

UCLA

UCLA Electronic Theses and Dissertations

Title

Multi-Channel High-Dynamic-Range Implantable VCO-Based Neural-Sensing System

Permalink

<https://escholarship.org/uc/item/3482z7nh>

Author

Jiang, Wenlong

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Multi-Channel High-Dynamic-Range Implantable
VCO-Based Neural-Sensing System

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

Wenlong Jiang

2017

© Copyright by
Wenlong Jiang
2017

ABSTRACT OF THE DISSERTATION

Multi-Channel High-Dynamic-Range Implantable VCO-Based Neural-Sensing System

by

Wenlong Jiang

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2017

Professor Dejan Marković, Co-chair

Professor Asad A. Abidi, Co-chair

Neuromodulation is the alternation of nerve activity through targeted delivery of a stimulus, such as electrical stimulation, to specific sites in the body. Deep brain stimulation (DBS) is a commonly-used neuromodulation treatment for neurological ailments when traditional methods, such as surgery, medication or psychotherapy, fail. DBS is performed by sending controlled electrical pulses into the brain to evoke the desired response. However, existing DBS systems can only administer open-loop stimulation over a limited number of channels. Future neuromodulation systems require a multi-channel *closed-loop* platform that can provide high spatial precision, and automatically adjust stimulation parameters based on feedback from recorded neural signals. This multi-channel, closed-loop system poses new challenges for brain-sensing circuit and system design. To enable closed-loop operation, the sensing circuit needs to work with concurrent stimulation. Therefore, it needs to provide a large input range to prevent saturation under stimulation artifacts. In addition, the sensing circuit should simultaneously meet device/patient safety constraints. Current state-of-the-art neural sensing circuits, however, do not meet these requirements.

This work presents an implantable VCO-based neural-sensing front-end design intended for multi-channel, closed-loop neuromodulation applications. Specifically, it converts the input voltage into the phase domain, and performs direct digitization without any voltage-

domain amplification, thus preventing saturation. The phase-domain processing allows a large input range that can comprise both stimulation artifacts and the neural signals. Four techniques have been implemented to overcome design challenges: (1) in the high-pass filter, we utilize a multi-rate duty-cycled resistor as a reliable solution to attenuate electrode-offsets; (2) inside of the VCO, chopping is applied to lower circuit noise; (3) at the analog-digital interface, we employ a new glitch-free quantizer; and (4) after digitization, circuit linearity is restored through the digital non-linearity correction. With these techniques, the design achieves $10\times$ linear range and 2-3 bit ENOB improvement over prior-art with comparable power and noise performance.

This work also presents a 32/64-channel sensing chip based on the proposed front-end design. The chip is assembled on a miniaturized PCB to achieve a fully integrated neuromodulation system. Sensing performance and function under concurrent stimulation have been verified in bench-top and *in-vitro* environments. This allows further development of a complete multi-channel closed-loop neuromodulation implant.

The dissertation of Wenlong Jiang is approved.

Nanthia A. Suthana

Sudhakar Pamarti

Asad A. Abidi, Committee Co-chair

Dejan Marković, Committee Co-chair

University of California, Los Angeles

2017

TABLE OF CONTENTS

Acronyms	1
1 Introduction	3
1.1 Motivation	3
1.2 Dissertation Outline	8
2 Neural-Sensing Front-End Design	10
2.1 Specifications of the Front-End Design	10
2.2 Review of Prior Art	13
2.3 VCO-Based Structure	17
2.4 Multi-Rate Duty-Cycled-Resistor Based HPF	22
2.4.1 HPF Requirements and Prior Art	22
2.4.2 Duty-Cycled-Resistor (DCR) Based HPF	27
2.4.3 Multi-Rate HPF Design	30
2.5 VCO Design	32
2.5.1 VCO Schematic	32
2.5.2 Chopping Inside of the VCO	35
2.6 Quantizer Design	36
2.6.1 Signal Processing of the Quantizer	36
2.6.2 Coarse-Fine Quantization of the RO Phase	40
2.6.3 Glitch-Free RO Sub-Quantizer Design	43
2.7 VCO Nonlinearity Correction	53
2.8 Front-End Implementation	57

3	Front-End Measurement Results	61
3.1	Noise Measurement	61
3.2	Linearity Measurement	64
3.2.1	Single-Tone and Two-Tone Test	64
3.2.2	Linearity Stability Under Temperature Variation	69
3.3	Power Measurement	70
3.4	Front-End Input-Interface Measurement	71
4	Sensing-System Integration	74
4.1	32/64-Channel Neural-Sensing Chip	74
4.2	Miniaturized Neuromodulation Unit	78
4.3	Concurrent Sensing and Stimulation Measurement	81
4.4	Neural-Sensing with ASAR	85
5	Conclusion	87
5.1	Comparison and Research Contributions	87
5.2	Future Work	89
	References	93

LIST OF FIGURES

1.1	Traditional DBS system (Medtronic Activa SC system shown).	5
1.2	Diagram showing the closed-loop neuromodulation concept.	6
1.3	Frequency and amplitude of neural signals.	7
1.4	This work bridges the gap between prior implants and wall-powered systems. . .	8
2.1	Neuro amplifier schematic by Harrison [1].	14
2.2	Chopper-stabilized instrumentation amplifier [2].	15
2.3	HermesE system: reset upon detecting large signals [3].	15
2.4	Block diagram for FDM-based neuromodulation system [4].	16
2.5	Direct digitization at the input with 1st order $\Sigma\Delta$ modulator [5].	17
2.6	VCO-based ADC architecture [6].	18
2.7	VCO-based front-end does not saturate and allows signal recovery [7].	19
2.8	Previous VCO-based front-end design [8] does not address existing challenges. .	21
2.9	The implemented VCO-based front-end contains VCO, quantizer, and NLC block.	22
2.10	Output noise PSD of the HPF.	24
2.11	Pseudo resistor schematic.	25
2.12	Schematic and intuitive explanation for duty-cycled-resistor based HPF.	28
2.13	Generation circuit for the <i>swclk</i> signal.	30
2.14	Multi-rate duty-cycled-resistor based HPF.	31
2.15	The VCO schematic for the front-end.	33
2.16	Chopping operation of the VCO. (a) connection at the positive phase; (b) con- nection at the negative phase.	35
2.17	Spectrum-domain representation of the chopping operation.	36
2.18	The post-chopping operation in the quantizer.	38

2.19	Quantizer schematic and timing diagram.	39
2.20	Schematic for the phase decoder.	41
2.21	Example of unwrapping RO phase with cycle count and phase decoded output.	43
2.22	Ideal RO sub-quantizer structure.	44
2.23	Practical timing issues with RO sub-quantizer.	45
2.24	Example of output glitch due to timing issues.	46
2.25	Dual-edge retiming in ADPLL resorts to redundancy from TDC [9].	47
2.26	Two counters at dual edges of the oscillation to remove glitch condition [10].	47
2.27	Multi-Latching technique on resolving the timing issue.	48
2.28	Example scenarios for the arbitration of the multi-latching technique.	50
2.29	Retiming of the <i>Counter Enable</i> signal.	51
2.30	Voltage interfacing between RO and the buffer leading to the half-cycle glitch.	52
2.31	Voltage-to-frequency transfer function of the VCO.	56
2.32	NLC implementation: Horner's method for polynomial correction.	57
2.33	Spectrum-domain inspection of the VCO based front-end.	58
2.34	Front-end silicon micrograph.	59
3.1	Test board for the front-end.	62
3.2	The input-referred noise PSD of the front-end.	63
3.3	Single-tone test for 7 Hz, 100 mV _{pp} input, without NLC.	65
3.4	Single-tone test for 7 Hz, 100 mV _{pp} input, with NLC.	66
3.5	Single-tone test for 203 Hz, 50 mV _p input, without NLC.	66
3.6	Single-tone test for 203 Hz, 50 mV _p input, with NLC.	67
3.7	Two-tone test for input of 50 mV _p at 103 Hz and 10 mV _p at 90 Hz, without NLC.	68
3.8	Two-tone test for input of 50 mV _p at 103 Hz and 10 mV _p at 90 Hz, with NLC.	69

3.9	Temperature test for front-end linearity performance.	70
3.10	Measurement of front-end low frequency response.	72
3.11	Measurement of DC input impedance or leakage current.	73
4.1	Sensing chip tape-out iterations.	75
4.2	Schematic for the sensing-system chip.	77
4.3	Layout for the sensing-system chip.	78
4.4	Front-end cluster.	79
4.5	Proposed neuromodulation system.	80
4.6	Fabricated and assembled NM unit.	81
4.7	Diagram of the test on concurrent sensing and stimulation.	82
4.8	Test set-up for concurrent sensing and stimulation [11].	83
4.9	The NM unit GUI for the test [12].	84
4.10	Waveform and spectrum plot for the concurrent-sensing-and-stimulation test. . .	85
5.1	A possible way to deal with large input common-mode fluctuation.	90

LIST OF TABLES

2.1	Front-end specifications.	13
2.2	Example for phase decoder outputs.	41
3.1	Front-end input-referred RMS noise performance.	64
3.2	Linearity performance result across temperatures.	70
3.3	Front-end power consumption measurement result.	71
5.1	Comparison with prior art.	88

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisors, Prof. Marković and Prof. Abidi. Prof. Marković has given me a rare opportunity of joining the research on neuromodulation circuit and system design. He has always been supportive of my endeavors and accessible for any of my questions or concerns. I have also been given the privilege of gaining advice from Prof. Abidi. His pursuit of high-quality teaching and research inspired me all along my PhD study. His knowledge and foresight has often enlightened me. I would also like to express sincere gratitude to the other committee members. Prof. Pamarti has been instructive in the class, TA work, and helping evaluate my research. In addition, the collaboration with Prof. Suthana's group started this research project.

I have been very fortunate to work at the Integrated Circuit and System Laboratory (ICSL), where people are both welcoming and brilliant. I have enjoyed learning from everyone and sharing delightful experiences with them: Dr. Vaibhav Karkare, Hariprasad Chandrakumar, Dr. Vahagn Hokinikyan, Sina Basir-Kazeruni, Dr. Dejan Rozgic, Hao Xu, Hsin-hung Chen, Alireza Yousefi, Dr. Neha Sinha, Dr. Jeffery Lee, Jiacheng Pan, Dihang Yang, Weiyu Leng, Wenhao Yu, and many others. Some other friends have also been very helpful in my research and life, including but certainly not limited to: Dr. Boyu Hu, Dr. Zuow-Zun Chen, Dr. Long Kong, Jinxi Guo and Dr. Linqi Song.

I am grateful to all of the staff in the Electrical and Computer Engineering Department for their invaluable assistance behind the scenes, particularly: Kyle Jung, Deona Columbia, Ryo Arreola and Mandy Smith.

My thanks also extend to the UCLA Jiu-Jitsu Club, and especially my coach Kris Martin. It always feels like home when I enter the club for practice. The difficult work and class discipline has taught me to enhance my resilience and composure under stress. I would also like to thank Peter Stamatopoulos, whose enthusiastic instruction in Pilates has allowed me to remain healthy and energetic under the hectic PhD research schedule.

Most of all, I want to thank my parents for all of the love and support they have provided to me during my life. Their constant belief in me has given me the courage to overcome

all obstacles while embarking on my work thousands of miles away from them and across different continents.

VITA

- 2005–2009 B.E. (Electrical Engineering), Tsinghua University, Beijing, China
- 2009–2011 M.Sc. (Microelectronics), Delft University of Technology, Delft, the Netherlands
- 2011–2012 Trainee, Broadcom Corporation, San Diego, California, U.S.A.

PUBLICATIONS

W. Jiang, V. Hokinikyan, H. Chandrakumar, V. Karkare, and D. Marković, “A $\pm 50\text{mV}$ linear-input-range VCO-based neural-recording front-end with digital nonlinearity correction,” *IEEE J. Solid-State Circuits*, vol. 52, no. 1, pp. 173-184, Jan. 2017.

W. Jiang, V. Hokinikyan, H. Chandrakumar, V. Karkare, and D. Marković, “A $\pm 50\text{mV}$ linear-input-range VCO-based neural-recording front-end with digital nonlinearity correction,” in *IEEE ISSCC Dig. 2016*, pp. 484-485, Feb. 2016.

D. Rozgic, V. Hokinikyan, W. Jiang, S. Basir-Kazeruni, H. Chandrakumar, W. Leng, and D. Marković, “A True Full-Duplex 32-Channel 0.135 cm^3 Neural Interface,” in *The 13th IEEE Biomedical Circuit and System Conference (BioCAS)*, Oct. 2017.

Acronyms

ADC analog-to-digital converter.

AP action potential.

ASAR adaptive stimulation artifact rejection.

CMRR common-mode rejection ratio.

CSF cerebrospinal fluid.

DARPA Defense Advanced Research Projects Agency.

DBS deep brain stimulation.

DCR duty-cycled resistor.

DNL differential nonlinearity.

ENOB effective number of bits.

FDA U.S. Food and Drug Administration.

FDM frequency-division multiplexing.

FPGA field-programmable gate array.

GUI graphical user interface.

HPF high-pass filter.

INL integral nonlinearity.

IPG implantable pulse generator.

LDO low-dropout regulator.

LFP local field potential.

MAC multiplication and accumulation.

NLC nonlinearity correction.

NM neuromodulation.

PBS phosphate-buffered saline.

PCB printed circuit board.

PSD power spectrum density.

PSRR power-supply rejection ratio.

RMS root mean square.

RO ring oscillator.

SNR signal-to-noise ratio.

SPI Serial Peripheral Interface.

VCO voltage-controlled oscillator.

XO crystal oscillator.

CHAPTER 1

Introduction

1.1 Motivation

Neuropsychiatric disorders are reported as the third-leading cause of disability globally and the foremost-leading cause in the U.S. [13]. They include, but are not limited to, debilitating diseases, such as Parkinson's, Alzheimer's, epilepsy, and depression. Such diseases hinder patients in their daily lives and pose high economic costs on patient care. They are often, at least partially, caused by aberrant behavior of certain brain neural circuit(s). Unlike other well-investigated diseases, such as diabetes or cancer, a full understanding of the mechanisms behind these brain diseases remains undefined. This challenge has motivated increasing numbers of initiatives worldwide on studying and treating brain disorders [14].

Among the extant studies, neuromodulation technology constitutes a major focus. It serves as an alternative treatment method when traditional ones such as surgery, medication or psychotherapy, reach their limits. For example, neurosurgery may lead to irreversible brain damage; long-term medication may trigger drug resistance; and psychotherapy is ineffective in curing neurodegenerative diseases. Instead, the neuromodulation method targets certain brain regions and regulates their behavior with stimuli. With a controllable and precise stimulus injected directly into the brain, the neuromodulation method may overcome the obstacles encountered by the traditional methods.

A primary branch of neuromodulation is deep brain stimulation, in which electric pulses are injected as the stimulus. It has been demonstrated to offer symptomatic relief/cure for some diseases [15][16]. Based on clinical trial results, the FDA approved it for treating essential tremor in 1997, Parkinson's disease in 2003, and epilepsy in 2014. An example of

the DBS system, the Medtronic Activa system, is shown in Fig. 1.1. In this system, one end of the probe is inserted into the target region of the brain. The other end of the probe is connected via a cable to an implantable pulse generator (IPG), which generates pulses for stimulation. However, the following shortcomings exist in current DBS systems:

- The stimulation is open-loop, i.e. without real-time information on the brain state. Therefore, it is typically always kept-on. Since the patient may have fluctuating levels of pathological activity, the stimulation can sometimes lead to major side effects, such as impairment of speech, gait, or balance [17].
- The stimulation therapy's details, such as location, interval and duration, need be highly personalized for optimum results in certain applications. The required tuning into these details can sometimes take weeks, requiring several visits to the hospital, which is inconvenient for patients.
- Current stimulation systems have only a limited number of electrodes on the probe, leading to very coarse stimulation precision. The brain neural circuit can become less responsive to the stimulus over time, which is known as "habituation". Under habituation, a change of stimulation location may be requisite, but probes with a limited channel number cannot provide this capability.

Neuromodulation is now limited by these shortcomings of current DBS systems, and its future development requires a system that has a large channel count, performs closed-loop modulation, and is implantable. A large channel count, which could range from 20 to more than 100, would enhance the spatial resolution of the stimulation for optimizing treatment, and provide the possibility of refining the stimulation upon brain habituation. The concept of closed-loop modulation is presented in Fig. 1.2. The sensing circuit records the brain activity. Recorded information is then sent to a signal-processing algorithm for brain-state detection. The recording and processing enables an adaptive adjustment of the stimulation circuit parameters, such as location, interval and duration. Therefore, the whole system closes the loop with the human brain. With an appropriate algorithm, closed-loop

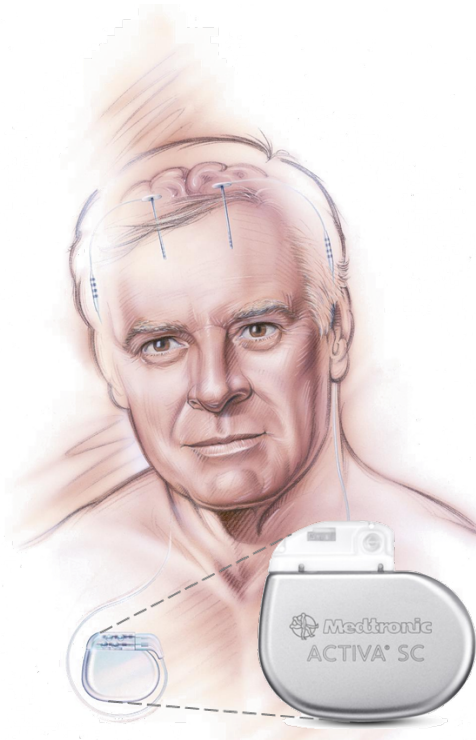


Figure 1.1: Traditional DBS system (Medtronic Activa SC system shown).

modulation can automatically turn the stimulation on or off, and thus minimize the side effects and reduce power consumption of the stimulation circuit [18]. Coping with a large channel count, the closed-loop neuromodulation can detect the habituation effect and change the stimulation site to an optimal location. Moreover, the implant implementation could eliminate the constraint from tethered cables of current wall-mounted systems, and make the treatment more accessible to patients.

In a multi-channel, closed-loop implant neuromodulation system, one essential component is the neural-sensing circuit, or the sensing front-end. Although neural-sensing circuits have been designed ever since investigations into brain signals began[19], the new system concept imposes two additional challenges. Firstly, the major focus of the sensing front-end so far regards how to suppress circuit noise for detecting weak neural signals only. However, the closed-loop neuromodulation would require concurrent sensing with stimulation, resulting in a much higher input range. Secondly, recording of only a few channels, as in traditional systems, allows the off-chip passives to ease the device/brain interface design. However, the

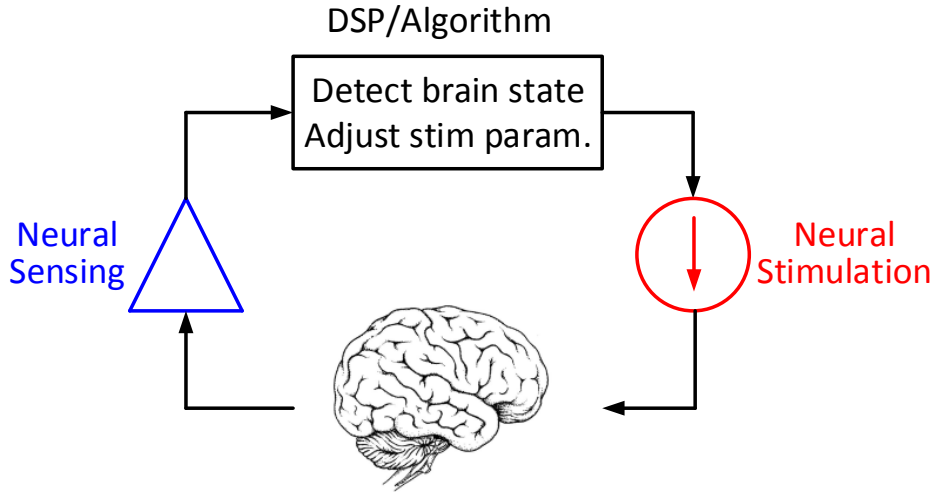


Figure 1.2: Diagram showing the closed-loop neuromodulation concept.

large channel count at the implant scale, as required in the future system, indicates that the circuit needs to solve the interface issue with a minimum number of external passives, and must meet tight power constraints.

An elaboration of the input-range requirement under concurrent sensing and stimulation is shown in Fig. 1.3. The brain neural signals comprise local field potential (LFP) signals ($\sim 1 - 250 \text{ Hz}$) and action potential (AP) signals (to $\sim 5 \text{ kHz}$). Although AP signals are instructive for some neuroscience research, medical researchers are focusing more on recording LFP signals for closed-loop operations in long-term treatment. The LFP signal level is approximately 10s of μV to several mV , while the biologic noise limits the detection sensitivity to approximately several μV . Therefore, prior sensing front-end has an input-range requirement of less than 10 mV , and a dynamic range of approximately 10 bits. However, the stimulation induces a large voltage disturbance in the target brain region, i.e. stimulation artifact, up to $\pm 50 \text{ mV}$. To guarantee real-time feedback for closing the loop, the front-end should record stimulation artifacts with neural signals, allowing the following algorithm to cancel artifacts and retain the neural signals. Such requirement increases the front-end input range by 10-fold and the dynamic range to 13 bits.

This dissertation presents a sensing front-end circuit and system design that aims to

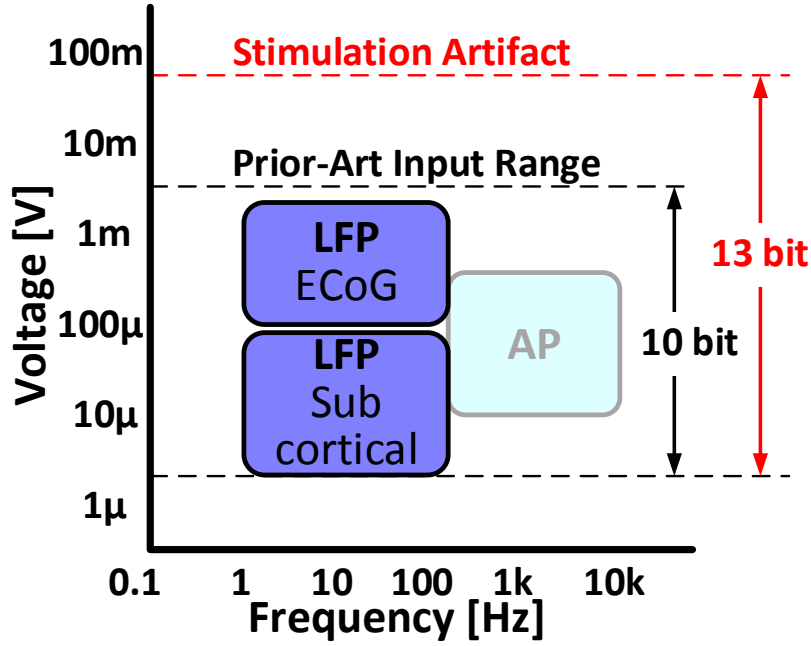


Figure 1.3: Frequency and amplitude of neural signals.

enable the entire multi-channel, closed-loop neural implant system. Thus, this constitutes a pioneering work with a high input range and an implant-scale power consumption, as well as noise performance that is on par with state-of-the-art designs. This design also bridges the gap between prior implants and wall-mounted systems, as presented in Fig 1.4. It also includes the integration of the front-end into a 32/64-channel sensing chip that is assembled on an implant-scale board for concurrent-sensing-and-stimulation tests.

The work presented in this dissertation offers four main innovations:

1. A fully-functional VCO-based neural-sensing front-end is designed. It transcends the gain-range trade-off in classical neural-recording circuit architecture and provides 10 times the input range compared to prior designs.
2. A multi-rate duty-cycled-resistor based high-pass filter is employed inside of the front-end to provide a reliable input interface that rejects the input DC offset and delivers very high DC input impedance.
3. A new quantizer across the analog-digital asynchronous interface is developed for the

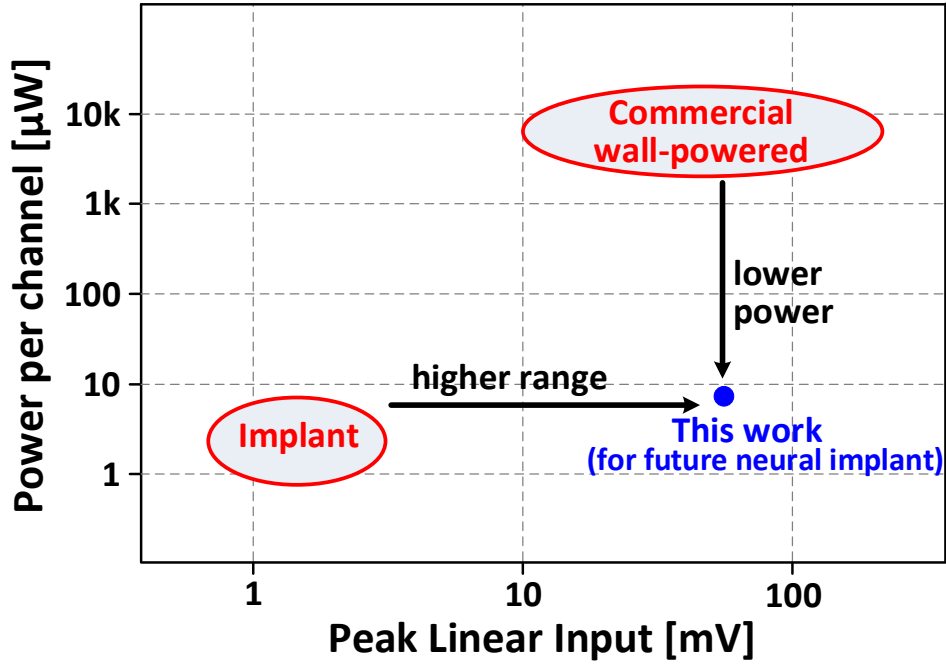


Figure 1.4: This work bridges the gap between prior implants and wall-powered systems.

VCO-based front-end. This quantizer guarantees glitch-free recording of the input signal.

4. A multi-channel neural-sensing chip is implemented with the proposed front-end as the core circuit. This chip enables the first concurrent-stimulation-and-sensing operation with implant-scale electronics, as demonstrated in the *in-vitro* test.

1.2 Dissertation Outline

Chapter 2 describes details of the sensing front-end design. The quantified specifications for the front-end are presented with a literature review on prior state-of-the-art systems. In addition, the advantage of the voltage-controlled oscillator (VCO) based structure is discussed, followed by an elaboration of the circuit techniques.

Chapter 3 shows the circuit measurement set-up and measurement results of the front-end.

Chapter 4 explains a multi-channel sensing chip design based on this front-end, and further integration of the chip into a neuromodulation (NM) unit at an implant scale. The concurrent-stimulation-and-sensing measurement set-up and results are also included for concept verification.

Chapter 5 concludes this work and discusses further research directions.

CHAPTER 2

Neural-Sensing Front-End Design

2.1 Specifications of the Front-End Design

In this section, we discuss the various design requirements that must be satisfied for a viable closed-loop neural recording front-end. Although existing literature has summarized some of the specifications [2][20], they are not tailored for front-ends in a closed-loop neuromodulation system, and thus need to be updated.

As mentioned in section 1.1, the front-end should be sufficiently sensitive to detect weak neural signals. This sensitivity sets the noise requirement of the front-end. Since the neuroscience community is still exploring better approaches to extract information from recorded brain signals, a definite value for the noise requirement has not yet been determined. However, two observations are helpful in estimating an approximate range for the required RMS noise. Firstly, the lower bound on the noise of the complete sensing system is inherently set by biological noise and electrode-electrolyte interface noise. Despite certain variance, these noise sources are reported to be at the level of $5 - 10 \mu V$ in the literature [21][22]. Design for a RMS noise much lower than this noise level does not improve performance, however it takes excessive power. Secondly, laboratory experiments show that LFP signals have an oscillation of $10 \mu V$ to a few mV [2][21][22]. Since excessive circuit noise is detrimental to signal quality, an RMS noise of $3 \mu V$ is set to provide sufficient SNR for brain signal processing.

The input range should be linear up to $\pm 50 mV$ as mentioned in section 1.1, and this range is determined by the amplitude of stimulation artifacts. Traditionally, single-ended brain stimulation scheme is performed, i.e. the current is injected from a current source and

returns to a common reference. In this scheme, the current has to pass through the reference electrode and the cable to complete the return path. Since the current can be up to a mA level and the return path can have $k\Omega$ resistance, single-ended stimulation triggers a large disturbance in the brain tissue, i.e. stimulation artifact, which can be up to several volts. A more optimal stimulation scheme is implemented in our system, which suppresses artifacts with better current-steering capability [11]. It adopts a differential stimulation scheme, in which two electrodes are connected to current sources with a current of the same magnitude, but opposite polarity. These two current sources serve as a current source and sink. If they are ideally matched, there is no current flowing through the reference electrode, and thus the CSF voltage disturbance is kept to a minimum. With this stimulation scheme, the sensing input range is defined by the voltage drop inside of the brain, caused by the stimulation current flowing through the equivalent path-resistance of the CSF. This CSF path resistance is much smaller than the electrode and cable resistance, and the stimulation artifacts are much smaller, usually confined within $\pm 50 mV$. With the targeted noise performance, this input range translates to approximately 80 dB in dynamic range or 13 bits in ENOB. A front-end with this linear input range can capture stimulation artifacts and neural signals with distortion below the noise level, so that subsequent artifact-rejection algorithms can remove the artifact and retain neural signals for closed-loop operation.

The circuit should consume low enough power. There are two main reasons for this constraint:

- To remove tethered constraints on patients and provide constant recording for patient care, the implant needs to be recharged wirelessly. The recharge frequency should be approximately once every few days or weeks, to minimize inconvenience due to frequent recharging. This, in turn, requires low power consumption for the implant system.
- The medical-device standard (ISO 14708 [23]) has a strict requirement on the temperature change at the surface of the implant, no more than a few degrees, because overheating can be a bio-hazard for brain tissue. This temperature-change requirement also sets the limit on implant system power consumption. Studies in [24] and

[25] suggest that power dissipation of 10 mW for a $6 \times 6 \times 2\text{mm}^3$ implant is low enough to avoid tissue damage due to thermal effect. Assuming 50% power for stimulation and 50% for power conversion efficiency, and another 50% design margins, the power consumption per channel should be less than $10 \mu\text{W}$ for a neural-recording front-end of 50 to 100 recording channels.

The implant interface has two additional constraints: input DC offset rejection and high DC input impedance. The input DC offset derives from the electrochemical effects at the electrode-electrolyte interface. The sensing front-ends connect to electrodes, which interface with the CSF (electrolyte). Theoretically, if the electrodes were perfectly matched and the electrolyte was homogeneous, the potential difference between these electrodes would be zero. Nevertheless, this assumption is not true in real scenarios. Electrodes have mismatches on material properties, size, and the microstructure; and the CSF is never homogeneous. Therefore, electrodes have different offset voltages, similar to the effect inside of a battery. These voltage differences result in an input DC offset to the front-end. This DC offset can vary between $\pm 50 \text{ mV}$, as reported in literature [2]. Such a large offset (relative to the neural signal amplitude) means that the DC offset can eat into the available input range and reduce the acceptable signal level for the front-end. Therefore, it is desirable to reject this input DC offset. Considering that LFP signals go down to 1 Hz , the high-pass corner for offset rejection should be kept below 1 Hz to avoid attenuation of the signals of interest. The high DC input impedance is a subsequent requirement for safety under a large DC offset. A finite DC input impedance leads to an input DC current (in the presence of a DC offset) that flows through the electrode-electrolyte interface. A large DC current could lead to electrode corrosion and is unsafe for a brain implant. Although the safety threshold for this DC current varies widely in the literature, we choose a conservative figure of 100 pA for our design. This value with a $\pm 50 \text{ mV}$ offset implies that the DC input impedance should be higher than $1 \text{ G}\Omega$.

The specifications of the front-end are summarized in Table 2.1.

Table 2.1: Front-end specifications.

RMS noise (1 – 250 Hz)	$< 3 \mu V$
Linear input range	$\pm 50 mV$
Power consumption	$< 10 \mu W/ch$
DC offset rejection corner	$< 1 Hz$
Input DC impedance	$> 1 G\Omega$

2.2 Review of Prior Art

The previous neural-sensing front-ends do not meet all of the specifications presented in Table 2.1, in particular, the linear input range and the interface requirements, including DC offset rejection and high DC input impedance. An inspection of some typical and widely-adopted designs can help summarize issues in these designs.

The traditional architecture of the neural-sensing front-end comprises an instrumentation amplifier, an optional filter, and an ADC. The most essential part in this architecture is the amplifier, since it determines the noise performance, the input range, and the input interface. A pioneering design of the front-end amplifier by Harrison is presented in [1], with the schematic shown in Fig 2.1. An AC-coupled input presents an infinite DC input impedance and limits the input DC current to zero. The capacitance ratio of C_1/C_2 determines the in-band gain (40 dB), and DC offset rejection is achieved by pseudo resistors formed by $M_a - M_b$, and $M_c - M_d$. While this circuit works adequately with a small input, a large input can saturate the amplifier and lead to a clipped output due to the voltage-rail limit. Even if clipping is avoided, the voltage swing across the pseudo resistor formed by $M_a - M_b$ goes beyond 0.2 V with an input amplitude of $> 2 mV$, leading to all the issues discussed further in section 2.4. This design has a THD of 1% for an input of 16.7 mV_{pp} , and clearly does not meet our specifications.

Another classical design by Denison [2] employs a chopper-stabilized instrumentation amplifier, as presented in Fig 2.2. Chopping is used to up-modulate the input voltage signal to the vicinity of the chopping frequency, and separate it from the flicker noise contributed by

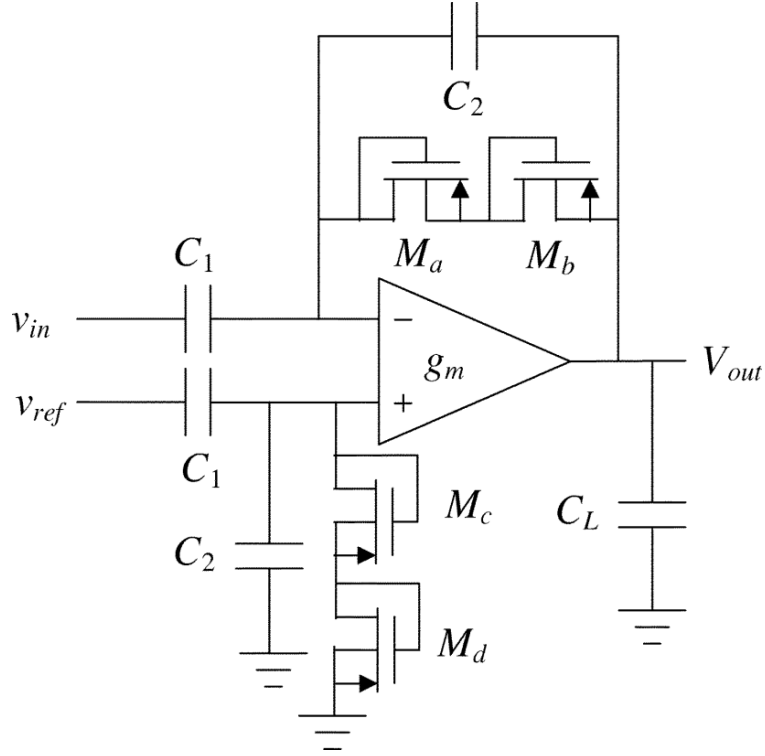


Figure 2.1: Neuro amplifier schematic by Harrison [1].

the OpAmp [26]. The amplifier design entails two feedback loops: the 1st feedback path from C_{fb} to C_i , which determines the in-band gain; and the 2nd feedback path from C_{hp} to C_i , for input DC offset rejection. The switched-capacitor integrator before C_{hp} forces the amplifier output DC level to zero. However, this front-end amplifier has a limited input range, up to 5 mV. In addition, this design has a low DC input impedance due to chopping. Given an input DC offset, chopping leads to charging / discharging cycles on C_i at the chopping frequency, and this charging / discharging of C_i draws a DC current from the input V_{in} . This effect is similar to the switched-capacitor resistor, and the resulting DC input impedance is:

$$R_{in,DC} = \frac{1}{4f_{chop}C_i} \quad (2.1)$$

From equation 2.1, it is easily found that $R_{in,DC}$ would be limited to approximately 10 M Ω with f_{chop} of a few kHz and C_i of 10 pF.

Several other designs have attempted to avoid saturation or hard clipping of the front-end with large stimulation artifacts. One way is to reset the input to avoid front-end saturation,

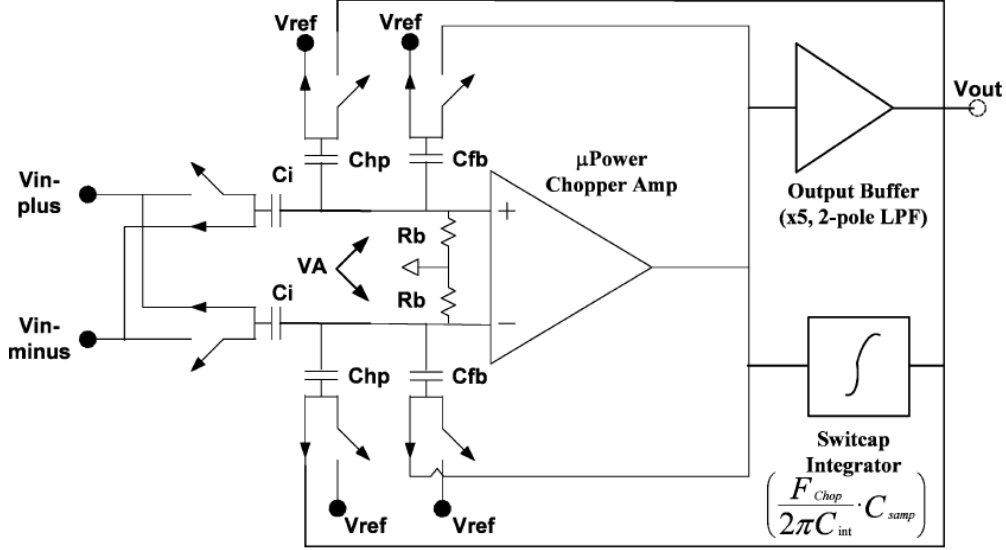


Figure 2.2: Chopper-stabilized instrumentation amplifier [2].

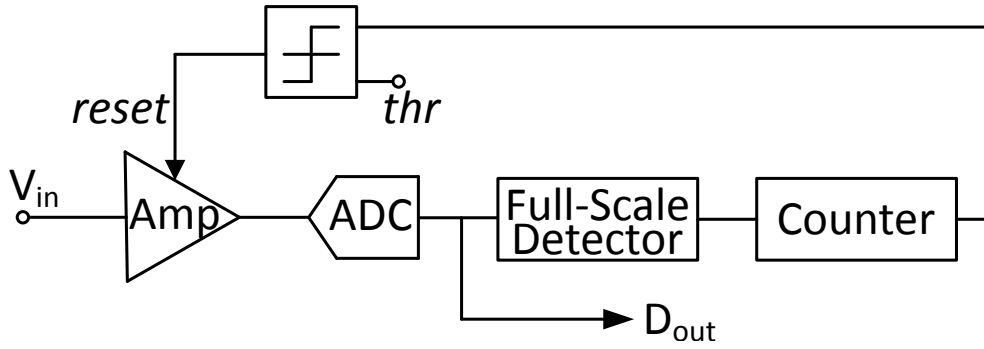


Figure 2.3: HermesE system: reset upon detecting large signals [3].

as implemented in the HermesE system [3]. Fig 2.3 shows a simplified schematic, in which the reset signal is triggered when the count of full-scale ADC outputs is beyond the threshold. Although this reset avoids saturation, it also blanks the front-end during stimulation events. Thus, neural signal information is lost when the amplifier is under reset, preventing timely feedbacks for closed-loop neuromodulation.

Motivated by the wireless transceiver principle, another design employs the frequency-division multiplexing (FDM) concept to filter out stimulation artifacts [4] to prevent saturation. The design selects a specific stimulation frequency, so that artifacts are kept away from the bio-marker in the frequency domain. Therefore, a frequency-selective filter can be

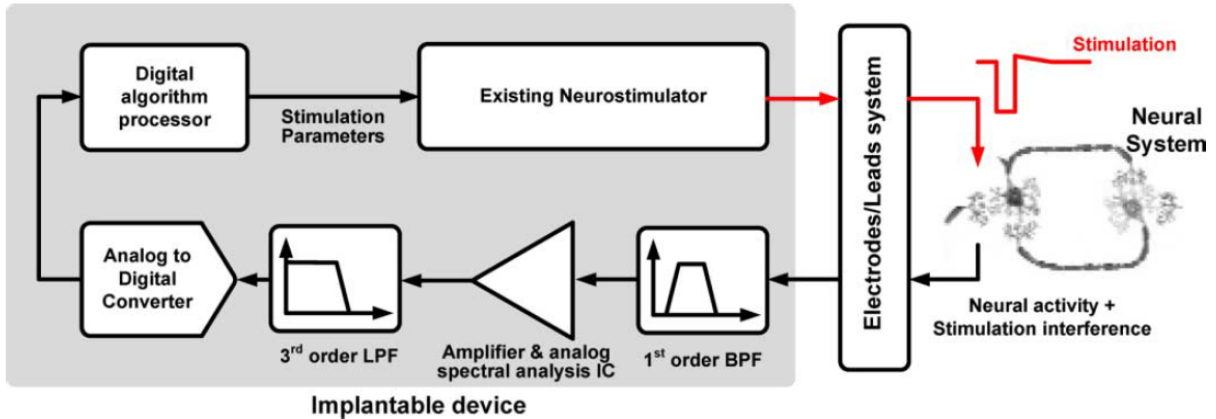


Figure 2.4: Block diagram for FDM-based neuromodulation system [4].

placed at the input of sensing circuit to remove artifacts, such as the band-pass filter shown in Fig 2.4. Nevertheless, this FDM concept only works well when we are concerned with bio-markers in a specific range of frequencies. It becomes quite challenging, or even impossible, to select a stimulation frequency and monitor bio-markers in two or more frequency ranges.

A recent published design bypasses the amplifier stage and directly digitizes the input with an oversampled data converter [5], as shown in Fig 2.5. The converter is a 1st order $\Sigma\Delta$ modulator, and the integrator is formed by the transconductance stage, the OpAmp and the capacitors C_{int} . To reduce the oversampling ratio, a 5-bit SAR ADC is used as the quantizer inside of the modulator, along with a 5-bit capacitance DAC (CDAC) bank in the feedback path. This modulator also chops the input signal to suppress the contribution of flicker noise from the circuit components. Although scaling-up transistor area can reduce flicker noise as well, it is not preferred because the large parasitic capacitance at the input requires higher driving strength in the feedback path, and thus leads to extra power consumption. Ideally, with noise-shaping in the modulator and chopping of the input, this design should achieve high resolution.

Two major obstacles prevent the application of this design in closed-loop neuromodulation systems. Firstly, the capacitors inside of the CDAC bank are not perfectly matched, and this mismatch limits the linearity of the converter. The front-end only achieves 10.2-bit

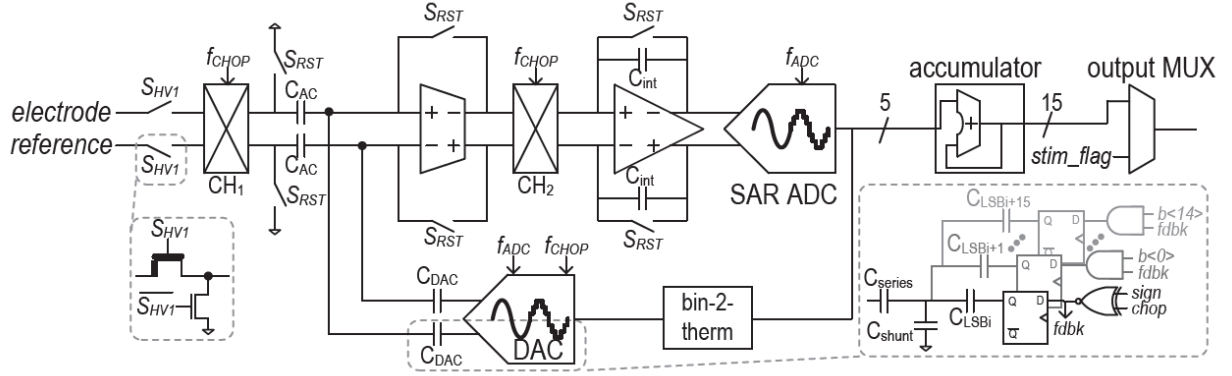


Figure 2.5: Direct digitization at the input with 1st order $\Sigma\Delta$ modulator [5].

ENOB, which is insufficient for our specification. Secondly, as discussed above, chopping at the input results in a switched-capacitor resistance that limits the DC input impedance. Therefore, the input interface is not acceptable in terms of patient safety.

In summary, most prior-art are intended for a sensing-only scenario, in which the stimulation artifact is absent and a small input range of $< 10\text{ mV}$ is sufficient. They are designed with a high-gain amplifier (40-80 dB), and hence not appropriate for the concurrent stimulation and sensing. Moreover, their input interfaces cannot meet all requirements: either the DC input impedance is low; or it does not reliably reject the DC offset without attenuating the LFP signal. Other methods try to bypass the challenge of achieving a large input range, yet remain unsuitable for a closed-loop neuromodulation implant system.

2.3 VCO-Based Structure

Recently, VCO-based signal-processing has attracted substantial attention of circuit designers due to its digitally-intensive approach. A VCO converts an input voltage into the oscillator frequency with a conversion gain of K_{VCO} . It is traditionally used as a clock reference, and has been extensively studied in phased-lock loop (PLL), where the VCO phase Φ_{VCO} is detected and fed back for control. The VCO phase Φ_{VCO} is the integral of the frequency f_{VCO} : $\Phi_{VCO} = \int 2\pi f_{VCO} dt$. This inherent integration property of the VCO has been exploited in the VCO-based ADC architecture [6][27]. As shown in Fig 2.6, these ADCs adopt the VCO

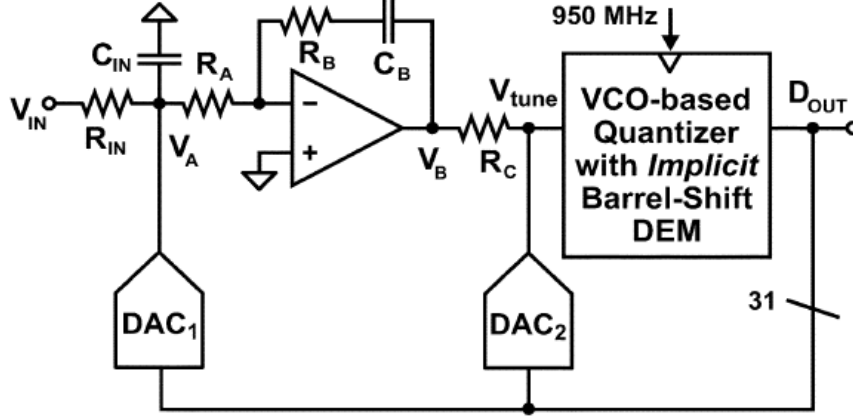


Figure 2.6: VCO-based ADC architecture [6].

as the quantizer inside of a continuous-time $\Sigma\Delta$ modulator loop, and also take advantage of the implicit dynamic-element matching inside of the VCO. By using a ring oscillator as the VCO, this structure will also benefit from technology scaling. However, although these ADCs achieve 13-bit ENOB and ADC figure-of-merit (FOM) that are comparable to the traditional pure-analog architectures, they cannot be used directly as a neural-sensing front-end. Firstly, it still requires a gain stage to amplify the input to full-scale for power-efficient conversion. Secondly, the electrode interface in the neural-sensing system cannot directly drive the resistive input-impedance of this architecture, requiring an additional stage for isolation.

In light of the above-mentioned problems, a different architecture with the VCO has been proposed as the neural-sensing front-end [8][28]. The VCO is placed directly at the input (without any voltage-domain pre-amplification) to convert the sensed voltage into the frequency/phase domain. This architecture offers the following benefits:

- The VCO-based front-end can allow a large input range for recording both stimulation artifacts and neural signals. Unlike voltage-domain processing, this architecture has no voltage amplification, and hence the output is not limited by a voltage rail. With appropriate design, the VCO will not saturate within the specified range, allowing the capture of both stimulation artifacts and neural signals. Therefore, a suitable

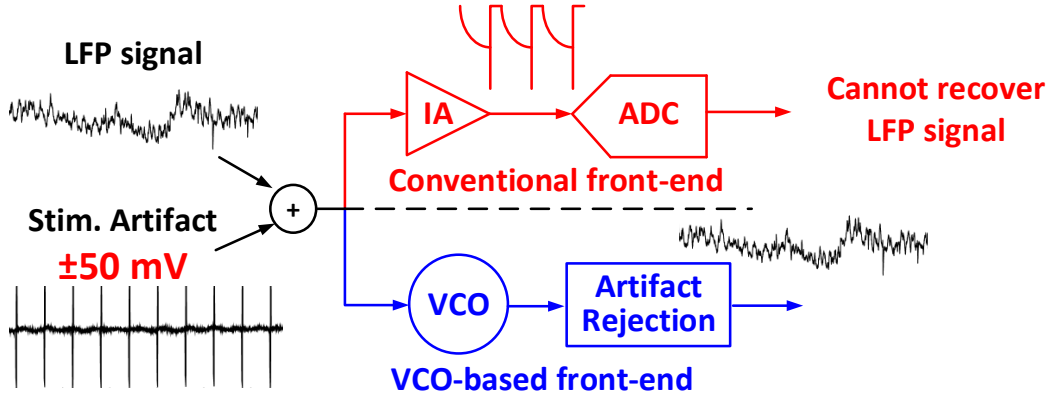


Figure 2.7: VCO-based front-end does not saturate and allows signal recovery [7].

algorithm downstream can suppress artifacts while retaining neural signals for closed-loop neuromodulation [29]. In contrast, the recovery of neural signals is not possible in a traditional front-end once a large input saturates the front-end. This comparison is shown in Fig 2.7.

- The VCO front-end can deliver high resolution with modern CMOS technology. This is mainly because of reduced gate delay in advanced technology nodes ($\sim 10 - 20ps$ with minimum length), resulting in a high K_{VCO} . Even at a μA -level power-consumption constraint, the VCO still provides enough K_{VCO} for the required recording resolution. In our front-end, the voltage-to-frequency gain K_{VCO} is $70 MHz/V$ and one VCO cycles contains 58 phases, which will be explained in more detail in section 2.6. The phase sampling window within one period is approximately $100 \mu s$. An input range of $\pm 50 mV$ corresponds to:

$$70 MHz/V \times 58 \times 100 \mu s \times 100 mV = 40600 = 2^{15.3} bits$$

Therefore, the provided resolution is 2 bits higher than the required ENOB., i.e. the quantizer noise does not limit front-end noise performance.

- Although the VCO frequency is subject to supply voltage or temperature variations, this dependency is not critical since the bio-implant system can minimize supply/temperature

variations. For the supply voltage, a well-regulated supply can be available in the system to limit the VCO frequency change with the supply voltage. Also, the normal human body temperature is stable, within a range of 35-38°C[30]. The temperature variation inside of the implant is also within a few degrees Celsius because of low power consumption in the implant and the regulation standard [23]. Therefore, the VCO frequency-drift due to temperature is also limited.

While the VCO-based structure is promising in achieving the high-input-range recording requirement, three challenges need to be addressed:

- The interface requirements, i.e. the DC input impedance and the DC offset rejection, need to be satisfied by appropriate choice of circuit topology and design techniques. For example, since $Z_{in,DC}$ needs to be higher than 1 G Ω , we cannot use the input directly as the VCO supply.
- The circuit noise needs to be kept low to meet the the noise specification in Table 2.1. Since LFP signals are at low frequencies, flicker noise can be the dominant noise contributor in the signal band, and the front-end design needs to address it.
- The VCO nonlinearity has to be corrected to ensure high linearity over the specified input range. The nonlinearity of the front-end can degrade signal quality despite the capability of capturing both stimulation artifacts and neural signals. During the occurrence of stimulation artifacts, neural signals may experience time-varying gains, which leads to distortion. In the frequency-domain perspective, nonlinearity can lead to distortion terms, such as harmonics, intermodulation and cross-modulation. All of these effects can spread the spectrum and contaminate weak neural signals. Therefore, nonlinearity correction/calibration is requisite.

A previous design of the VCO-based front-end [8], as presented in Fig 2.8, is unable to meet these three challenges. Firstly, although it adopted a pseudo-resistor structure as the input interface to reject the DC offset, this structure does not provide a reliable offset-rejection corner and will be discussed further in section 2.4. Secondly, the VCO phase sampling and

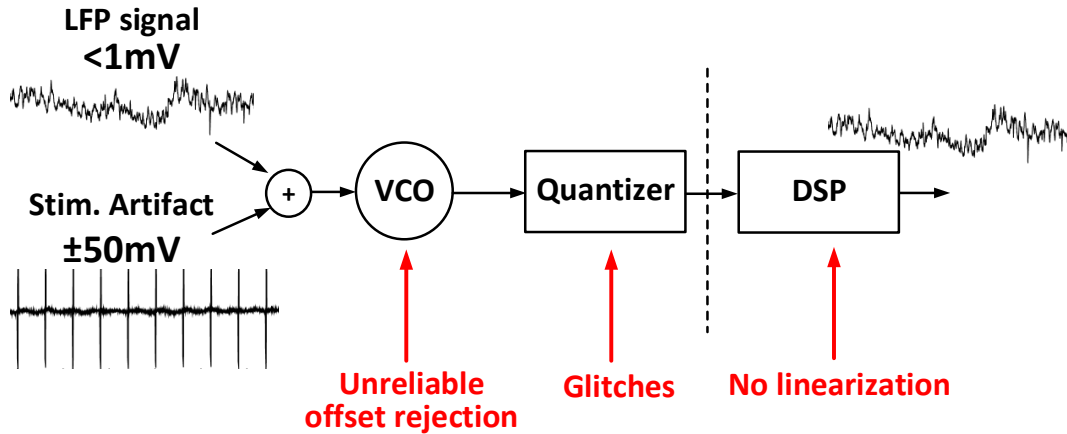


Figure 2.8: Previous VCO-based front-end design [8] does not address existing challenges.

digitization block, referred to as the “quantizer” in the figure, does not account for the asynchronous analog-digital interface. Therefore, the output exhibits excessive glitches, at the level of a few mV, which severely degrades circuit noise performance. Moreover, these glitches can render the downstream algorithm(s) inoperable. Thirdly, no hardware for non-linearity correction is implemented. Neural signals are thus distorted in the presence of stimulation artifacts and cannot be fully recovered. In summary, the design is not functional for the implant closed-loop neuromodulation system.

The proposed VCO-based front-end is designed to solve all of these challenges. As shown in Fig 2.9, it consists of three blocks: VCO, quantizer, and nonlinearity correction (NLC) block. The VCO converts the electrode voltage into the oscillator frequency. It employs the multi-rate duty-cycled-resistor based high-pass filter (HPF) to achieve the interface requirement, and adopts design choices for suppressing flicker noise. These details are presented in section 2.4 and 2.5. The quantizer samples and quantizes the phase traversed by the VCO, while accounting for the asynchronous analog-digital interface discussed in section 2.6. The NLC block corrects for the VCO nonlinearity, with the principle explained in section 2.7.

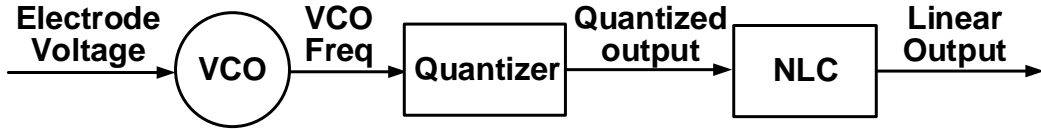


Figure 2.9: The implemented VCO-based front-end contains VCO, quantizer, and NLC block.

2.4 Multi-Rate Duty-Cycled-Resistor Based HPF

2.4.1 HPF Requirements and Prior Art

As stated in section 2.1, the front-end needs to suppress the input DC offset to avoid saturation. The front-end should also pass the input signal unattenuated, or with minimal attenuation, to prevent the degradation of sensitivity. This essentially requires a HPF to reject the offset with a corner frequency of $< 1 \text{ Hz}$. Besides this requirement, the HPF also needs to conform to other specifications of the front-end: it should neither introduce noise over the RMS noise bound set in Table 2.1, nor degrade the DC input impedance of the front-end.

A straightforward solution is to utilize an R-C based HPF. The AC coupling interface has infinite DC input impedance, thus preventing the flow of input DC current. Realizing an off-chip filter with sub-Hz corner is not difficult, since it only requires a resistance of $\sim 10 \text{ k}\Omega$ and a capacitance of $\sim 100 \mu\text{F}$. This has been realized in microphones, in which there are only a few sensing channels and the system has sufficient area on the PCB to accommodate these passives. However, such a solution is not affordable in a multi-channel implant, in which board area is limited. This necessitates an on-chip solution.

Achieving a sub-Hz corner by an R-C implementation on chip is infeasible considering the area cost. Although a capacitance of 100 pF is feasible, this requires a resistor of $> 1 \text{ G}\Omega$, which is an astronomical value for an on-chip passive resistor, especially when it has to be replicated for a multi-channel system. Thus, we need a different way of achieving the low frequency corner.

Another constraint on the resistance value derives from the noise specification. The integrated noise across the entire spectrum is equal to $\frac{kT}{C}$, independent of the resistance value. Hence, a capacitance value of at least 1 nF is needed to ensure that the noise contribution is less than $2 \mu V_{RMS}$. However, noise from the resistor is low-pass filtered with the same corner frequency as that of the high-pass filter. Consequently, a larger resistance would lead to less in-band noise and reduce the required capacitor value. This is graphically shown in Fig 2.10, where f_{Lo} and f_{Hi} annotate the frequency bound of the sensed signal band. The integrated noise power of the HPF with a larger resistance (area enclosed by two black dashed lines at f_{Lo} and f_{Hi} , and the blue line) is smaller than the integrated noise power with a smaller resistance (area enclosed by the same two dashed lines and the red line).

The integrated noise is calculated as:

$$\begin{aligned} \langle V_{n,HPF}^2 \rangle &= \int_{f_{Lo}}^{f_{Hi}} \frac{4kTR}{1 + (2\pi fRC)^2} df \approx \int_{f_{Lo}}^{f_{Hi}} \frac{4kTR}{(2\pi fRC)^2} df = \int_{f_{Lo}}^{f_{Hi}} \frac{4kT}{(2\pi fC)^2} \frac{1}{R} df \\ &= \frac{2kT}{\pi C} f_{HPF} \left(\frac{1}{f_{Lo}} - \frac{1}{f_{Hi}} \right) \approx \frac{2kT}{\pi C} f_{HPF} \frac{1}{f_{Lo}} \end{aligned} \quad (2.2)$$

It is clear that, in equation 2.2, the noise contribution from HPF is inversely proportional to the resistance value, or equivalently proportional to the corner frequency. With 100 pF capacitance, a $10 \text{ G}\Omega$ resistance is needed to limit the in-band noise to $2 \mu V_{RMS}$. Therefore, the noise specification sets a higher resistance requirement.

Some solutions have been proposed to realize such a large resistance. A typical way is to use a pseudo resistor, which is a stack of diode-connected transistors, with the gate shorted to the drain and the source connected to the bulk, as shown in Fig 2.11 and also embedded in Fig 2.1. The transistors are typically chosen as PMOS so that the bulk of the transistor can be shorted to the source. Assuming that there is no leakage to the substrate, the transistors will equally divide the voltage between the point A and B. The current in the subthreshold region is:

$$I = I_S \exp\left(\frac{V_P}{V_T}\right) \left[\exp\left(\frac{-V_S}{V_T}\right) - \exp\left(\frac{-V_D}{V_T}\right) \right] \quad (2.3)$$

where V_P is the pinch-off voltage, and I_S is the specific current [31]. Assuming $V_D = V_S$ for

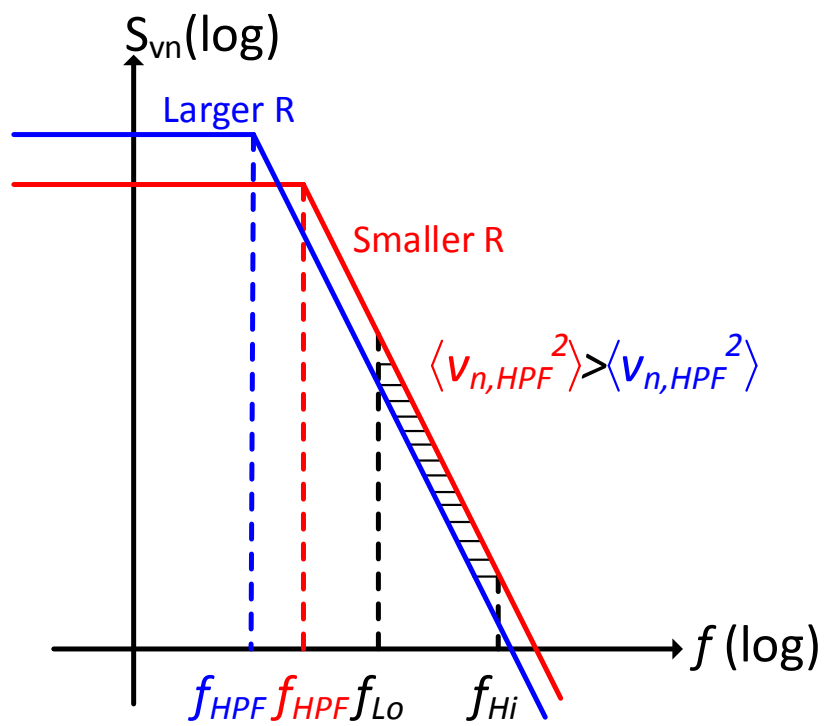


Figure 2.10: Output noise PSD of the HPF.

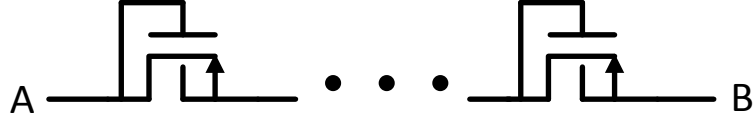


Figure 2.11: Pseudo resistor schematic.

the static case, the small-signal equivalent resistance of one diode-connected transistor is

$$\begin{aligned} \frac{dI}{dV_D} &= I_S \exp\left(\frac{V_P}{V_T}\right) \left[\exp\left(\frac{-V_S}{V_T}\right) - \exp\left(\frac{-V_D}{V_T}\right) \right] \frac{dV_P}{dV_D} + I_S \exp\left(\frac{V_P}{V_T}\right) \exp\left(\frac{-V_D}{V_T}\right) \frac{1}{V_T} \\ &= I_S \exp\left(\frac{V_P}{V_T}\right) \exp\left(\frac{-V_D}{V_T}\right) \frac{1}{V_T} \end{aligned} \quad (2.4)$$

A further simplified equation is shown below if we set $V_D = V_S = V_B = 0$, or equivalently, refer V_G or V_P with respect to V_B :

$$r_{ds} = \frac{V_T}{I_S} \exp\left(\frac{-V_P}{V_T}\right) = \frac{V_T}{I_S} \exp\left(\frac{V_{to} - V_G}{nV_T}\right) = \frac{V_T}{I_S} \exp\left(\frac{V_{to}}{nV_T}\right) \quad (2.5)$$

It is clear that the equivalent resistance of one diode-connected transistor is defined by the transistor process and basic semiconductor physics. The pseudo-resistor equivalent resistance is N times this value, where N is the number of stacked transistors. This large resistance can be attributed to the effort on reducing leakage power consumption in digital logic for CMOS technology. The pseudo resistor is adopted by many published works in the literature with a resistance larger than $100 \text{ G}\Omega$ [1][32][8].

Nevertheless, the pseudo resistor possesses some drawbacks that render it unreliable in practice:

- It can be seen in equation 2.5 that the resistance has a strong dependence on the CMOS technology process, as it is a function of I_S and V_{to} . Moreover, the resistance is also a strong function of V_T , thus exhibiting a high sensitivity to temperature. While the temperature is quite stable inside of the implant, the resistance value is still difficult to predict and can vary by $100\times$ across the corners according to simulations.
- The previous derivation assumes that transistors are the only available path for the flow of current. However, the bulk of the PMOS transistor is a N-doped well (NW). The

NW forms a parasitic reverse-biased diode with the substrate. The current through this reverse-biased diode is small, yet sufficient to cause a large voltage offset across the pseudo resistor given the large equivalent resistance. Assuming an equivalent resistance of $100\text{ G}\Omega$, a tiny current of 100 fA can lead to 10 mV offset, which can severely degrade the common-mode rejection ratio (CMRR) of the front-end. As reported in [8], the measured CMRR is only 50 dB and this low CMRR value is believed to be caused by the offset due to the well-leakage current.

- Apart from CMRR degradation, the pseudo resistor can also lead to significant distortion in the circuit. Equation 2.3 indicates that the current is the difference of two exponential components. Moreover, while the source and the drain can be treated symmetrically in the MOSFET, the gate and the bulk are not interchangeable. Therefore, the current versus voltage characteristic is not only nonlinear, but also asymmetric. Under a large input swing, the pseudo-resistor-based HPF can lead to output DC-level drift, i.e. the rectification of the input signal. This is difficult to correct, especially given the unpredictable pseudo-resistor characteristic and the long time-constant of the HPF.

Other solutions for a large equivalent resistance include linearization of the MOS transistor with the gate voltage controlled by the drain and source voltages [33], where additional control circuitry adds much more noise compared to an equivalent resistor; or an OTA-based Gm cell with input/output shorted together [34], which uses degeneration to improve linearity, but at a cost of noise performance degradation, with no consideration to the additional flicker noise contribution from the transistors.

Given the above challenges of building the high-pass filter with linear time-invariant (LTI) circuits, some designs have explored the possibility of applying periodic-switching circuits. The low-frequency resistance in a switched-capacitor resistor is defined as $\frac{1}{f_{sw}C_{sw}}$. Therefore, with an AC-coupling capacitance C_{HPF} , the corner is set as

$$f_{HPF} = \frac{1}{2\pi C_{HPF} \frac{1}{f_{sw}C_{sw}}} = \frac{1}{2\pi} \frac{f_{sw}C_{sw}}{C_{HPF}} \quad (2.6)$$

Although f_{sw} is subject to the Nyquist-rate limit for input signals, a very small value of $\frac{C_{sw}}{C_{HPF}}$ can establish a low enough frequency corner. More complicated switched-capacitor structures can realize the desired corner with a lower capacitance ratio, or equivalently, better area efficiency [35]. However, due to the noise aliasing effect of the switched-capacitor circuit, the filter noise contribution is white instead of low-pass filtered, and hence subject to $\frac{kT}{C}$ limit. In [36], it is shown that the switch-capacitor implementation has elevated the input-referred integrated in-band noise from $0.6 \mu V_{rms}$ to $6.7 \mu V_{rms}$.

2.4.2 Duty-Cycled-Resistor (DCR) Based HPF

The concept of the duty-cycled resistor (DCR) was proposed and analyzed in the 1960s [37], and also reused in the loop filter design of a PLL [38]. It increases the equivalent resistance by duty-cycling the resistor, and thus realizes a large time constant. The adoption of the DCR in our HPF is shown in Fig 2.12. The resistor is connected to the output node on one side and to a switch on the other side, and the switch is connected to a suitable bias voltage. The switch is controlled by the signal $swclk$ and is switched on periodically with a duty-cycle ratio D . The periodical switching leads to an equivalent resistance of:

$$R_{eq} = \frac{R}{D} \quad (2.7)$$

and the corresponding corner of:

$$f_{HPF} = \frac{1}{2\pi} \frac{D}{RC} \quad (2.8)$$

An intuitive explanation of R_{eq} is shown in Fig 2.12 by applying a step input and inspecting the output transient, a method used in [37]. Similar to the output of a continuous-time HPF, the output will first instantaneously step up and eventually settle back to V_{bias} . When $swclk$ is high, i.e. the switch is ON, the output settles the same way as a continuous-time filter with R and C only; when $swclk$ is low and the switch is OFF, the output is held constant until the next switching cycle. Therefore, on average, the slope of the output response is D times smaller than the R-C filter and the time constant becomes $\frac{RC}{D}$, indicating a lower frequency corner for the HPF.

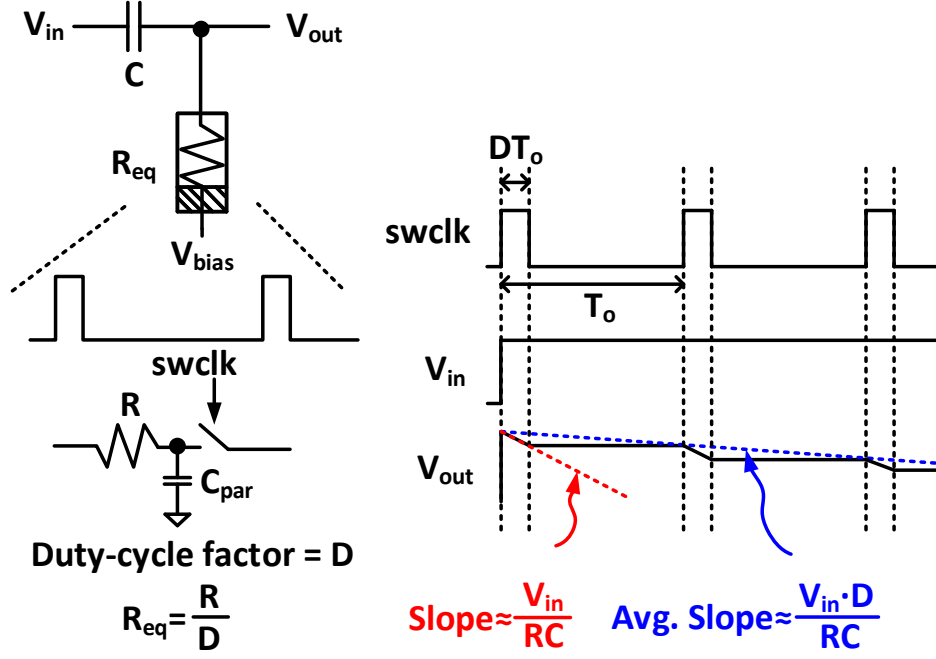


Figure 2.12: Schematic and intuitive explanation for duty-cycled-resistor based HPF.

A more rigorous analysis follows the approach in equations (1)-(3) in [39], where $G(t)$ represents the time-varying conductance of the duty-cycled resistor. From KCL, we have:

$$C \frac{dV_{in}(t)}{dt} = C \frac{dV_{out}(t)}{dt} + G(t)V_{out}(t) \quad (2.9)$$

and then

$$j\omega C V_{in}(j\omega) = j\omega C V_{out}(j\omega) + g_0 V_{out}(j\omega) + \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} g_k V_{out}(j(\omega - k\omega_o)) \quad (2.10)$$

where g_k represents the Fourier series of the periodic conductance $G(t)$ and specifically $g_0 = \frac{D}{R}$, and ω_o is the frequency of the *swclk* signal.

Some assumptions are needed for further simplification:

1. The input signal is band-limited and the input frequency is much lower than ω_o .
2. The time constant RC is much larger than the ON duration of the signal *swclk*, i.e. $RC \gg DT_o$.

These assumptions are also indicated in Table 1 of [37], where the error with a single-tone input is defined and analyzed in the time domain. In our frequency-domain analysis, these two assumptions guarantee that the input signal and the output signal are both at a low frequency. Consequently, we can define the following transfer function:

$$\frac{V_{out}(j\omega)}{V_{in}(j\omega)} = \frac{j\omega C}{j\omega C + g_0} = \frac{j\omega RC/D}{j\omega RC/D + 1} \quad (2.11)$$

which shows a corner frequency $f_{HPF} = \frac{D}{RC}$.

Since the noise from the HPF is contributed by the resistor, the output noise PSD is the same as the value derived in [38][39]:

$$S_{vn}^2 = \frac{4kT \frac{R}{D}}{|1 + j\omega C \frac{R}{D}|^2} \quad (2.12)$$

indicating a profile that is exactly the same as Fig 2.10.

The above two equations show that a DCR-based HPF performs like a continuous-time HPF with $R_{eq} = R/D$. Ideally, with $D < 1 \times 10^{-4}$, we can increase R_{eq} to be $> 10 \text{ G}\Omega$ with a $1 \text{ M}\Omega$ physical resistor. The *swclk* signal generation circuit for this low duty-cycle ratio is shown in Fig 2.13. Signal *A* has a frequency of $\sim 10 \text{ kHz}$ and is generated by digital logic. It is then inverted to generate signal *B*, with a delay tunable to a resolution of 1 ns . The AND operation of *A* and *B* generates the pulse signal *C*, with the same frequency as *A* ($\sim 10 \text{ kHz}$), and an ON-duration of a few nanoseconds, i.e. a duty cycle of less than 1×10^{-4} . Signal *C* is used as the *swclk* signal, and the switch is implemented as a transistor with an off-resistance much larger than R_{eq} .

Nevertheless, the achievable equivalent resistance is limited by the parasitic capacitance C_{par} , as presented in Fig 2.12. C_{par} includes the parasitic capacitance from the resistor and the switch. The output settling behavior with a step input is different from Fig 2.12 for a non-zero value of C_{par} . When the switch is ON, the output settling behavior is similar to an R-C HPF. However, when the switch is OFF, C_{par} is initially charged to V_{bias} and will redistribute charge with the filtering capacitor C , leading to a change of V_{out} . This effect is equivalent to a parasitic switched-capacitor resistance, with the value of $\frac{1}{f_{swclk} C_{par}}$ that

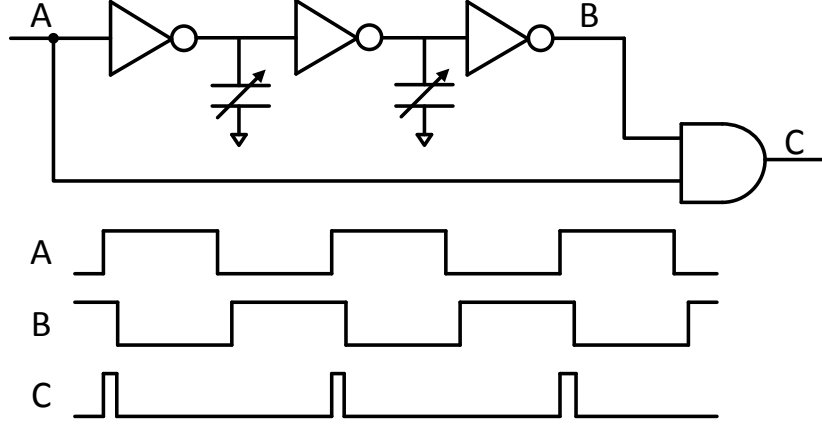


Figure 2.13: Generation circuit for the *swclk* signal.

appears in parallel with R_{eq} . It can be seen that with a small parasitic capacitance of 10 fF and switching clock frequency of 10 kHz , this resistance value reaches $10\text{ G}\Omega$, establishing a limit for the corner frequency of the HPF. Therefore, further improvement is needed to reduce the effect of C_{par} for larger R_{eq} , to lower in-band noise and reduce chip area.

2.4.3 Multi-Rate HPF Design

The multi-rate filter concept was utilized to overcome the limit of the parasitic switched-capacitor resistance in a DCR-based low-pass filter [40]. It is adapted in this HPF design and the schematic is shown in Fig 2.14. The components C_1 , R_1 and S_1 forms the DCR-based HPF as discussed above. For a band-limited input signal, $\frac{V_{out1}}{V_{in}}$ has a high-pass transfer function with a corner f_{HPF1} set to $\frac{1}{2\pi} \frac{D_1}{R_1 C_1}$ when ignoring the parasitic capacitance. This also means that $\frac{V_{in} - V_{out1}}{V_{in}}$ has a low-pass transfer function with the same corner f_{HPF1} . While f_{HPF1} is limited by the parasitic, it is not difficult to obtain a corner frequency of $\sim 10\text{ Hz}$. Thus, we can guarantee that $\frac{V_{in} - V_{out1}}{V_{in}}$ is suppressed by 40 dB at frequencies of a few hundred Hz. This suppression allows us to stack on C_1 a second DCR-based HPF stage with a clock f_2 (approximately 500 Hz), which is much lower than f_1 , and define a transfer function for $\frac{V_{out} - V_{out1}}{V_{in} - V_{out1}}$. This is a high-pass transfer function with the corner defined as

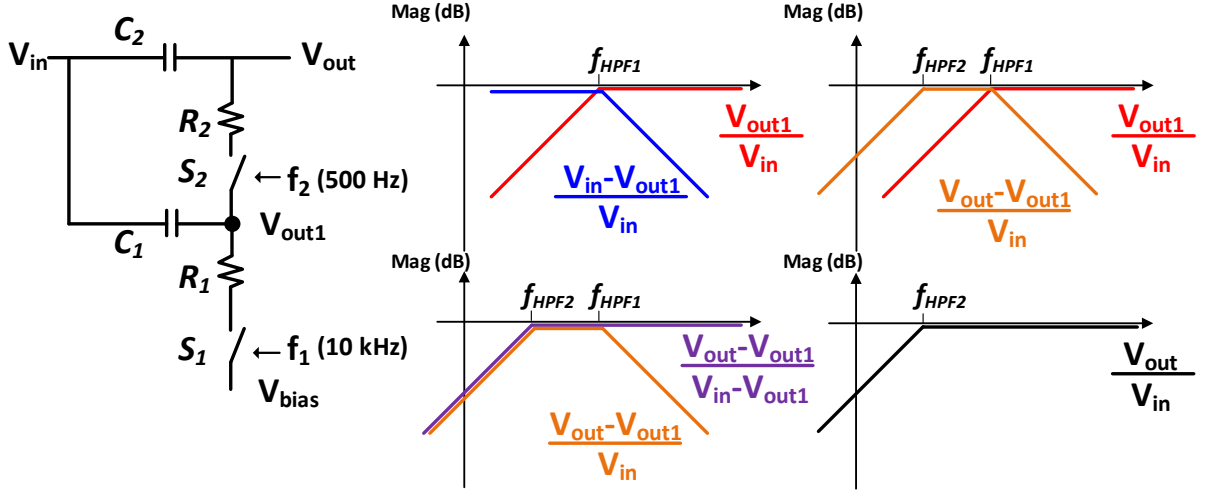


Figure 2.14: Multi-rate duty-cycled-resistor based HPF.

f_{HPF2} , ideally set to $\frac{1}{2\pi} \frac{D_2}{R_2 C_2}$. Since $f_2 \ll f_1$, the parasitic switched-capacitor resistance is much larger, and thus f_{HPF2} can be set lower than 0.1 Hz . The transfer function of $\frac{V_{out} - V_{out1}}{V_{in}}$ is band-pass with the low-pass corner at f_{HPF1} and the high-pass corner at f_{HPF2} . Therefore, combining $\frac{V_{out} - V_{out1}}{V_{in}}$ with $\frac{V_{out1}}{V_{in}}$, we have a high-pass transfer function for $\frac{V_{out}}{V_{in}}$ with the corner of f_{HPF2} without the issue of aliasing.

The implementation of this multi-rate DCR-based HPF requires twice the number of resistors and capacitors as compared to a single DCR-based HPF; however, C_1 and R_1 can be much smaller for area efficiency. The switch-control signal generation circuit is replicated, with the digital logic generating a higher frequency signal for S_1 and a lower frequency signal for S_2 . The edges of these two signals are intentionally misaligned so that S_1 and S_2 will not conduct at the same time.

This HPF design is able to meet multiple requirements: noise performance is the same as an R-C HPF; linearity is high since the components are linear passives except a MOS-transistor based switch, which performs close to an ideal switch with appropriate sizing; and since R_{eq} is well controlled and no transistor-stacking is employed, the voltage offset due to parasitic leakage current is much smaller as compared to the pseudo-resistor approach and

measured to be less than 1 mV . Consequently, the multi-rate DCR based HPF provides a reliable input interface for our front-end.

2.5 VCO Design

2.5.1 VCO Schematic

As shown in Fig 2.9, the VCO directly connects to the electrode and translates the voltage into the frequency/phase domain. It also needs to integrate the multi-rate DCR based HPF as discussed in the previous section. Moreover, since it is the first block in the signal chain, it is the bottleneck for the front-end performance, i.e. the noise, the raw linearity and the CMRR. These performance requirements dictate the architecture of the VCO.

The VCO schematic is shown in Fig 2.15. The differential pair follows the HPFs and translates the filtered input voltage into a differential current I_{diff} . I_{diff} is then commutated through a chopper to result in the current I_{osc} , which supplies two ring oscillators (ROs). Ignoring the chopper for now, one can see that the input voltage modulates the current and thus the RO frequencies. Therefore, this structure forms a VCO when the frequency difference between the two ROs is treated as the output.

The differential pair (diff pair) provides the requisite interface for the HPF. Since the HPF outputs are connected to the gates of the input diff-pair transistors, the HPF transfer function is maintained. The input differential-mode voltage is converted to I_{diff} with the transconductance of the diff pair. The input common-mode voltage fluctuation is rejected, with sufficient voltage headroom for the diff pair and the tail current source. Simulation reveals that the bias current does not change by more than 0.6% for a $\pm 50 mV$ common-mode voltage change.

The diff-pair transistors are sized according to noise, CMRR and linearity considerations. The flicker noise of the diff pair directly contributes to the front-end input noise. It is thus desirable to size-up the transistors to suppress the flicker noise. Larger input devices also reduce the transistor DC offset and improve CMRR. However, sizing-up transistors

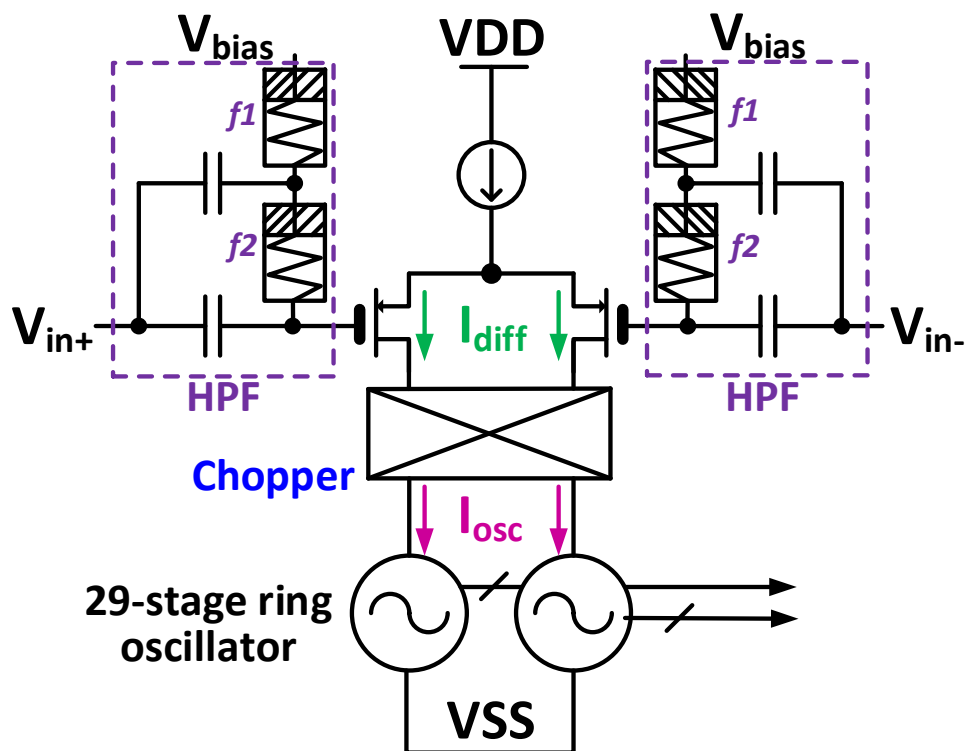


Figure 2.15: The VCO schematic for the front-end.

implies increasing capacitance at transistor gates. Hence, the capacitive division between the HPF capacitor and the transistor gate-capacitor will attenuate the input signal, leading to higher input-referred noise. The input-transistor size needs to achieve a good trade-off between these factors. On the other hand, the width/length ratio of the diff-pair transistor determines its overdrive voltage, and hence its transconductance g_m . A larger width/length ratio leads to a smaller overdrive voltage and a higher g_m , resulting in lower input-referred thermal noise. This benefit diminishes when the transistors enter into the deep-subthreshold region. Moreover, with a low overdrive voltage, the diff pair would be more nonlinear for a given input swing, and hence would require a greater effort in nonlinearity correction. An overdrive voltage of approximately $80 - 90 \text{ mV}$ proves a good trade off between the thermal noise and the raw linearity. Given all these considerations, the size of the transistors are chosen as $\frac{W}{L} = \frac{15 \mu\text{m}}{30 \mu\text{m}}$. The offset is below 1 mV , leading to 60 dB+ CMRR. The gate capacitance of the diff-pair transistors is approximately 10 pF , incurring $< 10\%$ noise performance degradation with 100 pF HPF capacitor.

The ROs each comprises 29 stages of inverters. From section 2.6, it is known that the inverter delay is the unit of phase count, and thus the number of stages in the RO does not influence the resolution of the front-end. Our choice of the number of stages (29) is based on two concerns: too few RO stages leads to a very high RO frequency that is challenging for the quantizer logic to meet the timing constraint; however, too many stages leads to accumulated mismatches within the RO which degrades DNL/INL of the front-end. The buffers following the ROs isolate the RO oscillation operation from the quantizer digital logic. The buffers also sharpen the slope of the RO outputs and shift the level to the supply rail of the digital logic. Additional discussion of the interface is presented in section 2.6.

The ROs also introduce noise in the front-end. Its thermal noise and flicker noise contribute to the phase noise or jitter of the RO output edges, and thus add noise to the sampled phase information in the quantizer. The thermal noise is from both the PMOS and NMOS devices of the inverter. A PMOS/NMOS transistor has a current noise density of $4kT\gamma g_m$ when charging the inverter output capacitance. It is shown in equation (33) of [41] that the thermal noise contribution depends on the bias current and the RO supply voltage, both of

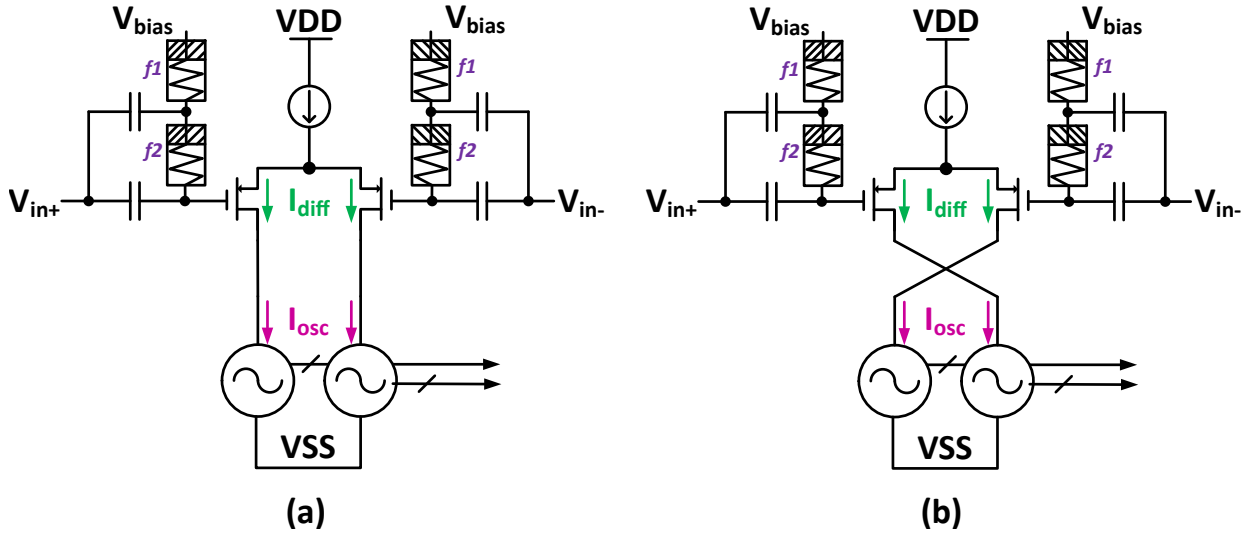


Figure 2.16: Chopping operation of the VCO. (a) connection at the positive phase; (b) connection at the negative phase.

which do not have much room for optimization. The flicker noise of the ROs can be suppressed by sizing-up the inverters, as in equation (45) of [41]. However, the noise reduction leads to longer delays, i.e. reduced phase resolution. Therefore, a chopper is inserted in between the diff pair and the ROs to separate the input signal and the RO-induced flicker noise in the frequency domain to avoid this trade-off. The chopping is further discussed below.

2.5.2 Chopping Inside of the VCO

The chopping operation in the VCO is presented in Fig 2.16, in which I_{diff} is commutated, resulting in the current I_{osc} . The current commutation via switches in the chopper is controlled by the digitally-timed signals S_p and S_n , the timing of which is shown in Fig 2.19. The small overlap between S_p and S_n ensures the availability of a current path for a quick transition between the two chopping phases.

Ignoring the overlap between S_p and S_n , chopping leads to the following relationship between currents:

$$I_{osc}(t) = I_{diff}(t) \times S\left(\frac{t}{T_{chop}}\right) \quad (2.13)$$

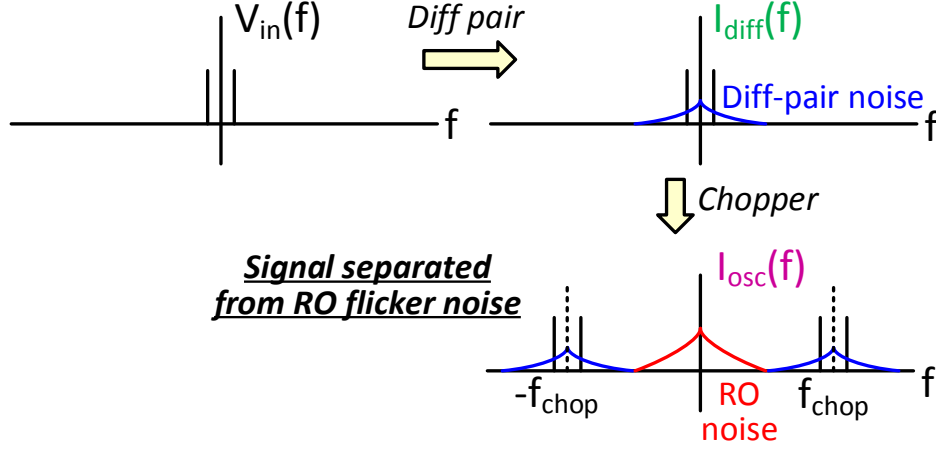


Figure 2.17: Spectrum-domain representation of the chopping operation.

where $S(\frac{t}{T_{chop}})$ is a square-wave function with two levels (+1 and -1) and a period of T_{chop} . This equation can be mapped to the Fourier domain as follows:

$$I_{osc}(f) = I_{diff}(f) * \mathcal{F}[S(\frac{t}{T_{chop}})] \quad (2.14)$$

where the latter item in this equation contains an impulse at $\pm f_{chop}$ and a lower impulse at $\pm 3f_{chop}$, etc. This convolution translates the current I_{diff} from low frequencies to frequencies centered at the odd harmonics of f_{chop} . Hence, the input signal is kept away from the flicker noise contributed by the ROs, as shown in Fig 2.17. Signal components that are up-modulated to $\pm 3f_{chop}$ or beyond are not shown for visual clarity. The RO current I_{osc} then controls the VCO output frequency, which will be digitized in the quantizer block.

2.6 Quantizer Design

2.6.1 Signal Processing of the Quantizer

As discussed previously, the VCO output information is in the frequency/phase domain, not in the voltage domain. This information needs to be sampled and quantized for further signal processing. Moreover, since the input signal and the RO-induced flicker noise are separated in the frequency domain, it is necessary to suppress the latter and bring the chopped signal

back to baseband, i.e. invert the chopping.

We will first intuitively discuss how to suppress the RO-induced flicker noise and invert the chopping at the same time. Sitting at low frequencies, the flicker noise perturbation from the RO stages remains almost unchanged over the positive phase and the negative phase in one chopping period. Thus, the RO output noise in the two phases are highly correlated. On the other hand, the VCO output frequency is modulated by $V_{in+} - V_{in-}$ in the positive phase and $V_{in-} - V_{in+}$ in the negative phase because of chopping. Thus, a subtraction of the RO outputs resulting from the two chopping phases can eliminate the RO-induced flicker noise while restoring the input signal to baseband.

A more rigorous analysis is achieved by inspecting the spectrum of the frequency/phase modulation information. The discussion in this part tries not to use the term “frequency domain” to avoid confusion. As presented in Fig 2.18, the first step is to obtain the traversed phase information in the two chopping phases. This corresponds to sampling at $2f_{chop}$, which transforms the information into the digital domain, noted as Φ_{OSC} , and does not filter out either the signal or RO-induced noise. The sampling of the traversed phase is a windowed integration of the RO frequency, and thus introduces a sinc function (not shown here). It does not distort the chopped signal, but filters out the aliasing components beyond $2f_{chop}$. Therefore, the spectrum is limited to $-f_{chop}$ to f_{chop} . Following this sampling is the difference operation, i.e. a subtraction of the results from the two chopping phases. The difference introduces a high-pass filtering of $(1 - Z^{-1})$ that notches out the low frequency component and retains the signal component around f_{chop} . The apparent gain of 2 at the chopping frequency can be explained as a realignment and addition of the input-modulation component at both phases. The result of subtraction, $\Delta\Phi_{OSC}$, with significantly suppressed RO-induced flicker noise, can be resampled at f_{chop} . This resampling translates the up-modulated input signal back to baseband.

The quantizer architecture, as shown in Fig 2.19, implements the above-mentioned operations. the RO sub-quantizers calculate the traversed phase of each ring oscillator within the time windows defined by logic high of the *Count* signal. There are two windows within one chopping period, one in the positive phase and the other one in the negative phase. The

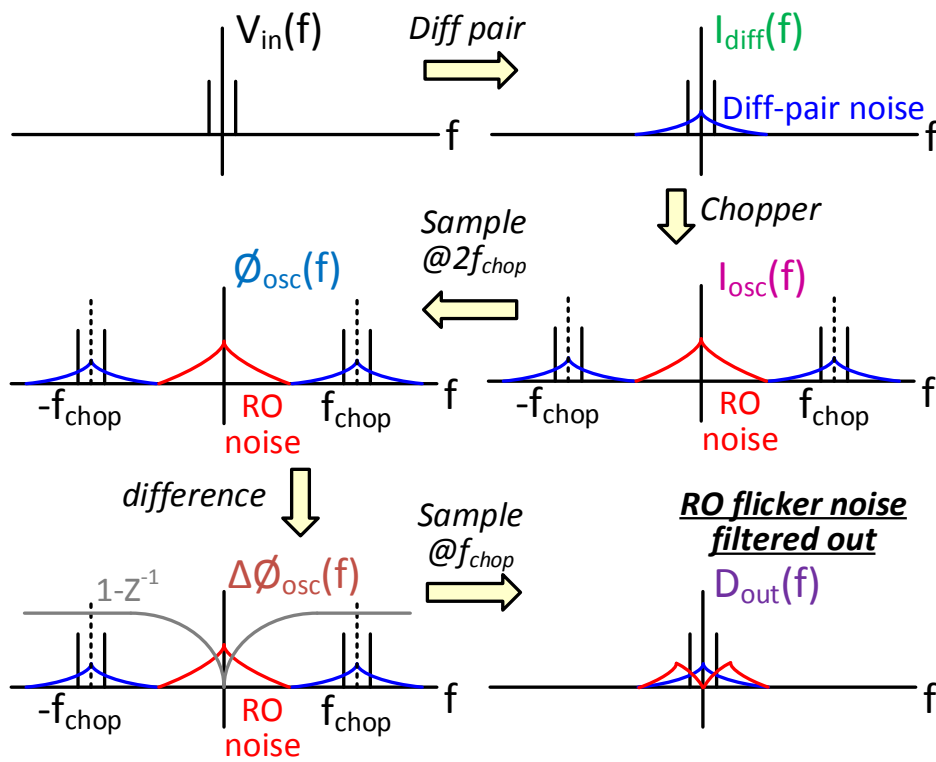


Figure 2.18: The post-chopping operation in the quantizer.

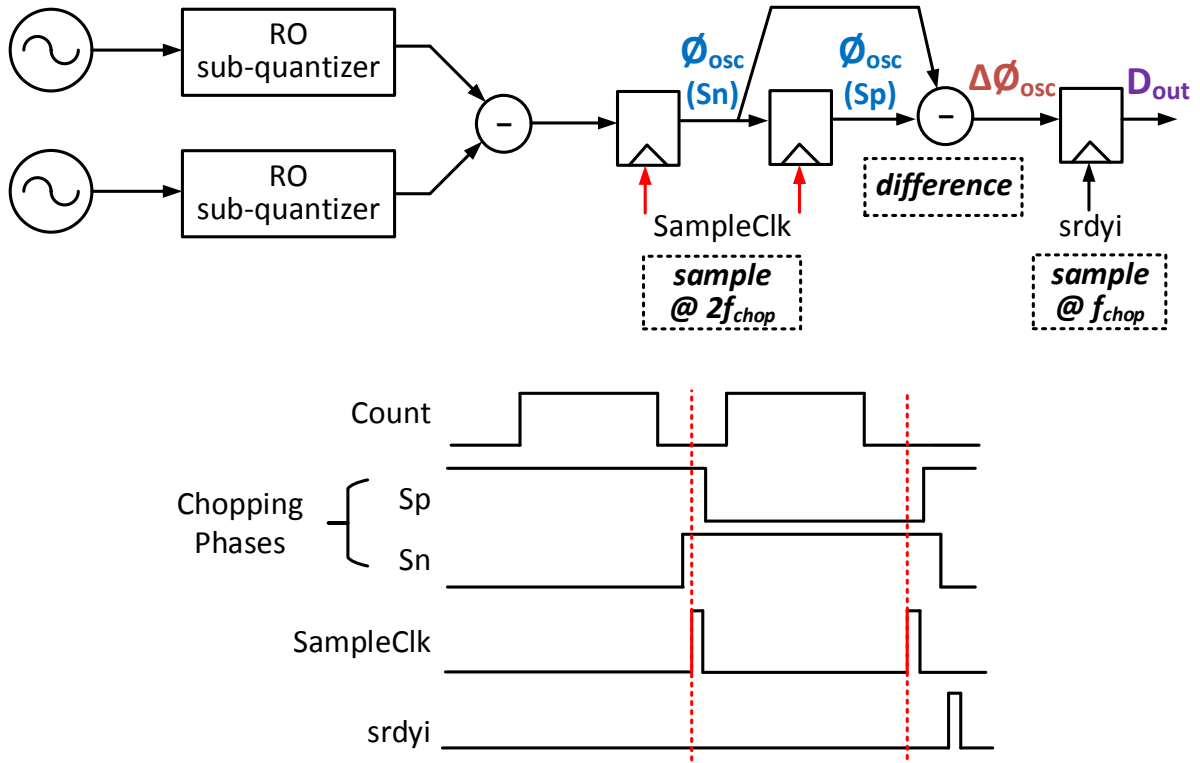


Figure 2.19: Quantizer schematic and timing diagram.

rising edge of *Count* is sufficiently away from the phase transition, so that the oscillator frequency has already settled for the phase sampling. The output of the RO sub-quantizers subtract each other, resulting in the differential phase information. The subtracted output is then sampled at $2f_{chop}$ at rising edges of the *SampleClk* signal. The sampled phases, annotated as Φ_{OSC} for phases S_p (positive phase) and S_n (negative phase), are subtracted to obtain the difference $\Delta\Phi_{OSC}$. Finally, $\Delta\Phi_{OSC}$ is resampled by the rising edge of the signal *srdyi* and sent to the NLC block for further processing.

The design of the RO sub-quantizers in Fig 2.19 needs to consider the interface with the RO outputs, and interpret their transitions as appropriate phase information. In the following subsections, the RO sub-quantizer is presented in detail and then details of the glitch issue, which is due to the asynchronous analog-digital interface, are discussed.

2.6.2 Coarse-Fine Quantization of the RO Phase

As discussed in section 2.3, the VCO can provide very high resolution with high K_{VCO} and low sampling rate. In a ring-oscillator based VCO, the phase is quantized with the resolution of a unit delay of one stage, i.e. one inverter delay in our design. Although further resolution enhancement is possible with added circuit complexity, e.g. inter-stage interpolation [42][43] or ADC assistance for sub-gate delay resolution [44], these techniques consume additional power, and enhanced resolution is unnecessary for our application.

The oscillator phase is decoded by inspecting inverters' output logic levels. A ring oscillator with an N-stage inverter chain (N being an odd number) has an oscillation period of 2N inverter delays, i.e. 2N phases. The output transition ripples from one inverter, along all stages and returns to the starting inverter after N inverter delays, but with opposite polarity. It takes another N inverter delays to ripple back to the starting inverter with the original polarity. Since the oscillator behaves in a rippling operation, there is only one inverter at transition at any time, which is defined as "active". For the active inverter, the transition would complete when the output reaches from the same logic level of the input to the opposite logic level of the input. Therefore, the input and output logic levels are the same, either 2'b00 or 2'b11, before the active operation of this inverter. A simple XOR operation between the input and output of every inverter can detect which inverter is in the active state. Considering that the oscillation ripples through the same inverter twice within one period, phase decoding should also consider the logic level of the inverter in transition.

For our 29-stage RO, the phase decoder implementation is presented in Fig 2.20. First, the RO outputs are buffered to avoid excessive loading on the RO stages. Buffer outputs *ro_buffer* are then latched by the *Count* signal. To decode both the initial phase and the final phase, they are latched on both edges. Only the latching at the positive edge is shown here for clarity. The latched outputs *InitState* are passed to the XOR gate array for the decoding operation. The output of this XOR array can be viewed as the real-time phase information. The left-most inverter is chosen as the starting stage in a RO. If this starting inverter is active at the *Count* rising edge, *ro_buffer[28]* and *ro_buffer[0]* will be latched at

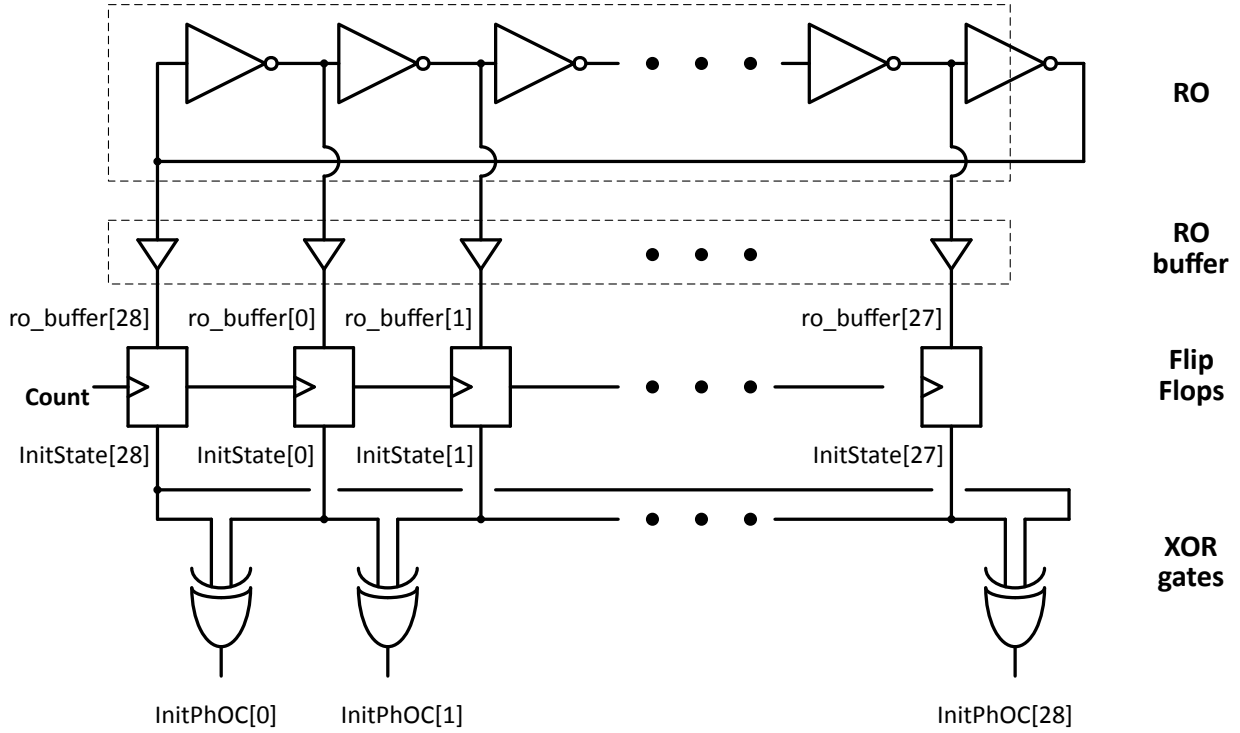


Figure 2.20: Schematic for the phase decoder.

Table 2.2: Example for phase decoder outputs.

$InitState[0:28]$	10101...0101	01010...1010	10010...1010	00101...0101
$InitPhOC[0:28]$	01111...1111	01111...1111	11011...1111	10111...1111
Φ_{INIT}	0	29	2	1

the same logic level. Therefore, the XOR operation will give $InitPhOC[0]$ as logic 0 and the rest of the $InitPhOC$ word as logic 1s. If we define phase 0 as when $InitPhOC[0]$ is 0 and $InitState[0]$ is 1, due to the logic level inversion in every stage, phase 1 will correspond to the case in which $InitPhOC[1]$ is 0 and $InitState[1]$ is 0 and phase 2 will correspond to the case in which $InitPhOC[2]$ is 0 and $InitState[2]$ is 1, etc. This continues until phase 29, where $InitPhOC[0]$ is 0 and $InitState[0]$ is 0. Phase 30 is then defined as the phase in which $InitPhOC[1]$ is 0 and $InitState[1]$ is 1. This continues until the last phase (phase 57). The decoding is presented in Table 2.2.

The phase detection as discussed above provides a phase value with a range only within

one oscillation period, or up to $2N$ phases. When traversing beyond $2N$ phases, the RO would finish one cycle and the phase would wrap around. This wrapping behavior requires us to record the traversed cycle as well, to achieve the desired high resolution. A counter with any RO output (any RO buffer output in practice) as the input can record oscillator cycles. A combination of cycle count and phase information delivers the unwrapped VCO phase output. This unwrapping concept of the phase with the cycle count is shown in Fig 2.21. The figure shows that the phase increases from phase 0 to phase 57 while the cycle count remains unchanged. Then in the event when the phase wraps back from phase 57 to phase 0, the cycle count increases by 1. This process continues for the duration of phase integration.

Fig 2.21 shows an example of the phase-unwrapping operation. The sampling window for the traversed phase in our VCO is defined by the *Count* signal, in which its rising edge latches the RO initial phase, and its falling edge latches the RO final phase. Therefore, the initial cycle count is 1 and the phase decoder output Φ_{INIT} is 2; whereas, the final cycle count is 10 and the phase decoder output Φ_{FINAL} is 0. The unwrapped phase calculation is:

$$\text{unwrap phase} = 2N \times \text{cycle count} + \text{decoded phase}$$

which leads to the unwrapped initial phase of $1 \times 58 + 2 = 60$ and the unwrapped final phase of $10 \times 58 + 0 = 580$.

This phase calculation is similar to the coarse-fine quantization scheme in voltage-domain data conversion. Thus, it also requires alignment of two quantization levels, which means that the phase wrapping and the cycle increment should occur at the same time. In the phase decoder logic, we have chosen an inverter as the starting inverter and define phase 0 as when its input and output are both logic-high (2'b11 for the digital designer). Every time when the oscillation returns to phase 0, the starting inverter input would have experienced a rising edge. Therefore, the starting inverter input can be utilized as the input of the cycle counter, and the counter should be positive-edge triggered.

The ideal RO sub-quantizer structure is shown in Fig 2.22. It splits into the phase path and the cycle path for the operations discussed above. It is again noted here that the RO

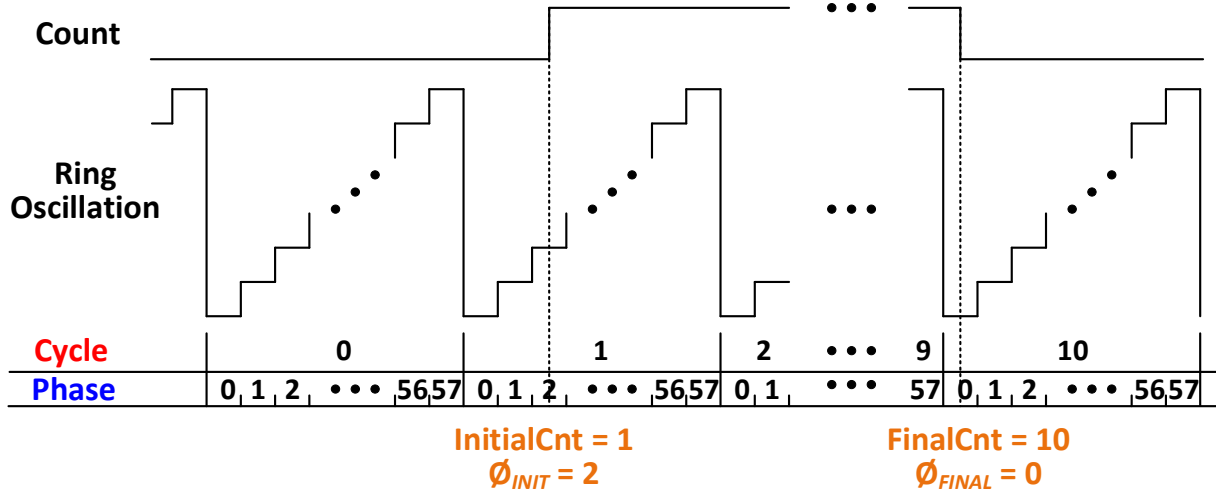


Figure 2.21: Example of unwrapping RO phase with cycle count and phase decoded output.

outputs are buffered first prior to going to the RO sub-quantizer, to isolate the RO operation from the sub-quantizer dynamics. The phase path contains the phase decoder, as shown in Fig 2.20, and results in wrapped phases Φ_{INIT} and Φ_{FINAL} . In the cycle path, *ro_buffer[28]* clocks the counter increment at its rising edge, when the counter is enabled by the *Count Enable* signal. The counter output is also latched at the edges of *Count*, as InitialCnt and FinalCnt for the phase-unwrapping operation.

2.6.3 Glitch-Free RO Sub-Quantizer Design

While the principle of this cycle-phase coarse-fine quantization is straightforward, the implementation requires understanding the intricacies of crossing the analog-digital interface, where timing is a general concern:

- The RO buffer output signals, *ro_buffer[0:28]*, are asynchronous with the signals *Count* and *Counter Enable*. Although ring oscillators in the VCO are turned on and off by the control signal, their frequencies depend on the input voltage and therefore are uncorrelated with the system clock for digital logic, resulting in asynchronicity. In the phase path of Fig 2.22, it can lead to metastability at a certain flip-flop output in the decoder. This metastability can be solved by giving enough time for the decoder to

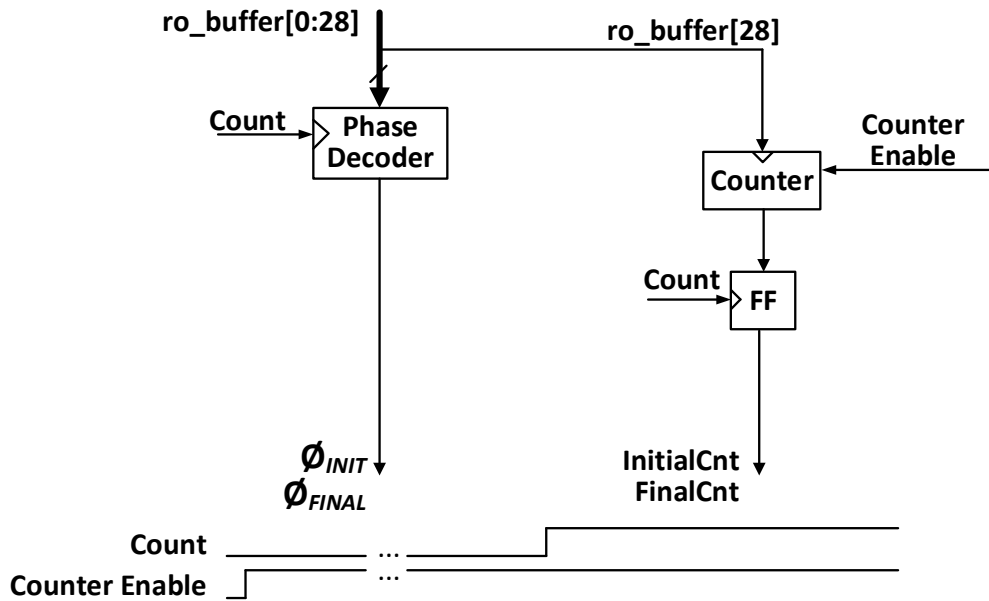


Figure 2.22: Ideal RO sub-quantizer structure.

resolve its output. Since the decoding logic uses the XOR operation of every stage output, the worst consequence of this metastability is a unit-error in one phase count. This does not affect overall performance, since the VCO has a sufficiently high resolution. Nevertheless, the metastability is more critical in the cycle path, in which the counter output might be latched by $Count$ edges while in transition, and the latched output can be off by more than one due to the binary coding of the counter. Additionally, the asynchronicity between $ro_buffer[28]$ and $Count Enable$ can result in an undetermined starting state of the counter.

- The phase path and the cycle path cannot be aligned perfectly in practice. The phase-decoder logic and the cycle counter may exercise different threshold voltage levels. In addition, the $Count$ signal for these two paths cannot be exactly matched, leading to clock skews. While the input information from the phase decoder is simply latched in flip-flops, the cycle counter requires a certain delay in propagating to the new counter output, i.e. the counter transition time is not zero. All of these factors add to the timing mismatch between the two paths.

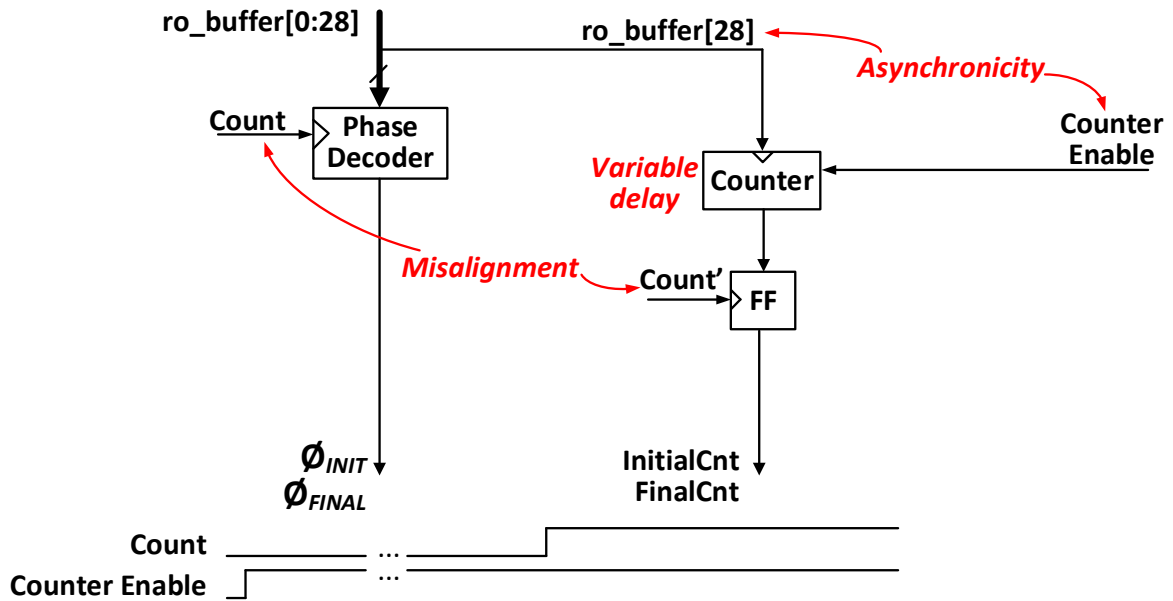


Figure 2.23: Practical timing issues with RO sub-quantizer.

All of these timing issues are highlighted in Fig 2.23. The latch signal of the cycle count path is noted as $Count'$, indicating the mismatch between the latching signals of each path. The unwrapped phase output can have glitch errors if these issues are not solved. An example is shown in Fig 2.24 for the latching of the final-phase at the falling edge of $Count$. The figure shows the scenario in which the phase wraps from 57 to 0 while the cycle count increases from 9 to 10. This is the same case as the final-phase latching in Fig 2.21, except that here the phase path is slightly ahead of the cycle path and the phase latching slightly lags behind the cycle-counter latching. Hence, the counter output $FinalCnt$ is 9 and the phase-decoder output Φ_{FINAL} is 0, resulting in an unwrapped phase of $9 \times 58 + 0 = 522$. This is 58 less than the correct value 580, and corresponds to a full-cycle glitch.

The mixed-signal design community has examined similar challenges in dealing with glitches due to timing issues across the asynchronous interface. Some redundancies are needed in either path to correct for possible glitches. For example, in the design of all-digital PLLs, the retiming of reference clock $FREF$ with the oscillator clock CKV is necessary, but can cause metastability and lead to an error in the phase-error calculation. To solve

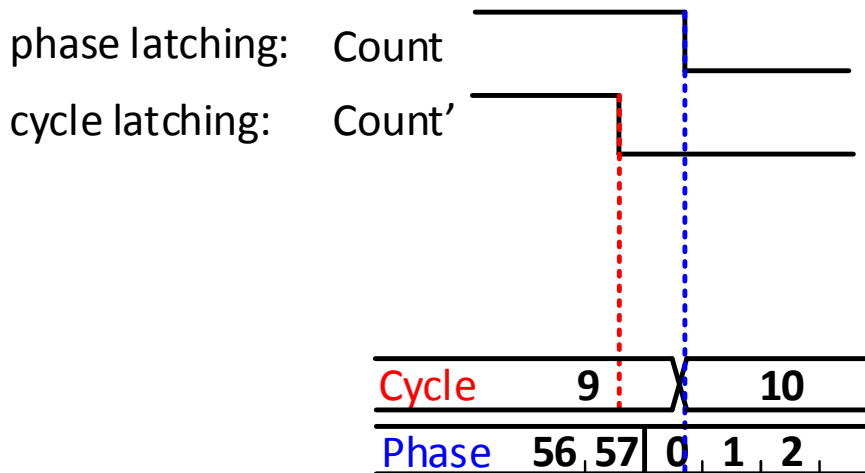


Figure 2.24: Example of output glitch due to timing issues.

this issue, a scheme of dual-edge retiming, where *FREF* is latched by both the positive and negative edges of *CKV*, is adopted [45][9]. The TDC output is employed to arbitrate which edge is appropriate for the use of phase error calculation, as presented in Fig 2.25. The VCO output *CKV* in the PLL is well-buffered and can drive the extra loads added by the dual-edge retiming scheme.

Another option is to have two counters, one clocking at the input rising edge and the other clocking at the input falling edge. The output is similarly arbitrated by the decoded phase information, as shown in Fig 2.26 [10].

While it is feasible to implement the above-mentioned approaches for our system, they incur some disadvantages: the dual-edge retiming used in the ADPLL would add significant load to the clock signal, *ro_buffer[28]* in our case, and lead to inter-stage delay mismatch; and the dual-counter approach doubles the logic that is switching at the RO frequency, leading to higher power consumption. Therefore, we implement a different method to solve the glitch issue: multiple latching of the cycle counter output by signals generated from a delay line. In our implementation, the signal *Count'* is now split into three signals: *Count1*, *Count2* and *Count3*. *Count1* is ahead of *Count* while *Count3* lags behind *Count*, and *Count2* is between

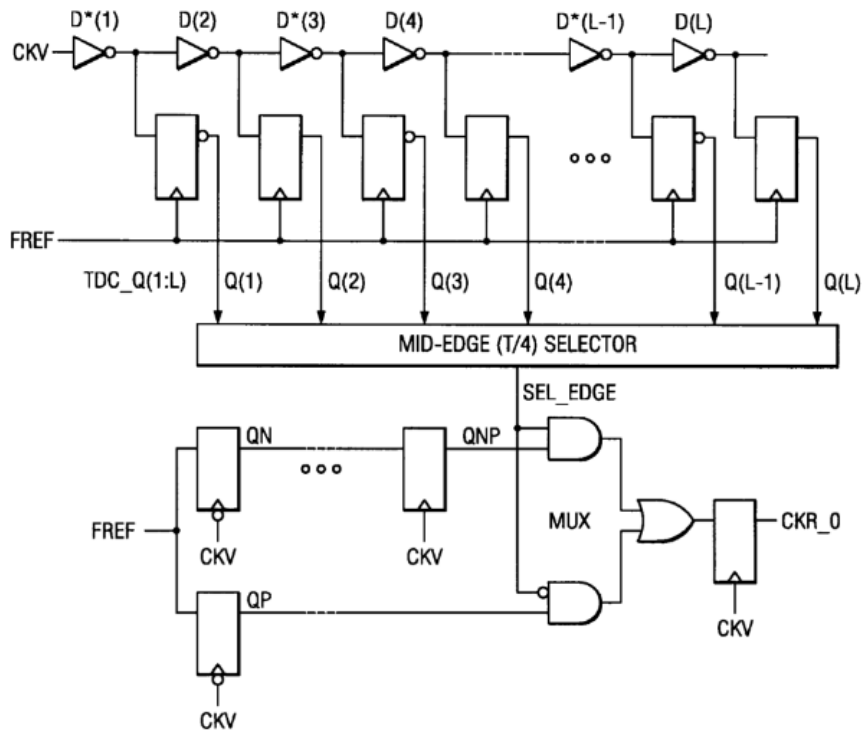


Figure 2.25: Dual-edge retiming in ADPLL resorts to redundancy from TDC [9].

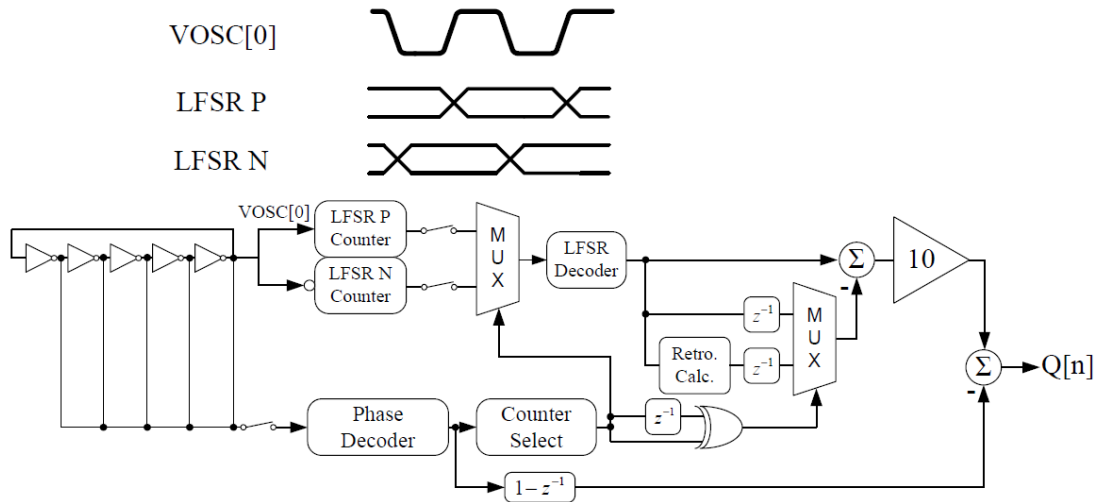


Figure 2.26: Two counters at dual edges of the oscillation to remove glitch condition [10].

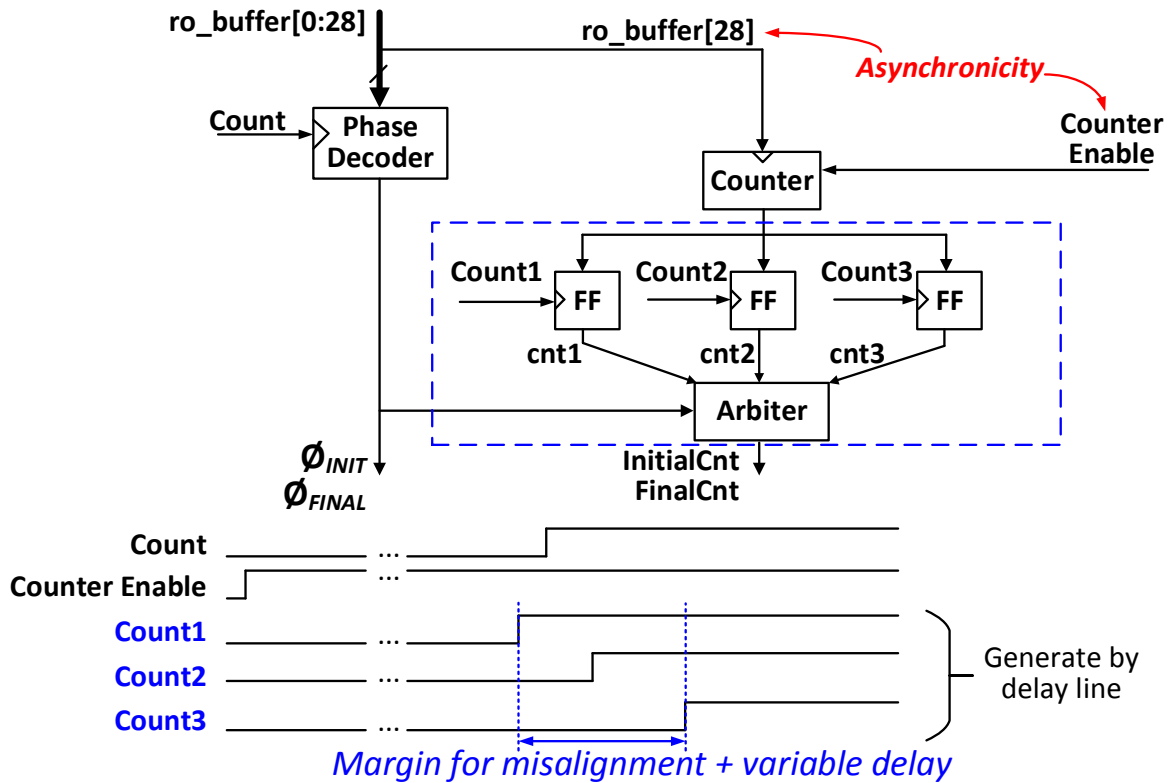


Figure 2.27: Multi-Latching technique on resolving the timing issue.

Count1 and *Count3*. The latched outputs (*cnt1*, *cnt2*, *cnt3*) are sent to an arbiter which makes the final decision based on the phase-decoder output. In this multi-latching scheme, the delay between *Count1* and *Count3* serves as the margin for the path misalignment and the variable delay incurred in the counter transition.

The arbitration process is presented with four examples shown in Fig 2.28. It is noted that if $cnt1 = cnt2 = cnt3$, no phase-wrapping transition occurs within the edges of these three signals, and thus no arbitration is needed. In scenario (a), where $cnt1 = cnt2 \neq cnt3$ and $\Phi_{FINAL} = 56$, we recognize that the phase has not wrapped back in the phase path. Thus *FinalCnt* should be the value of *cnt1*, which is 9. In scenario (b), where $cnt1 = cnt2 \neq cnt3$ yet $\Phi_{FINAL} = 0$, the phase has wrapped back to a small value, and therefore a large value is chosen for *FinalCnt* as $cnt1 + 1 = 10$ (note that we do not directly choose *cnt3* since this value may be incorrect if *Count3* latches at the counter output transition). In scenario (c), *cnt2* is a number that is distinct from *cnt1* and *cnt3*, and $\Phi_{FINAL} = 0$. This indicates that

Count2's falling edge is close to the counter output transition, and the phase has wrapped back. Therefore, we choose FinalCnt as cnt3, which is 10. In the scenario (d), we have $\text{cnt1} \neq \text{cnt2} = \text{cnt3}$, and $\Phi_{FINAL} = 57$. The phase decoder output indicates that the phase has not wrapped back, and thus FinalCnt is arbitrated as $\text{cnt3} - 1 = 9$. The phase-unwrapping calculations yield numbers that are close to 580 in all four scenarios, and thus generate no glitch.

With the multi-latching technique, the only remaining issue is the asynchronicity between the cycle counter input, *ro_buffer[28]*, and the *Count Enable* signal. This is resolved by retiming *Count Enable* with *ro_buffer[28]*. The final RO sub-quantizer structure, as shown in Fig 2.29, implements double-latch retiming, a standard technique in the digital circuit. The one RO cycle in between two flip-flops allows for resolving any possible metastability.

Although the above discussion resolves the timing concerns in the RO sub-quantizer, some implementation details need be considered in the voltage-domain crossing at the analog-digital interface. This crossing occurs between the RO and the RO buffer. Since the supply current of the RO is from the diff pair, the RO supply voltage $V_{ro,H}$ depends on the front-end input. On the other hand, the RO buffer is supplied by the digital supply with a fixed voltage level VDD_{buffer} , for interfacing with the following digital logic. When $V_{ro,H}$ is significantly lower than VDD_{buffer} , the decision threshold of the RO buffer is close to $V_{ro,H}$. With this high threshold, a RO buffer will have a delayed logic-high output (1'b1) and an advanced logic-low output (1'b0) compared to a half- $V_{ro,H}$ threshold. Consequently, the duration of a phase defined by detecting 2'b11 in the decoder output is much shorter than that defined by detecting 2'b00. This phase imbalance, together with the delay mismatch in between RO buffer outputs, can cause half-cycle (29 LSB) glitch errors. This glitch condition is demonstrated in Fig 2.30. Shown in (a) are O output waveforms with a high decision threshold (the dashed lines) due to a high VDD_{buffer} . The ideal buffer outputs are shown in (b), where *ro_buffer[28]* and *ro_buffer[0]* are both high at the *Count* rising edge. Therefore, the phase decoder determines that it is a rising-edge condition for *ro_buffer[28]* and thus Φ_{INIT} is decoded as 0, yet for a much shorter duration compared to the phase decoded as 1 or 57. However, adding a timing mismatch between *ro_buffer[28]* and *ro_buffer[0]*, as

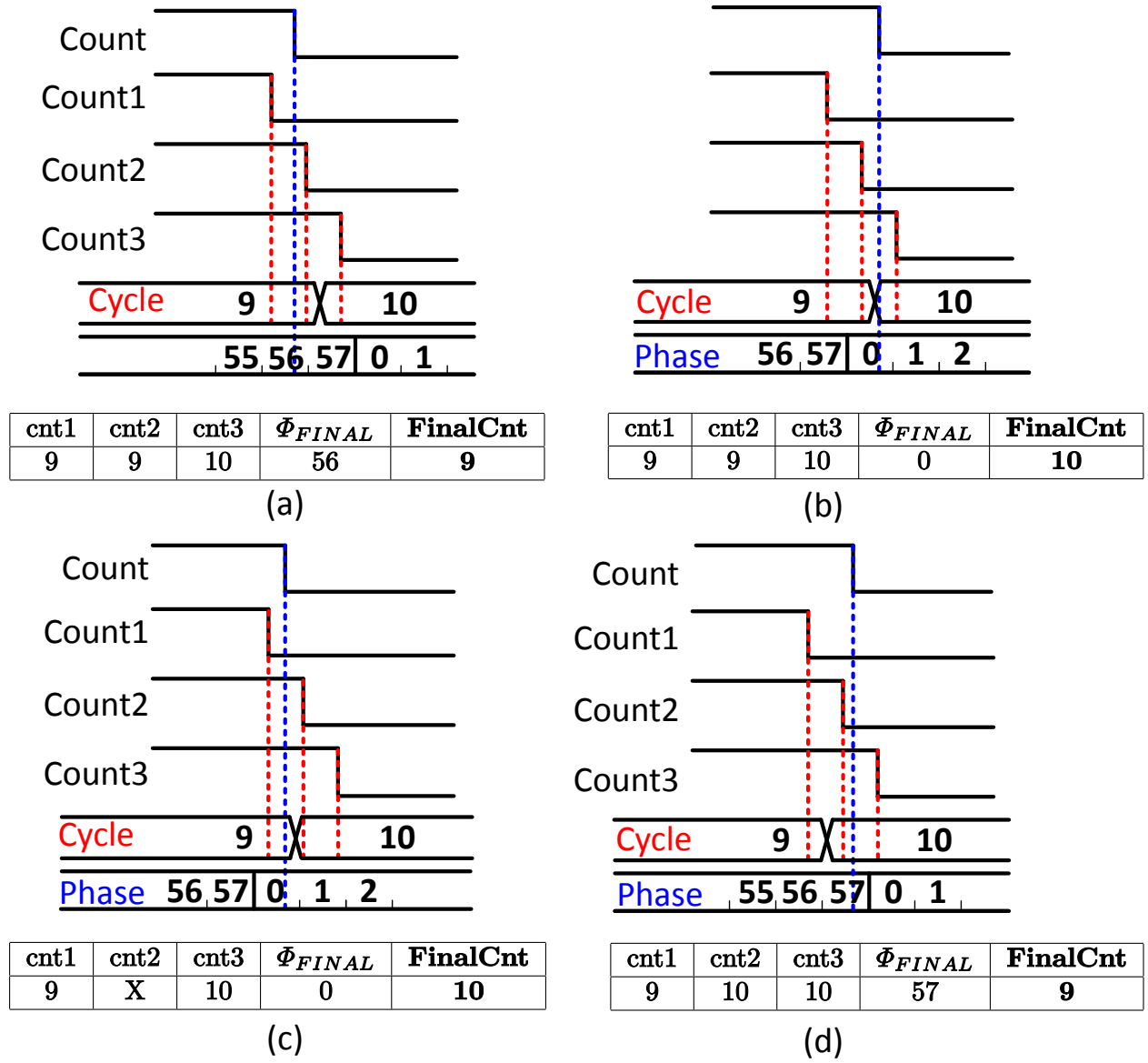


Figure 2.28: Example scenarios for the arbitration of the multi-latching technique.

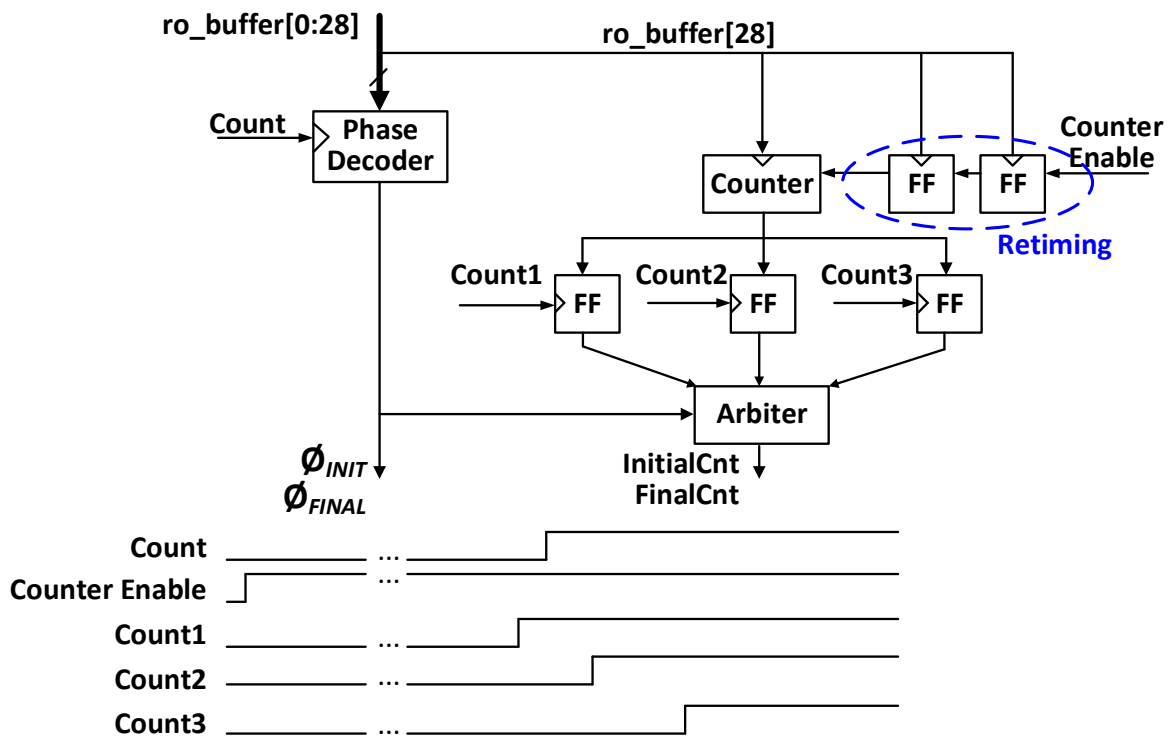


Figure 2.29: Retiming of the *Counter Enable* signal.

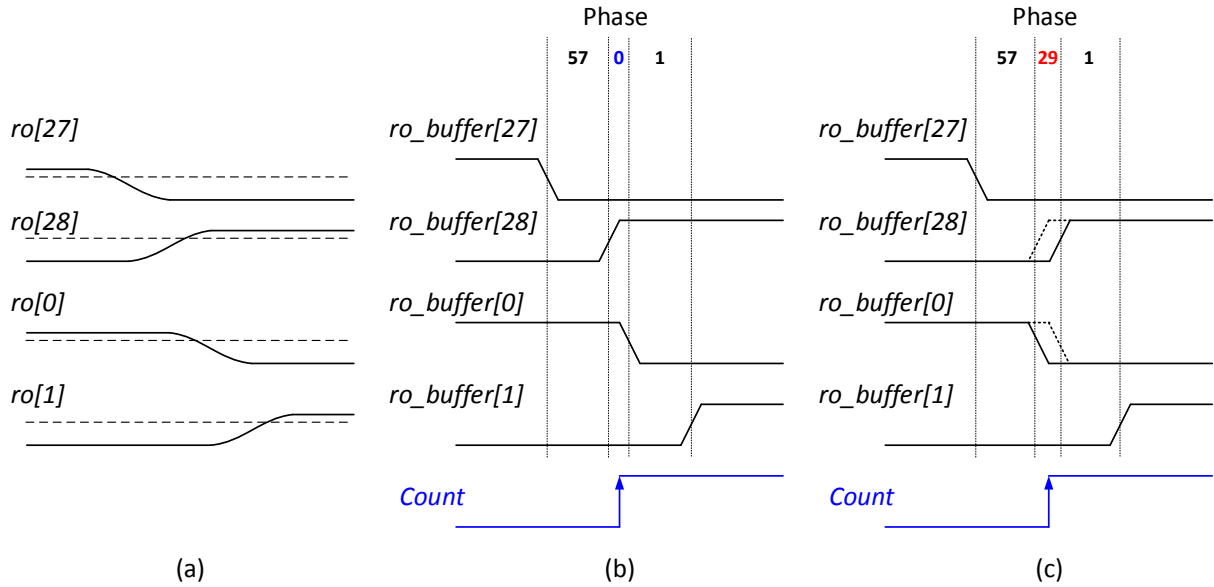


Figure 2.30: Voltage interfacing between RO and the buffer leading to the half-cycle glitch.

presented in (c), *Count* latches $ro_buffer[28]$ and $ro_buffer[0]$ as both logic-low. The phase decoder will incorrectly determine that it is a falling-edge condition for $ro_buffer[28]$ and decode it as phase 29, and thus introduce a glitch of 29 phases (half a cycle).

The circuit implementation avoids this half-cycle glitch error through a joint effort on the analog domain and the digital domain. Firstly, the RO inverters are scaled so that at the full-scale ($\pm 50\text{ mV}$) input, $V_{ro,H}$ is not too low compared to VDD_{buffer} . This minimizes the phase imbalance and provides more margin against the RO buffer output mismatch. Secondly, the phase decoder adopts an edge detection that correctly detects the transition edge even in the condition of Fig 2.30(c). The edge detection inspects the logic level of stable outputs from the other inverters in the RO instead of the input/output of the transitioning inverter. In the case of Fig 2.30(c), the decoder uses $ro_buffer[24]$ and $ro_buffer[4]$ (not shown here) for edge-detection. They are both logic-high despite all the dynamics in $ro_buffer[0][28]$ at the *Count* rising edge, and thus correctly indicate the phase as 0 instead of 29.

All of these techniques for eliminating the glitch issue should be verified in the simulation prior to implementation. The full-precision transistor-level simulation is ideally desired for capturing the details at the interface. However, since the glitch rate can be very low (once

per a few thousand samples), a simulation duration of at least tens of seconds is needed, which is impossible in practice. Therefore, a pure digital simulation with the VCO simplified as a behavioral model is used instead. Some extra efforts are then needed since:

- The digital simulation alone over-simplifies the analog-digital interface and no voltage-crossing issue can be revealed.
- The standard-cell timing definition for the digital simulation is too cumbersome. It is good for some fully-synchronous designs, but not intuitive for debugging in an inherent asynchronous interface.

Therefore, some model customization is requisite, particularly in the behavioral-level simulation for early verification. We model the flip-flop specifically such that within the metastability region (set-up/hold time error regions) it gives a random output initially, and another random output after a long time compared to a typical flip-flop clock-to-Q delay. These two outputs can have opposite logic levels. In addition, we have added the option to inject the phase latching error, as shown in Fig 2.30(c). Digital simulation with these models can easily accommodate a fully automatic or semi-automatic set-up in an effort to exhaust all conditions. These extra efforts ensure that the final RO sub-quantizer design has successfully addressed all of the issues discussed above and delivers a correct unwrapped phase output.

2.7 VCO Nonlinearity Correction

The quantizer delivers a low-noise and glitch-free digitized signal. At this point, however, the VCO nonlinearity has not been dealt with. This nonlinearity is the limiting factor of the signal quality at large input swings, or more specifically, it limits the signal quality during concurrent sensing and simulation, in the presence of stimulation artifacts. The NLC block has been designed and implemented to correct for this nonlinearity, and to restore an accurate digitized input for the following signal-processing blocks. It thus allows extracting meaningful information under concurrent sensing and stimulation.

It is essential to understand the sources of VCO nonlinearity for choosing an appropriate

correction technique:

- The diff-pair voltage-to-current transfer function dominates VCO nonlinearity. The diff pair has a compressive $V - I$ curve, in which the asymptote remains linear till V_{ov} and then flattens beyond that point. Without considering the subthreshold operation of transistors, the ideal cut-off would be at $\sqrt{2}V_{ov}$. This implies that V_{ov} should be larger than the signal swing, as presented in section 2.5 where diff-pair sizing was discussed; this is to avoid deep compression of the curve and to ease the correction. This also ensures sufficiently high resolution throughout the whole input range. With the chosen V_{ov} , the compression can be captured as 3rd or 5th order nonlinearity and will be corrected for.
- The RO current-to-frequency function is also nonlinear because the RO supply voltage changes with its current. The RO conducts current in a relay fashion, in which the NMOS in one inverter stage discharges the inverter output after the PMOS in the previous stage almost finishes the charging. After that discharge, the PMOS of the next stage charges its output, and this continues. Consequently, in a steady oscillation state, the average current by NMOS transistors is proportional to the supply current. For a higher supply current, the transistors need a higher V_{ov} and thus a higher supply voltage. Given supply voltage V_{ro} , the current I_{ro} , and the inverter stage capacitance C_{inv} , we have the following:

$$f_{ro} = \frac{1}{T_{ro}} = \frac{1}{2N} \frac{I_{ro}}{V_{ro}C_{inv}} \quad (2.15)$$

It is clear that V_{ro} reliance on I_{ro} breaks the linear relationship between f_{ro} and I_{ro} . The RO transistors are sized that when turned on, they operate mostly in the weak-inversion region, where V_{ov} is a logarithmic function of the conducted current. Therefore, this $V_{ro} - I_{ro}$ reliance is weak, and consequently, this effect is smaller compared to the diff-pair $V - I$ nonlinearity. However, it is significant enough to be corrected for in order to achieve our desired resolution.

Although these two sources are deterministic and static (or memoryless), a comparison with the dynamic component is still needed to determine if the memory effect of the VCO,

in particular, the diff pair, should be considered for the correction. For the diff pair, the capacitive current is $\omega C_{gs}V_{in}$, while the transconductance current is $g_m V_{in}$. The Miller capacitance is ignored here because C_{gd} is small in the saturation region, and the weak reliance of V_{ro} on I_{ro} yields almost no voltage gain at the diff-pair drain nodes. The ratio between these two currents is exactly the ratio between the maximum input signal frequency and f_T of the transistor. A quantitative comparison has:

$$\omega C_{gs} \leq 2\pi \times 250 \text{ Hz} \times 5 \text{ pF} = 7.85 \times 10^{-9} \Omega^{-1} \quad (2.16)$$

and

$$g_m \approx \frac{2I}{V_{ov}} = \frac{2 \times 1.5 \times 10^{-6}}{80 \times 10^{-3}} = 3.75 \times 10^{-5} \Omega^{-1} \quad (2.17)$$

where we use the previously-mentioned number of 5 pF for C_{gs} , 80 mV for V_{ov} of diff-pair transistors, and 1.5 μ A for each diff-pair branch. These two equations show that the capacitive current is approximately -74 dB lower than the transconductance current of the transistor. Nonlinearity due to capacitance nonlinearity would be at least another $10-20 \text{ dB}$ lower. Therefore, the memory effect of the VCO can be neglected when considering the VCO nonlinearity correction.

Since VCO nonlinearity can be considered static, the NLC only needs to invert the overall voltage-to-frequency function, as shown by the red line of Fig 2.31, to linearize the entire front-end. The monotonous V-to-F function makes post-digitization correction feasible by mapping the digital output back into the desired point, shown by the blue dashed line.

Considering the low sampling rate of the front-end, the digital correction is also energy efficient. The correction logic only consumes a few μ W with the available advanced CMOS technology. More area and energy saving techniques can also be applied, such as interleaving or voltage scaling.

The NLC can be implemented in two categories: background correction or foreground correction. The background correction aims to automatically discover the VCO nonlinearity and then feed it back to the correction parameters or coefficients; although it is robust and accurate against the process, temperature and supply variations, it is complex and incurs potential area/power cost. On the other hand, the foreground correction typically requires

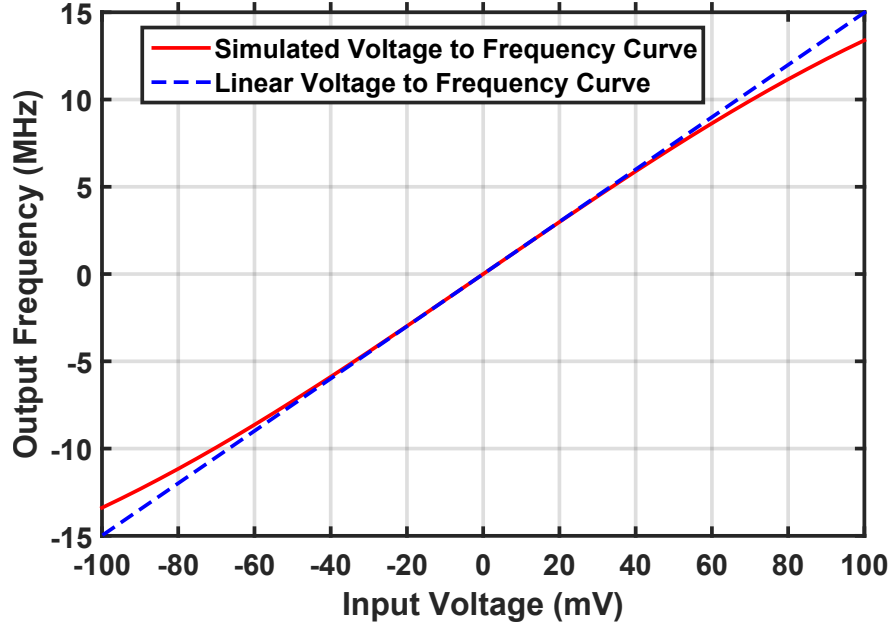


Figure 2.31: Voltage-to-frequency transfer function of the VCO.

an external configuration of the correction parameters/coefficients but is often simpler to implement. We choose the foreground-correction approach because a stable environment is available for a consistent VCO nonlinearity in our application. In the neuro-implant, the supply is provided by the LDO, as presented in section 4.1, thus it is well regulated. The body temperature range is typically within $\pm 2^\circ\text{C}$ and given the low power consumption of the implant system, it is expected that the chip temperature does not vary beyond 5°C . As shown in section 3.2, VCO gain and nonlinearity remain stable in this range. Therefore, a foreground correction is sufficient to serve our purpose.

The detailed implementation of the NLC block is described in [12] and outside the scope of this dissertation. This implementation is briefly shown here, to present a complete system. A polynomial-correction scheme is adopted to invert the V-to-F mapping up to the 5th order. Given our sensing input is sampled at a low data rate, Horner’s method is utilized to trade the calculation time for the area.

Horner’s method calculates the polynomial in an iterative way:

$$\begin{aligned}
 y[n] &= a_0 + a_1x[n] + a_2x[n]^2 + a_3x[n]^3 + a_4x[n]^4 + a_5x[n]^5 \\
 &= a_0 + x[n](a_1 + x[n](a_2 + x[n](a_3 + x[n](a_4 + x[n](a_5 + x[n] \times 0))))
 \end{aligned} \tag{2.18}$$

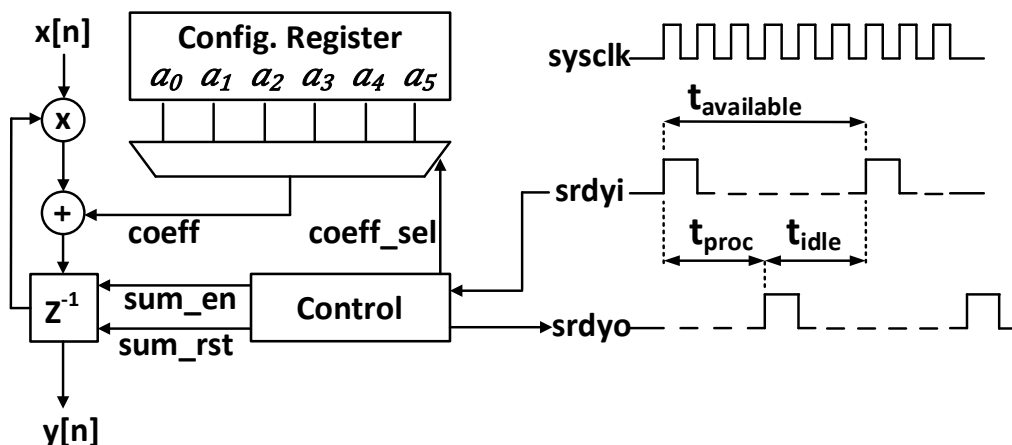


Figure 2.32: NLC implementation: Horner's method for polynomial correction.

where the calculation starts with the innermost multiplication and accumulation (MAC) operation, $a_5 + x[n] \times 0$, and uses the result as one operand for the outer MAC operation with a_4 and $x[n]$. This continues until the outermost layer with a_0 and $x[n]$. In the hardware implementation, as presented in Fig 2.32, this reduces multiple multipliers into one MAC unit and one sequencer for scheduling the iteration, leading to great area savings.

2.8 Front-End Implementation

It is worth revisiting the entire front-end structure prior to presenting the silicon prototype. The spectrum-domain presentation is added to Fig 2.9 and is presented in Fig 2.33. While the input voltage, shown as the blue bin in the figure, modulates the VCO frequency, the distortion terms due to the VCO nonlinearity are generated as well, shown as red bins. The quantizer samples and quantizes the traversed VCO phase within the window (i.e. window-integration of the VCO frequency). The following NLC module then restores the system linearity and suppresses all distortion terms to below the noise level.

The sampling rate of the front-end has to be higher than the Nyquist rate of the input signal $2f_{in,max}$ for two reasons:

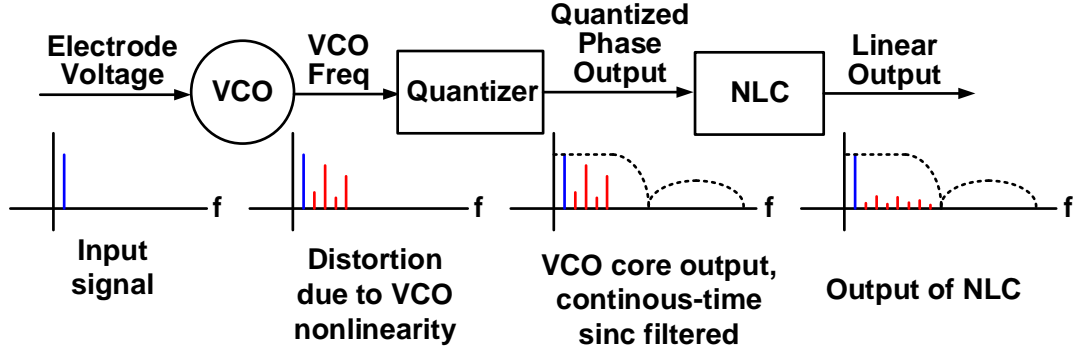


Figure 2.33: Spectrum-domain inspection of the VCO based front-end.

1. The windowed integration during the quantization of phase also introduces a sinc filter. Although this sinc filter suppresses noise aliasing and eliminates the need for an explicit anti-aliasing filter, it attenuates any signal at a frequency comparable to its notch frequency. In addition, the NLC requires a flat response for the LFP signal and its main distortion terms to achieve high correction accuracy, although the sinc filter is linear phase, and could be compensated exactly. Consequently, a sampling rate of approximately several kHz is needed to make sure that distortion terms, up to the 5th harmonic in this case, are lower than the notching frequency.
2. The downstream adaptive-stimulation-artifact-rejection (ASAR) algorithm prefers a finer time step for better rejection results [29]. It is desirable to have a sampling rate that is no less than $5 - 6 kHz$.

The implemented sampling rate is $6 kHz$ with a system clock of $12 MHz$ derived from a crystal oscillator and a division ratio of $\sim 2^{11}$.

The front-end is implemented with 40nm CMOS technology. The VCO and the quantizer are combined together as one unit, while the NLC block is combined with the SPI interface and the controller in the layout. The VCO is supplied with $1.2 V$ for sufficient headroom, while the quantizer works at $0.6 V$ to interface with the RO outputs. The quantizer output is then level-shifted to $\sim 1.0 V$ for the NLC and the following processing. The micrograph of the front-end, shown in Fig 2.34, does not include the NLC block. The VCO is laid out manually,

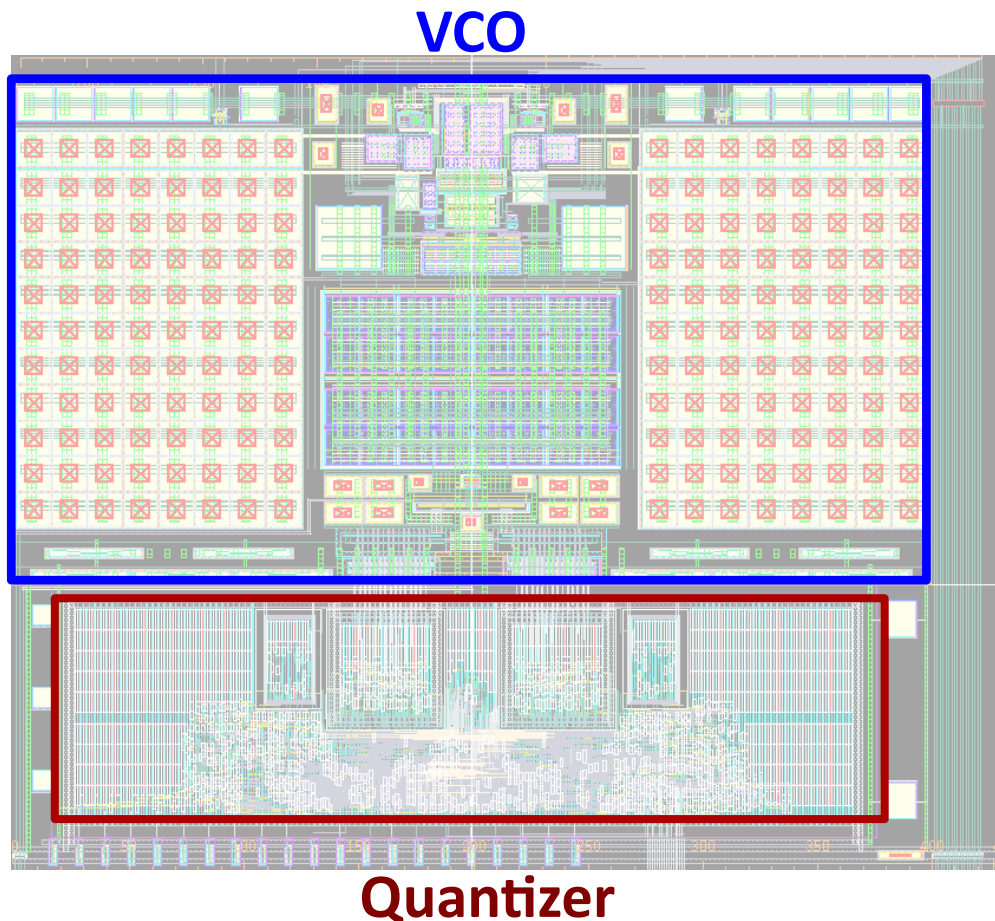


Figure 2.34: Front-end silicon micrograph.

while the quantizer is implemented in digital design flow. The buffers for the control and the output data are placed around the periphery. Its dimension is approximately $0.38\text{ mm} \times 0.31\text{ mm}$, while the NLC area in this implementation is estimated to be $\sim 0.005\text{ mm}^2$.

The area of the VCO is dominated by the HPF capacitors on the left and right side. The capacitor in the 1st stage of HPF is $\sim 8\text{ pF}$, while the capacitor in the 2nd stage is $\sim 100\text{ pF}$. The prototype aims to guarantee performance first; as a result, the capacitance is not aggressively reduced, although this is theoretically feasible.

Additional care is needed in the design flow of the quantizer. The digital design tools tend to insert buffers for proper driving strength or balancing of the clock delay, which can cause timing issues at the asynchronous interface. In particular, the RO buffer outputs should be directly fed to the phase decoder, to avoid unequal delays between the outputs. Additionally,

the delay line that generates the multiple latching signals, *Count1*, *Count2* and *Count3*, are also constrained so that the tool does not insert undesired buffers/inverters.

CHAPTER 3

Front-End Measurement Results

The implemented front-end is measured with the bench-top test set-up to fully evaluate its performance against the specifications in Table 2.1. The test board, shown in Fig 3.1, is composed of four parts: (a) the regulators that generate the supplies for the board components, as well as chip supply voltages (1.8 V, 1.2 V, 1.0 V, 0.6 V) for bypassing chip internal supplies if necessary, have been included on the right side; (b) the input conditioning circuit, where low-noise OpAmps convert a single-ended input into differential signals, and filter out-of-band signal/noise, are shown on the top and bottom side; (c) the digital interface on the left side serves to provide a variety of communication interfaces, from pattern generator to NI PXI digital IO to FPGA control, for different test/debug needs; (d) socket housing the chip packaged in a 180-pin PGA is located at the center of this test board. In addition to providing a complete range of functional test coverage, this test board also offers debugging options for the chip. The measurement results in this chapter, unless otherwise noted, are obtained when the board is connected to the FPGA and configured under the normal functionality test mode.

3.1 Noise Measurement

The noise of the front-end is measured by recording the output at zero input voltage, and then dividing it by the front-end gain for the input referred RMS noise and the noise power spectrum density (PSD). The input-referred noise PSD is shown in Fig 3.2. At low frequency, the noise is dominated by the HPF noise, which has a $\frac{1}{f^2}$ shape, as shown in equation 2.12. The flicker noise is not observable and the noise PSD directly transfers from the $\frac{1}{f^2}$ region to

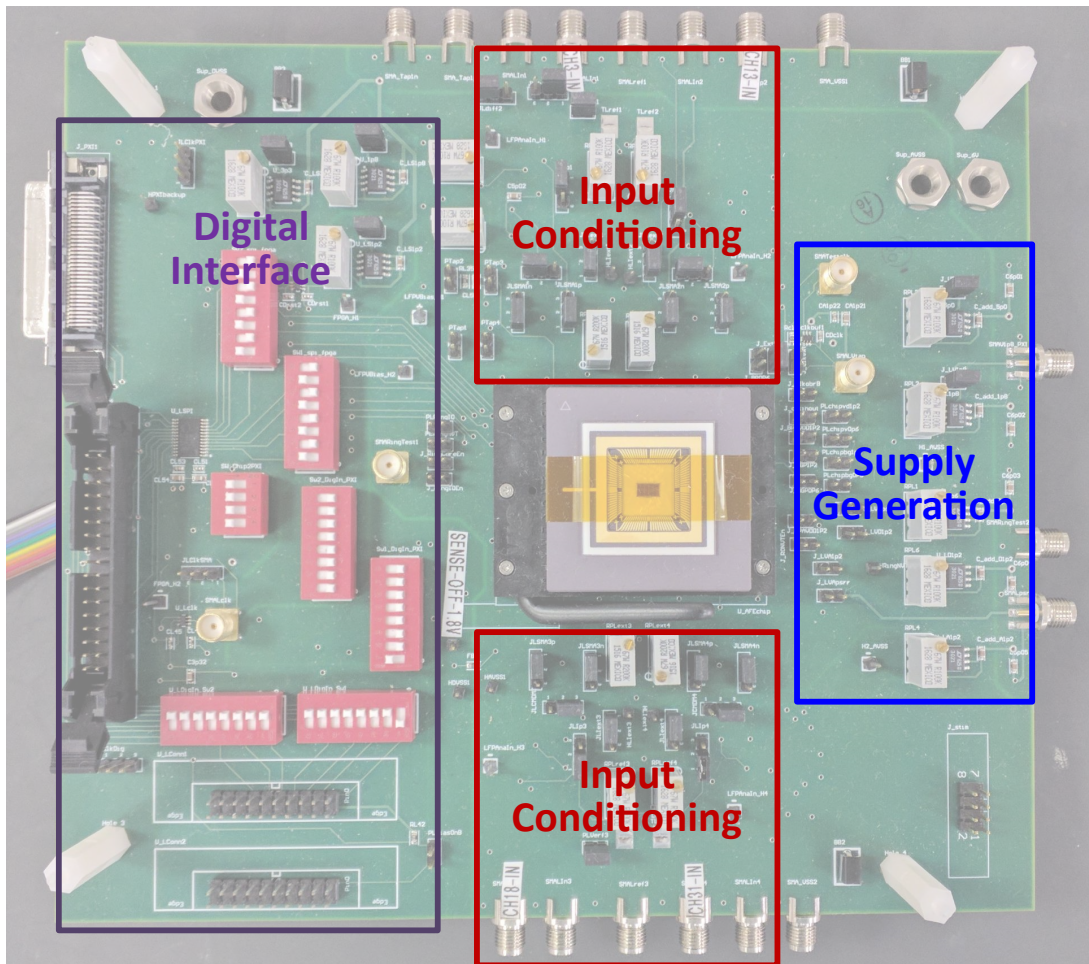


Figure 3.1: Test board for the front-end.

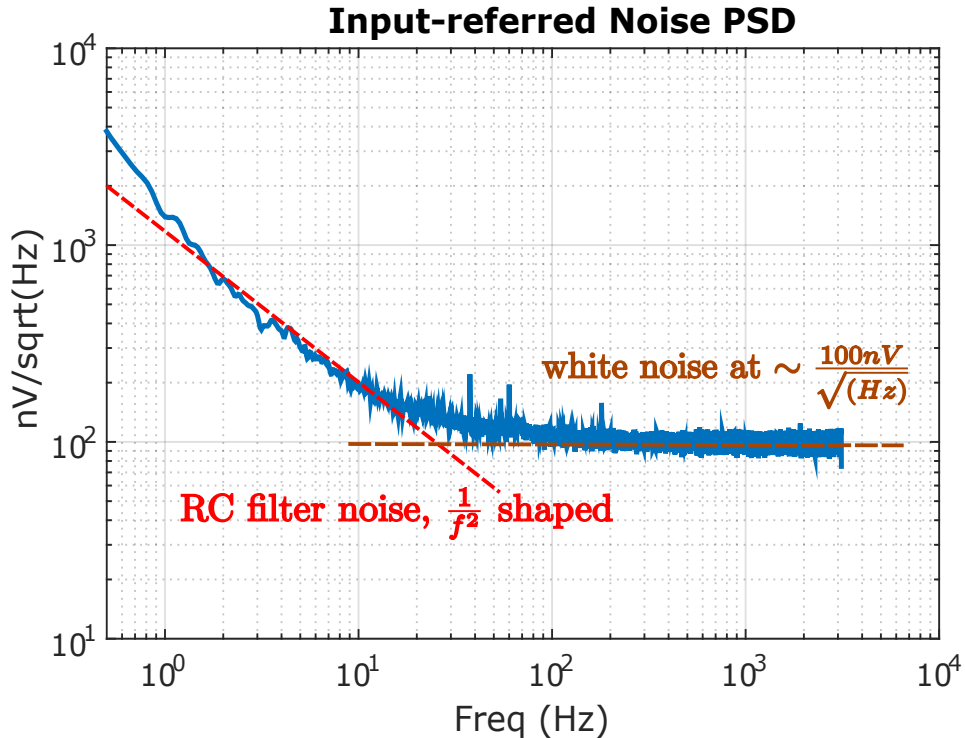


Figure 3.2: The input-referred noise PSD of the front-end.

the white-noise region, due to rigorous design efforts in the diff-pair sizing and the chopping inside VCO. The white noise derives from the thermal noise of both the diff pair and the inverter stages in the VCO, given that the quantization noise is sufficiently low with high K_{VCO} . The diff pair contributes approximately one third of the white noise, while the inverter stages contribute the rest of the white noise.

The integrated RMS noise is $< 3 \mu V$ for the LFP band. Table 3.1 shows the measurement results of the RMS noise over different iterations of the chip with several techniques that were discussed in the previous chapter: the chopping inside VCO decreases the contribution from the RO flicker noise; the RO sub-quantizer eliminates the full-cycle/half-cycle glitches during digitization, and hence reduces the noise floor; the multi-rate HPFs at the input push the corner lower for less filter-noise contribution. The combined effort has led to a more than 6 dB reduction in noise power, leading to better sensitivity of the front-end.

The front-end noise performance degrades at a large input signal compared to the zero-input scenario. The measurement shows that the input-referred RMS noise at $20 mV_{pp}$ is

Table 3.1: Front-end input-referred RMS noise performance.

Iteration	RAM P1	RAM P1	SUBNETS P1	SUBNETS P2
Setting	w/o chopping	w chopping	glitch reduction	multi-rate HPF
Noise (μV)	6.4	5.2	4	2-3

approximately $3 \mu V$, yet increases to approximately $7 \mu V$ at $100 mV_{pp}$. This degradation is because of the noise from tail current bias and reference-current generation circuits. At zero input, the noise current generated by these circuits cancel each other out at two diff-pair branches, because they are in common mode, and do not propagate to the output. However, the diff-pair is unbalanced with a large input, and thus a significant portion of this noise current leaks to the output. In particular, the flicker noise increases dramatically in such scenario. In spite of this noise performance degradation, the system performance of the front-end is not significantly jeopardized; mainly because the stimulation is injected as bursts of pulses and for most of the time, the front-end receives an input at a mV level. Additional flicker-noise reduction techniques can be applied to the bias and the reference if a consistent noise performance over the input range is required (e.g. the proposed front-end is adopted as a stand-alone neural-sensing IP).

3.2 Linearity Measurement

The front-end linearity performance constitutes a critical difference from the prior art, and its significant improvement enables the possibility of concurrent sensing and stimulation. To evaluate this, a single-tone test is performed across the LFP band, and a modified two-tone test is designed to mimic targeted applications. A temperature chamber was used to test the stability of front-end linearity under varying temperatures.

3.2.1 Single-Tone and Two-Tone Test

The single-tone test inspects the output spectrum with a sine-wave input. To validate the effectiveness of the NLC implementation, the outputs without and with NLC are both saved,

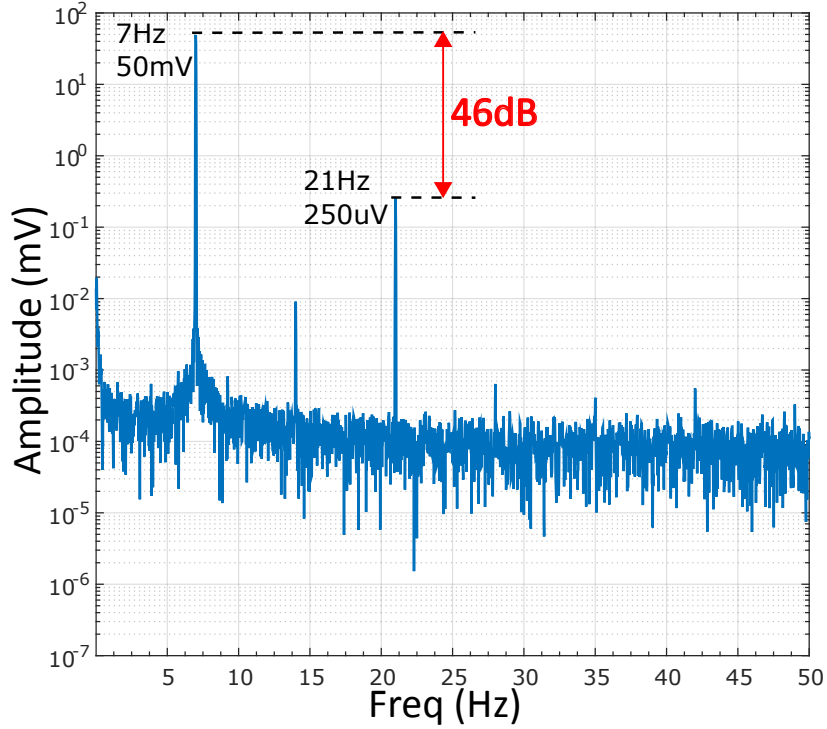


Figure 3.3: Single-tone test for 7 Hz , 100 mV_{pp} input, without NLC.

and the spectrums are compared with each other. The test has swept frequencies across the LFP band with a *fixed* NLC coefficient set. The results for two representative frequencies, 7 Hz and 203 Hz , are shown in this section. The input for both frequencies has a full swing of 50 mV . The output spectrum for 7 Hz input without NLC is shown in Fig 3.3. The dominant distortion term is the 3rd harmonic with an amplitude of 250 μV and equivalently $HD_3 = -46 \text{ dBc}$. The 2nd harmonic is approximately 10 μV , and all harmonics beyond the 3rd harmonic is below 1 μV . In contrast, the output spectrum with NLC is shown in Fig 3.4, where the highest distortion term is still the 3rd harmonic but greatly suppressed to -87 dBc .

The output spectrum for 203 Hz without NLC is shown in Fig 3.5. Similar to the 7 Hz case, the dominant distortion term at the output is the 3rd harmonic $HD_3 = -46 \text{ dBc}$. The 2nd harmonic is approximately 10 μV , again. With NLC activated, the highest distortion, the 3rd harmonic, is successfully suppressed to -80 dBc , as shown in Fig 3.6.

The similar levels of the distortion term for 7 Hz input and 203 Hz input reaffirm that

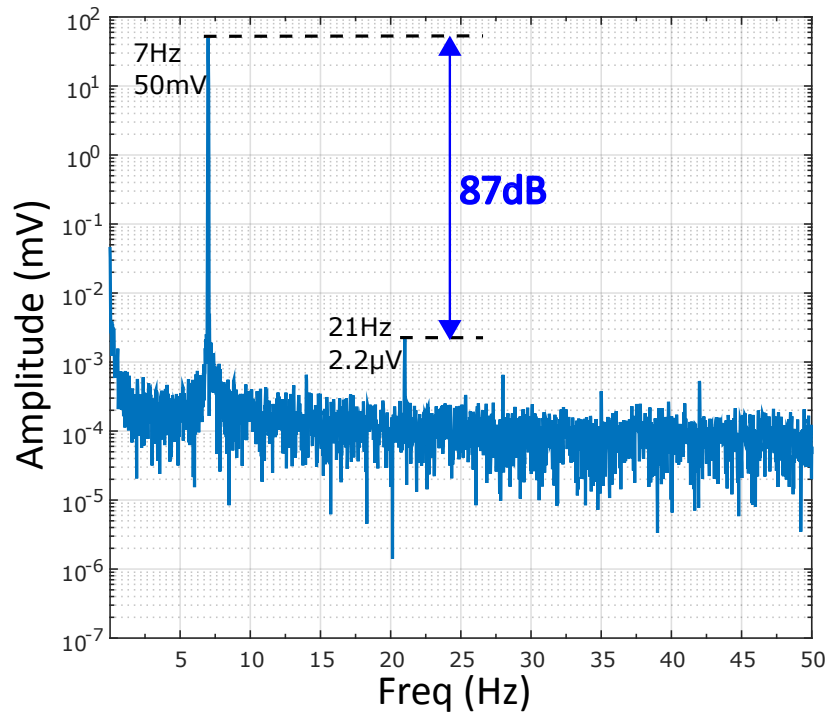


Figure 3.4: Single-tone test for 7 Hz, 100 mV_{pp} input, with NLC.

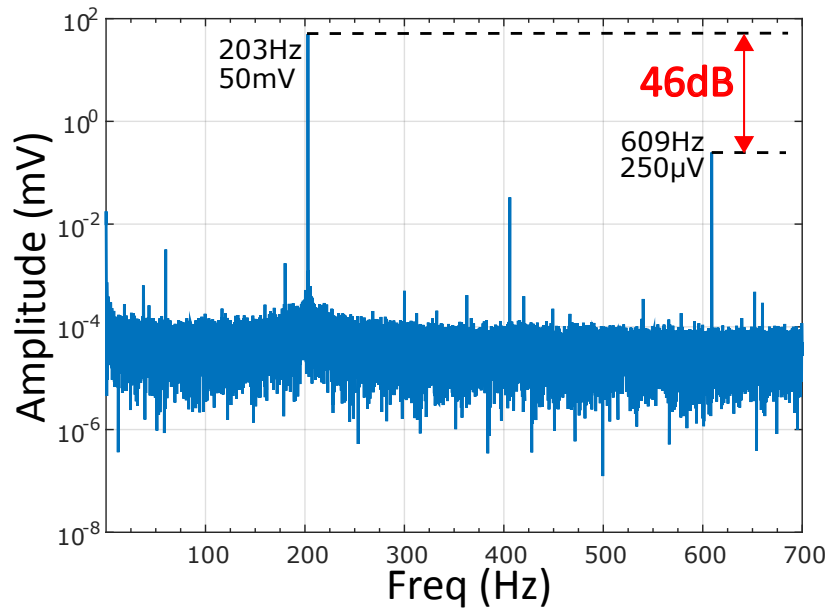


Figure 3.5: Single-tone test for 203 Hz, 50 mV_p input, without NLC.

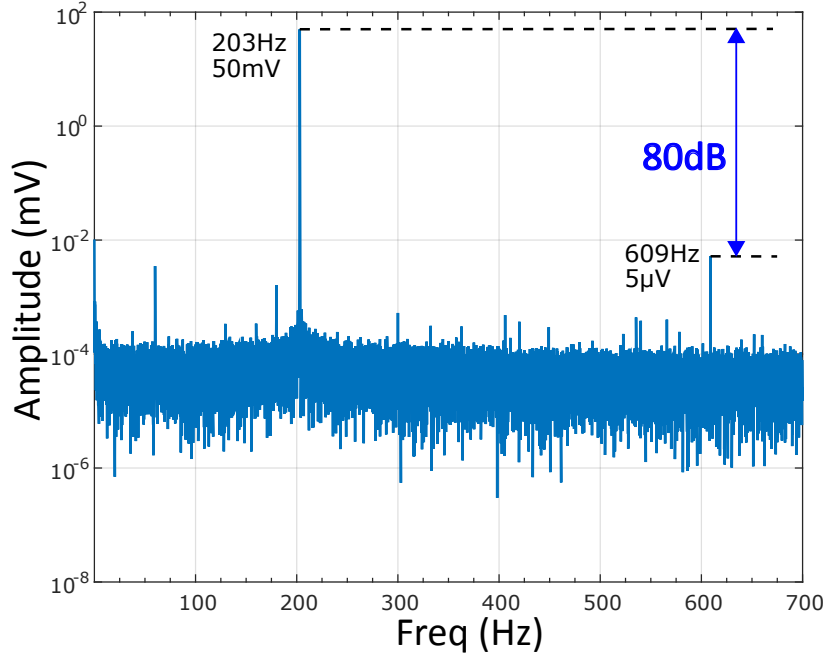


Figure 3.6: Single-tone test for 203 Hz , 50 mV_p input, with NLC.

the major VCO nonlinearity is static or memory-less. The difference in the levels of the residue terms post-NLC can be due to the frequency-dependent component from C_{gs} as discussed in section 2.7, or the slight attenuation of the high-frequency harmonics from the inherent sinc filtering in the quantizer operation.

The two-tone test input is modified as one full-swing sine wave and another small sine wave at nearby frequencies, instead of two equal-amplitude sine waves. This modification aims to emulate the practical scenario of the LFP signal with the stimulation artifact. An example of front-end output spectrum without NLC for the modified two-tone test is shown in Fig 3.7. One sine wave input is at 103 Hz with 50 mV amplitude and the other sine wave is at 93 Hz with 10 mV amplitude. The output IM3 terms are as high as 160 μV . For comparison, the front-end output spectrum with NLC for the same inputs shows that all the IM terms and the harmonic terms are below 3 μV , as in Fig 3.8, which is approximately 40 dB improvement.

All single-tone test results and two-tone test results were obtained with the same set of NLC coefficients. This demonstrates that NLC is able to improve system linearity across the

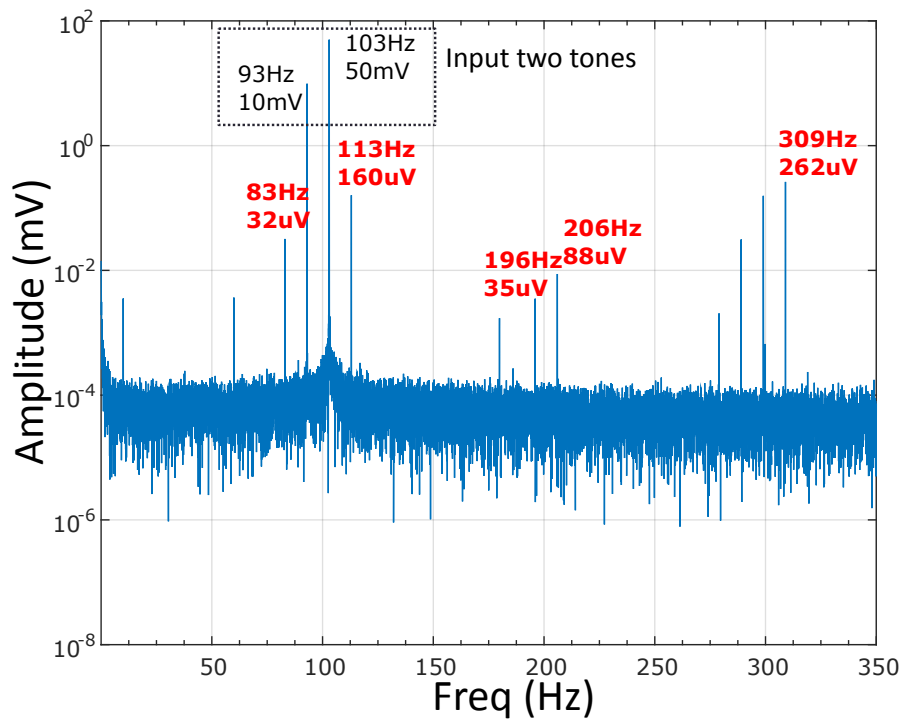


Figure 3.7: Two-tone test for input of 50 mV_p at 103 Hz and 10 mV_p at 90 Hz , without NLC.

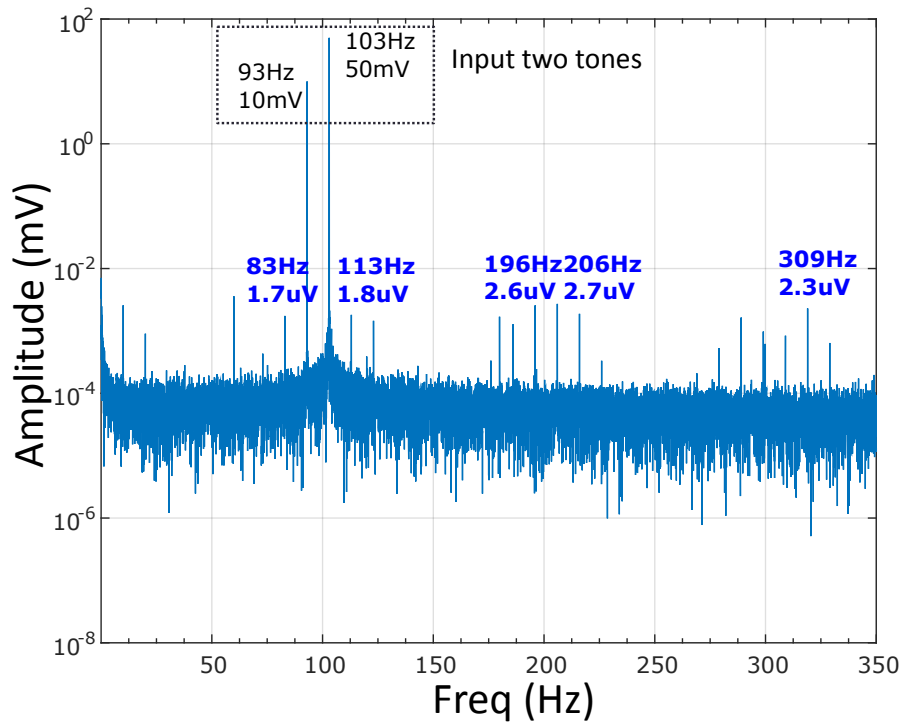


Figure 3.8: Two-tone test for input of 50 mV_p at 103 Hz and 10 mV_p at 90 Hz , with NLC.

entire LFP band. In order to verify the stability of linearity performance, within the implant temperature range, additional verification is performed; these measurements are presented in the next section.

3.2.2 Linearity Stability Under Temperature Variation

We measure the linearity stability of the front-end under temperature variation, by placing the chip inside of a temperature chamber, and then inspecting the output spectrum for a full-swing input signal across different temperatures (i.e., within the implant temperature range). The system set-up is shown in Fig 3.9, where the test board and the FPGA dongle are inside the chamber and are connected to the external PC terminal. Due to the test constraint, we use the NLC coefficients extracted for room temperature ($21 \text{ }^\circ\text{C}$). For input frequencies across the LFP band, the front-end linearity shows similar stability across various temperatures; a representative result is shown in Table 3.2. The distortion is dominated



Figure 3.9: Temperature test for front-end linearity performance.

Table 3.2: Linearity performance result across temperatures.

Temp ($^{\circ}\text{C}$)	-5.3	-0.5	4.9	9.6	14.2	19.2	24.4	29.9	34.9	40.2	45.5
HD3 (dBc)	-62.0	-63.9	-67.1	-72.6	-81.5	-79.0	-80.0	-73.6	-68.8	-68.9	-64.5

by the 3rd order harmonic, and it can be seen that HD3 remains approximately -80 dBc within $\pm 5^{\circ}\text{C}$ away from room temperature, and better than -70 dBc within $\pm 10^{\circ}\text{C}$ away from room temperature. Considering that the implant environment temperature is very stable and $\pm 5^{\circ}\text{C}$ can be regarded as the upper bound from the regulation standard [23], the stability of the linearity is sufficient for implant application requirement. The front-end gain also remains stable across the implant temperature range and only varies within $\pm 0.15\text{ dB}$.

3.3 Power Measurement

The front-end is supplied by internal regulators in the normal mode, as shown in section 4.1. To measure the power consumption from every block, the regulators are bypassed and the supplies are provided from on-board regulators. The test result is shown in Table 3.3.

Table 3.3: Front-end power consumption measurement result.

Current	Analog (1.2V)	Digital (0.6V)	Digital (1.0V)	Overall
Static Current (μA)	1.6	1	≈ 0	2.6
Dynamic Current (μA)	1.4	3.9	1.6	6.9
Sum Current (μA)	3.0	4.9	1.6	9.5

Although the current consumption is approximately $9.5 \mu A$, further effort can effectively bring the power down. For example, the static current in the analog supply (1.2 V) is employed to generate the current reference and the voltage bias; this can be shared between multiple front-ends given more time for the integration effort. The dynamic current from the digital supply (0.6 V) is not fully optimized because a major design effort aims toward resolving the glitches in the quantizer. It is anticipated that the power consumption can be reduced to below $7 \mu A$ for further iterations.

3.4 Front-End Input-Interface Measurement

The front-end's low-frequency response is measured to confirm its DC-offset-rejection functionality. Both cases of enabling the full multi-rate HPF and enabling only the 1st stage are covered. The results are shown in Fig 3.10, clearly illustrating the high-pass characteristic. The response below the corner is at a slope of approximately 20 dB/dec, as expected for a first-order R-C filter. The corner with the enabled multi-rate HPF is approximately 0.05 Hz while the corner with only the 1st stage on is approximately 1 Hz. Therefore, the multi-rate HPF is effective in pushing the corner further lower for our front-end requirements.

DC input impedance or leakage current is also measured to resolve the concern of any parasitic conductance through the HPF capacitor or non-zero current through the ESD diode. The set-up is shown in Fig 3.11. The front-end input is in series with a $270 M\Omega$ resistor, and the voltage across the resistor (V_{res}) is observed via the multimeter. To avoid loading effect, the output terminal of the resistor that is close to the front-end is buffered by a

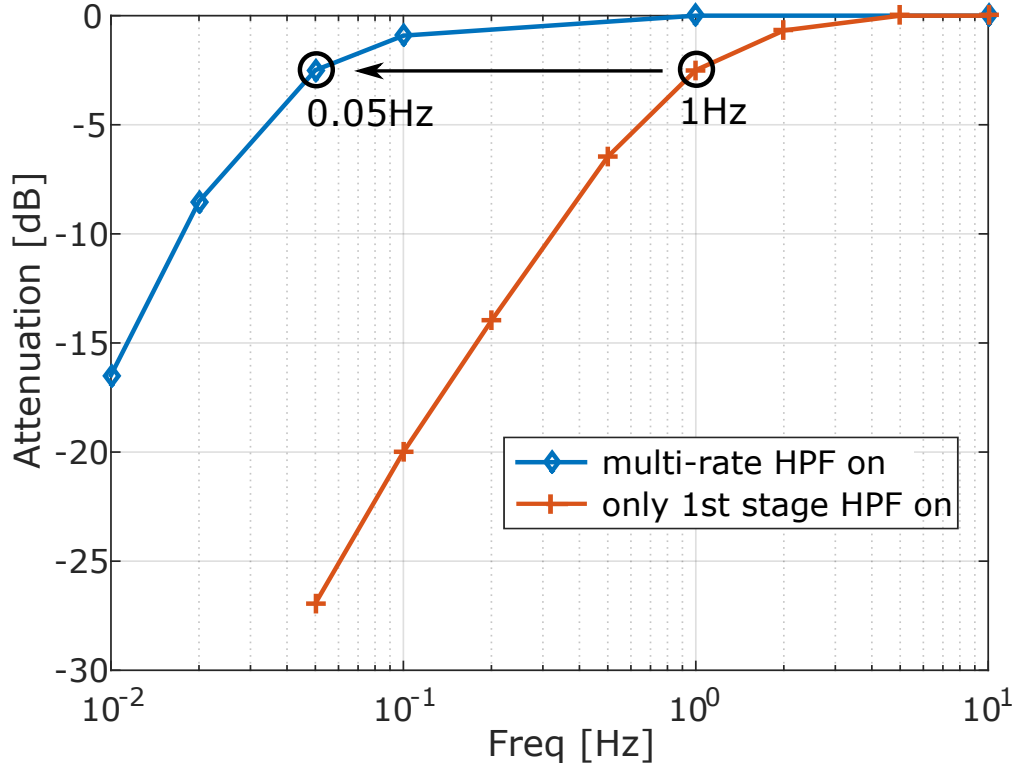


Figure 3.10: Measurement of front-end low frequency response.

low-offset-current OpAmp. The input voltage V_{in} is set to 400 mV for the test. The voltage across the resistor without the chip is measured first and the absolute value is $< 4\text{ mV}$, which can be because of the OpAmp offset current, the resistor noise (due to the multimeter non-zero bandwidth), or the OpAmp noise. With the chip, the voltage difference is $< 10\text{ mV}$ for all five chips being tested. Consequently, the current through the resistor (I_{res}) is less than $\frac{10\text{ mV} + 4\text{ mV}}{270\text{ M}\Omega} \approx 52\text{ pA}$. This sets an upper bound for the leakage current and a lower bound for the DC input impedance. The leakage current should be less than I_{res} , within the range of 100 pA as mentioned in section 2.1. Assuming no leakage current and V_{res} is only due to the finite front-end DC input impedance, the DC input impedance is no less than

$$\frac{400 - 14}{14} \times 270\text{ M}\Omega \approx 7.4\text{ G}\Omega$$

meeting the specification introduced in Table 2.1.

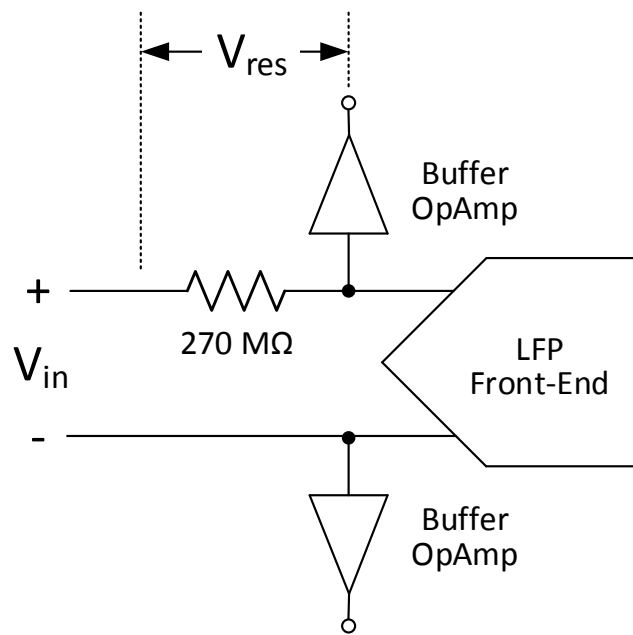


Figure 3.11: Measurement of DC input impedance or leakage current.

CHAPTER 4

Sensing-System Integration

Our front-end circuit is designed to sense brain signals with concurrent stimulation, which is needed for a multi-channel closed-loop neuromodulation implant. Therefore, it is a critical effort to integrate multiple front-ends into a chip for the implant requirement, and verify its function and performance for concurrent sensing and stimulation.

This sensing-chip design and system-integration trial has evolved over several tape-out iterations, from 4-channel front-end-only prototype to 32/64-channel full sensing system, and spans two DARPA projects (RAM [46] and SUBNETS [47]), as presented in Fig 4.1. This chapter focuses on the latest iteration, i.e. the chip on the right-most side. The multi-channel sensing system that is implemented in this iteration is demonstrated. In addition, the miniaturized neuromodulation unit housing both the sensing and stimulation chip as an implantable system is shown. We then present *in-vitro* concurrent stimulation and sensing measurement result to verify the system design concept.

4.1 32/64-Channel Neural-Sensing Chip

Since our system is intended for an invasive implant, it is necessary to miniaturize the entire system assembly to facilitate surgical implantation and reduce tissue damage. In order to realize this miniaturization, the chip must be maximally self-contained, i.e. with minimal external components. While the power supply and the data communication have to come off-chip, some other components can be integrated into the same chip as follows:

- The system control and the data packetization are tightly coupled with the front-ends. An on-chip control allows internal sequencing of the front-end sampling and

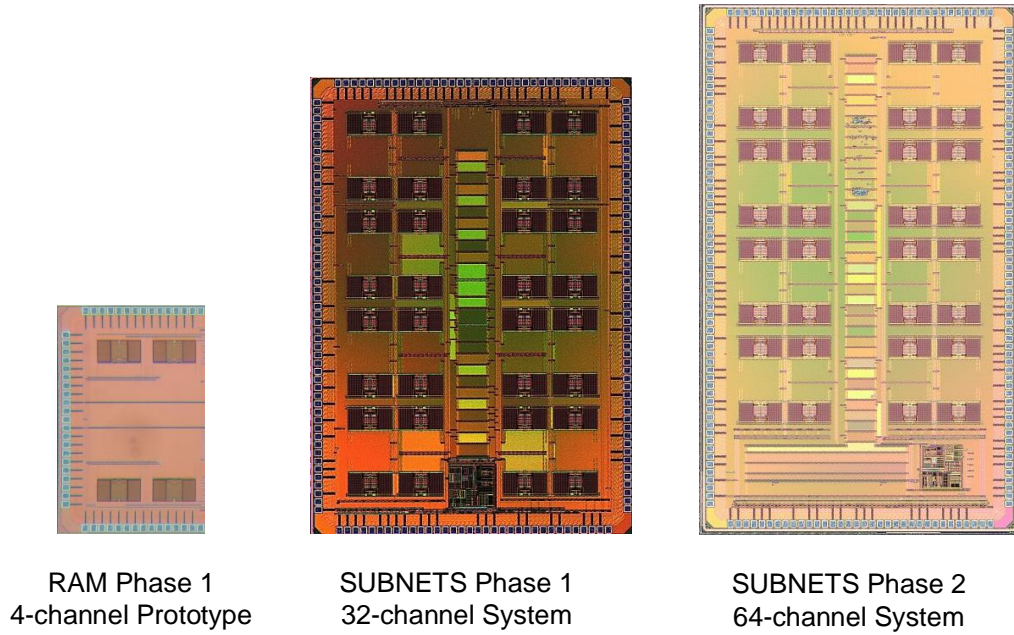


Figure 4.1: Sensing chip tape-out iterations.

processing, and thus reduces the chip pin count. The data packetization can pack the multiple channel data into one frame and add requisite information for diagnosis and information logging, such as time-stamps or specific triggers. Then, the packets are sent under the SPI protocol to downstream modules in the system.

- Voltage regulators are desired in the chip. The front-end requires several supply voltages for various blocks. While it is possible to have regulated supplies externally, their rejection of the interferences is limited by the assembly parasitics, such as coupling to bond-wires, and thus inferior compared to integrated regulators. On-chip regulators provide a more stable environment for the front-end and a higher power-supply rejection ratio (PSRR).
- On-chip clock generation is preferred for power/area saving. The external crystal oscillator (XO) is not only bulky but also power-consuming (at a mW level). In contrast, on-chip clock generation for system timing only consumes power at a μW level.

- Inclusion of ASAR engines in the sensing chip can provide data-speed and power benefits to the neuromodulation system. The ASAR engine outputs have a much reduced dynamic range (less number of bits per sample) and allow a reduced sampling rate to the Nyquist rate of the input signal. Consequently, the data rate can be easily reduced by more than $10\times$. The reduced data rate helps cut down the additional power consumption of the following signal processing or possible wireless data communication [48].

The integration is demonstrated in Fig 4.2. The external 1.8V supply is regulated to several voltages for various blocks. The system clock is provided by an oscillator with an external crystal. The sensing core, which contains 32 front-ends, provides capabilities of either single-ended recordings for 32 electrodes, or differential recordings for 64 electrodes. The front-end outputs are sent to interleaved NLC blocks. The data after NLC is sent to ASAR engines, packetized and the sent to the output via an SPI interface.

The chip is built with the system test in mind. The LDOs can all be bypassed by external supplies, which allows direct power-consumption measurement. The XO can also be overridden by an external clock source for system timing. For a long signal-processing chain, from the VCO, through to quantizer and NLC, to ASAR, every block should be able to be tested individually prior to the overall system verification. Therefore, the NLC and ASAR both have bypass switches from system configurations. In addition, NLC inputs can be routed from an external digital stream instead of front-ends. All of these options allow sweeping the test patterns of all relevant combinations, such as “NLC + ASAR” and “sensing core + NLC + ASAR”, for maximum flexibility in system diagnosis and debugging.

The chip layout is shown in Fig 4.3 and the floor-planning is highlighted. The 32 front-ends occupy the majority of the area, while digital blocks, NLC/SPI/controller/ASAR, are placed and routed as one entire block. The LDOs and XO are located at the top right corner. It also includes the function of power-on-reset (PoR), which resets the system status to default until fully powered-up. The PoR avoids possible initial erratic output until the supplies are stable and data communication is set up. This chip is a system-on-chip (SoC)

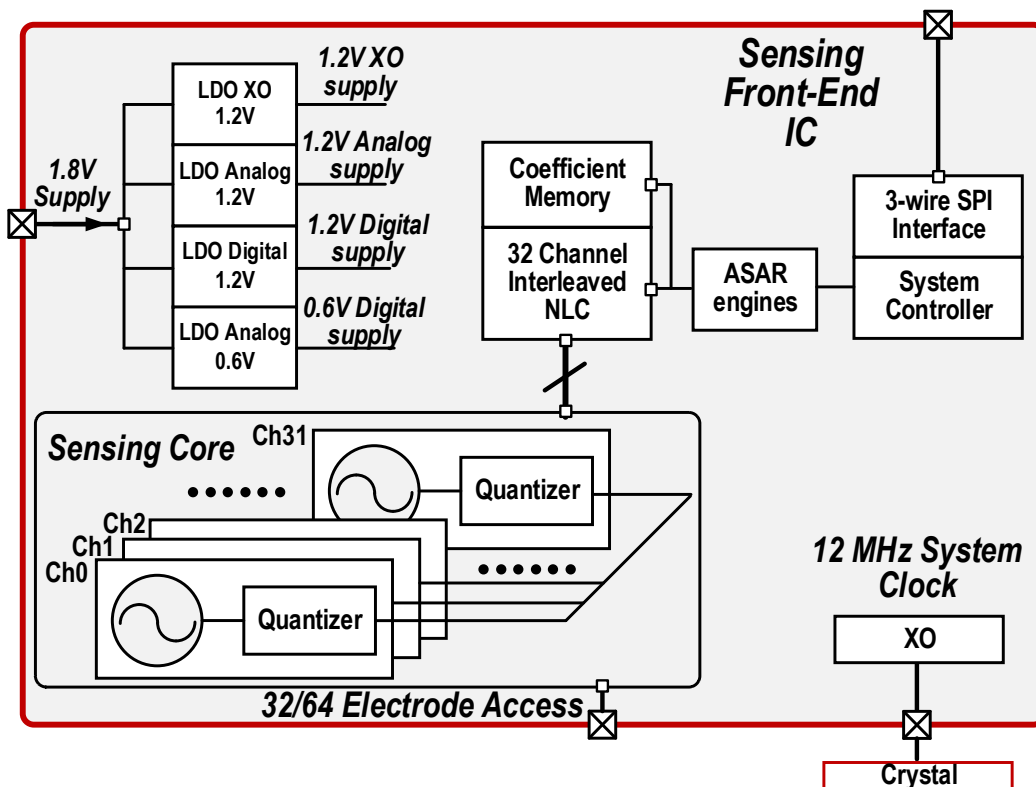


Figure 4.2: Schematic for the sensing-system chip.

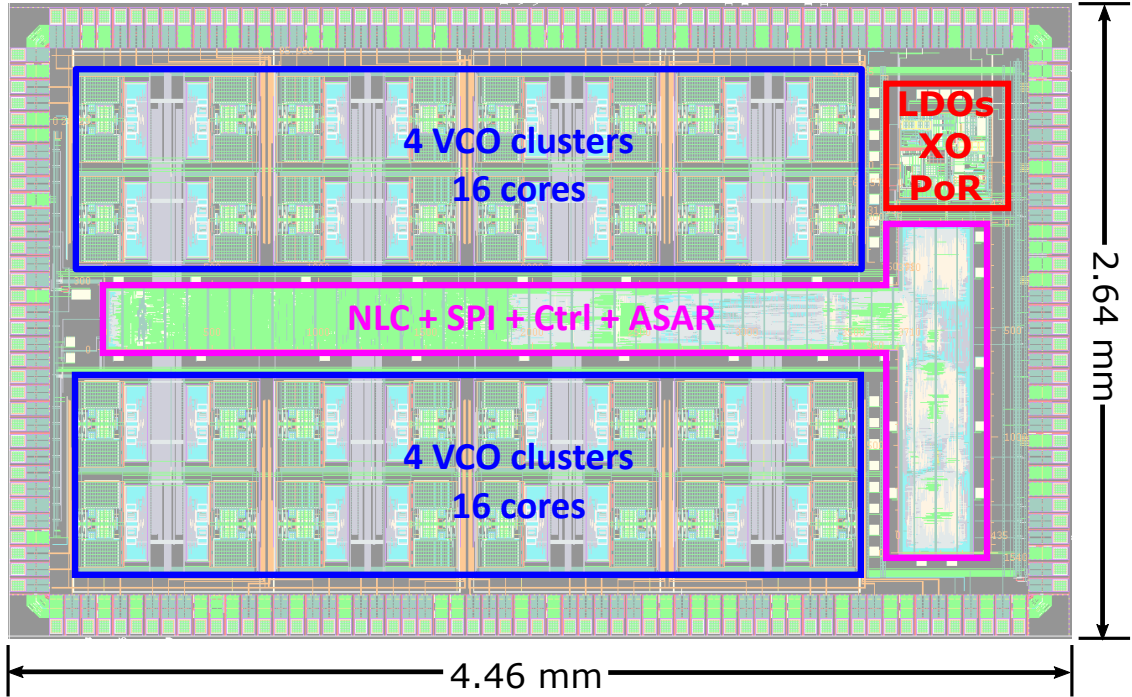


Figure 4.3: Layout for the sensing-system chip.

and requires team effort for completion. The major contribution of the presented research is leading the integration, as well as the placement and routing at the top level. The major credit for the center digital blocks belongs to Vahagn Hokyikyan and Sina Basir-Kazeruni, while the credit for the LDOs and XO belongs to Hariprasad Chandrakumar.

The 32 front-ends are grouped in eight clusters, as shown in Fig 4.3. The detailed layout of one cluster is presented in Fig 4.4. The digital IO signals are included together, buffered and routed in the vertical way, while the power supplies are routed in the horizontal way with top-level metals. This cluster hierarchy simplifies the digital interface, as well as the power routing for all front-ends.

4.2 Miniaturized Neuromodulation Unit

The multiple-channel sensing chip is housed together with the stimulation chip as a neuromodulation (NM) unit. This NM unit serves as the electronics directly interfacing with the probes in an envisioned closed-loop neuromodulation system, as shown in Fig 4.5. NM

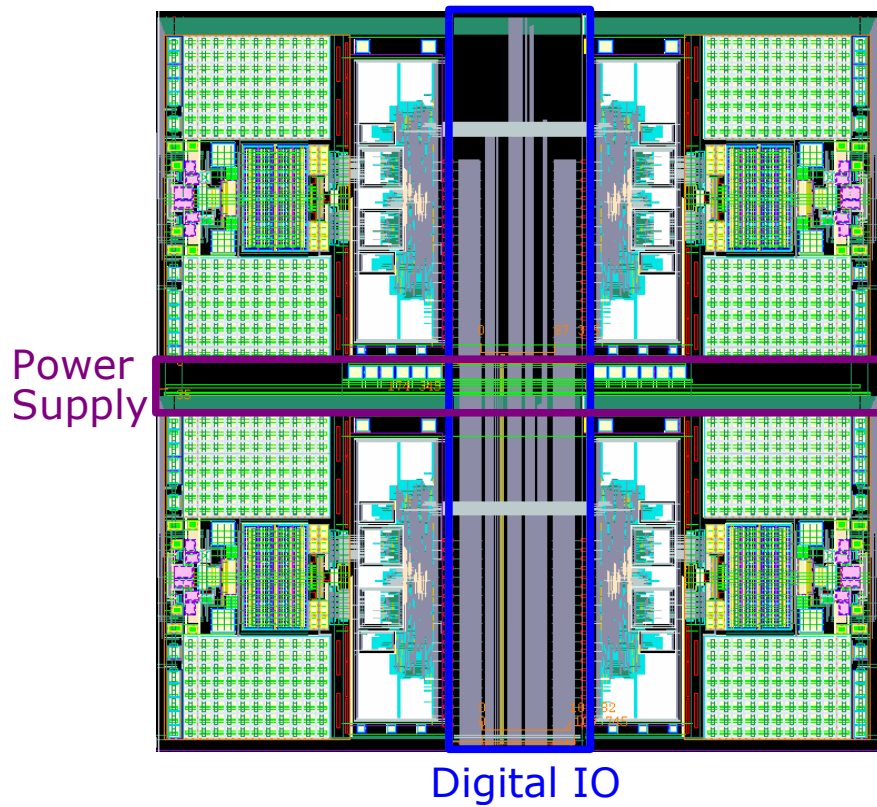


Figure 4.4: Front-end cluster.

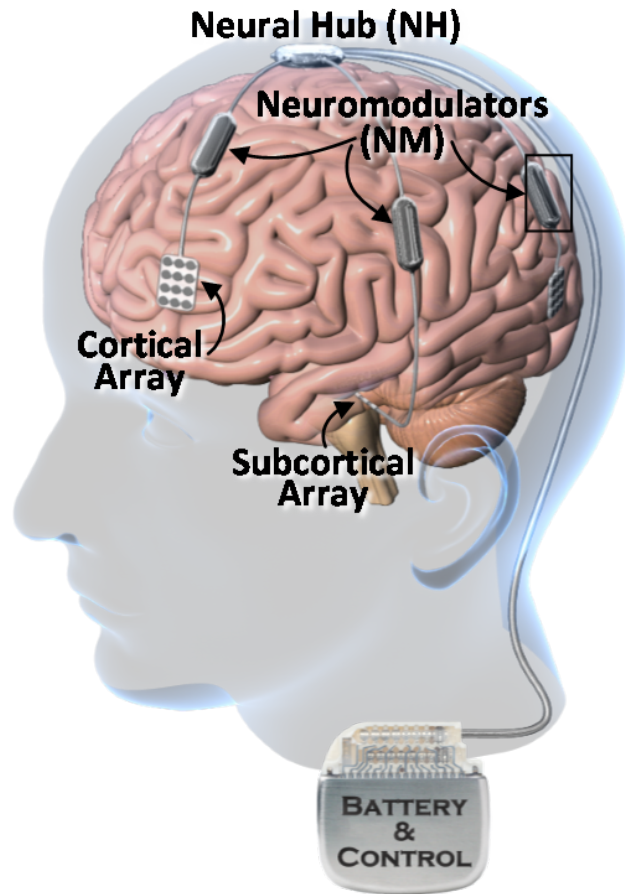


Figure 4.5: Proposed neuromodulation system.

units are placed inside of the skull, and the bone is intentionally thinned at the locations where these units are accommodated. A chest unit receives power and control commands wirelessly from an external module outside the body, and a neural hub routes the power and the data/control between this chest unit and all NMs. It is advantageous to place the NM unit close to the target region, which reduces the voltage drop across the probe during stimulation, and thus allows a more power-efficient stimulator design. Moreover, the locality also reduces interference on the sensed signals. Nevertheless, this requires the NM unit to be at a low volume in order to minimize surgery operation difficulty and invasiveness.

The fabricated and assembled NM unit is shown in Fig 4.6. It has the stimulation chip, the sensing chip and the caps on the top side, while provides all of the necessary probe/cable contacts on the bottom side. The stimulation chip is mainly designed by Dejan Rozgić,

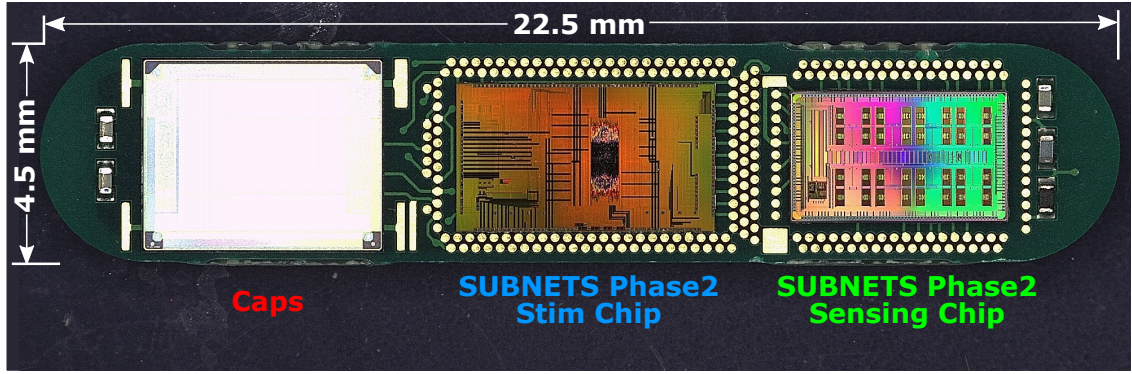


Figure 4.6: Fabricated and assembled NM unit.

Vahagn Hokyikyan, Ipei Akita and Sina Basir-Kazeruni. The caps on the left side are used for charge pumps in the stimulation chip. The crystal, as presented in Fig 4.2 for the sensing chip clock generation, sits on top of the sensing chip. The overall NM volume is $22.5 \text{ mm} \times 4.5 \text{ mm} \times 2 \text{ mm}$. It will be encapsulated within a titanium shell as an implant and inserted into the desired place via a guide tube during the surgery.

The power chain and the communication protocol are planned for the system initialization. The stimulation chip receives AC power (to avoid probe material polarization) from the neural hub, and then rectifies it to a DC voltage approximately $2V+$ for pulse generation. Meanwhile, a regulator at the stimulation chip provides $1.8V$ for the sensing chip supply. The communication protocol proceeds in the reverse way. When the sensing chip is powered up, it receives the command via the SPI interface and then streams necessary information down to the stimulation chip. Meanwhile, the sensing chip starts to transmit the sensed data back to the neural hub. The protocol helps to set the system to the default state initially, and ensures that the communication can only start when all aspects are powered up.

4.3 Concurrent Sensing and Stimulation Measurement

The concurrent-sensing-and-stimulation test was performed to verify the function of the NM unit. The diagram of the test bench is shown in Fig 4.7. The electrodes are submerged into the PBS solution inside of the beaker. A signal generator injects a small-amplitude signal

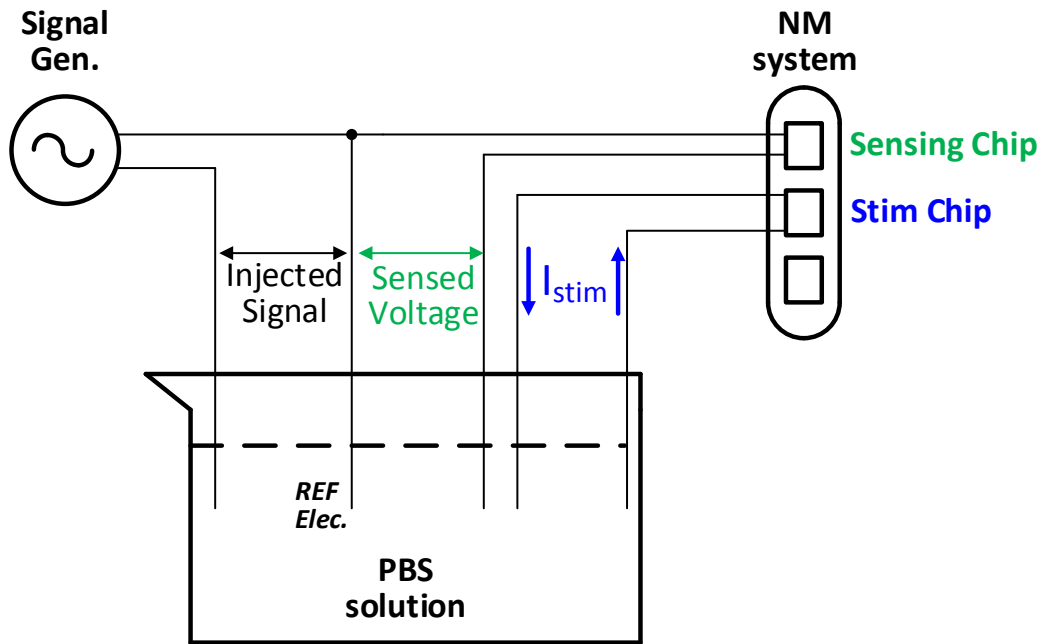


Figure 4.7: Diagram of the test on concurrent sensing and stimulation.

into the solution to mimic the neural signal. Meanwhile, the stimulation chip sends differential current pulses via the designated electrodes as neuromodulation therapy. Concurrently, the sensing chip records the voltage across the reference electrode and a non-stimulating electrode. As we have stated in chapter 1, the front-end should not be saturated by stimulation artifacts, and the recorded output should not contain distortion terms for the neural-signal component.

The test set-up is shown in Fig 4.8. The NM unit is housed in the pogo socket of the testbed. The data and control are streamed to/from a FPGA dongle, which connects to the PC via USB. The oscilloscope can evaluate the stimulation current with a small detection resistor on the testbed. The power supplies provide DC current for the testbed components, and AC power to the NM unit. The right-most signal generator injects the small-amplitude signal. A more detailed explanation of the set-up can be found in the following link: <https://youtu.be/rDq5y2qej4I>.

A customized GUI, as shown in Fig 4.9, is also developed to facilitate the control of the

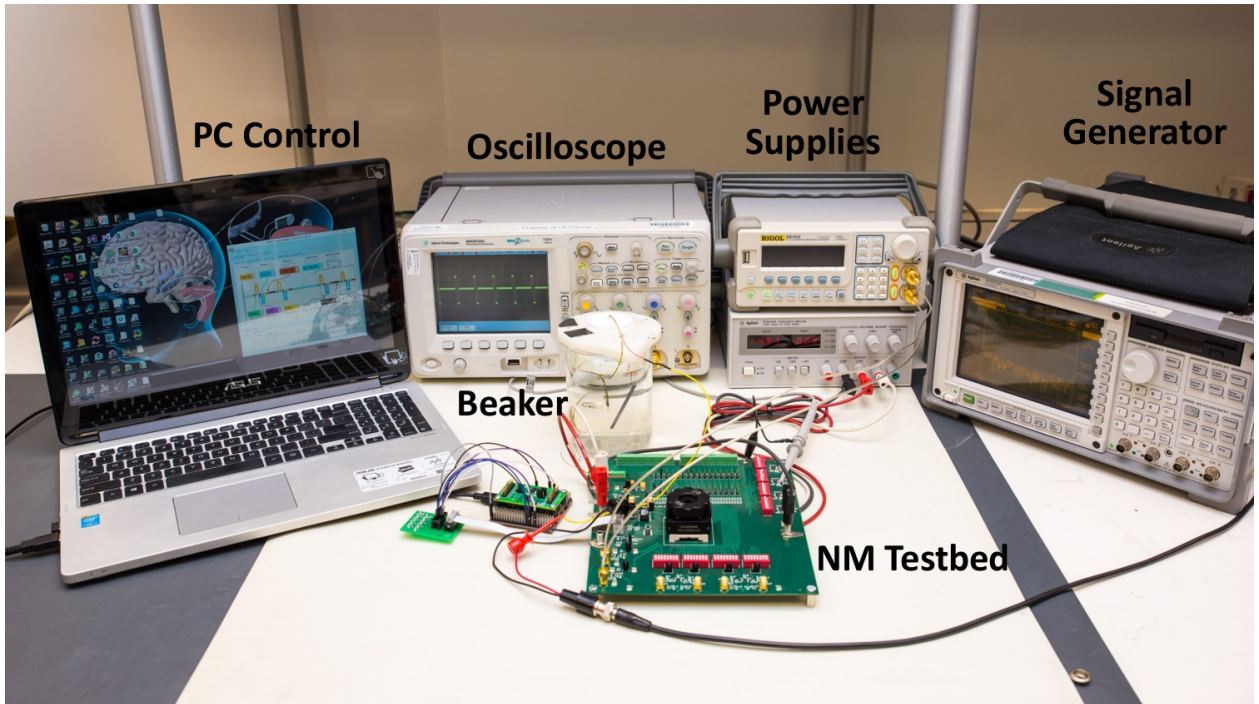


Figure 4.8: Test set-up for concurrent sensing and stimulation [11].

stimulation/sensing chip in the test, as well as data streaming/logging. It allows on-site tuning of the supply current, sampling rate, sampling-window timing, and power-gating of every individual front-end. This GUI interface renders great flexibility and convenience for the test.

For better visibility in the waveform and the spectrum plot, in the concurrent sensing and stimulation test, the injected signal is a single tone within the LFP band. For the stimulation settings, cases of various stimulation amplitudes (up to two stimulation engines with 3 mA each) and stimulation-duration settings are tested. In addition, the frequency of the injected single tone is also swept across the LFP band. The test shows that the sensing front-end never saturates for all conditions, while the output shows no distortion terms with the injected single tone. An example is presented in Fig 4.10, where a 7 Hz signal is injected and the stimulation current is at a 2 mA amplitude. The waveform on the top clearly shows a superposition of the stimulation pulses and the 7 Hz signal, with no observable saturation. The spectrum plot at the bottom exhibits no harmonics for the signal, and the only undesirable tone in the low frequency is 60 Hz coupling from the wall-powered supply.

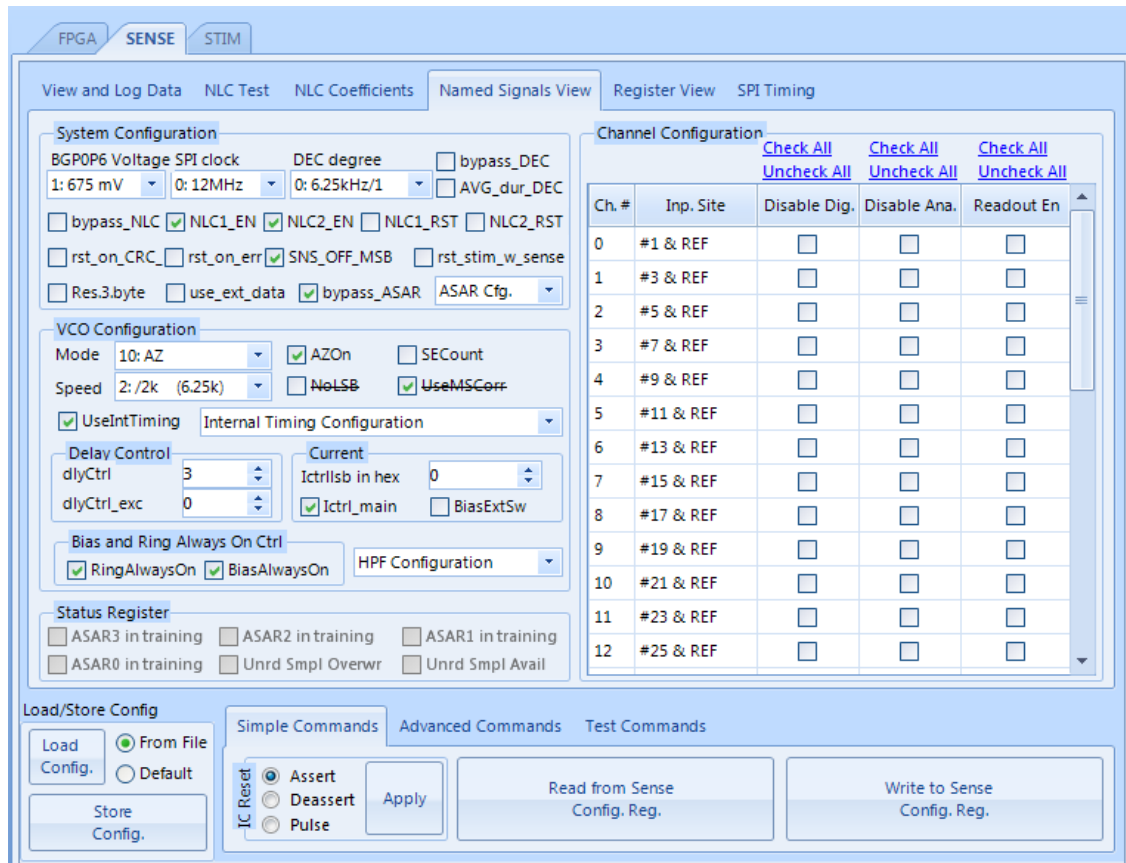


Figure 4.9: The NM unit GUI for the test [12].

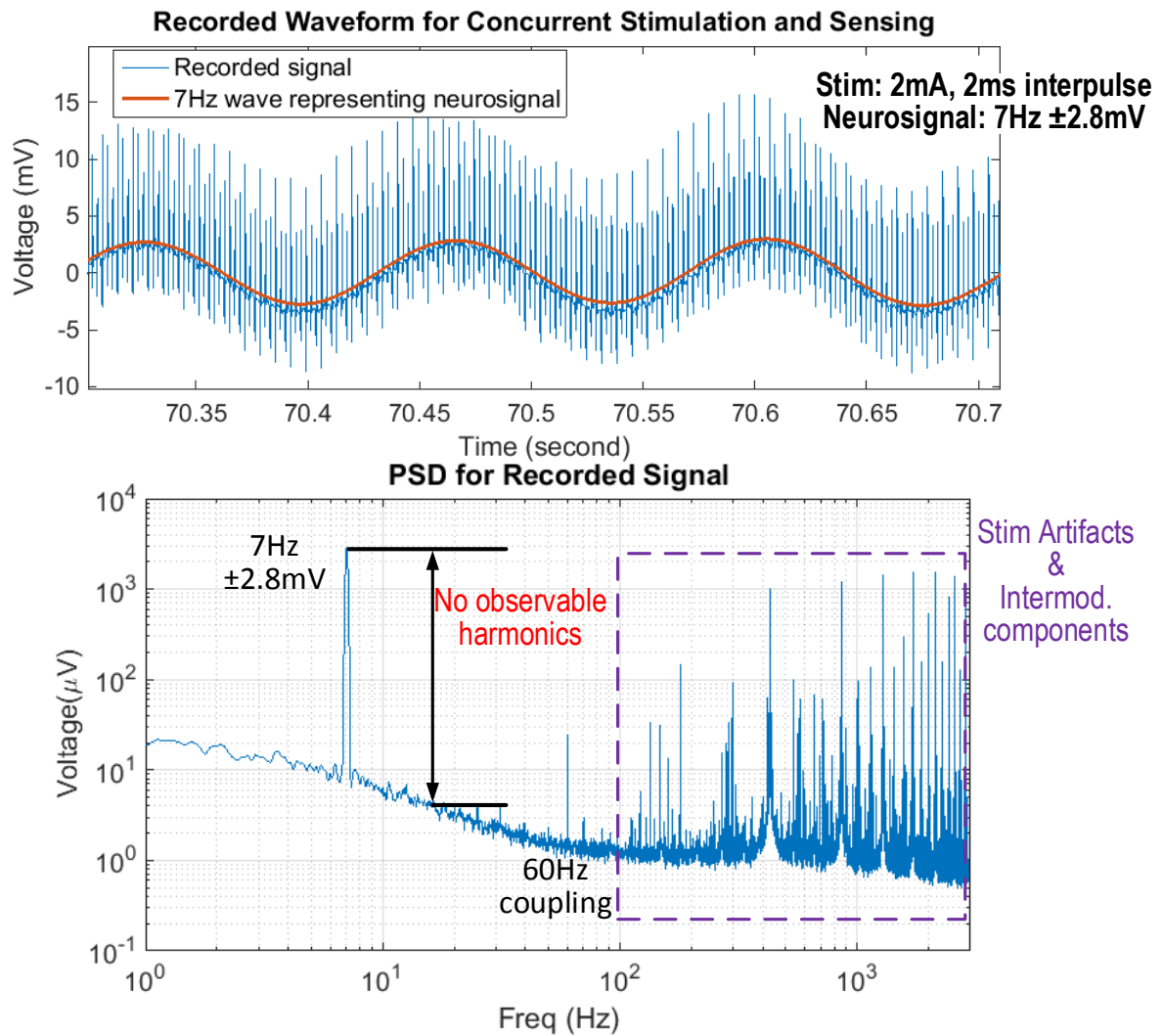


Figure 4.10: Waveform and spectrum plot for the concurrent-sensing-and-stimulation test.

The spurs at the high frequencies are the stimulation artifacts and their intermodulation with 60 Hz harmonics. These spurs can be notched out or rejected by following signal-processing blocks [29].

4.4 Neural-Sensing with ASAR

As mentioned in section 4.1, the integrated sensing chip includes ASAR engines to extract neural signals from stimulation artifacts, and to reduce the communication data rate. This

is not easy to verify in the *in-vitro* concurrent-sensing-and-stimulation environment, due to difficulty in generating local neural signals around every electrode in the beaker. For the set-up in Fig 4.7, the sensed voltage due to the injected signal at different electrode sites would be strongly correlated, and thus makes the validation of our ASAR algorithm [29] a challenge.

To bypass this challenge, ASAR within our sensing system is verified by using a bench-top test set-up with no beaker. The front-ends receive input signals which are generated from a high-precision instrument (NI PXI dynamic signal generator card), by using data from real patient recording. While it is ideal to verify the artifact rejection up to 100 mV_{pp} , the artifact amplitude in the prerecorded data is limited to 29 mV_{pp} . The recorded ASAR achieves an artifact rejection up to 37 dB as reported in the measurement [29].

CHAPTER 5

Conclusion

5.1 Comparison and Research Contributions

The comparison of our front-end with prior art is shown in Table 5.1. The power consumption of the complete front-end is calculated from Table 3.3: $1.2 \times 3.0 + 0.6 \times 4.9 + 1.0 \times 1.6 = 8.14 \mu W$. This is comparable to state-of-the-art front-ends except [2], which only presents the amplifier without ADC. The noise performance of our front-end is also on par with the state-of-the-art. The advantages of our front-end are highlighted. Our front-end increases the input range by $10\times$ as compared to others except the recent work in [5]. Although the work in [5] reports an input range of $\pm 50 mV$, the linearity is limited to 63 dB. This results in an ENOB of 10.2b in [5], which is insufficient for closed-loop neural recording. However, our front-end maintains high linearity even up to $\pm 50 mV$ input-swings, which results in an ENOB of 13b. This ENOB is higher than the current state-of-the-art by at least 2 bits, and is sufficient for closed-loop neural recording. The input impedance is theoretically infinite at DC, and the measured result is no less than $7.4 G\Omega$.

To summarize, the main contributions of this work are as follows:

- This work defines the challenges and specifications of the neural-sensing circuit for a multiple-channel, closed loop neuromodulation system. The limitations of prior front-end implementations are discussed and put in perspective to establish the importance of this research.
- We have explored phase-domain acquisition and processing structure for low-frequency, high-resolution application in general, and brain-sensing circuits in particular. We have

Table 5.1: Comparison with prior art.

	[2] JSSC'07	[49] JSSC'12	[50] JSSC'15	[5] VLSI'17	This work
Topology	CCIA (no ADC)	DiffAmp + ADC	Chop.Amp + ADC	Direct Conv. $\Sigma\Delta$	VCO with NLC
Signals	LFP	Spike/LFP	LFP	LFP	LFP
Technology	0.8 μm	65 nm	65 nm	180 nm	40 nm
Supply	1.8 V	0.5 V	0.5 V	1 V	1.2 V (Analog) 0.6 V (Digital)
Area/ch	1.7 mm^2	0.013 mm^2	0.025 mm^2	N/A	0.135 mm^2
Power/ch	2 μW	5.04 μW	2.3 μW	8 μW	8.2 μW
In-ref. Noise (RMS)	1 μV	4.3 μV	1.3 μV	1.2 μV	2-3 μV
Peak input	5 mV_p	3.5 mV	$\pm 0.5 mV$	$\pm 50 mV$	$\pm 50 mV$
ENOB	11.0	8.0	7.8	10.2	13.0
$Z_{in}(DC)$	8 $M\Omega$	∞	28 $M\Omega$	30 $M\Omega$	∞

presented the VCO-base structure as a viable solution to avoid the signal-gain / input-range trade-off in the voltage domain, in order to achieve the desired high dynamic range.

- This work has presented and implemented a neural-sensing interface design using a multi-rate duty-cycled-resistor based HPF. Through employing a duty-cycling control with the linear passive components, the interface provides a reliable HPF corner of $< 0.1 \text{ Hz}$ and an infinite DC input impedance for device/patient safety.
- We have presented a glitch-free implementation of the quantizer for phase-domain processing. The effects of asynchronicity and voltage misalignment across the analog-digital interface are explained. Associated techniques to eliminate the glitches, i.e. multiple-latching algorithm and robust phase decoder design, are also explained.
- The integration practice of the sensing front-end into a multi-channel sensing system on silicon is discussed, with the emphasis on minimizing off-chip components to meet the requirements of an implant application. Further construction with the stimulation chip in a neuromodulation (NM) unit is also presented and a primitive *in-vitro* concurrent sensing-stimulation test has been performed to verify the functionality of the entire system.

5.2 Future Work

The VCO-based architecture constitutes a new structure for brain-sensing front-end design. While some of the major issues in this structure are resolved in this work, there are other limitations that merit further research:

- Our front-end does not feature a specific provision to tolerate the large input common-mode (CM) fluctuation. The $60 \text{ dB}+$ CMRR as mentioned in section 2.5 is for a specific range. In our tests with the stimulation chip, the CM voltage fluctuation is less than 50 mV , which does not significantly degrade the front-end performance. However, to

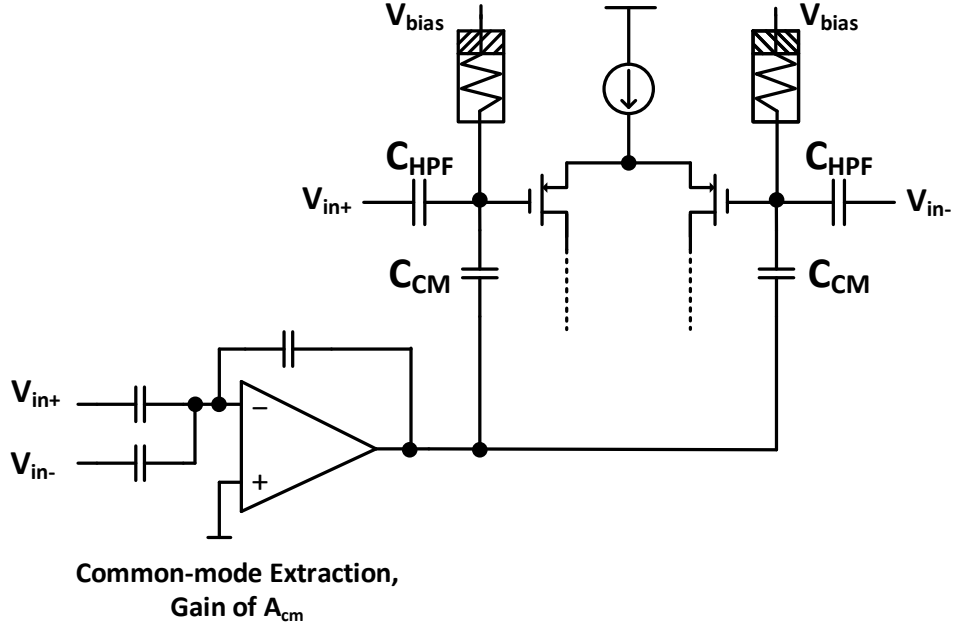


Figure 5.1: A possible way to deal with large input common-mode fluctuation.

make the front-end universally applicable to pair with other stimulators, it is desirable to deal with large-signal input CM fluctuations.

One way of solving this issue is proposed in [40][51] and shown in Fig 5.1. It uses an amplifier to extract the CM voltage from the electrodes. The amplifier output, which is the amplified CM fluctuation, is capacitively coupled to the diff-pair input nodes. When the capacitance value is set as: $C_{CM} = \frac{C_{HPF}}{A_{CM}}$, the diff-pair inputs are free from the CM fluctuation. Therefore, the front-end can maintain a stable performance. The power consumption of this amplifier can be low since its output noise will be suppressed as the diff pair CM input. However, due to the capacitive division between C_{HPF} and C_{CM} , the input signal will experience some attenuation. A smaller C_{CM} reduces signal attenuation, however it comes at the cost of a smaller CM cancellation range as limited by $\frac{C_{CM}}{C_{HPF}}V_{DD}$. Therefore, an appropriate trade-off is needed between the CM range and the sensitivity of the front-end.

- The NLC is currently implemented as a foreground calibration, and the coefficients are derived off-line. While this derivation is feasible at a small scale, it can be time-

consuming when a large-scale experiment/deployment is needed. Moreover, this limits the potential of applying our front-end in more generic environments, i.e. in non-implant scenarios. Therefore, it is desirable to develop a simple and reliable background calibration algorithm.

Since the calibration relies on an accurate perception of VCO nonlinearity, we can inspect either the spectrum or the time-domain waveform to determine the coefficients, with the NLC turned off and the input provided from an on/off-chip high-accuracy signal source. The spectrum domain processing requires an accurate estimation of the harmonic distortion. Therefore, a sine-wave input is preferred as the vector for calibration, with a narrow-band filter or a complex filter in the system for isolating the harmonics. On the other hand, we can directly evaluate the VCO V-to-F curve by examining the time-domain waveform. The input is preferred to be a square wave with variable amplitudes at a period that covers more than a few VCO sampling samples. Since the HPF corner is very low, the front-end output is almost stable after the initial settling at square-wave edges. These outputs at different input amplitudes can also be used for estimating the V-to-F curve and then calculating the NLC coefficients.

- The front-end has sufficiently suppressed the flicker noise, such that it is not the dominant source of the noise within the signal band. Further power efficiency improvement requires an effort to reduce the front-end white noise. As discussed, the diff pair and the oscillator stages both contribute to white noise, while only the diff pair provides transconductance for the signal gain. Consequently, to provide better power efficiency, a design that maximizes the possible transconductance with a given current is needed.
- The front-end bandwidth is limited by linearity performance degradation at higher frequencies, primarily because of the frequency-dependent nonlinearity or the sinc filtering inherent in the VCO phase-processing operation. In case of a need to cover higher-frequency AP signals, it is worth exploring a way to equalize or compensate for the frequency-dependent effect in order to extend the signal bandwidth.
- Lastly, the *in-vivo* neural recording from our front-end is essential in validating the

overall design. This can be accomplished in several ways:

The sensing chip can record in parallel with an existing rack-mounted instrument. This concordance test provides a baseline to quantify the quality of live recording from our front-end.

The NM unit can be implanted and tested in animal models first. Since the concordance test is impossible in such a scenario, we have to use a well-established bio-marker for validation. This requires close cooperation with neurologists to define the test plan and implement surgical procedure.

The final step involves an *in-vivo* concurrent sensing and stimulation test in a real patient. This step aims primarily to examine the efficacy of the overall system, and to provide us with feedback for further refining the system.

REFERENCES

- [1] R. R. Harrison and C. Charles, “A Low-Power Low-Noise CMOS Amplifier for Neural Recording Applications,” *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 958–965, June 2003.
- [2] T. Denison, K. Consoer, W. Santa, A. T. Avestruz, J. Cooley, and A. Kelly, “A $2\mu\text{W}$ 100 nV/rtHz Chopper-Stabilized Instrumentation Amplifier for Chronic Measurement of Neural Field Potentials,” *IEEE Journal of Solid-State Circuits*, vol. 42, pp. 2934–2945, Dec 2007.
- [3] H. Gao, R. M. Walker, P. Nuyujukian, K. A. Makinwa, K. V. Shenoy, B. Murmann, and T. H. Meng, “HermesE: A 96-Channel Full Data Rate Direct Neural Interface in $0.13\ \mu\text{m}$ CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 47, no. 4, pp. 1043–1055, 2012.
- [4] S. Stanslaski, P. Afshar, P. Cong, J. Giftakis, P. Stypulkowski, D. Carlson, D. Linde, D. Ullestad, A. T. Avestruz, and T. Denison, “Design and Validation of a Fully Implantable, Chronic, Closed-Loop Neuromodulation Device With Concurrent Sensing and Stimulation,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, pp. 410–421, July 2012.
- [5] B. C. Johnson, S. Gambini, I. Izyumin, A. Moin, A. Zhou, G. Alexandrov, S. R. Santacruz, J. M. Rabaey, J. M. Carmena, and R. Muller, “An implantable $700\ \mu\text{W}$ 64-channel neuromodulation IC for simultaneous recording and stimulation with rapid artifact recovery,” in *2017 Symposium on VLSI Circuits*, pp. C48–C49, June 2017.
- [6] M. Z. Straayer and M. H. Perrott, “A 12-Bit, 10-MHz Bandwidth, Continuous-Time $\Sigma\Delta$ ADC With a 5-Bit, 950-MS/s VCO-Based Quantizer,” *IEEE Journal of Solid-State Circuits*, vol. 43, pp. 805–814, April 2008.
- [7] W. Jiang, V. Hokyikyan, H. Chandrakumar, V. Karkare, and D. Marković, “A $\pm 50\text{-mV}$ Linear-Input-Range VCO-Based Neural-Recording Front-End With Digital Nonlinearity Correction,” *IEEE Journal of Solid-State Circuits*, vol. 52, pp. 173–184, Jan 2017.
- [8] V. Karkare, “Robust, Reconfigurable, and Power-Efficient Electrophysiological Recording Systems,” 2014.
- [9] R. B. Staszewski, K. J. Maggio, and D. D. Leipold, “Method and apparatus for asynchronous clock retiming,” Apr. 10 2012. US Patent 8,155,256.
- [10] J. Hamilton, S. Yan, and T. R. Viswanathan, “An Uncalibrated 2MHz, 6mW, 63.5dB SNDR Discrete-Time Input VCO-Based $\Delta\Sigma$ ADC,” in *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference*, pp. 1–4, Sept 2012.
- [11] D. Rozgić, V. Hokyikyan, W. Jiang, S. Basir-Kazeruni, H. Chandrakumar, W. Leng, and D. Marković, “A True Full-Duplex $32\text{-Channel } 0.135\text{cm}^3$ Neural Interface,” in *The 13th IEEE BioCAS*, Oct 2017.

- [12] V. Hokyhyan, “Design and Verification of a Closed-loop-ready High-channel-count Neuromodulation Unit,” 2017.
- [13] “U.S. Leading Categories of Diseases/Disorders.” <https://www.nimh.nih.gov/health/statistics/disability/us-leading-categories-of-diseases-disorders.shtml>. Accessed: 2017-10-20.
- [14] “Worldwide Neuroscience Initiatives.” <http://www.sfn.org/advocacy/neuroscience-funding/worldwide-neuroscience-initiatives>. Accessed: 2017-10-21.
- [15] M. K. Lyons, “Deep Brain Stimulation: Current and Future Clinical Applications,” in *Mayo Clinic Proceedings*, vol. 86, pp. 662–672, Elsevier, 2011.
- [16] P. Krack, A. Batir, N. Van Blercom, S. Chabardes, V. Fraix, C. Ardouin, A. Koudsie, P. D. Limousin, A. Benazzouz, J. F. LeBas, *et al.*, “Five-Year Follow-up of Bilateral Stimulation of the Subthalamic Nucleus in Advanced Parkinson’s Disease,” *New England Journal of Medicine*, vol. 349, no. 20, pp. 1925–1934, 2003.
- [17] P. Gubellini, P. Salin, L. Kerkerian-Le Goff, and C. Baunez, “Deep brain stimulation in neurological diseases and experimental models: from molecule to complex behavior,” *Progress in neurobiology*, vol. 89, no. 1, pp. 79–123, 2009.
- [18] P. J. Rossi, A. Gunduz, J. Judy, L. Wilson, A. Machado, J. J. Giordano, W. J. Elias, M. A. Rossi, C. L. Butson, M. D. Fox, *et al.*, “Proceedings of the Third Annual Deep Brain Stimulation Think Tank: A Review of Emerging Issues and Technologies,” *Frontiers in neuroscience*, vol. 10, 2016.
- [19] K. Wise and J. Angell, “A Microprobe with Integrated Amplifiers for Neurophysiology,” in *Solid-State Circuits Conference. Digest of Technical Papers. 1971 IEEE International*, vol. 14, pp. 100–101, IEEE, 1971.
- [20] T. Jochum, T. Denison, and P. Wolf, “Integrated circuit amplifiers for multi-electrode intracortical recording,” *Journal of neural engineering*, vol. 6, no. 1, p. 012001, 2009.
- [21] W. Wattanapanitch, M. Fee, and R. Sarpeshkar, “An Energy-Efficient Micropower Neural Recording Amplifier,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 1, no. 2, pp. 136–147, 2007.
- [22] K. Guillory and R. Normann, “A 100-channel system for real time detection and storage of extracellular spike waveforms,” *Journal of neuroscience methods*, vol. 91, no. 1, pp. 21–29, 1999.
- [23] “Implants for surgery – Active implantable medical devices – Part 1: General requirements for safety, marking and for information to be provided by the manufacturer,” standard, International Organization for Standardization, Aug. 2014.

- [24] S. Kim, R. A. Normann, R. Harrison, and F. Solzbacher, "Preliminary Study of the Thermal Impact of a Microelectrode Array Implanted in the Brain," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pp. 2986–2989, IEEE, 2006.
- [25] S. Kim, P. Tathireddy, R. A. Normann, and F. Solzbacher, "*In vitro* and *in vivo* study of temperature increases in the brain due to a neural implant," in *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pp. 163–166, IEEE, 2007.
- [26] C. C. Enz and G. C. Temes, "Circuit Techniques for Reducing the Effects of Op-Amp Imperfections: Autozeroing, Correlated Double Sampling, and Chopper Stabilization," *Proceedings of the IEEE*, vol. 84, no. 11, pp. 1584–1614, 1996.
- [27] K. Reddy, S. Rao, R. Inti, B. Young, A. Elshazly, M. Talegaonkar, and P. K. Hanumolu, "A 16-mW 78-dB SNDR 10-MHz BW CT $\Delta\Sigma$ ADC Using Residue-Cancelling VCO-Based Quantizer," *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 2916–2927, Dec 2012.
- [28] V. Karkare, H. Chandrakumar, D. Rozgić, and D. Marković, "Robust, Reconfigurable, and Power-Efficient Biosignal Recording Systems," in *Proceedings of the IEEE 2014 Custom Integrated Circuits Conference*, pp. 1–8, Sept 2014.
- [29] S. Basir-Kazeruni, S. Vlaski, H. Salami, A. H. Sayed, and D. Marković, "A Blind Adaptive Stimulation Artifact Rejection (ASAR) Engine for Closed-Loop Implantable Neuromodulation Systems," in *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 186–189, May 2017.
- [30] P. A. Mackowiak, S. S. Wasserman, and M. M. Levine, "A Critical Appraisal of 98.6 F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich," *Jama*, vol. 268, no. 12, pp. 1578–1580, 1992.
- [31] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications," *Analog integrated circuits and signal processing*, vol. 8, no. 1, pp. 83–114, 1995.
- [32] R. Sarpeshkar, W. Wattanapanitch, S. K. Arfin, B. I. Rapoport, S. Mandal, M. W. Baker, M. S. Fee, S. Musallam, and R. A. Andersen, "Low-Power Circuits for Brain–Machine Interfaces," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 2, no. 3, pp. 173–183, 2008.
- [33] K. H. Wee and R. Sarpeshkar, "An Electronically Tunable Linear or Nonlinear MOS Resistor," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 9, pp. 2573–2583, 2008.
- [34] S. Chen and R. Rieger, "Linear Low-Frequency Filter using On-Chip Giga-ohm Resistance," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pp. 1256–1259, May 2010.

- [35] K. Nagaraj, “A Parasitic-Insensitive Area-Efficient Approach to Realizing Very Large Time Constants in Switched-Capacitor Circuits,” *IEEE Transactions on Circuits and Systems*, vol. 36, no. 9, pp. 1210–1216, 1989.
- [36] Q. Fan, F. Sebastiano, J. H. Huijsing, and K. A. A. Makinwa, “A $1.8 \mu\text{W}$ $60 \text{ nV}/\sqrt{\text{Hz}}$ Capacitively-Coupled Chopper Instrumentation Amplifier in 65 nm CMOS for Wireless Sensor Nodes,” *IEEE Journal of Solid-State Circuits*, vol. 46, pp. 1534–1543, July 2011.
- [37] J. A. Kaehler, “Periodic-Switched Filter Networks—A Means of Amplifying and Varying Transfer Functions,” *IEEE Journal of Solid-State Circuits*, vol. 4, pp. 225–230, Aug 1969.
- [38] M. H. Perrott, S. Pamarti, E. G. Hoffman, F. S. Lee, S. Mukherjee, C. Lee, V. Tsinker, S. Perumal, B. T. Soto, N. Arumugam, *et al.*, “A Low Area, Switched-Resistor Based Fractional-N Synthesizer Applied to a MEMS-Based Programmable Oscillator,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2566–2581, 2010.
- [39] H. Chandrakumar and D. Marković, “A High Dynamic-Range Neural Recording Chopper Amplifier for Simultaneous Neural Recording and Stimulation,” *IEEE Journal of Solid-State Circuits*, vol. 52, no. 3, pp. 645–656, 2017.
- [40] H. Chandrakumar and D. Marković, “A $2.8\mu\text{W}$ 80mVpp Linear-Input-Range $1.6\text{G}\Omega$ Input Impedance, Bio-signal Chopper Amplifier Tolerant to Common-Mode Interference up to 650mVpp ,” in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 448–449, Feb 2017.
- [41] A. A. Abidi, “Phase Noise and Jitter in CMOS Ring Oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1803–1816, 2006.
- [42] Y.-H. Choi, B. Kim, J.-Y. Sim, and H.-J. Park, “A Phase-Interpolator-Based Fractional Counter for All-Digital Fractional-N Phase-Locked Loop,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, no. 3, pp. 249–253, 2017.
- [43] M. Lee and A. A. Abidi, “A 9 b, 1.25 ps Resolution Coarse–Fine Time-to-Digital Converter in 90 nm CMOS that Amplifies a Time Residue,” *IEEE Journal of solid-state circuits*, vol. 43, no. 4, pp. 769–777, 2008.
- [44] Z.-Z. Chen, Y.-H. Wang, J. Shin, Y. Zhao, S. A. Mirhaj, Y.-C. Kuan, H.-N. Chen, C.-P. Jou, M.-H. Tsai, F.-L. Hsueh, *et al.*, “A Sub-Sampling All-Digital Fractional-N Frequency Synthesizer with -111dBc/Hz In-Band Phase Noise and an FOM of -242dB ,” in *Solid-State Circuits Conference-(ISSCC), 2015 IEEE International*, pp. 1–3, IEEE, 2015.
- [45] R. B. Staszewski and P. T. Balsara, *All-Digital Frequency Synthesizer in Deep-Submicron CMOS*. John Wiley & Sons, 2006.
- [46] “DARPA Restoring Active Memory.” <https://www.darpa.mil/program/restoring-active-memory>. Accessed: 2017-11-1.

- [47] “DARPA System-Based Neurotechnology for Emerging Therapies.” <https://www.darpa.mil/program/systems-based-neurotechnology-for-emerging-therapies>. Accessed: 2017-11-1.
- [48] A. Yousefi, D. Yang, A. A. Abidi, and D. Marković, “A Distance-Immune Low-Power 4-Mbps Inductively-Coupled Bidirectional Data Link,” in *VLSI Circuits, 2017 Symposium on*, pp. C60–C61, IEEE, 2017.
- [49] R. Muller, S. Gambini, and J. M. Rabaey, “A 0.013 mm^2 , $5\text{ }\mu\text{W}$, DC-Coupled Neural Signal Acquisition IC With 0.5 V Supply,” *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 232–243, Jan 2012.
- [50] R. Muller, H.-P. Le, W. Li, P. Ledochowitsch, S. Gambini, T. Bjorninen, A. Koralek, J. M. Carmena, M. M. Maharbiz, E. Alon, *et al.*, “A Minimally Invasive 64-Channel Wireless μECoG Implant,” *IEEE Journal of Solid-State Circuits*, vol. 50, no. 1, pp. 344–359, 2015.
- [51] H. Chandrakumar and D. Marković, “An 80mVpp Linear-Input Range, $1.6\text{G}\Omega$ Input Impedance, Low-Power Chopper Amplifier for Closed-Loop Neural Recording That Is Tolerant to 650mVpp Common-Mode Interference,” *IEEE Journal of Solid-State Circuits*, vol. 52, no. 11, pp. 2811–2828, 2017.