

# A 0.5–1.7 V Efficient and PVT-Invariant Constant Subthreshold $g_m$ Reference Circuit in CMOS

Martin Lee<sup>1</sup> and Kambiz Moez<sup>1</sup>, *Senior Member, IEEE*

**Abstract**—This paper presents a low-power constant subthreshold transconductance reference circuit that produces a constant transconductance by subtracting the output currents of two independent transconductance references. By taking the difference between the output currents of the two independent transconductance references, process, voltage and temperature variations are reduced by minimizing the effects of channel length modulation and drain-induced barrier lowering without relying on feedback to regulate the drain voltage. The proposed reference, fabricated in TSMC’s 130 nm process and testing with off-chip temperature-insensitive resistors, can provide a constant transconductance over a temperature range of  $-30$  to  $120^\circ\text{C}$ , and a supply voltage range of 0.5 to 1.7 V. The experimental result of the proposed transconductance reference shows an average variation of  $\pm 0.8\%$  over temperature for the entire operating voltage range and an average variation of  $\pm 1.97\%$  over voltage for the entire operating temperature range. Through the subtraction, the proposed circuit also shows less variation across process corners compared to the conventional constant transconductance reference. The proposed constant transconductance reference exhibits the highest power efficiency (transconductance over power consumption) amongst the reported constant transconductance references. At 0.5 V, the proposed transconductance reference produces a transconductance of  $21.46 \mu\text{S}$ , while consuming  $1.95 \mu\text{W}$ .

**Index Terms**—Constant  $g_m$ , reference circuit, beta-multiplier, process, voltage, temperature (PVT) variation, subthreshold, low voltage, low power.

## I. INTRODUCTION

THE number of connected Internet of Things (IoT) devices is forecasted to exceed 30 billion by 2025 [1]. The rapid growth in the number of IoT devices, which in many cases will be battery operated, presents a considerable challenge to achieve sustainable and reliable operation because the finite energy capacity of batteries forces a trade-off between functionality and battery life [2]. Other methods such as energy harvesting can be used to supplement or replace batteries, however, these methods also have constraints on the available energy [3]. Ultra-low power operation is essential for IoT devices to reduce the effect of this trade-off [4]. One of the

Manuscript received 23 February 2022; revised 4 May 2022 and 12 June 2022; accepted 24 June 2022. This work was supported in part by the Natural Sciences and Engineering Research (NSERC) and Alberta Innovates. This article was recommended by Associate Editor R. Rieger. (*Corresponding author: Martin Lee.*)

The authors are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada (e-mail: mkleee@ualberta.ca; kambiz@ualberta.ca).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCSI.2022.3188707>.

Digital Object Identifier 10.1109/TCSI.2022.3188707

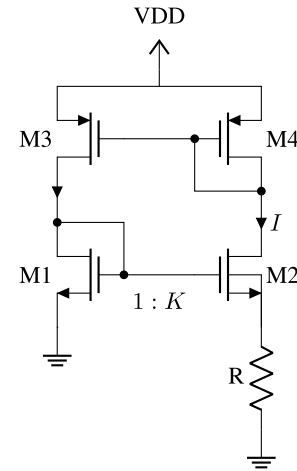


Fig. 1. Conventional beta-multiplier.

methods to achieve low power operation is to operate circuits in the subthreshold region for CMOS integrated circuits [5]. This can lead to substantially lower power consumption due to the increase in power efficiency since the ratio of the transistor’s transconductance to its drain-source current ( $g_m/I_{ds}$ ) is maximized in the subthreshold region. The energy saving in the subthreshold operation comes at the expense of increased sensitivity of parameters to process, voltage and temperature (PVT) variations. The increased sensitivity is due to the exponential dependency of the drain-source current of transistors to PVT parameters, which will translate to variations in the device transconductance ( $g_m$ ) that often determine the main circuit performance parameters. For instance, the gain, frequency response, input matching of a low-noise amplifier are strongly dependent on the transistor’s  $g_m$ . Therefore, to produce reliable circuits operating in the subthreshold region, it is critical to keep the  $g_m$  constant so that the operation of the biased circuit remains consistent for all operating conditions.

Constant  $g_m$  reference circuits are widely used to produce invariant  $g_m$  over the desired temperature and supply voltage range. As shown in Fig. 1, the conventional circuit of maintaining a constant  $g_m$  using MOSFETs, known as the beta-multiplier, was proposed in [6]. By creating an approximately proportional-to-temperature (PTAT) current, a constant  $g_m$  can be produced [7]. However, channel length modulation (CLM) and drain-induced barrier lowering (DIBL) in this circuit causes  $g_m$  to vary significantly with temperature and voltage because of differences in  $V_{ds}$ . Cascode current mirrors can be used to reduce the difference in  $V_{ds}$ , however, at the cost

of a higher minimum operating voltage [7]–[16]. Therefore, some reference circuits utilize negative feedback to regulate the drain voltages to be equal, for instance with the use of op-amps [7], [8], [10], [12]–[14], [16]–[22]. This matches the drain voltages of all four transistors and forces equal currents in both branches of the beta-multiplier. While this method can effectively reduce variations over VDD, it will not eliminate variations over temperature. The difference in source voltage between M1 and M2, due to the voltage drop across the resistor, will cause the  $V_{ds}$  of M1 and M2 to vary differently over temperature. Reference [19] proposed to minimize the effects of CLM and DIBL by making  $V_{ds}$  equal at the expense of increased reliance on a precise resistor and a separate current reference. Reference [23] uses a long series of self-cascode transistors to set the gate voltage separately. While this eliminates the differences in source voltage, it introduces the effects of CLM and DIBL at the gate of M1 and M2. Furthermore, both these works still require the use of cascode or op-amps to minimize CLM and DIBL. Reference [24] also makes  $V_{ds}$  equal by defining the gate voltage externally by using switched capacitors. This develops a reliance on a clock signal. Reference [14] proposes to have the core device within the  $g_m$  biasing circuit, and to use differential signals to minimize the second-order effects. This limits its potential applications. Reference [25] describes an approach to reducing the sensitivity to VDD by taking the difference between the outputs of two variants of the beta-multiplier for a voltage reference. However, it uses long-channel transistors and it does not account for the temperature and voltage variations of the subthreshold slope factor, which will result in increased sensitivity to VDD and temperature for a constant  $g_m$  reference.

In this paper, we propose a circuit that addresses the second-order effects of the MOSFET in the subthreshold region by subtracting the output currents of two independent and closely matched references. By closely matching the two references, the independent references will have a similar response across variations in process corners, voltage and temperature introduced by CLM, DIBL and the subthreshold slope factor, thereby, allowing the second-order effects to cancel out by subtraction. As a result, the sensitivity of the output current of the proposed  $g_m$  reference circuit to subthreshold temperature, voltage and process variations are significantly reduced across a wide range of operating voltages and temperatures in sub-micron technologies. This paper is organized as follows; to provide the background and motivation behind this paper, Section II will introduce PVT variations of subthreshold MOSFETs, and Section III will formally detail the conventional MOSFET  $g_m$  reference. Section IV will describe the proposed constant  $g_m$  reference and its design, while the measured results of the circuit will be shown in Section V. Section VI will summarize the results and conclude this paper.

## II. PVT VARIATIONS OF SUBTHRESHOLD MOSFETs

The sensitivity of subthreshold MOSFETs to PVT variations can be explained by its I-V characteristic in

the subthreshold region [5]:

$$I_{ds} = I_0 \exp\left(\frac{V_{gs} - V_{th} + \eta V_{ds}}{nV_T}\right) \times \left(1 - \exp\left(\frac{-V_{ds}}{V_T}\right)\right)(1 + \lambda V_{ds}), \quad (1)$$

where the gate-source voltage ( $V_{gs}$ ) is less than the threshold voltage ( $V_{th}$ ) of the MOSFET,  $V_{ds}$  is the drain-source voltage,  $\eta$  is the DIBL coefficient,  $V_T$  is the thermal voltage,  $n$  is the subthreshold slope factor,  $\lambda$  is the channel length modulation coefficient and the characteristic current,  $I_0$ , is defined as

$$I_0 = \mu C_{ox} \frac{W}{L} (n-1) V_T^2, \quad (2)$$

where  $\mu$  is the charge mobility,  $C_{ox}$  is the gate oxide capacitance per unit area and  $W/L$  is the dimensions of the MOSFET. From (1), process variations affect  $W/L$ ,  $\mu$  and  $V_{th}$ , variations in voltage affect  $n$ ,  $V_{gs}$  and  $V_{ds}$ , and temperature variations affect  $n$ ,  $\mu$ ,  $V_{th}$  and  $V_T = k_B T/q$ . While the temperature variations of  $n$  will be detailed in Section IV, the impact of temperature on  $\mu$  and  $V_{th}$  can be explicitly shown through the following equations [26]:

$$V_{th}(T) = V_{th}(T_0) - A(T - T_0), \quad (3)$$

$$\mu(T) = \mu(T_0) \left(\frac{T}{T_0}\right)^\alpha, \quad (4)$$

where  $T_0$  is the reference temperature,  $\alpha$  is the mobility temperature parameter and  $A$  is the threshold voltage temperature coefficient. As a result of the exponential relationship between the current and the overdrive voltage ( $V_{gs} - V_{th}$ ), PVT variations have a significant impact on the subthreshold current. Subsequently, these variations to current will translate to variations in  $g_m$ , since the subthreshold  $g_m$  is calculated as

$$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} = \frac{I_{ds}}{nV_T}. \quad (5)$$

## III. CONVENTIONAL MOSFET $g_m$ REFERENCE

If  $I_{ds}$  is made to vary proportionally to  $nT$ , a constant  $g_m$  that is insensitive to PVT variations can be created. This is conventionally done using the aforementioned beta-multiplier shown in Fig. 1 [6]. By operating all MOSFETs in the subthreshold region to save power, the ideal current produced by the beta-multiplier can be derived by finding the difference in  $V_{gs}$  between M1 and M2. This can be shown by ignoring the effects of CLM and DIBL in (1). Furthermore, for  $V_{ds} \geq 4V_T$  MOSFETs operate in subthreshold saturation and the third term in (1) becomes negligible. Therefore, (1) simplifies to (6), and  $V_{gs}$  can be calculated as a function of the subthreshold current as shown in (7):

$$I_{ds} = I_0 \exp\left(\frac{V_{gs} - V_{th}}{nV_T}\right), \quad (6)$$

$$V_{gs} = nV_T \ln\left(\frac{I_{ds}}{I_0}\right) + V_{th}. \quad (7)$$

Equations (6) and (7) are derived by also assuming the body of M2 is connected to its source, as shown in Fig. 1, which allows the current to be proportional to  $n$ , and to avoid the body effect of M2. This gives the ideal current of the beta-multiplier

shown below that results in a  $g_m$  that is only a function of  $K$  and  $R$ :

$$I = \frac{nV_T \ln(K)}{R}. \quad (8)$$

The effects of CLM and DIBL appear due to the mismatch in the  $V_{ds}$  of the current mirrors. We can mathematically show the impact of these second-order effects on the beta-multiplier. Taking into account the differences in  $V_{ds}$  of M1, M2, M3 and M4, (8) can be rewritten as

$$I = \frac{n_n V_T}{R} \left[ \ln(K) + \ln\left(\frac{1 + \lambda V_{ds2}}{1 + \lambda V_{ds1}}\right) + \ln\left(\frac{1 + \lambda |V_{ds3}|}{1 + \lambda |V_{ds4}|}\right) \right] + \frac{\eta}{R} \left[ (V_{ds2} - V_{ds1}) + \frac{n_n}{n_p} (|V_{ds3}| - |V_{ds4}|) \right]. \quad (9)$$

The suffix appended to  $V_{ds}$  in (9) represents the  $V_{ds}$  of the respective transistors in Fig. 1. The subthreshold slope factor is also differentiated for a PMOS and NMOS transistor as  $n_p$  and  $n_n$ , respectively. The  $V_{ds}$  mismatch between the NMOS and PMOS transistors introduces four additional terms in the current. The temperature dependence of  $V_{ds}$  for the four transistors can be determined by deriving the  $V_{ds}$  of the diode-connected MOSFETs, M1 and M4. There are several assumptions made to simplify the derivation for the temperature dependency of  $V_{ds}$ . Second-order effects are neglected, and the beta-multiplier is assumed to operate ideally as shown in (8). While  $n$  does increase with temperature, it is assumed to be temperature-independent in this derivation. Since M1 is diode-connected, the  $V_{ds}$  of M1 is equal to its  $V_{gs}$  as shown in (7). The temperature dependencies shown in (3), (4) and  $V_T$  along with the current in (8) can be substituted into (7). All the temperature-independent terms in (2) and (6) are grouped into a term  $C$  aside from  $K$ . This results in the following equation:

$$\begin{aligned} V_{ds1} &= nV_T \ln\left(\frac{C \ln(K)}{T^{\alpha+1}}\right) + V_{th}(T_0) - A(T - T_0), \\ &= V_{th}(T_0) + AT_0 + \left(\frac{nk_B \ln(C \ln(K))}{q} - A\right)T \\ &\quad - \frac{nk_B(\alpha + 1)}{q} T \ln(T). \end{aligned} \quad (10)$$

If the current through both branches of the beta-multiplier are equal,  $V_{ds4}$  will have the same relationship with temperature shown in (10), aside from the differences between PMOS and NMOS transistors. It should be noted that in (10)  $A$  is a positive coefficient with a typical value of 2 mV/°C and  $\alpha$  is a negative parameter with a typical value of -1.5. [26].  $C$  is found to be positive but is less than 1; therefore,  $nk_B \ln(C)/q$  is negative. Equation (10) shows that as temperature increases, the third term will cause the  $V_{ds}$  of M1 and M4 to fall. Although the fourth term of (10) increases with temperature since  $\alpha$  is negative, its coefficient is much smaller than  $A$ ; therefore,  $V_{ds1}$  and  $V_{ds4}$  will decrease with temperature. Since the total voltage drop in the left branch of the beta-multiplier is equal to  $V_{ds1}$  and  $V_{ds3}$ , as  $V_{ds1}$  decreases with temperature,  $V_{ds3}$  must increase with temperature, provided that VDD remains constant. For the right branch of the beta-multiplier, the total voltage drop is the summation of  $V_{ds2}$ ,  $V_{ds4}$  and the voltage drop across  $R$  ( $V_R$ ). While  $V_R$ ,  $nk_B T \ln(K)/q$ ,

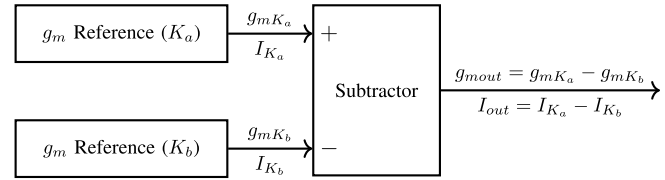


Fig. 2. Block diagram of proposed circuit.

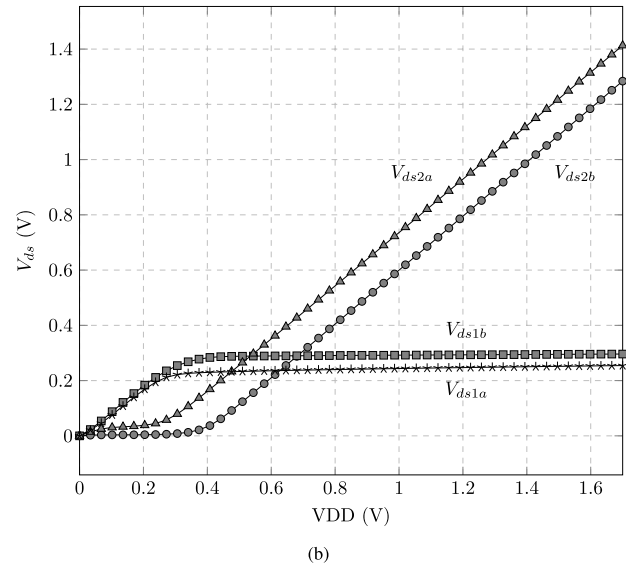
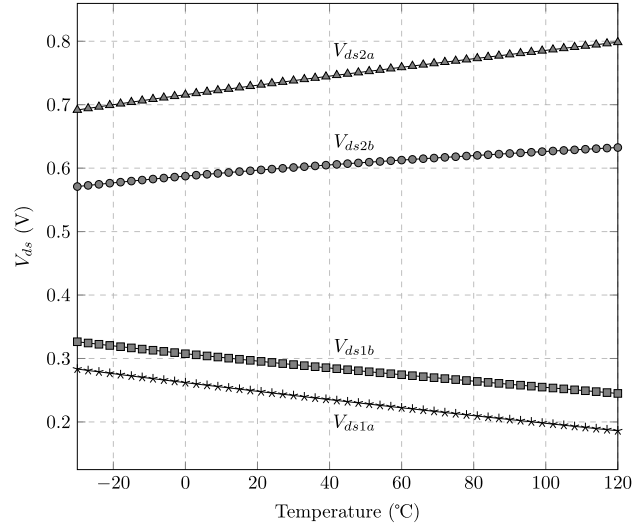


Fig. 3. Variation of  $V_{ds}$  over (a) temperature (b) supply voltage for  $K_a = 18$  and  $K_b = 2$ .

increases with temperature, the decrease in  $V_{ds4}$  with temperature is larger because the coefficient of the third term in (10) is larger than  $nk_B \ln(K)/q$ , therefore,  $V_{ds2}$  must increase with temperature. Overall, as the temperature increases, the  $V_{ds}$  of M1 and M4 will decrease while the  $V_{ds}$  of M2 and M3 will increase.

Variations of  $V_{ds}$  across VDD can be determined by finding the output impedance of the MOSFETs in the beta-multiplier. The output impedance of the diode-connected MOSFETs,

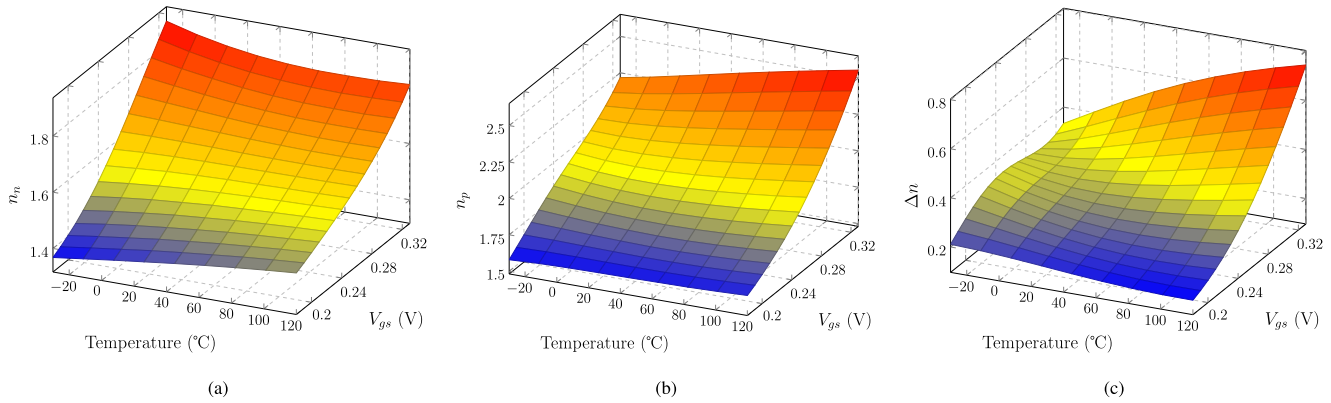


Fig. 4. Surface plot of subthreshold slope factor over temperature and  $V_{gs}$  (a) NMOS ( $n_n$ ) (b) PMOS ( $n_p$ ) (c) difference between PMOS and NMOS ( $\Delta n$ ).

M1 and M4, are equal to  $1/g_m$ , while the output impedance of M3 is  $r_o$ . M2 is a source-degenerated MOSFET and it has the equivalent output resistance of  $r_o + g_m r_o R + R$ . M2 and M3 are high impedance, therefore, as VDD increases the majority of VDD will fall across M2 and M3. Since M1 and M4 are diode-connected, the current determines the  $V_{ds}$  of these transistors and this limits changes to  $V_{ds}$  to M2 and M3. Therefore, according to (9), the additional four terms will exhibit a temperature and voltage dependence causing variations in the output  $g_m$ .

#### IV. PROPOSED CONSTANT $g_m$ REFERENCE

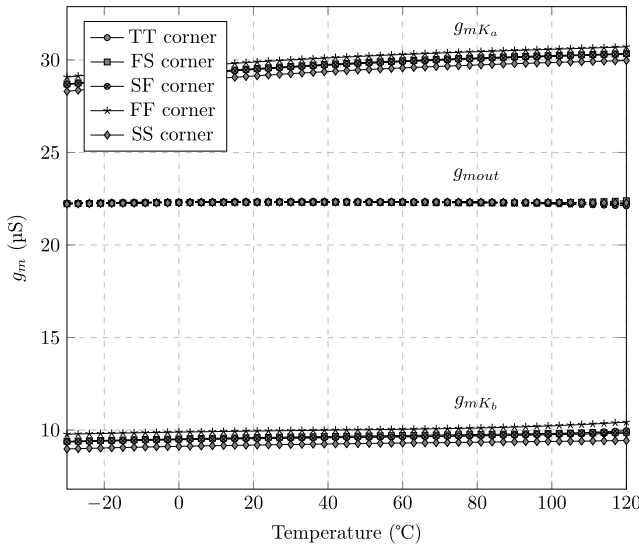
The principle of operation of the proposed circuit can be explained with the help of the illustration in Fig. 2. The circuit is developed to replace existing conventional reference circuits. The proposed circuit utilizes two independent  $g_m$  references with the only difference in the  $K$  value (size ratio of two bottom transistor in the beta-multiplier), where  $K_a > K_b$ , but are otherwise identical. By using two closely matched references, we can ensure the two references share a similar response to PVT variations. This is exploited to minimize the effects of CLM and DIBL by taking the difference in current between the two beta-multipliers to obtain the output  $g_m$ . To demonstrate this, using (9), the output current of the proposed reference can be shown as

$$\begin{aligned}
 I_{out} = & \frac{n_n V_T}{R} \left[ \ln\left(\frac{K_a}{K_b}\right) + \ln\left(\frac{1 + \lambda V_{ds2a}}{1 + \lambda V_{ds2b}} \frac{1 + \lambda V_{ds1b}}{1 + \lambda V_{ds1a}}\right) \right. \\
 & \left. + \ln\left(\frac{1 + \lambda |V_{ds3a}|}{1 + \lambda |V_{ds3b}|} \frac{1 + \lambda |V_{ds4a}|}{1 + \lambda |V_{ds4b}|}\right) \right] \\
 & + \frac{\eta}{R} [(V_{ds2a} - V_{ds1a} - V_{ds2b} + V_{ds1b}) \\
 & + \frac{n_n}{n_p} (|V_{ds3a}| - |V_{ds4a}| - |V_{ds3b}| + |V_{ds4b}|)]. \quad (11)
 \end{aligned}$$

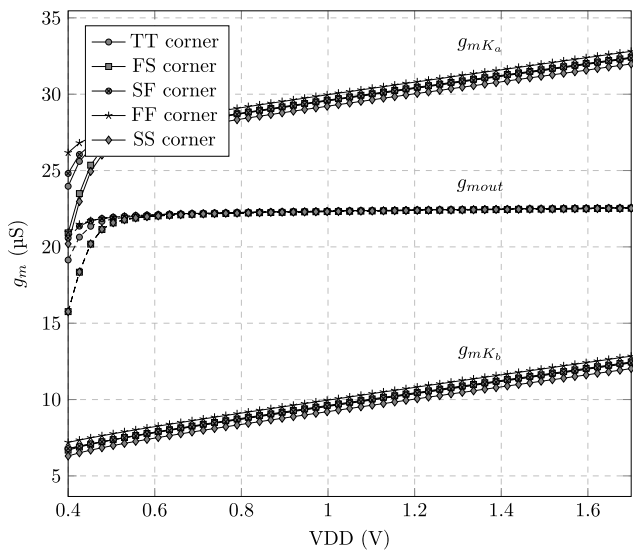
To achieve a PVT invariant  $g_m$  reference, the logarithmic terms that contain  $V_{ds}$  must remain constant, while the terms that show up due to DIBL must vary proportionally to temperature. Through (10), it can be found that the  $V_{ds}$  of the same transistor across two beta-multipliers vary similarly over PVT variations. In (10), larger  $K$  values increases the third term, which causes  $V_{ds}$  of M1 and M4 to increase with  $K$ , and the  $V_{ds}$  of M2 and M3 to decrease with  $K$ . This is confirmed

by the simulation results in Fig. 3, where the number in the subscript indicates the corresponding MOSFET in Fig. 1 and the subscript  $a$  and  $b$  represents the beta-multiplier designed with a  $K$  value of  $K_a$  and  $K_b$ , respectively. Since the effect of PVT variation on  $K$  can be neglected, changing the  $K$  value creates a shift in the temperature and voltage plots and does not significantly affect the relationship of  $V_{ds}$  to VDD and temperature in (10). The simulation results shown in Fig. 3 further confirms that the  $V_{ds}$  of the same transistor across two beta-multiplier, designed with different  $K$  values, vary with a similar relationship across VDD and temperature. The variations of  $V_{ds}$  produce two significant results. First, the second and third logarithmic terms of (11), which contain  $V_{ds}$ , will remain relatively constant over PVT variations due to the similar relationship of  $V_{ds a}$  and  $V_{ds b}$  over VDD and temperature. This will minimize the effects of CLM on the proposed reference. Second, since modifying the  $K$  value shifts the  $V_{ds}$  curves over temperature and voltage, the difference in  $V_{ds}$  of the same transistor across the beta-multiplier will exhibit lower second-order temperature variations. The summation of the  $V_{ds}$  terms within the round brackets of (11) produces an approximately proportional relationship to temperature and due to the subtraction, second-order temperature variations are reduced, while the combined  $V_{ds}$  terms are relatively constant across voltage variations. Therefore, PVT variations due to DIBL are also minimized.

As briefly mentioned in Section II, the subthreshold slope factor depends on temperature and voltage, therefore, requires further consideration.  $n$  can be determined from the subthreshold slope of the MOSFET and is found to increase with temperature and voltage as shown in Fig. 4a and 4b. Furthermore,  $n$  for a PMOS and NMOS transistor shows slightly different dependencies to temperature and voltage as shown in Fig. 4c. The PVT dependency of the subthreshold slope factor and the differences in its dependency between a PMOS and NMOS transistor will introduce PVT variations through the  $n_n$  and  $n_n/n_p$  factors in (11). These slight temperature and voltage variations limit the precision of the output  $g_m$ . To address the effects of the  $n_n$  factor, the dependency on the NMOS subthreshold slope factor is minimized only when the output of both independent references are proportional to  $n_n$ . This requires the two independent to share the same design and to



(a)



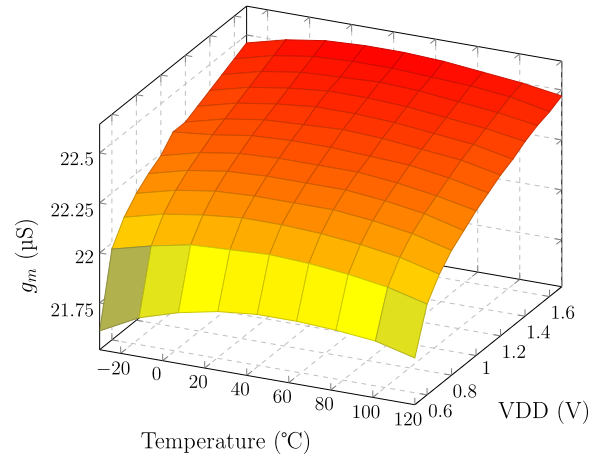
(b)

 Fig. 5.  $g_m$  over process corners vs. (a) temperature and (b) VDD.

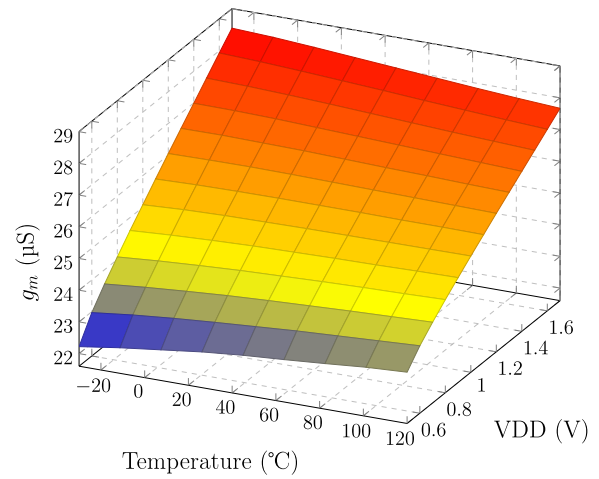
be matched by enforcing the only difference between the two references to be the value of  $K$ . This allows the output  $g_m$  to be insensitive to  $n_n$  when the output is taken through an NMOS transistor, which minimizes the dependency to  $n_n$  as evident in (5). As for the  $n_n/n_p$  factor, by making the current in the two branches of the beta-multiplier equal, this will match the  $V_{ds}$  of the PMOS transistors, therefore, isolating the PVT variations of the  $n_n/n_p$  factor and removing the contribution of  $V_{ds3a/4a}$  and  $V_{ds3b/4b}$ . The resulting  $g_m$  that is produced by the above analysis reduces (11) to:

$$g_{mout} = \frac{1}{R} \left[ \ln\left(\frac{K_a}{K_b}\right) + \ln\left(\frac{1 + \lambda V_{ds2a}}{1 + \lambda V_{ds2b}} \frac{1 + \lambda V_{ds1b}}{1 + \lambda V_{ds1a}}\right) \right] + \frac{\eta}{n_n V_T R} (V_{ds2a} - V_{ds1a} - V_{ds2b} + V_{ds1b}). \quad (12)$$

Since the logarithmic terms are insensitive to PVT variations, while the last term varies proportionally to temperature, this produces the desired PVT-invariant  $g_m$ . Fig. 5 and 6



(a)



(b)

 Fig. 6. Surface plot of simulated  $g_m$  over temperature and VDD (a) proposed (b) conventional.

demonstrate the reduced sensitivity of the proposed reference to PVT variations that show the simulated  $g_m$  across temperature, voltage and process corners. As shown in Fig. 5, the divergence of  $V_{ds}$  of the PMOS and NMOS transistor pairs in (9) causes  $g_m K_b$  and  $g_m K_a$  to increase across both temperature and voltage. By taking the difference between the  $g_m$  in the two beta-multiplier,  $g_{mout}$ , the  $g_m$  of Mbias, has a flatter response over both temperature and voltage. Furthermore, inter-die process variations are reduced by taking advantage of the similar response of the beta-multipliers to process corners, a result of the matched design of the beta-multipliers, indicated by the tighter grouping of  $g_m$  curves.

The improvements in the simulated  $g_m$  are further observed in Fig. 6, which shows a surface plot of  $g_m$  of the proposed reference and a single beta-multiplier in the form shown in Fig. 9, with a  $K$  value of 9 to produce a similar output, across temperature and VDD. The proposed reference shows a significantly more stable  $g_m$  output, with a lot of the variations occurring from 0.5 to 0.6 V since this region is close to the minimum operating voltage. Whereas the single beta-multiplier shows significant variations across VDD and temperature, with the only exception from 1 to 1.2 V where

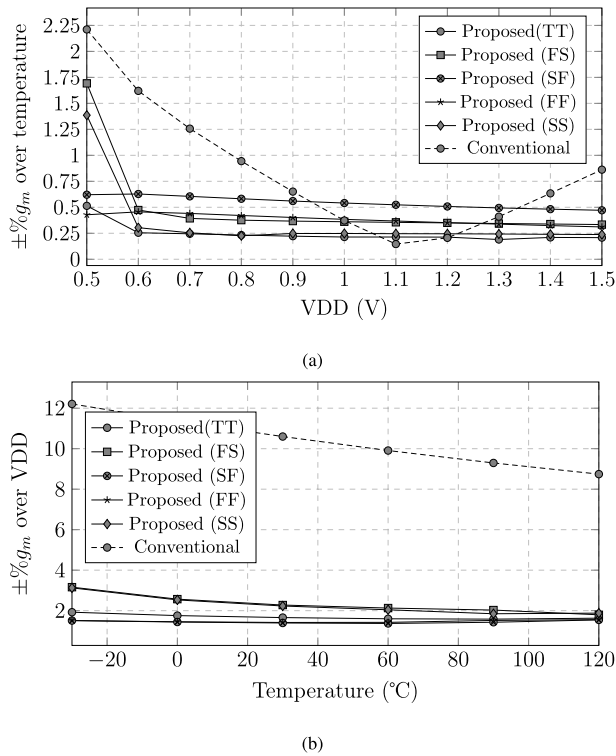


Fig. 7.  $g_m$  dependence over (a) temperature for a given VDD (b) VDD for a given temperature.

the second-order effects of the MOSFET cancel out across temperature by coincidence.

The variation of  $g_m$  in the surface plots can be quantified through the maximum positive/negative variation ( $\pm\%$ ) of  $g_m$  with respect to the mean. Fig. 7a and 7b plots the  $\pm\%$  error over temperature at indicated VDDs and VDD at indicated temperatures, respectively. In addition,  $\pm\%g_m$  is also evaluated for the proposed circuit across the five process corners. The  $g_m$  of the proposed reference shows a minimum variation of  $\pm 0.19\%$  over temperature at 1.3 V VDD and a minimum variation of  $\pm 1.41\%$  over VDD at 30  $^{\circ}\text{C}$ . The only occurrence where  $\pm\%g_m$  over temperature exceeds  $\pm 1\%$  is at 0.5 V VDD for FS and SS process corners due to the higher minimum operating voltage at these corners. Likewise, the increase in operating voltage at these corners also increases the  $\pm\%g_m$  over VDD. Without accounting for the increase in operating voltage, the simulation of the proposed circuit shows a maximum variation of  $\pm 0.62\%$  over temperature and  $\pm 1.92\%$  over VDD. While the conventional circuit was able to achieve a minimum variation of  $\pm 0.15\%$  over temperature at a supply voltage of 1.1 V, due to the second-order effects of the MOSFET cancelling out by coincidence as mentioned, it shows significant degradation over the entire VDD range, with a maximum variation of  $\pm 2.2\%$  over temperature at a supply voltage of 0.5 V. Variations over VDD is significantly worse with a minimum variation of  $\pm 8.74\%$  over VDD. These simulations showcase the potential improvements of the proposed reference over the conventional reference.

The above simulations and analysis are obtained by shorting the body and source of M2a and M2b. This requires a

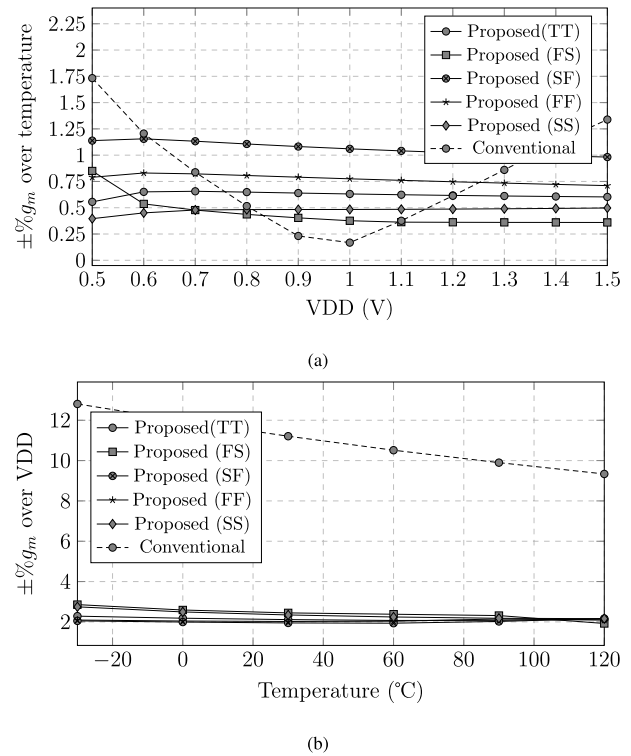


Fig. 8.  $g_m$  dependence over (a) temperature for a given VDD (b) VDD for a given temperature when the body of M2a/b is connected to ground.

deep n-well transistor that may not be feasible in a standard CMOS process. Fig. 8 demonstrates the PVT variations of the proposed and conventional circuits if they are implemented in a standard CMOS process, where the body of all PMOS and NMOS transistors are connected to VDD and ground, respectively. In this configuration, while the proposed circuit does exhibit higher PVT variations due to the body effect and increased sensitivity to  $n$ , the results are still promising.

#### A. Circuit Design

As described above, reducing the PVT variations of  $n$  partially requires the current in both branches of the beta-multiplier to be equal. This should be enforced by the PMOS current mirror of the beta-multiplier, however, a simple current mirror cannot maintain equal current in the two branches because of the low output impedance of short-channel MOSFETs. Using a cascode configuration or an op-amp to equalize the currents in both branches is not desirable due to a higher minimum required voltage and lower power efficiency. Therefore, these configurations are excluded to prevent additional power consumption in the design. The self-cascode transistor, otherwise called a composite transistor, where two transistors are connected in series with a shared gate connection can be used in this situation. The use of the self-cascode transistor was first demonstrated in [27] to work above the threshold voltage, however, it has been shown to work in the subthreshold region as well [28]–[30]. By sizing the top (bottom) transistor to be  $m$  times wider than the bottom (top) transistor for NMOS (PMOS) transistors, a high

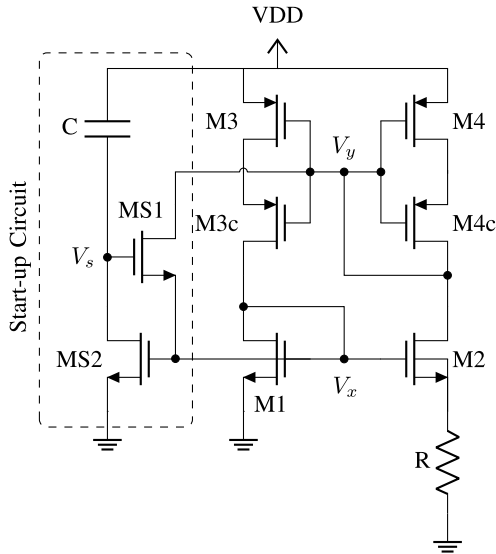


Fig. 9. Beta-multiplier with self-cascode PMOS current mirror and start-up circuit.

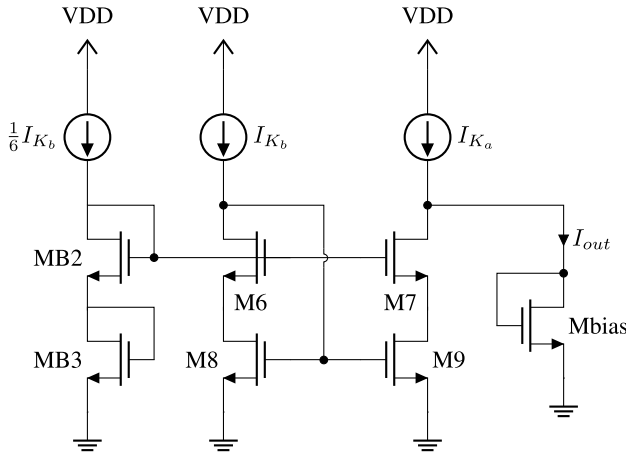


Fig. 10. PTAT current subtractor.

output impedance can be obtained, while enabling low voltage operation. The larger the  $m$  value, the higher the output impedance that can be obtained. This configuration is used for the top PMOS current mirror to improve current tracking in the beta-multiplier, as shown in Fig. 9.

To subtract the outputs of the two beta-multipliers, their currents are first copied out using the self-cascode PMOS current mirror shown in Fig. 9. The output currents of the two beta-multipliers are shown as  $I_{K_a}$  and  $I_{K_b}$  in Fig. 10. An additional current mirror, shown in Fig. 10 as M6 to M9, is used to mirror  $I_{K_b}$  into the same branch as  $I_{K_a}$  to take the difference in the current [31]. This allows one of the beta-multiplier to source the current and the other beta-multiplier to sink the current. The difference in current between the two beta-multipliers is passed through a transistor, Mbias, that can bias the desired device in a similar fashion to conventional biasing methods.

The current mirror for the subtraction operation is implemented using a low voltage cascode current mirror due to

TABLE I  
SIZE OF TRANSISTORS IN FIG. 11

Component	Size
M1a, M1b	2.5/0.24 $\mu\text{m}$
M2a	45/0.24 $\mu\text{m}$
M2b	5/0.24 $\mu\text{m}$
M3a-M5a, M3b-M5b	8/0.24 $\mu\text{m}$
M3ca-M5ca, M3cb-M5cb	160/0.24 $\mu\text{m}$
M6-M9	3/0.24 $\mu\text{m}$
MB1	1.2/0.24 $\mu\text{m}$
MB2	2.4/0.3 $\mu\text{m}$
MB3	3.6/0.3 $\mu\text{m}$
MB1c	24/0.24 $\mu\text{m}$
MS1a, MS2a, MS1b, MS2b	1.2/0.3 $\mu\text{m}$
Mbias	5/0.24 $\mu\text{m}$

the low voltage requirement and the high output impedance associated with the cascode design. The bias voltage for the top transistors in the low voltage cascode is provided by using a fraction of the current,  $I_{K_b}$ , by sizing the PMOS transistors to be approximately 1/6 of the original. A series of two diode-connected transistors, MB2 and MB3, is used to set the bias voltage. The resulting current subtractor is shown in Fig. 10.

A start-up circuit is necessary to ensure that the beta-multiplier does not remain in the zero-current operation point and operates at the desired operation point as the circuit is powered on [8]. The start-up circuit proposed in [9] is implemented in this design and is composed of two transistors and a capacitor, as shown in Fig. 9. The capacitor in the start-up circuit acts as a short circuit as the supply voltage ramps up, this causes  $V_s$  to rise with the supply voltage. The transistor connected between the gates of the PMOS and NMOS mirror turns on as a result of the applied gate bias. Current flows between  $V_x$  and  $V_y$ , pulling  $V_x$  and  $V_y$  up and down respectively. This causes the transistors to turn on, allowing the beta-multiplier to operate at the desired operating point. At steady-state, the capacitor isolates the beta-multiplier from the start-up circuit to limit power draw and to prevent the start-up circuit from affecting the PVT variations of the beta-multipliers.

The complete proposed constant  $g_m$  reference circuit is shown in Fig. 11. The proposed  $g_m$  reference is a standalone circuit and does not require external current references, clock signals and op-amps. The output current is proportional to  $\ln(K_a/K_b)$ , therefore, the selection of  $K_a/K_b$  will determine the output current. A small  $K_a/K_b$  will reduce the overall power consumption of the circuit at the expense of power efficiency. The opposite is true if  $K_a/K_b$  is large. To balance the total power consumption with power efficiency,  $K_a = 18$  and  $K_b = 2$  were chosen for the final design. It should be noted that while these  $K$  values were selected for fabrication, the proposed design can minimize PVT variations across various  $K_a/K_b$  and  $R$  values.

The sizes of the transistors in Fig. 11 are listed in Table. I. Since this design requires the two beta-multipliers to be





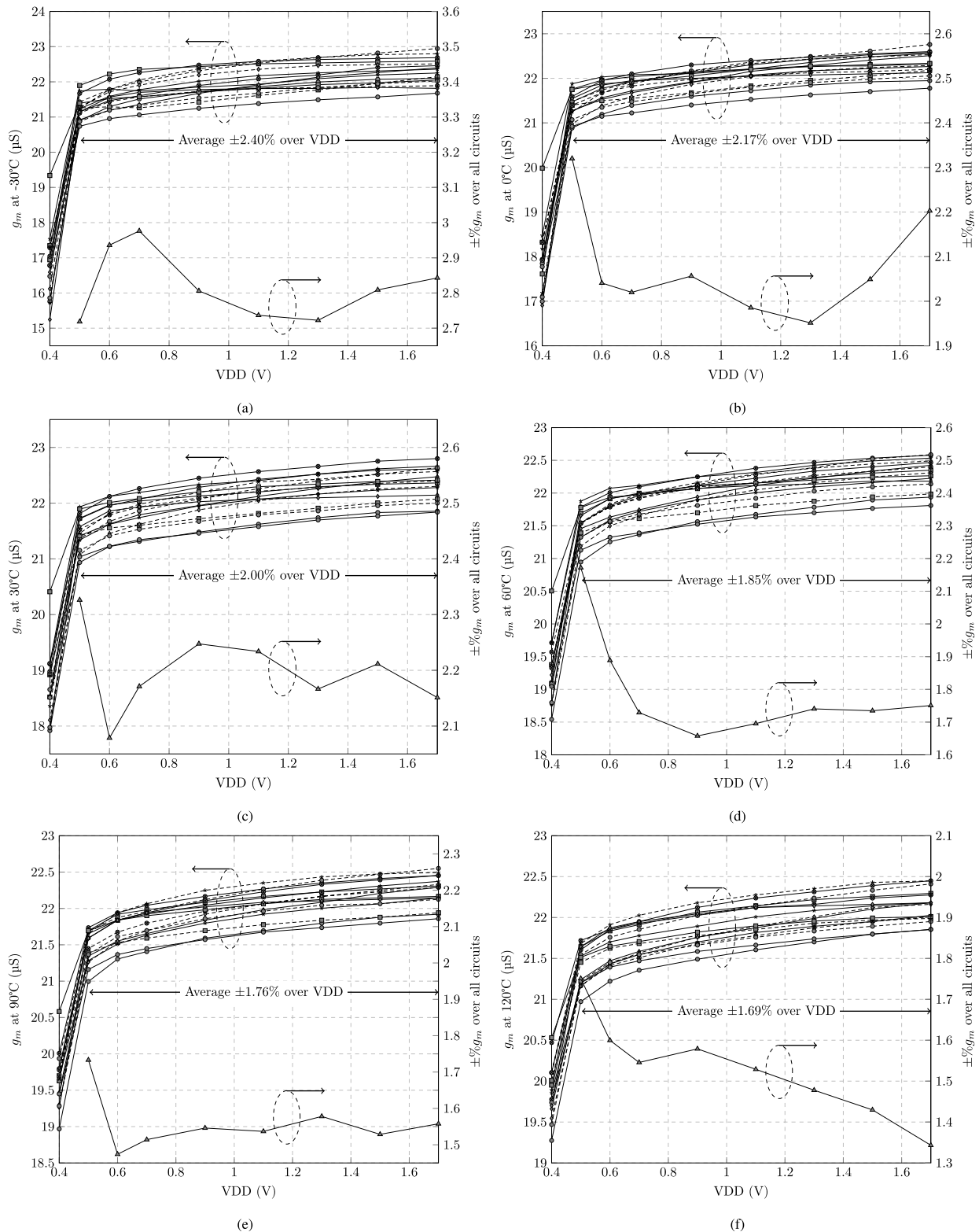


Fig. 13. Measured  $g_m$  plotted against VDD for all circuits with average  $\pm\%g_m$  over VDD and the  $\pm\%g_m$  variation across the circuits shown at (a)  $-30^\circ\text{C}$  (b)  $0^\circ\text{C}$  (c)  $30^\circ\text{C}$  (d)  $60^\circ\text{C}$  (e)  $90^\circ\text{C}$  (f)  $120^\circ\text{C}$ .

variation over VDD evaluated at six temperatures,  $g_m$  shows smaller variations over VDD at higher temperatures. This is in agreement with the simulation result shown in Fig. 7b. As for the variation found between the fabricated references,

a maximum and minimum variation of  $\pm 2.98\%$  and  $\pm 1.34\%$  are found. The measurements verify the stability of the output  $g_m$  found in simulation across temperature, VDD and process. As mentioned in Section IV, most of the variation over VDD

TABLE II  
COMPARISON TO OTHER WORKS

Ref.	Process	Sim./Meas.	Temp. Range (°C)	Voltage Range (V)	$\pm\%g_m$ over VDD	Nom. $g_m$	$\pm\%g_m$ over Temp.	Power	Area (mm <sup>2</sup> )	PSRR (dB)	$\frac{g_m}{\text{power}}$
This Work	130 nm	Meas.	-30 to 120	0.5 to 1.7	2 (1.41 <sup>c</sup> )	21.46 $\mu\text{S}^a$	0.87 <sup>a</sup>	1.95 $\mu\text{W}^a$	0.028	-103 <sup>d</sup>	11.01 <sup>a</sup>
						22.12 $\mu\text{S}^b$	0.77 <sup>b</sup>	4.85 $\mu\text{W}^b$			4.56 <sup>b</sup>
[10]	180 nm	Sim.	20 to 80	1.2 to 1.8	1	-	1.5	1.275 mW	-	-	-
[11]	180 nm	Meas.	25 to 100	1.8	-	1.41 $\mu\text{S}^g$	1.4	1.44 $\mu\text{W}$	0.012	-106	0.98
[12]	0.35 $\mu\text{m}$	Meas.	20 to 80	2.7 to 3.3	1.3	284 $\mu\text{S}^g$	1.1	19 mW	0.64	-	0.015
[13]	65 nm	Sim.	-20 to 110	0.4 to 1.2	6.15 <sup>e</sup>	10.5 mS <sup>g</sup>	2.6 <sup>f</sup>	-	-	-17	-
[14]	0.5 $\mu\text{m}$	Sim.	-30 to 110	1.1 to 1.7	-	69.2 nS <sup>g</sup>	0.76 <sup>g</sup>	110.6 nW	-	-	0.63
[19]	65 nm	Sim.	-60 to 130	2.5	-	5.002 mS	0.22	-	-	-	-
[20]	180 nm	Sim.	-25 to 125	1.0 to 2.0	-	83.5 $\mu\text{S}$	1.02	486 $\mu\text{W}$	0.264	-	0.17
[22]	180 nm	Sim.	-40 to 85	1.5 to 2.5	0.55 <sup>g</sup>	165 $\mu\text{S}$	3.55	149.4 $\mu\text{W}$	-	-	1.1
[32]	180 nm	Meas.	25 to 100	1.8	-	210 $\mu\text{S}^g$	3.4 <sup>f</sup>	136 $\mu\text{W}$	0.023	-	1.54
[33]	150 nm	Sim.	-40 to 125	1.52 to 1.68	0.8	860 pS <sup>g</sup>	0.63	9 nW	-	-	0.096

<sup>a</sup> Operating at 0.5 V and 30 °C temperature.

<sup>b</sup> Operating at 1.1 V and 30 °C temperature.

<sup>c</sup> Calculated for voltage range from 0.6 to 1.5 V.

<sup>d</sup> Simulated result.

<sup>e</sup> Maximum  $\pm\%g_m$  vs. VDD and process for a voltage range from 0.54 to 0.66 V.

<sup>f</sup> Maximum  $\pm\%g_m$  vs. temperature and process.

<sup>g</sup> Extrapolated from given results.

occurs below 0.6 V because the minimum operating voltage is experimentally determined to be approximately 0.5 V before the output  $g_m$  settles. The minimum operating voltage of the proposed circuit is determined by the required diode-connected voltage of M1a/M1b and to place M3a/M3b in the subthreshold saturation region in Fig. 11. The variation across the 16 circuits is found to be higher than the variation seen across process corners because of the introduction of intra-die variations from process mismatch. While the proposed circuit reduces the variation across process corners by design, process mismatch is only accounted for in the layout of the circuit. Further improvements in the layout can reduce this variation. As a result of mismatch, as well as errors introduced in testing such as variations in temperature, current leakage through the LPF filter at the output and slight differences in  $R$  over temperature, both the variation of  $g_m$  over temperature and VDD, most notably temperature, shows an increase in comparison to the results found in simulation. The passive components for testing were carefully selected, however, there were still small differences in the resistance over temperature. The distribution of the output  $g_m$  of the proposed reference over temperature and voltage is shown in Fig. 14a. Using the same analysis performed in simulation, the  $\pm\%$  error is calculated over temperature and VDD and is reported in Fig. 14b and 14c for the fabricated circuits. A nominal  $g_m$  of 21.93  $\mu\text{S}$  is measured over 768 data points with a standard deviation of 0.38  $\mu\text{S}$ . An average variation of  $\pm 0.8\%$  with a standard deviation of  $\pm 0.23\%$  over temperature and  $\pm 1.97\%$  with a standard deviation of  $\pm 0.56\%$  over VDD is found.

Table II compares this proposed  $g_m$  reference to other prior works. While some works on  $g_m$  references have prioritized the lowest power consumption, in Table II, the efficiency of  $g_m$  reference circuits was evaluated using a  $g_m/\text{power}$  metric, because of the focus on  $g_m$  efficiency in the subthreshold region for MOSFETs. This metric for transistors is taken and applied to the total power consumed by reference circuits to evaluate their efficiency. The results of the proposed circuit in Table II are shown when operating nominally at 0.5 and 1.1 V, and the average  $\pm\%g_m$  and power consumption across 16 circuits are reported here. Supply variations of the proposed circuit show higher  $\pm\%g_m$  variations compared to other works. However, the proposed circuit was evaluated over a larger voltage range than most works, and most of the variation is attributed to operating from 0.5 to 0.6 V. If the voltage range is redefined from 0.6 to 1.7 V, there is an average  $g_m$  variation of  $\pm 1.41\%$ . The proposed  $g_m$  reference shows the least variations over temperature against other prior works with measurement results, while again maintaining a larger temperature range than most works. Comparison over simulations, also shows that our proposed reference has the least variation over temperature. A simulated dc power supply rejection ripple (PSRR) for the proposed circuit was found to be  $-103$  dB. Most notable is the efficiency in generating the output  $g_m$  compared to other works at both operating voltages; the output  $g_m$  produce for the given power is higher than other reported  $g_m$  references. The high efficiency of the proposed reference circuit is helped by operating the proposed reference in the subthreshold region, the low operating voltages and by using passive resistors rather than the active resistors and

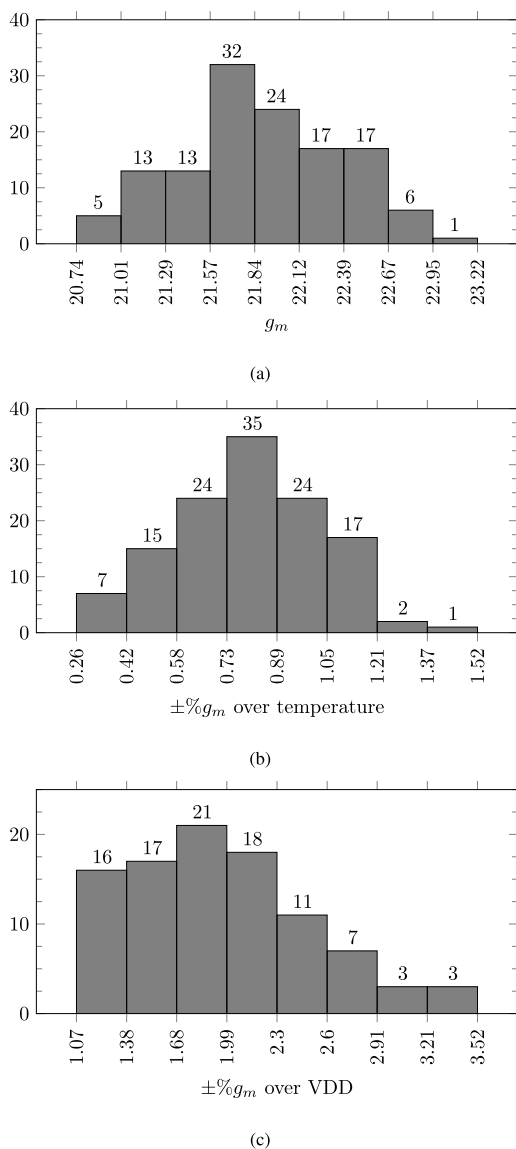


Fig. 14. Distribution of measured results (a)  $g_m$  (b)  $\pm\%g_m$  vs. temp (c)  $\pm\%g_m$  vs. VDD.

trimming circuits that other works have explored to reduce the PVT variations of on-chip resistors [9]–[12], [22]–[25], [32]–[34].

## VI. CONCLUSION

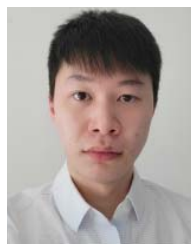
Constant subthreshold  $g_m$  reference circuits show increased susceptibility to PVT variations because of second-order effects, such as CLM and DIBL. In this paper, the effects of CLM and DIBL are discussed. The proposed constant  $g_m$  reference circuit addresses these issues by subtracting two independent references to cancel out the effects of CLM and DIBL without using long-channel transistors. By closely matching the references, the PVT variations of  $n$  are also minimized. The resultant  $g_m$  reference circuit was able to produce a stable  $g_m$  over a wide voltage and temperature range of 0.5 to 1.7 V and  $-30$  to  $120^\circ\text{C}$ , respectively. The proposed circuit demonstrates the smallest variation over temperature

and the highest efficiency compared to other works that have been fabricated and tested. While the proposed reference can produce an insensitive  $g_m$  through Mbias, it does not guarantee that the device biased by Mbias will also have an insensitive  $g_m$  due to differences in  $V_{ds}$  with the driven device. In this case, Mbias can be replaced with a cascode transistor, or other methods need to be found to control the  $V_{ds}$  of the driven device. Further work can be done to eliminate the use of external resistors by exploring methods to replace them with integrated resistors realized by active resistor circuits or work to further minimize the effects of mismatch.

## REFERENCES

- [1] (Nov. 2020). *Internet of Things (IoT) and Non-IoT Active Device Connections Worldwide From 2010 to 2025 (in Billions)*. Accessed: Feb. 22, 2022. [Online]. Available: <https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/>
- [2] S. Chen, H. Xu, D. Liu, B. Hu, and H. Wang, "A vision of IoT: Applications, challenges, and opportunities with China perspective," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 349–359, Aug. 2014.
- [3] U. Muncuk, K. Alemdar, J. D. Sarode, and K. R. Chowdhury, "Multi-band ambient RF energy harvesting circuit design for enabling batteryless sensors and IoT," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2700–2714, Aug. 2018.
- [4] M. Silva, D. Cerdeira, S. Pinto, and T. Gomes, "Operating systems for Internet of Things low-end devices: Analysis and benchmarking," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10375–10383, Dec. 2019.
- [5] N. Reynders and W. Dehaene, *Ultra-Low-Voltage Design of Energy-Efficient Digital Circuits*. Cham, Switzerland: Springer, 2015.
- [6] E. Vittoz and J. Fellrath, "CMOS analog integrated circuits based on weak inversion operations," *IEEE J. Solid-State Circuits*, vol. SSC-12, no. 3, pp. 224–231, Jun. 1977.
- [7] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*. Hoboken, NJ, USA: Wiley, 2019.
- [8] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York, NY, USA: McGraw-Hill, 2002.
- [9] S. Mandal, S. Arfin, and R. Sarpeshkar, "Fast startup CMOS current references," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, pp. 2845–2848.
- [10] N. Talebdeyokhti, P. K. Hanumolu, P. Kurahashi, and U.-K. Moon, "Constant transconductance bias circuit with an on-chip resistor," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, pp. 2857–2860.
- [11] A. Amaravati, M. Dave, M. S. Baghini, and D. K. Sharma, "800-nA process-and-voltage-invariant 106-dB PSRR PTAT current reference," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 60, no. 9, pp. 577–581, Sep. 2013.
- [12] A. McLaren and K. Martin, "Generation of accurate on-chip time constants and stable transconductances," *IEEE J. Solid-State Circuits*, vol. 36, no. 4, pp. 691–695, Apr. 2001.
- [13] M. M. Vinaya, R. Paily, and A. Mahanta, "A new PVT compensation technique based on current comparison for low-voltage, near sub-threshold LNA," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 12, pp. 2908–2919, Dec. 2015.
- [14] V. Agarwal and S. Sonkusale, "Ultra low power PVT independent sub-threshold gm-C filters for low frequency biomedical applications," *Anal. Integr. Circuits Signal Process.*, vol. 66, no. 2, pp. 285–291, Feb. 2011.
- [15] M. Shahghasemi and K. M. Odame, "A constant  $g_m$  current reference generator with pseudo resistor-based compensation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 3, pp. 1115–1124, Mar. 2022.
- [16] N. Mirchandani and A. Shrivastava, "High stability gain structure and filter realization with less than 50 ppm/ $^\circ\text{C}$  temperature variation with ultra-low power consumption using switched-capacitor and sub-threshold biasing," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.
- [17] F. Serra-Graells and J. L. Huertas, "Sub-1-V CMOS proportional-to-absolute temperature references," *IEEE J. Solid-State Circuits*, vol. 38, no. 1, pp. 84–88, Jan. 2003.
- [18] H. C. Nauta and E. H. Nordholt, "New class of high-performance PTAT current sources," *Electron. Lett.*, vol. 21, no. 9, pp. 384–386, 1985.
- [19] C. Chu and Y. Wang, "A PVT-independent constant- $G_m$  bias technique based on analog computation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 61, no. 10, pp. 768–772, Oct. 2014.

- [20] J. Chen and B. Shi, "Circuit design of an on-chip temperature-compensated constant transconductance reference," *Anal. Integr. Circuits Signal Process.*, vol. 37, no. 3, pp. 215–222, 2003.
- [21] S. Pavan, "Systematic development of CMOS fixed-transconductance bias circuits," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 5, pp. 2394–2397, May 2022.
- [22] Y. Shen and R. Zhang, "Constant-gm bias circuit without off-chip components," in *Proc. 13th IEEE Int. Conf. Solid-State Integr. Circuit Technol. (ICSICT)*, Oct. 2016, pp. 1309–1311.
- [23] W. M. Sansen, F. O. Eynde, and M. Steyaert, "A CMOS temperature-compensated current reference," *IEEE J. Solid-State Circuits*, vol. SSC-23, no. 3, pp. 821–824, Jun. 1988.
- [24] B. R. Gregoire and U. Moon, "A sub 1-V constant  $G_m - C$  switched-capacitor current source," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 54, no. 3, pp. 222–226, Mar. 2007.
- [25] Z. Zhu, J. Hu, and Y. Wang, "A 0.45 V, nano-watt 0.033% line sensitivity MOSFET-only sub-threshold voltage reference with no amplifiers," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 9, pp. 1370–1380, Sep. 2016.
- [26] 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.
- [27] C. Galup-Montoro, M. C. Schneider, and I. J. B. Loss, "Series-parallel association of FET's for high gain and high frequency applications," *IEEE J. Solid-State Circuits*, vol. 29, no. 9, pp. 1094–1101, Sep. 1994.
- [28] D. J. Comer, D. T. Comer, and R. P. Singh, "A high-gain, low-power CMOS op amp using composite cascode stages," in *Proc. 53rd IEEE Int. Midwest Symp. Circuits Syst.*, Aug. 2010, pp. 600–603.
- [29] D. J. Comer, D. T. Comer, and C. S. Petrie, "The utility of the composite cascode in analog CMOS design," *Int. J. Electron.*, vol. 91, no. 8, pp. 491–502, Aug. 2004.
- [30] L. H. C. Ferreira, T. C. Pimenta, and R. L. Moreno, "An ultra-low-voltage ultra-low-power weak inversion composite MOS transistor: Concept and applications," *IEICE Trans. Electron.*, vol. E91-C, no. 4, pp. 662–665, Apr. 2008.
- [31] M. G. Degrauwe, J. Rijmenants, E. A. Vittoz, and H. J. De Man, "Adaptive biasing CMOS amplifiers," *IEEE J. Solid-State Circuits*, vol. SSC-17, no. 3, pp. 522–528, Jun. 1982.
- [32] A. Amaravati, M. Dave, M. S. Baghini, and D. K. Sharma, "A fully on-chip PT-invariant transconductor," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 9, pp. 1961–1964, Sep. 2015.
- [33] U. A. Antao, J. C. M. Hsieh, and T. W. Berger, "A 9-nW on-chip constant subthreshold CMOS transconductance bias with fine-tuning," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2017, pp. 1–4.
- [34] R. Nagulapalli, K. Hayatleh, S. Barker, S. Zourob, N. Yassine, and S. Sridevi, "A PVT insensitive programmable amplifier for biomedical applications," in *Proc. Int. Conf. Microelectron. Devices, Circuits Syst. (ICMDCS)*, Aug. 2017, pp. 1–5.



**Martin Lee** received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Alberta, Edmonton, AB, Canada, in 2017 and 2020, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include the design of analog, radio frequency, PVT-invariant, and subthreshold CMOS integrated circuits and systems.



**Kambiz Moez** (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 1999, and the M.Sc. and Ph.D. degrees from the University of Waterloo, Waterloo, ON, Canada, in 2002 and 2006, respectively. Since January 2007, he has been with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, where he is currently working as a Professor. His research interests include the analysis and design of analog, radio frequency, and millimeter-wave CMOS integrated circuits and systems for variety of applications, including wired/wireless communications, biomedical imaging, instrumentations, radars, and power electronics. He is a Registered Professional Engineer in the province of Alberta. He is serving as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS and *Electronics Letters* (IET).