2011

Analog Signal Processing Elements for Energy-Constrained Platforms

Kyle Robert McMillan
West Virginia University

Analog Signal Processing Elements for Energy-Constrained Platforms

by

Kyle Robert McMillan

Thesis submitted to the
College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Electrical Engineering

Afzel Noore, Ph.D.
Lawrence A. Hornak, Ph.D.
David W. Graham, Ph.D., Chair

Lane Department of Computer Science and Electrical Engineering

Morgantown, West Virginia
2011

Keywords: Low-power signal processing, analog, derivative, floating gates, AGC, VGA, subthreshold

Abstract

Analog Signal Processing Elements for Energy-Constrained Platforms

by

Kyle Robert McMillan

Energy constrained processing poses a number of challenges that have resulted in tremendous innovations over the past decade. Shrinking supply voltages and limited clock speeds have placed an emphasis on processing efficiency over the raw throughput of a processor. One of the approaches to increase processing efficiency is to use parallel processing with slower, lower resolution processing elements. By utilizing this parallel approach, power consumption can be decreased while maintaining data throughput relative to other more power-hungry architectures.

This low resolution / parallel architecture has direct application in the analog as well as the digital domain. Indeed, research shows that as the resolution of a signal processor falls below a system-dependent threshold, it is almost always more efficient to preform the processing in the analog domain. These continuous-time circuits have long been used in the most energy-constrained applications, ranging from pacemakers and cochlear implants to wireless sensor "motes" designed to run autonomously for months in the field.

Most audio processing techniques utilize spectral decomposition as the first step of their algorithms, whether by a FFT/DFT in the digital domain or a bank of bandpass filters in the analog domain. The work presented here is designed to function within the parallel, array-based environment of a bank of bandpass filters. Work to improve the simulation of programmable analog storage elements (Floating-Gate transistors) in typical SPICE-based simulators is presented, along with a novel method of harnessing the unique properties of these Floating-Gate (FG) transistors to extend the linear range of a differential pair. These improvements in simulation and linearity are demonstrated in a Variable-Gain Amplifier (VGA) to compress large differential inputs into small single-ended outputs suitable for processing by other analog elements. Finally, a novel circuit composed of only six transistors is proposed to compute the continuous-time derivative of a signal within the sub-banded architecture of the bandpass filter bank.

Acknowledgements

Many thanks to my family - Mom, Dad and Erin. Without your support, I never would have had so many of the opportunities that have led me here. I love you all, and I can't express how much you mean to me.

Many thanks to my friends - too many to name. Without your support, I never would have pursued so many of the opportunities that have kept me here.

Finally, a quote:

If we knew what it was we were doing, it would not be called research, would it?
- Albert Einstein

Contents

List of Figures

Chapter 1

Introduction

In many areas of our society, there is an unmet need for signal processing on energy-constrained processing platforms. Whether they be cellular phones, pacemakers [1], hearing aids and cochlear implants [2], speech recognition systems [3], speaker recognition systems [4] or autonomous sensor platforms [5], these devices can realize significant improvement through energy-efficient processing. Current implementations of many low-power systems rely heavily on Digital Signal Processing (DSP), an inefficient technique for low-to-moderate resolution signals [6].

Analog (Continuous-Time) processing, possesses some inherent advantages over Digital (Discrete-Time) processing. Perhaps the largest benefits stem from the fact that all real-world signals are continuous-time; an ASP can perform computations on the signal directly and produce a continuous time output, suitable for direct interfacing to an eternal system. By contrast, a DSP must first discretize a continuous time signal via an Analog-to-Digital Converter (ADC), process the discrete-time signal in a digital core, then convert the output back to a continuous-time signal via a Digital-to-Analog Converter (DAC). These two extra steps, the DAC and the ADC, draw relatively large amounts of power, consume valuable on-chip area, and introduce additional noise into the system.

However, DSP's are flexible, versatile tools that are easy to simulate, reconfig-

urable and ubiquitous. ASP's are not easily reconfigurable, difficult to simulate and, due to the lack of reconfigurability, must be designed to meet specific design criteria on a per-case basis. Therefore, many designers sacrifice the native efficiency of the ASP for the rapid development and low design cost of the DSP. For many applications, this trade-off is acceptable. However, in the field of energy-constrained platforms, efficiency becomes paramount and the design challenges of reconfigurability, simulation and development must be addressed.

To that end, an ASP for audio-band signals has been developed [7] to implement several common signal processing operations and incorporate this new research into easing the design of ASP's. This particular processor implements one of the most common audio signal processing functions - spectral decomposition. In the digital domain, this is most commonly implemented through a DFT or FFT. However, previous work [7] has accomplished the same goals by implementing an array of continuous-time analog Band-pass Filters (BPFs) with adjustable passband gain, quality factor and frequency spacing; all programmable via Floating-Gate (FG) transistors and consuming only micro-watts of power. By providing both temporal and spectral representations of an analog signal, this array of filters (a "filterbank") provides the basis for further signal processing within each sub-band of the original signal.

This work presents three innovations to address the un-met needs of reconfigurability, system simulation and processing limitations within this sub-banded filterbank environment (fig. 1.1). First, a method of simulating floating-gate transistors is presented to increase the flexibility, reconfigurability and programmability of not only the sub-banded processing elements, but the BPF array itself. Elements of this research also serve to increase the limited dynamic range of many sub-banded signal processing elements, leading to an investigation of other methods to further increase the dynamic range. Coupled with the work on the FG model, a novel technique to increase the dynamic range of many circuits based around the differential pair structure (Operational Transconductance Amplifiers (OTAs)) is presented. Finally,

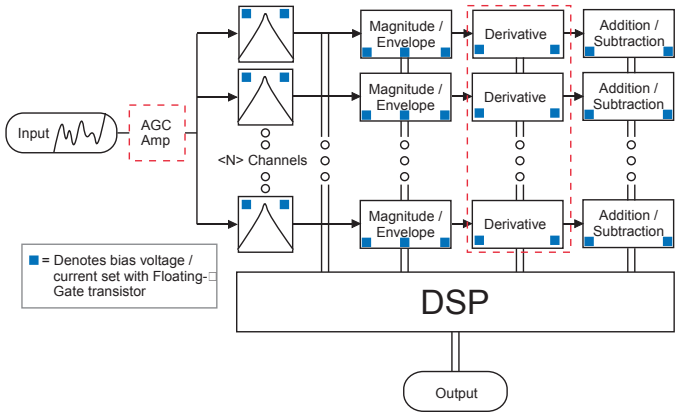


Figure 1.1: The proposed analog/digital signal processor architecture, where signal feature extraction is preformed in the analog domain and high-order processing is preformed in the digital domain. Biasing through Floating-Gate transistors is shown in blue, major processing elements covered here are outlined in red.

to address the need for a small derivative circuit with a high array fill-factor and low power consumption, a novel circuit consisting of only six transistors to implement the derivative function on a sub-banded signal is introduced and analyzed.

The following sections provide a brief overview of each of these accomplishments and provide a context for their introduction.

1.1 Programmability and Flexibility - FG Modeling

One of the traditional drawbacks to implementing an analog signal processing system is the flexibility of a design once it has been fabricated. Common biasing techniques for processing elements utilize resistive dividers or Digital-to-Analog Converters (DACs). Of the two, DACs consume a tremendous amount of power and area relative to the analog circuits they bias, while resistive dividers also carry a high area cost and are hard to fabricate with precision on many CMOS processes. These factors make these biasing methods them very unsuitable for array-based applications, where

each processing element may require three or four biases each.

Over the last decade, Floating-Gate transistors have been used in place of resistive dividers in many analog applications due their programmability and long-term storage of continuously-valued variables. FG transistors have been used as the core element of FLASH memory for several decades, and digital designers have developed several ways to program them to operate in binary states. However, compared to a regular MOSFET, these floating-gate transistors demonstrate non-ideal behavior, such as a large capacitive coupling between the gate and drain and an effective reduction in the gate-channel voltage coupling term, κ. While these non-idealities are of little consequence to digital operations (and are not modeled in many digital simulators), they are critical when the FG devices are operated in the analog domain.

Thus, there has been an un-met need for a simulation model to accurately describe floating-gate transistor behavior during programming in DC, AC, and Transient analyses over the open range of analog states. In Chapter 2, such a model is presented, along with simulation results compared to data taken from a circuit fabricated in a 0.5μm CMOS process. These tests demonstrate proper modeling of several key characteristics of FG transistors and the benefits of the Multiple-Input Floating Gate (MIFG) transistor. Specifically, charge modification, voltage storage, simulation flexibility, and increased linear range in a differential pair via capacitive division are demonstrated [8].

1.2 Increasing Linear Range - Ohmic Biasing

In many analog filterbanks, especially those designed to model biological processes, traditional low-power CMOS design techniques restrict the dynamic range of filters and other components to 40-60 dB [9]. However, biological systems, such as the human cochlea, have dynamic ranges of $\sim$120dB. To address this difference, a novel method of increasing the dynamic range through an expansion of input linear range

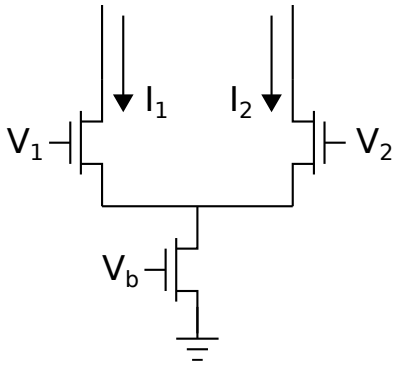


Figure 1.2: Standard nFET based differential pair. Output current is typically taken as the difference between I_1 and I_2 such that $I_{out} = I_1 - I_2$.

of a circuit known as the differential pair (fig. 1.2) is presented here.

The differential pair is one of the most fundamental circuits in analog signal processing - it is the common building block of Operational Transconductance Amplifiers, Operational Amplifiers, and the Gilbert multiplier cell. Since the differential pair is an ubiquitous foundation for most differential-mode analog processing elements, the fixed input linear range ($\sim$74 mV) of a differential pair biased for low-power operation with a subthreshold current presents a design challenge. Low power, open-loop implementations of these differential pairs require a designer to apply special techniques to address this small linear range. One technique, an effective reduction of the gate-channel voltage coupling term κ, is incorporated into the Floating Gate model. In this section, an additional technique called Ohmic biasing is presented to further increase the linear range.

Ohmic biasing is a novel technique to increase the linear range of a subthreshold-biased differential pair by moving the bias transistor out of saturation and into the Ohmic regime. This decreases the bias current to the differential pair for differential inputs about a specified input common mode. Unlike previously proposed techniques to extend linear range, this technique does not require high supply voltages [10], large device sizes [11], or above-threshold CMOS designs [12]. This technique also

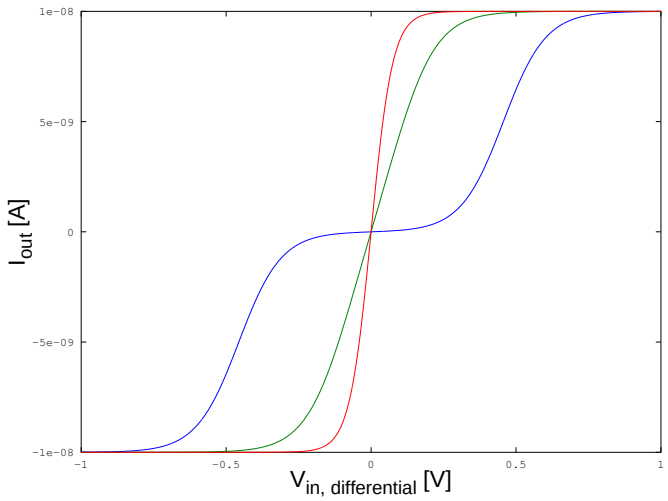


Figure 1.3: Output current of a differential pair as a function of differential input voltage for a circuit biased in three different operating regimes. Here, red trace corresponds to a normal differential pair with a saturated bias transistor, the green trace represents the expanded linear range and linearity of our technique, and the blue trace represents an "over-linearized" differential pair with an essentially "cutoff" bias transistor for small differential voltages.

introduces a novel design capability: to dynamically change the linear range of the differential pair at run-time. An illustration of this capability is shown in fig. 1.3

This chapter provides an analysis of the circuit, a design guide for biasing the circuit for a desired linear range, a comparison of the technique with other techniques, and an application to a Variable-Gain Amplifier.

1.3 Sub-banded Processing - Derivative Circuit

Finally, to provide an example of a standalone sub-banded processing element, Chapter 4 covers a new implementation of a temporal derivative circuit and sets forth three criteria for developing a continuous-time derivative operating in a noisy environment. This six transistor element processes a band-limited continuous-time

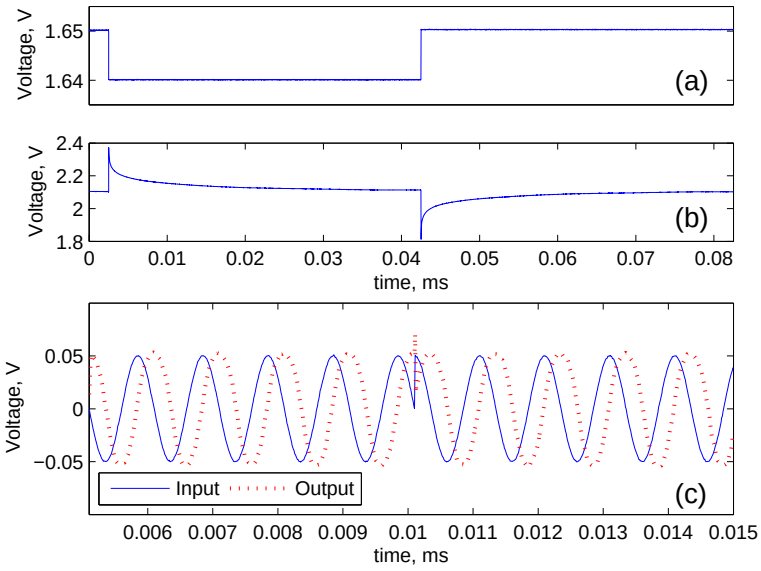


Figure 1.4: Transient Response of the derivative circuit. (a) A step input with small magnitude. (b) Transient response of the derivative circuit showing large signal changes for each discontinuity in (a). (c) A sine wave undergoing a hard transition to a cosine wave. Note the spike in the output voltage associated with the step function in the transition from sin(x) to cos(x).

derivative and can be programmed for use throughout the human audio spectrum, providing a method for onset detection and first-order feature extraction in addition to computing the standard analog mathematical function.

Due to the inherent properties of computing a wide-band derivative, developing a robust wide-band derivative circuit capable of operating on a signal with any amount of noise is not possible. The circuit presented here works on a band-limited signal to keep high-frequency noise from dominating the circuit output. As with the Ohmically-biased differential pair, a circuit level analysis is performed and experimental results are provided from a circuit fabricated in a 0.5μm CMOS process are provided (fig. 1.4).

1.4 Conclusion

Each of these novel advancements fills a role within a sub-banded filterbank platform. Issues regarding flexibility and reconfigurability of an analog system are addressed by the Floating-Gate model. An increase in dynamic range for filters and signal-processing elements under a variety of conditions is demonstrated with the newly-developed Floating-Gate model and the technique of Ohmic biasing. Finally, a sub-banded signal processing circuit operating on small signals (such as those compressed by the VGA structure proposed in Chapter 3) is presented, analyzed and demonstrated on a 0.5μm process. Each of these facets of an analog signal processing system work in concert to lower the power consumption of existing systems and leverage the inherent efficiency of analog signal processing.

Chapter 2

Floating Gate Modeling

2.1 Introduction

Floating-Gate (FG) transistors have been used for decades as a digital memory element in EEPROM and Flash memory devices. They consist of a standard CMOS transistor with an electrically isolated gate capacitively coupled to an external voltage as shown in Fig. 2.1. This isolated gate structure is considered to be "floating" with no DC path to ground for charge to move along. Thus, in the absence of charge modification through Fowler-Nordheim tunneling, hot-electron injection or UV photoinjection, the charge deposited on the gate at fabrication will remained "trapped" there for the life of the device.

Since there is no DC path to ground from the gate of the transistor, standard SPICE-based circuit simulators cannot model the operation of this device; the equations describing circuit behavior do not converge to a steady-state solution. This prohibits designers from simulating the performance of these devices and limits them to extrapolating the predicted operation of new structures from previously fabricated devices. More critically, the current field of simulation models for FG devices does not provide an instance capable of modeling charge modification for simulating programming processes at runtime. It is the goal of this research to provide a single, extensible

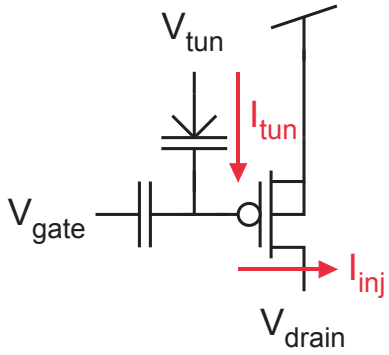


Figure 2.1: Electrical schematic of a Floating-Gate transistor implemented with a pFET device. Programming is accomplished by Fowler-Nordheim tunneling I_{tun} or hot-electron injection I_{inj}.

model of FG transistors that incorporates charge modification on the floating node for DC, AC and transient circuit operation.

2.2 Background

Digital designers have long used FG transistors as the digital storage element in FLASH memory, where the devices are modeled as either "on" or "off." As analog designers have adopted FG devices, several SPICE-compatible models have been presented to address the simulation of these structures through different means. Some provide DC operating points and enable circuit simulation through capacitive coupling and some provide models of programming processes, but none cover all aspects of charge modification and transistor operation. These processes are outlined in the sections below.

2.2.1 Fowler-Nordheim Tunneling

Fowler-Nordheim tunneling (known as "tunneling") is the process of moving electrons **off** of the floating node. Through this process, the electrons trapped on the FG move across the oxide of a specially designed capacitor. This tunneling capacitor

is made with a thin oxide, and is usually implemented with a varactor or MOSCAP structure. As the electric field across the oxide increases (up to the point of dielectric breakdown), the effective thickness of the oxide is decreased, and trapped electrons on the FG node gain enough energy to tunnel through the barrier via the Fowler-Nordheim process [13].

Producing the high electric field necessary to tunnel through a thin oxide typically requires a high voltage on the non-FG side of the device (e.g. 14V for a standard 0.5 μm process rated for 3.3V). Once the effective width of the oxide has been decreased enough to allow tunneling to take place, it can happen very rapidly and is difficult to control. For these reasons, tunneling is typically used as an "erase" function for programming FG devices.

2.2.2 Hot-Electron Injection

Hot-electron injection (known simply as "injection") is the process of moving electrons **on** to the floating node, in this case from the drain-to-source channel and across the gate oxide. This lowers the effective FG voltage and, in the case of a pMOS device, increases the amount of current through the channel of the transistor for a given V_{DS}.

Injection works via impact ionization in the channel under high-field conditions between the drain and source. As V_{DS} is increased past some process-dependent threshold (e.g. 5.5V for a standard 0.5 μm process rated for 3.3V), high energy collisions between free electrons and the lattice produce electrons via impact ionization with enough energy to tunnel across the gate oxide. This causes a very small (femto-Ampere) current to flow from the gate, across the oxide and into the drain of the transistor. In the case of the pFET device, this process continues until enough charge has moved across the oxide to bring the floating node voltage down closer to the drain voltage, thus decreasing the lateral field strength across the oxide and "turning

off" the injection process. Once V_{DS} has been lowered below the injection voltage, the trapped charge on the FG remains there indefinitely or until removed through tunneling.

2.2.3 Existing Simulation Models

The authors of [14, 15, 16] have proposed several simulation models for simple circuit operation with no charge modification. These models include capacitive coupling onto the FG node, and establish an initial FG charge to establish a DC operating point and permit circuit analysis. However, these techniques do not model charge modification, and thus cannot be used to depict the programming characteristics of FG transistors or associated feedback structures.

The authors of [17, 18] present methods of simulating charge modification, specifically through Fowler-Nordheim tunneling and hot-electron injection, depicted in Fig. 2.1. These processes allow accurate simulation of the programming characteristics of the FG transistor, but lack an implementation of capacitive coupling to enable DC or AC circuit simulation.

The model presented here addresses these deficiencies by modeling of both capacitive coupling onto the FG node and charge modification processes. This SPICE-compatible model properly captures capacitive coupling effects in DC sweeps and programming and adaptation effects in transient analyses. This model is also flexible and extensible - different mathematical models for calculating tunneling or injection currents may be used and any standard SPICE transistor model (BSIM, EKV, PSP, etc.) may be used to model basic circuit elements.

2.3 Description of the Model

To provide compatibility with SPICE simulators, we have developed the model of Fig. 2.2. This model is composed of two parts:

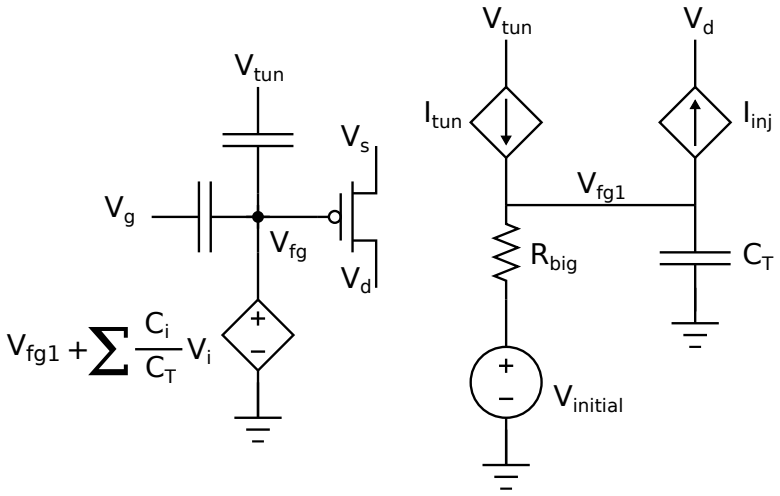


Figure 2.2: Schematic of the FG model, including charge modification current sources and the "dummy" integration node. The V-I equations of I_{tun} and I_{inj} are given in Equns. 2.2a and 2.2b, respectively.

1. The actual transistor and associated voltage sources to model capacitive coupling through the transistor.

2. An electrically isolated "dummy" node for charge storage and modification for run-time programming.

2.3.1 Capacitive Coupling and Transistor Modeling

In the first part the model, the capacitive inputs to the floating gate are modeled as voltage-controlled voltage sources (VCVS) connected in series between the floating node and a unity-gain voltage source connected to the dummy node. Each VCVS has a gain determined by the ratio of the size of the input capacitor to the total capacitance attached to the floating node. Mathematically, this structure can be represented as Eq. 2.1. This structure couples the inputs from multiple capacitors, both drawn and parasitic, onto the FG node while maintaining electrical isolation by nature of the infinite output impedance of the VCVSes. It also couples the contribution of

the charge modification section onto the floating gate node, once again providing electrical isolation.

$$V_{fg} = V_{dummy} + \sum_i \frac{C_i}{C_{total}} V_i \tag{2.1}$$

2.3.2 Charge Modification

The second part of the model includes the Voltage-Controlled Current Sources (VCCS) necessary to simulate charge modification processes, such as tunneling and injection, and the initial charge on the floating gate, set by R_{big}, C_T and $V_{initial}$.

Tunneling and Injection

Various models of these processes have been proposed (tunneling in [17] and injection in [17, 19]), and any of these models can be incorporated with either SPICE primitives or Verilog-A. The key to permitting DC convergence is to nest the expressions for I_{tun} and I_{inj} within an IF() statement that ensures no current flows when programming is not "turned on." Example flow-control SPICE expressions are given in eqns. 2.2.

$$\text{Gtun Vfg1 0 value=IF}((V(V_{tun}) - V(V_{fg})) > V_{dd},\ \textit{Tun Expr}, 0) \tag{2.2a}$$

$$\text{Ginj Vfg1 0 value=IF}((V(V_s) - V(V_d)) > V_{dd},\ \textit{Inj Expr}, 0) \tag{2.2b}$$

The tunneling and injection models used in this implementation of the FG architecture are taken from [17] and given in Eq. 2.3a and 2.3b, respectively. Each of the parameters here is fit from empirical data. Implementation details are found in [8], but are beyond the scope of this research.

$$I_{tun} = -I_{tun0} W L e^{\frac{-V_f}{V_{ox}}} \tag{2.3a}$$

$$I_{inj} = \alpha I_s e^{-\frac{\beta}{(V_{gd}+\delta)^2}+\lambda V_{sd}} \tag{2.3b}$$

Initial Charge

As mentioned, the initial charge on the floating gate is set by $V_{initial}$. Since it is connected to the dummy node by R_{big} (the largest resistor size available in the simulator, typically on the order of TΩ), any charge integrated across C_T from the current sources is effectively disconnected from $V_{initial}$. This results in a very, very small leakage current through R_{big}, but the RC time constant of the circuit is on the order of days or weeks for standard MOS capacitances and thus negligible for most transient simulations.

2.4 Testing and Evaluation

To test the model over DC, AC and Transient analyses, three different test structures were used. The first is a standard PMOS Floating Gate transistor, where gate sweeps are used to show the effect of changing the initial charge on the floating gate. A drain sweep of this device is also provided to illustrate the capacitive coupling difference between a non-FG device and the FG cell. An Operational Transconductance Amplifier (OTA) with Multiple-Input Floating-Gate (MIFG) transistors was fabricated and is used to illustrate the effects of capacitive division on the linear range of the OTA through further DC analysis. Finally, a Gm-C lowpass filter was fabricated with an indirectly programmed bias transistor to illustrate an adjustable corner frequency through AC analyses and programming to a target through transient analyses.

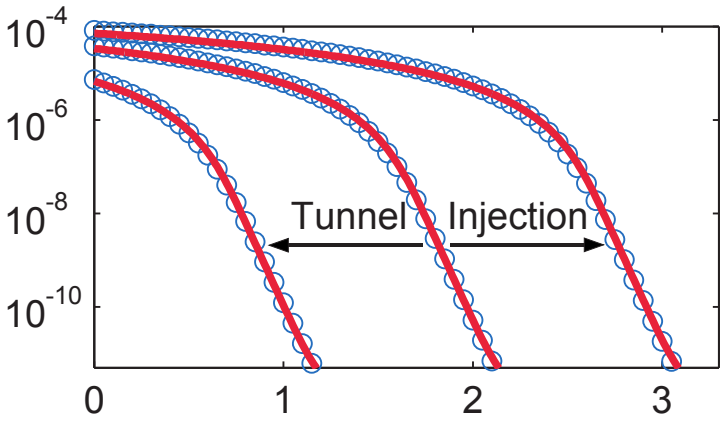


Figure 2.3: Gate sweeps of a Floating-gate PMOS transistor at three different levels of charge on the FG node. The charge was modified in between sweeps by tunneling and injection processes. Simulation data is shown in red, testing data is shown as blue circles.

2.4.1 DC Analyses

Changes to Transistor Operation

Perhaps the most noticeable effect of using a FG transistor is the effect of changing the initial charge on the FG node. Fig. 2.3 illustrates the effect of changing this effective voltage through tunneling and injection.

In a FG transistor, capacitive coupling between the drain and the FG node drastically decreases the Early voltage (and the effective resistance) of the device. This effect is demonstrated in Fig. 2.4, where a standard PMOS transistor is shown along with the FG transistor.

Capacitively Divided OTA

One common variant of the standard FG transistor studied previously is the Multiple Input Floating-Gate (MIFG) transistor shown in Fig. 2.5. These transistors are functionally equivalent to a single input FG transistor when all of the inputs are

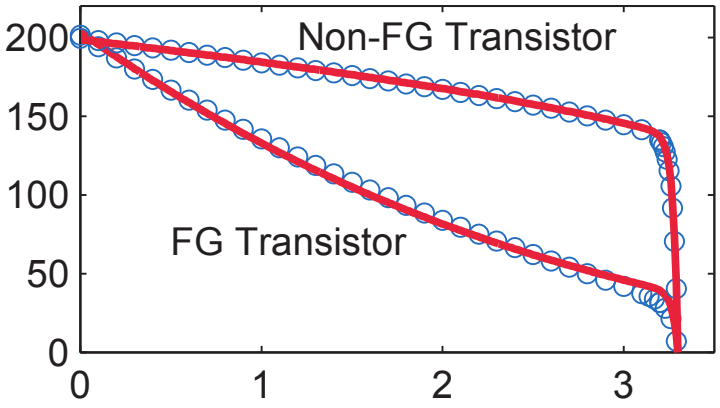


Figure 2.4: Drain sweeps of FG and standard PMOS transistors. The capacitive coupling between the gate and the drain serves to dramatically decrease the Early voltage of the device. Simulation data is shown in red, testing data is shown as blue circles.

tied to a common voltage - there is no impedance division and the input voltage is coupled directly to the FG node. If the inputs are held to different voltages, the capacitors form an impedance-based voltage divider and the FG node voltage is a weighted summation of the inputs.

One novel application of MIFG transistors is to increase the linear range of Operational Transconductance Amplifiers (OTAs). This is done by decreasing the effective κ of the input transistors through capacitive division, where the linear range of the OTA (V_L) is approximated by Eq. 2.4. The inputs of an OTAs with various capacitive input ratios were swept symmetrically and the resulting transconductance curves are shown in Fig. 2.6

$$V_L = \frac{C_{total}}{C_{in}} \frac{2U_T}{\kappa} \tag{2.4}$$

Another useful application of floating gate inputs to OTAs is offset removal through programming. All OTAs suffer a small amount of DC offset in their output; for a differential input of zero volts, there is usually a small amount of current that will

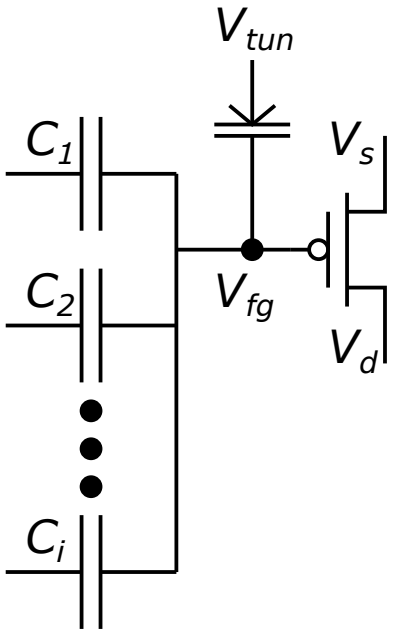


Figure 2.5: Multiple-Input Floating Gate (MIFG) transistor. During programming all inputs are tied to a common voltage for faster tunneling / injection. At run-time the inputs are moved independently to modify the FG voltage through impedance division.

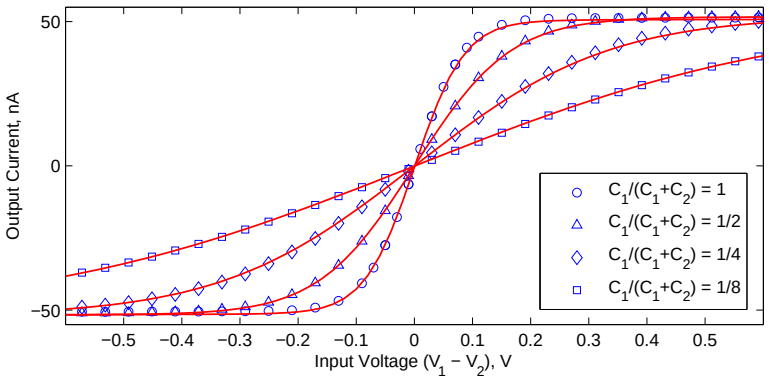


Figure 2.6: V-I curves of an OTA with capacitively divided inputs. As the capacitance division ratio is decreased from 1 to 1/8, the effective κ of the input pair decreases and the linear range increases.

flow through the output. By changing the amount of charge on the FG node of each input transistor, this offset can be removed to yield a current of zero Amperes for a differential input of zero volts.

2.4.2 AC and Transient Analyses

To demonstrate operation in the AC and Transient domains, a simple Gm-C low-pass filter was fabricated. With a FG transistor providing the bias current, this filter is programmably tunable over the filter's operational range via run-time programming via an Indirectly Programmed Floating-Gate (IPFG) transistor. The details of this structure are discussed, as well as the benefits of using this topology to demonstrate the minutiae of the model.

Indirectly Programmed Floating-Gate Transistors

An Indirectly Programmed Floating-Gate (IPFG) transistor consists of a regular FG structure with not one, but two or more transistor gates connected to the FG node. When possible, IPFG structures are very useful in large-scale layouts due to two factors:

1. IPFG transistors enable run-time programming of a processing element. Using direct programming requires the floating gate transistor to be isolated from its processing circuit and switched to the programming infrastructure. IPFGs remain connected to the programming lines at all times, enabling a user to change the stored charge on the FG node at any time during normal circuit operation.

2. Simplified programming infrastructure: T-gates and switching structures are not necessary to program each FG in a large-scale layout.

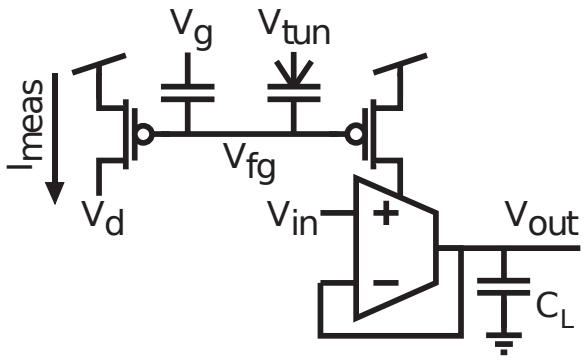


Figure 2.7: A schematic of the Gm-C filter as fabricated.

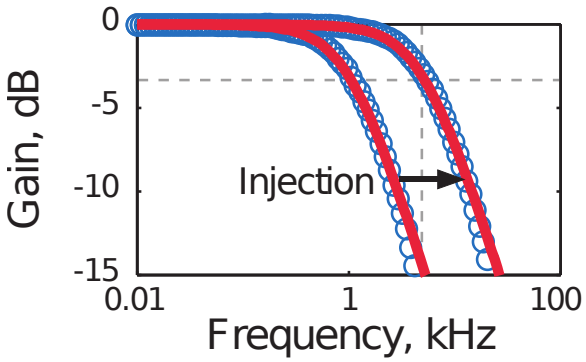


Figure 2.8: Filter gain characteristics measured after programming to target corner frequencies.

Run-Time Programming of a Gm-C filter

To test the accuracy of the injection model, the Gm-C filter of Fig. 2.7 was fabricated on a standard 0.5μm process. The charge on the floating gate was removed via tunneling, then replaced in phases targeted to increase the filter's corner frequency one decade via injection. After each phase, the AC response of the circuit was taken and compared to the simulation model. The results of these experiments are shown in Fig. 2.8 and demonstrate the close correlation between theory and practice.

To further illustrate the flexibility of the model, very short (1 ms) injection pulses were used to program the filter from a starting corner frequency of 1 kHz up to 5 kHz. After each pulse, the filter's gain at 5 kHz was measured. This experiment

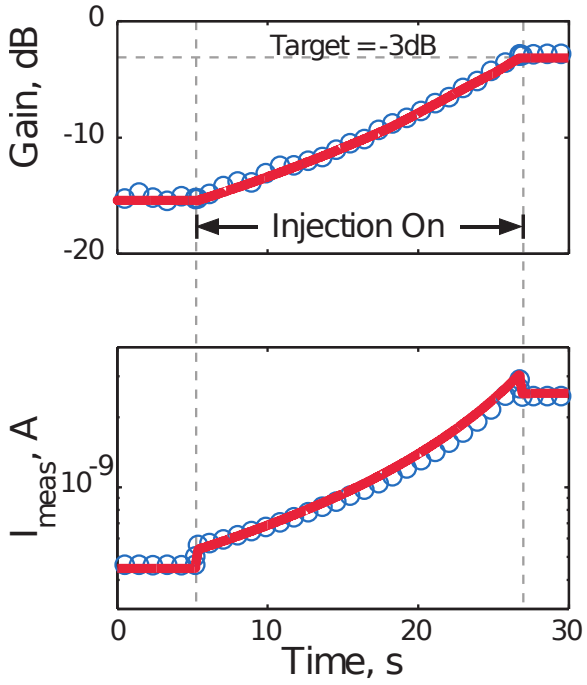


Figure 2.9: Filter gain and associated bias current during the programming process. As injection begins by lowering V_D, the capacitive coupling via the parasitic C_{gd} can be seen to increase the filter gain by decreasing V_{FG} and increasing the amount of bias current. After a period of time the target gain of -3 dB is reached and injection ceases. Once again, the effect of coupling via C_{gd} is seen to affect the filter gain.

is illustrated in Fig. 2.9. As shown in the figure, the simulation model tracks the transient profile of the fabricated circuit with very little deviation.

2.5 Conclusion and Further Work

The usefulness of the model architecture described here has been shown in DC, AC and transient simulations with real-world implementations. Though the injection model used works only in the subthreshold domain, the results of these experiments confirm good correlation between the chosen mathematical models and fabricated circuits within that domain.

Future work is planned to implement other injection models to properly demonstrate injection with higher currents. This would greatly expand the usefulness of the model, since subthreshold programming is limiting when capacitor sizes are large or low-impedance loads must be driven. Further improvements to the model may also come from a more streamlined characterization process for existing floating-gate structures and investigation of the effects of changing process parameters.

Chapter 3

Ohmic Biasing of Sub-V_T Differential Pairs

3.1 Introduction

A common task in analog signal processing is to convert a voltage-mode signal to a current-mode signal. This can be done through a resistor or single transistor, but input impedance, output impedance, and local feedback issues make these unattractive options. Instead, designers turn to the differential pair circuit of fig. 3.1. This three transistor circuit provides output currents linearly proportional to the input voltages. In this chapter, the differential pair is used to convert a difference between the input voltages $(V_1 - V_2)$ to a single output current, $I_{out} = (I_1 - I_2)$. In this case, the term that relates the amount of output current to a given differential voltage is G_m, such that $I_{out} = G_m(V_1 - V_2)$. This variable is called *transconductance*, but it is only constant over a small range of differential input voltages. Once an input moves outside that range, the differential pair is said to "saturate" - a further change in voltage produces no change in current and the circuit is rendered useless. The range of inputs where the amplifier output is determined by the transconductance is called the *linear range*, and this chapter describes a novel technique to expand it.

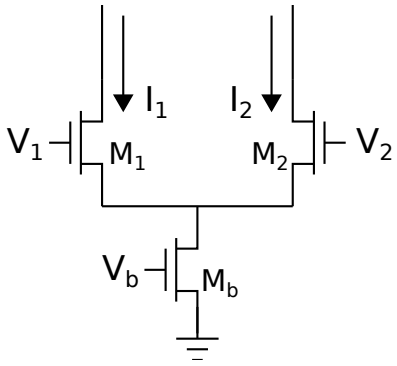


Figure 3.1: Standard nFET based differential pair.

Low-power circuits utilizing differential pairs biased in the subthreshold (sub-V_T) region have always suffered from a fixed linear range limited to ∼74mV when compared to those utilizing above-threshold bias currents (∼74mV and up). Several techniques have been proposed to increase this linear range, but all require design trade-offs with high power dissipation, high supply voltages [10], large device sizes [11], or above-threshold CMOS designs [12]. In our array-based processing application, high power consumption is undesirable and design size must be kept to a minimum. The method described here fulfills these requirements: it simultaneously increases the input linear range of a differential pair and reduces distortion within the linear range. This work provides an analysis of the proposed circuit, along with data from simulation and circuits fabricated in a standard 0.5μm process. Finally, an application of the technique to a Variable-Gain Amplifier (VGA) architecture is proposed.

3.2 Background

The size and power specifications of array processing elements lend themselves to a design built around subthreshold-biased differential pair structures (fig. 3.1). However, differential pairs biased in the subthreshold domain have a input linear range restricted to a few tens of millivolts, as shown analytically in their voltage -

current transfer function in eq. 3.1 [20].

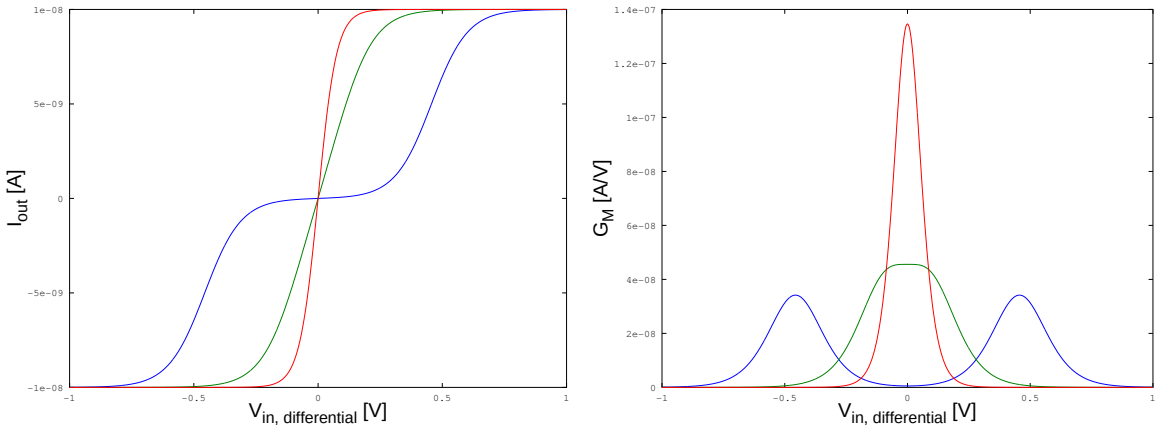
$$I_{out} = I_1 - I_2 = I_b \tanh\left(\frac{\kappa}{2U_T}\Delta V_{in}\right) \tag{3.1}$$

From this equation, we can see that the linear range is fixed and depends on the thermal voltage U_T and the gate-channel coupling term κ. Neither of these are traditionally regarded as designable parameters, but in this chapter we use a combination of techniques to create a wide linear range subthreshold OTA with a designable linear range.

Various techniques for increasing the linear range of a subthreshold differential pair have been proposed. Bump linearization [10, 11, 21, 22, 23] has long been popular and works by "stealing" current from the center node of the differential pair at small differential input voltages, thereby reducing the current flowing through each input transistor and extending the linear range. This technique changes the transfer function of the differential pair from eq. 3.1 to the form of eq. 3.2 [10], where the linear range is now dependent on w, the size of the central "bump" transistors relative to the size of the input transistors. The authors of [10] found that for the $\lim_{w\to 0}$, there was no appreciable increase in the linear range. However, the value $w = 2$ provides maximal linear range, while values larger than 2 introduce undesirable non-linearities in the transfer curve. Therefore, while this technique is extremely useful, it increases the size of the design with the addition of extra transistors and the linear range is fixed at fabrication through the designer's choice of w.

$$I_{out} = I_1 - I_2 = \frac{\sinh\left(\frac{\kappa}{U_T}(V_1 - V_2)\right)}{1 + \frac{w}{2} + \cosh\left(\frac{\kappa}{U_T}(V_1 - V_2)\right)} \tag{3.2}$$

Another technique for increasing the linear range is a reduction of the gate-channel coupling coefficient, κ. This has been explored in [10, 11, 24, 25, 26]. In [10], the researchers decreased κ by using a well-input transistor configured for "gate degeneration." In [11, 24, 26], the decrease in the effective κ was accomplished through



(a) Output Current (b) Transconductance

Figure 3.2: Output current and corresponding transconductance curves as a function of differential input voltage for an OTA biased in three different operating regimes. Here, red traces correspond to a normal OTA with a saturated bias transistor, the green traces represent the expanded linear range and linearity of our technique, and the blue traces represent an "over-linearized" OTA with an essentially "cutoff" bias transistor for small differential voltages.

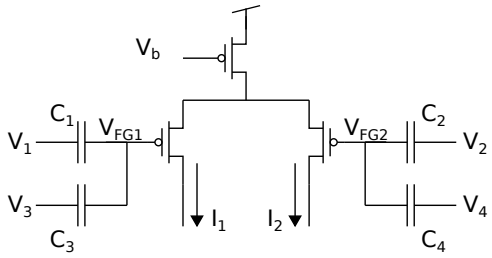


Figure 3.3: Capacitively divided differential pair. Note that there is no DC signal path to the gates of the input transistors.

capacitive division of the inputs, as shown in fig. 3.3.

In every case, capacitive division works well to increase linear range, but fabricating matched capacitors carries a high area cost; in many cases, the input capacitors of an OTA can consume more chip area that the rest of the design. The area cost of using well-input transistors to decrease κ is much smaller, but the trade-off comes with the restricted range of input voltages that will prevent forward-biasing the PN-junction of the drain or source to the well. Source degeneration presents a similar issue. For these reasons, we will avoid the use of well-inputs or source degeneration and use small capacitive division ratios when necessary to achieve a large linear range.

The technique covered in this chapter utilizes concepts from both bump linearization and capacitive division; critically, this technique does not require the extra transistors of bump linearization and uses small capacitive division ratios, decreasing design sizes from techniques previously proposed. It requires the designer to build a circuit biased around an arbitrary input common-mode voltage (ICMV) to a differential pair. This ICMV, combined with a programmed charge on the gate Floating-Gate transistor, pushes the bias transistor out of saturation and into the Ohmic region of operation. This produces an effect parallel to the "current stealing" of the bump linearization technique, but rather than draw current off of the common node, it limits the amount of current available to the differential pair for small differential voltages.

3.3 Analysis

In this section, we will show increases in both the linear range and linearity of the amplification of a differential pair. Additionally, techniques to compensate for changes in the ICMV are presented. The initial analysis is performed on a nFET based differential pair, then shown applied to a pFET based OTA with floating-gate inputs.

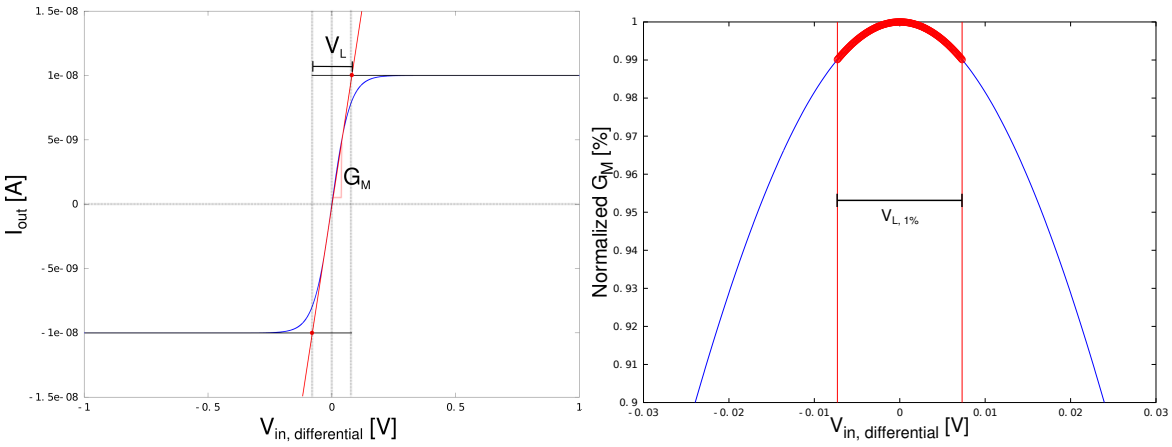
3.3.1 Performance metrics

We will define two variables, V_L and $V_{5\%}$, to describe the transconductance characteristics of a differential pair. V_L is a convenient analytical expression from [10] and given in eq. 3.3, loosely describing the linear range while neglecting the linearity of the amplification region. $V_{5\%}$ describes both the linear range and linearity by normalizing the transconductance curve and finding the differential input voltage at which the curve deviates 5% from the maximum value. Both of these quantities are illustrated in fig. 3.4.

$$V_L = 2\frac{I_{bias}}{G_m} \tag{3.3}$$

3.3.2 nFET Based Differential Pair

We will begin by using the standard differential pair (fig. 3.1) for analysis using the EKV model proposed in [27]. The standard analysis of this circuit assumes transistor M_b remains in saturation for all values of V_1 and V_2. However, this is not an entirely valid assumption. As the common mode voltage (V_{CM}) moves close to the source voltage V_S, the channel current of M_1, M_2 (denoted as I_t) decreases. Since the gate voltage of M_b is held constant, this decrease in available current forces M_b into the Ohmic region of operation. When M_b is biased in the Ohmic region, the original



(a) Output Current (b) Transconductance

Figure 3.4: Plots of the metrics used to determine performance gains in a differential pair. The analytical quantity V_L is illustrated in 3.4a, while $V_{5\%}$ is shown in 3.4b.

saturated bias current ($I_{b,sat}$) is decreased for a given gate voltage V_b, decreasing the transconductance of the differential pair. As shown in eq. 3.3, a decrease in G_m corresponds to an increase in the linear range as described by V_L.

As we move V_{CM} closer to V_S, this trend continues to a point where the G_m curve becomes maximally flat for small differential inputs. However, as we move past this point, distortion in the V-I transfer curve increases dramatically as V_{DS,M_b} becomes very close to zero, the bias current is cut off and the output current is pushed to 0 A. The effect of this decrease in V_{DS,M_b} is shown in the blue traces of fig. 3.2.

The output current as a function of differential input voltage is derived in Appendix B and presented in eq. 3.4. This function was evaluated against multiple test circuits with transistors of varying width and found to be accurate across all cases. As a corollary to eq. 3.4, the transconductance of an ohmically biased differential pair is presented in eq. 3.5.

$$I_{out} = \frac{\sinh \frac{\kappa \Delta V_{in}}{2U_T}}{\frac{I_{b,sat}}{2I_t} + \cosh \frac{\kappa \Delta V_{in}}{2U_T}} I_{b,sat} \tag{3.4}$$

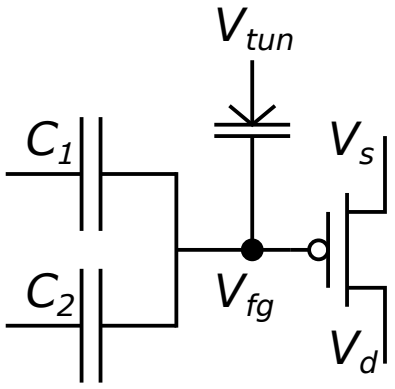


Figure 3.5: Multiple Input Floating Gate PMOS Device.

$$G_m = \frac{\kappa I_{b,sat} I_t \left(I_{b,sat} \cosh\left(\frac{\kappa \Delta V_{in}}{2U_T}\right) + 2I_t \right)}{U_T \left(I_{b,sat} + \cosh\left(\frac{\kappa \Delta V_{in}}{2U_T}\right) 2I_t \right)^2} \tag{3.5}$$

$$V_1 = V_{in,CM} + \Delta Vin$$

$$V_2 = V_{in,CM} - \Delta Vin$$

$$I_{b,sat} = I_0 e^{\frac{\kappa}{U_T} V_b}$$

$$I_t = I_0 e^{\frac{\kappa}{U_T} V_{in,CM}}$$

Appendix B also shows that eq. 3.4 achieves maximum linearity and linear range when $\frac{I_{b,sat}}{2I_t} = 2$. Therefore, the biasing condition for maximum linearity is

$$I_t = \frac{I_{b,sat}}{4} \tag{3.7}$$

3.3.3 Ohmic Biasing With Capacitive Division

Utilizing capacitive division on the inputs of an ohmically biased differential pair serves three purposes. 1) It decreases the effective κ of the input transistors, further extending the linear range. 2) It introduces an intermediate variable, V_{fg} to the

	Saturated bias	Ohmic bias	Capacitive Division with Ohmic Bias
V_L	$\frac{2U_T}{\kappa}$	$\frac{6U_T}{\kappa}\big\vert_{I_{b,sat}=4I_t}$	$\frac{C_{total}}{C_{in}}\frac{6U_T}{\kappa}\big\vert_{I_{b,sat}=4I_t}$

Table 3.1: Expressions of V_L and $V_{5\%}$ utilizing three different strategies for implementing a differential pair.

analysis of the differential pair, allowing the "target" current through the input transistors (I_t) to be set to desired value for an arbitrary input common mode voltage. 3) It enables compensation for changes in the input common mode at runtime through an additional capacitive input.

Decrease in effective κ

As shown in [28], a two input floating gate transistor has a floating-gate voltage described by eq 3.8, where $C_{total} = C_1 + C_2$. Since the input, V_1, is decreased by a factor of $\frac{C_1}{C_{total}}$, the effective gate-channel coupling coefficient, κ_{eff}, is given by eq. 3.9.

$$V_{fg} = V_{fg,initial} + \frac{C_1}{C_{total}}V_1 + \frac{C_2}{C_{total}}V_2 \tag{3.8}$$

$$\kappa_{eff} = \frac{C_1}{C_{total}}\kappa \tag{3.9}$$

This new value of κ_{eff} results in a new expression for V_L, which is compared in Table 3.1 to alternative methods of implementing the differential pair. As shown in the table, Ohmic biasing at the point of maximum linear range results in a three-fold increase in V_L. With the addition of capacitive division, this increase can be expanded, constrained only by the size constraints of a given design.

Programming for a given common mode voltage

Assuming V_2 is held constant, the expression for V_{CM} becomes eq. 3.10. By programming the initial voltage on the floating-gate, the common mode may be set at a desired voltage for an arbitrary signal common mode given by $V_{in,CM}$. One of the other implications of eq. 3.10 is that any change in the input common mode can be offset by an opposite change in V_2 scaled by $C_2/(C_1 + C_2)$.

$$V_{CM} = \frac{C_1}{C_{total}} V_{in,CM} + V_{FG,initial} + \frac{C_2}{C_{total}} V_2 \tag{3.10}$$

3.3.4 Testing and Evaluation

To check the accuracy of equations 3.4 and 3.5, they were compared to data taken from a differential pair with floating-gate inputs fabricated on a standard 0.5μm CMOS process. Of chief concern was to investigate the optimum bias point, $I_b = 4I_t$. For these trials, I_b was held constant at 8nA and the floating-gate transistors were programmed to have a current $I_t = 2$nA for $V_{CM} = 2$V. V_CM was swept to produce a family of curves illustrating the increase in both linearity and linear range. For the sake of clarity, the results presented here in figures 3.6 and 3.7 do not include over-linearized test cases.

The transistor parameters of the on-chip differential pair were inserted into eq. 3.5 and the results were overlaid onto fig. 3.7 to produce fig. 3.8. As shown in the figure, the analytical solution for the behavior of the differential pair closely matches the operation of the fabricated circuit.

Finally, the parameters V_L and $V_{5\%}$ were calculated for the families of curves in fig. 3.8. As shown in fig. 3.9, the theoretical results were found to trend closely with the measured performance of the chip. The results of these tests show that the analytical model proposed here has close correlation to the performance of an actual differential pair, leading to the further design work in section 3.4.

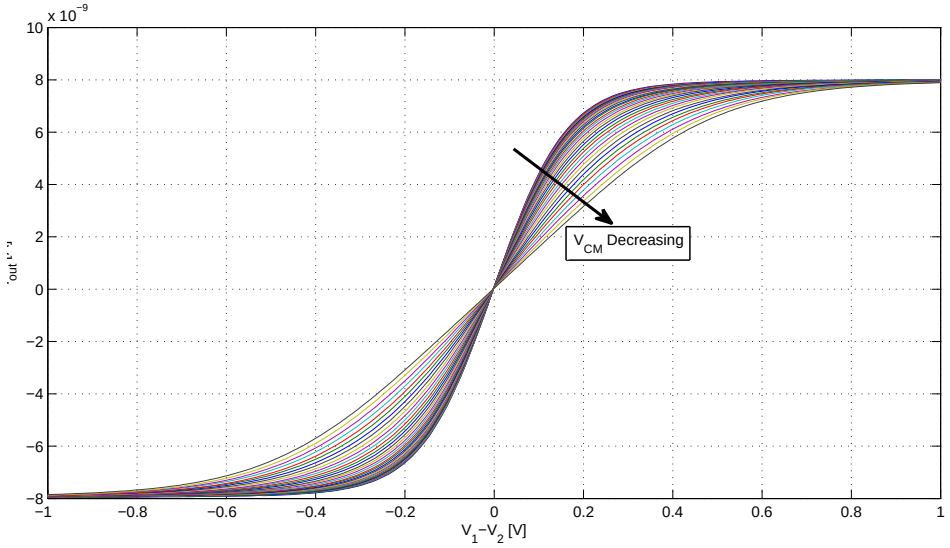


Figure 3.6: Output current of a differential pair as V_{CM} is swept to decrease I_t, demonstrating an increase in linear range. Only voltages within [-1 1] are shown for the sake of clarity.

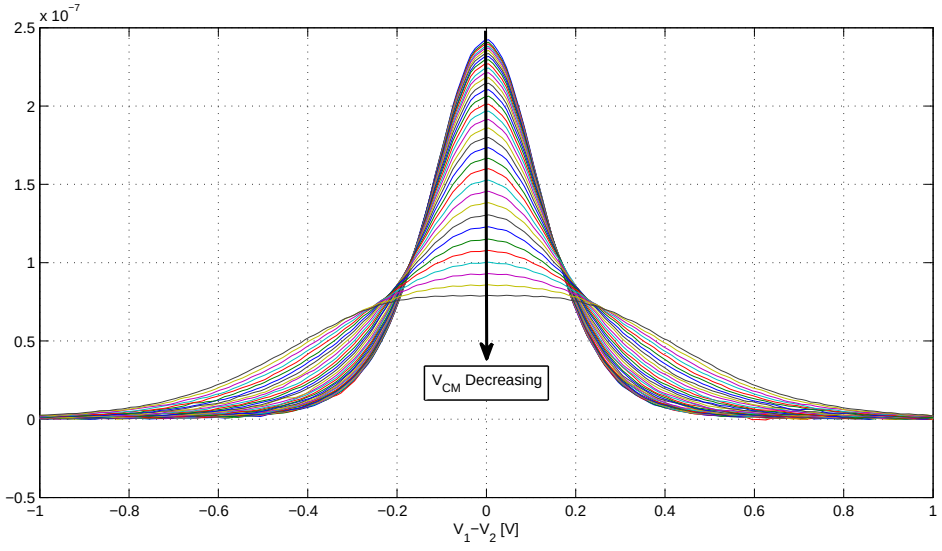


Figure 3.7: Transconductance of a differential pair as V_{CM} is swept to decrease I_t, demonstrating not only an increase in linear range, but also an increase in linearity.

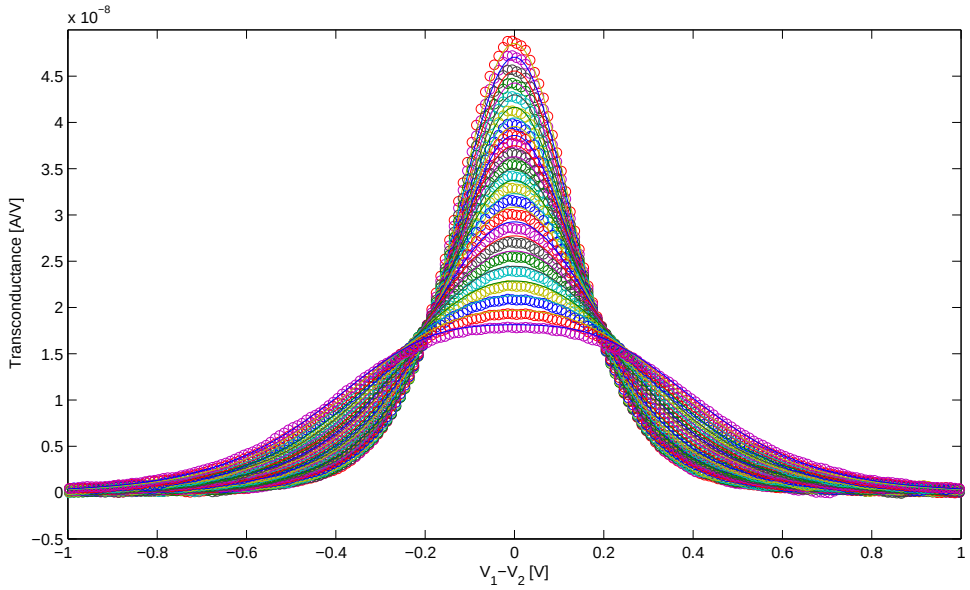
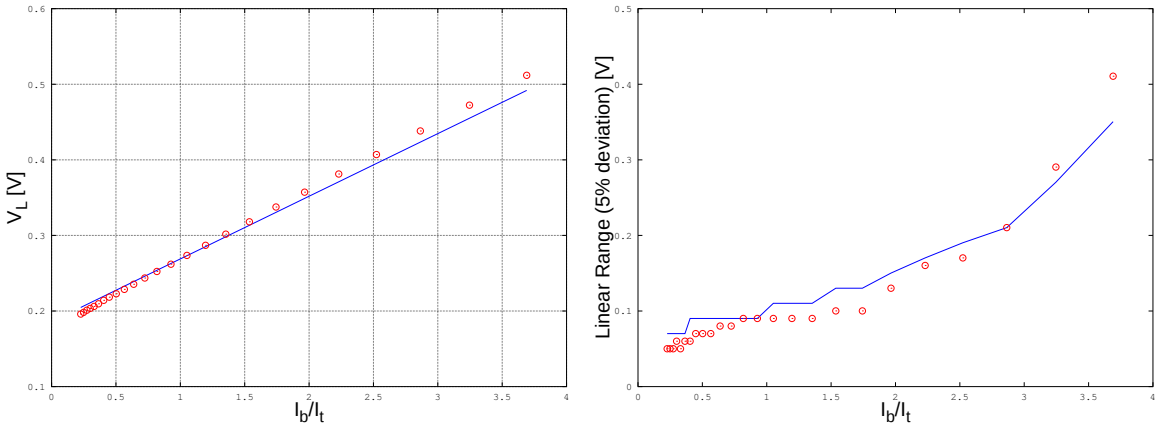


Figure 3.8: Transconductance of a differential pair as V_{CM} is swept to decrease I_t. Solid lines are analytical curves, while open circles indicate measured transconductance from the fabricated circuit.



(a) V_L (b) 5% Linear Range

Figure 3.9: 3.9a Plot of V_L as a function of the bias current (I_b) divided by the "target" current (I_t). 3.9b Plot of the actual linear range, $V_{5\%}$, also defined as a function of I_b/I_t. Theoretical data is shown as a solid line, while experimental data is shown with open circles.

3.3.5 Comparison With Standard Differential Pair

To analyze the increase in linearity at the maximum bias point described by eq. 3.7, a Taylor series expansion was performed to observe the high-order polynomial contributions to the output current for both a standard differential pair with a saturated bias transistor and a differential pair with an ohmic bias. The standard differential pair (eq. 3.11) has over a third of the signal in third-order or higher contributions. In comparison, the ohmically biased differential pair (eq. 3.12), has *no* third-order contributions and has much smaller higher-order coefficients.

$$\tanh x = x - \frac{x^3}{3} + \frac{2x^5}{150} - \frac{17x^7}{315} + \frac{62x^9}{2835} \tag{3.11}$$

$$\frac{\sinh x}{2 + \cosh x} = \frac{x}{3} - \frac{x^5}{540} + \frac{x^7}{4563} - \frac{x^9}{77760} \tag{3.12}$$

3.3.6 Drawbacks to Ohmic biasing

Maintaining this increased linear range requires $I_t = 4I_{b,sat}$ for all operating conditions. However, since the target current, I_t, is a function of the input common mode, any changes in this parameter must be compensated by either a change in I_b, or by feedback applied to the input to keep I_t constant. On initial inspection, changing I_b would appear to be the easiest method to maintain maximum linear range. However, most multi-stage amplifiers suffer large changes in performance if the supply currents are changed dynamically with a signal. For this reason, input feedback implemented through Floating-Gate transistors to maintain a constant I_t is presented here.

3.4 Application

To demonstrate the merits of this technique, a Variable Gain Amplifier (VGA) was designed utilizing two OTAs (fig. 3.10). This circuit is ideal for demonstrating

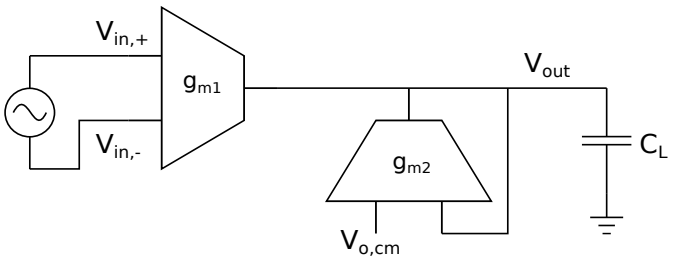


Figure 3.10: General schematic of a Differential-input, single-ended output VGA.

our technique due to the open-loop operation of the input OTA, labeled g_{m1}. This open-loop operation imposes a severe limitation on the input linear range for typical subthreshold CMOS OTAs. However, as we show here, applying an Ohmic bias can increase the input linear range dramatically.

As another consequence of this open-loop operation, a high voltage-mode gain is required to keep the output common mode fixed and decrease the steady-state output error. Therefore, a folded-cascode OTA design was developed with capacitively divided inputs and an ohmic bias on the input OTA (fig. 3.12).

3.4.1 Theory of Operation

This particular VGA topology implements both a VGA and a Low-Pass Filter (LPF). Passband gain a_v is determined by the ratio of transconductances, where

$$a_v = \frac{g_{m1}}{g_{m2}} \tag{3.13}$$

The corner frequency, f_c of the LPF is set by the load capacitance and g_{m2} such that

$$f_c = \frac{g_{m2}}{C_L} \tag{3.14}$$

3.4.2 Application of Ohmic Biasing

Due to the open loop nature of its operation and the differential-mode input signal, if the input common mode is known *a priori*, the input OTA can be programmed to operate with an Ohmically-biased tail transistor. However, the single-ended nature of the output signal requires the transimpedance element, g_{m2} to operate over a relatively large input common mode range. With this in mind, the second OTA has been fabricated with a large capacitive division ratio ($C_1/C_2 = 1/8$) to maintain a large linear range over all output common modes.

3.4.3 Evaluation

To evaluate the VGA, the maximum input amplitude of the differential signal that produced 1% THD was found under three test scenarios. 1) Both OTAs operated with saturated bias transistors and with no capacitive division. 2) Capacitive division was applied to both OTAs and the biases were kept in the saturation region. 3) The bias of the input OTA was programmed to operate in the Ohmic region with capacitive division, while the transimpedance OTA utilized a larger capacitive division ratio to increase linear range with a saturated bias transistor. The results of simulations of these three scenarios are shown in fig. 3.11.

To maintain a constant corner frequency and passband gain, the bias currents of each OTA had to be compensated for the addition of each linearization technique. Eq. 3.15 describes the biases of each OTA for the three test cases for a desired transconductance.

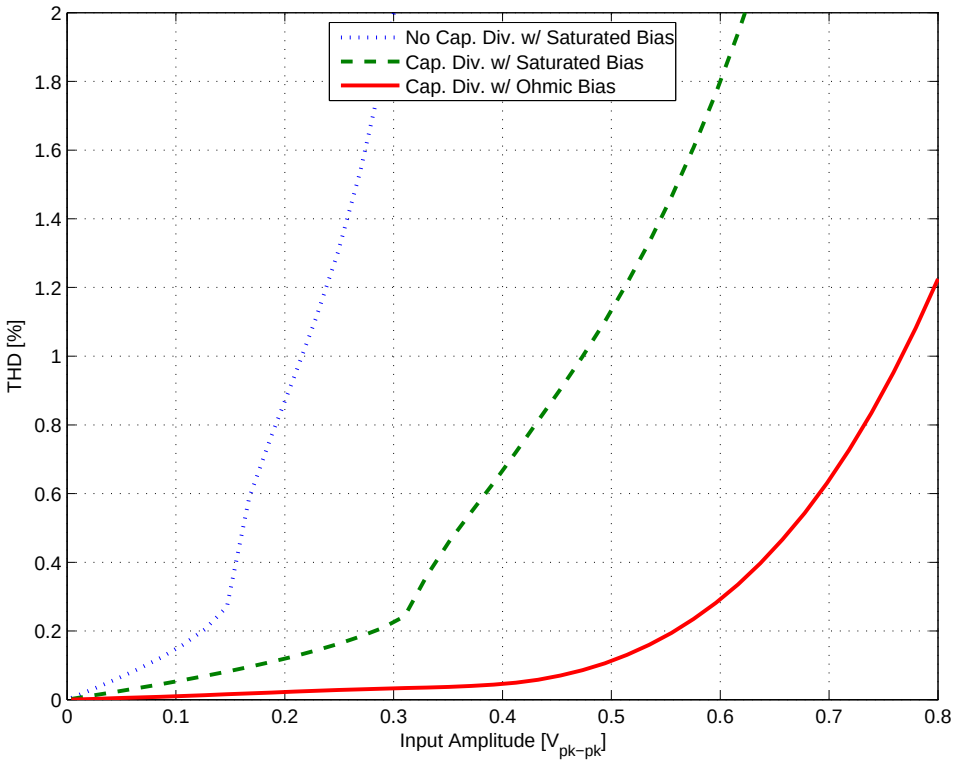


Figure 3.11: THD of the simulated VGA as a function of input amplitude. In this example, VGA gain was 0dB.

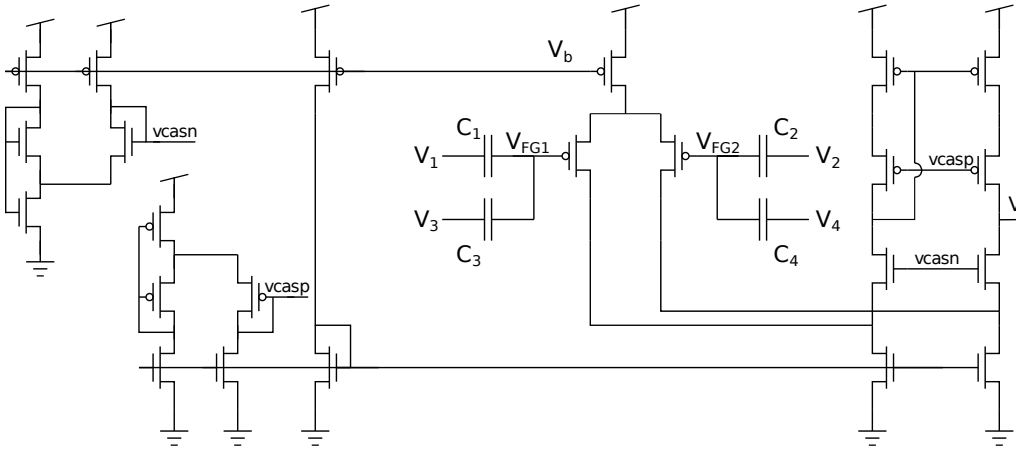


Figure 3.12: Simplified high gain OTA used in fig. 3.10. Floating Gate programming infrastructure has been omitted for the sake of clarity.

$$I_b|_{saturation} = \frac{2U_T}{\kappa} G_m \tag{3.15a}$$

$$I_b|_{capdiv} = \frac{2U_T}{\kappa_{eff}} G_m \tag{3.15b}$$

$$I_b|_{ohmic,capdiv} = \frac{6U_T}{\kappa_{eff}} G_m \tag{3.15c}$$

$$I_t|_{ohmic,capdiv} = \frac{3U_T}{2\kappa_{eff}} G_m \tag{3.15d}$$

3.4.4 Simulation Analysis

As shown in fig. 3.11, the maximum linear range under ohmic bias improves to 770 mV, an increase 350% from the baseline linear range of 220 mV. The results are summarized in table 3.2.

Unfortunately, using THD to measure linear range does not permit for a comparison with the previously discussed values of V_L or $V_{5\%}$. However, this simulation shows the merit of the Ohmic biasing method to increase the linear range of a low-power, array-based processing element.

Test Condition	$V_{in,pk-pk}$
Saturated bias	220 mV
Capacitive Division with Saturated Bias	470 mV
Capacitive Division with Ohmic Bias	770 mV

Table 3.2: Comparison of the 1% THD threshold for the simulated VGA structure

3.5 Conclusions and Future Work

We have presented an analysis of a novel technique to not only extend the linear range of a differential pair, but also increase the linearity within that range. Additionally, we have provided a practical method of input common-mode compensation and provided simulation results demonstrating an increase in the input common-mode range through our Ohmic biasing technique.

Immediate improvement in this linearization technique would come from the development of an efficient input common mode detection circuit and feedback mechanism to maintain a constant I_t despite a changing input common mode. This circuit would have to be compact, efficient and stable across the operational frequencies of the amplifier.

Additionally, this technique is directly applicable to fully differential OTA design. However, the implementation of an output common-mode feedback circuit was beyond the scope of this work. Realization of a fully differential OTA with common-mode feedback and appropriate compensation to maintain a constant I_t would greatly benefit the community.

Chapter 4

Derivative Circuit

4.1 Introduction

Determining how a signal pattern varies over time is important for many perceptual and sensory processing applications. For example, temporal derivatives are used for motion detection within pixel arrays [29, 30] and for speech processing on sub-banded audio signals [3, 31]. The circuit proposed here is compact, consumes little power, and is suitable for inter- and intra-pixel processing in addition to in-band processing within an array of sub-banded signals.

The design of analog derivative circuits has long been an unexplored field; the concept of an analog derivative is simple, but the implementation of a real-world system that can process derivatives in the presence of noise has proven difficult. A perfect wideband derivative will have an infinitely large gain for arbitrarily high frequencies (i.e. the faster the signal changes, the larger the derivative), pushing the effective SNR to 0.

To overcome this fundamental difficulty, the circuit proposed here operates in the context of a band-limited system. This accomplishes two goals: (1) it attenuates the effects of high-frequency noise and (2) it provides a method of "normalizing" the derivative with respect to a frequency of interest. Both of these unique characteristics

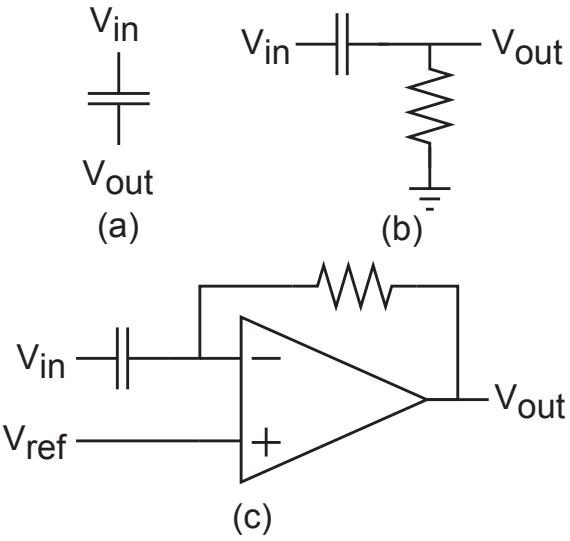


Figure 4.1: Previously proposed methods of computing a continuous-time derivative. (a) Current through a capacitor, (b) R-C voltage-mode passive differentiator, (c) Clamped capacitor active differentiator.

will be explored in depth in this chapter.

4.2 Background

To better understand the operation of our circuit, let us first consider several existing analog elements for computing the derivative.

The most basic continuous-time derivative circuit is the current-mode derivative formed by the voltage across a capacitor, given by $I = C\frac{dV}{dt}$, shown in Fig. 4.1(a). This simple circuit works well in theory, but issues such as output impedance, DC coupling, and fixed gain (given by C) prevent simple implementations of the derivative with the capacitor.

By placing a resistance between the capacitor and ground, a R-C network forming a high-pass filter, as shown in Fig. 4.1(b) can form a voltage-mode differentiator. For a sinusoidal input, this high-pass transfer function can satisfy the trigonometric derivative $\frac{d}{dt}\sin(\omega t) = \omega\cos(\omega t)$ up to a the frequency where $\omega > 1/10\omega_c$. The point

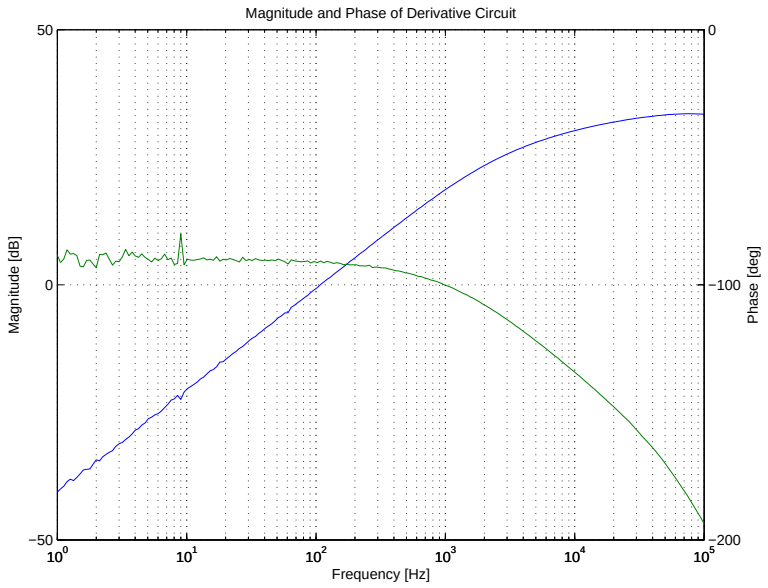


Figure 4.2: Magnitude and phase of system to compute a continuous time derivative. In this instance, $\omega_d = 1$kHz. Note the 20 dB/decade slope and 90°phase shift.

when the high-pass filter is operating in the 20 dB/decade rolloff region with 90°of phase shift, as shown in Fig. 4.2 is referred to as the "differentiator slope."

This R-C filter has a number of issues that prevent it from being practical. Foremost among them is the issue of noise. While signals within the "differentiator slope" pass properly, any higher frequency content, such as noise, that falls in the passband of the filter is emphasized over the lower-frequency content. Fabrication constraints also play a role - for processing biological signals with time constants on the order of 10 ms and maximum capacitance values on the order of picofarads, resistance values would be on the order of gigaohms [30].

Finally, the authors of [30] proposed the Clamped-capacitor architecture of Fig. 4.1(c). This circuit provides the necessary differentiator slope and phase shift and, unlike the R-C filter, is tunable over a range of frequencies. However, it computes the derivative over a small range of frequencies and with a very high passband gain. With this high gain, THD becomes a significant issue as harmonic amplification increases

exponentially with frequency.

4.3 Frequency-Normalized Derivative Criteria

To address these issues, it becomes apparent a band-limited derivative is necessary. By limiting the range of frequencies each derivative circuit covers, both issues are solved. Within this band of frequencies, the derivative function can be normalized with respect to a "derivative frequency" (ω_d), typically the center frequency of the band. This changes the output of the derivative function to eq. 4.1.

$$\frac{d}{dt}\sin\left(\frac{\omega}{\omega_d}t\right) = \frac{\omega}{\omega_d}\cos\left(\frac{\omega}{\omega_d}t\right) \tag{4.1}$$

Since high-frequency contributions are filtered out by the nature of the band-limited environment and the frequency ω_d can be chosen to produce moderate gain within the band (e.g. unity gain at ω_d), the output of the circuit will not exceed the dynamic range of the system. In addition to the 90°phase shift necessary to produce the derivative of sinusoids, the slope of the frequency response at ω_d must be 20 dB/decade to allow for the presence of instantaneous magnitude changes in the input (hard transients).

These criteria for the band-limited derivative operation centered around ω_d are summarized as:

1. The gain at ω_d equals one
2. The slope of the magnitude frequency response within the band is 20 dB/decade
3. The phase is a constant 90°in the frequency band of interest

From these criteria, it can be seen that a band-limited, frequency-normalized derivative can be implemented with either a high-pass or band-pass filter if the filter has a gain $>$ 20dB and the lowest corner frequency is $> 10\omega_d$.

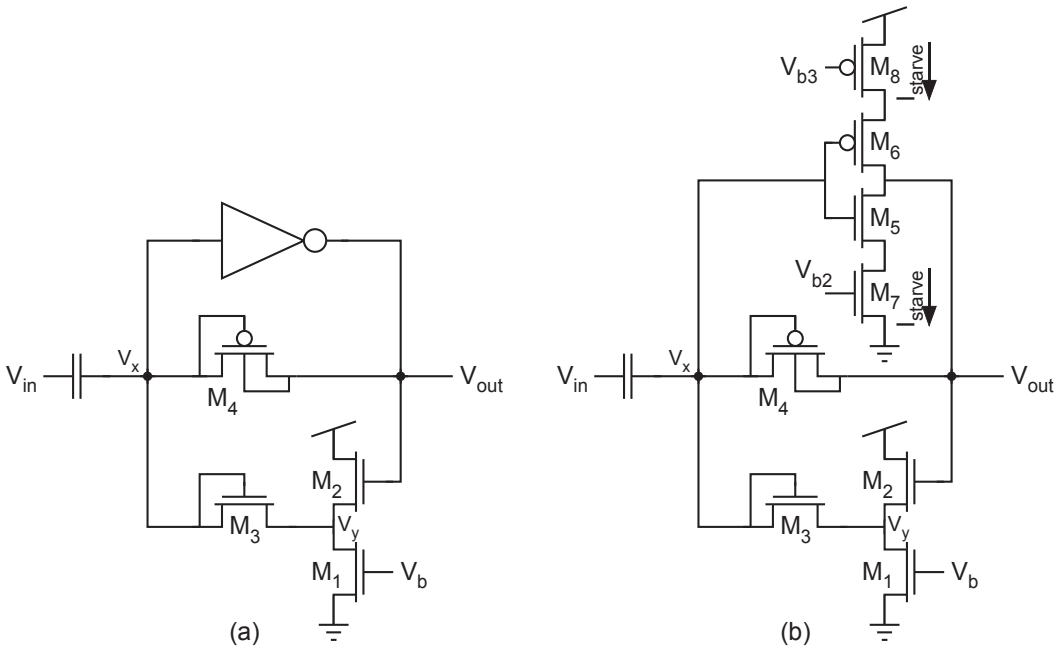


Figure 4.3: Schematics of the two versions of the derivative circuit **(a)** Standard version and **(b)** Lower-power version

4.4 Circuit Description

To address the issues with these previous implementations, we have developed the circuits of Fig. 4.3 in a commercially available 0.5 μm CMOS process. Composed of only six (eight for the low-power version), it satisfies the three criteria for a band-limited derivative circuit while consuming both little power and little area on chip.

This circuit consists of (1) a capacitor to perform the voltage-to-current derivative, (2) a high-gain inverting amplifier constructed from a digital inverter, (3) a source follower (M_{1-2}), (4) a diode-connected transistor (M_3), and (5) an impedance element for feedback (M_4) commonly referred to as the Tobi element [32]. The Tobi element acts as both a very large resistive element (GΩ) for small differential voltages and also as a current limiter for larger voltages across it due to the bi-directional exponential current-voltage relationship [32]. The Voltage-Current transfer function of this device is shown in fig. 4.4.

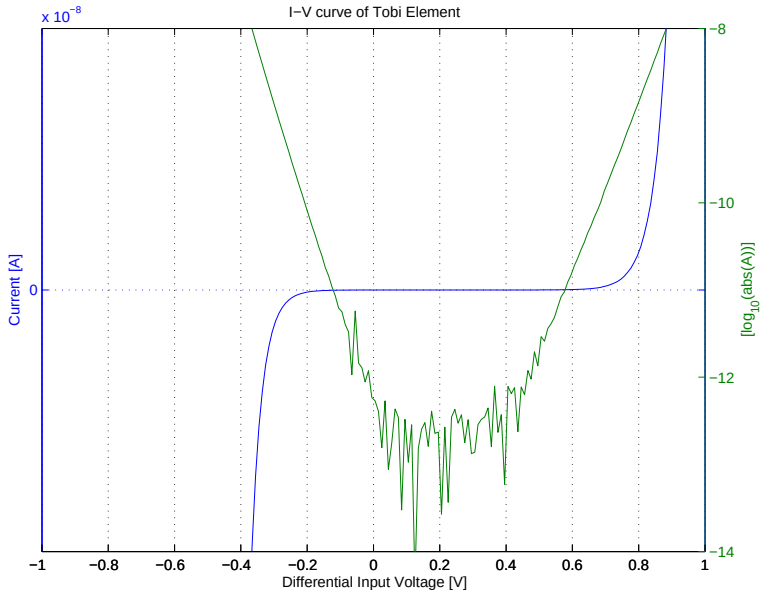


Figure 4.4: I-V characteristics of the Tobi element. The superimposed semi-logarithmic plot of the current illustrates the bi-directional exponential V-I characteristic of this device.

4.5 DC Analysis

The digital inverter is sized such that its threshold voltage is at approximately mid-rail. Negative feedback is accomplished through the large resistive element of M_4, which ensures that node V_x lies within the linear range of the inverting amplifier and, accordingly, maintains a nearly constant value. The source follower of M_{1-2} acts as a DC level shifter so that node V_y is less than V_{out} by an amount that is dependent upon the bias voltage, V_b. As a result, there will always be a bias-dependant voltage across the diode-connected M_3, generating a current that flows out of the inverter and then through the series combination of M_4 and M_3. Linear changes in V_b produce linear changes in the offset between V_x and V_{out} which, in turn, translate into exponential changes in the current through M_3 since it operates in weak inversion.

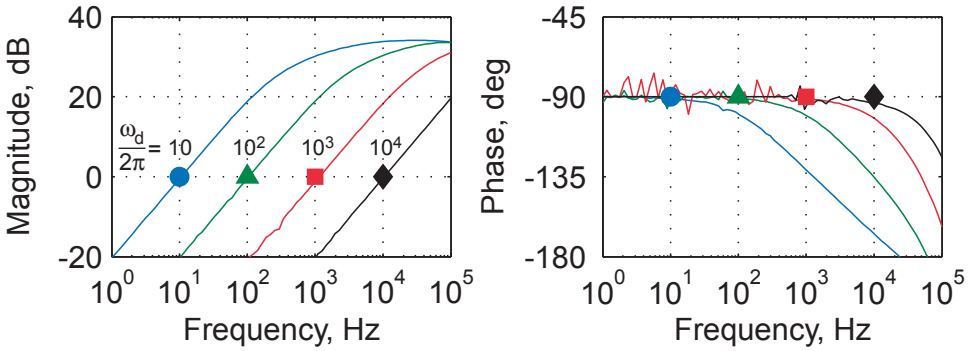


Figure 4.5: Frequency response of standard derivative circuit under four different bias conditions.

4.6 AC Analysis

It should be noted that in the steady-state condition of $V_{out} > V_{mid}$, the differential pair can be thought of as a source-follower amplifier with a diode-connected NFET on the output. Using this assumption, the simplified small signal model of Eq. 4.2 was developed.

$$A_v = \frac{(g_{m4} + g_{Mi})}{g_{m4}}$$
$$H(s) = -A_v \frac{sC/(A_v(g_{m3} + g_{m4}))}{sC/(A_v(g_{m3} + g_{m4})) + 1} \tag{4.2}$$

This analysis neglects the output resistance of the inverter, but it provides a first-order model of the circuit's operation and sheds light on how the circuit may be tuned for operation.

Tuning

The corner frequency of this circuit is tunable through several decades by changing the voltage on M_b as shown in Fig. 4.5. Though the differential pair is operating in moderate to strong inversion, this change in frequency is still exponential for a linear change in bias current. This is due to the exponential V-I relationship of the Tobi

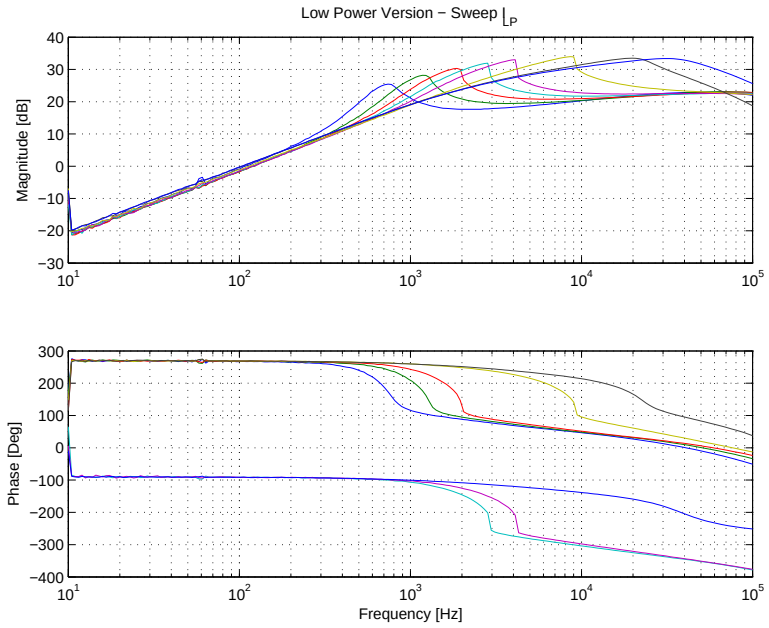


Figure 4.6: Frequency response of low power derivative circuit under different I_{starve} bias conditions while V_{bias} is held constant.

element - as the voltage across it increases linearly for a linear change in V_{bias}, the current through it changes exponentially and thus changes G_{M_T} exponentially.

4.7 Low Power Derivative Circuit

The previously described circuit suffers several drawbacks that prevent it from finding immediate application - namely, the quiescent current draw through the inverter and the high gain in the passband. These issues increase power consumption and lead to high THD and noise accumulation.

To remedy both these issues, source degeneration was introduced to the inverter. By limiting the amount of quiescent current available, we immediately decrease the power consumption - this version of the circuit draws only only 1.45-20.13μW in the audio range of 20Hz to 20kHz. This also decreases the transconductance of the inverter and introduces another pole to our small signal model. This lower transcon-

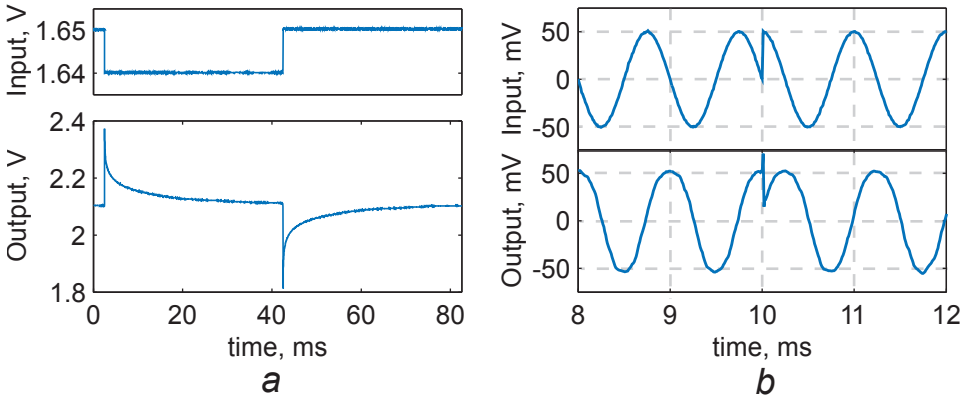


Figure 4.7: Time-domain responses of the derivative circuit. **(a)** Step input and response and **(b)** Sinusoidal input with a sudden transition.

ductance can be used to tune an upper corner frequency (Fig. 4.6) to the circuit and decrease the amount of high-frequency integrated noise.

4.8 Experimental Results

Both the standard and low-power versions of this circuit were implemented in a standard 0.5μm CMOS process. To decrease the design size, the input capacitor was implemented using a MOSCAP - testing the MOSCAP alongside standard poly-poly caps yielded identical circuit performance.

Figures 4.5 and 4.6 demonstrate the frequency response characteristics of the circuit. As seen in the figures, these circuits effectively implement a high-pass filter with a passband gain $>$ 20dB/decade and an electronically tunable lower corner frequency that is $> 10\omega_d$. As a result, this circuit meets all three criteria for a frequency-normalized derivative (with a -90°phase shift rather than +90°due to the inverting gain).

To demonstrate the temporal characteristics of our circuit, Fig. 4.7 illustrates two important test cases. In Fig. 4.7a, a step input of 10 mV was applied to illustrate the differentiation achieved when the input changes instantaneously. Accordingly,

the output 'jumps' when an input step occurs and then returns to an equilibrium value when the input remains constant, as expected of a derivative operation. In the demonstration of Fig. 4.7b, we biased the circuit to perform a derivative for a signal at 1kHz (i.e. $\omega_d = 2\pi(1\text{kHz})$). A 1kHz sine wave was applied until t=0.01 seconds, and then the input signal instantaneously transitions to a cosine at the same frequency. Fig. 4.7b shows that output of the circuit provides the negative derivative of the input with unity gain, i.e. the steady-state output for a sine-wave input is a cosine wave (with a gain of -1). The derivative circuit emphasizes the discontinuities in the input signal, as expected (see t=0.01sec), and quickly returns to the steady-state conditions. Results of the lower-power version of the derivative circuit (Fig. 4.3(b)) are identical.

4.9 Conclusion

This work has outlined criteria for implementing practical analog derivative circuits and presented a new, compact circuit able to meet these criteria. Because of this circuit's very small size (6-8 transistors) and low power consumption (only 1.45-20.13μW for the low-power version in the audio range of 20Hz to 20kHz), it is well-suited for implementation within an array-based processing environment. Example applications include sub-banded audio processing on energy-constrained platforms, such as silicon cochlear models [29, 33] and motion detectors implemented on-chip in CMOS imagers [30].

Chapter 5

Conclusion

In the effort to improve energy efficiency in low-power systems, engineers are increasingly turning to analog signal processing for audio-band signals. However, implementing these analog systems on energy-constrained platforms (such as hearing aids, cochlear implants, autonomous sensor platforms, cellular phones, etc.) poses a number of challenges. System reconfigurability, limited dynamic range, and a lack of basic signal processing elements suitable for implementation in a compact, low-power architecture have all hampered the adoption of continuous-time processing as a solution for common processing tasks.

This work first covered the development of a flexible, extensible, SPICE-compatible simulation model for a Floating-Gate transistor. This model is the first reported in literature to incorporate the non-ideal effects of capacitive coupling and charge modification on the floating node in DC, AC and transient analyses. Furthermore, this model is extensible - different mathematical models of charge modification (Fowler-Nordheim tunneling, hot-electron injection, UV photoinjection) may be substituted to optimize the model's performance under different operating regimes. Finally, the model's description of capacitive division is leveraged to demonstrate a reduction of the basic transistor parameter κ via a designable process.

The designable nature of κ on a floating-gate transistor can be used to increase

the dynamic range of many devices based on the differential pair circuit. Another method , Ohmic biasing, uses the stored charge on the floating node to change the common-mode voltage of the differential pair's inputs, forcing the biasing transistor out of saturation and into the Ohmic regime. Analog designers have traditionally avoided Ohmic biasing in differential pairs, but here it is shown to decrease the transconductance of the circuit for small differential inputs and increase the dynamic range by extending the input linear range. This decrease in transconductance is designable, leading to a variable input linear range that is larger than a standard floating-gate differential pair by a factor of three. The increased dynamic range, at no additional area cost or power consumption, is a powerful new tool to increase the efficiency of continuous-time circuits. To demonstrate these benefits, a Variable-Gain Amplifier (VGA) was designed and simulated, showing a 350% increase in linear range compared to a VGA implemented with the same components operating in saturation without floating-gate inputs.

Finally, a derivative circuit suitable for inclusion in a band-limited system is presented. Due to the very nature of the derivative operation, designing a circuit to perform a continuous-time derivative has been met with little success. Three criteria for developing a band-limited, continuous-time derivative circuit are introduced and are implemented in a novel new design. Operating on only micro-watts of power and composed of 6-8 transistors, this compact, low-power circuit is shown to compute the continuous-time derivative of not only sinusoidal signals, but also handles step or impulse inputs as well. The band-limited, low-power, small-footprint aspects of this implementation make it ideal for inclusion in an array-based analog processing environment.

5.1 Future Work

The broad nature of this research presents myriad avenues for future work. Using the floating-gate model for future designs will speed the development time of new circuits considerably. It was instrumental in the development of the Ohmic biasing technique, and can be used with different tunneling or injection models to investigate run-time modeling of FG programming processes. Immediate applications of the model tie in with development work on the Ohmic biasing technique.

The Ohmic biasing technique was demonstrated on a simulation of a VGA with a single-ended output, but the technique is much better suited for use in a fully differential system. Utilizing the Ohmic biasing technique in a fully differential VGA, rather than the single-ended system shown here, will mitigate many of the impacts of changing the common-mode and make Ohmic biasing a more flexible method of extending linear range.

In a larger context, the VGA can be incorporated with a gain control circuit to develop an Automatic Gain Control (AGC) amplifier. Used as an input stage to a filterbank prior to spectral decomposition, an AGC amplifier could increase the dynamic range of the system considerably by compressing signals too large for processing in sub-banded elements, or amplifying signals that fall below the noise floor of others, such as the derivative circuit.

Finally, the derivative circuit lends itself to immediate inclusion in a signal processing system. Applications such as visual processing [30], cardiac monitoring [1], and speech recognition [3] all have an immediate need for low-power, continuous-time differentiator circuits.

All of the future work discussed here is targeted at the common goal of increasing computational efficiency while decreasing component cost. A fully-differential, ohmically biased AGC could improve the dynamic range of existing filterbank structures dramatically. The FG simulation model enables advanced testing and design

of the entire filterbank, while the derivative circuit improves the flexibility of the filterbank for implementing innovative signal processing algorithms. Taken together, these advancements in low-power analog signal processing open new applications to sub-banded signal processing and enhance the state-of-the art.

Appendix A

Derivative Circuit Small-Signal Derivation

A.1 Abstraction

To reach a simplified small-signal transfer function, three assumptions have been made about the operation of the derivative circuit in fig. A.1.

- The current through M_3 is much, much smaller than the current through M_2, so the combination of M_1 and M_2 can be modeled as a Common Drain Amplifier.
- The inverter formed by M_5 and M_6 is operating in its linear region.
- Back-gate effects can be neglected and M_3 can be modeled as a simple resistance.

Performing KCL at node V_x:

$$(V_{in} - V_x)\, sC + (V_{out} - V_x)\, g_{M4} + (V_{out} - V_x)\, g_{M3} = 0 \tag{A.1}$$

Performing KCL at node V_{out}

$$V_x g_{Mi} + (V_{out} - V_x)\, g_{M4} = 0 \tag{A.2}$$

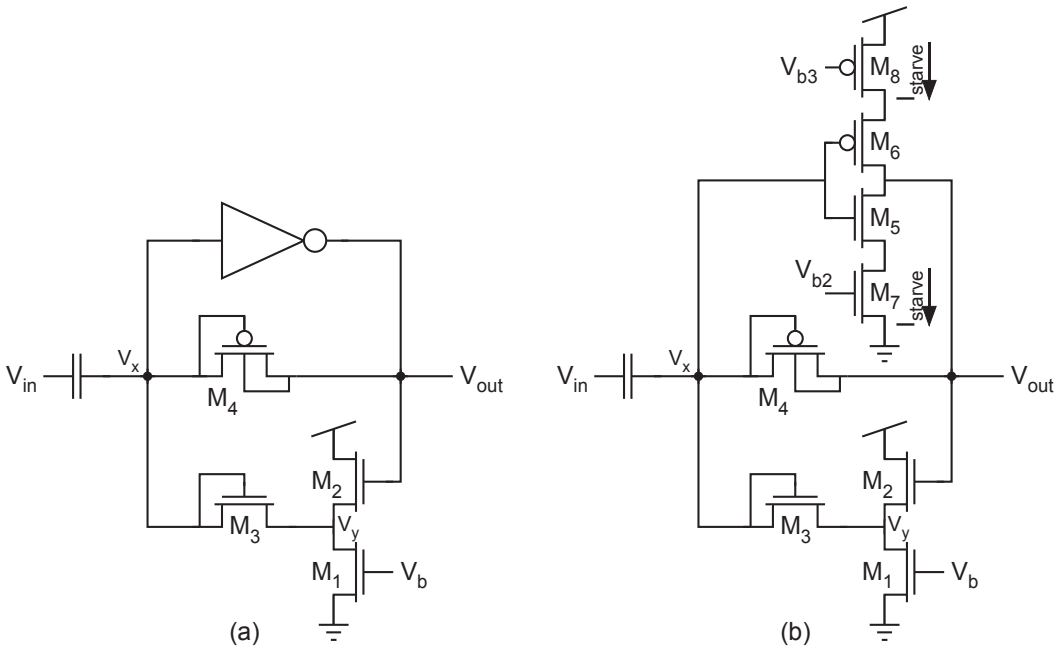


Figure A.1: Schematics of the two versions of the derivative circuit **(a)** Standard version and **(b)** Lower-power version

Separate the voltage terms of Eq. A.1:

$$V_{in}sC - V_xsC + V_{out}g_{M4} - V_xg_{M4} + V_{out}g_{M3} - V_xg_{M3} = 0 \tag{A.3a}$$

$$V_{in}sC + V_{out}\left(g_{M3} + g_{M4}\right) = V_x\left(sC + g_{M4} + g_{M3}\right) \tag{A.3b}$$

$$V_{in}\frac{sC}{g_{M3} + g_{M4}} + V_{out} = V_x\frac{sC + g_{M4} + g_{M3}}{g_{M3} + g_{M4}} \tag{A.3c}$$

$$V_{in}\frac{sC}{g_{M3} + g_{M4}} + V_{out} = V_x\left[\frac{sC}{g_{M3} + g_{M4}} + 1\right] \tag{A.3d}$$

Do the same to the voltage terms of Eq. A.2:

$$V_x g_{Mi} + V_{out} g_{M4} - V_x g_{M4} = 0 \tag{A.4a}$$

$$V_x \left(g_{Mi} - g_{M4}\right) = -V_{out} g_{M4} \tag{A.4b}$$

$$V_x \left(g_{M4} - g_{Mi}\right) = V_{out} g_{M4} \tag{A.4c}$$

$$V_x = V_{out} \frac{g_{M4}}{g_{M4} - g_{Mi}} \tag{A.4d}$$

$$V_x = \frac{V_{out}}{1 - \frac{g_{Mi}}{g_{M4}}} \tag{A.4e}$$

Now combine Equations A.3d and A.4e and solve for V_{out}/V_{in}.

$$V_{in} \frac{sC}{g_{M3} + g_{M4}} + V_{out} = \frac{V_{out}}{1 - \frac{g_{Mi}}{g_{M4}}} \left[\frac{sC}{g_{M3} + g_{M4}} + 1\right] \tag{A.5a}$$

$$\text{Define: } g_{mx} = g_{m3} + g_{m4} \tag{A.5b}$$

$$V_{in} \frac{sC}{g_{mx}} = \frac{V_{out}}{1 - \frac{g_{Mi}}{g_{M4}}} \left[\frac{sC}{g_{mx}} + 1\right] - V_{out} \tag{A.5c}$$

$$V_{in} \frac{sC}{g_{mx}} = V_{out} \left[\frac{sC}{g_{mx}\left(1 - \frac{g_{Mi}}{g_{M4}}\right)} + \frac{1}{1 - \frac{g_{Mi}}{g_{M4}}} - \frac{1 - \frac{g_{Mi}}{g_{M4}}}{1 - \frac{g_{Mi}}{g_{M4}}}\right] \tag{A.5d}$$

$$V_{in} \frac{sC}{g_{mx}} = V_{out} \left[\frac{sC}{g_{mx}\left(1 - \frac{g_{Mi}}{g_{M4}}\right)} + \frac{\frac{g_{Mi}}{g_{M4}}}{1 - \frac{g_{Mi}}{g_{M4}}}\right] \tag{A.5e}$$

$$V_{in} \frac{sC}{g_{mx}} \left[1 - \frac{g_{Mi}}{g_{M4}}\right] = V_{out} \left[\frac{sC}{g_{mx}} + \frac{g_{Mi}}{g_{M4}}\right] \tag{A.5f}$$

$$V_{in} \frac{sC}{g_{mx}} \frac{g_{m4}}{g_{mi}} \left[1 - \frac{g_{Mi}}{g_{M4}}\right] = V_{out} \left[\frac{sC}{g_{mx}} \frac{g_{m4}}{g_{mi}} + 1\right] \tag{A.5g}$$

$$\tag{A.5h}$$

Now define the intermediate variable τ:

$$\tau = \frac{C}{g_{mx}} \frac{g_{m4}}{g_{mi}} \tag{A.6a}$$

$$V_{in} s\tau \left[1 - \frac{g_{mi}}{g_{m4}}\right] = V_{out} \left[s\tau + 1\right] \tag{A.6b}$$

$$\frac{V_{out}}{V_{in}} = \left[1 - \frac{g_{mi}}{g_{m4}}\right] \frac{s\tau}{s\tau + 1} \tag{A.6c}$$

$$\text{Assume: } \left[1 - \frac{g_{mi}}{g_{m4}}\right] >> 1 \tag{A.6d}$$

$$H(s) = -A_v \frac{s\tau}{s\tau + 1}, \text{ where } A_v = \frac{g_{mi}}{g_{m4}} \tag{A.6e}$$

Finally, substitute A_v into the expression for τ:

$$H(s) = -A_v \frac{\frac{sC}{A_v(g_{m3}+g_{m4})}}{\frac{sC}{A_v(g_{m3}+g_{m4})} + 1} \tag{A.7}$$

Appendix B

Analysis of the Ohmic Differential Pair

Differential pairs, such as the one in Fig. B.1 are typically modeled as a function of the hyperbolic tangent [20] as shown in Eq B.1. In this model, I_{out} is the differential output current, I_b is the bias current through the "tail" transistor, κ is the gate-channel coupling coefficient of M_1 and M_2, and ΔV_{in} is the differential input voltage described by $V_1 - V_2$.

$$I_{out} = I_1 - I_2 = I_b \tanh\left(\frac{\kappa}{2U_T}\Delta V_{in}\right) \tag{B.1}$$

In this derivation we will use the EKV model [27] for MOSFET devices and neglect the body effect for simplicity. Transistors M_1 and M_2 have the aspect ratio $\frac{W_{in}}{L_{in}}$ and M_b has the aspect ratio $\frac{W_b}{L_b}$.

Interestingly, when the tail transistor operates in the Ohmic regime rather than in saturation, the linear range of the differential pair is extended and this simple model breaks down. This section provides the derivation of a new transfer function that better describes a differential pair with extended linear range through Ohmic biasing of the tail transistor. This model includes distortion terms induced through ohmic biasing and gives rise to a simple expression for finding the approximate linear range

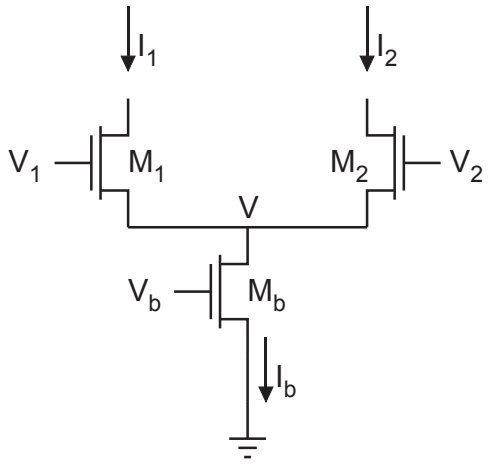


Figure B.1: Differential pair used in the derivation of the transfer function.

of a differential pair under subthreshold biasing conditions.

B.1 Standard Differential Pair In Saturation

To better understand the operation of the differential pair with Ohmic biasing, first look at the derivation for the output current defined as $I_{out} = I_1 - I_2$ in a normal differential pair, given in equations B.2

First define the current through each transistor, assuming that the bias transistor is always in saturation.

$$I_1 = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_1/U_T} e^{-V/U_T} \tag{B.2a}$$

$$I_2 = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_2/U_T} e^{-V/U_T} \tag{B.2b}$$

$$I_b = I_0' \frac{W_b}{L_b} e^{\kappa V_b/U_T} \tag{B.2c}$$

Then solve to eliminate the dependency on the node V.

$$I_b = I_1 + I_2 \tag{B.2d}$$

$$I_0' \frac{W_b}{L_b} e^{\kappa V_b/U_T} = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_1/U_T} e^{-V/U_T} + I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_2/U_T} e^{-V/U_T} \tag{B.2e}$$

$$\frac{W_{in}}{L_{in}} e^{-V/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T} \right) = \frac{W_b}{L_b} e^{\kappa V_b/U_T} \tag{B.2f}$$

Define a variable w to simplify.

$$w = \frac{\frac{W_b}{L_b}}{\frac{W_{in}}{L_{in}}} \tag{B.2g}$$

$$e^{-V/U_T} = \frac{w e^{\kappa V_b/U_T}}{e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}} \tag{B.2h}$$

$$e^{-V/U_T} = \frac{w e^{\kappa V_b/U_T}}{e^{\kappa \left(V_{cm} + \frac{\Delta V_{in}}{2}\right)/U_T} + e^{\kappa \left(V_{cm} - \frac{\Delta V_{in}}{2}\right)/U_T}} \tag{B.2i}$$

$$e^{-V/U_T} = \frac{w e^{\kappa V_b/U_T}}{e^{\kappa V_{cm}} \left(e^{\frac{\kappa \Delta V_{in}}{2U_T}} + e^{-\frac{\kappa \Delta V_{in}}{2U_T}} \right)} \tag{B.2j}$$

Now use this value of e^{-V/U_T} to solve for the differential output current, I_{out}.

$$I_{out} = I_1 - I_2 \tag{B.2k}$$

$$I_{out} = I_0' \frac{W_{in}}{L_{in}} \left(e^{\kappa V_1/U_T} - e^{\kappa V_2/U_T} \right) e^{-V/U_T} \tag{B.2l}$$

$$I_{out} = \frac{I_0' \frac{W_{in}}{L_{in}} w \left(e^{\frac{\kappa \Delta V_{in}}{2U_T}} - e^{-\frac{\kappa \Delta V_{in}}{2U_T}} \right) e^{\kappa V_b/U_T}}{e^{\frac{\kappa \Delta V_{in}}{2U_T}} + e^{-\frac{\kappa \Delta V_{in}}{2U_T}}} \tag{B.2m}$$

$$I_{out} = \frac{I_0' \frac{W_b}{L_b} e^{\kappa V_b/U_T} \sinh \left(\frac{\kappa \Delta V_{in}}{2U_T} \right)}{\cosh \left(\frac{\kappa \Delta V_{in}}{2U_T} \right)} \tag{B.2n}$$

$$I_{out} = I_b \tanh \left(\frac{\kappa \Delta V_{in}}{2U_T} \right) \tag{B.2o}$$

B.2 Standard Differential Pair With Ohmic Bias

The analysis of the differential pair with ohmic biasing follows the same methodology as the analysis with a saturated bias transistor, with the exception that the current I_b through the tail transistor is now defined as

$$I_b = I_0' \frac{W_b}{L_b} e^{\kappa V_b/U_T} \left(1 - e^{-V/U_T}\right) \tag{B.3a}$$

I_1 and I_2 are the same as the previous derivation.

$$I_1 = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_1/U_T} e^{-V/U_T} \tag{B.3b}$$

$$I_2 = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_2/U_T} e^{-V/U_T} \tag{B.3c}$$

Using KCL around node V,

$$I_b = I_1 + I_2 \tag{B.4}$$

By inserting Eqs. B.3 into Eq. B.4,

$$I_0' \frac{W_b}{L_b} e^{\kappa V_b/U_T} \left(1 - e^{-V/U_T}\right) = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_1/U_T} e^{-V/U_T} + I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_2/U_T} e^{-V/U_T} \tag{B.5a}$$

$$\frac{W_{in}}{L_{in}} e^{-V/U_T} \left(e^{\kappa V_1/U_T} + \frac{W_{in}}{L_{in}} e^{\kappa V_2/U_T}\right) = \frac{W_b}{L_b} e^{\kappa V_b/U_T} - e^{\kappa V_b/U_T} e^{-V/U_T} \tag{B.5b}$$

$$\frac{1}{w} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) = e^{\kappa V_b/U_T} e^{V/U_T} - e^{\kappa V_b/U_T} \tag{B.5c}$$

$$\frac{1}{w} e^{-\kappa V_b/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) = e^{V/U_T} - 1 \tag{B.5d}$$

$$e^{V/U_T} = \frac{1}{w} e^{-\kappa V_b/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) + 1 \tag{B.5e}$$

$$e^{-V/U_T} = \frac{1}{\frac{1}{w} e^{-\kappa V_b/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) + 1} \tag{B.5f}$$

In this derivation, the intermediate variable w is defined as:

$$w = \frac{\frac{W_b}{L_b}}{\frac{W_{in}}{L_{in}}} \tag{B.6}$$

Eq. B.5f can then be inserted back into Eqs. B.4 and B.3 to solve for the currents independently of the variable V.

$$I_1 = \frac{I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_1/U_T}}{\frac{1}{w} e^{-\kappa V_b/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) + 1} \tag{B.7a}$$

$$I_2 = \frac{I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_2/U_T}}{\frac{1}{w} e^{-\kappa V_b/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) + 1} \tag{B.7b}$$

$$I_{out} = I_1 - I_2 = \frac{I_0' \frac{W_{in}}{L_{in}} \left(e^{\kappa V_1/U_T} - e^{\kappa V_2/U_T}\right)}{\frac{1}{w} e^{-\kappa V_b/U_T} \left(e^{\kappa V_1/U_T} + e^{\kappa V_2/U_T}\right) + 1} \tag{B.7c}$$

Now comes a little sleight of hand. We define a variable I_t to describe the "target current" through an input transistor for a given common mode voltage across the differential pair. This variable is derived by separating V_1 and V_2 into their respective common mode V_{CM} and differential mode ΔV_{in} voltages. We also define the variable $I_{b,sat}$ to describe the maximum current thorugh the bias transistor for a given value of V_b.

$$I_t = I_{t1} = I_{t2} = I_0' \frac{W_{in}}{L_{in}} e^{\kappa V_{CM}/U_T} \tag{B.8a}$$

$$I_{b,sat} = I_0' \frac{W_b}{L_b} e^{\kappa V_b/U_T} \tag{B.8b}$$

Substituting these expressions back into Eqs. B.3b and B.3c and then combining with Eq. B.5f, we obtain the following expression for I_{out}.

$$I_1 = I_0' \frac{W_{in}}{L_{in}} e^{\kappa\left(V_{CM} + \frac{\Delta V_{in}}{2}\right)/U_T} e^{-V/U_T} \tag{B.9a}$$

$$I_2 = I_0' \frac{W_{in}}{L_{in}} e^{\kappa\left(V_{CM} - \frac{\Delta V_{in}}{2}\right)/U_T} e^{-V/U_T} \tag{B.9b}$$

$$I_{out} = \frac{I_t \left(e^{\kappa \frac{\Delta V_{in}}{2}/U_T} - e^{-\kappa \frac{\Delta V_{in}}{2}/U_T}\right)}{1 + \frac{I_t}{I_{b,sat}} \left(e^{\kappa \frac{\Delta V_{in}}{2}/U_T} + e^{-\kappa \frac{\Delta V_{in}}{2}/U_T}\right)} \tag{B.9c}$$

The expression $I_{b,sat}$ assumes M_b is operating in saturation and the length is sufficiently long to provide a large Early voltage. This value is useful for the analysis,

but during circuit operation for small values of ΔV_{in} M_b will be operating in the Ohmic regime and sourcing a smaller current.

The expression for I_{out} can be simplified by recognizing the definitions of the hyperbolic sine and cosine:

$$\sinh = \frac{1}{2}\left(e^{x} - e^{-x}\right) \tag{B.10a}$$

$$\cosh = \frac{1}{2}\left(e^{x} + e^{-x}\right) \tag{B.10b}$$

By defining two variables for substitution and simplification such that

$$x = \frac{\kappa \Delta Vin}{2U_T} \tag{B.11a}$$

$$\beta = \frac{I_{b,sat}}{2I_t} \tag{B.11b}$$

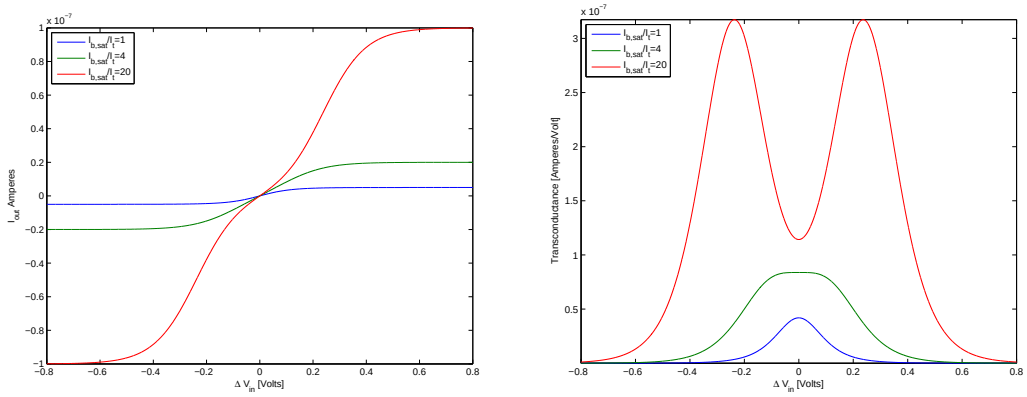
The expression for I_{out} then becomes:

$$I_{out} = \frac{2I_t \sinh\left(\frac{\kappa \Delta Vin}{2U_T}\right)}{1 + \frac{2I_t}{I_{b,sat}} \cosh\left(\frac{\kappa \Delta Vin}{2U_T}\right)} \tag{B.12a}$$

$$I_{out} = \frac{I_{b,sat} \sinh\left(\frac{\kappa \Delta Vin}{2U_T}\right)}{\frac{I_{b,sat}}{2I_t} + \cosh\left(\frac{\kappa \Delta Vin}{2U_T}\right)} \tag{B.12b}$$

$$I_{out} = \frac{I_{b,sat} \sinh\left(x\right)}{\beta + \cosh\left(x\right)} \tag{B.12c}$$

At this point, it is interesting to note that Eq. B.12c is the same as Eq. (6) in [10]. In that publication, the authors used a different method of extending the linear range, which nonetheless has the same effect on the DC operation of the circuit. It is also interesting to note that the output current not depend on the relative sizes of the input pair and the bias transistor.



(a) (b)

Figure B.2: (a) Eq. B.12c plotted for three different values of I_t, illustrating the importance of maximizing the linear range while minimizing distortion. (b) Eq. B.13 plotted for three different values of I_t, illustrating the distortion when the ratio of $I_{b,sat}/I_t$ is too large.

B.3 Biasing for Maximum Linear Range / Minimum Distortion

As shown in Fig. B.2a, the expression for I_{out} from Eq. B.12c can take on three distinct shapes as the ratio between $I_{b,sat}$ and I_t changes. It is therefore necessary to define an ideal operating point to maximize the linear range while simultaneously minimizing nonlinearities in the amplification region (this also minimizes THD).

This operating point is defined as the bias condition that yields maximal flatness in the first derivative of I_{out}. This first derivative is also the transconductance of the differential pair, such that

$$g_m = \frac{dI_{out}}{d\Delta V_{in}} = \frac{\kappa I_{b,sat} I_t}{U_T} \frac{(I_{b,sat}\cosh(x) + 2I_t)}{(I_{b,sat} + 2I_t\cosh(x))^2} \tag{B.13}$$

The nonlinearities in the amplification region are especially emphasized in these curves, illustrated in Fig. B.2b.

The boundary condition between maximum flatness and distortion can be described as the point where I_{out} is always concave for positive values of ΔV_{in} and convex for negative values of ΔV_{in}. More precisely, $\frac{dg_m}{d\Delta V_{in}} \leq 0$ for $\Delta V_{in} \geq 0$ and $\frac{dg_m}{d\Delta V_{in}} \geq 0$ for $\Delta V_{in} \leq 0$.

$$\frac{dg_m}{d\Delta V_{in}} = \frac{\kappa^2 I_{b,sat} I_t}{2U_T^2} \frac{sinh(x)\left(I_{b,sat}^2 - 2I_{b,sat}I_t\cosh(x) - 8I_t^2\right)}{\left(I_{b,sat} + 2I_t\cosh(x)\right)^3} \tag{B.14}$$

From B.14, it can be seen that since the denominator will be positive for all values of ΔV_{in} and the sign of $\sinh(x)$ only changes at $x = 0$, the expression

$$\left(I_{b,sat}^2 - 2I_{b,sat}I_t\cosh(x) - 8I_t^2\right) < 0 \tag{B.15}$$

to satisfy the boundary condition. The maximum value of Eq. B.15 occurs at $x = 0$. Therefore, we can solve to find the the ratio of $I_{b,sat}/I_t$ that satisfies our condition for increased linear range with no distortion.

$$\left(I_{b,sat}^2 - 2I_{b,sat}I_t\cosh(x) - 8I_t^2\right)|_{x=0} \leq 0 \tag{B.16a}$$

$$\left(I_{b,sat}^2 - 2I_{b,sat}I_t - 8I_t^2\right) \leq 0 \tag{B.16b}$$

$$-2 \leq \frac{I_{b,sat}}{I_t} \leq 4 \tag{B.16c}$$

Since $\frac{I_{b,sat}}{I_t}$ can never take on negative values, the root at -2 is invalid. Therefore,

$$\frac{I_{b,sat}}{I_t} \leq 4 \tag{B.17}$$

Therefore, the maximum linear range occurs when $\frac{I_{b,sat}}{I_t} = 4$.

B.4 Solving for Linear Range

Let V_L be the voltage at which the line defined by the maximum transconductance equals $I_{b,sat}$.

$$I_{b,sat} = g_{m,max} V_L \tag{B.18}$$

It can be seen that g_m takes on its maximum value at $\Delta V_{in} = 0$. Solving with the analytical solution for g_m,

$$V_L = I_{b,sat} \frac{1}{g_{m,max}} \tag{B.19a}$$

$$V_L = I_{b,sat} \frac{U_T}{\kappa I_{b,sat} I_t} \frac{\left(I_{b,sat} + 2I_t \cosh(x)\right)^2}{\left(I_{b,sat} \cosh(x) + 2I_t\right)}\Big|_{x=0} \tag{B.19b}$$

$$V_L = \frac{U_T}{\kappa I_t} \left(I_{b,sat} + 2I_t\right) \tag{B.19c}$$

$$V_L = \frac{U_T}{\kappa} \left(\frac{I_{b,sat}}{I_t} + 2\right) \tag{B.19d}$$

$$V_L = \frac{U_T}{\kappa} \frac{I_{b,sat}}{I_t} + \frac{2U_T}{\kappa} \tag{B.19e}$$

References

[1] R. Kher, D. Vala, T. Pawar, and V. Thakar, "Implementation of derivative based qrs complex detection methods," in *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, vol. 3, oct 2010, pp. 927 –931.

[2] D. Anderson, R. Harris, and D. Chabries, "Evaluation of a hearing compensation algorithm," in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 5, May 1995, pp. 3531 –3533 vol.5.

[3] J. Wilpon, C.-H. Lee, and L. Rabiner, "Improvements in connected digit recognition using higher order spectral and energy features," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, apr 1991, pp. 349 –352 vol.1.

[4] L. Wang, S. Ohtsuka, and S. Nakagawa, "High improvement of speaker identification and verification by combining mfcc and phase information," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, April 2009, pp. 4529 –4532.

[5] B. Rumberg, D. W. Graham, and V. Kulathumani, "Hibernets: energy-efficient sensor networks using analog signal processing," in *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, ser. IPSN '10. New York, NY, USA: ACM, 2010, pp. 129–139. [Online]. Available: http://doi.acm.org/10.1145/1791212.1791229

[6] P. Hasler, P. Smith, D. Graham, R. Ellis, and D. Anderson, "Analog floating-gate, on-chip auditory sensing system interfaces," *Sensors Journal, IEEE*, vol. 5, no. 5, pp. 1027 – 1034, Cctober 2005.

[7] B. Rumberg, K. McMillan, C. Rea, and D. Graham, "Lateral coupling in silicon cochlear models," in *Circuits and Systems, 2008. MWSCAS 2008. 51st Midwest Symposium on*, August 2008, pp. 25 –28.

[8] S. J. Rapp, "A comprehensive simulation model for floating gate transistors," Master's thesis, West Virginia University, December 2010.

[9] B. Rumberg, "Speech processing front-end in low-power hardware," Master's thesis, West Virginia University, May 2009.

[10] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A low-power wide-linear-range transconductance amplifier," in *Analog Integrated Circuits and Signal Processing*, 1997, pp. 123–151.

[11] A. E. Mourabit, G.-N. Lu, and P. Pittet, "A low-frequency, sub 1.5-v micropower gm-c filter based on subthreshold mifg mos transistors," in *Solid-State Circuits Conference, 2005. ESSCIRC 2005. Proceedings of the 31st European*, sept. 2005, pp. 331 – 334.

[12] M. Wang, T. Mayhugh, S. H. J. Embabi, and E. Sanchez-Sinencio, "Constant-gm rail-to-rail cmos op-amp input stage with overlapped transition regions," *Solid-State Circuits, IEEE Journal of*, vol. 34, no. 2, February 1999.

[13] R. H. Fowler and L. Nordheim, "Electron emission in intense electric fields," *Proceedings of the Royal Society of London. Series A*, vol. 119, no. 781, pp. 173–181, 1928. [Online]. Available: http://rspa.royalsocietypublishing.org/content/119/781/173.short

[14] L. Yin, S. Embabi, and E. Sanchez-Sinencio, "A floating-gate mosfet d/a converter," in *Circuits and Systems, 1997. ISCAS '97., Proceedings of 1997 IEEE International Symposium on*, vol. 1, June 1997, pp. 409 –412 vol.1.

[15] A. Mondragon-Torres, M. Schneider, and E. Sanchez-Sinencio, "Extraction of electrical parameters of floating gate devices for circuit analysis, simulation, and design," in *Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on*, vol. 1, August 2002, pp. I – 311–14 vol.1.

[16] L. Larcher, P. Pavan, S. Pietri, L. Albani, and A. Marmiroli, "A new compact dc model of floating gate memory cells without capacitive coupling coefficients," *Electron Devices, IEEE Transactions on*, vol. 49, no. 2, pp. 301 –307, Feb 2002.

[17] K. Rahimi, C. Diorio, C. Hernandez, and M. Brockhausen, "A simulation model for floating-gate mos synapse transistors," in *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*, vol. 2, 2002, pp. II–532 – II–535 vol.2.

[18] J. Gray, R. Robucci, and P. Hasler, "The design and simulation model of an analog floating-gate computational element for use in large-scale analog reconfigurable systems," in *Circuits and Systems, 2008. MWSCAS 2008. 51st Midwest Symposium on*, August 2008, pp. 253 –256.

[19] P. Hasler, A. Basu, and S. Kozil, "Above threshold pfet injectionmodeling intended for programmingfloating-gate systems," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, May 2007, pp. 1557 –1560.

[20] S. C. Liu, J. Kramer, G. Indiveri, T. Delbruck, and R. Douglas, *Analog VLSI: Circuits and Principles*. The MIT Press, 2002.

[21] T. Delbruck, "'bump' circuits for computing similarity and dissimilarity of analog voltages," in *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, vol. i, jul 1991, pp. 475 –479 vol.1.

[22] M. Cohen and A. Andreou, "Analog cmos integration and experimentation with an autoadaptive independent component analyzer," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 42, no. 2, pp. 65 –77, feb 1995.

[23] P. Furth and H. Ommani, "Low-voltage highly-linear transconductor design in subthreshold cmos," in *Circuits and Systems, 1997. Proceedings of the 40th Midwest Symposium on*, vol. 1, aug 1997, pp. 156 –159 vol.1.

[24] K. Odame, D. Anderson, and P. Hasler, "A bandpass filter with inherent gain adaptation for hearing applications," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 55, no. 3, pp. 786 –795, april 2008.

[25] B. Minch, "Evolution of a folded floating-gate differential pair," in *Circuits and Systems, 2000. Proceedings of the 43rd IEEE Midwest Symposium on*, vol. 3, 2000, pp. 1052 –1056 vol.3.

[26] J. Ramirez-Angulo, R. Carvajal, J. Tombs, and A. Torralba, "Low-voltage cmos op-amp with rail-to-rail input and output signal swing for continuous-time signal processing using multiple-input floating-gate transistors," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 48, no. 1, pp. 111 –116, jan 2001.

[27] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical mos transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog Integrated Circuits and Signal Processing*, vol. 8, pp. 83–114, 1995, 10.1007/BF01239381. [Online]. Available: http://dx.doi.org/10.1007/BF01239381

[28] S. Rapp, K. McMillan, and D. Graham, "Spice-compatible modelling technique for simulating floating-gate transistors," *Electronics Letters*, vol. 47, no. 8, pp. 483–485, 2011. [Online]. Available: http://link.aip.org/link/?ELL/47/483/1

[29] C. Mead, *Analog VLSI and Neural Systems*. Addison-Wesley, 1989.

[30] A. Moini, A. Bouzerdoum, A. Yakovleff, D. Abbott, O. Kim, K. Eshraghian, and R. Bogner, "An analog implementation of early visual processing in insects," in *VLSI Technology, Systems, and Applications, 1993. Proceedings of Technical Papers. 1993 International Symposium on*, 1993, pp. 283 –287.

[31] J. Wilpon, C.-H. Lee, and L. Rabiner, "Improvements in connected digit recognition using higher order spectral and energy features," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, apr 1991, pp. 349 –352 vol.1.

[32] T. Delbruck and C. Mead, "Adaptive photoreceptor with wide dynamic range," in *Circuits and Systems, 1994. ISCAS '94., 1994 IEEE International Symposium on*, vol. 4, may 1994, pp. 339 –342 vol.4.

[33] T. Hamilton, J. Tapson, C. Jin, and A. van Schaik, "Analogue vlsi implementations of two dimensional, nonlinear, active cochlea models," in *Biomedical Circuits and Systems Conference, 2008. BioCAS 2008. IEEE*, nov. 2008, pp. 153 –156.

Index