A 405-nW CMOS Temperature Sensor Based on Linear MOS Operation

Man Kay Law, *Student Member, IEEE*, and A. Bermak, *Senior Member, IEEE*

Abstract—This brief presents a CMOS temperature sensor suitable for ultralow-power applications. With a MOS transistor operating in the linear region, a linear relationship between delay and temperature can be obtained. A differential sensing architecture is utilized to reduce the signal offset and increase the effective signal-to-noise ratio. A design methodology concerning power optimization and improved sensor linearity is also presented. The sensor, which occupies 0.0324 mm², is fabricated using the TSMC 0.18- μ m one-polysilicon six-metal (1P6M) process. Measurement results show that the sensor consumes 405 nW with a 1-V supply at 1 ksample/s at room temperature. An inaccuracy value of -0.8°C to $+1^\circ\text{C}$ from 0°C to 100°C after calibration is achieved.

Index Terms—CMOS temperature sensor, ultralow-power applications.

I. INTRODUCTION

RECENTLY, various kinds of on-chip CMOS temperature sensors that perform thermal sensing for enhanced system reliability and performance have been reported, including power and temperature control on processors [1], dynamic temperature management [2], dynamic random access memory refresh-rate control [3], and wireless temperature-sensing applications [4]–[6]. The uprising of on-chip temperature sensing is mainly due to the recent rapid development of process technology. The combinational effect of increased functionality, high speed, and small area results in increased self-heating and increased leakage current, which are detrimental to system-level reliability and performance. As a result, on-chip thermal management schemes that enable the realization of enhanced thermal solutions both locally and globally becomes attractive. These embedded sensors usually feature small-area, low-power consumption (for minimal heat dissipation) and digital outputs [2], to achieve minimal cost and perform accurate temperature measurement. Moreover, the inclusion of embedded temperature sensors in passive radio-frequency identification (RFID) tags has recently become more and more attractive toward wireless temperature-sensing applications [4], [5]. As there is only limited power available, temperature sensors that consume submicrowatt power are necessary while maintaining the tag

Fig. 1. Block diagram of the proposed temperature sensor.

performance. In [7], a time-to-digital-converter-based CMOS temperature sensor that utilizes delay lines and a time-to-digital converter (TDC) is presented. This sensor eliminates the use of bipolar junction transistors (BJTs) and the power-consuming analog-to-digital converters (ADCs) [12] while achieving an inaccuracy value of -0.7°C to $+0.9^\circ\text{C}$ from 0°C to 100°C , and consuming 10 μW at 10 ksamples/s. Although this time-domain sensor features more compact size and smaller power overhead, it has relatively poor supply rejection due to the use of simple inverter chains, which also result in relatively large area overhead. A large enough supply voltage is required to maintain high linearity with respect to temperature [7]. In [8], the way of improving the supply sensitivity of the time-domain temperature sensor is introduced. However, a high supply voltage is still required. This, as in the case of [7], makes the sensor not applicable to deep-submicrometer processes.

In this brief, we introduce our proposed time-domain CMOS temperature sensor that is based on the delay generated from a MOSFET operating in the linear region, with the overall system block diagram as shown in Fig. 1. Differential temperature sensing using both proportional-to-absolute temperature (PTAT) and complementary-to-absolute temperature (CTAT) is utilized to reduce the signal offset and to increase the effective signal power at the same time. Optimization on the delay lines to achieve high linearity with respect to temperature, as well as low-power operation, is also presented. Supply voltage change is sensed through the supply voltage sensor, and the signal is feed-forwarded to reduce the output variation.

II. LINEARITY IN DELAY VERSUS TEMPERATURE

The utilization of the propagation delay of simple inverters has been proposed in [7] for temperature sensing instead of the conventional BJT-based implementations due to its highly linear relationship with respect to temperature changes. Apart from that, it also exhibits relatively lower power consumption and smaller area than its ADC counterpart. The average

Manuscript received May 6, 2009; revised September 2, 2009. Current version published December 16, 2009. This work was supported by the Hong Kong Innovation and Technology Fund (ITF), Hong Kong R&D Center for Logistic Supply Chain Management, under Grant ITP/027/07LP. This paper was recommended by Associate Editor B. Otis.

The authors are with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: eeml@ust.hk; eebarmak@ust.hk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSIL.2009.2034201

propagation delay of a simple CMOS inverter, as indicated in [9], can be expressed as

$$t_p = \frac{2C_L V_T}{\mu C_{ox} \left(\frac{W}{L}\right) (V_{DD} - V_T)^2} + \frac{C_L}{\mu C_{ox} \left(\frac{W}{L}\right) (V_{DD} - V_T)} \times \ln \left(\frac{1.5V_{DD} - 2V_T}{0.5V_{DD}} \right) \quad (1)$$

where $\mu C_{ox}(W/L)$ is the transconductance parameter, V_T is the transistor threshold voltage, and C_L is the effective load capacitance. Note that (1) shows the average propagation delay of simple inverters with step input, without considering higher order nonlinearities such as channel length modulation. The temperature dependence of (1), as stated in [10], is determined by the variation of mobility and threshold voltage parameters with respect to temperature, which are expressed as

$$\mu = \mu_0 \left(\frac{T}{T_0} \right)^{km} \quad (2)$$

$$V_T = V_T(T_0) + \alpha(T - T_0) \quad (3)$$

where T_0 is the reference temperature, and μ_0 and $V_T(T_0)$ are the corresponding mobility and threshold voltage at the reference temperature T_0 , respectively. The values of km and α are process dependent and are roughly in the range of -1.2 to 2.42 and -1 to -4 mV/K, respectively [10].

Notice that (1) is composed of two regions: one is charging/discharging through the linear-region current, whereas the other is through the saturation-region current. As the sum of the delays are highly linear to temperature, the corresponding charging/discharging time of the load transistor C_L to 50% of the supply voltage swing using either a linear or a saturation current should also be linearly dependent to temperature. In particular, the corresponding charging/discharging time required by either a linear or a saturation current bias with load C_L and supply voltage V_{DD} with respect to temperature are

$$t_{lin} = \frac{C_L V_{DD}}{2\mu C_{ox} \left(\frac{W}{L}\right) \left((V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right)} \quad (4)$$

$$t_{sat} = \frac{C_L V_{DD}}{\mu C_{ox} \frac{W}{L} (V_{GS} - V_T)^2} \quad (5)$$

Fig. 2 shows the simulation result comparing the linearity of the delay generated by (1), (4), and (5). The simulation is carried out using the TSMC 0.18- μm 1P6M process and with C_L , sizing, and supply voltage equal to 10 pF, 5 $\mu\text{m}/0.5 \mu\text{m}$, and 1.8 V, respectively. As can be observed, both delays induced by linear and saturation currents have comparable inaccuracy to the inverter delay (within $\pm 1^\circ\text{C}$) after two-end-point calibration. As (1) is differentiable over the sensing range from 270 to 370 K, the corresponding Taylor-series expansion can be used to explain the relationship between linearity and supply voltage. It can readily be observed that, for a large enough supply voltage, the high-order terms that contribute to the overall nonlinearity can greatly be suppressed, resulting in high linearity. However, this leads to three major problems: First, this method is not scalable to advanced processes. As the supply voltage is lowered, the linearity with respect to temperature is worse. Second, as the supply voltage is increased for improved linearity, the temperature-induced delay is reduced, meaning a long delay chain is required to increase the signal power,

Fig. 2. Linearity error versus temperature due to inverter delay t_p , linear current delay t_{lin} , and saturation current delay t_{sat} , with $C_L = 10$ pF and $W/L = 5 \mu\text{m}/0.5 \mu\text{m}$, using the TSMC 0.18- μm process.

hence occupying a large area. Third, as the inverter propagation delay is highly dependent on the supply voltage change, the output temperature data can vary according to different supply voltages. This brief demonstrates that these limitations can be eliminated by using a linear-current-generated delay instead, resulting in an area-efficient low-power temperature sensor with high linearity. The optimization of the linear current to achieve even better linearity is justified in Section III-B, whereas the use of the linear current to improve supply sensitivity is explained in Section III-C.

III. PROPOSED TEMPERATURE SENSOR DESIGN

The temperature sensor implementation is shown in Fig. 3. Temperature sensing is accomplished using the PTAT and CTAT delay generators. When the start signal V_{ST} is exerted, both PTAT and CTAT delay lines discharge the corresponding load capacitors, and a signal is triggered upon reaching the switching threshold of the buffers. The signals from the delay lines are then XORed to generate a temperature-modulated pulsewidth (PW). V_{FB} , which indicates the end of the conversion, is fed back to shut down the analog building blocks to further save power. The capacitors $C_{CT1,2}$ and $C_{PT1,2}$ are used to compensate for process variation. In our proposed implementation, the corresponding temperature-modulated output PW t_{PW} is

$$t_{PW}(T) = (t_{PT}(T_0) - t_{CT}(T_0)) + (k_P - k_C)(T - T_0) \quad (6)$$

where $t_{PT}(T_0)$ and $t_{CT}(T_0)$ are the corresponding PTAT and CTAT delay offsets at reference temperature T_0 , and k_C and k_P are the corresponding first-order temperature proportional constants. It can be observed that the first term, which represents the signal offset, is reduced after the XOR operation, whereas the second term, which is temperature dependent, is increased (as k_C is negative). In that case, the effective signal-to-noise ratio at the sensor output is improved. The implementation of the PTAT and CTAT delay lines is discussed as follows.

A. CTAT Delay Line

The proposed CTAT delay generator, as indicated in Fig. 3, is a voltage-current (V - I) converter. If M_{CT1} is operating in the subthreshold region and with a large enough amplifier gain A_{CT} , the corresponding voltage V_R across resistor R_{CT} can be approximated as

$$V_R = V_- \approx V_+ - \frac{V_T}{A_{CT}} \quad (7)$$

Fig. 3. Proposed temperature sensor architecture.

where V_T is the transistor threshold voltage. Equation (7) shows that the voltage across the resistor depends on the threshold voltage. As $V_B = V_+$ is generated by identical PMOS transistors and is almost temperature independent, the corresponding temperature dependence of V_R can be obtained after differentiating (7) with respect to temperature, i.e.,

$$\frac{\partial V_R}{\partial T} \approx -\frac{\alpha}{A_{CT}} \quad (8)$$

where α is process dependent but generally negative, as stated in [10]. From (8), V_R is PTAT. For a particular process, the temperature behavior of on-chip resistors is characterized by the first- and second-order temperature dependences. Assume that V_R and R_{CT} can be expressed as

$$R_{CT} = R_0(1 + k_{1R}T + k_{2R}T^2) \quad (9)$$

$$V_R = V_0(1 + k_{1V}T) \quad (10)$$

where k_{1R} and k_{2R} are the first- and second-order temperature coefficients of R_{CT} , respectively, and k_{1V} is the first-order temperature coefficient of V_R . The CTAT delay generated is

$$t_{CT} \propto g(T) = \frac{1}{I_{CT}} = \frac{R_{CT}}{V_R}. \quad (11)$$

From (9)–(11)

$$g(T) \approx \frac{R_0}{V_0}(1 + C_1T + C_2T^2) \quad (12)$$

where

$$C_1 = k_{1R} - k_{1V} \quad (13)$$

$$C_2 = k_{2R} - k_{1V}k_{1R} + k_{1V}^2. \quad (14)$$

Assume that the CTAT delay can be represented by (12) and that higher order terms can be neglected. As shown in (8), k_{1V} is positive. Considering (13), for the maximum temperature signal, $|C_1|$ should be maximized, meaning k_{1R} should be negative. On the other hand, for best linearity with respect to temperature, C_2 should be minimized. As the second and third terms of (14) are positive, k_{2R} should be chosen to be negative. This means that the choice of resistor is limited to the one that has negative first- and second-order temperature

dependences. In our process, only the P^+ polyresistor fits in the aforementioned requirements. The expected resultant nonlinearity is about 0.3°C .

B. PTAT Delay Line

The proposed PTAT delay generator is also shown in Fig. 3. As discussed in the previous section, the PTAT delay can be generated using a MOS transistor operating in the linear region. We outline the linearity optimization and power consumption reduction methodologies in the following discussion. The drain current for a MOSFET operating in the linear region can be expressed as

$$I_{PT} = I_{lin} = \mu C_{ox} \left(\frac{W}{L} \right) \left((V_{gs} - V_T)V_{ds} - \frac{V_{ds}^2}{2} \right). \quad (15)$$

For a small enough V_{ds} , we first express (4) as

$$t_{PT} \propto f(T) = \frac{1}{I_{PT}} \approx \frac{A}{T^{km}(B - \alpha T)} \quad (16)$$

where

$$A = \frac{T_0^{km}}{\mu_0 C_{ox} \left(\frac{W}{L} \right) V_{ds}} \quad (17)$$

$$B = V_{gs} - V_{T0} + \alpha T_0. \quad (18)$$

If we further differentiate (16) with respect to temperature, the first- and second-order derivatives are

$$f'(T) = \frac{A}{T^{km+1}(B - \alpha T)} (-kmB + \alpha(km + 1)T) \quad (19)$$

$$f''(T) = \frac{Akm(km + 1)}{T^{km+2}(B - \alpha T)} - \frac{2A\alpha km}{T^{km+1}(B - \alpha T)^2} + \frac{2A\alpha^2}{T^{km}(B - \alpha T)^3}. \quad (20)$$

In the following discussion, we assume that (4) can accurately be approximated by its Taylor-series expansion up to the second-order derivatives and that higher order derivatives can be neglected. It can be observed that (20) can be minimized to

Fig. 4. Maximum absolute error versus temperature over different processes from 0 °C to 100 °C after two-end-point calibration, with $V_{ds} = 0.1$ and $W/L = 0.5 \mu\text{m}/10 \mu\text{m}$.

improve the overall linearity with respect to temperature. If we equate (20) to zero

$$B = \alpha T \left[\left(1 + \frac{1}{km+1} \right) \pm \sqrt{\left(1 + \frac{1}{km+1} \right)^2 - \left(1 + \frac{2}{km} \right)} \right]. \quad (21)$$

From (18) and (21)

$$V_{gs} = V_{T0} + \alpha \left\{ \left[\left(1 + \frac{1}{km+1} \right) \pm \sqrt{\left(1 + \frac{1}{km+1} \right)^2 - \left(1 + \frac{2}{km} \right)} \right] T - T_0 \right\} \quad (22)$$

where

$$km \approx -1.58 \quad (23)$$

$$\alpha \approx \frac{-0.224}{T_0} \quad (24)$$

in our process. By substituting (23) and (24) into (19), it can be shown that $f'(T)$ is always positive, verifying that the delay generated from a MOS operating in the linear region is always PTAT. On the other hand, by substituting (23) and (24) into (22), V_{gs} can be deduced as

$$V_{gs} \approx 0.7 \text{ or } 1. \quad (25)$$

The result in (25) shows the approximated values where the optimal linearity in delay with respect to temperature is obtained. As km and α are process dependent, this value may not exist. Fig. 4 shows the simulation results of the maximum absolute error with different gate bias V_g from 0 °C to 100 °C after two-end-point calibration. It can be seen that the optimal gate bias value indeed exists for all the three processes. In that case, a high supply voltage is not required for our sensor implementation to achieve high linearity. On the other hand, note that when V_g is high enough, the error reduces again, verifying that a high supply voltage can also improve linearity, as stated in [7].

From (22), it can be observed that this optimal V_g is process dependent. Fig. 5 shows the effect of process variation over the value of the optimal gate bias voltage for our process. It can be seen that a required error of less than 0.5 °C can be achieved through different process corners by using a V_g from 0.74 to

Fig. 5. Maximum absolute error versus temperature over different process corners for the TSMC 0.18- μm 1P6M process from 0 °C to 100 °C after two-end-point calibration, with $V_{ds} = 0.1$ and $W/L = 0.5 \mu\text{m}/10 \mu\text{m}$.

Fig. 6. Die micrograph of the proposed temperature sensor, occupying an active area of 0.0324 mm².

Fig. 7. Measured temperature error versus temperature over four samples.

Fig. 8. Measured relative temperature error versus temperature with $\pm 10\%$ variation in supply voltage.

0.91 V (gray region). This means that our proposed scheme can tolerate the effect of process variation and achieve an error well below the accuracy requirement. In our design, the overdrive voltage of M_{PT0} is designed to be larger than the corresponding V_{ds} to ensure that the transistor is in the linear region in all operating conditions.

TABLE I
PERFORMANCE SUMMARY OF RECENTLY PUBLISHED CMOS TEMPERATURE SENSORS

Sensor	Resolution (°C)	Error (°C)	Area (mm ²)	Temp. Range (°C)	Power Consumption (μW)	Sampling Rate	Technology
[12]	0.002~0.01	± 0.1	4.5	-55~125	247.5	1~10	0.7μm
[7]	0.16	-0.7~+0.9	0.175	0~100	10	10k	0.35μm
[11]	0.0908~0.0927	-0.25~+0.35	0.6	0~90	36.7	2	0.35μm
[4]	N/A	-1.8~+2.2	N/A	0~100	10	N/A	N/A
[5]	0.909	± 1	0.2	27~47	0.9	10k	0.18μm
[6]	0.3	-1.6~+3	0.05	0~100	0.22	100	0.18μm
Proposed	0.3	-0.8~+1	0.0324	0~100	0.405	1k	0.18μm

C. Reduced Supply Sensitivity

The propagation delay, which is defined as the time required for charging/discharging a load capacitor C_L to 50% of the supply voltage V_{DD} through a current I , can be described as

$$t_p = \frac{C_L V_{DD}}{2I}. \quad (26)$$

From the preceding equation, it can be observed that the delay is proportional to the supply voltage V_{DD} , meaning that the induced temperature data are supply voltage dependent. In addition, notice that the delay can be insensitive to supply variation if the charging/discharging current I is also proportional to V_{DD} . As shown in Fig. 3, for both PTAT and CTAT generators, biasing the voltage across R_{CT} and V_{ds} of M_{PT0} to be proportional to V_{DD} yields to currents proportional to V_{DD} . To achieve this, a supply voltage sensor (also shown in Fig. 3) is implemented. It is simply a potential divider composed of identical PMOS transistors with the body and the source tied together. In that case, V_B (refer to Fig. 3) is proportional to V_{DD} but insensitive to temperature. As the voltage across R_{CT} and V_{ds} of M_{PT0} are clamped to V_B by the amplifiers, I_{CT} and I_{PT} are proportional to V_{DD} , generating supply-insensitive delays.

IV. MEASUREMENT RESULTS

The proposed temperature sensor is implemented using the TSMC 0.18-μm 1P6M process. Fig. 6 shows the die micrograph of the proposed temperature sensor. Extensive matching is exercised to minimize the adverse effect due to process variation and mismatch. The sensor only occupies an active area of 0.0324 mm². With a supply voltage of 1 V and an external clock signal, the sensor dissipates a measured power of 405 nW at room temperature at 1 ksample/s.

The measured temperature error from 0 °C to 100 °C after two-end-point calibration, as shown in Fig. 7, is within -0.8 °C to +1 °C over four measured samples. The effective resolution is measured to be around 0.3 °C. The variation in measured temperature with respect to supply voltage change of ±10% is plotted in Fig. 8. It is observed that a measured temperature error of ±0.8 °C is achieved at room temperature. Note that, when the supply voltage is increased, a positive temperature error is resulted, and vice versa. This is mainly due to the second-order V_{ds} term in (15). The preceding negative sign results in a current level less than the designed linear-to- V_{ds} relationship, and the corresponding delay generated is longer than expected. This results in an increase in PW, and hence increases the corresponding digital output.

V. CONCLUSION

A CMOS temperature sensor with power consumption of 405 nW based on the linear MOS operation has been presented. By utilizing a linear MOS for temperature sensing and a differential sensing architecture for offset reduction and effective signal increase, and implementing the previously discussed optimization method, a measured temperature error of -0.8 °C to +1 °C from 0 °C and 100 °C and a resolution of 0.3 °C can be achieved. The sensor occupies an active area of only 0.0324 mm². The performances of recently reported temperature sensors fabricated using CMOS processes are summarized in Table I. It can be concluded that our proposed sensor is suitable for integrated thermal management applications and passive RFID tag applications as it features ultralow-power consumption and improved compactness compared with previously reported sensors.

REFERENCES

- [1] C. Poirier, R. McGowen, C. Bostak, and S. Naffziger, "Power and temperature control on a 90 nm Itanium-family processor," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 229–237, Jan. 2006.
- [2] P. Ituero, J. L. Ayala, and M. L. Vallejo, "A nanowatt smart temperature sensor for dynamic thermal management," *IEEE Sensor J.*, vol. 8, no. 12, pp. 2036–2043, Dec. 2008.
- [3] C. K. Kim, B. S. Kong, C. G. Lee, and Y. H. Jun, "CMOS temperature sensor with ring oscillator for mobile DRAM self-refresh control," in *Proc. ISCAS*, May 2008, pp. 3084–3097.
- [4] K. Opasjumruskit, T. Thanthipwan, O. Sathusen, P. Sirinamarattana, P. Gadmanee, E. Pootarapan, N. Wongkomet, A. Thanachayanont, and M. Thamsirianunt, "Self-powered wireless temperature sensor exploit RFID technology," *IEEE Pervasive Comput.*, vol. 5, no. 1, pp. 54–61, Jan.–Mar. 2006.
- [5] S. Zhou and N. Wu, "A novel ultra low power temperature sensor for UHF RFID tag chip," in *Proc. ASSCC*, Nov. 2007, pp. 464–467.
- [6] Y. S. Lin, D. Sylvester, and D. Blaauw, "An ultra low power 1 V, 220 nW temperature sensor for passive wireless applications," in *Proc. CICC*, Sep. 2008, pp. 507–510.
- [7] P. Chen, C. C. Chen, C. C. Tsai, and W. F. Lu, "A time-to-digital-converter-based CMOS smart temperature sensor," *IEEE J. Solid-State Circuits*, vol. 40, no. 8, pp. 1642–1648, Aug. 2005.
- [8] M. K. Law and A. Bermak, "A time domain differential CMOS temperature sensor with reduced supply sensitivity," in *Proc. ISCAS*, May 2008, pp. 2126–2129.
- [9] T. A. Demassa and Z. Ciccone, *Digital Integrated Circuits*. New York: Wiley, 1996.
- [10] I. M. Filanovsky and A. Allam, "Mutual compensation of mobility and threshold voltage temperature effects with applications in CMOS circuits," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 7, pp. 876–884, Jul. 2001.
- [11] P. Chen, K. M. Wang, Y. H. Peng, Y. S. Wang, and C. C. Chen, "A time-domain SAR smart temperature sensor with -0.25 ~ +0.35 °C inaccuracy for on-chip monitoring," in *Proc. Eur. Solid-State Circuits Conf.*, Sep. 2008, pp. 70–73.
- [12] M. A. P. Pertigas, K. A. A. Makinwa, and J. H. Huijsing, "A CMOS smart temperature sensor with a 3σ inaccuracy of ±0.1 °C from -55 °C to 125 °C," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2805–2815, Dec. 2005.