A sub-1V 78-nA bandgap reference with curvature compensation

Ziyang Luo, Yan Lu, Mo Huang, Junmin Jiang, Sai-Weng Sin, Seng-Pan U, Rui P. Martins

State Key Laboratory of Analog and Mixed-Signal VLSI and FST-ECE, University of Macau, Macao, China
State Key Laboratory of Analog and Mixed-Signal VLSI, University of Macau, Macao, China
School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China
ECE Department, The Hong Kong University of Science and Technology, Hong Kong

ARTICLE INFO

Keywords:
Bandgap reference
CMOS analog integrated circuits
Internet-of-things (IoT)
Ultra-low power
Curvature compensation
Temperature coefficient

ABSTRACT

This paper presents a low power bandgap reference (BGR) with 573 mV average output voltage and 0.95-V minimum supply voltage. Here, the resistor ratio of two types of resistors that have opposite temperature coefficient (TC) realizes a second-order curvature compensation. In addition, to reduce the supply voltage we adopt a voltage divider, and we also employ proportional to absolute temperature (PTAT) voltage generators to reduce the resistor values for low power applications. The proposed BGR fabricated in 65-nm CMOS achieves an average TC of 57 ppm/°C from 0 to 100 °C with 8 samples. The power dissipation at room temperature is 78 nA and the power supply ripple rejection (PSRR) is −67 dB at DC with 1.2 V supply.

1. Introduction

BGR has been a critical building block for data converters and power management ICs. It provides an accurate reference voltage for voltage sensing or supply voltage regulation. Meanwhile, it needs to be ultra-low power for the massively deployed internet-of-things (IoT) nodes. In conventional BGRs, resistors are used to generate a PTAT current and consequentially a PTAT voltage to cancel out the complementary to absolute temperature (CTAT) portion of the diode voltage. Therefore, large silicon area is required by the large resistors for low power in conventional BGRs. Thus, there is a compromise between power consumption and chip area [1–4]. For example, the BGR in [4] consumes 490 nA, while the resistors occupy about 80% of the layout area.

Several new circuit topologies [5–11] have been proposed to advance the tradeoff between power consumption and chip area. A resistor-less BGR was proposed in [6], where a nA-level current reference circuit and PTAT voltage generators based on overdrive voltage differences were used. However, it may not be suitable for high-precision application for its relatively larger TC. Another BGR in [7] used the switch capacitor circuits to eliminate resistors and used charge pumps to reduce the supply voltage, but the generated PTAT voltage was sensitive to parasitic capacitors and charge pump output variation, which degraded the TC. Some other resistor-less CMOS voltage references [8–10] achieved nW-level power consumption employing the threshold voltage of a MOSFET, but suffered from process variations and lost the absolute accuracy. For instance, a coefficient of variation (defined as the standard deviation over the mean value) of 7% along with the process was observed in [9]. A current mode BGR with curvature compensation was presented in [11] for better TC. However, the large resistance requirement for the current mode topology has limited the area reduction.

In this work, to overcome the above mentioned problems, we propose a BGR that employs the PTAT voltage generator based on overdrive voltage differences and uses relatively small resistors only for curvature compensation. This paper is organized as follows. Section 2 presents the design of the proposed BGR. Section 3 exhibits the experimental results and the related discussions. Finally, we draw the conclusions in Section 4.

2. Design of the proposed BGR

2.1. Architecture of the proposed BGR

Fig. 1 shows that the proposed BGR is composed by four parts: the start-up circuit, the core, the voltage divider, and the PTAT voltage generator.

Fig. 1: Architecture of the proposed BGR

The start-up circuit provides a current source to start up the core. The core part generates the second-order curvature compensated CTAT voltage V_{CTAT}. Then V_{CTAT} is divided into V_{CTAT}/2 by
is the elementary charge, and $V_{TH}$ is the threshold voltage of the MOSFET, and $\eta$ is the constant of the subthreshold slope factor.

As the bias current is small and the aspect ratio of MOSFET is relatively large, all the MOSFETs in Fig. 3 operate in the subthreshold region, where the generated PTAT voltage can be derived as:

$$V_{PTAT} = V_{OUT} - V_{IN} = V_{SG,M1} - V_{SG,M2} = \eta V_T \ln \left( \frac{K_2}{K_1} \right).$$  \(4\)

where $K_1$, $K_2$ correspond to the aspect ratios of $M_1$, $M_2$. Therefore, $V_{PTAT}$ is proportional to the absolute temperature with a first-order term only. Besides, the PTAT voltage generator is biased by $V_{BIAS}$, which is the output of the amplifier in the core.

2.3. Second-order curvature compensation

Fig. 4 presents the core circuit of the proposed BGR, in which the $V_{EB}$ is second-order compensated by the voltage across the p-diffusion resistor $R_3$. According to the simulation, the high-resistive poly resistor ($R_1$, $R_2$) has the negative TC as given in (5), while the p-diffusion resistor ($R_3$) has a positive TC as given in (6),

$$R_i(T) \approx R_i(T_0) \cdot \left[ 1 - K_i(T - T_0) \right],$$  \(5\)

$$R_3(T) \approx R_3(T_0) \cdot \left[ 1 + K_3(T - T_0) \right],$$  \(6\)

where $R_i(T_0)$ and $R_3(T_0)$ are the resistance at reference temperature $T_0$, $K_1$ and $K_3$ are TC constants. In the resistance-temperature simulation, $K_1$ is about $3.5 \times 10^{-4} /K$ and $K_3$ is about $1.5 \times 10^{-3} /K$. The BJT current in Fig. 4 is:

$$I = \frac{(V_{EB1} - V_{EB2})\cdot I_1}{R_1} = V_T \ln [N/R_1].$$  \(7\)

2.2. Principle of PTAT voltage generator

To save chip area, we adopt the circuit in Fig. 3 for the PTAT voltage generator [6] that provides a PTAT voltage without using resistor. When the drain-source current is small and the aspect ratio of MOSFET is relatively large, the MOSFET tends to operate in the subthreshold region. When the drain-source voltage $V_{DS}$ of a MOSFET is higher than 100 mV, the subthreshold drain current $I_D$ in the MOSFET is simplified as:

$$I_D = K_B \exp \left( \frac{V_{GS} - V_{TH}}{\eta V_T} \right),$$  \(2\)

$$I_S = \mu_C C_{OX} (\eta - 1) V_T^2,$$  \(3\)

where $K$ is the aspect ratio of the MOSFET, $\mu$ is the carrier mobility, $C_{OX}$ is the gate-oxide capacitance per unit area, $V_T = K_B T/q$ is the thermal voltage, $K_B$ is the Boltzmann constant, $T$ is the absolute temperature, $q$ is the charge of an electron. The subthreshold drain current is expressed as:

$$I_S = \frac{1}{2} \mu_C C_{OX} q \exp \left( \frac{V_{TH} - V_T}{\eta V_T} \right).$$  \(4\)

The subthreshold drain current is very small and can be neglected.

The voltage divider for realizing half of the bandgap voltage $V_{BG}/2$, with the advantage of lower supply voltage. Also, the voltage divider reduces the required number of PTAT stages, thus saves silicon area. Finally, the cascaded PTAT voltage generators achieve the first-order compensation for the $V_{CTAT}/2$ and outputs the reference voltage $V_{REF}$ (about $V_{BG}/2$).

Figs. 2(a) and (b) show the conceptual diagrams of the conventional BGR and the second-order compensated BGR, respectively. In the conventional BGR [12], $V_{EB}$ generated by the bipolar junction transistor (BJT) is first-order compensated by the PTAT voltage $\Delta V_{EB}$ across a resistor, as Fig. 2(a) shows. $V_{EB}$ is expressed as [13]:

$$V_{EB}(T) = V_{BG} - [V_{EB}(T_r) - V_{BG}] \frac{T}{T_r} - (\eta - m)V_T \ln \left( \frac{T}{T_r} \right)$$  \(1\)

where $V_{BG}$ is the bandgap voltage (about 1.2 V) of silicon extrapolated at 0 K, $T_r$ is the reference temperature, $\eta$ is a temperature constant depending on the process, $m$ is 1 when the bias current of BJT is a PTAT current. Whereas, due to the $V_{EB}$ nonlinear term $(\eta - m)V_T \ln(T/T_r)$ in (1), a first-order compensation is not adequate for high accuracy. The nonlinear term of $V_{EB}$ should be reduced so as to achieve a better TC. In this design, as shown in Fig. 2(b), we add a resistor voltage on top of $V_{EB}$ to realize the second-order curvature compensation. Since the impact of the $V_{EB}$ nonlinear term decreases, the TC gets better.

Fig. 1. Architecture of the proposed BGR.

Fig. 2. Voltage versus temperature curves with (a) first-order compensation and (b) second-order compensation.

Fig. 3. Schematic of the overdrive voltage based PTAT voltage generator.

Fig. 4. $V_{EB}$ with second-order curvature compensation.
The voltage across \( R_3 \) is derived as:

\[
V_{R3} = I_{D3} R_3 \approx V_T \ln \left[ N \frac{v_{GS1}}{v_{TH}} \left[ 1 + \frac{v_{GS1} - v_{TH}}{v_{TH}} \right] \frac{v_{GS1} - v_{TH}}{v_{TH}} \right] \\
\approx \frac{1}{q} \ln \left[ N \frac{v_{GS1}}{v_{TH}} \right] \left[ 1 - K_1 I_D - K_2 I_D^2 + (K_3 + K_4) I_D^3 \right]
\] (8)

As illustrated in Fig. 2(b), we use the second-order term of \( V_{R3} \) to compensate \( V_{EB1} \). The resistor \( R_3 \) can be trimmed for better second-order curvature compensation. The second-order compensated voltage, \( V_{CTAT} \), is connected to voltage divider for dividing the input into half.

Then the divided voltage, \( V_{CTAT/2} \), is first-order compensated by the PTAT voltage generators.

### 2.4. Voltage divider

Fig. 5(a) [6] shows a voltage divider adopted for reducing the supply voltage and power. Thus, the reference output voltage is half of the bandgap voltage (about 0.6 V). The two NMOS \( M_1 \) and \( M_2 \), fabricated in p-well (with deep n-well) with same aspect ratio, have their bulks connected to their sources to obtain the same threshold voltage. Assume the gate and substrate leakage currents are small enough to be neglected, the drain-source currents across \( M_1 \) and \( M_2 \) can be regarded as equal, and be expressed as:

\[
I_{DS,M1} = K I_D \exp \left( \frac{v_{GS1} - v_{TH}}{qV_T} \right) \left( 1 - \exp \left( \frac{v_{GS1} - v_{TH}}{qV_T} \right) \right)
\]

\[
I_{DS,M2} = K I_D \exp \left( \frac{v_{GS2} - v_{TH}}{qV_T} \right) \left( 1 - \exp \left( \frac{v_{GS2} - v_{TH}}{qV_T} \right) \right)
\] (9)

Nevertheless, since \( V_{GS2} \) and \( V_{DD} - V_{GS2} \) are not equal to each other and change with temperature and supply voltage, \( V_{GS1} \) is not exactly the same as \( V_{GS2} \). Therefore, the \( V_{OUT} \) is not exactly equal to half \( V_{IN} \). The structure in Fig. 5(a) degrades the TC performance.

We adopt the circuit in Fig. 5(b) to improve the precision of the voltage divider, by inserting a buffer between the gate and the drain of \( M_1 \). In this structure, the lengths and the widths of \( M_1 \) and \( M_2 \) are the same, with both of the \( V_{IN} \) equal to their \( V_{GS} \). According to the simulation, \( V_{GS1} \) and \( V_{GS2} \) change approximately from 350 mV to 250 mV from 0 °C to 100 °C. Thus, the relationship between \( M_1 \) and \( M_2 \) can be simplified as:

\[
K I_D \exp \left( \frac{V_{GS1} - V_{TH}}{qV_T} \right) = K I_D \exp \left( \frac{V_{GS2} - V_{TH}}{qV_T} \right).
\] (10)

\( V_{GS1} = V_{GS2} = V_{CTAT/2} \) can be deduced. The relationship is independent of temperature and supply voltage.

### 2.5. BGR output voltage

Fig. 6 shows the whole BGR, consisting of four parts as mentioned above. \( K_{M24}/K_{M1} \) is included in the logarithmic function as Eq. (4). Hence, \( K_{M24}/K_{M1} \) must be a large value (about one thousand in this work) to generate a sufficient PTAT voltage to compensate \( V_{CTAT/2} \). To avoid large size ratio between \( M_1 \) and \( M_2 \), three PTAT voltage generators are cascaded to realize the first-order compensation. The reference voltage is:

\[
V_{REF} = \frac{1}{2} \left( V_{EB1} + V_{EB2} \right) + V_{PTAT}
\]

\[
= \frac{V_{DD}}{2} + \frac{R_3}{3R_3} V_T \ln(N) + \eta V_T \ln \left( \frac{K_{M24}K_{M3}}{K_{M2}K_{M1}} \right) \approx \frac{V_{DD}}{2}.
\] (11)

The startup circuit detects the positive input of the amplifier \( A_1 \). Before the circuit starts up, \( M_{S23} \) will pull the \( V_{HAS} \) down to ground, which will generate sufficient current to turn on the BGR properly.

### 2.6. PSRR analysis

In Fig. 6, \( V_{HAS} \), which is the output of the amplifier \( A_1 \), biases the PTAT voltage generators. As mentioned in [14], the supply ripple at the output of the amplifier which has a PMOS differential-to-single-ended current mirror, has roughly the same amplitude at the supply. The supply ripple can be treated as a common-mode signal at the gate and the source of \( M_{P6} \), \( M_{P9} \), and \( M_{P10} \). Therefore, the structure reduces the feedthrough of the power ripple from the supply.

Fig. 7(a) presents the self-biased amplifier \( A_1 \) in the core circuit. \( M_{P1} \), \( M_{P2} \), and \( M_{P5} \) are high threshold voltage devices, while \( M_{P3} \) and \( M_{P4} \) are low threshold voltage devices. Since the ripple coming from the biasing branch affects \( V_{X1} \), \( V_{X2} \) and \( V_{Y1} \), \( V_{Y2} \) in a common-mode

![Fig. 6. Full schematic of the proposed BGR.](image-url)
form, then, the effect of the ripple from the biasing branch can be ignored. Fig. 7(b) shows the simplified schematic of A1. $R_a$ represents the series channel resistance of $M_{P2}$ and $M_{P4}$. $R_b$ represents the channel resistance of $M_{N4}$ and $M_{N5}$. Besides, $1/g_{m}$ is the diode-connected PMOS resistance, which is negligible when compared with $R_b$. Fig. 7(c) shows the small signal PSRR model of the amplifier. The current source $i_{Rb}$ models the current mirror. The small signal of the output of the amplifier, namely $V_{bias}$, is expressed in (12):

$$
V_{bias} = V_{dd} \left( \frac{R_a}{R_a + R_b} \right) + i_{Rb} \left( R_{dd} \right) \\
\approx V_{dd} \left( \frac{R_a}{R_a + R_b} \right) + \frac{V_{dd}}{R_b} \left( R_{dd} \right) = V_{dd}.
$$

(12)

Because the gate and the source voltages of $M_{P9}$, $M_{P9}$ and $M_{P10}$ change simultaneously in the same direction and with approximately the same amplitude, the feedthrough of the supply ripple is cancelled.

3. Experiment results

Since the current in the output branch is relatively small (about 6 nA) to drive the measurement equipment which has about 1 MΩ input impedance, an output buffer with independent supply voltage is essential for the measurement, as shown in Fig. 6. The buffer is carefully designed with large size for reducing the input offset voltage, and the TC of the buffer’s offset voltage only contributes a small portion to the TC of the BGR. All the measured and simulated data of the BGR include the effect of this buffer. Fig. 8 shows the micrograph of the proposed BGR. The overall BGR chip area is 0.036 mm² (excluding the buffer for testing). We measured eight samples from the same batch without trimming. At room temperature, the total current is about 78 nA. Besides, the average reference voltage of the eight chips is 573 mV, with a coefficient of variation of 0.82%.

Fig. 9 shows the sub-BGR voltage as a function of temperature ranging from 0 °C to 100 °C, with 0.95 V supply. The average TC of the eight chips is 57 ppm/°C and the worst case is 120 ppm/°C, which can be further improved by trimming. Trimming can be realized with two steps: nonlinear trimming and linear trimming. The first step is for the nonlinear trimming, which adjusts the resistor $R_3$ shown in Eq.(11) to reduce the nonlinear term of $V_{REF}$. The second step is for the linear trimming, by adjusting the size of the differential pairs to reduce the linear term.

To observe the effectiveness of the curvature compensation, we mathematically add a coefficient $\eta V_T \ln(\Delta K)$ to the $V_{REF}$ curve to imitate the linear trimming, where $\Delta K$ is the change of the size ratios of differential pairs in the PTAT voltage generator. In the differential pairs of the PTAT voltage generator, the finger of MOSFET that has the same aspect ratio is different, which generates the PTAT voltage. The normal size ratio of the differential pair $M_5$ and $M_6$ in Fig. 6 is $2/16$ (1/8). In the mathematical linear trimming, we assume the finger number of $M_5$ is adjustable from 2 to 4 with a step size of 1 (2 trimming bits) while the finger number of the $M_6$ is adjustable from 8 to 64 with a step size of 1 (6 trimming bits). Therefore, the size ratio of $M_5$ and $M_6$ could be changed from $4/8$ (1/2) to $2/64$ (1/32), which means the trimmable PTAT voltage generator could compensate the maximum voltage variation of about ± 18 mV from 0 °C to 100 °C. The mathematical
trimming is achieved by two temperatures points (the maximum and the minimum), and the trimmed TC in Fig. 10 shows that the average TC is 6.6 ppm/°C, which means that the proposed second-order compensation method indeed contributes to the reduction of the nonlinear term of $V_{REF}$.

Fig. 11 shows the TC at three different supply voltages (0.95, 1.2, and 1.8 V) from 0 °C to 100 °C. The BGR can operate properly when the supply voltage is larger than 0.95 V. Fig. 12(a) plots the measured PSRR of the BGR circuit with output buffer. To reduce the noise and improve the PSRR, a 12 pF on-chip decoupling capacitor is added to the node of the output of the PTAT voltage generator while considering a reasonable chip area and start-up time. With 1.2 V supply, the measured PSRR is about $-67$ dB at DC. The decoupling capacitor at the output of the BGR causes the falling of PSRR at 7 kHz. Meanwhile, the gate of the PMOS current mirror $V_{BIAS}$ in Fig. 7(a) is a high-resistive node, and the parasitic capacitor between the supply and $V_{BIAS}$ causes a PSRR rise at about 200 kHz, of which the effect coincides with the post-layout PSRR simulation shown in Fig. 12(b). Fig. 13 is the post-layout simulation result of noise with/without a 12 pF decoupling capacitor.

Table 1 exhibits the performances of the proposed BGR compared with that of the state-of-the-art works. The average TC of 57 ppm/°C from 0 °C to 100 °C without trimming is still a competitive performance, with such low power consumption.

and small chip area. Based on the mathematical calculation (with linear trimming), the TC can be only 6.6 ppm/°C. When compared with the CMOS voltage reference based on $V_{TH}$, the coefficient of variation (0.82%) in the proposed BGR is smaller.

Since the number of samples is limited, to further illustrate the process variation of $V_{REF}$ and the curvature compensation method, 300 runs of post-layout Monte Carlo simulation results of $V_{REF}$ and TC are provided in Fig. 14. The Monte Carlo simulation results shown in Fig. 14 include the process variation and mismatch of the whole circuit, which composes of the BGR circuit and the output buffer. Thus, the buffer offset voltage is also included in the results. Compared to the voltage reference based on $V_{TH}$, the coefficient of variation of this circuit has a better result of 1.78%. Besides, an untrimmed average TC of 17 ppm/°C with a standard deviation $\sigma$ of 11.5 ppm/°C is observed, which further verifies the effectiveness of curvature compensation.
For reducing the resistor values, three stages of the PTAT voltage generator are cascaded to realize the first-order compensation with less chip area. When compared with other curvature-compensated BGRs, the proposed BGR manages to use small values of resistors and realize the curvature compensation with low power. The measured average TC is 57 ppm/°C, which can be further improved by trimming the resistor and the aspect ratios of the differential pairs in the PTAT generators.

Acknowledgments

This work was supported by the Macao Science & Technology Development Fund (FDCT) under Grant No. 122/2014/A3 and the Research Committee of University of Macau.

References