points a, b, and c in Figure 48.

The hot electron reliability issue becomes even more important when the channel length of the nMOSFET is decreasing and the supply voltage of the circuit is increasing.

5.3 Phase Noise Modeling and Analysis

Phase noise is one of the most parameters of oscillators and many other modules in communication systems. It describes how pure the signal is and is usually characterized in the frequency domain. As illustrated in Figure 52, in ideal case, an oscillator operating at ω_0, the spectrum assumes the shape of an impulse and the phase noise is negative infinity. However, for a practical oscillator, the spectrum exhibits "skirts" around the center frequency. The phase noise is quantified as the noise power in a unit bandwidth at an offset $\Delta\omega$ divided by the carrier power at ω_0, its unit is dBc/Hz, meaning decibel below carrier for each 1 Hz bandwidth.

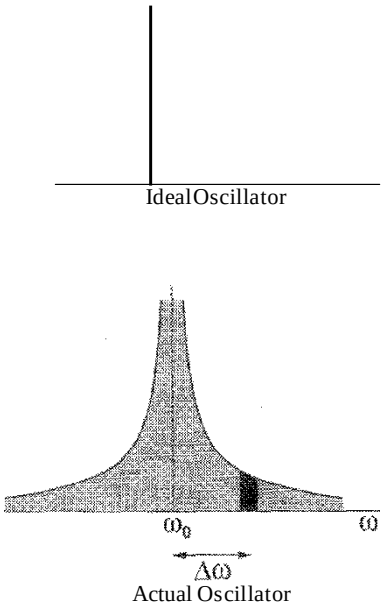


Figure 52: Illustration of Oscillator Phase Noise.

The phase noise of Colpitts oscillator shown in Figure 46 is analyzed for examining device parameter variations. The phase noise to account for MOS transistor parameter drift due to aging is expressed as[89]:

$$L(\Delta f) = 10\log\left(\frac{\overline{V}_n^2}{2\overline{V}_{\tan k}^2}\right)$$

$$= 10\log\left\{\left[\left(\frac{\left|g_{m(1)}\right|^2 K_f}{4C_{ox}WL\Delta f} + \sum_{n=1}^{\infty}\left|g_{m(n-1)} + g_{m(n+1)}\right|^2 \times \frac{kT\gamma}{\overline{g}_m}\right)\alpha + \frac{kT}{R}\right]\left(\frac{Rf_0}{Q\Delta f A_S}\right)^2\right\} \quad (13)$$

Here, V_n is the output noise voltage, V_{tank} is the signal voltage of the oscillator output, $g_{m(n)}$ is the n^{th} Fourier coefficients of transconductance, K_f is a process dependent constant on the order of 10^{-25} V^2F, C_{ox} is the gate oxide capacitance per unit area, W and L are the channel width and length of the MOS transistor, f_0 is the center frequency, Δf is the offset frequency from center frequency, K is the Boltzman's constant, T is the absolute temperature, γ is a coefficient

(about 2/3 for long-channel transistors and larger for submicron MOSFETs), $\overline{g}_m$ is the average transconductance of the transistor, α is is the transfer parameter from nonlinear network port to linear network port ($\alpha = (1\text{-}F)^2$, where F = C1/(C1+C2)), R is the parasitic resistance in the LC tank, Q is the quality factor of LC tank, and A_S is the amplitude of the AC voltage at the source of the transistor.

In (13),

$$\overline{g}_m = \beta A_S(\sin\theta - \theta\cos\theta)/\pi,\ \ \beta = \mu_{n0}C_{ox}W/L,\ \ \theta = \cos^{-1}[(V_T - V_G)/A_S\},\beta = \mu_{n0}C_{ox}W/L ,$$

$g_{m(n)} = g_{m(-n)}$,

and

$$g_{m(n)} = \begin{cases} \beta A_S[\dfrac{(\sin\theta - \theta\cos\theta)}{\pi}] & for\ n = 0 \\ \beta A_S[\dfrac{(\theta - \sin\theta\cos\theta)}{\pi}] & for\ n = 1 \\ 2\beta A_S[\dfrac{\sin n\theta\cos\theta - n\cos\theta\sin\theta}{n(n^2-1)\pi}] & for\ n \geq 2 \end{cases} \qquad (14)$$

ADS simulation has been done with the Colpitts oscillator shown in Figure 46. To repeat what has been done in the mixed-mode simulation, the same circuit element values as in mixed-mode simulation are used in the ADS circuit simulation. The simulated transient output waveform and its Spectral density are depicted in Figure 53 and Figure 54. The oscillation frequency measured from Figure 54 is 2.4 GHz and its fundamental signal spectral power is -4 dBm at 2.4 GHz.

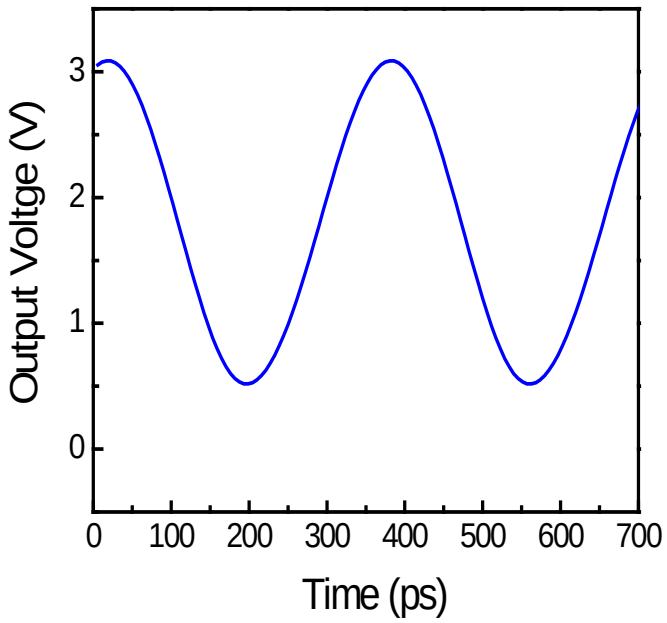


Figure 53: Simulated output waveform versus time.

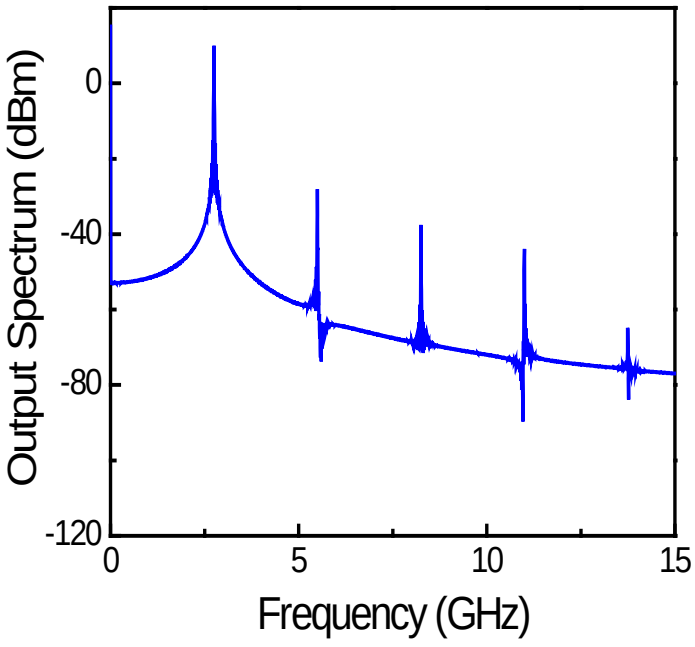


Figure 54: Simulated output power spectrum characteristics.

The phase noise predicted by the analytical equation in (13) is compared with the ADS simulation result in Figure 55. The solid circles in Figure 55 represent the model predictions and the solid line represents the ADS simulation. As seen in Figure 55, the phase noise decreases with offset frequency, as expected. A good agreement between the model predictions and ADS simulation results before hot electron stress is achieved. As shown in mixed mode simulation, there are noticeable I.I. effects when drain source voltage and drain current are both high. High I.I rate may generate lots of hot electrons and holes which will lead to HCI effect. Hot electrons may overcome the barrier at the interface layer and get trapped in the oxide, increasing the threshold voltage. On the other hand, hot holes near the drain area may generate excess interface states and degrade the SiO2 and Si interface. These interface traps can capture or release charge carriers from or to the channel. Since this is a random process with relatively long time constants, the resulting fluctuations in the drain current manifest themselves as low frequency noise with 1/f frequency dependence. Hot-carrier induced traps are known to produce the same kind of noise from original traps at the SiO2 and Si interface[90]. Furthermore, 1/f noise can be up-converted to the phase noise close to the carrier. Thus, the phase noise will increase after hot electron stress. K_f factor in (13) is a parameter to indicate the quality of interface layer. When the interface is degraded by HCI effects, K_f will go up. The HCI effect on the phase noise is also predicted by the analytical model using multiple K_f values. As seen in Figure 55 the phase noise increases with increasing K_f factor, which is related to the interface quality or interface states between the SiO2 and Si interface. As hot carriers generate more interface states at the SiO2 and Si interfacial layer, the K_f factor increases, thus the 1/f noise of the MOSFET and phase noise of the oscillator increase.

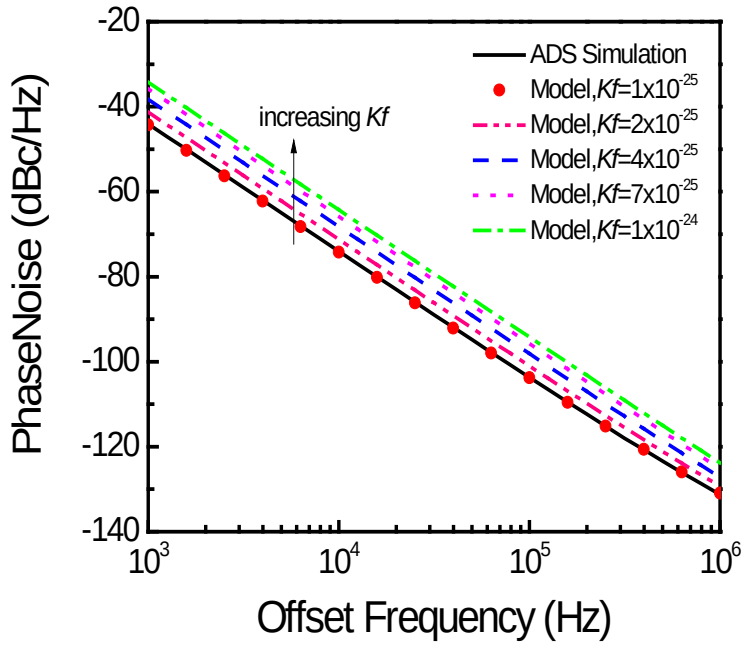


Figure 55: Phase noise modeling versus offset frequency including K_f effect.

5.4 Adaptive Body Biasing and Monte Carlo Simulation

5.4.1 Adaptive Body Biasing

An adaptive body bias scheme is introduced to reduce the process variation effect on this Colpitts oscillator. As shown in Figure 56 the body bias of M1 is determined by the adaptive body bias circuit in the dashed oval circle.

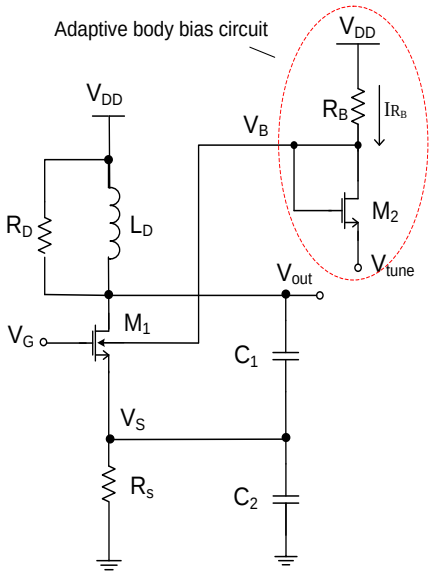


Figure 56: Colpitts oscillator with adaptive body bias.

To account for the body bias effect, the threshold voltage of M_1 can be written as

$$V_T = V_{T0} + \gamma(\sqrt{2\phi_F - V_{BS}} - \sqrt{2\phi_F}) \tag{15}$$

Here γ is the body effect factor, ϕ_F is the Fermi potential, and VBS is the body-source voltage. The drain current of the MOSFET including the body bias effect can be approximately as

$$I_{DS} \approx \frac{\mu_{n0} C_{ox} W}{2L[1+\theta_1(V_{GS}-V_T)+\theta_2 V_{BS}]}(V_{GS}-V_T)^2 \tag{16}$$

The transconductance based on the derivative of drain current with respect to gate-source voltage is derived as

$$g_m \equiv \frac{\partial I_{DS}}{\partial V_{GS}} = \frac{\mu_{n0} C_{ox} W}{2L} \frac{2(V_{GS}-V_T)(1+\theta_2 V_{BS})+\theta_1(V_{GS}-V_T)^2}{[1+\theta_1(V_{GS}-V_T)+\theta_2 V_{BS}]^2} \tag{17}$$

The transconductance equation taking the body bias into account in (17) is then used in the phase noise prediction in (13) for the oscillator with an adaptive body bias.

The use of adaptive body bias to reduce process variability effect on the Colpitts oscillator can be explained as follows: The threshold voltage shift including body bias effect can be expressed as

$$\Delta V_T = \Delta V_{T0} - \frac{\gamma \times \Delta V_{BS}}{2\sqrt{2\phi_F - V_{BS}}} \tag{18}$$

Where ΔV_{T0} is the threshold voltage shift from original process variations, it is assumed to be same for both M_1 and M_2. ΔV_{BS} is Body bias variation due to current variation of M_2 in body bias circuit from similar process variation. The minus sign of the second term in (18) indicates that the body bias effect provides a compensation effect for threshold voltage variations of main transistor M_1 from process uncertainties. For example, if the threshold voltage of M_1 in Figure 56 is increased by the process variations, the body bias V_B of M_1 increases due to less $I_{D2}R_B$ ohmic loss in the adaptive body bias circuit. The overall V_T in M1 is reduced to compensate the initial increase in V_T. On the other hand, when process variability decreases the V_T in M_1, the adaptive body bias circuit will decrease the V_B to M_1. This in turn results in V_T increase in M_1 to compensate the initial decrease in V_T.

5.4.2 Monte Carlo Simulation

Monte Carlo analysis is associated with simulating the design over a given number of trials. In each trial the yield variables have values that distribute randomly with specified probability distribution functions. To further examine the process variation and compensation results on Colpitts oscillator, Monte Carlo (MC) circuit simulation has been performed. In ADS the Monte Carlo simulation assumes statistical variations (Gaussian distribution) of transistor model parameters such as the threshold voltage, mobility, and oxide thickness.

Monte Carlo simulation obtains the overall performance variation by randomly varying network parameter values according to statistical distributions. Monte Carlo yield analysis methods have been widely utilized as an efficient approach to estimate yield. It performs a series of trials from randomly generating yield variable values according to statistical-distribution specifications. Simulation is performed and results evaluated against stated performance specifications.

In this work, the sample size is set as 1000, the initial values of V_{T0}, μ_0, and t_{ox} are 0.42 V, 491 $cm^2/V \cdot s$, and 1.85 nm, respectively. The statistical variations for V_{T0}, μ_0, and t_{ox} are set at +-10%, +-5%, and +-3% with Gaussian distribution from 65 nm technology node. The phase noise variation is displayed in Figure 57. In this histogram plot, the x-axis shows the phase noise distribution of the oscillator and y-axis displays the probability density of occurrence. The mean value of phase noise is -121 dBc/Hz and the standard deviation is 0.71 dBc/Hz.

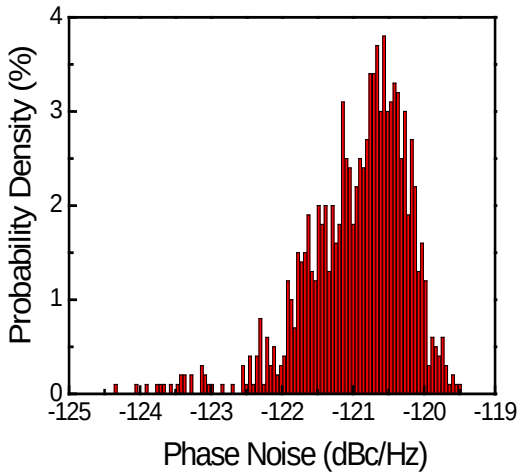


Figure 57: Phase noise distribution (Δf is at 400 kHz).

The Monte Carlo simulation result of the Colpitts oscillator including the body bias effect is shown in Figure 58. In this histogram plot the mean value of phase noise is also -121dBc/Hz

and the standard deviation is 0.18 dBc/Hz. The phase noise is evaluated at the offset frequency of 400 kHz.

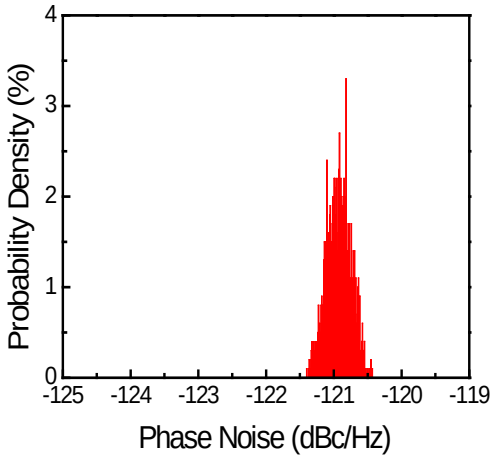


Figure 58: Phase noise distribution using the adaptive body bias scheme (Δf is at 400 kHz).

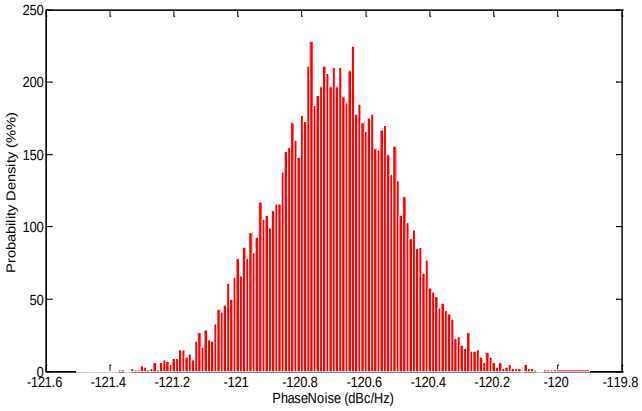


Figure 59: Matlab modelilng with only V_{TH} variation considered.

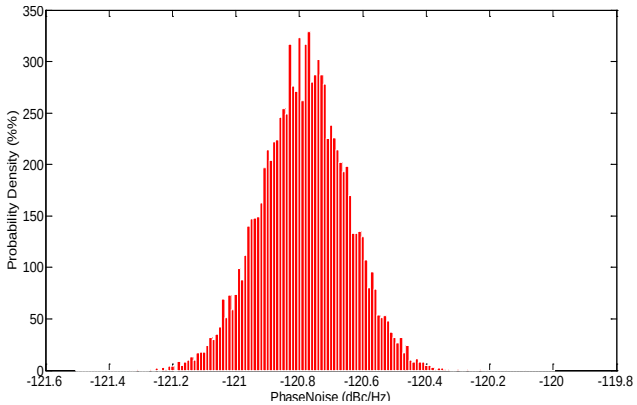


Figure 60: Matlab modelilng with only V_{TH} variation considered w/ body biasing.

To further verify the compensation results of adaptive body biasing technique, the phase noise model illustrated in equation 13 is utilized again, for simplicity, only V_{TH} is considered to have Gaussian distribution with the mean value of 0.42 V, and the standard deviation is set to be 5 %, that is 0.021 V, 10000 samples were calculated, and the probability density of phase noise with and without adaptive body biasing are plotted in Figure 59 and Figure 60. With the same mean value of phase noise, which is -120.7dBc/Hz, the standard deviation is reduced from 0.1944 to 0.136 with the added body biasing.

5.5 Summary

In this chapter, nanoscale CMOS reliability and process variation effect on Colpitts oscillator was studied. We have examined reliability issues such as hot electron effect on Colpitts oscillator by mixed-mode device and circuit simulation to investigate the physical insight of impact ionization. Analytical equation of phase noise was derived and modeled. The analytical model predicts that the phase noise increases after hot electron stress due to an increased interface states and 1/f noise up-conversion to the carrier. A good agreement between simulation and modeling is reached. Variability of Colpitts oscillator is simulated using Monte Carlo

algorithm. As evidenced by Monte Carlo simulation results, the Colpitts oscillator is also sensitive to process variations. An adaptive body bias to minimize process variation effect on the Colpitts oscillator was also proposed and is proved by ADS and Matlab simulation.

CHAPTER SIX：ON CHIP PVT SENSING USING PLL

This chapter is about a robust adaptive design technique to reduce PVT variation effects on RF circuits. The sensing part is accomplished by a PLL. By doing this, we assume there is already PLL existing in the same chip, so no extra design overhead is required. This idea has been applied with digital circuit [91], but not in RF circuit to the author's best knowledge.

6.1 PLL Design

6.1.1 PLL in the system

PLL is a significant part in modern communication system. It is widely used for carrier phase/frequency recovery and synchronization, frequency division and multiplication, demodulation, symbol/bit synchronization, clock recovery etc. PLL is a feedback system, which compares the phase and frequency of input signal with its output. It adjusts the control signal of VCO, makes its output frequency and phase to match with the input. Here is the diagram of a PLL shown in Figure 61. The core component includes PFD (phase frequency detector), CP (charge pump), LPF (low pass filter), VCO (voltage controlled oscillator), and a frequency divider.

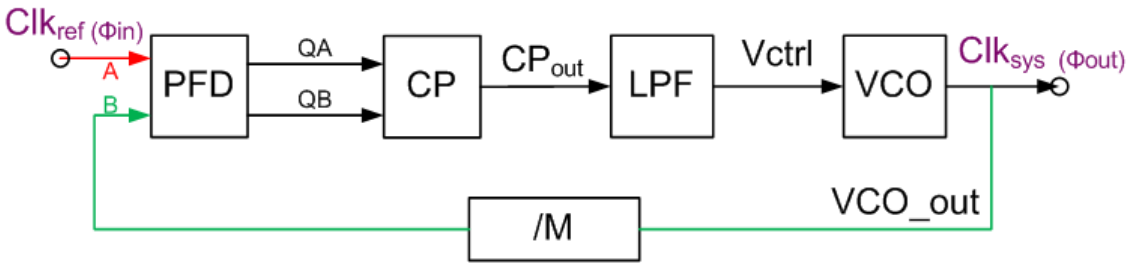


Figure 61: Diagram of PLL.

6.1.2 PFD and Charge pump

A phase frequency detector composed mainly by NAND gate is illustrated in Figure 62. It compares phases of two inputs and generates two outputs that are not complementary. The average output is linear to the phase difference of two inputs. The output is positive if the input has higher frequency, zero if lagging. The width is equal to the phase difference between the two inputs.

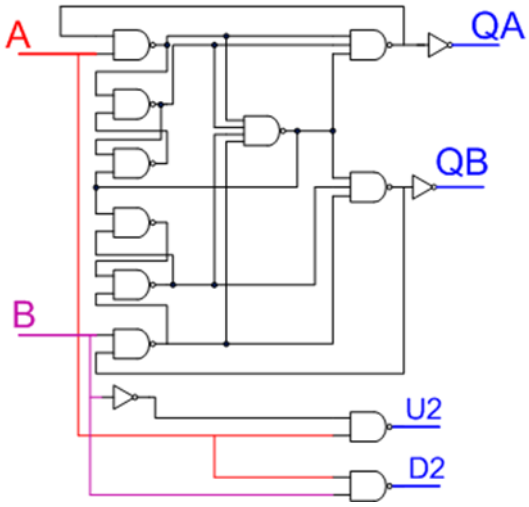


Figure 62: Architecture of Phase Frequency Detector.

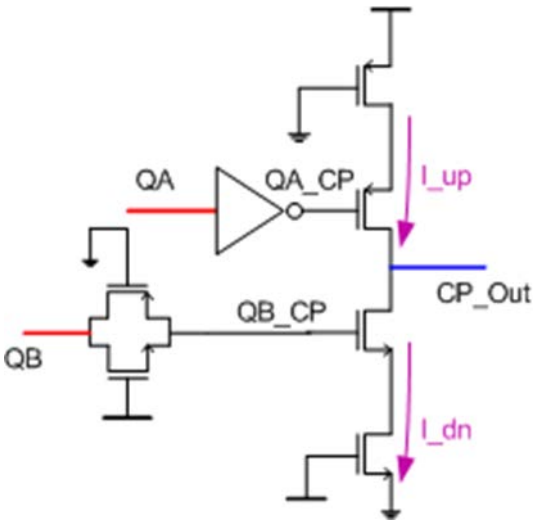


Figure 63: Architecture of charge pump.

Figure 63 shows the architecture of a simple charge pump which transfers the pulses of PFD output into a specific control voltage. [92] Two switched current source pumps charge into and out of loop filter according to two logical inputs. To balance the delay difference of these two inputs, a complementary pass gate is inserted.

6.1.3 VCO

A three stage current starved ring oscillator as shown in Figure 64 is used in this PLL. [93] The output voltage of charge pump is filtered by a second order low pass filter and used to control the output frequency. The output frequency can be expressed as [94]

$$f_{osc} = \frac{I_S}{NC_L V_{DD}} = \frac{\mu_n C_{ox} W_0 (V_{CTRL} - V_{T0})^2}{2 L_0 N C_L V_{DD}} \quad (19)$$

Where I_s is drain current of nMOS transistor, N is the number of stages in the oscillator, N=3 here. C_L is the load capacitance for each inverter stage, V_{DD} is supply voltage, μ_n is the mobility of nMOS transistor, V_{T0} is the threshold voltage of nMOS transistor, C_{ox} is gate oxide capacitance per unit area, W_0 and L_0 are the width and length of the nMOS transistor respectively.

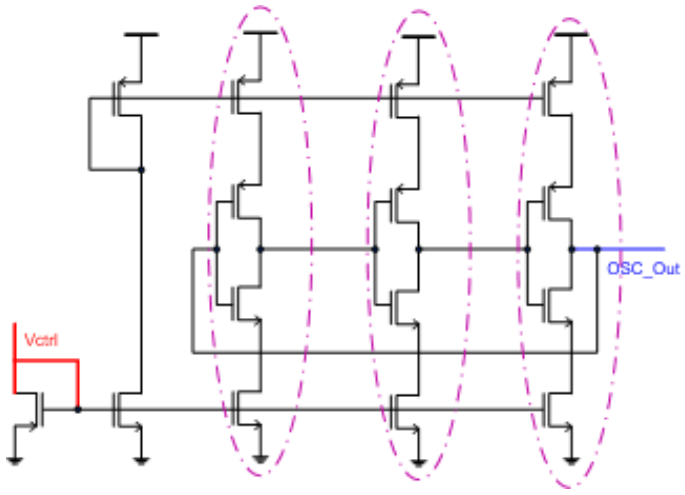


Figure 64: 3 stage current starve ring oscillator.

6.1.4 Simulation results of PLL

After properly tuning, the PLL is locked with the input signal set at a period of 0.45 ns. Figure 65 shows the time it takes the control signal V_{out} to stabilize which is about 200 ns. Figure 66 illustrates the two input signals as well as VCO output when PLL is unlocked (left) and locked (right). The phase error is (0.009/0.45)*2*pi ≈ 7.2 °.

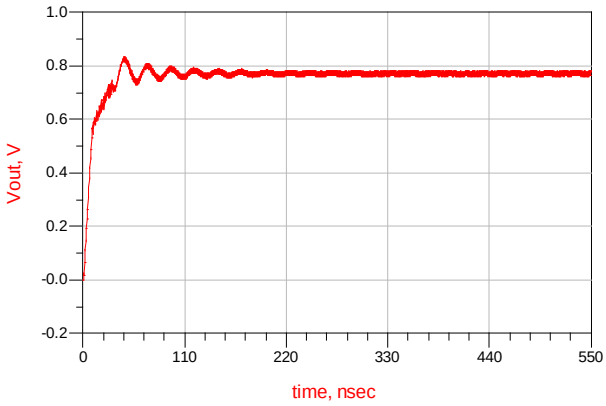


Figure 65: V_{ctrl} vs. Time.

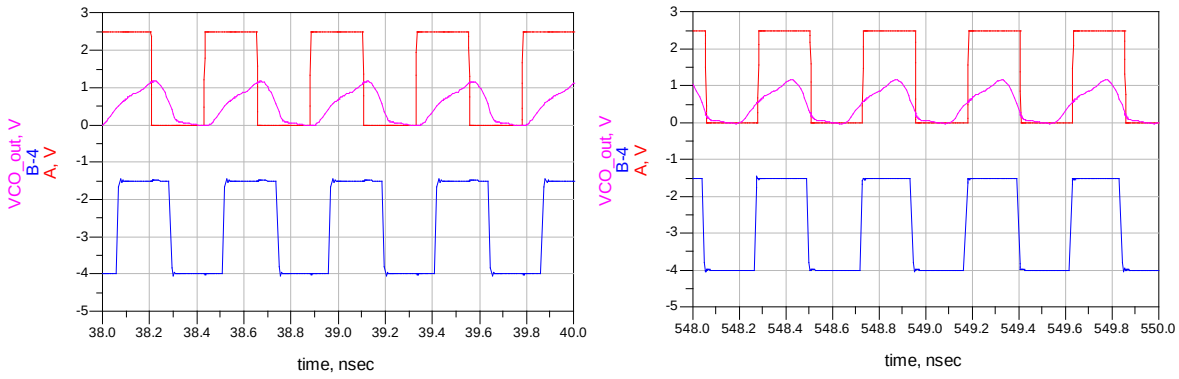


Figure 66: Unlocked and locked input/output.

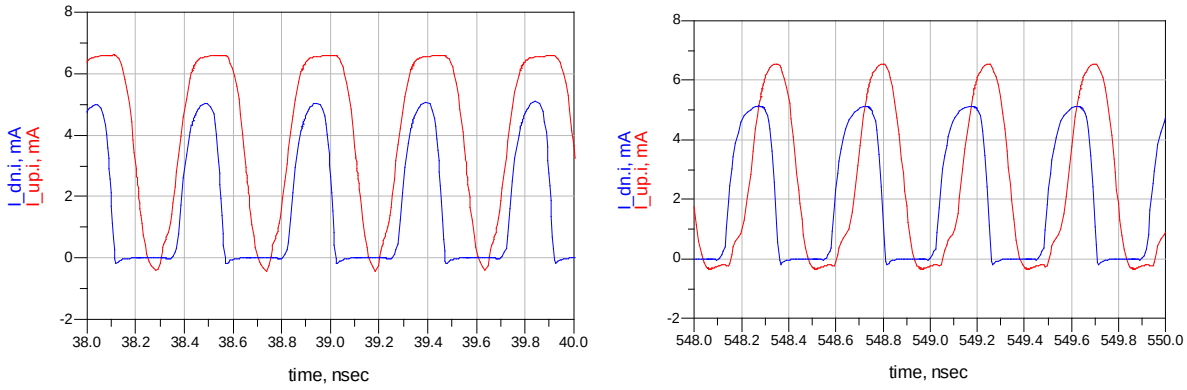


Figure 67: Unlocked and locked charge pump current.

A key effort in designing this PLL is choose the right charge pump structure and transistor sizing. When the PLL is locked, the current flow through the up and down transistor in each cycle should be equal. The left side of Figure 67 is the current through two switch transistors, one can see that there is big mismatch between these two and the total current in one cycle is not equal. This is during the time the PLL is unlocked. This is correspondent to the part when V_{out} is climbing in Figure 65. While the right side is the current when PLL is locked, the net current injected by the CP now is zero. The aforementioned phase error is created here since the average value of control signal V_{ctrl} needs to be constant, corresponding to the flat region in Figure 65.

6.2 PLL to sense PVT and Process variation

6.2.1 Why PLL can be used for sensing

A chip in operation is subjected to multiple sources of variations: process variation, supply voltage variation, temperature instability, transistor aging (δV_{th}, $\delta \mu_n$,). These variations will reflect themselves in circuit performances instability. For example, frequency of a free running ring oscillator which is described in equation (19) is affected by these variations through unstable mobility μ_n, threshold voltage V_{T0} , supply voltage V_{DD} etc.

However, for a PLL whose input frequency is constant during operation to keep locked, f_{osc} of its inside oscillator should be exactly the same as the input frequency. In this PLL designed, ring oscillator has a V_{ctrl} signal which shift with these variation mentioned above, so that it has stable output frequency no matter what the circuit suffers from. From equation (19), V_{ctrl} can be transformed into

$$V_{CTRL} = \sqrt{\frac{2L_0NC_LV_{DD}f_{osc}}{\mu_nC_{ox}W_0}} + V_{T0} = \sqrt{k\frac{V_{DD}}{\mu_n}} + V_{T0} \tag{20}$$

Here, $k = \frac{2L_0NC_Lf_{osc}}{C_{ox}W_0}$

To simplify the analysis, we assume V_{ctrl} signal is influenced by V_{dd}, V_{T0}, μ_n only, Thus δV_{ctrl} can be expressed as:

$$\delta V_{CTRL} = \frac{\partial V_{CTRL}}{\partial V_{DD}}\delta V_{DD} + \frac{\partial V_{CTRL}}{\partial \mu_n}\delta \mu_n + \frac{\partial V_{CTRL}}{\partial V_{T0}}\delta V_{T0} \tag{21}$$

And the three parts can be derived separately as:

$$\frac{\partial V_{CTRL}}{\partial V_{DD}} = \frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}}, \qquad \frac{\partial V_{CTRL}}{\partial \mu_n} = -\frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}}, \qquad \frac{\partial V_{CTRL}}{\partial V_{T0}} = 1$$

Plug these three derivatives into equation (21), it can be rewritten as

$$\delta V_{CTRL} = \frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}}\delta V_{DD} - \frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}}\delta \mu_n + \delta V_{T0} \tag{22}$$

From equation (22), one can see that PLL as an on-chip sensor accounts for all these variations mentioned above, using V_{ctrl} as the monitoring signal. As a result, V_{ctrl} signal can be used as a monitor to reflect how the circuit is suffered from: for a fresh chip in a good working condition, V_{ctrl} will be lower; (Since a free running oscillator has higher frequency in good condition, and oscillator output frequency increases with $V_{ctrl.}$) When the chip is degraded or in working in bad environment, V_{ctrl} has to be higher to lift the oscillator output frequency, so that it

can keep up with the input. Thus an on-chip PLL can be used as comprehensive monitor to reflect inside degradation and outside working condition.

6.2.2 PLL sensing capability

Here are a few simulation results showing how V_{ctrl} in a PLL and free-running ring oscillator change with different source of variations: supply voltage, temperature variations, V_{th} and mobility variations and technology corners. As demonstrated in Figure 68 when V_{dd} increases, free-running ring oscillator frequency increases while V_{ctrl} of PLL decreases.

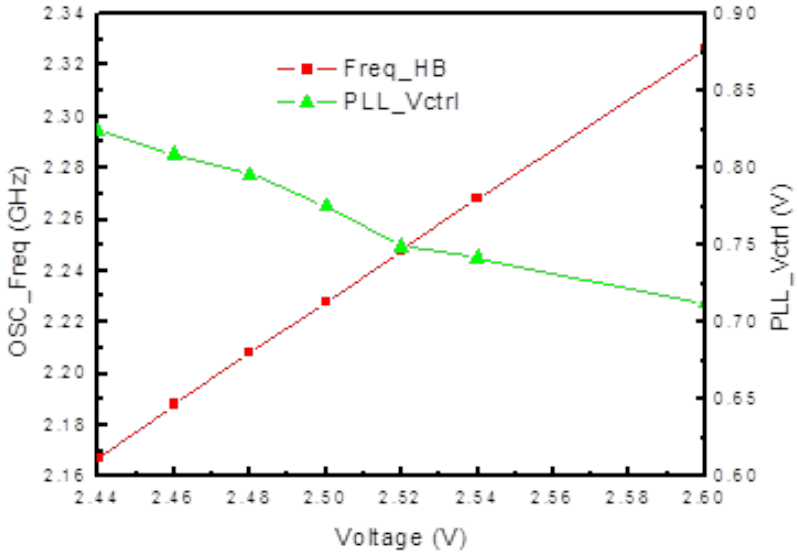


Figure 68: Voltage sensing capability.

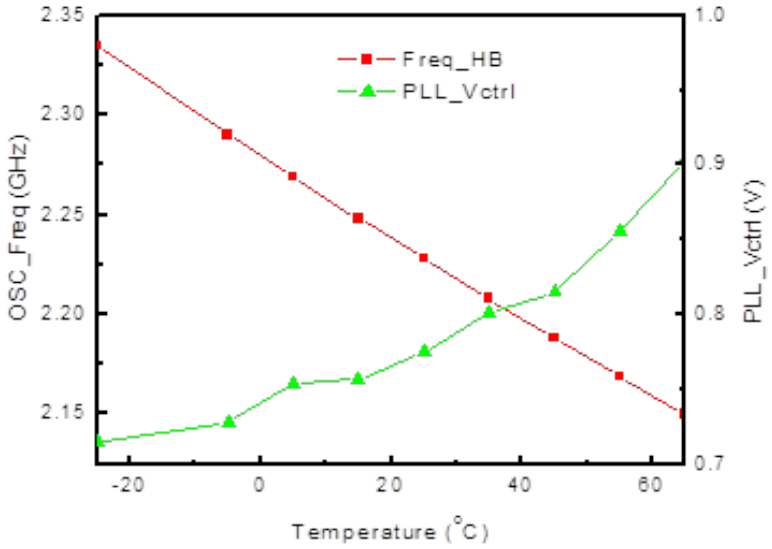


Figure 69: Temperature sensing capability.

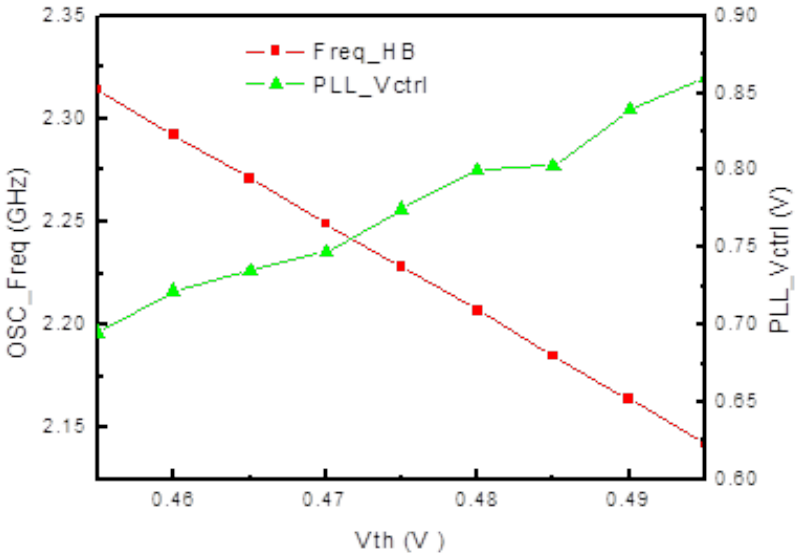


Figure 70: Vth sensing capability.

In figure 69, temperature incensement leads to lower free-running ring oscillator frequency and higer V_{ctrl} of PLL. When V_{th} increases from 0.455 V to 0.495V in Figure 70, free-running ring oscillator frequency decreases while Vctrl increases.

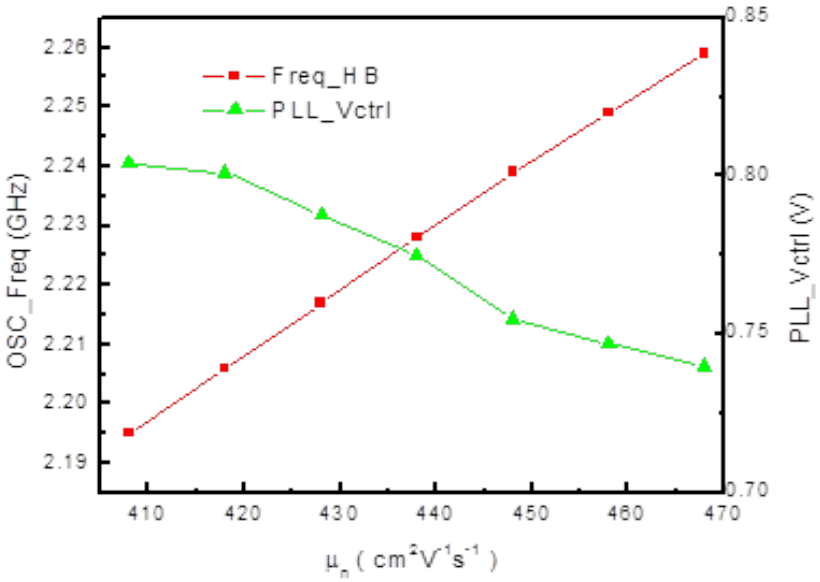


Figure 71: Mobility sensing capability.

Mobility is swept from 400 $cm^2V^{-1}s^{-1}$ to 470 $cm^2V^{-1}s^{-1}$in Figure 71. When the mobility is decreased due to any kind of degradation, the frequency of a free running oscillator decreases. The control signal in a PLL with the constant input increases so that the output frequency can keep up with the input signal.

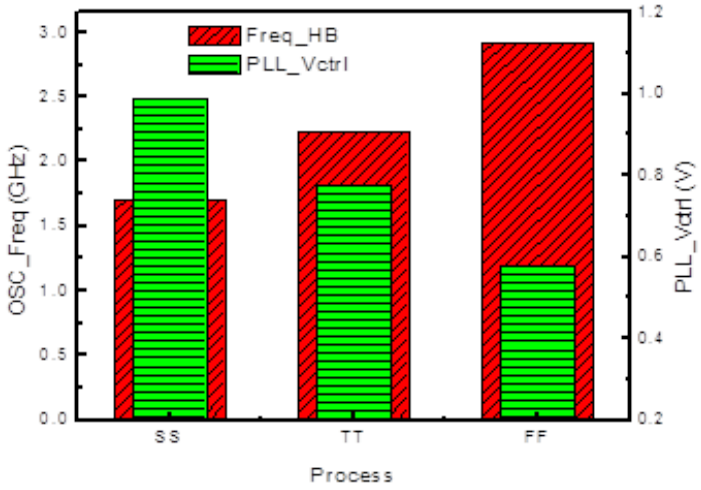


Figure 72: PLL sensing capability at three corners.

With three technology corners, SS, TT, and FF, the free running frequency of ring oscillator and control signal of PLL have the opposite trends too. (Simulation results shown in Figure 72) The faster the technology, higher oscillation frequency, lower control voltage for PLL, vice versa.

6.3 PLL compensated current injection PA

6.3.1 PA design

A simple class AB power amplifier is designed using load pull and source pull in TSMC 0.18 µm RF technology. The architecture is shown in Figure 73. Adaptive body biasing technique has been widely used for digital and RF applications [95], here it is adopted to compensate the output power degradations. With the body biasing effects of NMOS transistor, higher positive body biasing will decrease the threshold voltage, make the main transistor open more, let more drain current go through, thus output more power to the load.

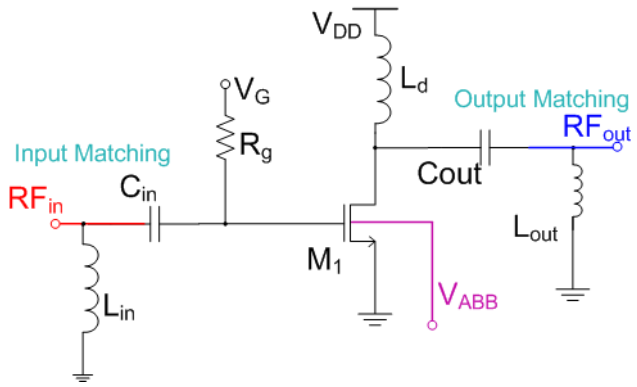


Figure 73: One Stage Class AB PA used for verification.

The following equations explain body effects

$$V_{TH} = V_{T0} + \gamma(\sqrt{2\emptyset_F - V_{BS}} - \sqrt{2\emptyset_F}) \quad (23)$$

Here, V_{T0} is the original threshold voltage of this NMOS transistor, γ is body effect coefficient lies between 0.3 and 0.4 $V^{0.5}$, $\emptyset_F = (kT/q)\ln(N_{sub}/n_{i)}$, N_{sub} is doping concentration of substrate, V_{BS} is body source potential difference. Similarly, if we simplify the analysis, suppose the transistor is only affected by V_{TH}, μ and V_{dd}, then δV_{th} with respect to these three variation sources can be written as:

$$\delta V_{TH} = \frac{\partial V_{TH}}{\partial V_{T0}}\delta V_{T0} + \frac{\partial V_{TH}}{\partial V_{BS}}\delta V_{BS} = \delta V_{T0} - \frac{\gamma \delta V_{BS}}{2\sqrt{2\emptyset_F - V_{BS}}} \quad (24)$$

From equation (24), the total threshold voltage variation comes from two sources. One is threshold variation due to inside degradation and outside environmental change that is shared by all transistors on the chip. The other one comes from the tuning of body bias voltage V_{BS}. If V_{BS} is properly designed that it is controlled by V_{ctrl} signal from PLL as previously described, it is possible that delta V_{th} can be much more flat with respect to all these variations.

With bad working conditions or severe device degradation, ring oscillator (w/o compensation) frequency decreases, output power of PA (w/o compensation) also decreases. However, V_{ctrl} signal from PLL increases, (so that ring oscillator frequency in PLL keeps unchanged), if V_{BS} also increases by certain range, the output power of PA can be kept constant too. On the other hand, in better working environment, both oscillator and PA tend to have higher output, however, the lower V_{ctrl} and V_{SB} signal will drag them down to a lower value, make the output trends with these variations more flat.

6.3.2 Mapping technique

To reach the goal described in the last paragraph of 6.3.1, we firstly simulated PLL with a variety of PVT changes and obtained a group of V_{ctrl} values with respect to these variations. For

example: condition 1 (V_{dd} =2.5V, V_{th} =0.475V,μ_n=435 $cm^2V^{-1}s^{-1}$, TT, 25 ^oC) V_{crtl} =0.775V, condition 2 (V_{dd} =2.5V, V_{th} =0.475V,μ_n=435 $cm^2V^{-1}s^{-1}$, TT, 45 ^oC) V_{crtl} =0.815V…

After that, with the exact same variations applied to PLL, each time PA is simulated with body bias voltage tuned so that output power of PA reaches a normalized value (V_{dd} =2.5V, V_{th} =0.475V,μ_n=435 $cm^2V^{-1}s^{-1}$, TT, 25 ^oC), this group of V_{ABB} are also recorded. For example: condition 1, (V_{dd} =2.5V, V_{th} =0.475V,μ_n=435 $cm^2V^{-1}s^{-1}$, TT, 25 ^oC) V_{ABB} =0.3V, condition 2 (V_{dd} =2.5V, Vth =0.475V,μn=435 $cm^2V^{-1}s^{-1}$, TT, 45 ^oC) V_{ABB}=0.375V…

Then these two groups of V_{ctrl} value and V_{ABB} value are mapped on one chart, the scattered dots were extrapolated linearly and the trend is obtained as shown in Figure 74.

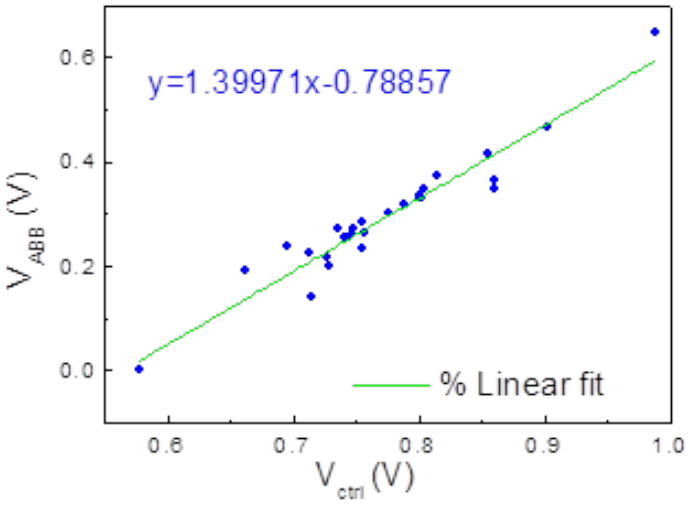


Figure 74: mapping from PLL to ABB.

As we already got V_{ctrl} signal and desired V_{ABB} from separate PLL and PA simulation, the last part is to conduct voltage shifting by a differential amplifier. This conversion is completed by a mapping circuit shown in Figure 75, and the body biasing voltage V_{ABB} is related to V_{CTRL} by

$$V_{ABB} = \frac{R_2}{R_1}(V_{CTRL} - V_{ref}) \quad (25)$$

V_{ref} is a constant DC biasing value set by calculation.

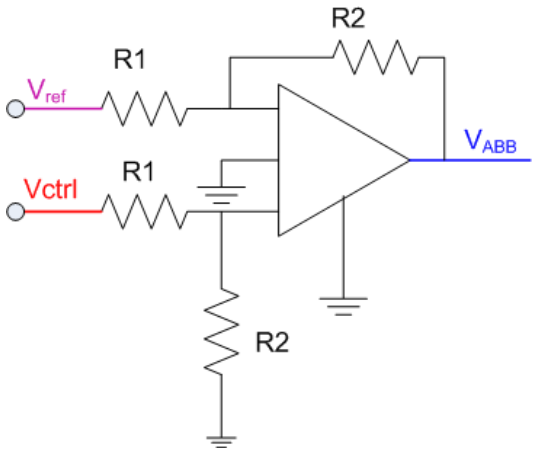


Figure 75: mapping circuit implementation.

V_{ABB} is equal to V_{BS} for the main transistor of PA since the source node is connected to ground here. The variation of V_{BS} can be derived as following:

$$\delta V_{BS} = \frac{\partial V_{BS}}{\partial V_{CTRL}} \delta V_{CTRL} = \frac{R_2}{R_1} \delta V_{CTRL} \tag{26}$$

If we plug in equation (22) from section 6.2.1, yields,

$$\delta V_{BS} = \frac{R_2}{R_1} \left(\frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}} \delta V_{DD} - \frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}} \delta \mu_n + \delta V_{T0} \right) \tag{27}$$

6.3.3 Compensated results

After all preparations described above, the PA circuit with compensation from PLL is simulated and the results are compared with PA without compensation. Before we list the comparison, let's combine equation (24) and equation (27) together, that's how we get the variation of threshold voltage of PA main transistor:

$$\delta V_{TH} = \delta V_{T0} - \frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}} \left(\frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}} \delta V_{DD} - \frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}} \delta \mu_n + \delta V_{T0} \right) \tag{28}$$

For the PA circuit shown in Figure 73, we make a few assumptions to simplify the analysis: (1) V_{DD} variation is linearly proportional to V_{GS}, $V_{GS} = \alpha\delta V_{DD}$,(2)Output power is linearly proportional to drain current of main transistor, $\frac{\Delta P_o}{P_o} = \frac{\Delta I_D}{I_D}$; (3) Main transistor is working in saturation region, $I_D = \frac{1}{2}\mu_n C_{ox}\frac{W_1}{L_1}(V_{GS} - V_{TH})^2$,(4) Drain current I_D fluctuation subjects to key transistor parametric drifts μ_n ,δV_{GS} and δV_{TH}.

Then delta I_D can be expressed as,

$$\delta I_D = \frac{\partial I_D}{\partial \mu_n}\delta\mu_n + \frac{\partial I_D}{\partial V_{GS}}\delta V_{GS} + \frac{\partial I_D}{\partial V_{TH}}\delta V_{TH} \tag{29}$$

And for these three derivatives,

$$\frac{\partial I_D}{\partial \mu_n} = \frac{1}{2}\beta(V_{GS} - V_{TH}) \qquad \frac{\partial I_D}{\partial V_{GS}} = \mu_n\beta \quad \frac{\partial I_D}{\partial V_{TH}} = -\mu_n\beta$$

Here,

$$\beta = C_{ox}\frac{W_1}{L_1}(V_{GS} - V_{TH})$$

Again, plug in these three derivatives into equation (29), which can be re-written as,

$$\delta I_D = \frac{1}{2}\beta(V_{GS} - V_{TH})\delta\mu_n + \mu_n\beta\alpha\delta V_{DD} - \mu_n\beta\left[\delta V_{T0} - \frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}}\left(\frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}}\delta V_{DD} - \frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}}\delta\mu_n + \delta V_{T0}\right)\right]$$

$$= \mu_n\beta\left(\alpha - \frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}}\frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}}\right)\delta V_{DD} + \left[\frac{1}{2}\beta(V_{GS} - V_{TH}) - \mu_n\beta\frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}}\frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}}\right]\delta\mu_n - \mu_n\beta\left(1 - \frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}}\right)\delta V_{T0} \tag{30}$$

If we normalize equation (30) to I_D, it yields:

$$\frac{\delta I_D}{I_D} = \frac{2}{V_{GS} - V_{TH}}\left(\alpha - \frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}}\frac{\sqrt{k}}{2\sqrt{\mu_n V_{DD}}}\right)\delta V_{DD} + \left[\frac{1}{\mu_n} - \frac{\gamma R_2}{R_1(V_{GS} - V_{TH})\sqrt{2\emptyset_F - V_{BS}}}\frac{\sqrt{kV_{DD}}}{2\mu_n^{\frac{3}{2}}}\right]\delta\mu_n - \frac{2}{V_{GS} - V_{TH}}\left(1 - \frac{\gamma R_2}{2R_1\sqrt{2\emptyset_F - V_{BS}}}\right)\delta V_{T0} \tag{31}$$

And if we compare this result to PA without any compensation using the same assumptions, we can similarly arrive at:

$$\delta I_{D0} = \mu_n \beta' \alpha \delta V_{DD} + \frac{1}{2}\beta'(V_{GS} - V_{T0})\delta\mu_n - \mu_n \beta \delta V_{T0}\ , \tag{32}$$

Here, $\beta' = C_{ox}\frac{W_1}{L_1}(V_{GS} - V_{T0})$

Normalized to I_{D0}, gives

$$\frac{\delta I_{D0}}{I_{D0}} = \frac{2}{V_{GS} - V_{T0}}\alpha \delta V_{DD} + \frac{1}{\mu_n}\delta\mu_n - -\frac{2}{V_{GS} - V_{T0}}\delta V_{T0} \tag{33}$$

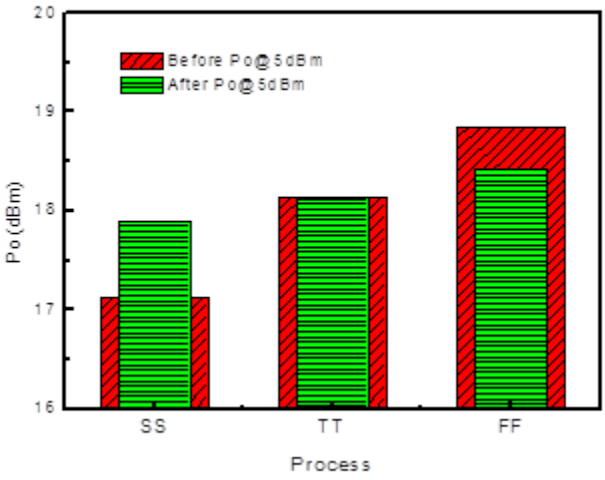


Figure 76: Compensated Pout @ 3 corners.

Finally, the compensated and non-compensated results are shown in Figure 76 to Figure 81. The compensation results turn out to be good. We can see that under each condition there is obvious improvement after the adaptive body biasing is applied. Figure 76 shows how process variation affects the PA performance. The one with compensation appears to be much more insensitive to technology corner change.

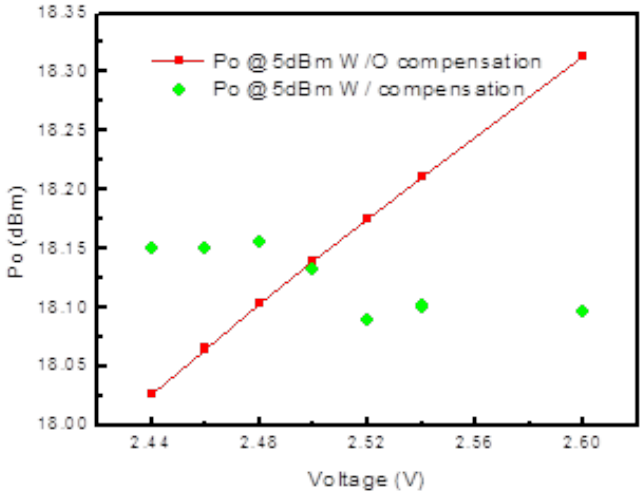


Figure 77: Compensated Pout with VDD variation.

V_{DD} variation is demonstrated in Figure 77. With 5 dBm input, the output power with compensation is more flat, compared to the PA output power without compensation.

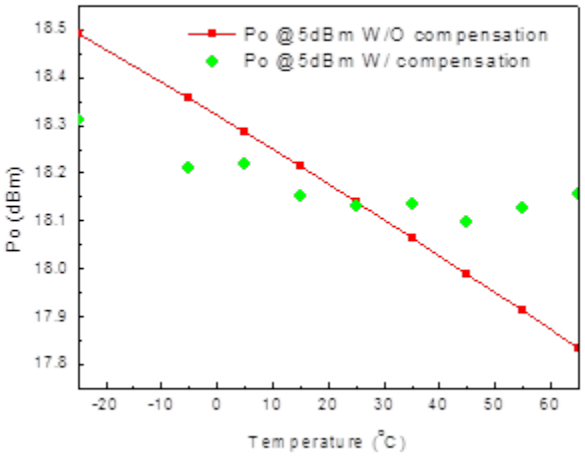


Figure 78: Compensated output power with temperature variation.

Figure 78 is the output power trend with temperature variation. Output power of PA decreases with increased temperature. However, the one with compensation drops much slower.

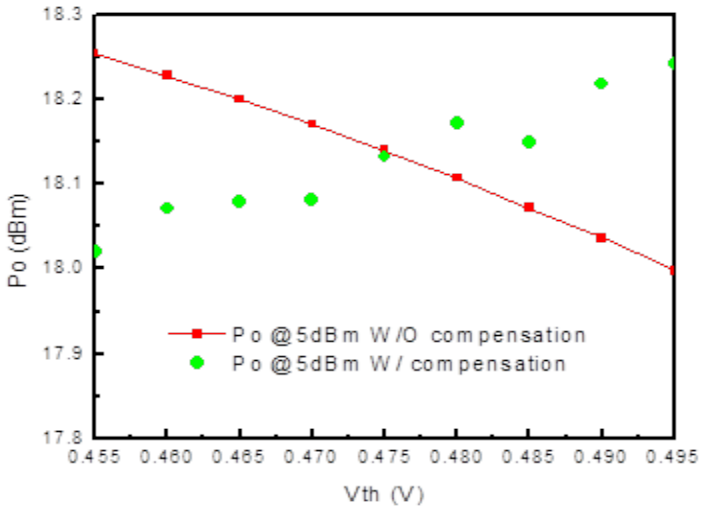


Figure 79: Compensated Pout with Vth variation.

V_{th} variation is presented in Figure 79, one may see that higher V_{th} will lead to less output power. Although the compensation effect is not as obvious as the others, the one with compensation does have more flat trend than the one without.

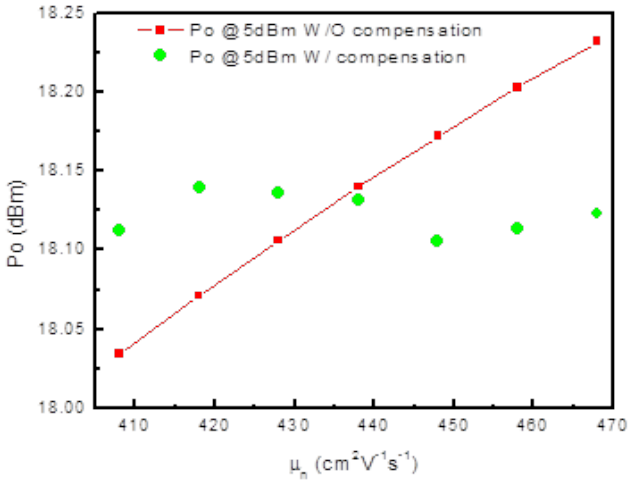


Figure 80: Compensated Pout with mobility variation.

Compensated Pout with mobility variation are plotted in Figure 80. Again, PA with compensation has more flat output power with respect to mobility variation.

If we compare equations (30, 31) with equations (32, 33), one can make at least two observations:

1. Three targeted trends: I_D vs. V_{dd}, I_D vs. μ_n, I_D vs. V_{th}, are consistent with these simulated results. As it is demonstrated that Vdd and mobility increase leads to higher output power, and V_{th} increase leads to lower output power.

2. The canceling effects with respect to each variation comes from the "minus a positive value part" in each bracket. One can predict that with other variations that are not included in this equation analysis, there exist similar canceling mechanisms. That's how PLL can be used as a comprehensive monitor, and V_{ctrl} signal can be applied for comprehensive compensation.

6.4 Summary

Based on the observations that control signal of ring oscillator varies with fluctuations from process, voltage shifts, temperature instability (PVT) to keep PLL locked, the author introduces this new method using V_{ctrl} from PLL to compensate variations in RF circuits such as PA. A free running ring oscillator is simulated against PVT variations; the oscillation frequency variation trends with these variations are obtained. A PLL circuit was designed to generate the desired Vctrl signal. Op amp is utilized to convert Vctrl signal to body biasing of a class AB PA. Simulation results shows that output power as a function of PVT with body biasing becomes more stable compared to the one with no compensation. This technique serves as a candidate for stabilizing RF circuit performances with neglect able design overhead.

CHAPTER SEVEN：CONCLUSIONS

7.1 Accomplishments

In this work, the author evaluated hot electron and oxide stress effects on Class E PA based on a designed chip and reliability experiments. A chip was fabricated using TSMC 0.18 μm mixed-signal RF process. Test results shows that under elevated DC stress for 10 hours, chips behaved different degradations in S21, output power, PAE due to gate oxide stress as well as hot carrier degradation. The more DC stress applied, more degradation observed. Experiments have been supported by mixed mode circuit simulation.

Several temperature compensation techniques established for RF PA, their compensation capability have been compared. Device as well as PA self-heating effects have been simulated in TCAD sentaurus, lattice temperature is proved to increase by 52K while source and drain stress is increased to 4.5 V for a single device, or increase by 55 K in PA circuit application when transient simulation is used. A simple gate biasing circuit turns out to be the most effective among them. The compensation mechanism has been verified with analytical equations;

A Colpitts oscillator was designed in ADS. Process variations and reliability issues featured by phase noise are examined using Monte Carlo simulation. Analytical equations were developed and model in Matlab to support the compensation body biasing circuit. Mixed-mode simulation was carried out to evaluate its impact ionization, field and current distribution inside CMOS transistor.

After that a PLL circuit was studied and a robust, adaptive design technique was introduced into RF circuit. PVT variation effects on RF circuits were analyzed with ADS simulation and analytical equations. PLL used as an on chip monitor can reflect the environment condition and automatically adjust the control voltage of body biasing of RF circuit. Circuit

compensated by this automatic bod biasing technique has more flat results compared with its counterpart without compensation.

7.2 Future Work

With all those contributions mentioned above, this work is not fully completed yet, more effort should be made by the following researchers. Since NBTI has become a field drawing more and more eyes, it needs to be supported by detailed mixed mode, at least device level simulation, provided that Sentaurus has this new feature added on and not enough experimental data handy. This simulation results be observed in mixed mode RF circuit can be applied in more RF circuits performance analysis with regards to NBTI effects.

Reliability research is about the study of different degradation mechanism; predict device as well as circuit lifetime; and propose feasible, effective solutions to increase product yield and eliminate possible loss. Current work has been focused on the study of different types of degradation, and the search for compensation circuits. However, lifetime study for different RF circuits due to different degradation mechanisms still requires more attention.

Besides these, it is still a big task to discover new comprehensive degradation monitor circuit and new compensation method. More RF circuits should be implemented and the monitor and compensation capability should be tested with real chip and compared between them.

LIST OF REFERENCES

[1] K. L. Pey, X. Wu W. H. Liu, X. Li, N. Raghavan, K. Shubhakar, and M. Bosman, "An overview of physical analysis of nanosize conductive path in ultrathin SiON and high-κ gate dielectrics in nanoelectronic devices"

[2] C. Yu and J. S. Yuan, "CMOS device and circuit degradations subject to HfO2 gate breakdown and transient charge-trapping effect," IEEE Trans. Electron Devices, vol. 54, no. 1, pp. 59–67, Jan. 2007.

[3] K. F. Schuegraf and C. Hu, "Metal-Oxide-Semiconductor Field-Effect-Transistor Substrate Current During Fowler-Nordheim Tunneling Stress and Silicon Dioxide Reliability," J. Appl. Phys., vol. 76, no. 6, pp. 3695-3700, 1994.

[4] E. Avni and J. Shappir, "A Model for Silicon-Oxide Breakdown Under High Field and Current Stress," J.Appl.Phys., vol. 64, no. 2, pp. 743-748, 1988.

[5] D. J. Dumin, J. R. Maddux, R. S. Scott, and R. Subramoniam, "A Model Relating Wearout to Breakdown in Thin Oxides," IEEE Trans. Electron Devices, vol. 41, no. 9, pp. 1570-1580, 1994.

[6] Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, and H. E. Maes, "New Insights in the Relation Between Electron Trap Generation and the Statistical Properties of Oxide Breakdown," IEEE Trans. Electron Devices, vol. 45, no. 4, pp. 904-911, 1998.

[7] E. Miranda and J. Sune, "Electron transport through broken down ultrathin SiO2 layers in MOS devices," Microelectron. Reliab., vol. 41, no. 1, pp. 1–23, Jan. 2004.

[8] J. H. Stathis, "Physical and predictive models of ultra thin gate reliability in CMOS devices and circuits," in Proc. IEEE IRPS, 2001, pp. 132–149.

[9] C. H. Tung, K. L. Pey, L. J. Tang, M. K. Radhakrishnan, W. H. Lin, F. Palumbo, and S. Lombardo, "Percolation path and dielectric-BD-induced- epitaxy (DBIE) evolution during ultrathin gate dielectric BD transient," Appl. Phys. Lett., vol. 83, no. 11, pp. 2223–2225, 2003.

[10] C-H Tung, K-L Pey, L. J. Tang, Y. Cao,M. K. Radhakrishnan, and W. H. Lin, " Fundamental Narrow MOSFET Gate Dielectric Breakdown Behaviors and Their Impacts on Device Performance" IEEE Trans. Electron Devices, Vol. 52, No. 4, April 2005

[11] S. Lombardo, F. Crupi, A. La Magna, and C. Spinella. Electrical and thermal transiet during dielectric breakdown of thin oxides in metal-SiO2-silicon capacitors. Journal of Applied Physics, 84(1):472–479, July 1998.

[12] A. Acovic, G. La Rosa and Y-C Sun, "A Review of Hot Carrier Degradation Mechanism In MOSFETs", Microelectron. Reliab., Vol. 36, No. 7/8, pp. 845-869, 1996

[13] S. Mahapatra, , C. D. Parikh, V. R. Rao, C. R. Viswanathan, and J. Vasi" Device Scaling Effects on Hot-Carrier Induced Interface and Oxide-Trapped Charge Distributions in MOSFET's" IEEE Trans. Electron Devices, Vol. 47, No. 4, April 2000

[14] S.Tam, P-K. Ko, and C. Hu,"Lucky-Electron Model of Channel Hot-Electron Injection in MOSFET's" Trans. Electron devices, Vol. ed-31, No. 9, September 1984

[15] C. Hu, SIMON C. TAM, F-C Hsu, P-K Ko,T-Y Chan, and K. W. Terril "Hot-Electron-Induced IMOSFET Degradation— Model, Monitor, and Improvement" IEEE J. of Solid-State Circuits, Vol. Sc-20, No. 1, February 1985

[16] S. Bampi, J.D. Plummet, "Modified LDD device structures for VLSI", in IEDM Tech. Dig., p.234 1985.

[17] N. Kimizuka, et al., "The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling," VLSI Symp. on Tech., pp. 73-74, 1999.

[18] S. Mahapatra, P. B. Kumar, and M. A. Alam, ``Investigation and Modeling of Interface and Bulk Trap Generation During Negative Bias Temperature Instability," Trans. Electron Devices, vol. 51, no. 9, pp. 1371-1379, 2004.

[19] Sentaurus Device User Guide", Version A-2007.12, December 2007

[20] Vattikonda, R.; Wenping Wang; Yu Cao," Modeling and minimization of PMOS NBTI effect for robust nanometer design" , Design Automation Conference, 2006 43rd ACM/IEEE , pp. 1047 – 1052,2006

[21] W. Abadeer and W. Ellis, "Behavior of NBTI under AC dynamic circuit conditions," IRPS, pp. 17-22, 2003.

[22] "Managing Process Variation in Intel's 45nm CMOS Technology" Intel Technology Journal, Vol 12, Issue 2,June 17, 2008

[23] A.R. Brown, G. Roy, A. Asenov, Poly-Si gate related variability in decananometre MOSFETs with conventional architecture. IEEE Trans. Electron. Dev. 54, 3056 (2007)

[24] A. Asenov and B. Cheng," Modeling and Simulation of Statistical Variability in Nanometer CMOS Technologies" www.springer.com/cda/content/.../9789400703902-c2.pdf

[25] Behzad razavi,"Design of Analog CMOS Integrated Circuits", int. edn. mcgraw-hill 2001.

[26] Behzad razavi,"RF microelectronics", 2nd edition.

[27] N.O. Sokal and A.D.Sokal," "Class E - A new class of high-efficiency tuned single-ended switching power amplifiers," IEEE J. Solid-State Circuits, pp. 168-176, 1975.

[28] F.H.Raab, " An introduction to Class F Power Amplifiers," RF Design, pp.79-84, May 1996.

[29] I. J. Bahl, Fundeamental of RF and Microwave Transistor Amplifiers, John Wiley & Sons: New York, 2009.

[30] E. Miranda and J. Sune, "Electron transport through broken down ultra-thin SiO2 layers in MOS devices," Microelectronics Reliability, pp. 1-23, 2004.

[31] L. Pantisano, D. Schreurs, B. Kaczer, W. Jeamsaksiri, R. Venegas, R. Degraeve, K. P. Cheung, G. Groeseneken, "RF performance vulnerability to hot carrier stress and consequent breakdown in low power 90nm RFCMOS," Tech. Dig. Int. Electron Devices Meetings, 2003, pp. 181-184.

[32] Z. Chen, K. Hess, J. Lee, J. W. Lyding, E. Rosenbaum, I. Kizilyalli, S. Chetlur, and R. Huang, "On the mechanism for interface trap generation in MOS transistors due to channel hot carrier stressing," IEEE Electron Device Lett., pp. 24-26, 2000.

[33] F. H. Raab, "Idealized operation of the class E tuned power amplifier," IEEE Trans. Circuits Syst., pp. 725-735, 1977.

[34] A. Mazzanti, L. Larcher, R. Brama, and F. Svelto, "Analysis of reliability and power efficiency in cascode class-E PAs," IEEE J. Solid-State Circuits, pp. 1222-1229, 2006.

[35] Y. Song, S. Lee, E. Cho, J. Lee, and S. Nam, "A CMOS class-E power amplifier with voltage stress relief and enhanced efficiency," IEEE Trans. Microwave Theory and Techniques, pp. 310-317, 2010.

[36] L. Larcher, D. Sanzogni, R. Brama, A. Mazzanti, and F. Svelto, "Oxide breakdown after RF stress: Experimental analysis and effects on power amplifier operation," Int. Rel. Phys. Symp., 2006, pp. 283-288.

[37] M. Apotolidou, M. P. van der Heijden, D. M. W. Leenaerts, J. Sonsky, A. Heringa, and I. Volokhine, "A 65 nm CMOS 30 dBm class-E RF power amplifier with 60% PAE and 40% PAE at 17 dB back-off," IEEE J. Solid-State Circuits, pp. 1372-1379, 2009.

[38] http://www.agilent.com/find/eesof-ads

[39] http://www.synopsys.com

[40] D. B. M. Klaassen, "A unified mobility model for device simulation – I: Model equations and concentration dependence," Solid-State Electron., pp. 953–959, 1992.

[41] T. Quémerais, L. Moquillon, V. Huard, J.-M. Fournier, P. Benech, N. Corrao, and X. Mescot, "Hot-carrier stress effect on a CMOS 65-nm 60-GHz one-stage power amplifier," IEEE Electron Device Lett., pp. 927-929, 2010.

[42] C.-H. Liu, R.-L. Wang, Y.-K. Su, C.-H. Tu, and Y.-Z. Juang, "Degradation of CMOS power cells after hot-carrier and load mismatch stresses," IEEE Electron Device Letts., pp. 1068-1070.

[43] P. Aminzadeh, M. Alavi, and D. Scharfetter, "Temperature dependence of substrate current and hot carrier-induced degradation at low drain bias," Symposium on VLSI Technology, 1998, pp. 178-179.

[44] M. Kimura, "Field and temperature acceleration model for time-dependent dielectric breakdown," IEEE Trans. Electron Devices, pp. 220-229, 1999.

[45] T. Pompl, H. Wurzer, M. Kerber, and I. Eisele "Investigation of ultra-thin gate oxide reliability behavior by separate characterization of soft breakdown and hard breakdown" Tech. Dig., Int. Phys. Symp., 2000, pp. 40-47.

[46] L. Pantisano and K.P Cheung, "The impact of postbreakdown gate leakage on MOSFET RF performances", IEEE Electron Device Lett., pp. 585-587, 2001.

[47] C.-H. Liu, R.-L. Wang, and Y.-K. Su, "DC and RF degradation induced by high RF power stresses in 0.18-μm nMOSFETs," IEEE Trans. Device and Materials Reliability, pp. 317-323, 2010.

[48] P. Roussel, R. Degraeve, C. van den Bosch, B. Kaczer, and G. Groeseneken, "Accurate and robust noise-based trigger algorithm for soft breakdown detection in ultra thin oxides," Tech. Dig. Int. Rel. Phys. Symp., 2001, pp. 386-392.

[49] A. Avellán and W. H. Krautschneider, "Impact of soft and hard breakdown on analog and digital circuits," IEEE Trans. Device and Materials Reliability, pp. 676-680, 2004.

[50] R. Rodríguez, J. H. Stathis, and B. P. Liner, "A model for gate-oxide breakdown in CMOS inverters," IEEE Electron Device Lett., pp. 114-116, 2003.

[51] O. Semenov, A. Vassighi, and M. Sachdev, "Impact of self-heating effect on long-term reliability and performance degradation in CMOS circuits," IEEE Trans. Device and Materials Reliability, pp. 17-27, 2006.

[52] R. Wang, J. Zhuge, R. Huang, D.-W. Kim, D. Park, and Y. Wang, "Investigation on self-heating effect in gate-all-around silicon nanowire MOSFETs from top-down approach," IEEE Electron Device Lett., pp. 559-561, 2009.

[53] P. Aminzadeh, M. Alavi, and D. Scharfetter, "Temperature dependence of substrate current and hot carrier-induced degradation at low drain bias," Symposium on VLSI Technology, 1998, pp. 178-179.

[54] P. Su, K. Goto, T. Sugii, and C. Hu, “Self-heating enhanced impact ionization in SOI MOSFETs,” IEEE Int. SOI Conf., 2001, pp. 31-32.

[55] N. S. Waldron, A. J. Pitera, M. L. Lee, A. Fitzgerald, and J. A. del Alamo, ”Positive temperature coefficient of impact ionization in strained-Si,” IEEE Trans. Electron Devices, pp. 1627-1633, 2005.

[56] D. Vasileska, “Self-heating in SOI nano devices,” IEEE Nanotechnology Materials and Devices Conf., 2010, pp. 389-394.

[57] M. Braccioli, G. Curatola, Y. Yang, E. Sangiorgi, and C. Fiegna, “Simulation of self-heating effects in 30nm gate length FinFET,” Int. Conference on ULSI, 2008, pp. 71-74.

[58] S. Boumaiza and F. M. Ghannouchi, “Thermal memory effects modeling and compensation in RF power amplifiers and predistortion linearizers,” IEEE Trans. Microwave Theory and Techniques, pp. 2427-243, 2003.

[59] W.-L. Chen, S.-F. Chang, K.-M. Chen, G.-W. Huang, and J.-C. Chang, “Temperature effect on Ku-band current-reused common-gate LNA in 0.13-µm CMOS technology,” IEEE Trans. Microwave Theory and Techniques, pp. 2131-2138, 2009.

[60] K. Yamauchi, Y. Iyama, M. Yamaguchi, Y. Ikeda, S. Urasaki, and T. Takagi, “X-band MMIC power amplifier with an on-chip temperature-compensation circuit,” IEEE Trans. Microwave Theory and Techniques, Vol. 49, No.12, pp. Dec 2001.

[61] D. Gómez, M. Sroka, and J. Jiménez. “Process and temperature compensation for RF low-noise amplifiers and mixers,” IEEE Trans. Circuits and Systems -I: Regular Papers, pp. 1204-1211, 2010.

[62] I. M. Filanovsky and A. Allam, “Mutual compensation of mobility and threshold voltage temperature effects with applications in CMOS circuits communications". IEEE Trans. Circuits and Systems -I: Fundamental Theory and Applications, pp. 876-884, 2001.

[63] C. Fiegna, Y. Yang, E. Sdngiorgi, and A. G. O’Neill, “Analysis of self-heating effects in ultrathin-body SOI MOSFETs by device simulation,” IEEE Trans. Electron Devices, pp. 233-244, 2008.

[64] C.-H. Choi, J.-H. Chun, and R. W. Dutton, “Electrothermal characteristics of strained-Si MOSFETs in high-current operation,” IEEE Trans. Electron Devices, pp. 1928-1931.

[65] M. Shrivastava, M. S. Baghini, D. K. Sharma, and V. R. Rao, “A novel bottom spacer FinFET structure for improved short-channel, power-delay, and therm performance,” IEEE Trans. Electron Devices, pp. 1287-1294, 2010.

[66] Synopsis website: http://www.synopsis.com, 2010.

[67] D. B. M. Klaassen, “A unified mobility model for device simulation – I: Model equations and concentration dependence,” Solid-State Electron., pp. 953–959, 1992.

[68] S. Sze, Physics of Semiconductor Devices, 2nd edition, Wiley: New York, 1981.

[69] http://edocs.soco.agilent.com/display/ads2009/BSIM4+Model-BSIM4+MOSFET+Model

[70] H. Kawasaki, V.S. Basker, T. Yamashita, C.-H. Lin, Y. Zhu, J. Faltermeier, S. Schmitz, J. Cummings, S. Kanakasabapathy, H. Adhikari, H. Jagannathan, A. Kumar, K. Maitra, J. Wang, C.-C. Yeh, C. Wang, M. Khater, M. Guillorn, N. Fuller, J. Chang, L. Chang, R. Muralidhar, A. Yagishita, R. Miller, Q. Ouyang, Y. Zhang, V.K. Paruchuri, H. Bu, B. Doris, M. Takayanagi, W. Haensch, D. McHerron, J. O'Neill, and K. Ishimaru, “Challenges and solutions of FinFET integration in an SRAM cell and a logic circuit for 22 nm node and beyond,” Proceedings of Int. Electron Devices Meet., 2009 pp. 1-4.

[71] N. Kumar, V. D. Maheta, S. Purawat, A. E. Islam, C. Olsen, K. Ahmed, M. A. Alam, and S. Mahapatra, “Material dependence of NBTI physical mechanism in silicon oxynitride (SiON) p-MOSFETs: A comprehensive study by ultra-fast on-the-fly (UF-OTF) IDLIN technique,” Tech. Dig., Int. Electron Devices Meet., 2007, pp. 809-812.

[72] A. Cathignol, B. Cheng, D. Chanemougame, A. R. Brown, K. Rochereau, G. Ghibaudo, and A. Asenov, “Quantitative evaluation of statistical variability sources in a 45-nm technological node LPN-MOSFET,” IEEE Electron Device Letters, pp. 609-611, 2008.

[73] S. K. Saha, “Modeling process variability in scaled CMOS technology,” IEEE Design & Test of Computers, pp. 8-16, March-April 2010.

[74] G. D. Panagopoulos and K. Roy, “A three-dimensional physical model for Vth variations considering the combined effect on NBTI and RDF,” IEEE Trans. Electron Devices, pp. 2337-2346, 2011.

[75] B. Vaidyanathan, S. Bai, and A. S. Oates, “The relationship between transistor-based and circuit-based reliability assessment for digital circuits,” Int. Rel. Phys. Symp., pp. 706-709, 2011.

[76] R. Rodriguez, J. H. Stathis, B. P. Linder, S. Kowalczyk, C. T. Chuang, R. V. Joshi, G. Northrop, K. Bernstein, A. J. Bhavnagarwala, and S. Lombardo, “The impact of gate-oxide breakdown on SRAM stability,” IEEE Electron Device Lett., pp. 559-561, 2002.

[77] S. Naseh, M. J. Deen, and C.-H. Chen, “Effects of hot-carrier stress on the performance of CMOS low-noise amplifiers,” IEEE Trans. Device and Materials Reliability, pp. 501-508, 2005.

[78] C. Yu and J. S. Yuan, “MOS RF reliability subject to dynamic voltage stress - modeling and analysis,” IEEE Trans. Electron Devices, vol. 52, pp. 1751-1758, 2005.

[79] S. Naseh, J. Deen, and O. Marinov, “Effects of HC stress on the performance of the LC-tank CMOS oscillatgors,” IEEE Trans. Electron Devices, pp. 1334-1339, 2003.

[80] Y. Li, C.-H. Hwang, and T.-Y. Li, “Random-dopant-induced variability in nano-CMOS devices and digital circuits,” IEEE Trans. Electron Devices, pp. 1588-1597, 2009.

[81] M. Hansson and A. Alvandpour, “Comparative analysis of process variation impact on flip-flop power performance,” IEEE Int. Symp. Circuits and Systems, 2007, pp: 3744-3747.

[82] R. Rao, K. A. Jenkins, and J.-J. Kim, “A local random variability detector with complete digital on-chip measurement circuitry,” IEEE J. Solid-State Circuits, pp. 2616-2623, 2009.

[83] S. Mukhopadhyay, K. Kim, and C. Chuang, “Device design and optimization methodology for leakage and variability reduction in sub-45-nm FD/SOI SRAM,” IEEE Trans. Electron Devices, pp.152-162, 2008.

[84] D. Gómez, M. Sroka, J. Luis, and G. Jiménez, “Process and temperature compensation for RF low-noise amplifiers and mixers,” IEEE Trans. Circuits and Systems—I: Regular Papers, pp. 1204-1211, 2010.

[85] D. Han, B. Kim, and A. Chatterjee, “DSP-driven self-tuning of RF circuits for process-induced performance variability,” IEEE Trans. Very Large Scale Integration Systems, pp. 305-314, 2010.

[86] Y. Liu and J. S. Yuan, “CMOS RF power amplifier variability and reliability resilient biasing design and analysis,” IEEE Trans. Electron Devices, pp. 540-546, 2011.

[87] M. C. Vecchi and M. Rudan, “Modeling electron and hole transport with full-band structure effects by means of the spherical-harmonics expansion of the BTE,” IEEE Trans. Electron Devices, pp. 230-238, 1998.

[88] M. Valdinoci et al., “Impact-ionization in silicon at large operating temperature,” in Int. Conf. Simulation of Semiconductor Processes and Devices (SISPAD), Kyoto, Japan, 1999, pp. 27-30.

[89] Z. Li, J. Bao, P. Tang, and F. Wang, “Analysis of phase noise spectrum in LC oscillator based on nonlinear method,” Int. Conference on Microwave and Millimeter Wave Technology, 2010, pp. 649-652.

[90] R. Brederlow, W. Weberm D, Scgnjutt-Landsiedel, and R. Thewes, “Hot-carrier degradation of the low-frequency noise in MOS transistors under analog and RF operating conditions,” IEEE Trans. Electron Devices, pp. 1588-1596, 2002.

[91] K. Kang, , S. Park, K. Kim, and K. Roy, " On-Chip Variability Sensor Using Phase-Locked Loop for Detecting and Correcting Parametric Timing Failures" IEEE Trans. Very Large Scale Integration Systems. Vol.18, No.2, 2010.

[92] Behzad Razavi, "Design of Analog CMOS Integrated Circuits", McGraw-Hill International edition.

[93] C. Zhang, A. Srivastava, H-C Wu, "Hot-Electron-Induced Effects on Noise and Jitter in Submicron CMOS Phase-Locked Loop Circuits", 48th Midwest Symposium on Circuits and Systems, pp. 507 - 510 Vol. 1 2005.

[94] Xuan Zhang and Alyssa B. Apsel "A Low-Power, Process-and-Temperature-Compensated Ring Oscillator With Addition-Based Current Source", IEEE Transactions on Circuit and System I: Regular Papers, Vol. 58, No. 5, May 2011.

[95] Yidong Liu, Jiann-Shiun Yuan, "CMOS RF Low-Noise Amplifier Design for Variability and Reliability," IEEE Transactions on Device and Materials Reliability. Vol 11, Issue 3, pp.450-457, 2011