A High-Linearity Low-Noise Reconfiguration-Based Programmable Gain Amplifier

Seok-Kyun Han, Huy-Hieu Nguyen, and Sang-Gug Lee

Abstract—This paper presents a high-linearity low-noise small-size programmable gain amplifier (PGA) based on a new low-noise low-distortion differential amplifier and a proposed reconfiguration technique. The proposed differential amplifier combines an inverter-based differential pair with an adaptive biasing circuit to reduce noise and distortion. The reconfiguration technique saves the chip size by half by utilizing the same differential pair for the input transconductance and load-stage, interchangeably. Fabricated in 0.18-µm CMOS, the proposed PGA shows a dB-linear control range of 21 dB in 16 steps from -11 dB to 10 dB with a gain error of less than ±0.33 dB, an IIP3 of 7.4~14.5 dBm, a P1dB of -7~1.2 dBm, a noise figure of 13 dB, and a 3-dB bandwidth of 270MHz at the maximum gain, respectively. The PGA occupies a chip area of 0.04 mm² and consumes only 1.3 mA from the 1.8 V supply.

Index Terms—Variable gain amplifier (VGA), programmable gain amplifier (PGA), dB-linear gain, linearity

I. INTRODUCTION

Variable gain amplifiers (VGAs) are key building blocks in many applications to accommodate a large dynamic range of signals, for example, in various wireless communication systems [1-7]. Beside the large dB-linear control range requirement, high linearity, low noise, and low power consumption are the main challenges in designing VGA. VGAs are usually implemented in a cascade of many gain stages to satisfy the large dB-linear gain range requirement [1-6]. The VGA architecture reported in [7] can provide an 84dB dB-linear gain range, while consuming the smaller power and chip area, compared to those of conventional topology by combining one variable and three fixed gain amplifier stages. According to [7], the gain of variable gain stage is controlled to follow the pseudo-exponential approximation function, $e^{2x} \approx (1+x)/(1-x)$, that can provide a dB-linear gain range of 20 dB with less than ±0.25 dB gain error. The schematic of the variable gain stage reported in [7] is shown in Fig. 1. In Fig. 1, the variable gain stage consists of a differential pair with diode-connected active loads. By simultaneously varying the size and bias current of the input and load transistors by the same ratio, the gain is controlled to follow the pseudo-exponential approximation function, $e^{2x} \approx (1+x)/(1-x)$. The simultaneous variation of the size and bias current keeps the current density of the transistors at the same value at all gain levels, leading to better linearity at a low gain compared to the case of the current density controlled VGA reported in [1]. However, the

Manuscript received Mar. 22, 2013; accepted May. 10, 2013
Dept. of Electrical Engineering, KAIST, Daejeon, Korea.
E-mail: sskk@kaist.ac.kr

Fig. 1. Schematic of the variable gain stage reported in [7].
linearity of the VGA is still limited by the inherent nonlinearity of the input and load transistors.

The variable gain stage shown in Fig. 1 can be considered a cascade of V-I and I-V converters. The distortion caused by the V-I converter (pre-distortion) can be eliminated by the inversion I-V conversion (distortion). Therefore, for unity gain condition, the nonlinearity generated by the V-I converter ($G_m$-stage) is cancelled by the I-V converter (load), leading to a very good linearity. However, for the voltage gain that is greater than one, the transistor size and the bias current of the $G_m$-stage becomes larger than that of the load. Therefore, the signal swing at the input of the I-V converter becomes larger than that of the V-I converter. Thus the distortion is not fully compensated, leading to linearity degradation. By the same principle, the linearity is degraded for the gain that is smaller than one. In summary, the variable gain stage shown in Fig. 1 shows the best linearity at unity gain, and at other gains, the linearity is determined by the combination of the V-I/I-V distortion compensation and the linearity of the $G_m$-load-stage. A simple way to improve the linearity is to increase the current densities of the $G_m$- and load-stages, which leads to a larger amount of current consumption. Therefore, an energy effective way to improve the linearity of VGA is to adopt transconductance linearization techniques.

In the conventional differential pair shown in Fig. 2(a), the common-source node does not stay constant, but fluctuates with respect to the input signal amplitude. This fluctuation of the common-source node voltage is the reason for the distortion [8, 9]. Fig. 2(b) shows a differential pair with adaptive biasing introduced in [8, 9]. In Fig. 2(b), the common-source node S is kept as a perfect virtual ground by varying the bias current of the differential pair $M_{1,2}$ through the adaptive biasing circuit which consists of $M_{3,4}$ and $M_{5,6}$, with respect to the amplitude variation of the input signal. However, the small-signal transconductance of the differential pair with adaptive biasing is smaller than that of the conventional differential pair. The reason is because a part of total bias current, $2I_{b0}$, is injected into adaptive biasing circuit that senses the input voltage but does not generate any output current. The larger the injected current is, the lower small-signal transconductance is provided by the differential pair with adaptive biasing, leading to higher input referred noise. Thus the differential pair with adaptive biasing trades small-signal transconductance and noise performance for low distortion performance.

Fig. 2(c) shows a differential amplifier based on a shunt-feedback inverter that is introduced in [12, 13]. In Fig. 2(c), the transconductance of the inverter-based differential pair can be twice that of the conventional differential pair, by properly sizing the PMOS transistor and for large enough value of feedback resistor $R_f$. Due to higher transconductance, the shunt-feedback inverter offers much smaller input referred noise compared to that of the conventional differential pair. This will be analyzed more detail in Section III. Therefore, the shunt-feedback inverter-based differential pair has advantages in transconductance and noise performance compared to those of the conventional differential pair. Moreover, by adopting shunt-feedback resistors, the output common mode voltage is defined without any additional circuit for biasing and common-mode feedback.

By applying adaptive biasing circuit introduced in [8, 9] to the inverter-based differential pair (see Fig. 2(c)), the disadvantages of adaptive biasing technique on transconductance and noise can be are partly compensated while low distortion performance is still achieved. The schematic details and linearity characteristics of the inverter-based differential pair with adaptive biasing are analyzed in Section II. The analysis and comparison of noise performance of three differential pairs will be given in Section III. Then, this paper reports a new PGA architecture that can offer high linearity and low noise performance by adopting the inverter-based differential pair with adaptive biasing. A new reconfiguration technique that utilizes the same differential pair for the input transconductance and load-

![Fig. 2.](image)
stage interchangeably is also proposed in this paper to save the chip size by half and improve the bandwidth of the PGA. The proposed reconfiguration technique and the implementation details of the proposed 4-bit PGA are described in Section IV. The experimental results are presented and discussed in Section V. Finally, the conclusions are given in Section VI. The idea of the inverter-based differential pair with adaptive biasing and reconfiguration technique have been introduced in [14] and [15], respectively by the authors. As an extended paper, this work will provide more in-depth analysis, simulation and measurement results.

II. LINEARITY ANALYSIS FOR INVERTER-BASED DIFFERENTIAL PAIR WITH ADAPTIVE BIASING

The highly linear transconductance stage that combines the inverter-based differential pair and adaptive biasing technique [14] is shown in Fig. 3. In Fig. 3, the inverter includes transistors $M_1$, $M_2$, $M_5$, and $M_6$, where transistors $M_1$ and $M_2$ are matched to $M_5$ and $M_6$, respectively. The adaptive biasing block consists of transistors $M_{1s}$, $M_{2s}$, $M_{3s}$, and $M_{4s}$ where transistors $M_{1s}$ and $M_{2s}$ are correspondingly matched to $M_{5s}$ and $M_{6s}$. The adaptive biasing block varies the bias current of the inverter by sensing the input signal amplitude. When the input signal, $v_{id}$ is small, the cross-coupled transistor pairs $M_1$ and $M_6$ operate in saturation mode while $M_5$, $M_6$ are in linear mode. When the input signal, $v_{id}$ is large enough, one of the transistor pairs $M_3$ and $M_6$ enters into the cutoff mode, while the other fall into saturation mode, reducing the total current through the adaptive biasing circuit [8]. For the given tail current of $2I_{in}$, the current reduction in the adaptive biasing circuit leads to a bias current increase in the inverter-based differential pair. This increase in bias current compensates the transconductance reduction of the differential pair at a large $v_{id}$. Therefore, the transconductance of the inverter-based differential pair in Fig. 3 stays constant over a wide range of input signal amplitude, thereby providing better linearity.

Applying the same quantitative analysis of the adaptive biasing block given in [8], but for a more general expression where the transconductance parameters of transistor pair $M_{1,3}$ to be $a$ times that of transistor pair $M_{5,6} (K_{7,8} = aK_{5,6})$, the total current dissipated by the adaptive biasing circuit can be given by

$$I_{DS} + I_{D6} = 2K_{7,8} \frac{V_C^2}{1 + \alpha} - 4\sqrt{1 + \alpha + 3\alpha - 5} \left( \frac{v_{id}}{2} \right)^2$$  (1)

where $I_{DS}$ and $I_{D6}$ are the drain current of $M_5$ and $M_6$, respectively, and $V_C = V_{CM} - V_S - V_{TN}$. The total current dissipated by the main transistors $M_{1,2}$ and $M_{3,4}$ can be given by

$$I_{D1} + I_{D2} = K_{1,2} \left( \frac{V_C}{2} + V_{CM} - V_S - V_{TN} \right)^2$$

$$+ K_{1,2} \left( -\frac{v_{id}}{2} + V_{CM} - V_S - V_{TN} \right)^2 = 2K_{1,2} [V_C^2 + (\frac{v_{id}}{2})^2]$$  (2)

$$I_{D3} + I_{D4} = -K_{3,4} \left( \frac{v_{id}}{2} + V_{CM} - V_{DD} - V_{TP} \right)^2$$

$$-K_{3,4} \left( -\frac{v_{id}}{2} + V_{CM} - V_{DD} - V_{TP} \right)^2$$

$$= -2K_{3,4} [(V_{CM} - V_{DD} - V_{TP})^2 + (\frac{v_{id}}{2})^2]$$  (3)

where $K_{1,2}$ and $K_{3,4}$ are the transconductance parameters of transistor pair $M_{1,2}$ and $M_{3,4}$, $V_{DD}$ supply voltage, and $V_{TP}$ the threshold voltage of PMOS transistors. By adding (1) and (2), the total supply current is given by

$$2I_{0} = I_{D1} + I_{D2} + I_{D3} + I_{D6} = 2 \left( \frac{K_{7,8}}{1 + \alpha} + K_{1} V_C^2 \right) + 2 [K_{1,2} - K_{7,8} \frac{4\sqrt{1 + \alpha + 3\alpha - 5}}{(1 + \alpha)^2} (\frac{v_{id}}{2})^2]$$  (4)

Then $V_C$ is expressed as

![Fig. 3. Inverter-based transconductance stage with adaptive biasing.](image)
\[ V_C = A \sqrt{1 + BV_{id}^2} \]  
(5)

where \( A = \sqrt{(1+\alpha)I_0 \over K_{7,8} + (1+\alpha)K_{1,2}} \) and

\[ B = {1 \over 4I_0} \left[ -K_{1,2} + K_{7,8} {4\sqrt{1+\alpha + 3\alpha - 5} \over (1+\alpha)^2} \right] \]

are coefficients that depend only on transconductor parameters of transistors and the total bias current. Expanding the equation of \( V_C \) in Taylor series and considering the first three terms, \( V_C \) is expressed as

\[ V_C = A \left[ 1 + {BV_{id}^2 \over 2} - {B^2V_{id}^4 \over 8} \right] \]  
(6)

At dc condition, \( V_{id}=0 \) and then \( I_{D1} + I_{D2} + I_{D3} + I_{D4} = 0 \). Thus from (2) and (3), the relation between \( V_{CM} \) and \( V_S \) is given by

\[ K_{3,4}(V_{CM} - V_{DD} - V_{TP})^2 = K_{1,2} V_C^2 \bigg|_{V_{id}=0} = K_{1,2}A^2 \]  
(7)

In Fig. 3, it can be seen that by neglecting the current through the feedback resistor \( R_F \), the differential incremental output current of the proposed transconductance stage can be given by

\[ i_d = i_0 - i_0 = (I_{D3} + I_{D1}) - (I_{D4} + I_{D2}) = 2V_{id}[K_{1,2}V_C + K_{3,4}(V_{DD} + V_{TP} - V_{CM})] \]  
(8)

where \( K_{3,4} \) is the transconductance parameter of the transistor pair \( M_{1,4} \). Substituting \( V_C \) from (6) and \( (V_{DD} + V_{TP} - V_{CM}) \) into (7) into (8), the differential incremental output current is expressed as

\[ i_d = 2(K_{1,2} + \sqrt{K_{1,2}K_{3,4}})AV_{id} + 2K_{1,2} \left[ {ABV_{id}^3 \over 2} - {AB^2V_{id}^5 \over 8} \right] \]  
(9)

As can be seen from (9), the third and fifth harmonic distortions can be removed from the incremental output current if \( B = 0 \) or

\[ K_{1,2} = {4\sqrt{1+\alpha + 3\alpha - 5} \over (1+\alpha)^2} K_{7,8} \]  
(10)

Then, the differential incremental output current is a linear function of input signal, \( V_{id} \), and the transconductance of the inverter-based differential pair with adaptive biasing (IDPAB) shown in Fig. 3 can be given by

\[ G_m[\text{IDPAB}] = \left. {di_d \over dv_{id}} \right|_{V_{id}=0} = 2(K_{1,2} + \sqrt{K_{1,2}K_{3,4}})A \]

\[ = 2\left(\sqrt{K_{1,2}I_0} + \sqrt{K_{3,4}I_0}\right) \left[ {4\sqrt{1+\alpha + 3\alpha - 5} \over 4\sqrt{1+4\alpha - 4} \right] \]  
(11)

Meanwhile, with similar calculations, the output current of the inverter-based differential pair (IDP) shown in Fig. 2(c) and the conventional differential pair (CDP) shown in Fig. 2(a) can be expressed as

\[ i_0[\text{IDP}] = 2\left(\sqrt{K_{1,2}[I_0]} + \sqrt{K_{3,4}[I_0]}\right)V_{id} \]

\[ - K_{1,2}[I_0]\left(\sqrt{K_{1,2}[I_0]} + \sqrt{K_{3,4}[I_0]}\right) \left. \right|_{V_{id}=0} \]  
(12)

\[ i_0[\text{CDP}] = 2\left(\sqrt{K_{1,2}[I_0]} V_{id} - {K_{1,2}[I_0] \over 2} \right) \left. \right|_{V_{id}=0} \]  
(13)

Then, the small-signal transconductance of the IDP and CDP are given by

\[ G_m[\text{IDP}] = 2\left(\sqrt{K_{1,2}[I_0]} + \sqrt{K_{3,4}[I_0]}\right) \]  
(14)

\[ G_m[\text{CDP}] = 2\sqrt{K_{1,2}[I_0]} \]  
(15)

where \( K_{1,2}[I_0] \) are transconductance parameters of NMOS pairs in Fig. 2(a), \( K_{1,2}[I_0] \) and \( K_{3,4}[I_0] \) transconductance parameters of NMOS and PMOS pairs in Fig. 2(c), respectively. The IDP and IDPAB are normally designed so that \( K_{1,2}[I_0] = K_{3,4}[I_0] \) and \( K_{1,2} = K_{3,4} \), respectively. With the same amount of total bias current, \( 2I_g \), to keep the current densities through transistors \( M_{1,2} \) and \( M_{3,4} \) in the three differential pairs equivalent, \( K_{1,2}[I_0] = K_{3,4}[I_0] = K_0 \) and \( K_{1,2} < K_p \). In this simulation, \( \alpha \) is chosen to be 1.1 and \( K_{1,2} = 1/2K_0 \). Then \( K_{7,8} = 1.1K_{5,6} \) and \( K_{1,2} = 1.02K_{5,6} \). Then the small-signal
transconductance of the CDP, IDP and IDPAB are 
\[ 2\sqrt{K_0 I_0}, 4\sqrt{K_0 I_0}, \text{ and } 2.3\sqrt{K_0 I_0}, \]
respectively. From (11), (14) and (15), it is clear that small-signal transconductance of IDPAB is smaller than that of IDP but slightly larger than that of CDP.

Generally, the linearity of a transconductance stage can be enhanced by minimizing the third-order coefficient of the Taylor series, \( g_m^{\alpha} \), as the nonlinearity is dominated by the third harmonic distortion [18]. As can be seen from (9), (12) and (13), if the condition (10) is satisfied, the third and fifth harmonic distortion at the output of IDPAB can be cancelled out, while at the output of IDP and CDP, the third harmonic distortion always exists. Thus compared to IDP, IDPAB trades current efficiency \( (G_m/I_{bias}) \) for low distortion performance. The above theoretical analysis can be verified by \( g_m \) simulation results shown in Fig. 4 and \( g_m^{\alpha} \) simulation results shown in Fig. 5. Fig. 4 shows the transconductance characteristics versus the input signal amplitude of four differential pairs: CDP, IDP, and IDPAB that consume the same amount of bias current. In this simulation, \( \alpha \) is chosen to be 1.1. Then \( K_{7,8} = 1.1K_{5,6} \) and \( K_{1,2} = 1.02K_{5,6} \). As can be seen in Fig. 4, the transconductance of IDP is higher than that of CDP due to the additional transconductance provided by the PMOS transistor pair. By adopting adaptive biasing technique, the small-signal transconductance of IDPAB is smaller than that of IDP but slightly larger than that of CDP. However, the transconductance of CDP and IDP decrease with an increase in the input signal amplitude while IDPAB show nearly constant transconductance over a wide input range. The comparison of \( g_m^{\alpha} \) among the CDP, IDP and IDPAB is shown in Fig. 5. As can be seen in Fig. 5, the IDPAB shows \( g_m^{\alpha} \) values nearly zero over the input signal ranges up to ±240 mV. The calculation in (12) ignores the third-order nonlinearity contributed from PMOS pseudo-differential pair because a fully-differential pair exhibits higher third-order nonlinearity than a grounded-source biased at the same current and device dimensions [16]. Therefore, the third-order term of IDP output current shown in (12) equals to that of CDP output current shown in (11). But in practice, nonlinear channel length modulation and mobility degradation with the vertical field in the channel increase the third-order distortion of a pseudo-differential pair. Thus the absolute value of \( g_m^{\alpha} \) of IDP is larger than that of CDP. This result can be observed in Fig. 5. The linearity performance is confirmed one more time by simulation results of total harmonic distortions (THD) versus input signal amplitude that is shown in Fig. 6. As can be seen in Fig. 6, the IDP shows a slightly better THD than the CDP due to an additional transconductance from pseudo-differential PMOS transistor pair that contributes less distortion than a fully-differential pair [16] while the IDPAB shows the THD much smaller than the other architecture.

The flat range of transconductance is limited by the condition where the transistor pairs \( M_{7,8} \) and \( M_{6,5} \) operate in saturation and linear mode, respectively. This condition guarantees that the above derivations, from (1) to (11), are valid. The value of \( \alpha \) will affect the input linear range and the transconductance flatness of the \( G_m^{\alpha} \)-cell. For example, the \( V_{DS} \) in Fig. 3 becomes smaller with a larger value of \( \alpha \). Therefore, transistors \( M_{6,5} \) easily fall into saturation mode, leading to a smaller input linear range. Otherwise, the \( V_{DS} \) becomes larger with a smaller value of \( \alpha \), leading to a larger input linear range but with the transconductance flatness reduction of the \( G_m^{\alpha} \)-cell. Fig. 7 shows the transconductance characteristics versus the input signal amplitude of the inverter-based differential pair with adaptive biasing (IDPAB) with three different values of \( \alpha \). As can be seen in Fig. 7, the

---

**Fig. 4.** Simulated \( g_m \) vs. input signal amplitude of the CDP, IDP and IDPAB.

**Fig. 5.** Simulated \( g_m^{\alpha} \) vs. input signal amplitude of the CDP, IDP and IDPAB.
larger $\alpha$ is, the smaller the input linear range but the better the transconductor flatness can be achieved. The value of $\alpha$ is chosen with the trade off being between the input linear range and transconductance flatness. The simulation results shown in Figs. 4-7 agree well with the analytical observation. In summary, IDPAB provides a wider flat input ranges compared to that of CDP and IDP with the cost of transconductance value compared to IDP. The following Section will discuss and compare noise performance of three differential pairs.

III. NOISE ANALYSIS

Fig. 8 shows the equivalent circuit for conventional differential pair, wherein the thermal and $1/f$ noises of each transistor are modeled as voltage source in series with the transistor gate $V_{n,j}^2$. The noise source $V_{n,j}^2$ can be expressed by (16)

$$
\overline{V_{n,j}^2} = \frac{16kT}{3g_{m,j}} \frac{2K_{N,P}}{C_{ox}(WL)} \alpha_i f
$$

(16)

where $k$ is Bonzeman constants, $C_{ox}$ and $K_{N,P}$ the process-dependent constants, $T$ the temperature, $g_{m,j}$ and $(WL)_i$

Fig. 6. Simulated total harmonic distortion versus input signal amplitude of the CDP, IDP and IDPAB.

Fig. 7. Simulated gm vs. input signal amplitude IDPAB with 3 different value of $\alpha$. transconductance and dimension of transistor $i$, respectively, and $f$ the frequency. Assuming the transconductance of transistor $M_1$ and $M_3$ are equal, $g_{m1}=g_{m3}$, the input-referred noise of CDP is given by

$$
\overline{V_{n,in}^2}_{[\text{CDP}]} = \frac{(2V_{n1}^2g_{m1}^2 + 2V_{n3}^2g_{m3}^2)}{g_{m1}^2} = 2(V_{n1}^2 + V_{n3}^2)
$$

(17)

Similarly, assuming $R_f$ is larger enough, the input-referred noise of IDP is given by

$$
\overline{V_{n,in}^2}_{[\text{IDP}]} = \frac{(2V_{n1}^2g_{m1}^2 + 2g_{m3}^2V_{n3}^2)}{(g_{m1} + g_{m3})^2}
$$

$$(V_{n1}^2 + V_{n3}^2)/2$$

(18)

It can be seen from Fig. 3, transistors pairs $M_{5,6}$ and $M_{7,8}$ do not contribute any noise to the output of the IDPAB. Thus input-referred noise calculation of IDPAB is similar to that of IDP. Note that as mentioned in Section II, with the same amount of bias current $2I_o$, the sizes of transistors $M_1-M_4$ in IDPAB are chosen to be half of that in IDP. Therefore, $(W/W_{IDPAB})=4$ and from (11) and (14), then $g_{m,i,IDPAB}=0.575g_{m,i}$. The modeled noise source at the gate of transistor $i$ ($i=1-4$) in IDPAB can be expressed by

$$
\overline{V_{n,in}^2}_{[\text{IDPAB}]} = \frac{16kT}{3g_{m,i,IDPAB}} \frac{2K_{N,P}}{C_{ox}(WL)_{i,IDPAB}} f
$$

$$
= \frac{1.74 \cdot 16kT}{3g_{m,i}} \frac{2 \cdot 2K_{N,P}}{C_{ox}(WL)_{i}} < 2\overline{V_{n,j}^2}
$$

(19)

Then the input-referred noise of IDPAB is given by

$$
\overline{V_{n,in}^2}_{[\text{CDP}]} = \frac{(V_{n1}^2 + V_{n3}^2)}{2} = \frac{1}{2}\overline{V_{n,in}^2}_{[\text{CDP}]}$

(20)
As can be seen from (17), (18) and (20), IDPAB shows larger input-referred noise than IDP but still smaller than CDP. Based on the IDPAB, together with a newly proposed reconfiguration technique, a low-noise, low-distortion, and small size PGA is proposed that is discussed in the following section.

**IV. PROPOSED PGA WITH A NEW RECONFIGURATION TECHNIQUE**

Fig. 9 shows the circuit details of the PGA reported in [14] where the gain control scheme introduced in [7] is applied but the CDP is replaced with the IDPAB for both the Gm-stage and diode connected load. Fig. 9(a) shows the conceptual constituting details of the PGA topology that includes input Gm-stage and diode-connected load arrays separately. Fig. 9(b) shows the schematic of the reconfigurable Gm/load-cell, while Fig. 9(c) shows the actual PGA schematic based on the reconfigurable Gm-load-cell. In Fig. 9(a), both the Gm-stage and diode-connected load arrays contain IDPAB cells with weighted size factors of a geometric sequence, 20, 21, 23...2^n-1, respectively. It means that the transistor sizes of IDPAB in the two arrays are varied in a binary weighted sequence while the current density is kept constant. By properly combining the control signal, the input transconductance and the load impedance can be varied in a binary sequence. Similar to the gain control scheme reported in [7], the gain of the PGA shown in Fig. 9(a) can be varied by switching the Gm- and load-cells of the corresponding arrays ON or OFF. For example, when the control word \(a_1 \ldots a_2 a_0\) is set to 0...000, all of the Gm-cells in the Gm-stage array turn OFF, except for the “Gm-cell null”, while all of the load-cells in the diode-connected load array turn ON, including the “load-cell null”, leading to the smallest value of input transconductance and load resistance. Thus the lowest level of voltage gain is achieved. When the control word \(a_1 \ldots a_2 a_0\) is changed to 0...001, the “Gm-cell 0” and “Gm-cell null” turn ON, while the “load-cell 0” turns OFF. Thus the input transconductance and the load resistance increase slightly, leading to one step increase in voltage gain. Note that in each gain step, when the \(i\)-th Gm-cell is ON, the \(i\)-th load-cell is OFF and vice versa. Therefore, each \(i\)-th Gm- and the corresponding load-cell can be combined into one reconfigurable Gm-load-cell, as shown in Fig. 9(b), where the cell can be configured as Gm-cell \((a_1 = 1)\) or load cell \((a_0 = 0)\) depending on the states of \(a\). Fig. 9(c) shows the final circuit schematic of the proposed PGA based on the reconfigurable IDPAB-based Gm-load-cell, which occupies approximately half the amount of chip area compared to the original PGA shown in Fig. 9(a).

The 3-dB bandwidth of PGA shown in Fig. 9 depends

![Fig. 9. (a) The proposed PGA with separate Gm- and load-cells, (b) Schematic of reconfigurable Gm-load-cell, (c) The final PGA schematic based on the reconfigurable Gm-load-cell.](image-url)
on parasitic capacitance at the output nodes and load resistance. The parasitic capacitors at the output nodes of $G_m$-cell include $C_{ds}$, $C_{gb}$ and $C_{db}$, while the parasitic capacitors at the output nodes of load-cells include $C_{gs}$, $C_{ds}$, $C_{gb}$ and $C_{db}$. Because $C_{ds}$, $C_{gb}$ and $C_{db}$ are much smaller than $C_{gs}$, they can be neglected. Thus the parasitic capacitance at the output node of the PGA is dominated by parasitic capacitors $C_{gs}$ of load-cells. The parasitic capacitors of each load-cell are shown in Fig. 10. In the PGA shown in Fig. 9(a), at high gain mode, some load-cells in the diode-connected load array are turned OFF. The additional parasitic capacitance of each deactivated load-cells degrades the 3-dB bandwidth of the PGA. Whereas, in the PGA shown in Fig. 9(c), the reconfigurable-$G_m$/load cells always turn ON. Thus there is no additional parasitic capacitance from deactivated cells. However, different from the PGA shown in Fig. 9(a), the PGA shown in Fig. 9(c) uses reconfiguring switches in the signal path and in connections between the input and output nodes that will affect the performance of the PGA. In the reconfigurable-$G_m$/load cell $i$ shown in Fig. 9(b), the nonlinearity from switches $a_i$ will diminish the linearity performance of the PGA; and the switches $a_i$ contribute additional parasitic capacitance to the output nodes, degrading the bandwidth of the PGA. The sizes of switches $a_i$ should be small to minimize their effect on 3-dB bandwidth of the PGA.

Fig. 11 shows the simulated ac response at highest gain of the PGA without reconfiguration, the PGA without reconfiguration but with all the deactivated load-cells disconnected from output nodes, and the PGA with reconfiguration. Due to the additional parasitic capacitance from deactivated load-cells, ac response of the PGA without reconfiguration is the dash-dash black line. While in the ideal case, if all the deactivated load-cells in the PGA without reconfiguration are disconnected from output nodes, the ac response of the PGA is the gray line that shows much better 3-dB bandwidth. The ac response of the PGA with reconfiguration is solid black line. As can be seen in Fig. 11, by minimizing the sizes of switches $a_i$, the 3 dB bandwidth at the highest gain of the PGA with reconfiguration shown in Fig. 9(c) is still larger than that of the PGA without reconfiguration shown in Fig. 9(a).

To diminish the effect of switches $a_i$ on linearity performance, the large size transmission gate switches should be used to reduce on-resistance. The sizes of switches $a_i$ can be chosen so that the on-resistance as small as 100Ω that is much smaller than input impedance of the $G_m$-cell at frequency up to few hundred MHz. Thus the nonlinearity contributed from switches $a_i$ can be ignored. From Fig. 9, since the transconductances of the $G_m$/load-cells in the arrays are geometrically binary weighted, the transconductances of the $G_m$-stage and diode-connected load can be given by

$$G_{m\text{-input}} = g_{m0}(2^0a_0 + 2^1a_1 + 2^2a_2 + 2^3a_3 + ... + 2^{n-1}a_{n-1} + k)$$

(21)

$$G_{m\text{-load}} = g_{m0}(2^0\bar{a}_0 + 2^1\bar{a}_1 + 2^2\bar{a}_2 + 2^3\bar{a}_3 + ... + 2^{n-1}\bar{a}_{n-1} + k)$$

(22)

where $a_i$ (with a value of 0 or 1) is the digital control bit, $n$ the number of control bit, $g_{m0}$ the transconductance of $G_m$/load- cell 0, and $k$ the size factor of “$G_m$-cell null”. From (21) and (22), the voltage gains of the PGA shown in Fig. 9(c) can be given by

$$A_v = \frac{G_{m\text{-input}}}{G_{m\text{-load}}} = \frac{2^0a_0 + 2^1a_1 + 2^2a_2 + 2^3a_3 + ... + 2^{n-1}a_{n-1} + k}{2^0\bar{a}_0 + 2^1\bar{a}_1 + 2^2\bar{a}_2 + 2^3\bar{a}_3 + ... + 2^{n-1}\bar{a}_{n-1} + k}$$

$$= \frac{x + k}{k + 2^n - 1 - x}$$

(23)
where \( x = 2^n a_0 + 2^{n-1} a_1 + 2^{n-2} a_2 + ... + 2^1 a_2 + a_1 \). By defining \( r = (x-2^{n-1}+0.5)/(k+2^{n-1}+0.5) \), (23) can be expressed as \( A = (1+r)/(1-r) \approx e^{2x} \), which is a pseudo-exponential function that can provide the dB-linear range of 20dB with a gain error of less than \( \pm 0.25\)dB. The constant \( k \) in (23) can be adjusted to control the gain range of the PGA while trading off the gain error values. From (23), the dB-linear range \( (GR) \) can be expressed as a function of the constant \( k \)

\[
GR = 40\log \frac{2^n - 1 + k}{k} \tag{24}
\]

The voltage gain versus control word based on (23) is shown in Fig. 12 for various \( k \) values. As shown in Fig. 12, the dB-linear range varies as a function of \( k \) value. For a gain error limit of less than \( \pm 0.85 \)dB, the dB-linear range can be extended up to 31dB.

The gain step can be reduced easily by inserting one or more binary weighted cells in the arrays with an equivalent number of increases in the control bits. From (24), when one more binary weighted cell is inserted into the arrays, the constant \( k \) needs to be increased to keep the same gain range. As a result, the input transconductance increases and load resistance reduces. Thus noise performance is even better at the cost of higher power consumption. The linearity performance does not degrade since the current densities of input transconductance and diode-connected resistance are kept the same. Furthermore, when one more cell is added, parasitic capacitance at the output node increase but load resistance reduce leading to the same 3-dB bandwidth of the PGA. Thus one or more control bits can be added without any performance degradation.

Moreover, Eq. (23) shows that the voltage gain of the proposed PGA does not depend on any component parameters but the ratios of the transconductance of reconfigurable-\( G_m \)/load cells. The mismatches of the configurable-\( G_m \)/load cells can be minimized by drawing the layout carefully. Thus the proposed PGA is very insensitive to temperature and process variations. Fig. 13(a) shows the Monte-Carlo simulation result at the highest gain of the proposed PGA. The standard deviation is only around 0.02dB over 1000 samples. Fig. 13(b) shows the simulated gain versus temperature. The proposed PGA shows the gain deviations of 0.2 dB over the temperature range of from -30°C to 100°C.

V. MEASUREMENT RESULTS AND COMPARISONS

The PGA shown in Fig. 9(c) is designed for 4 control bits to obtain the step size of around 1.4 dB while the constant \( k \) is set to 6 in order to achieve the dB-linear range of 21 dB. The proposed PGA is implemented in 0.18 \( \mu \)m CMOS technology with a 1.8V supply. For comparison, a 4-bit PGA based on the CDP reported in [7] and a 4-bit PGA based on IDP are also implemented. In the PGA circuit design, the lengths of PMOS and NMOS transistors are chosen to be 0.5 \( \mu \)m and 1 \( \mu \)m, respectively, for a good matching performance, minimizing short-channel effects, and a reasonable 3-dB bandwidth. The IDP- and IDPAB-based PGAs are designed to dissipate the same amount of current (1.3 mA) while the CDP-based PGA is designed to consume 0.1 mA more for the bias and common mode feedback circuits. All three PGAs are designed to provide a dB-linear control range of 21 dB from -11 dB to 10 dB. Fig. 14 shows the measured voltage gain and gain error versus the control word of the CDP-, IDP-, and IDPAB-based PGAs at a frequency of 100MHz. In Fig. 14, the...
three PGAs show almost the same dB-linear gain range of 21 dB from -11 to 10 dB with a gain error of less than ±0.33 dB.

Fig. 15 shows the measured frequency responses of the IDPAB-based PGA for a number of gain values. As can be seen in Fig. 15, the PGA shows a nearly constant bandwidth of 270 MHz over the gain variation.

A constant bandwidth is considered important to obtain a constant settling time for the automatic gain control feedback loop. Measurements also show the 3-dB bandwidth as being 290 MHz and 470 MHz at the maximum gain for the CDP- and IDP-based PGAs, respectively. As discussed above, the pole at the output nodes of the PGAs dominates the 3-dB bandwidth. As can be seen in Fig. 4, due to the additional transconductance from PMOS pair, IDP provides the higher small-signal transconductance, leading to lower impedance at the output node of IDP-based PGA compared to that of CDP and IDPAB-based PGA. Additionally, by adopting reconfiguration technique, the parasitic capacitance at the output nodes of IDP-based PGA is smaller than that of CDP-based PGA. Thus the IDP-based PGA shows a larger 3-dB bandwidth compared to that of CDP-based PGA. From Fig. 4, IDPAB provides slightly higher transconductance compared to CDP. However, due to the additional parasitic capacitance from the adaptive biasing circuits of load-stage contributing to the output of the PGA, 3-dB bandwidth of IDPAB-based PGA is still smaller than that of CDP-based PGA even though the new reconfiguration technique is adopted to IDPAB-based PGA.

The measured input IP3 and input P1dB of the all three PGAs are shown in Fig. 16 and 17, respectively. As can be seen in Fig. 16, the proposed IDPAB-based PGA shows more than 10 dB improvement in input P1dB compared to that of the CDP-based PGA at low gains, while in Fig. 17, the IDPAB-based PGA shows nearly 10 dB improvement in input IP3 at high gains compared to that of IDP- and CDP-based PGAs.

Fig. 18 shows the measured spectrum of three PGA outputs when 99 MHz and 101 MHz tones (-18 dBm each) are applied, while the voltage gain of three PGAs is set at maximum of 10 dB. In Fig. 18, the IDPAB-based PGA shows the value of the third inter-modulation product smaller by 12.6 and 17.4 dB compared to that of the IDP- and CDP-based PGAs, respectively.

The noise figures of the three PGAs at the maximum gain of 10 dB are shown in Fig. 19. As mentioned before, due to the higher transconductance, the IDP-based PGA shows about 4 dB better noise figure than that of the CDP-based PGA. The noise figure of IDPAB-based PGA
is higher than that of the IDP-based PGA, but it is still 2.5 dB smaller than that of the CDP-based PGA. This is because IDPAB shows a input-referred noise higher than that of IDP but smaller than that of CDP does as discussed in Section III. Noise figure measurement is carried out with a series noise figure analyzer Agilent N8975A, while other measurements are conducted with a signal generator Agilent E4438C and a series spectrum Analyzer Agilent E4440A.

Table 1. Performance Comparison

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Supply</td>
<td>20.52</td>
<td>5.4</td>
<td>27.5</td>
<td>2.52</td>
<td>2.34</td>
<td>2.34</td>
<td>mW</td>
</tr>
<tr>
<td>Gain range/gain step</td>
<td>(-39–55)</td>
<td>(-42–42)</td>
<td>(-35–55)</td>
<td>(-11.2–10.6)</td>
<td>(-11–10.5)</td>
<td>(-11–10.2)</td>
<td>dB</td>
</tr>
<tr>
<td>Gain error</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>&lt;±0.27</td>
<td>&lt;±0.33</td>
<td>&lt;±0.33</td>
<td>dB</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>900</td>
<td>350</td>
<td>210</td>
<td>330</td>
<td>470</td>
<td>270</td>
<td>MHz</td>
</tr>
<tr>
<td>IIP3</td>
<td>--</td>
<td>-22–20</td>
<td>-28–7</td>
<td>-3–6.3</td>
<td>-1.5–12.3</td>
<td>7.4–14.5</td>
<td>dBm</td>
</tr>
<tr>
<td>P1dB</td>
<td>-59–11</td>
<td>-40–8</td>
<td>-10.5–5.2</td>
<td>-7.5–0.7</td>
<td>-7–1.2</td>
<td>0.4–4</td>
<td>dBm</td>
</tr>
<tr>
<td>NF</td>
<td>6.8</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>11</td>
<td>13</td>
<td>dB</td>
</tr>
<tr>
<td>Die area</td>
<td>0.42</td>
<td>0.18</td>
<td>0.49</td>
<td>0.05</td>
<td>0.036</td>
<td>0.04</td>
<td>mm²</td>
</tr>
<tr>
<td>FoM</td>
<td>0.3*</td>
<td>12.8</td>
<td>1.3</td>
<td>7.3</td>
<td>41.1</td>
<td>108.3</td>
<td></td>
</tr>
</tbody>
</table>

* FoM with estimated IIP3
linearity performance compared to IDP-based PGA. By adopting adaptive biasing technique, the IDPAB-based PGA trades noise figure and bandwidth for linearity performance. To compare the amplifiers in terms of linearity, gain, bandwidth, noise, and power consumption, a figure of merit (FoM) that is defined as follows is estimated

\[ \text{FoM} = \frac{\text{IIP}_3[mW] \times \text{Av} \times \text{BW}[MHz]}{\text{P}_{dc}[mW] \times (F - 1) \times f[MHz]} \quad (25) \]

where \( \text{Av} \) is the voltage gain, \( \text{BW} \) the bandwidth, \( F \) noise factor, and \( \text{P}_{dc} \) the power consumption of the amplifiers. The FoM shown in Table 1 is calculated at the highest gain. From Table 1, IDP-based and IDPAB-based PGAs show better FoMs compared to others. Due to much higher IIP3 at highest gain, IDPAB-based PGA still archives much better FoM even with smaller bandwidth and higher noise figure compared to that of IDP-base PGA.

VI. CONCLUSIONS

An all-CMOS programmable gain amplifier that is based on a proposed inverter-based differential pair with adaptive biasing and a new reconfiguration technique is introduced. By adopting the proposed schemes, the proposed PGA shows about 10 dB higher IIP3, 2 dB lower noise figure and smaller chip size, compared to those of the conventional PGA while consuming about the same amount of current. The details of the linearity and noise analysis for the proposed PGA are desribed. The measurement results of the proposed PGA demonstrates 7.4~14.5 dBm of linearity, 13dB of noise at 10.2 dB of voltage gain, and 270 MHz of 3-dB bandwidth while consuming 1.3 mA from 1.8V supply and occupying 0.04 mm² of chip area.

ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. R11-2005-029-06001-0).

REFERENCES


