Amplifier Design for Differential Temperature Sensors in Built-In Testing Applications

A Thesis Presented

by

Junpeng Feng

to

The Department of Electrical and Computer Engineering

in partial fulfillment of the requirement

for the degree of

Master of Science

in

Electrical Engineering

in the field of

Electronic Circuits and Semiconductor Devices

Northeastern University
Boston, Massachusetts

November, 2013
© Copyright 2013 by Junpeng Feng
All Rights Reserved
Abstract

Low-cost radio frequency (RF) communication circuits are frequently fabricated in complementary metal-oxide-semiconductor (CMOS) technology. As system-on-chip and system-in-package designs are becoming increasingly complicated, the manufacturing test cost is continuing to make up a larger portion of the total product cost while the need for built-in testing and calibration methods rises to ensure reliable operation during the lifetimes of chips. Accordingly, on-chip performance monitoring of individual analog blocks within RF chips is beneficial for identification of faulty devices and system-level self-calibration.

A non-intrusive built-in testing method for on-chip performance monitoring of analog/RF integrated circuits is advanced in this research. A sensor circuit was designed to monitor signal power dissipation and linearity characteristics based on electro-thermal coupling instead of an electrical connection in order to avoid impact on the performance of the device under test (DUT). With this approach, a bipolar junction transistor is placed in the vicinity of the DUT as thermal detector. The sensor circuit transmits a detected temperature change that reflects the power dissipation in the DUT. Two amplifiers are used in the sensor core to process the currents from the temperature-sensing transistors. In this thesis, a class AB output stage was developed for these amplifiers, which allowed to extend the dynamic range and enhance the sensitivity of the sensor circuit. Designed in 0.18µm CMOS technology, one version of the temperature sensor has a simulated sensitivity up to 207.1mV/ºC while consuming 2.23mW of power. A low-power version has a simulated sensitivity of 185.7mV/ºC with 1.13mW power consumption. The sensor was re-designed in 0.11µm CMOS technology for fabrication of a prototype chip with a power consumption of 0.6mW. This sensor has a measured sensitivity of 395.4mV/mW to power dissipation in a DUT that is located in its vicinity on the chip.
Acknowledgements

First of all, I thank my research advisor Prof. Marvin Onabajo. His constructive suggestions and encouragements lead me to make progress with this research. In addition, his guidance helped me to realize where my passion is and what research area I am going to concentrate on in the future. I would also like to thank the members of my thesis committee (Prof. Yong-Bin Kim, Prof. Fabrizio Lombardi) for discussions of my research results and their valuable advices, as well as Analog & Mixed-Signal Integrated Circuit Research (AMSIC) Laboratory members Chun-hsiang Chang, Hari Chauhan, Yuchi Ni, and Alireza Zahrai. I would like to thank Prof. Yong-Bin Kim for providing access to the silicon area for the prototype chip fabrication in 0.11μm CMOS technology through his research sponsor Techwin Corp.

Junpeng Feng

Boston, MA
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 MOTIVATION</td>
<td>1</td>
</tr>
<tr>
<td>1.2 BACKGROUND</td>
<td>4</td>
</tr>
<tr>
<td>1.2.1 Electrothermal Coupling</td>
<td>4</td>
</tr>
<tr>
<td>1.2.2 RF Detectors</td>
<td>8</td>
</tr>
<tr>
<td>1.3 OUTLINE OF THIS THESIS</td>
<td>12</td>
</tr>
<tr>
<td>2. BUILT-IN TESTING OVERVIEW</td>
<td>14</td>
</tr>
<tr>
<td>2.1 PROCESS VARIATIONS</td>
<td>14</td>
</tr>
<tr>
<td>2.2 RF CIRCUIT PERFORMANCE SENSORS</td>
<td>17</td>
</tr>
<tr>
<td>2.2.1 Power Detectors</td>
<td>17</td>
</tr>
<tr>
<td>2.2.2 Non-intrusive Sensors</td>
<td>22</td>
</tr>
<tr>
<td>2.3 NON-INFRINGEMENT TESTING BASED ON ELECTRO-THERMAL COUPLING</td>
<td>26</td>
</tr>
<tr>
<td>2.3.1 Thermal Coupling</td>
<td>26</td>
</tr>
<tr>
<td>2.3.2 Bipolar Junction Transistor Model</td>
<td>29</td>
</tr>
<tr>
<td>2.3.3 RF Power Detection with On-Chip Temperature Sensors</td>
<td>31</td>
</tr>
<tr>
<td>3. DIFFERENTIAL TEMPERATURE SENSOR DESIGN</td>
<td>34</td>
</tr>
<tr>
<td>3.1 INTRODUCTION</td>
<td>34</td>
</tr>
<tr>
<td>3.1.1 Existing Temperature Sensors</td>
<td>34</td>
</tr>
<tr>
<td>3.1.2 Reference Sensor Design</td>
<td>39</td>
</tr>
<tr>
<td>3.2 PROPOSED DESIGN</td>
<td>44</td>
</tr>
<tr>
<td>3.2.1 Class AB Output Stage</td>
<td>44</td>
</tr>
<tr>
<td>3.2.2 Common-mode Feedback Circuit Design</td>
<td>48</td>
</tr>
<tr>
<td>3.2.3 Stability</td>
<td>52</td>
</tr>
<tr>
<td>3.2.4 Low-Power Sensor Design and Comparison</td>
<td>64</td>
</tr>
<tr>
<td>3.2.5 Fabricated Sensor Version</td>
<td>66</td>
</tr>
<tr>
<td>4. SIMULATION RESULTS</td>
<td>71</td>
</tr>
<tr>
<td>4.1 COMMON-MODE FEEDBACK STABILITY</td>
<td>71</td>
</tr>
<tr>
<td>4.2 AMPLIFIER NOISE</td>
<td>73</td>
</tr>
<tr>
<td>4.3 SENSOR PERFORMANCE</td>
<td>75</td>
</tr>
<tr>
<td>4.4 CALIBRATION</td>
<td>89</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 1.</td>
<td>Target application example: typical transceiver with thermal detectors.</td>
<td>2</td>
</tr>
<tr>
<td>Fig. 2.</td>
<td>Block diagram of a receiver with thermal sensing for calibration.</td>
<td>2</td>
</tr>
<tr>
<td>Fig. 3.</td>
<td>Thermal impedance model for a unit volume of the silicon die (substrate).</td>
<td>4</td>
</tr>
<tr>
<td>Fig. 4.</td>
<td>Electrothermal coupling between a device under test and the sensor.</td>
<td>6</td>
</tr>
<tr>
<td>Fig. 5.</td>
<td>Block diagram of a power amplifier with integrated power detector.</td>
<td>10</td>
</tr>
<tr>
<td>Fig. 6.</td>
<td>Experimental setup for the on-chip characterization of an LNA [20].</td>
<td>12</td>
</tr>
<tr>
<td>Fig. 7.</td>
<td>Specification variation impact on product yield [26].</td>
<td>16</td>
</tr>
<tr>
<td>Fig. 8.</td>
<td>Typical RMS detector design approach (from [20]).</td>
<td>19</td>
</tr>
<tr>
<td>Fig. 9.</td>
<td>Simplified diagram of the RMS detector in [20].</td>
<td>19</td>
</tr>
<tr>
<td>Fig. 10.</td>
<td>Characteristic curves of RMS detectors at the input and output of an LNA.        (conceptual summary of the results from [20])</td>
<td>21</td>
</tr>
<tr>
<td>Fig. 11.</td>
<td>Simplified one-dimensional thermal equivalent circuit for modeling heat diffusion [6].</td>
<td>29</td>
</tr>
<tr>
<td>Fig. 12.</td>
<td>Simplified PNP hybrid-π small-signal model for electro-thermal coupling.</td>
<td>30</td>
</tr>
<tr>
<td>Fig. 13.</td>
<td>Class A power amplifier.</td>
<td>33</td>
</tr>
<tr>
<td>Fig. 14.</td>
<td>Simplified schematic of a differential temperature sensor.</td>
<td>35</td>
</tr>
<tr>
<td>Fig. 15.</td>
<td>Temperature sensor with Darlington-pair PNP transistors.</td>
<td>37</td>
</tr>
<tr>
<td>Fig. 16.</td>
<td>Reference sensor architecture.</td>
<td>39</td>
</tr>
<tr>
<td>Fig. 17.</td>
<td>Schematic of amplifier $A_1$ in [5].</td>
<td>41</td>
</tr>
<tr>
<td>Fig. 18.</td>
<td>CMFB circuit.</td>
<td>42</td>
</tr>
</tbody>
</table>
Fig. 19. (a) Typical class B push-pull output stage, and (b) typical class AB output stage. ................................................................. 45

Fig. 20. Class AB output stage used in the proposed design. ............................................ 46

Fig. 21. Small-signal model of the class AB output stage. ................................................ 47

Fig. 22. Proposed amplifier for each stage (TSMC 0.18μm technology). ................. 47

Fig. 23. High-gain amplifier and a simplified model......................................................... 49

Fig. 24. Example: differential amplifier stage with different currents (I_A, I_B) in each branch due to device mismatches. ................................................................. 51

Fig. 25. Example CMFB circuit. ......................................................................................... 52

Fig. 26. Simulated open-loop frequency response of the uncompensated amplifier. 53

Fig. 27. Locations (A, B, C, D) of significant poles in the signal path of the proposed amplifier. ........................................................................................................ 54

Fig. 28. Equivalent circuit of Fig. 27 for transfer function analysis. ......................... 57

Fig. 29. Pole and zero locations of the uncompensated amplifier.............................. 58

Fig. 30. Pole and zero summary of the uncompensated amplifier............................. 59

Fig. 31. Simulated open-loop frequency response of the amplifier with a Miller capacitor (MC) between nodes A and C. ................................................................. 60

Fig. 32. Simulated open-loop frequency response of the amplifier with a Miller capacitor (MC) between nodes C and D................................................................. 61

Fig. 33. Simulated open-loop frequency response of the amplifier after adding R_2 in series with the Miller capacitor (C_C\text{-}D)........................................................................ 62

Fig. 34. Pole and zero plot after compensation................................................................. 63
Fig. 35. Pole and zero summary after compensation. .............................................63
Fig. 36. Schematic of the low-power amplifier version..............................................65
Fig. 37. Schematic of the amplifier fabricated with Dongbu 0.11μm CMOS technology.................................................................66
Fig. 38. Simulated open-loop frequency response of the amplifier fabricated in 0.11μm Dongbu CMOS technology.................................................................67
Fig. 39. Layout of the first amplifier stage (fabricated version). ...............................69
Fig. 40. Prototype chip layout ...................................................................................69
Fig. 41. Screenshot of the Cadence schematic (fabricated design in Fig. 38). .........71
Fig. 42. Screenshot of the CMFB loop stability summary............................................72
Fig. 43. Gain and phase of the CMFB loop. ..............................................................72
Fig. 44. Noise simulation summary for the proposed amplifier (including CMFB). 73
Fig. 45. Noise simulation summary for the low-power design (including CMFB). 74
Fig. 46. Noise simulation summary for the fabricated design (including CMFB) ....75
Fig. 47. Simplified small-signal equivalent circuit of the sensing PNP devices.......77
Fig. 48. Closed-loop input impedance, gain and phase of the reference design. ......79
Fig. 49. Closed-loop input impedance, gain and phase of the proposed design. ......80
Fig. 50. Closed-loop input impedance, gain and phase of the low-power design.....81
Fig. 51. Closed-loop input impedance, gain and phase of the fabricated design.....82
Fig. 52. ΔV_o vs. ΔT of the sensors using the highest sensitivity setting and different amplifier designs..........................................................83
Fig. 53. Tuning of I_{core} to adjust the sensitivity of the proposed design. ..............85
Fig. 54. Sensitivity vs. $I_{\text{core}}$ plots. .................................................................86
Fig. 55. Simulated effect of a calibration current..............................................90
Fig. 56. Bonding diagram vs. photo of the packaged die...................................91
Fig. 57. Chip micrographs and layout...............................................................92
Fig. 58. PCB schematic.......................................................................................94
Fig. 59. Assembled PCB for the prototype chip..................................................95
Fig. 60. Measurement setup. ...............................................................................96
Fig. 61. Measured differential output voltage vs. power dissipation in diode- connected transistors ($\text{DIO}_1$ and $\text{DIO}_2$) close to the PNP devices on each side of the sensor ($I_{\text{core}} = 100\mu\text{A}$)........................................98
Fig. 62. Measured differential output voltage vs. power dissipation in test resistors $R_{\text{t1}}$ and $R_{\text{t2}}$ close to the PNP devices on each side of the sensor ($I_{\text{core}} = 300\mu\text{A}$)..........................................................100
Fig. 63. Asymmetry between test resistor $R_{\text{t1}}$ and $R_{\text{t2}}$..............................100
Fig. 64. Measured sensitivity vs. $I_{\text{core}}$........................................................101
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table I</td>
<td>RF detector performance comparison</td>
<td>21</td>
</tr>
<tr>
<td>Table II</td>
<td>Amplifier component values</td>
<td>68</td>
</tr>
<tr>
<td>Table III</td>
<td>Simulated specifications of the amplifier (A_1) designs</td>
<td>87</td>
</tr>
<tr>
<td>Table IV</td>
<td>Dynamic range comparison with maximum sensitivity</td>
<td>88</td>
</tr>
<tr>
<td>Table V</td>
<td>Comparison over equal wide dynamic range with reduced I_{core}</td>
<td>89</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1 MOTIVATION

Low-cost telecommunication circuits with multi-GHz are widely used, but the test costs have become a large portion (about 40%) of the total product costs for complex mixed-signal integrated circuits. A major challenge is that when System-on-Chip (SoC) or System-on-Package (SOP) designs are highly integrated, many internal blocks are not directly accessible for verification during high-volume manufacturing testing with automatic test equipment (ATE). Consequently, monitoring performances of individual analog blocks within radio frequency (RF) chips is beneficial for identification of faulty devices and self-calibration [1]-[5]. In this thesis, the typical RF receiver chain is used as an example. Fig. 1 visualizes one envisioned application for the built-in temperature sensor that was developed in this research for the purpose of measuring RF circuit performance parameters in integrated transceivers, such as on smart phone boards. As demonstrated in [6], on-chip temperature gradients affect the performances of sensitive RF circuits such as mixers. Other possible applications include RF chips in medical and military communication devices that require high reliability.
Typically, an RF receiver chain consists of at least a low-noise amplifier (LNA), mixer, low-pass filter, and analog-to-digital converter (ADC). The design of RF transceiver front-ends entails numerous challenges, particularly in the case of ultra-wideband (UWB) transceivers due to stringent requirements over wide bandwidths. For instance, the LNA in the receiver path must achieve good performance (i.e., low noise figure and
3

high gain) across the complete frequency spectrum from 3.1GHz to 10.6GHz, regardless of the type of UWB system (i.e., impulse radio or multi-band) being used [7]. High performance requirements give incentives to monitor parameters such as the gain and 1-dB compression point of analog blocks on the chip in order to perform automatic tuning for enhancements in the presence of manufacturing process variations and environmental changes. Fig. 2 depicts the typical block diagram of a single-chip transceiver, which has to be designed to cover the particular bandwidth specified by the given communication standard. Even though receiver implementations vary depending on the application (e.g., for cell phones, Bluetooth, Wireless Local Area Network (WLAN), etc.), the purpose of this simplified general block diagram is to identify the common analog circuits at RF frequencies and baseband frequencies after the down-conversion of the received signal by the mixer, where the filter or an additional baseband amplifier might have variable gain.

Semiconductor device fabrication progress entails continuous shrinking of devices for high-speed operation. However, with each technology node, both process variations and die-to-die variability become worse. [8]. A main consequence is the reduction of yield. Traditionally, automatic-test equipment (ATE) is used to measure the performances of chips after packaging for quality control purposes, but high-volume manufacturing test is more costly as faults and process variations increase. Built-in test (BIT) methods reduce the complexity and cost of external hardware for testing. Moreover, they improve
testability and aid test time reduction ([3]-[4]) for further cost savings. They also reduce the number of extra pins that are necessary to verify a chip, which allows to use smaller and less costly packages. Some alternative built-in test methods are discussed in Chapter 2 where the advantages and disadvantages of ATE and built-in testing are addressed with more details. In the remainder of this chapter, the relationship between built-in testing methods and the use of the proposed on-chip temperature sensor architecture will be introduced.

1.2 BACKGROUND

1.2.1 Electrothermal Coupling

Fig. 3. Thermal impedance model for a unit volume of the silicon die (substrate).
The general built-in test approach in this thesis is based on electro-thermal coupling effects through the silicon substrate of a chip, which can be exploited to monitor the power dissipation in a circuit under test (CUT). The silicon substrate can be modeled with a resistor-capacitor (RC) network for coupled electro-thermal simulations of the CUT [5]-[6]. Fig. 3 provides the RC network model for a location (unit volume) on the silicon substrate in 3-dimensions. In [5]-[6], the cube size was selected with unit dimensions \((x_u \times y_u \times z_u)\) of \(10\mu\text{m} \times 10\mu\text{m} \times 10\mu\text{m}\) to approximate the distance between devices on the chip. Resistors \(R_x, R_y,\) and \(R_z\) in Fig. 3 are thermal resistances, and \(C_E\) is a thermal capacitance. These thermal impedance elements can be calculated to ensure equivalence between the thermal and electrical domains during simulations, as will be elaborated in Section 2.3. In particular, the temperature at a location on the die corresponds to the voltage at the node in the unit element, and the power dissipation of a device at the location of the unit element is modeled by injecting a thermal current [6]. The basic model for the representation in Fig. 3 is discussed in [6], which allows simulations with a conventional electrical simulator to model thermal coupling with equivalent dissipated power (current in the electrical domain), temperature changes (voltage changes in the electrical domain), and thermal impedances (impedances in the electrical domain) using a 3-D model of the die. The power dissipated by any circuit device is injected in the thermal circuit as a current at the node corresponding to its location on the surface layer of the grid. The resulting temperature change (in degrees) at any other grid node can be read as the voltage change at the node of interest in the circuit...
simulation. Such equivalence can also be analyzed with partial differential equations [6] that can be related to the discretized heat equation modeled by the unit elements in the 3-D grid.

Fig. 4. Electrothermal coupling between a device under test and the sensor.

With a model of the silicon die in the thermal domain it can be explained how power dissipation relates to temperature changes in the area close to the CUT, and how these changes can be used to monitor DC and signal power dissipation. In this thesis, the sensor employs vertical bipolar (PNP) transistors that are available in standard complementary metal-oxide semiconductor (CMOS) process technologies to measure the power dissipation of CUTs by sensing local temperature changes of the silicon die. Fig. 4 visualizes the electrothermal coupling that enables the measurement of power
dissipation at the surface of silicon die. For instance, when the shown typical broadband LNA operates, it causes a temperature change in the vicinity of each power-dissipating transistor and resistor, which is injected as thermal current into the corresponding location of the grid in the model. The propagation of this current through the network models the heat flow across the die, such that the voltage change at the test point with the PNP transistor (Q₁) corresponds to the temperature change at this sensing point. An ideal voltage-controlled voltage source with a gain of k (between -1.5mV/Kelvin and -2.0mV/Kelvin) can be used to modulate the base-emitter voltage (V_{be}) of the PNP transistor according to its temperature sensitivity [5]-[6]. Thus, the sensitivity of differential emitter current can be defined as \( S_T = \Delta I_e / \Delta T = k \times g_{mQ} \), where \( \Delta I_e \) is the difference between the emitter currents of Q₁ and Q_{ref}, and \( \Delta T \) is the temperature difference between the locations at which Q₁ and Q_{ref} are placed. Notice that Q_{ref} is ideally located at a reference location that represents the average chip temperature, but with a calibration step it is only required that the reference temperature does not change during the measurement, as elaborated in Section 4.3. After further amplification in the sensor core, the output voltage reflects the measured temperature difference that is proportional to the change in the power dissipation of the CUT between the end of the calibration step and the new condition (e.g., activation of DC bias circuitry or application of a signal).
1.2.2 RF Detectors

RF power, peak, and amplitude detectors play a vital role as signal level meters for automatic gain control and built-in testing, enabling various measurements such as gain, compression, selectivity [9], and impedance matching [10]. Alternatively, on-chip temperature sensors can be used in lieu of electrical RF signal detectors, such as demonstrated in [5] where a method is presented to measure an LNA’s 1-dB compression point through temperature changes as a result of power dissipation.

RF power detectors can be categorized into several groups. References [12], [13] and [14] describe a group of on-chip RF signal detectors, while Schottky diode detectors ([15]-[16]) make up another group. Although, the second group has the advantages of simplicity and high-frequency capability, the drawbacks of power dissipation, impedance-matched design, and large chip area become critical for detection in the gigahertz range [17]. Furthermore, Schottky diodes fabricated in modified CMOS processes are not performing well at gigahertz frequencies due to the lossy substrate and parasitic capacitances [18].

On-chip RF signal level detectors that contain active devices [19]-[25] to perform rectification have been reported in the literature. The characteristics of the detectors depend on their target applications. For example, the work in [19] uses a real-time temperature-compensated CMOS RF detector on the chip to test a power amplifier (PA).
Power detectors that are used in automatic gain control schemes to control the transmit power more precisely require high linearity to detect a wide range of power levels within high-performance systems. Some of the main incentives to integrate power detectors on the same chip as the block being monitored are related to cost reductions and the avoidance of sensitive RF interfaces to external components. However, such power detectors are affected by process variations and the temperature changes on the chip due to self-heating of high-power circuits on the chip [19]. Fig. 5 displays the block diagram for a typical combination of a PA with an on-chip power detector. The input of the power detector is the output of the PA, which requires that the power detector has a high input impedance to avoid loading effects. Since the on-chip power detector accuracy is affected by self-heating of the die from the high power dissipation in the PA, advanced temperature-compensation methods such as in [19] become more important as the level of integration tends to increase in systems-on-a-chip, leading to higher on-chip power densities and heating effects. However, since the electrical power detector is connected to the circuit under test, some loading effects due to the detector’s finite input impedance are unavoidable, especially as operating frequencies continue to increase.
For built-in RF circuit performance testing the integrated RF detector must comply with the following essential characteristics: 1) minimum silicon area overhead; 2) a high input impedance in the frequency range of interest to avoid undesired loading of the RF circuit; 3) a dynamic range suitable for the test of the RF CUT; and 4) a wide band of operation to test at the different frequencies of interest for the CUT/system [20]. Besides the aforementioned matters, the power consumption also has to be considered even though the detectors often only operate during test modes and not continuously as the CUT. For example, the area of the detector in [20] is only 0.031mm$^2$ in a 0.35µm CMOS process, which is small enough for many applications. Usually, small transistors are utilized in the input stage to minimize the parasitic capacitance at the interface to the CUT, but larger transistor dimensions are used within the detector to increase its gain (sensitivity) together with larger capacitors to filter out the DC component after the
rectification stage. In general, DC detector outputs that represent the peak amplitude or root-mean-square (RMS) value of the signal being measured are preferable over AC outputs because they enable the use efficient on-chip analog-to-digital converters to capture the DC outputs. For this reason, many on-chip power detectors involve rectification, filtering, or mixing to down-convert the information to low frequencies or DC [22]-[25].

The temperature-sensing-based detectors in this thesis research are intended for built-in testing of RF and low-frequency analog circuits. Since the power levels at the LNA are low and difficult to detect, it is a good candidate to demonstrate the capabilities of RF detectors. Common on-chip performance measurements are gain and 1-dB compression point characterizations of LNAs. In [20], the authors placed RF amplitude detectors at the input and output of an RF LNA to measure the 1-dB compression point and the gain with adequate accuracy, where the DC output of the detectors monotonically varies with the signal amplitude at their inputs. Fig. 6 depicts this measurement method, which does not involve any switches in the signal path. The buffer is included at the output of the LNA to measure its performance with conventional off-chip equipment for correlation with the on-chip detector outputs. Since the buffer serves only for characterization purposes by driving the external 50Ω input impedance of the analyzer, it is not needed when the RF detectors are employed for built-in testing in the field. Typical electrical RF
detectors for on-chip testing applications operate in the 0.8-10GHz frequency range with a dynamic range of 27-63dB [9], [11], [13], [20].

Fig. 6. Experimental setup for the on-chip characterization of an LNA [20].

1.3 OUTLINE OF THIS THESIS

Chapter 1 briefly introduced the motivation and background of this thesis research. It was discussed why built-in testing techniques are employed in general, and how electrothermal effects can be modeled as well as utilized to perform non-invasive on-chip signal power measurements by monitoring temperature changes. A main goal in this work is the design of an on-chip temperature sensor with wide dynamic range to enable its use as an on-chip RF power detector that does not make electrical contact with the device under test and therefore does not impact RF performance. Chapter 2 describes the built-in testing application to identify design requirements and reference specifications.
for the detector based on temperature sensing. This chapter also elaborates on electro-thermal effects and how they are modeled during the design of on-chip temperature sensors. Chapter 3 presents the proposed sensor design with a focus on the class AB output stage that improves the dynamic range in comparison to the previous sensor architecture. Other design considerations such as common-mode voltage control and stability are also discussed. Chapter 4 includes schematic simulation results for the temperature sensor design with extended dynamic range. Measurement setup and printed circuit board (PCB) design aspects are discussed in Chapter 5, which also contains test chip fabrication and measurements results. Chapter 6 provides conclusions for this research.
2. BUILT-IN TESTING OVERVIEW

2.1 PROCESS VARIATIONS

Process variations naturally occur during the fabrication of integrated circuits. For instance, the length, width and thickness of the transistors may vary according to a Gaussian distribution. This is especially significant for the smaller process nodes (≤65nm) because the percent-variations of parameters are higher compared process nodes with larger minimum feature sizes. Scaling down the sizes of the transistors and other devices has been a trend over the past decades to improve speed/power performance and to save chip area. Meanwhile, the intra-die variability has also increased as a consequence. Since most high-performance analog circuits depend on matched devices and differential signal paths, process variations and die-to-die variability have impacted the yields and reliabilities of chip productions [23], [26]. For instance, the ratio of \( \sigma\{V_{Th}\}/V_{Th} \) (where \( \sigma\{V_{Th}\} \) is the typical transistor threshold voltage standard deviation) increases inversely with the minimum feature size of the technology node. This ratio was 4.7% at the 250nm technology node, and has grown to 16% at the 45nm technology node. Furthermore, the threshold voltage \( V_{Th} \) experiences additional variations due to DC biasing conditions [21] and through the drain-induced-barrier-lowering (DIBL) effect under the large drain-source voltage bias conditions [26].

Another consequence of process variations and the scaling down of process nodes is the increasing mismatch between devices on the same chip. For most analog designs, precise
device matching is the key to ensure high performance [27]. Hence, designers often include substantial design margins to reduce yield loss [28] and failure risks.

Variations in device parameters such as voltage threshold and mismatch will decrease production yields because block-level and system-level parameters will show a corresponding increase in variations. The Gaussian distribution with a standard deviation \( \sigma \) around the mean value \( \mu \) shown in Fig. 7 can illustrate the relationship between variation and yield. Guardbands are often used to identify uncertain chips during manufacturing testing, which require further inspection using methods such as repeating the same test or adopting other more comprehensive tests. These uncertain chips result in extra test costs [26]. Furthermore, more chips will need to be discarded or retested as the standard deviation \( \sigma \) increases. The sensor design proposed in this thesis is intended for on-chip testing of analog/RF circuit performance parameters of RF circuits, such that tuning mechanisms can be used to compensate for parameter variations. Effectively, on-chip built-in testing and tuning result in decreased specification parameters variations and increased yields.
Automatic Test Equipment (ATE) is ubiquitously used to test chips after fabrication. However, when the level of integration in systems-on-chips increases, it becomes more difficult to access test points internal to the chip with ATE for identification of particular faults in specific blocks. In addition, ATE cost also increases with design complexities and test frequencies. Given these limitations of ATE, the demand for built-in testing (BIT) and built-in self-testing (BIST) is continuing to rise. BIT provides both adequate coverage for device components and the capabilities to identify specific components at fault. Moreover, with a proper BIT solution, production costs will decline as a result of production volume and overall life cycle management savings [2], [29]. In short, the ability to quickly and easily identify a specific component at fault is the fundamental promise of an effective BIT system [30]. Besides fault detection, BIT can also be used to measure the performance of specific circuits, such as with the RF signal power detector in [2] and the PLL stability monitoring circuit in [29]. For reasons discussed above, it is
important to adopt BIT techniques to find the specific faults of chips and to measure their performances for on-chip corrections and calibrations.

2.2 RF CIRCUIT PERFORMANCE SENSORS

2.2.1 Power Detectors

Various power detectors have been designed to measure and optimize the signal power at the terminals of RF and baseband circuits, to enhance the linearity of power amplifiers, or to adjust the gain of receiver and transmitter paths. Often, the RF power is measured by extracting the peak or root-mean-square (RMS) values of the signal. Peak detection is typically used on constant-envelope or low peak-to-average ratio signals, while the RMS method is suitable for high peak-to-average ratio signals to obtain a better accuracy [31]. In many cases, the latter method uses larger output impedance to improve the conversion gain. It has been shown that a successive-approximation logarithmic amplifier can enlarge the dynamic range [32] of the measurement. Generally, RMS detection is more widely used than peak detection. A conventional RMS power detector can be modeled as a half-wave rectifier with nonlinearity [33]. In the literature, the frequency of operation, sensitivity, dynamic range, chip area, and power consumption of the detector are parameters that are commonly optimized and used for comparison. For instance, current-mode rectification [34] permits good linearity but it has limitations at high frequencies. Piecewise-linear (PWL) detectors can detect
the signal amplitude with a wide dynamic range, but are also restricted to lower frequencies.

Fig. 8 illustrates the process of RMS detection, which mainly consists of three steps. First, the input RF voltage signal is often converted to a current signal, which is then amplified. Second, the amplified AC current signal is rectified to obtain a DC current whose average value represents the RMS value of the input signal. Finally, the rectified current is converted back to a voltage signal and passed through a passive low-pass filter (LPF) to extract the DC component. Fig. 9 displays the simplified diagram of the RMS detector in [20], where transistor M₁ has high input impedance and converts the RF input voltage signal into current. Current source Iₘ₈₉₃ sets the operating point of M₁, and capacitor C₁ reduces the degeneration impedance at RF frequencies. Limited by the small size of M₁, the drain current is not adequately high for rectification. To solve this issue, this design amplifies the drain current of M₁ with a current mirror. The RF current flows through the capacitor C₂ and enters the class-AB rectifier that realizes full-wave rectification. The DC component is then extracted through the low-pass filter formed by R and C₃.
On-chip power detectors permit to characterize the gain and linearity of RF circuits. For example, when one detector is placed at the input of an LNA and a second detector at the output of an LNA, then a single-tone RF input signal can be applied with various power levels to determine the corresponding gains through which the 1-dB compression point can be identified. Fig. 10 visualizes typical measurement results of such a built-in test.
with swept amplitude level at the LNA input. When the power (in dBm) of the RF input signal changes from $X_1$ to $X_2$, the RMS power detector at the LNA input (solid line) converts the power change to a DC output voltage change. Within the dynamic range of the power detector, the output voltage of the detector at the input exhibits an almost linear relationship with the input power. In the same manner, the output voltage of the detector at the LNA output (dashed line) changes linearly with the output power within the detector’s dynamic range as also shown in Fig. 10. Notice that at each test point the difference between the detector output voltages represents the gain of the LNA. The LNA has a gain ($\text{Gain}_1$) with an input power of $X_1$, but a slightly different gain ($\text{Gain}_2$) with an input power of $X_2$. Thus, as the input power is increased, the measured voltage gain of the LNA will eventually decrease by 1dB due to the gain compression characteristics of the active devices in the amplifier, allowing to measure the large-signal behavior based on the conventional 1-dB compression point. In [20], the dynamic range of the detector is 30dB (approximately from -25dBm to +5dBm) for input frequencies in the 0.9 GHz to 2.4 GHz range. From the comparison in Table I, it can be observed that among the RMS RF power detectors, [20] provides the highest experimental sensitivity with 0.35μm CMOS technology but also consumes more than 8mW of power.
Fig. 10. Characteristic curves of RMS detectors at the input and output of an LNA. (conceptual summary of the results from [20])

Table I. RF detector performance comparison

<table>
<thead>
<tr>
<th>Reference</th>
<th>[11]</th>
<th>[19]</th>
<th>[20]</th>
<th>[22]</th>
<th>[32]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simulation Results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic Range (dBm)</td>
<td>27</td>
<td>30</td>
<td>&gt;30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Power Consumption (mW)</td>
<td>1.2</td>
<td>0.63</td>
<td>8.6</td>
<td>0.176</td>
<td>0.9</td>
</tr>
<tr>
<td>Technology (μm)</td>
<td>0.13</td>
<td>0.09</td>
<td>0.35</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Measurement Results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic Range (dBm)</td>
<td>-</td>
<td>35</td>
<td>&gt;30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sensitivity (mV/mW)</td>
<td>-</td>
<td>57.2</td>
<td>84.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Maximum Frequency (GHZ)</td>
<td>-</td>
<td>2</td>
<td>2.4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
2.2.2 Non-intrusive Sensors

Even carefully designed electrical power detectors degrade the performance of the circuit under test (CUT) to some extent, particularly as the frequency of operation is increased. To avoid this drawback, some non-intrusive sensors have been proposed in the literature [8], [35].

In general, non-intrusive process sensors can be formed with basic analog stages called “dummy circuits” that are placed in vicinity of the CUT or with process control monitors (PCMs) that are placed between dies on the wafer [35]. The fundamental idea of this strategy is that the performances of both the dummy circuit and CUT are likely to approximate one another if they are placed close to each other on the same die. The proximity of their placement allows for process variations to have a similar impacts on both. Under this premise, the sensor measurement can implicitly predict shifts in performance of the CUT [35]-[36]. In particular, these non-intrusive sensors are designed with some analog stages that have a similar structure as the CUT. Ideally, after fabrication with the same process variations, the outputs of these sensors will reflect the degradation of the actual CUT performances [36]. Thus, this approach only requires the measurement of the dummy circuit instead of the CUT itself. The expected situation is that both dummy circuits and PCMs can track the CUT performance changes in the CUT due to process variations with reasonable errors based on high device parameter correlations. The RMS prediction errors for several RF front-end performance
parameters are provided in [36] for different type of sensors. For example, the conversion gain of the RF front-end in [36] was predicted with 0.91% error using dummy circuits and PCMs, while the 1-dB compression point and third-order intermodulation intercept point (IIP3) were predicted with errors of 0.21% and 2.02% respectively.

Two examples of dummy circuits are a typical bias circuit with a current mirror and MOS gain stages with different geometries [36]. In general, dummy circuits can provide DC output voltages corresponding to different gain parameters, such as transistor transconductances and passive resistor values. In [35], dummy circuits were used to test the performance of an LNA by predicting the gain with an accuracy of 0.18dB. In this case, the dummy circuit transistors were designed to be analogous to the LNA transistors with the exception that the width/length ratios were not exactly the same.

A PCM can be a basic component such as a capacitor or transistor that is used for lower level process parameter measurements such as the reverse saturation current of a bipolar junction transistor (BJT). PCM devices are often placed in the area of the wafer kerf that will be destroyed after chip dicing [36], which requires measurement of the PCM during wafer testing. In the LNA performance test from [35], for instance, a metal-insulator-metal (MIM) capacitor used as PCM is placed close to the input capacitor ($C_{in}$) of the LNA. The function of the PCM is to predict the influence of device process variations.
(such as the capacitance) on higher level parameters such as impedance matching conditions and gain. The DC outputs of dummy circuits and the values of MIM capacitors can be mapped to LNA performance metrics such as s-parameters, noise figure, 1-dB compression point, and IIP3 [35]. The testing process requires an adequate number of samples from different lots and sites on the same wafer. For each sample, the performance of the dummy circuit sensor is measured for representative instances on each wafer that are correlated to the performance parameters of the corresponding CUT. The results for the samples in [35] show that the regression function is valid for the prediction of defects. However, dummy circuits and nearby CUTs may have parameters that fail to match due to intra-die variations. Since intra-die variations affect the dummy circuits in different ways than the CUTs, they cause inaccuracies during the performance estimation. In order to detect random defects due to process variations, the use of additional sensors such as a DC probe, an envelope detector, and a current sensor along the RF circuit signal path have been proposed [36]. The DC probe is the simplest device used for detecting degradation and defects, which can be implemented by connecting a resistor with a high value at a critical node to minimize the impact on the current flow within the CUT. Nevertheless, a DC probe is not suitable for assessing the variation effects of inductive or capacitive components on RF performance. To address this problem, an envelope detector (ED) and a current sensor can be employed. The envelope detector designed in [36] consists of a half-wave rectification circuit and a low-pass filter. Compared with the rectification circuit of [20], this half-wave rectification circuit
has lower efficiency because the negative half-cycle of the RF signal is diverted to ground. The built-in current sensor in [36] is used to extract the dynamic current signature of the power supply by monitoring the current in the line between the CUT and power supply pad that has a series resistance of several ohms. The current flowing through the resistor between the supply voltage and CUT generates a small voltage drop which is converted into a voltage signal that is applied to the input of envelope detector. Thereafter, the RMS amplitude of the RF current is obtained at the output of the envelope detector. Nevertheless, the placement of the built-in current sensor in the supply line can impact the performance of the RF CUT.

While non-intrusive sensors avoid loading effects, they can only mimic the CUT to a limited extent and do not measure the CUT performance directly. Instead, these sensors measure certain parameters as indicators of process variations in the vicinity of the RF CUT in order to estimate RF performance changes. However, it is important to keep in mind that process parameters measured with DC tests cannot account for RF performance impacts due to high-frequency losses, coupling between circuits/paths, and high-frequency linearity degradations. Many non-intrusive methods involve replica circuits that are characterized in lieu of the CUT to avoid loading effects, where parameter correlations between the CUT and the replica are assumed based on the proximity of devices on the chip or based on layout matching techniques. Hence, the effectiveness of many non-intrusive methods relies on the actual device-level parameter
correlations. In the next section, a non-invasive built-in test method is discussed that directly measures RF performance through the measurement of temperature changes that result from RF signal power changes.

2.3 NON-INVASIVE TESTING BASED ON ELECTRO-THERMAL COUPLING

The high-frequency behavior of RF circuits can also be measured by monitoring the DC and low-frequency temperature changes close to the CUT. This approach avoids performance degradation because it does not require an electrical connection to the CUT. A brief overview of electro-thermal coupling will first be giving in the remainder of this chapter to explain how RF signal and DC power dissipation can be measured with temperature sensors to extract the gain and 1-dB compression point of a CUT as with conventional power detectors. The next chapter will focus on the proposed amplifier architecture to improve the sensitivity and dynamic range of this on-chip built-in testing approach.

2.3.1 Thermal Coupling

On-chip thermal coupling causes temperature changes in the silicon area close to any device that dissipates power, and this phenomenon can be exploited for built-in testing of analog/RF circuits [37]. A basic thermal-electrical coupling model for on-chip interactions between circuits is discussed in [6], which illustrates the relationship between temperature changes on a silicon die due to power dissipation in a device and
the resulting impact on nearby devices. This model is elaborated further in this section and the following sections to explain the relevance to built-in testing based on the basic thermal model of the die shown in Fig. 3 of the previous chapter, the local interaction between thermal and electrical parameters at the CUT (in Fig. 4), and the sensing-device’s parameters. The settling time of the temperature changes, the frequency characteristics of thermal coupling, and the impact of the distance between the CUT and the sensing-device will also be considered. The discussion here will be based on the discretized heat equation derived from the partial differential equations for heat diffusion. Fig. 11 displays the simplified one-dimensional model for heat diffusion from [6]. For simplicity, the discussion below focuses on one dimension (i.e., the x-axis direction of the 3-D model in Fig. 3) since the corresponding equations can be obtained for the other directions. The thermal equation for the x-direction can be written as follows:

\[
\rho c \frac{\partial u(x)}{\partial t} = \kappa \frac{u(x - \Delta x) - u(x)}{\Delta x} + \frac{u(x) - u(x + \Delta x)}{\Delta x} + D(x),
\]

where \(\rho\) is the silicon mass density, which equals \(2.3 \times 10^6 \text{gm}^{-3}\), and \(c\) is the specific heat capacity that equals \(0.7 \text{Jg}^{-1}\text{K}^{-1}\), \(\kappa\) is the thermal conductivity that is equal to \(120 \text{Wm}^{-1}\text{K}^{-1}\) at 75 °C, and \(D(x)\) represents the dissipated power density in the x-direction [38]. \(u(x)\) is the local temperature on the x-axis, and \(\Delta x\) is the finite increment.
for the partial derivatives along the $x$-axis. If both sides are multiplied by the unit volume $\Delta cube = \Delta x\Delta y\Delta z$, then the resulting equation becomes:

$$
\Delta cube \times \rho c \frac{\partial u(x)}{\partial t} = \kappa \frac{\Delta y\Delta z}{\Delta x} [u(x - \Delta x) + u(x + \Delta x) - 2u(x)] + D(x)\Delta cube,
$$

where $D(x)\Delta cube$ is the power dissipated in the unit volume of the silicon substrate. Corresponding to Fig. 11 we can then define the equivalent electrical parameter as:

$$
C_E = \Delta cube \times \rho \times c = \Delta x\Delta y\Delta z \times \rho \times c,
$$

$$
R_x = \frac{\Delta x}{\kappa\Delta y\Delta z},
$$

$$
i_x = D(x)\Delta cube,
$$

$$
V_x = u(x).
$$

Equations (3) to (6) illustrate the premise of the thermal-electrical equivalent model and also provide the foundation for temperature sensor design. It is noteworthy to point out that conventional electrical circuit simulators can be used to simulate electro-thermal coupling based on the above relationships [5]-[6]. In such simulation, the instantaneous power at a device of interest can be obtained by multiplying the current through a device with its terminal voltage drop and injecting the result as thermal current into the resistor-capacitor grid using controlled sources with the appropriate scaling factors that depend on physical constants, layout dimensions, as well as the chosen $\Delta x$, $\Delta y$, and $\Delta z$ dimensions.
2.3.2 Bipolar Junction Transistor Model

As introduced in [37] and [39], bipolar junction transistors can be employed as thermal transducers for built-in testing applications. The temperature sensor circuit design in this thesis uses vertical PNP transistors as in [5], which are available in standard CMOS fabrication processes since they are commonly used in bandgap reference circuits. In general, bipolar transistors in BiCMOS processes can provide higher forward current gain ($\beta$) than lateral PNP devices because of their higher carrier mobility. However, the relatively low current gains of vertical PNP devices are sufficient for the design of temperature sensors with high precision and repeatability [5]. Fig. 12 displays a simplified schematic of a vertical PNP transistor for electro-thermal coupling analysis, which includes the collector (C) current component arising from temperature changes. In Fig. 12, $r_e$ is the input resistor from the base (B) to the emitter (E), $g_m$ is the transconductance parameter, $v_{be}$ is the base-emitter voltage, $\Delta T$ is temperature change.
difference between the PNP transistor and a reference location (i.e., the ambient chip temperature), and $S_T$ is the collector current’s temperature sensitivity defined as:

$$S_T = \frac{\partial I_c}{\partial T} \approx k \times g_{mq}.$$  \hspace{1cm} (7)

The above equation is a simplification of the one in Appendix I in [39]. In this equation, $k$ is the temperature sensitivity (negative temperature coefficient) of the base-emitter voltage, which is in the range of -1.5mV/K to -2.0mV/K; and $g_{mq}$ is the transconductance of the temperature-sensing PNP transistor.

![Fig. 12. Simplified PNP hybrid-π small-signal model for electro-thermal coupling.](image)

Based on the model in Fig. 12, when a temperature difference $\Delta T$ is generated due to power dissipation in a nearby device, an additional temperature-controlled current flows at the collector terminal of the temperature-sensing PNP transistor. As described in Section 1.2.1, an alternative model (Fig. 4) for the sensing process is to modulate the base-emitter voltage according to the PNP’s temperature sensitivity $k$. 
2.3.3 RF Power Detection with On-Chip Temperature Sensors

As outlined by the strategies in [5], [39] and [40], BJT devices can be used for power measurements based on DC and low-frequency thermal gradients on the chip. In [5] and [39], a differential measurement approach with two BJT devices is used: one as the monitor and the other as the reference device. The monitoring BJT is placed close to the CUT, typically at a distance of 25µm or less. The reference BJT is placed far away (>150µm) from the monitoring BJT to ensure that its temperature change to power dissipation in the CUT is negligible. Meanwhile, if no nearby heat sources are placed close to monitoring device or the effect from nearby constant power sources is removed with a calibration step, then the resulting temperature difference (ΔT) is proportional to the signal power when a test input signal is applied to the CUT. Thereafter, the differential output current of the two sensing devices is amplified, and a DC output voltage is generated that indicates the temperature difference. When an RF signal is applied to the input of a CUT (e. g. a power amplifier), only the temperature changes from the DC component of power dissipation create significant sensor output responses because the cutoff frequency of the thermal coupling in the silicon die is much lower than the RF frequency [39].

Both [5] and [39] provide analytical models for RF power dissipation estimation through electro-thermal sensing. In [5], an LNA was characterized with a built-in temperature sensor as an example, while the approach was validated with measurements of a basic
class A power amplifier (displayed in Fig. 13) in [37] and [39]. Since the DC component of the RF signal power causes DC temperature changes, the corresponding analytical expressions will be discussed next.

In Fig. 13, the input voltage includes both a DC component \( V_{\text{DC}} \) and an RF signal component that is assumed to be sinusoidal in this analysis: \( v_{\text{in}}(t) = A \cos(2\pi f t) \) with \( A \) being the signal amplitude. Inductor \( L \) allows DC current flow from \( V_{\text{DD}} \), but presents a high impedance at the RF operating frequency. Capacitor \( C \) blocks DC current flow to the load resistor \( R_L \). The current through the inductor \( (I_{\text{DC}}) \) is assumed to be a DC current, while the drain-to-source current of the transistor includes both a DC component \( (I_{\text{DC}}) \) and an RF component \( (i_{ds}) \). At the drain of the transistor, the voltage is \( V_{\text{DD}} + v_o \), where \( v_o \) is the amplified sinusoidal signal. The RF current can be expressed as:

\[
i_{ds} = g_m \cdot v_{\text{in}} = g_m \cdot A \cdot \cos(2\pi f t) = g_m \cdot A \cdot \cos(\omega t),
\]

where \( g_m \) is the transconductance of the transistor in Fig. 13. The power consumption of the load resistor \( R_L \) is:

\[
P_L = (g_m \cdot A)^2 \cdot \cos^2(\omega t) \cdot R_L = (g_m \cdot A)^2 \cdot R_L \cdot \frac{1 + \cos(2\omega t)}{2}
\]
Since the thermal coupling has a cutoff frequency that is much lower than the frequencies of the RF signal components ($f$ and $2f$ terms), only the DC components of the power dissipated by the transistors excites a measurable temperature change [37]. After dropping the high-frequency component of $P_L$ from equation (9), the DC term of the transistor’s power can be obtained as:

$$P_{dc} = V_{DD} \cdot I_{DC} - \frac{1}{2} g_m^2 \cdot A^2 \cdot R_L.$$  \hspace{1cm} (10)

The above equations indicate that the total power supplied by the source ($V_{DD}$) is $V_{DD} \cdot I_{DC}$, and that the power dissipation at the load resistor ($\frac{1}{2} g_m^2 \cdot A^2 \cdot R_L$) is independent of the RF signal frequency. This DC power term depends on the amplifier gain and input signal amplitude, and it can be measured with a nearby temperature-sensing device because the local silicon temperature change is proportional to the change...
in the power dissipation. Following the same inferential process, the DC power components of the previously discussed LNA in Fig. 4 can be obtained. The equations for the average power dissipation of the devices in this circuit can be found in [5].

Several on-chip temperature sensors for built-in testing applications have been demonstrated with experimental measurements that are published in the literature, but researchers today continue the pursuit of extending the dynamic range and increasing the sensitivity of these sensors to broaden their application range. The sensor design proposed in the next chapter is based on the architecture from [5], but includes a new amplifier with a class AB output stage to improve the dynamic range and sensitivity. The design also allows significant reduction of the sensor’s power dissipation, which will also be elaborated in the next chapter.

3. DIFFERENTIAL TEMPERATURE SENSOR DESIGN

3.1 INTRODUCTION

3.1.1 Existing Temperature Sensors

Based on the measurement process described in Section 2.3, the thermal signal generated by the DC component of the RF power needs to be converted into an electrical signal that can be measured. As Fig. 4 and Fig. 12 illustrate, a PNP device can sense the temperature change and convert it into a current difference. This current difference has
to be amplified to ease on-chip testing. A general expression for the output signal of built-in differential temperature sensors is provided in [39]:

$$ Signal_{out} = S_d (T_2 - T_1), $$

(11)

where $T_2$ and $T_1$ are the temperatures at two points on the silicon surface such as the input and output devices of an LNA. $S_d$ is the sensitivity of the sensor. The simplified circuit of a differential temperature sensor from [39] is displayed in Fig. 14, where $R_1 = R_2 = R$, BJTs $Q_1$ and $Q_2$ have surrounding silicon temperatures of $T_1 = T + \Delta T_1$ and $T_2 = T + \Delta T_2$ respectively when heat sources are activated, and transistor $Q_b$ is at the temperature $T$. In this analysis, the base current of $Q_b$ is approximated as zero, and it is assumed that $Q_1$ and $Q_2$ have the same operating point. Fig. 12 and equation (7) provide a simplified small-signal model of $Q_1$ and $Q_2$ to calculate the sensitivity $S_T$.

Fig. 14. Simplified schematic of a differential temperature sensor.
To describe the electrothermal behavior with more analytical detail, equation (11) can be rewritten as:

\[ \Delta V_{\text{out}} = S_{dT} (\Delta T_1 - \Delta T_2) + S_{cT} \left( T + \frac{\Delta T_1 + \Delta T_2}{2} \right), \]  

(12)

where \( S_{dT} \) is the sensitivity to differential temperature and \( S_{cT} \) is the sensitivity to common-mode temperature. Ideally, \( S_{dT} \) is very large while \( S_{cT} \) should be negligible. If \( Q_b \) remains at the temperature \( T \) and the output resistance of the current source \( I_e \) is assumed to be infinite, then equation (12) can be rearranged to:

\[ \Delta V_{\text{out}} = \frac{S_T \cdot r_o \cdot R}{2(r_o + R)} (\Delta T_1 - \Delta T_2), \]  

(13)

where \( r_o \) is the output resistance of \( Q_2 \). When \( R_1 \) and \( R_2 \) in Fig. 14 are replaced by a CMOS current mirror with high impedance, the differential output voltage of the resulting temperature sensor can be increased further, resulting in higher sensitivity.

In practice, the base current of \( Q_b \) in Fig. 14 cannot be zero because \( Q_b \) would not have any current flow at the emitter terminals to bias \( Q_1 \) and \( Q_2 \). For this reason, \( Q_1 \) and \( Q_2 \) must be designed asymmetrically to obtain roughly the same operating point. Furthermore, additional circuits are normally designed to compensate for offsets in the differential branches.
Another use of the negative temperature coefficient of the base-emitter voltage $V_{EB}$ of PNP devices was introduced in [41] using the temperature sensor design strategy displayed in Fig. 15. This circuit consists of a bandgap reference, a start-up circuit, and a sensor comprised of two bipolar PNP transistors ($Q_1$ and $Q_2$) connected in a Darlington configuration. The bandgap reference provides the bias voltage for the start-up circuit and sensor devices. Under a fixed collector bias current in a PNP device, the emitter-base voltage change $\Delta V_{EB}$ and the environmental temperature are related according to the following equation:

$$\frac{\partial V_{EB}}{\partial T} = \frac{V_{EB} - (4 + m)V_T - E_g / q}{T},$$  \hspace{1cm} (14)

where $T$ is the absolute temperature, $E_g$ is the electron energy gap of the semiconductor, and $m$ is a temperature parameter of $E_g$ that is often estimated as $-3/2$. $V_T$ is the thermal
voltage, which is equal to $kT/q$, where $k$ and $q$ are Boltzmann constant and electron charge quantity, respectively [41]. The current sources with bandgap reference are designed with a very small temperature coefficient to provide stable currents at the drains of $M_{p1}$ and $M_{p2}$. These bias currents are much larger than the collector current change $\Delta I_C$ of $Q_1$ and $Q_2$ in response to temperature variations in order to avoid that the current gain $\beta$ of the PNP transistors will be sensitive to temperature variations.

The output voltages $V_{o1}$ and $V_{o2}$ in Fig. 15 are equal to $V_{EB1} + V_{EB2}$ and $V_{EB2}$, respectively. Both of them have negative temperature coefficients. With the same temperature change ($\Delta T$) at both PNP transistors, the output $V_{o1}$ is larger than $V_{o2}$ because of the Darlington connection. Hence, the Darlington connection effectively increases the temperature sensitivity. The simulation and measurement results in [41] show that the bandgap reference current source is thermally stable and the temperature sensitivity is linear over a large temperature range of from 15°C to 110°C. The sensitivity is approximately 3.9mV/°C in this temperature range. The discussed design approach in [41] provides a method to extend the sensor’s dynamic range and to increase the sensitivity when the voltage headroom is available to use a Darlington connection. In the next subsection, a differential sensor architecture will be described that achieves higher sensitivity for built-in testing applications.
3.1.2 Reference Sensor Design

Fig. 16 displays the sensor architecture adapted from [5], which was used in this work with an improved amplifier for enhanced performance. In [5], amplifier $A_1$ was on the chip while $A_2$ was an off-chip component. Both amplification stages are designed to be on a single chip in the version presented at the IEEE North Atlantic Test Workshop [42] that is discussed in this thesis. This design is also based on the electro-thermal coupling effects at the silicon surface described in Section 2.3.

Since the proposed design was derived from the reference sensor introduced in [5], some key characteristics of the previous design are discussed first. As modeled in Fig. 16, the
dissipation of power at the device under test (DUT) changes the temperature of the nearby transistor Q₁, and thus changes the emitter current of Q₁. Transistors Q₁ and Q_ref were placed 420 μm apart in [5]. Both of them are biased with the same operation point. Q_ref in Fig. 16 is considered to be the reference device that is labeled as PNP_ref in Fig. 2. The emitter-base voltages (V_{EB}) of Q_ref and Q₁ are equal because of the virtual ground at the input of the first amplification stage. Hence, the collector current difference depends only on the temperature difference between the respective locations of Q₁ and Q_ref. In [5], the on-chip sensor components included the first amplifier, test and reference PNP transistors, resistors (R₁, R_L), the current mirrors of the calibration currents (I_cal₁ and I_cal₂), and base bias current I_core in Fig. 16. An initial output offset voltage typically exists even when the DUT is deactivated, which is caused by device mismatches from process variations in the sensor circuitry and thermal gradients from other power-dissipating devices on the chip. The calibration current sources I_cal₁ and I_cal₂ allow compensation for mismatches as well as for thermal gradients as elaborated in [5].

As modeled in Fig. 12, the temperature-dependent differential current $\Delta I_T = S_T \times \Delta T$ is the input of the first stage of amplifier A₁ in Fig. 16. To amplify this current, A₁ and its feedback network form a current amplifier in combination with A₂ configured as a transimpedance amplifier (TIA) to convert the current to a differential voltage at the output ports. The resistive load $R_L = \frac{1}{n} R_1$ leads to a current flow of $n \cdot \Delta I_T$ through $R_L$. 
In [5], the parameter $n$ was chosen to be 8. The resistor values $R_1 = R_2 = 8\, \Omega$ and $R_L = 1\, \Omega$ were also adopted for the design discussed in this thesis. The increase or decrease of $I_{\text{core}}$ leads to a corresponding change of the emitter currents of both PNP transistors. An increase of $I_{\text{core}}$ results in higher current-temperature sensitivity $S_T$ because the transconductance ($g_m Q$) of $Q_1$ and $Q_{\text{ref}}$ increases with higher base bias current. Since the sensitivity of the temperature sensor can be adjusted by changing the bias current source $I_{\text{core}}$, the same sensor design can be widely used to measure DUTs with different power dissipation levels in mixed-signal systems-on-a-chip. In the future, the adjustment of $I_{\text{core}}$ and the tuning of the calibration current sources could be automatically performed by a digital control loop as visualized in Fig. 2.

![Fig. 17. Schematic of amplifier A1 in [5].](image-url)
Fig. 17 displays the circuit schematic of the on-chip amplifier $A_1$ in Fig. 16. Inputs $IN_1$ and $IN_2$ of the amplifier are connected to the emitters of $Q_1$ and $Q_{ref}$, and $I_{bias}$ provides the bias current for $A_1$ through a current mirror comprised of two identical NMOS transistors ($M_1$). This amplifier structure consists of differential input pair $M_2$ loaded by PMOS transistors ($M_3$) and a source follower ($M_4$ and $M_5$) at the output. Large capacitors $C_1$ are used to obtain a frequency response that approximates a single-pole transfer function without stability concerns. The steadiness of the common-mode voltage at nodes $N_{X1}$ and $N_{X2}$ is important, which is why a common-mode feedback (CMFB) circuit is needed to regulate the DC level at $N_{X1}$ and $N_{X2}$ that is transferred to the output with a shift in the source follower. Fig. 18 shows the schematic of the CMFB circuit employed in [5].

![CMFB circuit](image)
It can be observed that the linear output voltage range of the amplifier in Fig. 17 has a significant limitation: The output DC level (1.55V in the reference design) at nodes $V_{\text{OUT1}}$ and $V_{\text{OUT2}}$ is determined by the voltages at nodes $N_{X1}$ and $N_{X2}$, and the gate-source voltage drop of $M_4$. Moreover, the drain-source voltage ($V_{\text{DS}}$) of $M_5$ must be maintained higher than its saturation voltage ($V_{\text{DSsat}}$) to ensure proper operation. This voltage headroom limitation (particularly if the second stage is to be implemented on the chip) is the primary constraint for the dynamic range of the sensor circuit.

For a comparison of the sensitivity and dynamic range with the proposed design discussed in the next section, the sensor from [5] was redesigned. In this sensor design, both amplification stages ($A_1$, $A_2$) were implemented with the on-chip amplifier in Fig. 17 to avoid that the second stage is external to the chip as in the previous version. The fact that this integration has an impact on the dynamic range and also requires amplifier design optimizations was the reason why the reference amplifier was redesigned in 0.18μm CMOS technology with a 1.8V supply for a proper comparison to the proposed amplifier. Since the variable of interest in this thermal monitoring approach is the sensor’s differential DC output voltage, the amplifier’s 3dB corner frequency is not critical and was designed to be 1.74MHz. Simulation results in comparison to the proposed design are provided in Chapter 4.
3.2 PROPOSED DESIGN

3.2.1 Class AB Output Stage

As explained in the previous section, the output stage of the reference sensor design has a voltage swing limitation that constraints the dynamic range. To alleviate this constraint, a class AB output stage was constructed in this thesis work to enhance the sensitivity and dynamic range of the amplifier.

One option to realize large output swing capability would be the use of a class B push-pull output stage as shown in generalized schematic in Fig. 19(a). However, main disadvantage of a class B push-pull output stage with input signals between $V_{DD}$ and ground is the limitation of the output swing from $V_{SGp}$ to $V_{DD} - V_{GSn}$, where $V_{SGp}$ and $V_{GSn}$ are the source-gate overdrive voltage of the PMOS transistor and the gate-source overdrive voltage of the NMOS transistor, respectively. Crossover distortion cannot be neglected when adopting a class B push-pull output stage. Since the proposed temperature sensor requires high resolution around zero temperature difference ($\Delta T = 0^\circ C$), the deficiency due to crossover distortion must be averted in the proposed design.
To avoid the issues mentioned above, one can consider the general class AB output stage depicted in Fig. 19(b). This structure enables almost rail-to-rail output voltage swing, and is especially useful with capacitive loads [43]. Referring back to the amplifier in Fig. 17, nodes $N_{X1}$ and $N_{X2}$ become the inputs to the differential class AB output stage. For the proposed design, the class AB output stage in Fig. 20 was formed based on similar class AB stages [44]-[47]. $I_s$ provides bias current, and $R_1$ is used to generate a different gate bias voltage for $M_P$ than the gate voltage of $M_N$. Resistor $R_2$ and capacitor $C$ are used to maintain the stability of this amplifier with the Miller effect. In addition to larger output voltage swing, this class AB design strategy also provides higher voltage gain compared to the reference amplifier. Fig. 21 shows the simplified small-signal model of the output stage in Fig. 20, from which equation (15) was derived to assess the voltage gain of the class AB structure:
\[ A_r \approx -g_{mX} \cdot r_{oX} \cdot (g_{mN} + g_{mP}) \cdot (r_{oN} || r_{oP}), \]  

(15)

where \( r_{oX} \) represents the output resistance of transistor \( M_X \), \( r_{oN} \) and \( r_{oP} \) represent the output resistance of transistors \( M_N \) and \( M_P \), respectively. Compared to the voltage gain of the source-follower output stage in Fig. 17, the improvement is obvious that the gain of the class AB output stage is much higher than the former, which only provides a gain of:

\[ A_v = \frac{g_{m4}(r_{o4} || r_{o5})}{1 + g_{m4}(r_{o4} || r_{o5})} < 1. \]  

(16)

Fig. 20. Class AB output stage used in the proposed design.
Fig. 21. Small-signal model of the class AB output stage.

Fig. 22. Proposed amplifier for each stage (TSMC 0.18μm technology).

Fig. 22 displays the proposed amplifier design that was used in both stages with different device dimensions to optimize performance. The output nodes are located between the
drains of transistors $M_5$ and $M_6$, which have approximately the same transconductance ($g_{m5} \approx g_{m6}$). Resistor $R_2$ and capacitor $C_2$ in Fig. 22 serve to stabilize the amplifier, which is explained in 3.2.3. A common-mode feedback circuit (CMFB) was employed to regulate nodes $N_{X1}$ and $N_{X2}$ as discussed in the next subsection.

3.2.2 Common-mode Feedback Circuit Design

The output common-mode level is quite sensitive to device properties and mismatches in high gain amplifiers [48]. The proposed amplifier is a differential high-gain amplifier, which requires a common-mode feedback (CMFB) loop to maintain the desired common-mode voltage level even if different currents flow into the high-impedance nodes of each branch due to device mismatches. Fig. 23 illustrates the schematic of a differential high-gain amplifier and a simplified model of one branch, where $R_N$ and $R_P$ represent the high drain-source resistances of the NMOS and PMOS transistors respectively. In the presence of process variations during fabrication, it is impossible to ensure that the currents of the NMOS and PMOS devices are perfectly matched. When $I_P$ and $I_N$ are unequal due to mismatches, the current difference will cause a voltage change of $(I_P - I_N) \cdot (R_P \parallel R_N)$ at the output [48]. Therefore, large output voltage offsets can be generated in amplifiers with high-impedance nodes for high gain.
To alleviate problems from mismatches in differential branches, a CMFB circuit can be added to sense the common-mode voltage level of the two outputs and to adjust the bias current through a feedback loop that regulates the common-mode voltage. Usually, the CMFB circuit can be divided into three sections: the sensing part, the comparison part, and the feedback components. The sensing part is used to extract the common-mode voltage level from the differential voltages in the branches. The output voltages at $N_{X1/X2}$ of first differential amplifier stage in Fig. 17 are connected to the input of CMFB in Fig. 18. Current source $I_{CMFB}$ in Fig. 18 provides the bias current for the CMFB amplifier through a current mirror. The double differential pair ($M_N$) compares the average of the voltages sensed at $N_{X1}$ and $N_{X2}$ to the reference voltage $V_{ref}$, which in this circuit should be equal to the desired common-mode level at $N_{X1/X2}$. The width/length ratio $W/L$ of
the diode-connected PMOS transistor \( M_O \) in Fig. 18 has twice the \( W/L \) ratio of \( M_P \) to balance the circuit by ensuring that the voltage drops across both transistors are approximately equal when the sensed common-mode voltage is equal to \( V_{ref} \) (when the current through \( M_O \) is twice as high as the current through \( M_P \)). Finally, the feedback voltage \( V_{FB} \) from the drain of \( M_O \) is sent back to the gate of the PMOS transistors \( M_3 \) and \( M_5 \) in Fig. 17 to complete the negative feedback loop that regulates the common-mode voltage of \( N_{X1} \) and \( N_{X2} \).

To further elaborate the CMFB operation, let us assume that a mismatch between the PMOS transistors \( M_3 \) in Fig. 17 exists due to fabrication process variations. The resulting quantities of interest are annotated in Fig. 24, where the drain currents (\( I_A \) and \( I_B \)) of PMOS transistors \( M_{3A} \) and \( M_{3B} \) are different; i.e., \( I_A \neq I_B \). Consider \( I_A < I_B \) for example, let \( I_B - I_A = \Delta I \), and assume \( I_B \) is equal to the expected value in the branch but \( I_A \) is lower than the expected value. Thus, the voltages \( N_{X1} \) and \( N_{X2} \) in the first amplifier stage is related as follows: \( V_{X1} < V_{X2} \). As a consequence, the sensed common-mode voltage \( V_{CM} = (V_{X1} + V_{X2})/2 \) is lower than the desired common-mode level (\( V_{ref} \)). As a result, the drain current of transistor \( M_{NA} \) in Fig. 25 decreases due to the reduction of \( V_{xI} \) relative to \( V_{ref} \). With the constant shared tail current \( I_{CMbias} \), the drain current of transistor \( M_{NB1} \) must increase, which also increases the drain current of transistor \( M_O \). Since the PMOS transistor \( M_O \) is diode-connected, it always remains in the saturation
region and the source-drain voltage drop increases from the higher current flow through it, leading to a lower feedback voltage ($V_{FB}$) value. Since the feedback voltage $V_{FB}$ is applied to the gate of transistors $M_{3A}$ and $M_{3B}$ in Fig. 24, the lower gate voltage (higher source-gate overdrive voltage $V_{SG}$) of $M_{3A}$ and $M_{3B}$ leads to higher currents ($I_A$ and $I_B$) and thereby increases the common-mode voltage at nodes $N_{X1}$ and $N_{X2}$. Thus, this negative feedback loop realizes the aim of driving the common-mode voltage closer to the desired value.

Fig. 24. Example: differential amplifier stage with different currents ($I_A$, $I_B$) in each branch due to device mismatches.
3.2.3 Stability

Stability is an important design consideration for the proposed multi-stage amplifier with class AB output stage and high voltage gain. For instance, Fig. 26 displays the open-loop frequency response of the proposed amplifier before frequency compensation. It is obvious that the class AB output stage enables a low-frequency gain increase from approximately 30dB in the reference design [5] to 94.14dB. However, the phase shift is $-223^\circ$ at the unity-gain (0dB) frequency, failing to meet the requirements for unconditional stability.
Fig. 26. Simulated open-loop frequency response of the uncompensated amplifier.

Fig. 27 displays a single branch of the symmetric amplifier, in which the most significant pole locations (A - D) are annotated. The value of resistor $R_1$ affects the frequency separation between poles B and C. It should also be noted that transistors $M_4$ and $M_7$ have the same current flow, but $M_7$ is sized with larger width to maintain approximately the same transconductance despite of the lower carrier mobility of the PMOS device. Thus, the parasitic capacitance associated with pole C is higher due to the larger PMOS device area. However, as evident in the following intuitive analysis, the estimation of the pole locations must be done in consideration of the equivalent resistances at each node and the impact of Miller effects.
Fig. 27. Locations (A, B, C, D) of significant poles in the signal path of the proposed amplifier.

According to Fig. 27, the equivalent capacitance and resistance at each node can be estimated to approximate the corresponding pole frequencies: $\omega_{\text{pole}} \approx 1/(C_{\text{eq}} R_{\text{eq}})$. At pole location A, the equivalent resistance $R_{\text{eq}}$ is approximately equal to $r_{\text{ON2}} || r_{\text{OP3}}$. Capacitors $C_{DB2}$, $C_{GD2} \left[1 + \frac{1}{g_{mN2}(r_{\text{ON2}} || r_{\text{OP3}})}\right]$, $C_{DB3}$, $C_{GD3}$, $C_{GB4}$, $C_{GS4}$, $C_{GD4} \{1 + g_{mN4}[r_{\text{ON4}} || (R_1 + r_{\text{OP7}})]\}$ are connected to ground, which can be summed up to approximate $C_{eq}$ at node A. $C_{DB2}$ is the drain-bulk capacitor of transistor M2 in Fig. 27, $C_{GD2} \left[1 + \frac{1}{g_{mN2}(r_{\text{ON2}} || r_{\text{OP3}})}\right]$ and $C_{GD4} \{1 + g_{mN4}[r_{\text{ON4}} || (R_1 + r_{\text{OP7}})]\}$ are effective Miller capacitances, where $C_{GD2}$ and $C_{GD4}$ are the gate-drain capacitor of transistor M2 and M4,
respectively. $C_{GB4}$ is the gate-bulk capacitance of transistor $M_4$. The term $g_{mN2}(r_{ON2} \parallel r_{OP3})$ is the magnitude of the low-frequency voltage gain of the gain from the gate to the drain of $M_2$, while $g_{mN4}[r_{ON4} \parallel (R_1 + r_{OP7})]$ is the magnitude of the low-frequency voltage gain from the gate of transistor $M_4$ to its drain. $g_{mN2}$ and $g_{mN4}$ represent the transconductances of transistor $M_2$ and $M_4$; and $r_{ON2}$, $r_{ON4}$, $r_{OP3}$, and $r_{OP7}$ are the drain-source resistances of transistors $M_2$, $M_4$, $M_3$, and $M_7$ respectively. $C_{DB3}$ and $C_{GD3}$ are drain-bulk capacitance and gate-drain capacitance of transistor $M_3$. Since the drain-source resistances and parasitic capacitances have strong dependencies on the foundry-supplied device models for a given process as well as on the operating points of the transistors, circuit simulations were mainly used in this work to obtain the above operating point parameters and to estimate the pole locations.

Similarly, at pole location B, capacitors $C_{DB4}$, $C_{GD4}\left[1 + \frac{1}{\frac{1}{1 + g_{mN4}[r_{ON4} \parallel (R_1 + r_{OP7})]}}, \right.$

$C_{GS6}$, $C_{GD6}[1 + g_{mN6} \cdot (r_{OP5} \parallel r_{ON6})]$ are tied to the ground, forming a basis to approximate $C_{eq}$. Furthermore, $R_{eq}$ can be estimated as $r_{ON4} \parallel (R_1 + r_{OP7})$ to obtain $\omega_{pole} \approx 1/(C_{eq} \cdot R_{eq})$ based on simplified intuitive analysis.
At pole location C, capacitors $C_{GD7}$, $C_{DB7}$, $C_{GS5}$, $C_{GB5}$, and $C_{GD5} \left[ 1 + g_{mP5} \cdot \left( r_{OP5} \parallel r_{ON6} \right) \right]$ are connect in parallel to ground, yielding $C_{eq}$; while $R_{eq}$ is comprised of $r_{OP7} \parallel (R_1 + r_{ON4})$.

At pole location D, capacitors $C_{GD5} \left[ 1 + \frac{1}{g_{mP5} \cdot \left( r_{OP5} \parallel r_{ON6} \right)} \right]$, $C_{DB5}$, $C_{DB6}$, and $C_{GD6} \left[ 1 + \frac{1}{g_{mN6} \cdot \left( r_{OP5} \parallel r_{ON6} \right)} \right]$ form $C_{eq}$; and $R_{eq}$ is approximately equal to $r_{OP5} \parallel r_{ON6}$.

For simplicity, the following capacitances can be defined:

\[
C_A = C_{DB2} + C_{DB3} + C_{GD3} + C_{GD2} \left[ 1 + \frac{1}{g_{mN2} \cdot \left( r_{ON2} \parallel r_{OP3} \right)} \right] + C_{GB4} + C_{GS4},
\]
\[
+ C_{GD4} \left[ 1 + g_{mN4} \cdot \left( r_{ON4} \parallel (R_1 + r_{OP7}) \right) \right],
\]

(17)

\[
C_B = C_{DB4} + C_{GS6} + C_{GD4} \left[ 1 + \frac{1}{g_{mN4} \cdot \left( r_{ON4} \parallel (R_1 + r_{OP7}) \right)} \right],
\]
\[
+ C_{GD6} \left[ 1 + g_{mN6} \cdot \left( r_{OP5} \parallel r_{ON6} \right) \right] + C_{GB6},
\]

(18)

\[
C_C = C_{GD7} + C_{DB7} + C_{GB5} + C_{GS5} + C_{GD5} \left[ 1 + g_{mP5} \cdot \left( r_{OP5} \parallel r_{ON6} \right) \right],
\]

(19)
Based on the simplified single-ended circuit for stability analysis in Fig. 27, the following transfer function can be derived from the equivalent circuit shown in Fig. 28:

\[
\frac{v_{out}}{v_{in}}(s) = \left( \frac{g_{mn6}r_{OP7} + g_{mn6}R_1 + g_{mn6}sc_{C}r_{OP7}R_1 + g_{mp5}r_{OP7}}{R_1 + sc_{C}r_{OP7}R_1 + r_{OP7}} \right) \times \left( \frac{r_{ON5}r_{OP5}}{r_{OP5} + r_{ON5} + sc_{D}r_{ON5}r_{OP5}} \right) \times \left( \frac{g_{mn4}(R_1r_{ON4} + sc_{C}R_1r_{ON4}r_{OP7} + r_{ON4}r_{OP7})}{r_{ON4} + r_{OP7} + sr_{ON4}r_{OP7}(c_{B} + c_{C}) + R_1(1 + sc_{C}r_{OP7} + sc_{B}r_{ON4} - c_{B}c_{C}r_{ON4}r_{OP7})} \right) \times \left( \frac{g_{mn2}r_{ON2}r_{OP3}}{sc_{A}r_{ON2}r_{OP3} + r_{ON2} + r_{OP3}} \right).
\]
From equation (21), it is evident that there are at least 4 poles and 2 zeros in the frequency response of the amplifier. Further poles at high frequencies are introduced due to the intricate device models with interconnected internal resistances and capacitances. Thus, simulations provide more accurate estimations of pole and zero locations than hand calculations. However, equations (17) to (21) permit rough calculations to gain insights into the design tradeoffs as well as to estimate the relative positions on the frequency axis. The uncompensated amplifier with the output stage in Fig. 27 was simulated with the pole-zero analysis feature of Cadence Spectre to obtain the pole and zero locations. Fig. 29 and Fig. 30 summarize the results.

Fig. 29. Pole and zero locations of the uncompensated amplifier.
Fig. 30. Pole and zero summary of the uncompensated amplifier.

A method to stabilize an amplifier with the aforementioned problem is to move the dominant pole and the first nondominant pole further away from each other. This method is called “pole splitting” [48]. To realize this compensation, a Miller capacitor is employed. The placement considerations for a Miller compensation capacitor will be discussed next. One might assume that the dominant pole is at node A. To test this hypothesis, a Miller capacitor was added between node A and node C in Fig. 27, resulting in the simulated frequency responses in Fig. 31 for Miller capacitor (MC) values from 0.5pF to 1.5pF. The results are in disagreement with the assumption because the phase margin degrades after the insertion of a Miller capacitor $C_{A-C}$ between nodes A and node C. Compared to the result in Fig. 26, the capacitor $C_{A-C}$ lowers the phase
margin from $-42.9^\circ$ to nearly $-180^\circ$. While increasing the value of $C_{A-C}$ moves the
dominant pole closer to the origin, it also creates a zero in the right-half plane that
worsens the phase margin.

Fig. 31. Simulated open-loop frequency response of the amplifier with a Miller
capacitor (MC) between nodes A and C.

For the circuit in Fig. 27, there is a clear benefit of placing a Miller compensation
capacitor $C_{C-D}$ between nodes C and D. The corresponding simulation results in Fig. 32
exhibit the improvement of the phase margin for Miller capacitor (MC) values from
0.5pF to 1.5pF. Although the phase margin improves to approximately $26^\circ$ with
$C_{C-D} = 1.5\, pF$, a more conservative value of more than $45^\circ$ or even $60^\circ$ is preferred for
enhanced stability and settling of the transient response. However, increasing the
capacitance of the Miller capacitor is associated with a tradeoff between bandwidth and “pole splitting” as can be observed from Fig. 26 and Fig. 32. After moving the dominant pole from the initial position in Fig. 26 to the new position shown in Fig. 32, the -3dB corner frequency (~260kHz) is located closer to the origin, which implies a loss of bandwidth and a slower transient settling response. It is also desirable to limit the capacitance values since large capacitors consume significant die area.

Fig. 32. Simulated open-loop frequency response of the amplifier with a Miller capacitor (MC) between nodes C and D.

As shown in the complete version of the proposed amplifier output stage in Fig. 20, a resistor ($R_2$) can be added in series with $C_{C-D}$ to move the zero in the right-half plane from the Miller compensation towards the left-half plane. Fig. 33 displays the simulated
effect of such an additional resistor (R₂ = 20kΩ), which helps to increase the phase margin to 86.86° for stable operation. Fig. 34 and Fig. 35 summarize the pole and zero locations of this compensated amplifier in the complex plane.

Fig. 33. Simulated open-loop frequency response of the amplifier after adding R₂ in series with the Miller capacitor (C_{C-D}).
Fig. 34. Pole and zero plot after compensation.

<table>
<thead>
<tr>
<th>Poles (Hz)</th>
<th>Real</th>
<th>Imaginary</th>
<th>Qfactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pole_1</td>
<td>-9.85072e+00</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Pole_2</td>
<td>-9.85891e+00</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Pole_3</td>
<td>-2.05424e+03</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Pole_4</td>
<td>-5.29836e+06</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Pole_5</td>
<td>-4.51059e+08</td>
<td>5.62855e+03</td>
<td>8.16971e-01</td>
</tr>
<tr>
<td>Pole_6</td>
<td>-4.51059e+08</td>
<td>-5.82855e+08</td>
<td>8.16971e-01</td>
</tr>
<tr>
<td>Pole_7</td>
<td>-3.86616e+09</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zeros (Hz)</th>
<th>Real</th>
<th>Imaginary</th>
<th>Qfactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero_1</td>
<td>-9.85072e+00</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Zero_2</td>
<td>-9.85891e+00</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Zero_3</td>
<td>-1.50615e+07</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Zero_4</td>
<td>-1.11313e+09</td>
<td>0.00000e+00</td>
<td>5.00000e-01</td>
</tr>
<tr>
<td>Zero_5</td>
<td>2.00420e+09</td>
<td>0.00000e+00</td>
<td>-5.00000e-01</td>
</tr>
<tr>
<td>Zero_6</td>
<td>7.13881e+09</td>
<td>0.00000e+00</td>
<td>-5.00000e-01</td>
</tr>
<tr>
<td>Zero_7</td>
<td>5.51559e+10</td>
<td>0.00000e+00</td>
<td>-5.00000e-01</td>
</tr>
</tbody>
</table>

Fig. 35. Pole and zero summary after compensation.
3.2.4 Low-Power Sensor Design and Comparison

A reduction of the sensor’s power consumption is generally desirable for low-power applications, but it is also associated with sensitivity and dynamic range tradeoffs. The amplifier discussed in the previous subsection was designed with equal power consumption as the reference amplifier to evaluate the sensitivity and dynamic range improvements. In addition, a low-power version of the amplifier was designed, which is displayed in Fig. 36. Table II lists the component values of all designs in standard CMOS processes, where the layout area estimations are only based on the dimensions of the active and passive components (without routing).

Taking the low-power version of the proposed amplifier as example, some key design observations are of help to optimize the linearity and dynamic range in general: The DC current flowing through the output branch (M5, M6 in Fig. 36) of the second amplifier (A2 in Fig. 16) should be larger relative to the same current of A1 in the first stage. This is due to the fact that the output voltage swing that has to be tolerated at A1 is significantly lower than at A2. The bias current I_{B1} in Fig. 36 is decreased to reduce the currents in the amplifier core branches of both stages in the power-optimized sensor version. However, the current in the output branch of the first stage was scaled down more aggressively such that it consumes approximately 40% of the total power. Moreover, the channel lengths of transistors M2 and M3 in the low-power amplifier version have been shortened to 180nm compared to the other designs (360nm). Overall,
the DC gain of the low-power amplifier stage is still above 90dB, which is primarily thanks to the high drain-source resistances of the transistors with small bias currents.

Since the proposed design and the low-power design have high impedances at nodes \( N_{o1} \) and \( N_{o2} \) in Fig. 36, it is beneficial to regulate the common-mode of the output stage with a CMFB circuit. Without CMFB2, device mismatches from process variations may cause offset current between \( M_4 \) and \( M_7 \), which could lead to a significant operating point change for \( M_5 \) or \( M_6 \). To reduce the sensitivity to process variations, designers may
decide to add CMFB2 at nodes No1 and No2, which was done for the low-power design version (Fig. 36) as a proof-of-concept using the CMFB circuit in Fig. 18. The stability of CMFB1 and CMFB2 have both been assured during simulations, which have phase margins of 80.2° and 99.9°, respectively.

3.2.5 Fabricated Sensor Version

The sensor designs discussed in this thesis so far were designed with TSMC 1.8V 0.18μm CMOS technology since the initial focus was to improve the performance compared to the existing reference design and to perform comparative simulations. A redesign was completed after access has become available to fabricate a prototype chip in Dongbu 1.2 V 0.11μm CMOS technology. Additional amplification stages were used to maintain high voltage gain with the lower supply voltage and lower voltage gain per stage due to channel length reductions. Fig. 37 displays the schematic of the amplifier that was designed for chip fabrication with 0.11μm CMOS technology.

Fig. 37. Schematic of the amplifier fabricated with Dongbu 0.11μm CMOS technology.
The simulated frequency response of the amplifier designed in 0.11μm CMOS technology is displayed in Fig. 38, indicating a phase margin of 68.02°.

Fig. 38. Simulated open-loop frequency response of the amplifier fabricated in 0.11μm Dongbu CMOS technology.
### Table II. Amplifier component values

<table>
<thead>
<tr>
<th>Main Amplifier</th>
<th>Reference Design</th>
<th>Proposed Amplifier</th>
<th>Low-Power Version</th>
<th>Fabricated Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>TSMC 1.8V 0.18μm CMOS</td>
<td>TSMC 1.8V 0.18μm CMOS</td>
<td>TSMC 1.8V 0.18μm CMOS</td>
<td>Dongbu 1.2V 0.11μm CMOS</td>
</tr>
<tr>
<td>Figure</td>
<td>Fig. 17</td>
<td>Fig. 22</td>
<td>Fig. 36</td>
<td>Fig. 37</td>
</tr>
<tr>
<td>M₁ (width/length)</td>
<td>12μm/360nm</td>
<td>11.9μm/405nm</td>
<td>9.65μm/180nm</td>
<td>20μm/130nm</td>
</tr>
<tr>
<td>M₂ (width/length)</td>
<td>4.5μm/360nm</td>
<td>5.2μm/360nm</td>
<td>7μm/180nm</td>
<td>2.4μm/400nm</td>
</tr>
<tr>
<td>M₃ (width/length)</td>
<td>4.64μm/360nm</td>
<td>4.5μm/360nm</td>
<td>1.3μm/180nm</td>
<td>1.2μm/400nm</td>
</tr>
<tr>
<td>M₄ (width/length)</td>
<td>90μm/180nm</td>
<td>1.09μm/405nm</td>
<td>1μm/12μm</td>
<td>1μm/5.2μm</td>
</tr>
<tr>
<td>M₅ (width/length)</td>
<td>9μm/180nm</td>
<td>6.6μm/360nm</td>
<td>8.2μm/900nm</td>
<td>8μm/130nm</td>
</tr>
<tr>
<td>M₆ (width/length)</td>
<td>\</td>
<td>1.3μm/360nm</td>
<td>1.6μm/900nm</td>
<td>5.4μm/520nm</td>
</tr>
<tr>
<td>M₇ (width/length)</td>
<td>\</td>
<td>7.1μm/360nm</td>
<td>12μm/450nm</td>
<td>8μm/130nm</td>
</tr>
<tr>
<td>M₈ (width/length)</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>1.2μm/600nm</td>
</tr>
<tr>
<td>M₉ (width/length)</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>5.6μm/300nm</td>
</tr>
<tr>
<td>M₁₀ (width/length)</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>1.2μm/3μm</td>
</tr>
<tr>
<td>M₁₁ (width/length)</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>60μm/200nm</td>
</tr>
<tr>
<td>C₁</td>
<td>\</td>
<td>500fF</td>
<td>40fF</td>
<td>\</td>
</tr>
<tr>
<td>C₂</td>
<td>\</td>
<td>1.5pF</td>
<td>1.5pF</td>
<td>1pF</td>
</tr>
<tr>
<td>R₁</td>
<td>\</td>
<td>2.272KΩ</td>
<td>4.272KΩ</td>
<td>\</td>
</tr>
<tr>
<td>R₂</td>
<td>\</td>
<td>15KΩ</td>
<td>20KΩ</td>
<td>12KΩ</td>
</tr>
<tr>
<td>Iᴮ₁</td>
<td>100µA</td>
<td>100µA</td>
<td>10µA</td>
<td>5µA</td>
</tr>
<tr>
<td>Iᴮ₂</td>
<td>100µA</td>
<td>\</td>
<td>\</td>
<td>5µA</td>
</tr>
<tr>
<td>Layout Area Estimate</td>
<td>0.0035mm²</td>
<td>0.0049mm²</td>
<td>0.00595mm²</td>
<td>0.01201mm²</td>
</tr>
<tr>
<td>CMFB₁ (Fig. 18)</td>
<td>Reference Design</td>
<td>Proposed Amplifier</td>
<td>Low-Power Version</td>
<td>Fabricated Design</td>
</tr>
<tr>
<td>Mᴮ</td>
<td>4μm/360nm</td>
<td>4μm/360nm</td>
<td>14.7μm/180nm</td>
<td>2.1μm/200nm</td>
</tr>
<tr>
<td>Mᴺ</td>
<td>10μm/360nm</td>
<td>10μm/360nm</td>
<td>14μm/180nm</td>
<td>1.2μm/300nm</td>
</tr>
<tr>
<td>Mᴾ</td>
<td>6μm/360nm</td>
<td>5μm/360nm</td>
<td>11.1μm/180nm</td>
<td>400nm/130nm</td>
</tr>
<tr>
<td>Mᴼ</td>
<td>12μm/360nm</td>
<td>10μm/360nm</td>
<td>2.2μm/180nm</td>
<td>800nm/130mm</td>
</tr>
<tr>
<td>I CMFB⁰</td>
<td>91.8µA</td>
<td>116.8µA</td>
<td>8µA</td>
<td>5.88µA</td>
</tr>
<tr>
<td>Layout Area Estimate</td>
<td>0.00026mm²</td>
<td>0.0002mm²</td>
<td>0.00026mm²</td>
<td>0.000539mm²</td>
</tr>
<tr>
<td>CMFB₂ (Fig. 18)</td>
<td>Reference Design</td>
<td>Proposed Amplifier</td>
<td>Low-Power Version</td>
<td>Fabricated Design</td>
</tr>
<tr>
<td>Mᴮ</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>42μm/180nm</td>
</tr>
<tr>
<td>Mᴺ</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>24µm/180nm</td>
</tr>
<tr>
<td>Mᴾ</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>6µm/450nm</td>
</tr>
<tr>
<td>Mᴼ</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>13µm/450nm</td>
</tr>
<tr>
<td>I CMFB⁰</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>5µA</td>
</tr>
<tr>
<td>Layout Area Estimate</td>
<td>\</td>
<td>\</td>
<td>\</td>
<td>0.000432mm²</td>
</tr>
</tbody>
</table>

Note: The values in the table represent specific component sizes and values for different designs and versions of amplifiers. The table includes technology details, amplifier designs, and layout area estimates for various components and versions.
Fig. 39. Layout of the first amplifier stage (fabricated version).

Fig. 40. Prototype chip layout.
Table II summarizes the key component values and device dimensions of each amplifier design. Fig. 39 displays the layout of one amplifier stage in the fabricated design. Even though the transistor dimensions of the first stage and second stage amplifiers are different, their layout areas are approximately the same.
4. SIMULATION RESULTS

4.1 COMMON-MODE FEEDBACK STABILITY

As described in Section 3.2.2, a high-gain differential amplifier with multiple stages requires one or multiple common-mode feedback (CMFB) circuits to ensure proper DC operating points in the presence of mismatches after fabrication.

The “stb” analysis tool in Cadence was adopted to assess the stability of the CMFB loop. Both, the amplifier stage of interest and the CMFB circuit, were placed in the same

Fig. 41. Screenshot of the Cadence schematic (fabricated design in Fig. 38).
schematic (as shown in Fig. 41) to verify the gain and phase margin of the loop. The “iprobe” component was inserted in the feedback path from the CMFB to the main amplifier. Fig. 42 and Fig. 43 illustrate the stability and loop gain of the CMFB, showing that the simulated phase margin is high enough to ensure stability.

![Screenshot of the CMFB loop stability summary.](image1)

**Fig. 42.** Screenshot of the CMFB loop stability summary.

![Gain and phase of the CMFB loop.](image2)

**Fig. 43.** Gain and phase of the CMFB loop.
4.2 AMPLIFIER NOISE

For comparison purposes, the noise analysis tool in Cadence was used to simulate the noise performance of each amplifier design by integrating the noise from 1Hz up to the unity-gain frequency of each amplifier. Fig. 44, Fig. 45 and Fig. 46 display the noise summaries after the simulations of the proposed design, low-power version, and fabricated design, respectively. It can be observed from the results that the total input-referred noise strongly depends on the transistors of the differential pair amplification stage (M2 and M3 in Fig. 22, Fig. 36 and Fig. 37).

![Noise simulation summary for the proposed amplifier (including CMFB).]
Fig. 45. Noise simulation summary for the low-power design (including CMFB).
Fig. 46. Noise simulation summary for the fabricated design (including CMFB).

4.3 SENSOR PERFORMANCE

Fig. 16 illustrates the topology of the reference design, which requires a high-gain off-chip amplifier $A_2$. The calibration current and the supply current $I_{\text{core}}$ are provided on the printed circuit board (PCB). As a result of mismatches and process variations after fabrication, the differential output voltage is normally nonzero even without the activation of a heat source close to the sensing devices. The calibration current sources $I_{\text{cal1}}$ and $I_{\text{calref}}$ in Fig. 16 allow compensation for electrical mismatches from process
variations in the sensor circuitry as well as for thermal gradients from other nearby power-dissipating devices on the chip. The measurement results in [5] demonstrate that the calibration current sources in this sensor architecture allow to counteract the interference from a 15.9mW power source directly next to one of the PNP transistors. As elaborated in [5], the calibration currents must be adjusted prior to the measurements until \( \Delta V_o = V_{o2} - V_{o1} \approx 0 \). Afterwards, the test signal can be applied to the circuit under test to measure its power dissipation using the on-chip temperature sensor. If other circuits (DC bias circuits, the sensor core amplifiers, etc.) are very close to the PNP transistors, then it is required that their DC power dissipation does not change during the measurement because a temperature difference due to such a DC power dissipation change would require a new calibration.

From Fig. 16, the current source \( I_{\text{core}} \) is transferred through a 10:1 current mirror that is integrated on the chip. This current determines the base bias current of the PNP devices. The electro-thermal coupling can be modeled with a voltage-controlled voltage source that modulates the base-emitter voltage \( (V_{be}) \) with a gain of \( k \approx -1.8\text{mV/K} \), such that

\[ \Delta V_{be} = k \times \Delta T. \]

The resulting change of the emitter current is

\[ \Delta I_e = I_{e1} - I_{e2} = \Delta V_{be} \times g_{mQ}, \]

where \( g_{mQ} \) is the transconductance of the PNP device. Hence, the sensitivity of the differential current can be defined as

\[ S_T = \frac{\Delta I_e}{\Delta T} = k \times g_{mQ}. \]

Correspondingly, the sensitivity of the differential output voltage \( (\Delta V_o = V_{o2} - V_{o1}) \) is

\[ S_V = \frac{\Delta V_o}{\Delta T} = S_T \times (R_f / R_L) \times R_2. \]

When the base current \( I_B \) of the PNP transistor is increased, the collector
current $I_C = \beta I_B$ and the emitter current $I_E = (1+\beta)I_B$ both increase, and the transconductance $g_{mQ} = I_C/V_T$ of the PNP transistor increases as well. Thus, increasing $I_{core}$ generally yields sensitivity improvements if the devices in the amplifier operate in the linear range. However, the transconductance gain of the PNP device is not the only factor that affects the sensitivity of the temperature sensor. The sensitivity also depends on the input impedance ($Z_{in}$) of the first amplification stage in the sensor, which is approximately inversely proportional to the voltage gain ($A_v$) of the amplifier when this gain is high enough.

As depicted in Fig. 47, the electro-thermal sensing can be represented with a simplified small-signal equivalent model, where the temperature-controlled current in gray ($=S_T\cdot1/2\Delta T$) is generated through electro-thermal coupling. $Z_{in}$ represents the input impedance of the first amplifier stage in the sensor. A smaller input impedance improves the sensitivity because higher fractions of the “small-signal” emitter currents ($I_{\Delta T}$ in

---

**Fig. 47.** Simplified small-signal equivalent circuit of the sensing PNP devices.
response to $\Delta T$) flow into or out of the amplifier instead of through the emitter-collector resistances of the BJTs. The amplifier input impedance $Z_{in}$ can be expressed as [5]:

$$Z_{in} \approx \frac{R_i}{1 + A_{v1} \left( \frac{R_L}{R_L || R_i} \right)} \approx \frac{R_i}{1 + A_{v1}}, \quad (22)$$

where $r_o$ and $A_{v1}$ are the output resistance and loaded voltage gain of amplifier $A_1$ in Fig. 16 (with the values listed in Table II). Fig. 48 to Fig. 51 show the plots of the input impedance magnitudes, loaded voltage gains, and phase responses vs. frequency for the different amplifier designs. Equation (22) illustrates that increasing $A_{v1}$ reduces the input impedance of the sensor and thereby improves its sensitivity. As can be observed in Fig. 48 the reference design has an input impedance of $Z_{in} = 1052\Omega$. In comparison, the input impedance is significantly lower for the proposed design in Fig. 49 ($2.201\Omega$) and the low-power design in Fig. 50 ($5.31\Omega$).
Fig. 48. Closed-loop input impedance, gain and phase of the reference design.
Fig. 49. Closed-loop input impedance, gain and phase of the proposed design.
Fig. 50. Closed-loop input impedance, gain and phase of the low-power design.
Fig. 51. Closed-loop input impedance, gain and phase of the fabricated design.
Fig. 52 shows the simulated differential output voltage of the sensor designs with different amplifiers as a function of the temperature difference between the two PNP sensing devices. The temperature difference ($\Delta T$) was swept by sweeping the SPICE parameter $\text{trise}$ for one of the PNP transistors. In each case, the maximum sensitivity setting was used by choosing the maximum bias current $I_{\text{core}}$ (in Fig. 16) for each design with which all MOSFET transistors in the amplifiers can operate in the saturation region (reference amplifier: $I_{\text{core MAX}} = 800 \mu$A, proposed amplifier: $I_{\text{core MAX}} = 740 \mu$A, low-power amplifier: $I_{\text{core MAX}} = 480 \mu$A, fabricated amplifier: $I_{\text{core MAX}} = 200 \mu$A). The fabricated design was simulated with Dongbu 0.11$\mu$m CMOS technology models, while other designs were simulated with TSMC 0.18$\mu$m CMOS technology models. Fig. 52 illustrates the significant sensitivity and dynamic range improvements of the sensors.
with the proposed amplifier design in contrast to the reference design. Note that the saturation voltage $V_{DS\text{sat}}$ of the transistors in the class AB output stage limits the output voltage swing. In the version of the proposed design with higher power consumption, the drain current through the output transistors is larger than in the low-power version. Hence, a higher drain-source voltage is needed to keep the transistors in saturation when they are driving the output node. As a result of the smaller saturation voltage, the low-power version can tolerate a larger output voltage swing but has less linearity in the linear range than the proposed design with higher power consumption. Note that the fabricated amplifier (0.11μm CMOS technology, 1.2V supply) has a smaller dynamic range than the proposed design (0.18μm CMOS technology, 1.8V supply) due to the voltage headroom limitations with a lower supply voltage. In general, the coupling of the power dissipation from the device under test to the temperature change ($\Delta T$) at the test point depends on the distance of the sensing PNP device among other factors. From the experimental measurements of the test chip in [5] for example, a coupling of approximately 0.667°C/mW can be deduced for the presented case. With such an electro-thermal coupling, even the low-power version of the sensor (with ±4°C range) would be able to monitor the power dissipation changes in a single device up to ±2.7mW.

The sensitivity can be tuned by changing the value of $I_{\text{core}}$. As shown for the proposed design in Fig. 53, the differential output voltage ($\Delta V_o$) has the same swing from -1.5V to
1.5V as the sensitivity varies from 207.1mV/°C ($I_{core} = 740\mu A$) to 55.94mV/°C ($I_{core} = 100\mu A$). Hence, the current $I_{core}$ can be set for the appropriate dynamic range to monitor power dissipation at various devices on the chip.

Fig. 53. Tuning of $I_{core}$ to adjust the sensitivity of the proposed design.

Fig. 54 provides a visual comparison of the $I_{core}$ tuning ranges of the different designs. In this figure, the maximum sensitivity values were extrapolated from the -1°C to 1°C range. When $I_{core}$ is lower than 280μA, the sensitivity of the low-power version is higher than that of the proposed design because the low-power amplifier has a higher gain due to the larger output resistance with lower bias currents. Consequently, the impedance ($Z_{in} \approx R_1 / [1 + A_1]$) seen at the input of the amplifier in Fig. 16 (where the emitters of $Q_1$ and $Q_{ref}$ are connected) is smaller in the low-power version, which improves the
sensitivity because higher fractions of the “small-signal” emitter currents (ΔI_e in response to ΔT) flow into or out of the amplifier instead of through the emitter-collector resistances of the BJTs. On the other hand, as I_{core} is increased above 280μA, the DC currents flowing through the feedback resistors in Fig. 16 become larger, requiring the amplifiers to provide larger DC output currents. Since the proposed design with higher power consumption has larger bias currents, it can operate properly with larger values of I_{core} before internal transistors saturate due to current-handling limitations.

Table III summarizes the simulated specification parameters of the amplifier designs (A_1) under evaluation. The gains of the proposed amplifier and its low-power version are over 90dB, while the reference design only has a gain close to 30dB. Since the input resistance looking into amplifier A_1 with a voltage gain of A_{v1} is approximately equal to R_j/(1+A_{v1}), the increased gain allows more current to flow from the PNP sensing devices
(with finite output resistance) into the amplifiers of the sensor. Hence, the high gain improves the sensor’s sensitivity. Note that the low-power version of the proposed amplifier also has a high voltage gain while consuming less than half of the power. The related reduction of the 3dB frequency is not a drawback because the variable of interest in this thermal monitoring approach is the sensor’s differential DC output voltage.

Table III. Simulated specifications of the amplifier (A₁) designs

<table>
<thead>
<tr>
<th>Version</th>
<th>Ref. Design</th>
<th>Proposed Amplifier</th>
<th>Low-Power Version</th>
<th>Fabricated Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>0.18μm CMOS</td>
<td>0.18μm CMOS</td>
<td>0.18μm CMOS</td>
<td>0.11μm CMOS</td>
</tr>
<tr>
<td>Supply Voltage</td>
<td>1.8V</td>
<td>1.8V</td>
<td>1.8V</td>
<td>1.2V</td>
</tr>
<tr>
<td>DC Gain (dB)</td>
<td>30.8</td>
<td>92.3</td>
<td>94.14</td>
<td>126.69</td>
</tr>
<tr>
<td>f₃dB (MHz)</td>
<td>1.76</td>
<td>0.043</td>
<td>6.93×10⁻⁴</td>
<td>5.4×10⁻⁵</td>
</tr>
<tr>
<td>Unity Gain Frequency, fᵤ (MHz)</td>
<td>60.90</td>
<td>819.0</td>
<td>36.32</td>
<td>45.94</td>
</tr>
<tr>
<td>Phase Margin (Degree)</td>
<td>91.0</td>
<td>60.2</td>
<td>86.73</td>
<td>68.02</td>
</tr>
<tr>
<td>Input-Referred Noise (µV) [integrated over fᵤ]</td>
<td>80.3</td>
<td>1207.5</td>
<td>187.9</td>
<td>239.6</td>
</tr>
<tr>
<td>CMRR (dB at 10KHz)</td>
<td>20.6</td>
<td>24.6</td>
<td>63.63</td>
<td>43.64</td>
</tr>
<tr>
<td>PSRR (dB at 10KHz)</td>
<td>32.2</td>
<td>92.7</td>
<td>30.2</td>
<td>37.6</td>
</tr>
<tr>
<td>Output Resistance (kΩ)</td>
<td>0.256</td>
<td>86.14</td>
<td>260</td>
<td>73.24</td>
</tr>
<tr>
<td>CMFB1 Loop: DC Gain/Phase Margin (dB/Degree)</td>
<td>33.9/89.6</td>
<td>36.5/85.3</td>
<td>31.36/80.2</td>
<td>30.10/84.55</td>
</tr>
<tr>
<td>Power Consumption (mW) [A₁ + CMFB]</td>
<td>1.062</td>
<td>1.308</td>
<td>0.398</td>
<td>0.231</td>
</tr>
<tr>
<td>Total Power* (mW)</td>
<td>2.22</td>
<td>2.23</td>
<td>1.125</td>
<td>0.752</td>
</tr>
</tbody>
</table>

Table IV summarizes sensitivities and the dynamic ranges from the plots. The listed values are the actual maximum sensitivity values over the complete dynamic range.
Compared to the sensor with the reference amplifier, the sensitivity of the output voltage is 10 times higher for the sensor with the proposed amplifier with same power consumption. However, in the maximum sensitivity setting, the dynamic range with the proposed amplifier is ±3.3°C narrower.

Table IV. Dynamic range comparison with maximum sensitivity

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity ($S_V = \Delta V_o/\Delta T$)</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. Design ($I_{core} = 800\mu A$)</td>
<td>18.0mV/°C</td>
<td>±10.0°C</td>
</tr>
<tr>
<td>Proposed ($I_{core} = 740\mu A$)</td>
<td>207.1mV/°C</td>
<td>±6.7°C</td>
</tr>
<tr>
<td>Low-Power Version ($I_{core} = 480\mu A$)</td>
<td>185.7mV/°C</td>
<td>±4.0°C</td>
</tr>
<tr>
<td>Fabricated Version ($I_{core} = 200 \mu A$)</td>
<td>115.6mV/°C</td>
<td>±8.4°C</td>
</tr>
</tbody>
</table>

To compare the designs over the same dynamic range of ±25°C, they were simulated with the $I_{core}$ values listed in Table V. With an assumed coupling of 0.667°C/mW (as in [5]), a ±25°C range would enable monitoring of an estimated ±16.7mW power dissipation change in a nearby transistor or resistor under test. In this reduced gain mode for equal dynamic range, the sensors with the proposed amplifier have an almost 10 times higher output voltage sensitivity than the reference design.
Table V. Comparison over equal wide dynamic range with reduced $I_{\text{core}}$

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity $\left( S_V = \frac{\Delta V}{\Delta T} \right)$</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. Design ($I_{\text{core}} = 60\mu A$)</td>
<td>4.7mV/°C</td>
<td>±25°C</td>
</tr>
<tr>
<td>Proposed ($I_{\text{core}} = 60\mu A$)</td>
<td>37.3mV/°C</td>
<td>±25°C</td>
</tr>
<tr>
<td>Low-Power Version ($I_{\text{core}} = 60\mu A$)</td>
<td>43.5mV/°C</td>
<td>±25°C</td>
</tr>
<tr>
<td>Fabricated Version ($I_{\text{core}} = 60\mu A$)</td>
<td>41.1mV/°C</td>
<td>±25°C</td>
</tr>
</tbody>
</table>

4.4 CALIBRATION

Since process variations during fabrication are unavoidable, it is necessary to complete a sensor calibration procedure before activating the heat sources of interest and performing accurate measurements. In Fig. 16, two calibration currents are included at the emitter of each PNP device through the use of a 10:1 current mirror. As explained in the first paragraph of Section 4.3, these currents allow counteracting the effects of electrical mismatches in the sensor circuitry as well as from static temperature imbalances caused by DC biasing power dissipation of other nearby circuits on the chip. To verify the adjustment range of the differential output voltage (ideally: $\Delta V_o = V_{o2} - V_{o1} \approx 0$), the calibration currents can be swept during simulations. As an example, Fig. 55 illustrates the impact of the calibration current $I_{\text{cal1}}$ in Fig. 16. In this sensor design, a calibration current of 300μA permits to shift the sensor’s characteristic curve by 10°C. In general, the tuning range for this sensor is approximately equal to half of the dynamic range for a
given sensitivity setting, which is more than sufficient according to the measured results in [5] with the same calibration approach.

In this thesis project, the calibration process was performed manually. However, future work can be directed towards automating the calibration steps using a comparator, on-chip digital control logic, and digitally programmable calibration currents.

![Simulated effect of a calibration current.](image)

**Fig. 55.** Simulated effect of a calibration current.
5. PROTOTYPE CHIP AND PRINTED CIRCUIT BOARD DESIGN

5.1 FABRICATED CHIP AND BONDING DIAGRAM

Fig. 56 shows the bonding diagram of the chip and a photo of the bonding in the DIP20 package. Fig. 57 displays micrographs of the chip surface in comparison to the layout screenshot from Cadence Virtuoso. The automatic filling option for all layers was selected to meet the density requirements during fabrication. Only some parts of the fabricated circuit are observable in the zoomed-in view of the micrograph, but the features are more clear and annotated in the layout screenshots of Fig. 57 and Fig. 39.

#### Fig. 56. Bonding diagram vs. photo of the packaged die.
5.2 MEASUREMENT SETUP

To obtain the experimental results, a printed circuit board (PCB) was designed and assembled to provide off-chip bias and supply voltages and bias currents for the prototype chip. The KiCad software was utilized for PCB design and layout. Fig. 58
displays the complete schematic of the PCB from KiCad. A voltage regulator (Texas Instruments LM150) provides a supply voltage ($V_{DD}$) of 1.2V. The test heat sources on the chip are 50Ω resistors and diode-connected transistors, which can be activated by injecting DC currents that are generated on the PCB. The test resistors and diode-connected transistors on the chip are placed in the vicinity of the PNP sensor devices with distances of 0.18μm. Furthermore, the PCB was designed such AC signals can be coupled to the test heat sources through SMA connectors. Large 100μF capacitors are connected at the differential outputs of the sensor to filter out noise. Noise bypass capacitors were also added at bias voltage and current inputs to the chip.
Fig. 58. PCB schematic.
Fig. 59. Assembled PCB for the prototype chip.

Fig. 59 displays the assembled PCB and Fig. 60 illustrates the measurement setup. The heat sources (test resistors or diode-connected transistors) can be activated by both DC currents and RF signals. In both cases, the DC component of the power applied at the test heat source (device under test) is of primary interest because this sensor was designed to measure only the DC component. If an RF signal is applied, then the down-converted DC power component is detected via electro-thermal coupling as described in Section 2.3.3. Bias voltages and currents can be generated on the PCB using
potentiometers and resistive dividers connected to the on-board supply voltage. The jumpers on the PCB can be used to activate or deactivate bias circuits, and to measure currents with a handheld multimeter. The power dissipation of the test heat sources can be measured from the voltages and currents at the input/output pin of the chip. In practice, the input DC power at a heat source causes a temperature change that can be detected by the nearby sensor and measured with the differential DC output voltage using a digital multimeter as shown in Fig. 60.

Fig. 60. Measurement setup.
5.3 MEASUREMENT RESULTS

As mentioned in the previous section, the prototype chip contains two kinds of test heat sources: diode-connected transistors and polysilicon resistors. First, the diode-connected transistors were activated with DC currents to generate temperature variations to be sensed by the PNP devices that are 0.18μm away from the diode-connected transistors. The sensor sensitivity to power dissipation in the two diode-connected transistors (abbreviated as DIO₁ and DIO₂) with $I_{\text{core}} = 100\mu\text{A}$ can be obtained from the measurement results in Fig. 61. While this sensitivity is equal to $395.4\text{mV/mW}$, the measured dynamic range is ±520μW. Note that this dynamic range is very narrow compared with the expectation based on the simulation results in Table IV with $I_{\text{core}} = 200\mu\text{A}$: ±8.4°C range with the temperature-to-power coupling of approximately $0.667\text{°C/mW}$ (from [5]) leads to an expected power range up to 12.59mW, which should be even larger with $I_{\text{core}} = 100\mu\text{A}$. The reason for the limited dynamic range of this design is that the operating points of several transistors in the amplifier had to be corrected to ensure that the transistors operate in the saturation region after fabrication process variations. In particular, the transistors in the output stage must have enough voltage headroom. The fabricated amplifier is very sensitive to small mismatches and variations due to the high impedance nodes at the drains of $M₄$-$M₉$ in Fig. 37. In retrospect, this multi-stage design can be improved by using multiple CMFB circuits to regulate the common-mode voltages at each stage. Alternatively, it can be explored to reduce the number of stages between the amplifier core ($M₁$-$M₃$) and the class AB output stage ($M₁₀$, $M₁₁$).
Fig. 61. Measured differential output voltage vs. power dissipation in diode-connected transistors (DIO₁ and DIO₂) close to the PNP devices on each side of the sensor (I_{core} = 100μA).

Fig. 62 displays the measured differential output voltage of the sensor (with I_{core} = 300μA) for different power dissipations in the test resistors that are close to the PNP devices on each side of the sensor. Contrary to the plots in Fig. 61 obtained by applying power to the diode-connected transistors, the plots in Fig. 62 reveal that the sensor output in response to power dissipation in the test resistors is not symmetrical. Even though process variations between resistors R_{t₁} and R_{t₂} site are a possible cause of the asymmetric response, a more significant factor is that the test resistor layout is not symmetric with respect to the position of the PNP sensing devices, as illustrated in Fig. 63. Compared to R_{t₁}, the resistor R_{t₂} has an offset of 5.2μm in the vertical direction, which was an oversight during the layout verification. This placement difference of R_{t₁} and R_{t₂} in relation to the respective PNP sensing devices affects the thermal coupling.
While the dynamic range for power detection in $R_{t2}$ is up to 16μW in this high sensitivity setting (with high $I_{\text{core}}$ value) according to Fig. 62, the sensor already saturates with 5.2μW of power dissipation in $R_{t1}$. However, the measurement results also indicate that the sensitivity at $R_{t2}$ is high (5841.65mV/mW.) The use of a large bias current ($I_{\text{core}}$) and the high sensitivity to variations in the amplifier described in the previous paragraph both lead to a significantly smaller dynamic range than expected from simulations. Significant improvements can be expected after the aforementioned amplifier design improvements and careful placement of devices in the layout, especially in consideration of the results in [5] where the responses to power dissipation in the test resistors were symmetrical because of symmetric layout. Although the measurement results for the fabricated sensor in this thesis have revealed some drawbacks, its sensitivity is still significantly higher (∼10 times larger) than that of the reference design for comparable values of $I_{\text{core}}$, which demonstrates the main merit of the fabricated design.

To evaluate the sensor sensitivity tuning with $I_{\text{core}}$, different $I_{\text{core}}$ values (25μA to 200μA) were used during testing. The DIO$_1$ test heat source was used for this characterization. Fig. 64 shows the comparison of the sensor sensitivity for different values of $I_{\text{core}}$. By changing $I_{\text{core}}$ between 25μA and 150μA, the sensitivity of temperature sensor can be adjusted from 286.9mV/mW to 427.42mV/mW. When $I_{\text{core}}$ is larger than 150μA, the sensitivity does not increase further because the transistors in the amplifier do not operate in the saturation region anymore.
Fig. 62. Measured differential output voltage vs. power dissipation in test resistors $R_{t1}$ and $R_{t2}$ close to the PNP devices on each side of the sensor ($I_{core} = 300\mu A$).

$$I_{R_{t1}} \approx 85\mu m$$
$$I_{R_{t2}} \approx 200\mu m$$
$$\approx 5.2\mu m$$

Fig. 63. Asymmetry between test resistor $R_{t1}$ and $R_{t2}$. 
Fig. 64. Measured sensitivity vs. $I_{\text{core}}$. 
6. CONCLUSION AND FUTURE WORK

6.1 SUMMARY

This research addressed the need for alternative sensors to evaluate changes of performance metrics of analog circuits due to CMOS fabrication process variations. As discussed in chapters 1 and 2, multiple methods are employed to monitor and control process variations in order to increase the yield of products, to reduce production costs, and to improve product reliability. Compared to conventional methods, the strategy of RF power detection employed in this thesis is non-intrusive and it avoids degradation of the RF signal path because the sensing involves electro-thermal coupling instead of an electrical connection to the circuit under test. The main contribution of this thesis research is the development of the differential temperature sensor design in Chapter 3, which has higher sensitivity and extended dynamic range than the previous version. A sensor with class AB amplification stages was designed in standard 0.18μm CMOS technology and compared to a prior version using the same sensor architecture and power consumption constraint. The simulation results in Chapter 4 indicate that the class AB output stage improves the maximum permissible output swing by a factor of 10, allowing to increase the sensor’s sensitivity by a factor of almost 10 over the same dynamic range. A low-power version of the sensor with class AB output stage has a simulated sensitivity up to 185.7mV/°C over an 8°C dynamic range. The complete sensor consumes a total of 1.13mW from a 1.8V supply. Chapter 5 provides an overview of the printed circuit board design and experimental setup to perform measurements.
using a version of the sensors fabricated in 0.11μm CMOS technology. The prototype chip measurements confirmed that the sensor has a high sensitivity of 395.4mV/mW. However, the extended output voltage range could not be verified due to the high sensitivity to process variations of this design because only the first stage in the amplifier is regulated with a common-mode feedback circuit. It is expected that the addition of more common-mode feedback circuits or the reduction of the number of amplifier stages can improve the design’s robustness to process variations and device mismatches.

6.2 FUTURE WORK
The measurement results suggest that an extended dynamic range will require a reduction of the amplifier’s sensitivity to process variations and device mismatches. Generally, it is preferable to avoid the stages between the amplifier core and the class AB output stage, or to regulate each stage with a common-mode feedback circuit. For example, the amplifier version in Fig. 36 would have been a better candidate for the test chip fabrication. This amplifier version can be designed with higher gain by increasing the channel lengths of the transistors to boost the impedances at internal nodes and at the output. In this work, manual calibration was performed prior to the measurements by adjusting the two calibration currents such that the differential output voltage is zero prior to the measurement. This process could be automated by comparing the two single-ended outputs with a comparator, and developing a digital control loop that tunes the reference currents until the differential output voltage is sufficiently close to zero.
REFERENCES


