

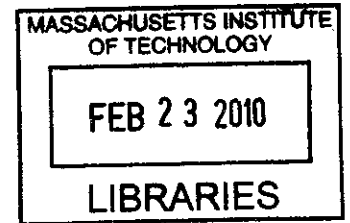
# Digitally-Assisted, Ultra-Low Power Circuits and Systems for Medical Applications

by

Jose L. Bohorquez

B.S., University of Florida (2002)

M.S., University of Florida (2004)



Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2010

**ARCHIVES**

© Massachusetts Institute of Technology 2010. All rights reserved.

Author .....  
Department of Electrical Engineering and Computer Science  
January 7, 2010

Certified by .....  
Joel L. Dawson  
Associate Professor  
Thesis Supervisor

Certified by .....  
Anantha P. Chandrakasan  
Professor  
Thesis Supervisor

Accepted by .....  
Terry Orlando  
Chairman, Department Committee on Graduate Theses



# Digitally-Assisted, Ultra-Low Power Circuits and Systems for Medical Applications

by

Jose L. Bohorquez

Submitted to the Department of Electrical Engineering and Computer Science  
on January 7, 2010, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Electrical Engineering

## Abstract

In recent years, trends in the medical industry have created a growing demand for a variety of implantable medical devices. At the same time, advances in integrated circuits techniques, particularly in CMOS, have opened possibilities for advanced implantable systems that are very small and consume minimal energy. Minimizing the volume of medical implants is important as it allows for less invasive procedures and greater comfort to patients. Minimizing energy consumption is imperative as batteries must last at least a decade without replacement. Two primary functions that consume energy in medical implants are sensor interfaces that collect information from biomedical signals, and radios that allow the implant to communicate with a base-station outside of the body. The general focus of this work was the development of circuits and systems that minimize the size and energy required to carry out these two functions. The first part of this work focuses on laying down the theoretical framework for an ultra-low power radio, including advances to the literature in the area of super-regeneration. The second part includes the design of a transceiver optimized for medical implants, and its implementation in a CMOS process. The final part describes the design of a sensor interface that leverages novel analog and digital techniques to reduce the system's size and improve its functionality. This final part was developed in conjunction with Marcus Yip.

Thesis Supervisor: Joel L. Dawson  
Title: Associate Professor

Thesis Supervisor: Anantha P. Chandrakasan  
Title: Professor





## Acknowledgments

Whatever challenges I have faced in life, they pale in comparison to the countless blessings God has granted me. I, therefore, thank Him above all for the great honor of having had the opportunity to study at MIT, and for giving me strength in times of difficulty.

Secondly, I thank Professors Joel Dawson and Anantha Chandrakasan for giving me the opportunity to be part of their research groups and the privilege of their guidance. Both their areas of expertise and their advising styles were complementary in a way that tremendously enriched my experience as a graduate student. They gave me the freedom to explore areas that I was interested in while guiding me to ensure my research was relevant. They were understanding when personal challenges required my absence; they were encouraging when I faced research difficulties. They also gave me access to two groups of exceptional research colleagues with a broad range of knowledge and experience that I was able to draw from to enhance my own research; a group of students who are as friendly and cooperative as they are bright.

I am very grateful to the members of both research groups for sharing their ideas and always being available to bounce my own off of. I am especially grateful to Patrick Mercier and Willie Sanchez. They reviewed my papers and often offered excellent advice on my research, but more importantly, they have been great friends whom I've enjoyed spending time with. I also owe a great debt of gratitude to Marcus Yip who was a pleasure to work with and whose efforts contributed directly and significantly to my thesis.

Over the last few years, my wife and I have been blessed with a handful of great friends whom we have shared much joy with. Jaime Andres and Marlyn Echeverry have been close friends for over a decade now, and we were extremely fortunate that fate brought them to Boston for two years during which we shared times of sorrow and incredible joy. We watched Gator football together and ate many steaks. We had the blessing to be there when their daughter, little Natalia, was born, and the gift of having them with us when our daughter, Isabela, was baptized. We had con-

tentious discussions on politics and edifying conversations on faith, all along fortifying a friendship that will surely last a lifetime.

During my first year at MIT we were also fortunate to meet Hoda Eydgahi, Miriam Makhoul, Michael Schnall-Levin, Samantha Green-Atchley, and Stephanie Gil. Despite being a little crazy, we all helped each other keep sane by sharing in social diversions from our academic obligations.

Although I have already thanked Willie Sanchez, I must thank him again along with his fiancée, Shirley Cardona, for being great friends over the past year. What began with their offer to baby-sit Isabela has grown into a wonderful friendship. Their love for Izzy means the world to us, and their friendship is a true blessing.

Before joining MIT and on multiple occasions during my studies, I had the good fortune to work with a team of exceptional engineers at BitWave Semiconductor. I am grateful to that entire team, especially Geoff Dawe, Russ Cyr, and Don Young Chang. They have been good friends and invaluable mentors.

Over the last year, I have also had the great fortune of meeting Al-Thaddeus Avestruz, Michael Rinehart, Seward Rutkove, and Lynne Levitsky. The exciting endeavor of starting a company with them has been an experience of great personal growth and new aspirations. I feel humbled to be part of such a talented team and look forward to our continued collaboration with great enthusiasm.

Finally, I thank my family for their love and support. All of the opportunities I have had in life were afforded by the many sacrifices my parents, Francisco and Clara, made. My brother, Jesus Alberto, and sister, Maria Carolina, have always been examples of faith and love in my life. Throughout my studies, their support and encouragement has been constant. My nieces, Natalie, Tatiana, and Daniela, have been sources of pride and joy since their birth. I also feel blessed to have married into a wonderful family and thank Carmen, Melissa, Sebastian, Humberto, Santiago, and Valentina for the support and joy they have brought into our lives.

Of course, the person who has sacrificed most throughout the last few years and whose unwavering support has been my greatest source of strength and comfort is my wife Natalia. She has been patient and understanding beyond reasonable expect-

tations; through wedding planning; through her pregnancy with Isabela; through our daughter's first two years; and, now, through her second pregnancy, she has always adjusted her life to lighten my load. My achievements are as much hers as they are mine, and I will always be grateful for her loving encouragement. I must also thank our daughter Isabela. Though she may not understand it now, her playful affection has brought me great peace, rest, and joy during the moments of greatest stress and frustration.



# Contents

<b>1</b>	<b>Introduction</b>	<b>25</b>
1.0.1	Trends in Medical Electronics . . . . .	25
1.0.2	Medical Implant Communication Service Band . . . . .	26
1.0.3	Biomedical Sensor Issues . . . . .	30
1.0.4	Research Goals and Thesis Organization . . . . .	31
<b>2</b>	<b>Frequency-Domain Analysis of Super-Regenerative Amplifiers</b>	<b>33</b>
2.1	Introduction . . . . .	33
2.2	General SRA Theory . . . . .	35
2.2.1	Circuit Model and Block Diagram for SRA . . . . .	35
2.2.2	SRA as Time-Varying, Second-Order Systems . . . . .	37
2.2.3	General SRA Solution for Linear Mode Operation . . . . .	38
2.2.4	SRA Solution for a Ramp and Sine Damping Functions . . . . .	40
2.3	Convolution Model of SRA Solution . . . . .	42
2.3.1	SRA Response to an Arbitrary Input . . . . .	45
2.3.2	SRA Response to a Sinusoidal Input . . . . .	47
2.3.3	SRA Response to Multiple Sinusoids . . . . .	50
2.3.4	SRA Response to a Pulse-Shaped Sinusoidal Input . . . . .	51
2.4	Receiver Sensitivity Analysis . . . . .	53
2.4.1	Noise Analysis using the Convolution Model . . . . .	56
2.4.2	Solving for the BER and Sensitivity . . . . .	57
2.4.3	Using a Time Random Variable for Detection . . . . .	60
2.5	Measurement Results . . . . .	63

2.6	Summary . . . . .	66
2.7	Appendix 1: Convolution yielding the variance of $\mathbf{I}_X$ . . . . .	68
2.8	Appendix 2: Solving for the PDF of $\mathbf{T}_T$ . . . . .	69
<b>3</b>	<b>Ultra-Low Power Transceiver for Medical Implant Communications</b>	<b>71</b>
3.1	Introduction . . . . .	72
3.2	Architecture Overview . . . . .	73
3.2.1	Exploiting the Unique Features of Medical Implants . . . . .	74
3.2.2	Link Analysis and Antenna Considerations . . . . .	75
3.3	FSK/MSK Transmitter . . . . .	77
3.3.1	FSK/MSK Theory . . . . .	78
3.3.2	Transmitter Implementation . . . . .	79
3.3.3	Capacitor Array with Predistortion . . . . .	79
3.3.4	Setting $f_C$ and $\Delta F$ . . . . .	82
3.4	Super-Regenerative Receiver . . . . .	82
3.4.1	Super-Regeneration Theory . . . . .	82
3.4.2	Receiver System Design . . . . .	85
3.5	Circuit Implementation . . . . .	87
3.5.1	DCO Implementation . . . . .	87
3.5.2	Programmable Ramp Quench Oscillator . . . . .	89
3.5.3	Envelope Detector . . . . .	92
3.5.4	Programmable Comparator . . . . .	93
3.6	Prototype Frequency Correction Loop . . . . .	94
3.7	Measurement Results . . . . .	96
3.7.1	Transmitter . . . . .	97
3.7.2	Receiver . . . . .	100
3.8	Summary . . . . .	103
<b>4</b>	<b>Digitally-Assisted Sensor Interface for Biomedical Applications</b>	<b>105</b>
4.1	Applications . . . . .	106
4.2	Top-Level Issues and Requirements . . . . .	108

4.2.1	Thermal Noise and the Noise Efficiency Factor . . . . .	110
4.2.2	Flicker Noise . . . . .	113
4.2.3	Using Chopping to Mitigate Flicker Noise . . . . .	114
4.2.4	Power Line Interference . . . . .	116
4.2.5	Techniques for Mitigating the Effects of PLI . . . . .	119
4.2.6	Electrode and IAMP DC Offset . . . . .	125
4.2.7	Dealing with Electrode DC Offset . . . . .	126
4.2.8	Top-Level Design Requirements . . . . .	127
4.3	Instrumentation Amplifier Design . . . . .	128
4.4	Sinc Anti-Aliasing Filter . . . . .	139
4.5	Mixed-Signal Feedback for Interference Cancelation . . . . .	151
4.5.1	Motivation . . . . .	151
4.5.2	PLI Considerations . . . . .	154
4.5.3	Prototype System Design . . . . .	155
4.6	Measurement Results . . . . .	169
4.6.1	Frequency Response . . . . .	173
4.6.2	Input-Referred Noise Measurement . . . . .	178
4.6.3	Total Harmonic Distortion . . . . .	181
4.6.4	IAMP Input Impedance . . . . .	183
4.6.5	Notch Filter . . . . .	186
4.7	Summary . . . . .	194
<b>5</b>	<b>Conclusion and Future Work</b>	<b>195</b>
5.1	Summary of Contributions . . . . .	195
5.2	Future Work . . . . .	198
<b>A</b>	<b>Chopper-Based IAMPs</b>	<b>201</b>
A.1	A Simple Chopper-Based IAMP . . . . .	201
A.2	Current Feedback IAMP with EDO Cancelation . . . . .	204
A.3	OpAmp-based IAMP with EDO Cancelation . . . . .	206
A.4	IAMP based on a Chopper-Stabilized OpAmp . . . . .	208





# List of Figures

1-1	Wireless link between patient and healthcare provider. . . . .	28
2-1	(a) SRA circuit model, and (b) SRA feedback loop model. . . . .	35
2-2	Time-varying pole/zero locations for an SRA as the damping function changes. . . . .	37
2-3	Sensitivity functions for sawtooth/ramp and sine damping functions ( $\gamma = 3$ ). . . . .	42
2-4	(a) Damping function, (b) sensitivity function, (c) time derivative of sinusoidal input, (d) windowed input, and (e) $k(t)$ . . . . .	44
2-5	Numerical simulations comparing the exact equation $k(t)$ with the convolution approximation $k_*(t)$ for $\gamma = 3$ and (a) $\omega_a = \omega_0$ , (b) $\omega_a = \omega_0 + 2\Omega_s$ , and (c) $\omega_a = \omega_0 + 4\Omega_s$ . . . . .	45
2-6	Graphical representation of SRA response to a sinusoidal input signal. . . . .	48
2-7	Frequency response using a sawtooth/ramp damping function and varying values of $\gamma$ . . . . .	50
2-8	Frequency response using a sawtooth/ramp and sine damping functions ( $\gamma = 3$ ). . . . .	51
2-9	SRA (a) transconductance $G_m(t)$ , (b) damping function $\zeta(t)$ , (c) input current $i_a(t)$ , (d) output voltage for linear system $v_o(t)$ , and (e) output voltage for compressive nonlinear system. . . . .	61
2-10	Schematic of SRA and envelope detector. . . . .	64

2-11	Theoretical and measured envelope detector waveforms for -75dBm and -85dBm CW input signals and two sample waveforms for no input signal (i.e. only noise). . . . .	67
2-12	Measured and theoretical probability density functions for the random variable $t_T$ . . . . .	67
2-13	Frequency response of the SRA for $\Omega_s/(2\pi) = 480\text{kHz}/\sqrt{\text{rad}}$ . . . . .	68
3-1	Transceiver block diagram. . . . .	73
3-2	(a) Loop antenna with FR4 substrate and superstrate, and copper patch below the substrate. (b) Loop antenna with metal patches above and below the substrate and superstrate. . . . .	76
3-3	(a) Direct modulation FSK transmitter and (b) simplified circuit model. . . . .	78
3-4	(a) Sub-ranging capacitor array and (b) piece-wise linear predistortion. . . . .	80
3-5	Frequency steps vs. $N_F$ using linear, predistorted, and piece-wise linear (PWL) predistorted capacitor banks. . . . .	81
3-6	(a) SRR circuit model, and (b) SRR feedback loop model. . . . .	83
3-7	(a) Differential Colpitts oscillator, (b) half circuit model, (c) equivalent half circuit, (d) noise model, and (e) equivalent noise model. . . . .	88
3-8	Differential Colpitts digitally-controlled oscillator with predistorted sub-ranging capacitor banks and loop antenna. . . . .	90
3-9	(a) Sawtooth oscillator, SRR bias generator, and baseband clock. (b) Output signals. . . . .	91
3-10	Schematic for differential envelope detector and comparator with programmable input offset voltage. . . . .	93
3-11	Measurements of the frequency correction loop including the transition of the DCO frequency, coarse tuning coefficient, and fine tuning coefficient as the channel is changed from (a) 7 to 8, (b) 1 to 10 incrementally, and (c) 1 to 10 in a single command. . . . .	95
3-12	Die photograph ( $1.0 \times 0.5\text{mm}$ active area). . . . .	97

3-13	DCO frequency for (a) coarse, (b) medium, and (c) fine tuning. Thermometer coding and predistortion of each capacitor bank lead to monotonic, linear tuning for each range. . . . .	98
3-14	Measured spectral mask of transmitter with a data rate of (a) 200kbps and (b) 120kbps. The signal is taken through an antenna 10cm from the transmitter. . . . .	99
3-15	Power received at an antenna placed 20cm from the transmitter and DCO frequency vs. bias current. . . . .	100
3-16	Receiver chain time domain signals. . . . .	101
3-17	(a) Far off and (b) close in CW blocker rejection. Input signal level set to sensitivity + 6dB = -93dBm with a carrier frequency of 403.35MHz and a bit rate of 40kbps. . . . .	102
4-1	Block diagram of a typical sensor interface including the off-chip and on-chip aggressors that must be dealt with. . . . .	108
4-2	Frequency-domain analysis of the signal and noise at different stages of a chopper-based sensor interface system. . . . .	115
4-3	Classical model for power line interference in two-electrode measurements ( $SW_{E3}$ off), and three-electrode measurements ( $SW_{E3}$ on). . .	118
4-4	Simplified driven-right-leg circuit for safe common-mode voltage reduction. . . . .	123
4-5	Typical connection of EEG amplifiers to their respective electrodes. .	124
4-6	Electrode configuration for standard 10-20 electrode system (black dots) with additional electrodes for 10-10 system (gray dots). Figure taken from [46] . . . . .	124
4-7	Instrumentation amplifier using MOS-bipolar pseudoresistor elements [25]. . . . .	127
4-8	Instrumentation amplifier architecture used for the work presented in this thesis. Although the chopping switches were not used, they are shown here to analyze their effect. . . . .	129

4-9	(A) Single ended equivalent of IAMP in Fig. 4-8 including noise sources, and (B) corresponding block diagram. . . . .	130
4-10	Schematic of fully differential operational amplifier including second stage choppers. These choppers are not used in normal operation but are included here for completeness. . . . .	132
4-11	Simplified schematic of fully differential opamp without choppers. . .	133
4-12	(A) Frequency-domain illustration of desired signal band (white) and aliasing components (gray) that would corrupt the signal after sampling. (B) A typical anti-aliasing filter amplitude response that filters aliasing components by some required amount. (C) The minimum required amplitude response of an anti-aliasing filter. The dark gray boxes represent bands that can be aliased without corrupting the signal band. (D) Discrete-time spectral component between $-f_s/2$ and $f_s/2$ illustrating how the desired channel is not corrupted and the non-filtered components of (C) are frequency translated to the bands between the channel edges and the Nyquist frequencies. . . . .	141
4-13	Simplified schematic of the instrumentation amplifier and SAFF. The output voltage of the SAFF, $V_{o2}$ is periodically reset to $V_{CM}$ using the clock signal $\phi_D$ . This <i>integrate-and-dump</i> or <i>charge-sampling</i> technique results in a sinc frequency response. . . . .	143
4-14	Schematic of fully differential, linear transconductor used in the filter.	146
4-15	Simplified schematic of transconductor used in the SAFF. The effective transconductance is $G_m = 2/R_s \times g_{m2}/g_{m1}$ , where $R_s$ is digitally tunable. The effective output resistance is a function of the second-stage transistors. Their value is made large through cascoding to prevent performance degradation. . . . .	147
4-16	Filter gain $G_\phi$ vs. $R_s$ setting $b_R$ for each $C_s$ setting $b_C$ . The top line corresponds to a value of $b_C = 0$ , and lower lines correspond to higher values of capacitance. . . . .	152

4-17	Filter gain $G_\phi$ for strategically selected values of $b_C$ and $b_R$ . As shown, the gain can be programmed in approximately 1 dB steps. Finer steps are possible at lower gain settings. . . . .	152
4-18	Block diagram of mixed-signal system including the IAMP+Filter+DAC ASIC, off-chip ADC, and FPGA for notch implementation. . . . .	156
4-19	Amplitude of the loop gain's frequency response for $G_f = 2^{-22}$ . As $f$ approaches $f_0$ , the amplitude approaches infinity. For large values of $ f  -  f_0 $ , however, the loop gain becomes very small. . . . .	161
4-20	Phase of the loop gain's frequency response for $G_f = 2^{-22}$ . . . . .	161
4-21	Amplitude of the closed-loop frequency response for $G_f = 2^{-22}$ . Clearly there is a very narrow notch at $f_0$ . The notch width can be easily programmed by changing $G_f$ . . . . .	163
4-22	Phase of the closed-loop frequency response for $G_f = 2^{-22}$ . . . . .	163
4-23	(A) Schematic of the combined instrumentation amplifier and charge redistribution DAC. The DAC is implemented with binary weighted capacitors and switches. A simplified diagram of the switches' implementation is shown in (B). (C) shows a simplified, single-ended model of the IAMP with with two inputs: one for the signal $v_i$ and another for an equivalent voltage DAC signal $v_{dac}$ . . . . .	168
4-24	Die photograph of IC with IAMP, filter, and feedback DAC. . . . .	170
4-25	Zoomed in version of die photograph. . . . .	170
4-26	Test board. . . . .	171
4-27	FPGA board connected to test board. . . . .	172
4-28	Simplified schematic of test board. . . . .	172
4-29	IAMP frequency response for various bias settings. . . . .	173
4-30	IAMP phase response for various bias settings. . . . .	174
4-31	Filter frequency response: $i_{SAAF} = 150$ nA, $b_C=31$ , $b_R=1$ , $i_{IAMP} = 1.6$ $\mu$ A, $G_\phi(0) = 1.4$ dB. . . . .	175
4-32	Filter frequency response plotted with a linear frequency axis: $i_{SAAF} = 150$ nA, $b_C=31$ , $b_R=1$ , $i_{IAMP} = 1.6$ $\mu$ A, $G_\phi(0) = 1.4$ dB. . . . .	176

4-33	Filter frequency response zoomed in near the notch frequency: $i_{SAAF} = 150$ nA, $b_C=31$ , $b_R=1$ , $i_{IAMP} = 1.6$ $\mu$ A, $G_\phi(0) = 1.4$ dB. . . . .	176
4-34	System frequency response (IAMP + filter) plotted with a linear frequency axis: $i_{SAAF} = 150$ nA, $b_C=31$ , $b_R=1$ , $i_{IAMP} = 400$ nA, $G_\phi(0) = 1.4$ dB. . . . .	177
4-35	System frequency response (IAMP + filter) plotted with a logarithm frequency axis: $i_{SAAF} = 150$ nA, $b_C=31$ , $b_R=1$ , $i_{IAMP} = 400$ nA, $G_\phi(0) = 1.4$ dB. . . . .	177
4-36	IAMP input-referred RMS noise voltage density. . . . .	178
4-37	Input-referred noise of the system (IAMP + filter) as measured from the output of the filter for four different IAMP bias settings and a constant filter bias current of $I_{SAAF} = 150$ nA. . . . .	180
4-38	Input-referred noise of IAMP alone, and IAMP plus filter for two IAMP bias settings for a constant filter bias setting of $I_{SAAF} = 150$ nA. . . . .	180
4-39	IAMP's output spectrum for an input signal of 20 mV <sub>p-p</sub> . The total harmonic distortion was 0.02% at 50 Hz and below 0.1% for frequencies between 1 Hz and 1 kHz. . . . .	182
4-40	Filter's output spectrum for an output signal of 620 mV <sub>rms</sub> and bias current of $i_{SAAF} = 150$ nA. $b_R = 0$ , $b_C = 10$ , $G_\phi = 1.75(4.8$ dB). The input signal was a 10.8 mV <sub>p-p</sub> sinusoid. The total harmonic distortion was below 0.1% for frequencies between 1 Hz and 1 kHz. . . . .	183
4-41	Modified IAMP gain when 1.2 pF capacitors are connected between $v_{im}$ and ground and between the DAC's output and $v_{ip}$ . This gain can be used to find the differential- and common-mode input impedances of the IAMP. . . . .	184
4-42	Extracted capacitance values that can be used to extract the the differential- and common-mode input impedances of the IAMP. . . . .	185
4-43	Filter gain versus $b_R$ setting for each $b_C$ setting. The higher lines represent low capacitance settings. . . . .	187

4-44	Filter gain versus $b_C$ setting for each $b_R$ setting. The higher lines represent low resistance settings. . . . .	187
4-45	Magnitude of the closed-loop frequency response of the system with the notch on (normalized to the passband gain to highlight the notch's attenuation). The gain of the digital block was varied to show the programmability of the notch width. The center frequency of the notch can be set arbitrarily. . . . .	188
4-46	Phase of the closed-loop frequency response of the system with the notch on. Smaller settings for $G_f$ result in a sharper notch and a narrower bandwidth of phase distortion. . . . .	188
4-47	Time-domain measurement of the output of the system when the notch is turned off (top) versus on (bottom). For this measurement, the notch frequency and the square wave fundamental frequency are both set to 50 Hz. The bias settings are: $I_{IAMP} = 400$ nA, $I_{SAAF} = 150$ nA, $V_{refD} = 10$ mV . . . . .	189
4-48	Input-referred RMS voltage spectral density when a 60 Hz, 1 mV signal is applied to the input of the system. The blue line shows that when the notch is turned on, the interference is eliminated, although other harmonics appear. The new harmonics are much smaller, however. The bias settings are $I_{IAMP} = 400$ nA, $I_{SAAF} = 150$ nA, $V_{refD} = 10$ mV	190
4-49	Input-referred RMS voltage spectral density when no input signal is applied. The blue line shows that any added broadband noise is negligible, although a small tone appears at 60 Hz. The bias settings are $I_{IAMP} = 400$ nA, $I_{SAAF} = 150$ nA, $V_{refD} = 10$ mV . . . . .	191
4-50	EKG measurement with the notch filter off. . . . .	192
4-51	EKG measurement with the notch filter on. . . . .	192
4-52	EKG measurement with the notch filter off. A $10\text{ M}\Omega$ resistor was placed between $v_{im}$ and ground for this measurement to create a mismatch between the common-mode input impedances and reduce the system's CMRR. This results in significantly more PLI corruption. . .	193

4-53	EKG measurement with the notch filter on showing how the notch filter is very effective at canceling PLI. . . . .	193
A-1	A simple representation of choppers (top), a symbol for a fully differential chopper (bottom left), and a switch-based implementation of a differential chopper, typically implemented with MOS switches (bottom right). Clocks signal $\phi_1$ and $\phi_2$ are often made non-overlapping, but can simply be inversions of each other. . . . .	202
A-2	Time-domain chopping waveforms. . . . .	202
A-3	A simple chopper-based IAMP. . . . .	203
A-4	A chopper-based IAMP that uses a linear transconductor instead of an OpAmp resulting in a higher input impedance [74]. A feedback path is used to attenuate the up-converted EDO. . . . .	205
A-5	A chopper-based instrumentation amplifier with an additional feedback path to cancel EDO [17]. Down-chopping is performed before the dominant pole to relax the bandwidth requirements and filter the up-converted DC offset of the OpAmp. . . . .	207
A-6	A chopper-based IAMP with the input choppers connected at the input of the OpAmp [65]. A second feedback path is used to reduce the DC offset caused by the parasitic switched-capacitor resistance created by the input choppers and $C_p$ . . . . .	209
A-7	Single-ended model of the chopper-stabilized IAMP in Fig. A-6 including noise sources and the parasitic switch-capacitor resistance $R_p$ . . .	209
A-8	(A) shows the block diagram of the system in Fig. A-7 including the choppers. (B) and (C) show how the position of the input chopper can be shifted without affecting the transfer function. (D) shows the simplified, equivalent block diagram with the choppers removed. . . .	210



A-9	Frequency response for the loop gain of the block diagram in Fig. A-8 for typical values that result in a stable system. These values are $f_{int} = 10$ mHz, $f_{eff} = 10$ mHz, $f_f = 159$ mHz, $f_0 = 1$ Hz, and $f_{pi} = 32$ Hz. . . . .	212
A-10	Closed-loop frequency response of the input-output transfer function for the block diagram in Fig. A-8 for typical values that result in a stable system. . . . .	213
A-11	Closed-loop frequency response of the resistor $R_p$ 's noise-to-output transfer function for the block diagram in Fig. A-8. . . . .	214
A-12	Closed-loop frequency response of the OpAmp's noise-to-output transfer function for the block diagram in Fig. A-8. . . . .	216



# List of Tables

- 1.1 Key specifications for the MICS band. . . . . 30
  
- 3.1 Summary of Measured Transceiver Performance . . . . . 103
  
- 4.1 Summary of Minimum Requirements for CMRR and IAMP Input  
Impedance . . . . . 121
- 4.2 Programmable  $R_s$  values . . . . . 150
- 4.3 Example Notch Filter Parameters . . . . . 162



# Chapter 1

## Introduction

Electronic medical implants have been around for over five decades and some enjoy widespread application. Pacemakers, perhaps the most famous and common, were first introduced in the 1960s and are currently used by more than 2 million Americans in a \$10 billion-a-year market that continues to grow quickly [6]. Other implants, such as deep brain stimulators, currently used by 190,000 patients, and cochlear implants, used by over 30,000 patients, continue to rise in popularity as they gain acceptance in the medical community and improve in functionality. Beyond these “common” medical implants, there is ongoing research and development in a wide array of systems, ranging from the artificial pancreas geared towards diabetes patients, to visual prosthetics that may return eyesight to the blind.

### 1.0.1 Trends in Medical Electronics

Although medical implants vary widely in their uses and architectures, they have some fundamental commonalities. First, all of these devices must be optimized for exceptional energy efficiency to improve battery life and prevent tissue damage due to heat. Many of these devices are life sustaining and must, therefore, have battery lives that are on the order of 10 years without the possibility of recharge. Other devices, such as cochlear implants, have a lower failure cost and can therefore be designed to use rechargeable batteries. However, they must still operate at low power levels

to prevent tissue damage, since increased power consumption implies greater heat dissipation if the circuits are not 100% efficient.

In addition to energy efficiency, the use of radio frequency (RF) telemetry to communicate with medical implants is becoming ubiquitous. RF telemetry allows doctors to reconfigure medical implants post-surgery, and it allows information extraction from implanted sensors. This is critical in applications such as deep-brain stimulation where, for example, the pulse frequency and amplitude used to stimulate different parts of the brain must be optimized to achieve the desired effect. Similar benefits are found in other implants such as pacemakers where sensors are used to record data on abnormal cardiac events that are later transferred wirelessly to a healthcare provider.

Circuit miniaturization is a third trend that is more specific to applications such as neural recording systems that require between 100–1000 sensors. Reducing the size of medical implants has always been a critical design goal, but the volume of the circuitry has hitherto been negligible compared with sensors, actuators, and batteries. Recent advances in micro-electro-mechanical systems (MEMS), however, have led to impressive reductions in sensor/actuator microelectrode arrays [45], [70]. Furthermore, improvements in energy efficiency and advances in battery technology are shrinking the volume of energy sources in implants. Combined, these two trends are placing pressure on the circuitry to shrink as well, and effectively mean that most of the electronics should be integrated on a single chip with minimal external components. This can be challenging for systems with many sensors, especially since instrumentation amplifiers often use large capacitors (nF –  $\mu$ F range) to AC couple to sensors or to implement ultra-low corner frequencies (<1 Hz).

### **1.0.2 Medical Implant Communication Service Band**

While electronic medical implants have been around for nearly fifty years, RF wireless communications with these devices was uncommon until recently. Inductive coupling was introduced in the 1970s as a means of communicating with medical implants and recharging implanted batteries. The capability of recharging an implanted battery is

critical for devices such as cochlear implants that consume relatively large amounts of power, while the ability to communicate with the device facilitates post-surgery reconfigurability and information extraction from implanted sensors. There are drawbacks to inductive coupling, however, including limited data rates ( 1–30kbps) and the requirement for physical contact and proper alignment between the programmer and the implant. To overcome some of these drawbacks, medical device manufacturers petitioned the Federal Communications Committee (FCC) in the mid-90s to allocate spectrum reserved for wireless RF telemetry with medical implants. In 1999, the FCC introduced the *medical implant communications service* (MICS) band, reserving the 402–405 MHz frequency range for the exclusive use of communication with and between medical implants.

RF telemetry in the MICS band was introduced to afford two main benefits: increased communication range of at least two meters, and increased data rates. Increasing the communication range between the implant and the base-station can have significant benefits for patients. For example, inductive coupling is a poor fit for applications that require continuous monitoring of an implanted sensor since the coil in the base-station must be well aligned with the coil in the implant. This effectively constrains the patient to a static location or requires that the base-station be physically attached on the skin. In contrast, if the base-station can be located anywhere within two meters of the implant, it can be carried around in the patient’s pocket or be placed near the patient while they sleep. This, in turn, opens the possibility of home monitoring since precise positioning of the base-station is unnecessary. In such a use case, the base-station can relay information from the implant to a healthcare provider through the internet, as shown in Fig. 1-1. This connection reduces the need for frequent doctor visits.

The second benefit of RF telemetry, increased data rates, can have various benefits. First, unlike inductive coupling, which is typically limited to  $\sim 50$  kbps, significantly higher data rates are possible in the MICS band since each channel is 300 kHz wide. Most biomedical signals vary slowly and, therefore, have small bandwidths and can be sampled at very low rates. Having the ability to transmit and receive information



Figure 1-1: Wireless link between patient and healthcare provider.

at high data rates, however, allows the implant to employ duty cycling. For example, if the rate at which data is being accumulated is 1 kbps and the data rate of the transmitter is 100 kbps, the implant can store 1 second worth of data and then transmit that data in a 10 ms burst. This means the transmitter can be turned off 99% of the time and its *average* power consumption is 100 times smaller than its power consumption when it is transmitting. This is beneficial since a major goal is to minimize the energy consumption to extend battery life, and energy is the time integral of power. If instead of transmitting at 100 kbps, the transmitter's data rate is 10 kbps and it consumes the same amount of power when transmitting, its average power consumption would be 10 times larger since its duty cycle would be 10% instead of 1%.

The assumption that a transmitter would consume the same amount of power at both data rates is reasonable for narrow-band RF transmitters using similar modulation schemes because the bulk of the power is consumed by RF components such as oscillators, mixers, and power amplifiers, whose power consumptions do not scale linearly with the transmitter's bandwidth. Instead, their power consumption is often set by other requirements such as linearity or noise, or they are limited by nonidealities in physical components such as the low  $Q$  of inductors. While this observation argues for using higher data rates, spectral mask requirements set limits on how fast data can be transmitted. Complex modulation schemes such as orthogonal frequency-division multiplexing (OFDM) or multi-level quadrature amplitude modulation enable high data rates for a given spectral mask requirement, but require complex hardware implementations or stringent requirements on RF blocks that lead to lower power



efficiency. For a narrow-band transmitter, therefore, there is a tradeoff between complex architectures that are more spectrally efficient but consume more power, and simpler topologies that consume less power but are less spectrally efficient [67]. With either approach, however, having the ability to use more bandwidth than inductive coupling should lead to better energy efficiency since the energy consumption does not, necessarily, increase linearly with utilized bandwidth.

In some applications, such as neural recorders that use hundreds of electrodes, the rate of data acquisition can be on the order of hundreds of kbps. These applications require higher data rates of data transmission and highlight the second benefit of using RF transmitters.

Table 1.1 summarizes the key requirements of the MICS band. In addition to these quantitative requirements, there are guidelines for operation to reduce the risk of interference between MICS transceivers [1]. For example, the transmitter in the implanted device should only transmit when asked to do so by the base-station, except in the case of a medical-implant event. Also, the base-station should operate in a *listen-before-speak* manner that first determines the channel with the lowest ambient noise and commands the implant to change to that channel before critical data exchanges occur. This last requirement necessitates frequency agility which dictates that the implant must have the ability to operate in any of the ten channels.

These basic requirements and guidelines outline how the MICS band should be used to achieve communication with implanted devices while reducing the probability of interference between MICS radios. They do not, however, go as far as creating a *standard* that would guarantee the inter-operability of devices from different manufacturers. For example, they do not require any particular modulation or encoding scheme. While there are commercial benefits to having standardized requirements, such as increased market penetration due to the benefits of inter-operability, there are also benefits to having greater design freedom. From a research point of view, the flexibility in MICS transceiver design opens the door to holistic optimization methods that should lead to ultra-low energy consumption.

Band of Operation	402–405 MHz
Channel Bandwidth	300 kHz
Maximum In-Channel Power Radiation	25 $\mu$ W (-16dBm) EIRP <sup>1</sup>
Maximum Out-of-Channel Power Radiation	20 dB below peak in-channel power
Frequency Stability	100 ppm ( $\sim$ 40 kHz)

Table 1.1: Key specifications for the MICS band.

### 1.0.3 Biomedical Sensor Issues

Electronic medical implants are generally used to either stimulate parts of the human body for therapy or to measure biomedical signals to perform diagnostics. More advanced systems use both sensors and actuators concurrently to create closed loop functionality that improve therapy [34]. A challenge that arises in signal detection is that some of these biomedical signals are very small. For example, neural field potentials (NFPs) can be as small as a few micro-volts. To accurately measure such small signals, the input-referred noise of a system must be on the order of 1  $\mu$ V, and it must be able to provide gain on the order of 40–80 dB. To achieve these goals, low-noise instrumentation amplifiers (IAMPs) and anti-aliasing filters are commonly used to interface between a sensor and an analog-to-digital converter (ADC) used to digitize the signals. The term *sensor interface* (SI) will be used throughout this thesis in reference to all circuitry necessary to interface between the sensor (typically a pair of electrodes) and the ADC.

There are several challenges that must be overcome to properly design such a sensor interface. First, the differential input impedance must be very large ( $\geq 10$  M $\Omega$ ) because the source impedance of the electrodes used to measure such signals can be very large. Second, the effects of flicker noise must be mitigated since many physiological signals occupy frequency bands below 100 Hz. Third, the SI must be robust against large DC offsets in the electrodes caused by charge accumulation at the skin-metal interface [40]. These offsets can be as large as hundreds of millivolts and will saturate the system if proper techniques are not employed. Fourth, the SI must be robust against power-line interference (PLI) at 50 or 60 Hz. The common-mode

amplitude of PLI can be on the order of tens of volts, and the differential-mode signal can be as large as a few millivolts even with an ideal IAMP. For many biomedical applications, PLI falls into the frequency bands of interest and its differential mode component can be significantly larger than the signal of interest, necessitating a wider dynamic range.

In addition to these performance requirements, the overall sensor interface must be very energy efficient, and for some applications, must consume minimal area and use few or no discrete components. This last requirement is particularly challenging because some of the techniques used to mitigate the effects of electrode DC offset or to notch out PLI use large discrete capacitors ( $\approx 1\mu F$ ) to obtain very low corner-frequencies ( $\ll 1$  Hz).

#### **1.0.4 Research Goals and Thesis Organization**

There were two fundamental goals to this research: the development of an energy efficient transceiver for medical implant communications, and the development of a small and ultra-low power sensor interface for biomedical applications. Although particular emphasis was placed on implantable devices, many of the techniques are applicable to non-invasive diagnostic systems. En route to developing the MICS transceiver, we developed a frequency-domain method for analyzing super-regenerative amplifiers (SRAs) and receivers (SRRs). We found that this method can be a powerful tool for predicting the sensitivity and selectivity of SRRs, and for selecting parameters for optimal performance. We also introduced a novel way of detecting on-off keying (OOK) with SRRs using time-domain measurements instead of amplitude measurements, and showed the benefits of this technique in relaxing linearity requirements of the system. Chapter 2 describes these methods and techniques.

Chapter 3 focuses on the development of a prototype MICS transceiver developed in 90-nm CMOS. The system includes a direct modulation MSK transmitter and super-regenerative OOK receiver, both using a single digitally-controlled oscillator with fine frequency resolution. We also introduce a frequency correction loop method that eliminates the need for a phase-locked loop.

Chapter 4 describes many of the issues surrounding biomedical sensor interface systems and proposes a novel variable-gain, anti-aliasing filter that is area- and power-efficient. It also introduces a novel mixed-signal notch that can be used to cancel PLI or other forms of interference, relaxing the dynamic range requirements of other blocks in the system. The notch filter's center frequency and bandwidth are digitally programmable.

Chapter 5 summarizes the contributions of this work and proposes new research projects to further the state of the art.

## Chapter 2

# Frequency-Domain Analysis of Super-Regenerative Amplifiers

In the early stages of our research, a survey of ultra-low power receiver architectures showed that super-regenerative receivers (SRRs) were among the most energy efficient and sensitive, particularly among narrow-band systems. When it came time to conduct sensitivity and selectivity analysis, however, we found that very few papers tackled the problem of noise analysis in super-regenerative amplifiers (SRAs). We also found that much of the theoretical groundwork laid in the 1940s and 50s restricted its analysis to AM demodulation through oversampling of the received signal's envelope. Since our goal was to use the SRR to synchronously detect OOK signals, we undertook the task of developing a mathematical model to accurately determine the sensitivity and selectivity of an SRA. This section describes the development of that model and how it can be used to analyze the response of SRAs to myriad input signals.

### 2.1 Introduction

The super-regenerative amplifier/receiver was first introduced by Edwin Armstrong in 1922 [5] as an exciting development in communications systems. Although it never gained the popularity of the superheterodyne receiver, recent growth in low-power,

short-range wireless links has reawakened an interest in SRAs due to their excellent sensitivity for small amounts of DC power consumption [12, 21, 43, 44, 48, 66]. The general concepts behind the operation of SRAs are intuitive, but the theory necessary for quantitative analysis tends to be mathematically tedious due to its nonlinear, time-varying nature. Thorough time-domain solutions have been available since the 1950s [71], and more recent work has focused on generalizing those results to generic SRAs [44]. Other recent work has shown the capacity to operate SRAs synchronously, improving the selectivity and data rate of SRA receivers [43], and the benefits of pulse-shaping OOK signals to optimize the sensitivity of an SRA receiver [42, 50]. Building on these works, we propose a convolution model that allows for *frequency domain* analysis of SRAs. We then show that frequency domain methods allow for straightforward analysis of arbitrary deterministic and stochastic input signals, and use various examples that lead to complete sensitivity equations.

As explained in [71], the SRA can be operated in four general modes combining the choices of slope-controlled versus step-controlled and linear versus logarithmic modes. The first choice describes the type of *quench signal* or *damping function* used, and the second choice describes whether the SRA output is limited to small values that prevent nonlinearities, or if its amplitude is permitted to grow to the point of compression. The subtleties of the slope- versus step-control modes will be explained in Section 2.2, but we note that the analysis presented here is restricted to the slope-controlled mode. This method has been of greatest interest in recent literature because it offers benefits in both sensitivity and selectivity. Also, we focus our analysis mostly on the linear mode of operation. However, we introduce the concept of a *trigger-time* in Section 2.4.3, which is relevant to both the linear and the logarithmic modes of operation.

This chapter is divided into five main sections. Section 2.2 briefly describes the general theory of the SRA and recounts the time-domain solution for its differential equation. Section 2.3 presents a convolution model of the SRA and uses it to find an SRA's response to various deterministic signals. Section 2.4 shows how the convolution model can be used to find the SRA's response to additive white Gaussian

noise (AWGN) and uses the results to calculate the expected bit error rate (BER) and sensitivity of an OOK receiver. It also introduces the concept of a trigger-time which can be used to accurately set the optimum threshold and detect signals in an OOK receiver while avoiding the problems usually caused by nonlinearity. Section 2.5 describes a test circuit used to verify the theory presented and compares measured results to those predicted using the convolution model. Section 2.6 summarizes the key concepts and concludes the paper.

## 2.2 General SRA Theory

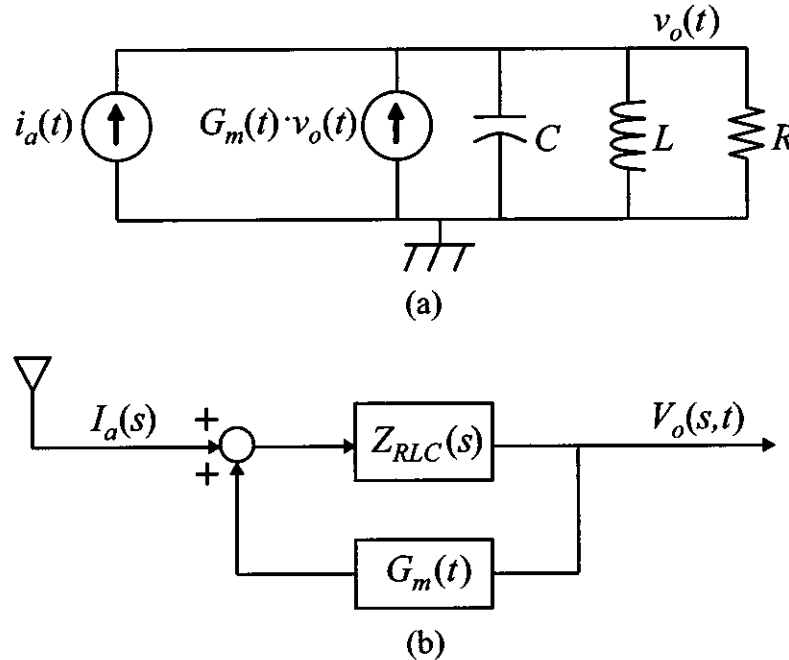


Figure 2-1: (a) SRA circuit model, and (b) SRA feedback loop model.

### 2.2.1 Circuit Model and Block Diagram for SRA

Fig. 2-1 shows the simplified (a) circuit model and (b) feedback model for an RLC-based SRA. The parameters of interest for the resonant RLC tank are:  $\omega_0$ , the resonant frequency,  $Z_0$ , the characteristic impedance,  $Q_0$ , the quality factor, and  $\zeta_0$ , the quiescent damping factor. The relationships between these parameters and the circuit

components are

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad (2.1)$$

$$Z_0 = \sqrt{\frac{L}{C}} = \omega_0 L = \frac{1}{\omega_0 C} \quad (2.2)$$

$$\zeta_0 = \frac{1}{2RC\omega_0} = \frac{1}{2Q_0} = \frac{1}{2} \frac{Z_0}{R}. \quad (2.3)$$

Using these parameters, the impedance for a parallel resonant tank can be written as

$$Z_{RLC}(s) = \frac{Z_0 \omega_0 s}{s^2 + 2\zeta_0 \omega_0 s + \omega_0^2}. \quad (2.4)$$

If  $G_m(t)$  varies slowly enough with respect to  $\omega_0$ , such that the system in Fig. 2-1 is quasi-static, we can define a time-varying transfer function for the feedback loop shown in Fig. 2-1(b) by

$$Z_{TV}(s, t) = \frac{V_o(s, t)}{I_a(s)} = \frac{Z_{RLC}(s)}{1 - G_m(t) \cdot Z_{RLC}(s)} \quad (2.5)$$

which can be rewritten as

$$Z_{TV}(s, t) = \frac{Z_0 \omega_0 s}{s^2 + 2\zeta(t) \omega_0 s + \omega_0^2} \quad (2.6)$$

where  $\zeta(t)$  is the *instantaneous damping factor* or *damping function* and is defined as

$$\zeta(t) = \zeta_0(1 - G_m(t)R). \quad (2.7)$$

Note that (2.6) only differs from (2.4) in the denominator where  $\zeta_0$  is replaced by  $\zeta(t)$ . This is because the positive feedback from transconductance  $G_m(t)$  can be modeled as a negative resistance that only affects the damping factor of the second-order system in Fig. 2-1.



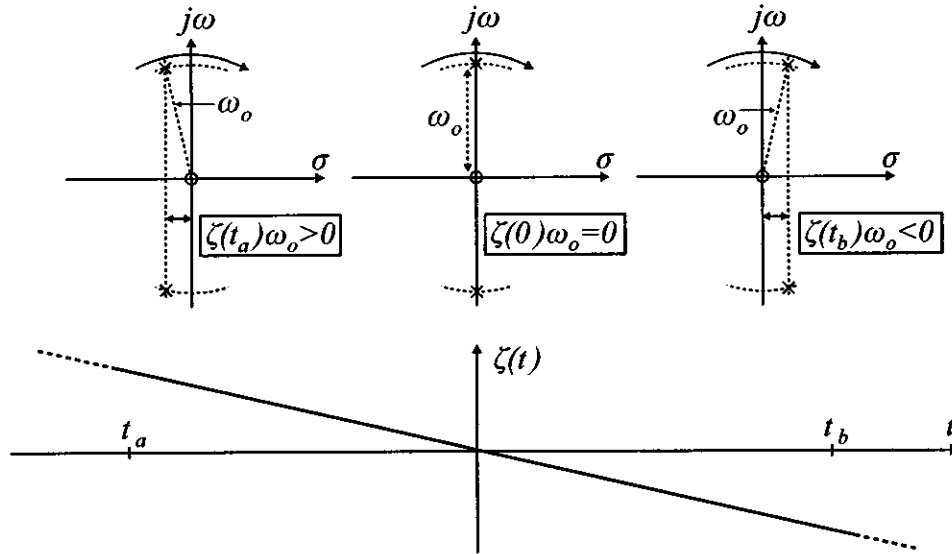


Figure 2-2: Time-varying pole/zero locations for an SRA as the damping function changes.

### 2.2.2 SRAs as Time-Varying, Second-Order Systems

An SRA differs from a linear time-invariant (LTI) system in that its poles are periodically shifted between the left-hand side and right-hand side of the complex plane by varying the damping function  $\zeta(t)$ . The exact function used to define  $\zeta(t)$  determines the characteristics of the SRA's response to an input signal, and various functions can and have been used [21], [66]. As an example, Fig. 2-2 shows the instantaneous value of the poles of an SRA as  $\zeta(t)$  is linearly varied during one cycle using a ramp function. As will be shown, for each period, the resulting time-varying system yields a filtered and amplified sample of its input signal's envelope. Unlike LTI systems whose filtering qualities are strictly dependent on the static location of their poles and zeros, the filtering qualities of SRAs additionally depend on the characteristics of the damping function used to vary their pole locations. Furthermore, in contrast with LTI systems, SRAs exploit the instability portion of their cycle to achieve very high gain despite using active components that provide relatively small gain.

### 2.2.3 General SRA Solution for Linear Mode Operation

Historically, SRAs have been used in either the *linear* or *logarithmic* mode [44, 71]. In the linear mode, the SRA is configured such that its *output* remains small enough throughout each quench cycle to prevent significant nonlinearities. As a result, the envelope of the SRA's output is proportional to the amplitude of the input signal. In the logarithmic mode, the SRA is configured such that its output saturates during each cycle. The integral of its envelope is then proportional to the logarithm of the input signal's amplitude. The following analysis is valid for the linear mode, but later discussions will show how the results can be used to accurately model an SRA that is allowed to enter compression.

Using (2.6) we can write the following differential equation to describe the LTV model of the SRA in Fig. 2-1

$$v_o''(t) + 2\zeta(t)\omega_0 v_o'(t) + \omega_0^2 v_o(t) = 2R\zeta_0\omega_0 i_a'(t). \quad (2.8)$$

A more general version of (2.8) and its thorough solution can be found in [44], where the general solution is broken down into the sum of the free response and the forced response. The following analysis assumes that the free response is zero. In a practical implementation, this means that any oscillation from a previous cycle is *quenched* before time  $t_a$  when a new cycle begins. This simplifies the mathematics, but more importantly, it improves the performance of the receiver since it ensures that each cycle is independent of all previous cycles. Mathematically, the quenching is done by allowing the poles to remain in the left-hand side long enough for the envelope of the output voltage to decay below the noise levels. To expedite the decay, the poles can be pushed far to the left, or equivalently, the damping function can be made a large positive value. Practically, this can be done by briefly shorting the tank as in [10], by reducing the transconductance,  $G_m(t)$ , to a very low value as in [12], or by making  $G_m(t)$  negative.

The general solution found in [44] can be modified for the specific case of an

RLC-based SRA with an input current  $i_a(t)$ , resulting in the output voltage

$$v_o(t) = Z_0 e^{-\omega_0 \int_0^t \zeta(\lambda) d\lambda} \times \int_{t_a}^t i_a'(\tau) e^{\omega_0 \int_0^\tau \zeta(\lambda) d\lambda} \sin[\omega_0(t - \tau)] d\tau. \quad (2.9)$$

In (2.9),  $\zeta(t)$  is defined such that it is positive for  $t_a \leq t < 0$  and negative for  $0 \leq t \leq t_b$  as in Fig 2-2. The solution can be broken down into the time-dependent gain  $\mu(t)$  and the filtering term  $k(t)$

$$v_o(t) = Z_0 \mu(t) k(t), \quad (2.10)$$

where

$$\mu(t) = e^{-\omega_0 \int_0^t \zeta(\lambda) d\lambda}, \quad (2.11)$$

$$k(t) = \int_{t_a}^t i_a'(\tau) g(\tau) \sin[\omega_0(t - \tau)] d\tau, \quad (2.12)$$

and

$$g(t) = e^{\omega_0 \int_0^t \zeta(\lambda) d\lambda}. \quad (2.13)$$

The gain component  $\mu(t)$  reaches its peak at  $t = t_b$  and its maximum value  $\mu(t_b)$  is referred to as the super-regenerative gain [71]. The term  $g(t)$  is referred to as the sensitivity function and has a peak value of unity at  $t = 0$ . For common damping functions it decays rapidly outside of a time window concentrated about  $t = 0$ , limiting the effect of the input signal  $i_a(t)$  outside of that window. This quality will be exploited in Section 2.3 to approximate  $k(t)$  by a convolution and perform frequency domain analysis of the SRA. For *slope-controlled* SRAs,  $\zeta(t)$  changes slowly enough that multiple periods of the input signal occur during the sensitivity period [71]. If  $\zeta(t)$  changes from positive to negative abruptly, the SRA is said to be operating in the *step-controlled* region and has a significantly different frequency response. The subsequent analysis assumes the SRA is operated in the slope-controlled mode, which is preferable as it achieves better sensitivity and selectivity [71].

## 2.2.4 SRA Solution for a Ramp and Sine Damping Functions

Almost any arbitrary shape can be used as the damping function  $\zeta(t)$  as long as it is positive for  $t_a \leq t < 0$  and negative for  $0 < t \leq t_b$ . Two common waveforms in slope-controlled SRAs are the ramp (or sawtooth) and the sine-wave [10,21,42–44,66]. The ramp damping function proves particularly useful for analysis since it leads to Gaussian equations that have closed-form solutions. Furthermore, it achieves higher gain and has a frequency response preferable to that of sine-wave damping as discussed in this section and Section 2.3.2.

For the time span  $t_a \leq t < t_b$ , the ramp damping function has the form

$$\zeta(t) = -\beta t \quad (2.14)$$

where  $\beta$  is its slope and has units of Hz. Substituting (2.14) into (2.11) and (2.13) results in

$$\mu_{ramp}(t) = e^{\frac{1}{2}\omega_0\beta t^2} = e^{\frac{t^2}{2\sigma_s^2}}, \quad (2.15)$$

and

$$g_{ramp}(t) = e^{-\frac{1}{2}\omega_0\beta t^2} = e^{-\frac{t^2}{2\sigma_s^2}}, \quad (2.16)$$

where

$$\sigma_s = \frac{1}{\sqrt{\omega_0\beta}} \quad (2.17)$$

has units of s/ $\sqrt{\text{rad}}$  and is defined as the *SRA time constant*.

The sine-wave damping function has the form

$$\zeta(t) = -\frac{\beta}{\omega_q} \sin(\omega_q t) \quad (2.18)$$

where

$$\omega_q = \frac{2\pi}{t_b - t_a}. \quad (2.19)$$

It is defined such that its slope is  $-\beta$  at  $t = 0$ . Substituting (2.18) into (2.11) and

(2.13) results in

$$\mu_{sin}(t) = \exp\left(\frac{1 - \cos(\omega_q t)}{\omega_q^2 \sigma_s^2}\right), \quad (2.20)$$

and

$$g_{sin}(t) = \exp\left(-\frac{1 - \cos(\omega_q t)}{\omega_q^2 \sigma_s^2}\right). \quad (2.21)$$

To facilitate the comparison between these two damping functions, we set  $t_b = |t_a| = \frac{1}{2}T_q$ , where  $T_q$  is the quench period, and define the ratio

$$\gamma = \frac{t_b}{\sigma_s} = \frac{T_q}{2\sigma_s}. \quad (2.22)$$

This allows us to make the substitution  $\omega_q = \pi/(\gamma\sigma_s)$ . Using (2.22) and solving (2.15) and (2.20) at  $t = t_b$  allows us to evaluate the super-regenerative gain for both damping functions as

$$\mu_{ramp}(t_b) = e^{\frac{1}{2}\gamma^2}, \quad (2.23)$$

and

$$\mu_{sin}(t_b) = e^{\frac{2}{\pi^2}\gamma^2}. \quad (2.24)$$

The ratio between the two gains is

$$\frac{\mu_{ramp}(t_b)}{\mu_{sin}(t_b)} = e^{(\frac{1}{2} - \frac{2}{\pi^2})\gamma^2} \approx e^{0.3\gamma^2}. \quad (2.25)$$

For a ratio of  $\gamma = 3$ , the gain of the SRA using a ramp damping function is  $14.5\times$  greater (23dB) compared with the sine-wave damping function. Of course the amplitude of the sine-wave damping function could be increased to increase its gain, but this would widen the bandwidth of the SRA, increasing noise and degrading selectivity. This will be shown in Section 2.3.2 along with the effects of  $\gamma$  on the frequency response of the SRA. Fig. 2-3 shows  $g_{ramp}(t)$  and  $g_{sin}(t)$  for  $\gamma = 3$ . For time values near zero, the two sensitivity functions are similar. However,  $g_{ramp}$  approaches zero more quickly and is reduced to 0.01 at  $t_b = 3\sigma_s$  while  $g_{sin}$  is only reduced to 0.161. Later sections show this gives the ramp damping function a superior frequency response. Generally speaking, the SRA's gain grows exponentially with  $\gamma$  and the fre-

quency response improves. For a given value of  $\sigma_s$ , this requires longer quench cycles and, therefore, lower bit rates creating a tradeoff. The tradeoff favors increasing  $\gamma$ , however, since the gain grows as  $e^{\gamma^2}$  whereas the bit rate is reduced linearly.

## 2.3 Convolution Model of SRA Solution

In this section, we show that (2.12) closely resembles a convolution and exploit this quality to perform frequency domain analysis on the SRA. Typically, we are interested in the value of the envelope of  $v_o(t)$  near the end of the cycle ( $t \approx t_b$ ), since that is when the maximum super-regenerative gain is achieved. Since the output is oscillatory, we are not interested in its value exactly at  $t_b$ , but rather at some time near  $t_b$  when the sinusoidal term is at its peak. In that time range, (2.12) can be rewritten as the nearly exact approximation

$$k(t) \approx \int_{-\infty}^{\infty} x(\tau) \Pi\left(\frac{\tau}{t_b - t_a}\right) \sin(\omega_0[t - \tau]) d\tau, \quad (2.26)$$

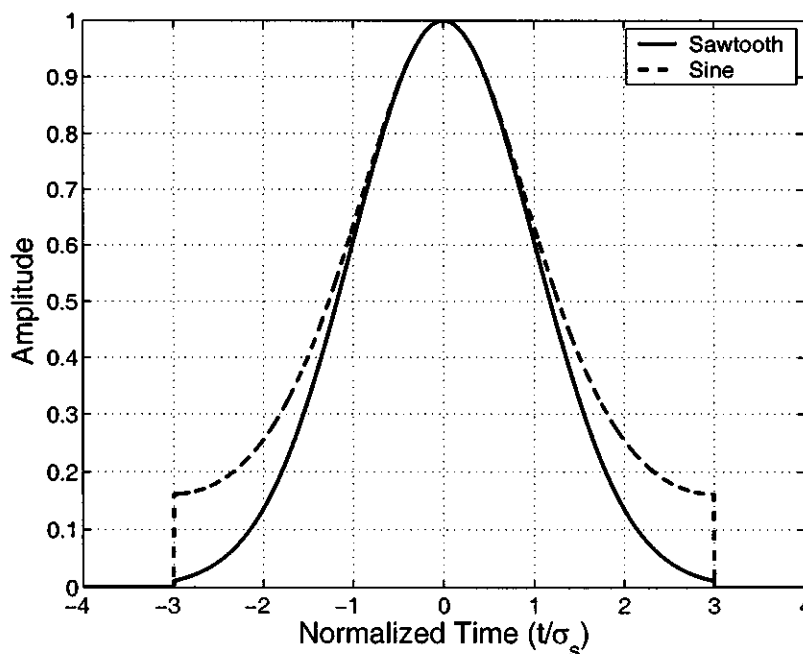


Figure 2-3: Sensitivity functions for sawtooth/ramp and sine damping functions ( $\gamma = 3$ ).

where

$$x(t) = i'_a(t)g(t). \quad (2.27)$$

and

$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < 1/2 \\ 0 & \text{otherwise.} \end{cases} \quad (2.28)$$

The approximation in (2.26) assumes  $t_b \approx -t_a$ , although this assumption is not necessary and can be avoided at the expense of increased complexity by modifying the argument of  $\Pi(t)$ . Equation (2.26) can be rewritten as the convolution

$$k_*(t) = x(t)\Pi\left(\frac{t}{t_b - t_a}\right) * \sin(\omega_0 t) \quad (2.29)$$

which is valid for time  $t \approx t_b$ . Note that if the value of  $k(t)$  is desired near some time other than  $t_b$ , that time instant can be substituted in the argument of  $\Pi(t)$ .

As discussed in Section 2.2.4,  $g(t)$  always has a maximum value of unity at  $t = 0$ , and typically drops sharply for  $|t| > 3\sigma_s$ . Fig. 2-4 illustrates the effects of this property on (2.12) when  $i_a(t)$  is the sinusoid

$$i_a(t) = I_a \sin(\omega_a t + \phi_a). \quad (2.30)$$

As shown in Fig. 2-4(a,b),  $g(t)$  grows to a maximum value of unity as the damping function approaches zero. Fig. 2-4(c) illustrates the time derivative of  $i_a(t)$ , and (d) shows  $x(t)$  which has the form of a time-windowed version  $i'_a(t)$ . As shown in Fig. 2-4(e),  $k(t)$  is oscillatory and grows for  $t < 3\sigma_s$ , but then flattens out. This occurs because  $x(t)$  becomes very small for values of  $t > 3\sigma_s$ . As mentioned previously,  $k_*(t)$  is only valid (and nearly exact) for  $t \approx t_b$ . However, Fig. 2-5 shows that it is generally a very good approximation for values of  $t \geq 3\sigma_s$ . In fact, if  $\gamma \geq 3$ , (2.29) can be simplified further by removing the  $\Pi(t)$  term without much loss to accuracy since  $x(t)$  is very nearly zero for  $|t| > 3\sigma_s$ :

$$k_*(t) \approx x(t) * \sin(\omega_0 t). \quad (2.31)$$

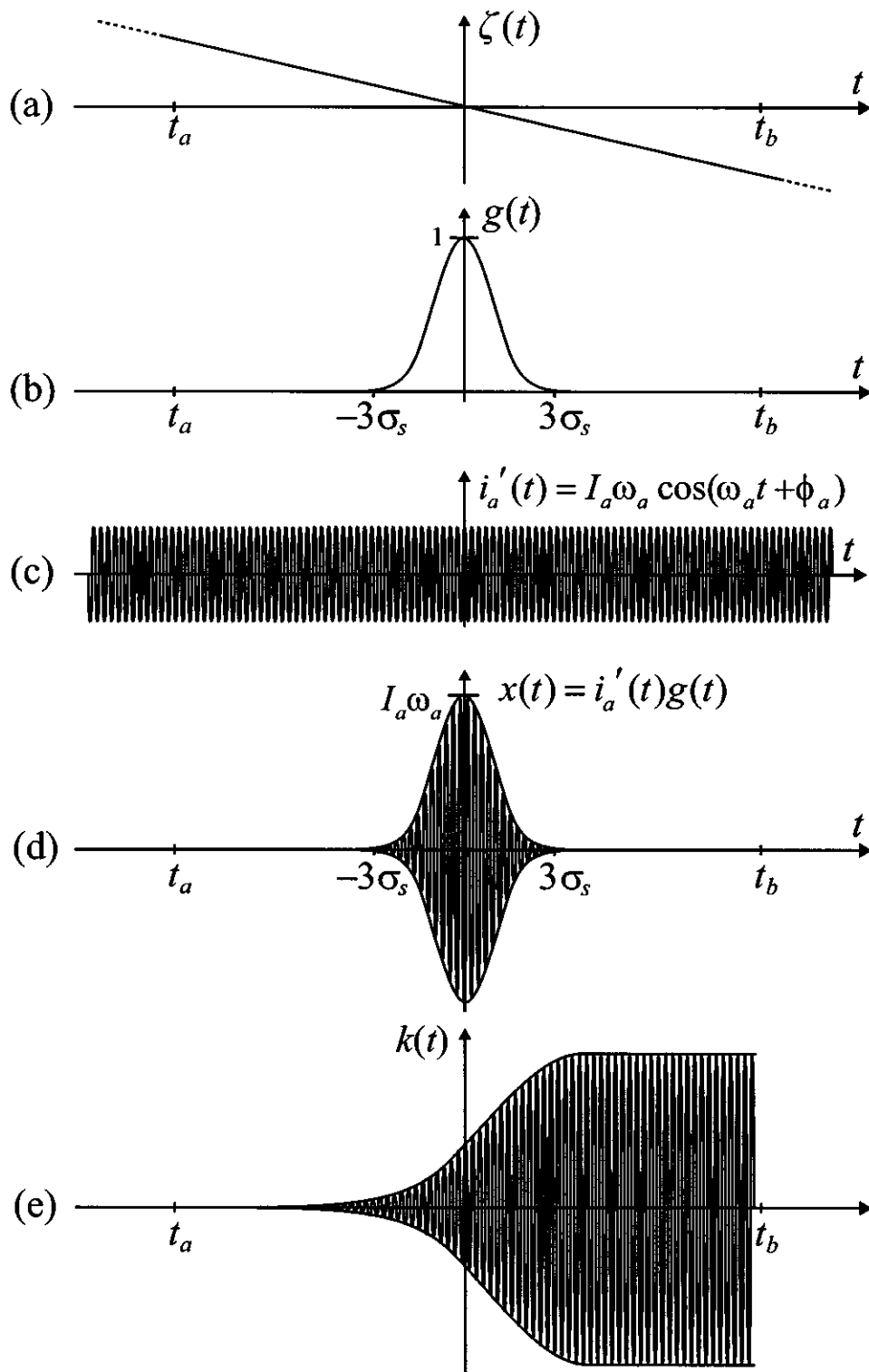


Figure 2-4: (a) Damping function, (b) sensitivity function, (c) time derivative of sinusoidal input, (d) windowed input, and (e)  $k(t)$ .



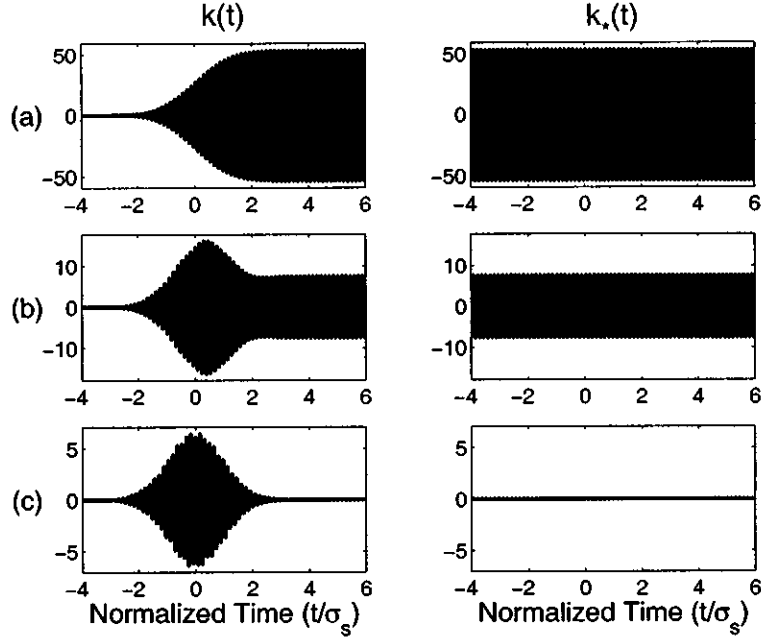


Figure 2-5: Numerical simulations comparing the exact equation  $k(t)$  with the convolution approximation  $k_*(t)$  for  $\gamma = 3$  and (a)  $\omega_a = \omega_0$ , (b)  $\omega_a = \omega_0 + 2\Omega_s$ , and (c)  $\omega_a = \omega_0 + 4\Omega_s$ .

This approximation has the same effect as setting  $\gamma = \infty$  (for  $k(t)$ , but not  $\mu(t)$ ) and results in an optimistic estimate of the frequency response. Nonetheless, we will use this approximation for the subsequent hand calculations and then show numerical examples that clarify its effects.

### 2.3.1 SRA Response to an Arbitrary Input

The key benefit to the convolution model is that it enables the use of frequency domain techniques that facilitate the analysis of an SRA's response to arbitrary input signals. Taking the Fourier transform of (2.31) yields the frequency domain signal

$$K_*(\omega) = j\pi (X(\omega_0)\delta(\omega - \omega_0) - X(-\omega_0)\delta(\omega + \omega_0)). \quad (2.32)$$

The time convolution of  $x(t)$  with a sinusoid yields the result that only the values of  $X(\omega)$  at  $\omega = \pm\omega_0$  are important. Taking the inverse Fourier transform of (2.32) gives

$k(t)$  for an arbitrary input signal  $i_a(t)$

$$k_*(t) = |X(\omega_0)| \sin(\omega_0 t + \angle X(\omega_0)). \quad (2.33)$$

Since  $x(t)$  is the time-domain product of  $i'_a(t)$  and  $g(t)$ , its Fourier transform is the convolution

$$X(\omega) = \frac{1}{2\pi} j\omega I_a(\omega) * G(\omega). \quad (2.34)$$

Substituting (2.33) into (2.10) yields the solution for the output voltage of the SRA which is usually accurate for  $t > 3\sigma_s$

$$\boxed{v_o(t) = Z_0 \mu(t) |X(\omega_0)| \sin(\omega_0 t + \angle X(\omega_0))}. \quad (2.35)$$

A more accurate solution can be used for numerical analysis that incorporates the  $\Pi(t)$  term in (2.29) by using the function

$$X_{\Pi}(\omega) = X(\omega) * (t_b - t_a) \text{sinc} \left( \frac{1}{2\omega(t_b - t_a)} \right) \quad (2.36)$$

in place of  $X(\omega)$  in (2.34).

Equations (2.35) and (2.34) define the general form of an SRA's output. For an input current  $i_a(t)$ , the output voltage is proportional to the resonator's characteristic impedance  $Z_0$ , its envelope grows exponentially, its filtering characteristics are determined by the sensitivity function  $g(t)$ , and it oscillates at its resonant frequency.<sup>1</sup> Typically, SRAs are used in amplitude modulation systems where the phase of the input signal carries no information. As a result, the sinusoidal term in (2.35) is often removed by connecting the output of the SRA to an envelope detector whose output is

$$v_e(t) = Z_0 \mu(t) |X(\omega_0)|. \quad (2.37)$$

---

<sup>1</sup>Since this is the result of linear analysis, the injection locking phenomenon is not modeled. For a large input signal whose frequency is close to  $\omega_0$ , the oscillation frequency may become the same as the input signal's as the output of the SRA grows and nonlinearities take effect. However, since it is the envelope of the output that is of interest, this does not affect the practical operation of the SRA.

### 2.3.2 SRA Response to a Sinusoidal Input

The filtering qualities of the SRA become more clear by analyzing its response to a sinusoidal input. For the input signal (2.30), (2.34) becomes

$$X(\omega) = \frac{\omega_a I_a}{2} \left( G(\omega + \omega_a) e^{-j\phi_a} + G(\omega - \omega_a) e^{j\phi_a} \right). \quad (2.38)$$

It is illustrative to use a specific example of a sensitivity function to appreciate the qualities of  $G(\omega)$ . For the ramp damping function described in Section 2.2.4,  $g(t)$  is Gaussian, such that its Fourier transform is also Gaussian:

$$G(\omega) = \frac{\sqrt{2\pi}}{\Omega_s} e^{-\frac{\omega^2}{2\Omega_s^2}}, \quad (2.39)$$

where

$$\Omega_s = \frac{1}{\sigma_s} = \sqrt{\omega_0 \beta} \quad (2.40)$$

is the *SRA frequency constant*. As a result, when  $X(\omega)$  is evaluated at  $\omega_0$ , the term  $G(\omega_0 + \omega_a)$  is practically zero and (2.38), evaluated at  $\omega_0$ , simplifies to

$$X(\omega_0) = \frac{\omega_a I_a}{2} G(\omega_0 - \omega_a) e^{j\phi_a}. \quad (2.41)$$

Substituting (2.41) into (2.35) yields the SRA's output for a sinusoidal input

$$\boxed{v_o(t) = \frac{Z_0 \omega_a I_a}{2} \mu(t) |G(\omega_0 - \omega_a)| \sin(\omega_0 t + \phi_a)}. \quad (2.42)$$

For the specific case of the ramp damping function (2.14), the SRA's output is

$$v_o(t) = \frac{I_a Z_0 \omega_a \sqrt{2\pi}}{2 \Omega_s} \cdot e^{-\frac{(\omega_0 - \omega_a)^2}{2\Omega_s^2}} \cdot e^{\frac{t^2}{2\sigma_s^2}} \sin(\omega_0 t + \phi_a). \quad (2.43)$$

Fig. 2-6 shows a graphical representation of the mathematics used to find (2.42) and (2.43). Since the input signal is sinusoidal, its time derivative,  $i'_a(t)$ , effectively “up-converts”  $g(t)$  and translates its spectrum to  $\pm\omega_a$ . The resulting bandpass signal  $x(t)$  is then convolved with  $\sin(\omega_0 t)$  which is equivalent to multiplying its spectrum

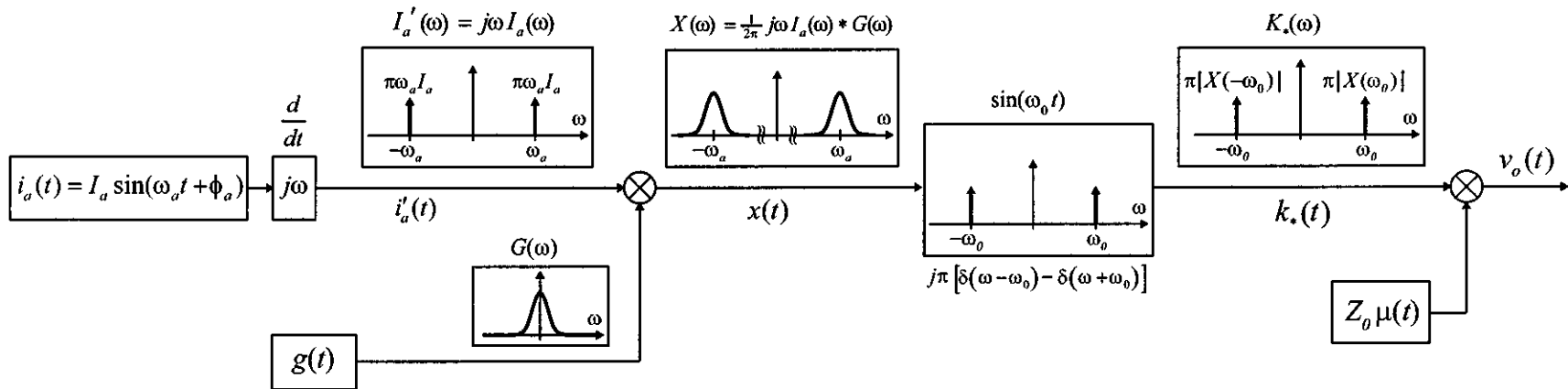


Figure 2-6: Graphical representation of SRA response to a sinusoidal input signal.

$X(\omega)$  with two Dirac delta functions. The result is a sinusoidal term  $k_*(t)$  whose amplitude depends on the magnitude of  $X(\omega)$  evaluated at  $\omega = \pm\omega_0$ . Finally, the sinusoidal term  $k_*(t)$  is multiplied by the SRA's time-dependent gain resulting in  $v_o(t)$ .

Solution (2.42) is valid for general damping functions while (2.43) describes the response of an SRA with a ramp damping function. There are five important observations that can be made from (2.43) that are generally true for other damping functions. First, the output voltage is linearly proportional to the amplitude of the input current and the characteristic impedance of the resonant tank. Second, there is a constant gain term that depends on the damping function and the input signal's frequency. Third, the SRA filters the input current with a Gaussian-shaped filter centered about the tank's resonant frequency,  $\omega_0$ , with a bandwidth defined by the frequency constant,  $\Omega_s$ . Fourth, the output voltage has an envelope that grows very rapidly (an exponential with a squared time exponent) and is dependent on  $\sigma_s$ . And fifth, the output voltage is oscillatory with a frequency equal to the tank's resonant frequency.

The constant  $\Omega_s$  and its inverse,  $\sigma_s$ , practically define the response of the system in terms of both frequency selectivity and gain. Since  $\omega_0$  is set by the application requirements, (2.43) shows that the only two design variables are  $\beta$ , the slope of the damping function, and  $Z_0$ . If the slope of the damping function is reduced, selectivity improves but gain is reduced. To achieve the same gain, each cycle must be longer (since the gain increases with time) resulting in a bit rate reduction. However, the tradeoff between gain, bandwidth, and bit rate are not linear as with most systems since the bit rate and bandwidth are inversely proportional to  $\sigma_s$ , whereas gain grows as  $e^{t^2/\sigma_s^2}$ . This means that bit rate and bandwidth are traded for the square root of the log of gain.

Fig. 2-7 illustrates the frequency response of the SRA for various values of  $\gamma$ . The  $\gamma = \infty$  waveform represents the result of the approximation used in (2.31) where the limits of integration are extended to  $\pm\infty$ . The other lines show that the primary effect of reducing  $\gamma$  is a degradation in the frequency response of the SRA manifested by

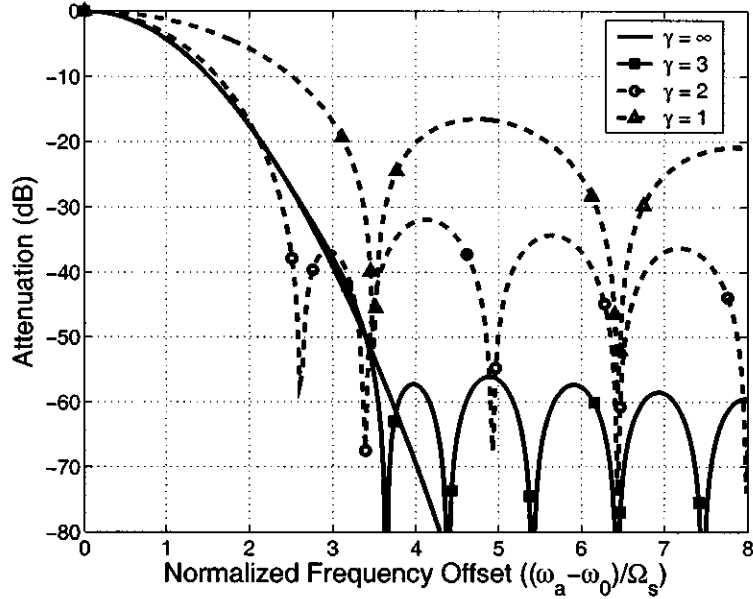


Figure 2-7: Frequency response using a sawtooth/ramp damping function and varying values of  $\gamma$ .

rising sidelobes. However, for  $\gamma \geq 3$  the SRA provides more than 55dB of attenuation in the stopband, confirming that (2.31) is a good approximation. Fig. 2-8 compares the frequency response of the ramp and sine-wave damping functions discussed in Section 2.2.4 and confirms that the ramp damping function is preferable since it achieves higher gain and lower sidebands for a given value of  $\gamma$ .

The fact that a system with a single active device with very small intrinsic gain is able to filter and amplify with almost limitless gain speaks to the SRA's compelling potential in low-power applications. The analysis above also shows a peculiar quality of the super-regenerative receiver: its filter bandwidth is not strictly a function of resistance, capacitance, and inductance, but also of the characteristics of its damping function. This means that high selectivity can be achieved despite limitations in the  $Q$  of an LC resonator.

### 2.3.3 SRA Response to Multiple Sinusoids

Analyzing the SRA's response to multiple sinusoidal inputs is simplified by the convolution model since superposition holds. For an input signal comprising the sum of

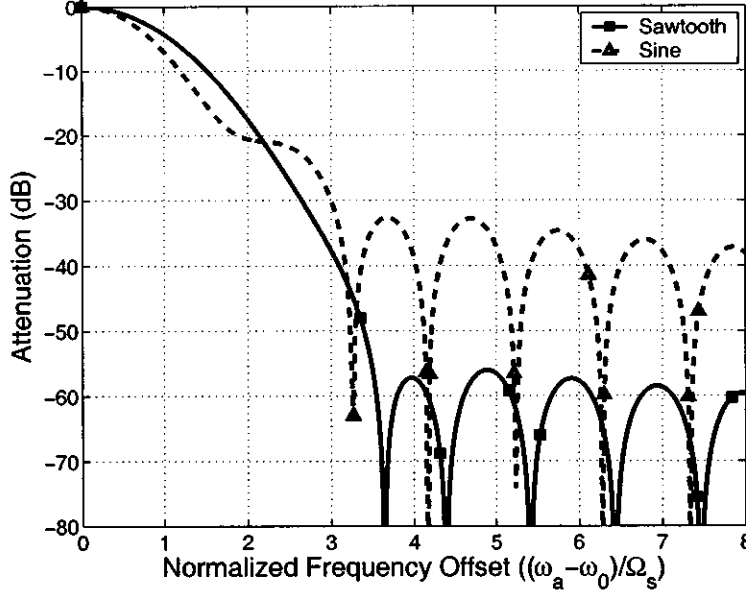


Figure 2-8: Frequency response using a sawtooth/ramp and sine damping functions ( $\gamma = 3$ ).

multiple sinusoidal inputs,

$$i_a(t) = \sum_n I_n \sin(\omega_n t), \quad (2.44)$$

$|X(\omega_0)|$  is evaluated using (2.34) and superposition

$$|X(\omega_0)| = \sum_n \frac{1}{2} \omega_n I_n |G(\omega_0 - \omega_n)|. \quad (2.45)$$

The  $|G(\omega_0 + \omega_n)|$  terms are ignored because  $G(\omega)$  is a baseband signal. This result can be combined with subsequent results to analyze the effect of blockers on the performance of the SRA.

### 2.3.4 SRA Response to a Pulse-Shaped Sinusoidal Input

Recent publications show that there is a benefit to using spread-spectrum techniques in super-regenerative receivers to improve their sensitivity [42, 54]. This can be done by shaping OOK pulses, such that a *one* is represented by the pulsed sinusoid

$$i_a(t) = \frac{T_q}{E_p} p(t) I_a \sin(\omega_0 t), \quad (2.46)$$

where  $T_q$  is the quench period and

$$E_p = \int_{-\infty}^{\infty} p^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |P(\omega)|^2 d\omega \quad (2.47)$$

is the energy of the normalized pulse. The pulse  $p(t)$  is defined to have a maximum value of unity at  $t = 0$  and is equal to zero for  $t < t_a$  and  $t > t_b$ . The term  $T_q/E_p$  is used such that the pulse-shaped input signal has the same energy over a quench period as a CW input signal. The input signal (2.46), has the spectrum

$$I_a(\omega) = \frac{jI_a T_q}{2E_p} (P(\omega + \omega_0) - P(\omega - \omega_0)). \quad (2.48)$$

Substituting (2.48) in (2.34) and evaluating  $|X(\omega)|$  at  $\omega_0$  yields

$$|X(\omega_0)| = \frac{I_a T_q}{4\pi E_p} \left| \int_{-\infty}^{\infty} (\omega_0 - \lambda) P(2\omega_0 - \lambda) G(\lambda) d\lambda + \int_{-\infty}^{\infty} (\omega_0 - \lambda) P(-\lambda) G(\lambda) d\lambda \right|. \quad (2.49)$$

Since both  $P(\omega)$  and  $G(\omega)$  are baseband signals, the first integral in (2.49) can be discarded, yielding

$$|X(\omega_0)| = \frac{I_a T_q}{4\pi E_p} \left| \int_{-\infty}^{\infty} \omega_0 G(\lambda) P(-\lambda) d\lambda - \int_{-\infty}^{\infty} \lambda G(\lambda) P(-\lambda) d\lambda \right|. \quad (2.50)$$

If  $g(t)$  and  $p(t)$  are even-symmetric, which is usually the case, their Fourier transforms are real and even-symmetric, making the second integral zero. Even without symmetry, this term is negligible since  $\lambda \ll \omega_0$  for values of  $G(\lambda)P(-\lambda)$  that are significant. This means that the second integral in (2.50) can be discarded, leaving

$$|X(\omega_0)| = \frac{\omega_0 I_a T_q}{4\pi E_p} \left| \int_{-\infty}^{\infty} G(\lambda) P(-\lambda) d\lambda \right|. \quad (2.51)$$

This can be interpreted to mean that  $|X(\omega_0)|$  is proportional to the cross-correlation of  $G(\omega)$  and  $P(-\omega)$ .

Solving this convolution and substituting its result in (2.42) gives the response of the SRA to a general pulse-shaped sinusoidal input. Time-domain techniques are



used in [42] to show that there is an optimum pulse shape that results in the maximum output signal energy for a given input signal energy. This can also be shown in the frequency domain by finding the pulse shape that maximizes  $|X(\omega_0)|$ . By Schwarz's inequality,

$$|X(\omega_0)|^2 = \left( \frac{\omega_0 I_a T_q}{4\pi E_p} \right)^2 \left| \int_{-\infty}^{\infty} P(-\lambda) G(\lambda) d\lambda \right|^2 \leq \left( \frac{\omega_0 I_a T_q}{2E_p} \right)^2 E_p E_g, \quad (2.52)$$

with equality occurring only if  $P(-\omega) = G(\omega)$ , a condition achieved if  $p(t) = g(t)$ .<sup>2</sup> In that case,  $|X(\omega_0)|$  is

$$|X(\omega_0)| = \frac{1}{2} \omega_0 I_a T_q. \quad (2.53)$$

This means that, for a given input signal energy, the output signal energy is maximized if  $p(t) = g(t)$ . In that case, the output signal energy is independent of the damping function and input pulse shape or energy. Instead it depends on the signal's amplitude, the SRA's resonant frequency, and the quench frequency. Note that using the optimal pulse shape comes at the expense of a wider transmitted signal. This is not always beneficial if a narrow transmission spectral mask is required. However, for spread spectrum (e.g. ultra-wide band) systems, this could be exploited to maximize the sensitivity of a super-regenerative receiver [42, 50].

## 2.4 Receiver Sensitivity Analysis

Modern SRA-based digital receivers are used almost exclusively to demodulate OOK signals. Their task is to detect whether a given transmitted bit was a *one* or a *zero*. For linear-mode SRA's, the actual detector implementation usually involves using a filter or envelope detector to measure the peak amplitude of the SRA's output, which occurs at time  $t_b$ . Since the phase information of the incoming signal is lost, the receiver is inherently non-coherent. When a *one* is sent, the signal at the input of

---

<sup>2</sup>Schwarz's inequality allows for an arbitrary constant of proportionality that is omitted here since  $p(t)$  and  $g(t)$  have normalized peak values of unity by definition.

the SRA is a pulsed sinusoid, and when a *zero* is sent, the input signal is modeled as additive white Gaussian noise (AWGN) (ignoring blockers). To calculate the sensitivity of the receiver, the probability density functions of these two cases must be solved. Any gain component that is common to both, such as  $\mu(t_b)$ , can be ignored since it does not affect the signal-to-noise ratio of the output signal. As a result, only the statistics of  $|X(\omega_0)|$  in (2.35) (or (2.37)) are needed. It is important to note that  $|X(\omega_0)|$  is not a function of time and can, therefore, be treated as a constant. To simplify the notation and clarify that it is a current, we define

$$I_X = |X(\omega_0)|. \quad (2.54)$$

Since the input signal contains a stochastic component,  $\mathbf{I}_X$  is a random variable.<sup>3</sup> To find the BER of the system for a given input signal amplitude, we must find the probability density functions for  $\mathbf{I}_X$  given that a *one* or a *zero* was transmitted; respectively  $p_{\mathbf{I}_X}(i_X|s_m = 1)$  and  $p_{\mathbf{I}_X}(i_X|s_m = 0)$ .

When a *zero* is transmitted, the incoming signal is strictly AWGN. Since  $\mathbf{X}(\omega)$  is the result of linear operations on Gaussian noise, it too has a Gaussian density function (though it is no longer white).  $\mathbf{X}(\omega_0)$  is, therefore, a Gaussian random variable, and  $\mathbf{I}_X$ , being its absolute value, has the Rayleigh density

$$p_{\mathbf{I}_X}(i_X|s_m = 0) = \frac{i_X}{\sigma_X^2} e^{-i_X^2/2\sigma_X^2}, \quad (2.55)$$

where  $\sigma_X^2$  is the variance of  $\mathbf{I}_X$  [33].

When a *one* is transmitted, the incoming signal can be modeled as the sum of a pulse-shaped sinusoid and AWGN, resulting in a rician density. However, for acceptable BER, the signal power must be considerably larger than the noise power, allowing a Gaussian density approximation for  $\mathbf{I}_X$ :

$$p_{\mathbf{I}_X}(i_X|s_m = 1) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-(i_X - I_s)^2/2\sigma_X^2}, \quad (2.56)$$

---

<sup>3</sup>Bold typeface is used to denote a random variable.

where  $I_s$  is the value of  $|X(\omega_0)|$  in response to a pulsed sinusoidal input signal of amplitude  $I_a$  and frequency  $\omega_0$ . For a general pulse shape and damping function,  $I_s$  can be calculated using (2.51) as

$$I_s = \frac{1}{2}\omega_0 I_a K_c, \quad (2.57)$$

where

$$K_c = \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} G(\lambda) P(-\lambda) d\lambda \right| \quad (2.58)$$

is defined as the *correlation coefficient* [33]. However, there are two specific cases of particular interest:

**Case 1,  $p(t) = 1$**

For the case with no pulse shaping,  $I_s$  can be derived using (2.41)

$$I_{s1} = \frac{1}{2} I_a \omega_0 G(0). \quad (2.59)$$

For the specific case of a ramp damping function,  $I_s$  is

$$I_{s1} = \frac{1}{2} I_a \omega_0 \sigma_s \sqrt{2\pi}. \quad (2.60)$$

**Case 2,  $p(t) = g(t)$**

When the pulse shape is matched to the sensitivity function,  $I_s$  is the same as (2.53)

$$I_{sg} = \frac{1}{2} I_a \omega_0 T_q. \quad (2.61)$$

From the analysis presented in Section 2.3.4, the signal energy should be maximized for  $p(t) = g(t)$ . For any sensitivity function, the benefit achieved by using that optimized pulse signal, therefore, is

$$\frac{I_{sg}}{I_{s1}} = \frac{T_q}{G(0)}. \quad (2.62)$$

Since the output noise power is the same in both cases, this gives a direct measure of the benefit in output signal-to-noise ratio. For the specific case of a ramp damping function, the benefit is

$$\frac{I_{sg}}{I_{s1}} = \frac{T_q}{\sigma_s \sqrt{2\pi}} = \frac{2\gamma}{\sqrt{2\pi}}, \quad (2.63)$$

making the substitution defined in (2.22). Recall, however, that (2.59) is based on the approximation in (2.53) which improves in accuracy for larger values of  $\gamma$  and becomes very accurate for  $\gamma \geq 3$ . For a value of  $\gamma = 3$ , the improvement in sensitivity is 7.6dB. It bears repeating that this benefit comes at the expense of a wider transmitted signal.

The final step required to find the sensitivity of an SRA is to solve for  $\sigma_X^2$ , the variance of  $\mathbf{I}_X$ .

### 2.4.1 Noise Analysis using the Convolution Model

The convolution model is particularly useful for analyzing an SRA's response to noise. The noise component of the input signal is modeled as a current  $i_n(t)$  with the power spectral density (PSD)

$$S_n(\omega) = \frac{N}{2}. \quad (2.64)$$

It is important to recall that noise and sinusoids are power signals with PSDs (i.e. they have infinite energy, but finite power). The sensitivity function  $g(t)$ , however, is an energy signal and does not have a PSD; instead it has an energy spectral density (ESD) [33]. Multiplying a power signal with an energy signal results in an energy signal, so  $x(t)$  is an energy signal and its ESD is the frequency-domain convolution of  $S'_n(\omega)$  and  $|G(\omega)|^2$ :

$$\overline{|X(\omega)|^2} = \frac{\omega^2}{2\pi} S_n(\omega) * |G(\omega)|^2 = \frac{N\omega^2}{4\pi} * |G(\omega)|^2. \quad (2.65)$$

The overline in (2.65) is used to signify *expected value* since  $x(t)$  is a random process and its Fourier transform,  $X(\omega)$ , is also random. Appendix A shows that this

convolution, evaluated at  $\omega_0$  simplifies to

$$\sigma_X^2 = \overline{|X(\omega_0)|^2} = \frac{1}{2}N\omega_0^2 E_g. \quad (2.66)$$

For the specific case of a ramp damping function,  $E_g = \sigma_s \sqrt{\pi}$ .

The variance  $\sigma_X^2$  is a measure of the output noise energy, whereas  $I_s^2$  is a measure of the output signal energy. The sensitivity of the receiver can be solved using these two quantities and the density functions (2.55) and (2.56).

## 2.4.2 Solving for the BER and Sensitivity

As shown in [33], the error probability for an OOK receiver with densities (2.55) and (2.56) can be closely approximated as

$$P_e = \frac{1}{2} e^{-\frac{I_s^2}{8\sigma_X^2}}. \quad (2.67)$$

Assuming that *ones* and *zeros* are equally likely means that the power received is

$$P_{Rx} = \frac{E_p I_a^2 R}{T_q 4}, \quad (2.68)$$

where  $R$  is the parallel resistance in Fig. 2-1(a). The noise density  $N$  can be written as the noise from  $R$  multiplied by some noise factor  $F$  which reflects the noise contribution from other sources (such as active devices) and depends on the actual topology of the SRA

$$N = \frac{4kTF}{R}. \quad (2.69)$$

Combining (2.57), (2.67)-(2.69), and (2.81) leads to the input signal power requirement for a given BER (or, equivalently,  $P_e$ )

$$P_{min} = \frac{-16kTF \ln(2P_e) E_g E_p}{K_c^2 T_q}. \quad (2.70)$$

This equation holds for general damping functions and pulse shapes. Using the definitions of  $E_g$ ,  $E_p$ , and  $K_c$ , it can be used to determine the sensitivity of a receiver that uses a generic linear-mode, slope-controlled SRA. As an example, the two important cases discussed previously are solve below for an SRA that uses a ramp damping function.

**Case 1,  $p(t) = 1$ , ramp damping**

For this case,  $K_c = \sigma_s \sqrt{2\pi}$ ,  $E_g = \sigma_s \sqrt{2\pi}$ , and  $E_p = T_q$ . Substituting these values in (2.70) yields

$$P_{min} = -8kTF \ln(2P_e) \frac{\Omega_s}{\sqrt{\pi}}. \quad (2.71)$$

For a desired BER of  $10^{-3}$ , (2.71) can be written in dBm as

$$P_{min,dBm} = -160\text{dBm} + 10 \log(F) + 10 \log(\Omega_s). \quad (2.72)$$

As might be expected, (2.72) shows that the sensitivity degrades with higher SRA bandwidths as is common with other receiver types. This equation is independent of  $\gamma$  and is accurate as long as  $\gamma$  is *big enough*. In previous sections it was shown that reducing  $\gamma$  affects the frequency response by increasing the side-lobes. But, as shown in Fig. 2-7, even for values of  $\gamma$  as small as 2, the numerical frequency response matches the estimate up to 20dB of attenuation, implying that the sensitivity functions presented are accurate for  $\gamma \geq 2$ .

The relationship between sensitivity and SRA bandwidth also affects the maximum data rate. The maximum data rate for an SRA receiver is  $1/T_q$ , and is achieved if the SRA is operated synchronously as in [43] and [10]. When synchronization is not used, the quench frequency must be greater than twice the bandwidth of the incoming signal since the SRA acts as a sampling device and must satisfy the Nyquist criterion. In either case, the bit rate is proportional to the quench frequency. For a particular value of  $\gamma$  (chosen to achieve a desired frequency response and gain),  $\Omega_s$  is proportional to the quench frequency and there is a direct tradeoff between sensitivity and bit rate [10].

**Case 2,  $p(t) = g(t)$ , ramp damping**

If the optimal pulse shape is used,  $K_c = T_q$  and  $E_g = E_p = \sigma_s \sqrt{\pi}$ , such that

$$P_{min} = -16kTF \ln(2P_e) \frac{\pi \sigma_s^2}{T_q^3}, \quad (2.73)$$

which can be written as

$$P_{min} = -8kTF \ln(2P_e) \frac{\pi \Omega_s}{4\gamma^3}. \quad (2.74)$$

In dBm, this is equivalent to

$$P_{min,dBm} = -160\text{dBm} + 10 \log(F) + 10 \log(\Omega_s) - 10 \log\left(\frac{4\gamma^3}{\pi^{3/2}}\right). \quad (2.75)$$

In this case, the sensitivity is a very strong function of  $\gamma$ . To appreciate the tradeoff, recall from (2.46) that the incoming current is proportional to  $T_q/E_p$ . When Gaussian pulses are used,  $T_q/E_p = 2\gamma/\sqrt{\pi}$ . This means the peak-to-average ratio of the transmitted signal's amplitude (which is regulated for some standards) is proportional to  $\gamma$ . Some spread spectrum standards, however, allow for high peak-to-average ratios, making this technique very attractive. As mentioned already, the bit rate is proportional to the quench frequency and, therefore, inversely proportional to  $\gamma$ . However, for the case of a matched pulse, the relationship is favorable since reducing the bit rate by  $2\times$  improves the sensitivity by  $8\times$  (9dB). Trading off sensitivity for bit rate by changing  $\Omega_s$  is also an option, but it is important to note that for a pulsed signal the transmitted spectrum depends on  $\Omega_s$ . When using Gaussian pulses, as in this analysis, the spectrum mask is set by  $P(\omega - \omega_0)$ , which would be Gaussian with a frequency standard deviation of  $\Omega_s$ .

### 2.4.3 Using a Time Random Variable for Detection

The probability density functions described by (2.55) and (2.56) are defined for the amplitude of the SRA's envelope at the end of each quench cycle. In OOK receivers, the optimum threshold for determining whether the received bit is a *one* or a *zero* is the point at which the two PDFs intersect [33]. This point is approximately equal to  $I_s/2$ . To achieve the minimum BER, the demodulator in the receiver should be able to determine this value accurately. Doing so is most important when the input signal is small since this is when there is the greatest amount of overlap between the PDFs. When the input signal is large, the accuracy of the threshold is less critical since there is a wider range of values that will yield an acceptable BER.

The analysis up to now assumes that the SRA is a linear system and, as a result,  $I_s$  is linearly proportional to the input signal. However, if the SRA is actually highly nonlinear (typically compressive),  $I_s$  may not be a good measure of the input signal's amplitude, and choosing the optimal detection threshold becomes challenging. This is illustrated in Fig. 2-9. If the system is linear, the SRA's envelope is larger at time  $t_b$  when a *one* is received and smaller when a *zero* is received as in Fig. 2-9(d). If the system is highly compressive, however, the difference in amplitude at time  $t_b$  could be largely independent of the input signal's amplitude, making detection very difficult (as shown in Fig. 2-9(e)).

Ensuring linearity in the SRA can be achieved practically by using feedback to actively limit its gain to a sufficiently low value. Such techniques have been used for many decades [71], but present two main problems. First, the gain-control loop adds complexity and power consumption. Second, the system still requires that the SRA maintain a significant linear range. In modern systems, the primary benefit of SRAs is their ultra-low power operation, and a primary means of keeping the power consumption at a minimum is by reducing the supply voltage. The requirement to maintain a wide linear voltage range threatens these benefits since it typically requires a high supply voltage. This motivates a desire to achieve the benefits of SRAs mentioned up to now while eliminating the need for a wide linear voltage



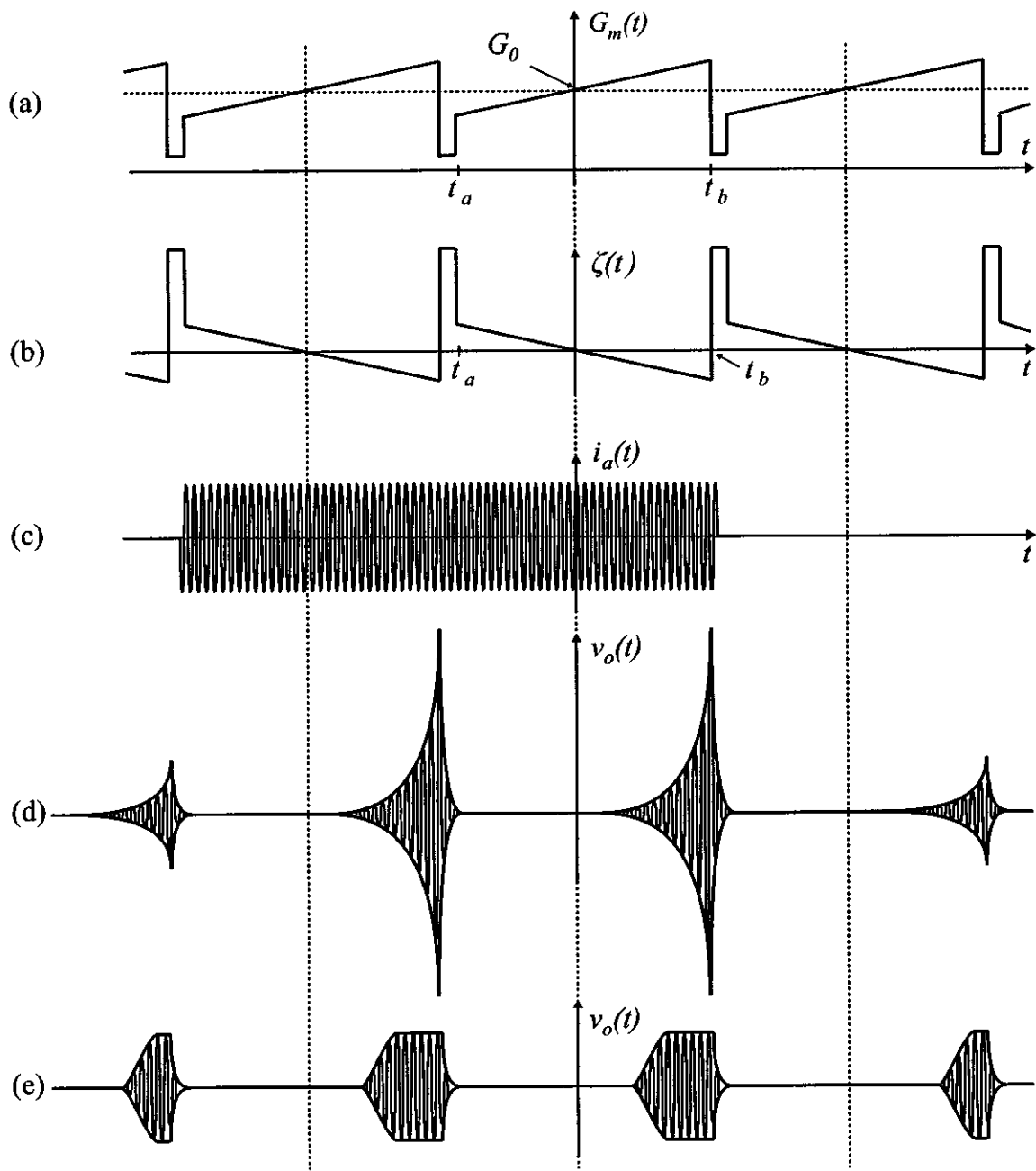


Figure 2-9: SRA (a) transconductance  $G_m(t)$ , (b) damping function  $\zeta(t)$ , (c) input current  $i_a(t)$ , (d) output voltage for linear system  $v_o(t)$ , and (e) output voltage for compressive nonlinear system.

range.

One way of doing this is to use a different means of detection that extracts information from the input signal *before* the SRA's output envelope grows to the point of becoming nonlinear. This can be done by setting a voltage limit,  $V_T$ , inside the linear range of the SRA and measuring the length of time required for the SRA's envelope to grow to that level. This is defined as the *trigger-time*, and it is a random variable [10]. For the case of a ramp damping function, setting  $v_e(t) = V_T$  in (2.37) and solving for  $t$  yields the random variable

$$\mathbf{T}_T = \sqrt{2\sigma_s^2 \ln \left( \frac{I_0}{\mathbf{I}_X} \right)}, \quad (2.76)$$

where

$$I_0 = \frac{V_T}{Z_0}. \quad (2.77)$$

As shown in Appendix B, the relationship between the probability density functions of  $\mathbf{I}_X$  and  $\mathbf{T}_T$  is

$$p_{\mathbf{T}}(t_T) = \left( I_0 \frac{t_T}{\sigma_s^2} e^{-\frac{t_T^2}{2\sigma_s^2}} \right) p_{\mathbf{I}_X} \left( I_0 e^{-\frac{t_T^2}{2\sigma_s^2}} \right). \quad (2.78)$$

Substituting (2.55) and (2.56) into (2.78) yields

$$p_{\mathbf{T}}(t_T | s_m = 0) = \left( \frac{I_0}{\sigma_X} \right)^2 \left( \frac{t_T}{\sigma_s^2} \right) e^{-\left( \frac{t_T^2}{\sigma_s^2} + \frac{I_0^2}{2\sigma_X^2} e^{-\frac{t_T^2}{\sigma_s^2}} \right)}, \quad (2.79)$$

and

$$p_{\mathbf{T}}(t_T | s_m = 1) = \left( \frac{I_0}{\sigma_X \sqrt{2\pi}} \right) \left( \frac{t_T}{\sigma_s^2} \right) e^{-\left( \frac{t_T^2}{2\sigma_s^2} + \frac{1}{2\sigma_X^2} \left( I_0 e^{\frac{-t_T^2}{2\sigma_s^2}} - I_s \right)^2 \right)}. \quad (2.80)$$

While the density functions of  $\mathbf{T}_T$  are far messier than those of  $\mathbf{I}_X$ , they allow accurate detection even when the SRA is highly nonlinear. This is because  $V_T$  can be set to a small value such that the trigger-time is extracted before the SRA's amplitude causes nonlinearities in the system. Furthermore, the probability of error in detection must, intuitively, be the same as if  $\mathbf{I}_X$  were used since there is a one-to-one mapping between their densities.

## 2.5 Measurement Results

Fig. 2-10 shows the schematic for a simple SRA based on a common-base Colpitts oscillator and an envelope detector. The resonant frequency was arbitrarily chosen in the 400MHz frequency range reserved for Medical Implant Communication Services (402–405MHz), but much higher frequencies could be used. Typically an LNA is used to isolate the antenna from the SRA so that its output signal is not radiated. But to better characterize the SRA and verify the theory presented, the LNA was excluded. A common-base configuration was chosen for the SRA to facilitate input matching, and the quench signal used to control the damping function was connected to the base through a simple low-pass filter. A high Q inductor was used in the resonant tank so that the effective parallel resistance is dominated by well controlled resistances (i.e.  $r_e$  of the transistor,  $50\Omega$  source impedance, and  $560\Omega$  bias resistor). While this is sub-optimal for performance, it allows for more accurate performance prediction. The envelope detector is similar to [43] and a sawtooth damping function similar to Fig 2-9(b) was generated using an arbitrary waveform generator. A low supply voltage of 1.0V was used to show that accurate detection is possible using the

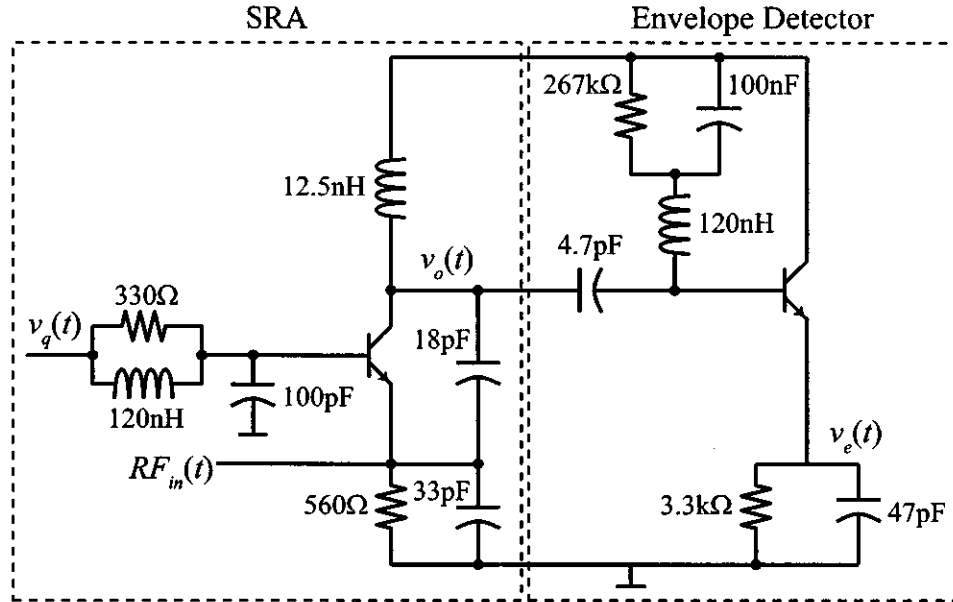


Figure 2-10: Schematic of SRA and envelope detector.

trigger-time technique despite a very narrow linear range.

Fig. 2-11 shows the output of the envelope detector for different input power levels of CW signals at the resonant frequency of the SRA. The dashed lines show the predicted shape of the envelope compared to the measurement results shown by the solid lines. Two sample signals are shown for the case when no input signal was used to show the random nature of the signal amplitude. Note that the measured results are in agreement with predictions when the SRA envelope is small, but diverge significantly as the nonlinearities of the system lead to signal compression. If this SRA had been designed to function strictly in the linear region, the SRA gain would have been set much lower (by reducing  $\gamma$  or  $\Omega_s$ , for example) or the supply voltage would have been made larger. Using the trigger-time technique eliminated the need to do so. Instead, the threshold voltage was set to a low value ( $V_T = 15\text{mV}$ ) and the trigger-time was measured using the histogram function of a digital oscilloscope. From Fig. 2-11 it is clear that there is good agreement between theory and measured results for values of  $v_e(t)$  below  $15\text{mV}$ , so the measured trigger-time PDFs should be in agreement with theory.

Fig. 2-12 shows the measured and theoretical probability density functions of

$T_T$  for varying input signal levels. For this measurement, the quench signal of the SRA was set to 300kHz, and its slope resulted in a value of  $4.4\text{MHz}\sqrt{\text{rad}}$  for  $\Omega_s$ . The density functions were extracted by weighting the time histograms such that they integrate to unity. The theoretical distribution functions of  $T_T$  are also plotted according to (2.79) and (2.80). Clearly, there is excellent agreement between the theoretical and measured signals for input power levels up to  $-80\text{dBm}$ . For larger input signals the theory becomes less accurate because  $t_T$  is no longer sufficiently larger than  $\sigma_s$ , as required for the convolution approximation to be accurate. This is not important, however, since selecting the detection threshold is trivial when the input signal is large (i.e. there is a wide range of values for which the BER will meet requirements). In contrast, accurately modeling the PDFs for low input signal levels is very important since it enables the selection of the optimum detection threshold, resulting in the best sensitivity. Fig. 2-12 shows there is excellent agreement between theory and measurements for small input signal levels. Furthermore, the  $-80\text{dBm}$  value is only specific to this design and only important in the sense that it represents an input signal significantly larger than the sensitivity level. For receiver designs with much lower or higher sensitivities, the theoretical PDFs are expected to match for a wide enough range of input signal values to allow for optimal threshold determination and, thereby, optimal sensitivity.

The BER, which is equivalent to the probability of error  $P_e$ , was calculated by finding the optimal threshold for a given input power and integrating the portion of each PDF that was on the *wrong side*. For example, for a  $-90\text{dBm}$  input signal, the optimum threshold was determined to be  $0.66\mu\text{s}$ . The probability of error was found by integrating the  $-90\text{dBm}$  PDF from  $t = 0.66\mu\text{s}$  to  $t = 1\mu\text{s}$  and adding the result to the integral of the noise PDF from 0 to  $0.66\mu\text{s}$ . Using this technique, it was determined that a BER of  $10^{-3}$  was achieved for a  $-87\text{dBm}$  CW input signal. This corresponds to an OOK modulated input signal level of  $-90\text{dBm}$  (since its average power would be 3dB lower), which matches the predicted value given by (2.72) for a noise factor of  $F = 2.3$  (3.6dB).

The DC current consumption of the SRA was  $500\mu\text{A}$  for a total power consumption

of  $500\mu\text{W}$ . The theoretical noise factor was calculated using techniques similar to those used in [10]. The effective resistance at the emitter of the SRA transistor is the parallel combination  $R_e = R_s || r_e || 560\Omega = 24.4\Omega$ , where  $R_s = 50\Omega$  is the source impedance of the signal generator and  $r_e = 25.9\text{mV}/500\mu\text{A} = 52\Omega$  is the small signal emitter resistance of the BJT. The effective resistance across the resonator is  $R_{eff} = R_e((18\text{pF} + 33\text{pF})/18\text{pF})^2 = 196\Omega$  [35]. The BJT's thermal noise density is given by  $N_{bjt} = 2qI_D = 160 \times 10^{-24}$  and the resistor's noise density is given by  $N_{res} = 4kT/R_{eff} = 84.5 \times 10^{-24}$ . The noise factor is, therefore,  $F = N_{tot}/N_{res} = 2.9$  (4.6dB). This means that the theoretical noise factor and the measured noise factor are within 1dB, confirming the accuracy of the theory.

Fig. 2-13 shows the filtering characteristics of the SRA resulting from the attenuating effect of frequency mismatch between the SRA's resonator and the input signal. For this measurement, the quench signal of the SRA was set to 300kHz, and its slope resulted in a value of  $2\pi \times 480\text{kHz}\sqrt{\text{rad}}$  for  $\Omega_s$ . This, in effect, describes the selectivity of the SRA. The measurement was made by finding the average value of  $t_T$  for a -80dBm CW input signal at 403MHz (the SRA's resonant frequency) and then sweeping the frequency and power level of the input signal. For each frequency, the input signal power was swept until the value of  $t_T$  was the same as for the reference signal (-80dBm, 403MHz). Measured results show excellent matching compared to (2.43) up to about 25dB of attenuation. Beyond such levels, other phenomena, including the effects of finite  $\gamma$ , begin to dominate, making the measurement less accurate.

## 2.6 Summary

A frequency-domain approach to analyzing super-regenerative amplifiers has been presented. Sensible approximations have been made that enable a convolution model to describe part of the time-varying solution. The convolution model was used to find the SRA's response to arbitrary deterministic and stochastic signals, with specific examples of its response to a single sinusoid, multiple sinusoids, a pulsed sinusoid, and AWGN. These solutions were then used to find the sensitivity of a synchronous

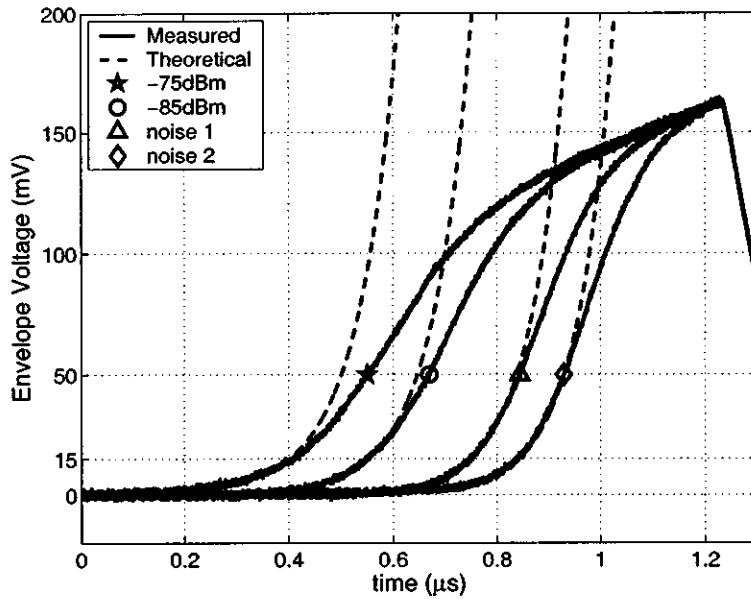


Figure 2-11: Theoretical and measured envelope detector waveforms for -75dBm and -85dBm CW input signals and two sample waveforms for no input signal (i.e. only noise).

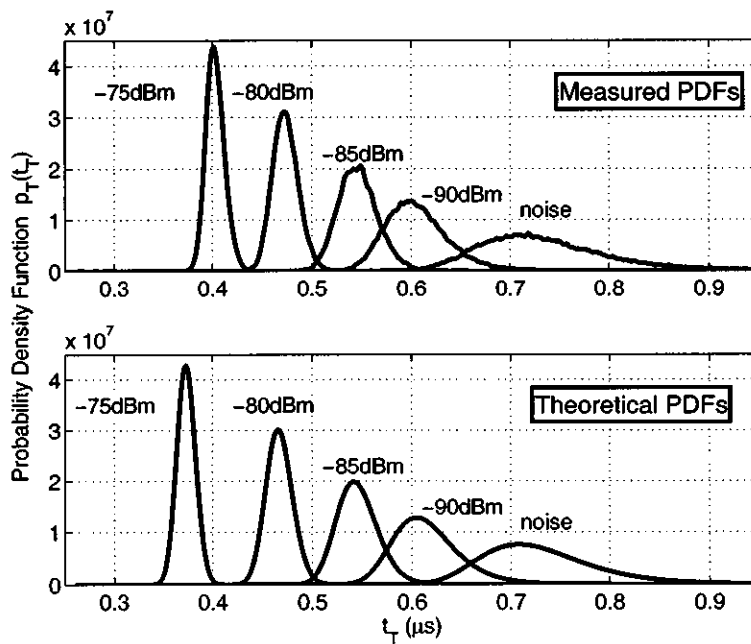


Figure 2-12: Measured and theoretical probability density functions for the random variable  $t_T$ .

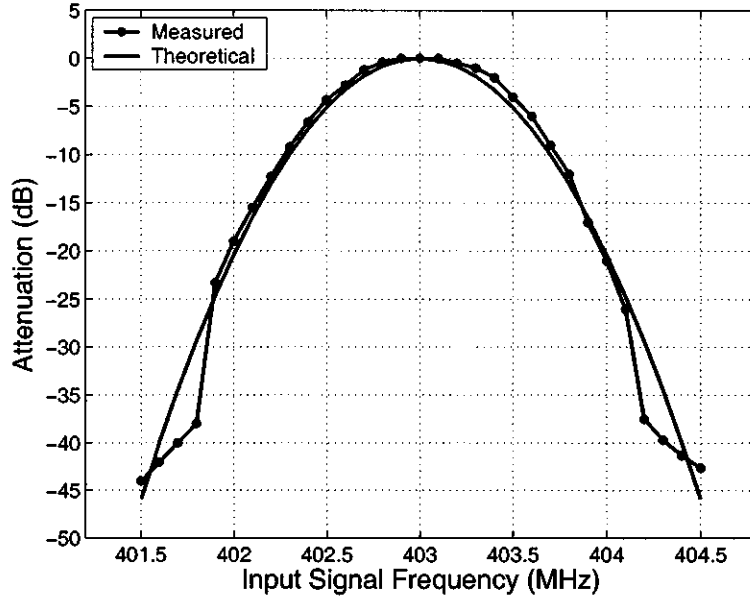


Figure 2-13: Frequency response of the SRA for  $\Omega_s/(2\pi) = 480\text{kHz}/\sqrt{\text{rad}}$ .

SRA receiver to an OOK signal with and without pulse shaping, and the benefits of both cases were discussed. The probability density functions were found for a trigger-time random variable that can be used for OOK detection and helps avoid the problems associated with SRA nonlinearity. Finally, experimental data was presented that matched the theory with excellent agreement.

## 2.7 Appendix 1: Convolution yielding the variance of $\mathbf{I}_X$

The first step to finding the variance of  $\mathbf{I}_X$  is to solve the convolution

$$\overline{|X(\omega)|^2} = \frac{N\omega^2}{4\pi} * |G(\omega)|^2 = \frac{N}{4\pi} \int_{-\infty}^{\infty} (\omega - \lambda)^2 |G(\lambda)|^2 d\lambda. \quad (2.81)$$



Expanding this solution and substituting  $\omega = \omega_0$  yields

$$\overline{|X(\omega_0)|^2} = \frac{N}{4\pi} \left( \int_{-\infty}^{\infty} \omega_0^2 |G(\lambda)|^2 d\lambda - \int_{-\infty}^{\infty} 2\omega_0 \lambda |G(\lambda)|^2 d\lambda + \int_{-\infty}^{\infty} \lambda^2 |G(\lambda)|^2 d\lambda \right). \quad (2.82)$$

Since  $|G(\omega)|$  is an even function, the second integral can be discarded. A further simplification can be made by observing that  $G(\omega)$  is a baseband signal, and therefore  $|G(\lambda)|$  becomes very small for values of  $\lambda$  well below  $\omega_0$ . As a result, over the range of  $\lambda$  for which  $|G(\lambda)|$  has a significant value, the  $\lambda^2$  term in the third integral is much smaller than the  $\omega_0^2$  term in the first integral. This means the third integral can also be discarded resulting in the simplified solution for the variance  $\sigma_X^2$

$$\sigma_X^2 = \overline{|X(\omega_0)|^2} = \frac{1}{2} N \omega_0^2 E_g, \quad (2.83)$$

where  $E_g$  is the energy of the sensitivity function defined as

$$E_g = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G(\omega)|^2 d\omega = \int_{-\infty}^{\infty} g^2(t) dt. \quad (2.84)$$

## 2.8 Appendix 2: Solving for the PDF of $\mathbf{T}_T$

In Section 2.4.3 the trigger-time random variable,  $\mathbf{T}_T$  was introduced to help alleviate some of the problems created by the nonlinearity of the SRA's active elements.  $\mathbf{T}_T$  is related to  $\mathbf{I}_X$  through (2.76), and its probability density function is the derivative

of its cumulative distribution function [36] given by

$$F_{\mathbf{T}}(t_T) = P \{ \mathbf{T}_{\mathbf{T}} \leq t_T \} \quad (2.85)$$

$$= P \left\{ \sqrt{2\sigma_s^2 \ln(I_0/\mathbf{I}_{\mathbf{X}})} \leq t_T \right\} \quad (2.86)$$

$$= P \left\{ \mathbf{I}_{\mathbf{X}} \geq I_0 e^{-\frac{t_T^2}{2\sigma_s^2}} \right\} \quad (2.87)$$

$$= 1 - F_{\mathbf{X}} \left( I_0 e^{-\frac{t_T^2}{2\sigma_s^2}} \right). \quad (2.88)$$

Differentiating (2.88) with respect to  $t_T$  yields the PDF of  $\mathbf{T}_{\mathbf{T}}$  with respect to the PDF of  $\mathbf{I}_{\mathbf{X}}$

$$p_{\mathbf{T}}(t_T) = \left( I_0 \frac{t_T}{\sigma_s^2} e^{-\frac{t_T^2}{2\sigma_s^2}} \right) p_{\mathbf{I}_{\mathbf{X}}} \left( I_0 e^{-\frac{t_T^2}{2\sigma_s^2}} \right). \quad (2.89)$$

## Chapter 3

# Ultra-Low Power Transceiver for Medical Implant Communications

Recent advances in the medical field are spurring the need for ultra-low power transceivers for wireless communication with medical implants. To deal with the growing demand for medical telemetry, the FCC commissioned the Medical Implant Communications Services (MICS) standard in 1999 in the 402-405MHz band. In this chapter, we leverage the SRA theory described in Chapter 2 to develop a  $350\mu\text{W}$  FSK/MSK direct modulation transmitter and a  $400\mu\text{W}$  OOK super-regenerative receiver (SRR) specifically optimized for medical implant communications. The transceiver is implemented in 90nm CMOS and digitally tunes 24MHz in frequency steps smaller than 2kHz. The transmitter meets MICS mask specifications with data rates up to 120kbps consuming only 2.9nJ/bit; the receiver has a sensitivity better than  $-99\text{dBm}$  with a data rate of 40kbps or  $-93\text{dBm}$  with a data rate of 120kbps consuming 3.3nJ/bit. A frequency correction loop incorporating the base-station is prototyped to eliminate the need for a frequency synthesizer in the implant while still achieving frequency stability of less than 3ppm.

### 3.1 Introduction

There are a few critical observations that motivate this work. The first is that the human body is an excellent temperature regulator, and the second is that the MICS standard features relaxed output power and frequency stability specifications. Together, these observations point to simplified transceiver architectures that consume less power than those commonly employed for other applications. Furthermore, while it is critical that the implant consume minimal power in order to preserve battery life, the corresponding base-station is free to consume much more power. This observation motivates shifting complexity in the wireless link from the implant to the base-station.

With these observations in mind, we propose a simple, low-power topology where a digitally-controlled oscillator (DCO) is directly modulated using frequency-shift keying (FSK) [9, 10]. Instead of using a PA to drive the loop antenna, the DCO incorporates it as its inductive element, radiating energy that would otherwise be lost as thermal heat [30]. To concurrently achieve an acceptably wide tuning range and fine frequency resolution, a sub-ranged capacitor array is used to divide 20 bits of frequency tuning into coarse, medium, and fine tuning capacitor banks. The capacitor banks are predistorted to achieve linear digital-to-frequency conversion and 14 effective bits of frequency resolution.

A challenge that arises from having the antenna attached to the DCO is that popular receiver architectures such as the super-heterodyne or homodyne topologies cannot be used. Instead, we propose a super-regenerative architecture to demodulate on-off keying (OOK), achieving excellent sensitivity and good selectivity while consuming less than  $400\mu\text{W}$ . By optimizing the system holistically, we achieve data transmission consuming  $2.9\text{nJ/bit}$  and reception consuming  $3.3\text{nJ/bit}$  while meeting the MICS 300kHz channel bandwidth requirements.

This chapter is organized as follows. Section 3.2 describes the transceiver’s architecture, Section 3.3 describes the FSK/MSK transmitter, Section 3.4 describes the OOK SRR, Section 3.5 details the circuit implementation of each block, Section 3.6

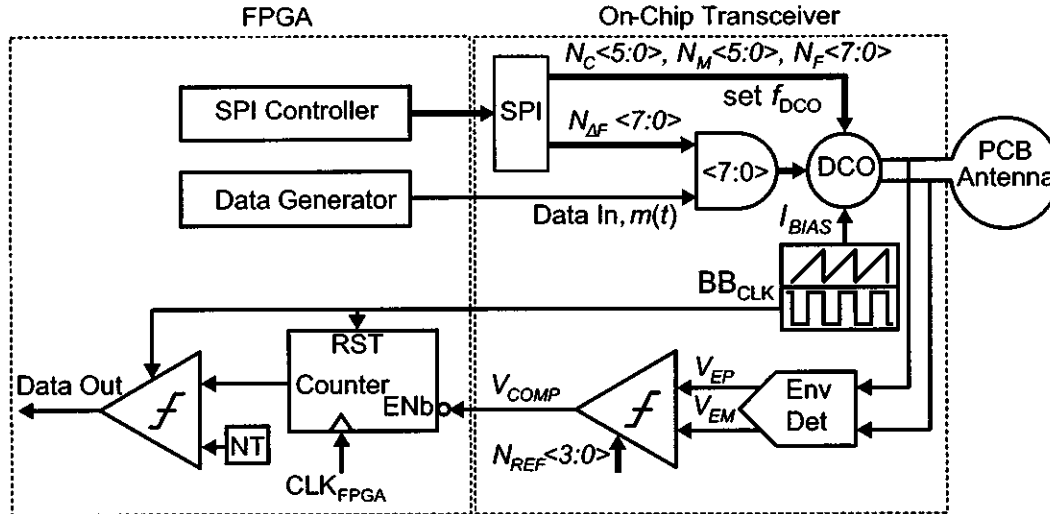


Figure 3-1: Transceiver block diagram.

describes a prototyped frequency correction loop used to set the DCO's center frequency, and Sections 3.7 and 3.8 discuss measurements of the the prototype and summarize the chapter.

## 3.2 Architecture Overview

Fig. 3-1 shows the prototype transceiver comprising a simple digital baseband implemented in an FPGA, a direct modulation FSK transmitter, and a super-regenerative receiver. The transmitter and receiver are time-division multiplexed and share an external loop antenna implemented on the prototype PCB. The low radiation power requirements of MICS are exploited in the transmitter by eliminating the PA and incorporating the antenna into the DCO<sup>1</sup>. On the receive side, the same DCO is used as a synchronous OOK super-regenerative receiver. As explained in Section IV, the SRR provides a tremendous amount of gain that results in relaxed input-referred noise specifications for subsequent stages. To exploit this benefit, the envelope detector and programmable comparator in Fig. 3-1 were biased in subthreshold and optimized for ultra-low power consumption.

<sup>1</sup>The maximum transmission power is  $25\mu\text{W}$  effective isotropic radiated power (EIRP), defined as the product of radiated power and the antenna gain.

The following subsections describe the unique features of medical implant communications that were exploited to achieve ultra-low power operation. Section 3.3 describes the theory and design of the direct modulation FSK transmitter. Section 3.4 explains the key elements of super-regeneration theory and describes the receiver implementation in detail. Section 3.5 describe the circuit implementations of the transceiver. Section 3.6 introduces a prototype frequency correction loop used to calibrate the DCO's frequency without the need for a frequency synthesizer in the implant. Finally, Sections 3.7 and 3.8 show measurement results for the transceiver and conclude the paper.

### **3.2.1 Exploiting the Unique Features of Medical Implants**

A fundamental difference between medical implant transceivers and more conventional radios is the environment in which they operate. Cellular phones, for example, must maintain a frequency stability of less than 1ppm in the temperature range of  $-40:+85^{\circ}\text{C}$ , while the MICS standard only requires 100ppm of stability in the range of  $25-45^{\circ}\text{C}$ . Furthermore, while MICS calls for proper implant operation over a  $20^{\circ}\text{C}$  range, the temperature of the human body rarely changes so drastically, and its moderate shifts occur very slowly. This quality of temperature regulation in the human body can be exploited to reduce the complexity, power consumption, and size of medical implants. For example, by choosing architectures that do not require stringent frequency accuracy in the implant, the classical frequency synthesizer can be replaced by a frequency-control loop that does not require a crystal oscillator in the implant. In section 3.6, such a system is described.

Another quality of medical implants that can be exploited is the low bandwidth of most biomedical signals. Since low sample rates can be used to digitize these signals, the transmitter can be off most of the time as data is accumulated. The transmitter is then turned on briefly, sending data packets in short bursts. To maximize the battery life, the transmitter should be designed to minimize the energy consumed per transmitted bit. This criteria creates a tradeoff between complex architectures that are more spectrally efficient but consume more power, and simpler topologies

that consume less power but are less spectrally efficient [67]. Direct modulation FSK transmitters have the advantage of consuming very little power due to their simplicity, while achieving good spectral efficiency when modulated using minimum-shift keying (MSK). Other modulation schemes (such as 256-QAM), while more spectrally efficient, require complex topologies that consume more power and may result in lower energy efficiency.

### 3.2.2 Link Analysis and Antenna Considerations

For MICS transceivers, antenna gain (the product of efficiency and directivity) inside the human body is considerably lower than in free-space for two main reasons. First, the size of the antenna is typically much smaller than the wavelength of the signal (75cm), and second, human tissue is conductive causing large amounts of loss. Fortunately, MICS data links are only intended to function within a 2 meter radius limiting the amount of path loss. Measurements on MICS signal propagation reported in [28] show the worst case path loss in a hospital room to be 47dB including fading margin. Using these results along with the MICS requirement limiting radiation to -16dBm EIRP and reasonable assumptions about the base-station lead to requirements for the medical implant.

Assuming the base-station transmits at the maximum -16dBm EIRP, means that at least -63dBm of power will be available at any location in the hospital room. This places a restriction on the receiver's sensitivity and antenna gain. For example, if the sensitivity is -92dBm, the antenna gain (including path loss in the body) must be at least -29dBi ( $P_{sens,imp} \leq P_{tx,bs} + G_{path} + G_{ant}$ , in dB). Using the same antenna to transmit from the implant and assuming the sensitivity of the base-station receiver is -105dBm (sensible for a 300kHz bandwidth channel), the power delivered to the antenna by the implanted transmitter must be at least -29dBm or  $1.3\mu\text{W}$  ( $P_{tx,imp} \geq P_{sens,bs} - G_{path} - G_{ant}$ ).

A *small* loop antenna (relative to its signal's wavelength) can be modeled as the series combination of an inductor and a resistor. The resistance is composed of two elements: radiation resistance and loss [41]. The radiation resistance is desirable since

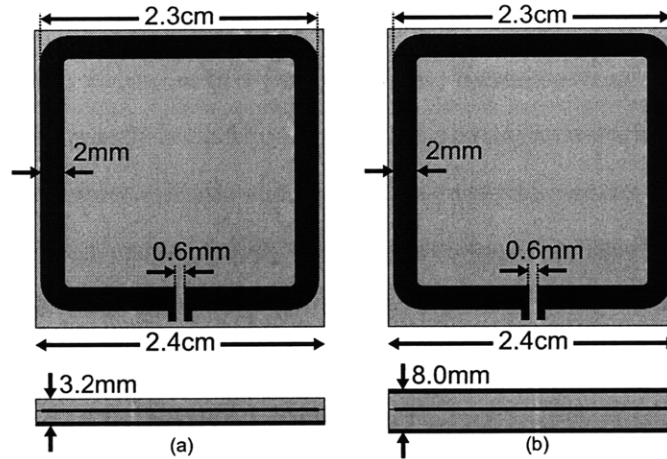


Figure 3-2: (a) Loop antenna with FR4 substrate and superstrate, and copper patch below the substrate. (b) Loop antenna with metal patches above and below the substrate and superstrate.

it models the conversion of energy into electromagnetic waves. The loss, expectedly, is not desirable since it reduces the antenna's efficiency:

$$\epsilon_{rad} = \frac{R_{rad}}{R_{rad} + R_{loss}}. \quad (3.1)$$

The poor efficiency of small loop antennas is manifested in their low radiation resistance and results in their impedance exhibiting a high quality factor (Q). To achieve high power transfer, the inductive element of the antenna must be resonated out by a capacitor. The high Q, however, poses challenges since the frequency response of the resonator is narrow and even small impedance variations can result in significant mismatch.

This last challenge can be overcome, and actually exploited. Since the power needed to drive the antenna in an MICS transmitter is only a small portion of the power budget, a smaller emphasis can be placed on PA efficiency. Using the antenna as the inductive element of an LC-oscillator eliminates the need for an explicit PA since the resulting power oscillator drives the antenna. Further benefits include inherent impedance matching, low-power consumption, and low-noise design since the small loop antenna has a high Q.

A drawback of using the antenna as part of the oscillator is that changes in the



antenna's environment lead to frequency pulling. To compensate for this effect, the bandwidth of the frequency-control loop (FCL) must be higher than the maximum rate of change of the implant's environment. Studies on human gait show that most human motion has frequency content of less than 10Hz [3]. Therefore, as long as the frequency pulling does not exceed the lock-in range of the FCL and the FCL bandwidth is greater than 10Hz, the effects of frequency pulling can be properly mitigated. To verify that frequency pulling does not pose a serious threat to the functionality of the system, we tested our chip with the simple antenna shown in Fig. 3-2(a). It uses a 1.6mm FR4 substrate and superstrate with a single patch on the bottom side. HFSS simulations show this structure to exhibit a Q of 115 in free space, an inductance of 23nH, and a radiation efficiency of 0.4%. We performed a few simple experiments such as waving a hand near the antenna and the FCL maintained correct center frequency with ease.

Perhaps a more serious threat to the functionality of the system is that the high conductance of human tissue makes a loop antenna lossy, lowering its Q and efficiency [31]. HFSS simulations show that high Q and acceptable efficiency can be maintained by using a substrate and superstrate with metal patches above and below as shown in Fig. 3-2(b). This antenna has a diameter of 2.3cm, uses a 4mm substrate and superstrate (with a relative permittivity of 3.5 and a dissipation factor of 0.0027). HFSS simulations of the structure immersed in a model of human tissue (relative permittivity of 42.8, conductivity of 0.65 S/m) show a worst-case Q of 130 and a radiation efficiency of 0.1%. Note that this Q is slightly better than the Q of the antenna used in the prototype. It follows that the difficulties of transmitting through human tissue will not compromise the performance of this system.

### 3.3 FSK/MSK Transmitter

Connecting the antenna directly to an oscillator precludes the use of conventional up-conversion transmitters. Fortunately, frequency and phase modulation techniques can be used by modulating the oscillator directly. The result is a simple transmitter

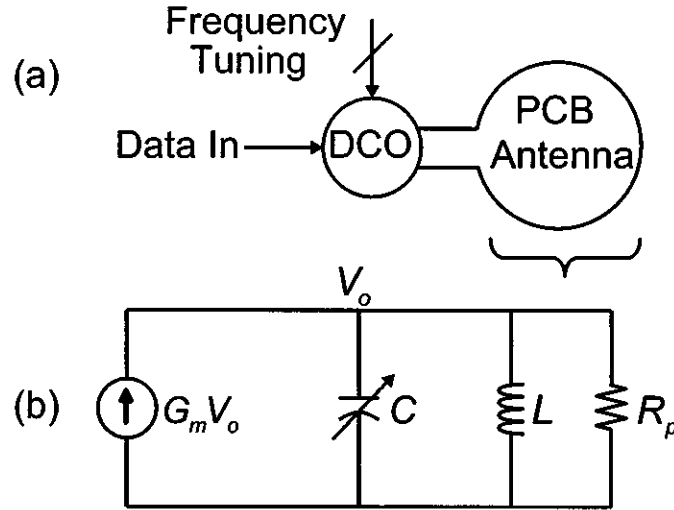


Figure 3-3: (a) Direct modulation FSK transmitter and (b) simplified circuit model.

architecture that is spectrally efficient and consumes very little power.

### 3.3.1 FSK/MSK Theory

FSK is a constant envelope form of digital modulation and can be easily implemented by directly modulating the instantaneous frequency of an oscillator:

$$f_i(t) = f_c + \Delta F \times m(t), \quad (3.2)$$

where  $f_c$  is the carrier frequency,  $\Delta F$  is the frequency deviation constant, and  $m(t) \in [-1, 1]$  is the digital modulating signal [14]. For a bit rate  $R$ , setting  $\Delta F = 0.25R$  results in MSK; the most spectrally efficient form of FSK that still produces orthogonal signaling and can be demodulated coherently. This relaxes SNR requirements in the base station receiver reducing output power requirements on the implanted transmitter. Theoretically, MSK can be used at a bit rate of 200kb/s without any pre-filtering and meet the MICS spectral mask since its first nulls occur at  $f_c \pm 0.75R = f_c \pm 150\text{kHz}$  and subsequent peaks are at least 25dB below the main lobe.

### 3.3.2 Transmitter Implementation

Fig. 3-3(a) shows a simple FSK transmitter comprising a DCO that incorporates a small loop antenna as its inductive element, and is modulated directly with digital data. Fig. 3-3(b) shows the simplified schematic of the DCO with an equivalent parallel resistance of

$$R_p \approx Q\omega_0 L \quad (3.3)$$

where  $\omega_0$  is the tank's resonant frequency. Frequency tuning and modulation are done using switched capacitors that have much lower temperature coefficients than varactors and allow for a fully digital implementation. A challenge that arises is that many bits are required to tune a wide frequency range while achieving small frequency steps. Furthermore, classical implementations of capacitor arrays would require impractically small capacitors. For example, to tune 24MHz (391-415MHz) in 2kHz steps requires 14 bits of resolution. Assuming an inductance value of 24nH, the capacitance must tune from 6.128pF to 6.904pF with a minimum capacitor step size of 0.047fF. Since it is impractical to implement such small capacitors in CMOS, a sub-ranging capacitor array was implemented to achieve very small effective capacitor step sizes while using practical capacitor values [52].

### 3.3.3 Capacitor Array with Predistortion

The digitally tunable capacitor in Fig. 3-3(b) is implemented using the four capacitor banks shown in Fig. 3-4(a). The carrier frequency ( $f_c$ ) of the DCO is tuned with coarse, medium, and fine-tuning capacitor banks  $C_C$ ,  $C_M$ , and  $C_F$  while the frequency deviation constant is set using  $C_{\Delta F}$ . Each capacitor bank is thermometer coded and predistorted as in [37] to improve linearity in digital-to-frequency conversion and guarantee monotonicity for each bank. This is critical since the DCO will be placed in a FCL and non-monotonicity could result in instability.

Capacitor bank  $C_C$  provides six bits of coarse frequency tuning and is directly connected across the inductor. As a result, any change in its capacitance results in an equal change in the total resonator capacitance. A small capacitor  $C_{SM}$  is

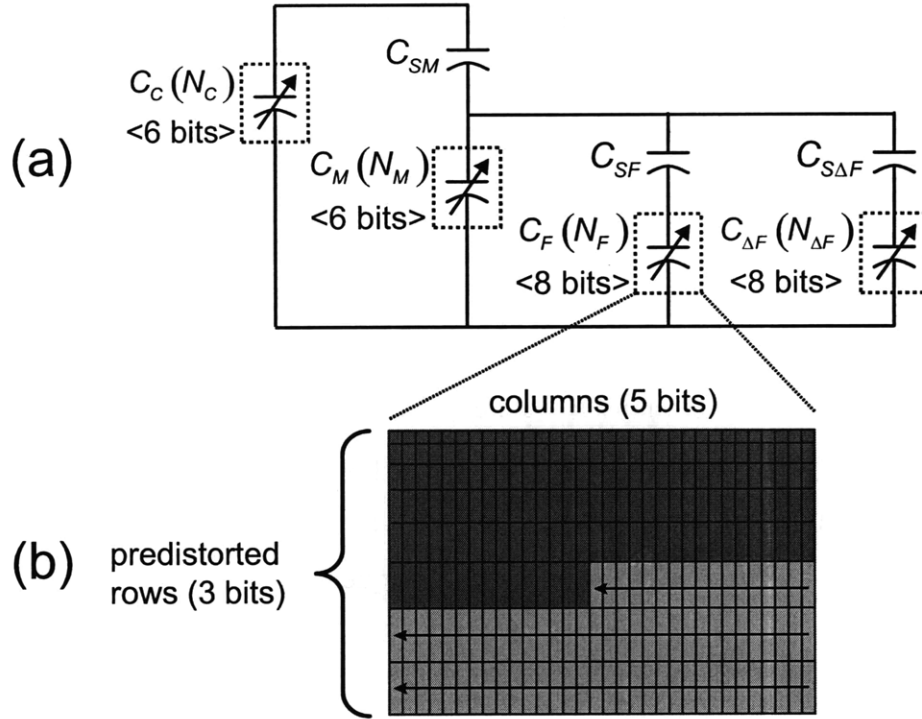


Figure 3-4: (a) Sub-ranging capacitor array and (b) piece-wise linear predistortion.

connected in series with capacitor bank  $C_M$  to provide six bits of medium frequency tuning. Since  $C_{SM}$  is much smaller than  $C_M$ , the effective change in the resonator's capacitance is much smaller than changes made to the capacitor bank  $C_M$ . This allows switch capacitors on the order of 10fF to be used while achieving incremental capacitance changes across the inductor on the order of 0.5fF. Similarly,  $C_{SF}$  is used to further reduce the effective capacitance changes across the inductor when the fine tuning capacitor bank  $C_F$  is tuned. The result is incremental resonator capacitance changes on the order of 30aF achieved using capacitors on the order of 10fF.

Since the DCO frequency is proportional to  $1/\sqrt{C}$ , even a linearly tuned capacitor array would result in nonlinear digital-to-frequency conversion. Using series capacitors to reduce the effective capacitance change across the inductor has the undesirable effect of adding further nonlinearity. Fortunately, capacitor bank predistortion can be used to mitigate nonlinearity. Choosing proper predistortion was done heuristically through simulation in [37], but the design cycle can be shortened by solving for the optimum digital-to-capacitance curves mathematically. For example, to find the

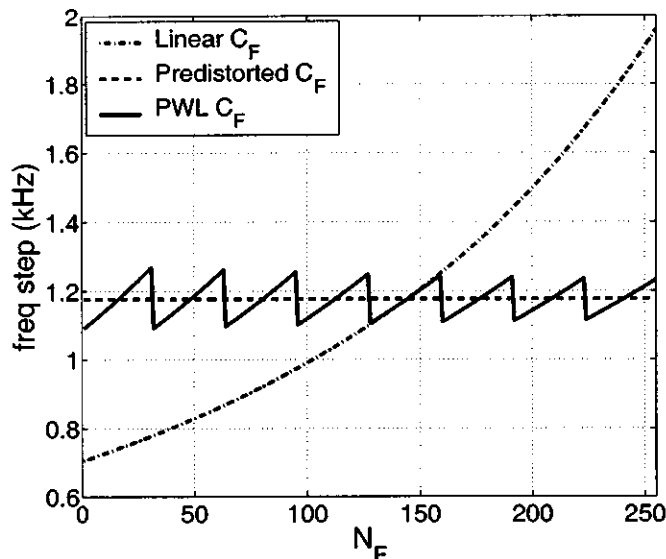


Figure 3-5: Frequency steps vs.  $N_F$  using linear, predistorted, and piece-wise linear (PWL) predistorted capacitor banks.

proper values of  $C_F$ , the other capacitor banks are set to their mid-range values and the relationship between  $C_F(N_F)$  and  $f_{DCO}(N_F)$  is found.

To achieve a perfectly linear digital-to-frequency relationship between  $N_F$  and  $f_{DCO}$ , the 256 desired frequencies are used to find the respective values of  $C_F$ . Implementing 256 unique incremental values of capacitance for  $C_F$ , however, would be impractical. Instead, a very good piece-wise linear approximation can be implemented using a thermometer-coded capacitor bank like the one shown in Fig. 3-4(b) where the capacitor bank is divided into 32 identical columns (5 bits) composed of 8 progressively larger capacitors (3 bits). Fig. 3-5 shows simulations of the frequency step sizes versus digital word  $N_F$ . Without predistortion (i.e. using linear  $C_F$ ), the frequency steps are small for low values of  $N_F$  and much larger for high values of  $N_F$ . This means that extra bits would be required to achieve a desired frequency resolution while covering the desired frequency range. If a perfectly predistorted capacitor bank were used, the frequency steps would all be equal, but the implementation would be impractical. As a compromise, a piece-wise linear (PWL) approximation, which is much easier to implement, is used to bound the frequency step sizes well within one least significant bit.

By building a mathematical model in MATLAB, the values for all four capacitor banks could be easily recalculated to include the effects of parasitics, resulting in fewer iterations of lengthy SPICE simulations.

### 3.3.4 Setting $f_C$ and $\Delta F$

The carrier frequency and frequency deviation constant are set in two steps. First,  $N_{\Delta F}$  is set to zero and the DCO frequency is calibrated to  $f_{DCO} = f_c - \Delta F$  using  $N_C$ ,  $N_M$ , and  $N_F$ . Then the DCO frequency is set to  $f_{DCO} = f_c + \Delta F$  using  $N_{\Delta F}$  and the value is stored in a register. To perform FSK modulation, each bit of  $N_{\Delta F}$  is applied to one input of an *and* gate, and unipolar nonreturn-to-zero modulation data  $m_u(t) \in [0, 1]$  is applied to the other input as shown in Fig. 3-1. The resulting instantaneous DCO frequency is

$$f_{DCO}(t) = f_C - \Delta F + 2\Delta F \times m_u(t) \in [f_C \pm \Delta F] \quad (3.4)$$

which is equivalent to (3.2).

## 3.4 Super-Regenerative Receiver

Using the loop antenna as the inductive element in the DCO has the benefits of reducing system complexity and power consumption for the transmitter. A challenge that arises, however, is that classical homodyne or superheterodyne receiver architectures cannot be used. In this section, the super-regenerative receiver (SRR) is introduced as an excellent low-power alternative that allows a direct connection between the antenna and the oscillator.

### 3.4.1 Super-Regeneration Theory

A thorough explanation of super-regeneration theory is given in Chapter 2. For convenience, we present a brief review here. Super-regeneration was first introduced in the 1920s and has had sporadic popularity in low power systems since then gaining

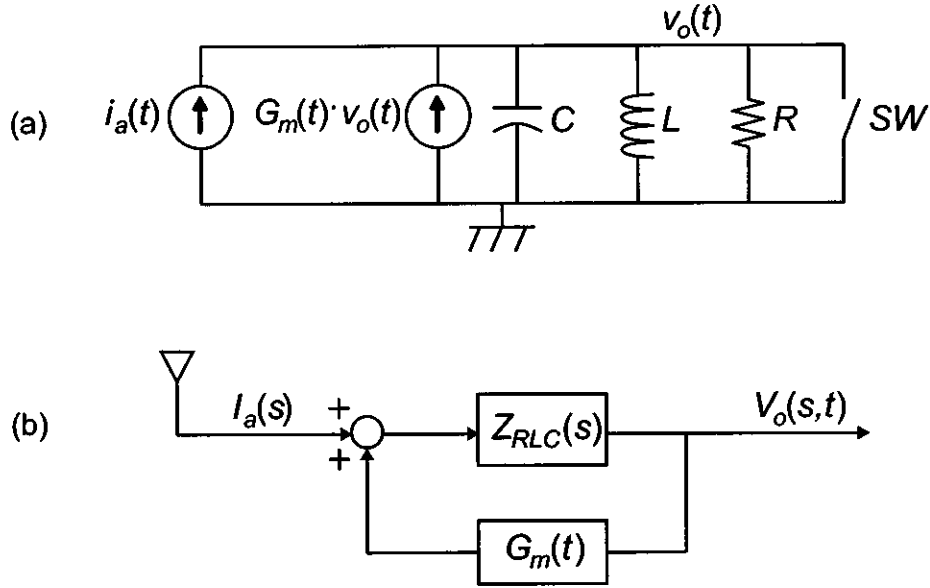


Figure 3-6: (a) SRR circuit model, and (b) SRR feedback loop model.

renewed interest in recent years [5, 12, 21, 43, 44, 48, 66]. The SRR in its simplest form comprises a resonator, time-varying positive feedback, and an input current. Fig. 3-6(a) shows a simplified SRR circuit with (b) its corresponding feedback loop model. The combination of the Laplace variable  $s$  and time variable  $t$  is acceptable only because the rate at which  $G_m(t)$  is varied is much slower than  $\omega_0$ , the resonant frequency of  $Z_{RLC}$  (i.e. the system is quasi-static). The current  $i_a(t)$  is induced by the antenna and serves as the input to the system. Using this model, the time-varying transfer function of the system can be written as

$$Z_{TV}(s, t) = \frac{V_o(s, t)}{I_a(s)} = \frac{Z_0 \omega_0 s}{s^2 + 2\zeta(t) \omega_0 s + \omega_0^2}, \quad (3.5)$$

where  $Z_0 = \sqrt{L/C}$  is the characteristic impedance of the resonant tank,  $\zeta_0 = 1/2Q_0$  is the quiescent damping factor, and

$$\zeta(t) = \zeta_0(1 - G_m(t)R) = -\beta t \quad (3.6)$$

is the damping function. This transfer function shows that the SRR is a second order system with time-varying poles. By changing  $\zeta(t)$ , the poles are periodically shifted

from the left-hand-side of the complex plane to the right-hand-side, forcing the system to become temporarily unstable. The resulting output is a growing oscillation whose amplitude depends on the characteristics of the input current  $i_a(t)$ .

A thorough and general solution to the differential equation describing this system can be found in [44]. For the specific case of the ramp damping function described by (3.6), and a sinusoidal input

$$i_a(t) = I_a \sin(\omega_a t + \phi_a) \quad (3.7)$$

the output voltage during each cycle is

$$v_o(t) = I_a Z_0 \sqrt{\frac{\pi}{2}} \frac{\omega_a}{\Omega_s} \cdot e^{-\frac{(\omega_0 - \omega_a)^2}{2\Omega_s^2}} \cdot e^{\frac{t^2}{2\sigma_s^2}} \sin(\omega_0 t + \phi_a). \quad (3.8)$$

The output voltage is a quickly growing oscillation that is proportional to the amplitude of the input signal  $I_a$ . The two exponential terms describe the filtering and time-dependent gain of the SRR. They are functions of  $\Omega_s$  and  $\sigma_s$ , the SRR frequency and time constants, defined as

$$\Omega_s = \frac{1}{\sigma_s} = \sqrt{\omega_0 \beta}. \quad (3.9)$$

The first exponential in (3.8) describes the Gaussian filtering quality of SRRs using a ramp damping function. The second exponential in (3.8) describes the tremendous time-dependent gain provided by SRRs.

Equation (3.8) is found using the assumption that the ramp damping function extends from  $-\infty$  to  $\infty$ . In reality, the ramp only extends from some time  $t_a$  to  $t_b$  and the process starts all over again resulting in a saw-tooth damping function. As long as  $t_a \ll -\sigma_s$  and  $t_b \gg \sigma_s$ , however, (3.8) is accurate because system is only sensitive to input currents during a narrow time window centered about the instant when  $\zeta(t) = 0$ . The rate at which this cycle is repeated is often referred to as the *quench frequency* because the output signal is usually quenched at the end of each cycle so that it does not affect the following cycle. The implementation presented here



includes a CMOS switch placed across the resonator used to quench all oscillations at the end of every cycle ( $SW$  in Fig. 3-6(a)).

### 3.4.2 Receiver System Design

To maximize the bit rate of the receiver, the SRR is used to synchronously receive OOK data as in [43]. To ensure proper operation, the quench oscillator and the incoming data must be synchronized so that the instant when  $\zeta(t) = 0$  occurs near the center of each bit period. As described in Section 3.6, the base-station has the task of synchronizing its baseband clock to the implant's baseband oscillator to reduce complexity in the implant.

As shown in Fig. 3-1, the SRR is followed by a fully differential envelope detector, a programmable comparator, a digital counter, and a digital comparator. At the start of each bit period (or quench cycle), the digital counter is reset and immediately begins to count. Receiving a *one* is equivalent to receiving a sinusoidal input with an oscillation frequency that is equal to the SRR's resonant frequency (i.e.  $\omega_a = \omega_0$ ). Passing the output of the SRR through an envelope detector results in the envelope signal

$$v_e(t) = I_a Z_0 \sqrt{\frac{\pi \omega_0}{2 \Omega_s}} \cdot e^{\frac{t^2}{2\sigma_s^2}}, \quad (3.10)$$

where  $I_a$  is the peak amplitude of the input current. The output of the envelope detector is connected to a comparator that has a programmable offset  $V_{OS}$ . When the amplitude of the envelope becomes larger than  $V_{OS}$ , its output changes states and disables the counter causing it to hold its value. This final value is a measure of the SRR's startup time (the time required for the SRR's envelope to reach  $V_{OS}$ ). For the receiver to have a low bit-error rate (BER), the amplitude of the current  $I_a$  must be significantly larger than the RMS value of equivalent input referred noise sources. This means that when a *one* is received, the SRR's startup time is faster and the counter's final count is lower. When a *zero* is received, the startup time is slower resulting in a higher count. In this manner, the counter functions as a time-integrating analog-to-digital converter that measures the startup time of the SRR

which is a function of the input signal's amplitude. At the end of the bit period, the final count is compared with a digital threshold number  $NT$ , and a decision is made as to whether a *one* or a *zero* was received.  $NT$  is set by averaging the counter output over a 32-bit preamble with an equal number of *ones* and *zeros*.

Note that the counter was implemented in an FPGA and is clocked by an external oscillator, but could easily be integrated and clocked using an amplified version of the SRR's output. In such an implementation, the counter would start counting once the SRR's output envelope reached a threshold and the final count would be higher when a *one* was received. The performance of the receiver would be the same since the important parameter is the difference between the counter's output when a *one* is received versus a *zero*. Also, the power consumption of the digital blocks would be minimal due to their simplicity.

When a *zero* is received, the output of the SRR is also a growing oscillation similar to (3.8) except its amplitude is a random variable with an RMS value that depends on the noise density of active and passive noise sources in the system along with their effective noise bandwidths. The sensitivity of the receiver, therefore, is optimized by reducing the amplitude of the noise sources and the bandwidth of the SRR which is proportional to  $\Omega_s$ . Since  $\omega_0$  is determined by the standard used (i.e.  $\omega_0 \approx 2\pi \times 400\text{MHz}$  for MICS), the only design variable that can be altered to reduce the noise bandwidth is  $\beta$  (the slope of the damping function). The tradeoff, however, is that reducing  $\Omega_s$  is equivalent to increasing  $\sigma_s$  which necessitates longer quench periods since the condition  $|t_{a,b}| \ll \sigma_s$  must be kept. This results in a tradeoff between sensitivity and bit rate.

A concern that arises from having the SRR directly connected to the antenna is that the receiver actually radiates power. In architectures such as super-heterodyne receivers this is undesirable since the radiated signals can interfere with other receivers. For an SRR, however, the radiated power occupies its own channel and does not interfere with other receivers. Furthermore, the radiated power is well below the  $25\mu\text{W}$  allowed for transmission and does not violate spectral mask requirements.

## 3.5 Circuit Implementation

The full transceiver is shown in Fig. 3-1, where a portion of the baseband section was implemented in an FPGA for maximum flexibility. The receiver comprises five major blocks: a DCO/SRR, a quench/baseband oscillator, a fully differential envelope detector, a comparator with programmable offset, and a digital counter. Exploiting the high gain provided by the SRR allows all analog blocks to be biased in subthreshold for low-power operation.

### 3.5.1 DCO Implementation

Fig. 3-7(a) shows a simplified schematic of the differential Colpitts oscillator with switching current source similar to [4]. The loop antenna is shown as two inductors and resistors to facilitate understanding. Fig. 3-7(b) shows the equivalent small-signal model of the half-circuit. The transconductance is twice  $G_m$  because two PMOS transistors provide positive feedback in each half-circuit. Fig. 3-7(c) shows a simplified version of the model where the current source and  $1/G_m$  resistor are placed in parallel with the antenna using capacitive impedance transformation [35]. The minimum value of  $G_m$  that will result in oscillation is

$$G_{m0} = \frac{2}{R_p n(2 - n)}, \quad (3.11)$$

where  $n$  is the impedance transformation ratio

$$n = \frac{C_1}{C_1 + C_2} \quad (3.12)$$

As shown in (3.6), the damping function  $\zeta(t)$  is controlled by varying  $G_m(t)$ , and  $\zeta(t) = 0$  occurs at the instant when  $G_m(t) = G_{m0}$ . Using a Colpitts topology requires a larger  $G_{m0}$  than the more common cross-coupled transistor topologies, but offers a noise benefit. Fig. 3-7(d) shows the equivalent combined noise source from the two PMOS transistors in each half-circuit, and Fig. 3-7(e) shows the effective noise source

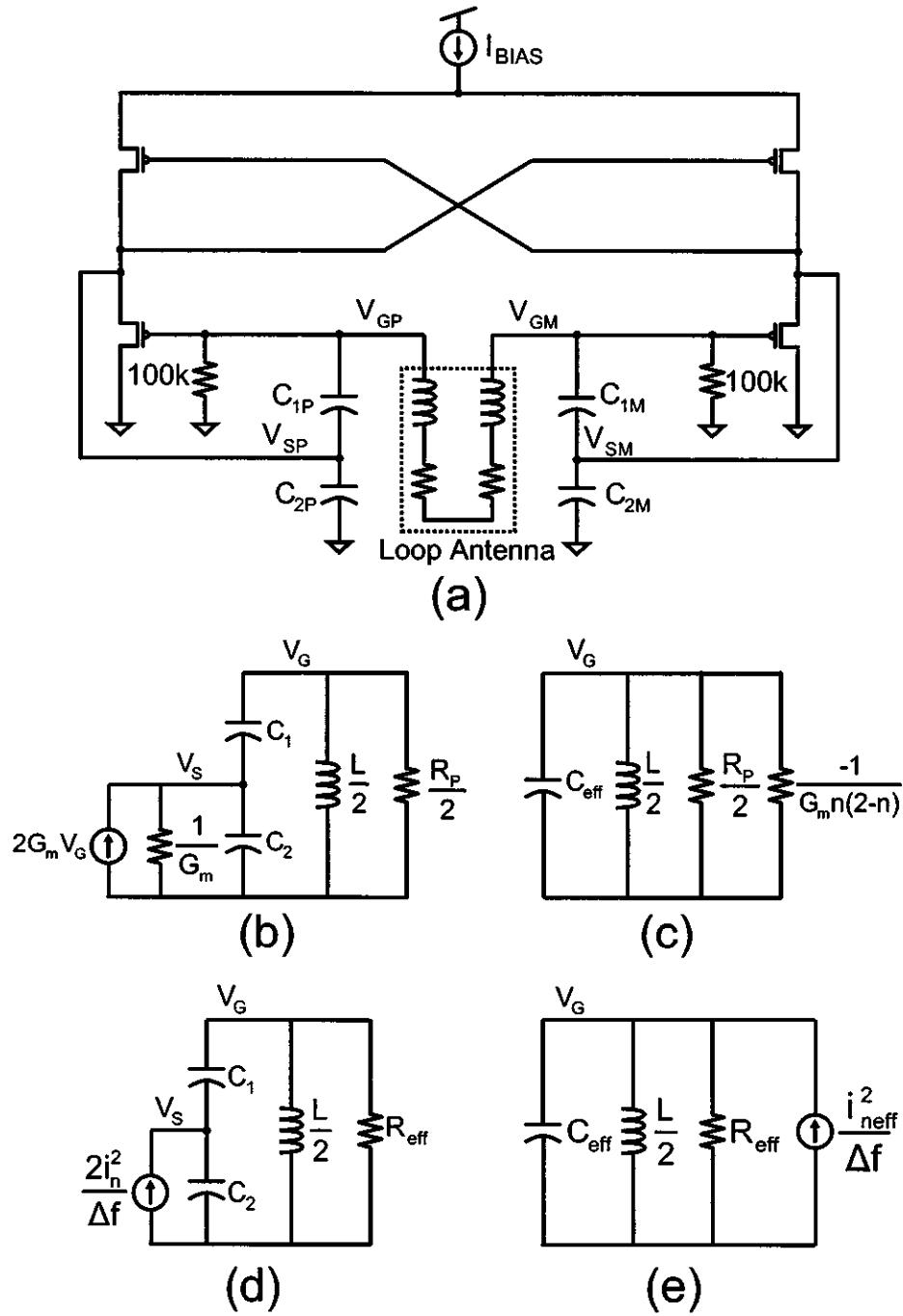


Figure 3-7: (a) Differential Colpitts oscillator, (b) half circuit model, (c) equivalent half circuit, (d) noise model, and (e) equivalent noise model.

in parallel with the resonator having a value of

$$\frac{i_{neff}^2}{\Delta f} = \frac{2n^2 i_n^2}{\Delta f}. \quad (3.13)$$

Since  $n < 1$ , the impedance transformation reduces the impact of transistor noise sources. The effective noise can be derived as

$$\frac{i_{neff}^2}{\Delta f} = \frac{4kT}{R_p} \left( \frac{n}{\kappa(2-n)} \right), \quad (3.14)$$

when the transistors are biased in subthreshold, where  $\kappa$  is a device parameter roughly equal to 0.7. By making  $n$  small, the effective noise from the transistors can be made much smaller than the noise from the passive components, improving the sensitivity of the receiver.

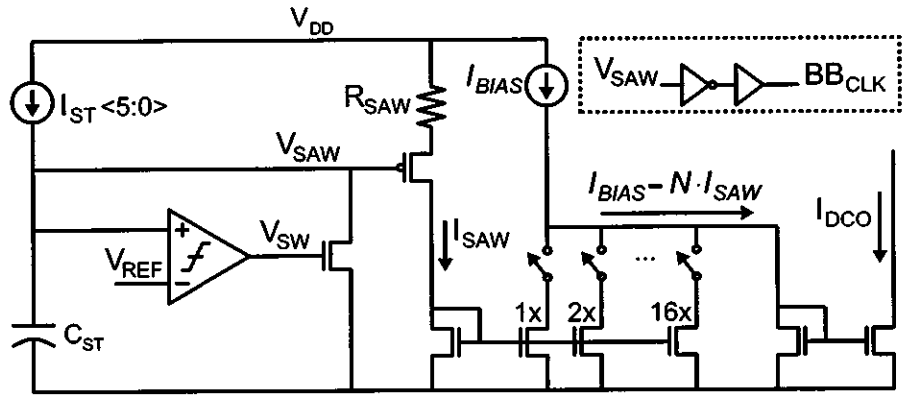
The differential Colpitts oscillator also has some power transfer advantages for the transmitter. By using a tapped resonator, the peak voltage across the antenna is not restricted to the power supply. This means that a lower supply voltage can be used while still delivering sufficient power to the antenna. As shown in [4], the cross-coupled transistors perform current switching once the signal amplitude is large, resulting in twice as much efficiency delivering power to the antenna.

The full DCO with predistorted capacitor banks is shown in Fig. 3-8. By dividing the total capacitance into coarse, medium, and fine tuning capacitor banks, more than 14 bits of frequency resolution are achieved. Capacitor bank predistortion leads to improved digital-to-frequency conversion, and thermometer coding guarantees monotonicity for each bank. The differential Colpitts topology results in improved sensitivity for the receiver and higher power transfer capabilities for the transmitter.

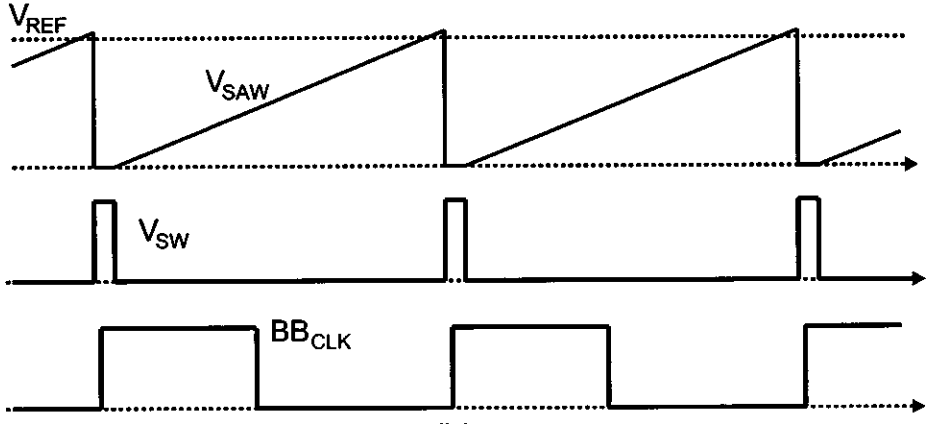
### 3.5.2 Programmable Ramp Quench Oscillator

Fig. 3-9(a) shows a simplified schematic of the sawtooth oscillator used to generate the quench signal for the SRR and the baseband clock. The sawtooth voltage waveform  $V_{SAW}$  results from capacitor  $C_{ST}$  integrating the current  $I_{ST}$  over time. Once  $V_{SAW}$





(a)



(b)

Figure 3-9: (a) Sawtooth oscillator, SRR bias generator, and baseband clock. (b) Output signals.

exceeds  $V_{REF}$ , a comparator turns on an NMOS switch that discharges the capacitor and forces  $V_{SAW}$  back to zero. This in turn causes the comparator to turn off the NMOS switch and restart the cycle. The comparator includes some delay stages that cause  $V_{SAW}$  to stay low and  $V_{SW}$  to stay high for a short period of time as shown in Fig. 3-9(b). The signal  $V_{SW}$  is used to drive the switch that briefly shorts the resonant tank of the DCO as shown in Fig. 3-6(a).  $V_{SAW}$  is passed through a pair of inverters to create a square wave with a duty cycle of roughly 50% that is used to clock the baseband circuits. The frequency of oscillation is

$$f_{BB} \approx \frac{I_{ST}}{V_{REF}C_{ST}}, \quad (3.15)$$

and is tuned digitally using the programmable current  $I_{ST}$ . The voltage waveform  $V_{SAW}$  is converted to a current  $I_{SAW}$  using a degenerated common-source PMOS transistor.  $I_{SAW}$  is then used as the reference for a multiplying current-mode digital-to-analog converter (IDAC) which allows for digital tuning of the quench signal's slope. The output of the IDAC is subtracted from the bias current  $I_{BIAS}$  and mirrored to create the bias current of the SRR. In transmit mode, the IDAC is disabled, and the DCO bias current is  $I_{BIAS}$ .

### 3.5.3 Envelope Detector

The task of the envelope detector is to create a one-to-one mapping between the SRR's output amplitude and the envelope detector's output voltage. This can be done by squaring the output of the SRR (using a Gilbert cell) and taking the average (using a low-pass filter). The Gilbert cell shown in Fig. 3-10(b) is biased in subthreshold at a very low bias current ( $10\mu\text{A}$ ), yet its input-referred noise has a negligible effect on the receiver's performance due to the high gain provided by the SRR. It is DC-coupled to the DCO/SRR in Fig. 3-8 and uses  $60\text{k}\Omega$  load resistors resulting in a common-mode output voltage of 300mV. A pair of 1pF capacitors are used to filter the  $2\omega_0$  term.



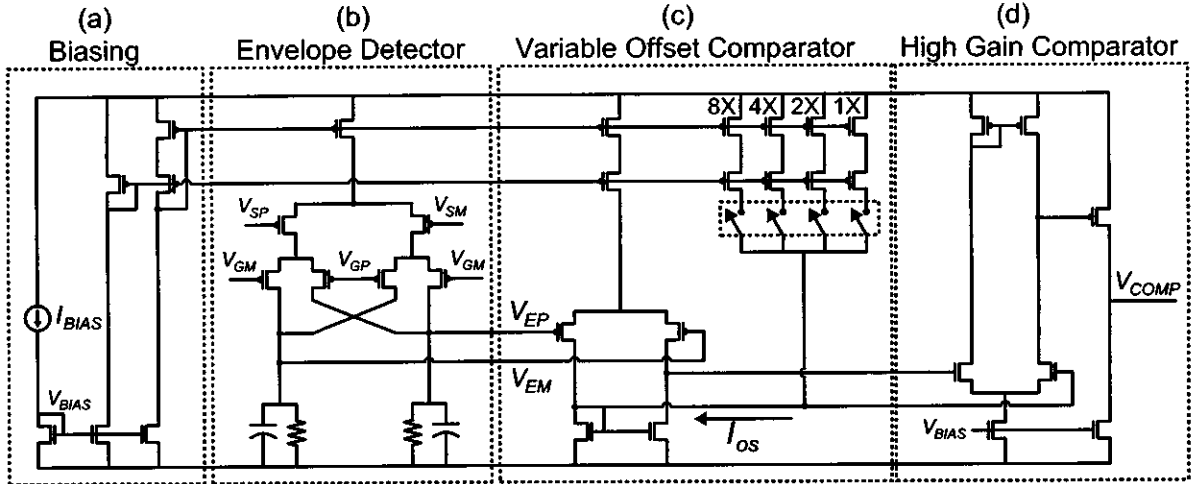


Figure 3-10: Schematic for differential envelope detector and comparator with programmable input offset voltage.

### 3.5.4 Programmable Comparator

The programmable comparator following the envelope detector in Fig. 3-1 has a digitally configurable offset that sets the differential input voltage threshold at which the comparator switches states. The offset is set by an IDAC and allows a fully differential connection between the envelope detector and the comparator, reducing the effects of common-mode noise. The programmable comparator comprises a low gain, variable offset comparator (Fig. 3-10(c)) and a two-stage op-amp used as a high gain comparator (Fig. 3-10(d)).

The input offset is created by feeding current into the low-impedance arm of the current mirror in (c). Since all of the circuits in Fig. 3-10 are biased in subthreshold, the transconductance of the input differential pair is

$$I_E = I_{EP} - I_{EM} = I_{DD} \tanh\left(\frac{\kappa V_E}{2\phi_t}\right). \quad (3.16)$$

$I_{DD}$  is the tail current of the differential pair and is set to  $40 \times I_{BIAS}$ , and  $\kappa$  is a process parameter approximately equal to 0.66 for this technology. Adding the current

$$I_{OS} = N_{OS} \times I_{BIAS}, \quad (3.17)$$

to the low impedance arm of the current mirror, creates the input voltage offset

$$V_{OS} = V_E = \frac{2\phi_t}{\kappa} \tanh^{-1} \left( \frac{1}{40} N_{OS} \right). \quad (3.18)$$

Since the maximum value of  $N_{OS}$  is 15,  $V_{OS}$  can be approximated at room temperature as

$$V_{OS} \approx \frac{2\phi_t}{\kappa} \times \frac{N_{OS}}{40} \approx 2.0mV \times N_{OS} \quad (3.19)$$

using the approximation  $\tanh^{-1}(x) \approx x$ .  $V_{OS}$  is proportional to absolute temperature and independent of  $I_{BIAS}$  as long as subthreshold operation is maintained. The comparator was designed to consume only  $10\mu W$ .

### 3.6 Prototype Frequency Correction Loop

Modern transceivers typically use frequency synthesizers to phase-lock a voltage-controlled oscillator (VCO) to a very stable reference such as a crystal oscillator. For ultra-low power medical implants, however, a frequency synthesizer may not be necessary for two reasons: first, the frequency stability requirement for MICS ( $\pm 100ppm$ ) is far more lax than typical systems and second, the human body provides excellent temperature regulation reducing both the amount of overall frequency drift and its rate of change. Instead of using a phase-lock loop, a frequency correction loop is proposed that pushes complexity from the implant to the base-station. In the proposed system, the base-station monitors the DCO's oscillation frequency and sends information to the implant to correct frequency errors. A crude implementation of such a system was prototyped using a spectrum analyzer, a PC, and an FPGA. The spectrum analyzer is used to measure the frequency error, and the information is sent to the PC. A simple MATLAB script determines what the new DCO settings ( $N$  values) should be and sends the information directly to the implant device using an FPGA. Fig. 3-11 shows the values of  $N_C$  and  $N_F$  for three scenarios ( $N_M$  did not change).

The first example shown in Fig. 3-11(a) uses the frequency calibration loop to

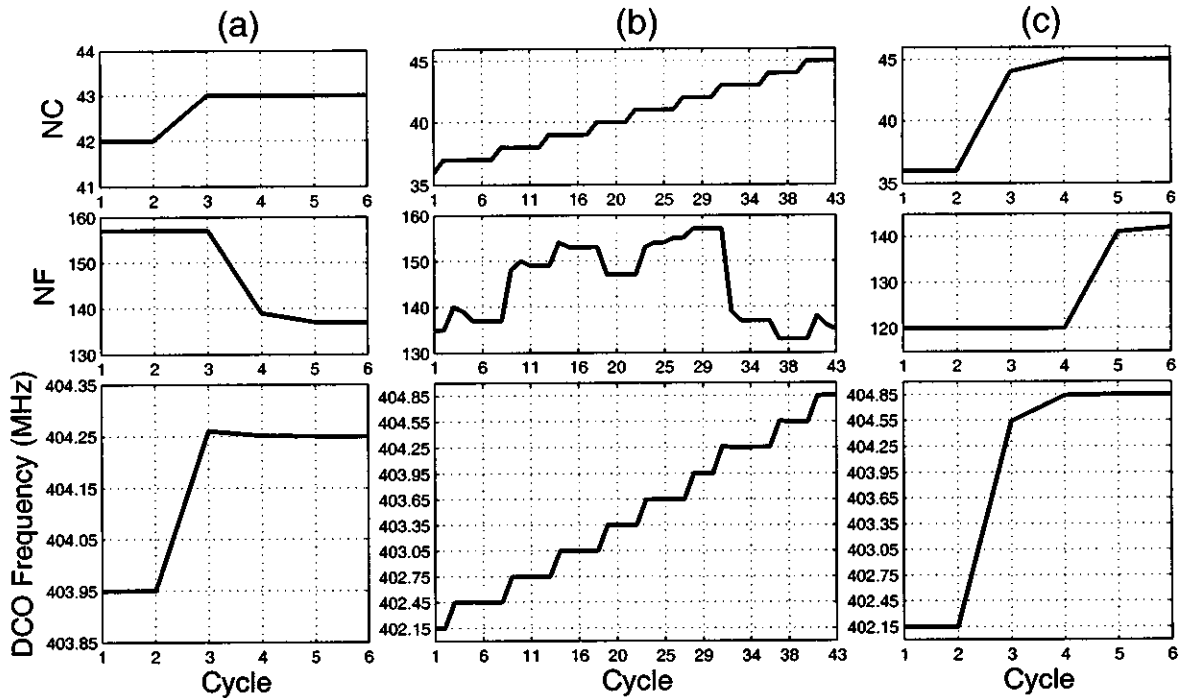


Figure 3-11: Measurements of the frequency correction loop including the transition of the DCO frequency, coarse tuning coefficient, and fine tuning coefficient as the channel is changed from (a) 7 to 8, (b) 1 to 10 incrementally, and (c) 1 to 10 in a single command.

jump from channel 7 to channel 8. The implant was initially set to 403.95MHz and a command was sent to change the frequency to 404.25MHz. First, the MATLAB script uses estimates of the coarse frequency step size to determine the correct value for  $N_C$  and the FPGA sends this information to the transceiver through a SPI bus. The spectrum analyzer determines the new DCO frequency and sends it to the PC. Since the new frequency error is small, a new value of  $N_F$  is sent to the transceiver and the cycle repeats until the frequency error is below a threshold of 1kHz. After only five cycles, the final frequency is 404.2499MHz (within 100Hz of the desired frequency). Fig. 3-11(b) shows the progression of  $N_C$ ,  $N_F$ , and the DCO frequency as the frequency control loop is used to change from one channel to the next. Finally, Fig. 3-11(c) shows the results from changing from channel 1 to channel 10 directly.

In all three of these examples, the frequency correction loop was used to correct large frequency errors and was able to do so within very few steps because of the

fairly linear digital-to-frequency relationship of the DCO. Furthermore, because the digital-to-frequency relationship of the DCO is monotonic, convergence was always achieved over many measurements. Typical frequency corrections would actually be much smaller than the examples shown (on the order of 1-50kHz), and even with this simple implementation, such corrections would require three cycles or less.

Calibrating the baseband oscillator is done similarly so that the SRR can synchronously demodulate OOK data. To do so, the implant transmits a short preamble prior to receiving data and the base-station uses it to lock to the implant. If the baseband clock frequency drifts beyond a limit, the transmitter sends a calibration command to correct it. Shifting the task of synchronization to the base-station results in power savings in the implant.

As explained in Section 3.2.2, some frequency drift can result from human motion and correcting it requires that calibration be repeated every 100ms. At a bit rate of 120kbps and assuming 128 bits are needed for each calibration cycle, the energy overhead is just over 1%. As a result, by taking advantage of the lax frequency stability requirements of MICS, the temperature regulation of the human body, and the robustness of the SRR to minor frequency mismatches, the frequency synthesizer can be completely replaced by a simple frequency correction loop that greatly reduces the power consumption of the implanted device.

### 3.7 Measurement Results

The transceiver was fabricated in 90nm CMOS and consumes a total of 0.5mm<sup>2</sup> of active area as shown in Fig. 3-12. Fig. 3-13 shows the coarse, medium, and fine tuning digital-to-frequency curves. The downward concavity of the tuning curves is due to a slight overcompensation in the predistortion resulting from unaccounted parasitics. The DCO tunes from 391–415MHz, and has an average frequency resolution of 590Hz with a maximum fine-tuning step size of 1.4kHz achieving 14 effective bits of resolution. The resolution is calculated as  $\log_2(24\text{MHz}/1.4\text{kHz}) = 14.1$  bits. Only 3MHz of tuning range is required by the MICS standard, but a 24MHz range is used to account

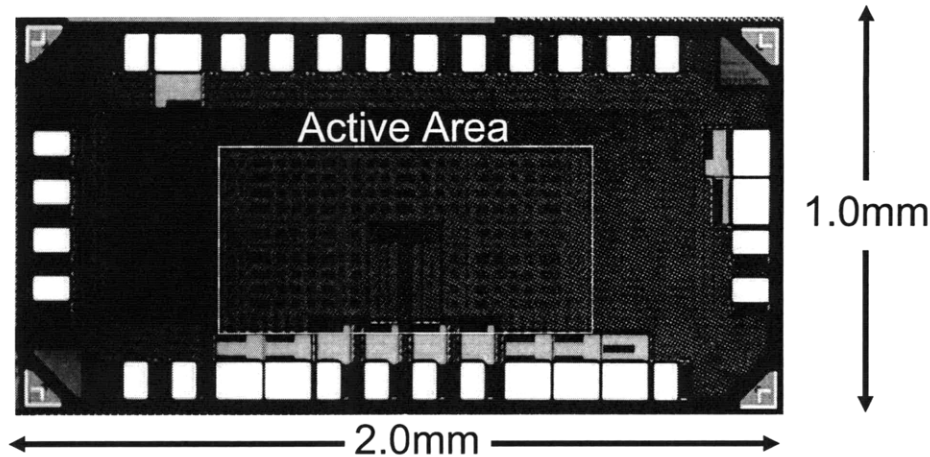


Figure 3-12: Die photograph ( $1.0 \times 0.5\text{mm}$  active area).

for process and temperature variations. Fig. 3-13(b) shows that  $C_M$  tunes the DCO 2.96MHz which is well above the maximum coarse frequency step (450kHz) providing ample overlap. The  $C_M$  tuning curve is plotted for the maximum and minimum temperature requirements and shows an average frequency shift of 923kHz over the entire temperature range, corresponding to a temperature coefficient of 46.2kHz/°C (115ppm/°C). Since the frequency stability requirements of MICS call for frequency stability of  $\pm 100\text{ppm}$ , the transceiver can tolerate temperature drifts of up to 0.87°C without calibration. Correcting for such variations should be simple with a calibration rate of 10 calibrations per second. However, if the implant is in sleep mode for a long period of time, it may drift beyond the 100ppm requirement. This may lead to a brief violation of the spectral mask and frequency accuracy requirements, but there is a very small possibility it would interfere with other implants due to the highly duty cycled nature of their operation.

### 3.7.1 Transmitter

Fig. 3-14(a) shows the output spectral mask of the transmitter measured through an antenna placed 10cm away from the device. The first side lobe is roughly 6dB higher than the theoretical value and violates the MICS spectral mask by 1.1dB at a bit rate of 200kbps. This could easily be corrected with a very simple and low-

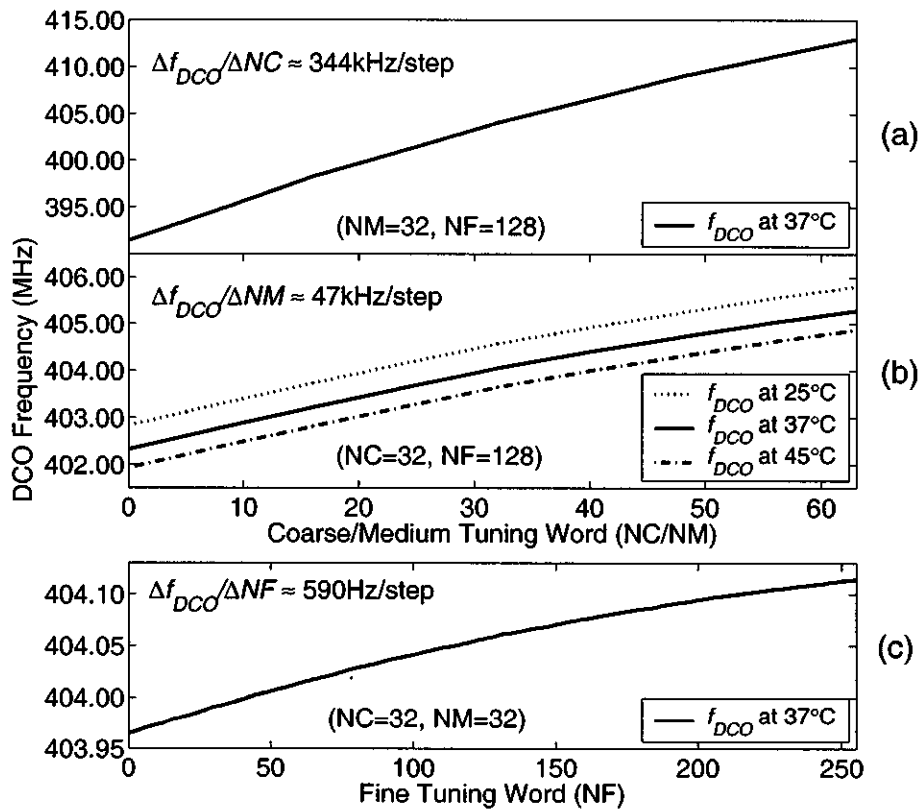


Figure 3-13: DCO frequency for (a) coarse, (b) medium, and (c) fine tuning. Thermometer coding and predistortion of each capacitor bank lead to monotonic, linear tuning for each range.

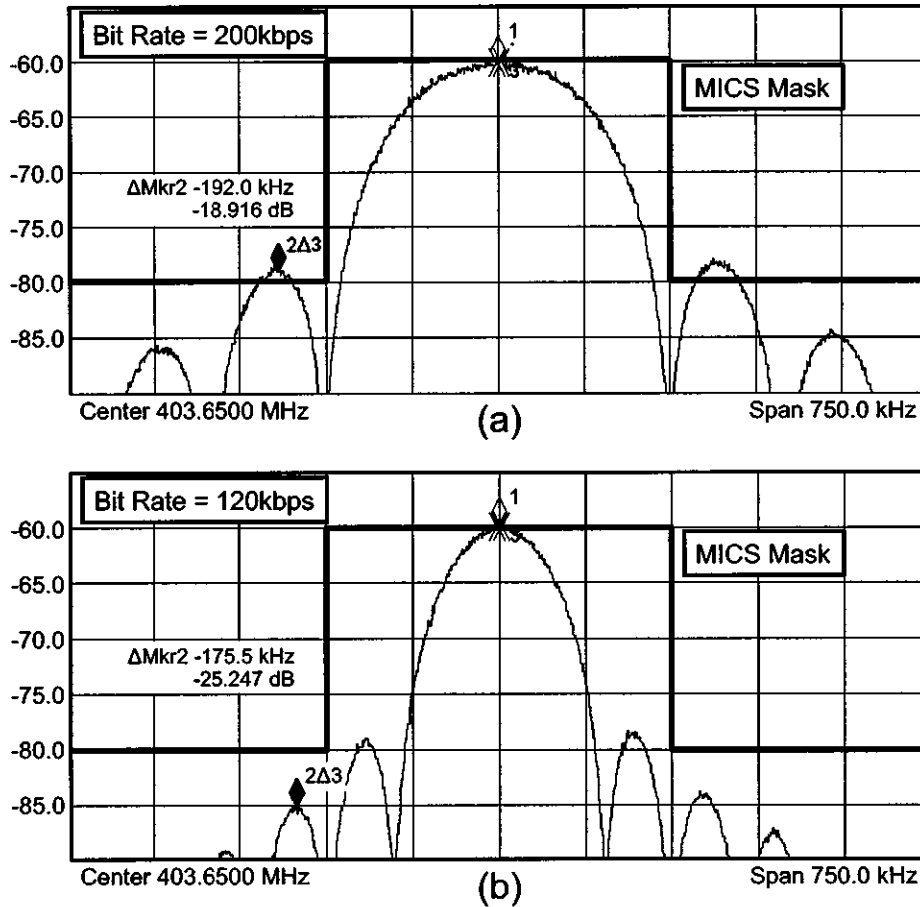


Figure 3-14: Measured spectral mask of transmitter with a data rate of (a) 200kbps and (b) 120kbps. The signal is taken through an antenna 10cm from the transmitter.

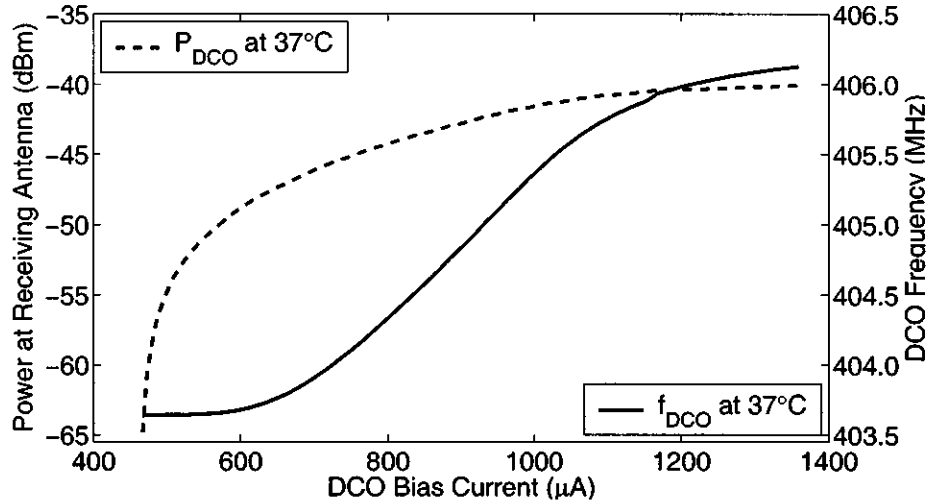


Figure 3-15: Power received at an antenna placed 20cm from the transmitter and DCO frequency vs. bias current.

power digital filter connected to the modulation capacitor bank  $C_{\Delta F}$ . At a bit rate of 120kbps and without pre-filtering, the MICS spectral mask requirement is met with 5dB of headroom as shown in Fig. 3-14(b). Fig. 3-15 shows the power received at an antenna placed 20cm from the transmitter and the DCO frequency as a function of bias current ( $V_{DD}=700\text{mV}$ ). As the DCO bias current is swept from  $450\mu\text{A}$  to  $600\mu\text{A}$ , the power at the receiving antenna varies from  $-65\text{dBm}$  to  $-48\text{dBm}$  with minimal variation in the DCO frequency. This ability to trade off power consumption for output power, means that as little as  $315\mu\text{W}$  of power can be consumed by the DCO when the path loss between the antennas is small. If the path loss is high, the base-station can send a command to the implant to increase the output power up to 17dB without considerable frequency drift. Furthermore, the ability to control both the frequency and amplitude of the DCO opens the possibility of employing more spectrally efficient modulation techniques such as polar modulation. Including all circuits, the transmitter consumes less than  $350\mu\text{W}$  under normal operation.

### 3.7.2 Receiver

Fig. 3-16 shows measured time-domain signals for the receiver chain. As expected, when a *one* is received, the startup time of the DCO/SRR is faster. When the



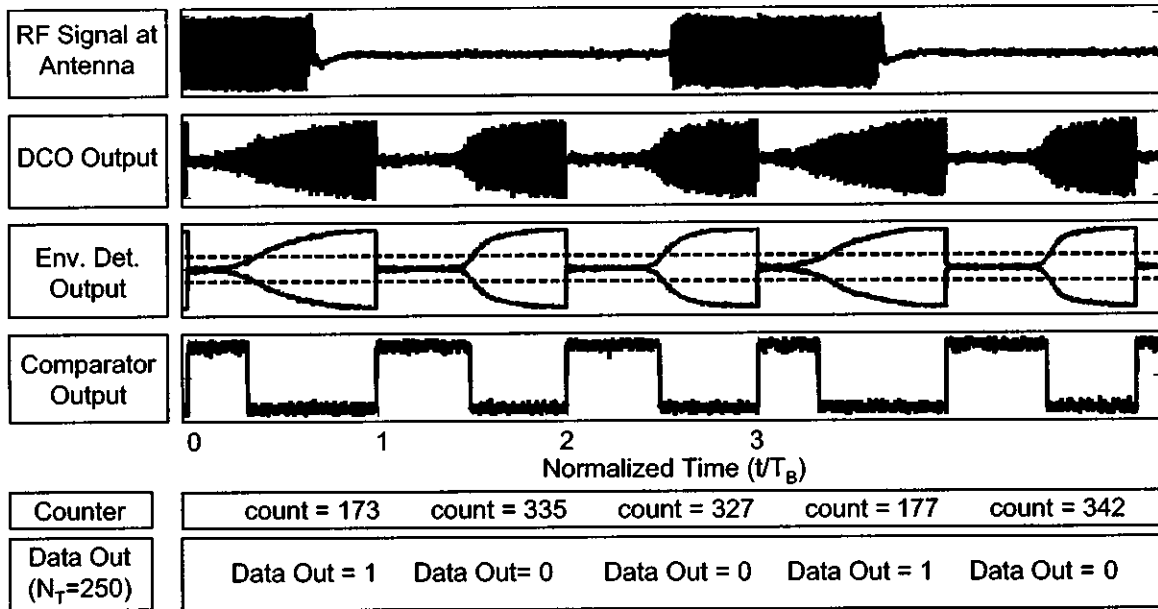


Figure 3-16: Receiver chain time domain signals.

envelope crosses the threshold voltage of the comparator, the counter is disabled and the final count is compared to the threshold  $NT$  (250 in this case). If the final count is smaller than this threshold, the decision is made that a *one* was received and vice-versa.

Since the transceiver was not designed to interface with a  $50\Omega$  source or load, measuring sensitivity was challenging. To do so, a matching network was connected to each side of the resonator. At resonance, where the receiver works, the input impedance is real, so the matching network is meant to step up the impedance of a signal generator from  $50\Omega$  to roughly  $2k\Omega$ . Each matching network comprised a coupling capacitor, shunt capacitor, series inductor, and shunt capacitor connected to an SMA connector. The network had a negligible effect on the resonant frequency, but loaded the DCO requiring more bias current. A signal source was connected to a power splitter with two outputs that were each connected to one of the matching networks. The required DCO bias current increased from  $450\mu A$  to  $650\mu A$  due to the additional loading.

For bit rates of 40kbps and 120kbps, the measured sensitivity was -99dBm and -93dBm respectively (BER=0.1%), showing how bit rate can be traded off for sensitiv-

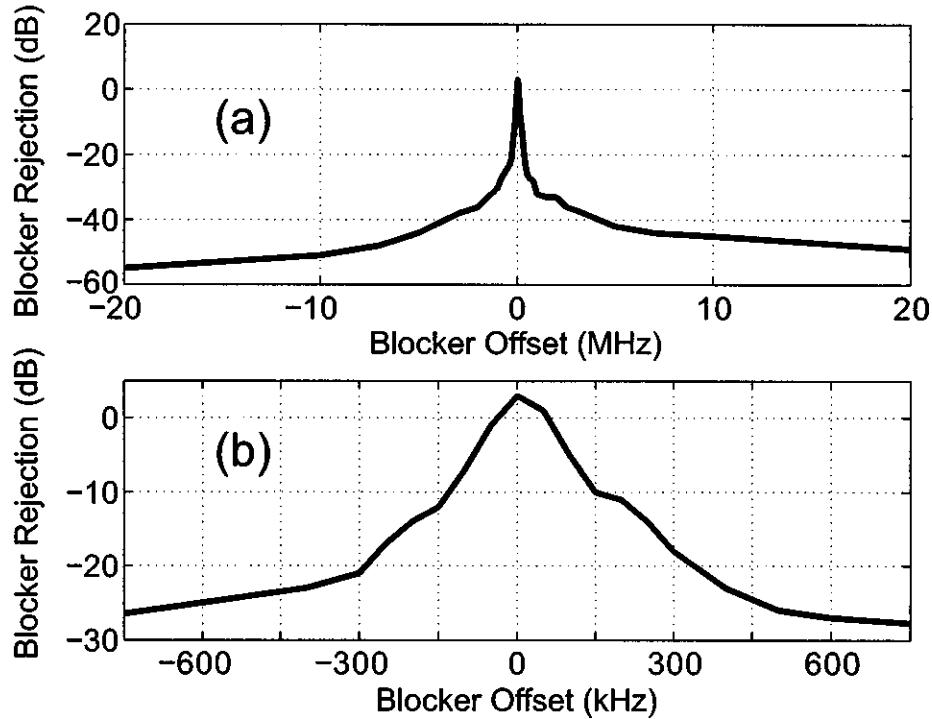


Figure 3-17: (a) Far off and (b) close in CW blocker rejection. Input signal level set to sensitivity + 6dB = -93dBm with a carrier frequency of 403.35MHz and a bit rate of 40kbps.

ity as discussed in Section 3.4. Without the matching networks, the power consumption in both cases was under  $400\mu\text{W}$  for the entire system resulting in an excellent energy-per-received-bit figure-of-merit  $3.3\text{nJ/bit}$  for a bit rate of 120kbps.

Fig. 3-17 shows the receiver's immunity to (a) far off and (b) close in blockers. For this measurement, the bit rate was set to 40kbps, the input signal to the SRR was set to sensitivity + 6dB (-93dBm), and a the power of a CW blocker was swept for each frequency until the BER was degraded to 0.1%. The receiver shows excellent selectivity, with the ability to reject far off blockers by almost 60dB and close in blockers by at least 10dB. For example, at the center frequency of the adjacent channel (300kHz offset from the desired channel), the receiver achieves more than 20dB of rejection and at least 27dB of rejection for the second adjacent channel (600kHz away). Table 3.1 summarizes the transceiver performance.

Table 3.1: Summary of Measured Transceiver Performance

RF Frequency Range	391–415MHz
Frequency Resolution	$\leq 1.4kHz$
Rx Modulation	Sync. OOK
Rx Sensitivity @ 40kbps	-99dBm
Rx Sensitivity @ 120kbps	-93dBm
Rx Power Consumption	$400\mu W$
Rx Energy per Bit @ 40kbps	3.3nJ/bit
Rx Energy per Bit @ 120kbps	10nJ/bit
Tx Modulation	MSK
Tx Bit Rate	120kbps
Tx Power Consumption	$350\mu W$
Tx Energy per Bit	2.9nJ/bit

### 3.8 Summary

A transceiver optimized for the MICS standard was presented that exploits unique features of medical implants to simplify its architecture and reduce power consumption. The central block in the transceiver is a digitally-controlled oscillator that uses the loop antenna as its inductive element and a predistorted and sub-ranged capacitor array to improve linearity in digital-to-frequency conversion. The DCO has a differential Colpitts topology that improves power transfer to the antenna and reduces the effects of active device noise sources. It is used as a super-regenerative receiver with excellent sensitivity and selectivity performance for minimal power consumption. The transmitter uses direct MSK modulation to meet the MICS spectral mask requirements with relatively high data rates for such a simple topology. The transmitter and receiver achieve an energy-per-received-bit of 2.9nJ/bit and 3.3nJ/bit respectively while meeting the 300kHz channel bandwidth requirements of MICS. Finally, a frequency correction loop was presented that eliminates the need for a frequency synthesizer in the implant, further reducing power consumption, device size, and system cost.



## Chapter 4

# Digitally-Assisted Sensor Interface for Biomedical Applications

As far back as the late 18th century, scientists were able to show that certain animals could generate electric waveforms [51]. By the late 19th century, scientists were already able to measure surface potentials for EKG, EEG, and EMG although these techniques did not become clinically common until the 1940s and 50s. With the advent of solid-state devices came an avalanche of research in the development of medical devices that made them more portable and affordable, and nearly ubiquitous in clinical settings. As far back as 1958, portable/wireless EKG devices were being investigated using BJT transistors and FM radios [8], creating a pathway towards wearable and implantable medical devices.

More recently, there has been a push towards minimizing the size and power consumption of medical electronics to allow for greater functionality, and in particular, for facilitating implantable devices that are minimally invasive. A typical goal for medical implants is a battery life of more than a decade. Since the energy stored in the battery is proportional to its volume, minimizing the energy consumption of the electronics is critical to reducing the size of the overall system. While battery recharging through inductive coupling and other means is also an option, power dissipation is still a concern due to the heat dissipation of the electronics. This is particularly important in applications such as multi-electrode neural recording systems where as

many as 1000 miniature electrodes are placed near the brain and elevated temperatures could damage surrounding tissue [25].

In the previous chapter, an implantable transceiver was discussed that allows communications between sensors and actuators in the body and an outside base-station. In this chapter, we explore the design of a small, ultra-low power system that can facilitate the integration of hundreds of sensor-interfaces on a single chip. The primary focus of this work was to leverage unconventional digital and mixed-signal techniques to holistically reduce the size and power of the system. Our two main contributions were the development of a novel programmable-gain, anti-aliasing filter that is area- and energy- efficient, and a mixed-signal control system that creates a digitally-programmable notch filter for interference cancelation. The anti-aliasing filter is well suited for systems that employ chopper-stabilized instrumentation amplifiers as it places notches precisely at the chopping frequency and harmonics where opamp DC offsets are frequency translated to.

The chapter is organized as follows: Section 4.1 describes several biomedical applications that require sensors. Section 4.2 discusses the main signal aggressors that must be dealt with when designing a sensor interface, common techniques used to deal with these, and the system specifications required to properly extract the desired physiological signals. Sections 4.3, 4.4, and 4.5 describe in detail the instrumentation amplifier, sinc anti-aliasing filter, and mixed-signal interference-cancelation notch filter we developed. Section 4.6 presents measurement results of the prototype system. This work was carried out in collaboration with Marcus Yip who implemented many of the system's sub-blocks.

## 4.1 Applications

Endogenic electrical signals are used throughout the human body to communicate desired actions. For example, the central nervous system sends action potentials to muscles in the extremities through motor neurons to induce movement. Conversely, sensory organs send information to the central nervous system through sensory neu-

rons, providing an essential feedback mechanism. Other examples include electrical signals sent throughout the cardiac system to induce pumping. In this case, the sinoatrial node sends rhythmic cardiac impulses to different parts of the heart without the involvement of the central nervous system. These electrical impulses lead to the most familiar electrical waveforms to laymen: those measured in electrocardiography (EKG or ECG), where action potentials are measured at the surface of the skin in the chest using noninvasive electrodes [69].

Two other common types of measurements are electromyography (EMG), where action potentials in various muscles are measured to assess neuromuscular function, and electroencephalography (EEG), where the electrical activity of brain cells are recorded on the surface of the skull's skin. Other non-invasive techniques for measuring action potentials include electro-nystagmography (ENG), electroretinography (ERG), and electrooculography (EOG), all associated with the eye. In addition to these non-invasive techniques, some measurements require needle electrodes that can make direct contact with various types of cells to access signals that are either too small or too high in frequency and therefore attenuated by the skin. These include electrocortigraphy (ECoG), which measures neural field potentials subdurally (on the surface of the cortex), and even single cell measurements where electrodes are injected into the cell to measure larger, higher frequency signals (although these measurements typically kill the cell within minutes) [17].

Finally, there are various types of sensors that can be used to extract non-electrical physiological signals (temperature, motion, electrolyte levels, glucose, etc.) both non-invasively and through medical implants. The system developed in this work is flexible enough to accommodate the requirements for many of these applications, both implantable and non-invasive. The overall goal is to minimize the energy consumption and area of a full system that can be integrated in CMOS technology.

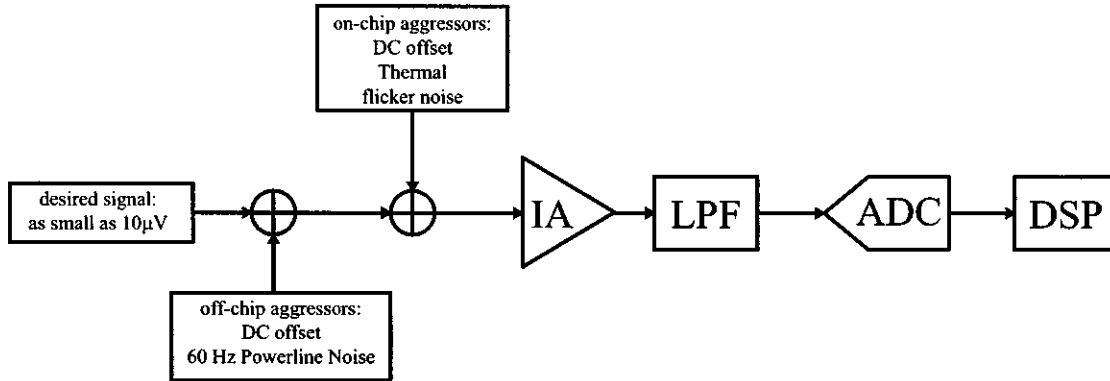


Figure 4-1: Block diagram of a typical sensor interface including the off-chip and on-chip aggressors that must be dealt with.

## 4.2 Top-Level Issues and Requirements

Fig. 4-1 shows a typical sensor interface including an instrumentation amplifier (IAMP), a low-pass filter (LPF), an analog-to-digital converter (ADC), and a digital signal processor (DSP). Many physiological signals lie within the amplitude range of  $10\ \mu\text{V} - 10\ \text{mV}$ , and reside in a frequency band between  $1\ \text{Hz} - 1\ \text{kHz}$ . Unfortunately, these signals are often corrupted by various off-chip aggressors including DC offset from the electrodes that can be hundreds of millivolts in amplitude and power-line interference (PLI) at 50 or 60 Hz, as well as on-chip aggressors such as DC offset, thermal noise, and flicker noise from the IAMP. The main task of the IAMP is to amplify the input signal without corrupting it excessively with noise so that the noise requirements of subsequent stages can be relaxed to save power. Unfortunately, aggressors including DC-offset and PLI are often significantly larger than the desired signal and within the band of interest or close to it. This limits the amount of amplification that can be used without saturating the system, or requires large supply voltages to be used increasing power consumption. Since an overriding goal is to minimize power consumption, it is imperative to use clever circuit and system techniques to minimize the current consumption of each block while achieving the necessary input-referred noise requirements and filtering aggressors before amplification to allow for reductions in supply voltage.



A secondary goal that has become important with the advent of applications that use large numbers of electrodes is system area consumption. Diagnostic EKG systems commonly use 12 electrodes [69], while modern EEG systems can use upward of 300 electrodes [46]. The new micro-electrode arrays (MEA) use as many as 1000 electrodes to measure neural activity with fine spatial precision [24]. The economics of semiconductors often create a one-to-one link between circuit area and cost creating a financial incentive to minimize area. In the case of implantable neural recorders, however, there is a further physical incentive to minimize area since the pitch of the electrodes is on the order of 100–400  $\mu\text{m}$ . This fine pitch makes it desirable for the electronics to be of the same size or smaller.

This dual goal of area and power efficiency drives the development of the design presented in this section. The design strategy is driven, in part, by recent advances in the state of the art in analog-to-digital converters and digital design. The figure of merit most commonly used for ADCs relates to the energy used per converted bit

$$\text{FOM} = 2^{-\text{ENOB}} P / f_s, \quad (4.1)$$

where  $\text{ENOB}$  is the effective number of bits,  $P$  is the DC power consumption, and  $f_s$  is the sampling rate. In the last three years, FOM values as low as 4.4 fJ/bit have been reported [62], and multiple designs have been reported with FOM values between 50–160 fJ/bit [2, 15, 64]. All of these designs used power supply voltages of 1–1.2 V and sampling rates between  $f_s=10$  kS/s – 1 MS/s with resolutions of 9–11 bits, and most of them had nearly constant FOMs down to data rates of 1 kS/s. At very low data rates, leakage currents and DC bias currents begin to hurt the FOM.

For many medical applications, sampling rates of a few hundred samples per second are sufficient. However, a quick calculation shows that the power consumption of ADCs has dropped to the point where it is a minor part of the overall power budget. Using the 4.4 fJ/bit FOM, for example, with a sampling rate of 10 kS/s and ENOB of 9 bits would only consume 22 nW. Even the far less aggressive performance of an ADC with an FOM of 60 fJ/bit would lead to a power consumption level of about

300 nW. This means that, from an energy perspective, it is relatively cheap to over-sample and potentially beneficial if it yields area and power savings in other blocks of the system.

#### 4.2.1 Thermal Noise and the Noise Efficiency Factor

Of all the aggressors shown in Fig. 4-1, the input-referred thermal noise of the IAMP is the most fundamental since it exists regardless of the technology used and is independent of environmental conditions. It is unique in that it is the only aggressor shown that is directly related to the current consumption of the IAMP, thereby creating a theoretical lower limit to the power consumption of the system dependent on noise spectral density requirements. To establish this theoretical limit, it is useful to analyze the input-referred noise of an amplifier that uses a single transistor. Since bipolar junction transistors (BJTs) achieve a lower input-referred noise spectral density than CMOS transistors for a given bias current, it is useful to use them as a reference.

For a common-emitter BJT, the short-circuit current noise density is

$$S_i(f) = \frac{i_{n,rms}^2}{\Delta f} = 2qI_C, \quad (4.2)$$

where  $i_{n,rms}$  is the RMS value of the current noise when integrated over a bandwidth  $\Delta f$ .  $I_C$  is the collector current and  $q$  is the charge on an electron. The input-referred voltage noise density is

$$S_v(f) = \frac{v_{ni,rms}^2}{\Delta f} = \frac{S_i(f)}{g_m^2}, \quad (4.3)$$

where  $v_{ni,rms}$  is the RMS value of the input-referred noise when integrated over a bandwidth  $\Delta f$  and  $g_m$  is the transconductance of the transistor. The transconductance is given by

$$g_m = \frac{I_C}{U_T}, \quad (4.4)$$

where  $U_T$  is the thermal voltage  $kT/q$ . Substituting (4.4) into (4.3), we find that the

input referred noise voltage can be written as

$$v_{ni,rms} = \sqrt{\frac{2qU_T^2 \Delta f}{I_C}}. \quad (4.5)$$

Clearly, there is a direct relationship between the input-referred noise voltage and the bias current. For a variety of reasons, most useful instrumentation amplifiers will not use a single transistor. However, practically all implementation will exhibit a similar tradeoff between noise and bias current, and it is useful to use the single BJT transistor as a reference to compare the performance of different IAMP implementations.

In the last decade, myriad publications have focused on the optimization of IAMP design. As a means for comparing the performance of these techniques, the *noise efficiency factor* (NEF) figure of merit was introduced in [60] and is commonly used to compare the bias current versus noise relation of an IAMP:

$$\text{NEF} = v_{ni,rms} \sqrt{\frac{2I_{tot}}{\pi \cdot U_T \cdot 4kT \cdot BW}}. \quad (4.6)$$

$I_{tot}$  is the total supply current consumed by the IAMP, and  $BW$  is the 3 dB bandwidth of the amplifier, assuming a single-pole roll-off (i.e. the effective noise bandwidth is assumed to be  $\Delta f = (\pi/2)BW$ ). The NEF FOM is meant to compare the performance of an amplifier to the theoretical limit of a single bipolar transistor. The derivation of (4.6) follows from (4.5), although there seems to be a discrepancy. In [60], the author states that the input-referred noise of a single bipolar transistor is

$$v_{ni,ref} = \sqrt{\frac{4kT \Delta f}{I_C}}. \quad (4.7)$$

However, since  $qU_T^2 = kTU_T$ , it follows from (4.5) that the correct equation for the input-referred noise is

$$v_{ni,rms} = \sqrt{\frac{2kT \Delta f}{I_C}}. \quad (4.8)$$

If (4.7) were correct, the NEF of a single BJT amplifier would be unity. Due to the small error, the NEF is actually  $1/\sqrt{2} \approx 0.707$ . Despite this error, the NEF is a useful

FOM since it is always true that a lower value of NEF corresponds to a more efficient amplifier. Since the NEF has been adopted in the form given by (4.6), we will use it to compare different implementations.

Since most IAMPs have a differential input, it is useful to calculate the theoretical limit of a differential pair. Assuming that each transistor in the differential pair carries the half as much current as the single transistor reference example such that the total currents are equal, it is straight forward to show that input-referred noise of a differential pair is  $v_{n,diff} = 2v_{ni,rms}$ , where  $v_{ni,rms}$  is defined by (4.8). Compared with the single transistor example, the noise of the differential pair is larger by a factor of  $\sqrt{2}$  due to the RMS sum of each transistor's contribution to the total noise, and another factor of  $\sqrt{2}$  because each transistor carries half as much current. This means that the theoretical minimum NEF of a differential pair is twice that of a single transistor. However, since the NEF of a single transistor is 0.707, **the NEF of a BJT differential pair is  $2\sqrt{2} = 1.41$ .**

When dealing with MOS transistors, the NEF is minimized when the input differential pair is biased in weak inversion [25]. The current noise density for a MOS transistor is also give by (4.2), but its transconductance is slightly smaller than that of a BJT transistor. Specifically,

$$g_m = \frac{\kappa I_C}{U_T}, \quad (4.9)$$

where  $\kappa$  is a process variable approximately equal to 0.7. While the exact value of  $\kappa$  depends on the specific technology, it is always smaller than unity. As a result, the transconductance of a MOS transistor biased in weak inversion is always smaller than that of a BJT. Since (4.3) also applies to MOS transistors, the input-referred noise of a MOS transistor biased in weak inversion is always larger than the noise of a BJT by a factor of  $1/\kappa$ . In the case of a differential pair, the theoretical limit is equal to that of a BJT divided by  $\kappa$ . Specifically, the lowest achievable NEF for a

MOS amplifier with a differential input pair is

$$\text{NEF}_{\text{mos,diff}} = \frac{\text{NEF}_{\text{bjt,diff}}}{\kappa} \approx \frac{1.41}{0.7} = 2.02, \quad (4.10)$$

as cited in [68].

Both [25] and [68] discuss techniques for minimizing the NEF of a differential amplifier using MOS transistors. They achieve respective NEF values of 4.0 and 2.67, both of which are among the lowest reported in the literature. A thorough discussion of these techniques is beyond the scope of this thesis, but we note that similar techniques were used in the implementation of our IAMP.

## 4.2.2 Flicker Noise

Flicker noise, sometimes referred to as  $1/f$  noise, exists in all active devices. In MOS devices, flicker noise is mostly caused by imperfections in the gate oxide that randomly capture and release charge carriers. The resulting noise spectral density of the current is

$$S_{i,1/f} = \frac{K_{1/f} I^a}{f^b}, \quad (4.11)$$

where  $K_{1/f}$  is a device-dependent constant and  $b \approx 1$ . For MOS devices,  $a \approx 2$  in subthreshold (weak inversion) and  $a \approx 1$  above threshold (strong inversion). To refer this current noise to the input of the device, (4.11) is divided by the square of the transistor's transconductance

$$S_{v,1/f} = \frac{S_{i,1/f}}{g_m^2}. \quad (4.12)$$

For subthreshold transistors,  $g_m$  is given by (4.9). Since  $a \approx 2$  in subthreshold, (4.12) can be rewritten as

$$S_{v,1/f} \approx \frac{K_{1/f} \phi_t^2}{\kappa^2 f}. \quad (4.13)$$

The  $1/f$  nature of flicker noise makes it particularly problematic at low frequencies, where biomedical signals often reside. To make matters worse, according to (4.13), the input-referred flicker noise of a MOS transistor cannot be reduced by increasing the current, as is the case with thermal voltage.  $K_{1/f}$  is inversely proportional to the

transistors' areas, so the noise can be reduced by sizing the transistors larger, but this increases their input capacitance which can also degrade noise performance by attenuating the input signal.

### 4.2.3 Using Chopping to Mitigate Flicker Noise

Flicker noise and amplifier DC offset are major aggressors that degrade the performance of biomedical IAMPs. Chopping is a technique commonly used to mitigate the effects of these aggressors by frequency translating the input signal before it is corrupted by flicker noise. Fig. 4-2 illustrates the effects of chopping on both the spectrum of the desired signal and on that of the unwanted flicker noise. The first step involves *up chopping* the incoming signal. Chopping is typically done using CMOS switches, such that minimal noise is introduced while translating the frequency spectrum of the incoming signal to a higher frequency. The input-referred noise and DC offset of the amplifier are shown as additive aggressors near DC. Since the input signal now lies at frequencies beyond the  $1/f$  corner frequency of the noise, it is only corrupted by thermal noise. Next, the signal and aggressors are amplified and the resulting signal is *down chopped* while the aggressors are up chopped. The next step involves passing the chopper's output through a low-pass filter to attenuate the frequency-translated aggressors. In modern systems, the filter's output is commonly digitized using an ADC.

Appendix A explains chopping in more detail and describes various IAMPs in recent literature that employ chopping. It details the benefits and drawbacks of each technique with special emphasis on the system described in [65], since we implemented a similar architecture. As Appendix A explains, there are major problems with this architecture that prevented us from using the chopping functionality. Nonetheless, with the choppers disabled, the IAMP works relatively well and we were able to demonstrate the functionality of the other blocks which were the real emphasis of this work.

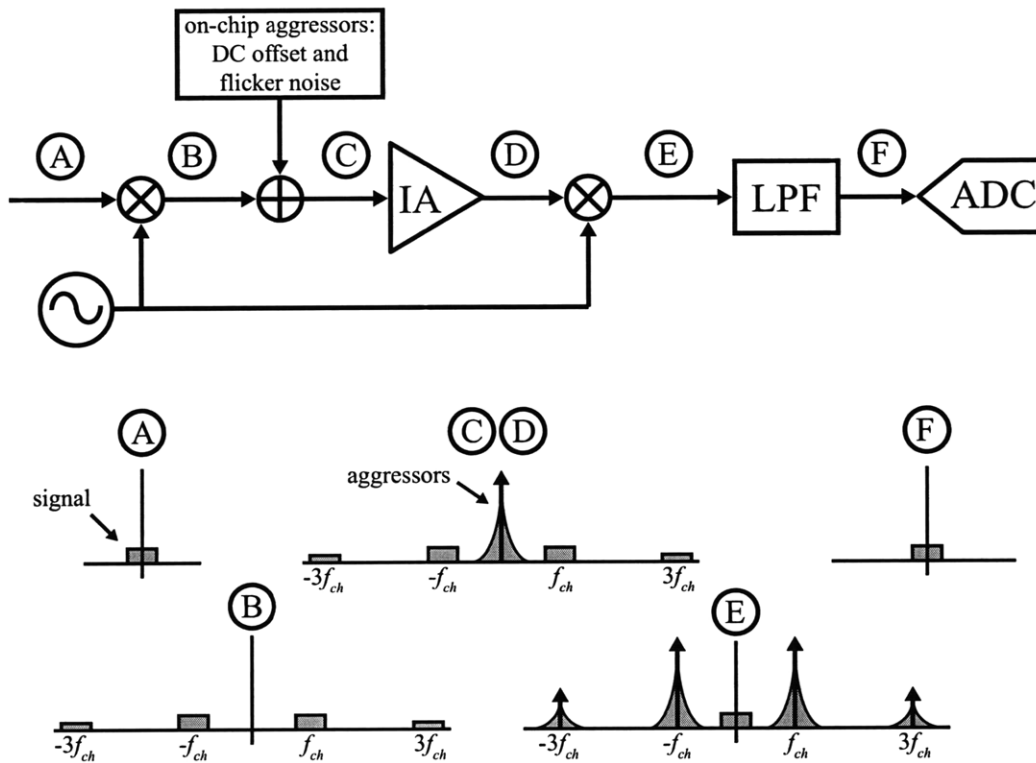


Figure 4-2: Frequency-domain analysis of the signal and noise at different stages of a chopper-based sensor interface system.

#### 4.2.4 Power Line Interference

In addition to stochastic noise such as thermal or flicker noise, unwanted periodic, or AC, aggressors often corrupt the desired signal being measured. The most common form of AC interference cited in the biomedical electronics literature is power line interference. It is caused by capacitive coupling between power lines that deliver electricity using either 50 or 60 Hz AC signals at voltage levels between 110 and 220  $V_{\text{rms}}$  (depending on the country). The earliest comprehensive model describing different coupling mechanisms for this type of aggressor is found in [26]. The four mechanisms listed are: 1) magnetic induction; 2) displacement currents into the electrode leads; 3) displacement currents into the body; and 4) equipment interconnect imperfections.

Magnetic induction occurs due to long lead connections between the patient and the instrumentation that create conductive loops. The changing magnetic flux density induces an electromotive force (potential) that depends on the area of the loop, its orientation, and the amplitude of the magnetic flux. This type of interference is significantly reduced by using short leads and twisting them as much as possible to minimize the area of the loop.

Displacement currents result from changes in the electric field intensity. Capacitive coupling between the power line and the leads or the body induce currents that are converted into voltages as they flow through the various impedances (most notably the capacitance between the patient and ground). Assuming that differential measurements are taken, most of this current leads to common-mode voltages that are rejected if the amplifier has a good common-mode rejection ration (CMRR). However, mismatches in the leads and/or contact impedances convert part of these currents into differential-mode signals that are amplified by the system. This effect can be largely eliminated in the case of electrode leads by shielding the leads. Displacement currents in the body, however, must be dealt with since it is typically not feasible to shield the patient.

The fourth type of interference is rare and can be remedied by proper instrumen-



tation use. It is caused, for example, if a patient is connected to two different instruments that are plugged into different power outlets creating a potential difference between the grounds of the two instruments and inducing leakage currents.

Of these four mechanisms, the third (displacement currents into the body) is the most prevalent and difficult to fix. Fig. 4-3 shows a typical model similar to the one introduced in [26] for two- and three-electrode measurements. This model has been widely used in subsequent literature, although other models are suggested in [59] for circuits with low common-mode input impedances and in [13] to include the effects of PLI coupling through the power supply.

Typical values for  $C_B$  range between 120 pF–550 pF, and for  $C_P$  between 0.2 pF–10 pF [57]. At 60 Hz, the impedance of  $C_P$  is significantly larger than impedances through the body, allowing  $i_P$ , the displacement current, to be modeled as the result of a current source. Its value can range between 10 nA–1  $\mu$ A (RMS) depending on  $C_P$  and the power line voltage. It can be easily measured by touching the tip of an oscilloscope probe with a finger being careful not to touch the ground conductor as this would provide a low impedance path to ground. Since the input impedance of the oscilloscope is known, the displacement current can be calculated from the measured voltage. For example, for a typical input impedance of 1 M $\Omega$ , a resulting voltage of 1 V<sub>p-p</sub> signifies a displacement current of  $i_P = 1 \mu\text{A}_{p-p}$ .

For isolated amplifiers whose reference ground is not the same as earth ground, a parasitic impedance results between the two. The isolation impedance  $Z_{ISO}$  is mostly important in the two-electrode case and is typically capacitive. A typical value of  $C_{iso}$  is on the order of 300 pF for bench-top equipment [63] and 1 pF for ambulatory devices [57].

There are three main ways in which  $i_P$  is converted into a differential-mode interference (DMI). The first results from part of  $i_P$  flowing through impedances internal to the body between the two points where the differential electrodes are connected (modeled as  $Z_T$  in Fig. 4-3). The resulting voltage,  $v_T$ , is inherently differential with respect to the measurement. The two other methods result from  $i_P$  flowing through various impedances (primarily  $C_B$ ) and creating a common-mode voltage,  $V_{CEMI}$ , that

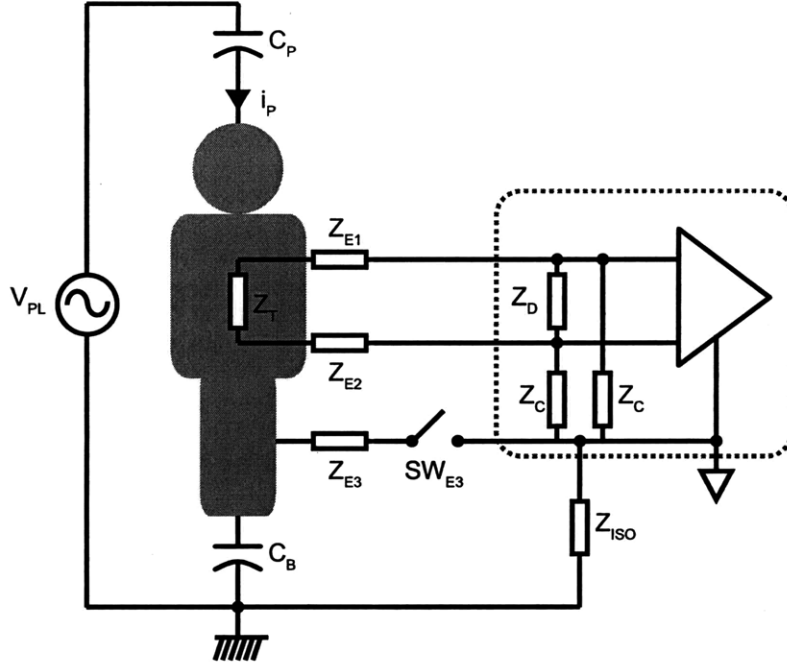


Figure 4-3: Classical model for power line interference in two-electrode measurements ( $SW_{E3}$  off), and three-electrode measurements ( $SW_{E3}$  on).

is converted into DMI through electrode mismatches or finite CMRR in the IAMP.

For finite common-mode input impedances in the IAMP,  $Z_C$ , mismatches in the electrode impedances  $Z_{E1,E2}$  convert  $V_{cEMI}$  into DMI. Assuming  $Z_C$  is significantly larger than  $Z_{E1,2}$  at 60 Hz, and that the percentage mismatch between the two  $Z_C$  impedances is significantly smaller than the percentage mismatch of the electrode impedances  $Z_E$ , the resulting DMI is

$$v_{\Delta Z} \approx v_{cEMI} \frac{\Delta Z_E}{Z_C}. \quad (4.14)$$

Common-mode PLI is also converted to DMI as a result of finite CMRR in the amplifier. The total DMI, therefore, is

$$v_{iEMI} = v_T + v_{cEMI} \left( \frac{\Delta Z_E}{Z_C} + \frac{1}{CMRR} \right), \quad (4.15)$$

where  $CMRR$  is the ratio between the amplifier's differential mode gain and its common-mode gain ( $A_{Vdm}/A_{Vcm}$ ). From (4.15) it is clear that large values of  $CMRR$

and input impedance are desirable.

Unfortunately,  $Z_T$  can be as large as 1 k $\Omega$  and could lead to values of  $v_T$  that dominate  $v_{iEMI}$  [57]. As a result, even an ideal IAMP with infinite input impedance and CMRR cannot eliminate all of the DMI. If all of  $i_P$  flows through  $Z_T$ , the resulting contribution to DMI is

$$v_T = i_P Z_T. \quad (4.16)$$

As noted in [26], this is unlikely and the actual voltage should be smaller. Nonetheless, using a worst-case value of 1  $\mu A$  for  $i_P$  and 1 k $\Omega$  for  $Z_T$ , we see that  $v_T$  values of order 1 mV<sub>rms</sub> are entirely possible.

#### 4.2.5 Techniques for Mitigating the Effects of PLI

According to (4.15), there are three terms responsible for the differential-mode interference that corrupts a signal amplified by an IAMP. The first term,  $v_T$ , results from displacement currents flowing through finite impedances in the body. Since it is independent of the amplifier's characteristics, it is very difficult to eliminate. The second and third terms, however, can be significantly attenuated by: 1) increasing the common-mode rejection ratio of the IAMP; 2) increasing the common-mode input impedance of the IAMP,  $Z_C$ ; and 3) reducing  $v_{cEMI}$ . The following sections describe techniques for doing each of these things.

##### Common-Mode Rejection and High Input Impedance

According to (4.15) the only design variables available to reduce interference due to PLI are  $V_{cEMI}$ ,  $Z_C$ , and  $CMRR$ . When measured with respect to earth ground, the worst-case value of  $V_{cEMI}$  occurs when  $C_P$  is large, and  $C_B$  is small. Using worst-case conditions of  $V_{PL} = 220$  V,  $C_P = 10$  pF, and  $C_B = 120$  pF yields a value of  $V_{cEMI} = 17$  V<sub>rms</sub> = 48 V<sub>p-p</sub>. While large values of  $CMRR$  and  $Z_C$  could, theoretically, be used to prevent this large common-mode signals from becoming DMI, ESD protection circuits would almost certainly clamp. Therefore, if ESD structures are used,  $V_{cEMI}$  must be limited to the supply rails' voltages.

$V_{cEMI}$  can be reduced by providing a low impedance path to ground for  $i_P$ . This is usually done by using a third electrode, although [19] gives an example of how this can be done with two electrodes (a common-mode signal is fed back to both electrodes such that  $Z_C$  is reduced while  $Z_D$  remains large). Assuming  $V_{cEMI}$  is reduced to a level of 1  $V_{p-p}$  to enable low power-supply operation, a common-mode rejection ratio of 60 dB would only add  $350 \mu V_{rms}$  to  $v_{iEMI}$  which is lower than the worst-case value from  $v_T$ . Similarly, a value of  $Z_C = 10^3 \times \Delta Z_E$  will produce the same amount of  $v_{iEMI}$ . While higher values of CMRR and  $Z_C$  can reduce the amount of common-mode voltage that is converted to a differential-mode aggressor, the benefits are diminished due to the interference from  $v_T$ . For applications where  $v_T$  is small because the electrodes are placed near each other, there is a benefit to having larger CMRR and  $Z_C$ . Otherwise, the differential-mode aggressor is dominated by  $v_T$  and increasing CMRR and  $Z_C$  yield minimal benefits.

To understand the value of  $Z_C$  that is necessary to minimize that amount of CMI that is converted to DMI, we must first understand what typical values of electrode impedance mismatch are. Mismatch in the electrode impedance varies significantly depending on the type of electrodes used. Skin surface electrodes commonly use conductive gel to reduce their contact impedance. Dry electrodes, on the other hand, do not use gel, resulting in much higher values of contact impedance and mismatch. Dry electrode impedance measurements conducted in [20] show that, at 60 Hz, the impedance between the electrode and the surface of the skin – the contact impedance – is almost entirely capacitive and larger than  $3.5 \mu F$  (measured over various conditions). The dominant impedance was actually found to be the impedance across the stratum corneum (the outermost layer of skin). This impedance is in series with the contact impedance and was measured to have a worst-case value of  $0.05 \mu F$ , or  $53 k\Omega$  at 60 Hz (a smaller value of capacitance is worse since this implies a larger impedance). Assuming that the worst-case mismatch is equal to the full value of the electrode impedance, a value of  $Z_C = 10^3 \times 53 k\Omega = 53 M\Omega$  would only add  $350 \mu V_{rms}$  to  $v_{iEMI}$ . While this value can be significantly larger than the desired signal, it is still smaller than the worst-case  $v_T$  value of  $1 mV_{rms}$ , and therefore reducing it

Table 4.1: Summary of Minimum Requirements for CMRR and IAMP Input Impedance

$V_{cEMI}$	$\leq 1V_{p-p}$
CMRR	$\geq 60\text{dB}$
$Z_C$	$\geq 60\text{M}\Omega$ @60 Hz
$Z_D$	$\geq 80\text{M}\Omega$ @1 Hz

further may not yield benefits. However, achieving an amplifier input impedance on the order of 50 M $\Omega$  at 60 Hz is not trivial and the designer must be mindful of this requirement.

The requirements for the differential input impedance  $Z_D$  are not linked to PLI, but instead, to signal attenuation. At low frequencies, the electrode impedance  $Z_E$  can be very large and the voltage divider between it and  $Z_D$  can attenuate the signal. To prevent this,  $Z_D$  should be significantly larger than  $Z_E$ , but not necessarily as large as  $Z_C$ . For example, a value of  $Z_D = 40 \times Z_E$  results in less than 5% attenuation (<0.5 dB), which is acceptable. According to [20], at 1 Hz, the lower bound of most biomedical signals of interest,  $Z_E$  is mostly resistive and has a worst-case value smaller than 2 M $\Omega$ . To limit attenuation,  $Z_D$  should be made larger than 80 M $\Omega$  at 1 Hz. Table 4.1 lists the minimum requirements necessary to keep PLI from causing clamping in ESD circuits and for keeping the DMI below 1 mV<sub>rms</sub>.

### Using a Third Electrode

As discussed in Section 4.2.4, it is usually necessary to reduce the common-mode PLI signal to less than 1 V<sub>p-p</sub> to prevent it from causing ESD protection circuits from clamping. A simple way of reducing  $v_{cEMI}$  is to use a third electrode as shown in Fig. 4-3 when  $SW_{E3}$  is on. The electrode impedances  $Z_{E1,2,3}$  are typically on the order of 10 k $\Omega$ –100 k $\Omega$  [55]. Compared with  $|X_B|$  which is on the order of 5 M $\Omega$ –20 M $\Omega$  at 60 Hz,  $Z_{E3}$  is much smaller and results in a significantly lower value of  $v_{cEMI}$  (<100 mV). Note here that the common-mode voltage of interest is between the two inputs to the amplifier and the amplifier’s reference voltage (not earth ground). In other words, if  $Z_{ISO}$  is large, the common-mode voltage may still be large with respect to

earth ground, but it will be small with respect to the amplifier's reference voltage (often referred to a ground also).

### The *Driven-Right-Leg* Technique

In EKG measurements, a common technique for further reducing  $v_{cEMI}$  while protecting the patient from electrical shock, is to drive the right leg through the third electrode using common-mode feedback. This *driven-right-leg* (DRL) technique shown in Fig. 4-4 dates back to the early 70s and is described in [72]. Compared with the simpler three-electrode measurement shown in Fig. 4-3 ( $SW_{E3}$  on), DRL has two benefits: first, it reduces the effective impedance to the amplifier's reference voltage (thereby reducing  $v_{cEMI}$ ); and second, it limits the current through the person which is important for safety reasons. The analysis in [72] shows that the effective impedance to the reference voltage is  $R_{cEMI} = Z_{E3}/(G + 1)$ , where  $G = 2R_F/R_A$ . Clearly, by sizing  $R_F$  and  $R_A$  appropriately, this can reduce  $v_{cEMI}$  significantly. Additionally, by making  $R_O$  sufficiently high, the current driven into the body can be limited to safe levels ( $< 20 \mu A$ ) while maintaining a low effective resistance to the reference voltage. In summary, the DRL technique uses negative feedback to create an effective common-mode, low impedance path between the patient and the amplifier's reference voltage. This attenuates the common-mode interference while limiting the amount of current injected into the patient.

In EEG, a *ground* electrode is commonly used to provide the same benefits as the third electrode and/or the DRL in EKG. It can be placed anywhere in the body and is often placed on the wrist or on the forehead. Recent publications, however, argue excluding a ground electrode can yield benefits in the spatial symmetry and signal-to-noise ratio of the measurements [49]. In addition to the ground electrode, a *reference* electrode is used as one input of each differential amplifier. It is commonly connected to one or both ear lobes, the nose, or the mastoids (behind the ears) [18]. Fig. 4-5 shows the typical connection between EEG electrodes and their respective amplifiers.

Fig. 4-6 shows the standard EEG *ten-twenty* electrode configuration using 21

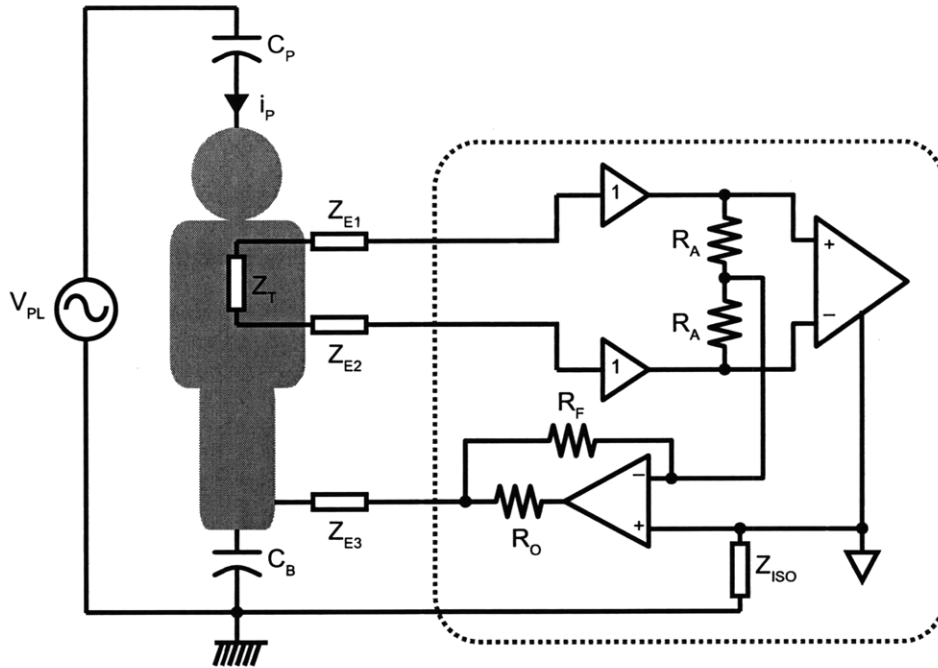


Figure 4-4: Simplified driven-right-leg circuit for safe common-mode voltage reduction.

electrodes along with the *ten-ten* electrode configuration that uses 74 electrode locations. The more recent *five percent* electrode system described in [46] allows for up to 345 electrodes.

EEG measurements share the same challenges regarding interference as EKG, with the added difficulties associated with having a greater number of electrodes, higher electrode contact impedances, and smaller signals. Whereas EKG signals are on the order of 1 mV, EEG signals range between 10-500  $\mu\text{V}$  [69]. As already noted, one of the ways common-mode interference is converted to DMI is through mismatches in the electrodes' contact impedances, and these are much higher in EEG due to hair and the head's curvature. Coupled with the need for greater amplification due to the signals small amplitude, it is clear that strong interference signals could saturate the analog blocks in the system or necessitate higher dynamic range operation. Further, the drive towards systems with significantly more electrodes motivate the development small low-power sensor interfaces.

In addition to noninvasive biopotential signals measured on the surface of the skin, there has been a recent surge in interest in developing implantable systems to

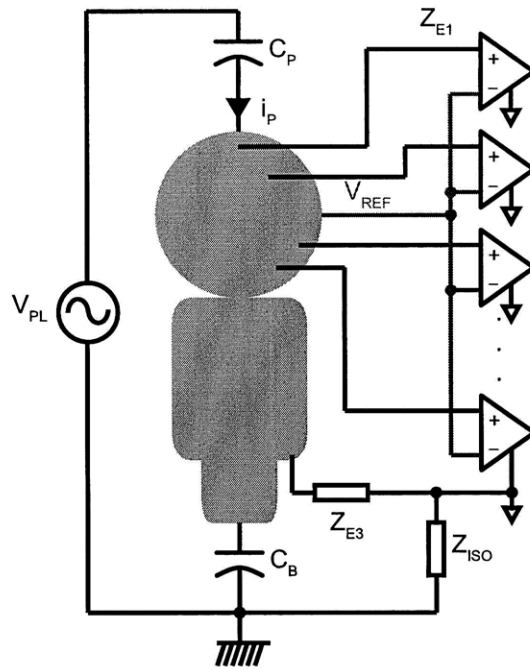


Figure 4-5: Typical connection of EEG amplifiers to their respective electrodes.

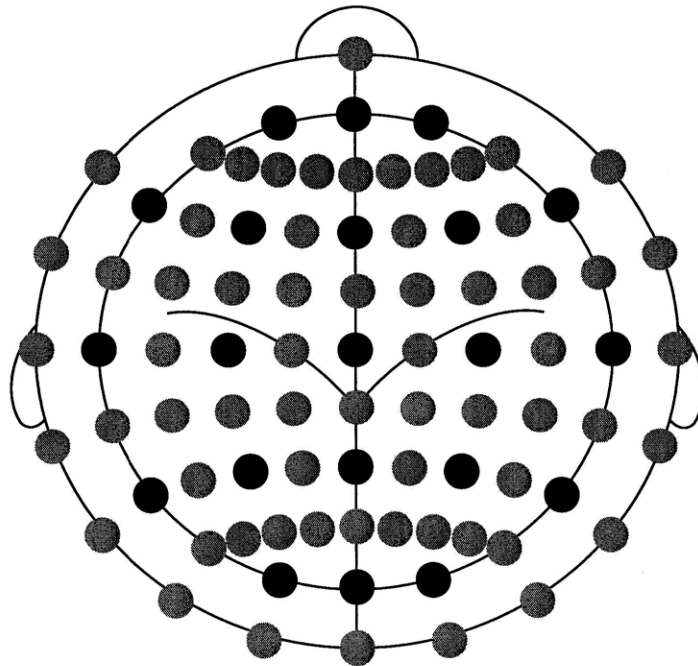


Figure 4-6: Electrode configuration for standard 10-20 electrode system (black dots) with additional electrodes for 10-10 system (gray dots). Figure taken from [46]



measure neural field potentials through electrocortigraphy (ECoG) on the surface of the cortex and local field potentials (LFP). To the author’s knowledge, there are no publications specifically discussing the issue of power line interference in medical implants. However, in [25], where neural activity in a rat’s olfactory cortex was measured using a CMOS IAMP and extracellular microelectrodes, the *Biological Test Results* section states that “due to the unshielded wires connecting the electrode array to the amplifier circuit, we observed strong interfering signals at 60 Hz and approximately 50 kHz”. It appears that power line interference is also a problem for medical implants.

#### 4.2.6 Electrode and IAMP DC Offset

PLI is particularly egregious because it typically lies within the desired band of interest. DC offset, on the other hand, while typically not in the band of interest, can cause significant problems because of its sheer size. There are two main types of DC offset: one is caused by electrode polarization (off-chip), and the other by mismatches in the electronics (on-chip). A thorough explanation of electrode DC offset (EDO) can be found in [69]. We present a brief synopsis here.

An electric potential called the *half-cell* potential arises from a double layer of charge that forms at any electrode-electrolyte interface. This potential can be larger than 1V, and can be positive or negative depending on the electrode material. When no current is flowing between the electrode and electrolyte, this voltage is constant. Therefore, since biopotential measurements are almost always made differentially, the EDO is typically much smaller than the half-cell potential. Unfortunately, different electro-chemical mechanisms cause electrodes to polarize, resulting in an *overpotential* that adds to the half-cell potential and creates a differential offset between two electrodes. These effects can be reduced by using *non-polarizable* electrodes, such as Ag/AgCl, where a silver electrode is coated with silver-chloride which is a slightly soluble ionic compound resulting in a more stable electrode. Even with coating, however, ionic concentrations across different membranes in the skin’s anatomy can cause potential differences on the order of tens of millivolts. Taking all of these mechanisms

into account, worst-case specifications for EDO of  $\pm 300$  mV are cited in [69] for skin surface electrodes. Platinum-iridium (PtIr) implantable electrodes were characterized in [17] to have a much smaller variation range of  $\pm 15$  mV ( $\pm 6\sigma$ ).

Electronic DC offset can also be on the order of millivolts and mostly arises from component mismatches. The IAMP at the front end of most sensor interface systems typically has significant gain, and as a result, its input-referred DC offset is the most important in the system. With proper design techniques, the systematic DC offset of an IAMP can be eliminated. However, variations in doping concentrations and geometry size often lead to random variations in the DC offset [23]. DC offset, therefore, can be modeled as a random variable with zero mean and some standard deviation. For many opamp-based IAMPs, the variance of the offset can be reduced by increasing the width and/or length of the input pair of transistors. Using reasonably sized transistors, these offsets can be made smaller than 1 mV.

#### 4.2.7 Dealing with Electrode DC Offset

In [25], Harrison introduced the use of the MOS-bipolar element as a means of integrating exceptionally high resistance values ( $10\text{ T}\Omega$ ) using minimal area. This *pseudoresistor* had been described in [16] for use in photoreceptors, and was used in [25] to create a sub-Hz high-pass cutoff frequency without the need for large, area-consuming capacitors. As explained in [16], the pseudoresistors, drawn as PMOS transistors in Fig. 4-7, act as diode connected PMOS devices when  $V_{GS}$  is negative. When  $V_{GS}$  is positive, the parasitic source-well-drain p-n-p BJT is activated and the device acts as a diode connected BJT.

Using this device and optimizing an operational amplifier for minimal power consumption under a given input-referred noise constraint, an IAMP consuming  $80\ \mu\text{W}$  was developed for neural recording applications, and a second device for EEG applications was designed consuming only  $900\ \text{nW}$ . This technique offered the benefits of small size and power, while rejecting electrode DC offsets. It also achieved good thermal noise performance, reasonable CMRR, and a high input impedance. The major deficiencies of this IAMP are its relatively high supply voltage (5 V) and  $1/f$

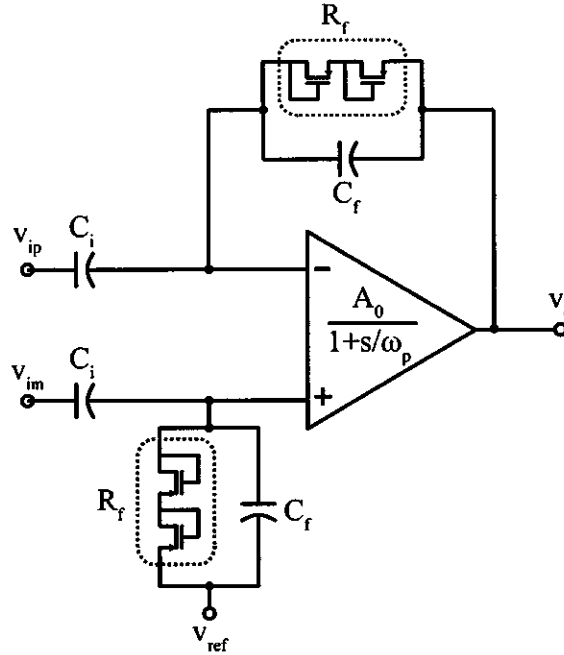


Figure 4-7: Instrumentation amplifier using MOS-bipolar pseudoresistor elements [25].

noise, along with limited rejection of 60 Hz.

In [68], a similar amplifier was built using using the same type of pseudoresistor, but further optimizing the operational amplifier to achieve a noise efficiency factor of 2.67. This is the lowest NEF reported in the literature, and close to the theoretical limit is 2.02 discussed in Section 4.2.1.

#### 4.2.8 Top-Level Design Requirements

As illustrated in Fig. 4-1 and discussed in the previous sections, many biomedical signals (including EKG, EEG, EMG, and neural field potentials) reside in the frequency range of 1 Hz-1 kHz and amplitude range of  $10 \mu\text{V}_{\text{p-p}}$ - $10 \text{mV}_{\text{p-p}}$ . Unfortunately, these signals can be corrupted by aggressors including PLI, DC offset,  $1/f$  noise, and thermal noise that threaten the usefulness of measurements.

The system-level requirements necessary to extract useful physiological signals in the midst of these aggressors can be summarized qualitatively as follows: 1) the input-referred noise should be small enough so as to not corrupt the measured signal,

2) the system should minimize the amount of common-mode PLI that is converted to DMI, 3) the system bandwidth should be at least 1 kHz wide or tunable to such a bandwidth, 4) the system must be able to withstand electrode DC offsets of up to  $\pm 300$  mV without saturating, and 5) the system should be able to withstand differential-mode PLI of up to  $1 \text{ mV}_{\text{rms}}$  with an option to notch it out.

Many applications with a bandwidth below 100 Hz (e.g. EKG, NFP) require accurate measurements of signals between  $20 \mu\text{V}$ – $1 \text{ mV}$ , and therefore require the input-referred noise to be below  $5 \mu\text{V}_{\text{rms}}$  [69]. EEG and ECoG signals are on the order of  $5 \mu\text{V}$ – $100 \mu\text{V}$  and therefore require the input-referred noise to be below  $1 \mu\text{V}_{\text{rms}}$ . EMG signals have a wider bandwidth, but are also larger requiring the input-referred noise to be below  $10 \mu\text{V}_{\text{rms}}$  in a frequency band of 10 Hz–1 kHz. In order to satisfy the requirements of as many applications as possible, the input referred noise spectral density of the system should be below  $100 \text{ nV}/\sqrt{\text{Hz}}$ , although many applications might only require a noise density between 300–500  $\text{nV}/\sqrt{\text{Hz}}$ .

Beyond these basic requirements the specifications of each block in the system depend largely on the design strategy. For example, adding more gain in the IAMP will relax the noise specifications of the filter, but may result in more demanding linearity specifications. In the following sections we discuss different strategies presented in the literature for dealing with each of the system requirements and later derive the specifications for each of our system’s blocks

### 4.3 Instrumentation Amplifier Design

The IAMP used in this work is shown in Fig. 4-8. Most of the circuits implemented in this section were designed, simulated, and implemented by Marcus Yip, but a their analysis is included here to facilitate system-level understanding. As explained in the previous section, our original intention was to use chopping to improve the 1/f noise rejection of the system. However, the switched-capacitor, parasitic resistance introduced by the input chopper ends up generating an unacceptable amount of noise that dwarfs any benefits achieved by chopping. As a result, we found the system

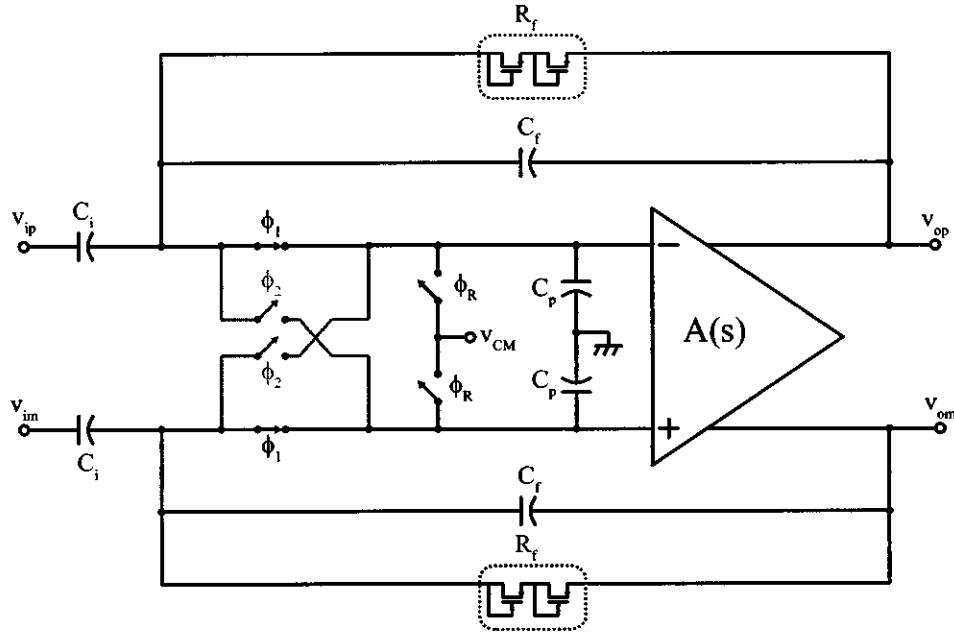


Figure 4-8: Instrumentation amplifier architecture used for the work presented in this thesis. Although the chopping switches were not used, they are shown here to analyze their effect.

performs much better with the choppers turned off. Since the chopping switches have their own *off* resistance, they are shown in the figure, but in gray to emphasize they are not used for chopping. A set of reset switches originally intended to make up part of the integrator feedback path are left in black, as they have proven to be useful in expediting the settling time of the opamp when it is first turned on.

Fig. 4-9 (A) shows the equivalent single-ended version of the IAMP in Fig. 4-8 including the *off* resistance of the chopper switches and all noise sources. Fig. 4-9 (B) shows the equivalent block diagram. The opamp is modeled with only its dominant pole at  $\omega_0$ . The loop gain for this system is

$$LG(s) = \frac{\omega_0 A_0 C_f (s + \omega_f)}{C_{eff} (s + \omega_{eff}) (s + \omega_0)}. \quad (4.17)$$

This gain is significantly larger than unity for typical component values and within the bandwidth of interest. We can, therefore, approximate the closed-loop transfer

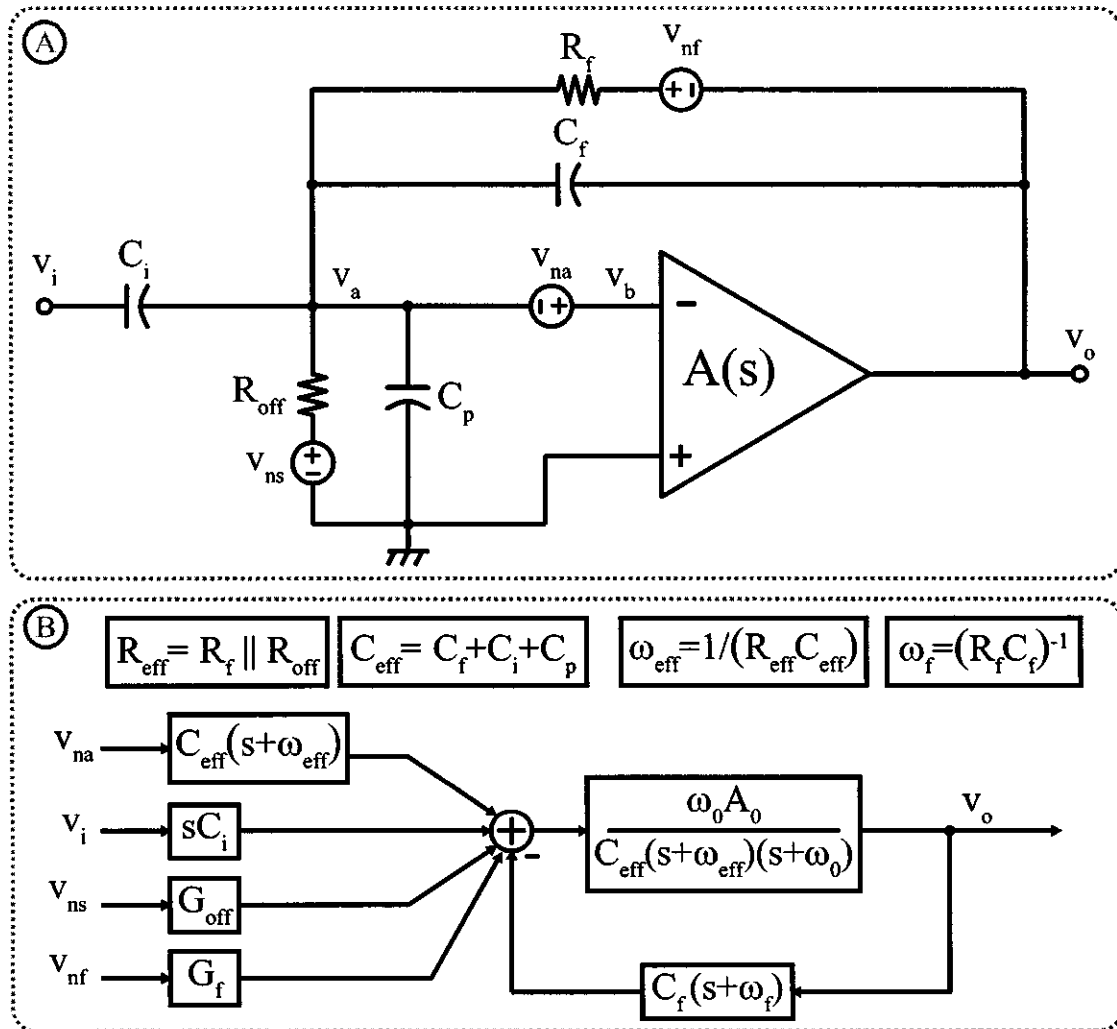


Figure 4-9: (A) Single ended equivalent of IAMP in Fig. 4-8 including noise sources, and (B) corresponding block diagram.

function of the feedback loop as the inverse of the feedback factor

$$H_{CL}(s) \approx \frac{1}{C_f(s + \omega_f)}. \quad (4.18)$$

The input-signal-to-output transfer function is therefore

$$H_{sig}(s) = \frac{C_i}{C_f} \frac{s}{(s + \omega_f)}, \quad (4.19)$$

which gives the desired high-pass characteristic necessary to filter out the electrode DC offset while achieving a mid-band gain of  $C_i/C_f$ . The corner frequency  $f_f$  is set by  $R_f$  and  $C_f$ . As in [25],  $R_f$  is implemented with back-to-back MOS-bipolar pseudoresistors with total value of  $R_f = 15 \text{ T}\Omega$ .  $C_f$  was designed to be 100 fF, and  $C_i$  is 10 pF, giving a high-pass corner frequency of 100 mHz and a mid-band gain of 100 V/V.

### Noise from opamp

The noise transfer function of the IAMP is

$$H_{na}(s) = \frac{C_{eff}(s + \omega_{eff})}{C_f(s + \omega_f)} = \frac{(C_i + C_f + C_p)(s + \omega_{eff})}{C_f(s + \omega_f)}. \quad (4.20)$$

For values of  $\omega \gg \omega_f, \omega_{eff}$ , the input-referred noise transfer function is

$$H_{na,i}(s) \approx \frac{(C_i + C_f + C_p)}{C_i}. \quad (4.21)$$

This means that, in the band of interest (i.e.  $\omega > \omega_f$ ), and as long as  $\omega_{eff} \leq \omega_f$ , the input referred noise of the IAMP is simply the product of  $v_{na}$  and the ratio of input capacitance to total capacitance. By design,  $C_i$  is made much larger than  $C_f$  and  $C_p$ , so the input referred noise can be further simplified as just the noise of the opamp.

Fig. 4-10 shows the full schematic of the opamp implemented in the prototype CMOS chip we designed. The chopping switches in the second stage are not used for chopping and instead biased in the triode/linear region or OFF. Specifically,  $V_{chp} =$

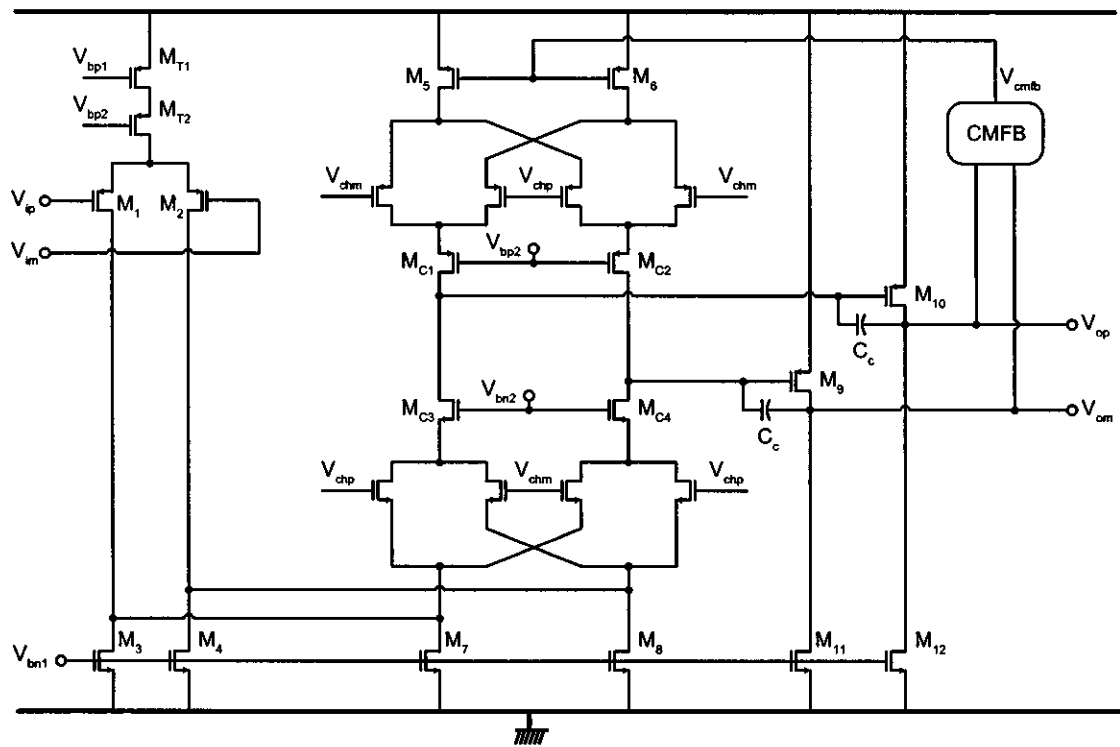


Figure 4-10: Schematic of fully differential operational amplifier including second stage choppers. These choppers are not used in normal operation but are included here for completeness.





noise power is, therefore,

$$\overline{v_{nth}^2} = \frac{4qI_D}{g_{m1}^2} \Delta f = \frac{4q\phi_t^2}{\kappa^2 I_D} \Delta f. \quad (4.24)$$

Since the only design variable is  $I_D$ , it is clear that beyond the design optimization necessary to minimize excess noise from other components, the only way to reduce the input referred thermal noise is by increasing  $I_D$ .

Depending on the transistors' sizing, the bias current, and the signal bandwidth, flicker noise in the opamp can be the most significant noise contributor in the biomedical IAMP. The input-referred flicker voltage noise of the input pair is

$$\frac{\overline{v_{nfl}^2}}{\Delta f} = \frac{K}{WLC_{ox}f}, \quad (4.25)$$

where  $K$  is a process constant dependent on various device characteristics and typically larger for NMOS devices than PMOS devices. The transistor's oxide capacitance  $C_{ox}$  is typically not a design variable available to the circuit designer for a given technology, and therefore, the only design variables are the width,  $W$ , and length  $L$  of the device. Note that the input referred flicker noise is independent of bias current. This means that increasing  $I_D$  reduces input-referred thermal noise, but not flicker noise, and therefore results in a higher  $1/f$  corner frequency. The total input referred noise of the opamp is, therefore

$$\frac{\overline{v_{na}^2}}{\Delta f} = \frac{4q\phi_t^2}{\kappa^2 I_D} \frac{\text{NEF}}{2.02} + \frac{K}{WLC_{ox}f} \alpha_f, \quad (4.26)$$

where the NEF is the *noise efficiency factor* introduced in Section 4.2.1, and the theoretical minimum for a CMOS, subthreshold opamp is 2.02 [68]. The ratio of NEF/2.02 gives a measure of the excess noise contributed by components other than the input pair and is has been reported as low as  $1.2\times$ , but is typically between  $2\times$  and  $10\times$ . The constant  $\alpha_f$  is introduced here to account for excess flicker noise contributed by components other than the input pair.

### Noise from $R_f$

The input-referred noise transfer function for the feedback resistor is

$$H_{f,i}(s) = \frac{1}{R_f C_i (s + \omega_f)}. \quad (4.27)$$

For  $f \gg f_f$ , (4.27) is approximately  $1/(\omega C_i)$ . Multiplying the square root of the resistor's noise spectral density with the absolute value of this transfer function gives the input-referred voltage noise PSD due to  $R_f$ :

$$|v_{ni,f}(f)| = \sqrt{\frac{4kT}{R_f} \frac{1}{2\pi f C_i}}. \quad (4.28)$$

Intuitively, we see that beyond its high-pass corner frequency, the impedance of the IAMP's feedback network is dominated by the capacitor  $C_f$ . Combined with the opamp,  $C_f$  integrates currents flowing into node  $v_a$  and therefore shapes the white noise of  $R_f$  giving it the appearance of  $1/f^2$  noise. Recalling the analysis in Sect. 4.2.8, the desired voltage noise spectral density should be less than  $100 \text{ nV}/\sqrt{\text{Hz}}$ . Rearranging (4.28) to solve for the minimum value of  $R_f$  that achieves this goal, and noting that the worst-case scenario occurs at the lower edge of the desired frequency band (i.e. 1 Hz), we find that  $R_f \gg 420 \text{ T}\Omega$ . Any value of  $R_f$  smaller than this will result in a  $1/f^2$  noise corner frequency larger than 1 Hz (e.g. for  $R_f = 15 \text{ T}\Omega$ ,  $f_{1/f^2} = 5.3 \text{ Hz}$ ).

The noise contribution of  $R_f$  was not discussed in [25] or [68], both of which use similar architectures and similar values of  $R_f$ . The noise PSD plots included in both papers, however, clearly show the  $1/f^2$  noise at low frequencies.

### Noise from $R_{off}$

The noise analysis for  $R_{off}$  is identical to that of  $R_f$ . The remaining question is what the off resistance is. The switches labeled  $\phi_2$  in Fig. 4-8 create a resistance between the inputs of the opamp of  $R_{off,2}/2$ , since the two switches are in parallel, and the  $\phi_R$  switches create an off resistance of  $R_{off,R}$  between each input and  $V_{CM}$ .

The equivalent resistance in the single-ended model of Fig. 4-9 is, therefore, equal to

$$R_{off} = R_{off,R} || (R_{off,2}/4) = R_{off,sw}/5. \quad (4.29)$$

This last equality assumes the off resistances are equal, which is reasonable since all of the switches are constructed with equally sized transistors and biased with the same voltages. Specifically, each switch comprises an NMOS and PMOS device in parallel, each with a width  $W = 2 \mu\text{m}$  and length  $L = 180 \text{ nm}$ . The common mode voltage connected to the  $\phi_R$  switches is the same as the CMFB voltage  $V_{CM}$ . The transistors are off, so the gate of each NMOS device is tied to ground and the gate of each PMOS device is tied to  $V_{DD}$ .

To determine the off resistance of each transistor, we must solve for the differential change in current through them in response to a differential change in voltage. For an NMOS transistor biased in subthreshold, the drain-to-source current is

$$I_D = I_0 e^{\kappa V_{GB}/\phi_t} (e^{-V_{SB}/\phi_t} - e^{-V_{DB}/\phi_t}). \quad (4.30)$$

For PMOS transistors the subscript of each voltage is swapped. The current  $I_0$  depends on various constants as well as the width and length of the transistor. For our process (0.18  $\mu\text{m}$  CMOS) and transistor widths,  $I_{0n} = 30 \text{ pA}$  for the NMOS devices and  $I_{0p} = 8.3 \text{ pA}$  for the PMOS devices. Since negative feedback forces the difference between the input voltages of the opamps to be approximately equal to the input-referred DC offset (typically less than a few millivolts for the transistor sizes used), the drain and source voltages of each device is approximately  $V_{CM}$ .

As a quantitative example, let us assume that  $V_{DD} = 1.5 \text{ V}$  and  $V_{CM} = 0.75 \text{ V}$ . For a differential drain-source voltage of  $V_{DS} = 2 \text{ mV}$ , the drain current through each switch is

$$I_{Dn} + I_{Dp} = 38.3 \text{ pA} (e^{-749 \text{ mV}/25.9 \text{ mV}} - e^{-751 \text{ mV}/25.9 \text{ mV}}) = 785 \times 10^{-27} \text{ A}. \quad (4.31)$$

Dividing the 2 mV drain-source voltage by this small current results in a theoretical

value of

$$R_{off,sw} = \frac{2\text{mV}}{785 \times 10^{-27}\text{A}} = 2.5 \times 10^{21}\Omega, \quad (4.32)$$

and

$$R_{off} = R_{off,sw}/5 = 500 \times 10^{18}\Omega. \quad (4.33)$$

This ultra-high value for off resistance was verified through simulation. The ultra-low current resulting in such a high resistance is due to the negative  $V_{GS}$  voltage (for the NMOS) that practically depletes the channel of all carriers (similarly for the PMOS). Fortunately, this means that the noise from the off resistance of these transistors is negligible.

### Total Input Referred Noise

From the above analysis, we find that the two main noise contributors are the opamp and the feedback resistor. The total input-referred noise is

$$\overline{v_{n,i}^2} = \frac{4q\phi_t^2}{\kappa^2 I_D} \frac{\text{NEF}}{2.02} + \frac{K}{WLC_{ox}f} \alpha_f + \frac{4kT}{R_f} \left( \frac{1}{2\pi f C_i} \right)^2. \quad (4.34)$$

There is a thermal noise component that is inversely proportional to current, a  $1/f$  noise component from the opamp that is independent of the drain current, and a  $1/f^2$  component from the feedback resistor that is a function of the input capacitance and the feedback resistance. Since the only variable that can be changed after the chip is fabricated is the drain current  $I_D$ , measurements should show that the colored noise at lower frequencies is constant while the broadband noise gets smaller with increasing  $I_D$ .

Given the foregoing analysis, we can determine bounds on our design variables in order to get the desired performance. The theoretical limit for the thermal noise component (for a MOS opamp in subthreshold) requires at least 88 nA of drain current to achieve the 100 nV/ $\sqrt{\text{Hz}}$  desired. Assuming an NEF of 2.5, which has been shown to be achievable, the bias current of each transistor in the input pair of the opamp should be biased with at least 110 nA of current. The opamp's flicker

noise is more difficult to predict since the process constants  $K$  and  $C_{ox}$  vary between technologies. Also, since the only design variable available to the designer that can be used to reduce this noise is the area of the input transistors, a tradeoff is created between  $1/f$  noise and the parasitic capacitance. Recall from (4.21) that the total input referred noise of the opamp increases with  $C_p$ , so the area of the input pair cannot be increased arbitrarily. In our design, these transistors were sized with a width of  $W = 100 \mu\text{m}$  with  $L = 1 \mu\text{m}$ , keeping in mind that we had intended to use chopping in which case the flicker noise of the opamp would not have affected us. In designs that do not use chopping, these should be sized larger. Finally, for the desired noise level given, the product of  $C_i\sqrt{R_f}$  should be less than  $205 \times 10^{-6}$  so that the third term in (4.34) is sufficiently small at 1 Hz. Some examples of component values that would achieve this result are 1)  $C_i = 52 \text{ pF}$ ,  $R_f = 15 \text{ T}\Omega$ , 2)  $C_i = 30 \text{ pF}$ ,  $R_f = 47 \text{ T}\Omega$ , 3)  $C_i = 20 \text{ pF}$ ,  $R_f = 105 \text{ T}\Omega$ , and 4)  $C_i = 10 \text{ pF}$ ,  $R_f = 420 \text{ T}\Omega$ .

### Common-Mode and Differential-Mode Input Impedance

As detailed in section 4.2.4, having a high common-mode differential impedance  $Z_C$  is important as it reduces the conversion of common-mode interference to differential-mode interference caused by electrode impedance mismatch. Specifically, per (4.15), the amount of common-mode PLI that is converted to differential-mode PLI is  $\Delta Z_E/Z_C$ . It can be shown that the common-mode input impedance of the IAMP in Fig. 4-8 is

$$Z_C = Z_i + Z_f, \quad (4.35)$$

where  $Z_i = 1/sC_i$  is the input impedance network and  $Z_f = R_f/(1 + sC_fR_f)$  is the feedback network. Since  $Z_f \approx 1/sC_f$  for the frequencies of interest (i.e. above the high-pass cutoff frequency), and  $C_i \gg C_f$ , the common-mode input impedance is

$$Z_C \approx \frac{1}{j\omega C_f}. \quad (4.36)$$

For a 100 fF feedback capacitor, this means  $Z_C \approx 29 \text{ G}\Omega$  at 60 Hz. Of course parasitic capacitances in the PCB traces and IC bondpads will probably cause this impedance to be smaller, but it is clear the IAMP itself will not contribute significantly.

Having a large differential-mode impedance  $Z_D$  is also important for some applications where the electrodes' impedance is large, such as dry electrodes. The differential-mode signal at the input of the IAMP is reduced by the voltage dividers between  $Z_D$  and the electrodes' impedance  $Z_E$ . Specifically,  $v_d = v_s(Z_D/(2Z_E + Z_D))$ . To limit this attenuation,  $Z_D$  should be at least  $10\times$  larger than  $Z_E$ . Since the input to the opamp in Fig. 4-8 presents a low differential impedance due to feedback, the differential impedance of the IAMP is the series combination of the two input capacitors:

$$Z_D = \frac{2}{j\omega C_i}. \quad (4.37)$$

Having a capacitive input impedance to the IAMP is beneficial since it is much larger at lower frequencies where the electrodes' impedance is typically largest (it can be modeled as a large capacitance down to low frequencies). For an input capacitance of  $C_i = 10 \text{ pF}$ ,  $Z_D = 32 \text{ G}\Omega$  at 1 Hz, which is typically the lowest frequency of interest for most physiological measurements. This is significantly larger than the required value of  $80 \text{ M}\Omega$  specified in Section 4.2.4.

## 4.4 Sinc Anti-Aliasing Filter

Almost all modern biomedical systems digitize the signal of interest using an ADC as shown in Fig. 4-1. To avoid corrupting the desired signal, an *anti-aliasing filter* is used to attenuate aggressors and noise components that are frequency translated to the signal band through the sampling process. Fig. 4-12 (A) shows an example continuous-time spectrum (before sampling) where the desired signal is illustrated in white and the spectral components that would be aliased onto the discrete-time signal band are in gray.

To avoid aliasing in the desired channel bandwidth, it is sufficient for the anti-aliasing filter attenuate signals inside of the bands illustrated as light gray boxes

in Fig. 4-12 (A). Fig. 4-12 (B) shows a typical anti-aliasing filter response, where attenuation is sufficient at the  $\pm f_H = \pm(f_s - f_L)$  band edges. Since for most analog low-pass filters the amplitude response continues to roll off beyond this frequency, there is typically more than enough filtering at the higher frequency aliasing terms than necessary. Fig. 4-12 (C) shows the bare minimum required anti-aliasing filter response. The dashed squares represent frequency bands that are not filtered; the signal band is in white, and the noise bands are in dark gray. Note that the frequency bands illustrated in light gray in Fig. 4-12 (A) are filtered out. Although the dark gray bands will be aliased, they will not corrupt the discrete-time signal since they will fall outside of the signal band. These bands can be filtered out in the digital domain, where sharper filters can be implemented more efficiently (with respect to power and area). The filter described in this section takes advantage of this relaxation in the anti-aliasing specifications by using notches that attenuate the aliasing terms. The result is a smaller filter with various beneficial features.

### **Taking Advantage of Over-Sampling**

As mentioned previously, recent advances in ADC design have led to ultra-low energy consumption per converted bit. One of the strategies used in this work to reduce the area and energy consumption of the whole system, is to take advantage of oversampling to relax the specifications of other blocks. The two main benefits to oversampling are a reduction in quantization noise of 3 dB per  $2\times$  increase in sampling rate and relaxed specifications for the anti-aliasing filter.

As discussed previously, the FOM (energy per conversion bit) of ADCs is often constant over a range of sampling frequencies. However, if the sampling rate is reduced too much, leakage currents and DC bias currents will emerge as the dominant energy loss mechanism and lead to worse FOM values. Also, to achieve higher accuracy (ENOB), it is often necessary to use larger capacitors to improve matching, but this leads to larger area and higher energy consumption. It can, therefore, be advantageous to achieve a given signal-to-noise ratio for the digitized signal by using a lower ADC ENOB while oversampling to maintain the same quantization noise. For example,



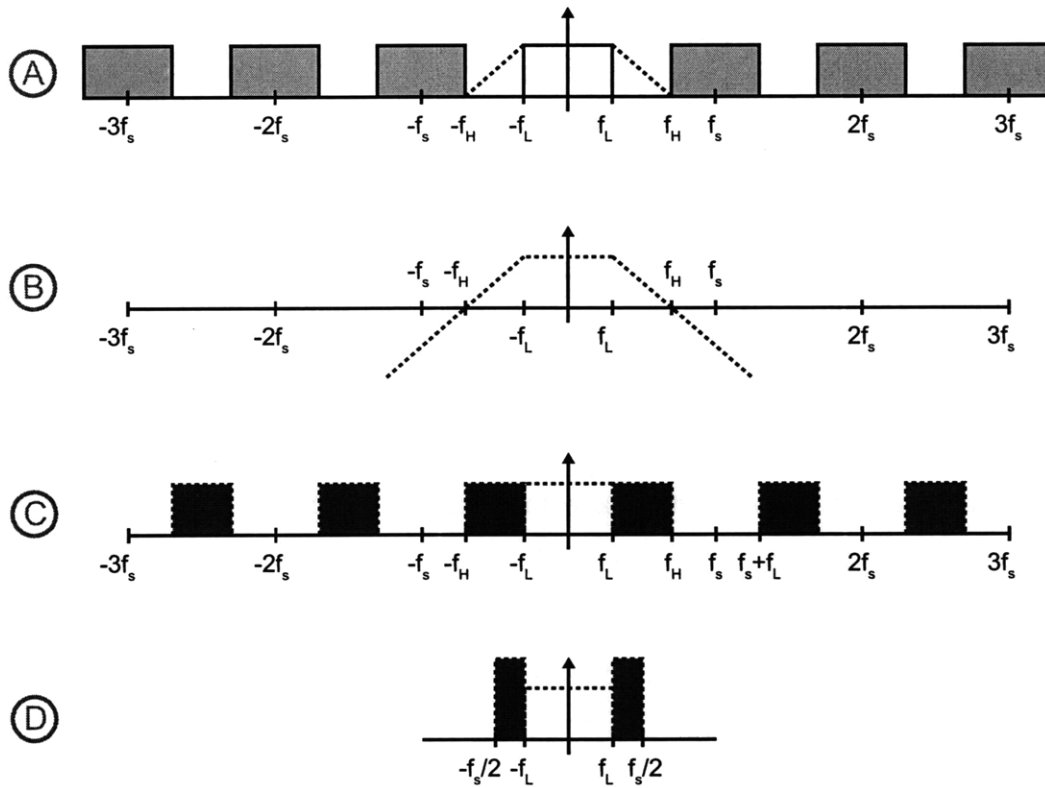


Figure 4-12: (A) Frequency-domain illustration of desired signal band (white) and aliasing components (gray) that would corrupt the signal after sampling. (B) A typical anti-aliasing filter amplitude response that filters aliasing components by some required amount. (C) The minimum required amplitude response of an anti-aliasing filter. The dark gray boxes represent bands that can be aliased without corrupting the signal band. (D) Discrete-time spectral component between  $-f_s/2$  and  $f_s/2$  illustrating how the desired channel is not corrupted and the non-filtered components of (C) are frequency translated to the bands between the channel edges and the Nyquist frequencies.

if a signal occupies a bandwidth of 100 Hz, and 10 bits of accuracy are desired for Nyquist sampling (200 S/s), sampling at  $64\times$  the Nyquist rate (12.8 kS/s) allows the use of an ADC with an ENOB of 7. An ADC with an ENOB of 7 can be significantly smaller than an ADC with an ENOB of 10. For an ADC with an FOM of 60 fJ/bit, an ENOB of 7, and a sampling rate of 12.8 kS/s, the total power consumption is only 100 nW, which is typically insignificant compared to the overall power budget

The second benefit of oversampling (relaxation of the anti-aliasing filter's requirements) can be exploited to reduce the area of the system. Some oversampling is always necessary to avoid aliasing, since analog anti-aliasing filters roll off at a finite rate. As an example of how oversampling can result in a smaller anti-aliasing filter, let us assume that the signal bandwidth is  $f_L = 100$  Hz and the sampling rate is  $f_s = 300$  S/s. Let us further assume that a minimum of 40 dB of attenuation is required at  $f_H = f_s - f_L = 200$  Hz with a maximum of 3 dB attenuation at the pass-band edge. These specifications would require a 7<sup>th</sup> order filter if a Butterworth response was desired. With common biquad implementations, this would require four stages, as well as a number of bulky passive components since the frequency constants of the filter would be on the order of 100 Hz.

Let us now consider an example where the sampling rate is  $f_s = 12.8$  kS/s, resulting in  $f_H = 12.7$  kHz. The same 40 dB of attenuation can now be achieved with a single pole at 127 Hz, two poles near 1.27 kHz, or three poles near 2.74 kHz. The first of these options has the benefit that it could be achieved with a passive RC filter, although the low corner frequency would require large components (for example,  $R = 12.5$  M $\Omega$ ,  $C = 100$  pF). The second and third options would require a biquad, which would add power consumption and more passive components, but their values would be significantly smaller resulting in smaller area. Clearly, any of these options would result in significantly less power and area consumption than if sampling were done at the lower rate of 300 S/s since fewer opamps and passive components would be necessary.

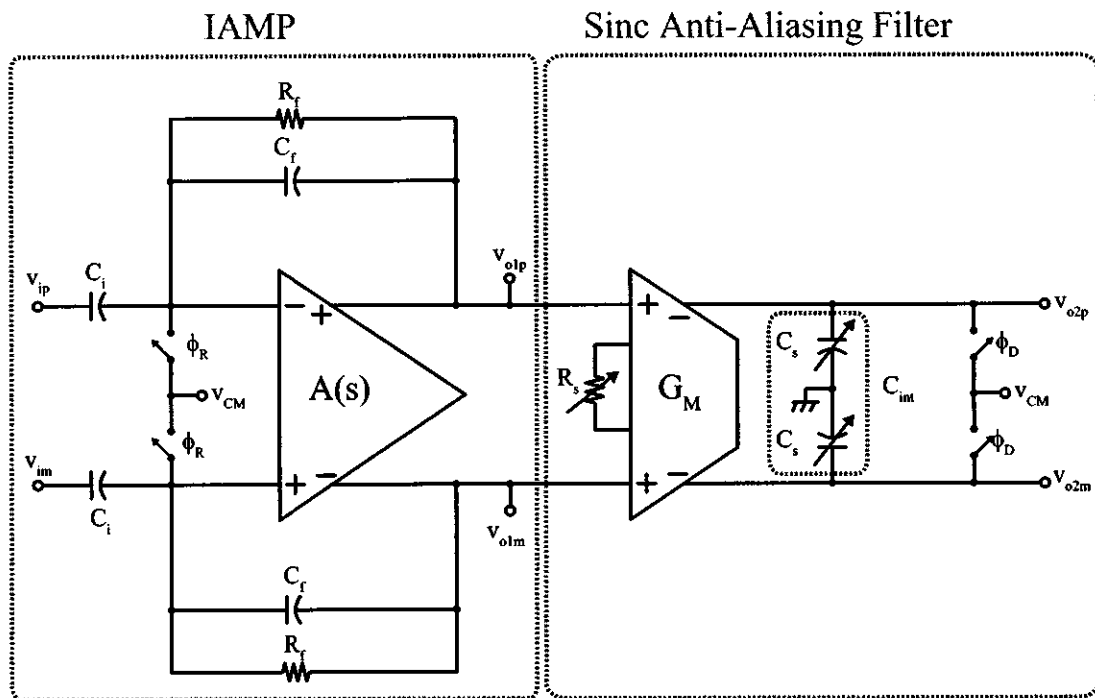


Figure 4-13: Simplified schematic of the instrumentation amplifier and SAFF. The output voltage of the SAFF,  $V_{o2}$  is periodically reset to  $V_{CM}$  using the clock signal  $\phi_D$ . This *integrate-and-dump* or *charge-sampling* technique results in a sinc frequency response.

## Using Charge-Sampling to Achieve Anti-Aliasing

As shown in Fig. 4-12 (C), anti-aliasing does not require a low-pass filter response. It only requires that the aliasing components of the signal be attenuated; a goal that can be achieved with strategically placed notches. While notches can be implemented with continuous-time filters, *charge-sampling* has the benefit that it concurrently effects a *sinc*-type frequency response while sampling the input signal [11], [73]. In contrast with voltage-sampling systems where a track-and-hold circuit is used, charge-sampling uses a transconductor to convert the input voltage into a current, integrates that current onto a capacitor for some time, and samples the resulting voltage at the end of the period. Fig. 4-13 shows an IAMP followed by a charge-sampling system.

The two capacitors, each labeled  $C_s$  are each connected to ground to keep the output common-mode signal from drifting significantly between samples. Their effective differential-mode capacitance is  $C_{int} = C_s/2$ . Each period starts by quickly resetting the voltage across each capacitor to the common-mode voltage  $V_{CM}$ , such that the differential voltage across  $C_{int}$  is zero. The transconductor, labeled  $G_M$ , converts the IAMP's output,  $v_{o1}$  to a current equal to  $i_o = v_{o1}G_m$ , and the current is integrated onto  $C_{int}$  ( $R_s$  is a physical resistor used to set  $G_m$  and  $G_m$  is inversely proportional to  $R_s$ ). At the end of the  $n^{th}$  cycle, ( $t = n \cdot T_s$ ), the voltage is

$$v_{o2}[n] = v_{o2}(n \cdot T_s) = \frac{G_m}{C_{int}} \int_{(n-1)T_s}^{nT_s} v_{o1}(\tau) d\tau. \quad (4.38)$$

To understand the effect of this process on the spectrum of  $v_{o2}$ , it is useful to think of it as a two-step process that is mathematically equivalent. In the first step,  $v_{o1}$  is convolved with a rectangular impulse response

$$h_{\Pi}(t) = \frac{1}{T_{int}} \Pi\left(\frac{t - T_s/2}{T_s}\right), \quad (4.39)$$

where  $\Pi(x) = 1$  for  $|x| < 1/2$ , and zero otherwise, and  $T_{int} = C_{int}/G_m$ . The spectrum

of the resulting intermediate signal is

$$V_{o2,cnt}(f) = V_{o1}(f)H_{\Pi}(f) = V_{o1}(f)\frac{T_s}{T_{int}}\text{sinc}(\pi T_s f)e^{-\pi f T_s}. \quad (4.40)$$

In the second step, the intermediate voltage is sampled at times  $n \cdot T_s$  ( $v_{o2}[n] = v_{o2,cnt}(n \cdot T_s)$ ), such that the discrete-time spectrum is

$$V_{o2}(f) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} V_{o2,cnt}(f - kf_s). \quad (4.41)$$

Charge sampling is, therefore, equivalent to passing a continuous-time signal through a *sinc* filter before sampling it. The notches of the *sinc* filter land at integer multiples of  $f_s$ , which is precisely what an anti-aliasing filter should do. Furthermore, they are independent of component values such as resistors and capacitors, mitigating the effects of component variations and allowing the use of smaller passives. In contrast with classical analog filters, variations of  $R_s$  and  $C_{int}$  only affect the gain of the block, but not the frequency response. While an accurate amount of gain may be important, it can often be calibrated and will not change significantly in implants due to the inherent temperature regulation of the human body.

Instead of affecting the frequency response,  $G_m$  and  $C_{int}$  set the gain of the *sinc* filter. Specifically, the gain of this block is  $T_s/T_{int}$ , which means that smaller values of  $R_s$  and  $C_{int}$  lead to higher gain. It also means that programmable resistors and capacitors can be used to implement variable gain. As a quantitative example, for a desired minimum gain of unity, and a sampling rate of 12.8 kS/s, values of  $G_m = 1/1.6 \text{ M}\Omega$  and  $C_{int} = 10 \text{ pF}$  can be used. Compared with previous examples, the passive components necessary for this technique are significantly smaller than the examples given earlier, even for oversampling systems.

Another benefit of the charge-sampling frequency response is that its only effect on the signal's phase is a delay of  $T_s/2$ . In other words, the group delay is constant. Recall also that for systems where chopping is used to reduce the effects of flicker noise, the DC offset of the opamp is up-converted to the chopping frequency  $f_{ch}$ . If



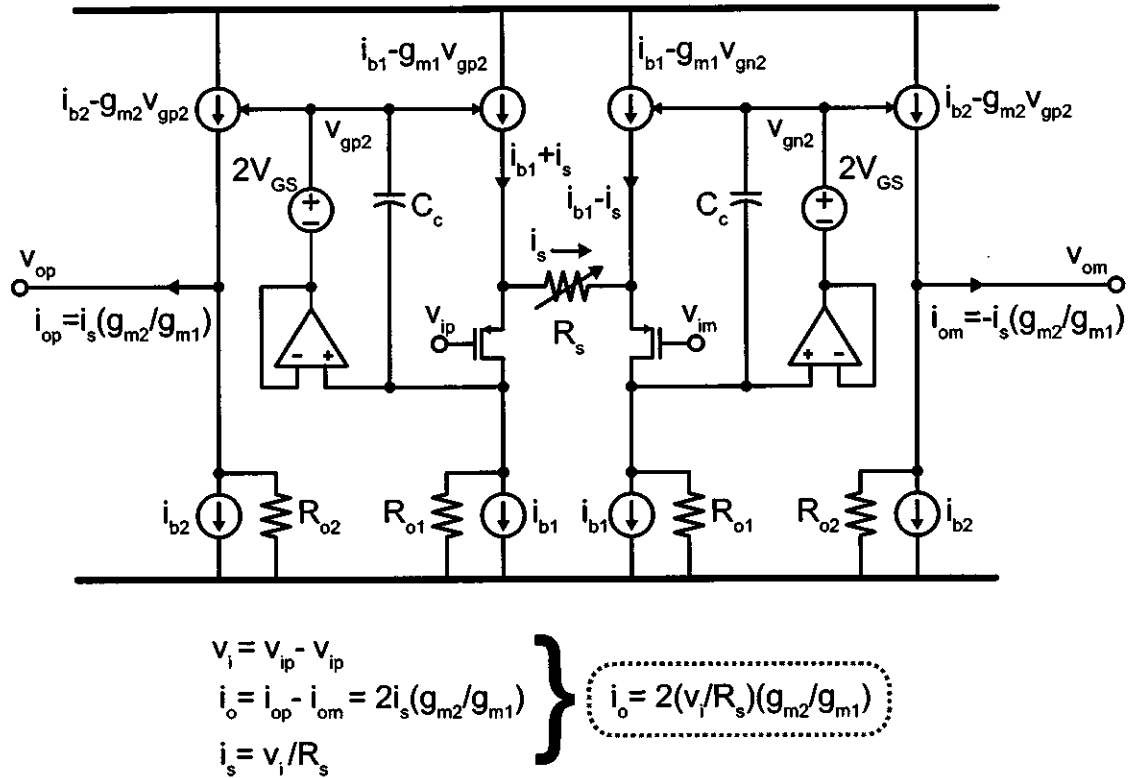


Figure 4-15: Simplified schematic of transconductor used in the SAFF. The effective transconductance is  $G_m = 2/R_s \times g_{m2}/g_{m1}$ , where  $R_s$  is digitally tunable. The effective output resistance is a function of the second-stage transistors. Their value is made large through cascoding to prevent performance degradation.

have the effect of lowering the source resistance of the  $M_1$  transistors which act as source followers. As a result, the voltage across  $R_s$  is  $v_i = v_{ip} - v_{im}$  and the current flowing through it is  $i_s = v_i/R_s$ . When  $v_i$  is increased, the feedback loop lowers the gate voltage of  $M_5$  to supply the necessary extra current, and in turn increases the current of  $M_9$  by a proportional amount. As shown in the figure the output current is

$$i_o = i_{op} - i_{om} = 2i_s(g_{m2}/g_{m1}) = 2(v_i/R_s)(g_{m2}/g_{m1}), \quad (4.42)$$

where the ratio  $g_{m2}/g_{m1}$  is there to illustrate that current multiplication can be done by proper sizing of  $M_5$  and  $M_9$ . The effective transconductance, therefore, is

$$G_m = \frac{i_o}{v_i} = \frac{2}{R_s} \frac{g_{m2}}{g_{m1}}. \quad (4.43)$$

In addition to the transconductance of this circuit, the output resistance is also important. Recall that the sinc filter analysis shown previously assumes that the transconductor,  $G_M$ , and capacitor,  $C_{int}$ , form an integrator. This is only true if the output resistance of the transconductor is infinite, which is not possible. However, as long as the time constant  $\tau_{int} = R_o C_{int}$  is significantly larger than  $T_s$ , the finite resistance will not affect the performance significantly. This can be illustrated with a simple example. The filter can be modeled as a current source in parallel with the output resistance  $R_o$  and capacitance  $C_{int}$ . Assuming  $v_o(t) = 0$  at time  $t = 0$ , an input step signal  $v_i(t) = V_I u(t)$ , results in the output voltage  $v_o(t) = V_I G_m R_o (1 - e^{-t/\tau_{int}})$ . For small values of  $t/\tau_{int}$ , the exponential term can be approximated as  $e^{-t/\tau_{int}} \approx 1 - t/\tau_{int}$ , such that

$$v_o(t) \approx \frac{V_I G_m R_o t}{\tau_{int}} = \frac{V_I G_m t}{C_{int}}, \quad (4.44)$$

which is what we would expect for an ideal integrator. Since the maximum amount of time that the filter integrates for is  $T_s$ , the condition  $t/\tau_{int} \ll 1$  requires that  $\tau_{int} \gg T_s$ .



The small-signal output resistance of a MOS transistor is

$$r_o = \frac{1}{\lambda I_D}. \quad (4.45)$$

By using cascoding, this resistance can be significantly increased, and the effective output resistance of the transconductor is

$$R_o \approx g_{m,p} r_{o,p}^2 \parallel g_{m,n} r_{o,n}^2 = \frac{\kappa}{I_D \phi_t} \left( \frac{1}{\lambda_p^2 + \lambda_n^2} \right), \quad (4.46)$$

where  $\lambda_p$  and  $\lambda_n$  are the respective output resistant constants of the PMOS and NMOS devices and relate to channel-length modulation. These can be made small by using long-channel devices. As a numerical example, for  $I_D = 20$  nA,  $\kappa = 0.7$ , and  $\lambda_p = \lambda_n = 0.1$ ,  $R_o = 68$  G $\Omega$ . For a value of  $C_{int} = 10$  pF,  $\tau_{int} = 0.68$  s, which is nearly four orders of magnitude larger than  $T_s$  when the sampling rate is  $f_s = 12.8$  kS/s.

### Choosing the Bias Current of the Filter

In our prototype implementation,  $g_{m1} = g_{m2}$ , such that  $G_m = 2/R_s$ . This also means that  $i_b = i_{b1} = i_{b2}$  in Fig. 4-15 and the total current is approximately

$$i_{Gm} \approx 4 \times i_b. \quad (4.47)$$

When designing the IAMP, the bias current was set to meet the input-referred noise requirements. In the filter, the input-referred noise is not a major issue because the IAMP provides so much gain. Instead, the minimum bias current is set by linearity requirements. This is because  $\max(i_s) = i_b$ , and therefore,  $\max(v_i) = i_b R_s$ .

Recall from (4.40) that the low-frequency gain of the filter is

$$G_\phi = \frac{T_s}{T_{int}} = \frac{G_m}{C_{int} f_s} = \frac{1}{R_s C_s f_s}, \quad (4.48)$$

where  $C_s = 2C_{int}$ . As just discussed, if  $v_i \geq i_b R_s$ , current clipping occurs in the

Table 4.2: Programmable  $R_s$  values

$b_R[2:0]$	$R_s$	$\min(G_\phi)$	$\max(G_\phi)$
000	20 M $\Omega$	0.277 (-11 dB)	13.0 (22 dB)
001	14 M $\Omega$	0.396 (-8.1 dB)	18.6 (25 dB)
010	10 M $\Omega$	0.554 (-5.1 dB)	26.0 (28 dB)
011	7 M $\Omega$	0.792 (-2.0 dB)	28.3 (31 dB)
100	5 M $\Omega$	1.11 (0.9 dB)	52.1 (34 dB)
101	3.5 M $\Omega$	1.58 (4.0 dB)	74.4 (37 dB)
110	2.5 M $\Omega$	2.22 (6.9 dB)	104 (40 dB)
111	2.0 M $\Omega$	2.77 (8.9 dB)	130 (42 dB)

filter's  $G_M$  stage. For a given filter gain setting,  $G_\phi$ , there is also a limit to how large  $v_i$  can be before the output of the filter saturates the ADC. Since the output of the filter is the input of the ADC, its output should be smaller than the full range of the ADC (i.e.  $|v_o| \leq V_{refA}$ ). On the other hand, it is beneficial to set the gain as large as possible so that as much of the ADC's range is used. The optimum setting for the filter's gain, therefore, is when it is as large as possible, but satisfies this condition:

$$G_\phi \leq \frac{V_{refA}}{\max(v_i)}, \quad (4.49)$$

where  $\max(v_i)$  is the largest input expected. This means that  $i_b$  should be slightly larger than  $\max(v_i)/R_s$ , since making it any larger will not help the overall system linearity. Combining these two observations leads conclusion that  $i_b$  should be made as small as possible, while still meeting this condition:

$$i_b \geq V_{refA}C_s f_s. \quad (4.50)$$

### Programming the Filter's Gain

As mentioned before, the filter's gain  $G_\phi$  can be made variable by programming  $C_s$  and  $R_s$ . In our prototype,  $R_s$  is programmed with a 3-bit digital word, and  $C_s$  is programmed with a 5-bit digital word for a total of 256 possible choices. The value

of  $C_s$  is set as follows

$$C_s = 7.8 \text{ pF} \cdot b_C[4] + 3.9 \text{ pF} \cdot b_C[3] \\ + 1.2 \text{ pF} \cdot b_C[2] + 0.6 \text{ pF} \cdot b_C[1] + 0.3 \text{ pF} \cdot b_C[0] + 0.3 \text{ pF}. \quad (4.51)$$

Its tuning range is 300 fF–14.1 pF. Table 4.2 shows the values of  $R_s$  that can be programmed in our prototype, along with the minimum and maximum gain that can be achieved for that setting assuming  $f_s = 12.8$  kS/s. From (4.47) and (4.50) we can approximate the range of current consumption of the filter using the minimum and maximum values for  $C_s$ . Assuming  $V_{refA} = 500$  mV and  $f_s = 12.8$  kS/s, we find that the current consumption of the filter is  $i_{Gm} \in [7.7 \text{ nA}, 361 \text{ nA}]$ . Fig. 4-16 shows the simulated gain of the filter for all  $C_s$  and  $R_s$  values available. The x-axis corresponds to the  $R_s$  selection. The top line corresponds to the smallest value of  $C_s$ , and lower lines correspond to incrementally larger selections of  $C_s$ . Since there is overlap in the amount of gain that different settings provide, we can carefully choose a selection that achieve a desired step size while minimizing the current.

Fig. 4-17 plots selected values for increasing gain in roughly 1 dB steps. All of the gain settings between  $G_\phi \in [4 \text{ dB}, 42 \text{ dB}]$  can be achieved with the 3 LSB of  $b_C$ , which means that the maximum capacitance they require is 2.4 pF. As a result, for most of the gain settings the total bias current is approximately  $i_{Gm} \leq 62$  nA. The settings between  $G_\phi \in [-3 \text{ dB}, 4 \text{ dB}]$  require a bit more power, but can be implemented with  $i_{Gm} \leq 150$  nA.

## 4.5 Mixed-Signal Feedback for Interference Cancellation

### 4.5.1 Motivation

Power consumption in a signal processing system is often determined by dynamic range requirements. The dynamic range is a measure of the ratio between the largest

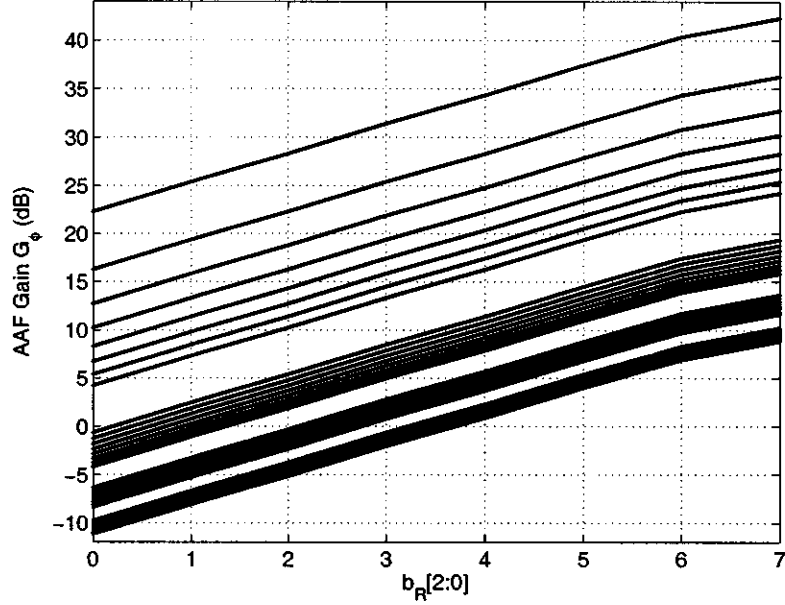


Figure 4-16: Filter gain  $G_\phi$  vs.  $R_s$  setting  $b_R$  for each  $C_s$  setting  $b_C$ . The top line corresponds to a value of  $b_C = 0$ , and lower lines correspond to higher values of capacitance.

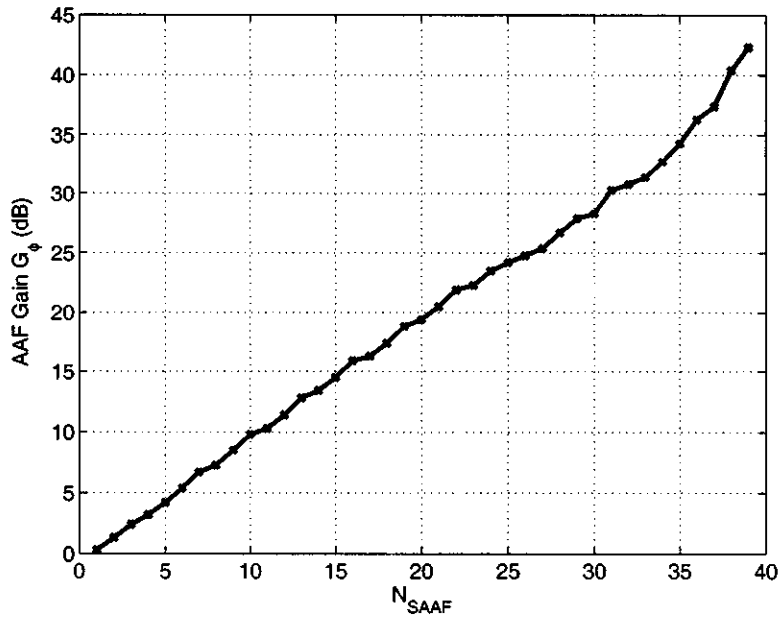


Figure 4-17: Filter gain  $G_\phi$  for strategically selected values of  $b_C$  and  $b_R$ . As shown, the gain can be programmed in approximately 1 dB steps. Finer steps are possible at lower gain settings.

signal that can be handled by the system without significant distortion and the minimum detectable signal set by the input-referred noise:

$$\text{DR}_{\text{dB}} = 20 \log \left( \frac{\max(v_{i,rms})}{v_{ni,rms}} \right). \quad (4.52)$$

The specifications for the minimum detectable signal are typically set by the signal being measured, but  $\max(v_{i,rms})$  it often set by interference. If that interference can be eliminated near the front end of the system, the dynamic range requirements of subsequent blocks can be relaxed and their power consumption reduced.

The term “significant distortion” can be defined in different ways depending on the application [29, 53]. While there is no standard protocol that defines “significant distortion” for all biopotential measurements, at least one commonly referenced paper uses *total harmonic distortion* (THD) as an appropriate measure [25].

THD relates to the amount of nonlinearity introduced by a system for a given input amplitude. For a sinusoidal input signal, the output spectrum should only include a tone at the signal’s frequency. The ratio between the power of all other harmonics introduced by nonlinearities and the fundamental tone is defined as

$$\text{THD} = \frac{P_2 + P_3 + \dots}{P_1}. \quad (4.53)$$

According to [25], a value of THD < 1% is sufficient for most biomedical applications, and the maximum signal that achieves this metric is considered the  $\max(v_{i,rms})$  term in (4.52).

Systems that employ negative feedback can achieve low distortion even for relatively large signals. For practically all systems, however, the maximum input and output signal must have a peak-to-peak amplitude that is smaller than the voltage supply range. As discussed in the previous section, the linearity of the filter is also limited by the bias current. This implies that, at least to some extent, the minimum supply voltage of the system and supply current of certain blocks are determined by  $\max(v_{i,rms})$ . This also implies that a reduction in  $\max(v_{i,rms})$  may result in power savings throughout the chain will

When optimizing a system for minimum power and area, it is important to consider whether  $\max(v_{i,rms})$  is set by the desired signal or some sort of interference. If  $\max(v_{i,rms})$  is set by an interfering signal and is filtered out near the front end, the dynamic range requirements of subsequent stages will be lower. As an example, EEG signals can range in amplitude between  $10\mu V_{p-p}$ – $500\mu V_{p-p}$ . PLI interference, on the other hand, can be larger than  $3\text{ mV}_{p-p}$ . Assuming that  $v_{ni,rms} = 1\mu V_{rms}$ , the dynamic range of the front-end must be better than 60 dB in order to handle the PLI ( $DR = 20\log(1\text{ mV}/1\mu V)$ ). On the other hand, if the PLI is filtered out near the front-end,  $\max(v_{i,rms})$  gets reduced to  $177\mu V_{rms}$ , and the dynamic range is reduced to 45 dB; a 15 dB difference.

These observations motivate the design of a system that filters out large interferers near the front-end, while minimally impacting the signal band of interest. In this work special emphasis is placed on PLI, but for certain applications other interferers may be present at different frequencies and of different bandwidths. It is therefore important for the interference cancelation method to be easily configurable.

## 4.5.2 PLI Considerations

Since PLI often resides in the frequency band of interest, filtering it usually requires a narrow notch that does not affect the important characteristics of the signal. This is often done digitally, since digital filters are often less area and power intensive [7, 38, 39, 58]. However, as discussed in the previous subsection, when the PLI is significantly larger than the signals of interest, eliminating it before digitization can yield benefits in power consumption. This motivates the development of a narrow notch filter that can filter PLI near the front end.

Analog implementations of notch filters are possible, but typically require multiple opamps and passive components [56]. Furthermore, since the notch must be narrow (i.e. less than 1 Hz wide at 60 Hz), relatively high Q is necessary leading to large component ratios. Further still, process, voltage, and temperature variations will almost certainly cause the notch to shift which will require periodic calibration or feedback of some sort.

A clever analog system is proposed in [27] that employs a quadrature feedback technique using analog multipliers, a phase-locked loop, and integrators to cancel PLI at the input of the system. This technique achieved about 54 dB of attenuation on the PLI with a notch bandwidth of 0.6 Hz. This implementation required the use of 10 M $\Omega$  resistors and 100  $\mu$ F capacitors, component values that are not compatible with integrated circuit design. In addition, the system requires the use of multiple opamps, linear multipliers, and a phase-locked loop. All of these are potentially power hungry circuits that are not compatible with ultra-low power design (actual power consumption numbers were not given).

In summary, neither digital post-processing nor purely analog solutions seem compatible with ultra-low-power, small area, fully integrated systems that cancel interference. The following section describes a mixed-signal system that achieves interference cancelation in an area and energy efficient way.

### 4.5.3 Prototype System Design

We proposed and prototyped a system that is conceptually similar to [27], but exploits the benefits of mixed-signal design to minimize size and power. Fig. 4-18 (A) shows a schematic of the system and (B) shows the block diagram. The system includes the IAMP and anti-aliasing filter described in previous sections along with an on-chip DAC, an off-chip ADC, and a field programmable gate array (FPGA). The first step to understanding the closed-loop response of this system is to derive the transfer function of the digital section implemented on the FPGA. The block diagram includes a transfer function for the digital section that will be derived in the next section.

#### Digital Section

The input to the digital section is labeled  $x[n]$  in Fig. 4-18 and the output is labeled  $d[n]$ . The top path of the digital block uses the  $I$  subscripts to denote that it is processing the “in-phase” component of  $x[n]$ , whereas the bottom path uses the  $Q$  subscript to denote it is processing the “quadrature” component. The in-phase





component is given by

$$x_I[n] = x[n] \cos(\Omega_0 n), \quad (4.54)$$

which is passed through an accumulator resulting in

$$y_I[n] = \sum_{k=N_0}^n x_I[k] = \sum_{k=N_0}^n x[k] \cos(\Omega_0 k). \quad (4.55)$$

$N_0$  is the sample time at which the system is turned on, and  $y_I[n] = 0$  at  $n = N_0$ .<sup>1</sup> The multiplication of  $x[n]$  with  $\cos(\Omega_0 n)$  effects frequency translation, while the accumulation effects low-pass filtering with infinite gain at DC. This can be shown in the frequency domain by using the Fourier transform. The spectrum of  $x_I[n]$  is

$$X_I(e^{j\Omega}) = \frac{1}{2\pi} X(e^{j\Omega}) * \mathfrak{F}\{\cos(\Omega_0 n)\}, \quad (4.56)$$

where

$$\mathfrak{F}\{\cos(\Omega_0 n)\} = \pi\delta(\Omega - \Omega_0) + \pi\delta(\Omega + \Omega_0), \quad (4.57)$$

is the Fourier transform of  $\cos(\Omega_0 n)$ . As a result,

$$X_I(e^{j\Omega}) = \frac{1}{2} X(e^{j(\Omega - \Omega_0)}) + \frac{1}{2} X(e^{j(\Omega + \Omega_0)}), \quad (4.58)$$

which is a frequency translated version of  $X(e^{j\Omega})$ . Next, we can rewrite  $y_I[n]$  as

$$y_I[n] = y_I[n - 1] + x_I[n], \quad (4.59)$$

which has the Z-transform

$$Y_I(z) = \frac{X_I(z)}{1 - z^{-1}}. \quad (4.60)$$

Evaluating (4.60) for  $z = e^{j\Omega}$  yields the Fourier transform of the transfer function

$$H_I(e^{j\Omega}) = \frac{Y_I(e^{j\Omega})}{X_I(e^{j\Omega})} = \frac{1}{1 - e^{-j\Omega}}. \quad (4.61)$$

---

<sup>1</sup>The convention in [47] is to use the  $\omega$  variable for discrete-time signals and  $\Omega$  for continuous-time signals. Here, we do the opposite because most of the analysis is done in the continuous time where the  $\omega$  variable is typically used to denote angular velocity.

For  $\Omega = 0$ , (4.61) evaluates to infinity. As  $\Omega$  increases,  $H_I(e^{j\Omega})$  becomes smaller, reaching a minimum of  $1/2$  at  $\Omega = \pi$ .

Combining (4.61) and (4.58) gives the result for  $Y_I(e^{j\Omega})$

$$Y_I(e^{j\Omega}) = \frac{1}{2} \frac{X(e^{j(\Omega-\Omega_0)}) + X(e^{j(\Omega+\Omega_0)})}{1 - e^{-j\Omega}}. \quad (4.62)$$

The output of the “in-phase” path,  $w_I[n]$  has a Fourier transform of

$$W_I(e^{j\Omega}) = \frac{G_f}{2\pi} Y_I(e^{j\Omega}) * \mathfrak{F}\{\cos(\Omega_0 n)\} = \frac{G_f}{2} \left( Y_I(e^{j(\Omega-\Omega_0)}) + Y_I(e^{j(\Omega+\Omega_0)}) \right). \quad (4.63)$$

Substituting (4.62) into (4.63) and rearranging terms gives relationship between the spectra of  $x[n]$  and  $w_I[n]$

$$W_I(e^{j\Omega}) = \frac{G_f}{4} \left[ X(e^{j\Omega}) \left( \frac{1}{1 - e^{-j(\Omega-\Omega_0)}} + \frac{1}{1 - e^{-j(\Omega+\Omega_0)}} \right) + \frac{X(e^{j(\Omega-2\Omega_0)})}{1 - e^{-j(\Omega-\Omega_0)}} + \frac{X(e^{j(\Omega+2\Omega_0)})}{1 - e^{-j(\Omega+\Omega_0)}} \right]. \quad (4.64)$$

Following the same steps for the “quadrature-phase” path, results in the following equation for the spectrum of  $w_Q[n]$

$$W_Q(e^{j\Omega}) = \frac{G_f}{4} \left[ X(e^{j\Omega}) \left( \frac{1}{1 - e^{-j(\Omega-\Omega_0)}} + \frac{1}{1 - e^{-j(\Omega+\Omega_0)}} \right) - \frac{X(e^{j(\Omega-2\Omega_0)})}{1 - e^{-j(\Omega-\Omega_0)}} - \frac{X(e^{j(\Omega+2\Omega_0)})}{1 - e^{-j(\Omega+\Omega_0)}} \right]. \quad (4.65)$$

Adding (4.64) and (4.65) and rearranging some terms gives the transfer function between  $x[n]$  and  $d[n]$

$$H_D(e^{j\Omega}) = \frac{D(e^{j\Omega})}{X(e^{j\Omega})} = \frac{G_f}{2} \left( \frac{1}{1 - e^{-j(\Omega-\Omega_0)}} + \frac{1}{1 - e^{-j(\Omega+\Omega_0)}} \right). \quad (4.66)$$

This exact transfer function can be simplified by making a simple approximation. Recall that this system employs oversampling with a high oversampling ratio ( $>100$ ).

For a sampling rate of  $f_s=10$  kS/s, all of the frequencies of interest are much smaller than the sampling rate. The relationship between the discrete-time frequency  $\Omega$  and the continuous-time frequency  $f$  is

$$\Omega = \frac{\omega}{f_s} = \frac{2\pi f}{f_s}. \quad (4.67)$$

This means that for frequencies  $f \ll f_s$ ,  $\Omega \pm \Omega_0 \ll 1$ , which allows us to approximate the exponential terms in (4.66) by the first two terms of their Taylor series expansions (i.e.  $e^{-jx} \approx 1 - jx$  for  $x \ll 1$ ). This simplifies (4.66) to

$$H_D(e^{j\Omega}) = \frac{G_f}{j} \left( \frac{\Omega}{\Omega^2 - \Omega_0^2} \right). \quad (4.68)$$

This transfer function is similar to that of an integrator, but it becomes infinite at  $\Omega = \pm\Omega_0$  instead of  $\Omega = 0$ , and decays to zero as  $\Omega$  approaches zero and infinity.

### Closed-Loop Transfer Function

Fig. 4-18 (B) shows the block diagram of the system in (A).  $G_I$  is the gain of the IAMP ( $\approx 100$ ) and  $G_\phi$  is the gain of the sinc anti-aliasing filter, which is programmable from 0.7 to 130. A 16-bit ADC (AD7684 from Analog Devices) was used and its gain,  $G_A$ , is defined as

$$G_A = \frac{2^{16} - 1}{2V_{refA}}, \quad (4.69)$$

where  $V_{refA}$  is the reference voltage used to determine the full-scale range of the ADC (i.e.  $-V_{refA} \leq v_a(t) \leq V_{refA}$ ). Although a lower resolution ADC can be used to meet the specifications discussed earlier, a high resolution ADC was used so that the rest of the system could be characterized without its performance being impacted by flaws in the ADC. An on-chip 8-bit DAC was used and its gain is

$$G_D = \frac{1}{4} \left( \frac{2V_{refD}}{2^8 - 1} \right). \quad (4.70)$$

The 1/4 term arises from the implementation of the DAC as will be discussed later.

In the previous section we found the discrete-time transfer function of the digital block. In the feedback loop, the AAF acts as both an anti-aliasing filter for the continuous-to-discrete-time conversion, and as a reconstruction filter for the discrete-to-continuous-time conversion. This allows us to treat the interface between the discrete-time and continuous time domains by simply substituting (4.67) into (4.68). After some algebra, the resulting equivalent continuous-time transfer function is given by

$$H_D(s) = -G_f f_s \left( \frac{s}{s^2 - \omega_0^2} \right), \quad (4.71)$$

The loop gain of the system is, therefore,

$$LG(s) \approx G_I G_\phi G_A G_D H_D(s) = \frac{sG_T}{s^2 - \omega_0^2}, \quad (4.72)$$

where  $G_T = G_I G_\phi G_A G_D G_f$ . The closed-loop transfer function from  $V_{in}(t)$  to  $x_o[n]$  is

$$H_{CL}(s) = \frac{G_I G_\phi G_A}{1 + LG(s)} \quad (4.73)$$

To simplify the analysis, the IAMP and anti-aliasing filter are modeled as gain stages with a constant frequency response. Obviously these blocks are frequency-dependent, but our goal here is to analyze the effect of the feedback path on the closed loop response of the system, and, as will be shown, for frequencies distant from  $\omega_0$ , the feedback path has a negligible effect on the closed-loop transfer function.

The 3 dB bandwidth of the notch can be solved for by finding the value of  $\omega$  for which the loop gain amplitude is unity. There are four frequencies at which this happens, as seen in Fig. 4-19, and some simple algebra shows that they are approximately located at

$$\omega_{3\text{dB}} \approx \pm (\omega_0 \pm G_T/2). \quad (4.74)$$

This approximation simply requires that  $\omega_0 \gg G_T$ , or equivalently, that the notch width be narrow compared with its center frequency. To make the notch narrow,  $G_f$  must be very small so that  $G_T$  is small.

Figs. 4-19 and 4-20 respectively show the magnitude and phase responses of the

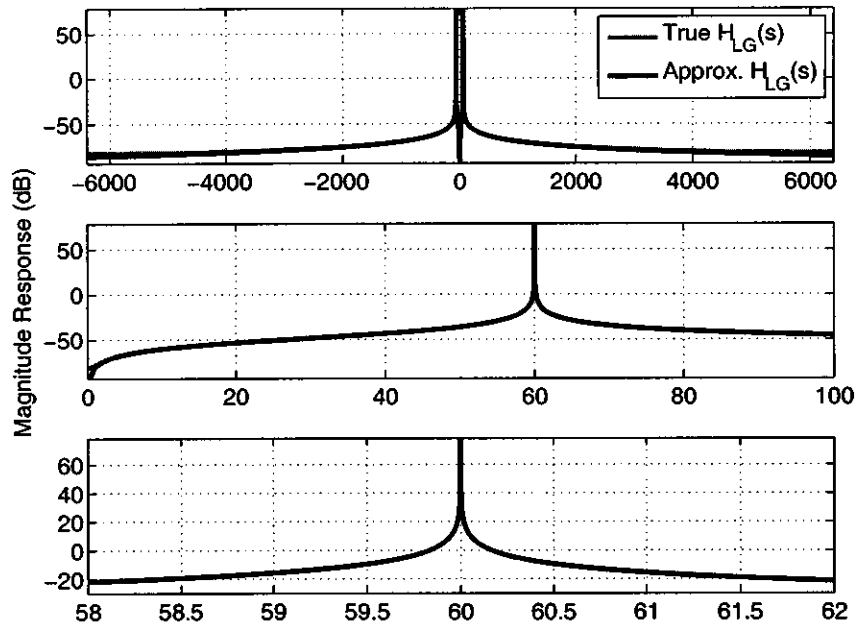


Figure 4-19: Amplitude of the loop gain's frequency response for  $G_f = 2^{-22}$ . As  $f$  approaches  $f_0$ , the amplitude approaches infinity. For large values of  $|f| - |f_0|$ , however, the loop gain becomes very small.

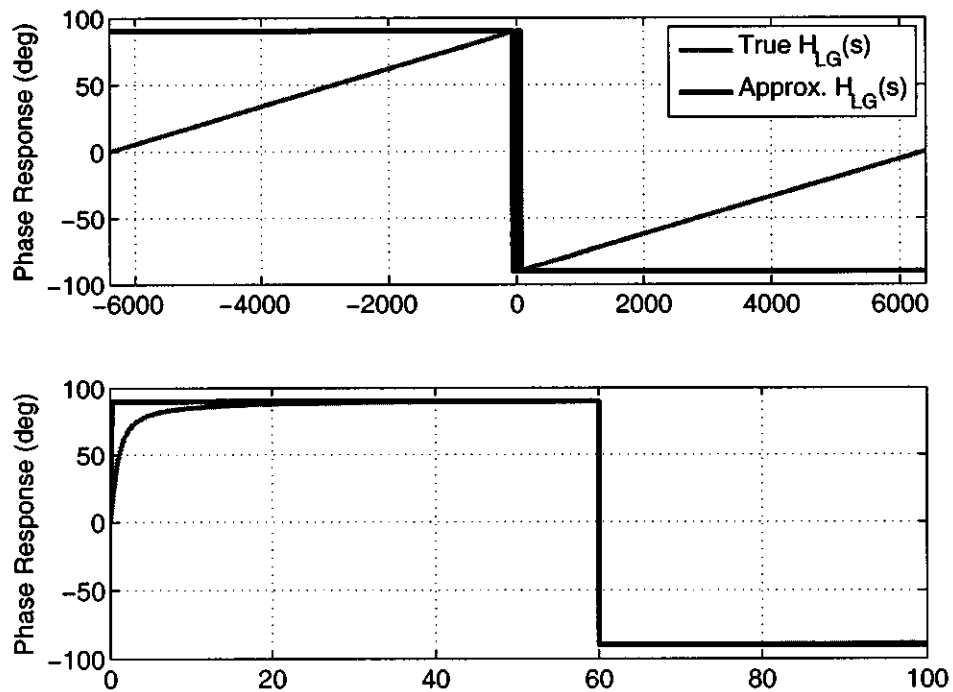


Figure 4-20: Phase of the loop gain's frequency response for  $G_f = 2^{-22}$ .

Table 4.3: Example Notch Filter Parameters

$G_I$	100
$G_\phi$	1
$G_f$	$2^{-22}$
$V_{refA}$	1.5 V
$V_{refD}$	40 mV
$\omega_0$	$2\pi \times 60$ Hz

loop gain using the parameters in Table 4.3 with different zoom levels. The top plot in Fig. 4-19 is zoomed out from  $-f_s/2$  to  $f_s/2$ . The gray line is the actual transfer function given by (4.66) and the black line is the approximation made in (4.68). Notice that the two plots are practically identical except for very high frequencies where the actual loop gain is slightly larger than the approximation. The plot shows that the loop gain is much smaller than 0 dB for most frequencies with the exception of two sharp spikes at  $\pm f_0 = \pm 60$  Hz. The middle and bottom plots in Fig. 4-19 show the amplitude response in the respective frequency ranges of 0–100 Hz and 58–62 Hz. Notice that the amplitude response crosses the 0 dB line at frequencies  $f = 60$  Hz  $\pm 0.2$  Hz. Inspection of (4.73) reveals that at  $f = f_0$ , where the amplitude of the loop gain is infinite, the closed loop response must be zero. In contrast, at frequency offsets beyond 1 Hz from  $f_0$ , the loop gain is very small and the closed-loop response is simply  $G_I G_\phi G_A$ . In other words, the feedback should notch out spectral components around 60 Hz without affecting the rest of the spectrum significantly.

Fig. 4-20 shows that the phase changes abruptly from  $+90^\circ$  to  $-90^\circ$  at  $f = f_0$ , as predicted by (4.68). From the top plot, it is clear that the approximation is less accurate for the phase at high frequencies, but the bottom plot shows that it is very good for  $f \approx f_0$ , which is the important frequency range where the amplitude is significant.

Figs. 4-21 and 4-22 respectively show the magnitude and phase responses of the closed-loop transfer function, except the output is taken at  $v_a(t)$  (before digitization). Clearly, the notch is very narrow, as expected, and has a 3 dB corner frequency at the same frequency that the loop gain crosses the 0 dB line. The transfer function adds about  $10^\circ$  of phase at 59 Hz and drops sharply at farther offsets ( $5^\circ$  at 58 Hz,

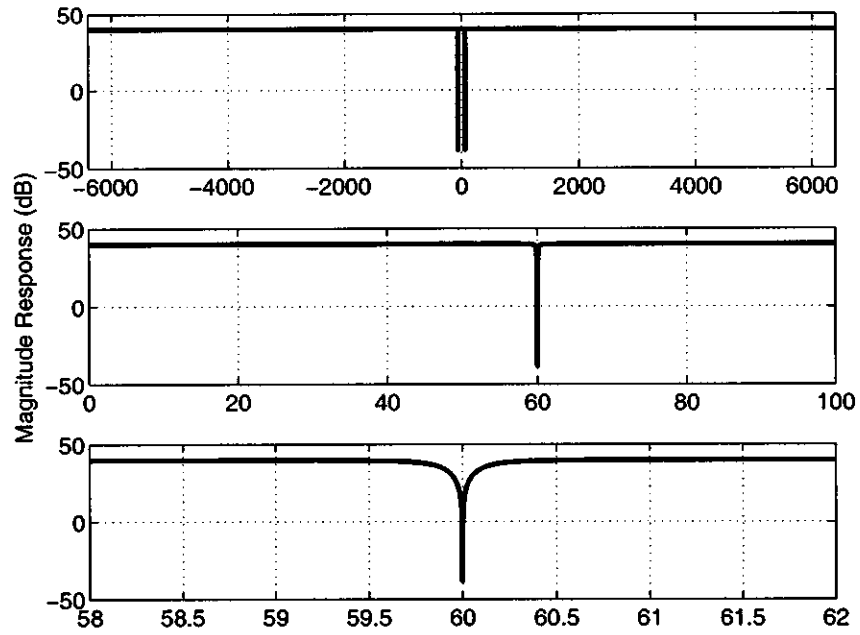


Figure 4-21: Amplitude of the closed-loop frequency response for  $G_f = 2^{-22}$ . Clearly there is a very narrow notch at  $f_0$ . The notch width can be easily programmed by changing  $G_f$ .

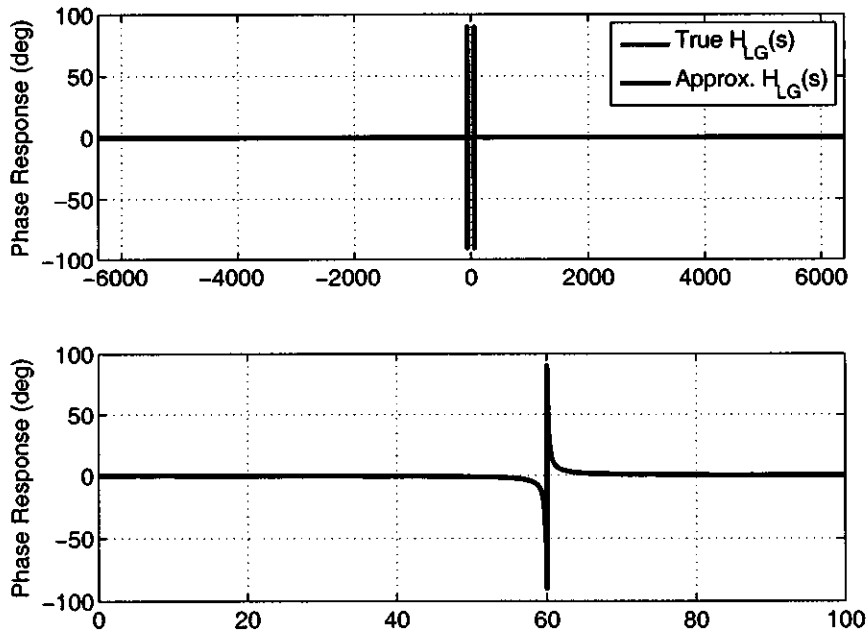


Figure 4-22: Phase of the closed-loop frequency response for  $G_f = 2^{-22}$ .

$2^\circ$  at 55 Hz, etc.).

There are four main benefits to this technique compared to the purely analog technique described in [27]. First of all, no large analog components are necessary, therefore reducing size. In the analog implementation, large resistors and capacitors are needed to create the large time constants necessary to achieve a narrow notch. In the mixed signal implementation, the width of the notch can be set by digitally by adjusting  $G_f$ , which can be done with a simple register. The hardware cost of making the notch arbitrarily narrow is practically insignificant as will be shown in the implementation section. This affects size, but it also highlights the second benefit, which is programmability. With the mixed-signal design, the notch's center frequency and width can be easily and arbitrarily reprogrammed. Although we have focused on filtering PLI, the main source of interference in some applications is set by some other signal. For example, in closed-loop, deep brain stimulators, the stimulation signal can interference with the sensor signal. Since stimulation occurs at a known rate, this technique could be used to notch out the interference it causes. Having the ability to easily reprogram the notch, therefore, can be a major benefit.

The third benefit of this technique is that it is significantly less power intensive than an analog approach. This is because the main blocks used to implement it are necessary anyway. The ADC is present in practically all systems, and its specifications are actually relaxed by this technique. Digital signal processing is practically always done, and the implementation of the digital block uses few resources clocked at a low rate. The only block that has to be added is the DAC. As will be shown, however, it can be made small, and its power consumption is negligible.

The fourth benefit of this technique is that additional digital blocks can be used in parallel to notch out multiple blockers. It is likely that more elaborate filtering is possible, although there are limitations due to stability requirements.

## Digital Block Implementation

As shown in Fig. 4-18, the digital block comprises four digital multipliers, two accumulators, a gain block ( $G_f$ ), a *direct-digital synthesizer* (DDS), and an adder. The



prototype uses 16-bit multipliers to minimize quantization noise and 32-bit accumulators to avoid overflow. Each accumulator is simply a 32-bit adder with the 16 LSBs of its output connected to one of its inputs.  $G_f$  is always much smaller than unity (e.g.  $2^{-22}$ ) and is implemented using two stages of binary division that are implemented by truncation or bit shifting. The first division is carried out by selecting which 16 of the 32 bits of the accumulator are used. For example, if  $z_I\{15 : 0\}[n] = y_I\{31 : 16\}[n]$ , a gain of  $2^{-16}$  is achieved, whereas if  $z_I\{15 : 0\}[n] = y_I\{27 : 12\}[n]$ , the gain is  $2^{-12}$ . This allows simple programmability of the gain by factors of  $2\times$  using multiplexors, and results in minimal hardware overhead. The second gain stage is implemented by truncating the 8 least significant bits (LSB) of  $d[n]$ . Since the DAC only uses 8-bits, this is necessary anyway and it effects a gain of  $2^{-8}$  with practically no hardware overhead. The truncation is done before the final addition so that an 8-bit adder can be used, as opposed to a 16-bit adder.

Direct-digital synthesis is a simple way of implementing a digital I/Q oscillator. The DDS simply comprises an accumulator and a look-up table (LUT) [61]. It exploits the fact that phase is the time integral of frequency and uses the accumulator to implement this integration. The input to the accumulator sets the frequency, and the output is equivalent to the oscillator's phase. Since the phase of a sinusoid maps directly to its amplitude, the mapping function can be stored in the LUT and the MSBs of the accumulator can be used as the LUT's address. Since the phase difference between sine and cosine is a constant, a simple adder can be used to get the LUT's address for the appropriate value of sine and cosine.

The frequency resolution of the DDS is set by the width of the accumulator and can be made very small. Specifically, the oscillation frequency is

$$f_0 = \frac{N_F f_s}{2^B}, \quad (4.75)$$

where  $B$  is the width of the accumulator and  $N_F \in [0, 2^B - 1]$  is the frequency tuning

word. The frequency resolution, therefore, is

$$\Delta f_0 = \frac{f_s}{2^B}. \quad (4.76)$$

In our implementation,  $B = 32$ . As an example, for a clock frequency of  $f_s = 12.8\text{ks/s}$ , the resolution is  $\Delta f_0 = 3\ \mu\text{Hz}$ . The look-up table was implemented using a 16-bit wide SRAM with 1024 addresses. The 10 MSBs of the accumulator are used as the LUT's address for the sine output. The sum of the accumulator's output and a constant ( $2^B/4$ ) is used to choose the LUT's cosine output.

## DAC Implementation

The DAC in Fig. 4-18 uses binary-weighted charge-redistribution and is implemented with switched capacitors. Fig. 4-23 (A) shows a simplified schematic of the DAC and how it connects to the IAMP, (B) shows the implementation of the switches, and (C) shows an equivalent single-ended model of the DAC and IAMP. Charge redistribution DACs usually require an opamp to convert the charge to a voltage [29]. Since the IAMP already uses an opamp, however, the DAC simply piggybacks on it. As a result, only the capacitor array and switches add to the total area, and only  $f_s CV^2$  power is added, which is very small because the reference voltage, sampling frequency, and capacitors are all small. As discussed in Sect. 4.3, shunt capacitance added at the input of the opamp attenuates the input signal resulting in degraded noise performance for the system. For this reason, it is important to minimize  $C_0$  such that total DAC capacitance  $C_{DAC} = 2^8 C_0$  is significantly smaller than  $C_i$ . In our implementation,  $C_0 = 10\ \text{fF}$ , such that  $C_{DAC} = 2.56\ \text{pF}$ , which is approximately  $C_i/4$ . The effective DAC voltage,  $v_{dac}(t)$  in Fig. 4-23 (C), is equal to

$$v_{dac}(t) = \frac{2V_{refD}(d(nT_s) - 128)}{255}, \quad (4.77)$$

where  $d(nT_s)$  is the value of the digital word  $d[n] \in [0, 255]$  at time  $t = nT_s$ . The equivalent input-referred DAC voltage is

$$v_d(t) = \frac{C_{DAC}}{C_i} v_{dac}(t) \approx \frac{v_{dac}(t)}{4} = \frac{V_{refD}}{512} (d(nT_s) - 128). \quad (4.78)$$

This explains why the gain of the DAC,  $G_D$  in (4.70) has a  $1/4$  term.

Since the DAC noise is directly added to the input before any amplification, it must be made significantly lower than the input-referred noise of the IAMP. There are two mechanisms through which the DAC can add noise: noise from the reference voltage  $V_{refD}$  and quantization noise (the switch resistances add negligible noise). The worst-case noise coupling from  $V_{refD}$  occurs when  $d[n] = 0$  or  $d[n] = 255$ , such that its noise is completely differential. In either case, the noise power is attenuated by a factor of  $16\times$ , which relaxes the noise specifications of  $V_{refD}$  significantly. In this prototype,  $V_{refD}$  is supplied with an external low-noise source, although in a final product, care should be taken to supply a clean supply without adding excessive power consumption to the system. Since noise from this source is attenuated by a factor of four when referred to the input, its noise density can be higher than the input-referred noise of the IAMP. As a result, its power consumption should be smaller than that of the IAMP. Further, for systems using multiple sensor-interface circuits, a single reference voltage can be used such that its power consumption becomes negligible compared to the power consumed by all of the IAMPs combined.

The more significant source of noise is the quantization noise introduced by the DAC. Assuming the DAC's ENOB is 8-bits, its input-referred quantization noise density is

$$S_q(f) = \left( \frac{V_{refD}}{512} \right)^2 \times \frac{1}{12f_s}. \quad (4.79)$$

The largest interference that can be canceled is one with a peak-to-peak value of  $V_{refD}/2$ , therefore (4.79) shows there is a direct tradeoff between the largest interferer that can be canceled and the amount of quantization noise that can be tolerated. As an example, if  $V_{refD} = 10$  mV and  $f_s = 12.8$  kS/s, the input referred quantization noise is  $v_{qni} = 50$  nV/ $\sqrt{\text{Hz}}$  which would only increase the total input referred noise by

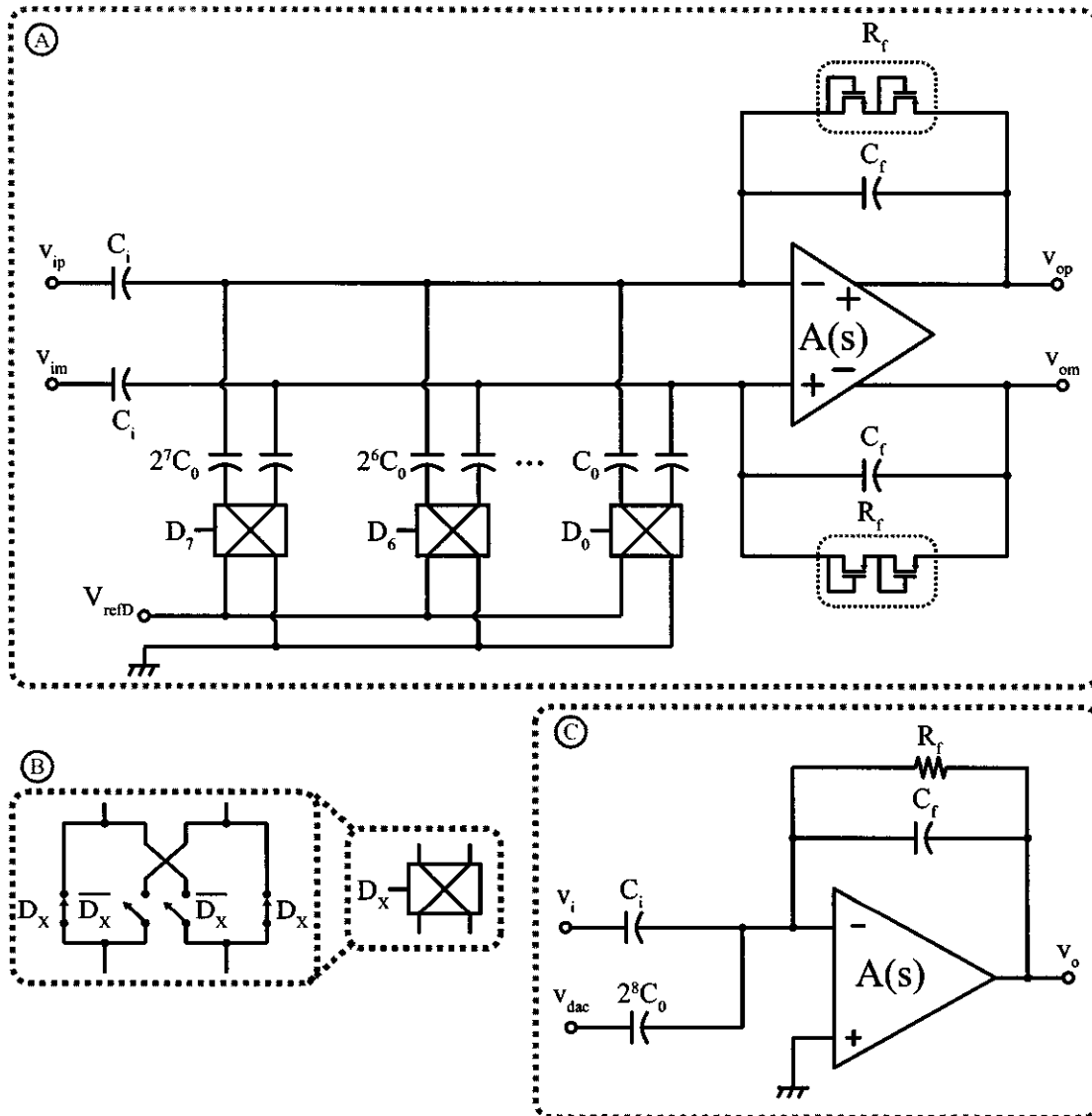


Figure 4-23: (A) Schematic of the combined instrumentation amplifier and charge redistribution DAC. The DAC is implemented with binary weighted capacitors and switches. A simplified diagram of the switches' implementation is shown in (B). (C) shows a simplified, single-ended model of the IAMP with two inputs: one for the signal  $v_i$  and another for an equivalent voltage DAC signal  $v_{dac}$ .

about 10% (assuming the input-referred noise of the IAMP is  $v_{na} = 100 \text{ nV}/\sqrt{\text{Hz}}$ ). The largest interferer that could be canceled would be  $5 \text{ mV}_{\text{p-p}} = 1.8 \text{ mV}_{\text{rms}}$  which should be more than sufficient for PLI.

For a signal bandwidth of 100 Hz, the integrated noise due to the DAC would be 500 nV, whereas the largest signal it could cancel is 1.8 mV. This is equivalent to a dynamic range of 71 dB. This relatively high dynamic range despite 8 bits of ENOB results from the oversampling nature of the DAC. The system-level benefit is that the dynamic range requirements of all of the blocks in the forward path (IAMP, AAF, and ADC) are relaxed, allowing lower-power operation.

## 4.6 Measurement Results

The IC including the IAMP, AAF, and DAC shown in Fig. 4-18 was implemented in a 0.18  $\mu\text{m}$  CMOS process from TSMC with no special layers. Fig. 4-24 shows a photograph of the full chip and Fig. 4-25 shows a zoomed in version of the circuits with annotations. Excluding the SPI and ADC driver, which are only used for testing purposes, the total area of the pertinent blocks is less than  $0.25 \text{ mm}^2$ .

The rest of the system was implemented on a PCB with off-the-shelf components including two 16-bit ADCs from Analog Devices (AD7684), a 12-bit DAC used as a signal source for testing (DAC7811 from Texas Instruments), high input-impedance opamps used for buffering (AD8603 from Analog Devices), and various voltage regulators (see Fig. 4-26). Fig. 4-27 shows the FPGA board manufactured by Opal Kelly (XEM3010) which includes a Xilinx Spartan-3 FPGA, a USB interface, and a series of supporting hardware components that facilitate the interface between the FPGA and a personal computer. A MATLAB graphical user interface and various scripts were used to perform analysis on the output data of the ADCs.

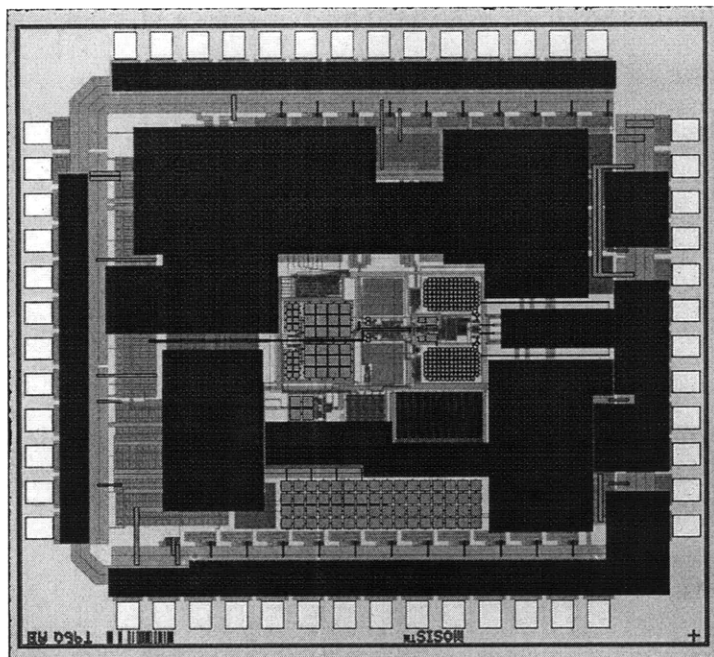


Figure 4-24: Die photograph of IC with IAMP, filter, and feedback DAC.

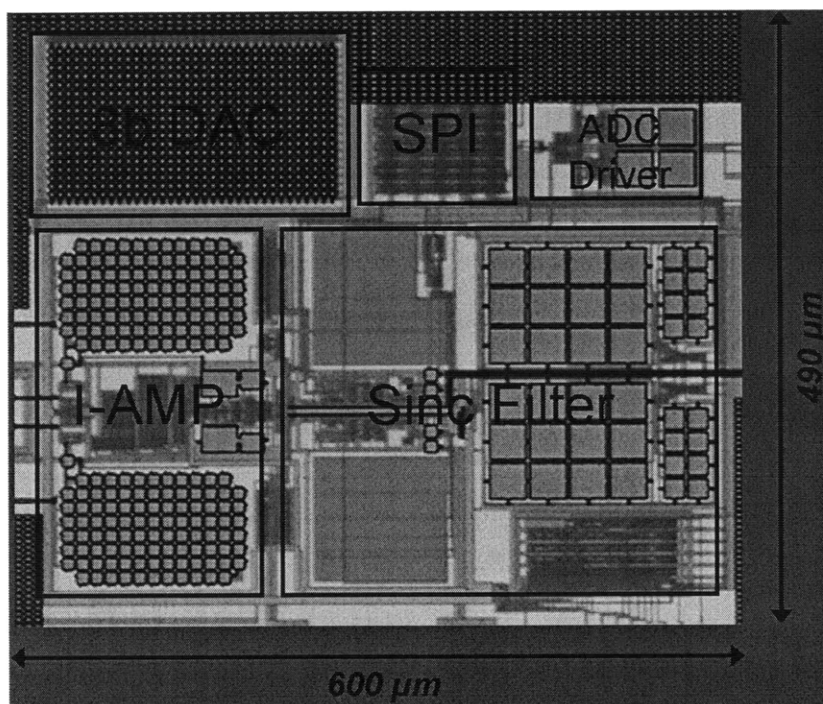


Figure 4-25: Zoomed in version of die photograph.

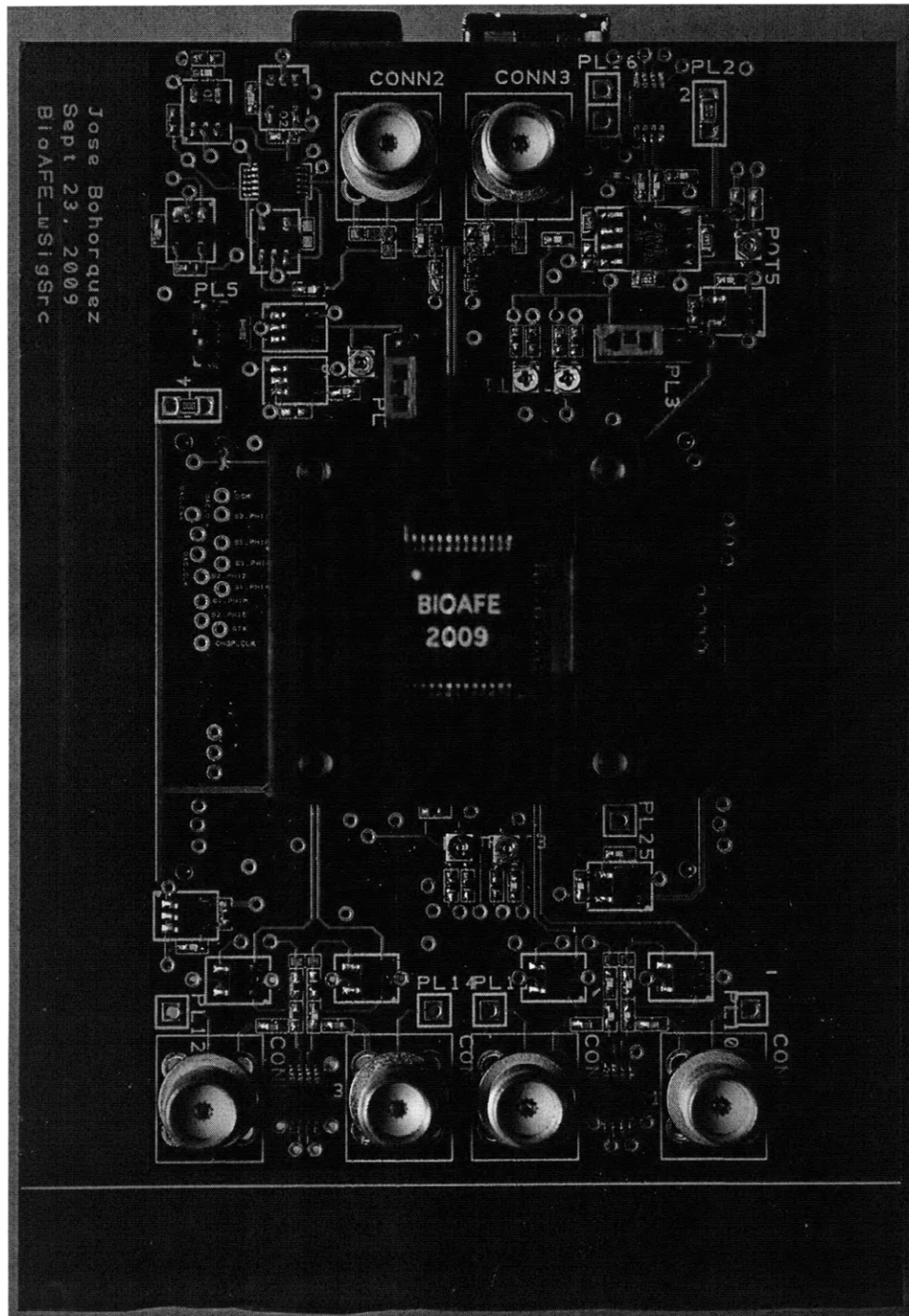


Figure 4-26: Test board.

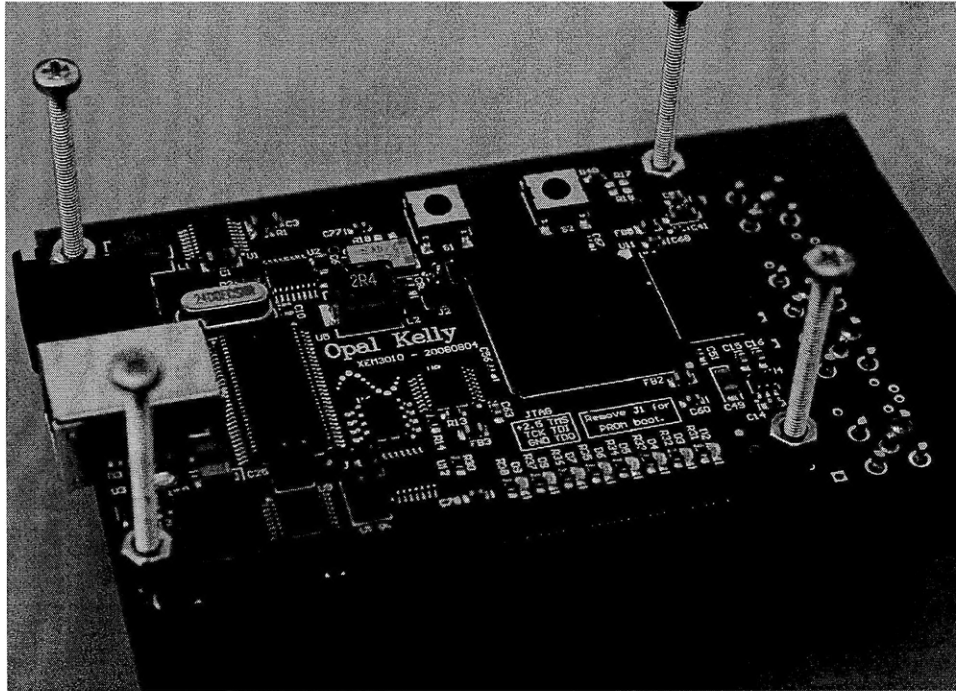


Figure 4-27: FPGA board connected to test board.

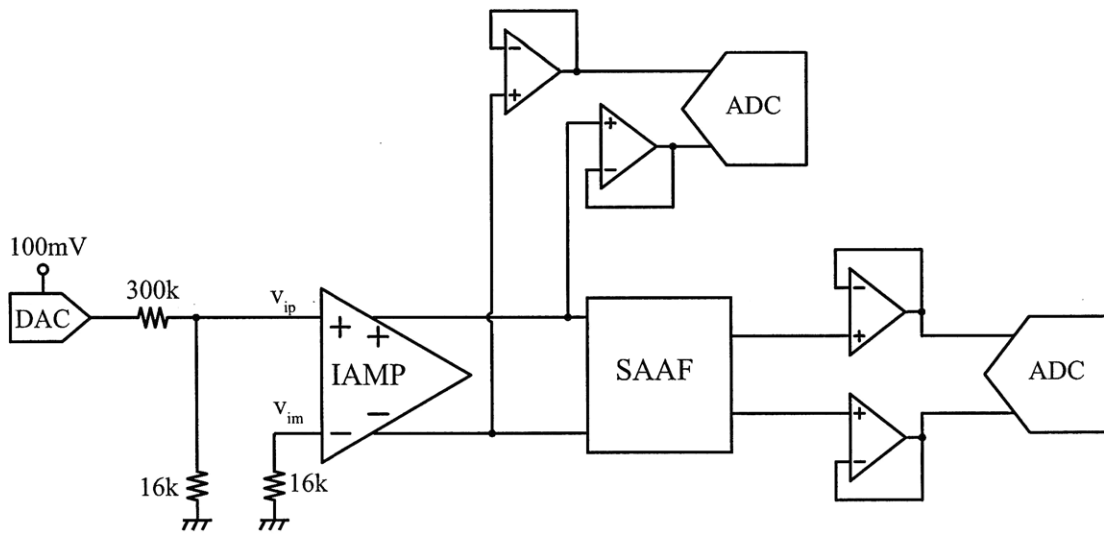


Figure 4-28: Simplified schematic of test board.



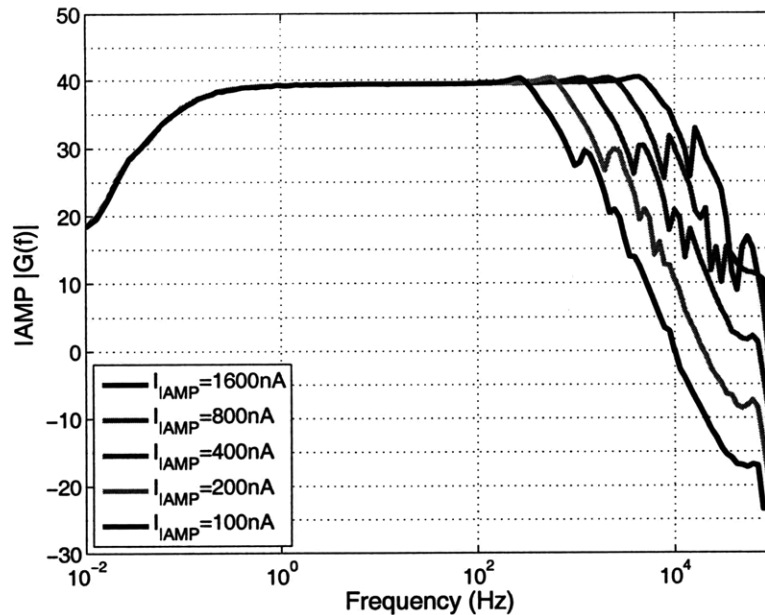


Figure 4-29: IAMP frequency response for various bias settings.

### 4.6.1 Frequency Response

#### IAMP

Fig. 4-28 shows the basic test setup used for most of the measurements performed. To calculate the IAMP's transfer function, the negative input  $v_{im}$  was connected to ground through a 16 k $\Omega$  resistor, and a 12-bit, single-ended DAC (DAC7811) was used as a signal source connected to  $v_{ip}$  through a voltage divider. The DAC's reference voltage was set to a nominal value of 100 mV with a voltage regulator. A sinusoid of maximum amplitude was measured to have a peak-to-peak amplitude of 96.8 mV. The voltage divider attenuates the signal by 19.75 $\times$ , such that the peak-to-peak signal at the IAMP's input is 4.90 mV. The IAMP's output was buffered using high input-impedance opamps and digitized using a 16-bit ADC. With a peak-to-peak input signal of 4.90 mV, the output of the IAMP was measured as 458 mV, so the gain is 93.5 $\times$ , or 39.4 dB, which is slightly lower than the designed value of 100 $\times$  and probably due to parasitic capacitance in the feedback path.

Fig. 4-29 shows the amplitude frequency response of the IAMP for five supply

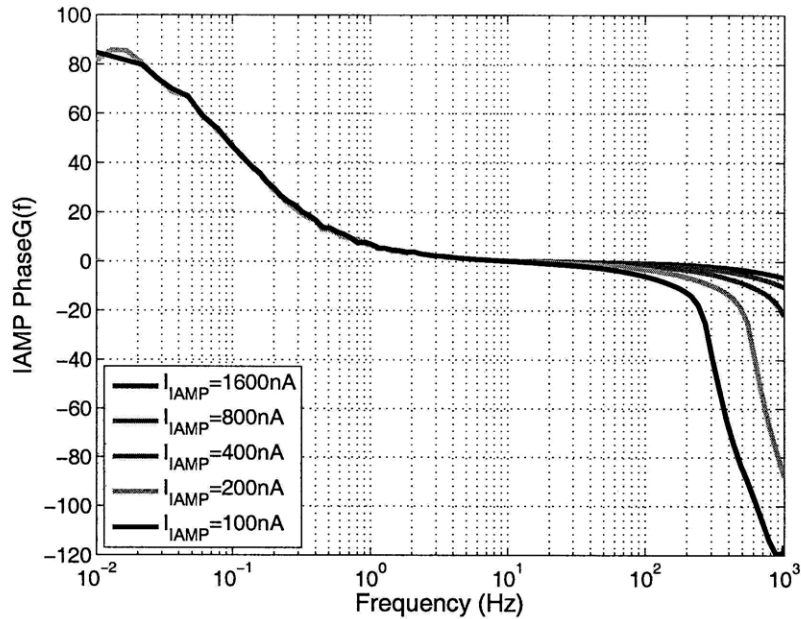


Figure 4-30: IAMP phase response for various bias settings.

current settings ( $I_{IAMP}$  = total current including biasing blocks). The power supply was set to 1.5 V and the filter bias current was set to 150 nA.<sup>2</sup> The bias current clearly does not affect the frequency response at low frequencies where the opamp's gain is very high. The high-pass corner frequency set by the feedback network is  $f_{HP} = 120$  mHz. The low-pass corner frequency is set by the opamp's poles and increases linearly with bias current ( $f_{LP} = 450$  Hz for  $I_{IAMP} = 100$  nA and  $f_{LP} = 7.6$  kHz for  $I_{IAMP} = 1600$  nA).

Fig. 4-30 shows the phase response of the IAMP. As with amplitude, the phase response is practically identical at low frequencies and nearly zero in the pass-band. The phase lead from the high-pass filter adds about  $7^\circ$  of phase at 1 Hz and the phase lag is  $7^\circ$  at 100 Hz for  $I_{IAMP} = 100$  nA and less than  $5^\circ$  at higher currents.

## Filter

Fig. 4-31 shows the frequency response of the filter with the following settings:  $i_{SAAF} = 150$  nA,  $b_C=31$ ,  $b_R=1$ ,  $i_{IAMP} = 1.6 \mu\text{A}$ ,  $G_\phi(0) = 1.4$  dB. The IAMP current

<sup>2</sup>The sinc anti-aliasing filter is referred to as the AAF or SINC in some of the figures.

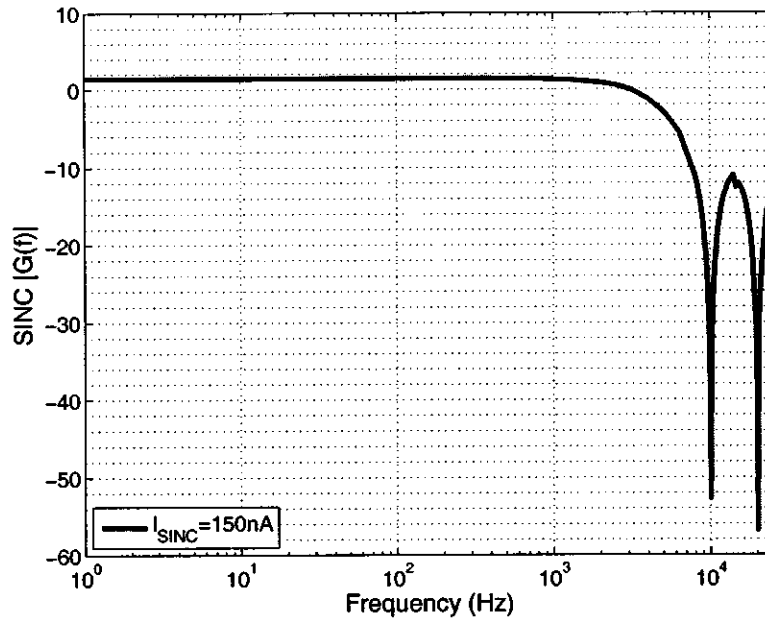


Figure 4-31: Filter frequency response:  $i_{SAAF} = 150 \text{ nA}$ ,  $b_C=31$ ,  $b_R=1$ ,  $i_{IAMP} = 1.6 \text{ }\mu\text{A}$ ,  $G_\phi(0) = 1.4 \text{ dB}$ .

was set relatively high to extend its bandwidth. Clearly, the filter amplitude response is extremely flat up to 1 kHz with less than 1 dB of attenuation, and has sharp notches at multiples of  $f_s = 10 \text{ kHz}$ . Fig. 4-32 shows the magnitude of the transfer function versus linear frequency to highlight the depth and width of the notches. Notice that while the 3 dB frequency is at 4.0 kHz, more than 50 dB of attenuation is achieved at 10 kHz. Fig. 4-33 zooms in more and shows that more than 17 dB of attenuation is achieved in the  $f_s \pm 1 \text{ kHz}$  band, and more than 35 dB of attenuation is achieved in the  $f_s \pm 100 \text{ Hz}$  band. Note that the notch's trough actually falls at a slightly higher frequency than  $f_s$ . This is because the short amount of time necessary to discharge the integrating capacitors leads to a time window slightly smaller than  $T_s$ .

Figs. 4-34 and 4-35 show the frequency response of the cascade between the IAMP and the filter ( $I_{IAMP} = 400 \text{ nA}$ ). Notice that their combined filtering leads to more than 40 dB of attenuation in the  $f_s \pm 1 \text{ kHz}$  band, and more than 50 dB of attenuation in the  $f_s \pm 100 \text{ Hz}$  band.

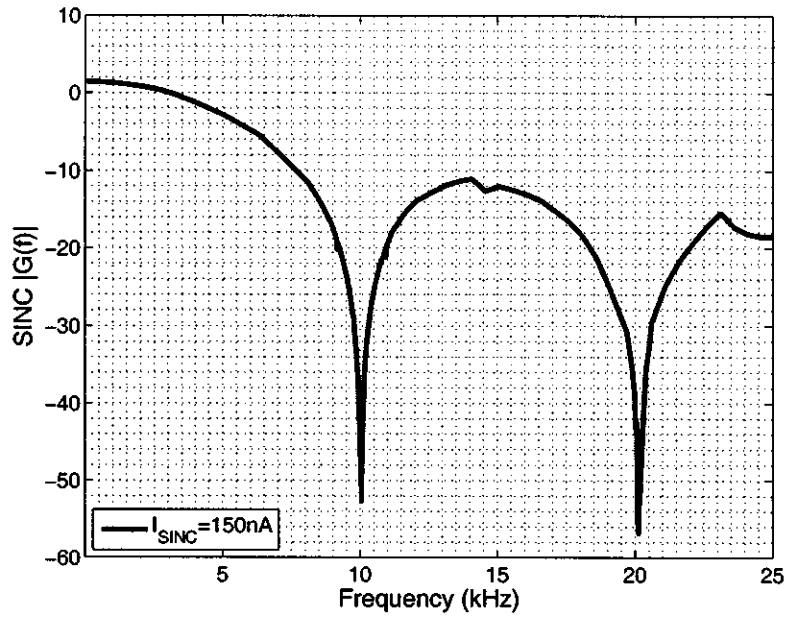


Figure 4-32: Filter frequency response plotted with a linear frequency axis:  $i_{SAAF} = 150 \text{ nA}$ ,  $b_C = 31$ ,  $b_R = 1$ ,  $i_{IAMP} = 1.6 \mu\text{A}$ ,  $G_\phi(0) = 1.4 \text{ dB}$ .

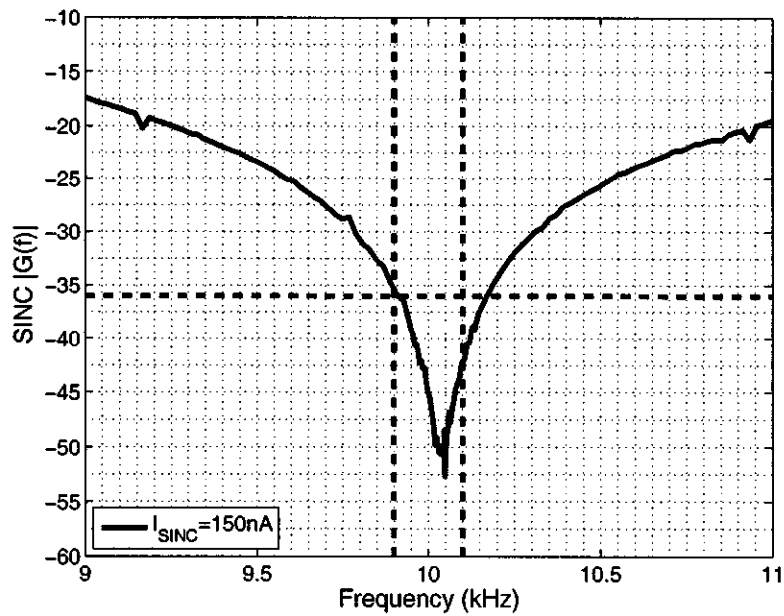


Figure 4-33: Filter frequency response zoomed in near the notch frequency:  $i_{SAAF} = 150 \text{ nA}$ ,  $b_C = 31$ ,  $b_R = 1$ ,  $i_{IAMP} = 1.6 \mu\text{A}$ ,  $G_\phi(0) = 1.4 \text{ dB}$ .

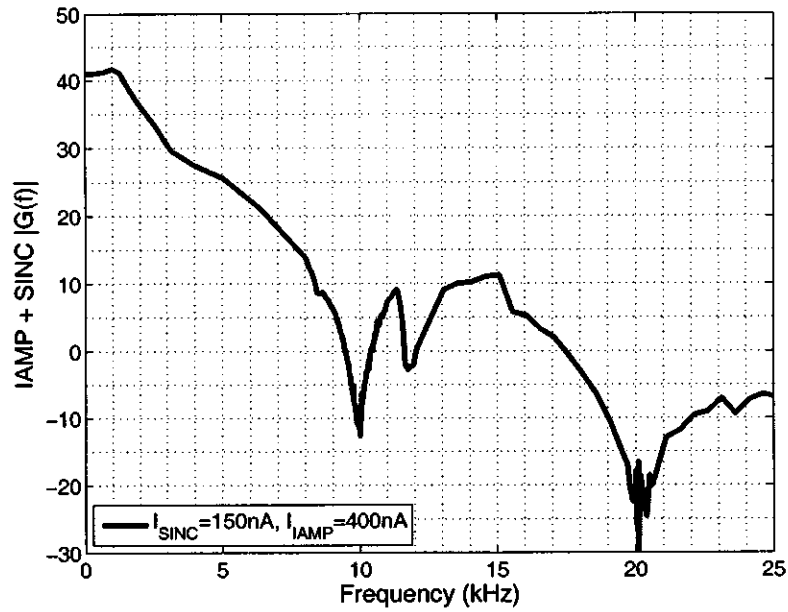


Figure 4-34: System frequency response (IAMP + filter) plotted with a linear frequency axis:  $i_{SAAF} = 150 \text{ nA}$ ,  $b_C=31$ ,  $b_R=1$ ,  $i_{IAMP} = 400 \text{ nA}$ ,  $G_\phi(0) = 1.4 \text{ dB}$ .

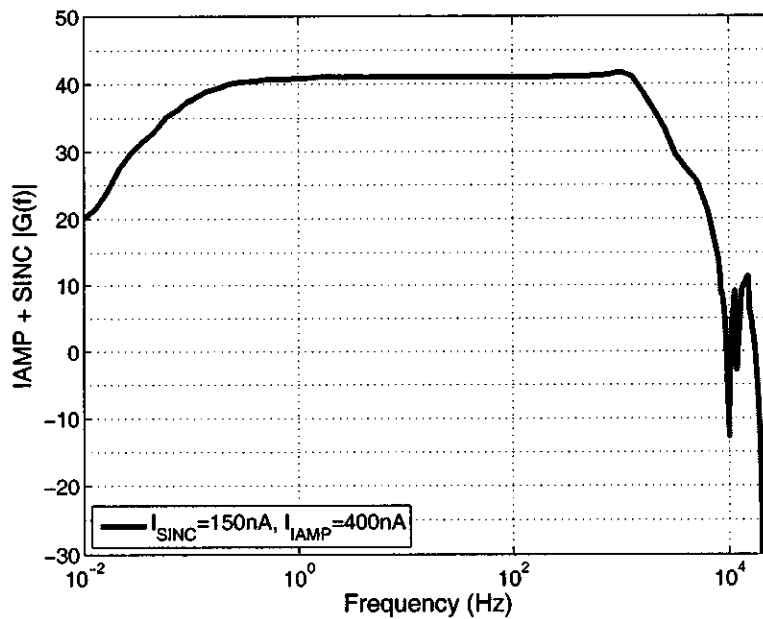


Figure 4-35: System frequency response (IAMP + filter) plotted with a logarithm frequency axis:  $i_{SAAF} = 150 \text{ nA}$ ,  $b_C=31$ ,  $b_R=1$ ,  $i_{IAMP} = 400 \text{ nA}$ ,  $G_\phi(0) = 1.4 \text{ dB}$ .

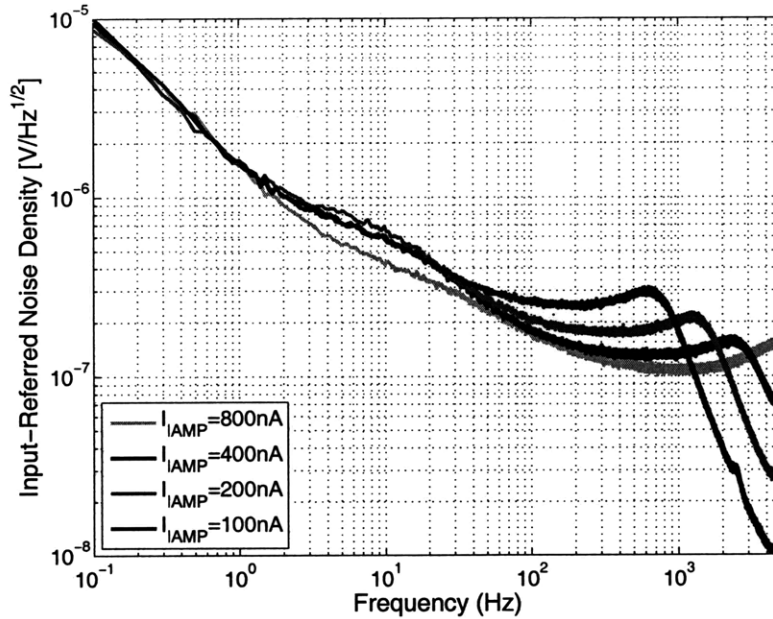


Figure 4-36: IAMP input-referred RMS noise voltage density.

## 4.6.2 Input-Referred Noise Measurement

### IAMP

Fig. 4-36 shows the input-referred noise of the IAMP when measured at the output of the top ADC in Fig. 4-28. The noise measurements were made by digitizing the output of the IAMP when the input was set to zero. Ten seconds worth of the ADC's output signal were stored ( $f_s = 10$  kS/s) and an FFT was performed on the signal. No anti-aliasing filter was used and the noise was referred to the input by dividing the voltage spectral densities by the mid-band gain of the IAMP.

Notice that the input-referred noise at low frequencies has a  $1/f^2$  shape as predicted by the noise analysis in Section 4.3 and is independent of bias setting.<sup>3</sup> This noise is due to the feedback resistor  $R_f$ , and theoretically should be about  $5.3 \mu\text{V}$  at 0.1 Hz. The measured value is about  $9.0 \mu\text{V}$  which is a bit higher. Beyond  $\approx 1$  Hz, the noise drops with a slope of  $1/f$  as also expected. This part of the spectrum is dominated by flicker noise in the IAMP and is largely independent of bias current,

<sup>3</sup>Note that the  $1/f^2$  term relates to power, whereas the noise figures plot RMS noise voltage density so the slope is  $1/f$  on the plot.

although it seems to be somewhat lower for  $I_{IAMP} = 800$  nA.

At higher frequencies, the noise density flattens and is reduced for higher bias settings as predicted in Section 4.3. This part of the spectrum is dominated by the IAMP's thermal noise. For  $I_{IAMP} = 800$  nA, the noise seems to be higher than expected at higher frequencies. As shown in the next subsection, this is caused by aliasing and the expected results occur when the signal is passed through the filter before digitization.

### Filter

Fig. 4-37 shows the input-referred noise of the system for four IAMP bias settings when the output is taken from the filter. The filter bias current was set to  $I_{SAAF} = 150$  nA for all four measurements. The results look nearly identical to those in Fig. 4-36, except the noise is lower at higher frequencies for  $I_{IAMP}=800$  nA. As expected, the thermal noise drops by 3 dB each time the bias current is doubled. For a bias setting of  $I_{IAMP} = 800$ nA, the input-referred noise density is  $\approx 90$  nV/ $\sqrt{\text{Hz}}$ , which is slightly better than the requirements. The power consumption for the IAMP and filter under these settings would be  $P_{IAMP} = 1.2$   $\mu\text{W}$  and  $P_{SAAF} = 225$  nW, for a total of  $P_{TOT} = 1.425$   $\mu\text{W}$ .

Fig. 4-38 compares the input-referred noise of the system when measured at the outputs of the IAMP and filter. For  $I_{IAMP} = 800$  nA, the noise is nearly identical at low frequencies, but for higher frequencies the noise is higher when measured at the output of the IAMP. The extra noise is caused by aliasing. Since, the IAMP's bandwidth is wider the Nyquist frequency, there is no anti-aliasing and noise beyond  $f_s/2$  folds back to the ADC's base band. As shown in the Fig. 4-32, the bulk of the filter's filtering occurs near multiples of  $f_s = 10$  kHz. However, even between  $5$  kHz  $\leq f \leq 9$  kHz, at least 4–20 dB of attenuation is achieved, which is enough to lower the aliased noise significantly in the base band between 1 kHz–5 kHz.

At the lower IAMP bias setting of  $I_{IAMP} = 100$  nA, the IAMP bandwidth is significantly lower and effects some anti-aliasing. For this reason, the noise is practically identical whether measured at the IAMP or filter output. For either bias setting, the

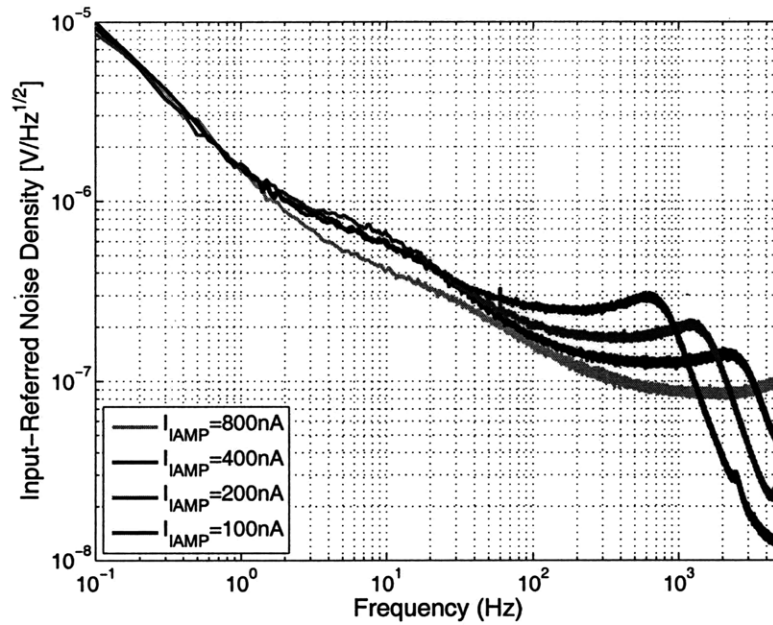


Figure 4-37: Input-referred noise of the system (IAMP + filter) as measured from the output of the filter for four different IAMP bias settings and a constant filter bias current of  $I_{SAAF} = 150$  nA.

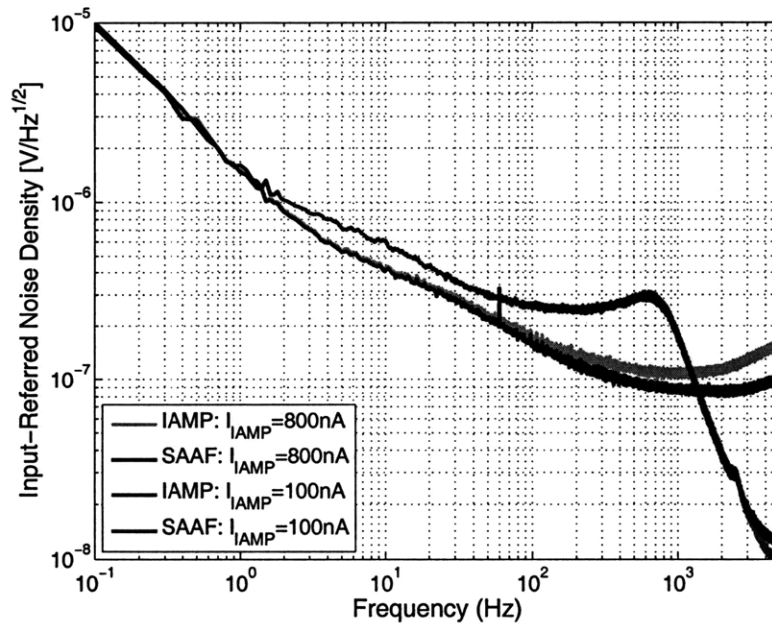


Figure 4-38: Input-referred noise of IAMP alone, and IAMP plus filter for two IAMP bias settings for a constant filter bias setting of  $I_{SAAF} = 150$  nA.



noise is nearly identical at low frequencies where aliasing is negligible. This confirms that the IAMP's high gain results in the filter's noise having a negligible effect on the overall noise performance of the system.

The significant amount of noise at lower frequencies confirms that chopping is practically indispensable for small, low-power, CMOS-based IAMPs. The techniques in [74] and [17] have some drawbacks, but result in significantly lower noise at low frequencies. Using similar techniques, or novel architectures that incorporate chopping, in conjunction with the filter architecture presented here and the notch filter, would result in a very attractive sensor interface system.

### 4.6.3 Total Harmonic Distortion

#### IAMP

For the THD measurements, a larger input signal was necessary, so the 300 k $\Omega$  resistor in Fig. 4-28 was changed to 62 k $\Omega$  such that the maximum signal at the IAMP's input could be 20.5 mV<sub>p-p</sub>. Fig. 4-39 shows the output spectrum of the IAMP with a 50 Hz, 20.5 mV<sub>p-p</sub> input signal. The IAMP was biased with  $I_{BIAS} = 400$  nA for this measurement. The THD was less than 0.02% at 50 Hz and less than 0.1% for all frequencies between 1 Hz and 1 kHz. This is significantly better than typical requirements of THD $\leq$ 1% for input signals smaller than 5 mV [17] (for a system gain of 40 dB).

#### Filter

Fig. 4-40 shows the output spectrum of the filter when its THD is about 1%. The input signal for this plot was a 10.8 mV<sub>p-p</sub> sinusoid, which means the input to the filter was 1.0 V<sub>p-p</sub>, since the IAMP's gain is 93.5. The bias current was set to  $i_{SAAF} = 150$  nA, and  $b_R$  was set to its lowest setting of 0 which corresponds to a nominal value of  $R_s = 20$  M $\Omega$ .

Compression in the filter can occur from current limitations or voltage limitations. Regardless of how much current is used, if the product of the input signal's peak-to-

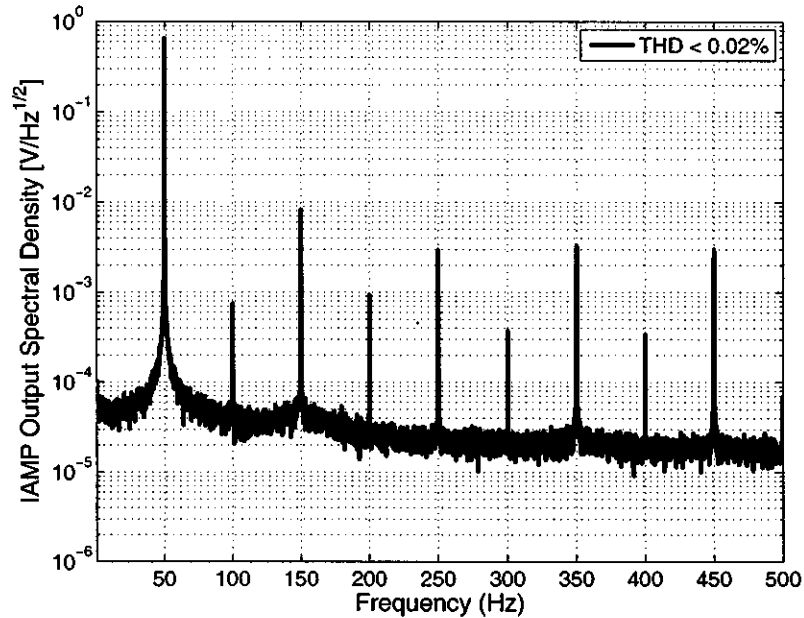


Figure 4-39: IAMP’s output spectrum for an input signal of 20 mV<sub>p-p</sub>. The total harmonic distortion was 0.02% at 50 Hz and below 0.1% for frequencies between 1 Hz and 1 kHz.

peak value and the filter’s gain is larger than about 1.75 V, voltage compression begins to occur. Of course, the ADC’s input range is most likely significantly smaller than this, so it is generally not a concern. This means compression will more likely occur due to bias current limitations. As discussed in Section 4.4, the filter’s transconductor current begins to clip if  $v_{in,pk}/R_s > i_b$ . We assumed earlier that  $i_b \approx I_{SAAF}/4$ , but it is actually about 20% smaller since some current is used in the biasing circuitry and the feedback branches. This means that for the measurement in Fig. 4-40,  $i_b \approx 30$  nA, and the theoretical maximum peak value for the input signal that avoids clipping is 600 mV. The measured value for 1% THD is 500 mV, which is very close.

The gain increases linearly with  $1/R_s$ , but the maximum input voltage that produces 1% THD is proportionately reduced. This results in a roughly one-to-one mapping between the output voltage limits and the bias current  $I_{SAAF}$  (for a given  $C_s$ ). Section 4.4 discusses appropriate strategies for biasing the filter properly, but generally speaking  $I_{SAAF} = 150$  nA is more than sufficient to avoid distortion for input signals as large as 10 mV<sub>pp</sub>.

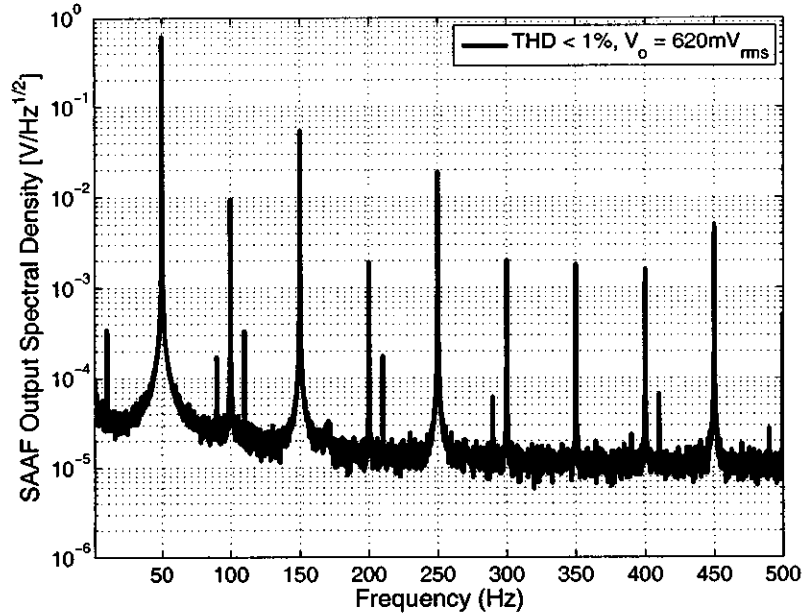


Figure 4-40: Filter's output spectrum for an output signal of  $620 \text{ mV}_{\text{rms}}$  and bias current of  $i_{SAAF} = 150 \text{ nA}$ .  $b_R = 0$ ,  $b_C = 10$ ,  $G_\phi = 1.75(4.8 \text{ dB})$ . The input signal was a  $10.8 \text{ mV}_{\text{p-p}}$  sinusoid. The total harmonic distortion was below 0.1% for frequencies between 1 Hz and 1 kHz.

#### 4.6.4 IAMP Input Impedance

To test the input impedance of the IAMP, the  $16 \text{ k}\Omega$  and  $300 \text{ k}\Omega$  resistors in Fig. 4-28 were removed and two capacitors were added ( $C_E = 1.2 \text{ pF}$ ); one between the DAC and  $v_{ip}$  and another between  $v_{im}$  and ground. These capacitors create a voltage divider with the input impedance of the IAMP such that the differential input voltage is

$$v_i = \frac{Y_E}{Y_E + Y_C + 2Y_D} v_{dac} = \frac{v_{dac}}{A_1}. \quad (4.80)$$

$Y_E = j\omega C_E$  is the admittance of each  $1.2 \text{ pF}$  capacitor.  $Y_C$  and  $Y_D$  are the common- and differential-mode admittances at the input of the IAMP. Fig. 4-41 shows the amplitude of the transfer function between the DAC's output and the IAMP's output. At low frequencies, the gain is  $19 \text{ dB}$ , which is  $20.4 \text{ dB}$  smaller than the differential gain of the IAMP. This means that the attenuation of the voltage divider in (4.80) is  $20.4 \text{ dB}$  or  $A_1 = 10.5\times$ . Our analysis in Section 4.3 shows that the theoretical values

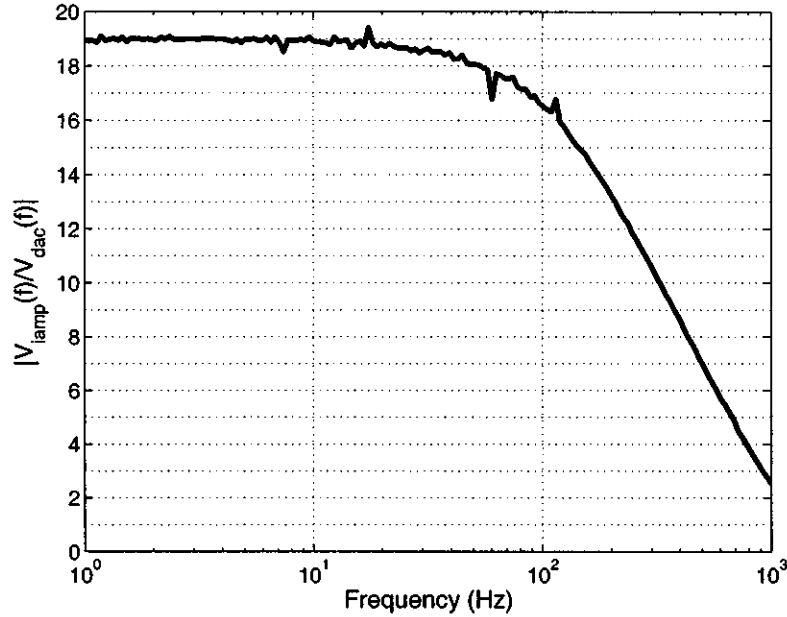


Figure 4-41: Modified IAMP gain when 1.2 pF capacitors are connected between  $v_{im}$  and ground and between the DAC's output and  $v_{ip}$ . This gain can be used to find the differential- and common-mode input impedances of the IAMP.

of the input impedances are  $Z_C = 1/(j\omega C_f)$  and  $Z_D = 2/(j\omega C_i)$ . Assuming this is true and doing some algebra, we find that

$$C_C + 2C_D \approx C_f + C_i = (A_1 - 1)C_E = 11.4 \text{ pF}. \quad (4.81)$$

Fig. 4-42 shows  $(Y_C + 2Y_D)/\omega$  which clearly shows that the input impedances are capacitive. The value of 11.4 pF is well within the tolerances of the designed values of  $C_i = 10 \text{ pF}$  and  $C_f = 100 \text{ fF}$ , and the additional 1.3 pF could easily be due to PCB trace and bondpad parasitic capacitance.

To distinguish between the differential-mode and common-mode input impedances, a second experiment is necessary. In Fig. 4-28, the three resistors are removed, the DAC is connected directly to  $v_{ip}$  and it is also connected to  $v_{im}$  through a 1.2 pF capacitor. With this setup, the input voltage of the IAMP is

$$v_i = \frac{v_{dac} C_C}{C_E + C_C + C_D} = \frac{v_{dac}}{A_2}. \quad (4.82)$$

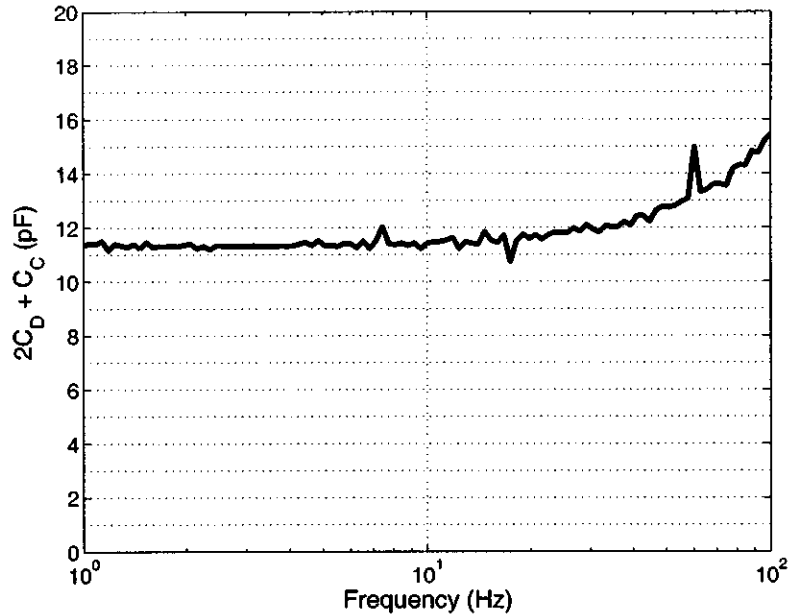


Figure 4-42: Extracted capacitance values that can be used to extract the the differential- and common-mode input impedances of the IAMP.

A frequency sweep of the IAMP's output voltage shows that the attenuation in this case is 24.7 dB, such that  $A_2 = 17.2\times$ . Applying algebra to (4.80) and (4.82) shows that

$$C_C = \frac{A_1 + 1}{2A_2 - 1} C_E = 413 \text{ fF}, \quad (4.83)$$

and from (4.81),  $C_D = 5.5 \text{ pF}$ . The theoretical values for these capacitances were  $C_D = C_i/2 = 5.0 \text{ pF}$ , and  $C_C = C_f = 100 \text{ fF}$ , although we noted earlier we expected the real values to be somewhat larger due to parasitic capacitances on the PCB and bondpads. Overall, these experiments confirm our analysis about the differential- and common-mode impedances. The resulting impedances are very large, just as desired.

### IAMP Common-Mode Rejection Ratio

The common-mode rejection ratio was measured using the same setup in Fig. 4-28, except the  $16 \text{ k}\Omega$  resistors were removed, and the two inputs to the IAMP were shorted together. This resulted in an input signal of  $100 \text{ mV}_{\text{p-p}}$ . A frequency sweep revealed that the IAMP's common-mode gain was very flat between 1 Hz–1 kHz

and  $G_{IAMP,CM} < -21$  dB for all frequencies in that range. Since the differential-mode gain is 39.4 dB, the CMRR is greater than 60 dB at all frequencies between 1 Hz–1 kHz. This is a relatively low value and probably due to mismatch in the feedback impedance networks. Nonetheless, as stated previously, the input common-mode voltage should be limited to  $v_{i,CM} \leq V_{DD} = 1.5 V_{p-p}$  using third-electrode techniques to avoid clamping from ESD protection circuits. This would result in an effective input-referred differential voltage of  $1.5 mV_{p-p}$ , which can be notched out with the mixed-signal notch.

### Filter Programmable Gain

Fig. 4-43 shows the measured gain of the filter versus resistance settings  $b_R$ , for each capacitance settings  $b_C$ . The highest line represents a setting of  $b_C = 0$  and the lowest represents  $b_C = 31$ . Fig. 4-44 shows the same data but the gain is plotted against  $b_C$ . The highest line represents  $b_R = 7$ . As discussed in Section 4.4, the filter's gain is  $G_\phi = 1/(R_s C_s f_s)$ . From Fig. 4-43 it is clear that for a given  $b_C$  setting, about 20 dB can be tuned using  $b_R$ . As also discussed in Section 4.4, predefined settings can be stored in a look-up table to tune the gain in approximately 1 dB steps.

### 4.6.5 Notch Filter

Figs. 4-45 and 4-46 show the magnitude and phase of the system transfer function including the IAMP, filter, ADC, digital block, and DAC. The magnitude plot was normalized to the mid-band gain to highlight the attenuation of the notch. The measurements were made for three different settings of  $G_f \in [2^{-18}, 2^{-20}, 2^{-22}]$ . For smaller values of  $G_f$ , the notch gets narrower, as expected, and the phase transition is also sharper. For this measurement the notch was tuned to 60 Hz, although it can easily be set to other frequencies.

Fig. 4-47 shows the system's output when a 50 Hz square wave is used for the input. The top and bottom plots respectively show the output when the notch is turned off and on.<sup>4</sup> Notice that, when the notch is turned on, the system's output

---

<sup>4</sup>Students who have taken 6.302 with Prof. Roberge will recognize this plot as a similar notch

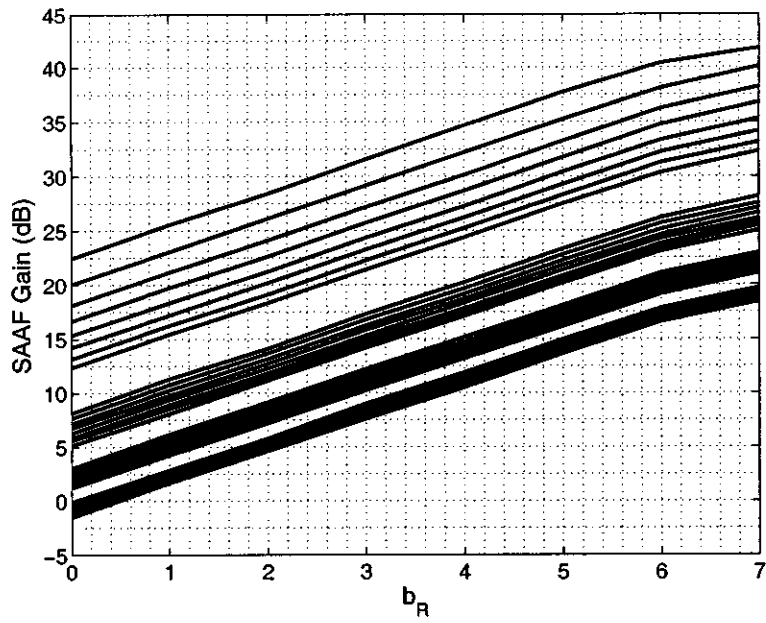


Figure 4-43: Filter gain versus  $b_R$  setting for each  $b_C$  setting. The higher lines represent low capacitance settings.

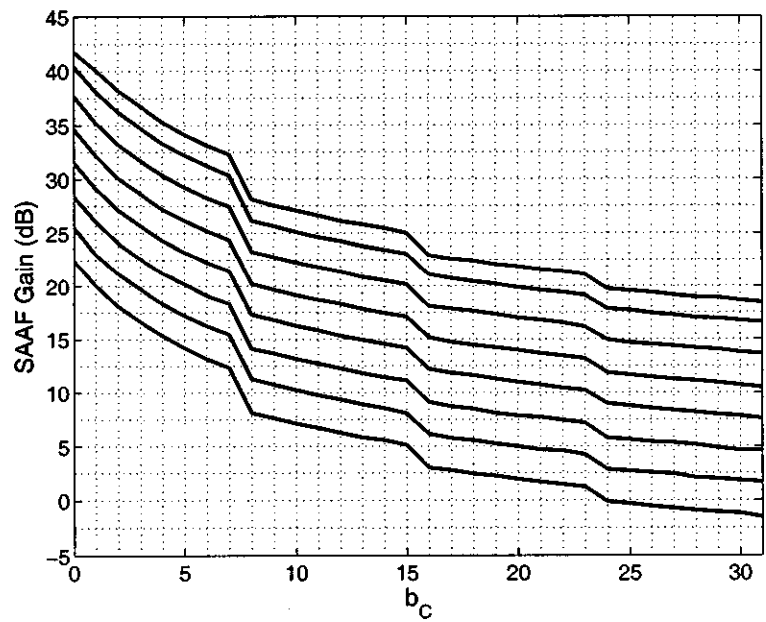


Figure 4-44: Filter gain versus  $b_C$  setting for each  $b_R$  setting. The higher lines represent low resistance settings.

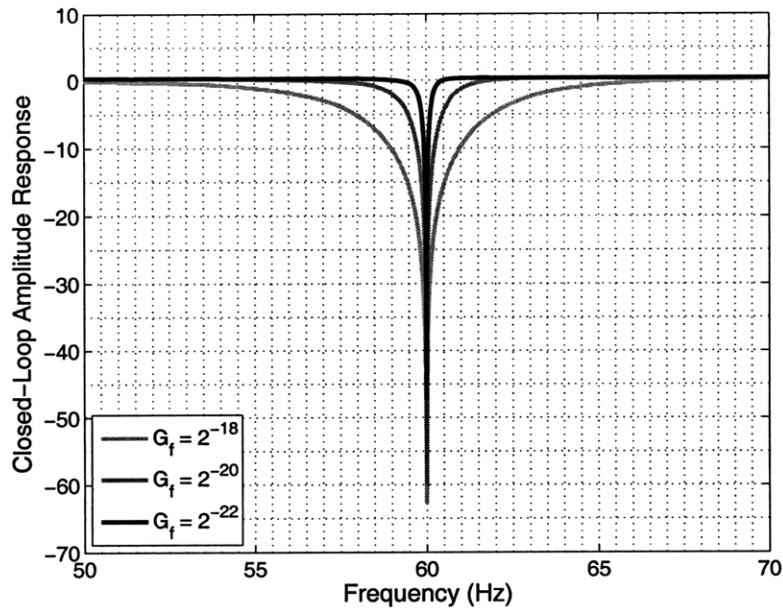


Figure 4-45: Magnitude of the closed-loop frequency response of the system with the notch on (normalized to the passband gain to highlight the notch's attenuation). The gain of the digital block was varied to show the programmability of the notch width. The center frequency of the notch can be set arbitrarily.

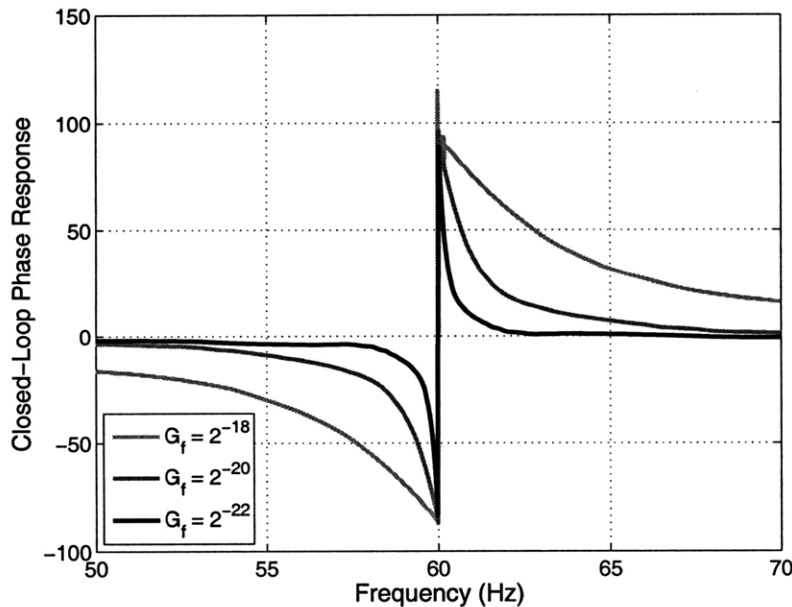


Figure 4-46: Phase of the closed-loop frequency response of the system with the notch on. Smaller settings for  $G_f$  result in a sharper notch and a narrower bandwidth of phase distortion.



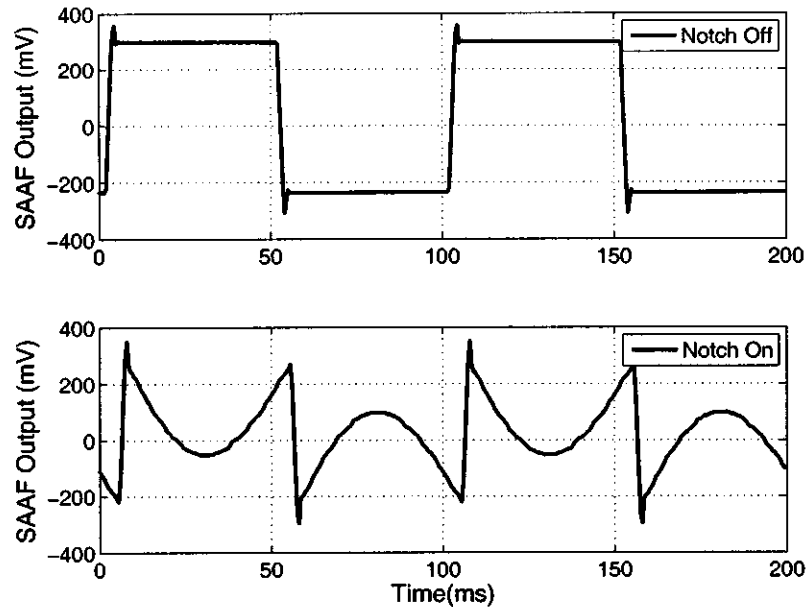


Figure 4-47: Time-domain measurement of the output of the system when the notch is turned off (top) versus on (bottom). For this measurement, the notch frequency and the square wave fundamental frequency are both set to 50 Hz. The bias settings are:  $I_{IAMP} = 400$  nA,  $I_{SAAF} = 150$  nA,  $V_{refD} = 10$  mV

looks like a square wave with the fundamental removed. Notice also that the step response of the system does not show significant peaking since the phase margin is always greater than  $90^\circ$ .

Fig. 4-48 shows the input-referred spectrum of the system's output when a relatively large input signal is applied to its input (60 Hz,  $1$  mV<sub>rms</sub>). The fundamental is attenuated by about 55 dB, but small distortion tones appear at multiples of 20 Hz caused by nonlinearities in the DAC. The largest in-band tone is smaller than  $3$   $\mu$ V<sub>rms</sub> and scrambling techniques such as sigma-delta modulation could probably be used spread their energy below the thermal noise level. Fig. 4-49 shows the input-referred noise spectrum of the system's output when no signal is applied. Note that some of the tones disappear, but the one at 60 Hz remains. This results from the system attempting to cancel out noise near 60 Hz by dithering the LSB of the DAC. It is also important to note that the thermal noise outside the notch bandwidth remains

---

experiment is done in one of his lectures.

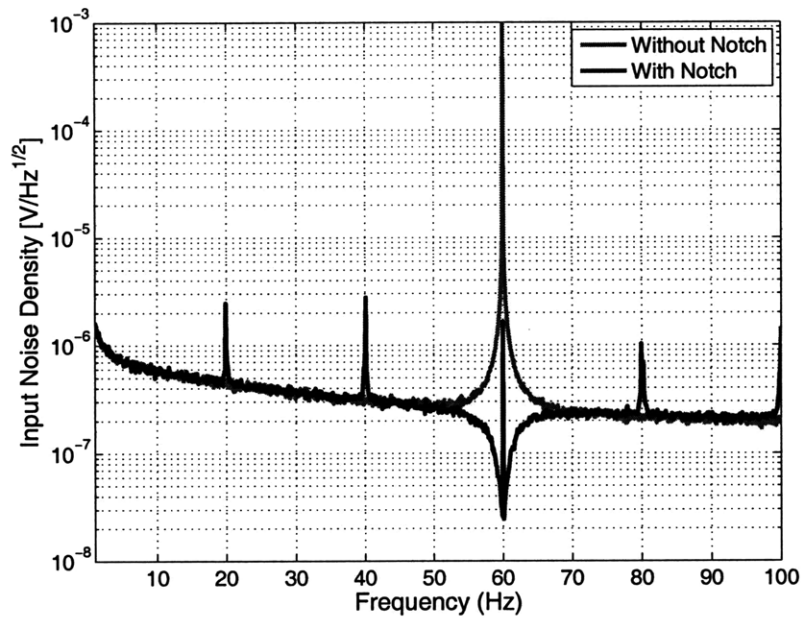


Figure 4-48: Input-referred RMS voltage spectral density when a 60 Hz, 1 mV signal is applied to the input of the system. The blue line shows that when the notch is turned on, the interference is eliminated, although other harmonics appear. The new harmonics are much smaller, however. The bias settings are  $I_{IAMP} = 400$  nA,  $I_{SAAF} = 150$  nA,  $V_{refD} = 10$  mV

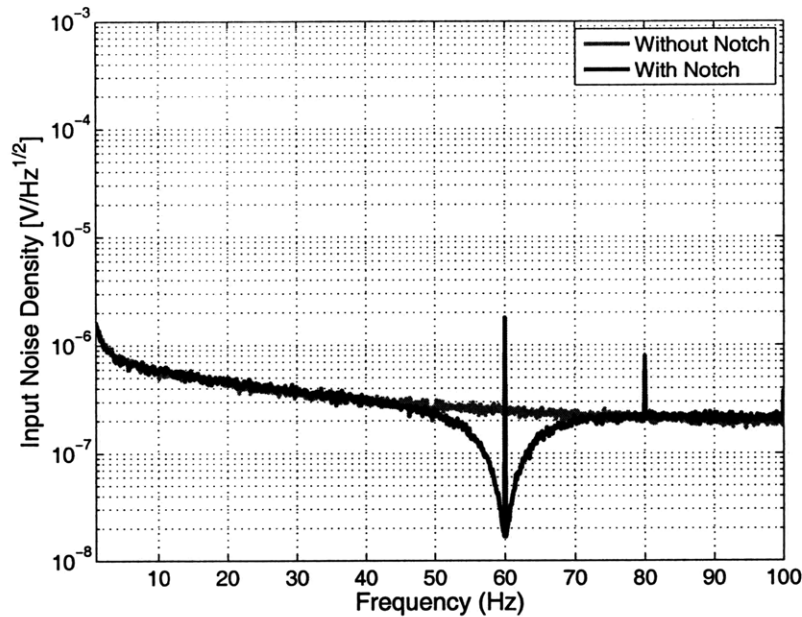


Figure 4-49: Input-referred RMS voltage spectral density when no input signal is applied. The blue line shows that any added broadband noise is negligible, although a small tone appears at 60 Hz. The bias settings are  $I_{IAMP} = 400$  nA,  $I_{SAAF} = 150$  nA,  $V_{refD} = 10$  mV

unchanged. In other words, the notch does not add significant broadband noise.

## Physiological Measurements

Fig. 4-50 shows an actual EKG measurements made with the prototyped system with the notch filter turned off and without the use of a third electrode. Clearly, a significant amount of 60 Hz noise corrupts the signal. Fig. 4-51 shows an EKG measurement when the notch filter is turned on, illustrating how effective it is at eliminating 60 Hz interference. Fig. 4-52 shows a second EKG measurement, where a  $10\text{ M}\Omega$  resistor was placed between  $v_{im}$  and ground to degrade the system's CMRR. Significantly more common-mode PLI is converted to differential-mode interference as expected. Fig. 4-53 shows the same measurement with the notch on.

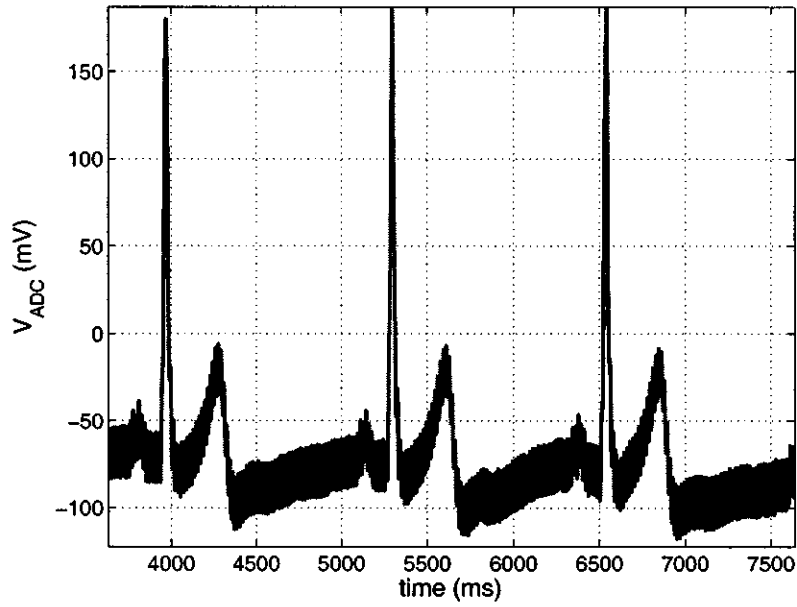


Figure 4-50: EKG measurement with the notch filter off.

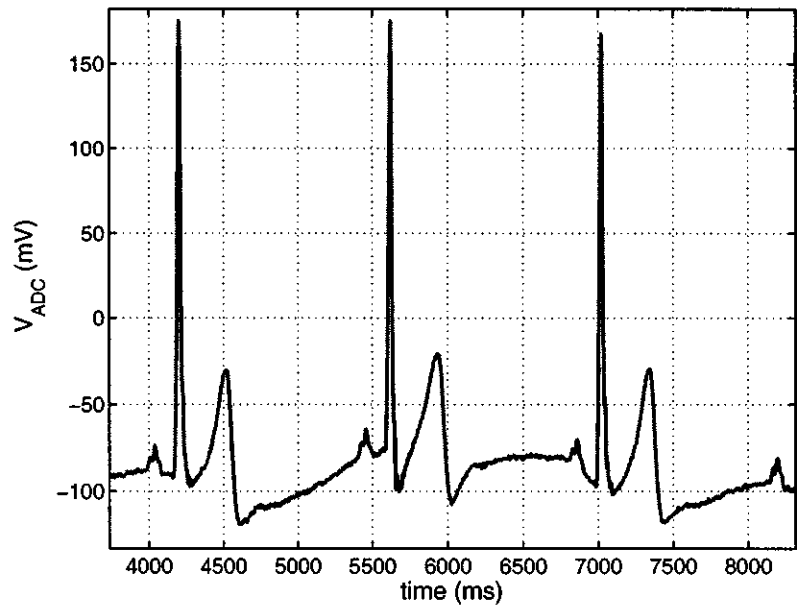


Figure 4-51: EKG measurement with the notch filter on.

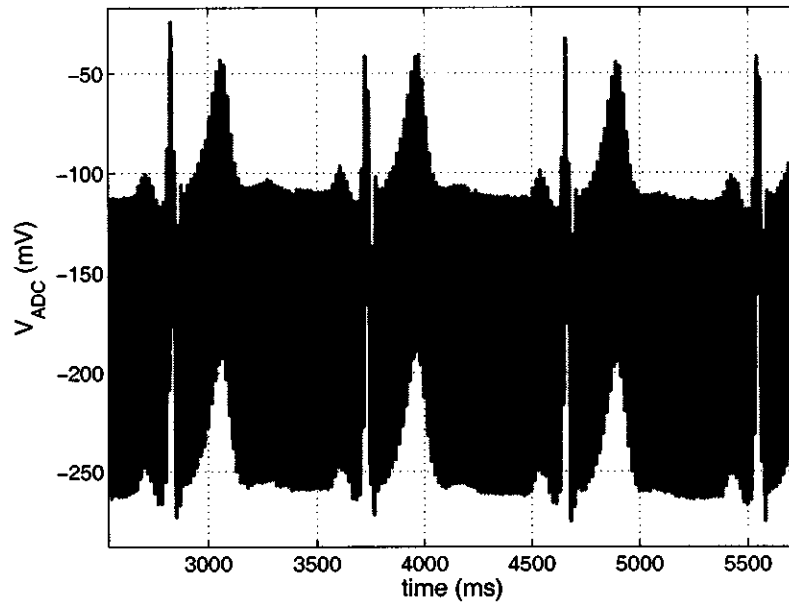


Figure 4-52: EKG measurement with the notch filter off. A  $10\text{ M}\Omega$  resistor was placed between  $v_{im}$  and ground for this measurement to create a mismatch between the common-mode input impedances and reduce the system's CMRR. This results in significantly more PLI corruption.

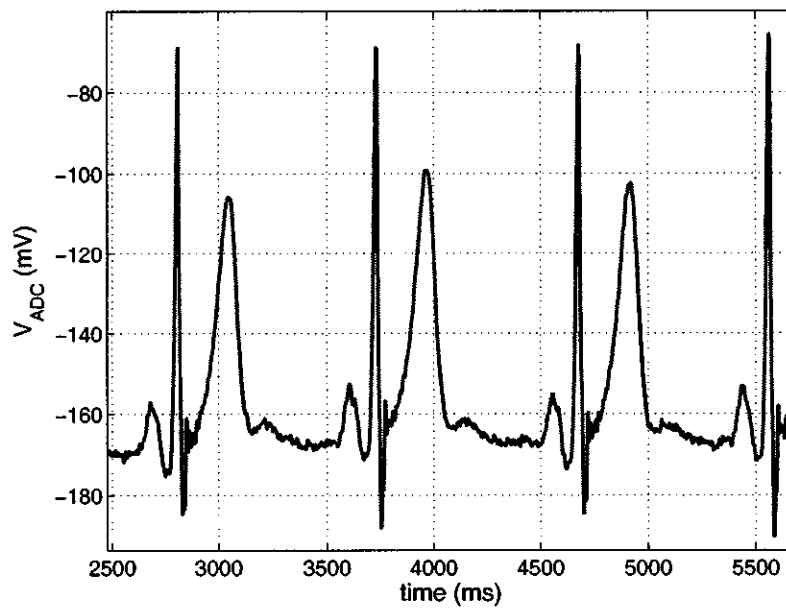


Figure 4-53: EKG measurement with the notch filter on showing how the notch filter is very effective at canceling PLI.

## 4.7 Summary

Recent advances in the energy efficiency of data converters make it possible to use oversampling when digitizing biomedical signals with minimal penalty to the overall power consumption of the system. A small, energy-efficient sensor-interface system has been presented that leverages oversampling to relax the specifications of other blocks and enable the use of unconventional technique for anti-aliasing filtering and interference cancelation. The result is a reduction in overall area and power consumption, while widening the dynamic range of the system.

Measurements of the prototype system prove the concept of performing anti-aliasing filtering using charge-sampling, which results in a sinc frequency response and programmable gain. The filter can be implemented using little area and consumes less than 225nW under most circumstances.

A notch filter that uses mixed-signal feedback was also implemented in the prototype. Its notch center frequency and width can be programmed digitally and made arbitrarily narrow without a significant area or power penalty. The additional hardware necessary to implement it is marginal because most of the blocks are necessary for open-loop operation, and the power consumption is also marginal because the feedback DAC can be implemented with passive components. For systems where the upper limit of the dynamic range is set by interferers, the notch filter can relax the specifications of the blocks in the forward path, including the ADC, resulting in smaller area and lower power.

# Chapter 5

## Conclusion and Future Work

### 5.1 Summary of Contributions

The overarching goal of the research presented in this thesis was to develop an ultra-low power transceiver for medical implant communications and a digitally-assisted sensor interface for biomedical applications. The transceiver was designed for the Medical Implant Communications Services (MICS) band, as it is the singular wireless band specifically reserved for medical implants. The super-regenerative architecture was selected for the receiver, but we found that sensitivity and selectivity analysis was lacking in the literature and determined it would be valuable to perform such analysis.

This endeavor led to the development of a novel frequency-domain model that enables the prediction of an SRA's response to arbitrary deterministic and stochastic input signals. The model, described in Chapter 2, facilitates the estimation of sensitivity and selectivity in super-regenerative receivers, and enables a performance comparison between different pulse-shaping techniques. After developing the theory, we designed a discrete-component prototype to validate the model and found there was excellent agreement between the theoretical and measured results. Specifically, we were able to predict the sensitivity of the receiver within 1 dB and the selectivity matched to within a fraction of a dB over a significant portion of the frequency response.

As described in Chapter 2, SRAs perform filtering, amplification, frequency translation, and envelope sampling with a single circuit. This eliminates the need for other RF blocks that typically consume significant power. While the SRA's selectivity is limited compared to the super-heterodyne and homodyne architectures that perform filtering at baseband frequencies, its power consumption can be significantly lower for similar sensitivity performance. Furthermore, since the input signal is injected directly into the SRA's oscillator, an inductive antenna can be used as part of its resonator. This is not possible with most other receiver architectures and enables the use of the simple transceiver described in Chapter 3.

The transceiver presented for MICS used a single digitally-controlled oscillator to implement a direct frequency-shift-keying transmitter and super-regenerative receiver. Both the transmitter and receiver achieved ultra-low power operation while meeting the specifications of the MICS band. At a bit rate of 120 kbps, the transmitter consumed less than  $350 \mu\text{W}$ , resulting in a  $2.9 \text{ nJ/bit}$  figure of merit, which is excellent for narrow band transmitters. Simple improvements to this architecture would allow a bit rate of 200 kbps that still meets MICS requirements and would lead to a  $1.8 \text{ nJ/bit}$  figure of merit.

The receiver was implemented using the same DCO as the transmitter, but configured to function as an SRA. Along with a fully differential envelope detector and a comparator with programmable DC offset, the SRA is used to synchronously demodulate on-off keyed input signals. The demodulation is performed using a digital counter to measure the startup time of the SRA. This novel technique simplifies the process for selecting an optimal demodulation threshold, and enables low supply voltage operation while bypassing the nonlinearity issues associated with linear-mode SRA receivers. With a supply voltage of 700 mV, the receiver achieved a sensitivity of  $-99 \text{ dBm}$  at a bit rate of 30 kbps and  $-93 \text{ dBm}$  at 120 kbps, while only consuming  $400 \mu\text{W}$  of power. At a bit rate of 120 kbps, this translates to a  $3.3 \text{ nJ/bit}$  FOM that is also excellent for narrow band receivers.

A frequency correction loop was also demonstrated that can be used to accurately set the carrier frequency of the transceiver without the use of a frequency synthesizer.



The FCL exploits temperature regulation in the body and incorporates the base station to eliminate the need for a crystal oscillator in the implanted device. A prototype of the system showed that it was stable under multiple conditions and converged very quickly.

In addition to the MICS transceiver, a digitally-assisted sensor-interface for biomedical applications was presented. The two main contributions of the DASIBA project were the development of a novel anti-aliasing filter and a mixed-signal feedback technique for interference cancelation. The DASIBA system leverages oversampling to relax anti-aliasing requirements and implements an area/power efficient, charge-sampling Sinc anti-aliasing filter with programmable gain. The full system includes the following on-chip components: a fully differential instrumentation amplifier, a Sinc anti-aliasing filter, and a charge-redistribution feedback DAC; along with the following off-chip components: a 16 bit ADC and an FPGA. The on-chip components were implemented on a 0.18  $\mu\text{m}$  CMOS process.

Using the feedback DAC and simple digital logic (implemented on the FPGA), a mixed-signal control loop was developed to cancel out interferers at the front end of the system. This mixed-signal feedback technique has multiple benefits. First, it relaxes the dynamic range requirements for all of the forward path blocks, including the IAMP, filter, and ADC. Second, it eliminates the need for bulky passive components used in purely analog implementations of architectures that are conceptually similar. This results in significant power and area savings. Third, the digital nature of the feedback path make the system significantly more flexible than analog implementations. For example, the center frequency of the notch and its width can be easily reconfigured digitally. Fourth, this implementation adds minimal area and power to typical open loops systems because the main blocks used to implement it are already present. The only additional hardware required is the feedback DAC and simple digital logic. Further, by implementing the feedback DAC using a charge redistribution technique, the increase in system power consumption is negligible. Finally, by adding parallel digital blocks, at minimal area and power costs, this technique can be used to notch out multiple interferers. By leveraging clever digital design, it is likely that

more elaborate filtering is possible using this technique.

The prototype developed to test many of these concepts was highly flexible and allowed tradeoffs between noise, bandwidth, and power. The anti-aliasing filter achieves more than 38 dB of attenuation in the aliasing bands while typically consuming less than 100 nW of power. It further allows for more than 40 dB of gain programmability and achieves excellent linearity for differential outputs as large as  $1.5 V_{pp}$ , while using a supply voltage of only 1.5 V. Bench-top measurements of the system showed that the mixed-signal feedback loop is capable of attenuating large interferes by more than 55 dB. Simple human tests showed the system could measure EKG signals accurately, and was effective at eliminating 60 Hz power line interference, even under simulated worst-case conditions.

## 5.2 Future Work

Despite the long history of super-regeneration circuits, they have garnered renewed interest for various reasons. First, they leverage positive feedback to effect very high gain while using components with relatively low gain. This has been exploited almost exclusively for RF receiver applications, but there may be baseband applications that can leverage carefully implemented positive feedback. So called “regenerative” comparators use positive feedback in a very similar way to SRAs, but there is little evidence in the literature that attempts have been made to use controlled quench signals to benefit from the filtering qualities of super-regeneration. As CMOS technologies scale, the intrinsic gain of transistors is degraded. Concurrently, reductions in supply voltages progressively restrict the use of cascoding. There is, therefore, a growing incentive to find innovative ways of extracting as much gain as possible from circuits with low inherent gain, and super-regeneration may play a significant role.

A second benefit to super-regeneration is that it breaks the linear tradeoff between gain and bandwidth typically associated with LTI systems. Learning how to exploit this idiosyncrasy in baseband circuits may be another interesting field of study.

For the MICS transceiver described in Chapter 3, future work may include the

fully wireless implementation of the frequency correction loop. The prototyped system used bench-top equipment to implement the frequency estimation, and the correction command was sent directly to the MICS chip through a wire-line connection. Working out synchronization issues and efficient frequency estimation would make for an interesting project. Also, the development of antennas that are small and efficient, while maintaining a relatively high Q will remain an interesting field of future study.

There are various improvements that can be made to the DASIBA system described in Chapter 4. First, the ADC should be implemented on-chip and should be co-designed with the filter. Most ADCs with low energy FOMs are successive approximation converters that sample a signal on a capacitor array and digitize its value periodically. The conversion process is usually very fast compared with the sampling period and most of the circuits are off during sampling. An area efficient implementation of the filter and ADC is possible if the filter's integrating capacitor is used as the ADC's sampling capacitor. The bulk of the SAR ADC's area is usually consumed by the capacitor array, so the area necessary to implement the ADC should be marginal. In such an implementation, a shared capacitor array would integrate the current generated by the filter's transconductor, and the ADC would digitize the output signal at the end of each period. This would eliminate the need for a buffer and reduce the area of the system.

Noise shaping can also be incorporated into the DASIBA system as a second improvement. In our current implementation, we exploited oversampling, but not noise shaping. We showed that for systems where the upper limit of the dynamic range is set by interference, the notch filter relaxes dynamic range requirements of other blocks. However, the dynamic range of the system then becomes limited by the dynamic range of the DAC. By using noise shaping through sigma-delta modulation, for example, the dynamic range of the DAC, and therefore the system, could be extended.

A third improvement to the DASIBA system would be to incorporate chopping, while leveraging the mixed-signal technique developed for the notch filter to reduce the overall area of the IAMP. As explained throughout Chapter 4 and Appendix

A, chopping is practically indispensable for low-power CMOS biomedical systems to limit the amount of corruption caused by high flicker noise at low frequencies. The architecture we implemented was flawed and prevented us from using the choppers, but other architectures are discussed in Appendix A that could be built upon. For such architectures, integrators with very high time constants are used to cancel out electrode DC offset components that otherwise saturate the system. As shown in Chapter 4, implementing these large time constants in the digital domain is far less area intensive and lends itself to greater flexibility.

# Appendix A

## Chopper-Based IAMPs

As discussed in Section ??, flicker noise and amplifier DC offset are major aggressors that degrade the performance of biomedical IAMPs. Chopping is a technique commonly used to mitigate the effects of these aggressors. Fig. A-1 shows a symbolic representation of an up-chopper and a down-chopper (top), the symbol for a fully differential chopper (bottom left), and the common switch-based implementation for such a chopper (bottom right). Fig. A-2 shows a time domain representation of what chopping does to a sinusoidal input signal  $v_a$ . The chopping waveform  $x_{ch}$  is not given the designation of a voltage to highlight the fact that switch-based choppers do not perform analog multiplication, but rather commutation (i.e. they reverse the polarity of the input signal periodically). Chopping a signal, therefore, is equivalent to multiplying it by a square-wave with unity amplitude and a DC value of zero. As shown in Fig. A-2, passing an up chopped signal through another chopper with the same phase returns the signal to its original shape (i.e. down-chops the signal).

### A.1 A Simple Chopper-Based IAMP

Fig. A-3 illustrates what a circuit implementation of a system might look like. The OpAmp is modeled as a single pole system (the dominant pole) with gain  $A_0$ . Unfortunately, there are three main weaknesses to this simple approach. First, since the signal is at a much higher frequency during amplification than before chopping, the

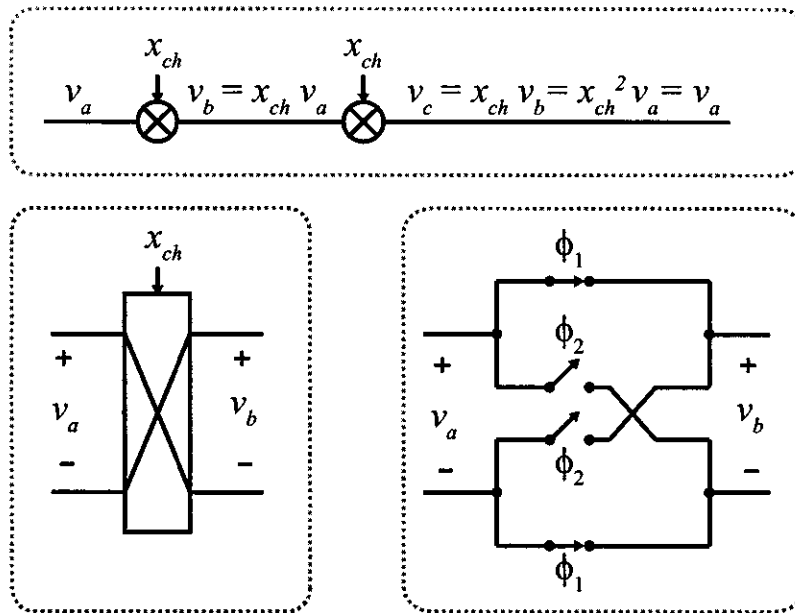


Figure A-1: A simple representation of choppers (top), a symbol for a fully differential chopper (bottom left), and a switch-based implementation of a differential chopper, typically implemented with MOS switches (bottom right). Clocks signal  $\phi_1$  and  $\phi_2$  are often made non-overlapping, but can simply be inversions of each other.

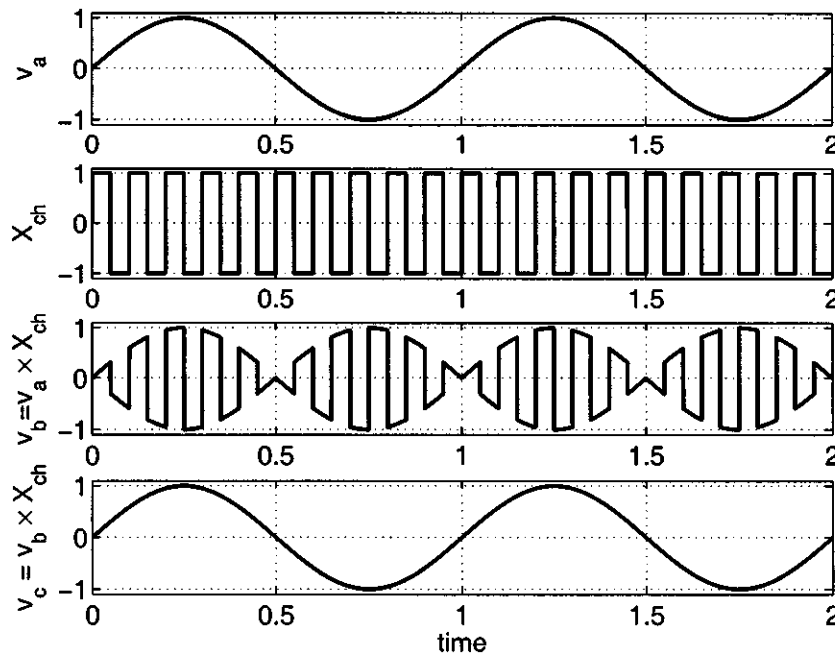


Figure A-2: Time-domain chopping waveforms.

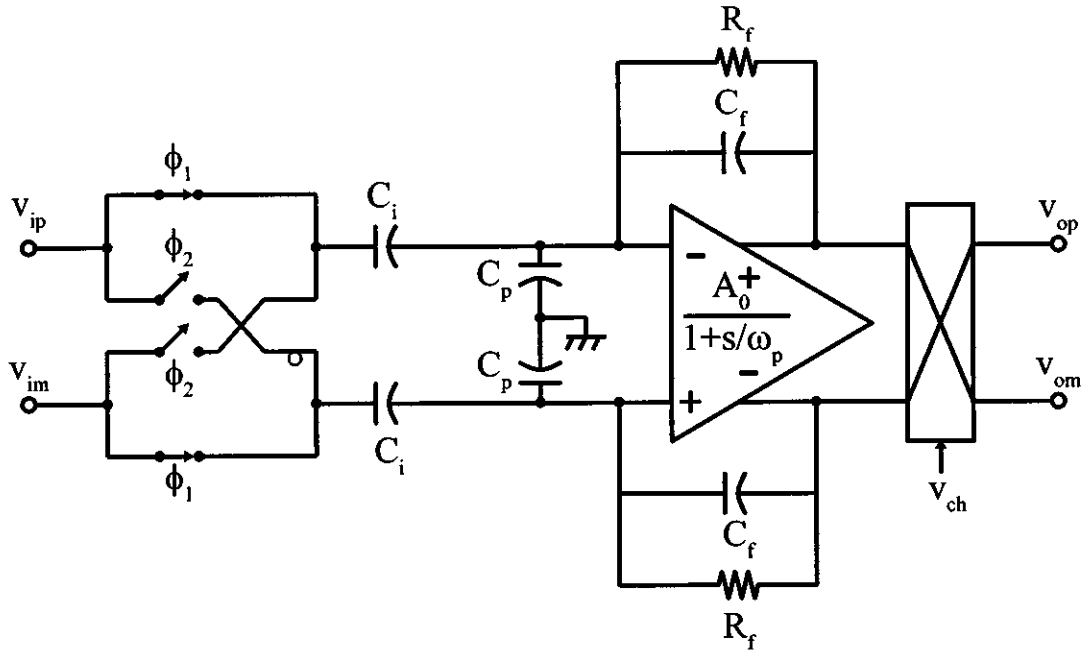


Figure A-3: A simple chopper-based IAMP.

dominant pole frequency of the OpAmp must be much higher than if chopping were not used if the benefits of feedback are needed.

For most types of OpAmps, the dominant pole frequency is proportional to the transconductance of the first stage which is proportional to the bias current (assuming weak inversion operation for CMOS amplifiers). Of course, the input-referred noise is also related to the bias current of the first stage, so the bandwidth of the OpAmp is often significantly higher than it needs to be to satisfy noise requirements. For practical IAMPs and chopping frequencies, however, the method of chopping shown in Fig. A-3 is not energy inefficient.

The second drawback to this approach is that the switches of the input chopper and the input capacitors  $C_i$  combine to create a switched capacitance resistor. The effective resistance across the inputs  $v_{ip}$  and  $v_{im}$  is

$$R_{eff} = \frac{1}{2C_i f_{ch}}. \quad (\text{A.1})$$

Since the mid-band, closed-loop gain of the IAMP is  $C_i/C_f$ ,  $C_i$  is usually much larger

than  $C_f$  and can result in a low input impedance. For example, if  $C_i=10\text{pF}$  and  $f_{ch}=10\text{kHz}$ ,  $R_{eff}=5\text{M}\Omega$ . While this value may seem large, applications such as EEG often require a differential input impedance greater than  $100\text{M}\Omega$  because the electrode impedances can be very high [74]. An input impedance of  $5\text{M}\Omega$  could attenuate the desired signal at the front end and ruin the signal-to-noise ratio.

The third drawback to this approach is that the electrode DC offset (EDO) can saturate the amplifier since it can be as large as  $\pm 300\text{mV}$  in some applications [69]. Unlike amplifier DC offset, EDO cannot be eliminated through chopping since it is added to the desired signal before the chopper. For an EDO of  $100\text{mV}$  and an IAMP gain requirement of 100, this would require a supply voltage greater than  $10\text{V}$  to avoid saturation. Put another way, for an input referred noise requirement of  $1\mu\text{V}_{\text{rms}}$ , this leads to a requirement of nearly  $100\text{dB}$  of dynamic range.

## A.2 Current Feedback IAMP with EDO Cancellation

To mitigate the effects of the second and third drawbacks (low input impedance and EDO saturation), [74] proposes using a current feedback IAMP with a feedback integrator to cancel out the up-converted EDO. Fig. A-4 shows a simplified schematic of the architecture. The mid-band gain of this amplifier is set by the ratio  $R_2/R_1$ . The transconductor uses internal feedback to achieve good linearity and a transconductance of  $G_{m1} \approx 1/R_1$ . Since there are no large capacitors connected to the input chopper, only the parasitic capacitance of the transconductor's input pair contribute to the parasitic switched capacitor resistance. Substituting  $C_p$  for  $C_i$  in (A.1) gives the parasitic resistance. As an example, for  $C_p=500\text{fF}$  and  $f_{ch}=10\text{kHz}$ ,  $R_{eff}=100\text{M}\Omega$  which meets the specifications of EEG amplifiers.

The amplifier  $A_2$ , integrator, and current source (shaded gray) in Fig. A-4 are used to create a closed-loop high-pass filter that attenuates the up-converted EDO. Intuitively, this is how the servo loop operates: for a large EDO, the output voltage



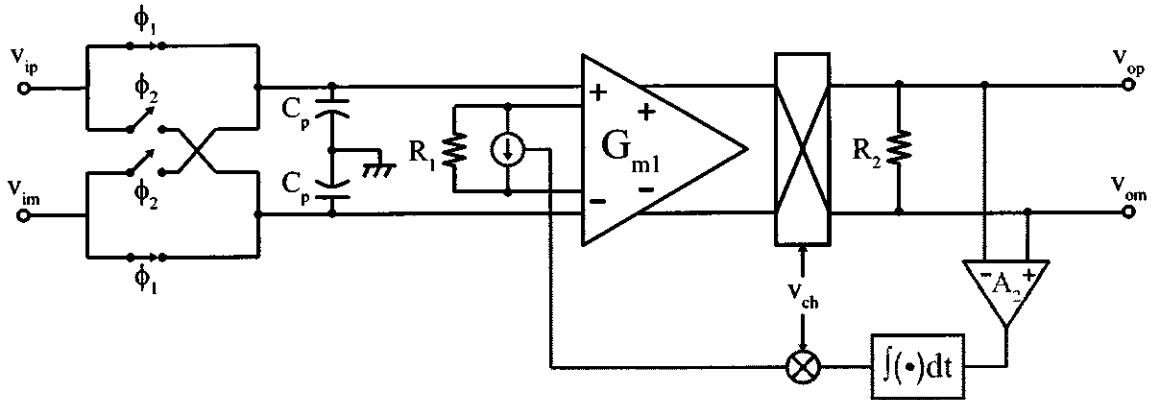


Figure A-4: A chopper-based IAMP that uses a linear transconductor instead of an OpAmp resulting in a higher input impedance [74]. A feedback path is used to attenuate the up-converted EDO.

is initially large and the transconductor saturates. This voltage is integrated, up chopped, and fed back into the transconductor via  $R_1$  in anti-phase to cancel out the up chopped EDO and prevents the transconductor from saturating while allowing the amplification of small signals. The high-pass corner frequency is set by  $R_1$ ,  $R_2$ , and  $A_2$  which is set via an external resistor and capacitor in [74]. There are limitations to how large an EDO this architecture can cancel out because the current fed back to the transconductor cannot exceed its bias current. Further, for EDO value beyond 50mV, the CMRR is degraded.

While this technique is effective at increasing the input impedance of the IAMP and filtering the EDO (for values below 50mV), it has a few of drawbacks. First, it requires the use of external components to set the integrator's gain, which make it less competitive for fully integrated, multi-electrode applications. Second, the NEF is 9.2 which is significantly worse than other designs and becomes worst for lower supply voltages. Third, a supply voltage of 3V is used which could be reduced with other architectures to minimize the power consumption. Fourth, the input common-mode DC voltage is limited to a range between 1.05V and 1.7V, which may not be guaranteed depending on the electrode setup. Finally, the bulk of the signal amplification occurs while the input signal is up chopped, which means that the bandwidth of the transconductor must be high, or the closed loop gain must be kept

relatively low. For applications where the input referred noise must be very low, this is not an issue since the transconductor's bias current must be relatively high to reduce noise, and bandwidth is extended as a result. However, for applications such as EKG where the noise specifications are less stringent, the bandwidth of the amplifier may become the dominant reason to use higher bias current, degrading the NEF of the system.

### A.3 OpAmp-based IAMP with EDO Cancelation

An alternative to the techniques of [74] was presented in [17] where chopping was also used to mitigate the effects of flicker noise and feedback was used to mitigate the effects of EDO saturation. Fig. A-5 illustrates a simplified version of the IAMP. A two stage OpAmp was used, but the down-chopping was done after the first stage as opposed to the more common method of chopping after the second stage. To highlight this, the two stages are drawn separately. As with most two-stage OpAmps, Miller-compensation is used to create a dominant pole at a much lower frequency than the non-dominant pole (illustrated here with the effective capacitance  $A_2C_c$ ). By chopping down the output current of  $G_{m1}$ , however, the desired signal is returned to base-band before it is filtered by the dominant pole. This alleviates the bandwidth requirements of the OpAmp since the desired signal is no longer at the chopping frequency. It also helps filter the up chopped DC offset of the OpAmp.

The top feedback path (black) in Fig. A-5 sets the mid-band gain similarly to classical feedback amplifiers, except that the signal must be chopped back up. A second feedback path (gray) is added to filter out the EDO using an integrator as in [74]. This path creates a closed-loop, high-pass response with a corner frequency of

$$\omega_{hp} = \frac{c_{hp}}{c_f} \omega_i, \quad (\text{A.2})$$

where  $\omega_i$  is set by a capacitor and resistor ( $\omega_i = 1/(C_{int}R_{int})$ ). Since this corner frequency must be very low ( $<0.1\text{Hz}$ ), the product of  $R_{int}$  and  $C_{int}$  must be extremely large relative to typical monolithic component values. Further, the maximum EDO

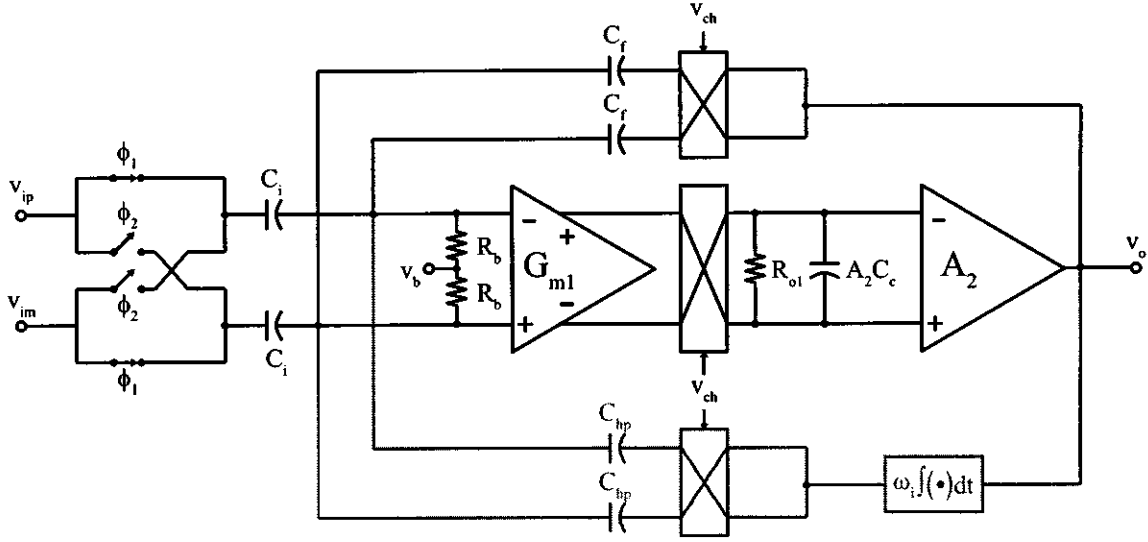


Figure A-5: A chopper-based instrumentation amplifier with an additional feedback path to cancel EDO [17]. Down-chopping is performed before the dominant pole to relax the bandwidth requirements and filter the up-converted DC offset of the OpAmp.

that can be eliminated is

$$|V_{EDO,max}| = \frac{C_{hp} V_{DD}}{C_i} \frac{1}{2}, \quad (\text{A.3})$$

creating a tradeoff between large values of  $C_{hp}$  for large EDO cancellation versus small values of  $C_{hp}$  to limit the size of  $R_{int}$  and  $C_{int}$ . Further still, the input referred noise is

$$v_{n,rti} = \left( \frac{C_i + C_{hp} + C_f}{C_i} \right) v_{n,amp}, \quad (\text{A.4})$$

creating a further incentive to minimize  $C_{hp}$ . As a result of these tradeoffs, the maximum EDO must be limited to a relatively low value. In [17], the maximum EDO that can be canceled is 50mV, and still an 800pF cap was necessary along with a 4.2G $\Omega$  switched-capacitor resistor. The integration capacitor alone consumed about 1mm<sup>2</sup> which more than doubled the total die area.

The main benefits of this technique were that it 1) reduced the 1/f corner frequency to about 1Hz, 2) allowed for up to 50mV of EDO to exist without saturating the amplifier, 3) achieved an impressive NEF of 4.6, and 4) does not suffer from the same limited input common-mode voltage range as [74]. The main drawbacks are that 1)

the input impedance is relatively low ( $8M\Omega$ ), 2) the system is relatively large ( $2\text{mm}^2$ ), 3) the CMRR is limited (80dB at 60Hz), and 4) the allowable EDO range is relatively low.

## A.4 IAMP based on a Chopper-Stabilized OpAmp

The final architecture to be examined is presented in [65] and shown in Fig. A-6. Here, an attempt was made to simultaneously correct the three weaknesses described earlier with the IAMP in Fig. A-3; namely: low input impedance, saturation due to large EDO, and need for a high bandwidth OpAmp. Note that with the exclusion of the integrator in the feedback path, this architecture is similar to the one shown in Fig. 4-7, but using a chopper-stabilized OpAmp. Since the inputs to the OpAmp are low-impedance nodes in the sense that feedback forces a virtual ground condition there, the input impedance to the IAMP is simply the series combination of the input capacitors (i.e.  $2/j\omega C_i$ ). As noted in the paper, however, the parasitic capacitors ( $C_p$ ) combine with the chopping switches to create a switched-capacitance resistor between nodes  $V_{ap}$  and  $V_{am}$ . The benefit of placing the chopper at this location is that this parasitic resistance does not affect the input impedance of the IAMP as in [17].

In [65], it is stated that the parasitic conductance created by the switched-capacitors multiplies with the amplifier's DC offset to generate a current  $I_{OS,CHOP}$ , resulting in a very large DC offset since  $R_f$  must be very large to create a sub-Hz high-pass corner frequency. The work presented in this thesis used a similar architecture, so we present a detailed analysis of this architecture.

### Detailed Analysis of an IAMP using a Chopper-Stabilized OpAmp

To simplify the analysis of the system in Fig. A-6, we first redraw the equivalent single-ended version and add the noise sources of each major element as shown in Fig. A-7. We can then create the block diagram for this schematic as shown in Fig. A-8 (A). The presence of the choppers in the block diagram can be distracting since these are generally reserved for LTI systems. With some simple manipulations, they

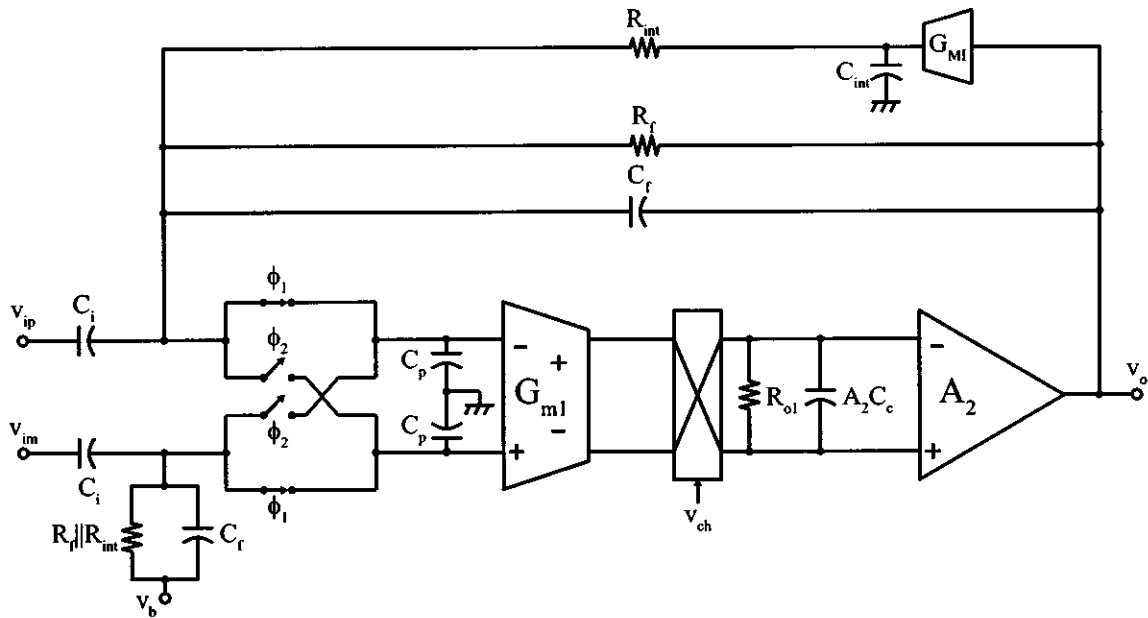


Figure A-6: A chopper-based IAMP with the input choppers connected at the input of the OpAmp [65]. A second feedback path is used to reduce the DC offset caused by the parasitic switched-capacitor resistance created by the input choppers and  $C_p$ .

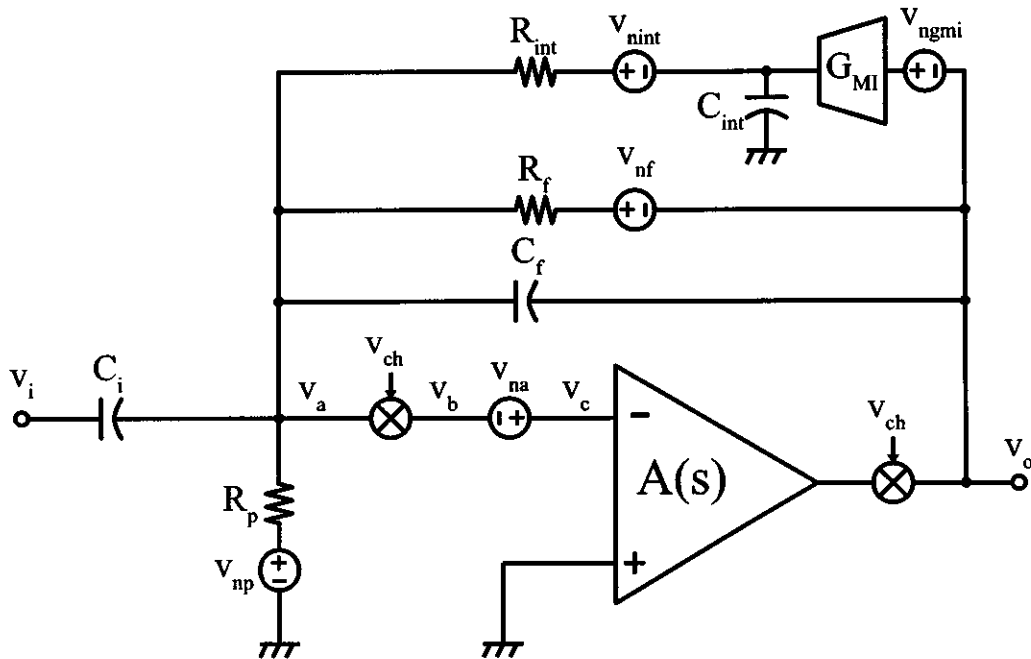


Figure A-7: Single-ended model of the chopper-stabilized IAMP in Fig. A-6 including noise sources and the parasitic switch-capacitor resistance  $R_p$ .

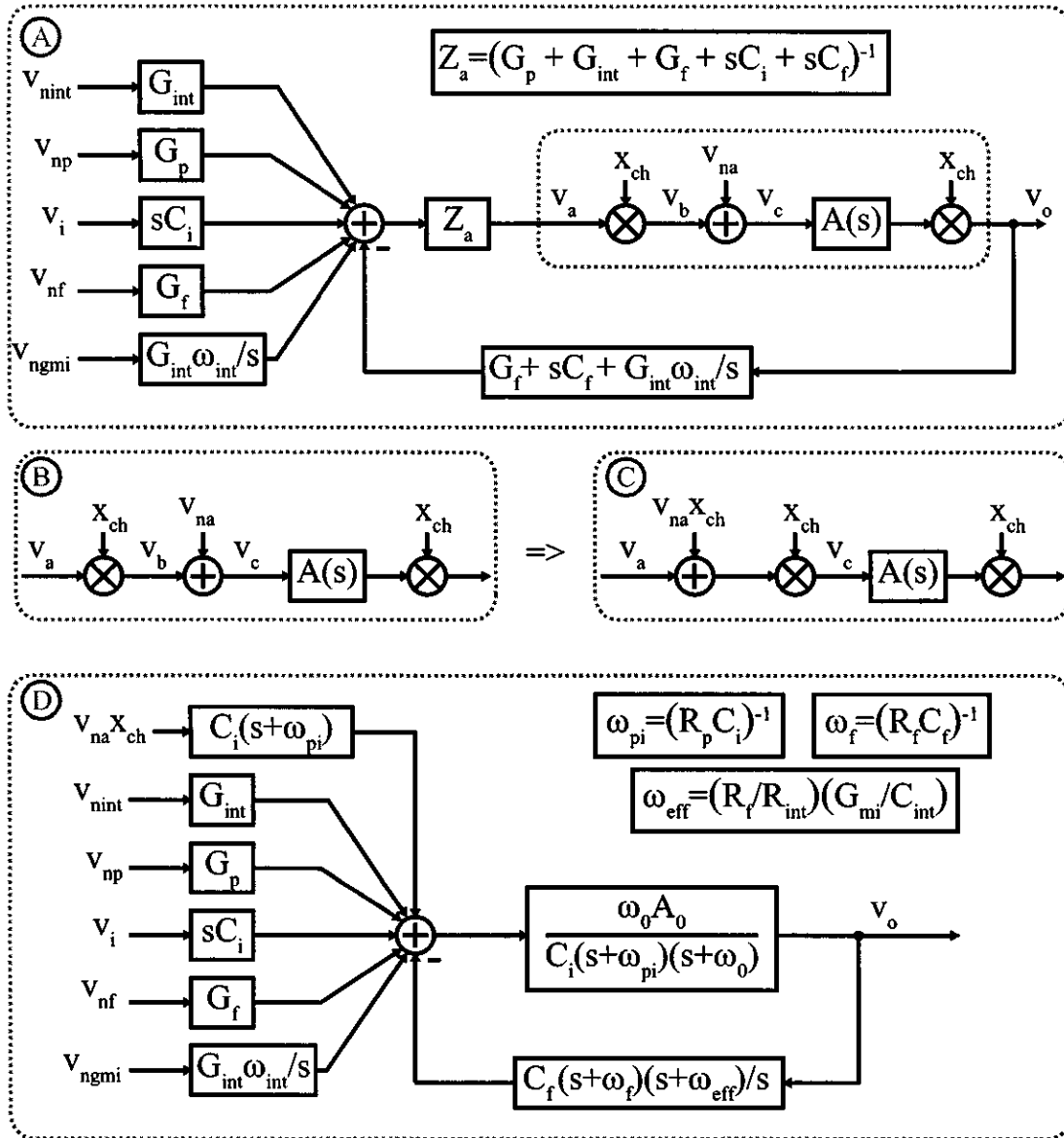


Figure A-8: (A) shows the block diagram of the system in Fig. A-7 including the choppers. (B) and (C) show how the position of the input chopper can be shifted without affecting the transfer function. (D) shows the simplified, equivalent block diagram with the choppers removed.

can be removed without changing the transfer functions of the system. First note that

$$v_c = v_a x_{ch} + v_{na} = v_a x_{ch} + v_{na} x_{ch}^2, \quad (\text{A.5})$$

as shown in Figs. A-1 and A-2. Eqn. (A.5) can be rewritten as

$$v_c = x_{ch} (v_a + v_{na} x_{ch}), \quad (\text{A.6})$$

which allows us to replace the portion of the block diagram illustrated in Fig. A-8 (B) with that in (C). We can further eliminate the two choppers in (C), since their product is unity. To further simplify the diagram, we move the  $v_{na} x_{ch}$  summer to the left of  $Z_a$  and multiply it by  $Y_a$ , which is the inverse of  $Z_a$ . Finally, we merge the two summers. The final block diagram is shown in Fig. A-8 (D) and allows us to easily determine the effect of each signal and noise source on the output voltage  $v_o$ . Fig. A-8 (D) includes the actual transfer functions of the amplifier (modeled as gain  $A_0$  with a single, dominant pole), and the impedances are rewritten as poles and zeros with the relevant frequencies defined as  $\omega_{pi}$ ,  $\omega_f$ ,  $\omega_{eff}$ , and  $\omega_0$  (the last one being the dominant pole of the OpAmp).

The loop gain is readily determined to be

$$LG(s) = \frac{\omega_0 A_0 C_f (s + \omega_f) (s + \omega_{eff})}{s C_i (s + \omega_0) (s + \omega_{pi})}. \quad (\text{A.7})$$

As an example illustrating the characteristics of this transfer function, Fig. A-9 shows the Bode plot for some typical component values and their resulting frequencies. Specifically,  $C_f = 100$  fF,  $C_i = 10$  pF,  $R_p = 500$  M $\Omega$ ,  $R_f = 10$  T $\Omega$ ,  $R_{int} = 10$  T $\Omega$ ,  $A_0 = 1\text{E}6$ ,  $\omega_{int} = 2\pi \times 10$  mHz, and  $\omega_0 = 2\pi \times 1$  Hz. The frequencies of interest are, therefore,  $f_{int} = 10$  mHz,  $f_{eff} = 10$  mHz,  $f_f = 159$  mHz,  $f_0 = 1$  Hz, and  $f_{pi} = 32$  Hz. Notice that there is no threat to stability since the phase of the loop gain is nowhere near  $\pm 180^\circ$ . Obviously the non-dominant poles of the OpAmp must be accounted for at the higher frequencies, and from the Bode plot, it is clear that they should lie near or beyond 10 kHz.

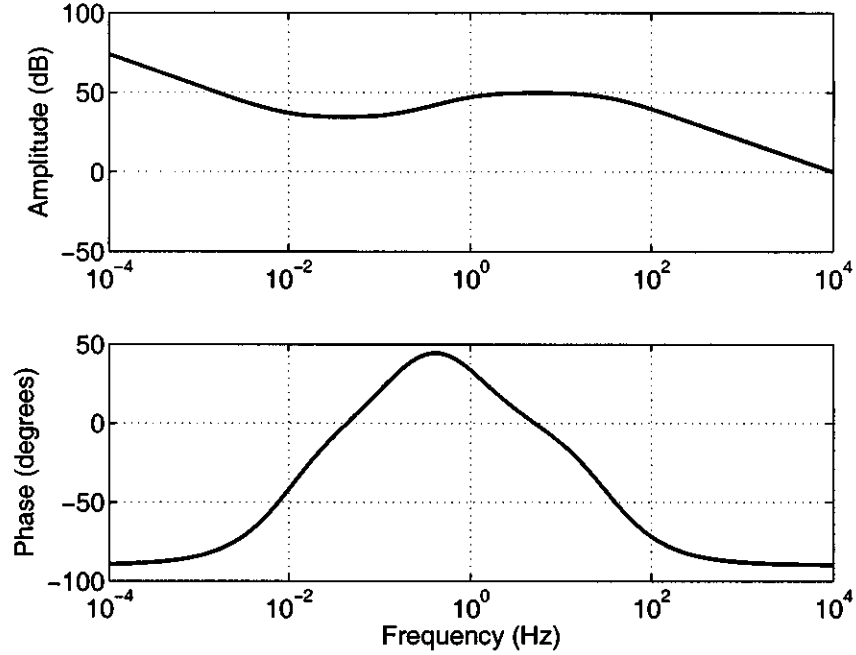


Figure A-9: Frequency response for the loop gain of the block diagram in Fig. A-8 for typical values that result in a stable system. These values are  $f_{int} = 10$  mHz,  $f_{eff} = 10$  mHz,  $f_f = 159$  mHz,  $f_0 = 1$  Hz, and  $f_{pi} = 32$  Hz.

Since the loop gain is much larger than unity for frequencies below 1 kHz, the closed loop response of the feedback loop can be accurately approximated as the inverse of the feedback factor  $\beta(s)$ :

$$H_{CL}(s) \approx \frac{1}{\beta(s)} = \frac{s}{C_f(s + \omega_f)(s + \omega_{eff})}. \quad (\text{A.8})$$

This allows us to easily determine the transfer function from the input signal and each noise source to the output voltage. For example, the signal transfer function is

$$H_{sig}(s) = \frac{s^2 C_i}{C_f(s + \omega_f)(s + \omega_{eff})}. \quad (\text{A.9})$$

The Bode plot for signal transfer function is shown in Fig. A-10. As shown, the desired high-pass characteristic is achieved and the mid-band gain is set by  $C_i/C_f = 100$ . The noise sources from each of the resistors ( $R_f$ ,  $R_p$ , and  $R_{int}$ ) each have similar transfer functions, only scaled by their conductances. For the parasitic resistor  $R_p$ , the transfer



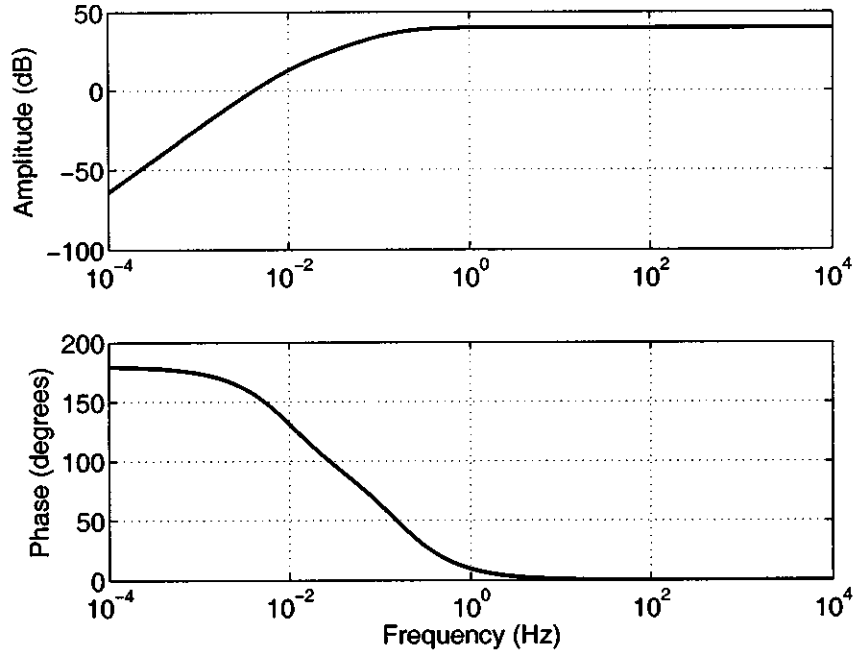


Figure A-10: Closed-loop frequency response of the input-output transfer function for the block diagram in Fig. A-8 for typical values that result in a stable system.

function is

$$H_p(s) = \frac{s}{R_p C_f (s + \omega_f)(s + \omega_{eff})}. \quad (\text{A.10})$$

Fig. A-11 shows the Bode plot for this transfer function. Unfortunately, although there is a zero at DC that eliminates any DC offset that would otherwise be amplified by the ratio  $R_f/R_p$ , the transfer function shows there is a large amount of gain at mid-band. At frequencies below 1Hz, this may not be critical since it is outside the bandwidth of the desired signal. However, at 1Hz (the typical lower bound of the signal bandwidth) the noise is amplified by 70dB. Beyond  $\omega_f$ , the amplitude frequency response is approximately

$$H_p(\omega) = \frac{1}{\omega C_f R_p}. \quad (\text{A.11})$$

Since the noise specifications of the IAMP are given with respect to the input, it is useful to refer all noise sources to the input. As shown, in the band of interest the IAMP gain is simply  $C_i/C_f$ , which means we can refer the output noise sources back to the input by dividing their transfer functions by the gain. For example, for  $v_{np}$ ,

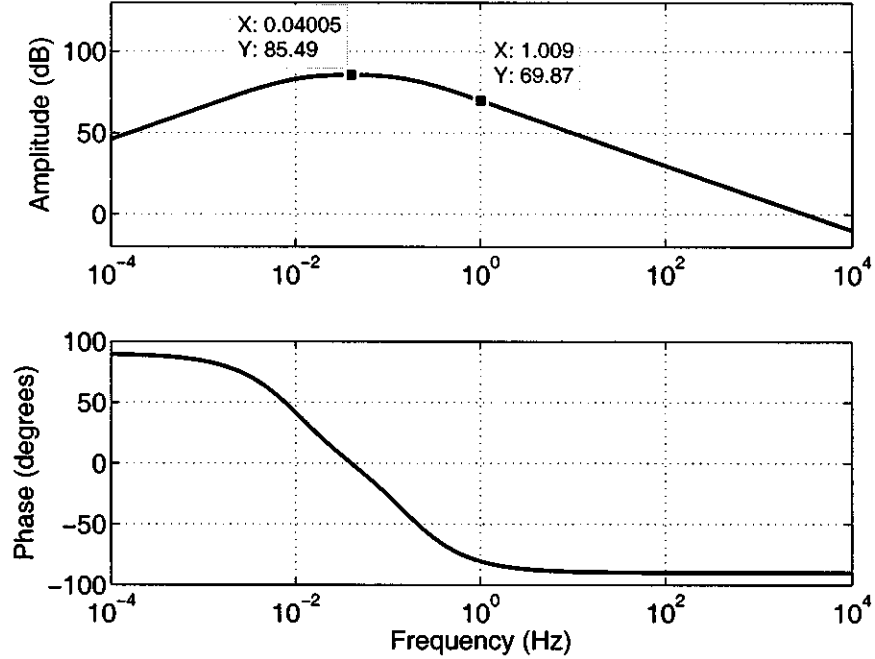


Figure A-11: Closed-loop frequency response of the resistor  $R_p$ 's noise-to-output transfer function for the block diagram in Fig. A-8.

the input referred noise transfer function can be approximated as

$$H_{p,i}(\omega) = \frac{1}{\omega C_i R_p}. \quad (\text{A.12})$$

Although  $R_p$  is a switched-capacitor resistance, it has a noise spectral density that can be modeled as white noise for frequencies well below the switching frequency and has the same value as the noise for a normal resistor of equal value; namely  $S_n(f) = 4kTR_p$  (for single-sided spectra) [22]. As a result, the input-referred noise spectral density resulting from the parasitic resistance is

$$v_{np,i}(f) = \sqrt{S_n(f)} |H_{p,i}(f)| \approx \frac{\sqrt{4kTR_p}}{2\pi f C_i R_p} = \sqrt{\frac{4kT}{R_p} \frac{1}{2\pi f C_i}}. \quad (\text{A.13})$$

As discussed in Chapter 4, the goal for input-referred voltage noise spectral density is  $100 \text{ nV}/\sqrt{\text{Hz}}$ . Using the values of  $R_p = 500 \text{ M}\Omega$  and  $C_f = 100 \text{ fF}$  used previously, we find that the input-referred noise spectral density at 1 Hz for  $R_p$  is  $92 \text{ }\mu\text{V}/\sqrt{\text{Hz}}$ . This is nearly three orders of magnitude larger than the desired value!

Note that, although  $R_p$ 's noise is white, its effective noise PSD is *colored*. In the band of interest, the input-referred noise PSD drops as  $1/f^2$  or 20dB/dec as opposed to the 10 dB/dec drop associated with flicker noise. Using the example above, this means that the noise will drop below our desired goal of  $100 \text{ nV}/\sqrt{\text{Hz}}$  for frequencies above 920 Hz.

From (A.13), it is clear that the noise effects of  $R_p$  are far worse than those of  $R_{int}$  and  $R_f$ , since those are much larger. Using the same analysis as above, yields an input-referred noise density of  $650 \text{ nV}/\sqrt{\text{Hz}}$  for each resistor, which is still larger than the desired amount, but only by a factor of 6.5 (i.e. the corner frequency would be at 6.5 Hz).

The parasitic capacitance of the OpAmp can be made only so small, and the chopping frequency must be larger than the  $1/f$  corner frequency of the OpAmp's noise PSD plus the signal bandwidth. This means that the value of  $R_p$  cannot be arbitrarily increased, and is limited to values smaller than  $10 \text{ G}\Omega$ , and more commonly  $\sim 500 \text{ M}\Omega$ . Assuming the latter, the input capacitance would have to be about  $10 \text{ nF}$  to achieve the desired noise performance, and should be even larger to reduce its noise contribution significantly below the OpAmp's. Such large capacitance values are not compatible with fully monolithic designs.

The two remaining noise sources to analyze are  $v_{na}$  and  $v_{ngmi}$ . The input referred noise of the OpAmp,  $v_{na}$ , is particularly interesting for three reasons: 1) by design, it is intended to be the dominant noise source, 2) it is up chopped in the block diagram of Fig. A-8, and 3) it contains a significant DC offset and  $1/f$  noise. As shown in Fig. 4-2, chopping a signal up (i.e. multiplying a signal by  $x_{ch}$ ) results in frequency translation. As a result, the spectral component of  $v_{na}$  that contributes to the overall output noise PSD is the thermal noise around  $f_{ch}$  and not the DC offset or flicker noise. Its transfer function is

$$H_{na}(\omega) = \frac{sC_i(s + \omega_{pi})}{C_f(s + \omega_f)(s + \omega_{eff})}, \quad (\text{A.14})$$

and its Bode plot is given by Fig. A-12. In the band of interest (i.e.  $\omega \gg \omega_f, \omega_{eff}$ ),

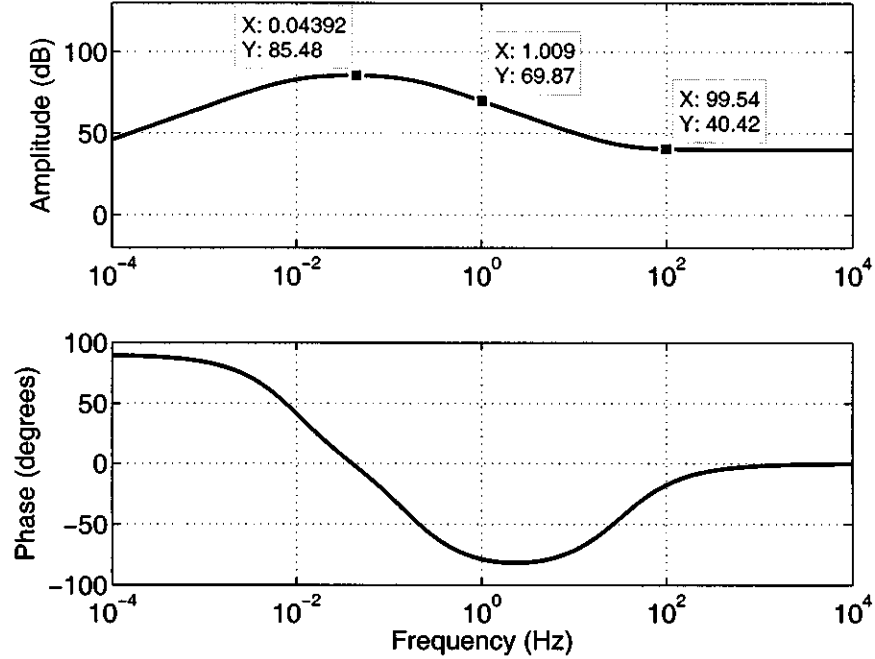


Figure A-12: Closed-loop frequency response of the OpAmp's noise-to-output transfer function for the block diagram in Fig. A-8.

and referred back to the input, (A.14) can be simplified to

$$H_{na,i}(\omega) \approx \frac{(s + \omega_{pi})}{s}. \quad (\text{A.15})$$

Recall that it is the thermal noise portion of  $v_{na}$  concentrated around  $f_{ch}$  that gets shaped by this transfer function. In other words, the noise source itself looks white. However, as with  $R_p$ , since this transfer function has a pole at  $\omega = 0$ , the effective input-referred noise is also colored and its PSD exhibits a 20dB/dec drop with frequency. The corner frequency of this  $1/f^2$  noise is  $f_{pi}$ , and beyond this frequency the noise PSD is flat.

Using the example values above, ( $R_p = 500 \text{ M}\Omega$ ,  $C_i = 10 \text{ pF}$ ), we find that the corner frequency is  $f_{pi} = 32 \text{ Hz}$ . This is significantly smaller than the corner frequency for  $R_p$ 's noise, but higher than the desired 1Hz cutoff. Note that this corner frequency is also reduced by increasing  $R_p$  and  $C_i$ .

As a final note on this architecture, recall that one of the primary goals is to increase the input impedance to prevent loading of the probes, and thereby, signal

attenuation. As with the IAMP presented in [25] and illustrated in Fig. 4-7, this IAMP has an input impedance of  $1/(j\omega 2C_i)$ . When using small capacitors amenable to chip integration (for example,  $C_i = 10$  pF), the input impedance ranges from  $80$  M $\Omega$ – $8.0$  G $\Omega$ , an acceptable range. However, as already shown, to meet the noise requirements of the system, the input capacitor must be more than 1000 times larger and the resulting input impedance would be on the order of  $80$  k $\Omega$ – $8.0$  M $\Omega$ .

This architecture has the benefits that it can be very energy efficient and withstand large electrode DC offsets. However, it has the drawbacks that it requires large external capacitors to limit low frequency noise and these capacitors reduce the input impedance. Fortunately, since the differential impedance of typical electrodes is largely capacitive, the input impedance requirements of the IAMP are frequency dependent. At higher frequencies, the impedance can be smaller, as would be the case in this architecture, without degrading the signal significantly. This means that the problem of having a lower input impedance is manageable, and tradeoffs can be made between  $1/f^2$  noise and input impedance.

In summary, for applications where noise requirements are somewhat lax at low frequencies and the use of external components is acceptable, this architecture may be adequate. It may even be desirable if a large electrode DC offset is expected and power consumption is critical. However, for architectures that are area sensitive and require a high input impedance along with low noise at low frequencies, this architecture would not be optimal.

### **Excuses, excuses**

Unfortunately, as stated earlier, we decided to use an architecture very similar to the one described in Sect. A.4 because we failed to do the detailed analysis presented here. Instead, we relied on noise simulations of the system where we removed the choppers and added the equivalent parasitic resistance  $R_p$  to model its noise. As (A.13) shows, the noise spectral density of this resistance is colored, and we mistook it for flicker noise. Since its corner frequency was roughly 1 kHz, and we were planning on chopping at  $\sim 10$  kHz, we assumed this noise would be up chopped and would

not corrupt the desired signal. Fortunately, by simply turning off the choppers, we were able to achieve good noise performance as shown in the measurements section, although without the benefits of chopping.

# Bibliography

- [1] MICS Band Plan. Part 95, January 2003.
- [2] A. Agnes, E. Bonizzoni, P. Malcovati, and F. Maloberti. A 9.4-ENOB 1V 3.8W 100kS/s SAR ADC with Time-Domain comparator. In *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, pages 246–610, 2008.
- [3] Cesare Angeloni, Patrick O. Riley, and David E. Krebs. Frequency content of whole body gait kinematic data. *IEEE Trans. Rehabil. Eng.*, 2(1):40–46, March 1994.
- [4] R. Aparicio and A. Hajimiri. A noise-shifting differential Colpitts VCO. *IEEE J. Solid-State Circuits*, 37(12):1728–1736, December 2002.
- [5] E. H. Armstrong. Some recent developments of regenerative circuits. *Proc. IRE*, 10:244–260, August 1922.
- [6] Michael Arndt. Rewiring the body; first came pacemakers. now exotic implants are bringing new hope to victims of epilepsy, paralysis, depression, and other disease. *Business Week*, 3923:74–82, March 2005.
- [7] L.D. Avendano-Valencia, L.E. Avendano, J.M. Ferrero, and G. Castellanos-Dominguez. Improvement of an extended kalman filter power line interference suppressor for ECG signals. In *Computers in Cardiology, 2007*, pages 553–556, 2007.
- [8] H. G. Beenken and F. L. Dunn. Short distance radio telemetering of physiological information. *Medical Electronics, IRE Transactions on*, PGME-12:53–57, 1958.
- [9] Jose L. Bohorquez, Joel L. Dawson, and Anantha P. Chandrakasan. A 350 $\mu$ W CMOS MSK transmitter and 400 $\mu$ W OOK super-regenerative receiver for medical implan communications. In *Proc. Symp. VLSI Circuits*, pages 32–33, Honolulu, HI, June 2008.
- [10] Jose L. Bohorquez, Joel L. Dawson, and Anantha P. Chandrakasan. A 350 $\mu$ W CMOS MSK transmitter and 400 $\mu$ W OOK super-regenerative receiver for medical implant communications. *IEEE J. Solid-State Circuits*, 44(4):1248–1259, April 2009.

- [11] L.R. Carley and T. Mukherjee. High-speed low-power integrating CMOS sample-and-hold amplifier architecture. In *Custom Integrated Circuits Conference, 1995., Proceedings of the IEEE 1995*, pages 543–546, 1995.
- [12] J. Y. Chen, Michael P. Flynn, and J. P. Hayes. A fully integrated auto-calibrated super-regenerative receiver in 0.13- $\mu\text{m}$  CMOS. *IEEE J. Solid-State Circuits*, 42(9):1976–1985, September 2007.
- [13] M.F. Chimene and R. Pallas-Areny. A comprehensive model for power line interference in biopotential measurements. *Instrumentation and Measurement, IEEE Transactions on*, 49(3):535–540, 2000.
- [14] L. W. Couch, II. *Digital and Analog Communications Systems*, page 643. Prentice Hall, Upper Saddle River, NJ, sixth edition, 2000.
- [15] J. Craninckx and G. Van der Plas. A 65fJ/Conversion-Step 0-to-50MS/s 0-to-0.7mW 9b Charge-Sharing SAR ADC in 90nm digital CMOS. In *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, pages 246–600, 2007.
- [16] T. Delbrück and C. A. Mead. Analog VLSI adaptive, logarithmic, wide-dynamic range photoreceptor. *Proc. IEEE Int. Symp. Circuits and Systems*, 4:339–342, 1994.
- [17] T. Denison, K. Consoer, W. Santa, A.-T. Avestruz, J. Cooley, and A. Kelly. A 2 W 100 nV/rtHz Chopper-Stabilized instrumentation amplifier for chronic measurement of neural field potentials. *Solid-State Circuits, IEEE Journal of*, 42(12):2934–2945, 2007.
- [18] Joseph Dien. Issues in the application of the average reference: Review, critiques, and recommendations. *Behav Res Methods Ins C*, 30(1):34–43, 1998.
- [19] D. Dobrev and I. Daskalov. Two-electrode biopotential amplifier with current-driven inputs. *Medical and Biological Engineering and Computing*, 40(1):122–127, 2002.
- [20] R. Dozio, A. Baba, C. Assambo, and M.J. Burke. Time based measurement of the impedance of the Skin-Electrode interface for dry electrode ECG recording. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5001–5004, 2007.
- [21] P. Favre, N. Joehl, A. Vouilloz, P. Deval, C. Dehollain, and M. J. Declercq. A 2-V 600- $\mu\text{A}$  1-GHz BiCMOS super-regenerative receiver for ISM applications. *IEEE J. Solid-State Circuits*, 33(12):2186–2196, December 1998.
- [22] C.-A. Gobet and A. Knob. Noise analysis of switched capacitor networks. *Circuits and Systems, IEEE Transactions on*, 30(1):37–43, 1983.



- [23] Paul R. Gray, Paul J. Hurst, Stephen H. Lewis, and Robert G. Meyer. *Analysis and Design of Analog Integrated Circuits*, pages 236–238. John Wiley and Sons, Inc., New York, NY, fourth edition, 2001.
- [24] R. R. Harrison, P. T. Watkins, R. J. Kier, R. O. Lovejoy, D. J. Black, B. Greger, and F. Solzbacher. A Low-Power integrated circuit for a wireless 100-Electrode neural recording system. *Solid-State Circuits, IEEE Journal of*, 42(1):123–133, 2007.
- [25] R.R. Harrison and C. Charles. A low-power low-noise CMOS amplifier for neural recording applications. *Solid-State Circuits, IEEE Journal of*, 38(6):958–965, 2003.
- [26] James C. Huhta and John G. Webster. 60-Hz interference in electrocardiography. *Biomedical Engineering, IEEE Transactions on*, BME-20(2):91–101, 1973.
- [27] In-Duk Hwang and J.G. Webster. Direct interference canceling for Two-Electrode biopotential amplifier. *Biomedical Engineering, IEEE Transactions on*, 55(11):2620–2627, 2008.
- [28] A. H. Johansson. Performance of a radio link between a base station and a medical implant utilizing the MICS standard. In *Proc. of IEMBS 26th Annual International Conference of the IEEE*, pages 2113–2116, September 2004.
- [29] David A. Johns and Ken Martin. *Analog Integrated Circuit Design*, chapter 4, pages 196 – 200. John Wiley and Sons, Inc., New York, NY, first edition, 1997.
- [30] Victor Karam, Peter H. R. Popplewell, Atif Shamim, John Rogers, and Calvin Plett. A 6.3 GHz BFSK transmitter with on-chip antenna for self-powered medical sensor applications. In *Proc. IEEE Radio Frequency Integrated Circuits Symposium*, pages 101–104, Honolulu, HI, June 2007.
- [31] Jaehoon Kim and Yahya Rahmat-Samii. Implantable antennas inside a human body: Simulations, designs, and characterizations. *IEEE Trans. Microw. Theory Tech.*, 52(8):1934–1943, August 2004.
- [32] M. Koyama, T. Arai, H. Tanimoto, and Y. Yoshida. A 2.5-V active low-pass filter using all-n-p-n gilbert cells with a 1-V<sub>pp</sub> range. *Solid-State Circuits, IEEE Journal of*, 28(12):1246–1253, 1993.
- [33] B. P. Lathi. *Modern Digital and Analog Communication Systems*, pages 116–129, 463, 595–597. Oxford University Press, New York, NY, third edition, 1998.
- [34] Jongwoo Lee, Hyo-Gyuem Rhew, Daryl Kipke, and Michael Flynn. A 64 channel programmable closed-loop deep brain stimulator with 8 channel neural amplifier and logarithmic ADC. In *Proc. Symp. VLSI Circuits*, pages 76–77, Honolulu, HI, June 2008.

- [35] Thomas H. Lee. *The Design of CMOS Radio-Frequency Integrated Circuits*, pages 610–620. Cambridge University Press, Cambridge, England, second edition, 2004.
- [36] Alberto Leon-Garcia. *Random Processes for Electrical Engineers*, pages 119–126. Addison-Wesley, second edition, 1993.
- [37] J. Lin. A low-phase-noise 0.004-ppm/step DCXO with guaranteed monotonicity in the 90-nm CMOS process. *IEEE J. Solid-State Circuits*, 40(12):2726–2734, December 2005.
- [38] S.M.M. Martens, M. Mischi, S.G. Oei, and J.W.M. Bergmans. An improved adaptive power line interference canceller for electrocardiography. *Biomedical Engineering, IEEE Transactions on*, 53(11):2220–2231, 2006.
- [39] J. Mateo, C. Sanchez, A. Tortes, R. Cervigon, and J.J. Rieta. Neural network based canceller for powerline interference in ECG signals. In *Computers in Cardiology, 2008*, pages 1073–1076, 2008.
- [40] Daniel R. Merrill, Marom Bikson, and John G. R. Jefferys. Electrical stimulation of excitable tissue: design of efficacious and safe protocols. *J. Neuroscience Methods*, 141(2):171–198, February 2005.
- [41] D. B. Miron. *Small Antenna Design*. Elsevier Science and Technology, Oxford, UK, first edition, 2006.
- [42] F. Xavier Moncunill-Geniz and Pere Palá-Schönwälder. Performance of a DSSS superregenerative receiver in the presence of noise and interference. In *IEEE Int. Symp. on Circuits and Systems (ISCAS) Proceedings*, pages 5684–5687, Island of Kos, May 2006.
- [43] F. Xavier Moncunill-Geniz, Pere Palá-Schönwälder, C. Dehollain, N. Joehl, and M. Declercq. An 11-Mb/s 2.1mW synchronous superregenerative receiver at 2.4GHz. *IEEE Trans. Microw. Theory Tech.*, 55(6):1355–1362, June 2007.
- [44] F. Xavier Moncunill-Geniz, Pere Palá-Schönwälder, and Orestes Mas-Casals. A generic approach to the theory of superregenerative reception. *IEEE Trans. Circuits Syst. I*, 52(1):54–70, January 2005.
- [45] C. T. Nordhausen, E. M. Maynard, and R. A. Norman. Single unit recording capabilities of a 100-microelectrode array. *Brain Res.*, 726:129–140, July 1996.
- [46] Robert Oostenveld and Peter Praamstra. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112(4):713–719, April 2001.
- [47] Alan V. Oppenheim, Ronald W. Schaffer, and John R. Buck. *Discrete-Time Signal Processing*, pages 55–62. Prentice Hall, Upper Saddle River, NJ, second edition, 1999.

- [48] B. Otis, Y. H. Chee, and J. Rabaey. A  $400\mu\text{W}$ -RX,  $1.6\text{mW}$ -TX super-regenerative transceiver for wireless sensor networks. In *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, pages 396,397,606, San Francisco, CA, February 2005.
- [49] Antti Kimmo Olavi Paukkunen and Raimo Sepponen. The effect of ground electrode on the sensitivity, symmetricity and technical feasibility of scalp EEG recordings. *Medical & Biological Engineering & Computing*, 46(9):933–938, 2008.
- [50] D. M. Pelissier and M. J. Soen. A new pulse detector based on super-regeneration for UWB low power applications. In *Proc. IEEE International Conference on Ultra-Wideband*, pages 639–644, Waltham, MA, September 2006.
- [51] Marco Piccolino and Marco Bresadola. Drawing a spark from darkness: John walsh and electric fish. *Trends in Neurosciences*, 25(1):51–57, 2002.
- [52] Nathan M. Pletcher and Jan M. Rabaey. A  $100\mu\text{W}$ ,  $1.9\text{GHz}$  oscillator with fully digital frequency tuning. In *Proc. of ESSCIRC*, pages 387–390, Grenoble, France, September 2005.
- [53] Behzad Razavi. *RF Microelectronics*, pages 49–50. Prentice Hall, Upper Saddle River, NJ, first edition, 1998.
- [54] U. L. Rohde and A. K. Poddar. Super-regenerative receiver. In *Electron Devices and Solid-State Circuits, EDSSC 2007, IEEE Conference on*, pages 263–266, Tainan, December 2007.
- [55] J. Rosell, J. Colominas, P. Riu, R. Pallas-Areny, and J.G. Webster. Skin impedance from 1 hz to 1 MHz. *Biomedical Engineering, IEEE Transactions on*, 35(8):649–651, 1988.
- [56] Adel S. Sedra and Kenneth C. Smith. *Microelectronic Circuits*, pages 915–938. Oxford University Press, Oxford, UK, fourth edition, 1998.
- [57] R.E. Serrano, M. Gasulla, O. Casas, and R. Pallas-Areny. Power line interference in ambulatory biopotential measurements. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 4, pages 3024–3027 Vol.4, 2003.
- [58] O. Sezer and M. Ferdjallah. Constrained RLS algorithm for narrow band interference rejection from EEG signal during CES. In *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, volume 1, pages 494–497, 2004.
- [59] E.M. Spinelli and M.A. Mayosky. Two-electrode biopotential measurements: power line interference analysis. *Biomedical Engineering, IEEE Transactions on*, 52(8):1436–1442, 2005.

- [60] M.S.J. Steyaert and W.M.C. Sansen. A micropower low-noise monolithic instrumentation amplifier for medical purposes. *Solid-State Circuits, IEEE Journal of*, 22(6):1163–1168, 1987.
- [61] J. Tierney, C. Rader, and B. Gold. A digital frequency synthesizer. *Audio and Electroacoustics, IEEE Transactions on*, 19(1):48–57, 1971.
- [62] M. van Elzакker, E. van Tuijl, P. Geraedts, D. Schinkel, E. Klumperink, and B. Nauta. A 1.9W 4.4fJ/Conversion-step 10b 1MS/s Charge-Redistribution ADC. In *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, pages 244–610, 2008.
- [63] A.C. Metting van Rijn, A. Peper, and C.A. Grimbergen. The isolation mode rejection ratio in bioelectric amplifiers. *Biomedical Engineering, IEEE Transactions on*, 38(11):1154–1157, 1991.
- [64] N. Verma and A.P. Chandrakasan. An ultra low energy 12-bit Rate-Resolution scalable SAR ADC for wireless sensor nodes. *Solid-State Circuits, IEEE Journal of*, 42(6):1196–1205, 2007.
- [65] Naveen Verma, Ali Shoeb, John V. Guttag, and Anantha P. Chandrakasan. A micro-power EEG acquisition SoC with integrated seizure detection processor for continuous patient monitoring. In *VLSI Circuits, 2009 Symposium on*, pages 62–63, 2009.
- [66] A. Vouilloz, M. Declercq, and C. Dehollain. A low-power CMOS super-regenerative receiver at 1GHz. *IEEE J. Solid-State Circuits*, 36(3):440–451, March 2001.
- [67] A. Wang, S. Cho, C.G. Sodini, and A. P. Chandrakasan. Energy efficient modulation and MAC for asymmetric RF microsensor systems. In *Proc. International Symposium on Low Power Electronics and Design*, pages 106–111, Huntington Beach, CA, August 2001.
- [68] W. Wattanapanitch, M. Fee, and R. Sarpeshkar. An Energy-Efficient micropower neural recording amplifier. *Biomedical Circuits and Systems, IEEE Transactions on*, 1(2):136–147, 2007.
- [69] John G. Webster. *Medical Instrumentation; Application and Design*, pages 186–194. John Wiley and Sons, Inc., Hoboken, NJ, third edition, 1998.
- [70] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. L. Nicolelis. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408:361–365, 2000.
- [71] John Reginald Whitehead. *Super-Regenerative Receivers*. Cambridge University Press, Cambridge, UK, first edition, 1950.

- [72] Bruce B. Winter and John G. Webster. Reduction of interference due to common mode voltage in biopotential amplifiers. *Biomedical Engineering, IEEE Transactions on*, BME-30(1):58–62, 1983.
- [73] Gang Xu and Jiren Yuan. Comparison of charge sampling and voltage sampling. In *Circuits and Systems, 2000. Proceedings of the 43rd IEEE Midwest Symposium on*, volume 1, pages 440–443 vol.1, 2000.
- [74] R.F. Yazicioglu, P. Merken, R. Puers, and C. Van Hoof. A  $60\mu\text{W}$   $60\text{nV}/\sqrt{\text{Hz}}$  readout Front-End for portable biopotential acquisition systems. *Solid-State Circuits, IEEE Journal of*, 42(5):1100–1110, 2007.