UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

High-Fidelity Spatial Signal Processing in Low-Power Mixed-Signal VLSI Arrays

Permalink

https://escholarship.org/uc/item/9xz56906

Author

Joshi, Siddharth

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

High-Fidelity Spatial Signal
Processing in Low-Power Mixed-Signal VLSI Arrays

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

in

Electrical Engineering (Computer Engineering)

by

Siddharth Joshi

Committee in charge:

> Professor Gert Cauwenberghs, Chair
> Professor Peter M. Asbeck, Co-Chair
> Professor William Hodgkiss
> Professor Bill Lin
> Professor Patrick P. Mercier

2017

The dissertation of Siddharth Joshi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____ Co-Chair

_____ Chair

University of California, San Diego

2017

DEDICATION

Dedicated to my parents, Anuradha and Rajiv Joshi,

and my brother Varun Joshi.

EPIGRAPH

The scientist describes what is; the engineer creates what never was.

– Theodore von Kármán.

Frequently the messages have meaning.

– Claude Shannon.

TABLE OF CONTENTS

vii

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This chapter is probably the most important contained in this thesis. Over time, the algorithms may be outdated, the circuits may become obsolete, and the analysis may become irrelevant, however, the warmth, help, and encouragement, provided by everyone will remain a constant source of support to me. This work is a testament to that.

Foremost, I would like to thank my advisor, Professor Gert Cauwenberghs, for his continuous support and expert guidance. Gert has been an exemplar with his advising. Not only did Gert manage to keep a sense of humor and optimism when I had lost mine. His maddening attention to detail drove me to finally improve my punctuation of prose (hopefully!). His ability to be both very creative and rigorous while designing some very beautiful circuits has been a wonder to watch.

I would also like to thank my committee: Professors Peter Asbeck, Patrick Mercier, William Hodgkiss and Bill Lin. Thank you for your guidance and advice, your questions shaped the direction of my research providing me with a better answer to the question- "Why Analog?".

From Bonner Hall to PFBH, I've always had the members of ISN helping me and advising me. None of this would have been possible without their contributions. I'll miss the many coffees, teas, talks, lunches, and dinners that we've had over the years. I'd like to thank Dr. Mike Chi, Dr. Theodore Yu, Dr. Jongkil Park and Steve Deiss Who welcomed me when I first joined the lab. I'd like to thank Dr. Chis Thomas, Dr. Sohmyung Ha, Abraham Akinin, Chul Kim, and Raj Kubendran, Jun Wang, and Dr. Christoph Maier, their advice and support was invaluable at all phases of IC design, from conception, to design to testing, and most especially that final week of layout before *tapeout*. The variety of work done at ISNL provided us with a fertile ground for many discussions. I'd like to thank Dr. Fred Broccard, Dr. Massoud Khraiche, Dr. Sadique Sheik, Bruno Pedroni, Cory Stevenson, Dr. Hesham Mustafa, and Dr. Emre Neftci, for their willingness to teach, advise, and debate the various aspects and intricacies of biology, politics, and neural networks (spiking or otherwise).

I was truly lucky to have made good friends during my time in San Diego and La Jolla, Aman, Arpit, Dilraj, Joshal, Mandar, Shams, and many more. I've learned a lot from you all, and it's been a wonderful experience. In addition, I'd also like to thank my friends at Texas Instruments: Dan Gerber, Jon Spaulding, Nachiket Desai, Ujwal Radhakrishna, and Qing He. It was a pleasure to work alongside you all.

I owe a particular debt of gratitude to my friends and mentors from my undergraduate days, Prof. RN Biswas at DAIICT, and Prof. Sachin Patkar at IIT-Bombay, both played an instrumental role in developing my skills in EECS. A special thanks to Harsha, Arijit, Ankur, and Sai Deep, I've been fortunate to have known you since DAIICT and I've probably spent as much time talking and complaining to you as is humanly possible, you guys are the best.

And finally, I thank my family. Thank you for understanding when I was always too busy to catch up, leading to weeks without a conversation. Without your belief in me this dissertation would not have been possible. I'm indebted to my grandparents, from the constant encouragement and love they provided me, to the many small projects we worked on together, you fed my curiosity and helped it grow. When I first arrived in San Diego, Davender, Barbara, and Brittney Agnihotri made me feel at home here, and have continued to do so ever since. Abha Ghani, Ejaz Ghani, and Ishaan, ensured that I felt at home regardless of where I traveled. Megha Unhelkar was my pillar. Thank you for your support, your optimism, and for lending me your strength. I cannot sufficiently express my gratitude for your endless encouragement and enthusiasm which served as my guide during this long, and sometimes arduous, journey. This work would have never been possible without the help and encouragement of my parents, Anuradha and Rajiv Joshi, and my brother Varun. I don't think I can put into words how much you've shaped me, and how much your unconditional love and support has meant to me, this thesis is dedicated to you.

Chapter Two is largely a combination of material that will appear in the 2017 Proceedings of The IEEE Custom Integrated Circuits Conference: Siddharth Joshi, Chul Kim, Sohmyung Ha, Gert Cauwenberghs, "From Algorithms to Devices: Enabling Machine Learning through Ultra-Low-Power VLSI Mixed-Signal Array Processing," to appear, *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, Apr. 2017. The author is the primary author and investigator of this work.

Chapter Three is largely a combination of material in the following two venues: Siddharth Joshi, Chul Kim, Sohmyung Ha, Gert Cauwenberghs, "A 6.5 μW/MHz Charge Buffer with 7 fF Input Capacitance in 65 nm CMOS for Non-contact Electropotential Sensing" *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 12, pp. 1161-1165, Dec. 2016. Siddharth Joshi, Chul Kim, Sohmyung Ha, Gert Cauwenberghs, "A 6.5μW/MHz Charge Buffer with 7fF Input Capacitance in 65nm CMOS for Non-Contact Electropotential Sensing" *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, Canada, pp. 2907-2907, 2016. The author is the primary author and investigator of this work.

Chapter Four is largely a reprint of material that appeared in 2017 ISSCC digest of technical papers: Siddharth Joshi, Chul Kim, Sohmyung Ha, Yu M Chi, Gert Cauwenberghs, "2pJ/MAC 14b 8×8 Linear Transform Mixed-Signal Spatial Filter in 65nm CMOS with 84dB Interference Suppression," to appear, *IEEE ISSCC Dig. Tech. Papers*, San Francisco, CA, Feb. 2017. The author is the primary author and investigator of these works.

Chapter Five is largely a selection of material that appeared in the IEEE Journal of Solid-State Circuits, 2015: Chul Kim, Siddharth Joshi, Chris M Thomas, Sohmyung Ha, Lawrence E Larson, Gert Cauwenberghs, "A 1.3 mW 48 MHz 4 Channel MIMO Baseband Receiver With 65 dB Harmonic Rejection and 48.5 dB Spatial Signal Separation," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 4, pp. 832-844, April 2016. The author is a primary author and investigator of this work, and the primary author and investigator of the part reprinted here.

Chapter Six is largely a reprint of material that is being prepared for pub-

lication: Siddharth Joshi, Chul Kim, Chris M Thomas, Gert Cauwenberghs, "Digitally Adaptive High-Fidelity Analog Signal Processing Insensitive to Capacitive Multiplying DAC Inter-Stage Gain Error," *In preparation*. The author is the primary author and investigator of this work.

VITA

2008	B.Tech. in Information and Communication Technology, Dhirubhai Ambani Institute of Information and Communication technology
2012	M.S. in Electrical Engineering (Computer Engineering), University of California, San Diego
2017	Ph.D. in Electrical Engineering (Computer Engineering), University of California, San Diego

PUBLICATIONS

Siddharth Joshi, Chul Kim, Chris M Thomas, Gert Cauwenberghs, "Digitally Adaptive High-Fidelity Analog Signal Processing Insensitive to Capacitive Multiplying DAC Inter-Stage Gain Error," *In preparation.*

Frederic Broccard, Siddharth Joshi, Jun Wang, Gert Cauwenberghs, "Neuromorphic neural interfaces: from neurophysiological inspiration to biohybrid coupling with nervous systems," *Journal of Neural Engineering*, vol. 14 (4), 041002, doi:10.1088/1741-2552/aa67a9, 2017.

Said Hamdioui, Shahar Kvatinsky, Gert Cauwenberghs, Lei Xie, Nimrod Wald, Siddharth Joshi, Hesham Mostafa Elsayed, Henk Corporaal, Koen Bertels. "Memristor for computing: Myth or reality?." *Proc. Design, Automation & Test in Europe (DATE)*. EDA Consortium (2017) pp 1729-1725, Mar-Apr. 2017.

Sukru Burc Eryilmaz, Emre Neftci, Siddharth Joshi, SangBum Kim, Matthew BrightSky, Hsiang-Lan Lung, Chung Lam, Gert Cauwenberghs, HS Philip Wong, "Training a Probabilistic Graphical Model With Resistive Switching Electronic Synapses," *IEEE Transactions on Electron Devices*, vol. 63 (12) pp 5004-5011 2016.

Siddharth Joshi, Chul Kim, Sohmyung Ha, Gert Cauwenberghs, "From Algorithms to Devices: Enabling Machine Learning through Ultra-Low-Power VLSI Mixed-Signal Array Processing," *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, Apr. 2017.

Siddharth Joshi, Chul Kim, Sohmyung Ha, Yu M Chi, Gert Cauwenberghs, "2pJ/MAC 14b 8×8 Linear Transform Mixed-Signal Spatial Filter in 65nm CMOS with 84dB Interference Suppression," *IEEE ISSCC Dig. Tech. Papers*, San Francisco, CA, Feb. 2017.

Jongkil Park, Theodore Yu, Siddharth Joshi, Christoph Maier, Gert Cauwenberghs, "Hierarchical Address Event Routing for Reconfigurable Large-Scale Neuromorphic Systems," to appear, *IEEE Transactions on Neural Networks and Learning Systems.*

Sohmyung Ha, Chul Kim, Jongkil Park, Siddharth Joshi, Gert Cauwenberghs, "Energy Recycling Telemetry IC With Simultaneous 11.5 mW Power and 6.78 Mb/s Backward Data Delivery Over a Single 13.56 MHz Inductive Link," *IEEE Journal of Solid-State Circuits*, vol. 51 (11), pp. 2664-2678, 2016.

Siddharth Joshi, Chul Kim, Sohmyung Ha, Gert Cauwenberghs, "A 6.5 μW/MHz Charge Buffer with 7 fF Input Capacitance in 65 nm CMOS for Non-contact Electropotential Sensing" *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 12, pp. 1161-1165, Dec. 2016.

Chul Kim*, Siddharth Joshi*, Chris M Thomas, Sohmyung Ha, Lawrence E Larson, Gert Cauwenberghs, "A 1.3 mW 48 MHz 4 Channel MIMO Baseband Receiver With 65 dB Harmonic Rejection and 48.5 dB Spatial Signal Separation," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 4, pp. 832-844, April 2016.*equal contribution

Emre O Neftci, Bruno U Pedroni, Siddharth Joshi, Maruan Al-Shedivat, G Cauwenberghs, "Stochastic Synapses Enable Efficient Brain-Inspired Learning Machines," *Frontiers in Neuroscience* vol. 10, pp. 241:1-16, 2016.

Bruno U Pedroni, Sadique Sheik, Siddharth Joshi, Georgios Detorakis, Somnath Paul, Charles Augustine, Emre Neftci, Gert Cauwenberghs, "Forward table-based presynaptic event-triggered spike-timing-dependent plasticity," *Proc. IEEE Biomedical Circuits and Systems (BioCAS)*, Shanghai, China, Oct. 17-19, 2016.

Siddharth Joshi, Chul Kim, Sohmyung Ha, Gert Cauwenberghs, "A 6.5μW/MHz Charge Buffer with 7fF Input Capacitance in 65nm CMOS for Non-Contact Electropotential Sensing" *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, Canada, pp. 2907-2907, 2016.

Sukuru B Eryilmaz, Siddharth Joshi, Emre O Neftci, Weir Wan, Gert Cauwenberghs, HS Philip Wong, "Neuromorphic architectures with electronic synapses," *2016 17th International Symposium on Quality Electronic Design (ISQED)*, Santa Clara, CA, 2016.

Chul Kim, Siddharth Joshi, Christopher M Thomas, Sohmyung Ha, Abraham Akinin, Lawrence E Larson, Gert Cauwenberghs, "A CMOS 4-channel MIMO baseband receiver with 65dB harmonic rejection over 48MHz and 50dB spatial signal separation over 3MHz at 1.3 mW," *Symposium on VLSI Circuits Digest of Technical Papers*, Kyoto, Japan, June 16-19, 2015.

Jongkil Park, Sohmyung Ha, Chul Kim, Siddharth Joshi, Theodore Yu, Wei Ma, Gert Cauwenberghs, "A 12.6 mW 8.3 Mevents/s contrast detection 128 128 imager with 75 dB intra-scene DR asynchronous random-access digital readout" *Proc. IEEE Biomedical Circuits and Systems (BioCAS)*, Lausanne, Switzerland, Oct. 22-24, 2014.

Chul Kim, Sohmyung Ha, Chris Thomas, Siddharth Joshi, Jongkil Park, Lawrence Larson, Gert Cauwenberghs, "A 7.86 mW +12.5 dBm in-band IIP3 8-to-320 MHz capacitive harmonic rejection mixer in 65nm CMOS" *Proc. IEEE European Solid-State Circuits Conference (ESSCIRC)*, Venice, Italy, Sept. 22-26, 2014.

Sohmyung Ha, Chul Kim, Jongkil Park, Siddharth Joshi, Gert Cauwenberghs, "Energy-recycling integrated 6.78-Mbps data 6.3-mW power telemetry over a single 13.56-MHz inductive link," *Symposium on VLSI Circuits Digest of Technical Papers*, Honolulu, HI, USA, Jun. 10-13, 2014.

Theodore Yu, Jongkil Park, Siddharth Joshi, Christoph Maier, Gert Cauwenberghs, "65k-neuron integrate-and-fire array transceiver with address-event reconfigurable synaptic routing," *Proc. IEEE Biomedical Circuits and Systems Conf. (BioCAS)*, Hsinchu Taiwan, Nov. 28-30, 2012.

Theodore Yu, Jongkil Park, Siddharth Joshi, Christoph Maier, Gert Cauwenberghs, "Event-driven neural integration and synchronicity in analog VLSI," *Proc. IEEE Engineering in Medicine and Biology Conf. (EMBC)*, San Diego CA, Aug. 28-Sept. 1, pp. 775-778, 2012.

Jongkil Park, Theodore Yu, Christoph Maier, Siddharth Joshi, Gert Cauwenberghs, "Live Demonstration: Hierarchical Address-Event Routing Architecture for Reconfigurable Large Scale Neuromorphic Systems," *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Seoul Korea, May 20-23, pp. 707-711, 2012.

Theodore Yu, Siddharth Joshi, Venkat Rangan, Gert Cauwenberghs, "Sub-threshold MOS Dynamic Translinear Neural and Synaptic Conductance," *Int. IEEE/EMBS Conf. Neural Engineering (NER)*, Cancun, Mexico, Apr. 27-May 1, pp. 68-71, 2011.

Siddharth Joshi, Stephen Deiss, Mike Arnold, Jongkil Park, Theodore Yu, Gert Cauwenberghs, "Scalable Event Routing in Hierarchical Neural Array Architecture with Gobal Synaptic Connectivity," *Proc. IEEE Int. Workshop Cellular Nanoscale Networks and Their Applications (CNNA)*, Berkeley CA, Febr. 3-5, 2010.

BY Vinay Kumar, Siddharth Joshi, Sachin Patkar, H Narayanan, "FPGA based High Performance Double-precision Matrix Multiplication," *Springer International journal of parallel programming (IJPP)*, vol. 38, no: 3-4, pp. 322-338, Feb. 2010.

BY Vinay Kumar, Siddharth Joshi, Sachin Patkar, H Narayanan, "FPGA based High Performance Double-precision Matrix Multiplication," *Proc. 22nd Int. Conference on VLSI Design.* New Delhi, India, 2009.

ABSTRACT OF THE DISSERTATION

High-Fidelity Spatial Signal
Processing in Low-Power Mixed-Signal VLSI Arrays

by

Siddharth Joshi

Doctor of Philosophy in Electrical Engineering (Computer Engineering)

University of California, San Diego, 2017

Professor Gert Cauwenberghs, Chair
Professor Peter M. Asbeck, Co-Chair

Machine learning and related statistical signal processing are expected to endow sensor networks with adaptive machine intelligence and greatly facilitate the Internet of Things (IoT). As such, architectures embedding adaptive and learning algorithms on-chip are oft-ignored by system architects and design engineers, and present a new set of design trade-offs. We focus on topologies efficiently implementing mixed-signal matrix-vector multiplication for applications in spatial filtering for IoT, where substantial processing gain in the analog domain alleviates the need for highly accurate and energy-consuming analog-to-digital conversion. We present a micropower, high-dynamic-range multichannel multiple-input multiple-

out (MIMO) mixed-signal linear transform system, with analog signal path and digital coefficient control, composed of an array of 14-bit Nested Thermometer Multiplying DACs (NTMDACs) implementing analog multiplication, and variable gain amplifier (VGA) implementing accumulation. Implemented in 65nm CMOS, the NTMDAC MISO system-on-chip measures 84 dB in interference suppression at 2 pJ of energy per mixed-signal multiply-accumulate. We demonstrate state-of-the art performance on two tasks, spectrally oblivious interference suppression in communication signals and EEG signal separation. We then provide experimental demonstration of the use of a MIMO mixed-signal linear-transform system within a radio-frequency receiver chain. Over-the-air experiments demonstrating signal separation for two broad-band modulated signals further validate the adaptive beamforming capabilities under severe multipath conditions even in the absence of line-of-sight communication path.

In order to mitigate adverse effects of radix errors and capacitive mismatch encountered in compact low-power realizations of high-resolution, high-dimensional MIMO analog processing systems, we introduce *Stochastic Successive Approximation*, or *S2A*, as an on-line adaptive optimization algorithm amenable to efficient implementation in massively parallel analog hardware. S2A offers a direct alternative to stochastic gradient descent overcoming several of its shortcomings, such as its sensitivity to analog mismatch model errors, while improving on the rate of convergence for high-dimensional analog computation. The S2A algorithm enables convergence to values closer to the optimal when facing non-convex optimization landscapes induced by mismatch in capacitive multiplying digital-to-analog converter components when applied to adaptive analog signal processing. We experimentally demonstrate, in fewer than 25 iteratations of S2A, 65 dB of processing gain in adaptive beamforming, over-the-air, multipath interferer suppression.

Chapter 1

Introduction

The remarkable capabilities recently demonstrated by machine intelligence at tasks long considered to be central to human cognition, have come at the cost of greater use of resources for computing, communication, storage, energy, and latency. When constraining the available resources, biological intelligence significantly outperforms machine intelligence. This performance gap is further exacerbated when complex unstructured interactions occur between the environment and the entity. Through the evolution of remarkable statistical computing and pattern recognition, biological entities have the capability of remarkable performance in regimes of low SNR, and incomplete information. This contrast starkly with the traditional realization of digital von Neumann computation. To truly enable ambient intelligence resource-constrained sensory nodes provide multi-modal sensory inputs to a distributed, robust, intelligence requires we reexamine how resources constrained systems might implement computation and communication. This is the key challenge I try to address in this dissertation.

1.1 Objectives

This dissertation focuses on improving the performance of intelligent systems at varying scales with an emphasis on resource-constrained ultra-low-power sensor nodes. The unprecedented growth in the capabilities of machine intelligence has placed greater demands on the computational capabilities and energy efficiency

of the underlying hardware [51,74]. This growth, spurred by the pervasiveness and ubiquitousness of electronics has had far reaching impacts from how data is sensed at the smallest scale, to how large data-centers analyze and use this data. This has led to the need for ever increasing computational capacity at minimal energy cost. We aim to improve the performance of intelligent and adaptive systems by focusing on the joint design of algorithms, architectures, circuits and systems.

This dissertation is unified around a focus on resource constrained, parallel silicon microsystems enabling real-time machine intelligence and sensory information processing. Such systems are widely applicable in areas like biomedical data acquisition, continuous infrastructure monitoring, intelligent sensor networks, and data analytics. Currently, large-scale, collective intelligence involves "dumb" nodes gathering data for a remote, centralized intelligence. In the absence of local processing, the latencies introduced by remote communication and processing renders low-energy autonomous systems impractical. Thus, low-power, on-chip intelligence is a prerequisite for autonomous systems interacting with the environment, making decisions, and taking required actions without human supervision. In order to effectively interface and actuate within an environment, next-generation sensory systems are seeing a push away from "dumb" signal acquisition, towards "smart" signal analysis. This move is further reinforced by the qualitative improvement in the informativeness and richness of the acquired data offered by self-contained, autonomous, "smart" sensory systems.

This inclusion of intelligence on chip entails a joint design of algorithms, architectures, circuits and systems to enable optimal trade-offs between power, speed, and quality of result. Thus this work traversed various levels of abstraction [73] from error tolerant algorithm development [57] for machine intelligence and analog signal processing to high-resolution efficient VLSI circuits for MIMO sensory and communication systems.

This work aims at building and designing energy efficient, scalable, highly parallel VLSI microsystems for applications in real-time signal processing, sensory interfaces, and, data analytics. By using techniques that enable approaching limits of energy efficiency, sensing and resolution by exploiting computational primitives

inherent in the physics of devices, sensors this work highlights methods that provide a principled approach to developing low-power analog computational systems. Despite its many advantages, analog processing naturally inherits the limitations of analog circuits, performance loss due to mismatch and process-variation, susceptibility to noise, offsets, and distortion. Consequently, analog processing based systems have to provide a principled approach to tackle these limitations, both at a hardware level and an algorithmic level. Thus the later chapters of this thesis are dedicated to algorithmic means to address these shortcomings.

1.2   Organization

Chapter 2 presents an overview of the fundamentals and state-of-the-art analog machine intelligence, with a focus on power efficient operation. Since the computational and energy burden imposed by emerging machine learning algorithms is the performance limiting factor, we focus on energy efficiency in this chapter. This chapter introduces energy performance trade-offs in the context of analog computation using passive components, discussing the advantages of digital computation over analog computation for various system requirements. We provide examples of state-of-the-art systems with the corresponding algorithmic references as well as some an introduction to emerging memory devices.

Chapter 3 presents a CMOS charge buffer with fF-range input capacitance for applications in capacitive electropotential sensing. We analyze and verify a feedback mechanism to negate parasitic capacitances seen at the input of a CMOS amplifier. Measurements are presented from a prototype fabricated in 65 nm CMOS occupying an active area of 193 μm^2 with an efficiency of 6.5 μW/MHz. Over-the-air measurements validate its applicability to electropotential sensing. This buffer forms a fundamental building block for the analog processing system that is introduced in chapter 4.

Chapter 4 builds upon the findings of the previous two chapters and introduces a custom IC designed based on the findings in the previous chapter. The chapter introduces both the mixed-signal spatial co-processor, and the *Nested*

Thermometer Multiplying Digital-to-Analog Converter (NTMDAC) which forms the fundamental building block for the co-processor. NTMDAC is developed using the *infinimp2* amplifier developed in Chapter 2. Fabricated in 65 nm CMOS, this mixed-signal spatial co-processor implements an 8×8 matrix-vector product at 14-bit analog resolution while consuming just 2 pJ per multiply accumulate.

Chapter 5 presents applications of a spatial filtering IC in the context of the Intermediate Frequency (IF) stage of a receiver Integrated Circuit (IC). We demonstrate both off-line source separation of two communication signals as well as over-the-air interferer suppression of two broad-band modulated sources in a multi-path RF environment. With 38 dB interferer suppression in over-the-air tests, and $\leq 2.5\%$ RMS EVM for spectrally oblivious separation of interfering QPSK and 16-QAM signals, we experimentally validate the high-resolution spatial processing capabilities of the micro-power MIMO spatial filtering IC.

Chapter 6 introduces the *successive stochastic approximation* and the *extended successive stochastic approximation* algorithms, modifications to stochastic gradient descent overcoming its shortcomings when applied to high-dimensional analog computation. S2A enables convergence to values close to the optimal in the presence of radix-errors introduced by mismatch in components for ASP. Conventional gradient descent proves to be sensitive to such analog mismatch model errors due to the effective non-convex optimization landscapes they typically induce. Experimental demonstrations in online over-the-air adaptive beamforming with the 25 iterations of the S2A algorithm achieves ≥ 65 dB of interferer suppression for narrow-band communication signals in multi-path environments.

Finally, Chapter 7 offers concluding remarks on the advances contributed in this thesis, their significance, and directions for future research.

Chapter 2

Enabling Machine Learning Through Ultra-Low-Power VLSI Mixed-Signal Array Processing

2.1   Introduction

Typical internet-of-things (IoT) connected intelligent integrated systems acquire sensory data, perform minimal local processing, and then offload more complex tasks to remote servers. Not only does such a system incur significant energy costs due to the need for constant communication [56], the increased latency makes real-time operation and response to change in the environment quite cost-prohibitive in the absence of any local processing [51]. Since energy efficiency, security and robustness are major factors driving the design of such sensor systems, there has been a move to shift the bulk of the processing closer to the sensory interface and hence drastically reduce demands on communication bandwidth [3]. Furthermore, on-chip intelligence can enable integrated systems to interact with their environment without constant remote input or monitoring, facilitating ubiquitous autonomy in IoT systems.

Achieving ultra-low-power operation, a critical requirement for autonomous IoT devices [51], entails a concerted effort at reducing power consumption at several
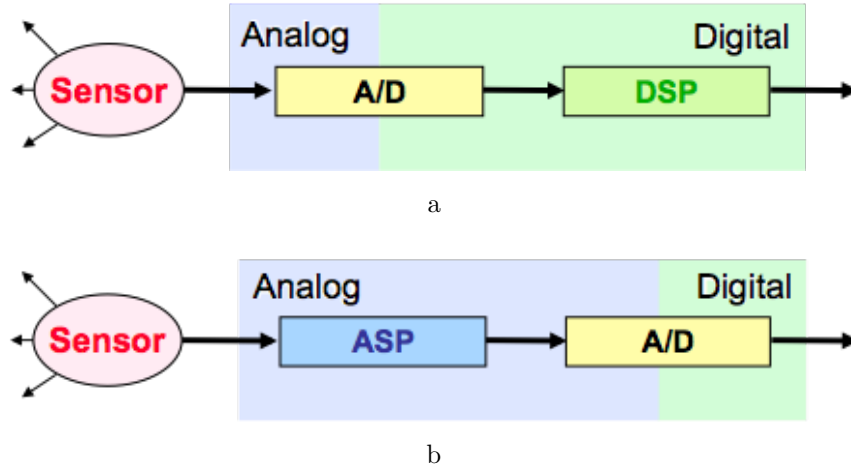
Figure 2.1: Signal processing flow a in conventional signal acquisition with digital signal processing (DSP), and b optimized for energy-efficient IoT with increased sensory-level analog signal processing (ASP) trading reduced analog-to-digital (A/D) conversion and DSP

levels in the design hierarchy as follows:

1. *Algorithmic and system level* — Analysis of sensor outputs should be robust to imprecision and noise, and algorithms amenable to local or distributed implementation lead to lower communication and power overhead;

2. *Architectural level* — Exploiting parallelism and pipelining as called for by the target application help restrict power expenditure;

3. *Circuit and logic level* — Appropriate use of sub-threshold vs. above-threshold MOS biasing, optimized mixing of logic styles, supply switching;

4. *Technology level* — Emerging devices and MEMS, increased reliability.

For intelligent sensory devices implementing processing on-chip, the signal chain conventionally consists of a sensor front-end providing inputs, followed by signal conditioning and filtering. Analog-to-digital converters (ADCs) then feed a digitized version of the output signal to a digital signal processor (DSP), illustrated in Fig. 2.1a. Embedding very low-power analog signal processing (ASP) subsystems near the sensory interface to remove redundancy in the input signal can help amortize the overhead of analog-to-digital conversion, and subsequent digital process-

ing, shown in Fig. 2.1b. Its versatility, general applicability to many algorithms, and amenability to low power implementation make matrix-vector multiplication (MVM) a natural choice as a primitive for such analog processing. Low-power and highly energy-efficient systems have been implemented with analog matrix-vector multipliers (aMVMs) for dimensionality reduction [7], linear classification [83], spatial filtering [37], support vector machines [39], neural networks [48] and many other settings in adaptive signal processing [43]. We shall focus this chapter on various implementations of the aMVM computational primitive in mixed-signal and analog circuits. Section 2.2 provides an overview of the algorithmic formulations. Section 2.3 provides bounds on energy consumption for analog processing as well as some examples of systems implementing aMVMs pre-digitization. Section 2.4 introduces systems where aMVM has been implemented post-digitization for energy minimizing optimization. Section 2.5 highlights latest advances in emerging resistive memory devices for massively parallel aMVM with applications in high-dimensional adaptive computation. We conclude with a look at future directions in Section 2.6.

2.2   Algorithmic Considerations

The canonical form of the narrowband beamformer shown in Fig. 2.2a bears remarkable resemblance to that of the perceptron Fig. 2.2b. This is further highlighted by beamformers implemented using the perceptron algorithm [76]. Exploiting the equivalence between adaptive beamformers and blind source separation [60] under the constraints of linear mixing enables the use of a wide range of machine-learning techniques for spatial signal processing. One such notable example is independent component analysis (ICA) [52] which uses the independence of signal statistics and spatial diversity in the measurements to separate and locate multiple sources from no more than the measurements alone. Crucially, ICA operates on the assumption that the set of signals received are a linear mixture of some underlying sources, i.e., for N observed signals $\boldsymbol{x}(t)$ there exists M source signals
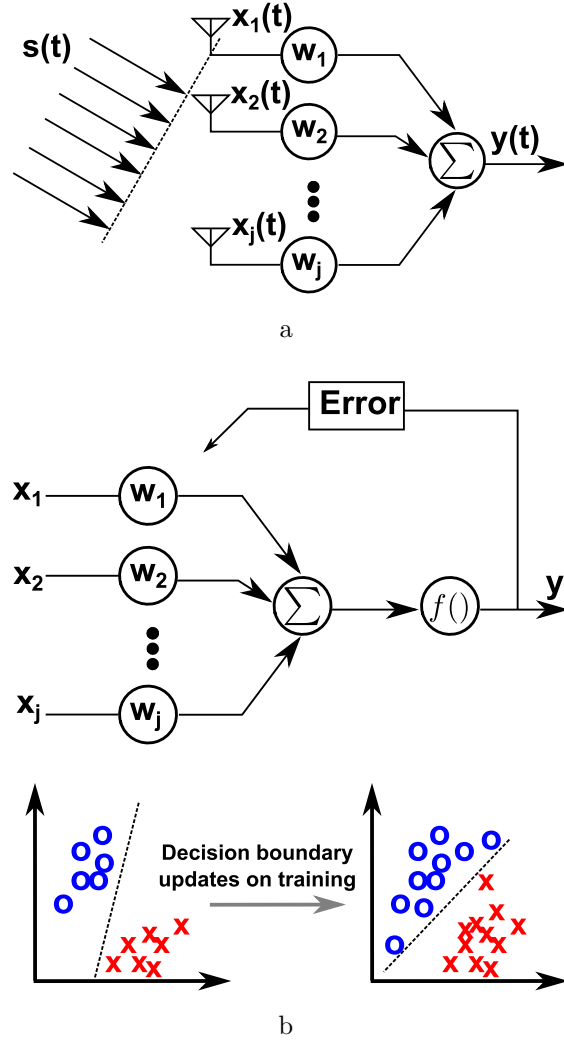
Figure 2.2: Structural similarity between a a canonical narrow-band beamformer and b the perceptron algorithm, highlighting the versatility and ubiquity of the matrix-vector multiplication (MVM) kernel.

$s(t)$ and an N-by-M matrix $\boldsymbol{A}$ such that, in vector form:

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t). \tag{2.1}$$

Thus, multiplying $\boldsymbol{x}(t)$ with the inverse of the mixing matrix $\boldsymbol{A}^{-1}$ results in the original signals, $\boldsymbol{s}(t)$, being recovered. The task in ICA is estimating $\boldsymbol{W} = \boldsymbol{A}^{-1}$ with minimal distortion, bearing in mind arbitrary permutations and scaling in the signal sources. However, a variety of signal conditions can result in ill-conditioned (almost singular) matrices requiring high-precision MVM to prevent errors (quan-
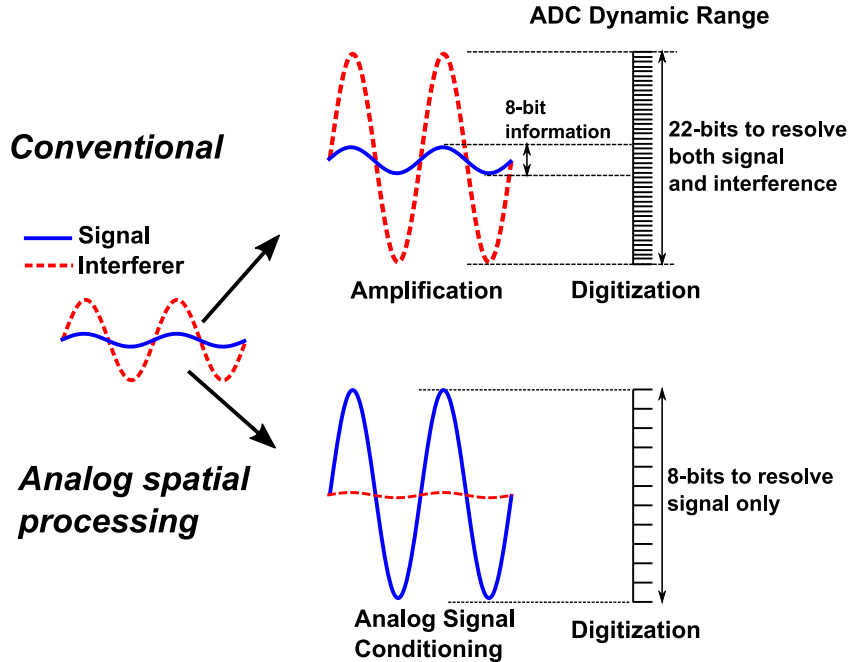
Figure 2.3: ASP can dramatically reduce dynamic range requirements prior to digitization. The processing gain of high-fidelity analog spatial filtering enables rejection of spectral interferers preventing subsequent stages in the ASP chain from saturating.

tified in Fig. 2.8). This is especially applicable for tasks such as beamforming separation of near-collinear sources, and other tasks incurring ICA or principal component analysis (PCA) [7].

Aside from limited resolution, linear systems comprising aMVMs incur performance loss due to limitations of linear transforms [82]. However, various aggregation techniques such as boosting have been introduced that enable collectives of linear maps to approximate more complex functions in piece-wise linear fashion [23, 70]. Such aggregation techniques are especially effective in overcoming limitations in analog hardware for classification, where multiple *weak* classifiers are pooled together to implement *strong* classification [67]. Indeed the use of such boosting techniques can provide a principled approach to alleviate process-voltage-temperatue (PVT) variations, mismatch, and noise at an algorithmic level.

The improvement in the signal-to-noise/signal-to-interferer ratio from the techniques highlighted in this Section come at increased energy and size costs of
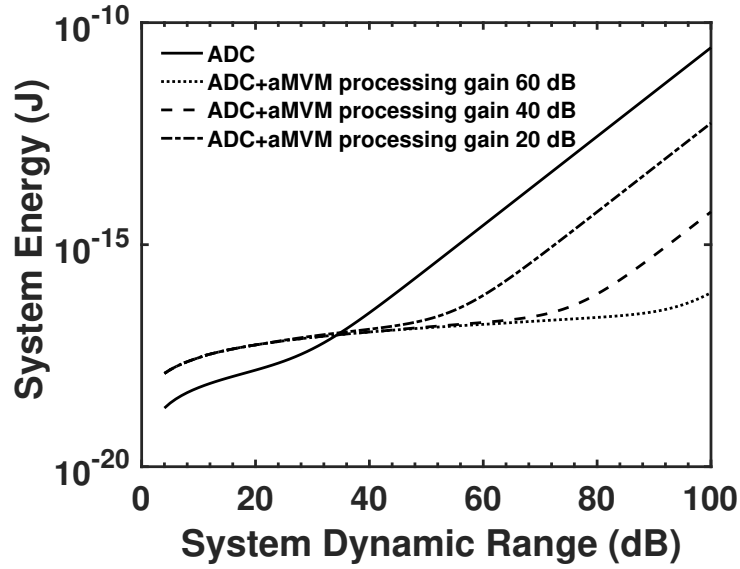
Figure 2.4: Minimum system energy limits $E_{\text{sys}} = P_{\text{sys}}/f_{\text{sig}}$ according to Eqs. (2.2)-(2.8) with $A = 8$, $\alpha = 2$ and processing gain $G = \text{DR}_{\text{prior}}/\text{DR}_{\text{post}} = 20$, 40, and 60 dB. At lower system dynamic range DR_{post} the energy of aMVM dominates that of SAR ADC, up to the cross-over point where the processing gain is limited to unity.

increased resolution in aMVM and increased analog processing through parallelism. We refer to this improvement in the signal-to-interferer ratio as processing gain, similar to processing gain in spread-spectrum techniques [63]. An example of *spatial* processing gain is illustrated in Fig. 2.3, where spatially selective filtering of the signal reduces the interferer power while maintaining the signal dynamic range, enabling lower resolution digitization and hence substantial energy savings. This increased analog processing in turn comes at an energetic cost, the balance of which is explored in Section 2.3.

2.3 Analog Signal Conditioning

Analog signal processing can facilitate a wide variety of sensory acquisition and emerging communications technologies. The central examples presented in this chapter focus on applications in the communication domain like full-duplex (FD) and cognitive radios (CR), which currently face many challenges due to high

dynamic range requirements [65]. However, it should be noted that analog signal processing has had wide applicability in machine learning, where nanowatt support vector machines have been demonstrated [24] and signal processing, where micropower implementations have enabled beamforming [11] and sound localization.

In what follows, we establish principles for analog processing to ensure overall energy savings compared to the conventional approach of directly quantizing the signal and operating upon it with DSP. For ease of notation, we consider a processing gain resulting from spatially filtering an interference source. This manifests as a reduction in the dynamic range specifications for a down-stream digitizer. The same principle applies to dynamic range reduction by feature extraction in other forms of signal processing.

2.3.1 Power Efficiency

Consider the power requirements for a successive approximation register (SAR) ADC with a binary weighted capacitive DAC with three main constituents:

$$P_{\text{SAR}} = P_{\text{driver}} + P_{\text{mean,switch}} + P_{\text{comp}}. \tag{2.2}$$

The power for the DAC driver P_{driver} is bounded by [53]

$$P_{\text{driver}} = 16 f_{\text{samp}} k_B T \, \text{DR} \tag{2.3}$$

where f_{samp} is the sampling frequency, k_B is the Boltzmann constant, T is absolute temperature, and DR is the dynamic range. The mean switching power over all codes $P_{\text{mean,switch}}$, assuming a *merged capacitor switching* based SAR [31], is

$$P_{\text{mean,switch}} = f_{\text{samp}} \sum_{i=1}^{n-1} 2^{n-3-2i} \left(2^i - 1\right) C_u V_{\text{ref}}^2$$

where C_u is the unit capacitor, and $n = (\text{DR}[dB] - 3)/6$ is the ADC number of bits. Minimum capacitor sizing for thermal noise[1] results in

$$P_{\text{mean,switch}} = 12 k_B T f_{\text{samp}} 2^n \sum_{i=1}^{n-1} 2^{n-3-2i} \left(2^i - 1\right). \tag{2.4}$$

---

[1] We size $C_u = 12 \, k_B T \, 2^n / V_{\text{ref}}^2$ to equate thermal and quantization noise, rather than sizing for mismatch, for a lower energy bound.
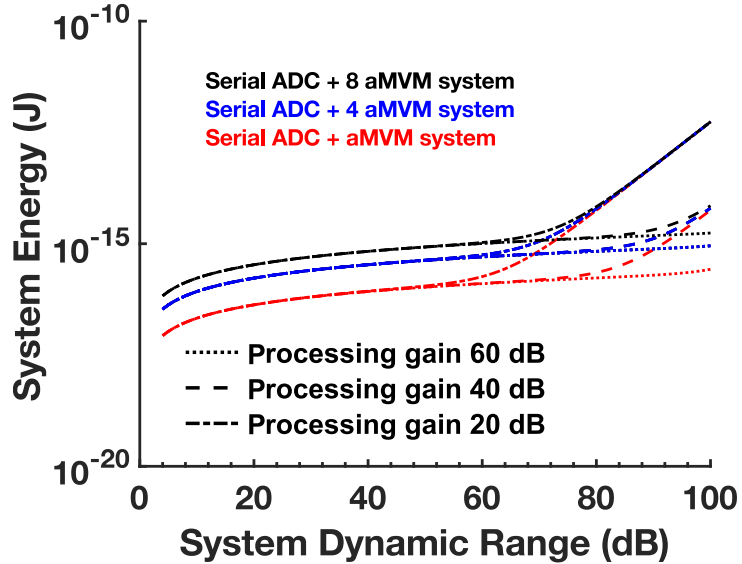
Figure 2.5: Minimum energy limits as in Fig. 2.4, with aMVM parallelism $N = 1$, 4, and 8 according to Eq. (2.7), at 10% parasitic capacitance ($\gamma = 0.1$). Amplifier gain A is increased in order to restore signal levels to full-scale for downstream ADC to counter the attenuation resulting from parallelism.

Finally, the switching power of the comparator P_{comp} is bounded by [54]:

$$P_{\mathrm{comp}} = 12 f_{\mathrm{samp}} \, k_B T \, n \, \mathrm{DR}. \tag{2.5}$$

Though greatly simplified, the resultant expression for P_{SAR} provides a lower-bound on power consumed for an n-bit SAR ADC.

Now, consider the presence of an interfering signal at signal-to-interference ratio SIR, which necessitates proportionally greater DR for the ADC to resolve the input signal amid the interferer without overload distortion. In turn, the greater ADC DR leads to higher ADC power consumption according to Eqs. (2.2)-(2.5). A suitable aMVM front-end subsystem capable of suppressing the interferer and restoring the signal to full strength prior to quantization can hence substantially reduce the ADC power consumption, albeit at some aMVM power cost.

Capacitive aMVM incurs power costs mainly for three operations: changing the capacitive weights P_{adapt}, driving the capacitor array P_{array}, and restoring the signal with gain P_{gain}. The minimum power required to drive a capacitor with a
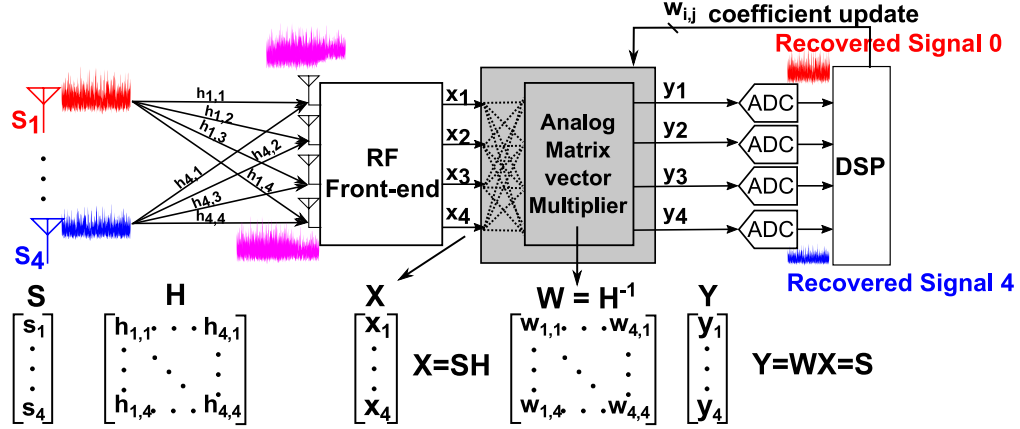
Figure 2.6: Analog signal processing can dramatically reduce the dynamic range prior to digitization. The implemented MIMO analog matrix-vector multiplier (aMVM) [37], reconditions the signal implementing high-fidelity analog beamforming.

sinusoidal signal with frequency $f_{\text{samp}} = 2f_{\text{sig}}$ at a given SNR is given by:

$$P_{\text{array}} = 8f_{\text{sig}}k_B T \, \text{SNR}. \tag{2.6}$$

Under the simplifying assumption relating the signal SNR to its dynamic range [54], the aMVM power reduces to

$$P_{\text{array}} + P_{\text{gain}} = 8f_{\text{sig}}(\text{DR}_{\text{prior}} + A\alpha \, \text{DR}_{\text{post}})k_B T \tag{2.7}$$

with closed loop gain A, amplifier inefficiency factor $\alpha \geq 1$, and dynamic range DR_{prior} prior to and DR_{post} post the aMVM gain stage. Continuous-time passive multiplication imposes a constant load on the drivers, in contrast to a switching structure which incurs an additional power cost. Due to the improved energy efficiency of passive multiplication, we choose that architecture over alternative switching architectures (discussed in Section 2.4.1). The net power of the combined aMVM-ADC system is then given by:

$$P_{\text{sys}} = P_{\text{array}} + P_{\text{gain}} + P_{\text{mean,switch}} + P_{\text{comp}} \tag{2.8}$$

in which the driver power (2.3), mean switching power (2.4), and comparator switching power (2.5) for the SAR ADC are incurred at the post dynamic range DR_{post}. Note that the ADC *post* driver power P_{driver} is subsumed by the aMVM

active gain power P_{gain} through the inefficiency factor α. The aMVM provides processing gain to boost the signal relative to interferer which relaxes the dynamic range accordingly, where the processing gain $G = \text{SIR}_{\text{post}}/\text{SIR}_{\text{prior}} = \text{DR}_{\text{prior}}/\text{DR}_{\text{post}}$. Thus, the combined aMVM-ADC system incurs a reduced cost for the ADC power P_{SAR} at the lower dynamic range DR_{post}, at the expense of aMVM power $P_{\text{array}} + P_{\text{gain}}$ providing the processing gain G. We normalize the power measures Eqs. (2.2)-(2.8) by f_{sig} and as such determine the minimum system energy limits $E_{\text{sys}} = P_{\text{sys}}/f_{\text{sig}}$ in Fig. 2.4. We show that the aMVM can reduce the cost of digitization, bounded by the processing gain G of the aMVM system. At lower system dynamic range the energy of aMVM dominates that of SAR ADC, where the cross-over point is determined by unity lower limit on processing gain. At higher system dynamic range, the benefits of aMVM are bounded by the processing gain, and it can be seen that the ADC energy cost once again dominates. A caveat to the analysis is that at higher system dynamic range, oversampling data-converters are more energy efficient and practical than SAR ADCs. Concurring with these findings, alternative analysis [80] suggests substantial (greater than 90%) power savings owing to analog preprocessing ($G = 40$ dB) within the context of a multiple-input-multiple output (MIMO) radio front-end under realistic channel conditions.

2.3.2 Limits of Parallelism

The inherent parallelism of analog computation offers several distinct advantages, such as the innate capability of accumulating charge from multiple sources onto a single wire shared connection [69]. The improved throughput from parallelism further benefits more computationally intensive applications [24]. Applications like CR and MIMO systems also benefit from parallelism, as do boosting algorithms as described above. Despite these advantages, recent work implementing aMVM through highly energy-efficient passive charge sharing [7,48,83] has not aggressively pursued parallelism. This is largely due to dynamic range limitations in massively parallel analog circuit architecture. In this Section we highlight some of these limitations along with methods to overcome them.

Highly parallel charge-redistribution capacitive arrays for aMVM suffer from gain error and signal level degradation as a result of parasitic capacitance as well as signal attenuation onto the parallel signal path. Consider the parallel connection of N capacitive multiplying DACs to compute the analog sum of N weighted inputs, $\sum_{j=1}^{N} W_{ij} x_j$ with digital weight coefficients W_{ij} and analog voltage inputs x_j. Passively connecting this aMVM output directly to the input of a SAR ADC, with another capacitive DAC for the ADC reference connected in parallel, results in charge-sharing attenuation of the voltage signal by a factor $C_{\text{DAC}} / (N C_{\text{DAC}} + C_{\text{samp}} + C_{\text{par}})$, where C_{DAC} is the Thevenin equivalent capacitance of each multiplying DAC, C_{samp} is the sampling capacitance of the ADC reference DAC, and C_{par} represents all parasitic capacitance on the shared aMVM-ADC node. Typically, the multiplying and reference DACs are identical, $C_{\text{DAC}} = C_{\text{samp}}$, and the parasitics result from bottom-plate capacitance $C_{\text{par}} = \gamma(N + 1) C_{\text{DAC}}$ where $\gamma \approx 0.1$. Thus, the attenuation factor can be approximately expressed as $1 / (N + 1)(1 + \gamma)$. The ADC reference is also similarly attenuated, exacerbating the effect of the accumulating noise degrading the SNR increasing the already stringent ADC specifications further.

A gain element following aMVM and doubling as ADC driver counters this attenuation at an increase in system energy as illustrated in Fig. 2.5. In particular, restoring signal levels back to full-scale to reduce the DR burden of the ADC comes at the cost of increased complexity and power (P_{gain} in Eq. (2.7)) of the aMVM active gain stage. This energy cost for the gain stage may be substantial where the aMVM costs dominate; however, the ADC cost dominates at higher system dynamic range, more than amply amortizing the cost of energy required to provide the restorative gain (Eq. (2.7)).

2.3.3 Circuit Architecture

Here we briefly describe an example aMVM system for spatial signal conditioning in adaptive beamforming for RF communication [43]. The system implements analog preprocessing on the outputs of a harmonic rejection mixer (HRM) receiver, providing substantial processing gain prior to digitization. The 8×8 aMVM

in Fig. 2.6 is composed of complementary capacitive multiplying digital-analog converters (MDAC) as detailed in Fig. 2.7. Beamforming is implemented through digitally programmed transform coefficients. The resulting capacitive weighting spatially filters the incident signal from four antennas at baseband, implementing 4×4 complex matrix-vector multiplication with the 8×8 real array. The achieved 68 dB processing gain is substantially larger than the conventional approach [27]. The same capacitive array structure is used for both the MDAC and the feedback capacitor in the OTA, resulting in consistent 48 dB of programmable gain in steps of 6 dB.

An improved aMVM system, designed for energy efficient operation using the principles highlighted in Sections 2.3.1 and 2.3.2 has recently been demonstrated [37]. To this end an alternate capacitive MDAC topology is introduced to reduce the effective C_{DAC} capacitive load on the driving circuitry while maintaining high-resolution. As shown in Fig. 2.8 the minimum resolvable angle and the dynamic range in the resolution are determined by the resolution of the aMVM system. This system demonstrates state-of-the art interference suppression of 84 dB, corresponding to the implemented 14-bit weight resolution. Measurements for base-band signal separation, shown in Fig. 2.9, demonstrate broad-band signal resolving capabilities. In addition to the MDAC topology, offset cancellation at the MDAC and the OTA implemented via correlated double-sampling (CDS) also contributes to the achieved precision. CDS at 500 Hz periodically sets the input DC bias point of the capacitively coupled differential amplifier.

2.3.4 Alternative Architectures

A wide range of mixed-signal circuit architectures have been pursued for aMVM systems in a wide range of adaptive signal processing and machine learning applications. Most of these tightly couple the matrix-vector multiplication with the digitization circuitry in order to reduce the system energy. Here we highlight two main directions of current developments.
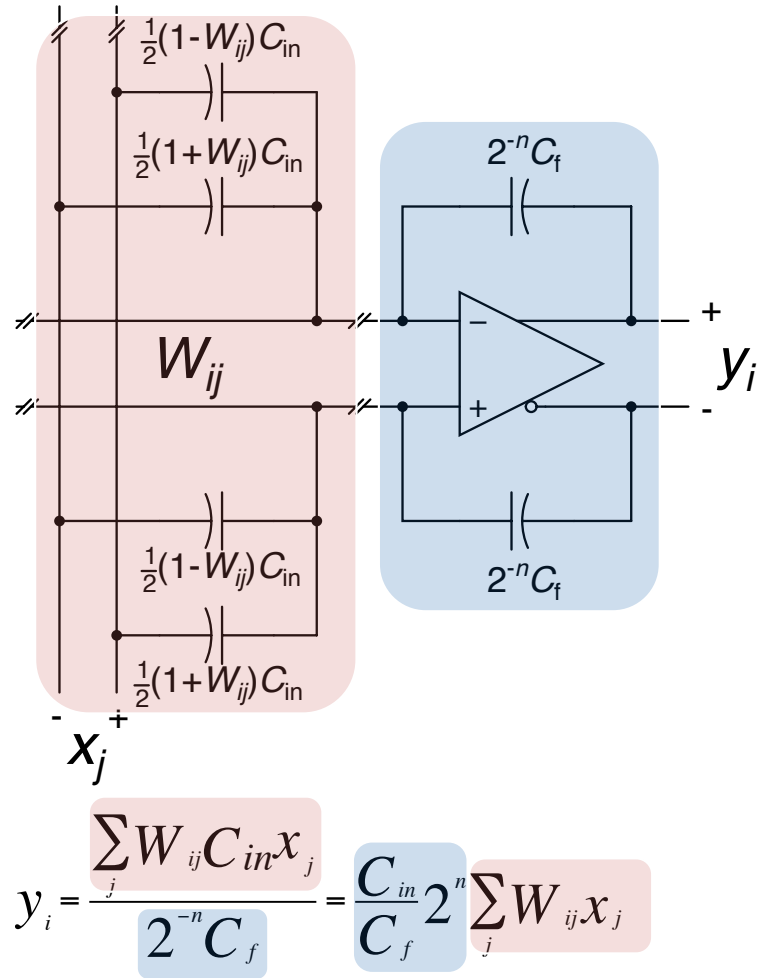
Figure 2.7: Capacitive high-resolution, high dynamic range digital-analog multi-plication. The parts of the circuit implementing weighting and gain in the linear transform are highlighted.

Nyquist Rate Systems

A multiplying-DAC embedded within a SAR feedback loop, targeting embedded sensing medical systems, is presented in [83]. A simplified diagram of the implemented architecture, performing consecutive linear-feature extraction and linear classification, is shown in Fig. 2.10. The recursive extraction-classification formulation results in a reduction in the number of operations required to implement ECG-based cardiac arrhythmia detection by a factor $85\times$, and image-pixel-based gender detection by a factor $200\times$. Employing *boosting* algorithms [82] to overcome limitations of linear classification as well as analog-mismatch, the improved
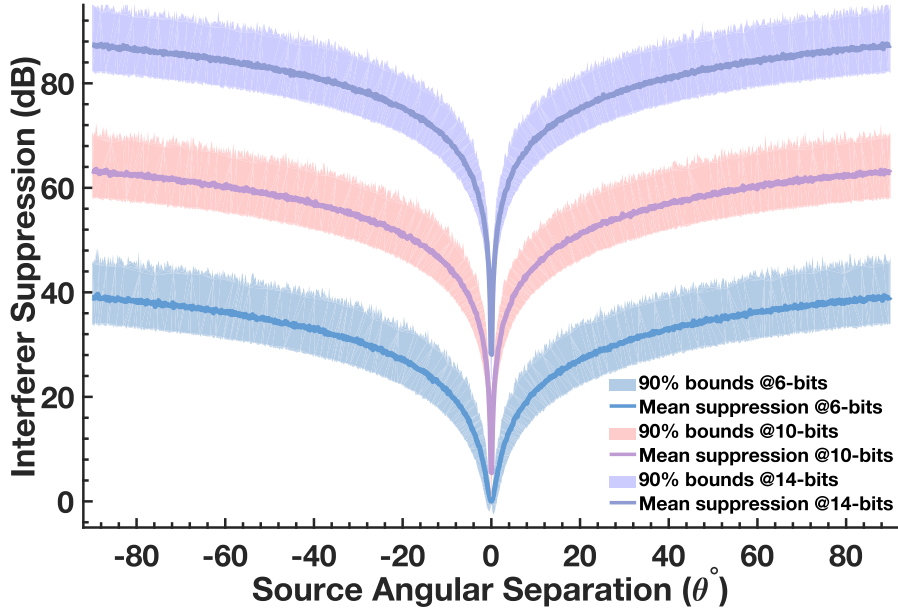
Figure 2.8: Effects of coefficient quantization on beamforming performance. Expected interferer suppression levels and 90% confidence bounds by Monte Carlo simulation at three levels of quantization: 6-bit, 10-bit, and 14-bit.

computational efficiency trades-off performance with energy.

The system implements a partial analog multiplication during the data-conversion step, with the residue multiplied digitally. Since analog multiplication results in an increase in the signal dynamic range, high-resolution multiplication is avoided. Furthermore, accumulation occurs digitally with no corresponding analog processing gain. This makes parallelism as described in Section 2.3.2 cost prohibitive, limiting the solution to a serial, albeit energy-efficient and highly configurable implementation.

Oversampling Systems

The principle of oversampling, widely used in overcoming the resolution limitations of Nyquist rate data-converters, has been extended to aMVM in a high-resolution, oversampling multiplying ADC presented in [7]. The prototype system embeds multiplication within a Delta-Sigma Modulator (DSM) at $100M$ 1-bit multiplications/s/channel. Owing to the oversampled respresentation, 14 effective bits of recognition accuracy have been achieved.
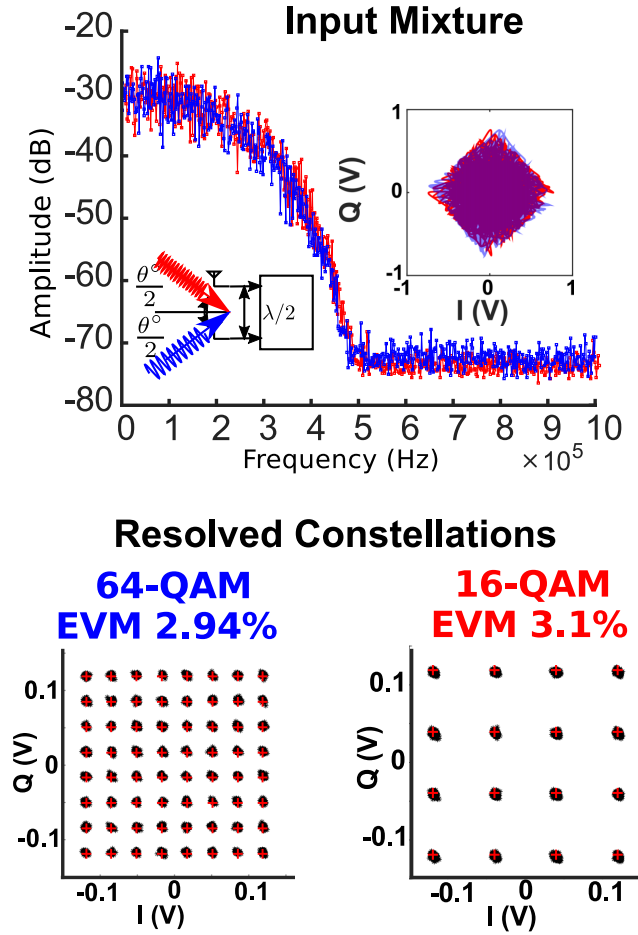
Figure 2.9: aMVM system [37] demonstrating state-of-the-art separation of signals with completely overlapping spectra. The system simultaneously separates 16-QAM and 64-QAM mixtures in two complex channels to less than 3.1% and 2.94% RMS-EVM.

A single-bit mixing sequence (multiplication with ± 1) is introduced within the DSM feedback loop as shown in Fig. 2.11, implementing a pass-through or inversion in the differential signal path. As such, the matrix-vector product is effectively this output sequence is accumulated digitally. In contrast to a conventional DSM, to first order the quantization error terms are canceled. Furthermore, arbitrary resolution can be achieved, in exchange for bandwidth, through appropriate sequencing of ± 1 values. The use of oversampling enables trading-off throughput with precision, via the appropriate oversampling factor.

Note that the energy efficiency analysis in Sections 2.3.1 and 2.3.2 assumed
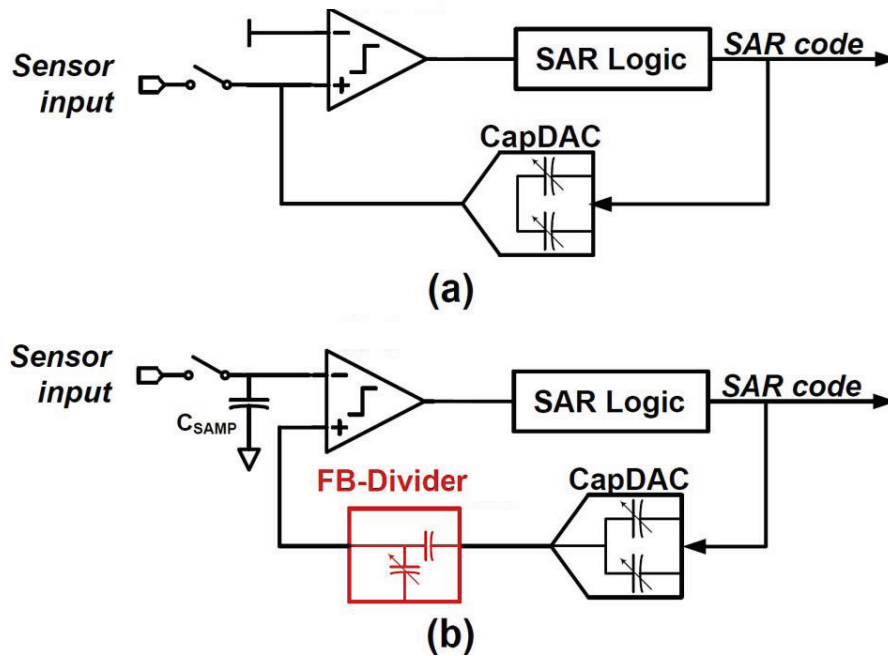
Figure 2.10: Block diagram showing (a) a conventional scalar SAR ADC and (b) the matrix-multiplying ADC (MMADC) system proposed in [83]. Loop-embedded passive capacitive division implements the product of a feedback factor and the input. This system implements partial multiplication in the analog domain, with multiplication on the residue and accumulation implemented in the digital domain.
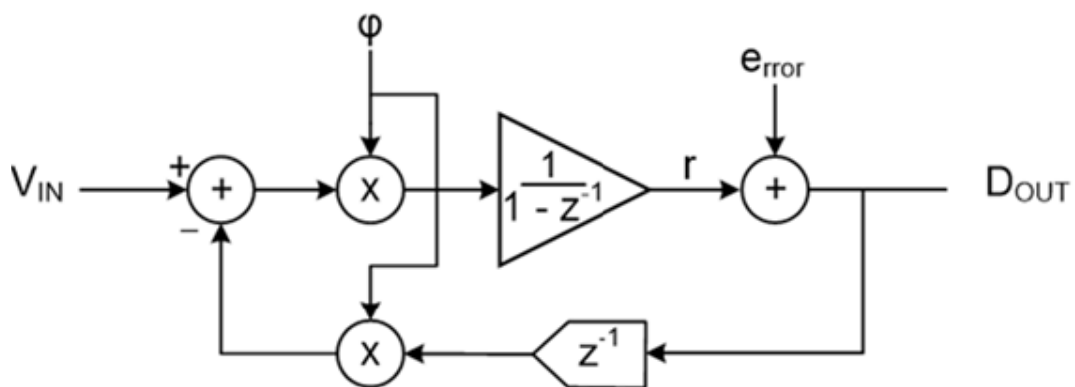
Figure 2.11: A first order model of the multiplying ADC proposed in [7].

Table 2.1: Comparison of analog signal conditioning systems

	Zhang et al. ISSCC 2015	Lee et al. ISSCC 2016	Buhler et al. VLSI 2016	Kim et al. JSSC 2016	Joshi et al. ISSCC 2017
Application	Feature Extraction	Sensor Classifier	Feature Extraction	Spatial Filtering	Linear Spatial Filtering
CMOS Technology (nm)	180	40	65	65	65
Number of channels	1^a	1^a	16^a	8	8
Area per MAC (mm²)	0.106	0.012	0.0594	0.045	0.021
Power (µW)	0.663	228	3856	1300	91
Signal Bandwidth (kHz)	10	10^6	100	1500	350
Power/Bandwidth (µW/MHz)	66.3	.228	38560	866	260
Effective Analog Multiplicand (bit)	4	3	14^b	8^c	14
Multiply Accumulate Efficiency (pJ/MAC)	16^d	.12	30000^d	6	2
Multiply Accumulate Efficiency /Multiplicand Level (fJ/MAC/Level)	1000	15	1830	23.4	0. 12

[a]Serial matrix-vector product. [b]Oversampled, 1-bit per sample.
[c]Reported 48 dB signal separation. [d]No analog accumulate.

Nyquist rate systems using SAR ADCs; it is however straightforward to extend the quantization model in the analysis to oversampling systems [6].

A comparison between the various aMVM systems presented above, with key performance metrics, is given in Table 2.1. The range of trade-offs in terms of energy, resolution, speed and efficiency demonstrates wide applicability of aMVM systems as alternatives to DSP in adaptive signal processing and machine intelligence.

2.4 Post-digitization

In a break from the trend of increasingly digital ICs, there has also been an argument for implementing charge-domain MVM for accelerating more conventional deep-learning systems as an alternative to DSP, GPU, and FPGA comput-

ing. The signal flow in these MVM systems operating on data in externally digital but internally analog form entails a DAC front-end feeding a capacitor array performing aMVM, followed by ADC digitization. High parallelism in the array in conjunction with low-resolution D/A and A/D conversion leads to marked energy benefits [69].

2.4.1 Analog Machine Learning Accelerators

Lee et. al. [48] present a switched capacitor matrix-multiplying ADC, that exploits matrix factorization to introduce a $64\times$ analog processing gain, enabling energy savings. 3-bit capacitive weights and 6-bit D/A and A/D interfaces accommodate the signal dynamic range modification after analog multiplication. Since analog accumulation occurs at a rate much slower than analog multiplication, this allows the digitization to occur at a rate much lower than the computation, hence lowering ADC energy and size costs. In contrast, a parallel switched capacitor medium-resolution system is demonstrated [4] for efficient implementation of machine learning algorithms. These two systems aim at accelerating deep learning algorithms and networks similar to [18, 45]. The focus on large-scale machine-learning as an application results in co-optimization and collocation of memory and computation as highlighted in Section 2.5.

2.4.2 Charge Recovery and Adiabatic Computing

A critical assumption in deriving the minimum energy bounds in (2.6) is that charge on a capacitor, and the associated $CV^2/2$ energy, cannot be recovered between samples.

Physical principles like adiabatic and reversible computing attempt to subvert this bound using energy recovery techniques that cannot be applied to conventional CMOS logic design. These circuits suffer from two sources of energy consumption beyond leakage: *adiabatic* losses, and *non-adiabatic* losses. *Non-adiabatic* losses are those due to incomplete recovery of supplied charge, while *adiabatic* losses are closely related to dynamic power consumption $P_{\mathrm{dyn}} = fCV^2$

in CMOS circuits. Optimal adiabatic charging implemented through a constant current source [61] aims at reducing dissipated energy, trading it with time and complexity. Consider that charging a capacitor C through a resistance R by applying a constant current I for a time T dissipates

$$E = RI^2T$$

energy through the resistor. This entails a net charge transfer of $Q = IT = CV$, resulting in net energy loss

$$E = \frac{R(CV)^2}{T} = \frac{\tau}{T}CV^2$$

with time constant $\tau = RC$. Conventional CMOS line drivers $T = \tau$ incur complete loss of CV^2 electrostatic energy in each charge cycle. Slowing the charging time $T \gg \tau$ results in substantial energy savings, tending to zero energy in the adiabatic limit $T \to \infty$. However, efficient generation of a slow current ramp and constant charging and discharging currents incurs losses in the supply generator. Instead, waveforms from a resonator [71] are used to provide this ramp, trading-off lower supply generation complexity with increased losses in the compute circuitry [38].

Reversible computing avoids *non-adiabatic* losses, ensuring preservation of information states to prevent the dissipation of energy upon erasure of information [22, 49]. High-density mixed-signal adiabatic processors [25, 38, 39] offering high-dimensional charge domain low-precision analog computation for kernel-based pattern recognition have been developed using these principles. These "Kerneltron" aMVM processors implement an externally digital, internally analog matrix-vector product for use in classification and pattern matching tasks [24]. Charge injection device (CID) arrays with DRAM storage elements store each bit of the matrix element as shared charge. The computed matrix-vector product occurs through non-destructive charge sensing on the bit/compute line (BCL) in a bit-serial, matrix-parallel fashion. Unlike conventional switched-capacitor charge-domain analog computing, the conservation of charge in the CID array throughout the computational cycle allows significant energy savings. A further boost in energy efficiency is obtained by a stochastic encoding and decoding scheme, which ensures a constant capacitive load of the CID array tuned for resonance in energy
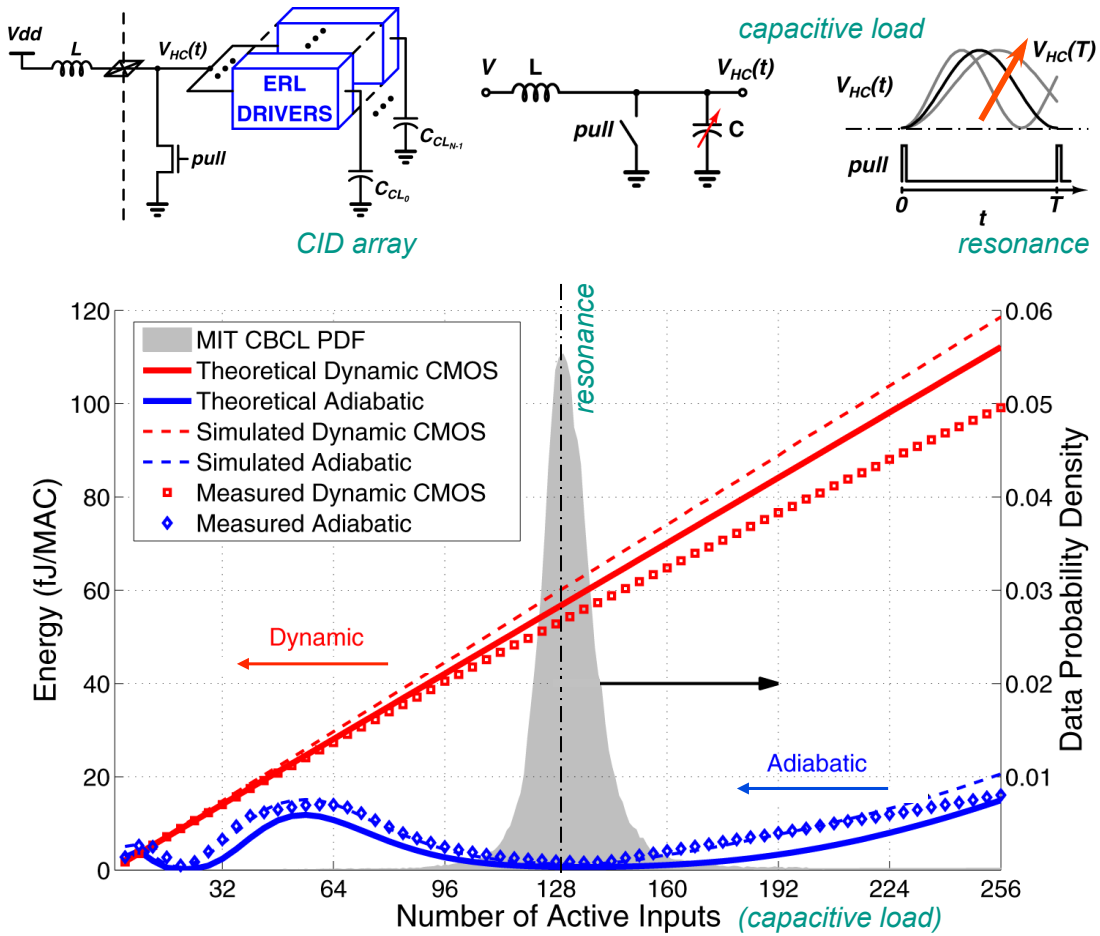
Figure 2.12: Summary of operating principles for the mixed-signal resonant adiabatic processor in [39]. Power is dissipated on charging and discharging the total capacitance of the array select lines, and the sum of the capacitances of the CID column. Conventional static CMOS logic drivers dissipate all fCV^2 power in driving the capacitive load of the CID array. This challenge is addressed through the use of energy recovery logic (ERL) resonant drivers coupling the CID capacitive load to an inductor to recirculate the energy in driving the CID array.

recovery in the array drivers. Fig. 2.12 summarizes the operation and performance of such adiabatic resonant energy-recycling in the Kerneltron, resulting in better than 1.1 TMACS/mW efficiency excluding on-chip digitization [39].

While there are several drawbacks of adiabatic energy-recovery aMVM in the complexity of the required circuitry, the low operational speed, the unaccounted for cost of digitization, and the susceptibility to variability, the prospects look

bright for the adoption of such techniques in specialized processors in the face of the inevitable end to CMOS scaling. In principle, fundamental limits on the efficiency of reversible computing extend beyond the *Landauer's principle* [78] which provides a link between the thermodynamic and information theoretic measures of entropy and the minimum bounds on energy required for *irreversible* computation, $E_{\min} = k_B T \ln 2$. Although the prospects of pushing fundamental energy limits are tempting, currently realized levels of energy efficiencies with adiabatic energy-recycling aMVM computing in the sub-fJ per operation range [39] are still far from the Landauer limit, which may only be achieved with probabilistic forms of computation able to operate effectively at near-unity signal-to-noise ratios.

2.5   Emerging Devices

Not only are machine learning algorithms compute and memory intensive [30], they face severe scaling penalties due to the Von-Neumann bottleneck. Emerging resistive memory technologies have the potential to enable processing-in-memory architectures bypassing this bottleneck [46]. Furthermore they facilitate 3D integration [72], with the gradual resistance change of these devices enabling a single cell to be used as analog memory.

Two-terminal resistive memory devices such as phase change memory (PCM) [46], resistive switching memory (RRAM) [59], conductive bridge memory (CBRAM) [58], and ferroelectric memory (FeRAM) [14], provide several advantages over conventional memory systems–superior scaling, low-energy programming, and non-volatile analog storage. Analog memory enables in-memory computing capabilities (similar to capacitive arrays in Section 2.3) where the matrix-vector product is performed in distributed parallel manner directly within the memory, avoiding the data movement typical in DSP, CPU and GPU systems and thus reaping major energy benefits [8].

Despite these advantages, several open problems and challenges remain in integration with conventional CMOS digital systems. Although $R_{\text{off}}/R_{\text{on}}$ ratios in the 10^3–10^4 range have been demonstrated, the performance severely degrades for
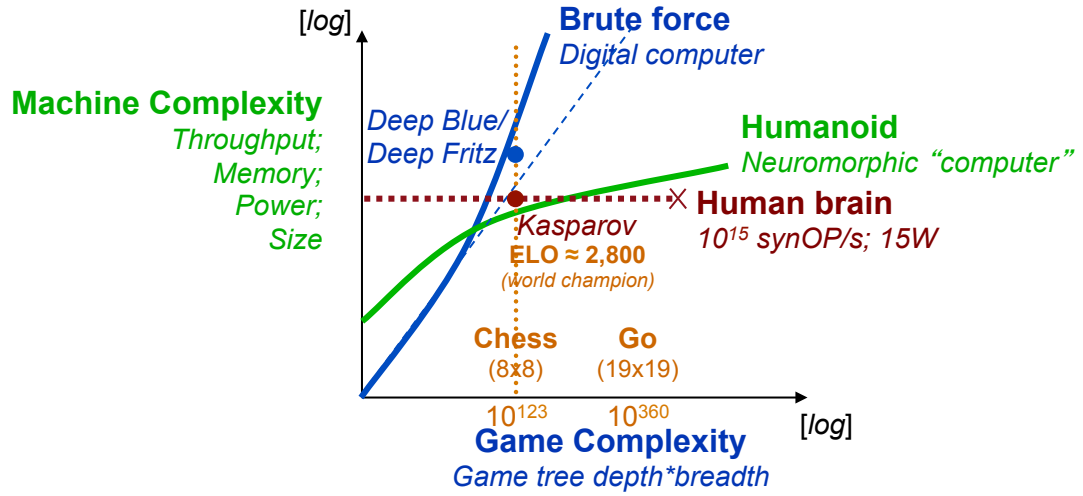
Figure 2.13: Advances in computing hardware will increase computing and communication efficiency; Similar advances in training methods, and supporting data, are required to adaptively reduce algorithmic complexity (adapted from [10]).

large arrays, with diminished yields [15]. Emerging memory devices also suffer from inherent device-to-device and cycle-to-cycle variations restricting their use to highly fault-tolerant algorithms. The large fan-out and fan-in in contemporary machine-learning algorithms necessitates large memory array sizes where energy to drive the array dominates the net energy [8]. Moreover, large arrays also suffer from increased IR drops along bit and word lines. Efforts to combat such IR drops with increased wire thickness result in decreased densities, increased capacitance and energy losses. Alternative architectures composed of smaller arrays require more communication between cores, yielding diminished energy savings [15].

Continuing research into materials [86], improved device modeling, and innovative use of circuit techniques introduced in Section 2.4.2, as well as architectural advances show promise in overcoming the aforementioned limitations.

2.6 Future Prospects

As we are seeing an unprecedented growth in the capabilities [74] of automated systems [21], increased autonomy will require more complex interactions with the environment through sensors and actuators as well as complex commu-

nication of information to remote locations, all while minimizing energy use. The circuit techniques and systems presented here have leveraged a diverse range of analog mixed-signal implementations of matrix-vector multiplication to enable a variety of adaptive signal processing and machine learning tasks. Such systems contribute at all levels in the signal pipeline, from sensory conditioning [7,37,83], local processing [39], and communication [43], to high-performance accelerators [8,48].

With increasing demands on such automated systems executing complex tasks, advances in neural networks and neuromorphic computing as well as innovations in computing architecture look poised to shift the trade-off between machine and task complexity in favor of greater efficiency and efficacy in computational systems [10] (Fig. 2.13). Noted by Von Neumann 50 years ago [81], computer design can draw inspiration from biology, where intelligence emerges from extremely efficient and resilient collectives of imprecise, and unreliable analog components. 50 years later and counting, this observation remains equally applicable to mixed-signal and analog systems designed for the next-generation of computational loads.

Acknowledgment

Chapter 3

A 6.5 μW/MHz Charge Buffer with 7 fF Input Capacitance in 65 nm CMOS for Non-contact Electropotential Sensing

3.1 Introduction

Emerging sensor technologies have greatly expanded our capability to sense and intelligently adapt to the environment [77,89]. Their ease of high density integration within a standard CMOS process makes capacitive sensing particularly attractive. Electric potential sensors for signal detection and communication can use such integrated capacitive sensors in a variety of contexts including bio-potential measurements [16,64], active proximity sensing, body channel communication [34] and capacitive near-field communication and human computer interaction [29]. However the design of interface circuits for such capacitive sensors is challenging for broadband operation especially when considering noise, parasitic loading, and power. Ideally electropotential sensors should not load the node where voltage is sensed and draw no current from that node disturbing the measurement. Thus, a major requirement from these interface circuits is a very high, ideally infinite, input
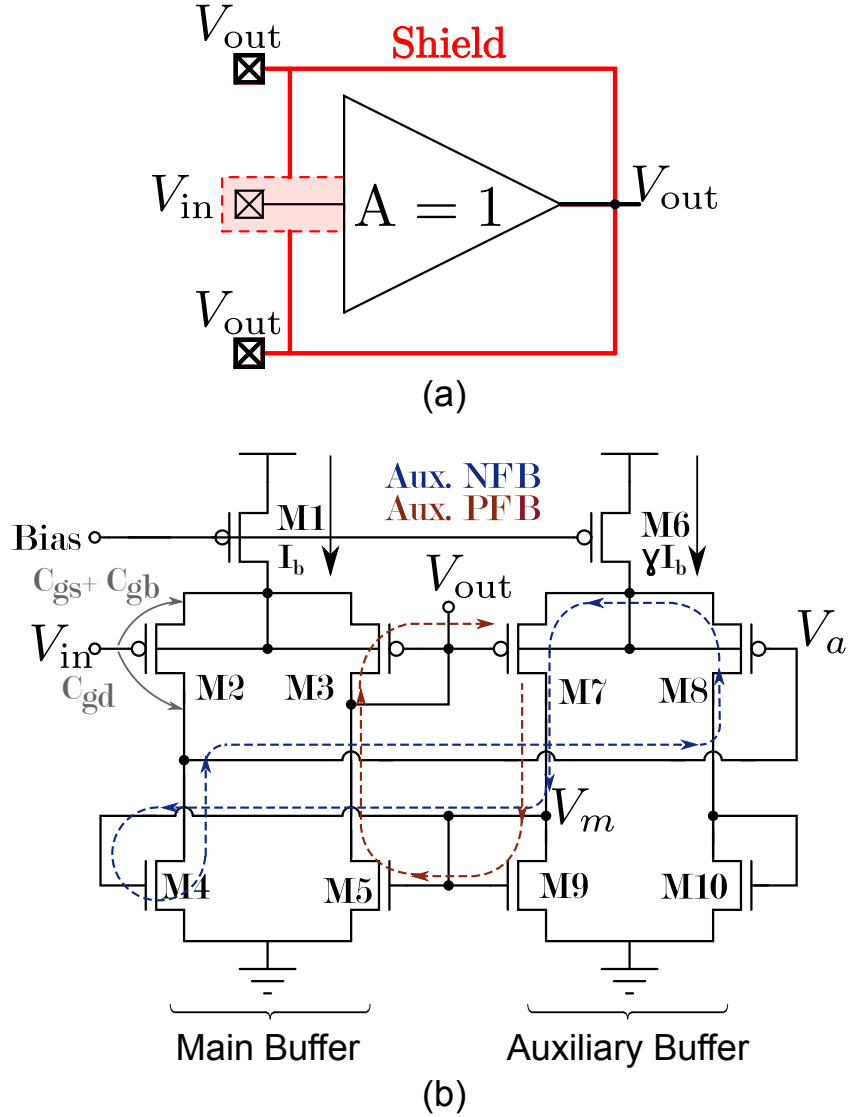
Figure 3.1: Unity gain charge buffer for capacitive non-contact electropotential sensing. (a) Unity gain active shielding of the entire input signal path to negate the effect of input capacitance. (b) Integrated implementation using nested main and auxiliary unity gain buffers.

impedance specification. Especially for capacitive sensing, this entails minimizing the input capacitance while maximizing the input resistance.

Although high input impedances, with input capacitances around 2-5 pF, have been reported, e.g., [29], these specifications are insufficient for broadband non-contact electropotential sensing. Additionally, discrete implementations report greatly reduced Signal-to-Noise and Distortion Ratios (SNDRs), as low as

10 dB [29] at cm distances, requiring system resources dedicated to de-noising the input. Since typical signals of interest are low magnitude, in the μV to lower mV range up to 20 MHz [47], the sensory front-ends must provide low parasitic input capacitance and high input resistances to enable accurate measurement.

MOSFET input impedances are typically degraded by parasitic effects at the gate, including leakage currents, the effect of parasitic capacitances at the input, and DC operating point biasing currents. The traditional solution to increasing the effective input impedance is to use bootstrapping [28], which exploits positive feedback in order to increase the effective impedance between two nodes at the expense of noise and stability. Integrated implementations enable alternative negative feedback to internal amplifier nodes that can vastly improve performance and stability. To this end, we present a unity-gain charge buffer for use in electric field sensing capable of driving a 2 pF load at 6.5 μW/MHz energy efficiency, occupying 193 μm^2 in 65 nm CMOS.

3.2 Circuit Design and Analysis

Fig. 3.1 shows the schematic of the presented charge buffer, which can directly implement a unity-gain active shield for non-contact biopotential sensing [16]. Active shielding enables cancellation of parasitics along the signal path as shown in Fig. 3.1 (a). However, active shielding loads the output of the buffer trading the cancellation of capacitances along the signal path against the bandwidth of unity-gain buffering. By virtue of unity gain, the bulk-source connected PMOS differential pair input stage leads to cancellation of the gate-to-source and gate-to-bulk input parasitic capacitances, C_{gs} and C_{gb} respectively. We mitigate the effect of the residual gate-to-drain input parasitic capacitance C_{gd} via a high loop-gain negative feedback through an auxiliary buffer effectively bootstrapping it, while also improving the main buffer voltage transfer function to better approach unity gain by negating the effect of drain conductance on the V_a node, which in turn improves the bootstrapping through the active shield.

Hence bootstrapping the input capacitance requires high gain to be effec-
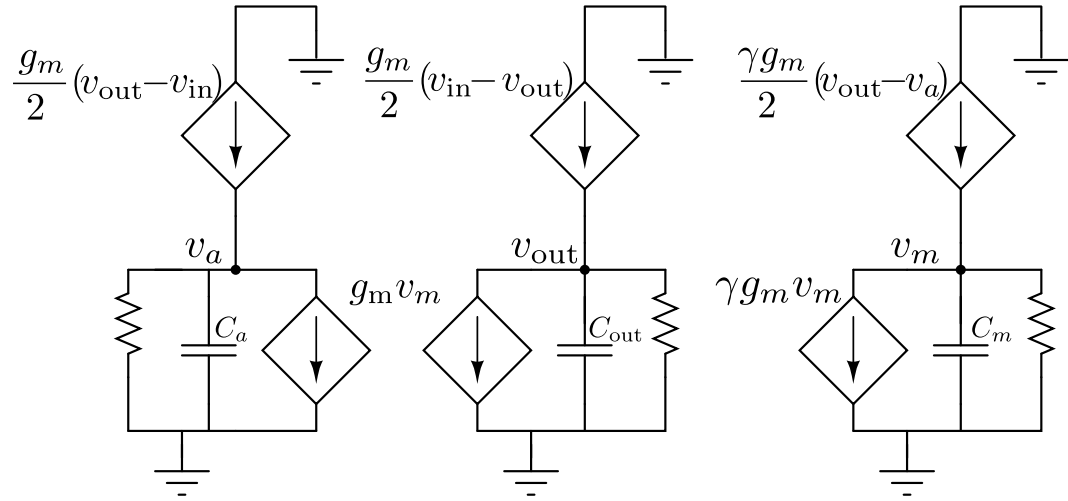
Figure 3.2: Small-signal model of the charge buffer of Fig. 3.1 (b). Here γ is the ratio of operating currents in the auxiliary and main buffers as annotated in Fig. 3.1 (b), C_a and C_m are parasitic capacitances at the nodes V_a and V_m respectively, and C_{out} is the buffer output load capacitance.
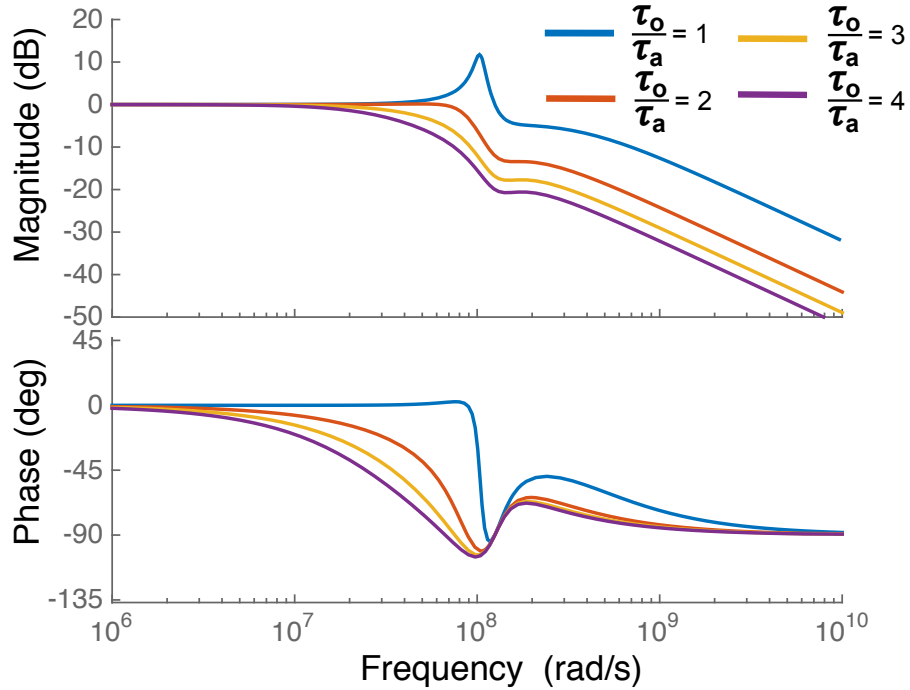
Figure 3.3: Bode plot demonstrating theoretically predicted variation of damping with varying ratios of time constants in the circuit. Simulation parameters used: $\gamma = \frac{1}{4}$, $\tau_a = 13.33$ ns, and $\tau_m = 4\tau_a$ (i.e., $C_m = C_a$).
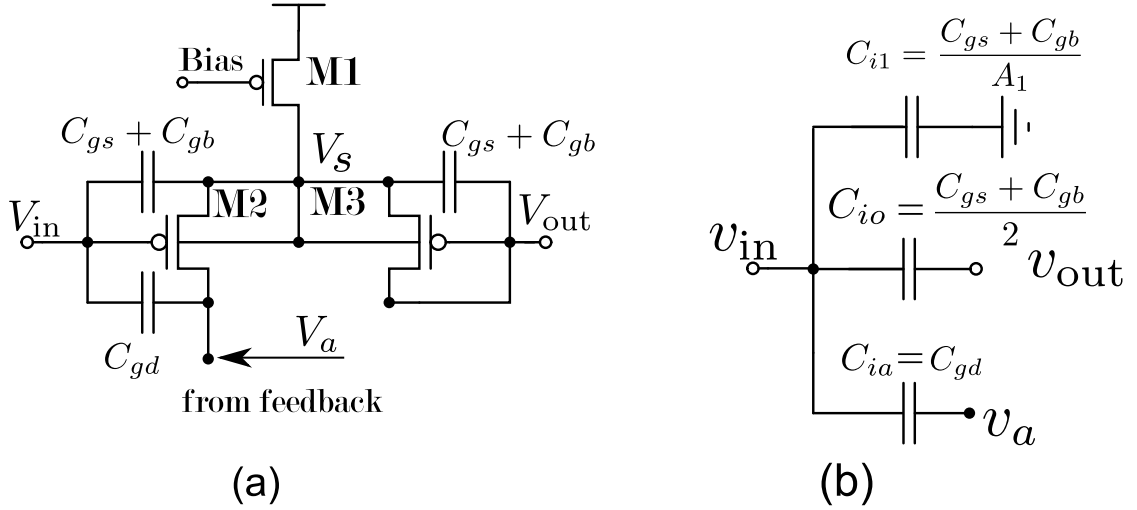
Figure 3.4: (a) The small signal model used for input impedance analysis. (b) Reduced Thevenin small-signal equivalent model with effective capacitances between the input v_{in}, output v_{out}, and feedback auxiliary node v_a.

tive [16]. This gain is provided by the negative feedback loop (NFB) composed of M8, M9, M4, M8 annotated in Fig 3.1 (b). This loop actively negates the parasitic input capacitance, C_{gd}, of the input differential pair M2-M3. Thus, the circuit buffers the input such that any change in voltage at the gate of M2 results in an equivalent change at both the source and the drain of the transistor preventing the flow of current from the gate to those nodes, maintaining high input impedance.

To quantify the effect of circuit parameters on DC and spectral response, we conduct a small-signal analysis on the nodes V_{out}, V_a and V_m, annotated in Fig. 3.1 (b) with small signal variables v_{out}, v_a and v_m. Here, γ represents the ratio of the current in the auxiliary buffer to that in the main buffer, where $\gamma \leq 1$ for power savings. In subthreshold, the transconductance of M2 and M3 is $g_m \approx \kappa q I_b / 2kT$ and the drain conductance of M4 and M5 is $g_d \approx I_b / 2V_{\mathrm{Early}}$, with back-gate coefficient κ and Early voltage V_{Early}, and where the corresponding conductances of M7-M10 are approximately scaled by γ. We ignore drain conductances of M1-M3 and M6-M8 which cancel out to first order (in the limit $v_{\mathrm{out}} \approx v_a \approx v_{\mathrm{in}}$). Considering the output load capacitance C_{out} and parasitic capacitances C_m and
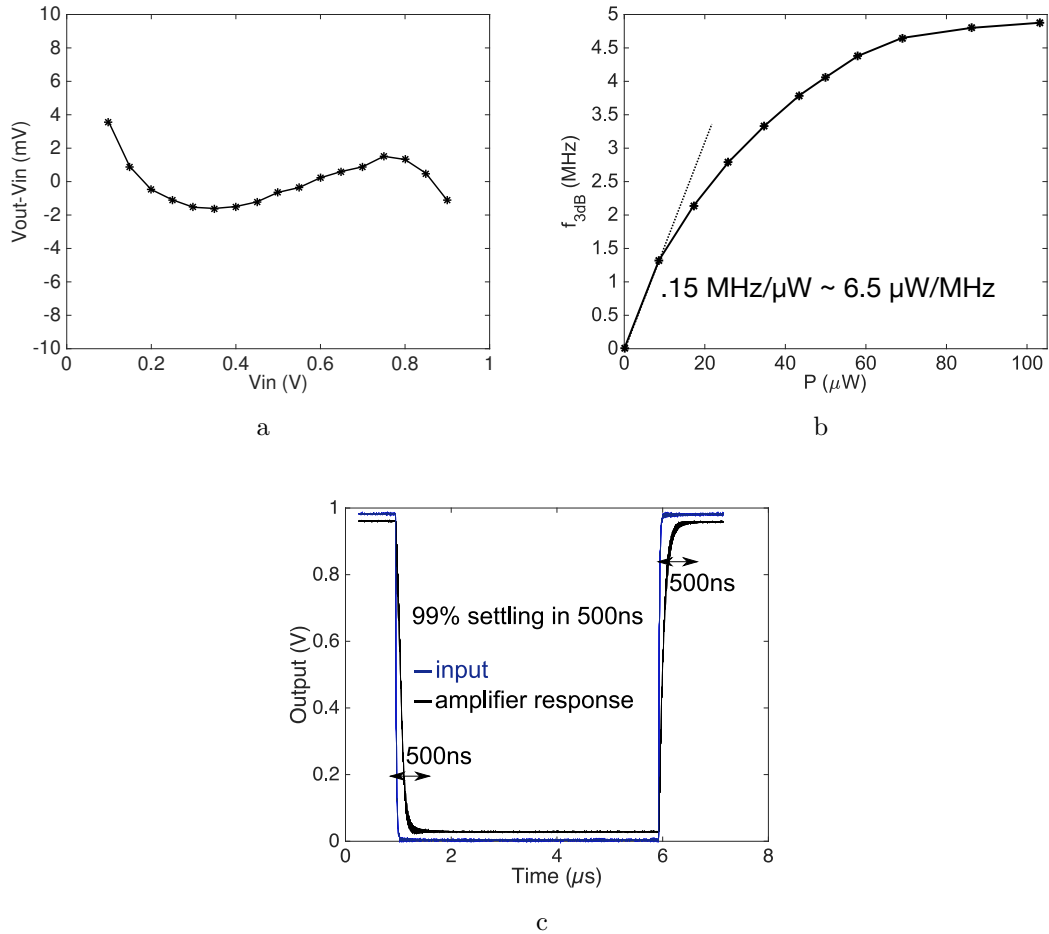
Figure 3.5: Measured linearity and dynamic response characteristics of the fabricated charge buffer. a The measured DC gain error is less than .6% over a .9 V range, while consuming less than 5 µA of current. b The -3 dB frequency achieved by the buffer at various levels of power consumption yields an efficiency figure-of-merit of 6.5 µW/MHz, at a midband frequency of 50 kHz. The linear bandwidth response saturates to 5 MHz for higher power levels. c First-order settling within 500 ns is demonstrated for a step response matching analytical results from expressions in (3.3).

C_a on the internal nodes V_a and V_m leads to three state equations

$$C_{\text{out}}\frac{\mathrm{d}v_{\text{out}}}{\mathrm{d}t} = \frac{1}{2}g_m(v_{\text{in}} - v_{\text{out}}) - g_m v_m - g_d v_{\text{out}}$$
$$C_a\frac{\mathrm{d}v_a}{\mathrm{d}t} = -\frac{1}{2}g_m(v_{\text{in}} - v_{\text{out}}) - g_m v_m - g_d v_a \qquad (3.1)$$
$$C_m\frac{\mathrm{d}v_m}{\mathrm{d}t} = -\frac{\gamma}{2}g_m(v_a - v_{\text{out}}) - \gamma g_m v_m - \gamma g_d v_m$$

with resulting DC gain

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{1 + \frac{1}{A_0}\left(1 + \frac{1}{A_0}\right)}{1 + \frac{1}{A_0}\left(1 + \frac{1}{A_0}\right)\left(1 + \frac{2}{A_0}\right)} \approx 1 - \frac{2}{A_0^2}, \qquad (3.2)$$

where $A_0 = g_m/g_d$ is the open-loop gain of the main and auxiliary buffers. We further define following characteristic time constants: $\tau_o = C_{\text{out}}/g_m$, $\tau_a = C_a/g_m$, and $\tau_m = C_m/\gamma g_m$. In the limit of infinite gain i.e., $g_d/g_m \to 0$ the AC transfer function reduces to

$$\frac{v_{\text{out}}(s)}{v_{\text{in}}(s)} = \frac{1 + \tau_a s + \tau_a \tau_m s^2}{1 + \tau_o s + (2\tau_o\tau_a + \tau_a\tau_m)\, s^2 + 2\tau_a\tau_m\tau_o s^3}. \qquad (3.3)$$

Since the coefficients of the denominator are all strictly positive, the poles are all contained in the left half plane and the amplifier is unconditionally stable. However, $\tau_o \simeq \tau_a, \tau_m$ can result in a poor phase margin as seen in Fig 3.3. The response for $\gamma = \frac{1}{4}$ is critically damped for $\tau_0 \simeq 2\tau_a = \frac{1}{2}\tau_m$, and overdamped for $\tau_o \gg \tau_a, \tau_m$. Hence first-order settling is observed for relatively large capacitive loads $C_{\text{out}} \gg C_a, C_m$.

According to the input equivalent circuit shown in Fig. 3.4 (a), the input admittance (the reciprocal of input impedance) can be expressed in terms of the parasitic capacitances of M2 in Fig. 3.1 as

$$Y_{\text{in}}(s) = \frac{s(C_{gs} + C_{gb})\,(v_{\text{in}} - v_s) + sC_{gd}\,(v_{\text{in}} - v_a)}{v_{\text{in}}}. \qquad (3.4)$$

The dynamics of V_s, the common source of M2 and M3 coupling to the drain of M1, can be modeled to first order as a small-signal dependence $v_s \approx \frac{1}{2}(v_{\text{in}} + v_{\text{out}})\,/\,(1 + g_{d1}/g_m)$, where g_{d1} represents the drain conductance of M1. For large source-coupling gain $A_1 = g_m/g_{d1}$ the input admittance (3.4) can be approximately

written as shown in Fig. 3.4 (b),

$$Y_{\text{in}}(s) \approx sC_{i1} + sC_{io}\left(1 - \frac{v_{\text{out}}}{v_{\text{in}}}\right) + sC_{ia}\left(1 - \frac{v_a}{v_{\text{in}}}\right) \tag{3.5}$$

with equivalent internal parasitic capacitances $C_{i1} = (C_{gs} + C_{gb})/A_1$ on the input, $C_{io} = (C_{gs}+C_{gb})/2$ between the input and output, and $C_{ia} = C_{gd}$ between the input and auxiliary node. Accounting for the additive effect of the parasitic capacitances C_{io} and C_{ia} onto the node capacitances C_{out} and C_a in the internal dynamics (3.1) gives similar simplified expressions for DC and AC gains as in (3.2) and (3.3) leading to

$$Y_{\text{in}}(s) \approx \left(C_{i1} + \frac{2}{A_0{}^2}(C_{io} + C_{ia})\right)s \\ + \frac{C_{io}(C_{\text{out}} - C_a) + C_{ia}(C_{io} - C_{ia})}{g_m}s^2 + \dots$$

with effective input capacitance (i.e., coefficient in s^1)

$$C_{\text{in}} \approx \left(\frac{1}{A_1} + \frac{1}{A_0{}^2}\right)(C_{gs} + C_{gb}) + \frac{2}{A_0{}^2}C_{gd}. \tag{3.6}$$

The resulting input impedance

$$llr\,Z_{\text{in}} \approx \frac{1}{sC_{\text{in}}}\frac{1}{1 + \tau_{\text{in}}s} \tag{3.7}$$

is predominantly a first-order capacitance response with cut-off frequency $1\,/\,2\pi\tau_{\text{in}}$ approximately $\left(1/A_1 + 1/A_0{}^2\right)/\pi\tau_o$, assuming $C_{\text{out}} \gg C_a, C_{io}, C_{ia}$ and $C_{gs} \gg C_{gd}$ and τ_{in} is the time constant associated with C_{in}. The effect is a substantial reduction, around 20 dB, in parasitic input capacitance at low frequencies through the source-coupling gain A_1 and cascaded open-loop feedback gain $A_0{}^2$.

3.3 Measurement Results

Benchtop measurements validating unity-gain functionality of the charge buffer, fabricated in 65 nm bulk CMOS with an active area of 9.1 μm × 21 μm, are presented in Figs. 3.5a, 3.5b and 3.5c. We used two Keithley 2400 source meters to measure the DC gain error. Providing highly accurate inputs with one while
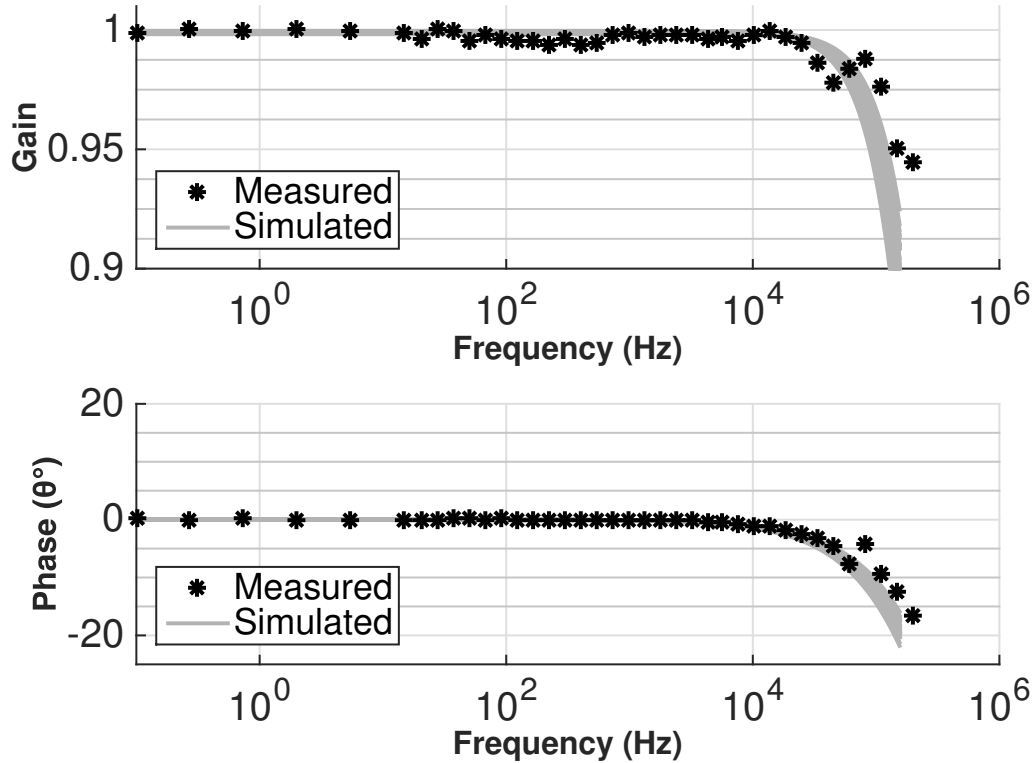
Figure 3.6: Correspondence between Monte Carlo simulated and measured transfer function of the designed amplifier. The AC gain and phase error is lower than .6% while consuming 2.5 μA from a 1.2 V supply. The -3 dB frequency lies beyond the measurement capabilities of the instrument.

measuring the output of the buffer with the other, a second set of measurements were performed to cancel offsets between the instruments. The measured gain error shown in Fig. 3.5a demonstrates greater than .6% accuracy over a voltage range of 900 mV from a 1.2 V supply. Fig. 3.5b highlights the energy efficiency of the buffer and Fig. 3.5c verifies the first order settling of the amplifier in response to a large input step. The buffer demonstrates 99% settling well within 500 ns matching theoretical results derived in (3.3). The transfer function measurement shown in Fig. 3.6 was performed using a Signal Recovery model 7265 (SR 7265) lock-in amplifier directly interfacing with the buffer. Due to limitations in the instrumentation, measurements are not available over the entire frequency range of the buffer.
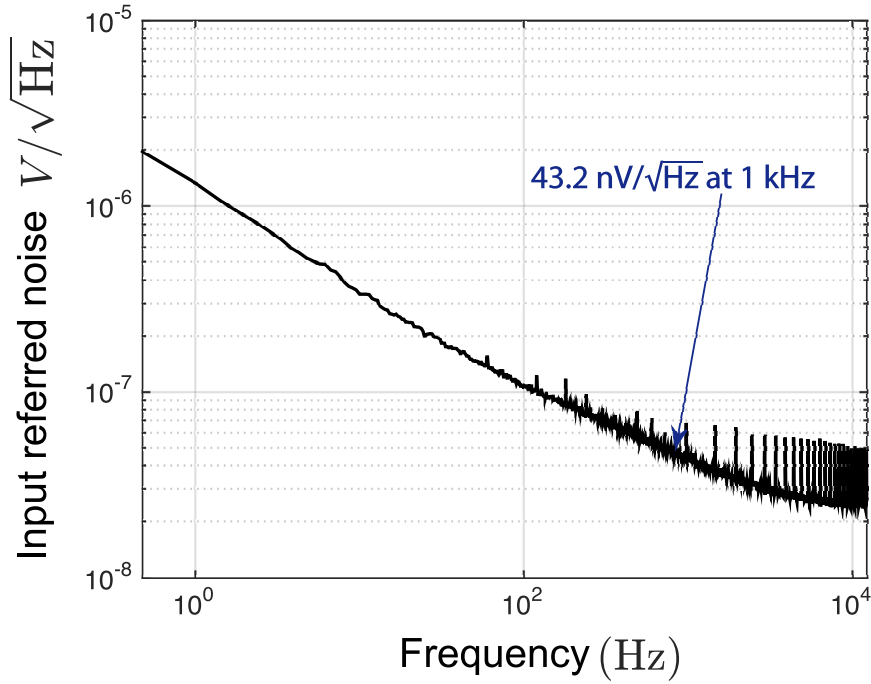
Figure 3.7: The noise power spectral density measured from 1 mHz-25 kHz. ADC measurement artifacts introduce additional tones, however, the measured spot noise at 1 kHz is 43.2 nV/$\sqrt{\text{Hz}}$.

The measured noise including noise from off-chip driving buffers was determined using a 4 Msps, 24-bit Analog to Digital Converter (ADC) (TI ADS-1675) with the sampling set to 125 ksps, maximizing the number of noise-free bits. The inputs of the buffer were driven to a known DC voltage and the ADC measured the resultant output, shown in Fig. 3.7. The corresponding Noise Efficiency Factor (NEF),

$$\text{NEF} = \frac{V_{\text{rms}}}{kT}\sqrt{\frac{q\,I_{\text{tot}}}{2\pi\,\text{BW}}}$$

is 3.13 over a bandwidth BW of 25 kHz. We observed 43.2 nV/$\sqrt{\text{Hz}}$ spot noise at 1 kHz comparing favorably with state-of-art integrated implementations [16]. These measurements were performed with the buffer power tuned to 5.3 μW.

We measure the input capacitance using the SR 7265 to provide an input to the buffer through an on-chip coupling capacitor designed to have a nominal value of 256 fF. Assuming a 10% tolerance on the capacitance, a worst case capacitance

Table 3.1: Measured characteristics and comparison of electropotential sensing amplifiers

		[16]	[64]	[62]	This work
Power	[μW]	4.95	70,000	-	8.8
Bandwidth efficiency	[μW/MHz]	49.5	-	-	6.5
Total integrated input-referred noise over bandwidth	[μV/$\sqrt{\text{Hz}}$]	0.0085	-	0.1	0.015
Input-referred noise spectral density at 1 kHz	[nV/$\sqrt{\text{Hz}}$]	45	100	-	43.2
Gain	[V/V]	1	0.7	0.01 to 1	0.995
Input capacitance C_{in}	[fF]	60	1000	30000	7
Technology (CMOS)	[nm]	500	discrete	350	65

of 307 fF can be assumed. We measured the gain by directly driving the buffers after bypassing the on-chip coupling capacitors. The ratio of the gain through the coupling capacitor in contrast with directly driving the buffer provides a measure of the capacitive division between the on-chip capacitor and the parasitic capacitance, resulting in 7fF of measured input capacitance.

The over-the-air test setup used for non-contact electric field measurements and for Near Field Communication (NFC) measurements is shown in Fig. 3.8. The potential difference between the sensing plates due to the field induced is buffered and digitized. Since the area of the sensing plates is much smaller than the area of the driven plates, the parasitics at the input can greatly degrade the signal. Measurements for a single sinusoid in Fig. 3.9 and multiple types of communication signals in Fig. 3.10 validate the broad-band sensing capabilities of this buffer. A 1.1 MHz Frequency Shift Keying (FSK) signal as well as a 1.0 MHz Amplitude Modulation (AM) signal were applied to the driven plates resulting in an induced field of 1.1 V/m. These distances are much shorter than the wavelengths of transmitted signal, and operation is entirely in the near-field regime. The received communication signal is given in Fig. 3.10 with >40 dB of SNR, demonstrating near-field capacitive communication capabilities. From these measurements we project an effective electric field sensitivity of 100 μV/m at unity (0 dB) SNR. In addition to non-contact electropotential sensing, the charge buffering unity-gain feature of the circuit has been further validated and tested as an integral part of communication systems [42].
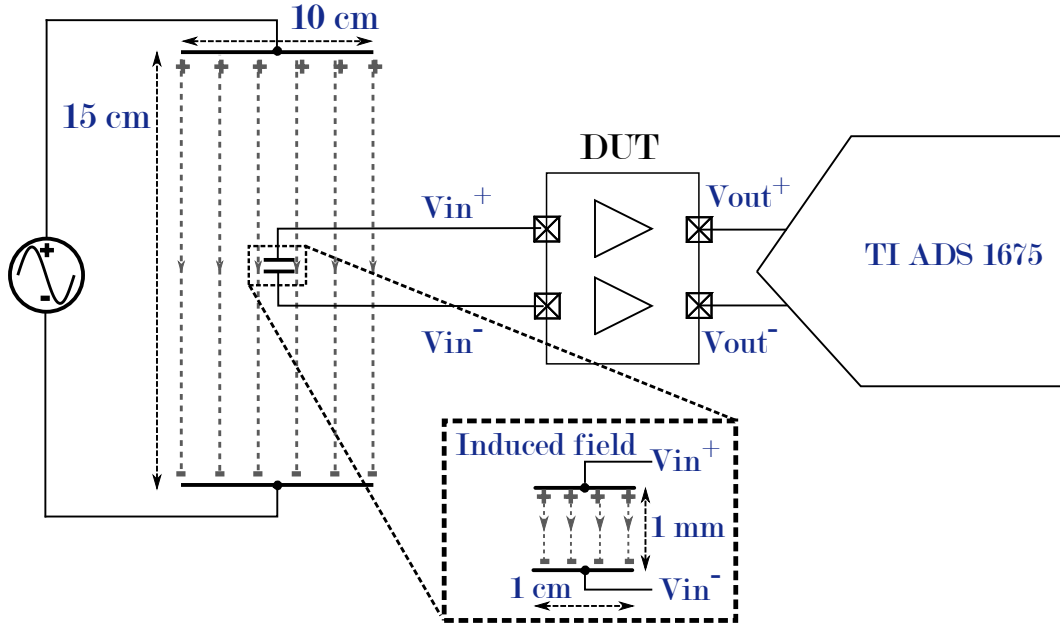
Figure 3.8: Experimental setup used for over-the-air non-contact electropotential sensing experiments, a high resolution 24 bit ADC digitizes the signals for further analysis.

3.4 Conclusions

The architecture and implementation of a fully integrated circuit for unity voltage-gain charge buffering has been presented. The architecture actively cancels the parasitic capacitances at the input of the buffer via two-stage negative feedback in unity-gain voltage buffering and active shielding. Measurements demonstrate state-of-the-art performance in input capacitance and power efficiency. Table 3.1 summarizes the performance of the charge buffer circuit in the context of related work [16,29,64]. We validated the non-contact electropotential sensing capabilities of the charge buffer, demonstrating >50 dB SFDR for a single tone demonstrating electric field sensitivity of 100 μV/m at a detection threshold of 0 dB SNR. The capacitive near-field communication capabilities of the system were validated via over-the-air experiments with both AM and FSK modulated communication signals received at high fidelity. Applications of the charge buffer in high density integrated circuits range from non-contact electropotential sensing to low-power capacitive near field communication.
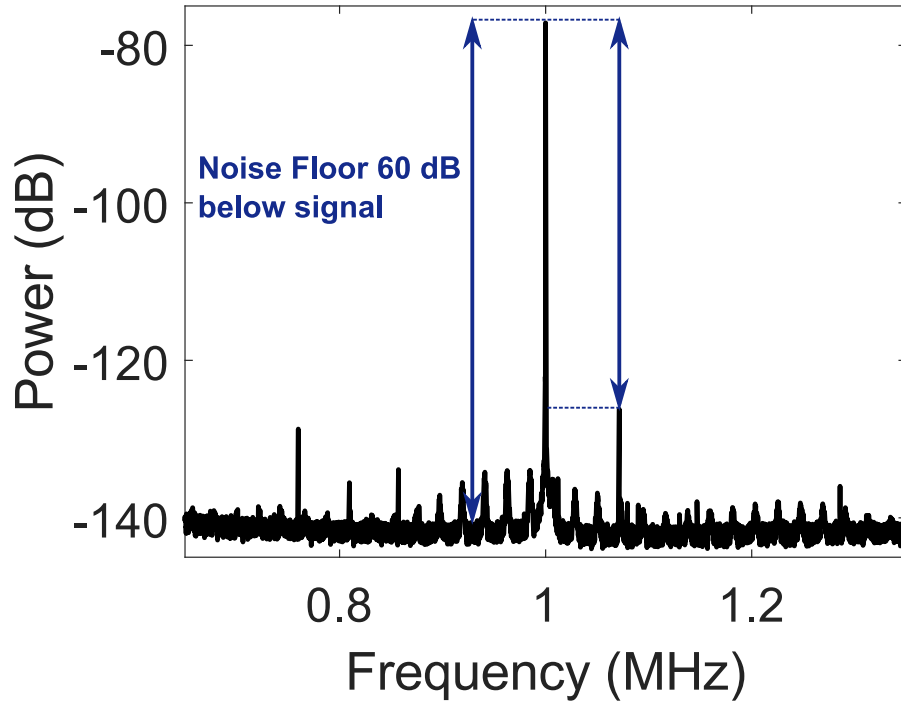
Figure 3.9: 1 MHz received tone sensing a field of 1.1 V/m with SNR >40 dB over a distance of 7.5 cm. The corresponding sensitivity of the non-contact sensor, as the extrapolated signal level at 0 dB SNR, is approximately 100 μV/m.

3.5   Acknowledgements

Figure 3.10: Received AM signal with f_c 1.0 MHz and modulation depth 25% and received FSK signal, with f_{hop} 0.831 MHz and 1.4 MHz with 100 kHz switching rate. These signals are received over a distance of 7.5 cm demonstrating suitability for use in capacitive near field communication (NFC).

Chapter 4

2 pJ/MAC 14-b 8×8 linear transform mixed-signal spatial filter in 65 nm CMOS with 84 dB interference suppression

4.1 Introduction

Advances in machine learning (ML) and the internet-of-things (IoT) have resulted in a renewed interest in analog matrix-vector multiplication (MvM) accelerators [7, 48, 83]. Classification based tasks have exploited low-to-medium resolution multiplication and accuracy boosting algorithms in order to compensate for the reduced resolution. Complementing classification, tasks like source separation and localization have diverse applications ranging from signal conditioning in communication [42] and ultrasound to electroencephalography (EEG) [52] source localization and spike sorting, and greatly benefit from similar algorithms. However, due to their lower resolution and limited channel count previously developed systems cannot be directly applied to this task. High-resolution analog multiplication introduces challenges that have limited prior work to less than 6-bit multiplication in the analog domain. Alternative approaches utilizing very high

oversampling result in very inefficient solutions. High precision in matrix multiplication can mitigate the effects of ill-conditioned (almost singular) matrices, as with beamforming separation of near-collinear sources, and other tasks incurring principal component analysis [3] or independent component analysis (ICA) [52]. As seen in 4.1, a large signal dynamic range at the input can result in an untenable dynamic range specification on the downstream dataconverters leading to greater than $10\times$ increase in power [4]. Thus, we present a multichannel multiple-input multiple-out (MIMO) mixed-signal linear transform system, with analog signal path and digital coefficient control, composed of an array of 14-bit Nested Thermometer Multiplying DACs (NTMDACs) implementing analog multiplication, and variable gain amplifier (VGA) implementing accumulation. We demonstrate state-of-the art performance on two tasks, spectrally oblivious interference suppression in communication signals and EEG signal separation.

4.2 Architecture

Figure. 4.2 presents the circuit architecture of one channel, or dot product unit (DPU), in the micro-power spatial processor (SP) with 8 such DPU implementing a general 8×8 real MvM. Each DPU is composed of 16 single-ended NTMDACs pairwise forming a pseudo differential structure. The outputs of these NTMDACs are summed onto the input node of a digitally controlled VGA. Owing to the large parallelism in passive capacitive multiplication, a gain stage is required in order to maintain SNR. The pseudo-differential structure of complementary pairs of NTMDACs presents a constant capacitive load to the input, and is used in conjunction with a digitally adjustable feedback capacitor across the VGA for precisely controlled variable gain independent of NTMDAC weights. Correlated double sampling serves the dual purpose of offset cancellation while periodically setting the DC operating point of the VGA, shown in Fig. 4.3. The measured transfer function of the full system demonstrating digitally programmable gain from -12dB to 30dB in steps of 6dB is shown in Fig. 4.2. Measured power vs. -3dB bandwidth indicates just 2 pJ of energy for each of the 64 multiply-accumulates
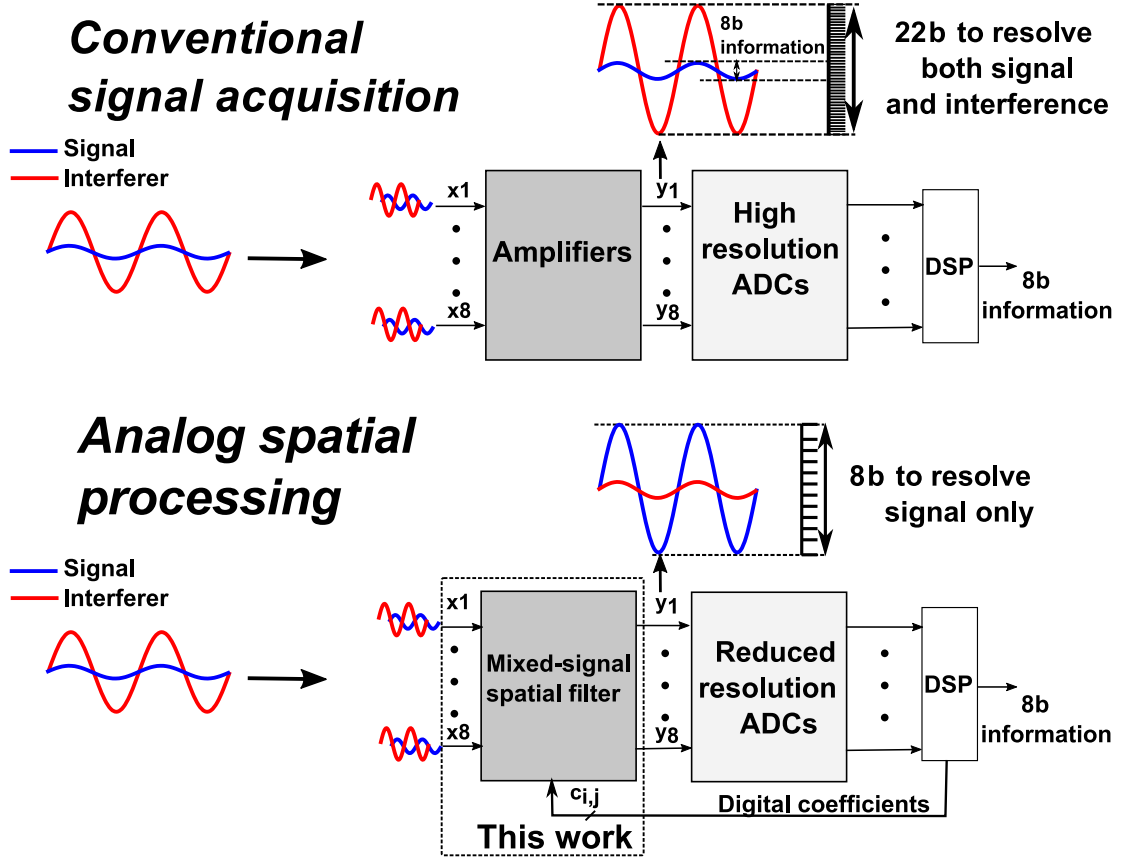
Figure 4.1: Interferer suppression and signal separation with analog signal processing and signal conditioning reduces the dynamic range requirements prior to digitization leading to system level energy benefits.

sampled at twice the bandwidth. Figure 4.3 illustrates the nested-thermometer coded operation of one single ended NTMDAC. The two-stage 7b+7b capacitive structure extends thermometer coding from the 7b LSB array to the 7b MSBs by stepping each MSB capacitor through all LSB levels prior to its full activation. This guarantees monotonicity across all intra-stage transitions, as required for use with online weight adaptation algorithms. A near-zero input capacitance unity gain buffer described in chapter 3, actively shields the LSB array feeding into the MSB array further enabling the use of a small unit capacitor, 2fF implemented as a $3.4\mu m \times 3.4\mu m$ custom shielded structure. Differential linearity measurements in Fig. 4.3 show NTMDAC monotonicity at 14b, as needed to support 84dB of interference suppression through spatial filtering.
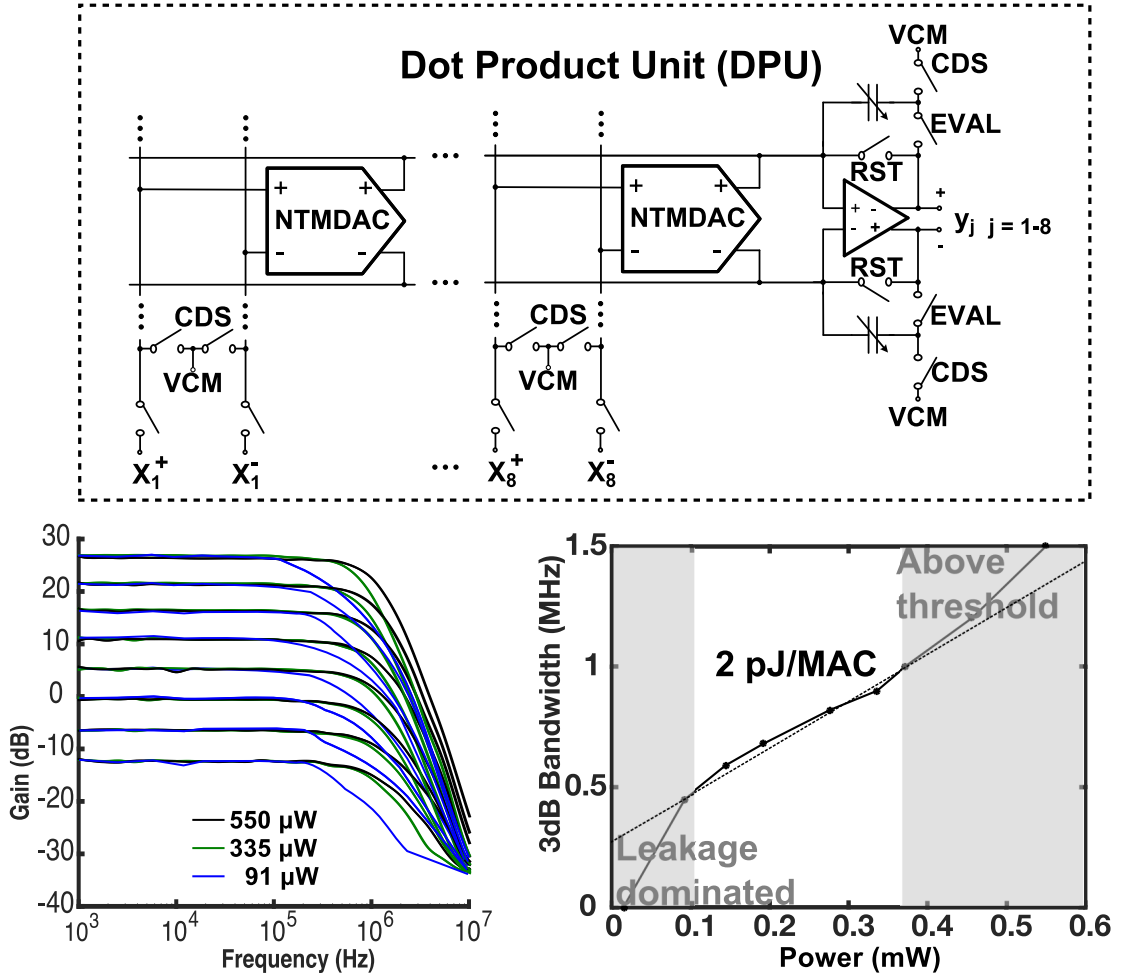
Figure 4.2: Circuit diagram of the proposed system and measured transfer function for various variable gain amplifier (VGA) gain and power settings. Bandwidth vs. power and energy efficiency per multiply-accumulate (MAC).

4.3 Measurements

To quantify spatial filtering performance in a typical beamforming application setting, we first observe the effect of beamforming coefficient quantization in Fig. 4.4. Figure. 4.5 then characterizes signal separation by the SP in the presence of ill-conditioned signals. The desired analog signal can be ill-conditioned due to two primary factors: signal to interferer ratio (SIR) of received power, and inter source angular spread. Measurements of SP performance against these validate the 14b analog multiplication while Fig. 4.4 demonstrates the need of such high
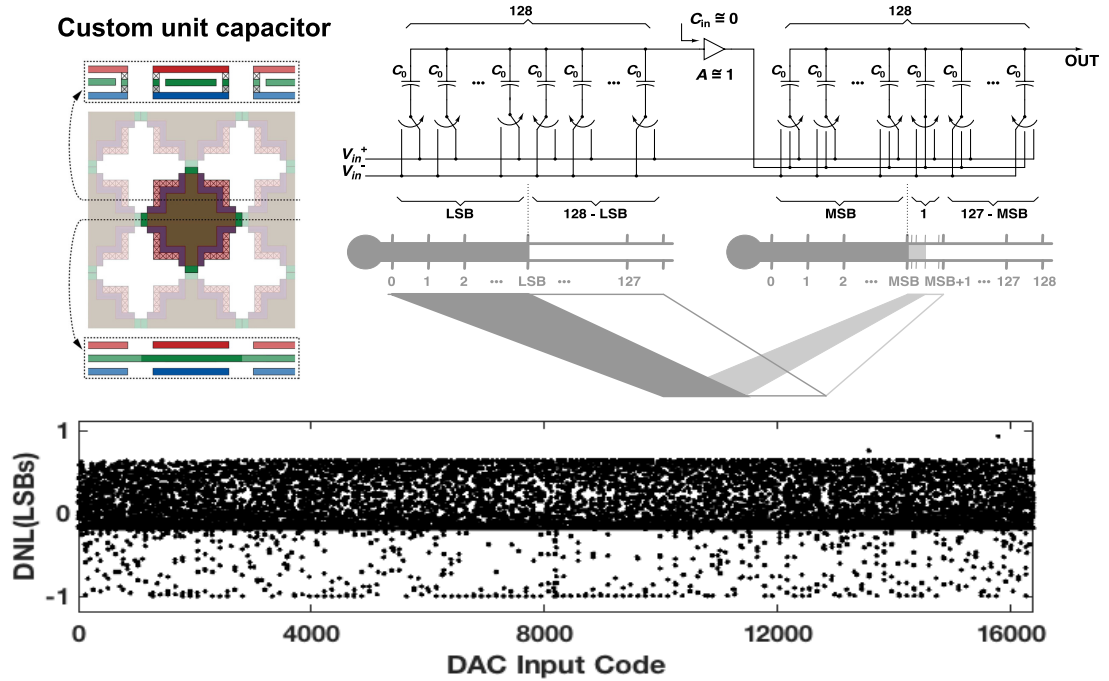
Figure 4.3: Nested-thermometer coded multiplying digital-to-analog converter (NTMDAC). Principle and circuit diagram along with top-view and cross-section view of custom 2 fF unit capacitor structure. DNL curve for the NTMDAC over all codes at the 14-bit level.
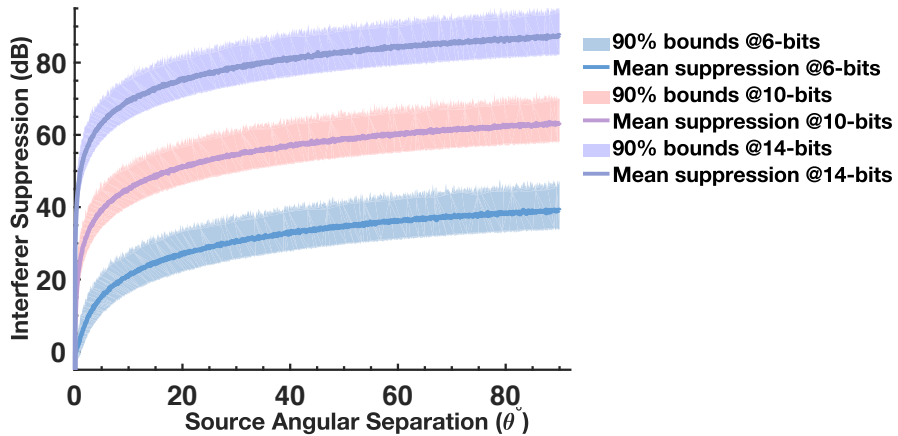
Figure 4.4: Effect of coefficient quantization on beamforming performance. Expected interferer suppression levels and 90% confidence bounds by Monte Carlo simulation at three levels of quantization: 6-bit, 10-bit, and 14-bit.
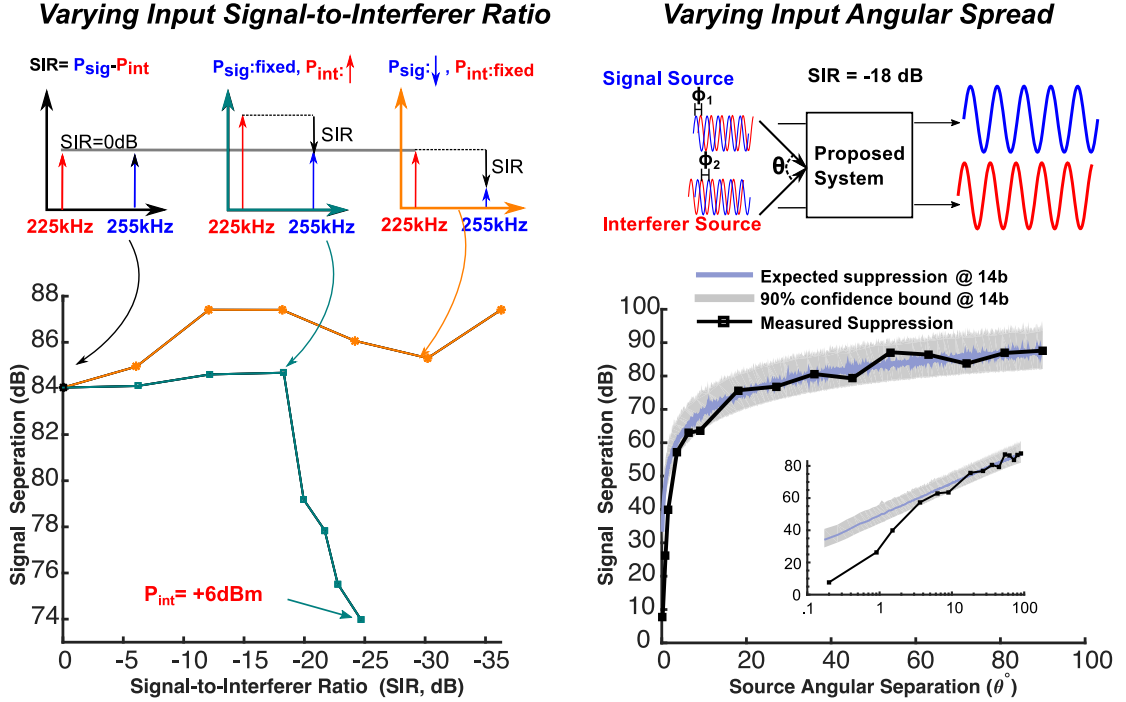
Figure 4.5: Measured interference suppression vs. signal-to-interferer ratio (SIR) and angular spread (θ) between signal and interferer sources. Comparison with Monte-Carlo simulated ideal 14-bit weights.

resolution when presented with an ill-conditioned mixture. Figure 4.5 shows the effect of varying SIR for complex mixtures of a sinusoidal interferer at 225kHz and a signal at 255kHz, at fixed 90 incident angle with weights for maximum suppression determined using an online algorithm. Consistent suppression $\geq$ 84dB is measured over 36dB of variation in signal power with varying SIR. Increasing the interferer power while keeping the signal power constant results in a decrease in performance due to clipping at the input switches. Even so, an interferer with an input power of +6dBm was suppressed to below -73.9dB with no effect on system gain. Also shown in Fig. 4.5, we measured interference suppression against source angular separation demonstrating high spatial resolution with $\geq$ 60dB suppression over angles ranging from 90° to 9°, and $\geq$ 25dB down to 1°, consistent with Monte Carlo simulated performance at ideal 14b weight resolution. A loss of performance at angles $<$ 1° is observed due to the finite gain of the system. Two example spatial filtering applications spanning the spectrum of the SP are highlighted in
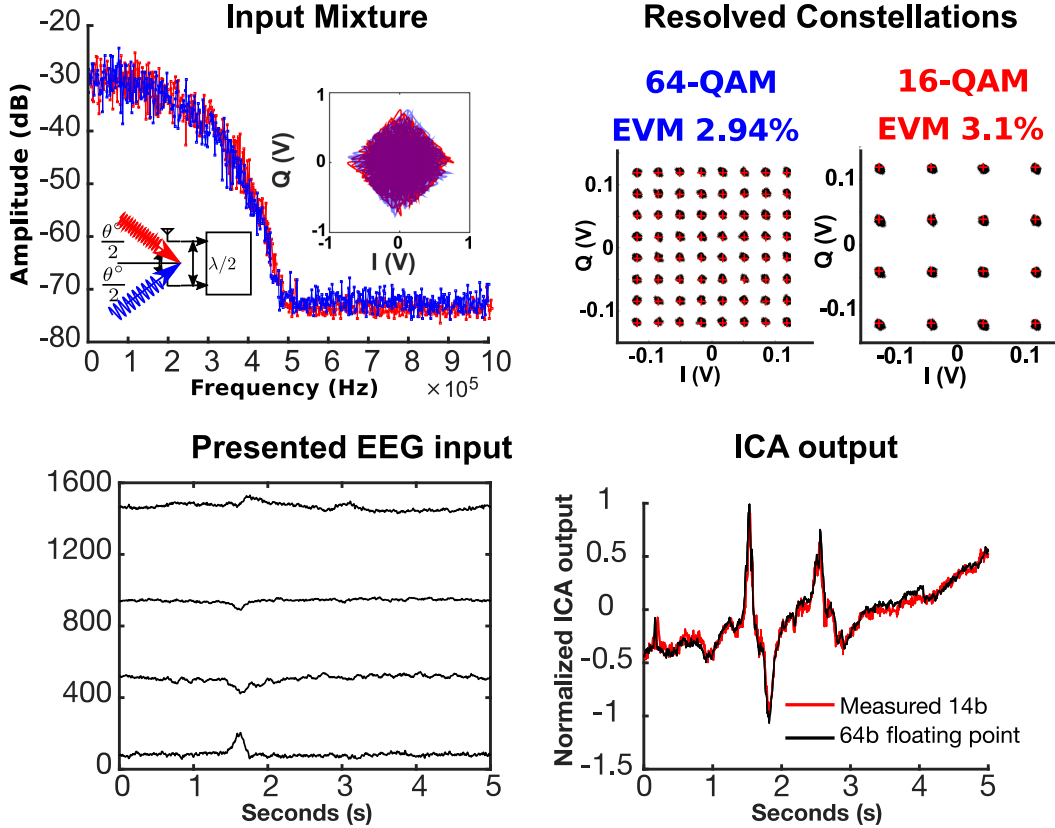
Figure 4.6: Simultaneous resolution of 64-QAM and 16-QAM modulated spectrally indistinguishable signals with a resultant Error Vector Magnitude (EVM) of 2.94% and 3.1%. Measured results from independent component analysis (ICA) on EEG signals with ICA along with the results from ICA with 64-bit floating point weights.

Fig. 4.6. Complex I/Q mixtures of 500 ksps 64-QAM, and 400 ksps 16-QAM signals emulating two-antenna RF spatial diversity were presented as four-channel input to the SP for baseband signal separation. Despite full spectral overlap, SP simultaneously resolved the two signals with a measured RMS EVM of 2.9% for 64-QAM and 3.1% for 16-QAM, corresponding to a BER better than 10^{-6}. This facilitates an increase in bandwidth, especially for in sensor networks where interference from nearby nodes might burden the already constrained power. We demonstrate the flexibility of SP by processing EEG signals. Reconstituted 500-samples/s 24b recordings of resting-state EEG from 4 channels of a dry-electrode headset were presented to the SP for spatially resolved separation of sources of

Table 4.1: Comparison of state of the art mixed-signal matrix-vector multiplication systems

	Zhang ISSCC 2015	Lee ISSCC 2016	Buhler et. al. VLSI 2016	Kim et. al. JSSC 2016	This work
Application	Feature Extraction	Sensor Classifier	Feature Extraction	Spatial Filtering	Linear Spatial Filtering
CMOS Technology (nm)	180	40	65	65	65
Number of channels	1[a]	1[a]	16[a]	8	8
Area per MAC (mm²)	0.106	0.012	0.0594	0.045	0.021
Power (μW)	0.663	228	3856	1300	91
Signal Bandwidth (kHz)	10	10^6	100	1500	350
Power/Bandwidth (μW/MHz)	66.3	.228	38560	866	260
Effective Analog Multiplicand (bit)	4	3	14[b]	8[c]	14
Multiply Accumulate Efficiency (pJ/MAC)	16[d]	.12	30000[d]	6	2
Multiply Accumulate Efficiency /Multiplicand Level (fJ/MAC/Level)	1000	15	1830	23.4	0. 12

[a]Serial matrix-vector product.   [b]Oversampled, 1-bit per sample.
[c]Reported 48 dB signal separation.   [d]No analog accumulate.

brain activity by ICA [52]. The differences between off-line computed ICA for one output component, and the SP output for 14b quantized digital weights, is well within 1%.

4.4   Conclusions

A comparison with state-of-the-art systems for mixed-signal matrix-vector multiplication is tabulated in Table 4.1, and spatial filtering and interference suppression is compiled in Table 4.2. A key advantage of the system is its low power consumption while maintaining high analog multiplication resolution without the use of costly oversampling techniques; this makes it suitable for adoption in emerging smart IoT devices. The micrograph of the system, implemented in 65nm CMOS with active area of 1.7 mm², is shown in Fig. 4.7.

Table 4.2: Comparison of state of the art spatial filtering and interference suppression systems

	Tseng et. al. JSSC 2010	Ghaffari et. al. JSSC 2014	Kim et. al. JSSC 2015	This work
Received EVM (dB)	-25	-	-28.8	-30.8
Effective number of bits	5	5	8	14
Angular Resolution (°)	22.5	22.5	<5[a]	<1[a]
Interferer Cancellation (dB)	30[b]	15[b,c]	48[b]	>80[b]
CMOS Technology (nm)	90	65	65	65
Power at Baseband (mW)	10[d]	68-195[e]	1.3	0.396
Bandwidth at Baseband (MHz)	20	5	3	2.4

[a]Greater than 15 dB cancellation, [b]Cancellation at 45° angular separation, [c]Out of beam, [d]LO power only, [e]Total power reported baseband power not reported
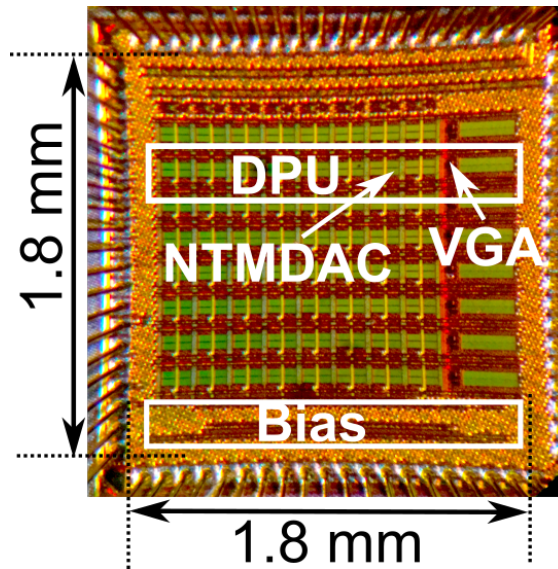
Figure 4.7: Die photograph (65nm CMOS).

4.5 Acknowledgements

Chapter Four is largely a reprint of material that appeared in 2017 ISSCC

digest of technical papers: Siddharth Joshi, Chul Kim, Sohmyung Ha, Yu M Chi, Gert Cauwenberghs, "2pJ/MAC 14b 8×8 Linear Transform Mixed-Signal Spatial Filter in 65nm CMOS with 84dB Interference Suppression," to appear, *IEEE ISSCC Dig. Tech. Papers*, San Francisco, CA, Feb. 2017. The author is the primary author and investigator of these works.

Chapter 5

Experimental Validation of Spatial Filtering Baseband Processor

5.1 Spatially Aware Cognitive Radio

Fig. 5.1 shows the proposed spectrally and spatially aware receiver for Cognitive Radio (CR) composed of four parts: an antenna array providing spatial diversity, a RF front-end acquiring broad-band RF signals, a down-converting mixer and lowpass filter, and MIMO signal separator, followed by ADCs and DSP for
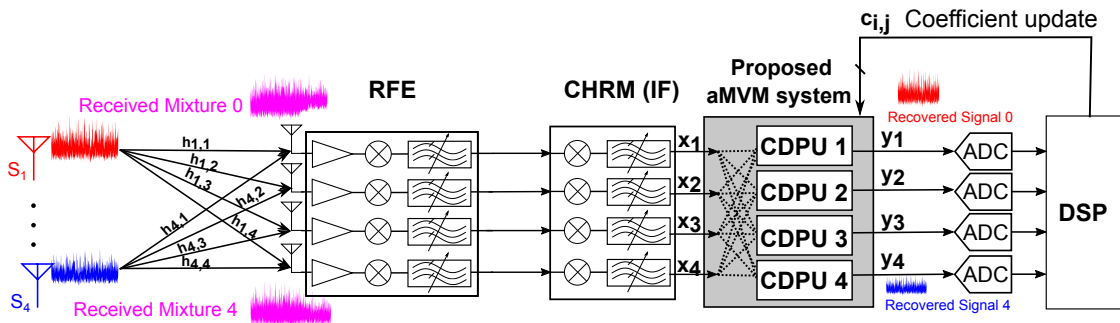
Figure 5.1: Combined spectrum and space aware cognitive radio with proposed MIMO baseband receiver. The highlighted spatial filter is the focus of this dissertation.

identifying signals, deciding usable RF bands, and updating digital weights for signal separation. This dissertation focuses on the MIMO spatial signal separator module.

A distributed geometry of multiple antennas receives a linear combination of signal sources $S_1, \ldots S_n$ present in the environment, each source and each antenna with a unique complex coefficient identifying mid-band amplitude gain and phase lag [32]. The amplitude and phase depend on the channel characteristics of each source as determined by attenuation and delay of its wavefront in relation to the antenna array. In turn, the amplitude and phase of the wavefront depend on frequency, and are approximately constant in a narrow frequency band. Hence, spectral and spatial diversity in the signal sources can be effectively leveraged by performing two-tiered spatio-spectral signal separation: first tuning to a subset of signal sources within a given spectral band, followed by MIMO signal separation specific to that band. The maximum number of narrowband sources that can thus be separated equals the product of the number of spectral bands (spectral multiplicity) and the number of antennas (spatial multiplicity). For broadband sources extending across multiple bands, the spectral components are separated separately in each band but can be identified and recombined for reconstruction, based on correspondence in angular or other spatial information derived from the MIMO weights.

The system in focus interfaces with an intermediate frequency (IF) down-converter that covers sixteen 3-MHz spectral bands spanning 48 MHz with a center frequency tunable from 100 MHz to 3 GHz, which in turn interfaces with a planar square array of four antennas. An RF front-end circuit at $f_{\text{LO\_RF}}$ is provided separately for each antenna, implemented on other integrated circuits using N-path tunable band-select filter [79], low-noise amplifier and RF quadrature mixer [40].

5.2 MIMO Baseband Receiver Architecture

We briefly introduce the harmonic rejection mixer architecture (HRM) used [43], before proceeding to describe the MIMO analog spatial filter (MAC).

5.2.1   Analog signal path

A 4-channel capacitive harmonic rejection mixer (HRM) receiver directly up/down-converts a selected 3 MHz band in the RF complex inputs I_{1-4} and Q_{1-4} (48 MHz bandwidth) to baseband complex signals $I_{\mathrm{HRM}1-4}$ and $Q_{\mathrm{HRM}1-4}$. Following the 4-channel HRM, the MAC implements 4×4 complex spatial filtering for signal separation. The MAC 4×4 complex linear transform is implemented as 8×8 real matrix-vector multiplication in the I and Q components, where redundancy in the real weighting W_{ij} can be harnessed to mitigate analog coefficient mismatch as needed. Relying just on spatial diversity, the MAC is capable of separating signals with completely overlapping spectra. For example, jammers appearing in-band due to the RF front-end's down-conversion of harmonic blockers f_{RFJ} at multiples of $f_{\mathrm{LO\_RF}}$ (*e.g.* $f_{\mathrm{IJ}} + N' f_{\mathrm{LO\_RF}}$ folding onto f_{I1} in-band) can still be separated by the MAC.

5.2.2   MAC resolution

To resolve residual mismatch in the HRM outputs and implement MAC signal separation over a wide range of angles, high accuracy is needed in the multiplying digital-to-analog converters (MDACs) for the digital weights W_{ij} in the MAC analog signal path.

These requirements must be met under power constraints while providing full programmability of the analog signal path by the DSP, adding further to the design challenge. The use of capacitive charge division to implement harmonic rejection in the HRM, as well as MDAC spatial filtering in the MAC, is crucial to the large reduction in power possible due to this architecture. To allow agile operation in dynamic CR environments, high-bandwidth ($>10^4$ updates/sec) programming of the MIMO receiver parameters (HRM frequency and gain parameters, and DAC digital weights W_{ij}) from external DSP is supported via a serial-peripheral-interface (SPI) bus having only 4 control lines.

5.3 MIMO Analog Core

The MAC implements analog preprocessing on the outputs of the HRM receiver for further preprocessing, prior to digitization. The 8×8 matrix composing MAC consists of complementary 14-bit split capacitor multiplying digital-analog converters (MDAC) shown in Fig. 5.2 (a), with an effective 10-bit resolution. In all experiments reported here, unless mentioned otherwise, the matrix coefficients are calculated from the angle of incidence of the RF signal and used to beamform the undesired incident signal at baseband before digitization. Compared to the conventional approaches that rely on LO-phase shifting [26] with N-path filtering techniques, we generate phase shifts by implementing a rotation matrix using complex matrix-vector multiplication with the 8×8 weight matrix. The minimum resolvable angle and the dynamic range in the resolution are, thus, determined by the accuracy of the MDAC. In order to ensure sufficient MDAC precision, offset cancellation at the MDAC and the OTA is implemented using Correlated Double Sampling (CDS) with the clock waveform in Fig. 5.2 (b), setting the input DC bias point of the capacitively coupled differential amplifier. The CDS frequency can be set to 500 Hz that is low enough to not disturb measurements. We create a custom shielded capacitive array structure for both the MDAC and the C_F in the OTA, resulting in a programmable gain range of -12 dB to 24 dB in steps of 6 dB at the output in addition to 14-b weighting of individual W_{ij} coefficients. The programmable gain is realized by digitally selecting different numbers of unit capacitors to constitute C_F in the feedback loop.   A standard fully-differential folded-cascode OTA with the same common-mode feedback as in Fig. 5.3 is employed for the MAC.

In comparison to the largely passive MDAC in the MAC, successive approximation ADCs have many more active components such as comparators, preamplifiers, track-and-hold circuits operating at higher speeds. They thus have higher energy costs for the same effective resolution and dynamic range. The MAC is used in conjunction with a base-band DSP that can evaluate and implement weight updates via an online blind spatial filtering algorithm [12].

(a)

(b)
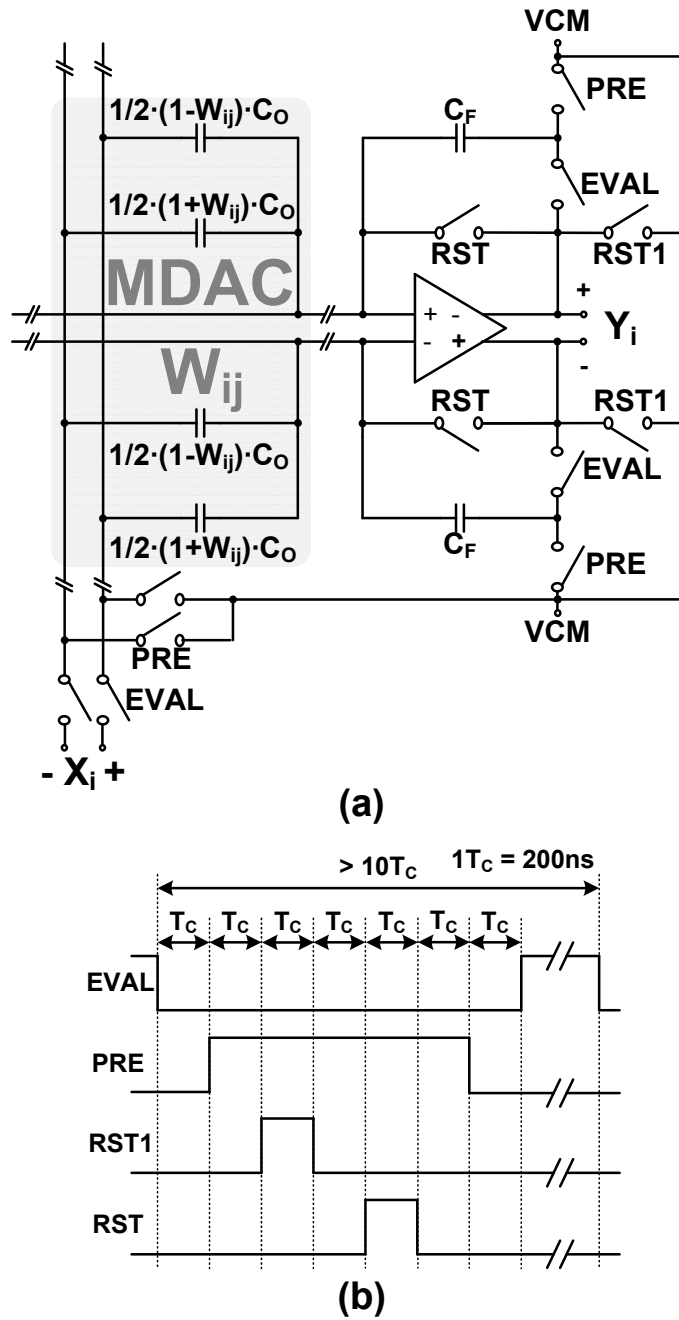
Figure 5.2: MIMO analog core (MAC) for signal separation by spatial filtering. (a) MAC circuit with multiplying digital-to-analog converter (MDAC) implementing digitally programmable analog linear weighting in the MAC signal path. To reduce the effect of offsets in the MDAC and the OTA, a correlated double sampling (CDS) scheme is employed. (b) Timing diagram for CDS.
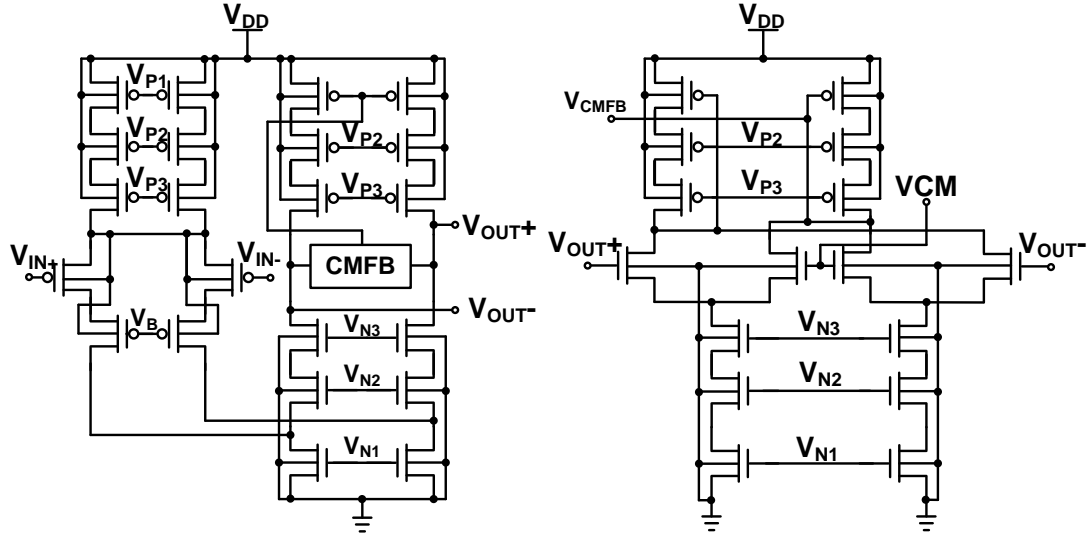
Figure 5.3: Circuits used to implement the variable gain amplifier. (a) Folded-cascode operational transconductance amplifier (OTA) in the VGA. All MOS transistors operate in sub-threshold for high-efficiency in g_m/I and large output swing range. (b) Common-mode feedback (CMFB) circuit.

5.4   Experimental Validation

The HRM and MAC functional blocks are tested independently, and in their intended cascaded succession. This dissertation focuses on results from the MAC. The complete system was validated for separation of RF sources with an antenna array and external front-end.

5.4.1   MAC characterization

The MAC experiments tested its capability to implement linear transforms of the analog input signal for beamforming of the input as well as compensating for linear gain errors in HRM channel properties. The measured two-tone separation capability of the MAC in isolation, shown in Fig. 5.4, demonstrates suppression of an in-band jammer signal $S2$ 20 dB above the signal tone $S1$ at the input, to 48.5 dB below the signal tone at the output, for a total of 68.5 dB jammer suppression. Synthetic mixtures with the algebraic sum and difference of the jammer $S2$ and 20 db attenuated signal $S1$ were presented through multichannel arbitrary waveform generators to two of the MAC inputs, with the other inputs grounded.
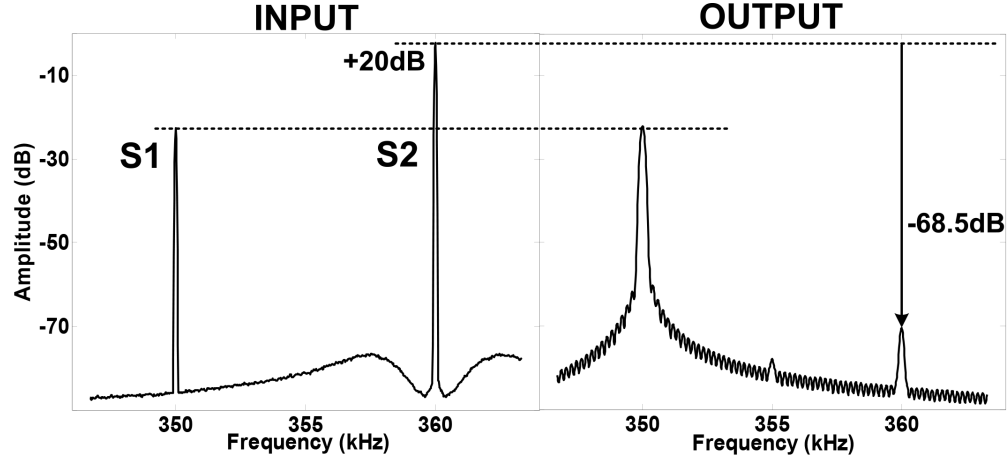
Figure 5.4: Measured in-band jammer rejection by the MAC for two synthesized inputs with linear mixtures of a signal S1 and an in-band jammer S2 +20 dB above S1, showing 68.5 dB jammer suppression at the MAC output.

The DAC digital weights W_{ij} were set to invert the synthesized linear mixing of the signals at the input, with fine adjustments for maximum nulling of the jammer.

5.4.2   Combined MIMO baseband receiver characterization

For demonstrating both spectral and spatial filtering capabilities, synthesized waveforms composed of spectrally fully overlapping mixtures of QAM and QPSK along with a 3rd harmonic blocker 24 dB above the signals are presented to the MIMO baseband receiver. Four single-to-differential amplifiers each with 6 dB gain are employed as shown Fig. 5.5 (a). Fig. 5.5 (b) and (c) show spectra and time domain waveforms at each stage of the signal chain, through the HRM and the MAC. The mixture of QAM-QPSK is shown separated by the MAC in two complex channels. Eye diagrams and I/Q constellations for the recovered 16-QAM and QPSK signals were acquired by synchronizing the oscilloscope readout of the MAC outputs with the generation of the HRM inputs and LO. The EVMs of the recovered signals, 2.69% for QPSK and 3.64% for 16-QAM, are obtained from the constellations shown in Fig. 5.6. The capability of the MAC to separate signals with completely overlapping spectra owes to its reliance on spatial rather than spectral diversity. Indeed, techniques of independent component analysis for blind
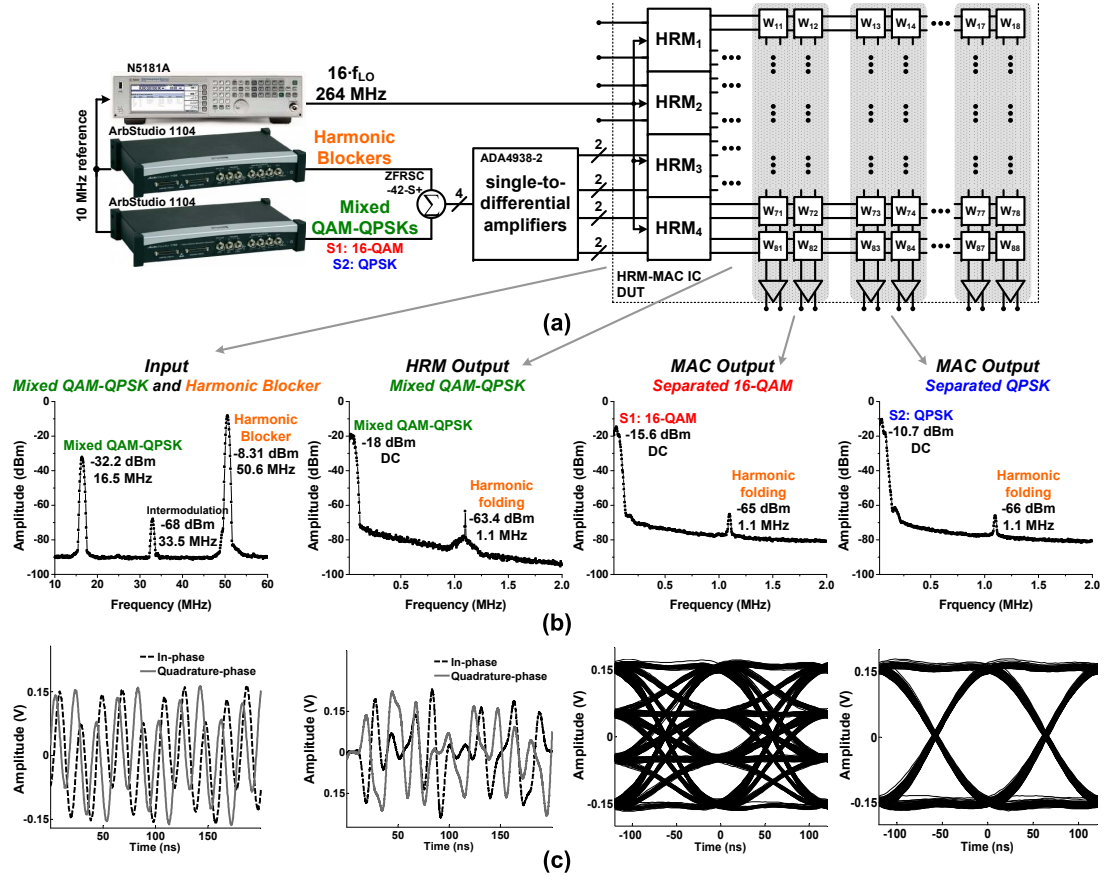
Figure 5.5: MIMO baseband receiver measurements demonstrating separation of signals with completely overlapping spectra in the presence of a strong harmonic blocker. (a) The measurement setup (b) Spectra obtained at each stage of the signal chain. The HRM output (the MAC input) contains the downconverted mixture while suppressing the blocker by 69 dB. The MAC simultaneously separates the 16-QAM and QPSK mixtures in two complex channels. (c) Time domain waveforms at each node. Eye diagrams are obtained by synchronization of oscilloscope readout with the signal inputs.

source separation (e.g., [12]) distinguish signals purely by statistical criteria and ignore their temporal spectral content.
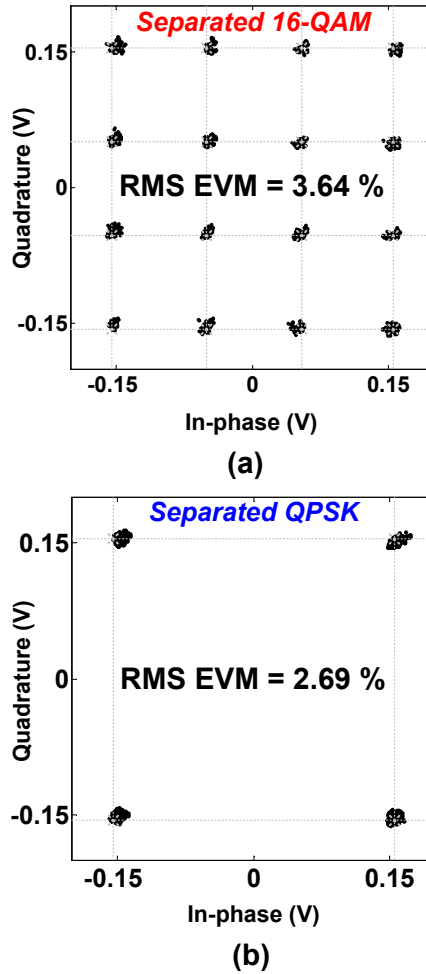
Figure 5.6: (a) Recovered 16-QAM with EVM of 3.64%. (b) Recovered QPSK with EVM of 2.69%. The constellation recovered after downconversion and signal separation by the baseband receiver clearly shows the capability of resolving complex modulated signals.

5.4.3 System validation with antenna array and RF front-end

To evaluate spectral and spatial separation performance in realistic RF conditions, measurements were conducted using an RF front-end with four antennas receiving spectrally overlapping and modulated 2.4 GHz signals from two transmitters in an un-controlled, non-line-of-sight, multi-path environment with the setup shown in Fig. 5.7 (a) and (b). The two TX antennas were positioned more than
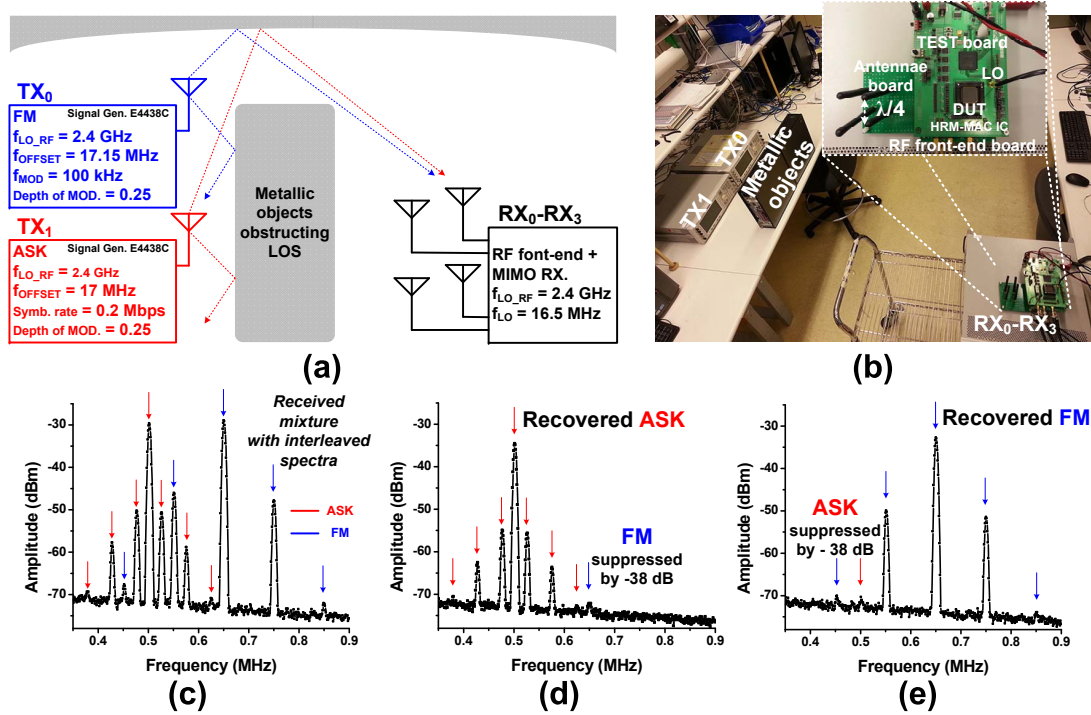
Figure 5.7: Proof-of-concept RF source separation in an uncontrolled open environment. The HRM-MAC IC outputs show recovery of non-line-of-sight RF sources with overlapping ASK and FM modulation spectra, with suppression of residual spectral components -38 dB below the signal of interest.

1 m away from the four RX antennas, with a metallic plate inserted in between to obstruct the line-of-sight path. Hence all received contributions are multi-path to emulate challenging real-world use cases. The MIMO baseband receiver (DUT) can be seen along with the RX antennas. An ASK signal at 2.417 GHz and a FM signal at 2.41715 GHz were chosen for the two RF sources, each with modulation depth of 25%. To quantify separation capability, first the down-converted signal was measured with the MAC weights set for no separation (identity weights, Fig. 5.7 (c)) and then the same test was repeated with MAC weights updated in a closed-loop fashion to separate out either of the two signals (Fig. 5.7 (d) for the ASK signal, Fig. 5.7 (e) for the FM signal). The digital weights were derived with an online greedy algorithm that maximizes the ratio of the peak power in the spectra of the down-converted, modulated signals. A net 38 dB of separation between the two signals was observed at the MAC outputs.

5.5   Acknowledgments

Chapter 6

Digitally Adaptive High-Fidelity Analog Signal Processing Insensitive to Capacitive Multiplying DAC Inter-Stage Gain Error

6.1 Introduction

There has been an unprecedented growth in the capabilities of machine-intelligence [74] and autonomous systems [21]. Moreover, autonomous systems are projected to have increasingly complex interactions with the environment, and increased communication with remote locations, all while minimizing energy use [51]. Contemporary architectures for such intelligent systems will are structured with sensor front-ends providing inputs, followed by signal conditioning and filtering, and analog-to-digital converters (ADCs), the outputs of which feed the digitized sensory information to a digital signal processing (DSP) to later be communicated to a remote server [56]. In the absence of local processing, the latencies introduced by remote communication and processing in conjunction with the inflexibility of
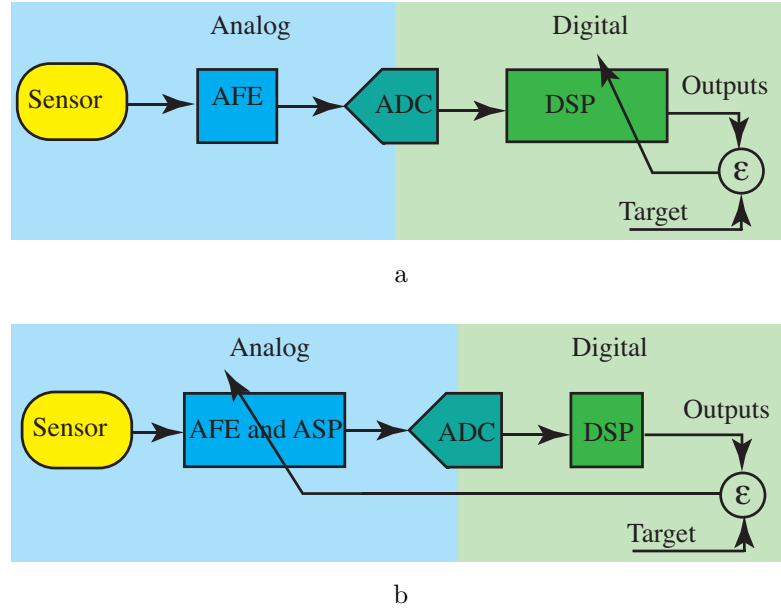
Figure 6.1: Adaptive signal processing flow in (a) conventional signal acquisition with adaptation implemented in digital signal processing (DSP), and (b) energy-efficient IoT with increased sensory-level adaptive ASP trading reduced analog-to-digital (A/D) conversion and DSP.

remote, offline learning, renders low-energy, autonomous systems impractical for use in complex environments. Thus, low-power on-chip intelligence is a prerequisite for autonomous systems interacting with the environment, making decisions, and taking required actions without human supervision. Low-power always-on reactive sensors for energy constrained applications are expected to incorporate machine learning algorithms [3] in order to boost their capabilities. A subset of such algorithms termed *online* adaptive algorithms are particularly well suited for use in autonomous devices. These algorithms receive data serially, updating their models after each new example so as to track changing conditions adapting their models over time. Since these algorithms place an increased computational burden on the underlying hardware, and thus heavily tax the system energy/power budget alternative architectures must be explored as a possible method of alleviating these increasingly stringent specifications.

Illustrated in Figure 6.1a, the traditional signal processing pipeline has adopted an architecture where available information is first digitized and then the
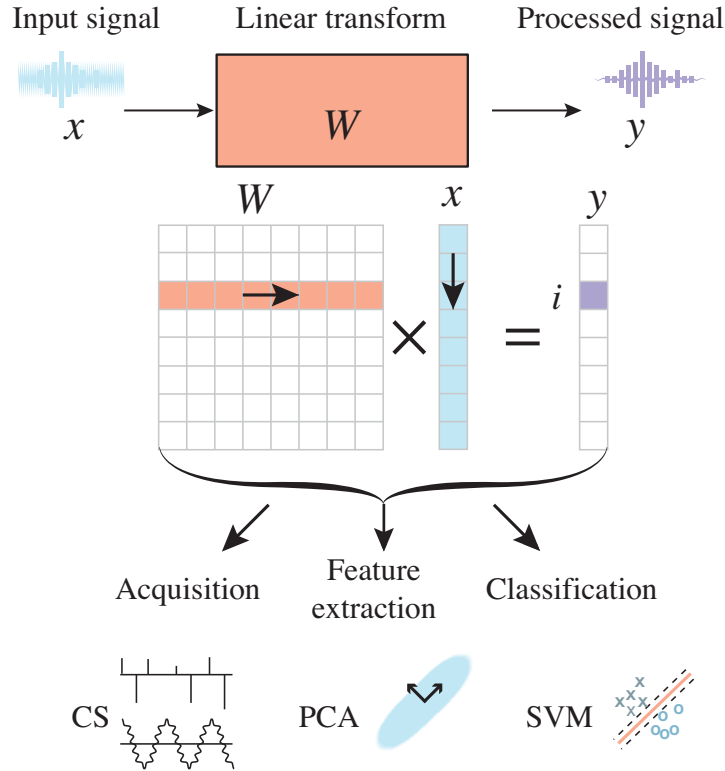
Figure 6.2: MVM is a central to signal processing at all stages of machine learning and signal processing. Various algorithms like compressive sensing (CS), principal component analysis (PCA), and support vector machines (SVMs) operate using linear transforms and linear maps. Multiplying digital-to-analog converters (MDACs) efficiently implement analog domain multiplication with digital precision.

primary information is extracted through multistage digital processing [33]. While general and widely applicable, this approach precludes analog preprocessing [37] and the resultant energy savings. Embedding very low power analog preprocessing subsystems, shown in Figure 6.1b, and thus removing irrelevant information can help amortize the overhead of analog-to-digital conversion, and the subsequent digital processing. Thus, analog preprocessing of the signal implementing some form of dimensionality or dynamic-range reduction can not only lead to energy savings, but also enable the adoption of many system level approaches previously considered unfeasible [5, 80].

MVM is a central computational primitive that is used across a large variety of tasks, as illustrated in Figure 6.2. While custom DSP accelerators can imple-

ment this primitive in a compact and efficient manner, further gains are possible due to the amenability of MVM to passive analog implementations. Thus, efforts aimed at increasing computational efficiency have resulted in Analog Matrix-Vector Multiplication (aMVM) systems designed to contribute at all levels in the signal pipeline, from sensory signal conditioning [7, 37, 83], local processing [39], and communication [43], to high-performance accelerators [48]. However, algorithmically overcoming the errors introduced by analog signal processing has remained under-explored. Thus, the main thrust of this chapter has been to propose algorithmic and circuit techniques for adoption in online algorithms, to facilitate high-resolution mixed-signal matrix-vector products in the presence of mismatch.

The central contributions of this chapter are two-fold. In Section 6.3 we analyze the energetic trade-offs associated with implementing high-resolution aMVM using capacitive MDACs. We show that the performance of many adaptive algorithms is dependent on MDAC non-linearity and capacitive matching. In Section 6.4 we propose an MDAC topology-aware algorithmic means of overcoming the effects of these nonlinearities in adaptive and online algorithms.

6.2   Background

An illustration of aMVM and its applications to a variety of tasks is provided in Figure 6.2. MVM operations multiply elements in a vector $\boldsymbol{x}$ with elements in the matrix $\boldsymbol{W}$ and accumulate the results to produce $\boldsymbol{y}$:

$$y_i = \sum_{j=1}^{n} W_{i,j} x_j. \tag{6.1}$$

Analog signal processing kernels implementing MVM occupy a range of niches, and thus must satisfy a variety of specifications regarding energy, speed, and resolution. To meet these specifications, various designs have been implemented using current-mode, voltage-mode, or charge-mode techniques. Current mode circuits [13, 50] can achieve high-dynamic range computation, using subthreshold and translinear circuits, this is particularly useful under reduced supply voltage conditions. However, these circuits are highly susceptible to PVT vari-

ations, which is exacerbated at low currents, especially in modern nano-meter processes. In contrast, capacitive MDACs face severe area and energy penalties in order to achieve high-dynamic range, but can be more robust to variations owing to the superior matching performance of capacitors.

6.2.1 Multiplying Digital-to-Analog Converters

Capacitor sizing in MDACs dictates performance, determining energy due to CV^2 driving, and switching losses, as well as accuracy due to mismatch induced errors. Capacitor sizing, thus trades-off energy with accuracy. Recent experiments on subfemto Farad capacitors in deep-submicron CMOS processes have demonstrated better than 1% matching, leading to approximately 6-bits of performance. However, a variety of signal processing algorithms require high-precision MVM to be useful (illustrated in Figure 6.6, and quantified in Figure 6.7). This is applicable in a variety of tasks ranging from beamforming separation of near-collinear sources, principal component analysis [7], independent component analysis [37] to adaptive filtering [1], and signal processing.

Illustrated in Figure 6.3, the effect of intra-array random mismatch is severe in the *C-2C* DAC. This is due to the MSB capacitor implemented using a unit capacitor, and exacerbated by the effect of parasitics on the coupling and attenuation capacitors. Thus, despite occupying minimal area, amenability to unit capacitor based implementations, and low switching energy the C-2C DAC has not seen widespread adoption in applications requiring high-resolution.

6.2.2 Processing Gain

The improvement in the signal-to-noise/signal-to-interferer ratio from employing aMVM for signal processing comes at the cost of increased area and resolution in aMVM, and increased analog processing through parallelism. We refer to this improvement in the signal-to-interferer ratio as processing gain, similar to processing gain in spread-spectrum techniques [63]. An example of *spatial* processing gain is provided in [37, 43], where spatially selective filtering of the signal
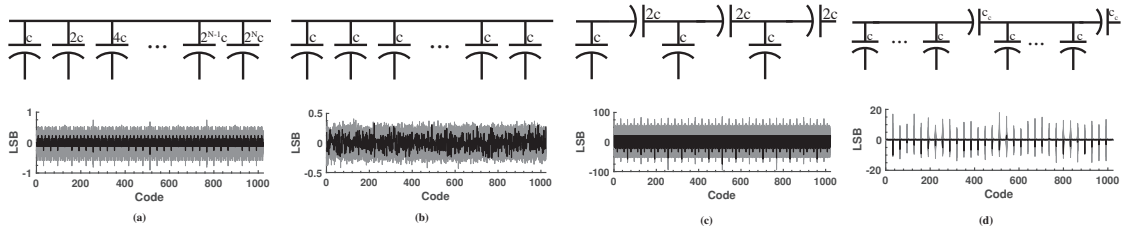
Figure 6.3: Common topologies for capacitive multiplying DACs, (a) a binary DAC, (b) thermometer DAC, (c) C-2C ladder DACs, and, (d) a segmented-DAC structure with each segment implementing a thermometer DAC. These topologies differ not only in their sensitivity to mismatch, and driving energy, but also in other properties like monotonic behavior and the frequency and location of the non-monotonicity. The simulations illustrate the effect of intra-array 10% capacitive mismatch on the static differential non-linearity (DNL).

reduces interferer power while maintaining the signal dynamic range, enabling lower resolution digitization and hence leading to substantial energy savings. This increased analog processing in turn comes at an energetic cost, the balance of which is explored in Section 6.3. Also deteriorating the performance of the ASP system, component mismatch can dramatically reduce the processing gain. To avoid this, ASP systems are over-engineered for the final performance, leading to an increase in the expended energy. This generally entails increasing the accuracy of the MDACs implementing the aMVM system, further increasing the energy expenditure. A more quantitative discussion of component accuracy on algorithmic performance is presented in Section 6.4.

It is crucial when extracting spatial processing gain from an ASP system that the analog system implement the entire linear transform step. Shown in Eq. (6.1), a linear transform entails parallel multiplies followed by accumulation. Multiplication of the signals ($\boldsymbol{x}$) with elements from $\boldsymbol{W}$ in the absence of the accumulation is identical to uniformly providing gain/attenuation to each channel with no resultant processing gain. Processing gain only occurs after accumulation of the resultant product. Thus, implementing processing gain via dynamic range reduction prior to digitization entails at least a partial analog accumulation.
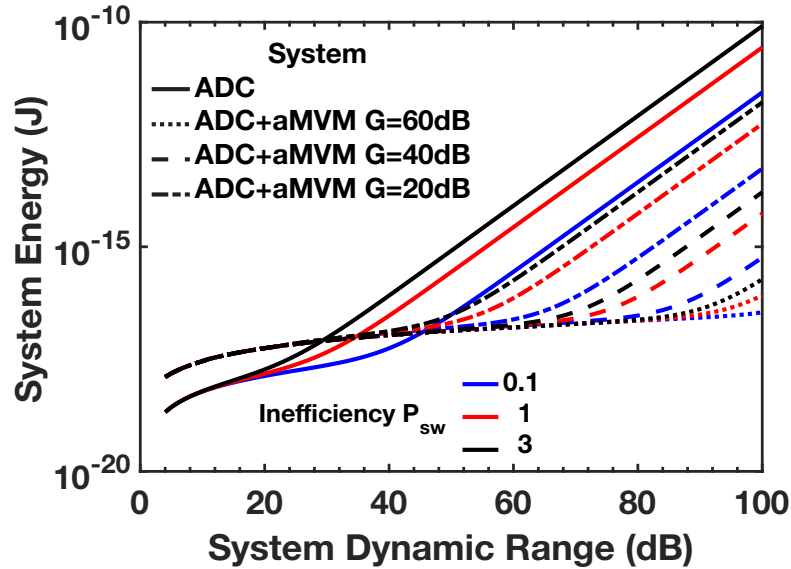
Figure 6.4: Variations to system energy limits $E_{\mathrm{sys}} = P_{\mathrm{sys}}/f_{\mathrm{sig}}$ according to Eqs. (2.2)-(2.8) with an inefficiency factor on $\eta = \{.1, 1, 3\}$, $A = 8$, $\alpha = 2$ and processing gain $G = \mathrm{DR}_{\mathrm{prior}}/\mathrm{DR}_{\mathrm{post}} = 20$, 40, and 60 dB. At lower system dynamic range $\mathrm{DR}_{\mathrm{post}}$ the energy of aMVM dominates that of SAR ADC, up to the crossover point where the processing gain is limited to unity.

6.3   Energy Costs of Capacitive aMVM

Analog signal processing can facilitate a wide variety of technologies for sensory acquisition and emerging communications. The central examples presented in this chapter focus on adaptive filtering, which has severe dynamic range requirements [75]. However, it should be noted that analog signal processing has had wide applicability in machine learning, where support vector machines have been demonstrated [24] and signal processing, where micropower implementations have enabled beamforming [11] and sound localization.

In what follows, we establish principles for analog processing to ensure overall energy savings compared to the conventional approach of directly quantizing the signal and operating upon it with DSP. For ease of notation, we consider a processing gain resulting from spatially filtering an interference source. This manifests as a reduction in the dynamic range specifications for a down-stream digitizer. The same principle applies to dynamic range reduction by feature extraction in

other forms of signal processing.

6.3.1 Power Efficiency

Consider the power requirements for a successive approximation register (SAR) ADC with a binary weighted capacitive DAC with three main constituents:

$$P_{\text{SAR}} = P_{\text{driver}} + P_{\text{mean,switch}} + P_{\text{comp}}. \tag{6.2}$$

The power for the DAC driver P_{driver} is bounded by [53]

$$P_{\text{driver}} = 16 f_{\text{samp}} k_B T \, \text{DR} \tag{6.3}$$

where f_{samp} is the sampling frequency, k_B is the Boltzmann constant, T is absolute temperature, and DR is the dynamic range. The mean switching power over all codes $P_{\text{mean,switch}}$, assuming a *merged capacitor switching* based SAR [31], is

$$P_{\text{mean,switch}} = \eta f_{\text{samp}} \sum_{i=1}^{n-1} 2^{n-3-2i} \left(2^i - 1 \right) C_u V_{\text{ref}}^2$$

where η is an inefficiency factor on the DAC switching, C_u is the unit capacitor, and $n = (\text{DR}[dB] - 3)/6$ is the ADC number of bits. Minimum capacitor sizing for thermal noise[1] results in

$$P_{\text{mean,switch}} = 12 k_B T f_{\text{samp}} 2^n \sum_{i=1}^{n-1} 2^{n-3-2i} \left(2^i - 1 \right). \tag{6.4}$$

Finally, the switching power of the comparator P_{comp} is bounded by [54]:

$$P_{\text{comp}} = 12 f_{\text{samp}} k_B T \, n \, \text{DR}. \tag{6.5}$$

Though greatly simplified, the resultant expression for P_{SAR} provides a lower-bound on power consumed for an n-bit SAR ADC.

Now, consider the presence of an interfering signal at signal-to-interference ratio SIR, which necessitates proportionally greater DR for the ADC to resolve the input signal amid the interferer without overload distortion. In turn, the greater

---

[1]We size $C_u = 12 \, k_B T \, 2^n / V_{\text{ref}}^2$ to equate thermal and quantization noise, rather than sizing for mismatch, for a lower energy bound.
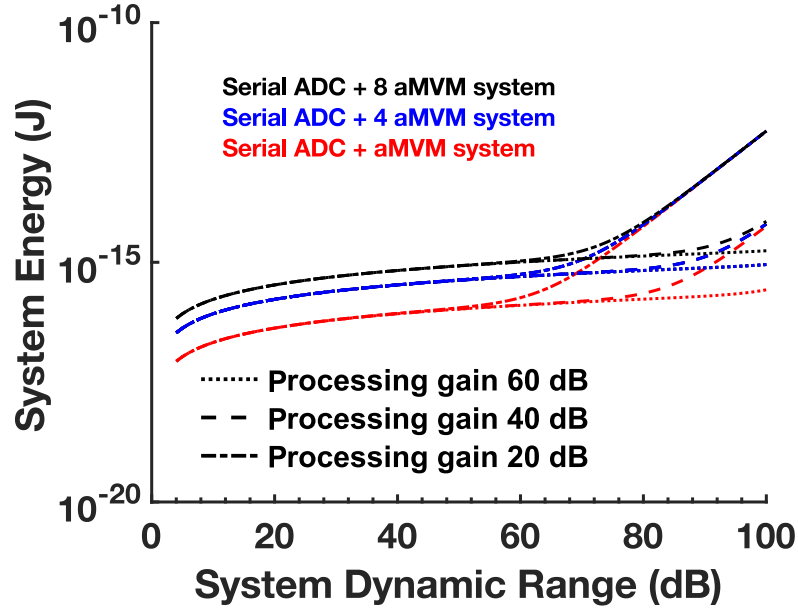
Figure 6.5: Minimum energy limits as in Fig. 6.4, with aMVM parallelism $N = 1$, 4, and 8 according to Eq. (6.7), at 10% parasitic capacitance ($\lambda = 0.1$). Amplifier gain A is increased in order to restore signal levels to full-scale for downstream ADC to counter the attenuation resulting from parallelism.

ADC DR leads to higher ADC power consumption according to Eqs. (6.2)-(6.5). A suitable aMVM front-end subsystem capable of suppressing the interferer and restoring the signal to full strength prior to quantization can hence substantially reduce the ADC power consumption, albeit at some aMVM power cost.

Capacitive aMVM incurs power costs mainly for three operations: changing the capacitive weights P_adapt, driving the capacitor array P_array, and restoring the signal with gain P_gain. The minimum power required to drive a capacitor with a sinusoidal signal with frequency $f_\text{samp} = 2f_\text{sig}$ at a given SNR is given by:

$$P_\text{array} = 8f_\text{sig}k_B T \,\text{SNR}. \tag{6.6}$$

Under the simplifying assumption relating the signal SNR to its dynamic range [54], the aMVM power reduces to

$$P_\text{array} + P_\text{gain} = 8f_\text{sig}(\text{DR}_\text{prior} + A\alpha\,\text{DR}_\text{post})k_B T \tag{6.7}$$

with closed loop gain A, amplifier inefficiency factor $\alpha \geq 1$, and dynamic range DR_prior prior to and DR_post post the aMVM gain stage. Continuous-time passive

multiplication imposes a constant load on the drivers, in contrast to a switching structure which incurs an additional power cost. Due to the improved energy efficiency of passive multiplication, we choose that architecture over alternative switching architectures [48]. The net power of the combined aMVM-ADC system is then given by:

$$P_{\text{sys}} = P_{\text{array}} + P_{\text{gain}} + P_{\text{mean,switch}} + P_{\text{comp}} \tag{6.8}$$

in which the driver power (6.3), mean switching power (6.4), and comparator switching power (6.5) for the SAR ADC are incurred at the post dynamic range DR_{post}. Note that the ADC *post* driver power P_{driver} is subsumed by the aMVM active gain power P_{gain} through the inefficiency factor α. The aMVM provides processing gain to boost the signal relative to interferer which relaxes the dynamic range accordingly, where the processing gain $G = \text{SIR}_{\text{post}}/\text{SIR}_{\text{prior}} = \text{DR}_{\text{prior}}/\text{DR}_{\text{post}}$. Thus, the combined aMVM-ADC system incurs a reduced cost for the ADC power P_{SAR} at the lower dynamic range DR_{post}, at the expense of aMVM power $P_{\text{array}} + P_{\text{gain}}$ providing the processing gain G. We normalize the power measures Eqs. (6.2)-(6.8) by f_{sig} and as such determine the minimum system energy limits $E_{\text{sys}} = P_{\text{sys}}/f_{\text{sig}}$ in Fig. 6.4. We show that the aMVM can reduce the cost of digitization, bounded by the processing gain G of the aMVM system. At lower system dynamic range the energy of aMVM dominates that of SAR ADC, where the crossover point is determined by unity lower limit on processing gain. At higher system dynamic range, the benefits of aMVM are bounded by the processing gain, and it can be seen that the ADC energy cost once again dominates. A caveat to the analysis is that at higher system dynamic range, oversampling data-converters are more energy efficient and practical than SAR ADCs.

6.3.2 Exploiting Parallelism for Analog Signal Processing

The inherent parallelism of analog computation offers several distinct advantages, such as the innate capability of accumulating charge from multiple sources onto a single wire through shared connection [69]. The improved throughput from parallelism further benefits more computationally intensive applications [24].

Furthermore, applications in spatial filtering and multichannel sensing require parallelism in order to derive spatial processing gain. Despite these advantages, recent work implementing aMVM through highly energy-efficient passive charge sharing [7, 48, 83] has not aggressively pursued parallelism. This is largely due to dynamic range limitations in massively parallel analog circuit architecture. In this Section we highlight some of these limitations along with methods to overcome them.

Highly parallel charge-redistribution capacitive arrays for aMVM suffer from gain error and signal level degradation as a result of parasitic capacitance as well as signal attenuation onto the parallel signal path. Consider the parallel connection of N capacitive multiplying DACs to compute the analog sum of N weighted inputs, $\sum_{j=1}^{N} W_{ij}x_j$ with digital weight coefficients W_{ij} and analog voltage inputs x_j. Passively connecting this aMVM output directly to the input of a SAR ADC, with another capacitive DAC for the ADC reference connected in parallel, results in charge-sharing attenuation of the voltage signal by a factor $C_{\mathrm{DAC}} / (N C_{\mathrm{DAC}} + C_{\mathrm{samp}} + C_{\mathrm{par}})$, where C_{DAC} is the Thevenin equivalent capacitance of each multiplying DAC, C_{samp} is the sampling capacitance of the ADC reference DAC, and C_{par} represents all parasitic capacitance on the shared aMVM-ADC node. Typically, the multiplying and reference DACs are identical, $C_{\mathrm{DAC}} = C_{\mathrm{samp}}$, and the parasitics result from bott om-plate capacitance $C_{\mathrm{par}} = \lambda(N + 1)C_{\mathrm{DAC}}$ where $\lambda \approx 0.1$. Thus, the attenuation factor can be approximately expressed as $1 / (N + 1)(1 + \lambda)$. The ADC reference is also similarly attenuated, exacerbating the effect of the accumulating noise degrading the SNR increasing the already stringent ADC specifications further.

A gain element following aMVM and doubling as ADC driver counters this attenuation at an increase in system energy as illustrated in Fig. 6.5. In particular, restoring signal levels back to full-scale to reduce the DR burden of the ADC comes at the cost of increased complexity and power (P_{gain} in Eq. (6.7)) of the aMVM active gain stage. This energy cost for the gain stage may be substantial where the aMVM costs dominate; however, the ADC cost dominates at higher system dynamic range, more than amply amortizing the cost of energy required to provide

the restorative gain (Eq. (6.7)).

6.3.3 Improving MDAC Efficiencies

Figure. 6.4 illustrates the effect of DAC switching efficiency on the ADC power. However, the equations described in Section 6.3 have been largely topology independent, albeit topology constrained since the topology indirectly constrains the signal dynamic range, signal-to-noise ratio (SNR), and other factors. Adaptive analog systems that entail multiple parameter updates may also benefit from energy and area efficient MDAC topologies such as the C2C DAC [68]. MDAC switching efficiency can have a major impact for an adaptive or online-learning system entailing many parameter updates. For such a system, multi-stage DAC topologies are well suited due to their amenability to compact high-resolution implementations and very low switching energy [68]. Although multi-stage MDAC topologies are area and energy efficient, inter-stage gain errors often introduce large static non-linearities as we shall demonstrate in Section 6.4.1. Thus, if the adoption of MDACs in adaptive systems is to provide an energetic advantage, the hardware and adaptive algorithms must work in concert [82].

6.4 Algorithms for High-Dimensional Analog Signal Processing

Designers conventionally compensate for capacitive mismatch in MDACs through oversizing the unit capacitor, leading to an increase in system power and area. An algorithmic approach compensating for these errors can improve the system efficiency if the overhead of this algorithm remains low. Thus, algorithms requiring extensive calculations [36,44] for optimization impose too large an energy burden and the recovered processing gain might not result in sufficient energy savings.
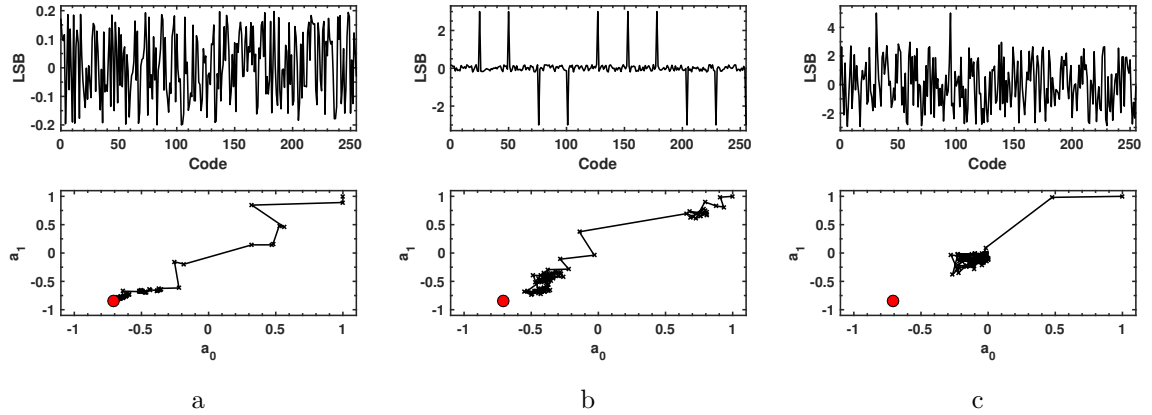
Figure 6.6: We use MDACs to implement the *Least Mean Squares* (LMS) algorithm for a two-parameter system over 100 iterations, as in Figure 2.1b. The estimate of the target is greatly affected by the MDAC component mismatch and non-monotonicity. For a well matched system a the algorithm quickly converges to the true value to within an arbitrarily small error, b the performance degrades slightly when MDAC components are well matched excepting a few non-monotonic codes (as in Figure 6.3(d)), and c the performance deteriorates significantly in the presence of very poor matching in MDAC components.

6.4.1 Algorithms for Adaptive Systems

Gradient descent is one of the most popular optimization algorithms in use in contemporary learning and adaptive systems. At its crux, gradient descent aims to minimize an objective function $\mathcal{E}(\omega)$, parameterized by a model's parameters $\omega \in \mathbb{R}^d$ by updating the parameters in the opposite direction of the gradient of the objective function. The parameter update for gradient descent is typically written as:

$$\Delta\omega = -\eta\nabla_\omega\mathcal{E},$$

where ω is the weight parameter, $\mathcal{E}$ the error functional, $\nabla_\omega\mathcal{E}$ the error gradient, and η is the step size, which needs to be small and positive to ensure convergence to a local minima. Under the conditions of convexity, this guarantees convergence to the global minima, since for a convex function the local minima is the global minima. The *LMS* algorithm, widely used in adaptive signal processing [75] is an example of stochastic gradient descent, a variation on gradient descent where the

explicit gradient isn't calculated.

For a broad class of problems like empirical risk minimization, LASSO minimization, and box constraint problems, where the objective function is separable or block separable, another class of methods termed coordinate descent (CD) can converge to the optima faster [85]. CD methods minimize the objective function by solving a set of minimization subproblems. This can provide an acceleration over gradient descent methods when the individual scalar minimization problems are simpler than the minimizing the composite. When considering multi-dimensional ASP implemented via MDACs, coordinate descent approaches prove to be amenable to mapping to a multi-MDAC approach with each MDAC typically implementing one coordinate/dimension.

In coordinate descent, generally, each coordinate is visited several times to reach a minimum, with the order of the visit called the sweep pattern. For a deterministic sweep pattern, we can write the algorithm for CD in an M-dimensional system with error functional $\mathcal{E}(\omega)$ as shown in Algorithm 1.

Algorithm 1 Coordinate Descent

  procedure CD$(\mathcal{E}, \omega, M)$ // CD on $\mathcal{E}$

    set $k \to 1$ and choose $x^0 \in \mathbb{R}^M$

    repeat

      choose index $i_k \in \{1, 2, 3, \ldots, M\}$

      $\omega^{k+1} \leftarrow \omega^k - \eta_k \left[\nabla \mathcal{E}(\omega^k) \right]_{i_k} e_{i_k}$ for some $\eta_k \geq 0$;

      $k \leftarrow k + 1$

    until Termination test is satisfied

  end procedure

Here, η_k is the step-size at the kth iteration, and e_{i_k} is the determined error. Algorithm 1 can be extended to block-CD algorithms in a straightforward way, by updating a block of coordinates rather than a single coordinate.

While these algorithms have long been analyzed for implementation in digital systems and digital signal processors, analysis has generally been restricted the effect of fixed-point operation [17] and quantization [19, 20]. However, adoption

within ASP systems composed of MDACs requires further analysis including the effect of mismatch and other MDAC errors which has remained largely unexplored. Furthermore, when optimizing some objective function $g()$, or interchangeably the error functional $\mathcal{E}$, gradient, and coordinate descent require that g be a smooth continuous function, with additional constrains on Lipschitz continuous differentiability and convexity required for analytical tractability. This places much more stringent constraints on the static non-linearity requirements on the implemented MDACs. To study these effects, we simulate the two-parameter adaptive system shown in Figure 2.1b implementing ASP with 8-bit MDACs. Figure 6.6 highlights the effect of the differential non-linearity (DNL) on such a system for three instances. For MDACs with contained static non-linearities as in a the algorithm quickly converges to the true value to within an arbitrarily small error. With limited instances of deviations in the DNL as in b, the performance degrades slightly, such scenarios would be encountered in topologies shown in Figure 6.3(d). When a highly compact and energy efficient topology like the C-2C MDAC (Figure 6.3(c)) is used the performance deteriorates significantly as shown in c. A more comprehensive quantification of this result is provided in Figure 6.7, with the peak signal-to-noise ratio of the reconstructed signal, serving as an indication of the error in the reconstruction filter coefficients estimated via LMS.

6.4.2   Errors in Multi-Stage Capacitive MDACs

Consider multi-stage capacitive MDACs as shown in Figure 6.3 (d). In this categorization a *thermometer* (Figure 6.3 (b)) is composed of a single segment, while an N-bit C-2C ladder DAC is composed of N segments (Figure 6.3 (c)).

Without loss of generality [9] we define inter-stage gain in terms of a radix, γ, in this formulation an ideal $N-$bit C-2C ladder MDAC $\gamma = 2$. For the general case, a multiplication code $\boldsymbol{b} = (b_1, b_2, b_3, \ \ldots \ , b_N)$, $b_i = \pm 1$ applied to an uncalibrated radix-γ C-2C MDAC results in an effective analog multiplicand W:

$$W = \sum_{i=1}^{N} b_i \gamma^{-i}, \quad \gamma \geq 1.$$

Analysis of this transfer function reveals that for $\gamma > 2$ the largest systematic
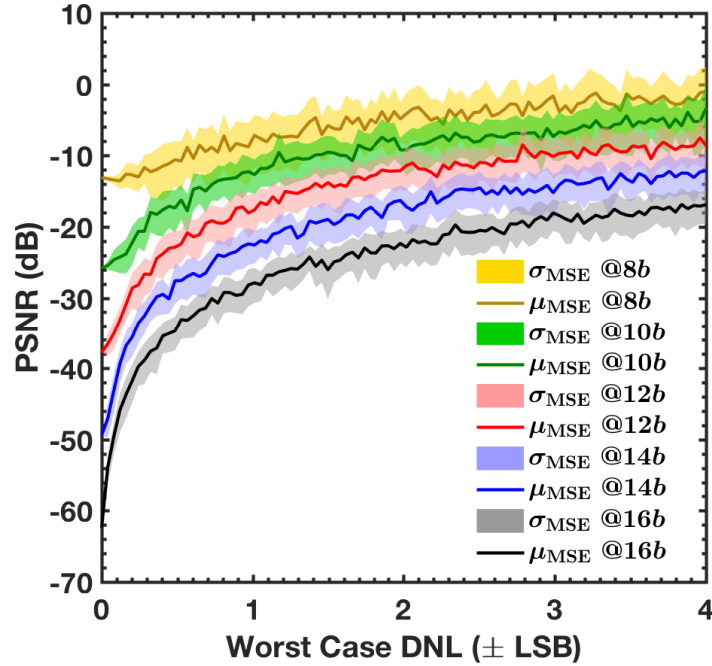
Figure 6.7: Effects of MDAC static non-linearity and resolution on LMS perfor-
mance. Expected PSNR between signal and reconstructed signal after 100 itera-
tions (μ_{MSE}) and the standard deviation (σ_{MSE}), as determined by 500 runs in a
Monte Carlo simulation at five levels of quantization: 8-bit, 10-bit, 12-bit, 14-bit,
and 16-bit. The signal reconstruction PSNR is significantly lowered in the presence
of component mismatch with ± 1 LSB errors resulting in as much as 30 dB loss in
performance.

errors occur at the mid-point of the transfer function with the code transitioning
from all *LSB* contributions to only the *MSB* contribution. An example for $\gamma = 2.4$
is shown in Figure 6.8a. For an N-bit DAC we denote this error by $e_{\mathrm{max},N}$, where

$$e_{\mathrm{max},N} = 2\left(\gamma^{-1} - \sum_{i=2}^{N-1}\gamma^{-i}\right)$$

$$e_{\mathrm{max},N} = 2\gamma^{-1}\left(1 - \gamma^{-1}\frac{1 - \gamma^{-(N-1)}}{1 - \gamma^{-1}}\right). \tag{6.9}$$

Shown in Figure 6.9a, the corroboration between the *Analytical* and *Simulated*
errors validates the analytical model. At $\gamma \leq 2$ the uncharacterized (uncalibrated)
MDAC transfer function has larger errors due to non monotonicities in the transfer
function. However, a full characterization and calibration involving a remapping
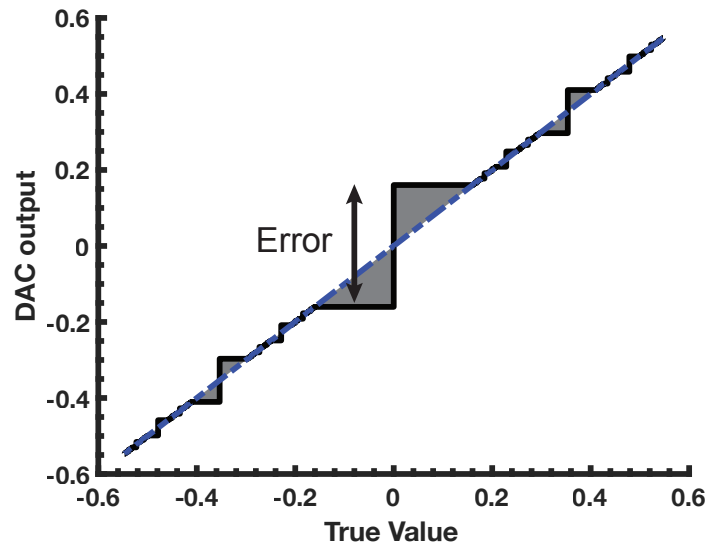
a
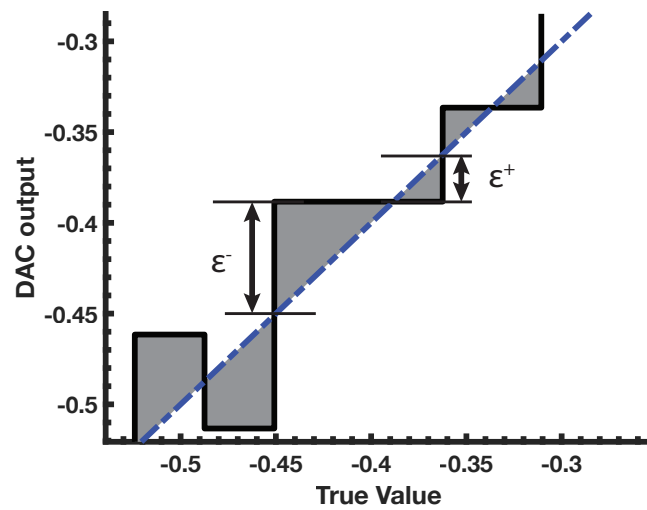
b

Figure 6.8: The maximum error in a DAC transfer function is shown normalized to DAC range, this error provides an estimate of the effect of inter-stage gain error in comparison to the number of stages in the DAC
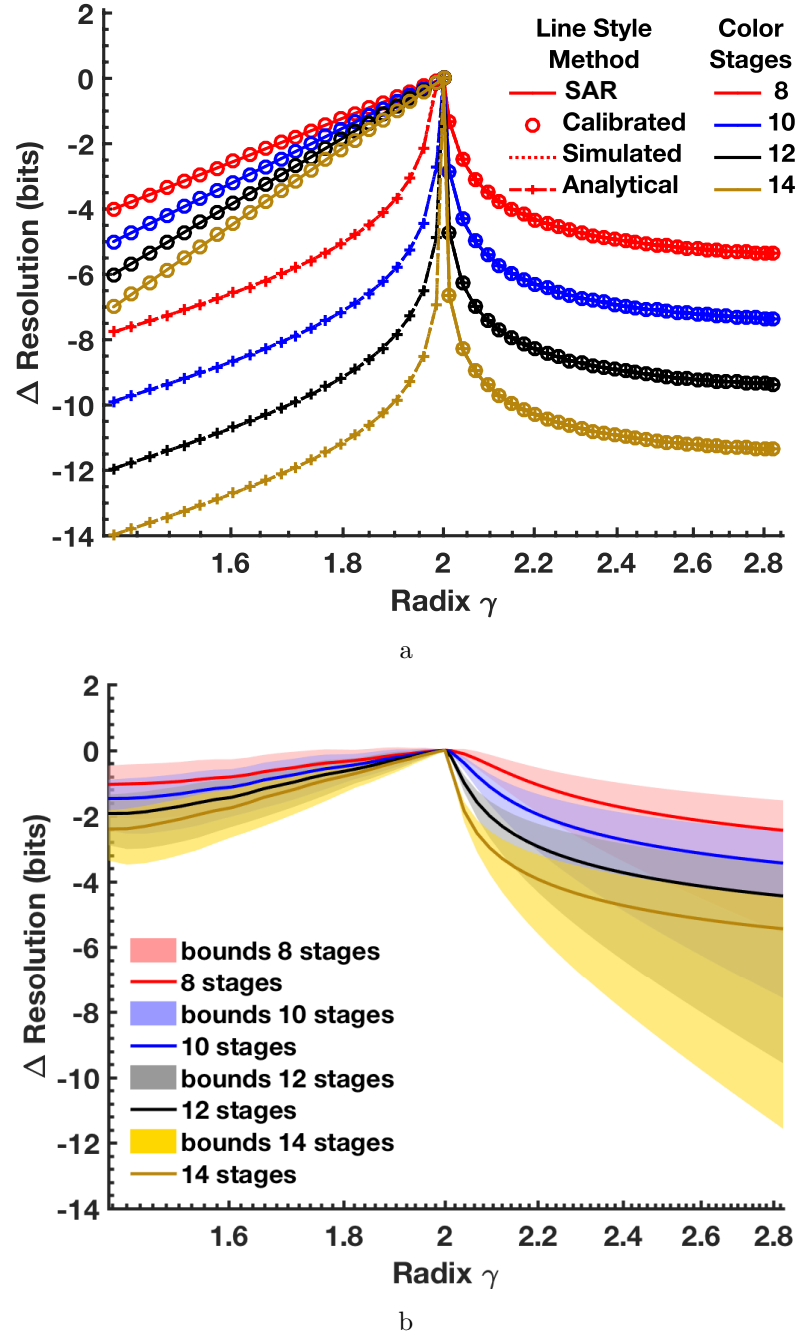
Figure 6.9: Effects of multi-stage gain for MDAC with radix γ on effective MDAC resolution. Two measures of resolution are provided, a shows the relative size of the maximum error according to Eq. (6.9) showing a dramatic reduction in resolution for $\gamma > 2$ and a milder effect for $\gamma < 2$, b illustrates the performance of the S2GD algorithm as measured by the mean squared error and its bounds.

of codes to their minimum error values dramatically reduces this error as shown in Figure 6.8. In contrast, the errors introduced by radix > 2 remain unchanged despite calibration since the transfer function is monotonic and any error cannot be corrected for through redundancy.

Since a *brute-force* based calibration is unfeasible in practice, we explore alternative adaptive strategies to improve upon the performance of multi-stage capacitive multiplying DACs when applied to ASP. The successive approximation algorithm, when applied using a radix-γ DAC, with single bit comparison per stage leads to performance on-par with a fully calibrated MDAC as seen in Figure 6.9a. This follows from a result introduced by Rényi [66] on β approximations, leading to applications in offset compensation in non-binary ADCs posited in [35], as well as [9]. With a foundation based on these results, we demonstrate the effect of radix-errors on the operation of the radix-γ bit-level successive approximation. During convergence to an analog target value x $(-1 \leq x \leq 1)$, determining the N output bits b_i closest to the target works as shown:

$$x_{\mathrm{SA},N} = \sum_{i=1}^{N} b_i \gamma^{-i}, \quad \gamma \geq 1 \ , \ b_i = \pm 1$$

$$b_i = \mathrm{sign}\left(x - x_{\mathrm{SA},i-1}\right)$$

$$x_i = \gamma^{i-1}\left(x - x_{\mathrm{SA},i-1}\right).$$

Here, the ith bit, $b_i = \pm 1$, is deterministically assigned at the ith comparison, $x_{\mathrm{SA},i}$ is the output after i successive approximation cycles, x_i is the residue at the end of $i-1$ conversion cycles. For ease of analysis we have assumed that all inter-stage gains γ are identical, a similar general analysis is possible if that condition doesn't hold. Variations of this result can also be applied to demonstrate the resilience of capacitive MDAC adaptation to offsets in measurements, as exploited by redundant sucecssive approximation register (SAR) ADCs [55, 88].

A more representative analysis of the errors over the DAC transfer function can be performed by decomposing the error metric outlined in Figure 6.8a into ε^+ and ε^- as formed by the intercepts with the line $y = x$. The distribution of these intercept values provide bounds on the errors over all codes. An evaluation of the SAR algorithm for this metric along with the error bounds is shown in fig. 6.9b.

These results indicate that radix non-idealities associated with high-resolution multi-stage capacitive MDACs can be compensated for through successive approximation-like iterative adaptation. Extending this result to higher dimensions results in successive approximation-like convergence to the true value for multiple dimensions. This forms the foundation of our work in creating successive stochastic approximation (S2A), a modification to stochastic coordinate descent that overcomes its shortcomings when applied to high-dimensional analog computation in the presence of component mismatch.

6.4.3 Successive Stochastic Approximation

Consider the task of determining the M-dimensional set of parameters $\boldsymbol{b}$ that minimize the error functional $\mathcal{E}(\boldsymbol{b})$. Here, $\mathcal{E}(.)$ provides a quantitative measure of the error between a desired state x and the current state x_{s2a}, where $x_{\mathrm{s2a}} = f(\boldsymbol{b})$, for some unknown function $f(.)$. In the restricted case of ASP implemented using N-bit MDACs, $x_{\mathrm{s2a}} = \sum_{i=1}^{N} b_i \gamma^{-i}$ is discrete and entirely determined by the ± 1 vector of codes $\boldsymbol{b}$, and the unknown *radix* γ.

Deriving inspiration from the *Metropolis-Hastings* (MH), and the *Simulated Annealing* algorithm, we extend the results we'd derived in the previous section to a higher-dimensional setting. The central iteration of the algorithm consists of greedily choosing between two proposed candidate steps at a resolution followed by successive approximation resolution increase.

To overcome inter-stage gain errors and mismatch in DAC weights, it is important that the successive adaptation stages are center-aligned so they have equal room to move in either direction. In a multi-stage MDAC this requires the *LSB* following the currently adapted stage to to be at $(1, -1, -1, -1, -1, \ldots)$ or $(-1, 1, 1, 1, 1, \ldots)$, (which we denote by *MID*) while the *MSB* candidates are proposed so that the subsequent LSB adaptation starts in the middle of the range, far away from cross-over distortion due to inter-stage DAC nonlinearity at MSB-LSB major transitions. Thus, for the ith successive adaptation, we propose the two candidates by flipping the ith bit while all along keeping the $(i + 1)^{\mathrm{th}}$ bit at 1 and the succeeding bits all at -1. For a single dimensional case the crux of the

algorithms effectiveness lies in two factors: the coarse-then-fine approach prevents the algorithm from getting stuck in local minimas far from the optimal (akin to simulated annealing), fixed-point operation due to ASP MDAC avoids slow-down from pathological curvature.

More formally, consider a M-dimensional adaptive system adapted through M, N-bit MDACs. After initializing the MDACs to *null* we uniformly randomly pick a dimension d_1^1 for the first resolution iteration (the superscript denoting the iteration number, and the resolution of adaptation). We then propose two candidate points $x_{\text{cand},\{1\}}^1, x_{\text{cand},\{2\}}^1$ generated by complementary updates to the existing coordinate at the current resolution (iteration number 1),

$$x_{\text{cand},\{1\}}^1 = x_{\text{s2a}}^1 + \gamma^{-1},$$
$$x_{\text{cand},\{2\}}^1 = x_{\text{s2a}}^1 - \gamma^{-1}.$$

We update the parameter by greedily choosing the candidate minimizing the error function $\mathcal{E}(f(\boldsymbol{b}))$, requiring two measurements of $\mathcal{E}(f(\boldsymbol{b}))$ under the complementary perturbations. Once this update has been performed for all dimensions $(d_1 \ldots d_M)$ at the current resolution, the resolution is updated, with the iteration proceeding until all bits $(1 \ldots N)$ have been adapted. This form of weight update can be easily implemented with digital circuits, without the need for explicit adders or counters enabling very low-cost adaptation of the MDACs. For analysis consider a 2-D convex function f:

$$f = (x - x_0)^2 + (y - y_0)^2 + \alpha(x - x_0)(y - y_0) \tag{6.10}$$

where, x_0 and y_0 are the target values, and thus the global optima. Figure 6.10 shows the *S2A* steps in converging to the target, $(x_0 = .1,\ y_0 = -.25)$ for an instanced ASP system given two MDACs with radix $\gamma = 1.6$. Despite the convex nature of the function described by (6.10), the MDAC non-monotonicity effectively results in a non-convex optimization landscape, thus making the adoption of general optimization algorithms unfeasible.

Algorithm 2 Successive Stochastic Approximation

   procedure S2A$(x, N, M, \gamma, \mathcal{E}(.))$ // S2A on x

     $\boldsymbol{b}^j = MID(1:N) \; \forall j$ // $j = \{1 \ldots M\}$

     $l \leftarrow list(1:M)$

     for $i := 1 \rightarrow N$ **do**   // Converged to LSB at i=N

       $D \leftarrow permute(l)$

       for each $j \in D$ **do**   // iterate over permuted l

         $f(\boldsymbol{b}_{\mathrm{cand},\{1\}}) \leftarrow \left(f(b^j_{\mathrm{s2A},i-1}) + \gamma^{-i}, MID(i+1:N) \right)$

         $f(\boldsymbol{b}_{\mathrm{cand},\{2\}}) \leftarrow \left(f(b^j_{\mathrm{s2A},i-1}) - \gamma^{-i}, MID(i+1:N) \right)$

         Note that: $b^j_{i,\{1\}} = +1$, $b^j_{i,\{2\}} = -1$

         $b^j_i \leftarrow \arg\min_{\boldsymbol{b}} \left(\mathcal{E} \left(f\left(\boldsymbol{b}_{\mathrm{cand}} \right) \right) \right)$

       end for

     end for

     return $\boldsymbol{b}$ // $\boldsymbol{b}$: vector of binary code for the min params

   end procedure

   procedure MID(MSB:LSB)

     $\boldsymbol{B} = (-1, 1, \ldots, 1)$ // $\boldsymbol{B}$ is a ± 1 vector of size MSB-LSB

     return $\boldsymbol{B}$

   end procedure

6.4.4   Extensions to Successive Stochastic Approximation

As with coordinate descent, strong coupling between the dimensions can lead to sub-optimal updates in *S2A* due to the greedily always increasing the resolution. An example highlighting this is shown in Figure 6.12. By always increasing the MDAC resolution and ensuring updates do not cross resolution boundaries the algorithm outlined in Algorithm 2 may have errors in the final output, where the results converge too quickly. To overcome these limitations, we modify the candidate generation to be exhaustively greedy as shown in Algorithm 3.
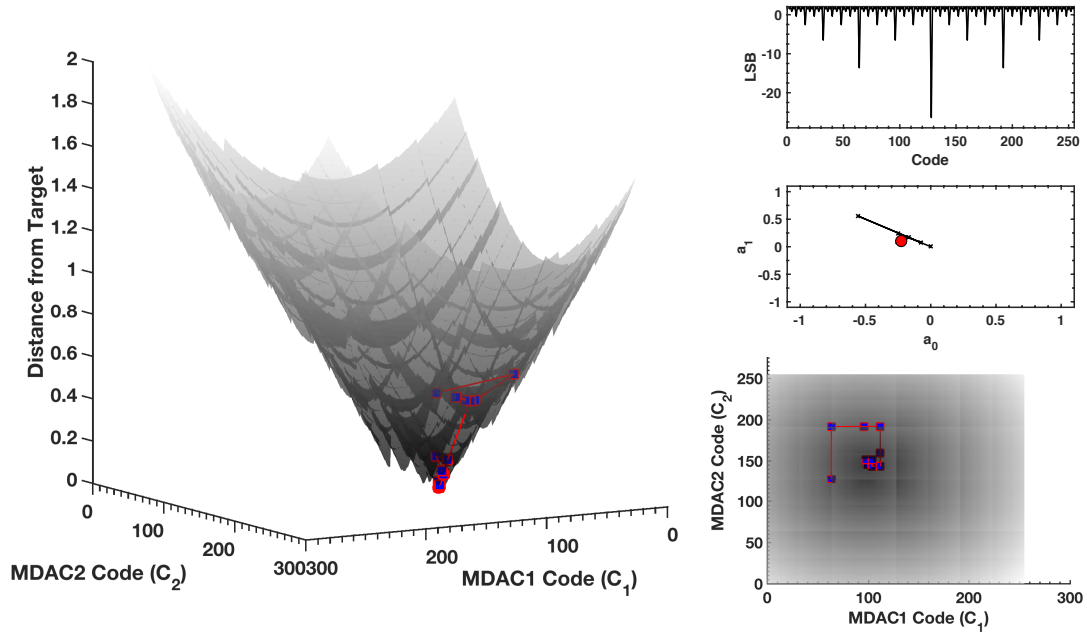
Figure 6.10: The absolute value distance from the target $(.01,,-.25)$ in a two-dimensional system with the objective function shown in eq. (6.10) and cross-coupling parameter $\alpha = .1$. The instanced MDACs had radix $\gamma = 1.6$ (resulting in a worst case DNL > 25), and no random mismatch.

Algorithm 3 Extended Successive Stochastic Approximation

    procedure xS2A$(x, N, M, \gamma, \mathcal{E}(.))$ // xS2A on x

      $\boldsymbol{b}^j = MID(1:N) \;\forall j$ // $j = \{1 \ldots M\}$

      for $i := 1 \to N$ **do**   // Converged to LSB at i=N

         generate all candidates at this resolution

         $b_i^j \leftarrow \arg\min_{\boldsymbol{b}} \left(\mathcal{E}\left(f\left(\boldsymbol{b}_{\text{cand}} \right) \right) \right)$

      end for

      return $\boldsymbol{b}$ // $\boldsymbol{b}$: vector of binary code for the min params

    end procedure

    procedure MID(MSB:LSB)

      $\boldsymbol{B} = (-1, 1, \ldots, 1)$ // $\boldsymbol{B}$ is a ± 1 vector of size MSB-LSB

      return $\boldsymbol{B}$

    end procedure
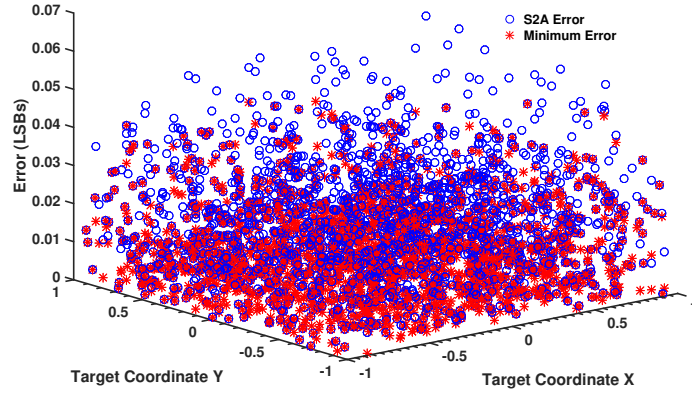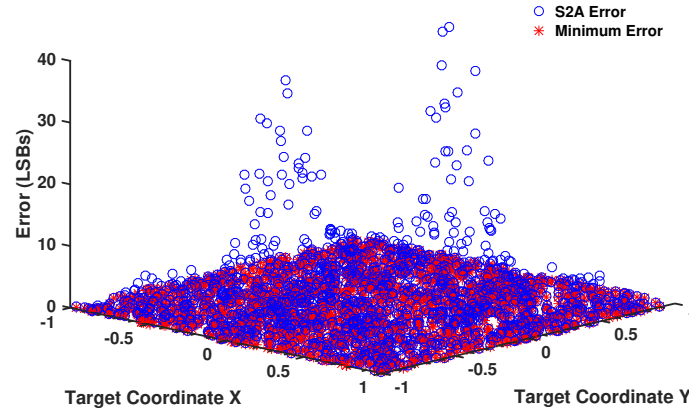
Figure 6.11: The effect of dimensional separability on the *S2A* algorithm instanced on 8-bit MDACs with radix $\gamma = 1.8$. When cross-dimensional coupling is weak, i.e., , $\alpha = 0$ in eq. (6.10) the system can be rewritten as two independent 1-D problems with successive approximation performed in each dimension, reducing to the problem in Section 6.4.2 (Figure 6.9). When there is stronger coupling systematic errors along the curvature are seen as outliers as explained in Section 6.4.4.
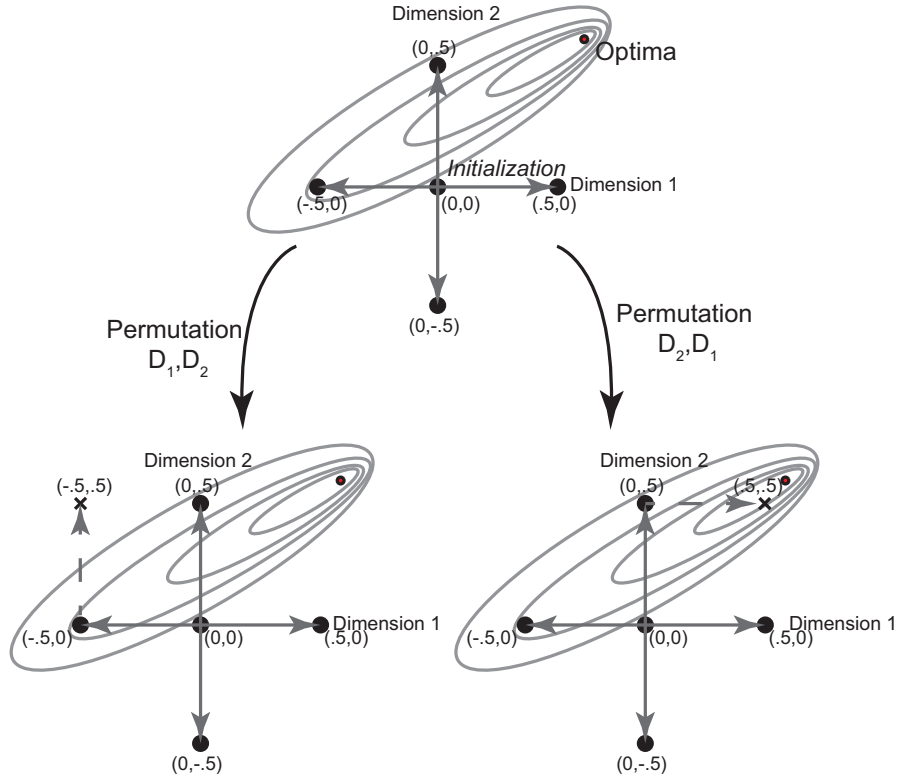
Figure 6.12: Illustration of possible steps of Algorithm 2 leading to a suboptimal decision. In the scenario shown, the correct decision would be to choose weights (.5,.5) in the first iteration. However, depending on the order of the traversal of dimensions,i.e., dimension 1, then dimension 2, or dimension 2, then dimension 1. The greedy parameter updates of the algorithm result in suboptimal decisions at the coarser resolutions. This highlights the effect of the separability of the underlying problems, a limitation often in in coordinate descent based algorithms.

Unlike $S2A$ highlighted in Algorithm 2 which takes $M2N$ steps to converge, Algorithm 3 results in $M2^N$ steps due to the generation of all possible candidates. This provides a greater explicit exploration of the parameter space, and is able to better overcome the limited non-monotonicities at the coarses resolutions for $\gamma < 2$.

6.4.5   Effects of Random Mismatch

In this section, we analyze the performance of the proposed algorithms when adapting 2-D, C-2C, MDACs systems. For the 2-D system we have been using, one

error metric is to normalize the error with the expected error for an ideal quantized system i.e., the average distance from the vertex of a square of size $E[DNL]$. This can be determined through the *square point picking problem* [84, 87] and is given by:

$$E[d_{\text{vertex,closest}}] = E[DNL]\frac{2 + \sqrt{2}\log(1 + \sqrt{2})}{24} \tag{6.11}$$

Dividing the errors by eq. (6.11) provides us with a normalized estimate of the errors in terms of LSBs in this 2D system. The two main factors influencing convergence for *S2A* and *xS2A* are the decomposibility into independent subproblems per constituent dimension, and the effect of random mismatch in addition to the radix errors introduced by systematic mismatch. Figures 6.11 and 6.13 highlight the effect of cross-dimensional coupling in two-dimensional system with 8-bit MDACs with radix error resulting in an effective radix 1.8, at $\alpha = 0$ and $\alpha = 1$ for the *S2A* (fig. 6.11), and *xS2A* (Figure 6.13).

Monte Carlo experiments over 1000 runs are performed on a two-dimensional system with random capacitive mismatch set to be within $\pm 1\%$ and linearly varying the cross-dimensional coupling α between 0 and 1, as well as varying the effective radix between $\sqrt{2}$ and 2. These results normalized as described by (6.11) are summarized in Figure 6.14. The error after calibration are used as reference for both the *S2A* and *xS2A* algorithm. As can be seen, strong coupling leads to a dramatic loss of performance, while separability results in the performance levels being maintained at dramatically decreased costs (2^{M-1} fewer steps).

Similarly we observe the effect of mismatch on the algorithm by varying the random mismatch between elements these results are summarized in Figure 6.15. With strong cross-dimensional coupling, increased random mismatch leads to an improvement in performance, behaving as a regularizer on *S2A*. This effect is negligible when the problem doesn't decompose into independent dimensions.

6.5   IC Measurements

An example aMVM system for spatial signal conditioning in adaptive beamforming for RF communication was described in [43]. The system implements
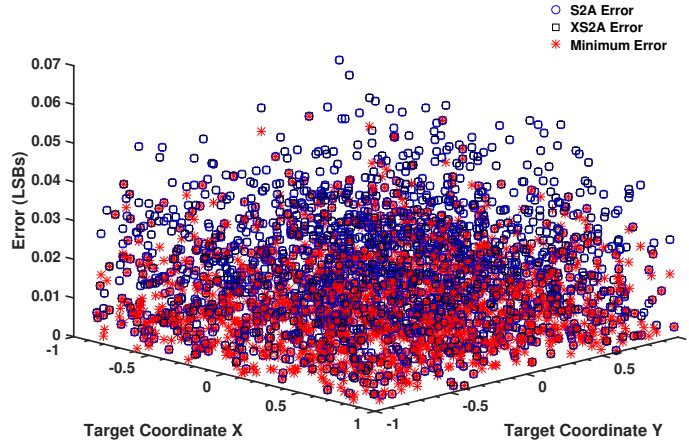
a
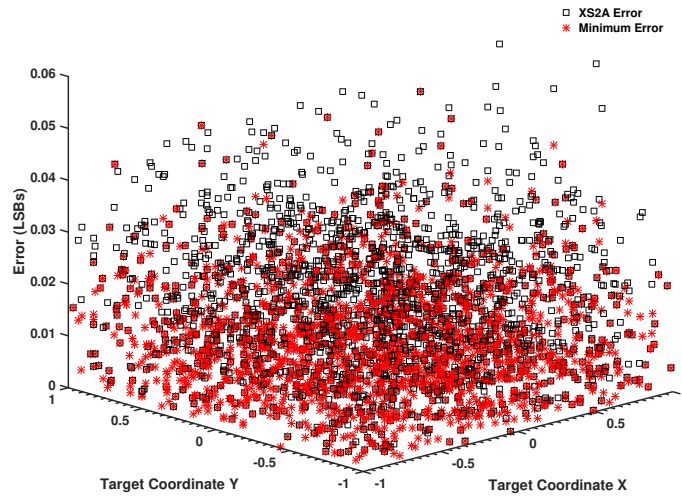
b

Figure 6.13: The *Extended S2A* algorithm instanced on 8-bit MDACs with radix $\gamma = 1.8$ better overcomes the effects of cross-dimensional coupling due to the exhaustive generation of candidates at each level of resolution. (a) illustrates the effect of target location on optimization performance for $\alpha = 0$, (b) illustrates the effect of target location on optimization performance for $\alpha = 1$
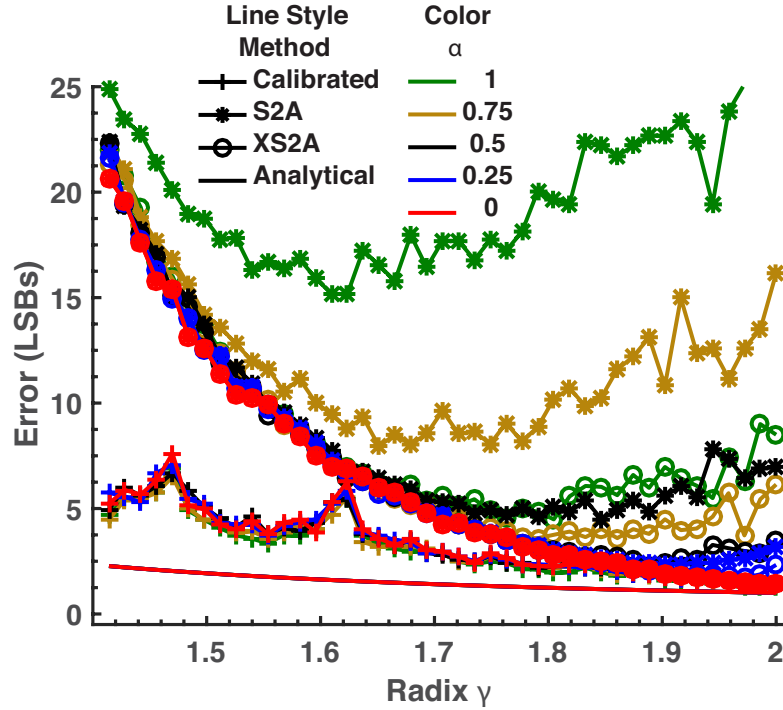
Figure 6.14: The effect of varying the coupling coefficient α in the objective function outlined in eq. (6.10). With no cross-coupling the algorithm performance is improved for both the coordinate descent algorithm outlined in Algorithm 2 and the extended version outlined in algorithm 3. The error hamming distance is maintained to within 16-LSBs for the *extended successive stochastic approximation* (xS2A) algorithm over the entire range of radixes.

aMVM preprocessing on the outputs of harmonic rejection channelization resulting in analog spatial processing gain prior to digitization. The 8×8 aMVM is composed of capacitive multiplying digital-analog converters (MDACs) implementing the linear transform. Beamforming is implemented through digitally programmed transform coefficients. The resulting capacitive weighting spatially filters the incident signal from four antennas at baseband, implementing 4×4 complex matrix-vector multiplication with the 8×8 real matrix as:

$$X = \begin{pmatrix} \mathbb{Re}(X) & -\mathbb{Im}(X) \\ \mathbb{Im}(X) & \mathbb{Re}(X) \end{pmatrix} \tag{6.12}$$

We implement the algorithm in conjunction with this system for use in a spatial separation task entailing the suppression of an interfering tone under

a

b
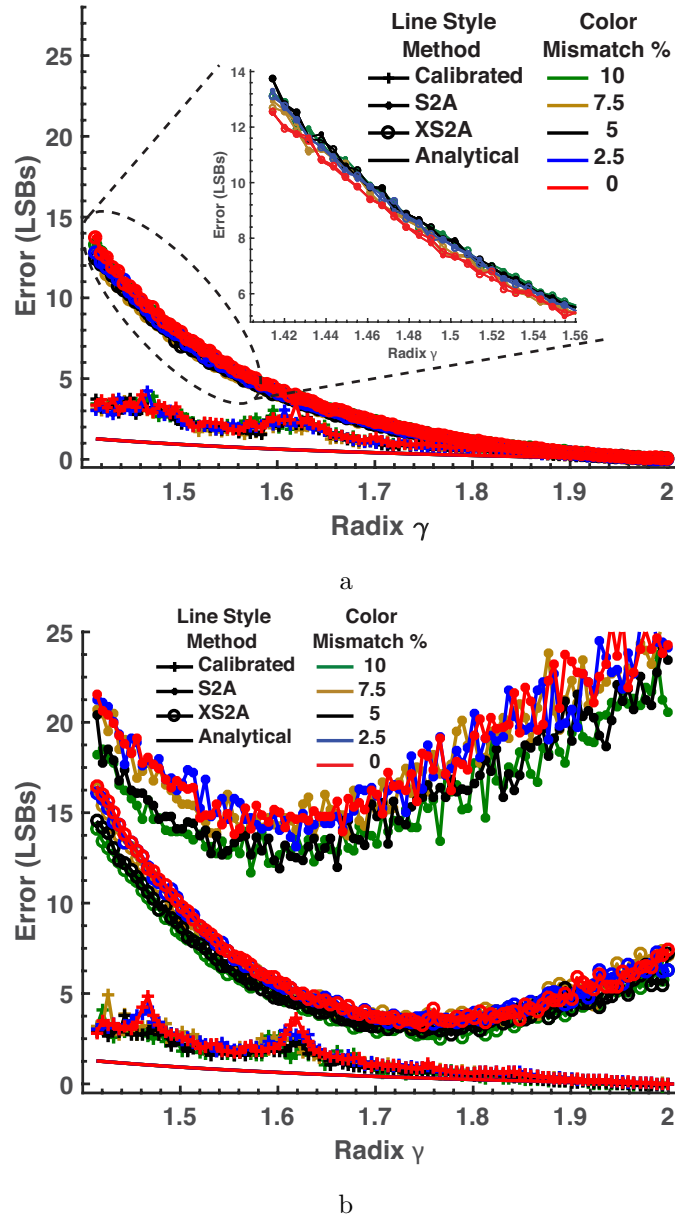
Figure 6.15: Effects of mismatch and the α term from eq. (6.10) on the performance of both the *S2A* and *xS2A* algorithm. (a) demonstrates the effect of a completely separable problem ($\alpha = 0$), with a negligible performance gap between *S2A* and *xS2A*. (b) demonstrates the regularizing effect of random mismatch when $\alpha = 1$, with increased mismatch leading to a minor improvement in normalized performance.
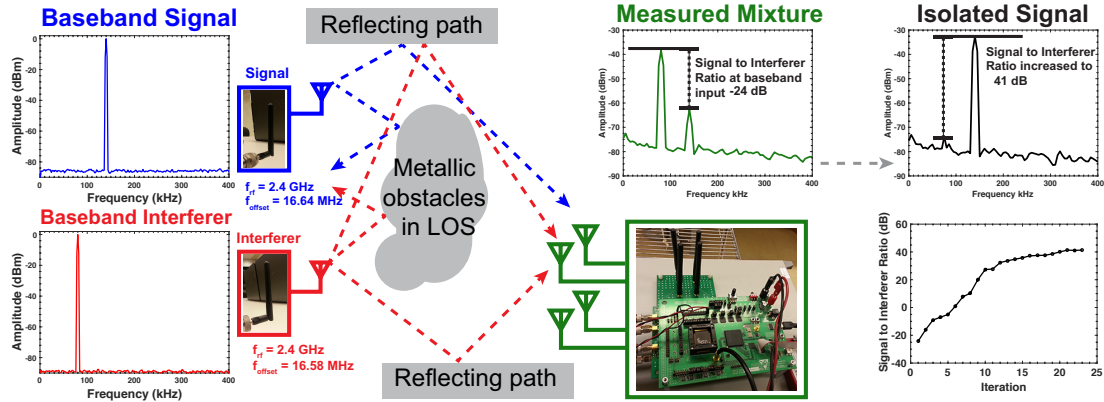
Figure 6.16: Over-the-air source separation in an uncontrolled environment [43] shows recovery of non-line-of-sight RF sources. Closely spaced sinusoidal tones are resolved with resulting residuals suppressed to the noise floor. Simulated spectra is are shown for baseband data at the transmission end, with measured spectra for the received signals shown on the right. The *S2A* as outlined in Algorithm 2 was applied to maximize $P_{\text{f}=140 \text{ kHz}}/P_{\text{f}=80 \text{ kHz}}$.

realistic RF conditions. An RF frontend with four antennas receives these tones modulated onto a 2.4 GHz carrier from two transmitters in an uncontrolled, non-line-of-sight, multipath environment with the setup shown in Figure 6.16. The two TX antennas were positioned with metallic obstacles obstructing the line-of-sight path to the four RX antennas more than 1 m away, creating multipath signal contributions emulating realistic channel conditions. To clearly demonstrate signal separation, a sinusoid at 2.41658 GHz and another tone at 2.41664 GHz are presented. The downconverted received mixture has two tones in the absence of spatial filtering as shown in Figure 6.16, this has an initial signal-to-interferer ratio of -24 dB.

The algorithm presented in Algorithm 2 greedily determines the digital weights to maximize the ratio of the peak power in the spectra of the downconverted, modulated signals. The signal separation performance over iteration count is shown in Figure 6.16, the resultant performance saturates to 41 dB of separation between the two signals Figure 6.16, which is at the measurement limit as seen in Figure 6.16. This results in a net 65 dB of interferer suppression within 24 iterations.

6.6 Conclusion

In this chapter, we have shown that analog signal processing can dramatically reduce system energy enabling communication and data acquisition by "smart" sensory systems. Analysis of the energetic limits of analog MVM systems and in conjunction with the compact size and energetic advantages of multi-segment capacitive DACs make them attractive for adoption for ASP. However, their greater susceptibility to radix errors due to capacitive mismatch can lead to non-monotonicities in their transfer function, and thus severe performance loss when used with adaptive algorithms. We introduced the Successive Stochastic Approximation algorithm as well as the Extended Successive Stochastic Approximation algorithm to overcome the effects of MDAC radix errors induced by static-nonlinearities. We analyzed the effect of random-capacitor mismatch on these algorithms demonstrating tolerance to high levels of mismatch. And finally in measurement results from over-the-air tests demonstrating the use of the presented algorithm in a system level application where aMVM is used to implement pre-digitization beamforming. Measured results show up to 65 dB of improvement in signal-to-interferer ratio over 24 iterations of the *S2A* algorithm when used on a multi-segment capacitive DAC.

Acknowledgment

Chapter 7

Conclusion

With the emergence of machine intelligence and ambient computing generally called the internet of things, research into intelligence at the "edge" of this network of devices. The key to decreasing the power consumed by such systems is to embed intelligence at all levels, from the sensory interface to communication. Appropriate optimization across the various levels of the design hierarchy from on-chip algorithms to the devices used for computation can address the challenges of increased local computation and communication at a reduced power budget. The main contribution of this dissertation has been towards the architecture and designing of high-fidelity, low-power, mixed-signal processing as well as the design of algorithms complementing these hardware efforts.

While high-resolution processing has traditionally remained the domain of digital processing systems [69], the central tenet of this dissertation has been to demonstrate in a principled fashion the means to achieve high-resolution analog processing. Chapter 2 of this dissertation lays out the energetic advantages of linear analog signal processing where high-dynamic range, high-fidelity analog signal processing proves beneficial over a more digital system. Through a reduction in the dynamic range of sensory signals prior to digitization and thus a reduction of power consumption in the digitization process, ASP systems can provide orders of magnitude greater energy efficiency. This ASP typically involves linear transforms implemented through passive capacitive weighting of signals, resulting in minimal energy expenditure.

Building upon these findings, we introduced a family of high-fidelity, capacitive, mixed-signal co-processors with applications to sensory data conditioning (Chapter 4) and aiding smart receivers in cognitive communication networks (Chapter 4 and 5). This work was used within the context of signal separation for a MIMO baseband processor for RF receivers demonstrating unprecedented performance for use in full-duplex radios, cognitive radios, and next generation adaptive communication systems. Further refinements to the design have led to a high-fidelity adaptive micro-power mixed-signal matrix-vector product integrated circuit (IC) enabling efficient implementations of independent component analysis (ICA), multiple signal classification (MUSIC), and other spatial processing algorithms. This has been the first reported micro-power Nyquist rate system demonstrating analog processing at 14 bit resolution. These multi-segment capacitive analog signal processors were able to achieve their high-resolution in part due to the the low-input capacitance unity gain buffer introduced in Chapter 3.

In order to improve upon the performance of such processors and achieve lower area, higher resolution, and increased energy efficiency, ASP systems must better exploit multisegment capacitive DACs. However, high-resolution systems built with multi-segment primitives incur major performance loss due to mismatch and process-variation resulting in radix errors. Consequently, Chapter 6 of this dissertation develops a class of algorithms, which we call *successive stochastic approximation*, aimed at overcoming these errors. This algorithm, S2A, is a modification to stochastic gradient descent that overcomes its shortcomings when applied to high-dimensional analog computation, specifically overcoming non-monotonicities introduced by radix errors due to capacitive mismatch. Using this we demonstrate 65 dB of sinusoidal interferer suppression in an over-the-air test with AM sources.

7.1   Outlook

This final chapter aims to provide a direction forward for this research. The algorithms and ICs developed in this dissertation integrate micro-watt power analog integrated circuits with advances in optimization and learning in order to gain

large reductions in system power required to perform adaptive signal processing functions. However, there remains significant work to be done in implementing various other kernels and accelerators to create a practical end-to-end system.

This dissertation focused on spatial filtering and adaptive signal processing for both sensory information processing and communication applications, we shall provide a brief overview of the direction of this study in both contexts.

7.1.1   Communication

Software oriented systems like traditional software-based cognitive radios, and software defined radios tend to put the burden of beamforming, waveform recognition, and symbol estimation on very high digital processing throughput requirements. High fidelity analog channelization [41] achieves large reduction in digital processing at the expense of having more complex and power inefficient analog filters. Building off the work presented in this dissertation, we can simplify the RF front-end channelization filtering, by using mixed-signal real-time ICA implemented on our analog co-processors to resolve multiple signals. These signals can reside across multiple sub-bands that relax RF channel bandwidths and filter roll-off design requirements. However, additional interference such as ISI can impact the channelization and classifier. Using the *S2A* algorithm and its extensions, can enable adaptive compensation for these sources of interference as well.

Architectural research into the interplay between these co-optimized hardware and algorithms in an end-to-end energy efficient, scalable, highly parallel microsystem remains under-explored. Further architectural research in this area should drive the development of receiver topologies better suited to emerging techniques like full-duplex communication, and cognitive radios. There has also been some research into the energetic advantages sub-Nyquist sampling and compressive sensing for communication applications [2]. These applications typically entail linear projection onto a random basis followed by reconstruction through complex optimization algorithms. Analog mismatch can degrade the performance of such systems, extending S2A to such compressive sensing based systems can lead to significant increase in system performance while meeting power constraints. In

parallel initiatives on mm-wave RADAR systems to enable super-resolution and direction sensitivity are being developed by exploiting the high-dynamic range exhibited by the mixed-signal spatial co-processor introduced in Chapters 4 and 5.

7.1.2 Sensory Signal Processing

A proliferation in IoT connected devices and sensors, in conjunction with advances in machine learning has led greater processing at the edge of the network. This has the benefit of enabling continuous learning on the device, tailoring each device to its environment, as well as preserving privacy, by limiting data sent to "the cloud". Thus there is a need for high-dynamic range spatial processing and multichannel feature extraction that can enable highly efficient sensory systems. Three fronts of exploration will need to be traversed when expanding this work for use in the IoT. First, at the level of systems and networking, there is a need for research into efficiencies gained from collectives of "smart" sensory nodes through coordinated co-operation. Second, an exploration of circuit topologies that can enable applications and technologies that heretofore not possible without analog processing, e.g., an implantable energy harvestingm multi-chip sensory system for brain-computer-interfaces. Finally, there is a need to develop more algorithms to improve the resilience and performance of analog processing.

A concrete example of research driving advances at the systems, circuits, and algorithmic levels would integrate signal acquisition and feature extraction with subsequent digital processing and data-telemetry. This would entail further research into high-fidelity analog spatial processing, MIMO communication, as well as further development in learning algorithms insensitive to analog-mismatch.

In concert with these developments, complementing them, there has been increasing effort from the devices community to develop computational primitives like oscillatory devices, as well as nonvolatile memories to overcome the memory bottlenecks associated with contemporary neurally inspired algorithms. These efforts provide a long term route to high-performance, low-power high-dimensional analog computational systems. The development of these computational devices has resulted in a parallel track of research into leveraging neurally inspired princi-

ples like stochastic, distributed computation to develop prototype neurally inspired learning hardware. While Chapter 2 briefly introduced non-volatile memories and their application to computation, there remains significant work to be done in order to realize large-scale systems exploiting such devices. These systems would target next-generation computational loads for large-scale data analysis and sensory processing. However, truly fundamental advances in computational efficiency of neurally inspired machine intelligence will entail a close collaboration between algorithm design, circuit design, and device development. While there has been some work enabling very efficient neural computation through resistive memories, these devices pose a significant power overhead in the peripheral circuits, preventing power savings. One means of overcoming these shortcomings would be to investigate charge-recovery and adiabatic techniques for efficiently driving large arrays of resistive memories.

7.2   Concluding Remarks

The current trends in technology are leading us towards distributed, large-scale intelligence through ambient adaptive sensory units providing input. Through continuous learning, and adaptation these units can be robust to environmental changes while ensuring they remain within resource constraints like energy, operating temperature, and delay. Through applications in bio-signal acquisition, brain-machine interfaces, autonomous systems, and large-scale data analytics, such systems will be directed at advancing pervasive and ambient intelligence with the capacity to vastly improve the quality of life.

Bibliography

[1] K. Abdelhalim, L. Kokarovtseva, J. L. P. Velazquez, and R. Genov. 915-MHz FSK/OOK wireless neural recording SoC with 64 mixed-signal FIR filters. *IEEE Journal of Solid-State Circuits*, 48(10):2478–2493, 2013.

[2] D. Adams, Y. C. Eldar, and B. Murmann. A mixer front end for a four-channel modulated wideband converter with 62-db blocker rejection. *IEEE Journal of Solid-State Circuits*, 52(5):1286–1294, 2017.

[3] A. Bahai. Ultra-low energy systems: Analog to information. In *2016 46th European Solid-State Device Research Conference (ESSDERC)*, pages 3–6, Sept 2016.

[4] D. Bankman and B. Murmann. An 8-Bit, 16 Input, 3.2 pJ/op Switched-Capacitor Dot Product Circuit in 28-nm FDSOI CMOS. In *IEEE Asian Solid-State Circuits Conference (A-SSCC)*, November 2016.

[5] D. Bharadia, E. McMilin, and S. Katti. Full duplex radios. *ACM SIGCOMM Computer Communication Review*, 43(4):375–386, 2013.

[6] S. Brenna, L. Bettini, A. Bonetti, A. Bonfanti, and A. L. Lacaita. Fundamental Power Limits of SAR and $\Delta\Sigma$ Analog-to-Digital Converters. In *Nordic Circuits and Systems Conference (NORCAS): NORCHIP & International Symposium on System-on-Chip (SoC), 2015*, pages 1–4. IEEE, 2015.

[7] F. N. Buhler, A. E. Mendrela, Y. Lim, J. A. Fredenburg, and M. P. Flynn. A 16-channel noise-shaping machine learning analog-digital interface. In *2016 IEEE Symposium on VLSI Circuits (VLSI-Circuits)*, pages 1–2, June 2016.

[8] G. W. Burr, R. M. Shelby, S. Sidler, C. di Nolfo, J. Jang, I. Boybat, R. S. Shenoy, P. Narayanan, K. Virwani, E. U. Giacometti, B. N. Kurdi, and H. Hwang. Experimental demonstration and tolerancing of a large-scale neural network (165 000 synapses) using phase-change memory as the synaptic weight element. *IEEE Transactions on Electron Devices*, 62(11):3498–3507, 2015.

[9] G. Cauwenberghs. Blind on-line digital calibration of multi-stage Nyquist-rate and oversampled A/D converters. In *Circuits and Systems, 1998. ISCAS '98. Proceedings of the 1998 IEEE International Symposium on*, volume 1, pages 508–511 vol.1, May 1998.

[10] G. Cauwenberghs. Reverse engineering the cognitive brain. *Proceedings of the National Academy of Sciences*, 110(39):15512–15513, 2013.

[11] A. Celik, M. Stanacevic, and G. Cauwenberghs. Gradient flow independent component analysis in micropower VLSI. In *Advances in Neural Information Processing Systems*, pages 187–194, 2005.

[12] A. Celik, M. Stanacevic, and G. Cauwenberghs. Gradient Flow Independent Component Analysis in Micropower VLSI. In *Adv. Neural Information Processing Systems (NIPS 2006)*, volume 8, pages 187–194. Cambridge: MIT Press, 2006.

[13] S. Chakrabartty and G. Cauwenberghs. Sub-microwatt analog vlsi trainable pattern classifier. *IEEE Journal of Solid-State Circuits*, 42(5):1169–1179, 2007.

[14] A. Chanthbouala, V. Garcia, R. O. Cherifi, K. Bouzehouane, S. Fusil, X. Moya, S. Xavier, H. Yamada, C. Deranlot, N. D. Mathur, M. Bibes, A. Barthélémy, and J. Grollier. A ferroelectric memristor. *Nature materials*, 11(10):860–864, 2012.

[15] P.-Y. Chen, D. Kadetotad, Z. Xu, A. Mohanty, B. Lin, J. Ye, S. Vrudhula, J.-s. Seo, Y. Cao, and S. Yu. Technology-design co-optimization of resistive cross-point array for accelerating learning algorithms on chip. In *Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition*, pages 854–859. EDA Consortium, 2015.

[16] Y. M. Chi, C. Maier, and G. Cauwenberghs. Ultra-high input impedance, low noise integrated amplifier for noncontact biopotential sensing. *IEEE JETCAS*, 1(4):526–535, Dec 2011.

[17] C. Y. Choo and H. Elabd. A memory reduction scheme for multi-channel echo canceller implementation. In *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, volume 5, pages 3301–3304 vol.5, 2001.

[18] M. Courbariaux, Y. Bengio, and J.-P. David. Binaryconnect: Training deep neural networks with binary weights during propagations. In *Advances in Neural Information Processing Systems*, pages 3123–3131, 2015.

[19] S. Douglas and T.-Y. Meng. Optimum error quantization for LMS adaptation. In *Communications, Computers and Signal Processing, 1991., IEEE Pacific Rim Conference on*, pages 704–708. IEEE, 1991.

[20] S. C. Douglas and T. H. Y. Meng. Normalized data nonlinearities for LMS adaptation. *IEEE Transactions on Signal Processing*, 42(6):1352–1365, Jun 1994.

[21] D. Floreano and R. J. Wood. Science, technology and the future of small autonomous drones. *Nature*, 521(7553):460–466, 2015.

[22] M. P. Frank. Introduction to reversible computing: motivation, progress, and challenges. In *Proceedings of the 2nd Conference on Computing Frontiers*, pages 385–390. ACM, 2005.

[23] Y. Freund. Boosting a weak learning algorithm by majority. In *COLT*, volume 90, pages 202–216, 1990.

[24] R. Genov and G. Cauwenberghs. Kerneltron: support vector "machine" in silicon. *IEEE Transactions on Neural Networks*, 14(5):1426–1434, 2003.

[25] R. Genov, G. Cauwenberghs, G. Mulliken, and F. Adil. A 5.9mW 6.5GMACS CID/DRAM array processor. In *Proceedings of the 28th European Solid-State Circuits Conference*, pages 715–718, Sept 2002.

[26] A. Ghaffari, E. Klumperink, F. van Vliet, and B. Nauta. Simultaneous spatial and frequency-domain filtering at the antenna inputs achieving up to +10 dBm out-of-band/beam P1dB. In *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 84–85, Feb 2013.

[27] A. Ghaffari, E. Klumperink, F. van Vliet, and B. Nauta. A 4-Element Phased-Array System With Simultaneous Spatial- and Frequency-Domain Filtering at the Antenna Inputs. *IEEE Journal of Solid-State Circuits*, 49(6):1303–1316, June 2014.

[28] R. J. Green and M. G. McNeill. Bootstrap transimpedance amplifier: a new configuration. *IEE Proceedings G - Circuits, Devices and Systems*, 136(2):57–61, Apr 1989.

[29] T. Grosse-Puppendahl, S. Herber, R. Wimmer, F. Englert, S. Beck, J. von Wilmsdorff, R. Wichert, and A. Kuijper. Capacitive near-field communication for ubiquitous interaction and perception. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Ubi-Comp '14, pages 231–242, New York, NY, USA, 2014. ACM.

[30] S. Han, X. Liu, H. Mao, J. Pu, A. Pedram, M. A. Horowitz, and W. J. Dally. EIE: Efficient Inference Engine on Compressed Deep Neural Network. In *2016 ACM/IEEE 43rd Annual International Symposium on Computer Architecture (ISCA)*, pages 243–254, June 2016.

[31] V. Hariprasath, J. Guerber, S. Lee, and U. Moon. Merged capacitor switching based SAR ADC with highest switching energy-efficiency. *Electronics Letters*, 46(9):620, 2010.

[32] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE Journal on elected Areas in Communications*, 23(2):201–220, Feb 2005.

[33] Q. He, G. W. Wornell, and W. Ma. An adaptive multi-band system for low power voice command recognition. In *Interspeech 2016*, pages 1888–1892, 2016.

[34] J. H. Hwang, T. W. Kang, Y. T. Kim, and S. O. Park. Measurement of transmission properties of HBC channel and its impulse response model. *IEEE Transactions on Instrumentation and Measurement*, 65(1):177–188, Jan 2016.

[35] I. Daubechies and R. DeVore and C. S. Gunturk and V. A. Vaishampayan. Beta expansions: a new approach to digitally corrected A/D conversion. In *2002 IEEE International Symposium on Circuits and Systems. Proceedings (Cat. No.02CH37353)*, volume 2, pages II–784–II–787 vol.2, 2002.

[36] M. Jabri and B. Flower. Weight perturbation: An optimal architecture and learning technique for analog VLSI feedforward and recurrent multilayer networks. *IEEE Transactions on Neural Networks*, 3(1):154–157, 1992.

[37] S. Joshi, C. Kim, S. Ha, M. Y. Chi, and G. Cauwenberghs. A 2pJ/MAC 14-b 8×8 Linear Transform Mixed-Signal Spatial Filter in 65 nm CMOS with 84 dB Interference Suppression. In *IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, February 2017.

[38] R. Karakiewicz, R. Genov, and G. Cauwenberghs. 480-GMACS/mW Resonant Adiabatic Mixed-Signal Processor Array for Charge-Based Pattern Recognition. *IEEE Journal of Solid-State Circuits*, 42(11):2573–2584, Nov 2007.

[39] R. Karakiewicz, R. Genov, and G. Cauwenberghs. 1.1 TMACS/mW Fine-Grained Stochastic Resonant Charge-Recycling Array Processor. *IEEE Sensors Journal*, 12(4):785–792, April 2012.

[40] C. Kim, S. Ha, C. Thomas, S. Joshi, J. Park, L. Larson, and G. Cauwenberghs. A 7.86 mW +12.5 dBm in-band IIP3 8-to-320 MHz capacitive harmonic rejection mixer in 65nm CMOS. In *European Solid State Circuits Conference (ESSCIRC)*, pages 227–230, Sept 2014.

[41] C. Kim, S. Joshi, C. Thomas, S. Ha, A. Akinin, L. Larson, and G. Cauwen-berghs. A CMOS 4-Channel MIMO Baseband Receiver with 65dB Harmonic Rejection over 48MHz and 50dB Spatial Signal Separation over 3MHz at 1.3mW. In *Symposium on VLSI Circuits (VLSIC)*, 2015.

[42] C. Kim, S. Joshi, C. M. Thomas, S. Ha, L. E. Larson, and G. Cauwenberghs. A 1.3 mW 48 MHz 4 channel MIMO baseband receiver with 65 db harmonic rejection and 48.5 dB spatial signal separation. *IEEE Journal of Solid-State Circuits*, 51(4):832–844, April 2016.

[43] C. Kim, S. Joshi, C. M. Thomas, S. Ha, L. E. Larson, and G. Cauwenberghs. A 1.3 mW 48 MHz 4 Channel MIMO Baseband Receiver With 65 dB Harmonic Rejection and 48.5 dB Spatial Signal Separation. *IEEE Journal of Solid-State Circuits*, 51(4):832–844, April 2016.

[44] D. B. Kirk, D. Kerns, K. Fleischer, and A. H. Barr. Analog VLSI imple-mentation of multi-dimensional gradient descent. *Adv. Neural Information Processing Systems (NIPS92)*, 5:789–796, 1993.

[45] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information pro-cessing systems*, pages 1097–1105, 2012.

[46] D. Kuzum, S. Yu, and H. P. Wong. Synaptic electronics: materials, devices and applications. *Nanotechnology*, 24(38):382001, 2013.

[47] G. Laput, C. Yang, R. Xiao, A. Sample, and C. Harrison. Em-sense: Touch recognition of uninstrumented, electrical and electromechanical objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software; Technology*, UIST '15, pages 157–166, New York, NY, USA, 2015. ACM.

[48] E. H. Lee and S. S. Wong. 24.2 A 2.5 GHz 7.7 TOPS/W switched-capacitor matrix multiplier with co-designed local memory in 40nm. In *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 418–419. IEEE, 2016.

[49] J. Lim, K. Kwon, and S.-I. Chae. Reversible energy recovery logic circuit without non-adiabatic energy loss. *Electronics Letters*, 34(4):344–346, Feb 1998.

[50] J. Lu, S. Young, I. Arel, and J. Holleman. A 1 tops/w analog deep machine-learning engine with floating-gate storage in 0.13 μm cmos. *IEEE Journal of Solid-State Circuits*, 50(1):270–281, 2015.

[51] K. Y. Ma, P. Chirarattananon, S. B. Fuller, and R. J. Wood. Controlled flight of a biologically inspired, insect-scale robot. *Science*, 340(6132):603–607, 2013.

[52] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski. Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, pages 145–151, 1996.

[53] B. Murmann. *Limits on ADC Power Dissipation*, pages 351–367. Springer Netherlands, Dordrecht, 2006.

[54] B. Murmann. Energy limits in A/D converters. In *Faible Tension Faible Consommation (FTFC), 2013 IEEE*, pages 1–4, June 2013.

[55] B. Murmann. On the use of redundancy in successive approximation A/D converters. In *Proc. IEEE Int. Conf. Sampling Theory and Applications (SampTA)*, pages 1–4, 2013.

[56] B. Murmann, D. Bankman, E. Chai, D. Miyashita, and L. Yang. Mixed-signal circuits for embedded machine-learning applications. In *2015 49th Asilomar Conference on Signals, Systems and Computers*, pages 1341–1345, Nov 2015.

[57] E. Neftci, B. Pedroni, S. Joshi, M. Al-Shedivat, and G. Cauwenberghs. Stochastic synapses enable efficient brain-inspired learning machines. *Frontiers in Neuroscience*, 10(3389):241:1–16, 2016.

[58] T. Ohno, T. Hasegawa, T. Tsuruoka, K. Terabe, J. K. Gimzewski, and M. Aono. Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. *Nature materials*, 10(8):591–595, 2011.

[59] F. Pan, S. Gao, C. Chen, C. Song, and F. Zeng. Recent progress in resistive random access memories: materials, switching mechanisms, and performance. *Materials Science and Engineering: R: Reports*, 83:1–59, 2014.

[60] L. Parra and C. Alvino. Geometric source separation: merging convolutive source separation with geometric beamforming. In *Neural Networks for Signal Processing XI, 2001. Proceedings of the 2001 IEEE Signal Processing Society Workshop*, pages 273–282, 2001.

[61] S. Paul, A. M. Schlaffer, and J. A. Nossek. Optimal charging of capacitors. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 47(7):1009–1016, Jul 2000.

[62] G. Peng, Z. Ignjatovic, and M. Bocko. Preamplifiers for non-contact capacitive biopotential measurements. In *IEEE EMBC, 2012*, pages 1482–1485, July 2013.

[63] R. Pickholtz, D. Schilling, and L. Milstein. Theory of spread-spectrum communications-a tutorial. *IEEE transactions on Communications*, 30(5):855–884, 1982.

[64] R. J. Prance, A. Debray, T. D. Clark, H. Prance, M. Nock, C. J. Harland, and A. J. Clippingdale. An ultra-low-noise electrical-potential probe for human-body scanning. *Measurement Science and Technology*, 11:291–297, Mar 2000.

[65] B. Razavi. Challenges in the design of cognitive radios. In *Custom Integrated Circuits Conference*, pages 391–398, Sept 2009.

[66] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Mathematica Hungarica*, 8(3-4):477–493, 1957.

[67] W. Rieutort-Louis, T. Moy, Z. Wang, S. Wagner, J. C. Sturm, and N. Verma. A large-area image sensing and detection system based on embedded thin-film classifiers. *IEEE Journal of Solid-State Circuits*, 51(1):281–290, Jan 2016.

[68] Saberi, Mehdi and Lotfi, Reza and Mafinezhad, Khalil and Serdijn, Wouter A. Analysis of power consumption and linearity in capacitive digital-to-analog converters used in successive approximation ADCs. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 58(8):1736–1748, 2011.

[69] R. Sarpeshkar. Analog versus digital: extrapolating from electronics to neurobiology. *Neural computation*, 10(7):1601–1638, 1998.

[70] R. E. Schapire. The strength of weak learnability. *Machine learning*, 5(2):197–227, 1990.

[71] C. L. Seitz, A. H. Frey, S. Mattisson, S. D. Rabin, D. A. Speck, and J. L. Van de Snepscheut. Hot clock nMOS. 1985.

[72] M. M. Shulaker, T. F. Wu, A. Pal, L. Zhao, Y. Nishi, K. Saraswat, H.-S. P. Wong, and S. Mitra. Monolithic 3d integration of logic and memory: Carbon nanotube fets, resistive ram, and silicon fets. In *2014 IEEE International Electron Devices Meeting*, pages 27–4. IEEE, 2014.

[73] Siddharth Joshi, Chul Kim, Sohmyung Ha and Gert Cauwenberghs. From Algorithms to Devices: Enabling Machine Learning through Ultra-Low-Power VLSI Mixed-Signal Array Processing. In *2017 IEEE Custom Integrated Circuits Conference*, April-May 2017, to appear.

[74] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[75] H. Simon. Adaptive filter theory. *Prentice Hall*, 2:478–481, 2002.

[76] A. B. Suksmono and A. Hirose. *Adaptive Beamforming by Using Complex-Valued Multi Layer Perceptron*, pages 959–966. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[77] T. J. Sullivan, S. R. Deiss, T.-P. Jung, and G. Cauwenberghs. A brain-machine interface using dry-contact, low-noise EEG sensors. In *2008 IEEE International Symposium on Circuits and Systems*, pages 1986–1989, May 2008.

[78] L. Szilard. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, 9(4):301–310, 1964.

[79] C. M. Thomas and L. E. Larson. Broadband synthetic transmission line N-path filter design. *IEEE Transactions on Microwave Theory and Techniques*, 63(10):3525–3536, Oct. 2015.

[80] J. van den Heuvel, J.-P. Linnartz, P. Baltus, and D. Cabric. Full MIMO Spatial Filtering Approach for Dynamic Range Reduction in Wideband Cognitive Radios. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 59(11):2761–2773, Nov 2012.

[81] J. Von Neumann. *The computer and the brain.* Yale University Press, 1958.

[82] Z. Wang, R. Schapire, and N. Verma. Error-adaptive classifier boosting (EACB): Exploiting data-driven training for highly fault-tolerant hardware. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3884–3888. IEEE, 2014.

[83] Z. Wang, J. Zhang, and N. Verma. Realizing Low-Energy Classification Systems by Implementing Matrix Multiplication Directly Within an ADC. *IEEE transactions on biomedical circuits and systems*, 9(6):825–837, 2015.

[84] E. W. Weisstein. Square point picking. 2004.

[85] S. J. Wright. Coordinate descent algorithms. *Mathematical Programming*, 151(1):3–34, 2015.

[86] S. Yu, P.-Y. Chen, Y. Cao, L. Xia, Y. Wang, and H. Wu. Scaling-up resistive synaptic arrays for neuro-inspired architecture: Challenges and prospect. In *2015 IEEE International Electron Devices Meeting (IEDM)*, pages 17–3. IEEE, 2015.

[87] Y. Yu, B. Krishnamachari, and V. K. Prasanna. Data Gathering with Tunable Compression in Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems*, 19(2):276–287, Feb 2008.

[88] D. Zhang and A. Alvandpour. A 12.5-ENOB 10-kS/s Redundant SAR ADC in 65-nm CMOS. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 63(3):244–248, 2016.

[89] Y. Zhang, R. Howver, B. Gogoi, and N. Yazdi. A high-sensitive ultra-thin MEMS capacitive pressure sensor. In *2011 16th International Solid-State Sensors, Actuators and Microsystems Conference*, pages 112–115, June 2011.