

A Programmable Piecewise Linear Large-Signal CMOS Amplifier

Axel Thomsen and Martin A. Brooke

Abstract—A high-speed CMOS piecewise linear approximation circuit is presented that can be programmed for correction of nonlinearity after fabrication. The basic building block generates a linear segment, for which slope and position can be adjusted. Adjustments to adapt to arbitrary functions are done with floating gate devices fabricated in standard CMOS technology. The circuit is a voltage-to-current converter with an input range of the full power-supply voltage swing. In an implementation with 18 linear segments less than 0.15% error over rail-to-rail input range was achieved for a linear transfer function. Examples of strongly nonlinear transfer functions approximated to 0.5% accuracy are shown. The large-signal 3-dB frequency is 10 MHz. The implementations are done solely with 2- μ m channel length devices.

I. INTRODUCTION

PIECEWISE linear approximation circuits have been used in various applications [1]–[3]. They are useful where arbitrary functions have to be realized that cannot be found in the device model equations, for example, as a waveform shaper. In short-channel MOS technology, like many other technologies, piecewise linear approximation of functions becomes important for two other reasons.

First, the traditional design approach for signal processing circuits does not yield sufficient accuracy. This approach relies on mathematical operations on the simple modeling equations to obtain a certain relationship [4]–[9]. An example would be a linear transconductance based on the square law relation between gate-source voltage and drain current of a MOSFET in saturation [6]. An implementation of such a circuit would not use minimum channel length devices. Short-channel MOS devices show strong second-order effects like mobility degradation and channel length modulation that will introduce high error and signal distortion to circuit designs based on simple modeling equations.

Second, most signal processing circuits have a limited range for which they operate with low error due to the limited range of validity of the modeling equations. For example, a recently presented large-signal transconductance [4] had a maximum input range of about 50% of the power supply. The distortion data given in most publications refer to input signals of 50% of the power supply or less. Piecewise linear approximators can increase the range of operation to the full power-supply voltage swing. Other problems that may require piecewise linear solutions are nonlinear load characteristics and nonlinear sensor characteristics.

Manuscript received June 28, 1992; revised August 17, 1992. This work was supported by NSF Research Initiation Award Grant MIP-90 111360 and Analog Devices Inc.

The authors are with the Microelectronics Research Center, Georgia Institute of Technology, Atlanta, GA 30332-0269.

IEEE Log Number 9204660.

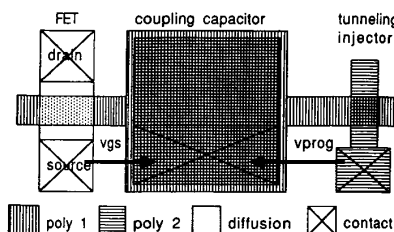


Fig. 1. Layout of the floating-gate device in a standard double-poly CMOS process tunneling occurs at the crossover of two polysilicon lines.

Existing CMOS implementations of piecewise linear approximators have utilized the transfer function of a differential pair [1] or used OTA's, diodes, and voltage or current sources [2], [3]. Implementations in faster and more compact short-channel technology show the additional problem of random effects that the existing designs cannot compensate for. Device parameters, device matching, sensor, or load characteristics vary from one fabrication run to the next. These problems require a more flexible piecewise linear implementation that allows post-fabrication adjustments.

Floating-gate devices have been successfully applied in analog circuits to compensate for mismatch in amplifiers [10], [11] as well as for weight storage in neural networks [12]. The fabrication of these devices in standard processes without additional processing steps as presented in [10] and [13] makes them a very interesting device that can have a wide range of applications. In this paper we describe their use for semi-permanent post-fabrication adjustment of bias currents.

II. BUILDING BLOCKS

A. The Floating-Gate Device

A special layout in a standard double-polysilicon CMOS process allows the fabrication of a floating-gate device with tunneling injector without additional processing steps [13] through MOSIS. The device layout is shown in Fig. 1. Tunneling occurs at a crossover of two polysilicon lines due to field enhancement and oxide thinning, when voltages of about 12 V are applied. While the properties of this device are not optimized in terms of compactness, read/write endurance, and programming voltage, the parameters significant for analog circuits applications, such as the achievable accuracy and the charge retention over time, are very promising. We have reported an estimated charge loss of less than 0.1% in ten years [13]. The achievable accuracy is only limited by the programming algorithm and the accuracy of measurements. Programming is done with pulses of fixed length and variable programming voltage.

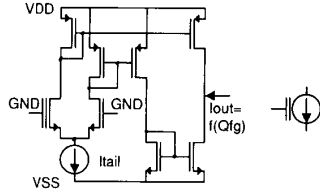


Fig. 2. Circuit diagram and symbol of an adjustable current source.

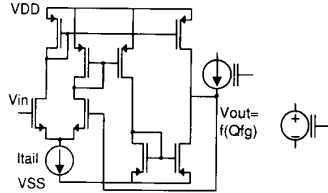


Fig. 3. Circuit diagram and symbol of an adjustable voltage source.

B. Adjustable Sources for Biasing

In this section adjustable sources are presented, the building blocks that allow the correction of a bias voltage or current to either direction. The first building block is an adjustable bidirectional current source. The circuit diagram is shown in Fig. 2. The floating-gate device is part of a differential pair. The charge on the floating gate Q_{fg} determines the input voltage V_{in} to this differential transconductance according to

$$V_{in} = \frac{Q_{fg}}{C_{cg}}$$

where C_{cg} is the coupling capacitance between floating gate and control gate. In this arrangement, thermal variation of the output current is low. To avoid damage during a positive programming pulse, the maximum current is limited to the tail current of the differential pair. The size of the tail current is set by the largest deviation to be corrected. The circuit is designed for output voltage in the following range:

$$V_{SS} + 0.5 \text{ V} < V_{out} < V_{DD} - 0.5 \text{ V}.$$

The other necessary building block is an adjustable floating voltage source. A transconductance in feedback can generate a voltage difference depending on the input current (Fig. 3). With the input current being programmable and bidirectional from the current source shown above, adjustment of voltage in either direction is possible.

C. A Circuit Module for a Piecewise Linear Output

The module to generate an adjustable piecewise linear segment is a transconductance. Voltage was chosen for the input because it needs to be distributed to all modules, and current for the output because it allows easy summing of all module outputs. It is a differential transconductance that has the voltage signal and a tunable reference voltage as inputs (Fig. 4). The output of this circuit gets clipped by two unidirectional current mirrors to a region of high linearity around the center point of the curve. Outside this region the output current is constant, the cutoff current I_{cut} . The current

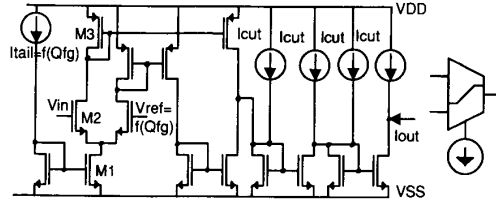


Fig. 4. Circuit diagram and symbol of a module for piecewise linear output.

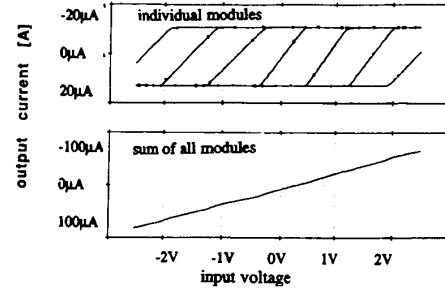


Fig. 5. Simulated output current of various modules and their sum.

output is shown in Fig. 5. The nonlinearity of the single module is determined by the ratio of tail current I_{tail} to I_{cut} . I_{tail} controls the transconductance g_m according to

$$g_m = 2 \cdot \sqrt{\beta \cdot I_{tail}}.$$

β is the transconductance parameter of the MOSFET. The input range of the module is given by

$$V_{min} = V_{SS} + \sqrt{\frac{I_{tail}}{\beta_1}} + \sqrt{\frac{I_{tail}}{2\beta_2}} + V_{t2}$$

and

$$V_{max} = V_{DD} + \sqrt{\frac{I_{tail}}{2\beta_3}} - |V_{t3}| + V_{t2}.$$

The indexes correspond to labels in Fig. 4. The maximum transconductance value is limited by the maximum tail current $\max(I_{tail})$ given either by

$$\max(I_{tail}) = I_{tail} + \max(I_{trim})$$

where I_{trim} is the output of a programmable current source, or by

$$\max(I_{tail}) = \frac{2\beta_1\beta_2}{\beta_1 + 2\beta_2} (V_{min} - V_{SS} + V_{t2})^2.$$

A minimum transconductance value cannot be specified, but with an increased ratio I_{cut}/I_{tail} the nonlinearity of the segment increases.

For input values near the power supply voltages, simple source-follower circuits are used as level shifters. Neglecting the bulk effect, the relationship between input voltage v_{in} and output voltage v_{out} is given by

$$v_{out} = v_{in} - v_t - \sqrt{\frac{I_{bias}}{\beta}} (\text{n-source follower}).$$

where v_t is the threshold voltage and I_{bias} the bias current for the source-follower circuit. This circuit can shift an input

voltage that exceeds the input range of the module into that range. A more detailed analysis has to include the bulk effect, which results in a voltage gain A_v slightly below unity given by the relation

$$A_v = \frac{g_m}{g_m + g_{mb}}$$

where g_m and g_{mb} are the small-signal transconductances from the gate and bulk, respectively. The reduced gain of these modules can be compensated for by increasing the tail current of the differential transconductance. Furthermore, adjustments of the tail current can compensate for mismatch in the transconductance parameter of the devices in the differential pairs.

This module has a transfer function that does not provide a smooth transition between slightly misaligned segments. Its advantages are its simplicity and the fact that the nonlinearity of the segment is low, thus allowing a precise approximation of a linear function. A smooth transfer function like a tanh function is more difficult to implement and introduces errors when approximating a linear relationship.

III. SYNTHESIS AND ADJUSTMENT OF THE CIRCUIT

A. Synthesis

The circuit diagram for the complete circuit is shown in Fig. 6. From left to right it consists of a tapped resistor to generate approximate reference voltages, p- and n-type level shifters to shift input voltage and reference voltage into the input range of the transconductance (PSF, NSF), an adjustable floating voltage source, the transconductance circuit to generate a piecewise linear segment, and an adjustable current source that is connected to the tail current of the transconductance. The number of modules depends on the accuracy needed. The number and kind of level shifters has to be determined from the input range of the module. The circuit requires a bias current for the transconductance element and reference voltages at the resistor. To achieve rail-to-rail input range the power supply voltages were chosen as references. Other choices are possible for different input ranges. An approximate generation of reference voltages and tail currents is important for various reasons: first, it simplifies trimming, since trimming now is only a slight correction to a good first guess and, second, it reduces any effect of charge loss from the floating gate because the trimming current provides only a small percentage of the bias current.

Certain restrictions apply to the functions that can be approximated: a function must be monotonic; the modules have a limited input range; the maximum gain is limited by the maximum tail current possible; and the achievable accuracy depends on the variation of the first derivative. A more flexible implementation for this purpose would include a programmable cutoff current. In general, the more knowledge about a function can be implemented in the approximate reference circuit, the more accurate a programmed function will be.

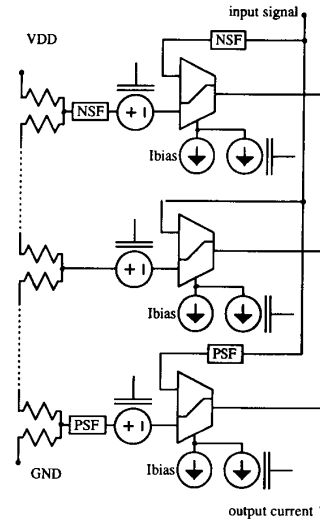


Fig. 6. Schematic of the complete circuit. The required number of modules depends on the desired error. See Section III for explanation.

B. Adjustment

Trimming is done only once after fabrication, but could be repeated if necessary. Changes in the nonlinearity or aging of MOS devices may require retrimming. It is done in the following manner: first, a general reset is done to set all devices into a known state. Here minimal reference voltage and maximum transconductance were used. Then beginning with the module with the highest input voltage, reference and gain are adjusted roughly first. After all modules are adjusted roughly, the trimming tolerances are reduced and all floating-gate charges readjusted. The accuracy of the reference voltages is most crucial to the overall accuracy of the transfer curve.

The programming algorithm used for the adjustment of floating-gate charge is as follows. Since the tunneling injectors are not very uniform, initial programming voltages have to be determined for each device. To do this, the programming voltage is increased by 1 V with every pulse until the measured value exceeds the target. Then the polarity of the programming voltage is reversed and the same procedure applied. The programming voltages found this way cause small changes in floating-gate charge. If the target is exceeded after only one pulse, the programming voltage for that direction is reduced by 0.5 V. Thus the programming voltages decrease as the target gets closer until finally the requested tolerance is met. This will take about 10 to 50 steps depending on the tolerance.

Adjustment of floating-gate devices for analog applications is a procedure that requires several cycles of measurement and programming pulse for every device. A system with 20 to 40 devices requires about 500 to 1000 cycles to achieve the accuracy stated here. The effort is much less than what is required for laser trimming and can be reduced by including more circuitry on chip. The algorithm explained above can be implemented on chip in a state machine. If the fabrication process supports voltages up to 25 V without drain-substrate breakdown, a full integration of the trimming process on chip is possible.

C. Accuracy

Errors get introduced to the transfer function in three places. First, the active region of a transconductance has a nonlinearity that causes a deviation from a straight line. The nonlinearity depends on the width of this active region determined by the cutoff current I_{cut} . The resulting deviation can be calculated from the simulation of a single module. A reduction in nonlinearity can be achieved by decreasing the cutoff current to reduce the width of the active region. The other sources of error are programming tolerances when fine-tuning the references and the transconductance values. Any tolerance for the target output current ΔI for a certain module will directly affect the output error. The tolerance in transconductance value Δg will appear as output error as

$$I_{\text{error}} = \Delta I + \frac{\Delta g}{g_m} \cdot I_{\text{cut}}.$$

These error sources can be reduced by decreasing tolerances. This will increase the required programming time. The error and distortion characteristics of this circuit are unlike other analog circuits that show increasing error for larger input signals and have a dominant second harmonic. The error is a random deviation from a straight line. The total harmonic distortion is approximately equivalent to the relative error

$$\text{THD} \approx \text{relative error} = \frac{I_{\text{error}}}{I_{\text{max}} - I_{\text{min}}}.$$

The harmonic distortion is best for large signals. The power distribution in the harmonics is very different from other analog circuits. The power is spread widely over many harmonics with less power than usual in the low-frequency harmonics.

IV. RESULTS

The circuit was implemented in a 2- μm double-polysilicon p-well process. All devices have a drawn channel length of 2 μm . Implementations with 12, 15, and 18 modules were tested. The circuits were identical except for the method of reference voltage generation in the 15-module version.

A. Linear Operation

The 18-module implementation was biased with a tail current of 30 μA and a cutoff current of 4 μA . The circuit was programmed to an arbitrarily chosen transfer curve of

$$I_{\text{out}} = 95 \mu\text{A} - 30 \frac{\mu\text{A}}{\text{V}} \cdot V_{\text{in}}$$

with tolerances of $\Delta g = 0.4 \mu\text{A/V}$ and $\Delta I = 100 \text{ nA}$. The transfer curve before and after trimming is shown in Fig. 7. It is obvious that the untrimmed circuit has insufficient accuracy for almost any application. The deviation of the trimmed curve from the desired curve is shown in Fig. 8. The accuracy observed was 0.2 μA or 0.15% of the output swing over the full input range. The measured large-signal 3-dB frequency for operation into a 50- Ω load is 10 MHz. The spectrum was of a circuit programmed to 1% accuracy was measured for a rail-to-rail input signal at 1 MHz. The third harmonic is approximately 40 dB below the signal. The output resistance

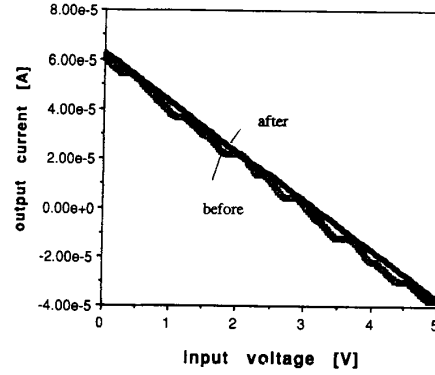


Fig. 7. Output characteristic before and after adjustments. The accuracy based only on matching (before) is not sufficient.

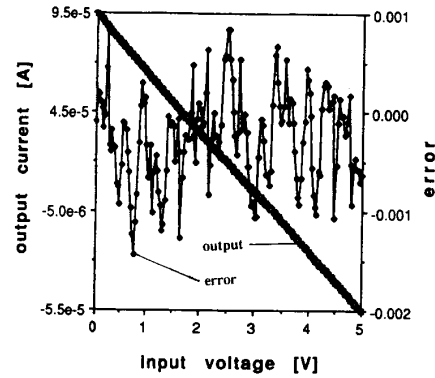


Fig. 8. Approximation of a linear function and remaining error. The error is less than 0.15% of the output swing. It is highest in the transition region between modules.

is in the order of 100 k Ω . The output voltage swing is 0.5 to 4.5 V. The speed is limited by the slew rate at the input nodes of the current mirrors that provide the current clipping, which have a large voltage swing between the active and the cutoff operation. Reducing this voltage swing or the capacitance at these nodes will improve the speed performance of the circuit.

B. Nonlinear Operation

Two circuits were programmed to nonlinear transfer functions given by

$$I_{\text{out}} = 20 \mu\text{A} - 70 \mu\text{A} \cdot \arctan \frac{V_{\text{in}} - 2.5 \text{ V}}{1.6 \text{ V}}, \quad 15 \text{ modules}$$

and

$$I_{\text{out}} = 100 \mu\text{A} \cdot e^{(V_{\text{in}} - 5 \text{ V})/7.21 \text{ V}} - 137 \mu\text{A}, \quad 12 \text{ modules}.$$

The parameters were chosen to fit the input and output ranges of the given bias point. The two transfer functions are shown in Fig. 9. The error was calculated from application of the inverse function to the output.

$$\Delta V = V_{\text{in}} - 2.5 \text{ V} + 1.6 \text{ V} \cdot \tan \frac{I_{\text{out}} - 20 \mu\text{A}}{70 \mu\text{A}}$$

$$\Delta V = V_{\text{in}} - 5 \text{ V} + 7.21 \text{ V} \cdot \log \frac{I_{\text{out}} + 137 \mu\text{A}}{100 \mu\text{A}}$$

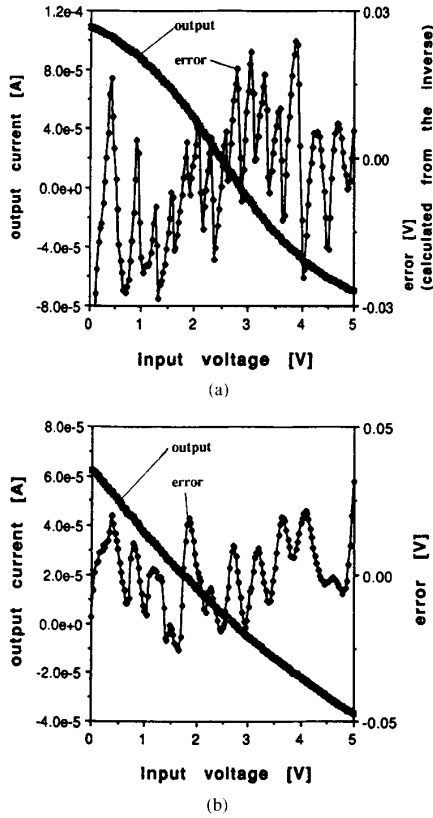


Fig. 9. (a) Approximation of an arctan function and remaining error. (b) Approximation of an exponential function and remaining error. The error is calculated from the inverse of the transfer function. It is less than 0.5%.

A deviation of 25 mV or 0.5% of the input swing was measured (Fig. 9). The first approximation was done with 15 modules, the second one with 12 modules.

In addition to the programming tolerances, an error is inherent in the piecewise linear approximation approach. It strongly depends on the shape of the function and is least for low variations in the first derivative. The error can be reduced by increasing the number of modules.

C. Noise

An analysis of the noise performance of this circuit has to be done for three operating points of the module: "saturated high," "saturated low," and "active." In a complete system there is always only one active module, all other modules are in the saturated state. Table I shows the noise power contributions of modules under various biasing conditions based on a SPICE simulation. Only the one active module shows a significant noise level that is comparable to a traditionally designed circuit. The noise contribution of the saturated modules is lower. Only a few transistors in the clipping current mirrors are noise contributors. The amplitude of their noise is determined only by the size of I_{cut} . If the number of modules n is increased to reduce error and I_{cut} is reduced to maintain the same maximum output current, the total noise remains constant

TABLE I
NORMALIZED NOISE POWER CONTRIBUTION OF MODULES
UNDER VARIOUS BIAS CONDITIONS. THE CONTRIBUTION OF
THE SATURATED MODULES IS MUCH LESS THAN ACTIVE MODULE

biasing	$I_{tail}/I_{cut}=$ 30 μ A/4 μ A	$I_{tail}/I_{cut}=$ 30 μ A/3 μ A	$I_{tail}/I_{cut}=$ 40 μ A/4 μ A
Vref=Vin (active)	1	0.940	1.128
Vref<<Vin (saturated)	0.317	0.249	0.317
Vref>>Vin (saturated)	0.077	0.059	0.077

because the noise contribution per module decreases with I_{cut} . Overall the noise will be slightly higher than in a traditionally designed circuit but the input range will be higher, yielding a competitive signal-to-noise ratio for this design method.

D. Temperature Behavior

Measurements and simulations of the temperature behavior of this circuit show that temperature dependence is significant where the circuit output depends on the transconductance parameter of a FET [14]. This parameter decreases with increasing temperature affecting the transconductance value g_m of each module as well as the floating voltage source values. According to simulations, the variation of g_m is 0.5%/°C, and the voltage variation is 0.25%/°C. The performance of the circuit is affected by this only locally, because the resistor chain as the main outline of the transfer function is temperature independent. Measurements of temperature dependence showed local variations of 0.8%, and an overall error of 2% for a temperature of 100°C for a circuit trimmed at room temperature. This thermal variation was due to an unsophisticated reference design. The temperature variation of the cutoff current caused an increased overall error. An improved reference implementation with temperature-stable cutoff current and temperature-controlled tail current would greatly improve the temperature behavior.

E. Performance Comparison

Due to the use of minimum-geometry devices in a short-channel technology and a very simple module design, the speed, cell size, and power consumption are considerably better than previously published designs of nonlinear function approximators. Table II shows a comparison of performance data. The crucial advantage of this design is its ability to compensate for the nonidealities introduced by the fabrication with on-chip storage of the required bias currents. In comparison to other linear transconductors this design shows an increased input range, higher area and power demand, and comparable speed and accuracy. As opposed to other analog design concepts, this circuit will not lose accuracy when implemented in faster technologies with even stronger second-order effects, so that the speed performance could be improved without reduction in accuracy.

V. CONCLUSION

A high-speed piecewise linear approximator circuit is presented. The circuit was programmed to approximate two nonlinear functions as well as a highly linear transfer function.

TABLE II
PERFORMANCE COMPARISON OF THIS CIRCUIT TO
OTHER NONLINEAR FUNCTION APPROXIMATORS

	this circuit	other designs
module power consumption	1mW	10mW [2]
module size	55 μ m * 520 μ m	220 μ m * 700 μ m [2]
chip size	825 μ m * 550 μ m (15 modules)	720 μ m * 1840 μ m [1]
3dB frequency	10MHz	1.1MHz [1]
distortion	0.15%	0.2% [1]

The input range is the full power-supply voltage swing. An error of 0.15% was measured for the linear case, and 0.5% error was measured for two nonlinear functions. The 3-dB frequency of the circuit is 10 MHz. It is shown that precision analog circuit design in short-channel technology is possible in continuous time and feedforward configuration with floating-gate circuit trimming. Not only does this circuit improve the possible input range and accuracy for high-speed large-signal transconductance circuits, but it also allows high-accuracy correction of nonlinearity after fabrication, which will improve performance, increase yield, or simplify the design of signal processing systems that include nonlinear amplifiers or sensors.

REFERENCES

- [1] J. W. Fattaruso and R. G. Meyer, "Triangle-to-sine wave conversion with MOS technology," *IEEE J. Solid-State Circuits*, vol. SC-20, pp. 623-631, 1985.
- [2] E. Sanchez-Sinencio, J. Ramirez-Angulo, B. Linares-Barranco, and A. Rodriguez-Vazquez, "Operational transconductance amplifier-based nonlinear function synthesis," *IEEE J. Solid-State Circuits*, vol. 24, pp. 1576-1586, 1989.
- [3] J. Ramirez-Angulo, E. Sanchez-Sinencio, and A. Rodriguez-Vazquez, "A piecewise-linear function approximation using current mode circuits," in *Proc. Int. Symp. Circ. Syst.* (San Diego), 1992, pp. 2025-2028.
- [4] J. Silva-Martinez, M. S. J. Steyaert, and W. M. C. Sansen, "A large-signal very low-distortion transconductor for high-frequency continuous-time filters," *IEEE J. Solid-State Circuits*, vol. 26, no. 7, pp. 946-954, 1991.
- [5] P. M. Van Peteghem, H. M. Fossati, G. L. Rice, and S. Y. Lee, "Design of a very linear CMOS transconductance input stage for continuous time filters," *IEEE J. Solid-State Circuits*, vol. 25, no. 2, pp. 497-501, 1990.
- [6] K. Bult and H. Wallinga, "A class of analog CMOS circuits based on the square law characteristics of a MOS transistor in saturation," *IEEE J. Solid-State Circuits*, vol. SC-22, pp. 357-365, June 1987.
- [7] E. Seevinck and R. F. Wassenaar, "A versatile CMOS linear transconductor/square law function circuit," *IEEE J. Solid-State Circuits*, vol. SC-22, pp. 366-377, June 1987.
- [8] R. R. Torrence, T. R. Viswanathan, and J. V. Hanson, "CMOS voltage to current transducers," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 1097-1104, Nov. 1985.
- [9] A. Nedungadi and T. R. Viswanathan, "Design of linear CMOS transconductance elements," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 891-894, Oct. 1984.
- [10] L. R. Carley, "Trimming analog circuits using floating-gate analog MOS-memory," *IEEE J. Solid-State Circuits*, vol. 24, no. 6, pp. 1569-1575, 1989.
- [11] E. Säckinger and W. Guggenbühl, "An analog trimming circuit based on a floating gate device," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, pp. 1437-1440, 1988.
- [12] M. Holler, S. Tam, H. Castro, and R. Benson, "An electrically trainable artificial neural network (ETANN) with 10240 floating gate synapses," in *Proc. Int. Joint Conf. Neural Networks*, vol. II (Washington, DC), 1989, pp. 191-196.
- [13] A. Thomsen and M. A. Brooke, "A floating gate MOSFET with tunneling injector fabricated using a standard double polysilicon CMOS process," *IEEE Electron Device Lett.*, vol. 12, no. 3, pp. 111-113, 1991.
- [14] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*. New York: Holt, Rinehart and Winston, 1987.