DESIGN AND IMPLEMENTATION OF ULTRA-LOW-POWER SENSOR INTERFACE CIRCUITS FOR ECG ACQUISITION

XU XIAOYUAN

NATIONAL UNIVERSITY OF SINGAPORE

2010

DESIGN AND IMPLEMENTATION OF ULTRA-LOW-POWER SENSOR INTERFACE CIRCUITS FOR ECG ACQUISITION

XU XIAOYUAN

(*B.Eng. (Hons.), NUS*)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2010

Acknowledgements

First, I would like to thank my supervisors Dr. Lian Yong and Dr. Yao Libin for their patient guidance, invaluable advice and consistent encouragement. Their profound knowledge, clear insights and inspiring foresights in the subject have guided me through the three-year journey of this work.

Second, I wish to express my gratitude to all the team members and in particular to Ms. Zou Xiaodan for her constant help and collaboration. My appreciation also goes to all the staff and students of the Signal Processing & VLSI lab, especially to Amit Bansal, Chen Xiaolei, Cheng San Jeow, Cheng Xiang, Hu Yingping, Li Yunlin, Muhammad Cassim Mahmud Munshi, Tan Jun, Wei Ying, Xue Chao, Yang Zhenglin, Yu Heng, Yu Rui, Zhang Jinghua and Zhu Youpan. Life with them has always been filled with joy and excitement.

I would also like to thank Mr. Teo Seow Miang and Ms. Zheng Huan Qun for their technical support, without which I would not have been able to make such smooth progress in my research.

This work is sponsored and coordinated by the Singapore agency for science, technology and research (A*STAR). My special thanks go to them for their financial and technical support.

Lastly, but most importantly, I am deeply indebted to my beloved father Xu Shu-wen and mother Sun Xiansu, whose love, remote support and constant confidence in me have always been my utmost motivation to overcome the obstacles and to dispel the clouds of confusion and frustration along the journey. I dedicate this thesis and all my accomplishments to them.

Contents

Summary

This work is about the design and implementation of ultra-low-power biomedical sensor interface circuits that are suitable for telemetric medical applications and in particular for wearable ECG devices. It is motivated by the increasing awareness and demand in pervasive and remote personal healthcare services due to population ageing; inspired and impelled by the rich options offered by today's microelectronic technology and material and biomedical sciences. Its preliminary outcome, as documented in the dissertation, is the world's first sub-μW ECG sensor interface chip.

The sensor interface chip integrates a low-noise frontend amplifier with programmable bandwidth and gain, and a 12-bit SAR ADC incorporating a dual-mode low-power clock module. The ultra-low power consumption is achieved through optimal system partitioning derived from the most efficient S/H duty ratio, and extensive applications of subthreshold circuit design techniques. A novel cross-coupled pseudo-resistor structure that favors both electrical balance and resistance tunability is proposed for onchip high-pass cutoff frequency tuning. The gain control is implemented by a novel "flip-over-capacitor" structure that eliminates the low frequency gain interruption due to the finite off-state resistance of the MOS switches. The dual-mode clock module offers options of both a more accurate crystal driver and a more power conserving relaxation oscillator, targeting applications with different power and accu-

racy requirements.

Fabricated in AMS 0.35-μm CMOS baseline process and operated at 1-V supply, the sensor interface chip features 0.6% of worst-case *THD*, 57 dB of dynamic range and 3.26 of *NEF* for the frontend amplifier; +0.8/−0.6 LSB of *DNL*, ±1.4 LSB of *INL* and 10.2 *ENOB* for the ADC. The power consumption for the entire chip is measured to be 445 nW in the minimum band *QRS* detection mode, and 895 nW in the full band ECG acquisition mode.

A miniature ECG plaster prototype based on the sensor interface chip and a commercial ZigBee transceiver is thereafter demonstrated. The captured ECG data are either stored locally to a Micro SD card or sent out to base stations or routers over ZigBee radio.

Also documented in the dissertation are some supportive information, considerations and analyses throughout the work. They include the introduction to the cardiac cycle, ECG signals and lead systems; the studies on the settling behavior and scalability of the first order S/H system, and on the static nonlinearity of the binary search capacitive DAC, etc.

List of Tables

List of Figures

List of Abbreviations

A/D	Analog-to-digital
ADC	Analog-to-digital converter
AV	Atrioventricular
CMOS	Complementary metal-oxide-semiconductor
CMRR	Common-mode rejection ratio
CT	The Wilson central terminal
D/A	Digital-to-analog
DAC	Digital-to-analog converter
DNL	Differential nonlinearity
DRL	Right-leg driver
DSP	Digital signal processor
ECG	Electrocardiogram
EEG	Electroencephalogram
ENOB	Effective number of bit
ESD	Electrostatic discharge
FFT	Fast Fourier transform
GB	Gain-bandwidth product
GE	Gain error
IC	Inversion coefficient
INL	Integral nonlinearity

LHP	Left-half-plane
LR	Linear mode voltage transition rate
LSB	Least significant bit
MSB	Most significant bit
NEF	Noise efficiency factor
OTA	Operational transconductance amplifier
PGA	Programmable gain amplifier
PSRR	Power supply rejection ratio
RHP	Right-half-plane
rms	Root-mean-square
SAR	Successive approximation register
SFDR	Spurious-free dynamic range
S/H	Sample-and-hold
SNDR	Signal-to-noise-plus-distortion ratio
SR	Slew rate
THD	Total harmonic distortion

List of Symbols

β	Feedback factor of a closed-loop system
η	Holding duty ratio of a S/H system
g_m	Transconductance of an active component

Chapter 1

Introduction

In recent years, personal telemetric medical system has attracted increasing attention as it reveals to be a promising solution to the overwhelming demand in healthcare industry due to population ageing. Based upon a prevention-oriented model and a pervasive, remote and continuous monitoring methodology, such system can buy doctors in-depth and real-time knowledge to patients' health conditions without much interference to their daily lives. As a direct benefit, precautionary measures and early treatments can be taken before serious disease attacks to save precious lives.

Similar to conventional biomedical devices, telemetric medical system needs to first of all capture and preprocess informative vital signs and physiological signals, and prepare them for further monitoring and diagnoses. This very frontend of the biomedical system chain is usually termed "sensor interface". At present, commonly used sensor interface circuits can capture bio-signals including body temperature, blood pressure, respiratory rate, electrocardiogram (ECG), electroencephalogram (EEG), etc. This work deals primarily with ECG signal and its corresponding sensor interface circuits that are tailored specifically for personal telemetric medical purposes.

However, many of the design techniques discussed here have been derived generically for ultra-low-power circuits, and can be readily applied to other biological forms.

The organization of this dissertation is as follows. Chapter 2 outlines a brief background of the ECG signal and its acquisition, provides an overview of the requirements and challenges in telemetric ECG sensor interface design, and reviews some of the popular solutions in the field. Chapter 3 describes the proposed system architecture that aims to achieve an optimal balance between performance and power consumption. Chapters 4 and 5 details the circuit level design challenges and the techniques proposed to hurdle them. The experimental results of the fabricated integrated circuit and the prototype wearable ECG device are demonstrated in Chapter 6. Chapter 7 concludes the work.

The results of this work were published and presented at 2008 Symposium on VLSI Circuits [1]; and published in the IEEE Journal of Solid-State Circuits [2]. Other publications include [3], [4] and [5].

Chapter 2

Overview of the ECG Signal and ECG Sensor Interface System

2.1 Background of the Human ECG and its Acquisition

2.1.1 Formation of the ECG Signal

The ECG signal reflects the electrical activities of a person's heart over time. Not only does it reflect his or her heartbeat, but it also provides greater insight to the detailed biological activities of the heart. Because it can be obtained through simple and nonintrusive procedures, the ECG signal has been one of the most sophistically studied and widely used indicators for diagnosing heart diseases.

Based on the early studies on dogs in the 1950s and the later similar studies on the human heart in the 1970s [6], it is commonly accepted that the ECG signal is essentially generated from the propagation of dipole wavefronts across the heart tissue that originate from the depolarization and repolarization processes in the heart cells. This is better understood from the illustrations in Fig. 2.1.

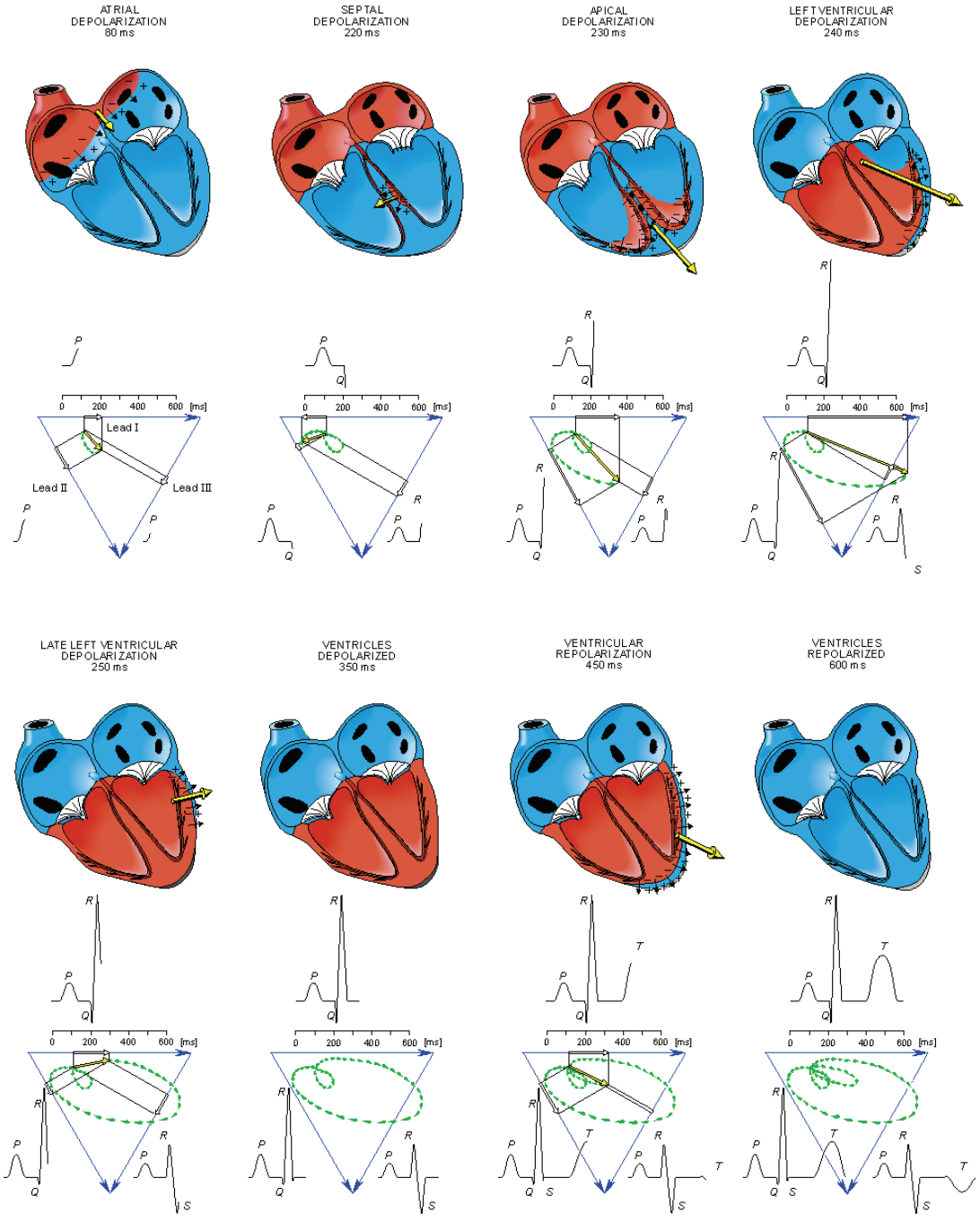


Figure 2.1: The formation of the ECG signal in the Einthoven limb leads [6].

The 3-vector triangle in each of the 8 phases represents the Einthoven limb leads configuration, which will be described later. The thick yellow vector denotes the resultant dipole from the depolarization/repolarization wavefronts. Assuming the human

body is a homogeneous medium, the projections of this dipole to the three limb leads form the actual voltage readouts obtained from the Einthoven configuration.

A brief description of the 8 phases in Fig. 2.1 is as follows. 1) The electric activation starts at the sinus node, and spreads along the atrial walls. The even propagation generates a positive *P* wave in all three limb leads. 2) After the depolarization wavefront has reached the atrioventricular (AV) node, it slows down and produces a few tens of milliseconds of flat response. Then the propagation proceeds along the inner walls of the ventricles and initiates the ventricular depolarization from the left side of the interventricular septum. This results in a negative *Q* wave in Leads I and II. 3) The ventricular depolarization now progresses on both sides of the septum, and produces a dipole pointing towards the apex, and in turn an upward *R* wave in all three leads. 4) The depolarization gradually propagates through the ventricular walls, with slower progress in the left ventricle due to thicker tissue. The resultant dipole vector turns leftwards, and the *R* wave in Leads I and II reaches maximum. 5) The depolarization in the left ventricle continues to the basal region. With the decrease of wavefront area, the dipole vector begins to drop and so does the *R* wave. 6) The ventricular depolarization now finishes. All leads return to rest state. 7) The ventricular repolarization starts from the epicardial surface of the left ventricular wall and diffuses inwards. This produces a positive *T* wave in Leads I and II and a negative one in Lead III. 8) The repolarization finishes and the heart is ready for the next cardiac cycle.

2.1.2 The ECG signal and the Cardiac Cycle

Fig. 2.2 depicts one cycle of the typical ECG signal obtained from Lead II and recorded on the standard ECG paper. The deflections are named in alphabetic order as *P*

wave, *QRS* complex, *T* wave and *U* wave respectively. The various segments and intervals are defined and used extensively in diagnoses.

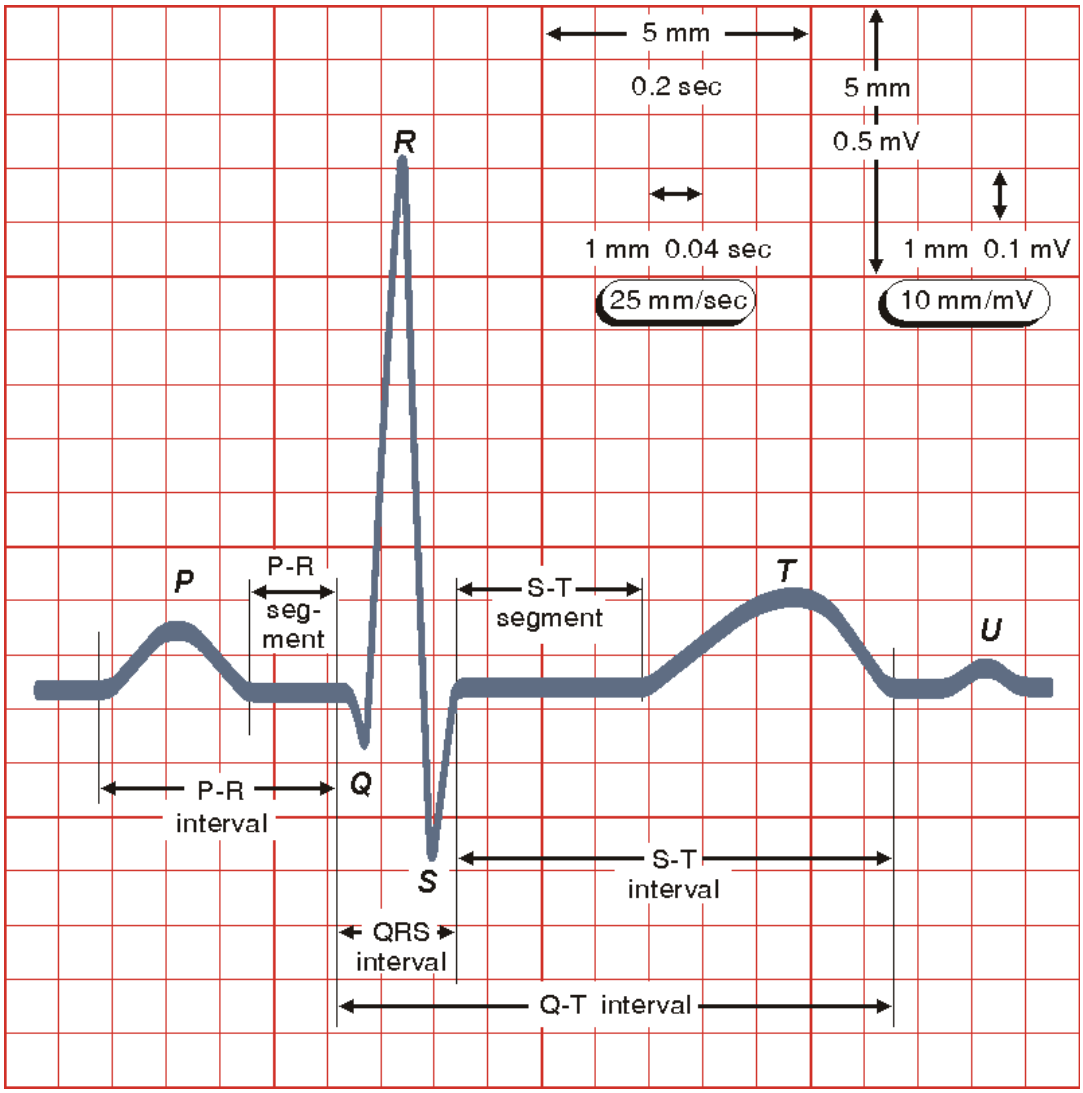


Figure 2.2: The normal ECG signal in one cardiac cycle [6].

The *P* wave corresponds to the atrial depolarization. The ventricular depolarization occurs during the *QRS* complex. The repolarization of the atria also takes place in this interval but is too small to be observed in the ECG. The *T* wave forms when the ventricles repolarize from activation. The formation of the *U* wave is not very clear

yet, and it is normally seen in 50% to 75% of ECGs [7].

In addition to direct profiling of the electric activities in the heart, the ECG signal also closely corresponds to other cardiac events and signals in each cardiac cycle, as illustrated in Fig. 2.3. Evidently, the ECG is essentially an *electric* view of the cardiac cycle.

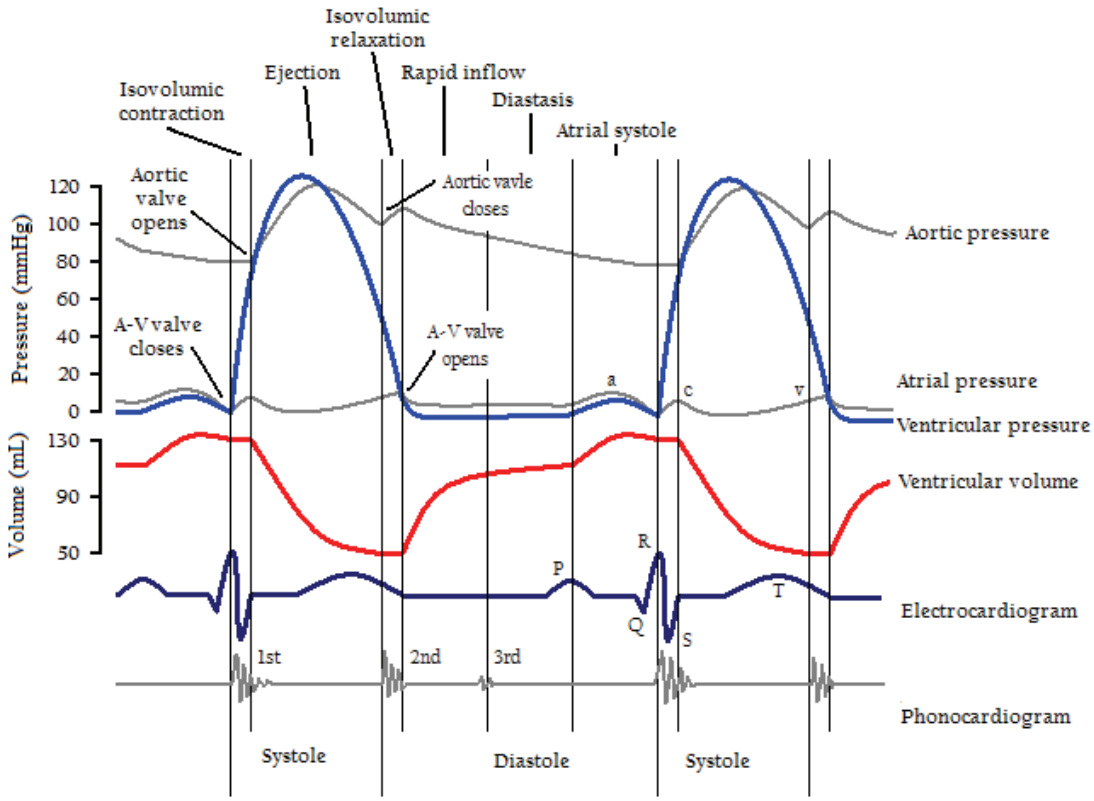


Figure 2.3: Two cycles of cardiac events in the left ventricle [8].

2.1.3 Lead Systems

The ECG signal is usually obtained from nonintrusive skin electrodes, and different probing sites and combinations can result in different lead configurations and different perspectives of the heart activities.

One of the most commonly applied lead systems in clinical diagnoses is the 12-lead configuration. It consists of 3 bipolar Einthoven limb leads, 3 unipolar augmented limb leads and 6 unipolar precordial leads.

The three Einthoven limb leads were proposed by Willem Einthoven in 1908 [6], and are formed by three electrodes attached to the right arm, the left arm and the left leg respectively. This is illustrated in Fig. 2.4, wherein the three lead vectors form the Einthoven triangle. Since all the three leads source their differential poles directly from the respective electrodes, they are termed bipolar leads.

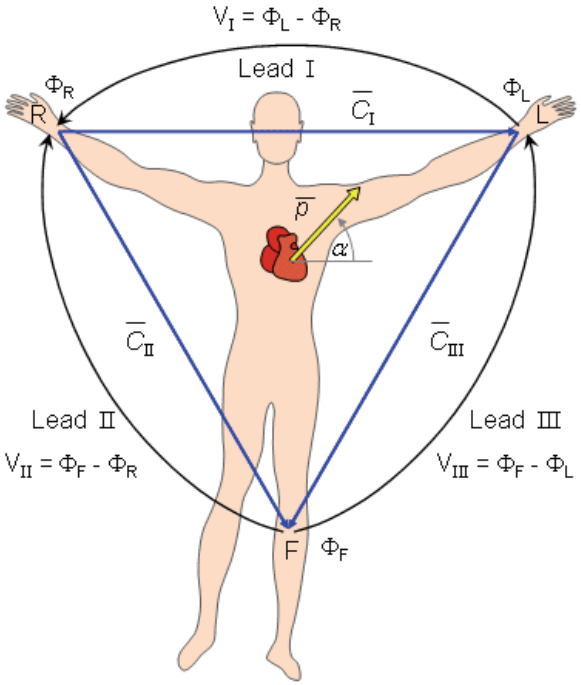


Figure 2.4: Einthoven limb leads and Einthoven triangle [6].

The rest nine leads are unipolar leads in the sense that each of them has only one true pole from one of the electrodes, with the other reference pole calculated from the signals acquired from many other electrodes.

One example of unipolar leads (not included in the 12-lead system) can be derived from the Einthoven triangle by averaging the potentials on the 3 limb electrodes to obtain the reference pole, as shown in Fig. 2.5. This reference pole is termed the Wilson central terminal (CT) after its inventor Frank Norman Wilson. The CT pole then pairs with the three limb electrodes/poles to form three unipolar limb leads.

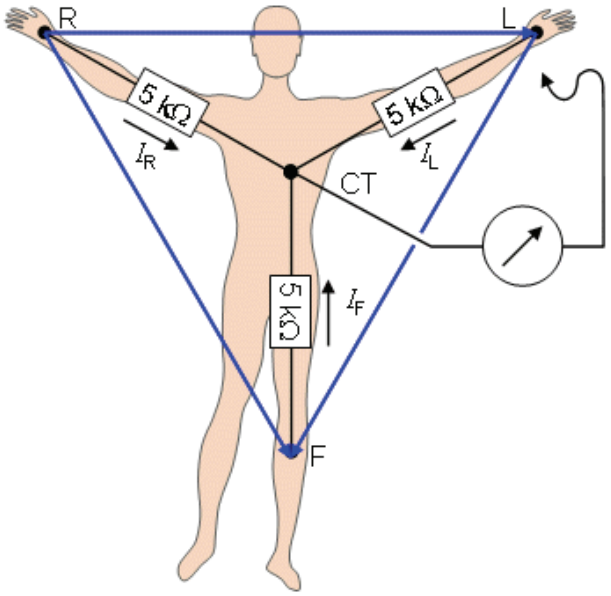


Figure 2.5: The Wilson central terminal (CT) [6].

In the 12-lead system, three unipolar limb leads are derived slightly differently, by omitting one of the three resistors in calculating the reference pole, as illustrated in Fig. 2.6. With the reference pole slightly bent towards the other two electrodes, the obtained unipolar leads aV_L, aV_F and aV_R are augmented version of the aforementioned unipolar limb leads. Therefore, they are termed augmented leads. It can be shown that the ECG signals obtained from the augmented leads are 50% higher than

their counterparts based on the CT reference pole.

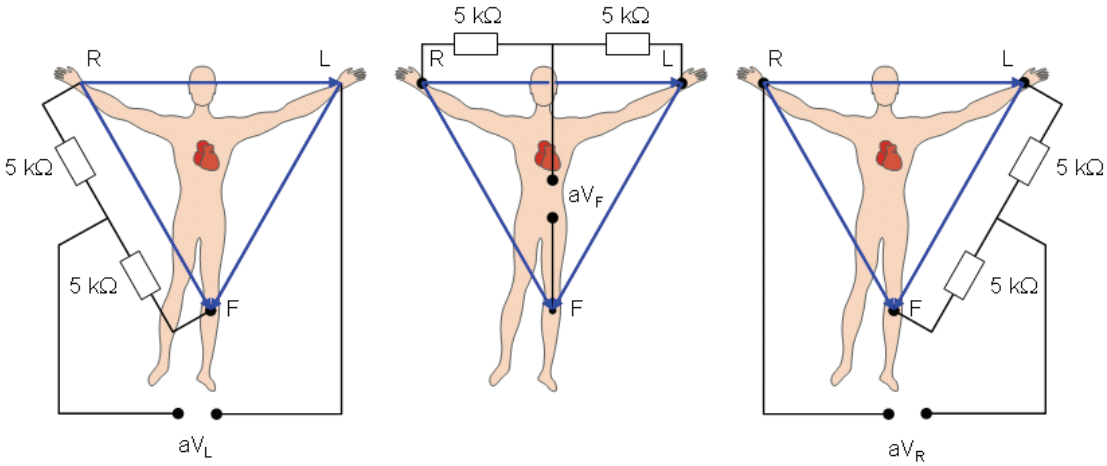


Figure 2.6: The three augmented limb leads in the 12-lead system [6].

The rest six precordial leads V_1 – V_6 in the 12-lead system are obtained from the chest electrodes as shown in Fig. 2.7.

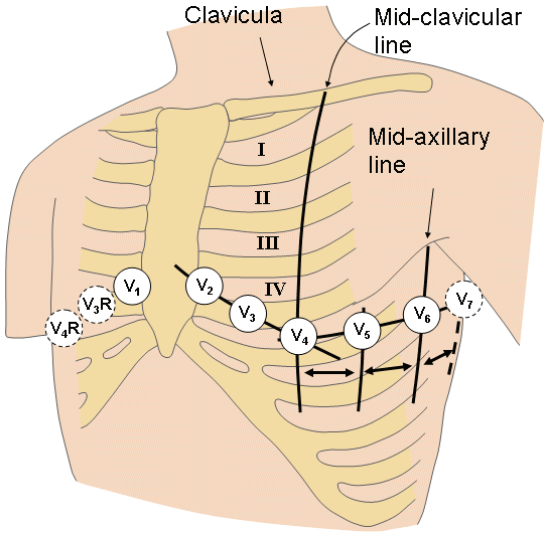


Figure 2.7: The precordial leads in the 12-lead system [6].

All six leads are unipolar leads that take the Wilson CT as the reference pole.

They provide a horizontal perspective of the heart activities, in contrast with the vertical views from the 6 limb leads.

Assuming the heart is an ideal dipole source and the human body is a homogeneous volume conductor, three vectors would be sufficient to describe all the heart activities. In other words, three independent leads, for instance, Leads I, II and V_2 can construct a complete heart model, whereas the rest nine leads are redundant. In reality, however, due to the distributed nature of cardiac sources and the inhomogeneity of body tissues, all the precordial leads are of diagnostic significance. Therefore, only four of the limb leads are redundant.

In telemetric ECG applications, especially in the context of wearable ECG devices, it is neither convenient nor necessary to have all the 12 leads or 10 electrodes (9 probing electrodes + 1 ground electrode) in most cases. Usually a limb lead, e.g. Lead II, or a precordial lead, e.g. Lead V_2, can tell much of the information the doctor needs for the patient monitoring.

2.2 Specifications of Telemetric ECG Sensor Interface

The general task of the ECG sensor interface system is to acquire ECG signals from the respective electrodes, filter and amplify them, and finally convert them into digital forms for easy storage, processing and lossless transmission. It is essentially an analog-to-digital frontend tailored for the ECG acquisition purpose.

2.2.1 General Requirements for ECG Sensor Interface

Like most application-specific systems, the ECG sensor interface system is usu-

ally customized to better cater the ECG signal. Some of the general considerations in customization are listed as follows.

a) Input range (differential mode)

The differential mode ECG signals acquired from the limb leads are normally in the range of a few hundred μV to one mV or slightly higher. The ones acquired from the precordial leads could be a bit higher, but still within a few mV. Typically, the differential input range of an ECG sensor interface is set to ±2.5 mV to ±5.0 mV.

b) Dynamic range

While the differential input range quantizes the upper signal rail of the dynamic range, the smallest feature size the sensor interface needs to resolve defines the lower rail. In a typical ECG recording, the smallest deflection that is of diagnostic significance, e.g. the *P* or *U* wave, can be well below 100 $\mu V_{p\text{-}p}$. It should be noted that in formal cardiac diagnoses, not only is the detection of such features alone useful, but the detailed resolution of the feature shapes is also of great importance. Therefore, the ECG sensor interface should provide at least one order of magnitude finer, i.e., lower than 10 $\mu V_{p\text{-}p}$, of effective resolution, which translates to a dynamic range of 54 dB to 60 dB. On the other hand, in coarse monitoring, where the larger characteristics such as the *QRS* complex are typically of greater interest, the requirement for the dynamic range is much relaxed to 42 dB to 48 dB.

c) Input range (common mode)

Depending on the architecture, the common mode input range or maximum DC

offset can be limited by various factors. In a conventional DC-coupled ECG frontend, it is mostly defined by the input range of the input amplifiers. In a complete AC-coupled circuit where input amplifiers are DC-isolated from the electrodes, it is usually limited by the DC limiting circuitry, e.g. the electrostatic discharge (ESD) protection module. Typically, this value can be safely set to a few hundred mV.

d) Common mode rejection ratio (*CMRR*)

The *CMRR* of a system is defined as the ratio of the differential mode gain over the common mode gain. This is a critical parameter in ECG sensor interface designs because the patients are often exposed to common mode interferences, among which the most common source comes from the power lines. Recall that in a standard 12-lead ECG system, in addition to the nine electrodes that form the twelve leads, a tenth electrode is required to level the common mode voltages of the human body and the sensor interface. If the contact resistance at this electrode is high, the sensor interface will have to take considerable amount of common mode injections produced by the interference current. Sometimes the common mode voltage drop across this tenth electrode due to the power line interference can be up to Volt level, approaching the supply voltage of the sensor interface system. If no other preventive measures are taken in this case, the sensor interface must keep the common mode gain below 1 to avoid extensive output saturation and signal distortions. For a system with 60-dB differential gain, this corresponds to at least 60-dB *CMRR* (*large* signal).

e) Input impedance

From the cardiac sources to any of the electrodes, the current path can be roughly

divided into two parts: the internal path and the skin-to-electrode contact. The internal resistance of the human body is usually in the range of 1 kΩ, which can be safely ignored. The skin-to-electrode contact resistance, on the other hand, can reach up to 100 kΩ according to [9]. This can create at least two problems on the sensor interface with poorly controlled input impedance. First, it degrades the signal seen by the sensor interface, as the contact resistance acts as the source resistance. The degradation could be "nonlinear" in the frequency domain, which may interfere with the bandwidth control. Second, when the two electrodes sourcing the differential input of the sensor interface differ considerably in contact resistance, the different source gains can transform any common mode interference, e.g. the power line noise, into differential mode signal. This will significantly degrade the *CMRR*. Therefore, the input impedance of the sensor interface must be orders of magnitude higher than the highest contact resistance. In practice, the typically used value is 10 MΩ @ 10 Hz.

f) Bandwidth and sampling frequency

At the lower end of the frequency domain, the ECG sensor interface system needs to filter out the DC offset and the baseline wander from the patients, which can originate from charge accumulation, perspiration, respiration, and body movements etc. A typical cutoff frequency used in diagnoses is 0.01 – 0.05 Hz. Sometimes a higher value of 0.5 Hz or above is chosen in *QRS* monitoring for better baseline filtering and faster settling. However, it should be noted that such a high cutoff point can distort the low frequency components in the ECG such as the *S-T* segment, and therefore should be avoided in formal diagnoses.

At the higher end of the frequency domain, the sensor interface system needs

about 150 – 300 Hz bandwidth to cover all the information in the ECG signal. However, since majority of the ECG energy resides below 30 Hz, it is also a common practice to set the cutoff at 30 Hz in the monitoring mode to save power.

If a Nyquist rate analog to digital converter (ADC) is used in the signal digitization, the filtered ECG must be sampled at over twice the signal bandwidth. Depending on the high pass cutoff frequencies, 500 S/s – 1 kS/s are common choices.

g) Gain adjustment

The ECG signals acquired in reality can differ considerably in magnitude. A tunable gain in this case helps to maintain the analog output level in a certain range, where the ADC resolution is fully utilized. This is especially true when the system resolution is bottlenecked by the ADC: the boost in gain for weak input reduces the *input-referred* quantization noise and hence counteracts the degradation of the effective dynamic range.

2.2.2 Special Requirements for Telemetric ECG Sensor Interface

For use in portable or wearable contexts, the telemetric ECG sensor interface system must be further optimized in the following aspects.

a) Battery life

One of the most desired features for a telemetric ECG sensor interface device, especially for a portable/wearable one, is ultralow power consumption. The ultra slim rechargeable batteries manufactured for good portability today usually have only a few hundred mAh of capacity. To operate the ECG device for weeks, the average cur-

rent consumption thereby should be strictly controlled within mA range. Because majority of the current has to go to the telemetry or storage circuit, the sensor interface module can only share some tens of μA or even lower. Fortunately, the sensor interface deals with low frequency and narrow bandwidth signals with medium dynamic range accuracy, which makes such low current consumption feasible.

b) Form factor

Small form factor is another essential feature for portable/wearable devices. One viable solution is to integrate as many functions as needed onto a single chip and to minimize the number of peripheral passive devices. For an ECG sensor interface system, this includes integrating the frontend amplifiers, the filters, the ADC, the reference generator, the clock generator, the standard I/O, simple digital signal processors (DSPs) and the local storage controller if possible.

2.3 Literature Review

Various micro-power ECG and other physiology sensor interface systems have been proposed and demonstrated in the past decades. While most of these designs have chosen subthreshold mode complementary metal-oxide-semiconductor (CMOS) circuits for best power efficiency and design compatibility, one common deficiency they are faced with is the need for bulky external RC components to implement the high cutoff function. Reid R. Harrison proposed a simple MOS-bipolar pseudo-resistor structure in [10], which uses small onchip active devices to generate huge resistance and eliminates such deficiency.

Fig. 2.8 shows a replica of Harrison's amplifier, wherein M_1, M_2, M_3 and M_4 form

four MOS-bipolar pseudo-resistors. According to [10], each pseudo-resistor can produce up to 10^{13} Ω incremental resistance at small signal level. Simple math can show that to obtain a high-pass cutoff frequency of 0.05 Hz, C_2 and C_4 need only to be in pF range or smaller. This means all the passive components in the amplifier loop can be economically integrated with today's CMOS processes.

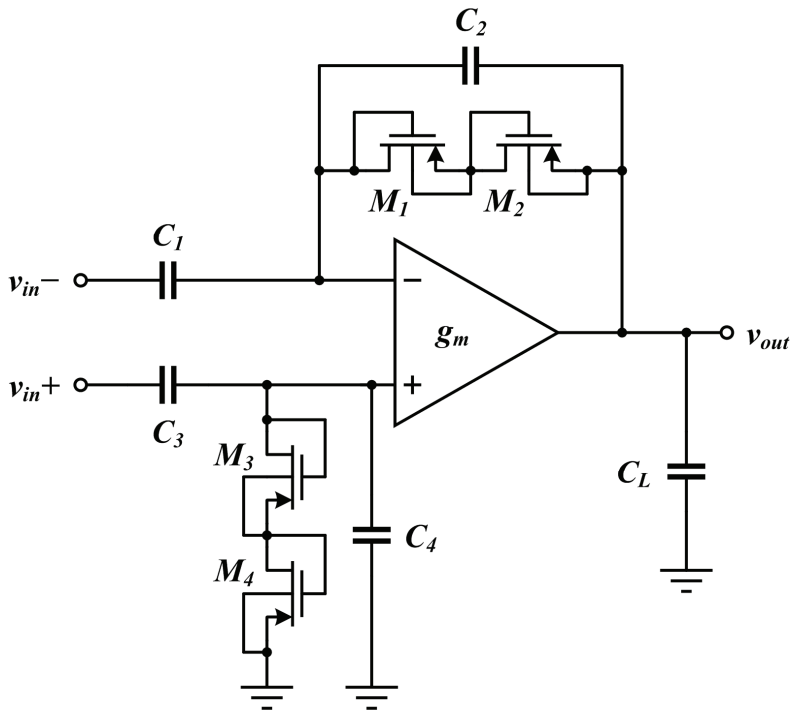


Figure 2.8: Harrison's neural amplifier with pseudo-resistors.

Due to its effectiveness and simplicity, Harrison's solution has been applied extensively in physiological amplifier designs such as [11] and [12]. It should be noted that the pseudo-resistor structure is not only effective in setting high-pass cutoff, but also suitable for most DC blocking circuits with proper customizations.

Later literature demonstrates efforts in integrating tunability into the pseudo-resistor structure in order to compensate its high dependence on process variations.

Ming Yin proposed in [13] a tunable pseudo-resistor structure composed of an n type transistor and a p type transistor, whose gate voltages are controlled by the bias circuitry, as shown in Fig. 2.9.

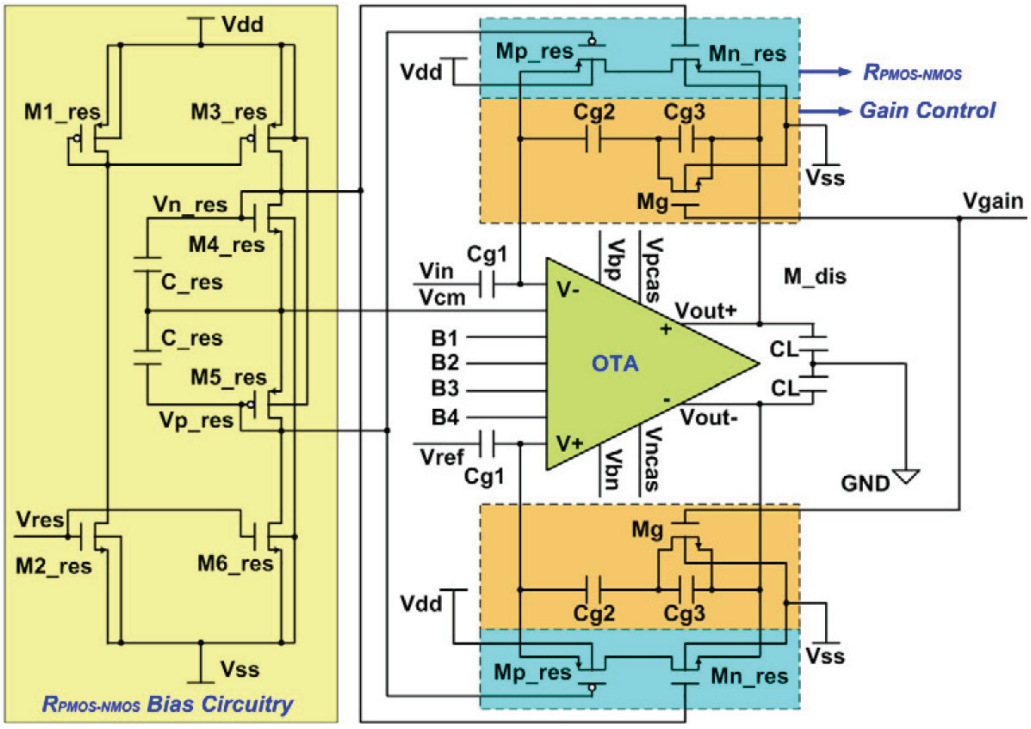


Figure 2.9: Ming Yin's amplifier with tunable pseudo-resistors [13].

M. Chae et al. used a simpler structure in [14], wherein the resistance is controlled by the gate bias V_B of the n transistors.

It is noticeable that most of the fixed pseudo-resistors and all of the tunable ones reported to date are topologically or electrically asymmetrical. This may introduce baseline drifting problems that can degrade the dynamic range. Detailed discussions can be found in Chapter 4.

To meet the requirements of the ECG sensor interface with maximum power effi-

ciency is another major concern. On top of various low power circuit techniques such as the aforementioned subthreshold approach, an efficient system level function and power breakdown can often be deterministic. This is also reflected in the previous examples [11] and [14]. Another example is Honglei Wu's design in [15], wherein a complete ECG sensor interface system is demonstrated (Fig. 2.10).

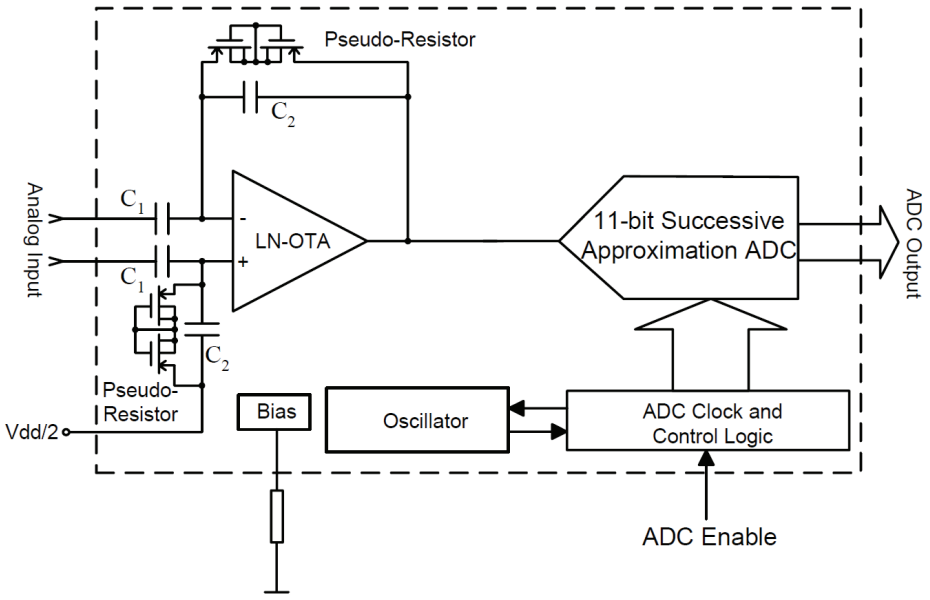


Figure 2.10: Honglei Wu's ECG sensor interface [15].

Here the low-pass cutoff and the sample and hold (S/H) functions are integrated into the frontend amplifier LN-OTA to conserve power. The tradeoff for the approach is substantially extended sampling period, which is then addressed by implementing an intermittent ADC conversion clock that is much higher than normally needed.

While [15] is a novel single channel design, it is not easy to port the approach to multi-channel ECG sensor interface systems. Moreover, the elevated clock speed can produce power overhead and reduce the efficiency. This dissertation will try to ad-

dress these problems from another approach, with a more flexible system architecture that is discussed in the next chapter.

Chapter 3

System Architecture Design[†]

3.1 The Settling Behavior of the First Order S/H System

Consider a S/H system whose higher order poles and zeros can be safely ignored. Assume that the S/H perturbation is injected to the input of the S/H amplifier through the feedback loop, and hence each sampling process can be considered as a step response of the amplifier.

3.1.1 Non-Return-to-Reference S/H without Slew

Let us first consider the non-return-to-reference S/H scheme, which initiates its value to the previous sampled result at the beginning of each sampling interval. If the S/H amplifier never falls into the slewing mode (whose criteria are discussed in Section 3.1.3), the step perturbation in each sampling interval is recovered by the linear settling that is controlled by the dominant pole of the amplifier. Typically, the tracking error at the end of each sampling process should be at least less than half of the least significant bit (LSB). Assuming that the maximum frequency of the input signal

[†] The discussions in this chapter assume linearity across the operation range unless otherwise stated.

is f_{sig}, the S/H amplifier has a −3 dB bandwidth f_{SHA}, and the S/H circuit works with an oversampling rate k, the above rule-of-thumb criterion can then be expressed against a rail-to-rail sine wave input $A\sin(2\pi f_{sig}t)$ as follows:

$$\left.\frac{dA\sin\left(2\pi f_{sig}t\right)}{dt}\right|_{\max}\cdot\frac{\eta}{2kf_{sig}}\cdot e^{-\frac{(1-\eta)}{2kf_{sig}}\cdot 2\pi f_{SHA}}<\frac{1}{2}\cdot\frac{2A}{2^n}, \tag{3.1}$$

where n is the resolution of the S/H system, and η denotes the duty ratio of the holding interval in each S/H cycle. Solving the inequality gives

$$\frac{f_{SHA}}{f_{sig}}>k\cdot\frac{n\ln 2+\ln\left(\pi\eta/k\right)}{\pi\left(1-\eta\right)}. \tag{3.2}$$

The tracking error calculated in Inequality 3.1 can be derived differently based on the fact that the step perturbation at the S/H amplifier output is no larger than $2A$.

$$2A\cdot e^{-\frac{(1-\eta)}{2kf_{sig}}\cdot 2\pi f_{SHA}}<\frac{1}{2}\cdot\frac{2A}{2^n}, \tag{3.3}$$

$$\Rightarrow\frac{f_{SHA}}{f_{sig}}>k\cdot\frac{\left(n+1\right)\ln 2}{\pi\left(1-\eta\right)}. \tag{3.4}$$

It should be noted that both Inequalities 3.2 and 3.4 are sufficient conditions for the precision criterion to hold. Comparing the difference between the two, one may conclude that the relation between the factor $\pi\eta/k$ and the constant 2 determines which requirement is more relaxed. Given that η is a fraction of unity, and k is usually greater than 1 in reality to avoid the aliasing problem, Inequality 3.2 serves well for single-channel designs in most cases (whereas Inequality 3.4 is a more suitable and convenient choice for multi-channel designs).

3.1.2 Return-to-Reference S/H without Slew

Another group of S/H systems use the return-to-reference scheme, whereby the captured value is reset to a fixed reference point at the beginning of each sampling interval. Assuming that one such system tracks the input in a linear mode, and the fixed reference point is set at 0, then a similar analysis can be applied to a rail-to-rail sine wave input $A\sin(2\pi f_{sig}t)$ as

$$A \cdot e^{-\frac{(1-\eta)}{2kf_{sig}} 2\pi f_{SHA}} < \frac{1}{2} \cdot \frac{2A}{2^n}, \tag{3.5}$$

where the worst case step perturbation is A. Solving the inequality gives

$$\frac{f_{SHA}}{f_{sig}} > k \cdot \frac{n \ln 2}{\pi(1-\eta)}. \tag{3.6}$$

Obviously, the requirement here is more relaxed than that in the non-return-to-reference case.

3.1.3 S/H with Slew

In most cases when the step perturbation is large enough, the S/H amplifier will be forced into slewing mode until the limited driving current can eventually support the linear settling. For simplicity, we will model the amplifier as a two-phase piecewise system, i.e., both the slewing and linear modes are constant by themselves, and the transition between the two is abrupt.

Consider a closed-loop S/H amplifier constructed by an ideal two-stage operational transconductance amplifier (OTA) and a feedback factor β. The OTA is com-

pensated by the Miller capacitor C_C, and its first stage current is I_{D1}. The slew rate (SR) is typically limited by

$$SR = I_{D1}/C_C \,. \tag{3.7}$$

In linear mode, on the other hand, the voltage transition rate (LR) is given by

$$LR = \beta v_{step} \cdot g_m\left(I_{D1}\right)/C_C \,, \tag{3.8}$$

where v_{step} is the step perturbation at the amplifier output, and $g_m()$ denotes the function of the transistor transconductance against its drain current.

From Equations 3.7 and 3.8, one may conclude that the transition point occurs at

$$\beta v_{step} \cdot g_m\left(I_{D1}\right) - I_{D1} = 0 \,. \tag{3.9}$$

In reality where the transconductance is also a function of the input step, the equation still holds, provided $g_m()$ used is the average value across βv_{step}.

It is evident from the above discussion that the mixed mode settling is slower than the pure linear settling. When the transconductance is large enough, we make the assumption that the settling process is dominated by the slew. Applying the same analysis as in Section 3.1.1 to a non-return-to-reference system gives

$$\left.\frac{dA\sin\left(2\pi f_{sig} t\right)}{dt}\right|_{\max} \cdot \frac{\eta}{2kf_{sig}} - SR \cdot \frac{1-\eta}{2kf_{sig}} < \frac{1}{2} \cdot \frac{2A}{2^n} \,, \tag{3.10}$$

which yields an *optimistic* requirement of *SR*

$$SR > 2Af_{sig} \cdot \left(\pi\eta - 2^{-n}k\right)/\left(1-\eta\right). \tag{3.11}$$

If $\pi\eta$ is much larger than $2^{-n}k$, the inequality simplifies to

$$SR > 2\pi A f_{sig} \cdot \frac{\eta}{1-\eta}. \tag{3.12}$$

Note that the *SR* should be even larger in reality to compensate for the slow linear settling that follows the slew. On the other hand, thanks to the nature of the S/H perturbation, some special effects may help to dynamically boost the *SR*, which will be discussed in Chapter 4.

3.2 The Proposed System Architecture

A further examination on the discussions in Section 3.1 reveals two important facts: 1) The static input stage current of the S/H amplifier, together with its resultant amplifier bandwidth and *SR*, are key factors that determine the S/H accuracy; 2) It is these parameters and the S/H duty ratio η among others that ultimately define the appropriate system architecture.

To elaborate on the second point, let us revisit the results in Section 3.1. For simplicity, we only consider the linear settling in the non-return-to-reference case, and assume k=1. Hence Inequality 3.2 becomes

$$\frac{f_{SHA}}{f_{sig}} > \frac{n\ln 2 + \ln(\pi\eta)}{\pi(1-\eta)}. \tag{3.13}$$

One approach to realize such a system is to implement f_{sig} and f_{SHA} in the same OTA so as to reduce the number of active components and hence the power consumption. However, for Inequality 3.13 to hold in this case, η must not exceed 0.007 for a

10-bit system. Typically, this means that the data conversion needs to finish within 0.7% of the whole S/H cycle. Hence, a much elevated conversion clock speed has to be applied.

One obvious drawback of such approach is the potential power overhead due to the high speed clock control circuit. Moreover, a further increase in resolution or multiplexing for two or more channels can make the implementation impractical or even impossible. Therefore, this approach is more suitable for single channel applications with medium resolution.

An alternative approach is to decouple f_{sig} and f_{SHA} into separate OTAs, so as to relax the requirement in η. For instance, when the S/H amplifier is designed to have f_{SHA} five times as wide as the signal bandwidth f_{sig}, η can be set to a more convenient value of 0.5. Moreover, options for higher resolution or multi channels become well feasible simply through scaling f_{SHA}.

Therefore, a scalable low-power sensor interface architecture is proposed based on this approach, of which a single channel instance is illustrated in Fig. 3.1. The analog frontend (of each channel) is divided into two steps: a low noise preamplifier that incorporates reconfigurable band-pass function, and a programmable gain buffer (PGA) that drives the S/H circuit (single or multiplexing switch). A successive approximation register (SAR) ADC is chosen as the quantization module due to its good trade-offs between power efficiency, conversion accuracy and design complexity in the biomedical context. The timing of the entire system is sourced by either the crystal oscillator or the relaxation oscillator. The former guarantees superior clock accuracy, but draws more current and requires an external quartz crystal; whereas the latter

draws less current at the cost of larger clock jitters. The programmable dual-clock scheme allows for flexible system-level power management to cater different power and accuracy requirements.

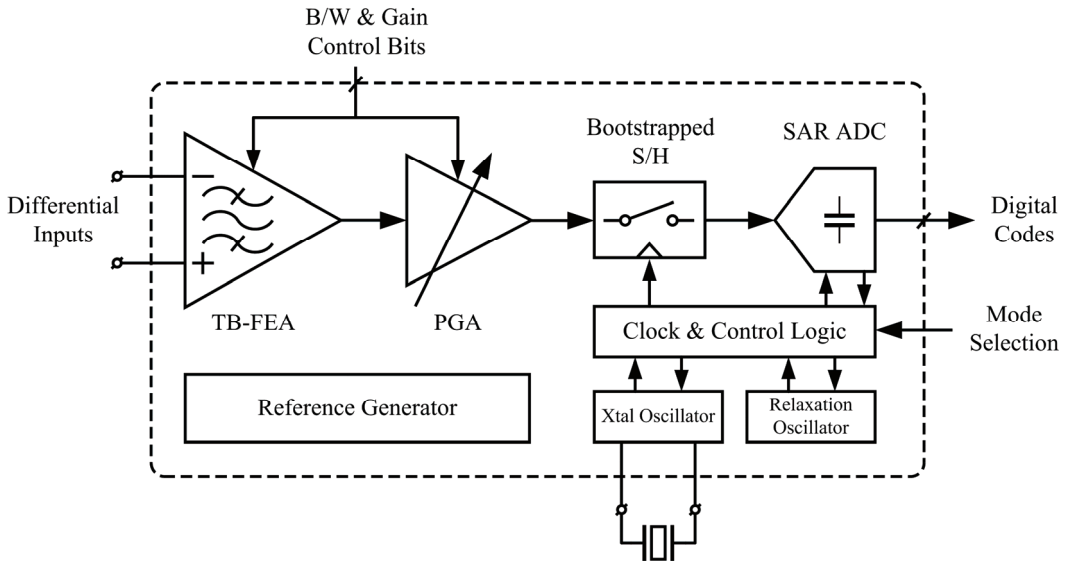


Figure 3.1: The proposed scalable low-power sensor interface architecture.

In addition to the aforementioned design flexibility and scalability, the proposed architecture also has the advantages of 1) The system pass-band no longer alters with different gain settings, as opposed to the single amplifier approach in [13] where the gain-bandwidth product is fixed; 2) The incorporation of the secondary gain stage (PGA) suppresses the signal swing of the preamplifier, hence reducing the risk of excessive distortions when nonlinear pseudo-resistors are employed in the preamplifier for bandwidth tuning.

3.3 System Level Power Optimization

Further studies on the function and parameter partitioning of the proposed archi-

tecture reveal the optimal specifications for each module, which yield the best overall power efficiency. As a starting point, for instance, given the desired input range of ±2.5 mV and power supply of 1 V, the minimum system gain is 200. Since in general the gain of the PGA G_{PGA} should be at least one order of magnitude smaller than that of the preamplifier G_{PRE}, it is reasonable to fix G_{PRE} at 100 and set G_{PGA} to be adjustable starting from 2.

Next, with the system dynamic range or resolution determined, the power consumptions of the preamplifier and the ADC are relatively fixed, while those of the PGA and the clock and timing sequence modules vary as a function of η. Hence, finding the optimal power efficiency becomes a mathematical problem of solving the global minimum of such function.

a) Finding the static current of the PGA

In the PGA, the gain-bandwidth product GB_{PGA} is determined by the input transconductance g_{mPGA} and the Miller capacitor C_C as

$$GB_{PGA} = g_{mPGA}/C_C \,. \tag{3.14}$$

By definition, it is also given by

$$GB_{PGA} = 2\pi \cdot G_{PGA} \cdot f_{SHA} \,. \tag{3.15}$$

Thus, one has

$$g_{mPGA} = 2\pi \cdot G_{PGA} \cdot C_C \cdot f_{SHA} \,, \tag{3.16}$$

where g_{mPGA} is a direct function of the input stage current I_{DPGA} of the PGA.

Since a telemetric ECG sensor interface system requires only medium resolution but superior power efficiency, a high transconductance efficiency is generally more desirable than a high transconductance itself. Therefore, it is preferable to bias the input transistors in subthreshold region where the transconductance efficiency tops. The drain current I_D of an n type transistor in this mode follows

$$I_D = I_0 \cdot \frac{W}{L} \cdot e^{\frac{V_{GS}-V_{TH}}{NU_T}} \cdot \left(1 - e^{-\frac{V_{DS}}{U_T}}\right), \tag{3.17}$$

where I_0 is a characteristic current defined by the process parameters, N is the subthreshold swing parameter, and U_T is the thermal voltage. Solving for the transconductance gives

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = I_0 \cdot \frac{W}{L} \cdot e^{\frac{V_{GS}-V_{TH}}{NU_T}} \cdot \left(1 - e^{-\frac{V_{DS}}{U_T}}\right) \cdot \frac{1}{NU_T} = \frac{I_D}{NU_T}. \tag{3.18}$$

Substituting Equation 3.18 into Equation 3.16 and solving for I_{DPGA}, one has

$$I_{DPGA} = 2\pi \cdot N \cdot U_T \cdot G_{PGA} \cdot C_C \cdot f_{SHA}. \tag{3.19}$$

b) Finding the average current of the clock and timing sequence circuits

The clock and timing sequence modules are mostly digital circuits. Assuming that the leakage and short circuit currents are negligible, the average current can be estimated as

$$I_{DCLK} = C_{CLK} \cdot V_{DD} \cdot f_{CLK}, \tag{3.20}$$

where C_{CLK} is the effective load capacitance seen by the driving clock f_{CLK}, and can be

readily extracted from a digital circuit once its topology is determined. For a *single* channel instance of the proposed architecture, where the clock is assumed to be running *continuously*, the clock frequency normally follows

$$f_{CLK} = 2nkf_{sig} / \eta \,. \tag{3.21}$$

Thus Equation 3.20 becomes

$$I_{DCLK} = 2nk \cdot C_{CLK} \cdot V_{DD} \cdot f_{sig} / \eta \,. \tag{3.22}$$

c) Finding the optimal overall current

The total variable current consumption I_{DVAR} includes both the static current from the PGA and the average current from the clock and timing sequence circuits. Hence,

$$I_{DVAR} = I_{DPGA} + I_{DCLK} \,. \tag{3.23}$$

Substituting Equations 3.19, 3.22 and the lower limit of Inequality 3.2 into the above equation, one obtains the function of I_{DVAR}, whose global minimum may be solved by taking its derivative against η.

The dependence of the total variable current and its components over η can be better visualized in Fig. 3.2. Evidently, for the single channel instance of ECG sensor interface system, the minimum I_{DVAR} occurs when η is approximately 0.82. This signifies the optimal partitioning point where the overall system achieves the best power efficiency. Also revealed by the graph is that the total current displays very little change when η varies from 0.7 to 0.9. Therefore, designing around 0.8 should produce a reliable result that is relatively resistant to non-idealities and variations.

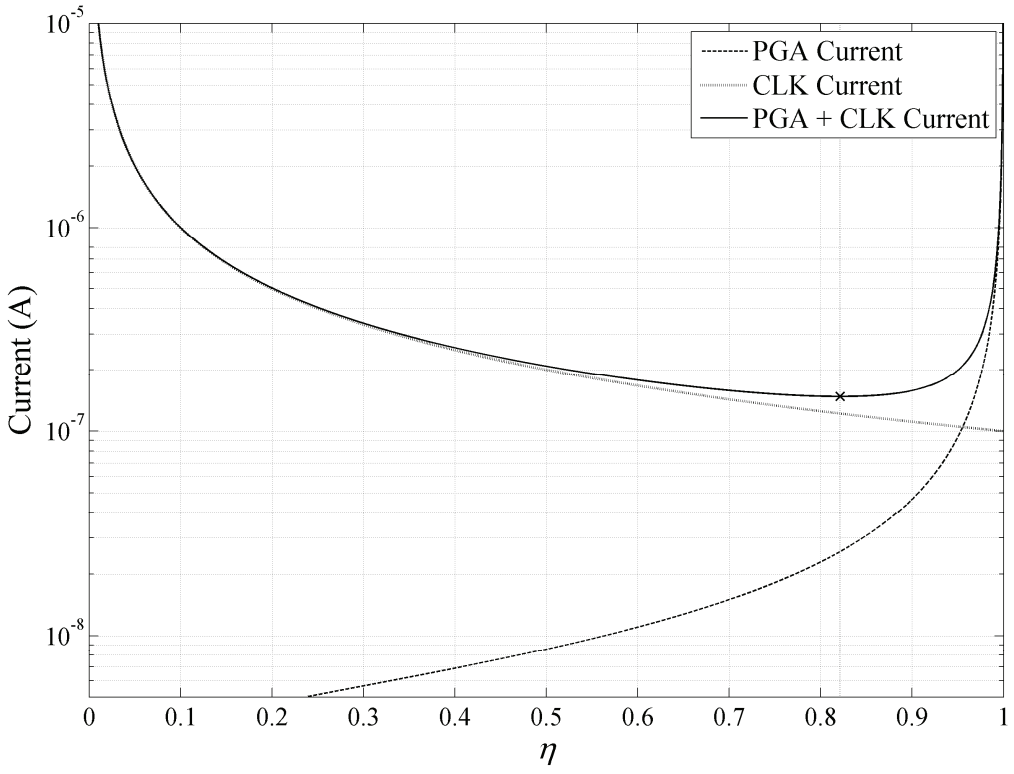


Figure 3.2: The total variable current and its components versus η.

With the architecture optimized for low power and scalable telemetric ECG sensor interface systems, the following chapters will deal with the circuit level design and optimization of individual modules.

Chapter 4

Frontend Design

The frontend of the proposed ECG sensor interface system includes a low noise preamplifier with integrated tunable band-pass function, a PGA, and other analog auxiliary circuits. In the following sections, the kernel modules that have been optimized for power efficiency, reconfigurability and scalability are described in details.

4.1 Balanced Tunable Pseudo-Resistor

4.1.1 Conventional Pseudo-Resistor Structures

a) Fixed imbalanced pseudo-resistor structures

The original form of pseudo-resistor, as described in Chapter 2, behaves as a MOS-bipolar hybrid during operation [10]. This is better understood from its cross-sectional view as illustrated in Fig. 4.1. When $V_A > V_B$, the device is equivalent to a diode-connected p type MOS transistor; When $V_A < V_B$, the device is equivalent to a diode-connected pnp bipolar transistor. Here cautions must be taken as when a diode-connected active device is weakly biased, its current tracks the bias voltage with the

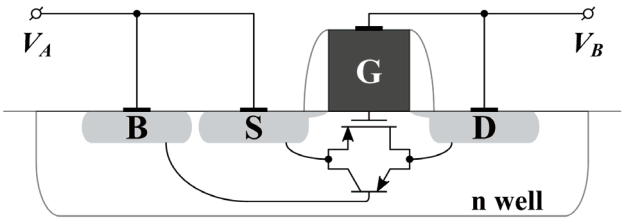


Figure 4.1: The cross-sectional view of a p type MOS-bipolar pseudo-resistor (not to scale).

exponential rule, and so does its effective resistance. Thereby the device will quickly fail as a high value resistor when the voltage drop across it reaches certain threshold. For a typical capacitively coupled amplifier as shown in Fig. 2.8, such threshold occurs when the effective resistance approaches the reactance of feedback capacitor in the useful frequency band.

Fig. 4.2 plots the simulated incremental resistance against the bias voltage for a 1-µm/8-µm p type pseudo-resistor in a standard 0.35-µm CMOS process. It should be noted that the flat top of the curve is not necessarily reflecting the saturation of resistance, but due to the precision setting of the SPICE simulator instead. The actual upper limit the pseudo-resistor can reach depends on multiple factors such as temperature, the aforementioned characteristic current I_0, and structural imperfections/defects among others.

Assume that a 1-pF feedback capacitor is employed in an application whose informative band extends down to 1 Hz. The maximum reactance of the capacitor hence is in the order of 10^{11} ohm. Any resistance degradation of the pseudo-resistor across this order of magnitude can cause the high-pass cutoff point to sweep over the useful band, and introduce harmonic distortions. Consequently, the voltage swing across the

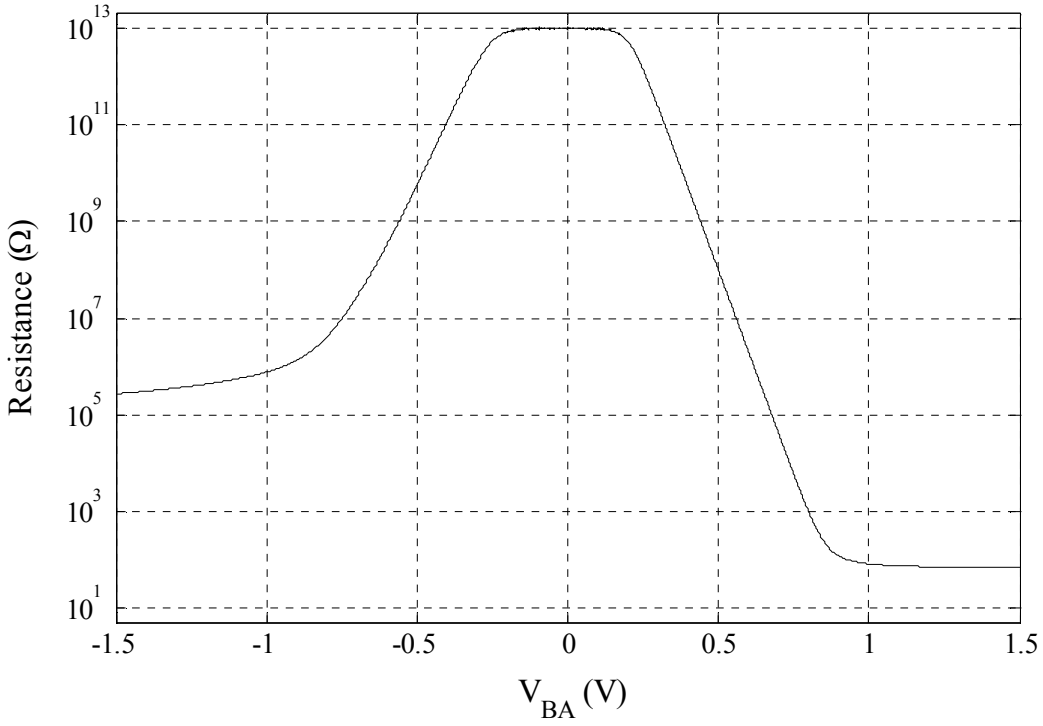


Figure 4.2: Simulated resistance of a p type MOS-bipolar pseudo-resistor.

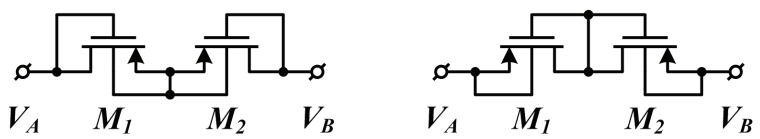
pseudo-resistor should be restricted such that the resistance never drops to the degree of being comparable to the reactance of the feedback capacitor in the informative band. Phrased another way, the high-pass cutoff point defined by the pseudo-resistor and the feedback capacitor must always stay well below the informative band even at full signal swing. Applying this rule to the pseudo-resistor in Fig. 4.2, it can be concluded that its useable swing for the said application is only approximately ±0.25 V.

To overcome this limit in practice, two or more pseudo-resistors are usually applied in series so as to sustain large voltage swings with acceptable level of distortion. In [10], two identical diode-connected transistors are paired up in the same orientation (Fig. 2.8), such that the amplifier can utilize the ±2.5-V supply range more effectively.

Also observed from Fig. 4.2 is that the incremental resistance is not symmetrically distributed along the bias condition where V_A=V_B (or V_{BA}=0). The pseudo-resistor structure in this case is said to be "imbalanced". The imbalance of resistance can cause the voltage drops in opposite directions to be different at any finite swing levels, if the net charge flow through the resistor is zero at steady state. One obvious outcome of such effect, when applied in the amplifier circuit in Fig. 2.8, is that the amplifier output will gradually deviate from the reference level and shift towards one of the power rails, as the signal swing increases. Eventually, the amplifier will see an early output clipping, which may degrade the dynamic range.

b) Fixed balanced pseudo-resistor structures

Another commonly used approach of stacking up pseudo-resistors, as opposed to that in [10], is to pair one MOS-bipolar pseudo-resistor with its mirrored replica, as shown in Fig. 4.3. In addition to the ability of sustaining wider swing range, such configuration also features the desired topological symmetry. It should be noted that p type transistors are generally preferred over their n type counterparts in constructing pseudo-resistors. This is because most standard CMOS processes do not provide the deep n well option, and all the n type transistors in this case have to share the same substrate that is tied to a fixed potential (ground). As a result, the bulk-to-source bias



(a) (b)

Figure 4.3: Two examples of fixed balanced pseudo-resistors.

combined with the leakage currents from the reversely biased junction diodes form a *third* terminal for the pseudo-resistor, which makes the structure electrically imbalanced inherently (even if it is topologically symmetrical).

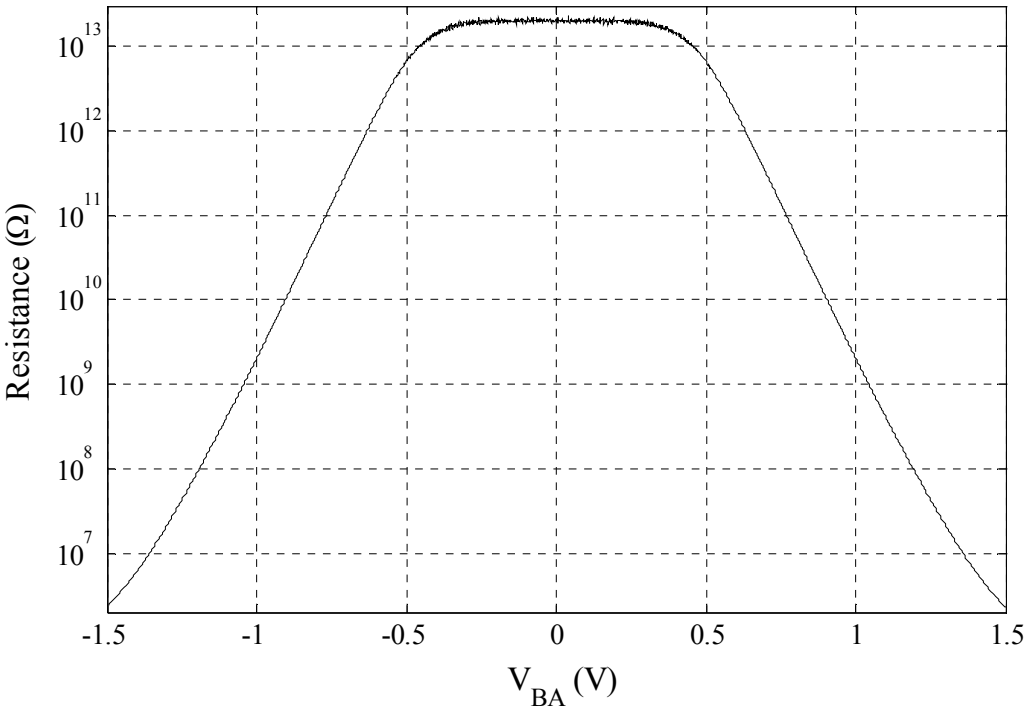


Figure 4.4: Simulated resistance of the fixed balanced pseudo-resistors in Fig. 4.3.

Fig. 4.4 shows the simulated resistance of the pseudo-resistors in Fig. 4.3. The two structures produce almost identical resistance across the ±1.5 V bias range. The useable range for the said application in Part a) is extended to ±0.65 V.

While the resistance curve is symmetrical along zero bias, it should be noted however that schematic level SPICE simulations do not automatically include the well-to-substrate diode leakage. Such leakage, in reality, poses a weak third terminal on devices such as that in Fig. 4.3 (a), and causes intrinsic DC offset (i.e. offset at zero sig-

nal swing) and swing dependent level shift. The structure in Fig. 4.3 (b) may be less affected if both terminals V_A and V_B are compensated by low impedance sources.

c) Tunable imbalanced pseudo-resistor structures

Tunability of pseudo-resistors is mostly realized by implementing external variable sources that can change the transistor inversion level and hence the incremental resistance. However, such approach may introduce at least one more *independent* terminal for external tuning, which automatically breaks the electrical balance if any.

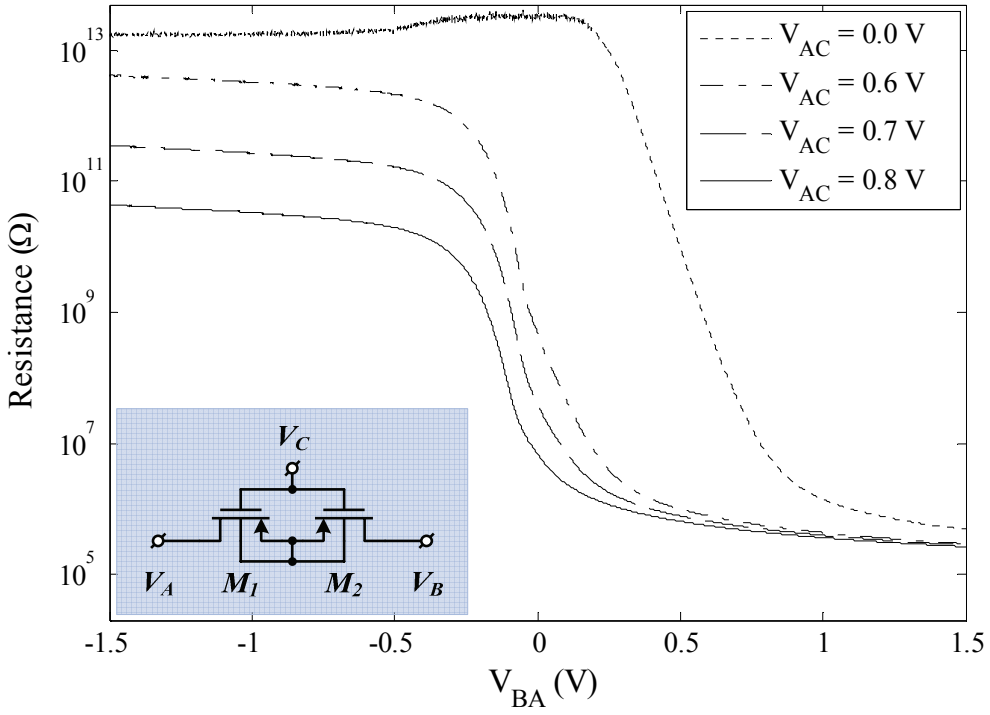


Figure 4.5: An example of tunable pseudo-resistor and its simulated resistance.

One example of commonly used tunable pseudo-resistors is illustrated in Fig. 4.5, where the resistance can be adjusted by varying the control voltage V_{AC}. Although the structure is topologically symmetrical, the incorporation of the third independent ter-

minal V_C makes it electrically imbalanced, as evidently reflected by the simulation data. When applied in a high-pass amplifier circuit such as the one in Fig. 2.8, the DC level of the output will eventually settle to the left half plane (V_{BA}<0) of the resistance plot, such that the net charge flow in the pseudo-resistor is zero. As the control voltage V_{AC} increases, the problem becomes more pronounced. In telemetric biomedical applications where the supply voltage is usually aggressively suppressed for power saving, such deviation of DC level can lead to early clipping of output and degradation of dynamic range.

4.1.2 The Proposed Cross-Coupled Tunable Pseudo-Resistor

From the above discussions, it can be concluded that electrical imbalance and resistance nonlinearity are two of the most significant drawbacks of conventional pseudo-resistors. While the latter may be rectified for a given signal swing range through stacking up replicas, the former generally gets worse with tunability incorporated. Therefore, it has become one of the goals of this work to design an *electrically* balanced tunable pseudo-resistor.

Strict electrical balance requires two necessary conditions. First, the device must be topologically symmetrical on its resistive components. Second, the device must be characterizable by a *two*-terminal model, resembling a pure resistor. Denoting the two terminals as V_A and V_B, and the difference between the two as V_{BA}, the incremental resistance r should follow

$$r = f\left(V_B - V_A, V_{CTRL}\right) = f\left(V_{BA}, V_{CTRL}\right), \tag{4.1}$$

where V_{CTRL} is the *internal* DC source that controls the transistor inversion level.

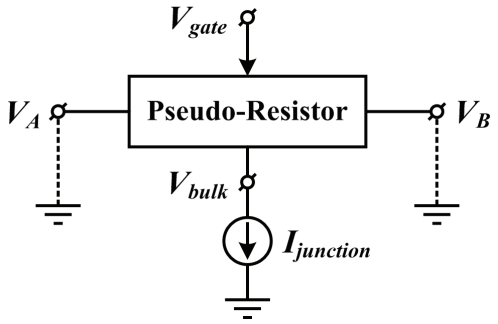


Figure 4.6: The 4-terminal model for a tunable pseudo-resistor.

Fig. 4.6 shows an abstract model for a generic tunable pseudo-resistor built on MOS transistors. V_{gate} and V_{bulk} denote the gate and body biases of the resistive MOS transistors. To ensure the four-terminal model is reducible to the two-terminal model, V_{gate} must be obtained *dependently* by

$$V_{gate} = g\left(V_A \; and \,/\, or \; V_B, V_{CTRL}\right); \tag{4.2}$$

V_{bulk} must be obtained *dependently* by

$$V_{bulk} = y\left(V_B \; and \,/\, or \; V_A\right). \tag{4.3}$$

The bulk junction leakage $I_{junction}$ needs to be compensated by low impedance sources in function y to avoid intrinsic DC offset. The resultant resistance control factors (e.g. the inversion levels or overdrives of transistors) must follow

$$rcf = z\left(V_{BA}, V_{CTRL}\right). \tag{4.4}$$

To make the design process more straightforward, one may consider the following guidelines. 1) Ensure topological and geometrical symmetry across the resistive structure; 2) Condition the impedance across the resistive structure boundary to compen-

sate for direct junction interface of local high impedance nodes with external sources; 3) Source V_{CTRL} solely from V_A or V_B or *local* low impedance nodes of the structure; 4) Through adaptively biasing V_{gate} by function g and V_{bulk} by function y, make sure the resistance control factors are decoupled from other external states than V_{BA}, and only dependent on V_{BA} and V_{CTRL} (preferably on V_{CTRL} alone for better linearity).

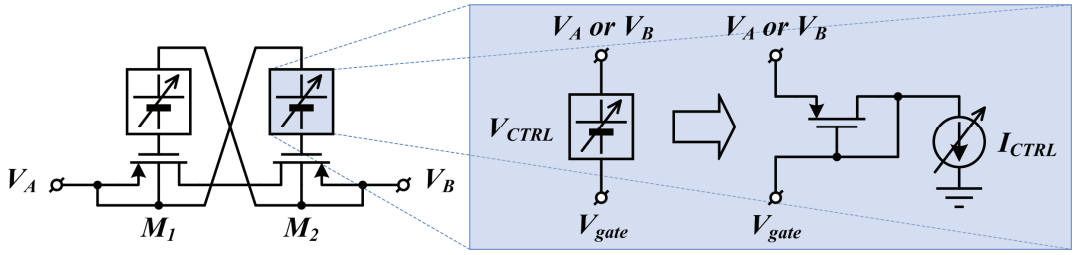


Figure 4.7: The proposed cross-coupled tunable pseudo-resistor.

Following the above conditions and guidelines, we proposed the fully balanced cross-coupled tunable pseudo-resistor structure as illustrated in Fig. 4.7. Two p type MOS transistors in series form the resistive path with two external terminals V_A and V_B. The gate biases V_{gate}s are generated by the internal variable DC sources V_{CTRL}s based on the terminal voltages in a cross-coupled configuration. The body biases V_{bulk}s are obtained locally from the two terminals. The variable voltage sources V_{CTRL}s are further mapped through diode-connected converters to variable current sources I_{CTRL}s, which are usually more accurately and linearly controllable.

When V_A and V_B are powered by low impedance sources, it is evident that the proposed structure is electrically balanced with no intrinsic DC offset or swing dependent level shift. Its operation when used in an amplifier feedback path as in Fig. 2.8 can be described as follows. Assume that V_A is connected to the negative input terminal of an

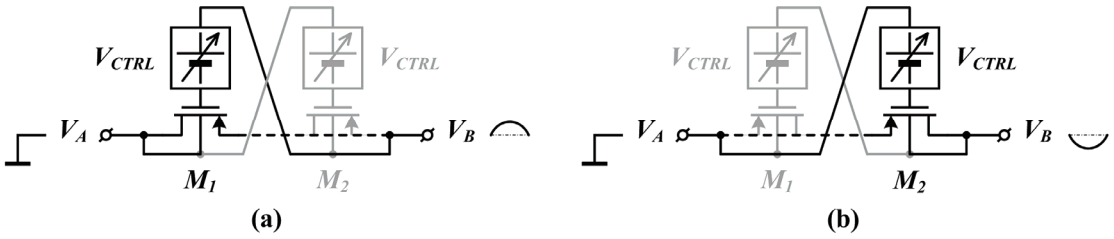


Figure 4.8: The operations of the proposed tunable pseudo-resistor during (a) positive and (b) negative halves of a sine wave swing at V_B.

ideal OTA and V_B to the output terminal of the OTA. If no signal clipping occurs, a sine wave output would have its entire swing projected on V_B, whereas V_A would be fixed at virtual ground. During the positive half of the sine wave (V_{BA}>0), as illustrated in Fig. 4.8 (a), the gate-to-source bias V_{gs2} of M_2 is discharged through V_A, which tends to turn on M_2, and transfers most of the signal swing to the source terminal of the then virtually shut M_1. During the negative half of the sine wave (V_{BA}<0), as illustrated in Fig. 4.8 (b), M_1 sees a large negative gate-to-source bias V_{gs1} and tends to turn on, leaving most of the signal swing to the now virtually shut M_2. Consequently, each transistor works as the active resistor for half of each output cycle, and the gate-to-source bias for the transistor in use is roughly defined by the respective V_{CTRL}. Such V_{gs} assignment helps to sustain the order of the incremental resistance across certain output range until the drain-to-source voltage drop V_{ds} of the transistor in use begins to considerably turn on the respective source-to-bulk p-n junction.

Fig. 4.9 shows the simulated resistance of the proposed cross-coupled tunable pseudo-resistor with different V_{CTRL}s. The symmetry of the resistance curves under different bias conditions evidently indicates the electrical balance obtained from the proposed structure. However, it should also be noted that the resistance quickly de-

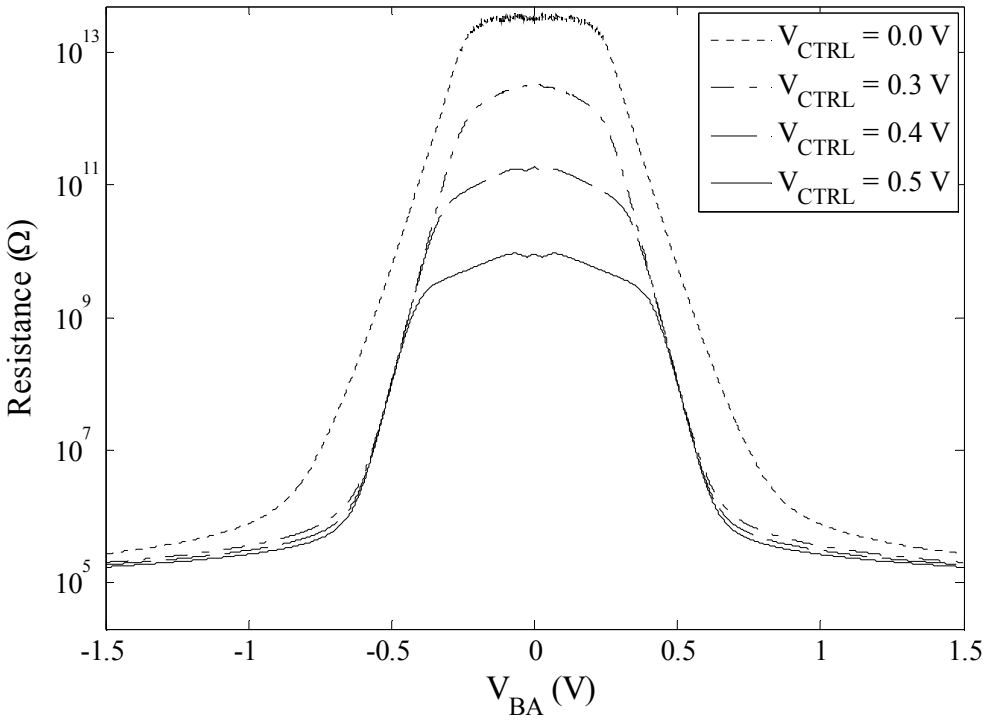


Figure 4.9: Simulated resistance of the proposed tunable pseudo-resistor.

grades once the leakage current through the aforementioned source-to-bulk junction begins to dominate, which limits the useful signal swing range of the proposed structure in certain applications.

4.2 Low Noise Preamplifier

4.2.1 Noise Efficiency

The system bandwidth and dynamic range are two major factors that restrict the minimum power investment into an analog system. For most biomedical applications, the informative bandwidth is rather low, posing no significant power requirement. The dynamic range, on the other hand, usually needs to be greater than 40 – 50 dB; some applications even require 60 dB or higher. As physiological signals are gener-

ally small in amplitude, e.g., most surface ECGs are no larger than a few mV, the smallest informative waves that need to be resolved can thus reach below 0.1 mV or still lower. Hence the input referred intrinsic noise of such systems must be well confined below these feature waves, which poses the dominant current-consuming factor in biomedical designs. With the emerging of telemetric version of many biomedical devices, it is thereby critical to enhance the current efficiency such that comparable noise figure is achievable with much reduced power consumption.

The noise efficiency factor (*NEF*) is a useful benchmark proposed in [16] that indicates a system's current efficiency in suppressing intrinsic noise. By comparing the input referred noise of the system to that of an ideal bipolar transistor, the *NEF* is defined as

$$NEF = V_{rms,in} \cdot \sqrt{\frac{2 \cdot I_{tot}}{\pi \cdot U_T \cdot 4kT \cdot BW}}, \tag{4.5}$$

where $V_{rms,in}$ is the root-mean-square (rms) voltage of the input referred noise, BW is the bandwidth, and I_{tot} is the total current drawn from the power supply. Given that the input referred noise voltage of an amplifier is generally inversely proportional to the square root of the input transconductance g_m, and assuming that the input stage current I_D dominates I_{tot} while other factors are ignored, one may conclude that the *NEF* is inversely proportional to the square root of the transconductance efficiency, which is defined by

$$TE = g_m / I_D . \tag{4.6}$$

Hence, keeping the *NEF* as small as possible is largely equivalent to finding the high-

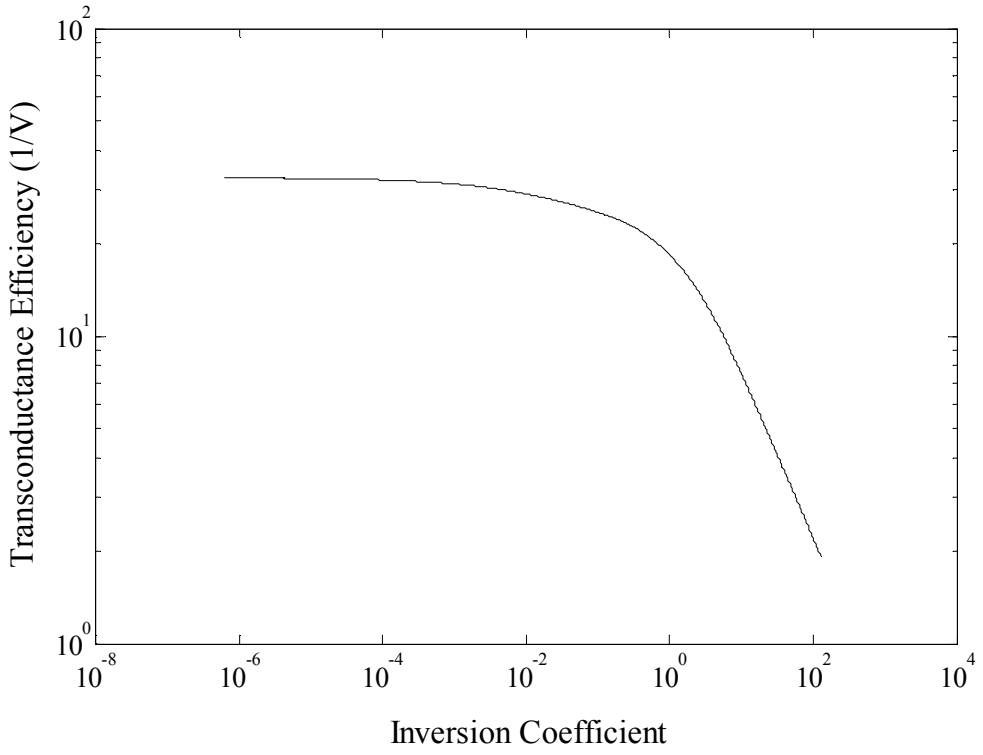


Figure 4.10: Simulated transconductance efficiency versus inversion coefficient for a long channel NMOS transistor.

est transconductance efficiency a technology can support.

Fig. 4.10 shows the transconductance efficiency of a long channel n type transistor in a standard 0.35-µm CMOS process versus its inversion coefficient (*IC*), which is defined as

$$IC = \frac{I_D}{2NK_{n0}(W/L)U_T^2}, \tag{4.7}$$

where K_{n0} is the process related gain factor for NMOS transistors. In general, the transistor is considered to be biased at weak inversion when $IC<0.1$, moderate inversion when $0.1<IC<10$, and strong inversion when $IC>10$ [17]. From Fig. 4.10, it is evident

that the transconductance efficiency increases when the inversion level of the transistor reduces, and gradually saturates after getting into the weak inversion region. Hence, it is desirable to bias the input stage at weak inversion when the power efficiency is a major concern.

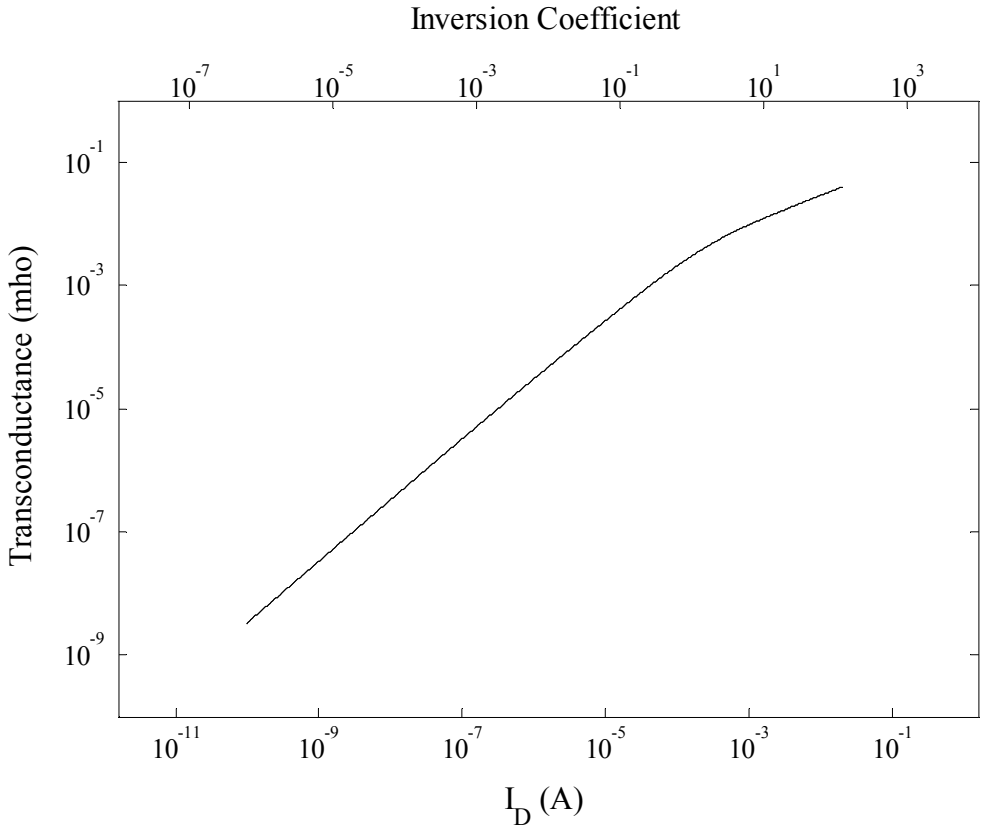


Figure 4.11: Simulated transconductance versus current and inversion coefficient for a long channel NMOS transistor.

On the other hand, with a given *W*/*L* ratio of the input stage transistor, the transconductance at weak inversion is much smaller than that at moderate or strong inversion due to the substantially weaker I_D. This is shown in Fig. 4.11, where the transconductance is plotted versus I_D and *IC*. In order to produce sufficient transconduc-

tance for the required dynamic range, while still maintaining the input stage at weak inversion mode for the desirable high current efficiency, one has to scale up I_D and W/L ratio simultaneously such that IC remains relatively unchanged below 0.1. However, it should be noted that such scaling is not unbounded due to potential degradation in frequency response, and chip area or process reliability restraint.

4.2.2 The Proposed OTA

Following the discussions in Section 4.2.1, we proposed the reconfigurable low-power low-noise OTA as shown in Fig. 4.12. It is essentially a variant of Miller-compensated two-stage OTAs that incorporates reconfigurability. M_1 and M_2 form the input differential pair and are made of large NMOS transistors (W/L = 1040 μm / 3 μm) biased in subthreshold (weak inversion) mode for best current efficiency in sup-

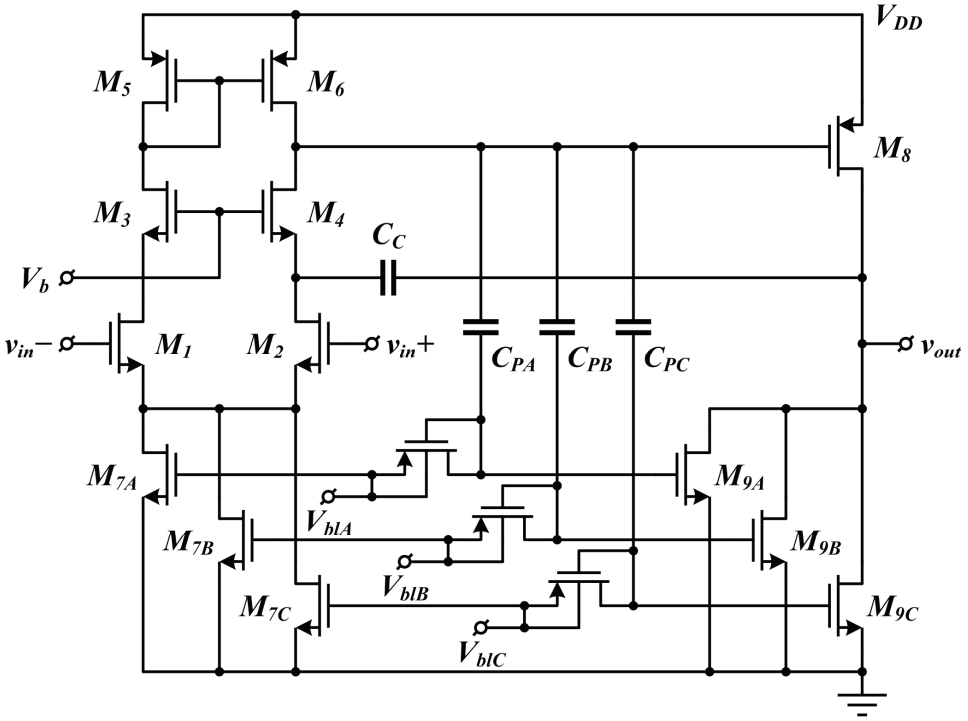


Figure 4.12: Circuit diagram of the proposed OTA.

pressing intrinsic noise. M_3 and M_4 serve to compensate for the right-half-plane (RHP) zero created by M_8 and the Miller capacitor C_C. M_8, C_Ps and M_9s form the push-pull output stage, which features enhanced *SR* when reacting to perturbations at *output*.

The gain-bandwidth product (*GB*) of the OTA is tuned based on the following equation

$$GB = g_{m1}/C_C \,, \tag{4.8}$$

where g_{m1} is the transconductance of the input differential pair, which changes linearly against the tail current in subthreshold mode. A 3-bit programming scheme by M_7s, M_9s and V_{bl}s is employed to implement the current adjustment. The bias control voltages V_{bl}s are generated by the circuit in Fig. 4.13, where $ctrl_{LPFA}$, $ctrl_{LPFB}$ and $ctrl_{LPFC}$ are digital control bits. By changing the control word value, V_{bl}s are switched between ground and the predefined bias voltage V_{bias}, producing multiple combinations. With a proper choice of M_7s, up to seven different values of tail current and hence *GB* can be obtained.

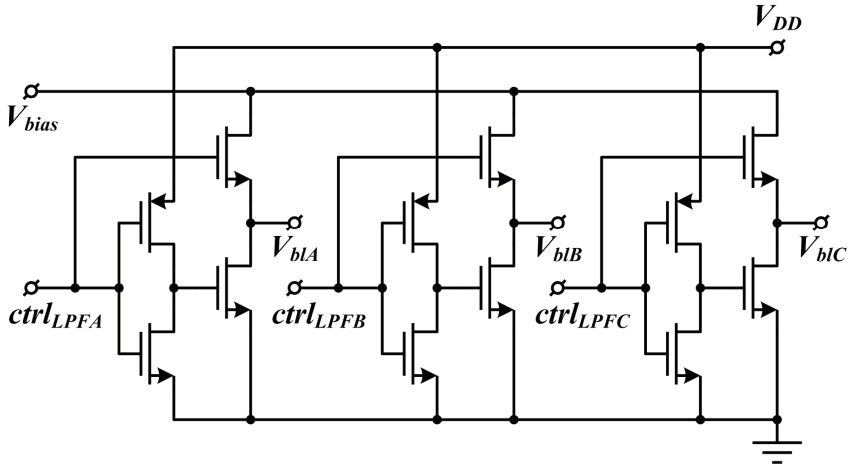


Figure 4.13: Circuit diagram of the 3-bit *GB* controller.

Some of the properties of the proposed OTA are discussed as follows.

a) Input-referred noise

The power spectral density of the input-referred thermal noise of the proposed OTA can be estimated by

$$\frac{\overline{v_n^2}}{\Delta f} = \frac{16kT}{3g_{m1,2}}\left[1 + \frac{1}{g_{m1,2}g_{m3,4}r_{O1,2}^2} + \frac{g_{m5,6}}{g_{m1,2}} + \frac{g_{m7}}{8g_{m1,2}g_{m5,6}^2r_{O5,6}^2} + \frac{1}{2g_{m1,2}(g_{m8}+g_{m9})r_{O5,6}^2}\right]. \quad (4.9)$$

It is evident that $g_{m1,2}$ is the key factor that determines the noise performance. With a given power budget at the input stage (hence a fixed tail current), it is desirable to use a large *W*/*L* ratio so as to force *IC* into weak inversion region, where $g_{m1,2}$ reaches its maximum. If the noise level is still beyond the design target, the only option then is to raise the power budget, and preferably to increase the *W*/*L* ratio accordingly to maintain the transconductance efficiency.

Also revealed in Equation 4.9 is that the push-pull output scheme helps to reduce the noise contribution of the output stage when referred to the input. The reduction is done by enhancing the output stage gain, where M_9s do not only serve as the current source, but also as a transconductance device. As a result, the g_{m9} factor appears in the denominator, as opposed to the case of a single transconductance output scheme, where the current source factor usually appears in the numerator.

As for the flicker noise, since its spectral density is generally inversely proportional to the gate capacitance of the respective transistor at the frequency of interest, it is desirable to use a large gate area ($W \times L$) for the input transistor pair.

b) RHP zero compensation

The Miller capacitor C_C provides a feedback path from the OTA output to the first stage output, and serves to improve the phase margin by pushing the dominant pole away from the non-dominant ones. However, the same path also channels feed forward signals, which produces a RHP zero. If not compensated, the zero of an ordinary OTA can be estimated by

$$z = g_{m8}/C_C \ , \tag{4.10}$$

where g_{m8} is the transconductance of the output stage. Because it is generally lower or comparable to the *GB* determined by Equation 4.8, such zero can significantly degrade the phase response.

In the proposed OTA, M_4 is used to compensate for the RHP zero. Given that zeros are *s* factors that make the OTA gain zero, one may take the following two approaches to estimate them in a "physical" perspective. First, the output of an amplifier is zero when the effective transconductance is cancelled out. In Fig. 4.12, this occurs when C_C becomes resistive and completely drains the transconductance currents by M_8 and M_9. In mathematical form, this is expressed as

$$sC_C \approx g_{m4}\left(r_{O4} \,\|\, r_{O6}\right)\left(g_{m8} + g_{m9}\right) \tag{4.11}$$

at $s=z$, which yields a RHP zero

$$z \approx g_{m4}\left(r_{O4} \,\|\, r_{O6}\right) \cdot \frac{g_{m8} + g_{m9}}{C_C} \ . \tag{4.12}$$

Compared to Equation 4.10, the RHP zero is effectively extended by $g_{m4}(R_{O4}||R_{O6})$, posing less a threat to the phase response.

Second, the output of an amplifier is zero when the output impedance is zero. In Fig. 4.12, this occurs when the output sees zero impedance via C_C to ground. In mathematical form, this is expressed as

$$\frac{1}{sC_C}+\left(\frac{1}{sC_4}\|\frac{r_{O4}+r_{O6}}{1+g_{m4}r_{O4}}\right)=0 \tag{4.13}$$

at s=z, where C_4 is the parasitic capacitance to ground seen at the source of M_4. Solving for s yields a left-half-plane (LHP) zero

$$z\approx-\frac{g_{m4}}{C_C+C_4}\cdot\frac{r_{O4}}{r_{O4}+r_{O6}}. \tag{4.14}$$

From the above discussions, it is seen that M_4 has extended the RHP zero to higher frequency and introduced a LHP zero. In general, it is desirable to implement an unbalanced bridge path between the first stage and the output, such that the feed-forward channel is substantially weaker than the feedback one (e.g. by feed-forward signal downscaling and unbalanced impedance).

c) V_b and power supply rejection ratio (*PSRR*)

The gates of M_3 and M_4 are biased at a predefined voltage V_b. Given that the OTA is designed to be operating primarily at 1-V supply, the useable V_b range is rather limited to close to the upper power rail in order to maintain a proper biasing of M_1 and M_2. For simplicity, we chose to source V_b from the power supply. However, a direct

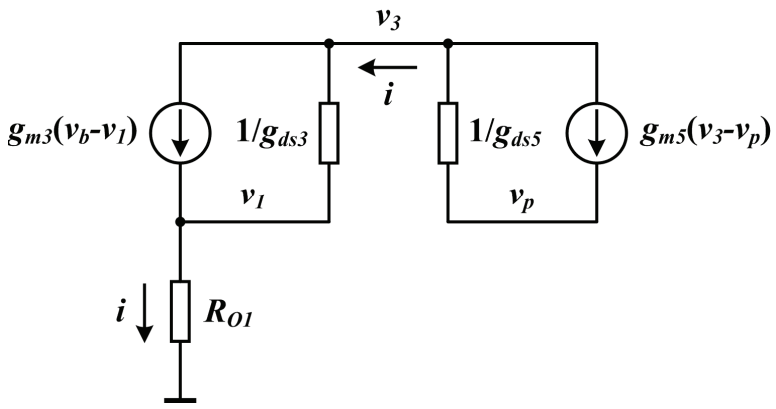


Figure 4.14: The small signal circuit of the OTA input stage when responding to close-to-DC power supply disturbance.

connection may degrade the *PSRR* of the OTA, which is explained as follows.

First, consider the close-to-DC small signal analysis in Fig. 4.14, where v_p is the power supply disturbance, v_b is the disturbance exerted on V_b, v_1 is the resultant drain voltage of M_1, and R_{O1} is the small signal resistance looking into the drain of M_1. Assuming the two branches of the differential pair are completely balanced, the first stage output can then be obtained by solving for v_3 on the drain of M_3 as

$$v_3 = \frac{(g_{m5}+g_{ds5})(g_{m3}R_{O1}+g_{ds3}R_{O1}+1)v_p - g_{m3}v_b}{(g_{m5}+g_{ds5})(g_{m3}R_{O1}+g_{ds3}R_{O1}+1)+g_{ds3}} \approx v_p - \frac{v_b}{g_{m5}R_{O1}}. \qquad (4.15)$$

Hence v_3 is approximately equal to v_p regardless of v_b, if R_{O1} is large. When applied to the push-pull output stage, such a voltage level can generate a fairly significant power gain provided it is effectively coupled to the input of M_9. Fortunately, the *PSRR* is still well suppressed by the huge differential gain of the OTA in this frequency range.

As the frequency of the power disturbance approaches the power mode dominant

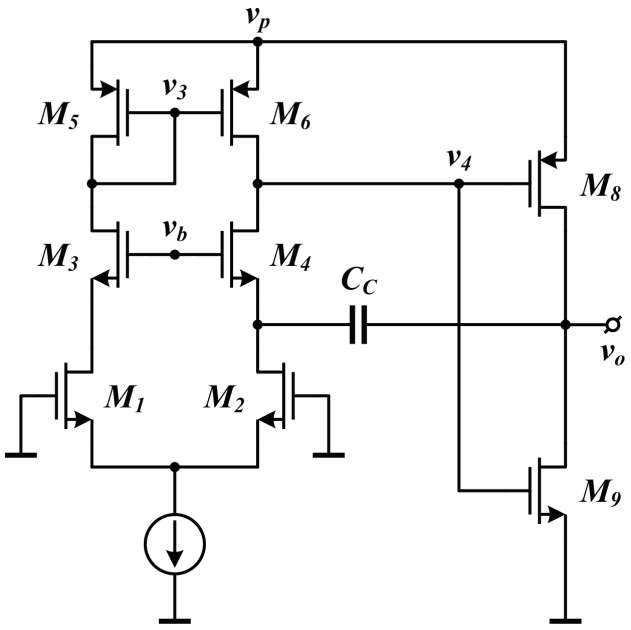


Figure 4.15: Simplified circuit diagram of the OTA when responding to power supply disturbance.

pole (close to the differential mode pole and less than 1 Hz in this design), the first stage output v_4 begins to deviate from v_3 due to the Miller feedback path of C_C and M_4 (see Fig. 4.15). If we consider the output stage of the OTA as a difference amplifier with two inputs v_p and v_4, the gradually increased fed back current through C_C will eventually overpower the original current inputs that generate the close-to-DC v_4 (equal to v_3), and force v_4 towards a defined voltage level. Assuming $g_{m8} \approx g_{m9}$, such level is approximately at half of v_p. During this process, the power gain at v_o keeps dropping at 20 dB/dec. Since the differential gain of the OTA also drops at 20 dB/dec in this frequency range, the *PSRR* does not alter much. The v_b value in this case poses no significant impact except some negligible bias to the dominant pole (to slightly lower frequency).

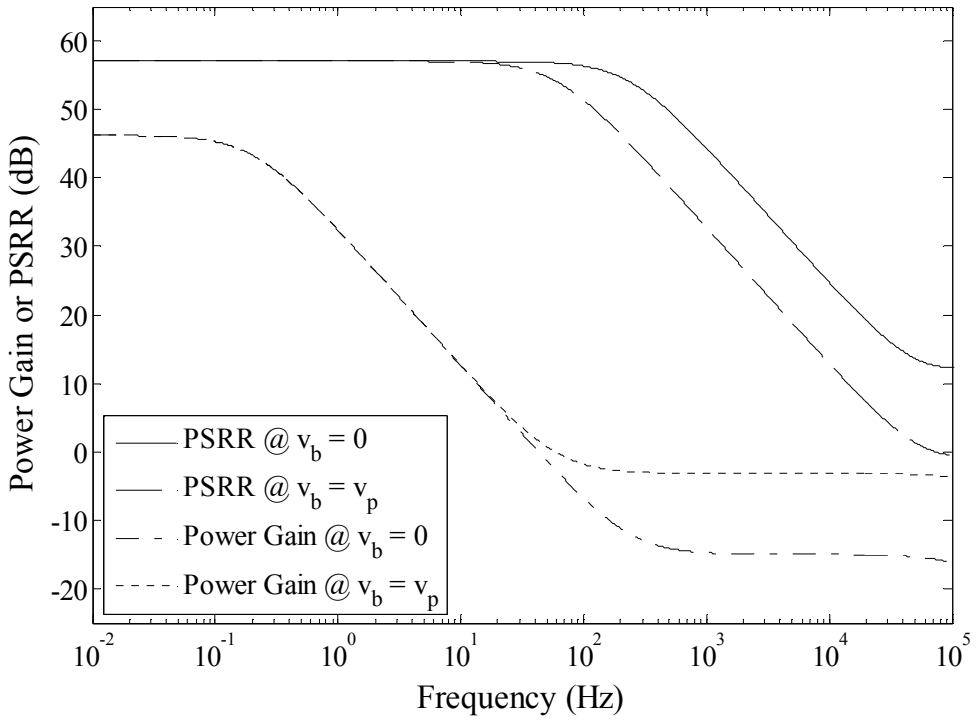


Figure 4.16: Simulated power gain and *PSRR* of the OTA at different v_b values.

As the frequency of the power disturbance further increases, the feed-forward effect of C_C begins to manifest. The power gain at v_o thereby turns flat due to the resultant zero. Since the differential gain of the OTA is still dropping at 20 dB/dec, the *PSRR* starts to degrade at 20 dB/dec. For v_o varying from 0 to v_p, the power disturbance at the source of M_4 increases, and so does the feed-forward effect. In the respective frequency response, this is equivalent to the zero moving towards lower frequency, and the flat power gain moving upwards. Therefore, it is desirable to make v_b small so as to invoke the zero at relatively higher frequency for a better *PSRR*.

The above discussions are evidently backed up by the simulated power gain and *PSRR* in Fig. 4.16, wherein the *PSRR* at v_b=0 is about 10 dB higher than that at v_b=v_p after each respective zero is invoked. Here the fixed V_b (v_b=0) is obtained by passing

the power supply through a pseudo-resistor-capacitor low-pass filter. For its simplicity and effectiveness, this scheme has been adopted in the proposed design.

4.2.3 The Proposed Preamplifier

The preamplifier employing the proposed OTA and tunable pseudo-resistor is shown in Fig. 4.17, wherein drawn at the lower right corner is the 3-bit controller that sets the high-pass corner frequency.

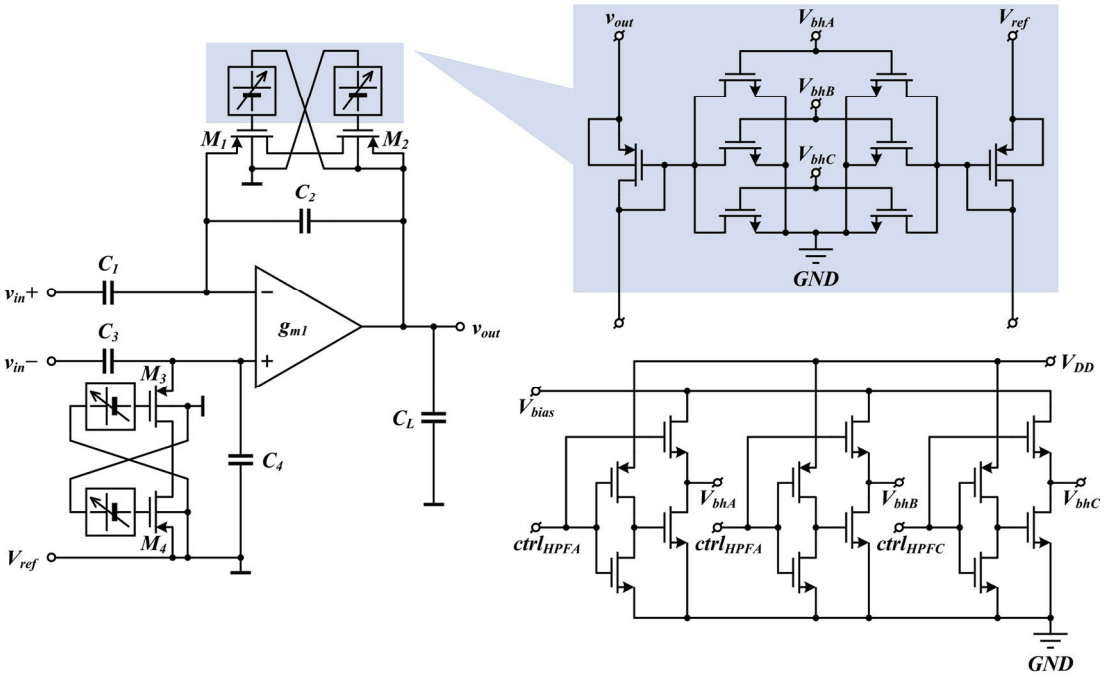


Figure 4.17: Circuit diagram of the proposed preamplifier.

Note that the implemented pseudo-resistor appears a bit different from its original form in Fig. 4.7, in that the M_1 (or M_3) bulk terminal and the M_2 (or M_4) V_{CTRL} source terminal are connected to a third external terminal V_{ref} instead of to the respective pseudo-resistor terminal. This is because the employed OTA features high impedance inputs, which are incapable of compensating for the bulk-to-substrate leakage of M_1

and M_3. The revised connection in Fig. 4.17 solves this problem via the low impedance V_{ref} terminal and yet still virtually maintains the 2-terminal topology and hence the electrical balance of the structure, thanks to the fact that the two inputs of the OTA are pinned at virtual ground (V_{ref}) at steady state in the frequency range of interest.

One drawback of pseudo-resistors when applied in "real-time" applications is the long settling time due to the large time constant. Depending on the initial circuit state, power up sequence, and perturbation distribution and characteristics among others, such circuits can take seconds or minutes or even longer to start up or settle down. Fortunately, with tunability embedded, an easy solution is viable for the proposed design: at circuit startup or upon significant circuit state perturbation, the pseudo-resistor can be switched to one of the lower resistance modes for a smaller time constant, and switched back to the desired value after the circuit settles.

The noise performance of the closed-loop preamplifier differs slightly from that of the open-loop OTA. Including the input-referred noise source of the OTA on its negative input terminal, we may represent the closed-loop system using the diagram in Fig.

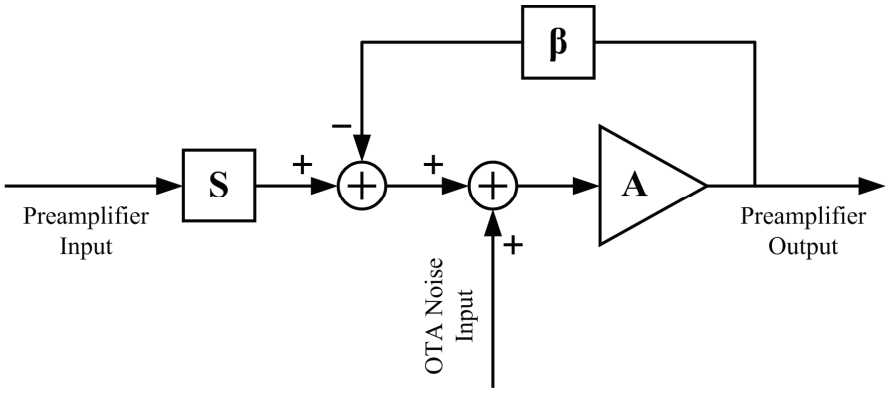


Figure 4.18: System diagram of the preamplifier including the OTA noise.

4.18, where block S is the noiseless input scaler for the preamplifier. Referring the OTA noise power spectral density to the preamplifier input yields

$$N_{in,PRE} = N_{in,OTA} \big/ S^2 \,. \tag{4.16}$$

In this design S is determined by

$$S = \frac{C_1}{C_1 + C_2 + C_{in,OTA}}, \tag{4.17}$$

where $C_{in,OTA}$ is the input capacitance of the OTA. Equation 4.16 then becomes

$$N_{in,PRE} = \left(\frac{C_1 + C_2 + C_{in,OTA}}{C_1} \right)^2 \cdot N_{in,OTA} \,. \tag{4.18}$$

Hence it is desirable to make C_1 much larger than C_2 and $C_{in,OTA}$ so as to reduce the input attenuation / noise amplification effect of scaler S.

4.3 PGA

The OTA in the proposed PGA adopts the same topology as that in the preamplifier, except that the current distribution now favors the output stage for better driving capacity. The gain adjustment is realized through varying the feedback factor, common implementations of which are shown in Fig. 4.19, wherein C_x is either connected or disconnected from the input or feedback path by switch S_x. Despite its simplicity, such approach may cause distortions in frequency domain due to the finite off-state resistance of S_x. Denoting such resistance as R_x, one may write the transfer function of the circuit in Fig. 4.19 (a) as

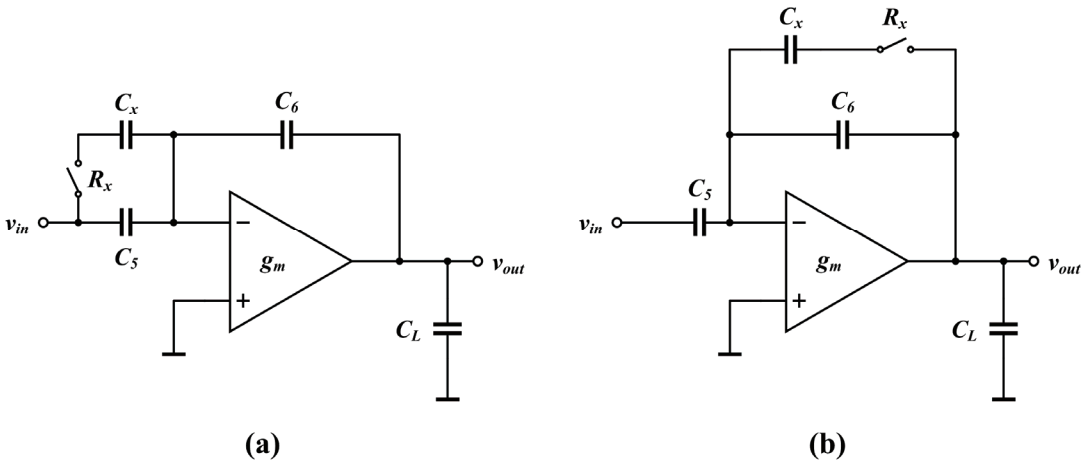


Figure 4.19: Two conventional gain adjustment schemes.

$$G = \frac{C_5}{C_6} \cdot \frac{sC_xR_x + 1 + C_x/C_5}{sC_xR_x + 1}, \tag{4.19}$$

and the transfer function of Fig. 4.19 (b) as

$$G = \frac{C_5}{C_6} \cdot \frac{sC_xR_x + 1}{sC_xR_x + 1 + C_x/C_6}. \tag{4.20}$$

Since C_x is typically comparable to C_5 or C_6, the resultant non-overlapping zero-pole pair in either case distorts the frequency response with a localized gain hump or dip. Given C_x is in pF range and R_x at approximately 10^{13} ohm, such distortion occurs at around sub-1 hertz to a few hertz, where majority of the ECG energy resides.

To correct this problem, we proposed the "flip-over-capacitor" scheme as illustrated in Fig. 4.20, wherein C_x is flipped between the input and feedback paths for gain tuning. It functions as part of the input capacitor in the high gain mode and part of the feedback capacitor in the low gain mode. R_x of the current open switch is effectively excluded from the input and feedback paths, and behaves only as a negligible

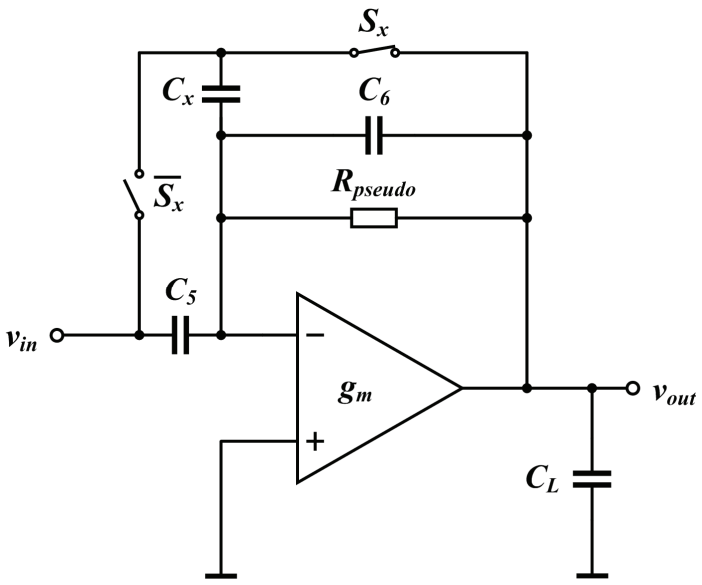


Figure 4.20: Simplified circuit diagram of the proposed "flip-over-capacitor" gain control scheme.

load to the previous stage and the PGA itself. Hence, the PGA gain can be accurately set throughout the entire frequency range of interest.

The PGA is designed to feature a much wider bandwidth than the preamplifier, such that the settling error in worst case S/H process is less than half LSB. It should be noted that achieving such bandwidth does not only require tuning the current and Miller capacitor in the OTA, but also involves adjusting the ratio of the input capacitance of the OTA and the external capacitor network. This is because in the context of ultra-low-power design, the OTA input pair employs large transistors to obtain the desired subthreshold mode biasing. The resultant large input capacitance $C_{in,OTA}$ may form a non-negligible leakage path for the input signal. As the OTA gain degrades with increasing frequency, the voltage swing at the negative input terminal of the OTA becomes larger, and so does the amount of leaked charges through $C_{in,OTA}$ in

each cycle. Consequently, less charges are delivered to C_6 for output regulation and the pole appears earlier than expected.

From the system perspective, the above process can be quantitatively explained by including $C_{in,OTA}$ in the transfer function. Since the feedback factor in this case is given by

$$\beta = \frac{C_6}{C_5 + C_6 + C_{in,OTA}}, \tag{4.21}$$

the bandwidth for the closed-loop system can be obtained as

$$BW_{PGA} = BW_{OTA} \times \left(1 + \frac{C_6}{C_5 + C_6 + C_{in,OTA}} A\right). \tag{4.22}$$

Given that C_5 and C_6 are comparable, they should be set at least one order higher than $C_{in,OTA}$ to eliminate its impact on bandwidth.

In addition to the wide bandwidth, the PGA would also require a sufficient *SR* to achieve the desired S/H accuracy. Fortunately, the rather limited *SR* of the Miller compensation scheme poses no bottleneck here, thanks to the fact that the S/H perturbation is primarily injected into the output terminal instead of the input terminal of the PGA. If the output restores fast, C_C remains virtually fully charged during the process and the *SR* problem is largely transferred to the output stage. With the push-pull output scheme employed, the perturbation is effectively coupled to the gate terminals of both output transistors. Hence, the output settles fast at the beginning with almost linear (large signal) behavior. When the PGA settling approaches the end, the normal input-induced linear settling begins to dominate. It works to remove the little charge

disturbance or error voltage on C_C that is created during the large signal settling process. Typically, such S/H mechanism facilitates considerable power saving, especially for the output stage of the OTA. This is because the quiescent current in the output branch neither needs to match the first stage current in order to support the Miller-capacitor-related *SR*, nor is it lower-bounded by the load-capacitor-related *SR* thanks to the efficient push-pull scheme.

Chapter 5

ADC Design

5.1 The ADC Architecture

The SAR data conversion scheme is known for its moderate accuracy, moderate speed, and low power overhead, and is therefore well suited for biomedical applications. In this design, we implemented a self-clocked 12-bit 1-kS/s SAR ADC, which provides sufficient dynamic range and bandwidth for purposes of single-channel ECG acquisition.

Fig. 5.1 depicts the proposed ADC architecture. In each S/H cycle, the analog input is sampled through the bootstrapped MOS switch, and thereafter held on the capacitive digital-to-analog converter (DAC), where a series of binary-weighted level conversions are conducted. The comparator compares each DAC output with a fixed reference voltage *REF* and registers the result to the SAR logic, which accordingly determines the control codes to the DAC. Taking the clock from the onchip oscillator (either a crystal driver or a relaxation oscillator), the timing sequence generator synchronizes each individual module and times the corresponding state transitions.

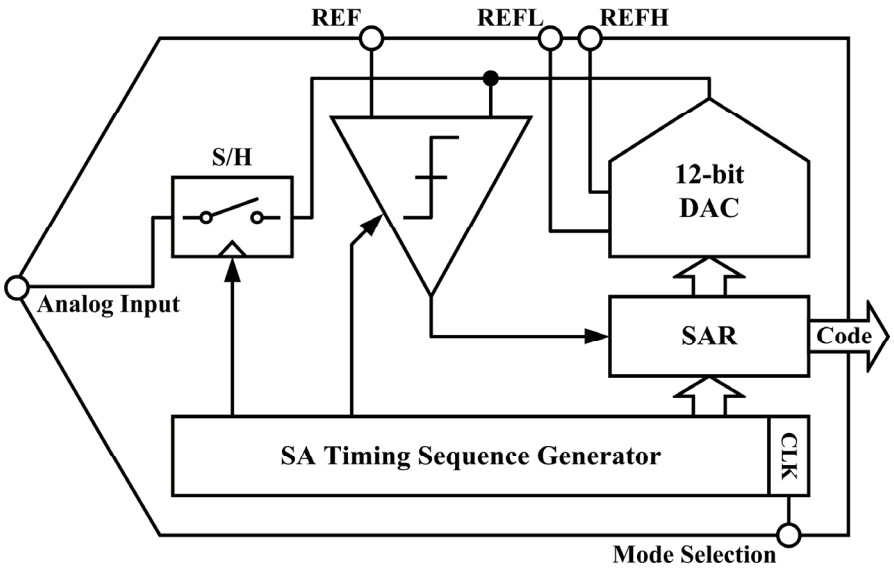


Figure 5.1: The proposed architecture of the SAR ADC.

5.2 The Bootstrapped S/H

Driven directly by the preceding PGA stage without the need of additional hold buffer, the employed open-loop S/H scheme enjoys fast settling and small offset error at low power cost. Instead of the standard transmission gate, the S/H switch is built on a single NMOS transistor valved by a bootstrapping controller. Such approach is derived from the following considerations.

First, the S/H switch must provide sufficient conductance across the entire input range. Since the PGA has been designed with sufficient bandwidth and *SR* to settle the DAC load within the sampling interval, the S/H bandwidth associated with the S/H switch alone must be orders of magnitude wider to make its impact negligible. Fig. 5.2 plots the simulated incremental resistance of a standard transmission gate operated at 1-V supply. Given $V_{thn}+V_{thp}\approx 1.15$ V for the 0.35-μm CMOS process used, the maximum resistance (approximately 5 MΩ under the typical process condition)

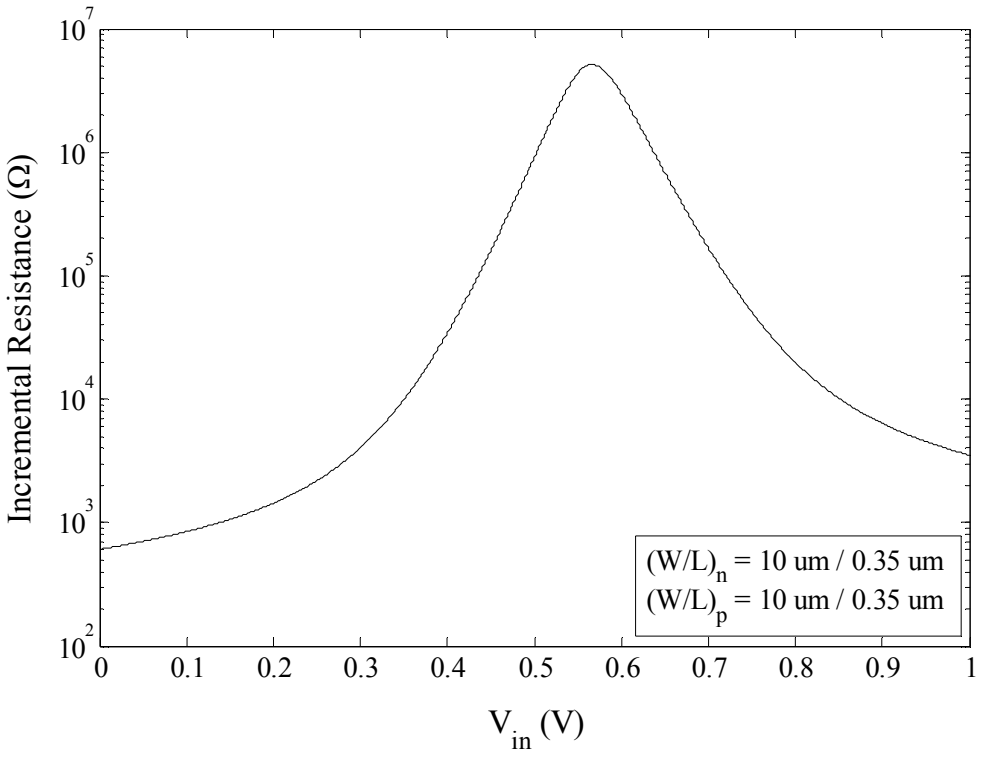


Figure 5.2: Simulated incremental resistance of a standard transmission gate.

occurs when the input falls into the mid range of the supply where both transistors tend to cut off. The resultant worst-case S/H bandwidth on a 100-pF DAC load is thereby only 318 Hz, much lower than the minimally required PGA bandwidth calculated from Inequality 3.13 (approximately 5 kHz).

One possible remedy is to increase the size of the transmission gate. Unfortunately, under the given bias and timing conditions, the bandwidth and accuracy requirements in ECG acquisition demand at least a few thousand or even tens of thousand for the W/L ratio of the complementary pass transistors, which would inevitably introduce large (close to pF range) parasitic capacitance. These nonlinear errors would not only make compensation for charge injection difficult, but would also cause severe inherent gain error and nonlinearity.

From the discussions above, one may agree that a more viable solution is to keep the pass transistors small while boosting the gate control voltage. Therefore, a standard bootstrapped NMOS switch with charge injection compensation is employed, where the pass transistor takes a small footprint of 2 μm / 0.35 μm. The simulated worst-case resistance (when both the input and output are close to 1 V) is less than 6.7 kΩ under the typical process condition, and less than 27 kΩ under the worst speed process condition, corresponding to over 240 kHz and over 60 kHz of S/H bandwidths, respectively, posing negligible impact to the overall sampling accuracy.

5.3 The 12-bit Capacitive DAC

5.3.1 DAC Structure

The 12-bit DAC is realized by a binary-weighted capacitor array, whose structure is illustrated in Fig. 5.3. The capacitor in the middle that bridges the two sub-arrays serves to isolate the exponential weight scaling of each sub-array and scale down the leverage of the LSB sub-array, such that good matching can be achieved with relatively small footprint.

Unlike the structure demonstrated in [18], the entire array is built on identical unit

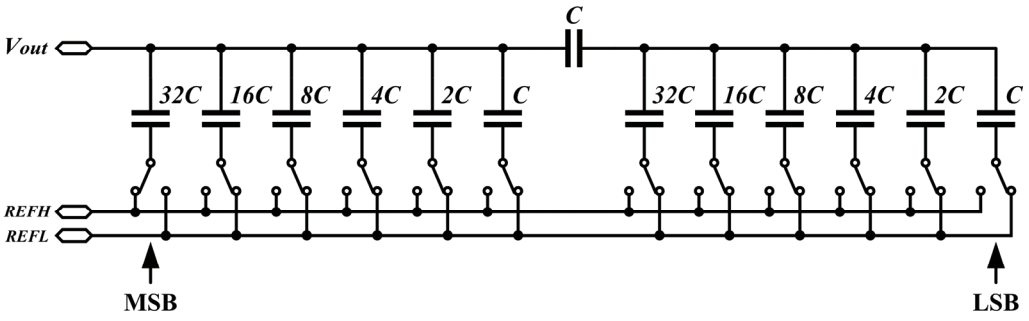


Figure 5.3: Simplified structure of the 12-bit binary-weighted capacitor array.

capacitors with no dummy capacitor at the tail of LSB sub-array. To evaluate the impact of such changes, let us consider the generalized form in Fig. 5.4.

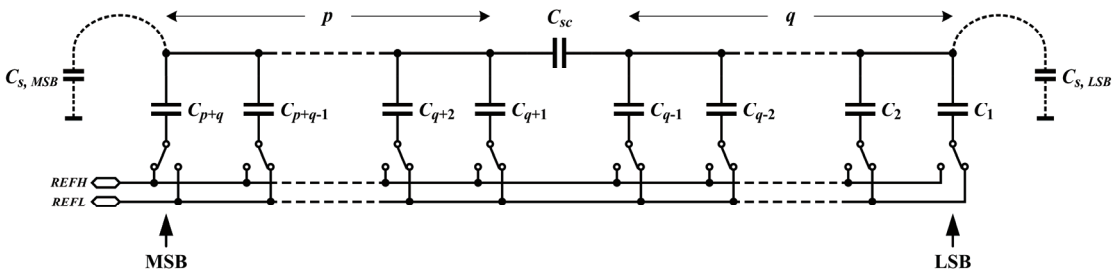


Figure 5.4: General structure of a scaled capacitor array.

In the generalized structure above, the most-significant-bit (MSB) and LSB sub-arrays are composed of p and q binary-weighted capacitors respectively. Hence the physical resolution n of the overall array is given by $p+q$. Assuming both sub-arrays are constructed with identical unit capacitance C, the j-th capacitor can then be expressed as

$$\begin{aligned} C_j &= k_j \cdot 2^{j-q-1} \cdot C \qquad if\ j > q, \\ C_j &= k_j \cdot 2^{j-1} \cdot C \qquad if\ j \le q, \end{aligned} \tag{5.1}$$

where k_j is the mismatch factor of the j-th capacitor. The scaling capacitor C_{sc} is given by $k_{sc} \cdot C$, where k_{sc} is the associated mismatch factor. $C_{s,\,MSB}$ and $C_{s,\,LSB}$ are stray capacitances seen at the common nodes of MSB and LSB sub-arrays respectively.

Ignoring all non-idealities, we may derive that regardless of its position, the flipping of j-th capacitor between *REFH* and *REFL* always generates the voltage change

$$V_j = \frac{2^{j-1}}{2^n - 1} \cdot V_{FS} \tag{5.2}$$

at the DAC output, where the absolute value of V_{FS} is given by *REFH−REFL*, and the sign is determined by the direction of flipping. It is evident that the proposed structure is equivalent to an intrinsic binary array with a missing dummy tail capacitor, hence producing a gain error close to ±½ LSB. In practice, this small error is generally negligible and sometimes even helpful as it can compensate for the inevitable $C_{s,\,MSB}$ with a negative 1-LSB capacitance.

5.3.2 Non-idealities and DAC Transfer Characteristics

All the aforementioned mismatch factors and stray parameters can affect the *static* transfer characteristic of the capacitor array. For simplicity, we will only consider one factor at a time (unless otherwise specified) and assume a constant value for the non-ideality of interest across its entire operation range in the following discussions.

a) Bit mismatch and stray error in an intrinsic binary array

Let us first examine an intrinsic binary array with *n*-bit resolution. When used to resolve an analog input with unity full scale range, the ideal analog-to-digital (A/D) step size is given by

$$U_{step,\,ideal} = \frac{1}{2^n} LSB^{-1}. \tag{5.3}$$

Assume that the mismatch occurs on the *i*-th bit, and the erroneous value is given by

$$(k_i - 1)\cdot C_i = e_i \cdot C_{LSB}, \tag{5.4}$$

where C_{LSB} in this case equals to C_1. Including the stray capacitance $e_s \cdot C_{LSB}$ on the common node, the total erroneous capacitance seen by the capacitor array can be writ-

ten as

$$e_{tot} \cdot C_{LSB} = \left(e_i + e_s\right) \cdot C_{LSB} . \tag{5.5}$$

The switching of *j*-th capacitor from *REFL* to *REFH* generates a voltage change V_j at the common node given by

$$\begin{aligned} V_j &= \frac{2^{j-1}}{2^n + e_{tot}} && if\ \ j \neq i, \\ V_j &= \frac{2^{j-1} + e_i}{2^n + e_{tot}} && if\ \ j = i. \end{aligned} \tag{5.6}$$

Consequently the *i*-th capacitor produces an additional error given by

$$V_{\Delta,i} = V_j\left(j = i\right) - V_j\left(j \neq i\right) = \frac{e_i}{2^n + e_{tot}} . \tag{5.7}$$

In a typical upward code transition, if the highest bit that toggles is the *j*-th bit, all the lower bits (from the (*j*−1)-th bit to the LSB) reset from 1 to 0. Therefore the combined effect on the common node can be calculated by

$$V_{step,j} = V_j - \left(V_{j-1} + \cdots + V_1\right) = \frac{1}{2^n + e_{tot}} \left(2^{j-1} - \sum_{m=1}^{j-1} 2^{m-1} \right) = \frac{1}{2^n + e_{tot}} , \tag{5.8}$$

without considering the effect of *i*-th capacitor in the transition. Comparing this result with the ideal A/D step size in Equation 5.3, we may obtain the instantaneous differential nonlinearity (*DNL*)

$$DNL = \frac{V_{step,j}}{U_{step,ideal}} - 1\ LSB = -\frac{e_{tot}}{2^n + e_{tot}}\ LSB . \tag{5.9}$$

If the *i*-th capacitor is involved in the code transition, we need to further include the correction factor defined by Equation 5.7. In the case of *i=j*, the erroneous capacitor switches upwards, and the resultant *DNL* is given by

$$DNL = -\frac{e_{tot}}{2^n + e_{tot}}\, LSB + \frac{e_i}{2^n + e_{tot}} \cdot U_{step,ideal}^{-1} = \frac{2^n e_i - e_{tot}}{2^n + e_{tot}}\, LSB\,; \qquad (5.10)$$

In the case of *i<j*, the erroneous capacitor switches downwards, and the resultant *DNL* is given by

$$DNL = -\frac{e_{tot}}{2^n + e_{tot}}\, LSB - \frac{e_i}{2^n + e_{tot}} \cdot U_{step,ideal}^{-1} = -\frac{2^n e_i + e_{tot}}{2^n + e_{tot}}\, LSB\,. \qquad (5.11)$$

Note that when *DNL*≤−1, the typical code density analysis on the A/D transfer characteristic would normally only reveal one or more missing codes and would thus give the minimum result of −1. Further insights can be obtained by studying the detailed A/D *and* digital-to-analog (D/A) transfer curves (instead of "blindly" by code statistics only).

All of the *DNL* values obtained above contain a common factor that contributes to the global integral nonlinearity (*INL*), or in a more commonly used term, the gain error (*GE*). Integrating this factor across the full scale range gives

$$GE = -\frac{e_{tot}}{2^n + e_{tot}}\, LSB \times \left(2^n - 2\right) \approx -e_{tot}\, LSB\,. \qquad (5.12)$$

After the *GE* compensation (see Part **b)** of this subsection for details), most of *DNL* values return to 0 except those associated with the transition of *i*-th capacitor, which are equal in magnitude but alter in sign in an interleaving manner. Hence, the

INL toggles between its two maxima as the input sweeps across the full scale range:

$$INL = \pm\frac{1}{2}\times\frac{2^n e_i}{2^n + e_{tot}}\ LSB \approx \pm\frac{1}{2} e_i\ LSB\ . \tag{5.13}$$

b) Stray error in a scaled array

Consider the scaled capacitor array in Fig. 5.4. The stray capacitance $C_{s,\,MSB}$ at the common node of MSB sub-array combined with the "missing" C_{LSB} as suggested by Equation 5.2 forms the total stray capacitance seen by the equivalent intrinsic binary array. Since all code transitions bear the same impact from this stray capacitance, the *DNL* stays constant across the full scale input range and is given by

$$DNL = -\frac{e_{tot}}{2^n + e_{tot}}\ LSB = -\frac{e_{s,\,MSB} - 1}{2^n + e_{s,\,MSB} - 1}\ LSB\ , \tag{5.14}$$

where the erroneous factor $e_{s,\,MSB}$ is defined by

$$e_{s,\,MSB} = \frac{C_{s,\,MSB}}{C_{LSB}}\ . \tag{5.15}$$

Hence the *INL* takes a linear global form and transforms to the *GE*, which is given by

$$GE = -\frac{e_{s,\,MSB} - 1}{2^n + e_{s,\,MSB} - 1}\ LSB \times \left(2^n - 2\right) \approx \left(1 - e_{s,\,MSB}\right) LSB\ . \tag{5.16}$$

Apparently when $C_{s,\,MSB}$=C_{LSB}, the scaled array is strictly equivalent to an intrinsic binary array.

The nonlinear characteristics associated with $C_{s,\,LSB}$ are a bit different due to the

leveraging effect of C_{sc} upon the LSB sub-array. Let us first consider the switching of a single capacitor from *REFL* to *REFH* in the LSB sub-array. The generated voltage change on the common node of the MSB sub-array can be expressed as

$$V_j = \frac{2^{j-1}C}{\left(1-2^{-p}\right)C + \left(2^q - 1\right)C + e_{s,\,LSB}C_{LSB}} \times \frac{C}{2^p C} = \frac{2^{j-1}}{2^n - 1 + 2^{p-q}e_{s,\,LSB}}, \quad (5.17)$$

where the erroneous factor $e_{s,\,LSB}$ is defined by

$$e_{s,\,LSB} = \frac{C_{s,\,LSB}}{C_{LSB}}, \quad (5.18)$$

and C_{LSB} satisfies

$$C_{LSB} = 2^{-q}C. \quad (5.19)$$

Similarly, we may derive the voltage change on the common node of the MSB sub-array generated by the switching of a single capacitor from *REFL* to *REFH* in the MSB sub-array:

$$V_j = \frac{2^{j-q-1}C}{\left(2^p - 1\right)C + \left(1 - \frac{1}{2^{-q}e_{s,\,LSB} + 2^q}\right)C} = \frac{2^{j-1} + 2^{j-2q-1}e_{s,\,LSB}}{2^n - 1 + 2^{p-q}e_{s,\,LSB}} \quad (5.20)$$

In a typical upward code transition, if the highest bit that toggles is the j-th bit, all the lower bits reset from 1 to 0. In the case of $j \le q$, the bit toggling only occurs in the LSB sub-array, and the A/D step size can be obtained as

$$V_{step,\,j} = \frac{1}{2^n - 1 + 2^{p-q}e_{s,\,LSB}}\left(2^{j-1} - \sum_{m=1}^{j-1} 2^{m-1}\right) = \frac{1}{2^n - 1 + 2^{p-q}e_{s,\,LSB}}. \quad (5.21)$$

Therefore, the instantaneous *DNL* can be calculated by

$$DNL = \frac{V_{step,\,j}}{U_{step,\,ideal}} - 1\ LSB = \frac{1 - 2^{p-q} e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}}\ LSB\,. \tag{5.22}$$

In the case of $j>q$, the bit toggling extends to the MSB sub-array, and thus both Equations 5.17 and 5.20 are involved in determining the A/D step size:

$$V_{step,\,j} = \frac{2^{j-1} - \sum_{m=1}^{j-1} 2^{m-1}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}} + \frac{\left(2^{j-1} - \sum_{m=q+1}^{j-1} 2^{m-1}\right) \times 2^{-2q} e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}} = \frac{1 + 2^{-q} e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}}\,. \tag{5.23}$$

The associated *DNL* can be calculated by

$$DNL = \frac{V_{step,\,j}}{U_{step,\,ideal}} - 1\ LSB = \frac{1 - 2^{p-q} e_{s,\,LSB} + 2^p e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}}\ LSB \tag{5.24}$$

Integrating the *DNL* values over the full scale input range, we may obtain the *GE* (the linear global *INL*) with end-point fit as

$$\begin{aligned} GE &= \frac{1 - 2^{p-q} e_{s,\,LSB} + 2^p e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}}\ LSB \times \left(2^p - 1\right) + \frac{1 - 2^{p-q} e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}}\ LSB \times \left(2^n - 2^p - 1\right) \\ &= \frac{2^n - 2 + \left(2^{p-q+1} - 2^p\right) e_{s,\,LSB}}{2^n - 1 + 2^{p-q} e_{s,\,LSB}}\ LSB. \end{aligned} \tag{5.25}$$

which can thereafter be rectified through *GE* correction. Note that in many cases when onchip calibration is not available, *GE* correction could involve backward fitting of the targeted linear A/D transfer function towards the measured gain curve and recalculation of the nonlinearity parameters against the new target. The analyses here,

on the other hand, aim to reveal the theoretical conversion characteristics in general. Hence the ideal A/D transfer function defined by Equation 5.3 is always used as the reference, and the *GE* factor is processed through another operation, whereby it is linearly subtracted from the nonlinearity parameters. Termed *GE* compensation in this work, the operation serves to shelve the linear global error so as to manifest local ones.

Observing the *DNL* values given by Equations 5.22 and 5.24, and applying the *GE* compensation, we may obtain the local *INL* as

$$INL = \pm\frac{1}{2}\times\frac{2^{p}\left(1-2^{-q}\right)e_{s,\,LSB}}{2^{n}-1+2^{p-q}e_{s,\,LSB}}\,LSB \approx \pm 2^{-(q+1)}e_{s,\,LSB}\,LSB\,, \tag{5.26}$$

where the approximation form is valid if p and q are close and $e_{s,\,LSB}$ is reasonably small. The *GE* compensation factor subtracted from each code is

$$GE_{\Delta/2^{n}} = \frac{1}{2^{n}-1+2^{p-q}e_{s,\,LSB}}\,LSB\,, \tag{5.27}$$

which is obtained by evenly distributing the *GE* (derived with end point discontinuities ignored) among all codes that take part in the integration.

From the above discussions, it can be concluded that unlike $C_{s,\,MSB}$ which only introduces a linear *GE*, $C_{s,\,LSB}$ can produce local *INL* errors and thereby should be minimized.

c) Bit mismatch in a scaled array

Assume that the mismatch occurs on the i-th bit of the scaled capacitor array in Fig. 5.4, and the erroneous value is given by

$$(k_i - 1) \cdot C_i = e_i \cdot C_{LSB} = e_i \cdot 2^{-q} C \,. \tag{5.28}$$

In the case of $i>q$, the result resembles that of the mismatch in an intrinsic binary array with a missing dummy tail capacitor. Therefore the *DNL* values are given by

$$DNL = -\frac{e_i - 1}{2^n - 1 + e_i} LSB \,, \tag{5.29}$$

if the i-th capacitor is not involved in the switching; and

$$DNL = \frac{(2^n - 1)e_i + 1}{2^n - 1 + e_i} LSB \approx e_i \; LSB \,, \tag{5.30}$$

if the *i*-th capacitor switches upwards; and

$$DNL = -\frac{(2^n + 1)e_i - 1}{2^n - 1 + e_i} LSB \approx -e_i \; LSB \,, \tag{5.31}$$

if the *i*-th capacitor switches downwards. The minimum *DNL* is bounded by −1 and will spread over to the subsequent codes if it exceeds this limit. The *GE* is given by

$$GE = -\frac{e_i - 1}{2^n - 1 + e_i} LSB \times (2^n - 2) \approx (1 - e_i) \, LSB \,. \tag{5.32}$$

The *INL* after *GE* compensation is given by

$$INL = \pm\frac{1}{2} \times \frac{2^n e_i}{2^n - 1 + e_i} LSB \approx \pm\frac{1}{2} e_i \; LSB \,. \tag{5.33}$$

In the case of $i \le q$, the array behaves similarly but the A/D transfer characteristic is partitioned by C_{sc} into repeating units, each of which takes the code span of the LSB

sub-array. Calculated through the same procedure, the *DNL* values are given by

$$DNL = \frac{1 - 2^{p-q} e_i}{2^n - 1 + 2^{p-q} e_i} LSB, \tag{5.34}$$

if the *i*-th bit is not involved in the switching, *or* the highest bit that toggles is in the MSB sub-array; and

$$DNL = \frac{1 + \left(2^p - 2^{p-q}\right) e_i}{2^n - 1 + 2^{p-q} e_i} LSB, \tag{5.35}$$

if the *i*-th bit switches upwards; and

$$DNL = \frac{1 - \left(2^p + 2^{p-q}\right) e_i}{2^n - 1 + 2^{p-q} e_i} LSB, \tag{5.36}$$

if the *i*-th bit switches downwards *and* the highest bit that toggles is in the LSB sub-array.

Note that in each code span of the LSB sub-array, there is one less *DNL* peak defined by Equation 5.36 than by Equation 5.35, the missing peak being canceled out by the switching of the MSB sub-array. Hence the *GE* contributed by each span can be calculated by

$$GE_s = \left(\frac{1 - 2^{p-q} e_i}{2^n - 1 + 2^{p-q} e_i} \times 2^q + \frac{2^p e_i}{2^n - 1 + 2^{p-q} e_i} \right) LSB = \frac{2^q}{2^n - 1 + 2^{p-q} e_i} LSB, \tag{5.37}$$

and the global *GE* can be obtained by

$$GE = GE_s \times 2^p = \frac{2^n}{2^n - 1 + 2^{p-q} e_i} LSB \approx 1\ LSB, \tag{5.38}$$

when ignoring the end code discontinuities.

The local *INL* value depends on the location of corresponding *DNL* peak within each code span of the LSB sub-array. After the global *GE* compensation, the maximum value can be estimated by

$$|INL| \le \left| \frac{2^p e_i}{2^n - 1 + 2^{p-q} e_i} \right| LSB \approx 2^{-q} |e_i| \, LSB \,. \tag{5.39}$$

It is worth mentioning that in Equations 5.34, 5.35, 5.36 and 5.39, the impact of the erroneous factor e_i to the numerators is modified by 2^{-q} due to the scaling topology. The unit capacitors in the LSB sub-array, on the hand, are scaled up from C_{LSB} by 2^q. Therefore, compared to an intrinsic binary array with the same C_n, the requirement for the mismatch factor k_i is not relaxed, but becomes easier to accomplish.

d) C_{sc} mismatch in a scaled array

Assume C_{sc} has the value of

$$C_{sc} = k_{sc} C = C + e_{sc} C_{LSB} = \left(1 + 2^{-q} e_{sc}\right) C \,. \tag{5.40}$$

Following the same procedure, we may obtained the *DNL* values as

$$DNL = \frac{1 + \left(2^p - 2^{p-q} + 2^{1-q} - 1\right) e_{sc}}{2^n - 1 + \left(2^{p-q} - 2^{1-q} + 1\right) e_{sc}} \, LSB \,, \tag{5.41}$$

if the highest bit that toggles is in the LSB sub-array; and

$$DNL = \frac{1 - \left(2^n - 2^{p+1} + 2^{p-q} - 2^{1-q} + 1\right) e_{sc}}{2^n - 1 + \left(2^{p-q} - 2^{1-q} + 1\right) e_{sc}} \, LSB \,, \tag{5.42}$$

if the highest bit that toggles is in the MSB sub-array.

The *GE* when ignoring the end code discontinuities is given by

$$GE = \frac{2^n - \left(2^n - 2^{p+1}\right)e_{sc}}{2^n - 1 + \left(2^{p-q} - 2^{1-q} + 1\right)e_{sc}} \, LSB \, . \tag{5.43}$$

Use this value for the *GE* compensation and we may obtain the local *INL* as

$$INL = \pm\frac{1}{2} \cdot \frac{\left(2^n - 2^{p+1} + 2^{p-q}\right)e_{sc}}{2^n - 1 + \left(2^{p-q} - 2^{1-q} + 1\right)e_{sc}} \, LSB \approx \pm\frac{1}{2} e_{sc} \, LSB \, , \tag{5.44}$$

where the approximation form is valid if p and q are close and e_{sc} is reasonably small.

5.3.3 Layout Considerations

It can be concluded from the analyses in the previous subsection that the scaling topology has effectively unloaded the matching challenges for the lower q bits. Most of the effort can then be put into the laying out of C_{sc} and the upper p bits. Since the bit size to be considered is much reduced, good matching result can be achieved with relatively small footprint.

In this work, the 12-bit capacitive DAC is divided into two 6-bit sub-arrays. The unit capacitor is chosen to be 1 pF, whose layer composition is depicted in Fig. 5.5. Given that the common nodes of both sub-arrays are sensitive to stray errors whereas the switching nodes are not, the former are constructed using Poly 2, and are sandwiched and shielded by the latter, which are formed by both Poly 1 and Metal 2. The irregular shape of the Metal 2 plate is due to the design rule restriction. Narrow Metal 1 traces route beneath the Metal 2 shields and connect the Poly 1 plates together to

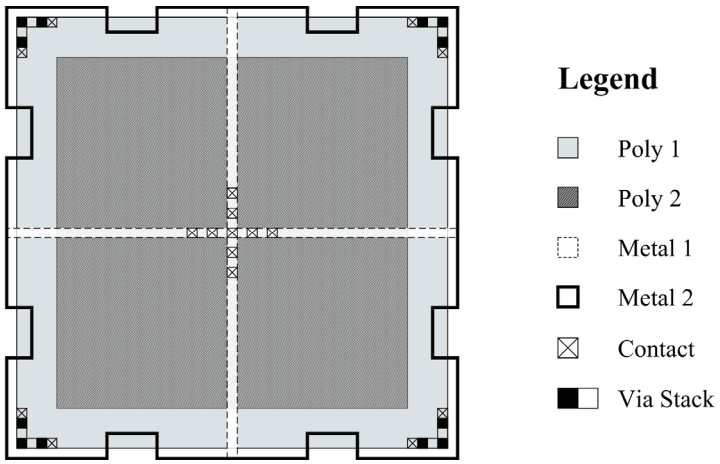


Figure 5.5: Layer composition of the unit capacitor (not to scale).

form the two common nodes. Via stacks at the four corners of each unit capacitor serve to stitch the top and bottom shielding plates together and provide upper layer taps that are used to wire the switching nodes above Metal 2. Hence, all sensitive nodes are laid out in a fixed pattern and isolated from the irregular interconnections.

The unit capacitors are then placed in a 16×8 matrix and connected in a common-

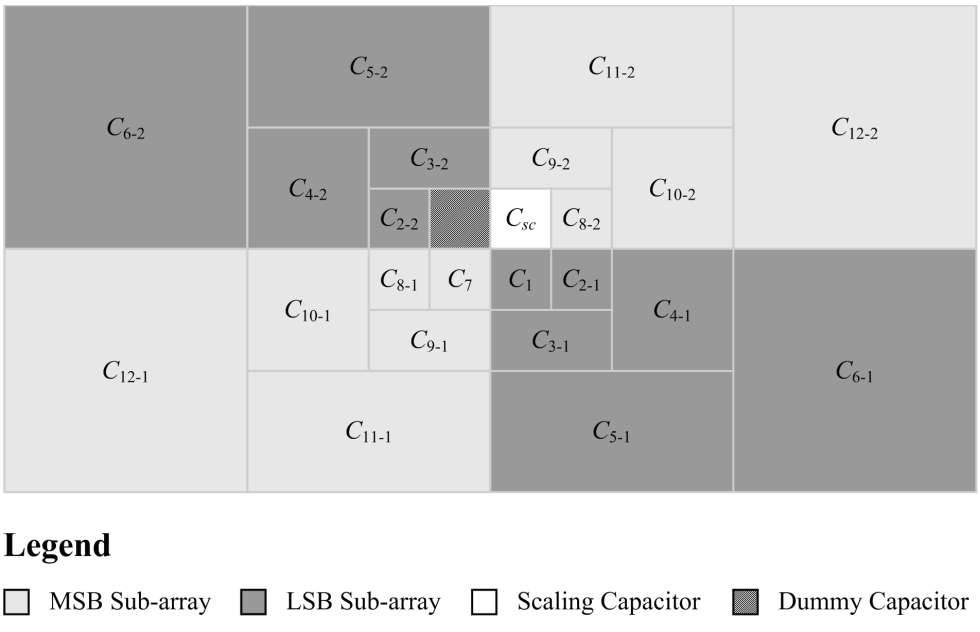


Figure 5.6: The common-centroid layout of the capacitor array.

centroid configuration to form the scaled array, as illustrated in Fig. 5.6. Placed next to the centroid is the scaling capacitor C_{sc}, whose shielded (Poly 2) plate is connected to the common node of the LSB sub-array, given it is better protected from parasitic errors than the shielding (Poly 1 and Metal 2) plates. The dummy capacitor is an un-connected unit capacitor for matching purpose. The perimeter of the entire matrix is enclosed by a ring of dummy capacitors to ensure topological continuity.

An improved approach could be enclosing the C_{sc} with the MSB sub-array matrix, and optimizing the matching of this matrix. The LSB sub-array, on the other hand, can be laid out in a less restricted manner. As aforementioned, this is because compared to an intrinsic binary array with the same C_n, the lower q capacitors in the scaled array are enlarged by 2^q while still retaining the k_i specifications, substantially relieving the matching challenges.

5.3.4 Static Behavioral Simulation

The operation of the capacitive DAC within the SAR ADC consists of a sequence of repeating events, which are equivalent to a series of numerical operations at very low frequency. It is thereby convenient to describe these *static* behaviors using a pure numerical model in MATLAB and evaluate the impact of *static* non-idealities to the transfer characteristics of the DAC. Since the produced results are fully consistent with the analytical ones in Subsection 5.3.2, they will not be discussed in details. Instead, the following text will only address one example non-ideality factor and its impact.

Consider the capacitor matrix in Fig. 5.6. The outer edges of C_{12}, C_{11}, C_6 and C_5

are extended by floating dummy capacitors to achieve uniform process loading. These dummy capacitors, however, create fringing capacitances that increase the switching capacitances of the affected capacitors. Furthermore, the proximity of the MSB and LSB sub-arrays also introduces fringing capacitance between the two that equivalently increases the size of C_{sc}, even if no mismatch has occurred. Fortunately, the fringing factor in this case is considerably smaller than the outer edge one for the floating dummy capacitors, because much of the fringing field is effectively blocked by the Metal 2 and Poly 1 shields that are tied to a fixed potential.

Assume the fringing factor for the floating dummy capacitors is 0.02 fF/μm, and for the sub-array interface is 0.004 fF/μm, and we may obtain the simulated static nonlinearities of the DAC as shown in Fig. 5.7. The peak *DNL* and *INL* are approximately ±0.2 LSB and ±0.3 LSB, respectively. Note that such nonlinearities are largely induced by perimeter discontinuities. An improved solution could be biasing the outer

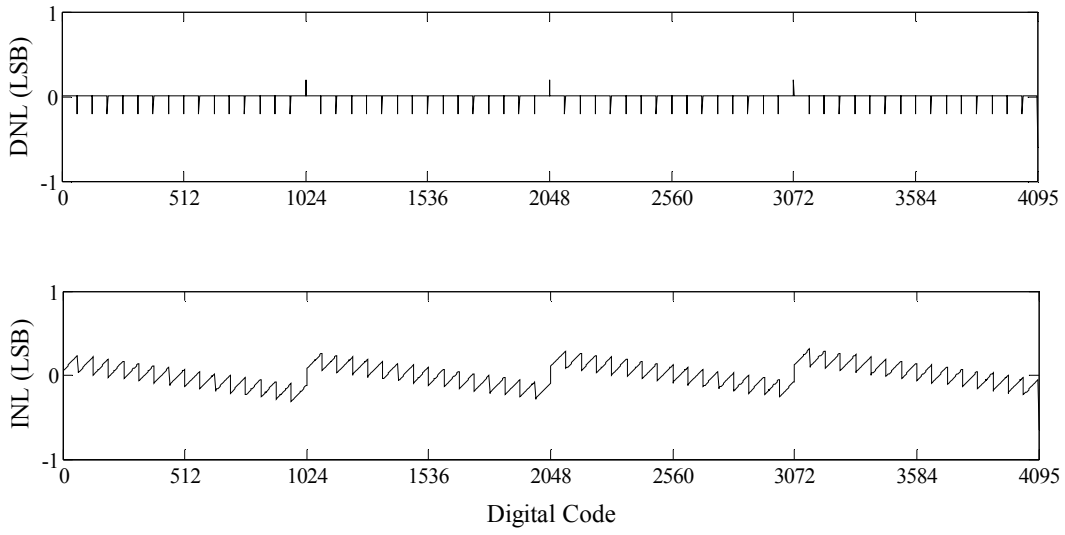


Figure 5.7: Simulated DAC nonlinearities due to fringing capacitances.

edge dummy capacitors to a fixed potential such that the associating fringing capacitances are no longer part of the switching capacitances, but only stray errors on the two common nodes of the capacitor array.

5.4 The SAR Logic and Timing Sequence Modules

The SAR logic circuit for two consecutive bits is shown in Fig. 5.8. The upper row of D flip-flops forms a shift register that activates the logic operations for every following bit at each rising edge of the CLK signal. When activated, the respective D flip-flop in the lower row calculates the current bit of digital input to the DAC based on the comparator output (CMP_OUT). The result is then latched to the DAC in the next CLK cycle.

The timing sequences that drive the SAR logic are generated onchip and shown in

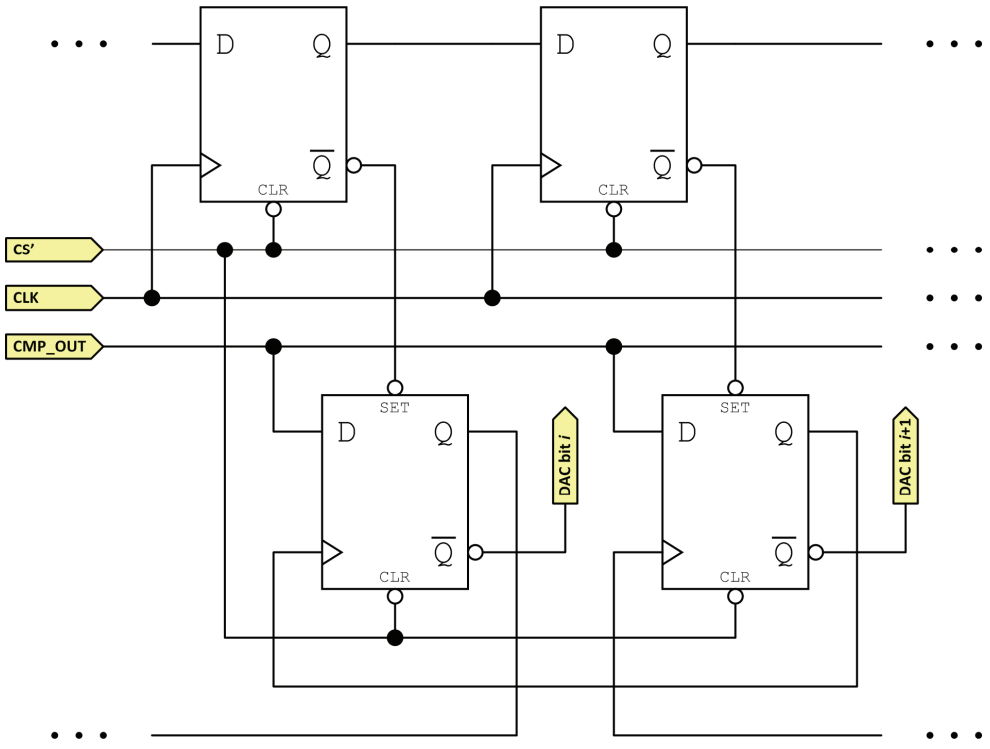


Figure 5.8: Simplified circuit diagram of the SAR logic module.

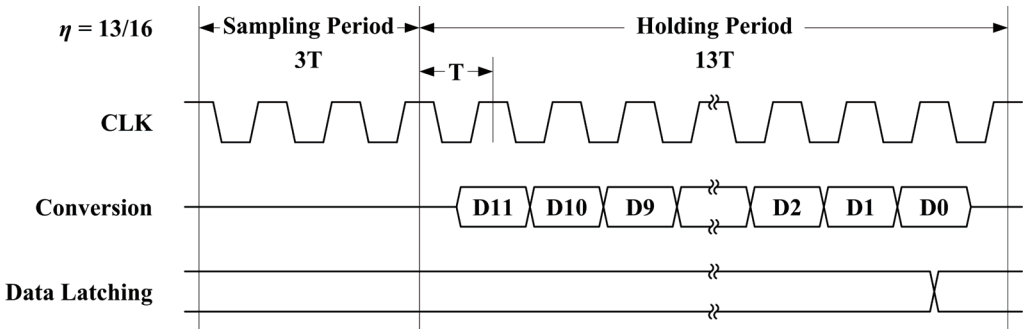


Figure 5.9: Implemented SAR timing sequences.

Fig. 5.9. 3 out of 16 CLK periods in each ADC cycle are dedicated to signal sampling, and the rest to data conversion. The resultant η (approximately 0.8) combined with Inequality 3.13 gives a minimally required PGA bandwidth of 5 kHz.

5.5 The Relaxation Oscillator

In addition to the crystal driver, a less accurate but more power conserving relaxation oscillator is included for the less demanding *QRS* detection applications. Resem-

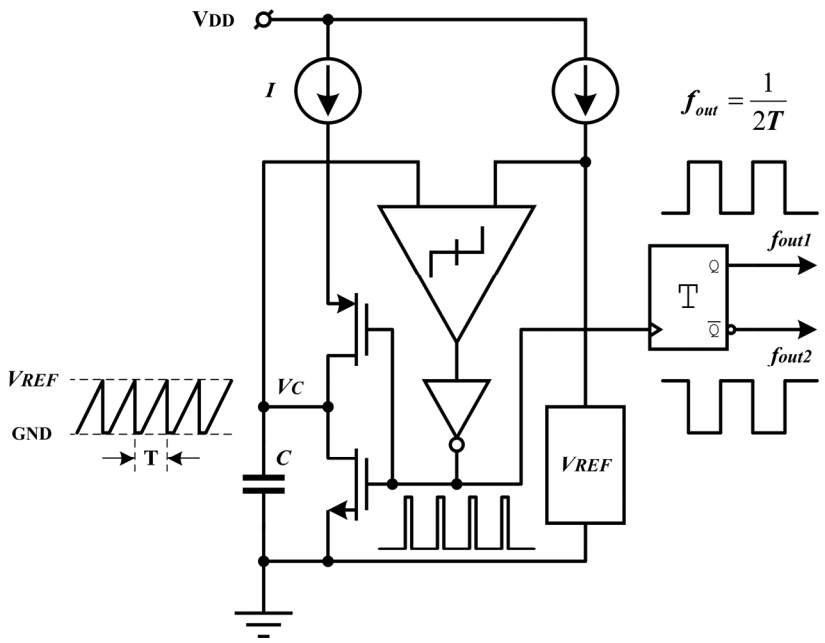


Figure 5.10: Simplified circuit diagram of the relaxation oscillator.

bling the operation of a 555 timer, the oscillator charges a capacitor with a fixed current and resets it once the voltage has reached the predefined threshold (see Fig. 5.10). Both the current source and the voltage threshold are sourced from an accurate reference generator, and the oscillator can operate at 30 kHz with a wide supply range of 0.8 to 3.3 V, and a small current consumption of 50 nA.

Chapter 6

Design Verification

6.1 Sensor Interface Circuits

The proposed sensor interface circuits were fabricated and verified in AMS 0.35-μm CMOS baseline process, chip microphotograph being shown in Fig. 6.1. The test results were published in [1] and [2], and are briefly summarized in this chapter.

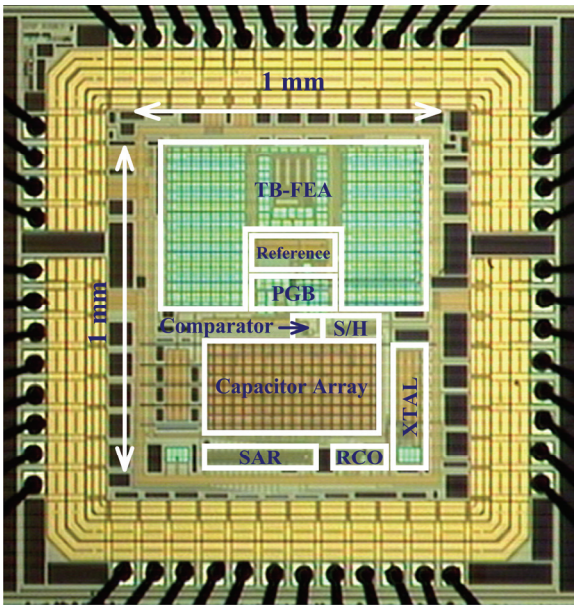


Figure 6.1: Microphotograph of the sensor interface chip.

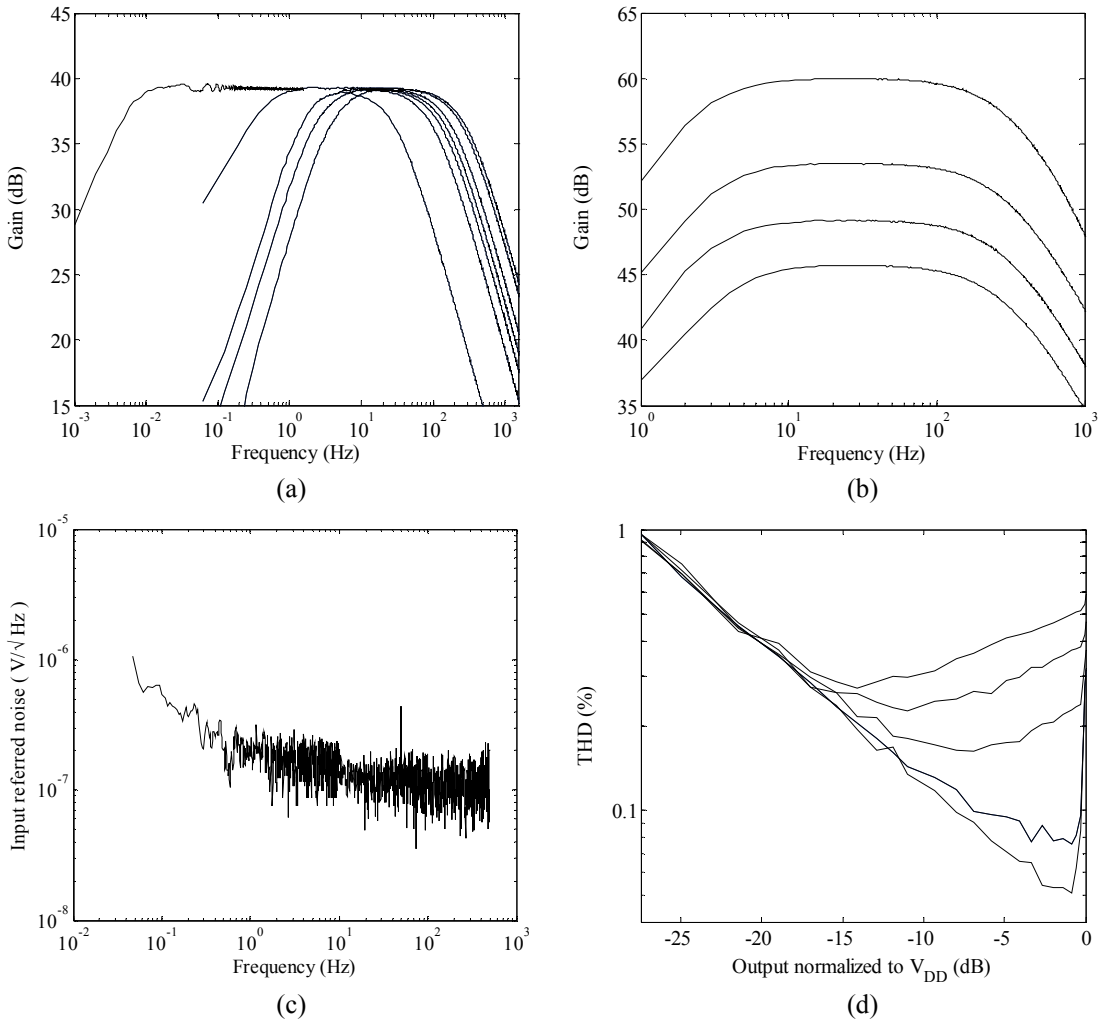
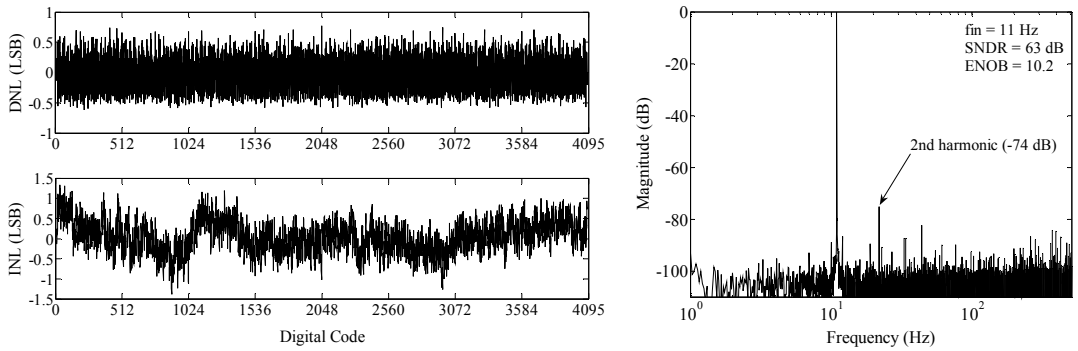


Figure 6.2: Measured performance of the frontend amplifiers.

Fig. 6.2 plots some of the measured parameters for the frontend amplifiers. In (a), frequency response curves with different bandwidth settings are shown. The high-pass corner frequency proves to be tunable from 4.5 mHz to 3.6 Hz, and the low-pass corner frequency from 31 Hz to 292 Hz; both of which provide sufficient tuning range to fulfill different application requirements and compensate for potential process variations. In (b), frequency response curves with different gain settings are shown. The mid-band gain values are measured to be 45.6 dB, 49 dB, 53.5 dB and 60 dB, respectively. No apparent low frequency gain hump/dip or bandwidth-gain dependence is

observed. (c) shows the input-referred noise spectrum when the low-pass cutoff is set at 292 Hz. Integrating it over the effective noise bandwidth (from 0.05 Hz to 460 Hz, assuming single pole system) gives 2.5 μV_{rms} of input-referred rms noise. (d) shows the total harmonic distortion (*THD*) in percentage against the normalized output swing. The curves from top to bottom correspond to high-pass cutoffs at 3.6 Hz, 2.2 Hz, 1.4 Hz, 0.25 Hz, and 4.5 mHz, respectively. Evidently, the proposed frontend architecture and the fully balanced pseudo-resistor structure have managed to keep the *THD* below 0.6% even at full output swing. Note that the measured harmonic power contains both that from the actual tones and that from noise. At small output level when distortions are overwhelmed by the noise floor, the *THD* curves display as a straight noise-to-signal line. At higher output level when distortions begin to manifest, the *THD* curves deviate from the straight line and fit towards the actual *THD* values. Also observed from the plot is that the deviation initiates at smaller output level when the cutoff frequency is set at higher values. This is because the pseudo-resistors in such cases are biased to produce smaller resistance, whose nonlinearity has more severe interference with the reactance of the loop capacitors in the frequency band of interest.

Fig. 6.3 plots some of the measured parameters for the SAR ADC. Summarized in (a) are the static nonlinearity test results, wherein the upper diagram shows the *DNL*, and the lower one shows the *INL*, both obtained from code density statistics. The worst-case *DNL* is within +0.8/−0.6 LSB, and *INL* within ±1.4 LSB. Though distorted by other nonlinearity factors, the *INL* plot still reveals the vague trend as predicted in Fig. 5.7. Shown in (b) is the 32768-point Fast Fourier transform (FFT) result when the ADC is fed with an 11-Hz, 0.95-$V_{p\text{-}p}$ sinusoidal input. The measured spurious-free



(a) (b)

Figure 6.3: Measured performance of the SAR ADC.

dynamic range (*SFDR*) is 74 dB. The calculated signal-to-noise-plus-distortion ratio (*SNDR*) is 63 dB, corresponding to 10.2 effective-number-of-bits (*ENOB*).

Table I summarizes the key design parameters of the implemented sensor interface circuits. Operated at 1-V supply, the single chip solution provides various operation options across the programmable range confined by two extremities: the mini-

Table 6.1: Design parameters of the sensor interface chip.

Parameter	Value
Process Technology	0.35-µm CMOS baseline
Supply Voltage	1.0 V
Frontend Current	33 ~ 337 nA
Frontend Input Impedance	> 100 MΩ @ 10 Hz
Frontend Gain	45.6 / 49 / 53.5 / 60 dB
−3 dB High-pass Cutoff	4.5 mHz ~ 3.6 Hz
−3 dB Low-pass Cutoff	31 ~ 292 Hz
Frontend Noise (Input Referred)	2.5 μV_{rms} (0.05 ~ 460 Hz)
Frontend *NEF*	3.26
Frontend *THD* @ FS Output	< 0.6%
Frontend Dynamic Range	57 dB
Frontend *CMRR*	≥ 71.2 dB (Below 300 Hz)
Frontend *PSRR*	≥ 84 dB (Below 300 Hz)
ADC Sampling Frequency	1 kS/s
ADC *DNL*	+0.8/−0.6 LSB
ADC *INL*	±1.4 LSB
ADC *SFDR*	74 dB
ADC *SNDR*	63 dB
ADC *ENOB*	10.2
Jitter (Relaxation Oscillator)	< 3 μs_{rms}
Total Power	445 ~ 895 nW

mum band *QRS* detection mode and the full band ECG acquisition mode. The former mode sets the frontend low-pass cutoff to 31 Hz and uses the power conserving relaxation oscillator as the clock source. The measured rms jitter from the relaxation oscillator is less than 3 μs_{rms}, 3 times of which produces less than −60 dBFS of tracking error, and less than 10 ppm of typical *QRS* timing error. The power consumption in this case is minimal and measured to be 445 nW. The latter mode sets the passband to its full capacity of 4.5 mHz to 292 Hz and uses the more accurate crystal driver as the clock source. Such configuration provides the highest fidelity for diagnostic grade ECG recording, and requires no higher than 895 nW of power consumption.

6.2 Wearable ECG Device Prototype

Upon successful design verification, the ECG sensor interface chip was used to build the ECG plaster prototype as shown in Fig. 6.4. On this wearable device, ECG signals are captured, conditioned and digitized by the sensor interface; the produced raw data are then preprocessed and either stored locally to a micro SD card or sent out

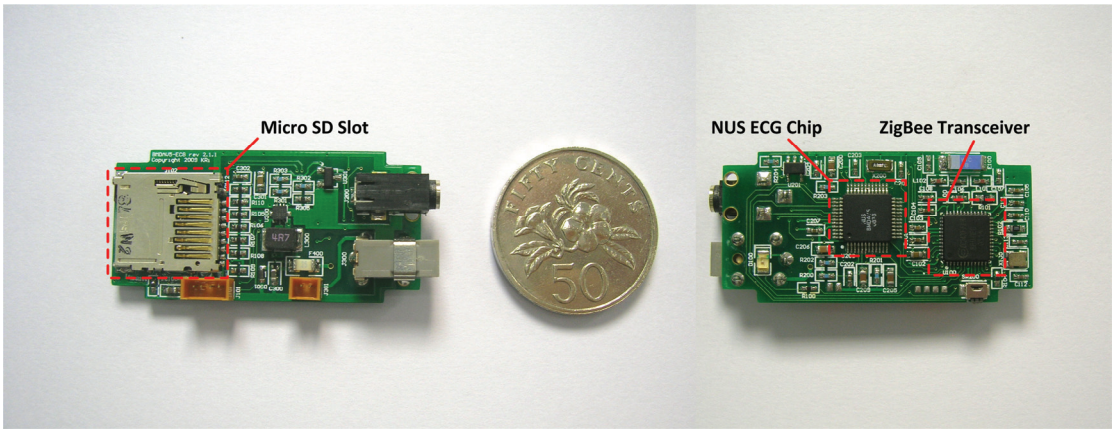


Figure 6.4: ECG plaster prototype with ECG sensor interface chip.

(a) Diagnostic Mode (4.5 mHz – 292 Hz)

(b) Monitoring Mode (2.2 Hz – 31 Hz)

Figure 6.5: Recorded Lead-II ECG over ZigBee radio.

to base stations or routers over ZigBee radio.

Fig. 6.5 shows two traces of Lead-II ECG recorded by the plaster device. Each square measures 1 mm by 1 mm. The upper trace was taken in the full band acquisition/diagnostic mode, and the lower one in one of the minimum band *QRS* detection/monitoring modes. Note how the high-pass setting in the latter case distorts the *T*, *U* and *P* waves. It should therefore be avoided in formal diagnoses. In *QRS* detection applications, however, an increased high-pass cutoff as such may be used with caution, since it offers better baseline filtering yet brings little impact to the *QRS* peaks. Comparing the two traces on high frequency tones, we may infer that majority of the ECG energy resides below 30 Hz, as no apparent information loss or distortion is observed on the minimum band trace.

Chapter 7

Conclusion

Good healthiness, both physically and mentally, has been one of the eternal quests ever since human civilization came into existence. While it may take different meanings for different individuals at different times, the advance of today's microelectronic and bioengineering technology has undoubtedly created more options that intend to provide means and convenience to such quest. In this work, we have demonstrated an example approach whereby ultra-portable personal telemetric medical devices, or more specifically wearable ECG sensors, become well feasible.

Operated at 1-V supply, the proposed single channel ECG chip consumes only 445 nW in the minimum band *QRS* detection mode, and 895 nW in the full band ECG acquisition mode, featuring the world's first sub-μW ECG sensor interface. The ultra-low power consumption is a combined product of the subthreshold biasing scheme that offers the highest transconductance efficiency, and the optimal system partitioning that is derived from the most efficient S/H duty ratio.

Some of the key innovations in this work include the fully balanced cross-coupled

tunable pseudo-resistor structure that overcomes the tradeoff between tunability and electrical balance as embedded in most of conventional pseudo-resistor solutions; and the "flip-over-capacitor" gain control scheme that rectifies the low frequency gain interruption problem due to the finite off-state resistance of the gain switch.

With the successful verification of the single channel ECG sensor interface design, the next step would be extending to the 12-lead ECG solution and further supporting other forms of physiological signals. Prior to that, a few improvements over the current design may be considered. First, the analog frontend could upgrade to a fully differential topology, which produces more balanced input loading. Second, each high impedance node biased by pseudo-resistors should be paired with a reset facility that assists fast and reliable settling. Third, a right leg driver (DRL) circuit could be added to actively suppress common-mode interferences, and in particular the power line noise.

As mentioned at the beginning of this dissertation, concerned with population ageing and life quality, people are becoming increasingly aware of prevention-oriented personal telemetric healthcare. The fast booming of microelectronic technology, material and biomedical sciences today has been actively paving the path into this new territory. Therefore, we can expect more and more innovations that may gradually reshape our way of life for the years to come.

Bibliography

[1] X. Y. Xu, X. D. Zou, L. B. Yao, and Y. Lian, "A 1-V 450-nW Fully Integrated Biomedical Sensor Interface System," in *Proc. 2008 Symposium on VLSI Circuits*, Jun. 2008, pp. 78–79.

[2] X. D. Zou, X. Y. Xu, L. B. Yao, and Y. Lian, "A 1-V 450-nW Fully Integrated Programmable Biomedical Sensor Interface Chip," *IEEE J. Solid-State Circuits*, vol. 44, pp. 1067–1077, Apr. 2009.

[3] X. D. Zou, X. Y. Xu, Y. Lian, and Y. J. Zheng, "A Low Power Sensor Interface IC for Wearable Wireless Biomedical Devices," in *Proc. ICST 2nd Int. Conf. on Body Area Networks*, Jun. 2007, pp. 22–23.

[4] X. D. Zou, X. Y. Xu, J. Tan, L. B. Yao, and Y. Lian, "A 1-V 1.1-μW Sensor Interface IC for Wearable Biomedical Devices," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2008)*, May 2008, pp. 2725–2728.

[5] M. C. Munshi, X. Y. Xu, X. D. Zou, E. Soetiono, C. S. Teo, and Y. Lian, "Wireless ECG Plaster for Body Sensor Network," in *Proc. 5th Int. Workshop on Wearable and Implantable Body Sensor Networks (BSN 2008)*, Jun. 2008, pp. 310–313.

[6] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*. New York: Oxford University Press, 1995.

[7] *Electrocardiography*. Wikipedia, the Free Encyclopedia. 17 Dec. 2009. 02 Jan. 2010. <http://en.wikipedia.org/wiki/Electrocardiography>

[8] *Cardiac Cycle*. Wikipedia, the Free Encyclopedia. 24 Nov. 2009. 02 Jan. 2010. <http://en.wikipedia.org/wiki/Cardiac_cycle>

[9] B. B. Winter and J. G. Webster, "Driven-Right-Leg Circuit Design," *IEEE Trans. on Biomedical Engineering*, vol. BME-30, pp. 62–66, Jan. 1983.

[10] R. R. Harrison and C. Charles, "A Low-Power Low-Noise CMOS Amplifier for Neural Recording Applications," *IEEE J. Solid-State Circuits*, vol. 38, pp. 958–965, Jun. 2003.

[11] R. F. Yazicioglu, P. Merken, R. Puers, and C. V. Hoof, "A 60 μW 60 nV/√Hz Readout Front-End for Portable Biopotential Acquisition Systems," *IEEE J. Solid-State Circuits*, vol. 42, pp. 1100–1110, May 2007.

[12] W. Wattanapanitch, M. Fee, and R. Sarpeshkar, "An Energy-Efficient Micro-power Neural Recording Amplifier," *IEEE Trans. on Biomedical Circuits and Systems*, vol. 1, pp. 136–147, Jun. 2007.

[13] M. Yin and M. Ghovanloo, "A Low-Noise Preamplifier with Adjustable Gain and Bandwidth for Biopotential Recording Applications," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2007)*, May 2007, pp. 321–324.

[14] M. Chae, W. Liu, Z. Yang, T. Chen, J. Kim, M. Sivaprakasam, and M. Yuce, "A 128-Channel 6mW Wireless Neural Recording IC with On-the-Fly Spike Sorting and UWB Transmitter," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2008, pp. 146–147.

[15] H. Wu and Y. P. Xu, "A 1V 2.3μW Biomedical Signal Acquisition IC," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2006, pp. 119–120.

[16] M. S. J. Steyaert, W. M. C. Sansen, and Z. Chang, "A Micropower Low-Noise Monolithic Instrumentation Amplifier For Medical Purposes," *IEEE J. Solid-State Circuits*, vol. 22, pp. 1163–1168, Dec. 1987.

[17] D. M. Binkley, B. J. Blalock, and J. M. Rochelle, "Optimizing Drain Current, Inversion Level, and Channel Length in Analog CMOS Design," *Analog Integrated Circuits and Signal Processing*, vol. 47, pp. 137–163, May 2006.

[18] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*. Hoboken: John Wiley & Sons, 2008.